ABSTRACT

ZHANG, KUANG. Linkage Map Construction and QTL Analyses for Fiber Quality, Yield and Agronomic Traits Using 63K Array-based SNPs in Upland Cotton (Gossypium hirsutum L.) (Under the direction of Dr. Vasu Kuraparthy).

Improving the fiber quality and yield are the primary research objectives in cotton breeding to improve the economic viability and sustainability of cotton production. Identifying the quantitative trait loci (QTLs) for fiber quality and yield traits using the high-density SNP based genetic maps allows for bridging genomics with plant breeding through marker assisted selection and genomic selection. Recently, a 63K SNP array has been developed in cotton. In this study, two recombinant inbred lines (RILs) populations, derived from a cross between two upland cottons (NC05AZ06 and NC11-2091) and a cross between two upland cottons (NC05AZ06 and NC11-2100), which represent broad allele diversity in upland cotton, were used to construct high density SNP based linkage maps and QTL mappings. Genotyping was performed using Cotton 63K SNP array and phenotyping was performed on six fiber quality traits, four yield traits and two agronomic traits evaluated in two-year field trials, which were used to identify and map the QTLs.

In NC05AZ06 × NC11-2091 RILs population, a high-density genetic map containing 3,129 polymorphic SNP markers was constructed. The constructed map had a total genetic length of 4,422.44 cM with an average marker density of 1.41 cM. QTL mapping identified a total of 106 QTLs which included 59 QTLs for six fiber quality traits, 38 QTLs for four yield traits and 9 QTLs for two agronomic traits. Among these QTLs, 22 stable QTLs were identified in both years. Sub-genome wide, 57 QTLs were identified in the A sub-genome which included favourable alleles of 43 QTLs originated from NC05AZ06 while the other 14 QTLs were from NC11-2091. In the D sub-genome, 49 QTLs were detected in the D sub-genome which included

favourable alleles of 37 QTLs contributed by NC05AZ06 while the other 12 QTLs were from NC11-2091. Forty-six of the mapped QTLs explained more than 10% of phenotypic variation. Further, mapping identified twenty-one QTL clusters in the Upland cotton where 12 QTL clusters were mapped in the A sub-genome and 9 clusters were mapped in the D sub-genome. 75% of the QTLs with improved alleles were contributed by the parent NC05AZ06. More than seventy percent of the mapped QTLs shared the similar linkage and physical position with previously reported QTLs.

In NC05AZ06 × NC11-2100 RILs population, 2,017 polymorphic SNPs markers were used to contrast the high-density genetic map, with a total genetic length of 3,837.46 cM and an average marker density of 1.9 cM. A total of 77 QTLs were detected, including 40 QTLs for six fiber quality traits, 27 QTLs for four yield traits and 10 QTLs for two agronomic traits. In the A sub-genome, 32 QTLs were identified, which included 24 QTLs with favourable alleles derived from NC05AZ06 while the other 8 were from NC11-2100. 45 QTLs were detected in the D sub-genome which included 27 QTLs with favourable alleles contributed by NC05AZ06 while the other 18 were from NC11-2100. Of these 77 QTLs, 39 were major QTLs, explaining more than 10% of phenotypic variation. 16 QTLs clusters were detected in this population, including 6 QTL clusters mapped in the A sub-genome and 10 clusters mapped in the D sub-genome. More than sixty percentage of the QTLs with improved alleles were contributed by the parent NC05AZ06. Almost 70% mapped QTLs shared the similar linkage and physical position with previously reported QTLs.

Linkage Map Construction and QTL Analyses for Fiber Quality, Yield and Agronomic Traits Using 63K Array-based SNPs in Upland Cotton (Gossypium hirsutum L.)

by
Kuang Zhang

A thesis submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the degree of
Master of Science

Crop Science

Raleigh, North Carolina
2019

APPROVED BY:

_____	_____
Vasu Kuraparthy Committee Chair	Randy Wells
_____	
Gina Brown-Guedira	

BIOGRAPHY

Kuang Zhang was born on May 2, 1992 in Hangzhou, Zhejiang, China. After graduating from high school at Hangzhou No.2 High School in Hangzhou, Zhejiang in 2010, Kuang obtained his bachelor's degree in Agronomy from Zhejiang University in Hangzhou, Zhejiang in 2014. In the summer 2014, Kuang came to NC State through "3+X" program between Zhejiang University and NC State, as a temporary non-degree student, and officially became a graduate student in Crop Science from 2014 and earned a master of Statistics as an additional degree from N.C State University in 2018.

ACKNOWLEDGMENTS

I would first want to thank to my major professor. Dr. Vasu Kuraparthy, for his support and guidance in my graduate studies. I would also like to express my appreciation to the other members of my committee, Dr. Randy Wells and Dr. Gina Brown-Guedira, for their comments and suggestions.

I am grateful to my mother, father and grandpa, who provided me both financial and emotional support in my life. I can't finish this study journey without their loving and encouragement.

Finally, last but by no means least, thanks to all my previous and current lab mates: Hui Fang, Linglong Zhu and Baljindar Kaur, who provided me a lot of help on my research work.

Thanks for all your encouragement!

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Literature review

Overview

Cotton (*Gossypium L.*) is one of the most important economic crops worldwide, accounting for 80% (118 of 146 million bales) of the world natural fibers production and 25% (118 of 472 million bales) of the total global fiber production (including manmade fiber) in 2018 [1]. By taxonomic classification, cotton (*Gossypium* L.) genus belongs to the tribe *Gossypieae*, in the *mallow* family, *Malvaceae*, native to the several isolated regions in the New World and Old World [2]. Four *Gossypium* species are cultivated: *G. arboreum, G. barbadense, G. herbaceum* and *G. hirsutum.* Among these four cultivated species, two tetraploid (2n = 4x = 52, AADD) species *G. hirsutum* (also called upland cotton) and *G. barbadense* (also known as pima cotton) are recently grown in the US, account for 97.6% and 2.4% of the total acres of harvested cotton in 2018 [3]. The United States is the third largest cotton fiber production country and the largest cotton fiber exporting country in the world. In recent years, the US fiber market is facing enormous challenges from the Asia cotton [4] and synthetic fibers [5]. Therefore, continued improvements in the fiber quality and yield is critical for future success of U.S. cotton.

With the development of molecular marker technology and linkage map construction, maker-assisted selection (MAS) [6] has been applied to the cotton breeding programs to fulfill the demand for quickly improving cotton fiber quality and yield simultaneously.

Cotton taxonomy and evolution

The cotton genus (*Gossypium*) was grouped into four subgenera and eight sections, comprising approximately 50 species, including 45 diploid (2n = 2x = 26) species and 5 tetraploid (2n = 4x = 52, AADD) species [7,8]. All these 50 species were classified into 9 genome groups with eight diploid genome groups (A through G plus K) and a tetraploid group

(AD) [2,9], which originated from worldwide areas including Africa (A, B, F), Asia (A), Arabia (E), Australia (C, G, K) and The New World (D, AD) [2,7,8,9]. In these regions, the use of the coarse hairs covering the seeds of four different cotton species: *G. arboreum* (A_2), *G. herbaceum* (A_1), *G. barbadense* (AD_2), and *G. hirsutum* (AD_2) were discovered inhabitants [2], which had evolved to become 4 cultivated species presently. It is believed that the tetraploid (AD) species are allopolyploids evolving from two diploid genomes, including an A genome species (*G. arboreum* or *G. herbaceum)* found in Africa or Asia and a D genome species (*G. raimondii*) from America [9,10].

Cotton fiber production

Because of their high yield performance and good fiber quality, the allotetraploid cotton cultivated species (especially *G. hirsutum*) have become the main cultivars grown in the world.Only small acreages of *G. arboreum* L. and *G. herbaceum* L. are grown in India and droughty regions of Africa and Asia, respectively [11] recently. *G. hirsutum* L. and *G. barbadense* L. account for approximately 90 and 8%of the global cotton production, respectively [11,12]. In the 2018/19 season, more than 118 million bales of cotton fiber was produced in approximately 30 countries [4].

Cotton production in the US

The United States is the third largest cotton fiber production country and the largest cotton fiber exporting country in the world. In 2018, the U.S. produced 18.91 million bales of upland cotton and 0.77 million bales of pima cotton [13]. Fifteen million bales of cotton were exported in the 2018 season [4]. In the US, cotton is grown and produced in 17 states, including Texas (6.53 million bales), Georgia (2.8 million bales), Mississippi (1.45 million bales), Arkansas (1.15 million bales), Alabama (1.1 million bales), Oklahoma (0.95 million bales),

California (0.88 million bales), Missouri (0.82 million bales), North Carolina (0.81 million bales), Tennessee (0.8 million bales), South Carolina (0.55 million bales), Arizona (0.49 million bales), Louisiana (0.42 million bales), Kansas (0.34 million bales), Florida (0.22 million bales), Virginia (0.22 million bales) and New Mexico (0.16 million bales) [13]. California is the main *G. barbadense* cotton-growing state, occupying more than 90% of the pima cotton production (0.71 of 0.77 million bales) in the U.S. [13]. The production, distribution and processing of cotton in United States provided more than 200,000 jobs and $27 billion direct business revenue [14].

Other major cotton producing countries

China is the country with the largest cotton production, importation, and utilization in the world. In 2018, China produced 27.5 million bales of cotton. In addition to this, 7.5 million bales of cotton were imported for a total of 40.5 million bales of cotton for textile processing [4].

The second largest cotton producing country is India. India is also the third largest cotton exporting country as well as the second largest in utilization. In 2018, India produced 27 million bales and imported 4.5 million bales of cotton. [4].

Brazil and Pakistan are two other major cotton production countries around the world, with cotton fiber production of 11.4 and 7.5 million bales, respectively, in 2018 season [4]. Brazil is also the second largest cotton exporting country. During 2018, Brazil exported 6.2 million bales of cotton [4].

Man-made fiber production

Synthetic fibers are the largest group of man-made fibers by far and are derived from chemical resources. In recent years, synthetic fibers have increased in popularity due to their low

price and high quality [15]. During 2018, 298 million bales of synthetic fibers were produced worldwide while the total production of cotton was only 118 million bales [1].

Future goal of cotton production in the US

The US cotton is always reliable in the world for its high quality. However, in order for US cotton to remain the competitive in domestic and international fiber markets, lower price and higher fiber quality is required [1, 4, 15]. One of the best ways to reduce the price is to increase the lint yield of the cotton. Therefore, improving the yield and fiber quality traits has been the important objectives of cotton breeders in the US.

Cotton breeding for traits improvement

Traditional breeding methods played an important role in the history of cotton cultivar development. Prior to 1900, the study of genetics was unknown [16]. The early breeders selected plants based on their phenotypes or visible plant characteristics and saved the seeds for the next year. By 1900, the concept of "the Mendel's laws" discovered by Mendel in 1865 was rediscovered by scientists, thus laying the foundation for modern plant breeding [17,18]. In conventional plant breeding, there are six basic types of cultivars developed by plant breeders. These are pure-line, open-pollinated, hybrid, clonal, apomictic and multiline cultivars [19]. The development of cotton cultivars has a long history in the U.S. The first recorded cotton culture was in 1607 by the British at Jamestown, Virginia in 1607 [2]. During 1700s, two cultivars, Georgia Green-seed, (*G. hirsutum*) and Creole Black-seed, (*G. hirsutum*) were widely grown in the US [2]. After 1850, large numbers of cultivars were developed across the US [2]. Starting from the early 1900s modern plant breeding methods led to a golden period of cultivar development [20-22]. Cotton breeding research was established at the University of Arkansas (UA). From 1920 to now, 27 cotton cultivars have been released by the UA cotton breeding

program, including several advanced cultivars (UA48, UA222, UA103, UA107, UA114, UA212ne) with high yield and excellent fiber quality [20]. The New Mexico cotton breeding program was initiated in 1926. The NM cotton breeding program is famous for its Acala 1517 cultivars which are tolerant or resistant to Verticillium wilt and bacterial blight [21]. From 1929 to now, 38 cotton cultivars have been released by the NM cotton breeding program, contributing to the development of 45% of the commercial cotton cultivars [21]. The Delta and Pine Land (D&PL), founded in 1886, is a leader company in cottonseed production [22]. In 1911, the company purchased one of the largest cotton plantations in the Mississippi Delta for cotton production. In the latter half of the 20th century, the cotton varieties developed by D&PL, such a Deltapine, Paymaster, and Sure-Grow, were popular and being planted by the cotton farmers throughout the U.S [22].

Bottlenecks of traditional cotton breeding

Cotton fiber quality traits and yield traits are known to inherit polygenically and influenced by environment [23, 24, 25]. Furthermore, fiber quality traits often have negative association with some yield traits due to the linkage and pleiotropic effects between lint yield components and fiber quality parameters [26]. In 1997, Smith et al. reported that fiber strength and length were negatively associated and with the most basic within-boll lint yield components [27], which was also reported by Basal et al in 2009 [26]. Therefore, further improvement in the trait values especially for the quantitative traits of yield and fiber quality using traditional breeding only approaches have been limited [28, 29].

Cotton breeding with maker-assisted selection (MAS)

With the advancement of molecular marker technology, maker-assisted selection (MAS) [28, 29] has been applied to the cotton breeding programs to overcome the aforementioned negative

associations. Marker-assisted selection allows identification and preservation of the best genotypes linked to the traits of interest [29].

Molecular markers and linkage mapping in cotton breeding

Genetic markers identify regions on chromosomes where genes of interest reside and are used for the construction of a linkage map [30]. Plant breeders select for the trait of interest by indirectly selecting for the markers instead of the exact genes thus lowering the level of difficulty. [30]. Genetic markers include morphological, cytological, biochemical and molecular markers [30]. In breeding programs, molecular markers are the most popular type of genetic markers with two advantages: (1) they are not be limited or influenced by environment, (2) they are large in number [30].

Genetic linkage occurs when DNA sequences on a chromosome are transmitted and tend to be inherited together. A linkage map, constructed by calculating recombination frequencies between genes or genetic markers on chromosomes, shows the position of genes or genetic markers relative to one another [30]. The basic calculation for linkage analysis is the logarithm of odds (LOD), which measures the likelihood of linkage between two markers. After calculating the LOD values, the linkage groups are established and the distance between a pair of markers within each group are estimated by their recombination frequencies. These in turn are converted to centiMorgans (cM) through the use of mapping functions (the Kosambi method or the Haldane method) [30].

Molecular markers in cotton linkage mapping

Based on the differences of the methods used for detecting the markers, the molecular markers can be classified into three groups, including (1) Hybridization-based markers, ("First generation molecular markers", RFLP) (2) PCR-based markers, ("Second generation molecular

markers: RAPD, AFLP, SRAP, SSR and others) and (3) sequence-based markers, ("New generation molecular markers" SNP, EST and many more) [30].

Restriction Fragment Length Polymorphism (RFLP): RFLP is the only hybridization-based molecular marker. The polymorphisms are the result of variations in insertion/deletions (known as InDels), point mutations, translocations, duplications and inversions yielding DNA fragments of different length [32,33,34]. RFLP was the first type of the markers used in the cotton improvement [35] and resulted in the first linkage map in cotton in 1994. The resulting genetic mapping population comprised 57 F_2 individuals from a cross between *G. hirsutum* L. race "palmeri" and *G. barbadense* L. acc. "K101" by Reinisch et al. [36]. Rong et al. (2004) [37] used 1884 RFLPs to form a high-density linkage map with the same population based on Reinisch's study. Another RFLP genetic linkage map on a F_2 population from a cross between *G. hirsutum* acc. TMS-22 and *G. tomentosum* acc. WT936 was constructed for improving the fiber quality from the wild type Hawaiian cotton (*G. tomentosum*) [38]. The major disadvantage of RFLP is that it requires large amount of DNA for restriction digest and Southern blotting, which is laborious and time consuming [30].

Random Amplified Polymorphic DNA (RAPD): RAPD is a PCR-based marker system utilizing the single, short and random primed regions to scan polymorphisms, developed independently by Williams et al. [39] and Welsh and Mcclelland [40]., It requires smaller quantities of DNA than RFLP. Amplification occurs when the positions are complementary to the primer's sequence during PCR [39, 40]. In 1995, Multani and Lyon [41] firstly fingerprinted the 14 cotton cultivars with RAPD markers. Since the number of RAPD markers is limited in cotton linkage mapping programs, RAPDs were used always with other types of molecular markers [42,43]. Zhang et al. (2002) reported a linkage map with 489 RAPDs and SSRs for a population from the cross

between *G. hirsutum* L. and *G. barbadense* L. [42]. In 2005, a linkage map for a *G. hirsutum* L. × *G. barbadense* L. F_2 mapping population was constructed by Lin et al. based on 62 RAPDs, 141 SSRs and 363 SRAP [42]. Use of RAPD markers in linkage mapping is limited because of its low reliability and reproducibility [33,34].

Amplified Fragment Length Polymorphism (AFLP): AFLP, developed by Zabeau and Vos [44], is visualized by selective PCR amplification of DNA restriction fragments, which combines the reliability of RFLP and the convenience of PCR, overcoming the limitations of RFLP and RAPD [30,33,34]. In 1997, an AFLP based linkage map was formed on a tri-species, F_2 population, with 194 AFLPs, 19 RAPDs and 3 morphological markers [45]. Yang et al. (2009) reported an AFLP-SSR based linkage map from a cross between CRI8 (*Gossypium hirsutum* L.) and Pima 90-53 (*Gossypium barbadense* L.) [46]. Although AFLP has more advantages than RFLP and RAPD, its high cost and time requirement dampened acceptance by the breeders [33,34].

Sequence-Related Amplified polymorphism (SRAP): Li and Quiros (2000) developed the SRAP, a simple marker technique aimed for the PCR amplification of open reading frames (ORFs), which is based on two-primer amplification [47]. SRAP is a simple and efficient marker system used widely in map construction, genomic and cDNA fingerprinting [33]. Lin et al. (2005) used 363 SRAPs with 114 SSRs and 62 RAPDs to construct the linkage map and map QTLs for cotton fiber quality on the cotton interspecific population [43]. In 2007, a high-density linkage map was formed by Yu et al. with 129 SRAPs, 697 SSRS, 171 TRAPs and 98 AFLPs for a F_2 mapping population from a cross between *G. barbadense* ("Hai 7124") and *G. hirsutum* ("CRI 36") [48].

Simple-Sequence Repeat (SSR): SSR, also called Microsatellites, are tandem repeat motifs of 1-8 nucleotides presenting abundantly and randomly in eukaryotic genomes, which are highly polymorphic [49]. Microsatellites are the most widely used PCR-based markers [30]. In recent years, SSR-based linkage mapping is still popular in cotton breeding programs [51-54]. Shang et al. (2015) constructed a genetic map on a RIL population of upland cotton with 581 SSR markers [51]. Tan et al. (2015) reported a high-density genetic map of an upland cotton intraspecific RIL population with 1333 polymorphic SSR markers, by screening 25313 SSR primer pairs [51]. Jamshed et al. (2016) used 851 SSR markers to form the genetic map of a RIL population from a cross between two upland cottons ("0–153" and "sGK9708") [53]. A genetic map for a maternal backcross population was reported by Ma et al (2017), with 623 SSR markers [54]. However, the marker numbers and time cost are still the limitations of PCR-based markers [33,34].

Single Nucleotide Polymorphism (SNP): SNPs are singe base changes exist in genome sequence, which provide the simplest and most abundant markers [33]. SNP genotyping techniques are being improved rapidly with the development of Next-generation sequencing (NGS) to achieve the demand for high throughput markers at a low cost [29,34]. In recent years, two types of NGS based SNPs genotyping methods are being prevalent in cotton genetic mapping studies: Genotyping-by-sequencing (GBS) and SNP array based genotyping[55-62]. GBS discovers the SNPs from sequencing of GBS DNA libraries based on reducing genome complexity with restriction enzymes [63]. Qi et al. (2017) constructed a high-density genetic map using GBS in a F_2 intraspecific population of upland cotton with 3187 SNP markers [55]. A high-density linkage map with 3978 GBS SNPs, which associated with a F_2 population from a cross between two upland cottons (ZHONG213 and LU28), was built by Li et al. (2017) [56]. 3557 SNPs discovered by GBS, were used to construct a high-density genetic map of pima cottons

(*Gossypium barbadense*) RIL population in 2018 by Fan et al [57]. Diouf et al. (2018) used 5178 GBS SNPs to form a high-density linkage map for a $F_{2:3}$ population developed from a cross of two upland cottons [58]. High-throughput genotyping arrays are designed for the discovery of SNPs. Recently, two sets of cotton SNPs array: CottonSNP63K and CottonSNP80k were developed and applied to the cotton breeding research [64,65]. Li et al (2016) built a linkage map with 2618 SNPs genotyped by the cotton 63K SNP array in the upland cotton RIL population [59]. Ulloa et al. (2017) used 3824 CottonSNP63K array-based SNPs to construct a high-density genetic consensus map for 3 intraspecific mapping populations of upland cotton [60]. A high-density genetic map in upland RIL population was formed by Tan et al. (2018) using the CottonSNP80K array with 12116 SNP markers [61]. 525 SNPs were used to construct a 63K SNP chip-based linkage map in pima cotton (*Gossypium barbadense* L.) F_2 population by Kumar et al. (2019) [62]. Various high-throughput genotyping methods make SNPs as the most attractive markers in recent breeding studies [33].

Previous high-density cotton linkage maps

Rong et al. (2004) constructed a high-density linkage map with a total of 2584 RFLP, SSR and SNP markers in a F_2 population from a cross between upland cotton race "palmeri" and pima cotton acc. "K101.", which spanned a total genetic length of 4447.9 cM with an average marker density of 1.72 cM [37].1308 markers were on the At subgenome with a total length of 2325.7 cM and an average marker density of 1.78 cM. 1276 markers were on the Dt subgenome with a total length of 2122.2 cM and an average marker density of 1.66 cM [37].

A high-density genetic map of a F_2 interspecific cotton population (*G. hirsutum* L. "CRI 36" × *G. barbadense* L. "Hai 7124"), based on 1097 genetic markers (697 SSRs, 171 TRAPs, 129 SRAPs, 98 AFLPs and two morphological markers), was built by Yu et al. (2007) [48]. This

resulted in genetic map length of 4536.8 cM with an average distance of 4.1 cM between markers. Of these 1097 markers, 590 markers were mapped to the A sub-genome with a genetic length of 2215.2 cM while 490 markers were mapped to the D sub-genome with a total length of 2203 cM [48].

Yu et al. (2011) used 2316 SSR markers to construct a high-density genetic map from a BC_1 population between *G. hirsutum* and *G. barbadense,* with a total genetic length of 4418.8 cM and 1.91 cM in average distance between adjacent markers [66]. A subgenome contained 1043 markers and spanned 2250.1 cM with an average density of 2.16 cM. D subgenome contained 1273 markers and spanned 2168.8 cM with an average density of 1.70 cM [66].

1274 SSRs were used to form a high-density linkage map in a RIL population from a cross between two upland cottons (CCRI 25 and Yumian 1) by Tan et al. (2015), which spanned 3076.4 cM with an average marker interval of 2.41 cM [52]. The total genetic lengths of the A subgenome and D subgenome were 1462.6 cM and 1613.8 cM, with 500 SSRs on A subgenome and 774 SSRs on D subgenome [52].

A high-density linkage map for a BC_1F_1 population of *G.hirsutum* × *G.barbadense* constructed using 2292 SSR markers was reported by Shi et al. (2015) [67]. This map covered 5115.16 cM of the total genome with an average marker density of 2.23 cM. 1025 SSRs were mapped on the A subgenome with a total coverage of 2420.71 cM and an average marker interval of 2.36 cM. D subgenome comprised 1267 SSRs with a total coverage of 2694.45 cM and an average marker interval of 2.12 cM [67].

Zhang et al. (2017) reported a high-density genetic map in an intraspecific RIL population of upland cotton built with 2316 SNPs and 77 SSRs, which spanned a total length of 2865.73 cM with an average distance of 1.20 cM between adjacent markers [68]. In the A

subgenome, 1072 markers were mapped with a total linkage distance of 1380.9 cM and an average marker interval of 1.29 Cm. In the D subgenme, 1321 markers were mapped, which spanned a genetic length of 1484.83 cM with an average marker density of 1.15 cM [68].

Tan et al. (2018) using the CottonSNP80K array to construct a high-density linkage map of an upland cotton RIL population with 12116 SNPs [61]. It covered 3741.81 cM with an average distance of 0.31 cM between consecutive SNPs. The A subgenome had 5975 SNP markers and spanned 1979.72cM with an average marker interval of 0.33 cM. 5141 SNPs covering 1752.09 cM of genetic distance were mapped on D subgenome with an average marker density of 0.29 cM [61].

5178 SNP markers were used for mapping a $F_{2:3}$ intraspecific population by Diouf et al. (2018) [58]. A genetic map spanned 4768.10 cM with an average distance of 0.92cM was generated, including 3313 SNPs in A subgenome with a coverage length of 2611.43 cM and 1865 SNPs in D subgenome with a total distance of 2156.67 cM, which accounted for the average marker interval of 0.79 cM and 1.156 cM, respectively [58].

Identification of QTL for cotton traits

A quantitative trait is a measurable phenotype varies from individual to individual according to genetic and environmental influences, which can be either continuous or discrete [69]. Quantitative trait loci (QTLs) are some genetic loci with the alleles affect the variation of the quantitative traits. In general, quantitative traits are affected by multiple polymorphic genes and environmental factors, so a quantitative trait can be influenced by one or many QTLs [69]. Since the exact gene locus of interest is usually hard to be observed, the mapping of genetic markers is used to detect the QTLs for trait of interest [30]. Several QTL mapping methods has been developed for identifying QTLs: (1) Single-marker analysis method (2) Simple interval

mapping method (3) Composite interval mapping method (4) Multiple interval mapping method (5) Bayesian interval mapping method [70].

QTLs reported for cotton fiber quality traits

Many QTLs for cotton fiber quality traits were reported in previous studies. Yang et al. (2009) detected 25 fiber quality QTLs in a F_2 population from a cross between upland cotton and pima cotton, with a range of phenotypic variation explained (PVE) from 10.2% to 52.8%, including 5 QTLs for fiber length (FL), 4 QTLs for uniformity index (UI), 7 QTLs for micronaire (MIC), 7 QTLs for fiber strength (FS) and 2 QTLs for fiber elongation (FE) [46]. 28 QTLs for fiber quality traits were reported by Zhang et al. (2011) from a F2 population of the cross between *G. hirsutum* (acc.TMS-22) and G. *tomentosum* (acc. WT936), with four QTLs for FE, eight for MIC, four for FL, four for FS and six for UI [38]. Sun et al. (2012) identified 50 QTLs for fiber quality in the F_2, $F_{2:3}$ and RIL population of upland cotton, including 10 for FS, 10 for FL, 10 for MIC, 8 for UI and 12 for FE, with 2 stable QTLs for FS on chromosomes A7 and D12 detected simultaneously among all 3 generations [71]. 3 different upland cotton populations were used as materials to detect 77 fiber quality QTLs by Shao et al. (2014), including 19 for FL, 14 for UI, 17 for MIC, 10 for FE and 17 for FS, with 5 stable QTLs identified across two or three populations [72]. In an interspecific (*G. hirsutum* L. × *G. barbadense* L.) RIL population, Yu et al. (2014) detected 60 significant fiber quality QTLs evenly distributed in A subgenome and D subgenome, of which 48.3% were stable QTLs identified in multiple environments [73]. A total of 62 QTLs for fiber quality explaining phenotypic variation from 5.0% to 28.1% in an upland cotton RIL population were reported by Tang et al. (2015), with 17 QTLs detected in more than one environment [74]. Jamshed et al. (2016) reported 165 fiber quality QTLs with 47 stable QTLs identified across multiple environments in a *Gossypium hirsutum* RIL population,

including 31 for FL, 35 for FS, 32 for FE, 32 for UI and 35 for MIC [53]. An intraspecific RIL population of upland cotton with phenotypic values measured across 5 years and 6 locations were used as materials to detect 63 QTLs for fiber strength, including 16 stable fiber strength QTLs detected in multiple environment by Zhang et al. (2017) [68]. Keerio et al. (2018) identified 30 fiber quality QTLs for a *G.hirsutum* × *G. tomentosum* introgression lines population with a range of PVE from 2.73% to 21.61%, including 4 QTLs for MIC, 10 QTLs for FL, 1 QTL for UI, 6 QTLs for FS and 9 QTLs for FE [75].

The CottonQTLdb is a public cotton QTL database (http://www.cottonqtldb.org; Release 2.3) aiming to collect all available QTL data in cotton, which contains 4892 QTLs from 156 publications currently (http://www2.cottonqtldb.org:8081/references) [23,24,76]. A total of 2070 QTLs for 6 fiber quality traits were obtained from this database, including 442 QTLs for MIC, 524 QTLs for upper half mean length (UHML) or FL, 289 QTLs for UI, 470 QTLs for FS, 287 QTLs for FE and 58 QTLs for SFC (Appendix A). Among these 2070 QTLs, 921 QTLs were mapped on A subgenome: A1(65), A2(45), A3(57), A4(50), A5(117), A6(71), A7(101), A8(86), A9(74), A10(67), A11(55), A12(86), A13(47), and 1149 QTLs were mapped on D subgenome: D1(110), D2(80), D3(80), D4(58), D5(79), D6(123), D7(50), D8(96), D9(27), D10(91), D11(144), D12(140), D13(71) (Appendix A) [76].

Micronaire (MIC) is an index of fiber fineness and maturity, which are critical for fiber processing and fiber quality. 442 QTLs for MIC were collected in CottonQTLdb, unevenly mapped on 26 chromosomes: A1(18), A2(10), A3(12), A4(9), A5(35), A6(18), A7(13), A8(14), A9(22), A10(11), A11(6), A12(18), A13(11), D1(17), D2(23), D3(24), D4(17), D5(13), D6(24), D7(14), D8(16), D9(6), D10(13), D11(28), D12(30), D13(20) (Appendix A) [76].

Upper half mean length (UHML) is a length parameter of cotton fiber, measuring the mean length of the longest 50% of a fiber sample. Fiber length is an important quality trait for cotton fiber because the longer fibers can produce stronger yarn during textile processing. There were totally 524 fiber length QTLs reported by the previous studies in CottonQTLdb: A1(27), A2(6), A3(23), A4(16), A5(28), A6(15), A7(26), A8(37), A9(12), A10(18), A11(15), A12(16), A13(9), D1(23), D2(16), D3(15), D4(10), D5(21), D6(36), D7(12), D8(32), D9(5), D10(22), D11(26), D12(33), D13(25) (Appendix A) [76].

Fiber uniformity index (UI), calculated by the ratio of the mean length to the upper half mean length, is a measure of the uniformity of fiber lengths of the sample. Fiber uniformity is important for ensuring overall mechanical properties in textile appliacations. 289 UI QTLs from previous publications were collected in CottonQTLdb: A1(3), A2(11), A3(6), A4(7), A5(14), A6(12), A7(12), A8(10), A9(11), A10(12), A11(6), A12(21), A13(10), D1(22), D2(8), D3(9), D4(14), D5(11), D6(16), D7(10), D8(8), D9(3), D10(12), D11(21), D12(10), D13(10) (Appendix A) [76].

Fiber strength (FS) is measured by how much force (in grams) is needed to break a bundle of fibers one tex unit in size (one tex = the weight in grams of 1000 meters of fiber), which is one of the most significant fiber quality traits for fiber processing and product application [2]. In CottonQTLdb, 470 QTLs for FS were obtained: A1(9), A2(5), A3(10), A4(10), A5(21), A6(11), A7(34), A8(20), A9(17), A10(10), A11(20), A12(14), A13(10), D1(26), D2(12), D3(23), D4(11), D5(25), D6(17), D7(7), D8(24), D9(5), D10(36), D11(42), D12(43), D13(8) (Appendix A) [76].

Fiber elongation (FE) is specified as a percentage of fiber stretch before break, which is of decisive importance for high quality textile products. 287 FE QTLs were collected in

CottonQTLdb: A1(7), A2(8), A3(4), A4(7), A5(19), A6(11), A7(12), A8(4), A9(12), A10(13), A11(6), A12(11), A13(7), D1(18), D2(20), D3(7), D4(3), D5(9), D6(24), D7(6), D8(14), D9(8), D10(8), D11(22), D12(22), D13(5) (Appendix A) [76].

Short fiber content (SFC) is the percentage of fibers shorter than 0.5 inches (12.7 mm). Fiber with high SFC may be adverse to its quality. In CottonQTLdb, 58 QTLs were reported by previous studies on 20 chromosomes: A1(1), A2(5), A3(2), A4(1), A6(4), A7(4), A8(1), A10(3), A11(2), A12(6), D1(4), D2(1), D3(2), D4(3), D6(6), D7(1), D8(2), D11(5), D12(2), D13(3) (Appendix A) [76].

QTLs reported for yield quality traits

Qin et al. (2009) detected 43 yield QTLs and 66 yield QTLs in two different F_2 upland cotton populations, respectively, with 12 QTLs found simultaneously in both populations [77]. 12 significant (LOD>3.0) QTLs and 48 putative QTLs for yield traits were reported by Yu et al. (2014) for an interspecific RIL population from a cross between *G. hirsutum* (acc. TM-1) and *G. barbadense* (acc. 3-79), including 46.2% of them on A subgenome and 53.8% of them on D subgenome [73]. 70 yield QTLs were detected independently in seven environments from a 178 RILs population of two upland cottons (acc. DH962 and cv. Jimian5) by Wang et al. (2015), explaining 4.4% to 9.93% of phenotypic variation [78]. Li et al. (2016) used an upland cotton RIL population of 188 individual lines as materials to detected 23 QTLs for yield traits, including 8 for boll weight (BW) with 4 major QTLs and 15 for lint percentage (LP) with 4 major QTLs [59]. An introgression lines population derived from a cross between *G.hirsutum* and G.*tomentosum* was used by Keerio et al. (2018) to identify 44 QTLs for yield traits with a range of PVE from 1.74% to 29.58% [75]. Liu et al. (2018) detected 122 yield QTLs with 1.68%-28.27% of PVE in an intraspecific upland cotton RILs population, including 53 QTLs for

BW, 39 QTLs for LP and 30 QTLs for seed index (SI) [79]. 13 QTLs were identified for yield traits in a pima cotton F_2 population (*G. barbadense "*Suvin" × *G. barbadense "*BCS 23-18-7") by Kumar et al. (2019) with PVE ranged from 1% to 28% [62].

A total of 692 QTLs for 4 yield traits were obtained from this database, including 176 QTLs for BW, 327 QTLs for LP, 147 QTLs for SI and 42 QTLs for lint index (LI) (Appendix A). Among these 692 QTLs, 339 QTLs were mapped on A subgenome: A1(21), A2(12), A3(37), A4(10), A5(32), A6(20), A7(54), A8(9), A9(23), A10(24), A11(17), A12(30), A13(50), and 353 QTLs were mapped on D subgenome: D1(27), D2(23), D3(42), D4(28), D5(23), D6(19), D7(4), D8(29), D9(52), D10(19), D11(22), D12(42), D13(23) (Appendix A) [76].

BW is the average weight in grams of a boll. 176 QTLs for BW were collected in CottonQTLdb on 24 chromosomes: A1(1), A2(5), A3(3), A4(2), A5(11), A6(10), A7(5), A9(2), A10(8), A11(5), A12(7), A13(5), D1(6), D2(4), D3(5), D4(4), D5(15), D6(2), D8(14), D9(18), D10(1), D11(3), D12(34), D13(6) (Appendix A) [76].

LP is a ratio between the total fiber weight and the total boll weight. In CottonQTLdb, 327 QTLs for LP were recorded on all chromosomes: A1(7), A2(6), A3(25), A4(7), A5(11), A6(9), A7(25), A8(5), A9(13), A10(8), A11(6), A12(17), A13(39), D1(9), D2(11), D3(36), D4(16), D5(4), D6(11), D7(2), D8(8), D9(13), D10(13), D11(13), D12(2), D13(11) (Appendix A) [76].

SI is the weight of 100 seeds in grams. 147 SI QTLs were reported by previous publications in CottonQTLdb: A1(10), A2(1), A3(8), A5(7), A6(1), A7(23), A8(4), A9(6), A10(5), A11(4), A12(4), A13(4), D1(8), D2(6), D3(1), D4(6), D5(4), D6(4), D7(1), D8(5), D9(19), D10(5), D11(6), D12(2), D13(3) (Appendix A) [76].

LI is the weight of lint in grams obtained from 100 seeds. 42 LI QTLs were collected in CottonQTLdb mapped on 19 chromosomes: A1(3), A3(1), A4(1), A5(3), A7(1), A9(2), A10(3), A11(2), A12(2), A13(2), D1(4), D2(2), D4(2), D6(2), D7(1), D8(2), D9(2), D12(4), D13(3) (Appendix A) [76].

QTLs reported for plant height

Qin et al. (2009) detected 3 and 11 QTLs for plant height explaining 3.58% to 8.45% of PV in two different F_2 upland cotton populations, respectively, with 2 stable QTLs identified in multiple environments [77]. Shang et al. (2015) reported 41 QTLs and 23 conditional QTLs for plant height traits at two experimental environments in an intraspecific RILs population of upland cotton, including 4 stable QTLs identified in both environments [51]. Seven PH QTLs were mapped in a F_2 upland cotton population on chromosomes A5 and D12 by Qi et al. (2017), explaining 3.49% to 21.51% of PVE, with a stable QTL on D12 detected in two environments [55].

Out of 122 QTLs for plant height collected in CottonQTLdb, 45 QTLs were mapped on A subgenome among 11 chromosomes: A1(3), A2(1), A3(4), A5(11), A6(1), A7(1), A8(6), A9(2), A10(4), A11(7), A13(5), and 77 QTLs were mapped on D subgenome in all 13 chromosomes: D1(14), D2(3), D3(2), D4(3), D5(1), D6(6), D7(1), D8(5), D9(11), D10(4), D11(6), D12(14), D13(7) (Appendix A) [76].

References

1. The Fiber Year Consulting. The Fiber Year 2018 World Survey on Textiles & Nonwovens. 2018; 18.
2. Smith C, Cothren J. Cotton**:** Origin, History, Technology, and Production. John Wiley & Sons, Inc. 1999.
3. U.S. Department of Agriculture. National Agricultural Statistics Service.2018. https://www.nass.usda.gov/Statistics_by_Subject/result.php?CDC03D5A-502B-344F-9014-F3F3729512C1&sector=CROPS&group=FIELD%20CROPS&comm=COTTON#skipnav.
4. U.S. Department of Agriculture. Cotton: World Markets and Trade. USDA Foreign Agricultural Service. 2018. https://apps.fas.usda.gov/psdonline/circulars/cotton.pdf
5. OECD/Food and Agriculture Organization of the United Nations. "Cotton", in OECD-FAO Agricultural Outlook 2018-2027, OECD Publishing, Paris/Food and Agriculture Organization of the United Nations, Rome. 2018. https://doi.org/10.1787/agr_outlook-2018-13-en
6. Ribaut JM and Hoisington D. Marker-assisted selection: new tools and strategies. Trends in Plant Science. 1998; 3(6), 236-238.
7. Fryxell PA. The natural history of the cotton tribe. Texas A and M University Press, College Station, TX. 1979.
8. Fryxell PA. A revised taxonomic interpretation of *Gossypium* L. (Malvaceae). Rheedea 2. 1992; 108–165.

9. Wendel J, Brubaker C, Alvarez I, et al. Evolution and Natural History of The Cotton Genus. *In Genetics and Genomics of Cotton.* Plant Genetics and Genomics: Crops and Models 3. 2009; DOI 10.1007/978-0-387-70810-2_1.
10. Beasley JO. The origin of American tetraploid *Gossypium* species. Amer. Nat. 1940; 74:285–286.
11. Jenkins J. Cotton. Traditional crop breeding practices: an historical review to serve as a baseline for assessing the role of modern biotechnology. OECD. 1993; 61-68.
12. Shim J, Mangat PK, Angeles-Shim RB. Natural variation in wild *Gossypium* species as a tool to broaden the genetic base of cultivated cotton. J. Plant Sci. Curr. 2018; Res 2: 005.
13. U.S. Department of Agriculture. Crop production. National Agricultural Statistics Service. 2018; https://www.nass.usda.gov/Publications/Todays_Reports/reports/crop0918.pdf.
14. National Cotton Council of America. Overview of the U.S. cotton industry. 2011; https://www.cotton.org/pubs/cottoncounts/upload/Cotton-Industry-Overview_Jan-19-2011.pdf
15. U.S. Department of Agriculture. Foreign Agricultural Service. Cotton Losing Share to Man-Made Fibers. 2012. https://www.fas.usda.gov/data/cotton-losing-share-man-made-fibers.
16. Monsanto. History of plant breeding. Innovations in plant breeding. 2017; https://monsanto.com/app/uploads/2017/06/history-of-plant-breeding.pdf
17. George Acquaah. Overview and historical perspectives. *In Principles of Plant Genetics and Breeding.* John Wiley&Sons, Ltd. 2012; 1-22.
18. Corcos AF, and Monaghan. FV. Gregor Mendel's Experiments on Plant Hybrids: A Guided Study. New Brunswick, NJ: Rutgers University Press. 1993.

19. George Acquaah. Selection methods. *In Principles of Plant Genetics and Breeding.* John Wiley&Sons, Ltd. 2012; 301-374.

20. Bourland M. History of Cotton Breeding and Genetics at the University of Arkansas. The Journal of Cotton Science. 2018; 22:171-182.

21. Zhang J. History and Progress in Cotton Breeding, Genetics, and Genomics in New Mexico. The Journal of Cotton Science. 2018; 22:191-210.

22. Delta and Pine Land Company records, Special Collections Department, Mississippi State University Libraries. http://library.msstate.edu/FindingAid/Delta_and_Pine_Land_Company_records_finding_aid_MSS.101.pdf

23. Said J, Lin Z, Zhang X, Song M and Zhang J. A comprehensive meta QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. BMC Genomics. 2013; 14(1): 776.

24. Said J, Song M, Wang H, Lin Z, Zhang X et al. A comparative meta-analysis of QTL between intraspecific *Gossypium hirsutum* and interspecific *G. hirsutum* × *G. barbadense* populations. Mol Genet Genomics. 2015; 290:1003–1025.

25. Lacape JM, Llewellyn D, Jacobs J, Arioli T, et al. Research article Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* x *G. barbadense* RIL population. BMC Plant Biology. 2010; 10:132.

26. Basal H, Unay A, Canavar O and Yavas I. Combining ability for fiber quality parameters and within-boll yield components in intraspecific and interspecific cotton populations. Span J Agric Res. 2009; 7(2): 364-374.

27. Smith CW, and Coyle GG. Association of Fiber Quality Parameters and Within-Boll Yield Components in Upland Cotton. Crop Sci. 1997; 37:1775-1779.

28. Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, et al. Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding. 1997; 3:87-103.

29. Boopathi NM, Thiyagu K, Urbi B, Santhoshkumar M, et al. Marker-Assisted Breeding as Next-Generation Strategy for Genetic Improvement of Productivity and Quality: Can It Be Realized in Cotton? International Journal of Plant Genomics. 2011; 2011: 16.

30. George Acquaah. Molecular breeding. *In Principles of Plant Genetics and Breeding*. John Wiley&Sons, Ltd. 2012; 383-470.

31. Cramer CS, Havey MJ. Morphological, Biochemical, and Molecular Markers in Onion. HortScience. 1999; 34(4): 589-593.

32. Botstein D, White RL, Skolnik M and Davis RW. Construction of a genetic linkage map in using restriction fragment length polymorphisms. Am. J. Hum. Genet. 1980; 32: 314–331.

33. Nadeem MA, Nawaz MA, Shahid MQ, et al. DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnology & Biotechnological Equipment.2018; 32(2): 261-285.

34. Rahman MU, Zafar Y, Paterson AH. *Gossypium* DNA Markers: Types, Numbers, and Uses. *In Genetics and Genomics of Cotton*. Plant Genetics and Genomics: Crops and Models 3. Springer Science Business Media, LLC. 2009; DOI: 10.1007/978-0-387-70810-2_5.

35. Meredith WR Jr. Contributions of introductions to cotton improvement, in "Use of plant introductions in cultivar development, Part 1," Shands HL and Wiesner LE eds. Crop Science Society of America. Madison, WI. 1991; 127–146.

36. Reinisch AJ, Dong J, Brubaker CL, Stelly DM, Wendel JF and Paterson AH. A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics. 1994; 138(3):829–847.

37. Rong J, Abbey C, Bowers JE, Brubaker CL, et al. A 3347-Locus Genetic Recombination Map of Sequence-Tagged Sites Reveals Features of Genome Organization, Transmission and Evolution of Cotton (*Gossypium*). Genetics. 2004; 166: 389-417.

38. Zhang Z, Rong J, Waghmare VN, Chee PW, et al. QTL alleles for improved fiber quality from a wild Hawaiian cotton, *Gossypium tomentosum.* Theor Appl Genet. 2011; 123:1075–1088.

39. Williams JG, Kubelik AR, Livak KJ, et al. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res. 1990;18(22):6531– 6535.

40. Welsh J, McClelland M. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res. 1990;18 (24):7213–7218.

41. Multani D, Lyon B. Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome. 1995; 38: 1005-1008.

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

43. Lin Z, He D, Zhang X, Nie Y, Guo X, et al. Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. Plant Breeding. 2005; 124: 180-187.

44. Zabeau M and Vos P. Selective restriction fragment amplification: a general method for DNA fingerprinting. 1993. European Patent Application number: 92402629.7, publication number 0 534 858 A1.

45. Altaf MK, Stewart JM, Wajahatullah MK et al. Molecular and Morphological Genetics of a Trispecies F_2 Population of Cotton. *Proceedings of the Beltwide Cotton Conference.* 1997; 1:448-452.

46. Yang X, Wang Z, Zhang G, Pan Y, et al. Construction of Molecular Genetic Map and QTL Analysis of Fiber Quality in Cotton. ACTA AGRONOMICA SINICA. 2009; 35(12): 2159-2166.

47. Li G, Quiros CF. Sequence-related amplified polymorphism (SRAP), a new marker system based on a simple PCR reaction: its application to mapping and gene tagging in Brassica. Theor Appl Genet. 2001;103(2–3):455–461.

48. Yu J, Yu S, Lu C, Wang W, Fan S, et al. High-density Linkage Map of Cultivated Allotetraploid Cotton Based on SSR, TRAP, SRAP and AFLP Markers. Journal of Integrative Plant Biology. 2007; 49(5): 716-724.

49. Powell W, Machray GC and Provan J. Polymorphism revealed by simple sequence repeats. Trends Plant Sci. 1996; 1: 215–222.

50. Vieira MLC, Santini L, Diniz AL and Munhoz CF. Microsatellite markers: what they mean and why they are so useful. Genet Mol Biol. 2016; 39(3): 312-328.

51. Shang L, Liu F, Wang Y, Abduwell A, Cai S, et al. Dynamic QTL mapping for plant height in Upland cotton (*Gossypium hirsutum*). Plant Breeding. 2015; 134(6): 703-712.

52. Tan Z, Fang X, Tang S, Zhang J, et al. Genetic map and QTL controlling fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Euphytica. 2015; 203:615-628.

53. Jamshed M, Jia F, Gong J, Palanga KK, et al. Identification of stable quantitative trait loci (QTLs) for fiber quality traits across multiple environments in *Gossypium hirsutum* recombinant inbred line population. BMC Genomics.2016. 17:197.

54. Ma L, Zhao Y, Wang Y, Shang L and Hua J. QTLs Analysis and Validation for Fiber Quality Traits Using Maternal Backcross Population in Upland Cotton. Front. Plant Sci. 2017; 8:2168.

55. Qi H, Wang N, Qiao W, Xu Q, et al. Construction of a high-density genetic map using genotyping by sequencing (GBS) for quantitative trait loci (QTL) analysis of three plant morphological traits in upland cotton (*Gossypium hirsutum* L.). Euphytica. 2017; 213:83.

56. Li L, Zhao S, Su J, Fan S, et al. High-density genetic linkage map construction by F2 populations and QTL analysis of early-maturity traits in upland cotton (*Gossypium hirsutum* L.). PLoS One. 2017; 12(8): e0182918.

57. Fan L, Wang L, Wang X, Zhang H, et al. A high-density genetic map of extra-long staple cotton (Gossypium barbadense) constructed using genotyping-bysequencing based single nucleotide polymorphic markers and identification of fiber traits-related QTL in a recombinant inbred line population. BMC Genomics. 2018; 19:489.

58. Diouf L, Magwanga RO, Gong W, He S, et al. QTL Mapping of Fiber Quality and Yield-Related Traits in an Intra-Specific Upland Cotton Using Genotype by Sequencing (GBS). Int. J. Mol. Sci. 2018; 19:441.

59. Li C, Dong Y, Zhao T, Li L, Li C, et al. Genome-Wide SNP Linkage Mapping and QTL Analysis for Fiber Quality and Yield Traits in the Upland Cotton Recombinant Inbred Lines Population. Front. Plant Sci. 2016; 7: 1356.

60. Ulloa M, Hulse-Kemp AM, Santiago LMD, Stelly DM and Burke JJ. Insights Into Upland Cotton (Gossypium hirsutum L.) Genetic Recombination Based on 3 High-Density Single-Nucleotide Polymorphism and a Consensus Map Developed Independently With Common Parents. Genomics Insights. 2017; 10: 1-15.

61. Tan Z, Zhang Z, Sun X, Li Q, Sun Y, et al. Genetic Map Construction and Fiber Quality QTL Mapping Using the CottonSNP80K Array in Upland Cotton. Front. Plant Sci. 2018; 9:225.

62. Kumar NVM, Katageri IS, Gowda SA, Adiger S, Yadava SK et al. 63K SNP chip based linkage mapping and QTL analysis for fiber quality and yield component traits in *Gossypium barbadense* L. cotton. Euphytica. 2019; 215: 6.

63. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. PLoS ONE. 2011; 6(5): e19379.

64. Hulse-Kemp, AM, Lemm J, Plieske J, Ashrafi H, Buyyarapu R, et al. Development of a 63K SNP array for cotton and high-density mapping of intraspecific and interspecific populations of *Gossypium spp.* G3 (Bethesda). 2015; 5(6): 1187–1209.

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

66. Yu Y, Yuan D, Liang S, Li X, et al. Genome structure of cotton revealed by a genome-wide SSR genetic map constructed from a BC1 population between *gossypium hirsutum* and *G. barbadense*. BMC Genomics. 2011; 12:15.

67. Shi Y, Li W, Li A, Ge R, at al. Constructing a high-density linkage map for *Gossypium hirsutum* × *Gossypium barbadense* and identifying QTLs for lint percentage. Journal of Integrative Plant Biology. 2015; 57(5): 450-467.

68. Zhang Z, Ge Q, Liu A, Li J, Gong J, et al. Construction of a High-Density Genetic Map and Its Application to QTL Identification for Fiber Strength in Upland Cotton. Crop Science. 2017; 57: 774-788.

69. Abiola O, Angel JM, Avner P, Bachmanov AA, Belknap JK, et al. The nature and identification of quantitative trait loci: a community's view. Nat Rev Genet. 2003; 4(11): 911-6;

70. Wang S, Basten CJ, and Zeng ZB. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. 2012.

71. Sun F, Zhang J, Wang S, Gong W, Shi Y, et al. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Mol Breeding. 2012; 30: 569-582.

72. Shao Q, Zhang F, Tang S, Liu Y, Fang X, et al. Identifying QTL for fiber quality traits with three upland cotton (*Gossypium hirsutum* L.) populations. Euphytica.2014; 198:43–58.

73. Yu JZ, Ulloa M, Hoffman SM, Kohel RJ, Pepper AE, et al. Mapping genomic loci for cotton plant architecture, yield components, and fiber properties in an interspecific

(*Gossypium hirsutum* L. × *G. barbadense* L.) RIL population. Mol Genet Genomics. 2014; 289(6): 1347-1367.

74. Tang S, Teng Z, Zhai T, Fang X, Liu F, et al. Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (*Gossypium hirsutum* L.). Euphytica. 2015; 201: 195-213.
75. Keerio AA, Shen C, Nie Y, Ahmed MM, Zhang X and Lin Z. QTL Mapping for Fiber Quality and Yield Traits Based on Introgression Lines Derived from *Gossypium hirsutum* × *G. tomentosum.* Int. J. Mol. Sci. 2018; 19: 243.
76. Said JI, Knapka JA, Song M, and Zhang J. Cotton QTLdb: a cotton QTL database for QTL analysis, visualization, and comparison between *Gossypium hirsutum* and *G. hirsutum*× *G. barbadense* populations. Mol. Genet. Genomics. 2015; 290(4):1615-1625.
77. Qin Y, Liu R, Mei H, Zhang T and Guo W. QTL Mapping for Yield Traits in Upland Cotton (*Gossypium hirsutum* L.). ACTA AGRONOMICA SINICA. 2009; 35(10): 1812-1821.
78. Wang H, Huang C, Guo H, Li X, Zhao W, Dai B, et al. QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments. PLOS ONE. 2015;10(6): e0130742.
79. Liu R, Gong J, Xiao X, Zhang Z, Li J, et al. GWAS Analysis and QTL Identification of Fiber Quality Traits and Yield Components in Upland Cotton Using Enriched High-Density SNP Markers. Front. Plant Sci. 2018; 9:1067.

CHAPTER 2: Introduction

Background

The cotton genus *Gossypium spp.* consists of at least 51 species, with 45 diploid (2n=2x=26) and six allotetraploid (2n=4x=52, AD) [1, 2] species. Of these only four are cultivated species: *G. hirsutum* L. (2n=4x, AADD), *G. barbadense* L. (2n=4x, AADD), *G. arboreum* L. (2n=2x, AA) and *G. herbaceum* L. (2n=2x, AA). The *G. hirsutum* L., also called Upland cotton, contributes to more than 90% of the global cotton production and acreage and *G. barbadense* L., known as sea island cotton, accounts for 8% of the cotton production in the world [3].

As the largest natural fiber source, cotton (*Gossypium* sps.) is one of the most important economic crops in the world. In 2018/19 season, cotton was primarily grown in around 30 countries, with more than 118 million bales of fiber produced [4]. In United States, which was the third largest cotton fiber production country as well as the largest cotton fiber exporting country in the world, 18.59 million bales of cotton fiber was produced with 15 million bales exported in 2018/19 season [4]. The production, distribution and processing of cotton in United States provided more than 200 thousand jobs and $27 billion direct business revenue [5]. However, the world cotton fiber market is recently under a lot of pressure because of the development of synthetic fibers [6]. In addition, the US cotton has to compete with hand-picked cotton from Asia. US cotton could compete in the international markets because of higher fiber quality. Therefore, improving the fiber quality has been an important objective of cotton breeders. Farm productivity and economic viability of cotton production directly related to the lint yields [5]. Therefore, continued improvements in the fiber quality and yield is critical for the U.S. cotton.

Plant height, a typical quantitatively inherited trait [7-9], can indirectly influence the yield of cotton fiber because appropriate plant height may contribute to machine harvesting and help achieve higher harvesting index [7]. Fuzziness seed trait, an important agronomic trait related to the cotton yield and fiber quality [10], was usually regarded as a binomial trait (fuzzy seed or fuzzless seed), however, some reports showed that fuzziness seed trait in majority of the cotton accessions was also controlled by multiple loci [10-13].

Cotton fiber quality traits and yield traits are known to inherit polygenically and influenced by environment [14, 15, 16]. Furthermore, fiber quality traits often have negative association with some yield traits [17]. Traditional breeding methods played an important role in the history of cotton cultivar development [18, 19]. However, further improvement in the trait values especially for the quantitative traits using traditional breeding only approaches have been limited [20, 21]. With the advancement of molecular marker technology, maker-assisted selection (MAS) [22] has been applied to the cotton breeding programs to overcome the bottlenecks. Restriction fragment length polymorphism (RFLP) was the first type of the markers used in the cotton improvement [23] and the first RFLP linkage map in cotton were constructed using RFLP in 1994 [24]. From then on, various types of the markers were used in the cotton breeding study, including: random amplified polymorphic DNA(RAPD) [25-28], amplified fragment length polymorphism (AFLP) [29, 30], target region amplified polymorphism (TRAP) [30], sequence related amplified polymorphism (SRAP) [27, 30], simple sequence repeats (SSR) [27, 28, 30-33], inter-simple sequence repeat (ISSR) [28, 34], Insertion-deletion polymorphisms (InDels) [33] and single nucleotide polymorphism (SNP) [32, 33, 35, 36]. With the progress in the number of traits mapped and diversity of genetic mapping populations utilized, higher density genetic maps with broadly adaptable markers are required for improving the efficiency of

detection and MAS based transfer of quantitative trait loci (QTLs) [33]. One way to construct high-density genetic maps is combining the different types of markers together to enrich the maps [27, 28, 30, 32, 33]. With the development of next-generation sequencing (NGS) technology, several efficient ways to discover large numbers of SNP-based markers for high-density linkage map construction has been applied in the breeding research, which have made the SNPs the most popular markers used in plant translational genomics [37].

Outline

Identifying the quantitative trait loci (QTLs) for fiber quality and yield traits using the high-density SNP based genetic maps allows for bridging genomics with plant breeding through marker assisted selection and genomic selection. Recently, two sets of cotton SNPs array: CottonSNP63K and CottonSNP80k were developed and applied to the QTL research [38,46].

In the present study, we used 63K SNP array [38] for genotyping two upland cotton recombinant inbred lines (RILs) populations. One RILs population was derived from a cross between two upland cottons (NC05AZ06 and NC11-2091), another RILs population was derived from a cross between two upland cottons (NC05AZ06 and NC11-2100). The genotyped SNPs for two RILs populations were used to construct the high-density linkage maps and to identify, map and analyze the QTLs for cotton fiber quality, yield and agronomic traits in cotton genome.

CHAPTER 3: Methods

Construction of the RILs populations

The G. hirsutum accessions NC05AZ06, NC11-2091 and NC11-2100 were used as parents to develop two RILs populations. NC05AZ06 is a sub-okra germplasm line with superior fiber quality and yield released by our program [65]. The landrace accession NC11-2091(TEX 2313; PI 607640) originating from Thailand and the landrace accession NC11-2100 (TEX 2324; PI 607650) originating from India, were obtained from the U.S. National Cotton Germplasm Collection (NCGC), USDA-ARS, College Station, Texas.

NC05AZ06 × NC11-2091 RILs population construction

In the summer of 2010, the parents NC05AZ06 and NC11-2091 were planted and crossed to develop F1 seeds at the Central Crops Research Station at Clayton, North Carolina. The F2 plants were grown and individual plants were manually selfed to obtain F3 seed in the summer nursery 2012. From 2013 to 2015, 107 of the 120 F2:3 lines were advanced to F5 generation by single seed decent method in the greenhouses. The 107 F5:6 lines were seed increased by manual self-pollination at Clayton in 2016. Seed cotton samples were ginned, acid delinted and used in the current study.

NC05AZ06 × NC11-2100 RILs population construction

In the summer of 2010, the parents NC05AZ06 and NC11-2100 were planted and crossed to develop F1 seeds at the Central Crops Research Station at Clayton, North Carolina. The F2 plants were grown and individual plants were manually selfed to obtain F3 seed in the summer nursery 2013. In the Summer of 2015, 112 of the 120 F2:3 lines were advanced to F4 generation by single seed decent method in the greenhouses. The 112 F4:5 lines were obtained by self-

pollination in the greenhouses in 2016. Seed cotton samples were ginned, acid delinted and used in the current study.

Field experiments and phenotyping

In the summer of 2016, the NC06AZ06 × NC11-2091 F_5 RILs population with parents and four checks UA-48, UA-222, DP-393, SG-747, were planted using an augmented randomized complete block design with seven blocks at Clayton. Each line was planted in the single row 10-ft plot with 2.5-3 seeds per ft. Fifty fully opened bolls from each plot were hand-harvested in November of each year of the trials. Four yield traits, including boll weight (BW), lint percentage (LP), seed index (SI), lint index (LT) were evaluated. Approximately 15 g of fiber sample ginned from each boll sample was evaluated for the fiber quality parameters using high-volume instrument (HVI) system at the Cotton Incorporated, Cary, North Carolina. The fiber quality traits evaluated were fiber elongation (FE), micronaire (MIC), short fiber content (SFC), fiber strength (FS), upper half mean length (UHML) and uniformity index (UI). MIC is an airflow measurement of fibers and indicates fiber fineness and maturity. UHML is the mean length of the longer half of the fibers in the sample, measured in hundredths of an inch. FS is the force in grams required to break a bundle of fibers one tex unit in size. FE is the amount a fiber sample will stretch prior to breakage. UI is a ratio between the mean length and the upper half mean length of the fibers, expressed as a percentage. It indicates the uniformity of fiber length in a sample. SFC is the percent by weight of fibers 0.5 inches (12.7 mm) long or less. BW is the average weight in grams of a boll. LP is a ratio between the total fiber weight and the total boll weight. SI is the weight of 100 seeds in grams. LI is the weight of lint in grams obtained from 100 seeds. An agronomic trait, fuzziness grade of seed (FG) was determined by four levels of the

seed fuzziness (0, 33.3%, 66.6%, 100%). Progressive numbers 0 to 100% indicate fuzz-free to fuzz-rich cotton seed.

In the summer of 2017, the NC06AZ06 × NC11-2091 F6 RILs population and the NC06AZ06 × NC11-2100 F5 RILs population with their parents and four checks were planted respectively at Clayton using the completely randomized block design with two replications. Each line was planted in the single 20-ft plot with a plant density of 2.5 seeds per foot. Forty fully opened bolls from each plot were hand-harvested in December 2017. Same phenotyping methods are used for evaluating the 11 cotton traits evaluated in 2016. And another agronomic trait, plant height (PH) was evaluated by taking the average of the manually-measured height of five randomly selected plants from each plot.

Marker genotyping and linkage map construction

Genomic DNA was extracted from young leaves of the RILs populations and their parents using DNeasy Plant Mini Kit (Qiagen). One hundred and four of the 107 phenotyped NC06AZ06 × NC11-2091 RILs and eighty-two of the 112 phenotyped NC06AZ06 × NC11-2100 RILs as well as their parents were genotyped with 63K cotton SNP array developed by Hulse-Kemp et al (2015) at Texas A&M Institute for Genome Sciences and Society. Candidate SNPs were filtered from the array with 63,058 SNPs based on the rules as follows: (1) SNPs with monomorphic genotypes were removed. (2) Poor-quality SNPs and SNPs with missing values more than 30% were removed. (3) Duplicate SNPs were removed [48].

The remaining candidate SNPs were used to construct the linkage maps by Joinmap 4.1 [66] based on *Kosambi's* mapping function [67] with logarithm of the odds (LOD) threshold of 7.0. Then the SNPs were aligned to the TM-1 (*G. hirsutum*) Genome NAU-NBI Assembly v1.1 & Annotation v1.1 database [39] by BLAST (https://www.cottongen.org/blast). Matching of the

linkage map groups with the physical map groups was performed with the circos plots by Circa (http://omgenomics.com/circa/).

Data analysis

All the trait phenotypic values of the RILs and parents were estimated using the linear mixed models in SAS version 9.4 (SAS Institute, Cary, NC). The SAS software was also used for calculating the study-related statistics, including the T-test of the difference between the value means of two parents, the broad-sense heritability of the traits, the genetic correlation between the traits and basic statistics.

In Joinmap 4.1, segregation of the markers from the Mendelian ratio 1:1 was tested using chi-square analysis ($P<0.05$) and a segregation distortion region (SDR) was identified when at least three adjacent markers showing a significant ($P<0.05$) segregation distortion [68].

All 12 traits related QTLs were detected using composite interval mapping (CIM) method [69] in WinWTLCart2.5 [70].

For the RIL population NC06AZ06 × NC11-2091, the genotype of alleles from NC05AZ06(P1) was coded as "AA" and the genotype of alleles from NC11-2091(P2) was coded as "aa". Based on the result of a 1000 times permutation procedure, logarithm of the odds (LOD) ≥2.5 with at least one year's phenotypic variation explained (PVE) ≥ 5.0 was used as the threshold for a QTL identified in both years with overlap region and LOD ≥ 3.0 with PVE ≥ 5.0 was the threshold to determine a QTL detected only in one year.

For the RIL population NC06AZ06 × NC11-2100, the genotype of alleles from NC05AZ06(P1) was coded as "AA" and the genotype of alleles from NC11-2100(P2) was coded as "aa". Based on the result of a 1000 times permutation procedure, LOD ≥ 3.0 with PVE ≥ 5.0 was the threshold to determine a QTL.

All the QTLs were labeled based on their population, trait type, and chromosome information. For example, QTLs for micronaire in population NC06AZ06 × NC11-2091 were labeled as qMIC-CH*-* (detected in both year), qMIC-16-CH*-* (detected in year 2016) or qMIC-17-CH*-* (detected in year 2017). QTLs for micronaire in population NC06AZ06 × NC11-2100 were labeled as q2MIC-CH*-* (* as a number).

The resulting linkage maps with identified QTLs were drawn using Mapchart Version 2.32 [71]. Further, the identified QTLs were used to detect the QTL clusters and meta QTL analysis was performed by comparing them with the QTLs reported in previous studies. All these previous reported QTLs' information was downloaded from the CottonQTLdb database (http://www.cottonqtldb.org; Release 2.3) developed by Said et al. [40].

CHAPTER 4: Results for NC05AZ06 × NC11-2091 RILs population

Analyses of the phenotypic traits

A summary of the statistical analysis for the phenotypic performance of the twelve traits are presented in Table 4.1. Among the six fiber quality traits measured, micronaire (MIC), upper half mean length (UHM), uniformity index (UI) and fiber strength (FS) of NC05AZ06 were significantly higher (13.0%~16.9%, 34.1%~36.6%, 4.4%~7.6%, 7.4%~8.1%, respectively) than those of NC11-2091 while the short fiber content (SFC) of NC11-2091 was much greater (26.3%~55.3%) than NC05AZ06. No significant difference was found between the two parents for the fiber elongation (FE). Among the four yield traits, boll weight (BW), lint percentage (LP), seed index (SI) and lint index (LI) were significantly higher (209.4%~222.8%, 137.2%~160.0%, 12.5% ~ 24.6%, 311.8%~317.9%, respectively) in NC05AZ06 than in NC11-2091. For agronomic traits, the plant height (PH) of NC05AZ06 was significantly lower (-32.5%) than NC11-2091. The seed fuzziness grade (FG) of NC05AZ06 is 100% (fuzz-rich) and the FG of NC11-2091 is 0 (fuzz-free). The broad-sense heritability of the traits calculated by the ratio of total genetic variance to total phenotypic variance for all the traits is listed in Table 4.2. Most traits except PH had high broad-sense heritability across two years with values ranging from 82%-96%. The broad-sense heritability of PH was only 56%, which indicated that the trait performance of PH was highly influenced by the environment.

Correlation coefficient analyses for the twelve traits is described in Table 4.3. For the fiber quality traits, UHM was significantly ($P<0.01$) positively correlated with UI, BW, LP, LI, FG and significantly($P<0.01$) negatively correlated with MIC, FE and SFC. FS showed significantly positive correlation with BW ($P<0.05$), SI ($P<0.01$), PH ($P<0.05$) and significantly negative correlation with FE ($P<0.05$), LP ($P<0.01$). SFC was significantly ($P<0.01$) positively

related to MIC, FE and it was significantly (P<0.01) negatively related to UI. FE was significantly (P<0.01) positively correlated with MIC and significantly negatively related to UI (P<0.01) and BW (P<0.05). Almost all the four yield traits (BW, LP, SI, LI) had a highly(P<0.01) positive correlation with each other, except the LP and SI, which were not correlated. The agronomic trait PH had a negative correlation with yield traits BW, LP, LI and a positive correlation with SI. Another agronomic trait fuzziness grade was highly positively correlated with all the four yield traits.

Construction of linkage maps

Out of 63,058 SNPs used in the genotyping, a total of 11,255 (17.8%) SNPs were polymorphic between the two parents. Finally, 3,129 SNPs were selected for linkage map construction after removing the bad-quality or duplicate SNPs. All the 3,129 markers were mapped to 26 linkage groups (26 chromosomes) (Fig 4.1 through Fig 4.7). This resulted in genetic map length of 4,422.44 cM with an average distance of 1.41 cM between markers (Table 4.4). Of these 3,129 SNPs, 1534 SNPs were mapped to the A sub-genome while 1,595 SNPs were mapped to the D sub-genome. The A sub-genome SNPs mapped in the genome with a total length of 2236.35 cM at an average density of 1.46 cM while 1595 D sub-genome SNPs were mapped with a total length of 2186.09 cM and an average density of 1.37 cM (Table 4.4). Genetic lengths of 26 linkage groups ranged from 103.9 cM to 252.5 cM. Number of markers mapped per chromosome range from 69 to 180 and average marker density ranging from 1.09 cM to 1.72 cM in each group (Table 4.4). Five gaps (adjacent markers distance >10 cM) were identified on 5 different linkage groups: Chr.03 (A3), Chr.08 (A09), Chr.09 (D5), Chr.26 (D6) and Chr.05 (D11) with the interval distances 11.02 cM, 11.30 cM, 14.59 cM, 10.01 cM and 10.01 cM, respectively (Table 4.4).

Of the 3,129 mapped SNPs, 175 (5.6%) SNP markers showed segregation distortion which spanned on 22 chromosomes, with the most distorted markers (34) and highest distortion rate (25.37%) on Chr.02 (A13) (Table 4.4). Seventeen segregation distortions region (SDR) were identified on 13 chromosomes, with 9 of the SDRs in A sub-genome and 8 SDRs in the D sub-genome (Table 4.4).

By comparing the genetic map with the physical map of the TM-1(*G.hirsutum*) reference genome sequence [39], the total 4422.44 cM length of the genetic map corresponded to 1911.76 Mb of the physical map, spanning 98.8% of the total genome length (Table 4.5). All chromosomes showed good coverage of the physical map with the range from 96.4% to 99.5% coveraged. Figure 4.8 were the circos plots showed strong collinearity between the genetic map and physical map, which suggested the high quality of the genetic map.

QTL analysis for cotton fiber quality, yield and agronomic traits

QTL analysis using composite interval mapping (CIM) identified a total of 106 QTLs, with 59 of QTLs for fiber quality traits, 38 for yield traits and 9 for agronomic traits. Overall the phenotypic variation explained by the QTLs ranged from 3.6% - 48.0% (Appendix B). Among the 106 QTLs, 22 were stable QTLs identified in both years, 40 QTLs were identified only in 2016 and 44 QTLs were identified only in 2017. By determining that SFC with lower value was favourable and other traits with higher value were favourable, the favourable alleles of 80 QTLs were derived from NC05AZ06 (P1) whereas 26 QTLs were contributed by NC11-2091 (P2). Of the 106 QTLs, 57 QTLs, were mapped in the A subgenome and 49 QTLs were in the D subgenome (Table 4.4). From the 57 A sub-genome QTLs, 43 QTLs with favourable alleles were from NC05AZ06 and 14 from NC11-2091 (Appendix B). In the D sub-genome 49 QTLs, 37 QTLs with favourable alleles were contributed by NC05AZ06 and 12 contributed by NC11-

2091. Overall, of the 106 mapped QTLs, 46 QTLs were major QTLs with PVE >10%. These included 29 QTLs for fiber quality traits (Table 4.6) (18 in the A sub-genome and 11 in the D sub-genome), 12 QTLs for yield traits (Table 4.7) (5 QTLs in the A sub-genome and 7 in the D sub-genome) and 5 QTLs for agronomic traits (one in A sub-genome and 4 in D su-bgenome (Table 4.8).

QTL for fiber quality traits

A total of 59 QTLs, including 15 stable QTLs, 23 QTLs in 2016 and 21 QTLs in 2017, were identified for six fiber quality traits, with the PVE ranging from 4.1% to 25.8% (Table 4.6, Appendix B). Parental accession NC05AZ06 contributed 43 QTLs' favourable alleles while NC11-2091 donated 16 of these 59 total QTLs. Sub-genome wide, from the 59 fiber quality QTLs, 31 QTLs were mapped in the A sub-genome (24 QTLs with favourable alleles from NC05AZ06 and 7 from NC11-2091) and 28 QTLs were mapped on the D sub-genome (19 QTLs with favourable alleles from NC05AZ06 and 9 from NC11-2091).

Micronaire (MIC): For fiber micronaire, seven QTLs explaining 4.1% to 25.8% of the phenotypic variance (PV) were identified, among which 5 are major QTLs (Table 4.6, Appendix B). Three major stable QTLs qMIC-CH10-1, qMIC-CH24-1, and qMIC-CH25-1 explained 16.2% ~ 16.2%, 23% ~ 25.8%, 4.1% ~ 10% of phenotypic variance, respectively. Two major QTLs qMIC-16-CH3-1 and qMIC-16-CH6-1 with the PVE 17.2% and 19.3%, respectively, were detected in the 2016 dataset. The qMIC-CH10-1 was the only QTL with favourable alleles derived from parental accession NC11-2091.

Upper half mean length (UHML): UHML length is a measure of fiber length. For UHML, 10 QTLs explaining 5.5% to 12.1% of PV were identified (Table 4.6, Appendix B). Five major QTLs, including 3 QTLs (qUHM-16-CH5-1, qUHM-16-CH7-1, qUHM-16-CH24-1) in 2016 and

2 QTLs (qUHM-17-CH7-1, qUHM-17-CH23-1) in 2017, with the PVE ranging from 10.1% to 12.1% were detected. Majority of the QTLs with favourable alleles are derived from the parent NC05AZ06. qUHM-16-CH5-1 was the only QTL with favourable alleles derived from NC11-2091.

Uniformity index (UI): Ten QTLs explaining 4.9% to 21% of PV were detected and mapped for UI in the genetic maps (Table 4.6, Appendix B). Seven QTL favourable alleles conferred by parental accession NC05AZ06. Of these six are major QTLs. These included 2 stable QTLs, qUI-CH3-1 and qUI-CH11-1 with 6% ~ 21%, 4.9% ~ 16.1%, respectively, of PVE and 4 single-year QTLs (qUI-16-CH4-1, qUI-16-CH10-1, qUI-17-CH21-1, qUI-17-CH26-1) explaining 10% - 13.1% of PV.

Fiber strength (FS): For fiber strength, 11 QTLs explaining 4.1% to 15.6% of PV, with 7 QTLs having favourable alleles conferred by NC05AZ06 were detected (Table 4.6, Appendix B). Of these four are major QTLs, including three stable QTLs (qFS-CH2-1, qFS-CH19-1, qFS-CH25-1) with PVE of 7.3% ~ 11.9%, 5.2% ~ 11.8%, 11.5% ~ 15.6%, respectively, and one QTL (qFS-17-CH14-1), detected only in 2017, explaining 11.9% of PV.

Elongation (FE): Nine QTLs explaining 5.7% to 13.5% of PV were mapped in the linkage maps (Table 4.6, Appendix B). Of these 9 QTLs for elongation, four are major QTLs which include three stable QTLs (qFE-CH4-1, qFE-CH8-1, qFE-CH19-1) with 7.1% ~ 13.5%, 7.5% ~ 12.3%, 8.1% ~ 12.4% of PVE and one QTL (qFE-17-CH23-1) found only in 2017, explained 11.2% of PV. Further, five of these mapped QTLs had favorable alleles from NC11-2091 for fiber elongation.

Short fiber content (SFC): For short fiber content, 12 QTLs explaining 4.9% to 20.6% of PV were identified (Table 4.6, Appendix B), with 5 major QTLs included. One major QTL (qSFC-CH3-1) was detected in both years with 7.9% ~ 18.4% of PVE, which was contributed by NC05AZ06. Another 4 major QTLs (qSFC-17-CH2-1, qSFC-16-CH4-1, qSFC-17-CH5-1, qSFC-17-CH18-1) with the PVE ranging from 12.4% to 20.6% were detected in single year environment. Most of the QTL favorable alleles were derived from NC05AZ06, except qSFC-CH3-1 and qSFC-17-CH26-1.

QTL for yield traits

A total of 38 QTLs, including 5 stable QTLs, 16 QTLs in 2016 and 17 QTLs in 2017, were identified for yield traits (BW, LP, SI, LI), with the PVE ranging from 4.2% to 30.4% (Table 4.7, Appendix B). NC05AZ06 contributed 35 QTLs with favourable alleles while NC11-2091 only donated 3 of these 38 total QTLs. Further, of the 38 yield QTLs, 22 QTLs were mapped in the A sub-genome, including 19 QTLs with favourable alleles contributed by NC05AZ06 and 3 contributed by NC11-2091 and 16 QTLs were mapped in the D sub-genome, of which the favourable alleles were all contributed by NC05AZ06.

Boll weigh (BW)t: For boll weight, 11 QTLs explaining 5% to 14% of PV were identified and all favourable alleles of the QTLs were derived from NC05AZ06 (Table 4.7, Appendix B). Two major QTLs (qBW-16-CH4-1, qBW-16-CH22-1) with 14%, 12.4% of PVE were detected only in year 2016 and another major QTL qBW-17-CH4-1 with 12.1% of PVE was identified in year 2017.

Lint percentage (LP): Eight QTLs explaining 5.9% to 17.7% of PV were identified for lint percentage which included 3 major QTLs (Table 4.7, Appendix B). Two major and stable QTLs qLP-CH24-1 and qLP-CH25-1 explained 8.8% ~ 16.6%, 5.9% ~ 17.7% of PV, respectively for

LP. Another major QTL (qLP-17-CH14-1) with 15.2% of PVE, were detected in 2017 dataset. All QTLs favorable alleles were derived from NC05AZ06.

Seed index (SI): For seed index, 9 QTLs explaining 5.8% to 30.4% of PV were detected. Among these 6 QTLs with favorable alleles were derived from NC05AZ06 (Table 4.7, Appendix B). Two major and stable QTLs qSI-CH12-1 and qSI-CH15-1 with 27.1% ~ 30.4%, 10.6% ~ 17.5% of PVE, respectively, were identified in both environments.

Lint index (LI): Ten QTLs explaining 4.2% to 21.1% of PV for lint index were identified and all their favourable alleles were contributed by NC05AZ06 (Table 4.7, Appendix B). Four major QTLs, including two QTLs detected only in the year 2016 (qLI-16-CH2-1, qLI-16-CH12-1) and other two QTLs detected only in year 2017 (qLI-17-CH22-1, qLI-17-CH24-1), explained 10.8%, 13.5%, 21.1%, 12.6% of PV, respectively.

QTL for agronomic traits

A total of 9 QTLs, including 2 stable QTLs, 1 QTLs in 2016 and 6 QTLs in 2017, were identified for agronomic traits (plant height and fuzziness grade), with the PVE from 3.6% to 48% (Table 4.8, Appendix B). NC05AZ06 contributed favorable alleles of 2 QTLs whereas NC11-2091 donated 7 of the 9 total QTLs (4 QTLs on the A sub-genome and 5 QTLs on the D sub-genome).

Plant height (PH): For plant height, 5 QTLs explaining 6.5% to 15.8% of PV were identified in year 2017 (Table 4.8, Appendix B). and all these QTLs with positive additive effect for plant height were derived from NC11-2091. Three major QTLs for plant height (qPH-17-CH8-1, qPH-17-CH9-1, qPH-17-CH19-1) explained 10.3%, 15.8% and 10.4% of PV, respectively.

Seed fuzziness (FG): For seed fuzziness grade, 4 QTLs explaining 3.6% to 48% of PV were identified (Table 4.8, Appendix B), of which 2 are major stable QTLs. The major stable QTL (qFG-CH22-1) was the only QTL with positive additive effect for seed fuzziness contributed by NC05AZ06, with 39.2% ~ 48% of PVE. Another major stable QTL (qFG-CH25-1) explained 3.6% ~ 19% of PV for seed fuzziness.

QTL clusters & meta-analysis

A QTL cluster was a short region (<30 cM) on the linkage map containing multiple QTLs [14]. In this population, 21 QTL Clusters (Table 4.9, Table 4.10) were identified in 16 different chromosomes (Chr3, Chr4, Chr7, Chr8, Chr9, Chr10, Chr11, Chr12, Chr15, Chr16, Chr19, Chr21, Chr22, Chr23, Chr24 and Chr25) after all the QTLs were mapped (Fig 4.1-Fig 4.7). 12 QTL clusters were detected in A sub-genome and 9 clusters were detected in D sub-genome, the names of the clusters depended on the trait categories of QTLs they contained. For example, Q-* meant a cluster contained only fiber quality traits QTLs; QY-* meant a cluster contained both fiber quality and yield traits QTLs; QYA-* meant a cluster contained fiber quality, yield and agronomic traits QTLs; etc.

Seven QTL Clusters (Q-1 to Q-7) contained multiple fiber quality traits QTLs. Cluster Q-1, Q-3, Q-4, Q-5, Q-6 were identified with QTLs from SFC and UI (Table 4.9). In each of these 5 Clusters, the favorable alleles of SFC and UI QTLs were contributed by same parents with different signs of additive effects.

For yield traits, four QTL Clusters (Y-1 to Y-4) were identified (Table 4.9). The favorable alleles for most of the yield QTLs in these Clusters were derived from NC05AZ06.

Other 10 Clusters contained multiple QTLs from different trait categories (Table 4.10). QYA-1 and QYA-2 were two clusters carrying multiple QTLs from all 3 trait categories. QYA-1

with a region in Chr.8(A9) from 56.8 cM to 78.78 cM, contained 4 QTLs for FG, FE, LP and PH. QYA-2 with a region in Chr.25(D12) from 97.24 cM to 108.63 cM, carried 4 QTLs for FS, MIC, LP and FG.

A total of 2884 cotton QTLs for 11 traits: MIC(442), UHM(524), UI(289), FS(470), FE(287), SFC(58), BW(176), LP(327), LI(42), SI(147), PH(122), which were collected by the CottonQTLdb [14, 15, 40] in different interspecific or intraspecific populations from 156 previous publications (http://www2.cottonqtldb.org:8081/references), were used for meta-analysis in recent study (Appendix A).

Out of 102 QTLs for 11 traits in this population, 74 QTLs were found to share the similar genetic and physical positions (genetic distance window of <20 cM) with previous reported QTLs, including 39 QTLs in A subgenome and 35 QTLs in D subgenome. All these 74 shared QTLs were separated in 11 different traits: FS(11), UI(10), UHM(7), MIC(7), FE(7), SFC(7), BW(7), SI(6), LP(5), LI(4) and PH(3), including 33 major QTLs. 13 of them were stable QTLs (qFE-CH8-1, qFS-CH2-1, qFS-CH19-1, qFS-CH25-1, qLP-CH24-1, qMIC-CH10-1, qMIC-CH24-1, qMIC-CH25-1, qSFC-CH3-1, qSI-CH12-1, qSI-CH15-1, qUI-CH3-1, qUI-CH11-1). 28 QTLs were unique QTLs with 17 QTLs in A subgenome and 11 QTLs in D subgenome, including 5 for SFC, 3 for UHM, 2 for FE, 6 for LI, 4 for BW, 3 for SI, 3 for LP, 2 for PH. Out of these 28 unique QTLs, 11 were major QTLs (qBW-17-CH4-1, qBW-16-CH4-1, qFE-CH4-1, qFE-CH19-1, qLI-17-CH24-1, qLI-16-CH12-1, qLP-CH25-1, qPH-17-CH9-1, qSFC-17-CH2-1, qUHM-17-CH23-1, qUHM-16-CH24-1). Three of them were stable QTLs: qFE-CH4-1, qFE-CH19-1, qLP-CH25-1.

Tables

Table 4. 1 Trait Performance of NC05AZ06 × NC11-2091 RILs population and their parents in year 2016 and 2017

Type	Trait[a]	Year	Parents			RILs			
			NC05AZ 06(P1)	NC11-2091(P2)	P1-P2	Min	Max	Mean	SD
Fiber Quality	MIC	16	4.90	4.19	0.71*	3.69	7.02	4.81	0.55
		17	4.78	4.23	0.55**	3.65	6.20	4.79	0.46
	UH	16	1.12	0.82	0.3**	0.73	1.11	0.92	0.08
	M	17	1.14	0.85	0.3**	0.74	1.13	0.93	0.1
	UI	16	82.25	76.43	5.82*	73.93	83.95	79.36	1.6
		17	83.20	79.73	3.48**	76.95	84.40	81.62	1.5
	FS	16	27.64	25.56	2.08*	22.31	32.57	26.82	2.13
		17	27.45	25.55	1.9*	20.90	30.55	25.49	2.23
	FE	16	6.92	6.74	0.19	3.85	12.54	7.26	1.28
		17	8.00	8.60	-0.6*	4.60	12.50	8.81	1.41
	SFC	16	8.32	12.92	-4.61*	7.16	17.90	10.63	1.67
		17	8.35	10.55	-2.2**	6.85	17.10	9.05	1.61
Yield	BW	16	4.92	1.59	3.33**	1.46	4.23	2.77	0.6
		17	6.23	1.93	4.3**	2.03	5.61	3.46	0.75
	LP	16	40.8	17.2	23.6**	16.39	39.22	28.63	4.98
		17	40.3	15.5	24.8**	19.00	38.50	28.21	4.44
	SI	16	10.19	8.18	2.01**	7.13	11.70	9.13	0.82
		17	11.15	9.91	1.24**	8.14	12.90	10.54	0.93
	LI	16	7.00	1.70	5.3**	1.86	5.62	3.69	1.06
		17	7.48	1.79	5.69**	2.31	6.58	4.18	1.18
Agronomic	PH	17	44.3	65.6	-21.3**	32.15	66.25	48.04	4.95
	FG	16	100	0	100**	0	100	66.7	38.2
		17	100	0	100**	0	100	43.8	37.9

*: 0.05 significant level; **: 0.01 significant level

[a] MIC micronaire; UHM upper half mean length; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzziness grade of seed

Table 4. 2 The broad-sense heritability of traits in NC05AZ06 × NC11-2091 RILs Population across two years

	MIC[a]	UHM	UI	FS	FE	SFC	BW	LP	SI	LI	PH[b]	FG
V_g	0.230	0.0077	2.507	4.593	1.675	3.004	0.430	21.032	0.797	1.197	24.517	15.60
V_p	0.258	0.0084	3.003	5.255	1.811	3.534	0.522	21.958	0.931	1.257	43.422	17.26
H^2	89%	92%	83%	87%	92%	85%	82%	96%	86%	95%	56%	90%

The broad-sense heritability (H^2) = genetic variance (V_g)/phenotypic variance (V_p)
[a] MIC micronaire; UHM upper half mean length; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzziness grade of seed
[b] PH with only year 2017 data used replication as environment effect

Table 4. 3 Correlation analysis between the traits in NC05AZ06 × NC11-2091 RILs Population across two years

	MIC[a]	UHM	UI	FS	FE	SFC	BW	LP	SI	LI	PH[b]
UHM	-0.36**										
UI	-0.28**	0.82**									
FS	0.18	-0.12	-0.08								
FE	0.3**	-0.62**	-0.51**	-0.2*							
SFC	0.24**	-0.79**	-0.93**	0.1	0.46**						
BW	0.12	0.29**	0.18	0.21*	-0.23*	-0.13					
LP	0.28**	0.25**	0.11	-0.27**	0.09	-0.09	0.46**				
SI	-0.06	0.12	0.19*	0.32**	-0.17	-0.15	0.37**	-0.04			
LI	0.21*	0.3**	0.18	-0.1	0	-0.14	0.61**	0.89**	0.4**		
PH	-0.06	-0.13	-0.12	0.21*	-0.11	0.11	-0.25*	-0.36**	0.21*	-0.22*	
FG	-0.14	0.27**	0.2*	-0.09	-0.11	-0.18	0.46**	0.31**	0.32**	0.42**	-0.15

*: 0.05 significant level; **:0.01 significant level
[a] MIC micronaire; UHM upper half mean length; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzziness grade of seed
[b] PH had only year 2017 data

Table 4. 4 Linkage Map information of NC05AZ06 × NC11-2091 RILs Population

Linkage Group	physical Map	Number of Markers	Length	Average Density	Gaps[a] (>10cM)	Number of QTLs	Distorted Markers[b]	No. SDR[c]
Chr16	A1	110	149.9	1.36	0	4	3	1
Chr23	A2	83	142.82	1.72	0	4	2	0
Chr03	A3	139	157.56	1.13	1	3	13	1
Chr13	A4	69	103.9	1.51	0	1	0	0
Chr10	A5	180	252.5	1.4	0	4	5	0
Chr18	A6	100	151.52	1.52	0	2	1	0
Chr07	A7	115	168.56	1.47	0	9	6	0
Chr14	A8	130	200.94	1.55	0	5	3	0
Chr08	A9	109	172.61	1.58	1	5	6	1
Chr11	A10	104	151.72	1.46	0	4	12	1
Chr04	A11	149	227.53	1.53	0	10	4	0
Chr15	A12	112	179.68	1.6	0	3	13	2
Chr02	A13	134	177.11	1.32	0	3	34	3
A-Subgenome		1534	2236.35	1.46	2	57	102	9
Chr19	D1	111	138.85	1.25	0	3	0	0
Chr21	D2	147	173.6	1.18	0	5	0	0
Chr24	D3	93	150.29	1.62	0	7	4	0
Chr20	D4	99	139.11	1.41	0	2	9	1
Chr09	D5	179	243.08	1.36	1	4	19	2
Chr26	D6	82	133.64	1.63	1	3	5	1
Chr06	D7	127	170.51	1.34	0	1	7	1
Chr17	D8	138	198.21	1.44	0	2	9	1
Chr01	D9	113	150.31	1.33	0	0	0	0
Chr12	D10	123	173.79	1.41	0	4	12	1
Chr05	D11	123	199.99	1.63	1	4	1	0
Chr25	D12	136	179.68	1.32	0	8	3	0
Chr22	D13	124	135.03	1.09	0	6	4	1
D-Subgenome		1595	2186.09	1.37	3	49	73	8
Total		3129	4422.44	1.41	5	106	175	17

[a]Gap: Distance between two adjacent markers > 10 cM
[b]Distorted Markers: Markers showing segregation distortion (chi-square test P<0.05)
[c]SDR segregation distortion region

Table 4. 5 Length of genome spanned by the linkage map for NC05AZ06 × NC11-2091 RILs population

Linkage Group	physical Map	Start site	End site	Genome spanned length	Reference genome length	Genome coverage ratio
Chr16	A1	567728	99870991	99303263	99884700	99.4%
Chr23	A2	518521	83029666	82511145	83447906	98.9%
Chr03	A3	536767	100143431	99606664	100263045	99.3%
Chr13	A4	237387	62124638	61887251	62913772	98.4%
Chr10	A5	1544314	91885184	90340870	92047023	98.1%
Chr18	A6	332742	102528427	102195685	103170444	99.1%
Chr07	A7	183539	77664102	77480563	78251018	99.0%
Chr14	A8	253013	102980194	102727181	103626341	99.1%
Chr08	A9	1897239	74776037	72878798	74999931	97.2%
Chr11	A10	810285	100817936	100007651	100866604	99.1%
Chr04	A11	725030	93108419	92383389	93316192	99.0%
Chr15	A12	230674	87226684	86996010	87484866	99.4%
Chr02	A13	212014	79758080	79546066	79961121	99.5%
A-Subgenome				1147864536	1160232963	98.9%
Chr19	D1	449547	61161478	60711931	61456009	98.8%
Chr21	D2	288235	66926440	66638205	67284553	99.0%
Chr24	D3	414604	46229908	45815304	46690656	98.1%
Chr20	D4	333688	51295622	50961934	51454130	99.0%
Chr09	D5	263755	61711575	61447820	61933047	99.2%
Chr26	D6	174860	64004405	63829545	64294643	99.3%
Chr06	D7	593089	54265422	53672333	55312611	97.0%
Chr17	D8	297535	65311191	65013656	65894135	98.7%
Chr01	D9	1280925	50444750	49163825	50995436	96.4%
Chr12	D10	831033	63348128	62517095	63374666	98.6%
Chr05	D11	631689	65974293	65342604	66087774	98.9%
Chr25	D12	561178	59087499	58526321	59109837	99.0%
Chr22	D13	81171	60339884	60258713	60534298	99.5%
D-Subgenome				763899286	774421795	98.6%
Total				1911763822	1934654758	98.8%

Table 4. 6 Major[a] QTLs of fiber quality traits identified in NC05AZ06 × NC11-2091 RILs population

Trait[b]	QTL	Year	Range	Peak	LOD	PVE[c]	AE[d]
MIC	qMIC-CH10-1*	16	233.43-252.01	244.08	11.4	16.2	-0.25
		17	233.43-252.01	243.58	9.1	16.2	-0.207
	qMIC-CH24-1*	16	71.47-87.7	81.76	16.3	25.8	0.335
		17	68.4-87.7	78.81	12.5	23	0.256
	qMIC-CH25-1*	16	102.67-109.11	104.67	3.3	4.1	0.128
		17	92.32-112.58	102.67	6.1	10	0.167
	qMIC-16-CH3-1	16	78.61-95.33	88.82	11.7	17.2	0.344
	qMIC-16-CH6-1	16	0-15.09	3	7	19.3	0.274
UHM	qUHM-16-CH5-1	16	0-4.58	0	6.5	11.2	-0.028
	qUHM-16-CH7-1	16	43.66-65.19	55.42	6.9	12.1	0.029
	qUHM-17-CH7-1	17	17.16-35.23	22.56	6.1	10.1	0.027
	qUHM-17-CH23-1	17	132.16-142.33	142.33	6.6	10.2	0.028
	qUHM-16-CH24-1	16	123.73-142.8	134.44	6.7	11.9	0.028
UI	qUI-CH3-1*	16	46.82-68.03	58.14	10.4	21	-0.921
		17	46.82-68.03	55.53	3.5	6	-0.386
	qUI-CH11-1*	16	30.34-52.37	42.28	8.4	16.1	0.9
		17	46-50.81	46	2.8	4.9	0.362
	qUI-16-CH4-1	16	207.88-227.03	220.88	7.4	13	0.72
	qUI-16-CH10-1	16	11.59-19.17	17.67	7.2	12.7	0.732
	qUI-17-CH21-1	17	6.05-24.36	17.12	6.1	10	0.488
	qUI-17-CH26-1	17	28.88-46.13	39.93	7	13.1	-0.596
FS	qFS-CH2-1*	16	107.73-118.93	112.84	4.9	7.3	0.639
		17	101.54-118.93	111.34	7.2	11.9	0.767
	qFS-CH19-1*	16	121.23-124.72	124.72	3.6	5.2	-0.519
		17	112.85-133.39	121.23	7.3	11.8	-0.726
	qFS-CH25-1*	16	98.22-115.15	108.63	9.4	15.6	-0.936
		17	101.68-121.15	110.59	7.3	11.5	-0.701
	qFS-17-CH14-1	17	8.8-28.94	21.46	7.5	11.9	0.736
FE	qFE-CH4-1*	16	156.02-172.58	167.72	4.7	7.1	-0.385
		17	156.02-181.46	172.58	8.7	13.5	-0.562
	qFE-CH8-1*	16	58.29-74.89	68.12	7.7	12.3	0.507
		17	65.15-85.78	71.03	5.3	7.5	0.437
	qFE-CH19-1*	16	124.72-137.87	136.4	7.7	12.4	0.499
		17	126.74-137.87	136.4	5.6	8.1	0.428
	qFE-17-CH23-1	17	132.16-142.33	142.33	7	11.2	-0.503
SFC	qSFC-CH3-1*	16	44.13-59.15	58.64	3.9	7.9	0.664
		17	46.82	46.82	2.5	18.4	0.807
	qSFC-17-CH2-1	17	36.99	36.99	3.6	20.6	-0.837
	qSFC-16-CH4-1	16	217.44-227.03	226.53	7.4	12.4	-0.811
	qSFC-17-CH5-1	17	176.37	176.37	3.1	20.1	-0.828
	qSFC-17-CH18-1	17	57.29	57.29	3.2	20.4	-0.827

Table 4. 6 (continued).

*QTLs identified in both years [a]Major QTLs: The QTLs with at least one year's PVE>10%
[b]MIC micronaire; UHM upper half mean; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content
[c]PVE phenotypic variation explained [d]AE additive effect

Table 4. 7 Major[a] QTLs of the yield traits identified in NC05AZ06 × NC11-2091 RILs population

Trait[b]	QTL	Year	Range	Peak	LOD	PVE[c]	AE[d]
BW	qBW-16-CH4-1	16	204.27-222.9	214.51	9.3	14	0.217
	qBW-17-CH4-1	17	156.02-181.46	171.61	8.9	12.1	0.251
	qBW-16-CH22-1	16	91.66-106.73	91.66	7.8	12.4	0.212
LP	qLP-CH24-1*	16	58-78.32	66.92	11.2	16.6	1.885
		17	58-76.37	65.94	7.3	8.8	1.208
	qLP-CH25-1*	16	86.34-105.64	97.24	4.5	5.9	1.077
		17	97.24-115.15	107.65	12.5	17.7	1.666
	qLP-17-CH14-1	17	61.88-75.77	68.84	11.1	15.2	1.561
SI	qSI-CH12-1*	16	80.19-98.29	91.65	14.8	27.1	0.712
		17	80.19-100.77	90.17	14.8	30.4	0.669
	qSI-CH15-1*	16	19.71-38.68	33.54	10.7	17.5	-0.432
		17	19.71-47.61	38.68	6.7	10.6	-0.332
LI	qLI-16-CH2-1	16	4.21-17.19	11.03	7	10.8	0.297
	qLI-16-CH12-1	16	79.71-99.78	90.17	8.9	13.5	0.324
	qLI-17-CH22-1	17	84.67-103.29	94.72	13.7	21.1	0.437
	qLI-17-CH24-1	17	55.56-76.86	66.92	9.2	12.6	0.337

*QTLs identified in both years [a]Major QTLs: The QTLs with at least one year's PVE>10%
[b]BW boll weight; LP lint percentage; SI seed index; LI lint index
[c]PVE phenotypic variation explained [d]AE additive effect

Table 4. 8 Major[a] QTLs of the agronomic traits identified in NC05AZ06 × NC11-2091 RILs population

Trait[b]	QTL	Year	Range	Peak	LOD	PVE[c]	AE[d]
PH	qPH-17-CH8-1	17	52.05-72.01	56.8	5.5	10.3	-2.557
	qPH-17-CH9-1	17	173.76-193.74	185.58	7.9	15.8	-2.625
	qPH-17-CH19-1	17	114.85-132.4	124.72	5.8	10.4	-2.204
FG	qFG-CH22-1*	16	124.58-134.55	134.55	23.5	39.2	24.361
		17	124.09-134.55	134.55	28.4	48	29.28
	qFG-CH25-1*	16	88.36-108.63	98.71	14.8	19	-15.112
		17	98.71-104.67	101.19	4	3.6	-7.972

*QTLs identified in both years [a]Major QTLs: The QTLs with at least one year's PVE>10%
[b]PH plant height; FG fuzziness grade of seed
[c]PVE phenotypic variation explained [d]AE additive effect

Table 4. 9 Clusters containing multiple QTLs from same trait category (fiber quality or yield) in NC05AZ06 × NC11-2091 RILs population

Cluster[a]	Chr.	Region	QTLs	Trait	Source of favorable alleles	PVE
Q-1	CH3	46.82-59.15	qSFC-CH3-1	SFC	NC11-2091	7.9
			qUI-CH3-1	UI	NC11-2091	21
Q-2	CH7	17.16-27.41	qUHM-17-CH7-1	UHM	NC05AZ06	10.1
			qFE-16-CH7-1	FE	NC11-2091	6.7
Q-3	CH10	12.09-19.17	qUI-16-CH10-1	UI	NC05AZ06	12.7
			qSFC-16-CH10-1	SFC	NC05AZ06	6
Q-4	CH11	30.34-52.37	qUI-CH11-1	UI	NC05AZ06	16.1
			qSFC-16-CH11-1	SFC	NC05AZ06	9.5
Q-5	CH21	11.56-18.62	qUI-17-CH21-1	UI	NC05AZ06	10
			qSFC-CH21-1	SFC	NC05AZ06	4.9
Q-6	CH22	68.73-79.76	qUI-17-CH22-1	UI	NC05AZ06	9.5
			qSFC-17-CH22-1	SFC	NC05AZ06	6.9
Q-7	CH23	133.16-142.33	qUHM-17-CH23-1	UHM	NC05AZ06	10.2
			qFE-17-CH23-1	FE	NC11-2091	11.2
			qSFC-16-CH23-1	SFC	NC05AZ06	6.4
Y-1	CH12	80.19-99.78	qLI-16-CH12-1	LI	NC05AZ06	13.5
			qBW-17-CH12-1	BW	NC05AZ06	8.6
			qSI-CH12-1	SI	NC05AZ06	27.1
Y-2	CH15	19.71-38.68	qLP-17-CH15-1	LP	NC05AZ06	6.1
			qSI-CH15-1	SI	NC11-2091	17.5
Y-3	CH16	44.41-53.76	qBW-16-CH16-1	BW	NC05AZ06	9.9
			qLI-16-CH16-1	LI	NC05AZ06	6.8
			qSI-16-CH16-1	SI	NC05AZ06	6
Y-4	CH22	91.66-103.29	qLI-17-CH22-1	LI	NC05AZ06	21.1
			qBW-16-CH22-1	BW	NC05AZ06	12.4

[a]Q: Cluster containing multiple QTLs for fiber quality traits
Y: Cluster containing multiple QTLs for yield traits

Table 4. 10 Clusters containing multiple QTLs from different trait categories (fiber quality, yield or agronomic) in NC05AZ06 × NC11-2091 RILs population

Cluster[a]	Chr.	Region	QTLs	Trait	Source of favorable alleles	PVE
QY-1	CH4	156.02-181.46	qFE-CH4-1	FE	NC11-2091	7.1
			qBW-17-CH4-1	BW	NC05AZ06	12.1
QY-2	CH4	214.51-226.53	qBW-16-CH4-1	BW	NC05AZ06	14
			qUI-16-CH4-1	UI	NC05AZ06	13
			qSFC-16-CH4-1	SFC	NC05AZ06	12.4
QY-3	CH7	126.63-139.78	qLI-CH7-1	LI	NC05AZ06	4.2
			qMIC-16-CH7-1	MIC	NC05AZ06	6.4
			qSI-17-CH7-1	SI	NC05AZ06	7
QY-4	CH24	60.96-81.76	qUHM-17-CH24-1	UHM	NC05AZ06	7.5
			qLI-17-CH24-1	LI	NC05AZ06	12.6
			qLP-CH24-1	LP	NC05AZ06	16.6
			qMIC-CH24-1	MIC	NC05AZ06	25.8
QY-5	CH24	119.17-141.3	qFS-17-CH24-1	FS	NC11-2091	5.8
			qUHM-16-CH24-1	UHM	NC05AZ06	11.9
			qLP-16-CH24-2	LP	NC05AZ06	6.7
QA-1	CH7	48.29-61.65	qUHM-16-CH7-1	UHM	NC05AZ06	12.1
			qUI-17-CH7-1	UI	NC05AZ06	8.9
			qPH-17-CH7-1	PH	NC11-2091	6.5
QA-2	CH19	124.72-136.4	qFS-CH19-1	FS	NC11-2091	5.2
			qPH-17-CH19-1	PH	NC11-2091	10.4
			qFE-CH19-1	FE	NC05AZ06	12.4
YA-1	CH9	183.45-193.74	qPH-17-CH9-1	PH	NC11-2091	15.8
			qLI-16-CH9-1	LI	NC05AZ06	8.9
QYA-1	CH8	56.8-72.49	qPH-17-CH8-1	PH	NC11-2091	10.3
			qFE-CH8-1	FE	NC05AZ06	12.3
			qLP-16-CH8-1	LP	NC05AZ06	7.6
			qFG-17-CH8-1	FG	NC11-2091	8.1
QYA-2	CH25	97.24-108.63	qLP-CH25-1	LP	NC05AZ06	5.9
			qFG-CH25-1	FG	NC11-2091	19
			qMIC-CH25-1	MIC	NC05AZ06	4.1
			qFS-CH25-1	FS	NC11-2091	15.6

[a]QY: Cluster containing multiple QTLs for fiber quality and yield traits
QA: Cluster containing multiple QTLs for fiber quality and agronomic traits
YA: Cluster containing multiple QTLs for yield and agronomic traits
QYA: Cluster containing multiple QTLs for fiber quality, yield and agronomic traits

Figures

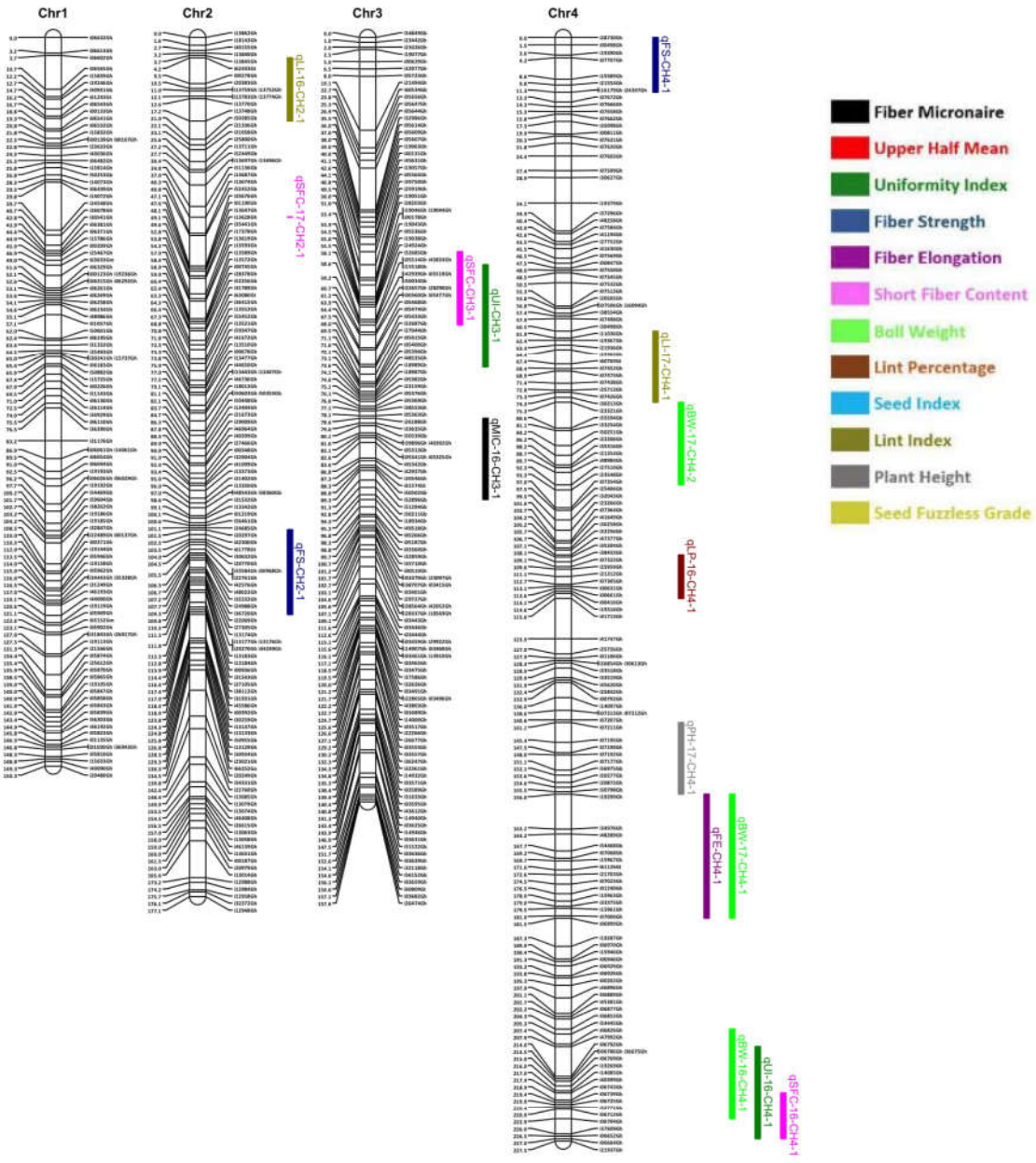

Figure 4. 1 Linkage maps of chromosomes Chr1(D9), Chr2(A13), Chr3(A3), Chr4(A11) along with detected QTLs for NC05AZ06 × NC11-2091 RILs population.

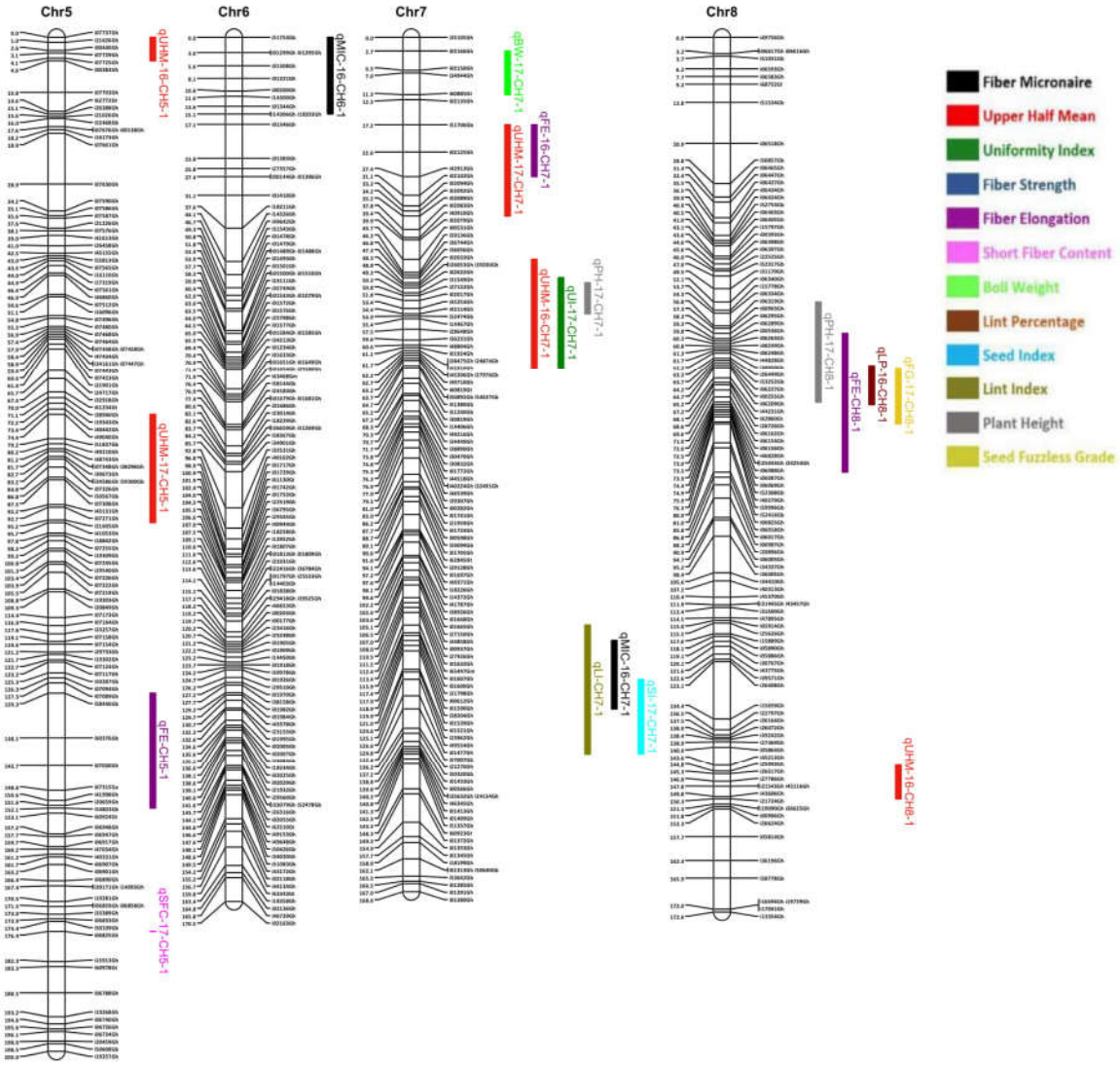

Figure 4. 2 Linkage maps of chromosomes Chr5(D11), Chr6(D7), Chr7(A7), Chr8(A9), along with detected QTLs for NC05AZ06 × NC11-2091 RILs population.

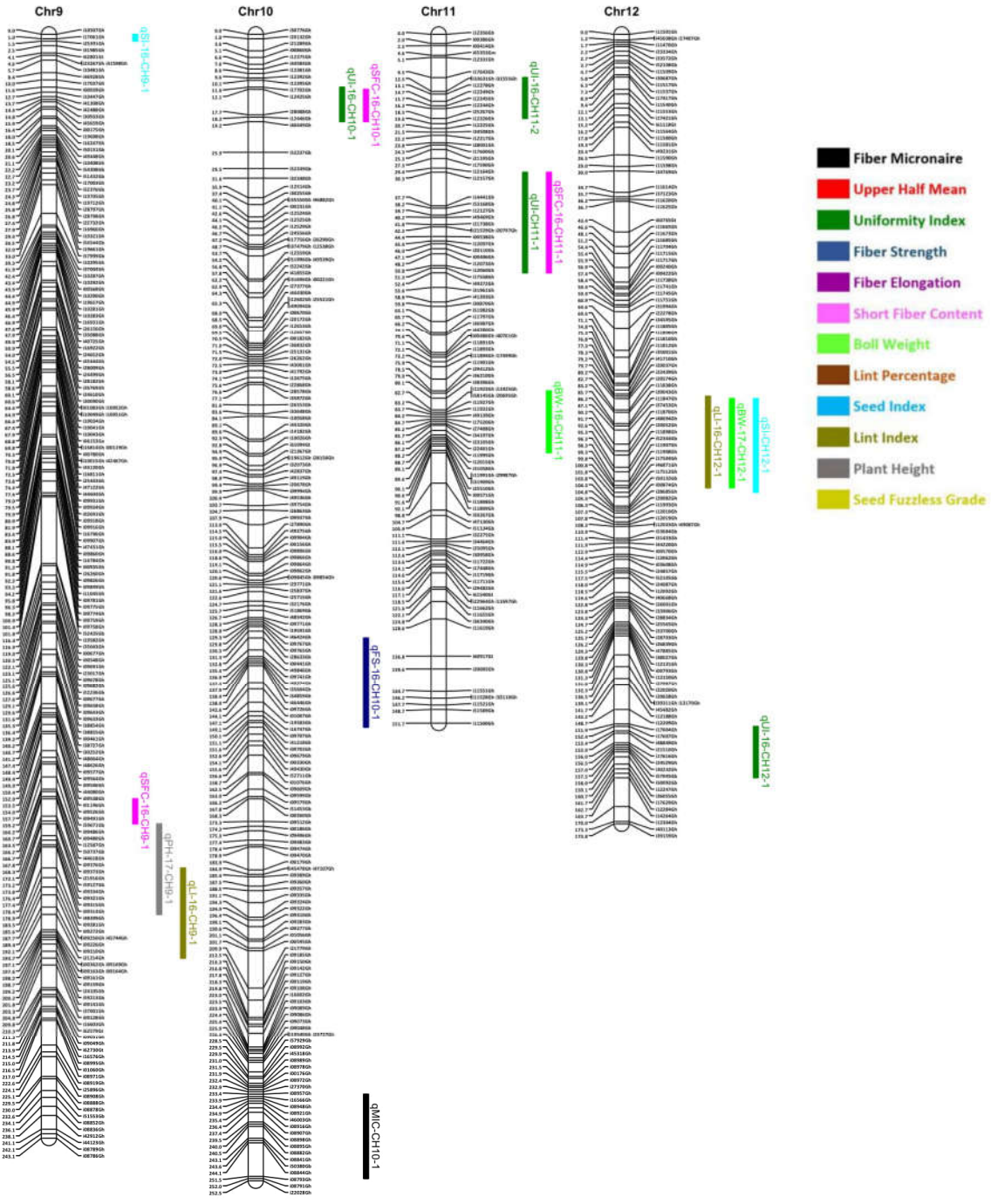

Figure 4. 3 Linkage maps of chromosomes Chr9(D5), Chr10(A5), Chr11(A10), Chr12(D10), along with detected QTLs for NC05AZ06 × NC11-2091 RILs population.

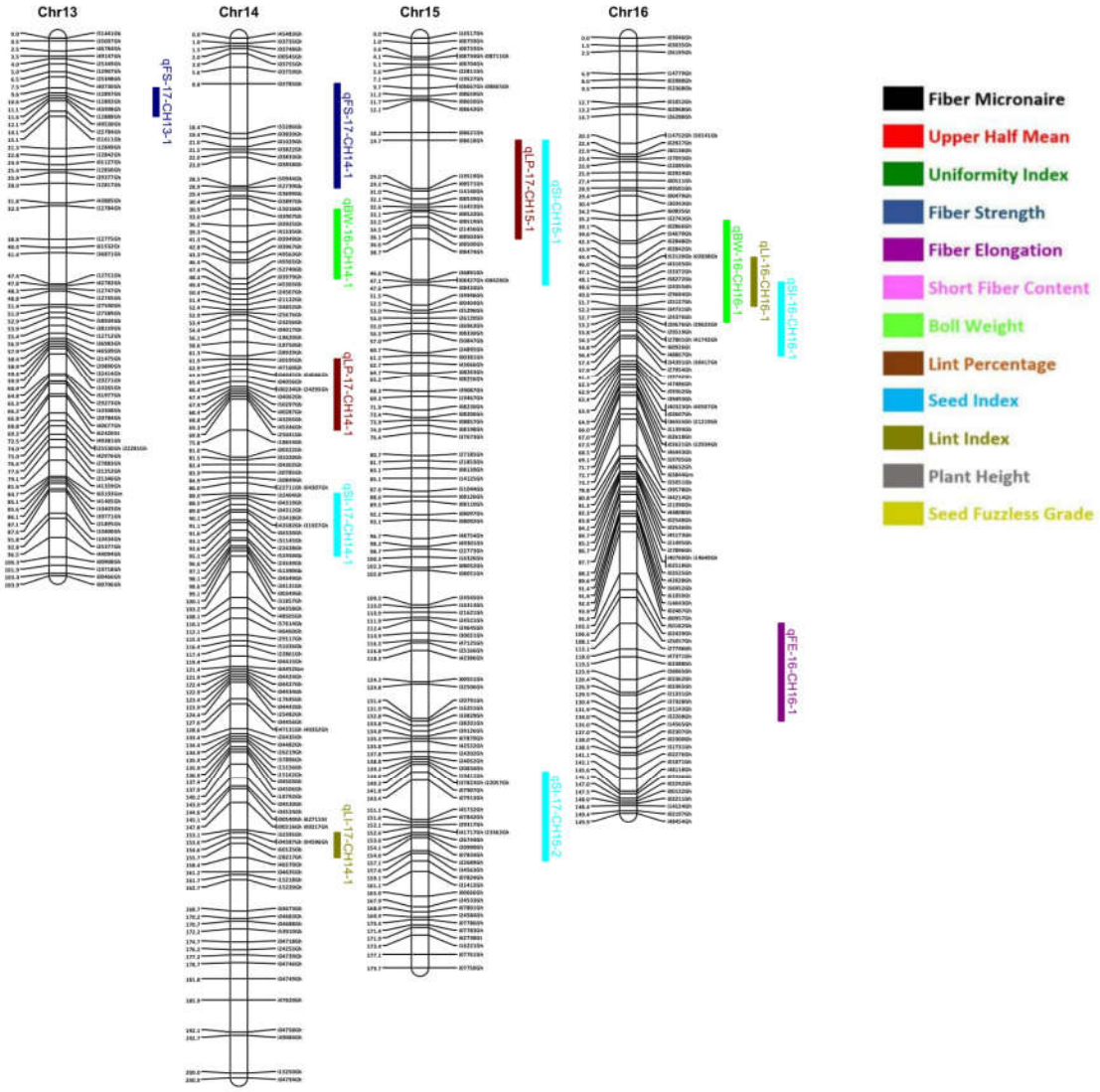

Figure. 4 Linkage maps of chromosomes Chr13(A4), Chr14(A8), Chr15(A12), Chr16(A1), along with detected QTLs for NC05AZ06 × NC11-2091 RILs population.

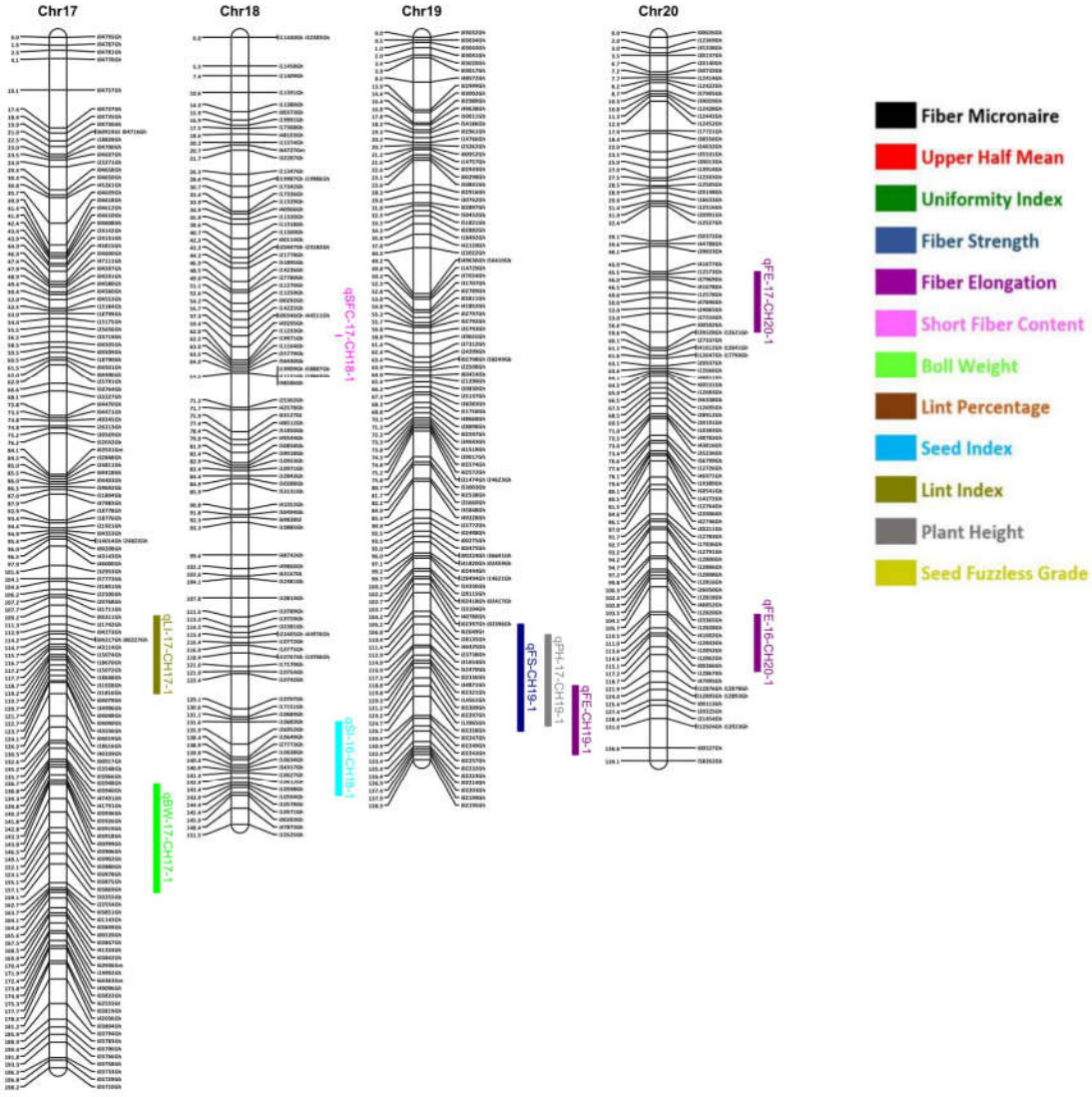

Figure 4. 5 Linkage maps of chromosomes Chr17(D8), Chr18(A6), Chr19(D1), Chr20(D4), along with detected QTLs for NC05AZ06 × NC11-2091 RILs population.

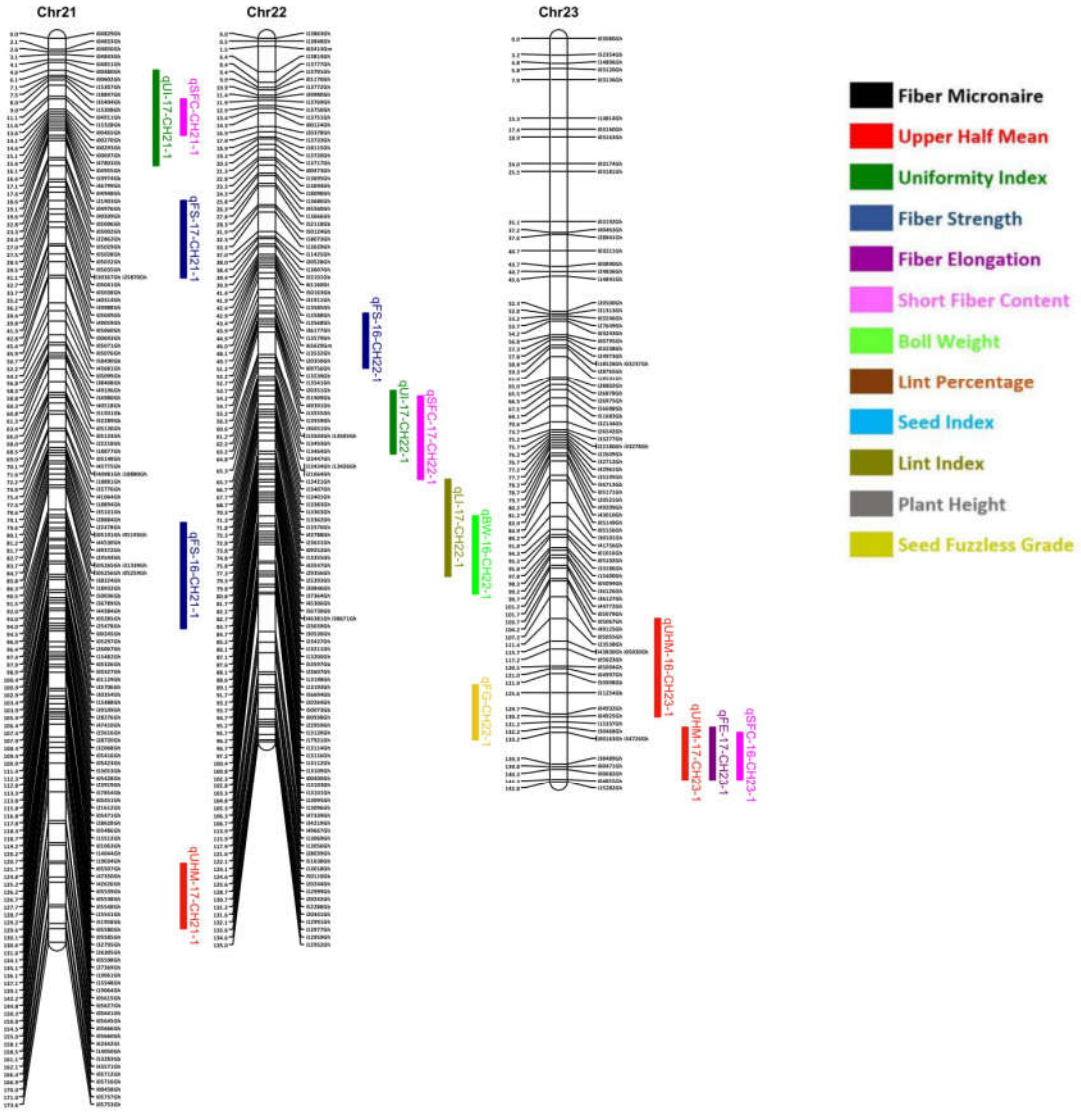

Figure 4. 6 Linkage maps of chromosomes Chr21(D2), Chr22(D13), Chr23(A2), along with detected QTLs for NC05AZ06 × NC11-2091 RILs population.

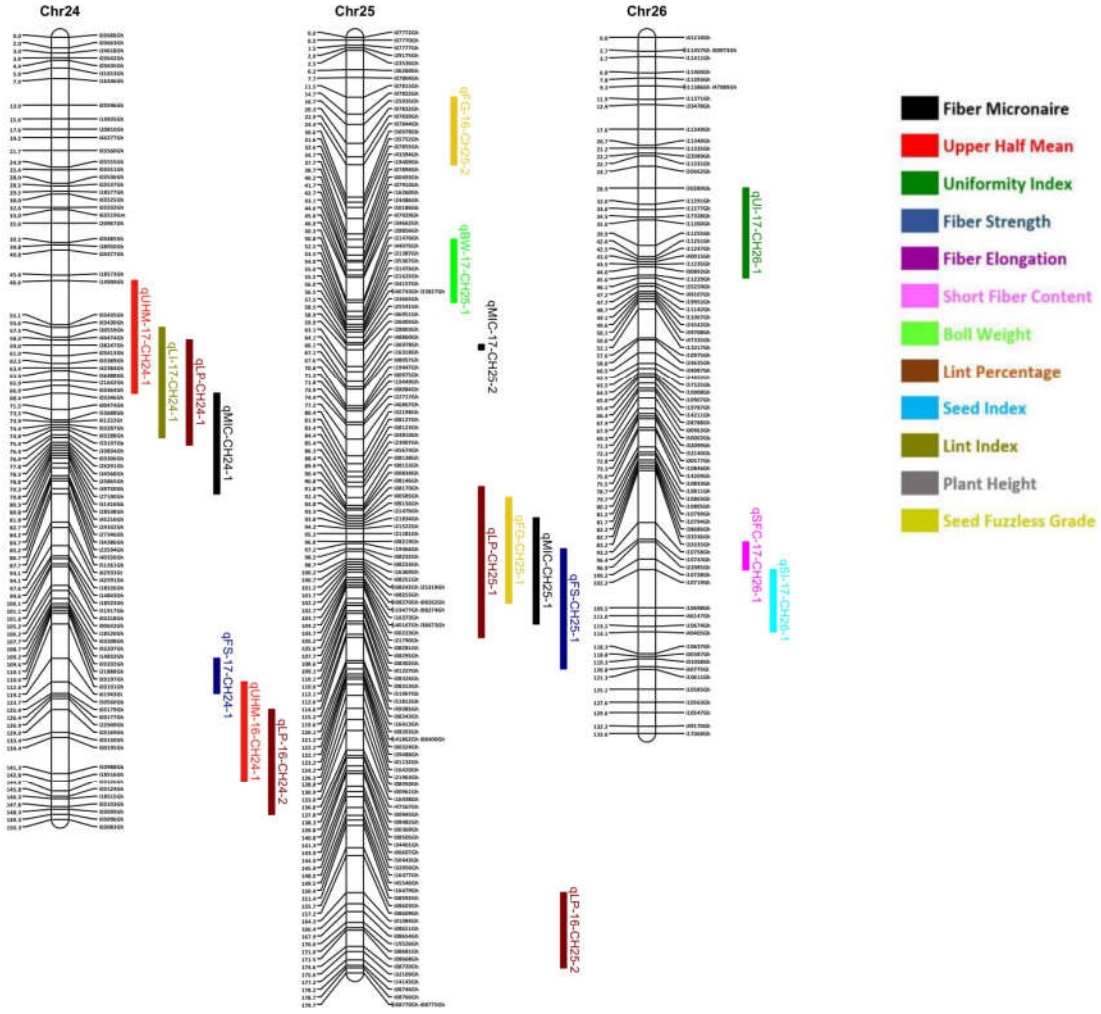

Figure 4. 7 Linkage maps of chromosomes Chr24(D3), Chr25(D12), Chr26(D6), along with detected QTLs for NC05AZ06 × NC11-2091 RILs population.

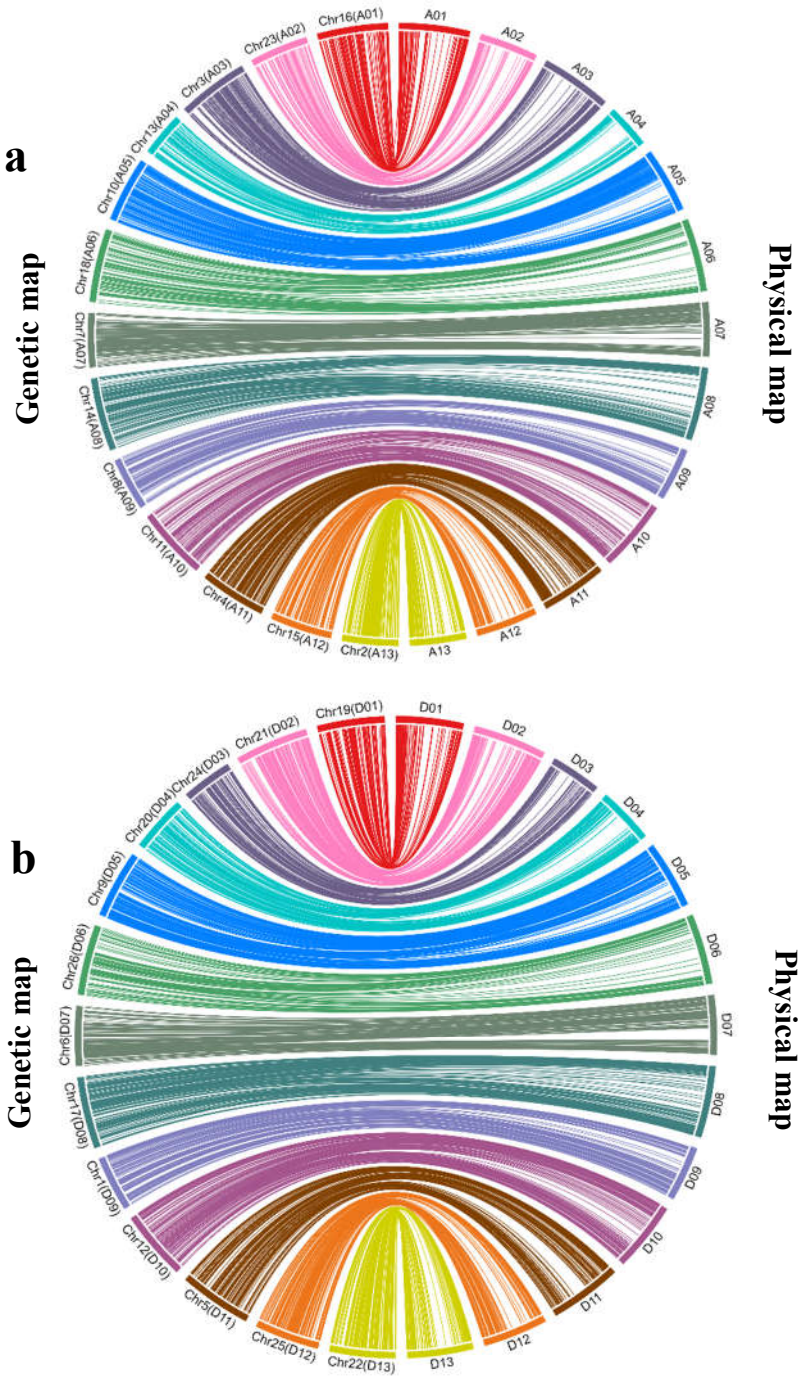

Figure 4. 8 Collinearity between genetic map and physical map in A subgenome and D subgenome for NC05AZ06 × NC11-2091 RILs population. (a) Collinearity for chrmosomes in A subgenome. (b) Collinearity for chrmosomes in D subgenome

CHAPTER 5: Results for NC05AZ06 × NC11-2100 RILs population

Analyses of the phenotypic traits

A summary of the statistical analysis for the phenotypic performance of the twelve traits are presented in Table 5.1. For the six fiber quality traits, micronaire (MIC), upper half mean length (UHM), uniformity index (UI), fiber strength (FS) and fiber elongation (FE) were significantly higher (18.7%, 16.2%, 1.8%, 6.7%, 52.1% respectively) in NC05AZ06 than in NC11-2100 while the short fiber content (SFC) of NC05AZ06 was lower (-18.2%) than NC11-2100. Among the four yield traits, boll weight (BW), lint percentage (LP), and lint index (LI) of NC05AZ06 were greater (35.9%, 11.6%, 25.8% respectively) than those of NC11-2100. No significant difference was found between the two parents for the seed index (SI). In the agronomic traits, the plant height (PH) of NC11-2100 was significantly taller (36.9%) than NC05AZ06. The seed fuzziness grade (FG) of NC05AZ06 is 100% (fuzz-rich) and the FG of NC11-2100 is 0 (fuzz-free). The broad-sense heritability of the traits calculated by the ratio of total genetic variance to total phenotypic variance for all the traits is listed in Table 5.2. Most traits except BW and PH showed high broad-sense heritability across two replications with values ranging from 77%-90%. The broad-sense heritability of BW and PH were 68% and 65%, which indicated that the trait performance of BW and PH was highly influenced by the environment.

Correlation coefficient analyses for the twelve traits is described in Table 5.3. For the fiber quality traits, MIC was significantly positively related to UI($P<0.01$), BW($P<0.05$), LP($P<0.01$), LI($P<0.01$) and significantly($P<0.01$) negatively related to UHM, SFC and PH. UHM was significantly ($P<0.01$) positively correlated with UI, FS, SI, FG and significantly negatively correlated with FE($P<0.01$), SFC($P<0.01$), LP($P<0.01$) and LI($P<0.05$). UI had

significantly($P<0.01$) positive correlation with FS, SI and significantly negative correlation with SFC($P<0.01$), LP($P<0.05$). FS was significantly positively correlated with BW($P<0.05$), SI($P<0.01$), FG($P<0.05$) and it was significantly($P<0.01$) negatively correlated with SFC and LP. FE and SFC showed significant($P<0.05$) positive correlation with LP. SFC showed significant($P<0.01$) negative correlation with SI. Among the yield traits, LI had significantly($P<0.01$) positive correlation with BW, LP and SI. SI was significantly($P<0.01$) positively related to BW and significantly($P<0.01$) negatively related to LP. Both agronomic traits PH and FG had negative correlation($P<0.05$) with yield traits LP and LI.

Construction of linkage maps

Out of 63,058 SNPs used in the genotyping, a total of 8,407 (13.3%) SNPs were polymorphic between the two parents. Finally, 2,017 SNPs were selected for linkage map construction after removing the bad-quality or duplicate SNPs. All the 2,017 markers were mapped to 26 linkage groups (26 chromosomes) (Fig 5.1 through Fig 5.7). This resulted in genetic map length of 3,837.46 cM with an average distance of 1.9 cM between markers (Table 5.4). Of these 2,017 SNPs, 954 SNPs were mapped to the A sub-genome while 1,063 SNPs were mapped to the D sub-genome. The 954 A sub-genome SNPs were mapped in the genome with a total length of 1974.73 cM at an average density of 2.07 cM while 1063 D sub-genome SNPs were mapped with a total length of 1862.73 cM and an average density of 1.75 cM (Table 5.4). Genetic lengths of 26 linkage groups ranged from 103.5 cM to 218.1 cM. Number of markers mapped per chromosome range from 52 to 120 and average marker density ranging from 1.38 cM to 2.66 cM in each group (Table 5.4). Thirty-one gaps (adjacent markers distance >10 cM) were identified on 18 different linkage groups (Table 5.4).

Of the 2,017 mapped SNPs, 100 (5.0%) SNP markers showed segregation distortion which spanned on 19 chromosomes, with the most distorted markers (16) and highest distortion rate (20%) on Chr.02 (A13) (Table 5.4). Thirteen segregation distortions region (SDR) were identified on 10 chromosomes, with 8 of the SDRs in A sub-genome and 5 SDRs in the D sub-genome (Table 5.4).

By comparing the genetic map with the physical map of the TM-1(*G.hirsutum*) reference genome sequence [39], the total 3837.46 cM length of the genetic map corresponded to 1907.20 Mb of the physical map, spanning 98.6% of the total genome length (Table 5.5). All chromosomes showed good coverage of the physical map with the range from 95.7% to 99.5% coveraged. A high degree of collinearity between the genetic map and physical map presented by circus plot confirmed the accuracy of the genetic map (Figure 5.8).

QTL analysis for cotton fiber quality, yield and agronomic traits

A total of 77 QTLs, including 40 QTLs for fiber quality traits, 27 QTLs for yield traits, 10 QTLs for agronomic traits, were detected based on composite interval mapping (CIM) analysis method, with the phenotypic variation explained by the QTLs ranged from 5% - 34.1% (Appendix C). By determining that SFC with lower value was favourable and other traits with higher value were favourable, the favourable alleles of 51 QTLs were derived from NC05AZ06 (P1) whereas 26 QTLs were contributed by NC11-2100 (P2). 32 QTLs were mapped in the A subgenome, including 24 QTLs with favourable alleles contributed by NC05AZ06 and 8 QTLs with favourable alleles contributed by NC11-2100. 45 QTLs were mapped in the D subgenome, including 27 QTLs with favourable alleles derived from NC05AZ06 and 18 QTLs with favourable alleles derived from NC11-2100 (Table 5.4, Appendix C). Of the total 77 mapped QTLs, 39 QTLs were major QTLs with PVE >10%, including 20 QTLs for fiber quality traits

(Table 5.6) (11 in the A sub-genome and 9 in the D sub-genome), 13 QTLs for yield traits (Table 5.7) (6 QTLs in the A sub-genome and 7 in the D sub-genome) and 6 QTLs for agronomic traits (2 in A sub-genome and 4 in D su-bgenome (Table 5.8).

QTL for fiber quality traits

A total of 40 QTLs, including 25 QTLs with favourable alleles contributed by NC05AZ06 and 15 QTLs with favourable alleles contributed by NC11-2100, were identified for six fiber quality traits, with the PVE ranging from 5.1% to 31.4% (Table 5.6, Appendix C). Out of these 40 fiber quality QTLs, 18 QTLs were mapped in the A sub-genome (13 QTLs with favourable alleles from NC05AZ06 and 5 from NC11-2100) and 22 QTLs were mapped on the D sub-genome (12 QTLs with favourable alleles from NC05AZ06 and 10 from NC11-2100).

Micronaire (MIC): Seven QTLs explaining 5.6% to 31.4% of the phenotypic variance (PV) were identified for fiber micronaire (Appendix C). Among these, 3 QTLs with favourable alleles were derived from parental accession NC11-2100. Three major QTLs q2MIC-CH7-1, q2MIC-CH15-1, q2MIC-CH26-1, explained 18.8%, 31.4%, 13% of phenotypic variance, respectively (Table 5.6).

Upper half mean length (UHML): 6 QTLs explaining 6.8% to 14% of PV were identified for UHML, including 5 major QTLs (Table 5.6, Appendix C). From these 6 QTLs, q2UHM-CH1-1 was the only QTL with favourable alleles derived from NC11-2100. Five major QTLs: q2UHM-CH1-1, q2UHM-CH5-1, q2UHM-CH14-1, q2UHM-CH21-1, q2UHM-CH23-1, explained 13.3%, 11.3%, 12.7%, 13.5%, 14% of PV, respectively.

Uniformity index (UI): For uniformity index, eight QTLs explaining 5.1% to 17.5% of PV were detected, with six QTLs' favourable alleles conferred by parental accession NC05AZ06 (Table

5.6, Appendix C). Only two major QTLs were identified: q2UI-CH14-1, q2UI-CH25-1, with 17.5% and 12.1% of PVE.

Fiber strength (FS): For fiber strength, six QTLs explaining 6.9% to 20.7% of PV were mapped in the linkage maps (Table 5.6, Appendix C). Of these 6 QTLs, 3 were identified with favourable alleles derived from NC05AZ06. Four are major QTLs: q2FS-CH2-1, q2FS-CH14-1, q2FS-CH18-1, q2FS-CH26-1, explaining 10.5%, 20.7%, 14% and 10% of PV.

Elongation (FE): Seven QTLs explaining 6.1% to 18.6% of PV were identified for fiber elongation, with 3 major QTLs included (Table 5.6, Appendix C). These 3 major QTLs: q2FE-CH3-1, q2FE-CH6-1, q2FE-CH20-1, explained 10.5%, 16.5% and 18.6% of PV. Among the 7 QTLs for fiber elongation, two carried favourable alleles contributed by NC11-2100.

Short fiber content (SFC): For short fiber content, 6 QTLs explaining 6.9% to 20.4% of PV were detected (Table 5.6, Appendix C), including two QTLs with favourable alleles derived from NC11-2100. 3 major QTL (q2SFC-CH5-1, q2SFC-CH11-1, q2SFC-CH14-1) were detected with 20.4%, 18.9%, 15.8% of PVE.

QTL for yield traits

A total of 27 QTLs, including 18 QTLs with favourable alleles contributed by NC05AZ06 and 9 QTLs with favourable alleles contributed by NC11-2100, were identified for yield traits (BW, LP, SI, LI), with the PVE ranging from 5% to 24.4% (Table 5.7, Appendix C). Among the 27 yield QTLs, 12 QTLs were mapped in the A sub-genome, including 10 QTLs with favourable alleles contributed by NC05AZ06 and 2 contributed by NC11-2100 and 15 QTLs were mapped in the D sub-genome, including 8 QTLs with favourable alleles conferred by NC05AZ06 and 7 conferred by NC11-2100.

Boll weight (BW): Seven QTLs explaining 6.6% to 16.9% of PV were identified and 2 QTLs' favourable alleles were derived from NC11-2100 (Table 5.7, Appendix C). Of these 7 QTLs for boll weight, five major QTLs (q2BW-CH11-1, q2BW-CH16-1, q2BW-CH21-1, q2BW-CH23-1, q2BW-CH26-1) were mapped with 10.5%, 16.9%, 12.9%, 13.6%, 11.4% of PVE respectively.

Lint percentage (LP): For lint percentage, 7 QTLs explaining 5% to 24.4% of PV were identified, which included 4 major QTLs (Table 5.7, Appendix C). These four major QTLs: q2LP-CH5-1, q2LP-CH14-1, q2LP-CH25-1, q2LP-CH26-1, explained 20.6%, 19.4%, 17.7%, 24.4% of PV, respectively for LP. Among the 7 QTLs, 3 QTLs' favourable alleles were derived from NC11-2100.

Seed index (SI): 6 QTLs explaining 7% to 18.5% of PV were detected for seed index with 2 major QTLs (Table 5.7, Appendix C). Two major QTLs q2SI-CH15-1 and q2SI-CH20-1, with 18.5%, 10.1% of PVE, were identified for SI. Of these 6 QTLs, three QTLs with favourable alleles were derived from NC05AZ06.

Lint index (LI): For lint index, seven QTLs explaining 6.3% to 16.3% of PV were identified and only one QTL(q2LI-CH9-1)'s favourable alleles were contributed by NC11-2100 (Table 5.7, Appendix C). Two major QTLs q2LI-CH7-1, q2LI-CH25-1, explained 13.4% and 16.3% of PV, respectively.

QTL for agronomic traits

A total of 10 QTLs, including 6 QTLs with favourable alleles contributed by NC05AZ06 and 4 from NC11-2100, were identified for agronomic traits (plant height and fuzziness grade), with the PVE from 5.6% to 34.1% (Table 5.8, Appendix C). 2 QTLs were mapped in the A sub-genome and 8 QTLs were mapped in the D sub-genome.

Plant height (PH): Six QTLs explaining 5.6% to 15.4% of PV were identified for plant height (Table 5.8, Appendix C). 3 of these six QTLs with positive additive effect for plant height were derived from NC11-2100. Three major QTLs for plant height (q2PH-CH5-1, q2PH-CH10-1, q2PH-CH19-1) explained 14.1%, 10.4% and 15.4% of PV, respectively.

Seed fuzziness (FG): For seed fuzziness grade, 4 QTLs explaining 9% to 34.1% of PV were identified (Table 5.8, Appendix C). Three major QTLs: q2FG-CH14-1, q2FG-CH22-1, q2FG-CH26-1, were detected with 14.6%, 34.1% and 10.4% of PVE. q2FG-CH26-1 was the only QTL with negative additive effect for seed fuzziness contributed by NC11-2100.

QTL clusters & meta-analysis

A QTL cluster was a short region (<30 cM) on the linkage map containing multiple QTLs [14]. In this population, 16 QTL Clusters (Table 5.9, Table 5.10) were identified in 14 different chromosomes (Chr1, Chr3, Chr5, Chr6, Chr7, Chr11, Chr14, Chr16, Chr20, Chr21, Chr23, Chr24, Chr25 and Chr26) after all the QTLs were mapped (Fig 5.1-Fig 5.7). 6 QTL clusters were detected in A sub-genome and 10 clusters were detected in D sub-genome, the names of the clusters depended on the trait categories of QTLs they contained. For example, Q2-* meant a cluster contained only fiber quality traits QTLs; QY2-* meant a cluster contained both fiber quality and yield traits QTLs; QYA2-* meant a cluster contained fiber quality, yield and agronomic traits QTLs; etc.

For fiber quality traits, 3 QTL Clusters (Q2-1 to Q2-3) containing multiple QTLs were detected (Table 5.9). Cluster Q2-1 contained QTLs from FE and UHM with favorable alleles conferred by NC11-2100 while cluster Q2-3 contained QTLs from FE and UHM with favorable alleles conferred by NC05AZ06. Q2-2 carried QTLs from FE, SFC and UI with favorable alleles contributed by NC05AZ06.

For yield traits, two QTL Clusters (Y2-1 to Y2-2) with QTLs for LI and BW were identified (Table 5.9). The favorable alleles for LI and BW QTLs in these two Clusters were derived from NC05AZ06.

Other eleven Clusters contained multiple QTLs from different trait categories (Table 5.10). 8 clusters (QY2-1 to QY2-8) with multiple QTLs in fiber quality and yield were detected, including 2 clusters (QY2-2, QY2-6) with favorable alleles from NC05AZ06, 2 clusters (QY2-3, QY2-7) with favorable alleles from NC11-2100 and 4 clusters (QY2-1, QY2-4, QY2-5, QY2-8) with favorable alleles from different parents. QYA2-1 and QYA2-2 were two clusters carrying multiple QTLs from all 3 trait categories. QYA2-1 with a region in Chr.14 (A8) from 60.3 cM to 68.82 cM, contained 6 QTLs for LP, FG, SFC, UI, FS and UHM. QYA2-2 with a region in Chr.26 (D6) from 39.59 cM to 56.49 cM, carried 5 QTLs for LP, FG, MIC, FS and BW.

A total of 2884 cotton QTLs for 11 traits: MIC(442), UHM(524), UI(289), FS(470), FE(287), SFC(58), BW(176), LP(327), LI(42), SI(147), PH(122), which were collected by the CottonQTLdb [14, 15, 40] in different interspecific or intraspecific populations from 156 previous publications (http://www2.cottonqtldb.org:8081/references), were used for meta-analysis in recent study (Appendix A).

Out of 73 QTLs for 11 traits in this population, 51 QTLs were found to share the similar genetic and physical positions (genetic distance window of <20 cM) with previous reported QTLs, including 19 QTLs in A subgenome and 32 QTLs in D subgenome. All these 51 shared QTLs were separated in 11 different traits: UI(6), UHM(6), MIC(6), FE(6), FS(5), SFC(2), LP(6), BW(4), SI(4), LI(2) and PH(4), including 26 major QTLs. 22 QTLs were unique QTLs with 12 QTLs in A subgenome and 10 QTLs in D subgenome, including 4 for SFC, 2 for UI, 1 for FE, 1 for FS, 1 for MIC, 5 for LI, 3 for BW, 2 for SI, 1 for LP, 2 for PH. Out of these 22

unique QTLs, 10 were major QTLs (q2BW-CH26-1, q2FS-CH18-1, q2LI-CH7-1, q2LI-CH25-1, q2LP-CH25-1, q2SFC-CH14-1, q2SFC-CH11-1, q2SFC-CH5-1, q2SI-CH15-1, q2UI-CH25-1).

Tables

Table 5. 1 Trait Performance of NC05AZ06 × NC11-2100 RILs population and their parents

Type	Trait	Parents			RILs			
		NC05AZ06(P1)	NC11-2100 (P2)	P1-P2	Min	Max	Mean	SD
Fiber Quality	MIC	5.08	4.28	0.8**	3.00	5.75	4.66	0.44
	UHM	1.14	0.99	0.16**	0.92	1.18	1.04	0.06
	UI	84.33	82.83	1.5**	80.00	85.45	83.11	0.99
	FS	27.85	26.10	1.75*	21.60	32.50	26.46	2.05
	FE	8.25	5.43	2.83**	4.80	9.05	6.80	0.82
	SFC	7.65	9.35	-1.7**	6.65	11.95	8.35	0.79
Yield	BW	5.18	3.82	1.37**	3.12	5.29	4.05	0.42
	LP	41.3	37.0	4.3**	32.5	45.8	38.9	2.27
	SI	11.15	10.55	0.6	7.66	12.52	9.81	0.78
	LI	7.84	6.24	1.61**	4.64	8.40	6.25	0.64
Agronomic	PH	34.4	47.1	-12.7**	26.2	56.8	39.7	5.52
	FG	100	0	100**	0	100	70.6	38.6

*: 0.05 significant level; **: 0.01 significant level

[a] MIC micronaire; UHM upper half mean length; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzziness grade of seed

Table 5. 2 The broad-sense heritability of traits in NC05AZ06 × NC11-2100 RILs Population

	MIC[a]	UHM	UI	FS	FE	SFC	BW	LP	SI	LI	PH	FG
V_g	0.193	0.003	0.989	4.191	0.666	0.630	0.180	5.171	0.608	0.414	30.451	11.70
V_p	0.241	0.004	1.292	4.719	0.750	0.806	0.265	5.741	0.770	0.517	46.879	13.36
H^2	80%	87%	77%	89%	89%	78%	68%	90%	79%	80%	65%	88%

The broad-sense heritability (H^2) = genetic variance (V_g)/phenotypic variance (V_p)

[a] MIC micronaire; UHM upper half mean length; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzziness grade of seed

Table 5. 3 Correlation analysis between the traits in NC05AZ06 × NC11-2100 RILs Population

	MIC[a]	UHM	UI	FS	FE	SFC	BW	LP	SI	LI	PH
UHM	-0.33**										
UI	0.28**	0.35**									
FS	0.12	0.59**	0.64**								
FE	0.06	-0.38**	0.04	-0.17							
SFC	-0.32**	-0.24**	-0.87**	-0.64**	-0.13						
BW	0.23*	0.08	0.11	0.21*	-0.01	-0.08					
LP	0.45**	-0.48**	-0.2*	-0.33**	0.24*	0.18*	0.11				
SI	0.1	0.3**	0.3**	0.35**	-0.12	-0.27**	0.49**	-0.36**			
LI	0.5**	-0.24*	0.04	-0.06	0.14	-0.03	0.5**	0.68**	0.43**		
PH	-0.42**	0.12	-0.15	-0.06	-0.15	0.17	-0.18	-0.22*	0.01	-0.2*	
FG	-0.06	0.26**	0.04	0.21*	0.01	0	-0.05	-0.2*	0.02	-0.19*	0.08

*: 0.05 significant level; **:0.01 significant level

[a] MIC micronaire; UHM upper half mean length; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzziness grade of seed

Table 5. 4 Linkage Map information of NC05AZ06 × NC11-2100 RILs Population

Linkage Group	physical Map	Number of Markers	Length	Average Density	Gaps[a] (>10cM)	Number of QTLs	Distorted Markers[b]	No. SDR[c]
Chr16	A1	67	117.59	1.76	0	2	7	1
Chr23	A2	62	103.5	1.67	0	4	0	0
Chr03	A3	75	140.71	1.88	1	2	1	0
Chr13	A4	57	122.7	2.15	1	2	3	1
Chr10	A5	111	213.64	1.92	2	1	1	0
Chr18	A6	52	134.39	2.58	3	1	5	1
Chr07	A7	67	144.34	2.15	1	4	0	0
Chr14	A8	90	163.85	1.82	2	7	3	1
Chr08	A9	65	151.58	2.33	3	0	0	0
Chr11	A10	76	162.27	2.14	2	4	7	1
Chr04	A11	82	218.05	2.66	3	1	2	0
Chr15	A12	70	166.97	2.39	1	2	4	0
Chr02	A13	80	135.14	1.69	0	2	16	3
A-Subgenome		954	1974.73	2.07	19	32	49	8
Chr19	D1	75	134.8	1.8	0	2	2	0
Chr21	D2	89	147.26	1.65	0	3	12	1
Chr24	D3	57	124.58	2.19	2	4	0	0
Chr20	D4	60	112.17	1.87	0	5	3	0
Chr09	D5	113	210.62	1.86	2	2	2	0
Chr26	D6	77	106	1.38	0	6	0	0
Chr06	D7	94	136.98	1.46	1	2	0	0
Chr17	D8	77	141.55	1.84	1	2	10	2
Chr01	D9	70	113.43	1.62	1	5	1	0
Chr12	D10	87	134.83	1.55	0	2	14	1
Chr05	D11	81	187.98	2.32	2	7	6	1
Chr25	D12	120	193.26	1.61	2	4	1	0
Chr22	D13	63	119.27	1.89	1	1	0	0
D-Subgenome		1063	1862.73	1.75	12	45	51	5
Total		2017	3837.46	1.9	31	77	100	13

[a]Gap: Distance between two adjacent markers > 10 cM
[b]Distorted Markers: Markers showing segregation distortion (chi-square test P<0.05)
[c]SDR segregation distortion region

Table 5. 5 Length of genome spanned by the linkage map for NC05AZ06 × NC11-2100 RILs population

Linkage Group	physical Map	Start site	End site	Genome spanned length	Reference genome length	Genome coverage ratio
Chr16	A01	656421	99708430	99052009	99884700	99.2%
Chr23	A02	430390	81755379	81324989	83447906	97.5%
Chr03	A03	417360	100112893	99695533	100263045	99.4%
Chr13	A04	2048781	62518933	60470152	62913772	96.1%
Chr10	A05	1290413	91618214	90327801	92047023	98.1%
Chr18	A06	332742	102920435	102587693	103170444	99.4%
Chr07	A07	53456	77180364	77126908	78251018	98.6%
Chr14	A08	509660	102782124	102272464	103626341	98.7%
Chr08	A09	311306	73259616	72948310	74999931	97.3%
Chr11	A10	810285	100576604	99766319	100866604	98.9%
Chr04	A11	792440	93108419	92315979	93316192	98.9%
Chr15	A12	407936	87211506	86803570	87484866	99.2%
Chr02	A13	406169	79475941	79069772	79961121	98.9%
A-Subgenome				1143761499	1160232963	98.6%
Chr19	D01	410926	61400232	60989306	61456009	99.2%
Chr21	D02	288235	67060530	66772295	67284553	99.2%
Chr24	D03	564364	46522952	45958588	46690656	98.4%
Chr20	D04	614836	51009055	50394219	51454130	97.9%
Chr09	D05	263755	61433131	61169376	61933047	98.8%
Chr26	D06	134070	64095362	63961292	64294643	99.5%
Chr06	D07	1202298	54641597	53439299	55312611	96.6%
Chr17	D08	391377	65489341	65097964	65894135	98.8%
Chr01	D09	1460627	50251477	48790850	50995436	95.7%
Chr12	D10	180800	62837647	62656847	63374666	98.9%
Chr05	D11	631689	65974293	65342604	66087774	98.9%
Chr25	D12	259476	59087499	58828023	59109837	99.5%
Chr22	D13	217310	60339884	60122574	60534298	99.3%
D-Subgenome				763523237	774421795	98.6%
Total				1907284736	1934654758	98.6%

Table 5. 6 Major[a] QTLs of the fiber quality traits identified in NC05AZ06 × NC11-2100 RILs population

Trait[b]	QTL	Range	Peak	LOD	PVE[c]	AE[d]
MIC	q2MIC-CH7-1	50.2-71.41	61.78	10.2	18.8	0.212
	q2MIC-CH15-1	33.78-49.73	33.78	12.2	31.4	0.308
	q2MIC-CH26-1	39.59-55.23	48.13	7.7	13	0.176
UHM	q2UHM-CH1-1	91.64-111.46	99.33	8.8	13.3	-0.022
	q2UHM-CH5-1	80.53-97.46	91.64	7.5	11.3	0.02
	q2UHM-CH14-1	58.36-77.94	68.82	8.4	12.7	0.02
	q2UHM-CH21-1	2.53-21.12	10.34	8.9	13.5	0.021
	q2UHM-CH23-1	13.6-33.77	23.35	9	14	0.022
UI	q2UI-CH14-1	52.88-72.3	64.1	8.7	17.5	0.465
	q2UI-CH25-1	182.76-191.94	183.99	6.1	12.1	0.396
FS	q2FS-CH2-1	2.6-20.28	8.23	6.8	10.5	0.662
	q2FS-CH14-1	54.96-77.31	68.16	12.2	20.7	-0.908
	q2FS-CH18-1	111.14-130.27	123.08	8.7	14	0.775
	q2FS-CH26-1	46.85-62.92	56.49	6.9	10	-0.623
FE	q2FE-CH3-1	90.4-107.43	96.25	6.2	10.5	0.302
	q2FE-CH6-1	70.48-80.33	70.48	3.2	16.5	-0.379
	q2FE-CH20-1	79.92-103.12	93.66	10.3	18.6	0.428
SFC	q2SFC-CH5-1	66.81-84.65	76.57	10.4	20.4	0.397
	q2SFC-CH11-1	60.67-73.61	73.61	4	18.9	-0.379
	q2SFC-CH14-1	52.88-73.59	64.1	8.5	15.8	-0.346

[a]Major QTLs: The QTLs' PVE>10%

[b]MIC micronaire; UHM upper half mean; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content

[c]PVE phenotypic variation explained [d]AE additive effect

Table 5. 7 Major[a] QTLs of the yield traits identified in NC05AZ06 × NC11-2100 RILs population

Trait[b]	QTL	Range	Peak	LOD	PVE[c]	AE[d]
BW	q2BW-CH11-1	109.87-127.67	120.42	5.9	10.5	0.158
	q2BW-CH16-1	79.46-94.89	82.58	8.7	16.9	0.201
	q2BW-CH21-1	2.53-21.12	3.77	7	12.9	0.174
	q2BW-CH23-1	47.38-64.11	55.8	7.4	13.6	0.181
	q2BW-CH26-1	53.26-71.21	62.92	6.3	11.4	0.165
LP	q2LP-CH5-1	118.9-140.93	128.11	10.5	20.6	1.281
	q2LP-CH14-1	50.22-68.82	60.3	10.2	19.4	1.192
	q2LP-CH25-1	96.82-109.09	109.09	7	17.7	-1.075
	q2LP-CH26-1	31.17-45.55	36.25	12.1	24.4	1.325
SI	q2SI-CH15-1	142.36-162.85	158.13	8.2	18.5	0.364
	q2SI-CH20-1	99.31-110.9	105.79	5.2	10.1	-0.277
LI	q2LI-CH7-1	35.03-50.2	46.49	6.3	13.4	0.275
	q2LI-CH25-1	104.65-116.01	109.09	6.7	16.3	0.291

[a]Major QTLs: The QTLs' PVE>10%
[b]BW boll weight; LP lint percentage; SI seed index; LI lint index
[c]PVE phenotypic variation explained [d]AE additive effect

Table 5. 8 Major[a] QTLs of the agronomic traits identified in NC05AZ06 × NC11-2100 RILs population

Trait[b]	QTL	Range	Peak	LOD	PVE[c]	AE[d]
PH	q2PH-CH5-1	52.54-72.62	60.07	7.8	14.1	2.609
	q2PH-CH10-1	161.7-187.01	179.01	5.7	10.4	-2.401
	q2PH-CH19-1	121.88-133.55	129.76	8.6	15.4	-2.835
FG	q2FG-CH14-1	52.88-64.76	52.88	3.2	14.6	15.823
	q2FG-CH22-1	105.95-118.66	118.04	19.8	34.1	25.519
	q2FG-CH26-1	32.42-53.26	45.55	8.2	10.4	-13.687

[a]Major QTLs: The QTLs' PVE>10%
[b]PH plant height; FG fuzziness grade of seed
[c]PVE phenotypic variation explained [d]AE additive effect

Table 5. 9 Clusters containing multiple QTLs from same trait category (fiber quality or yield) in NC05AZ06 × NC11-2100 RILs population

Cluster[a]	Chr.	Region	QTLs	Trait	Source of favorable alleles	PVE
Q2-1	CH1	91.64-97.4	q2FE-CH1-1	FE	NC11-2100	6.1
			q2UHM-CH1-1	UHM	NC11-2100	13.3
Q2-2	CH11	60.67-73.61	q2FE-CH11-1	FE	NC05AZ06	9.6
			q2SFC-CH11-1	SFC	NC05AZ06	18.9
			q2UI-CH11-1	UI	NC05AZ06	7.4
Q2-3	CH24	46.05-56.35	q2FS-CH24-1	FS	NC05AZ06	7.9
			q2UHM-CH24-1	UHM	NC05AZ06	6.8
Y2-1	CH16	70.66-82.58	q2LI-CH16-1	LI	NC05AZ06	7.5
			q2BW-CH16-1	BW	NC05AZ06	16.9
Y2-2	CH23	47.38-53.91	q2BW-CH23-1	BW	NC05AZ06	13.6
			q2LI-CH23-1	LI	NC05AZ06	6.3

[a]Q2: Cluster containing multiple QTLs for fiber quality traits
Y2: Cluster containing multiple QTLs for yield traits

Table 5. 10 Clusters containing multiple QTLs from different trait categories (fiber quality, yield or agronomic) in NC05AZ06 × NC11-2100 RILs population

Cluster[a]	Chr.	Region	QTLs	Trait	Source of favorable alleles	PVE
QY2-1	CH1	30.66-44.86	q2UI-CH1-1	UI	NC11-2100	5.4
			q2SI-CH1-1	SI	NC05AZ06	9.5
			q2MIC-CH1-1	MIC	NC11-2100	5.6
QY2-2	CH3	93.71-107.43	q2FE-CH3-1	FE	NC05AZ06	10.5
			q2LI-CH3-1	LI	NC05AZ06	9.9
QY2-3	CH6	61.05-70.48	q2LP-CH6-1	LP	NC11-2100	8.7
			q2FE-CH6-1	FE	NC11-2100	16.5
QY2-4	CH7	46.49-65.63	q2LI-CH7-1	LI	NC05AZ06	13.4
			q2MIC-CH7-1	MIC	NC05AZ06	18.8
			q2SFC-CH7-1	SFC	NC05AZ06	8
QY2-5	CH20	92.42-105.79	q2FE-CH20-1	FE	NC05AZ06	18.6
			q2FS-CH20-1	FS	NC11-2100	6.9
			q2SI-CH20-1	SI	NC11-2100	10.1
QY2-6	CH21	3.77-10.34	q2BW-CH21-1	BW	NC05AZ06	12.9
			q2UHM-CH21-1	UHM	NC05AZ06	13.5
QY2-7	CH24	114.44-117.55	q2LP-CH24-1	LP	NC11-2100	8.6
			q2MIC-CH24-1	MIC	NC11-2100	6.5
QY2-8	CH25	104.65-117.31	q2LP-CH25-1	LP	NC11-2100	17.7
			q2LI-CH25-1	LI	NC05AZ06	16.3
			q2SFC-CH25-1	SFC	NC05AZ06	7.4
QA2-1	CH5	66.81-86.62	q2PH-CH5-1	PH	NC05AZ06	14.1
			q2SFC-CH5-1	SFC	NC11-2100	20.4
			q2UI-CH5-1	UI	NC11-2100	7.7
			q2FG-CH5-1	FG	NC05AZ06	9
			q2UHM-CH5-1	UHM	NC05AZ06	11.3
QYA2-1	CH14	60.3-68.82	q2LP-CH14-1	LP	NC05AZ06	19.4
			q2FG-CH14-1	FG	NC05AZ06	14.6
			q2SFC-CH14-1	SFC	NC05AZ06	15.8
			q2UI-CH14-1	UI	NC05AZ06	17.5
			q2FS-CH14-1	FS	NC11-2100	20.7
			q2UHM-CH14-1	UHM	NC05AZ06	12.7
QYA2-2	CH26	39.59-56.49	q2LP-CH26-1	LP	NC05AZ06	24.4
			q2FG-CH26-1	FG	NC11-2100	10.4
			q2MIC-CH26-1	MIC	NC05AZ06	13
			q2FS-CH26-1	FS	NC11-2100	10
			q2BW-CH26-1	BW	NC05AZ06	11.4

[a]QY2: Cluster containing multiple QTLs for fiber quality and yield traits
QA2: Cluster containing multiple QTLs for fiber quality and agronomic traits
QYA2: Cluster containing multiple QTLs for fiber quality, yield and agronomic traits

Figures

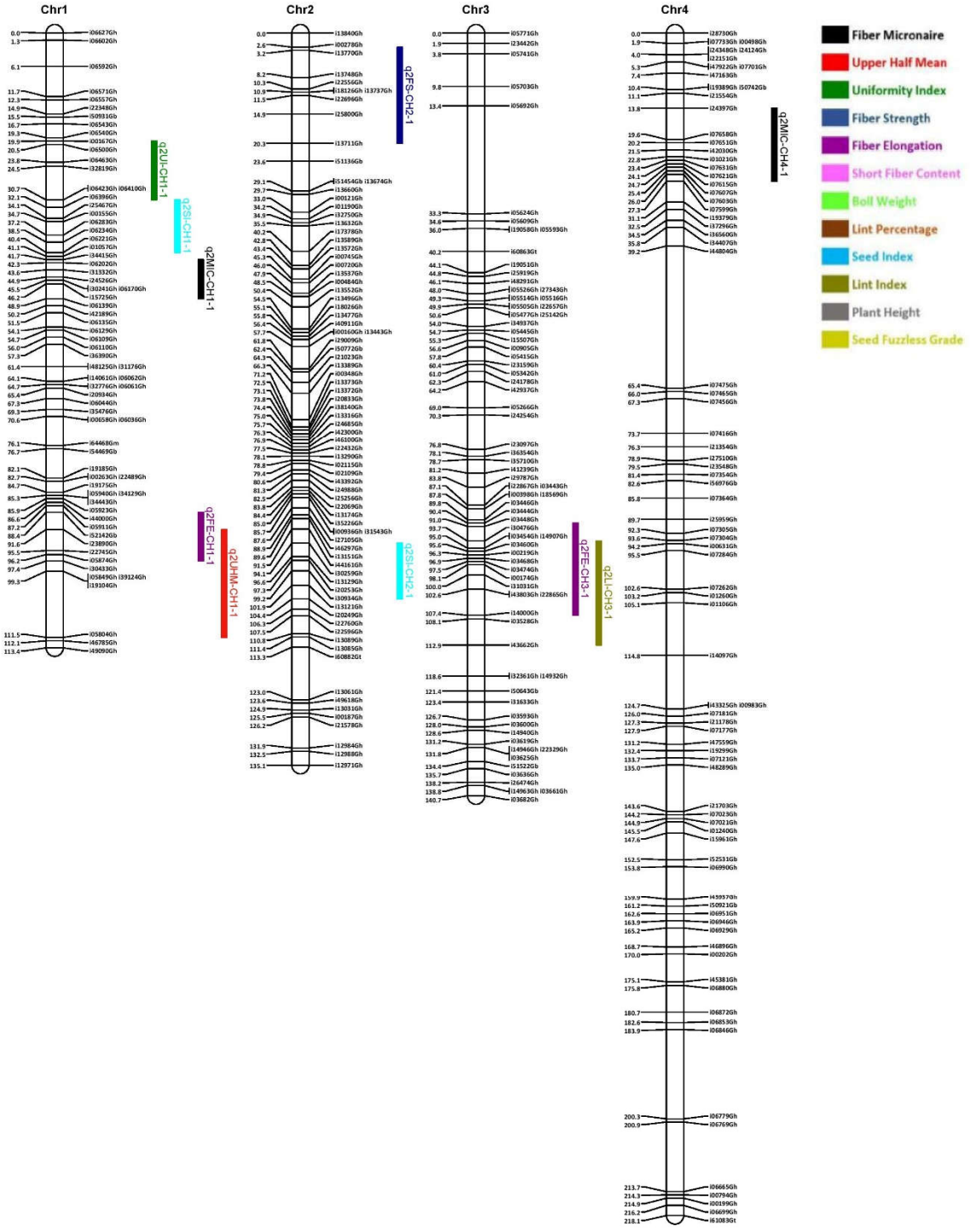

Figure 5. 1 Linkage maps of chromosomes Chr1(D9), Chr2(A13), Chr3(A3), Chr4(A11) along with detected QTLs for NC05AZ06 × NC11-2100 RILs population.

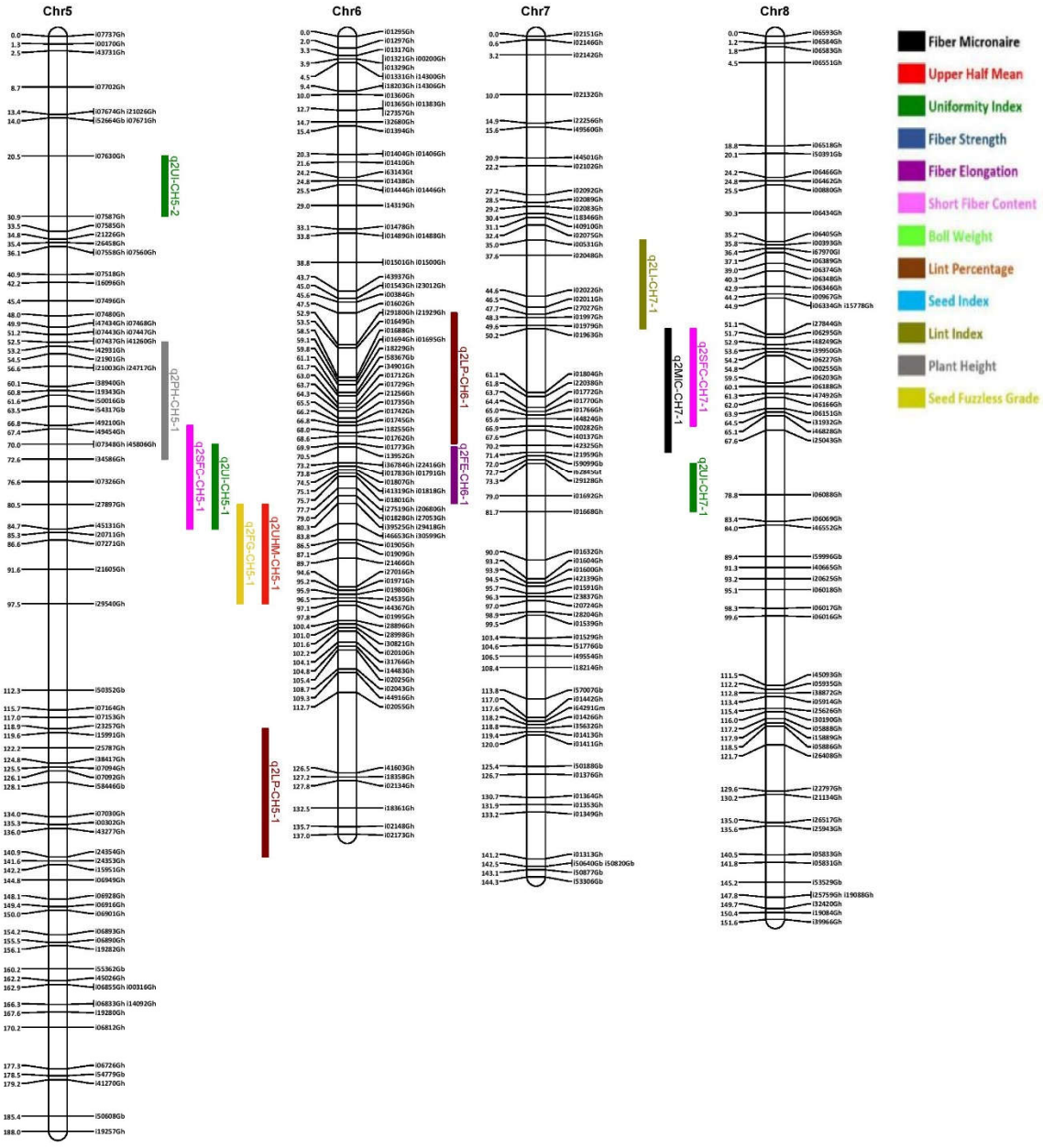

Figure 5. 2 Linkage maps of chromosomes Chr5(D11), Chr6(D7), Chr7(A7), Chr8(A9), along with detected QTLs for NC05AZ06 × NC11-2100 RILs population.

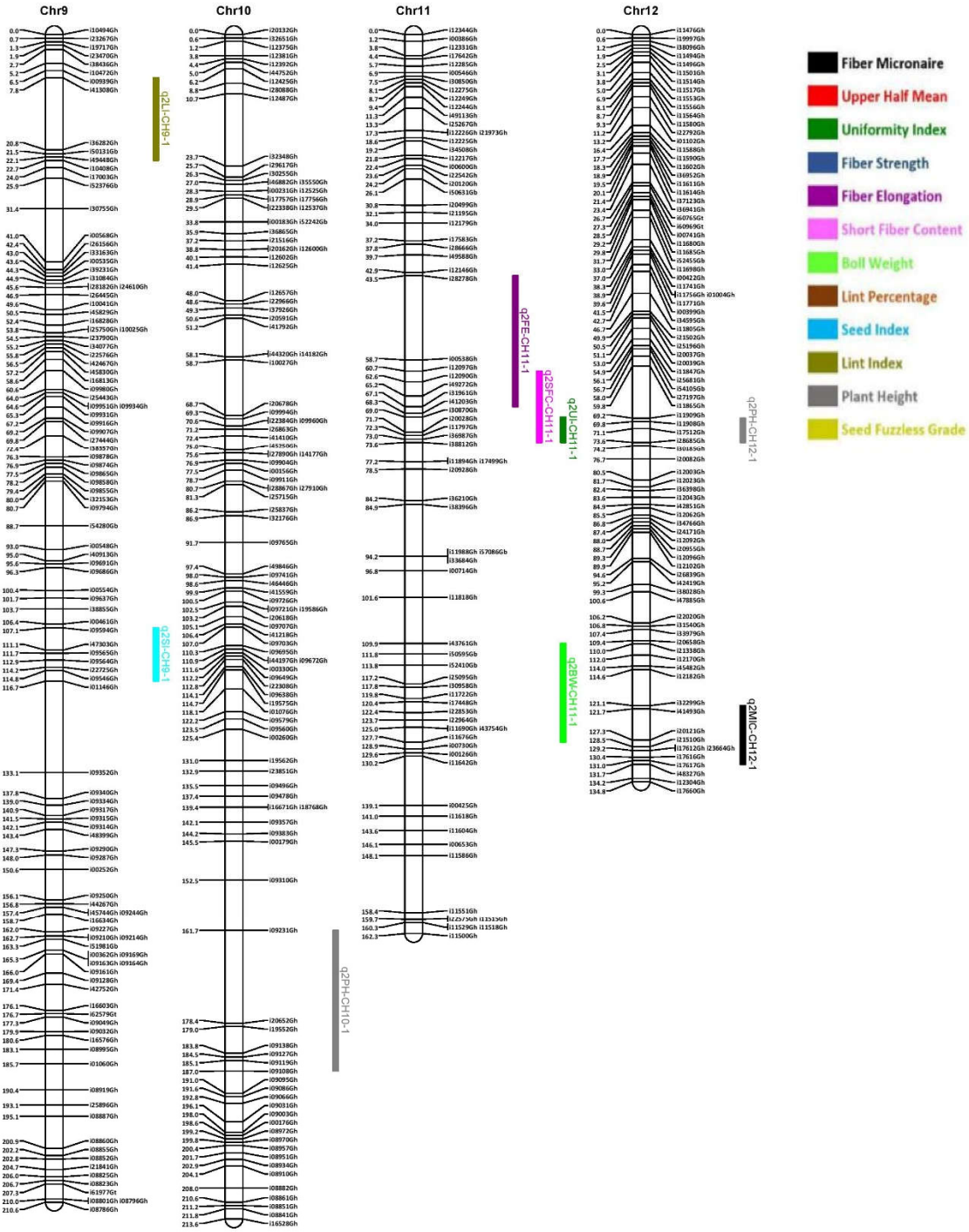

Figure 5. 3 Linkage maps of chromosomes Chr9(D5), Chr10(A5), Chr11(A10), Chr12(D10), along with detected QTLs for NC05AZ06 × NC11-2100 RILs population.

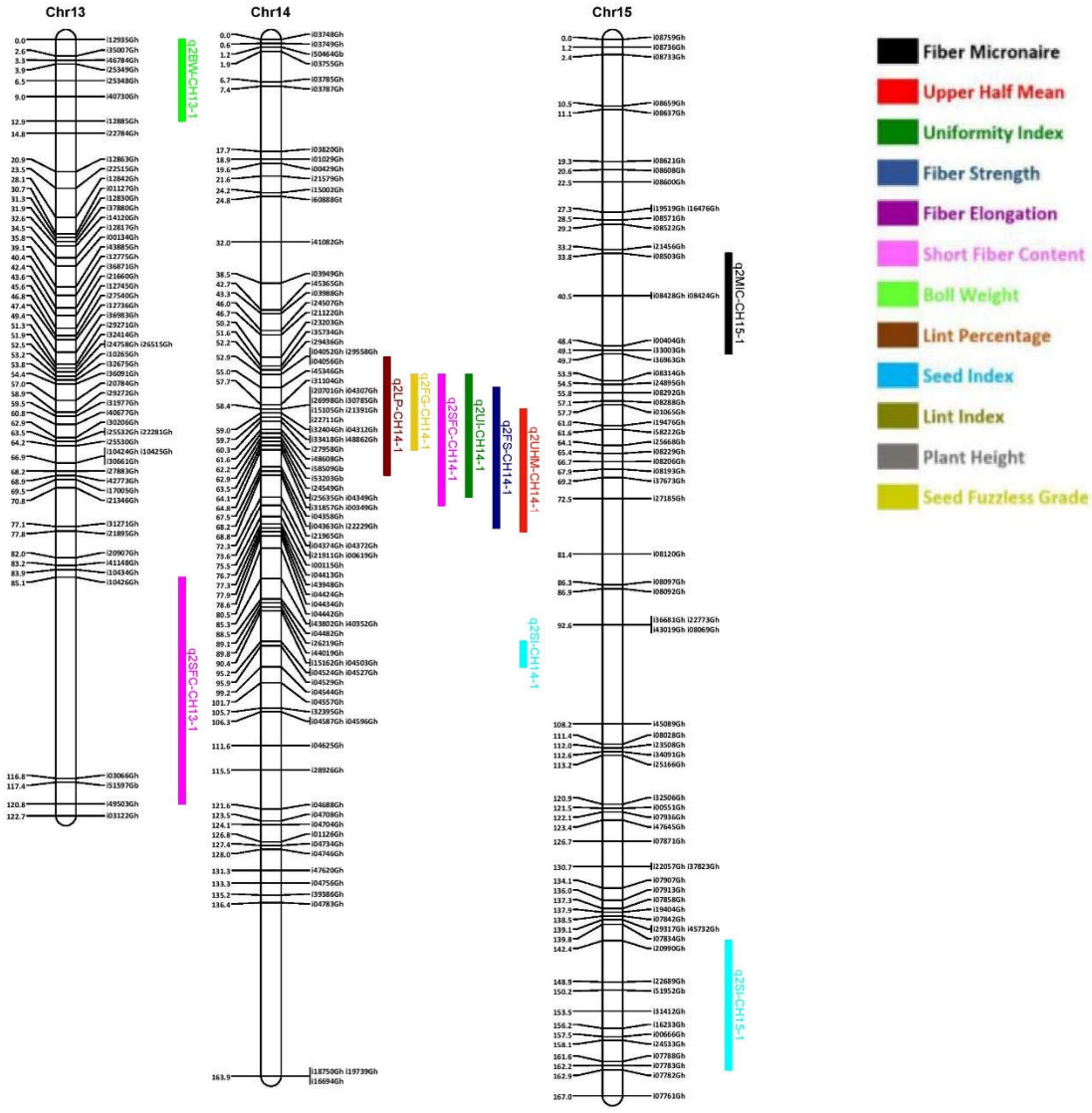

Figure 5. 4 Linkage maps of chromosomes Chr13(A4), Chr14(A8), Chr15(A12), along with detected QTLs for NC05AZ06 × NC11-2100 RILs population.

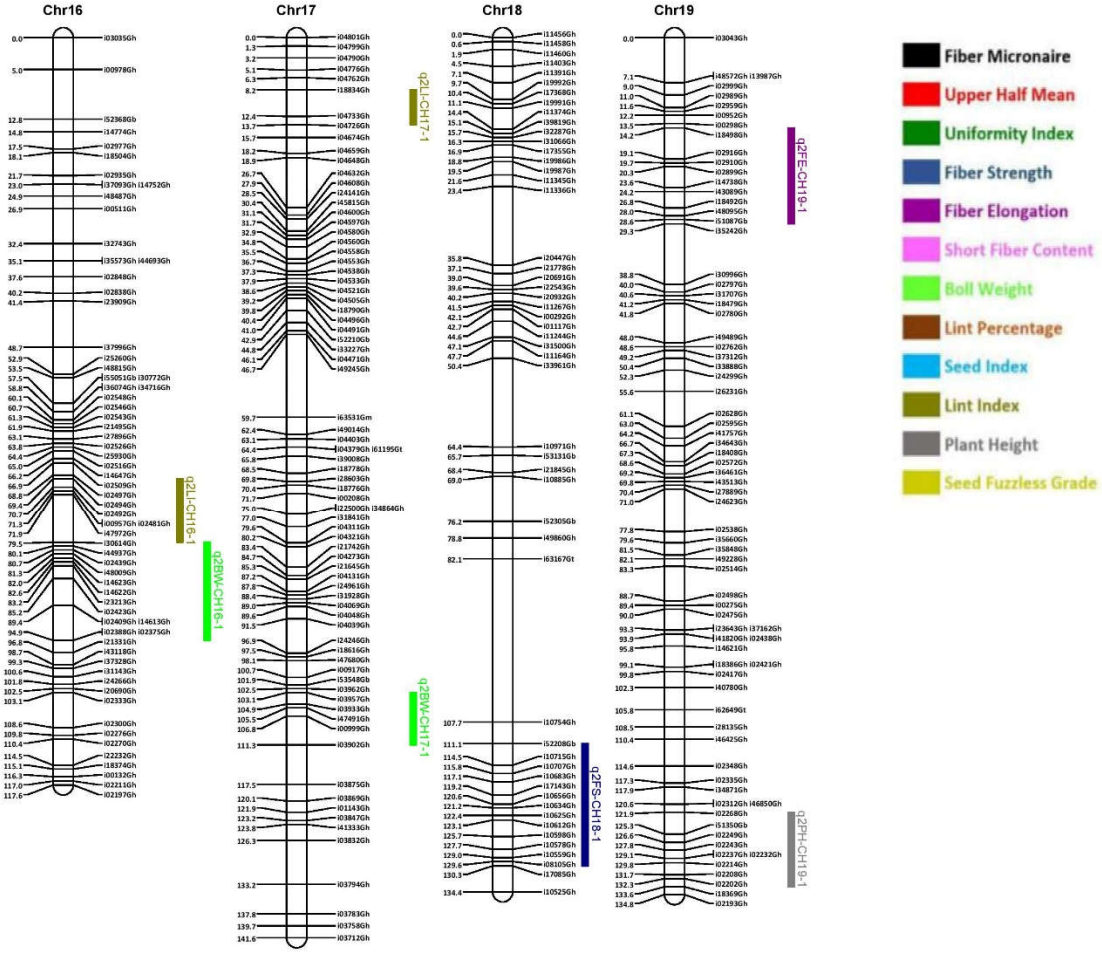

Figure 5. 5 Linkage maps of chromosomes Chr16(A1), Chr17(D8), Chr18(A6), Chr19(D1), along with detected QTLs for NC05AZ06 × NC11-2100 RILs population.

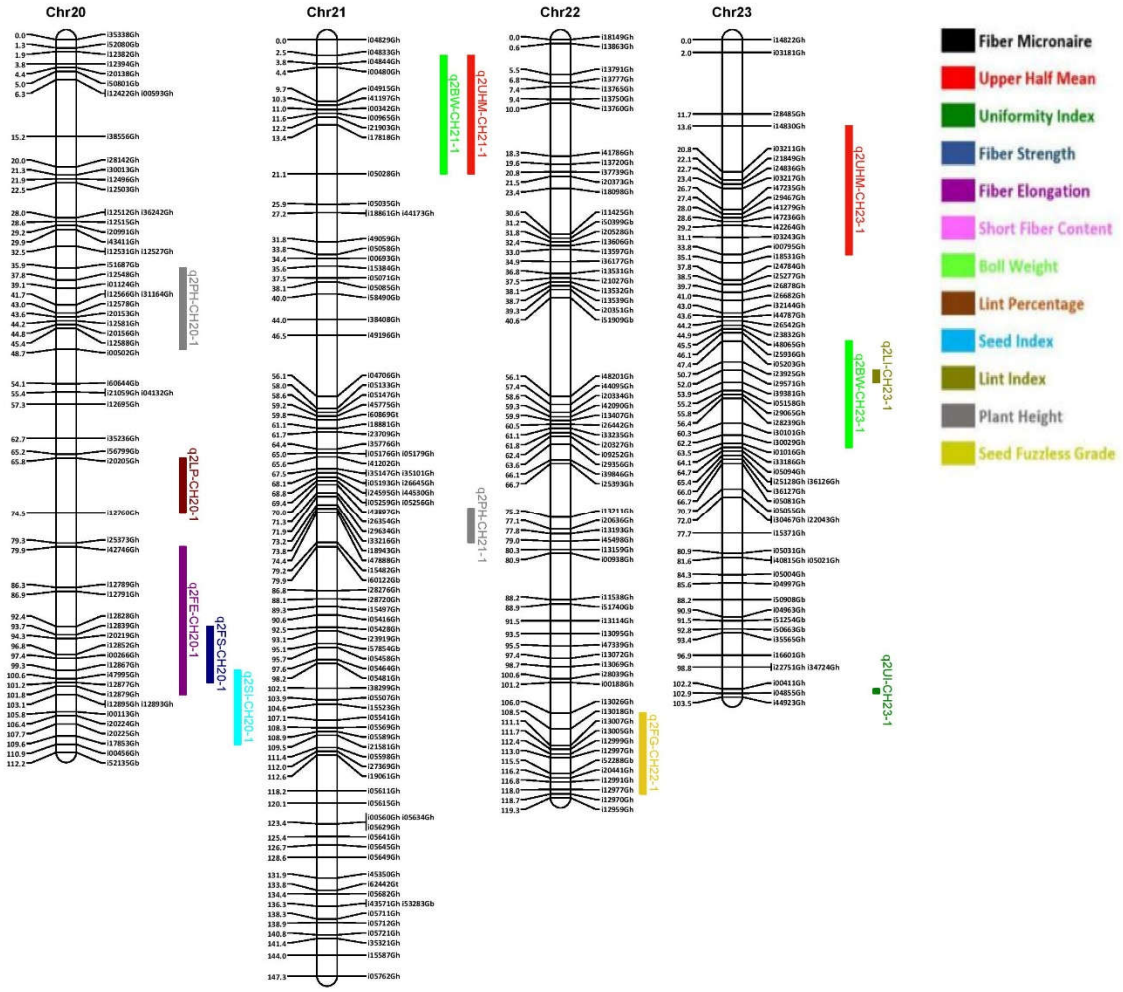

Figure 5. 6 Linkage maps of chromosomes Chr20(D4), Chr21(D2), Chr22(D13), Chr23(A2), along with detected QTLs for NC05AZ06 × NC11-2100 RILs population.

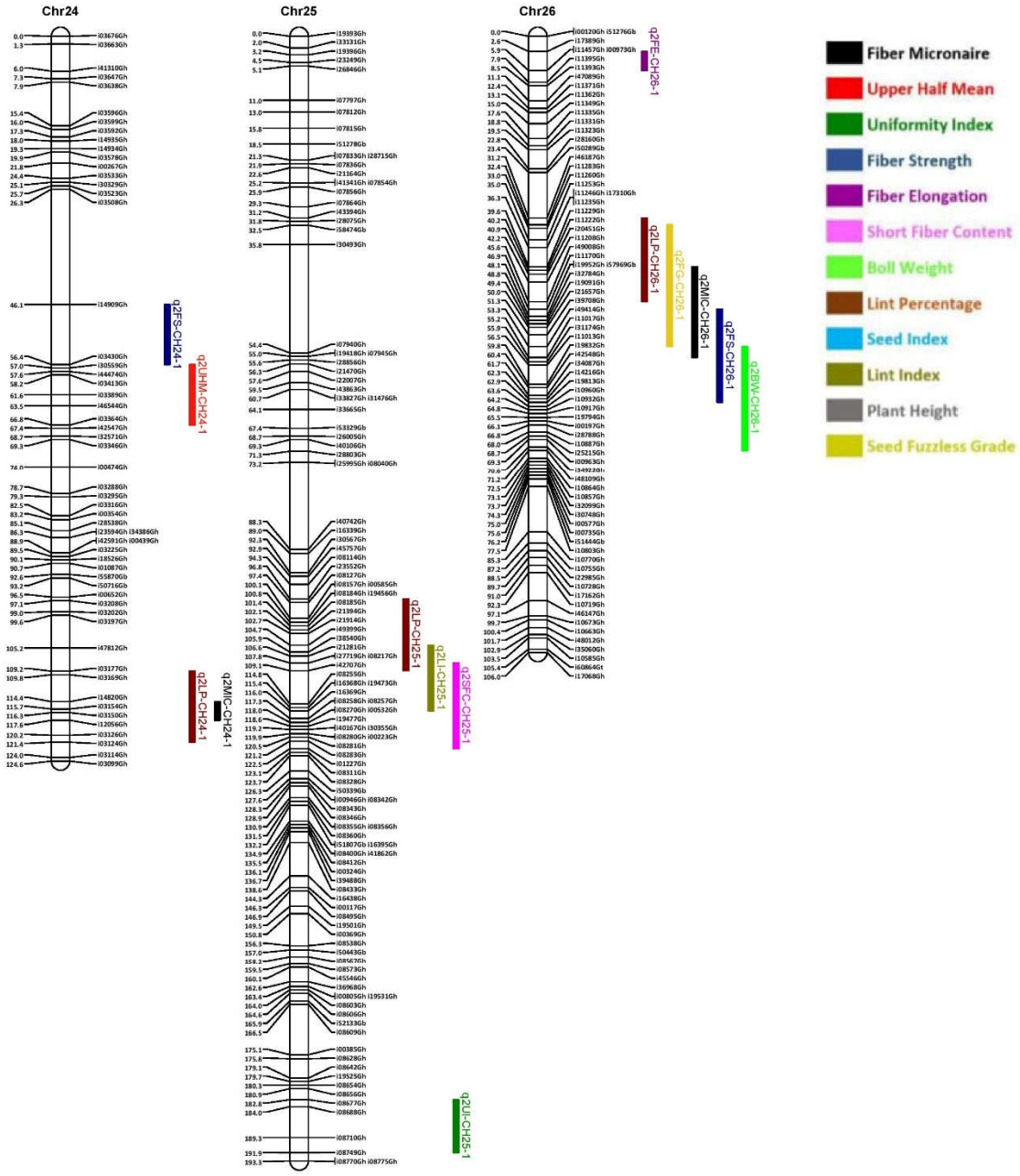

Figure 5. 7 Linkage maps of chromosomes Chr24(D3), Chr25(D12), Chr26(D6), along with detected QTLs for NC05AZ06 × NC11-2100 RILs population.

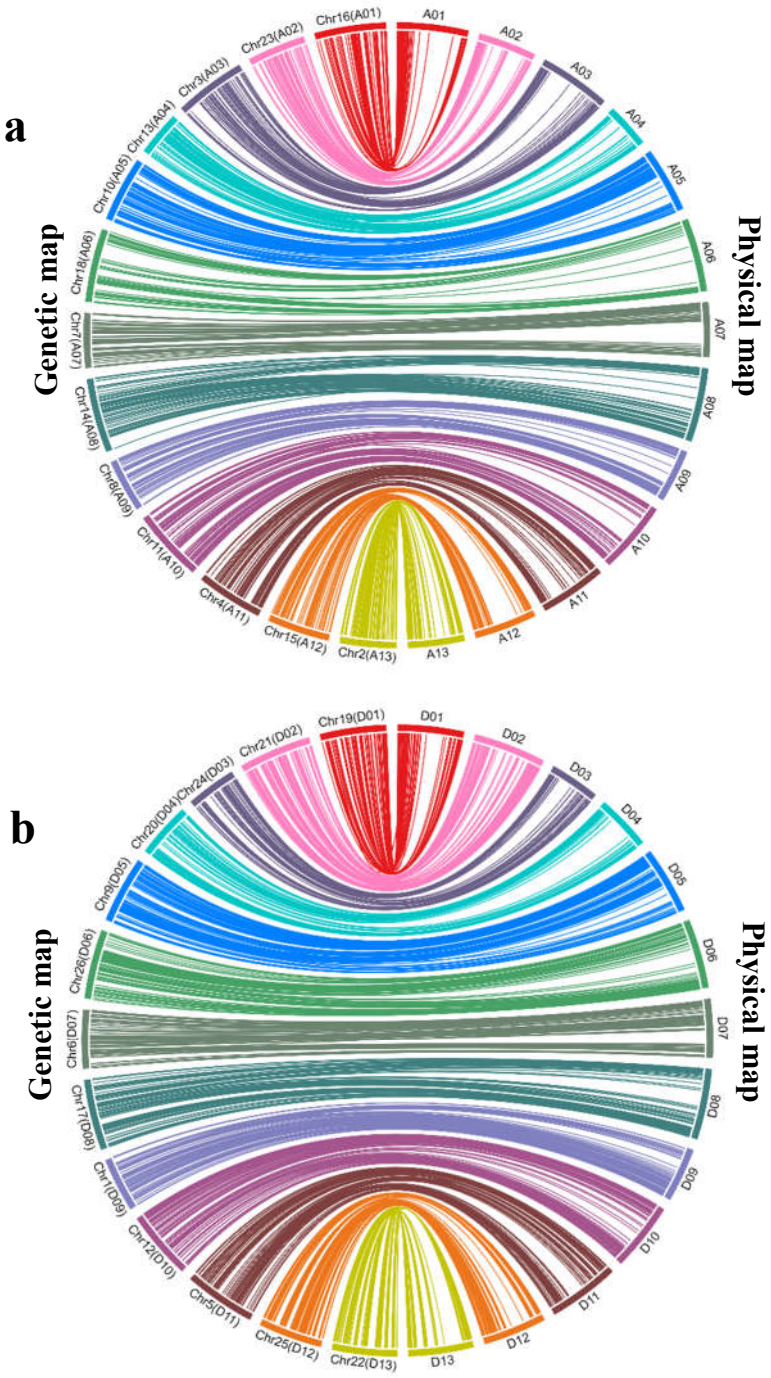

Figure 5. 8 Collinearity between genetic map and physical map in A subgenome and D subgenome for NC05AZ06 × NC11-2100 RILs population. (a) Collinearity for chrmosomes in A subgenome. (b) Collinearity for chrmosomes in D subgenome

CHAPTER 6: Discussion

Construction of high-density linkage map with SNP arrays

The limited quantity of the polymorphic markers selected for the studies were often limitations for the construction of high-density linkage maps in cotton [41, 42]. Due to the lack of the marker polymorphism in cotton, the linkage maps built by second-generation molecular markers such as SSR and AFLP, usually carried some disadvantages viz low coverage to the total cotton genome, poor marker density and large gaps [42-45]. SNPs, as the new generation DNA markers, provided abundant genetic variation and their loci distribute evenly along the whole genome. Hence, these have been becoming the most reliable markers for building high-density linkage maps and have been widely used in the QTL studies [38, 46, 47]. Recently, two sets of cotton SNPs array: CottonSNP63K and CottonSNP80k were developed and applied to the QTL research [38,46]. In the past three years, several high-density cotton genetic maps constructed successfully using the SNPs from cotton SNP arrays were reported [36, 47-51]. In the current study, the linkage maps of two upland cotton RILs populations were formed using SNPs selected from CottonSNP63K array.

For the RILs population NC05AZ06 × NC11-2091 (Named as “pop.2091”), It spanned a total genetic length of 4422.44 cM (Table 4.4), which was in correspondence with the estimated size of tetraploid cotton genome (4500 cM) [52]. The average marker density of the map was 1.41 cM and no large gap (>15 cM) was found which were significantly better than the SSR-based linkage maps in previous studies [27, 28, 30, 43-45, 53]. We identified 11,255 (17.8%) polymorphic SNPs between the parents from the 63,058 SNP markers and 3,129 (5.0%) of the polymorphic SNPs were unique. Based on the previous studies, the polymorphism rate of SSRs and SNPs for a RIL population was around 3% - 10% [43-50]. It meant that SNP markers had

the similar efficiency but much greater quantity than the SSR markers. Out of total 3129 SNP markers, 1534 SNPs were mapped in A sub-genome and 1595 SNPs were mapped in D sub-genome which showed a balanced distribution between the two sub-genomes. Genetic map lengths produced by these markers in A and D sub-genomes are 2236.35 cM and 2186.09 cM, respectively.

In the population NC05AZ06 × NC11-2100 (Named as "pop.2100"), we identified 8,407 (13.3%) polymorphic SNPs between the parents from the 63,058 SNP markers and 2,017 (3.2%) of the polymorphic SNPs were unique (Table 5.4). The genetic map spanned a total genetic length of 3837.46 cM, which was less than the estimated size of tetraploid cotton genome (4500 cM) [52]. The average marker density of the map was 1.9 cM and 31 gaps (>10 cM) were found. Among these 2017 SNP markers, 954 SNPs were mapped in A sub-genome and 1063 SNPs were mapped in D sub-genome, spanning 1974.73 cM and 1862.73 cM lengths of genetic map in A and D sub-genomes respectively.

By comparing the linkage maps of two populations: NC05AZ06 × NC11-2091 (pop.2091) and NC05AZ06 × NC11-2100 (pop.2100), we find that the quality of the linkage map for pop.2091 is much better than that in pop.2100, with higher marker density (1.41cM vs 1.9 cM), less markers interval gaps (5 gaps vs 31 gaps), and better linkage length (4422.44 cM vs 3837.46 cM). The accuracy of the linkage maps is guaranteed since both these two linkage maps showed good correlation with the physical maps (Table 4.5, Table 5.5, Figure 4.8, Figure 5.8). However, 3129 markers were selected to form the genetic map in pop.2091, while only 2017 SNPs were used in the genetic map of pop.2100. So, the main factor caused the difference of the quality between these two linkage maps is the number of the markers used for construction the maps since much more polymorphic SNPs were detected in the population pop.2091 than those

in the population pop.2100. Both circos plots confirms that the polymorphic SNPs detected in each of the chromosomes distributed unevenly (Figure 4.8, Figure 5.8), which suggests the point that the SNPs showed uneven distribution of polymorphism-rich and polymorphism-poor regions along each chromosome [54].

Segregation distortion

Segregation distortion (SD), a phenomenon that the genotypic frequencies of a marker against the expected Mendelian segregation ratio, is commonly observed in mapping populations. The similar phenomenon was mentioned in many previous reports [33, 36, 47, 48, 50], which could be influenced by plenty of factors, such as genetic drift, preferential fertilization by particular gametic and some other environmental factors [55-57]. For current study, 5.6% (175) of mapped markers unevenly distributed in two sub-genomes and 22 chromosomes showed segregation distortion in pop.2091 (Table 4.4), and 100 (5%) segregation distorted markers were found among 19 chromosomes in pop.2100 (Table 5.4), the percentage of which were lower than the previous maps of cotton (11.4%-32.8%) [33, 36, 47, 48, 50] . Wang et al. reported that the bigger the genetic difference between two parents was, the smaller the segregation distortion would be in the population [33], which may suggest our results since we chose the parents used in this study for their high genetic variation. In pop.2091, 34 distorted markers were observed in Chr.02 (A13) with a highest rate of distortion (25.37%), forming three SDRs (Table 4.4). 59% (103) of distorted markers were distributed on 6 chromosomes (A3, A10, A12, A13, D5, D10), which was consistent with the previous reports that the majority distorted markers were concentrated in a few chromosomes [33, 36, 47, 48, 50, 55-57]. In pop.2100, 52% (52) of distorted markers were mapped on 4 chromosomes (A13, D2, D8, D10) (Table 5.4). Same as the result in pop.2091, the highest rate of distortion (20%) was also observed in

chromosome A13 in pop.2100, with 16 distorted markers detected. Since these two populations both have a same parent NC05AZ06, the high rate of marker segregation distortion in A13 may cause by the preferential fertilization by the gametic from parent NC05AZ06. Usually the effects of segregation distortion can be ignored in the QTLs mapping since this influence is not significant [58, 59].

QTL mapping population

As we discussed before, the quality of a genetic map is directly related to the number of polymorphic markers selected for the map. The use of the cotton SNPs array provided us great numbers of the candidate markers screened for the polymorphic markers' selection. However, the number of candidate makers is not the only essential factor in polymorphic markers selection. The choice of mapping population is also a key factor. In order to potentially detect a broader array of polymorphic alleles, the parents of mapping population should have difference genetic backgrounds and high genetic variation which leads to the highly opposite traits performances on these two parents [60, 61]. Therefore, many mapping populations derived from the crosses between *G. hirsutum* and *G. barbadense* are reported in previous QTL studies [24, 31, 32, 53, 62, 63], since *G. hirsutum* is famous for its high yield performance while *G.barbadense* has much superior fiber quality. However, Said et al (2015) [15] reported that the QTL type and placement from a interspecific (*G.hirsutum* × *G.barbadense*) population was highly noncoincident with the *G.hirsutum* population, which may cause inconvenience to further marker-assisted selection (MAS) breeding on *G.hirsutum* cultivars. In this study, instead of an interspecific population, two intraspecific RILs populations were developed from a cross between a high fiber quality and yield line NC05AZ06 (*G.hirsutum*) and two landrace accessions NC11-2091(*G.hirsutum*), NC11-2100 (*G.hirsutum*) with poor yield and fiber quality.

QTLs with favorable alleles identification

For the RILs population NC05AZ06 × NC11-2091 (pop.2091), the performance of parent NC05AZ06 was superior to those of parent NC11-2091 for MIC, UHM, UI, FS and 4 yield traits (Table 4.1). Among the 106 QTLs, the favorable alleles of 80 QTLs originated from NC05AZ06 while other 26 from NC11-2091 (Appendix B). It was reasonable since there is a high chance that the QTLs with favorable alleles were contributed by the parents with better phenotypic performance for a given trait [64]. Only few of the QTLs with favorable alleles of a give trait were derived from the parents with significantly bad performance. However, these few QTLs would be very useful for future breeding program if we want to improve the cotton cultivar based on our NC05AZ06 line. For example, NC05AZ06 has a significant better performance on fiber micronaire (MIC) than NC11-2091. But the favorable alleles of qMIC-CH10-1, which was a stable major QTL of MIC, was contributed by NC11-2091. Therefore, the MIC of NC05AZ06 could theoretically be improved if the alleles of qMIC-CH10-1 on NC05AZ06 changed to that favorable alleles from NC11-2091. 15 QTLs that could improve the fiber quality traits on NC05AZ06 were identified and 4 QTLs that could improve the yield traits on NC05AZ06 were detected. Totally 9 of them are major QTLs: qMIC-CH10-1, qUHM-16-CH5-1, qUI-CH3-1, qFS-CH19-1, qFS-CH25-1, qFE-CH4-1, qFE-17-CH23-1, qSFC-CH3-1 and qSI-CH15-1.

In the RILs population NC05AZ06 × NC11-2100 (pop.2100), the performance of parent NC05AZ06 was superior to those of parent NC11-2100 for MIC, UHM, UI, FS, FE, BW, LP and LI (Table 5.1). Out of the 77 QTLs, the favorable alleles of 51 QTLs originated from NC05AZ06 while other 26 from NC11-2100 (Appendix C). 13 QTLs that could improve the fiber quality traits on NC05AZ06 were identified and 9 QTLs that could improve the yield traits on

NC05AZ06 were detected. Totally 7 of them are major QTLs: q2UHM-CH1-1, q2FS-CH14-1, q2FS-CH26-1, q2FE-CH6-1, q2SFC-CH5-1, q2LP-CH25-1, q2SI-CH20-1.

QTL location

Based on the reports from previous cotton QTL studies (http://www2.cottonqtldb.org:8081/references) [40], the QTLs for fiber quality traits and yield traits were widely distributed on most chromosomes, varying from population to population (Appendix A).

Of the 44 major QTLs for the 11 traits in the the RILs population NC05AZ06 × NC11-2091 (pop.2091), eleven QTLs are unique QTLs: 2 for FE (qFE-CH4-1, qFE-CH19-1), 2 for UHML (qUHM-17-CH23-1, qUHM-16-CH24-1), 1 for SFC (qSFC-17-CH2-1), 2 for BW (qBW-17-CH4-1, qBW-16-CH4-1), 2 for LI (qLI-17-CH24-1, qLI-16-CH12-1), 1 for LP (qLP-CH25-1) and 1 for PH (qPH-17-CH9-1) (Appendix A, Appendix B). The presence of the unique QTLs is reasonable since different populations result in different QTLs, and the considerable size of the SNP markers used in this study can help discovering more QTLs than those in previous studies. Out of these 11 unique major QTLs, 3 were stable QTLs (qFE-CH4-1, qFE-CH19-1, qLP-CH25-1), which should be a good addition to the previous QTLs database. Most of the QTLs were detected in the chromosomes where had QTLs for the corresponding traits reported before. Only 5 major QTLs were detected in the chromosomes where had no previous QTLs reported for the corresponding traits: qBW-16-CH4-1(A11), qLI-17-CH24-1(D03), qLI-16-CH12-1(D10), qSFC-17-CH2-1(A13), qUHM-17-CH23-1(A2). The previous study reported that there were two seed fuzzless trait loci on chromosomes A12 and D13 [11]. Here we found a seed

fuzzless trait locus qFG-CH22-1 on D13 (Chr.22) coincided with the report, and a new stable locus (qFG-CH25-1) on D12 (Chr.25) which might be the new locus for seed fuzzless trait.

In the RILs population NC05AZ06 × NC11-2100 (pop.2100), among the 36 major QTLs for the 11 traits, ten major QTLs are unique QTLs: 3 for SFC (q2SFC-CH14-1, q2SFC-CH11-1, q2SFC-CH5-1), 1 for FS (q2FS-CH18-1), 1 for UI (q2UI-CH25-1), 2 for LI (q2LI-CH7-1, q2LI-CH25-1), 1 for BW (q2BW-CH26-1), 1 for LP (q2LP-CH25-1) and 1 for SI (q2SI-CH15-1) (Appendix A, Appendix C). Only 2 of these 73 QTLs were detected in the chromosomes where had no previous QTLs reported for the corresponding traits: q2LI-CH23-1 (A2), q2LI-CH9-1 (D5). For the seed fuzzless trait, two new major QTLs q2FG-CH14-1, q2FG-CH26-1 were detected on A8 (Chr.14) and D6 (Chr.26).

Of the 3129 SNPs in pop.2091 and 2017 SNPs in pop.2100, 853 SNPs were identified in both populations. Among these 853 SNPs, 25 marker regions shared QTLs identified in both populations, including 86 SNPs and 90 QTLs (Table 6.1). Out of these 25 marker regions, 12 pairs of the QTLs on the same markers in two populations were detected for the same traits: qLP-17-CH14-1 and q2LP-CH14-1, qUI-CH11-1 and q2UI-CH11-1, qSFC-16-CH11-1 and q2SFC-CH11-1, qSI-17-CH15-2 and q2SI-CH15-1, qPH-17-CH19-1 and q2PH-CH19-1, qUHM-17-CH24-1 and q2UHM-CH24-1, qLP-16-CH24-2 and q2LP-CH24-1, qFE-16-CH20-1 and q2FE-CH20-1, qBW-17-CH17-1 and q2BW-CH17-1, qUHM-17-CH5-1 and q2UHM-CH5-1, qLP-CH25-1 and q2LP-CH25-1, qFG-CH22-1 and q2FG-CH22-1.

QTL clusters

The traits correlation analysis showed some high positive or negative relevance between different traits. Some of these correlations are stable even across the different populations. For example, a highly significant negative correlation between UI and SFC was observed in both

populations (-0.93 and -0.87). This stable relevance can be partially explained by the QTL clusters (Table 4.9 - 4.10, Table 5.9 - 5.10). For example, in pop.2091, Q-1, Q-3, Q-4, Q-5, Q-6 and QY-2 contained multiple QTLs from SFC and UI. However, the signs of additive effects of SFC and UI QTLs in each of these Clusters are opposite with favorable alleles from same parent. In this case, when we choose the favorable alleles of NC05AZ06 for this QTL, SFC will decrease and UI will increase. If we choose the other alleles for the QTL, the UI will decrease and SFC will increase. Same situations were observed in pop.2100. Q2-2, QA2-1, QYA2-1 contained QTLs for SFC and UI, of which the signs of additive effects are all opposite. That explained why UI and SFC would always show negative correlation rate. On the contrary, Y-1, Y-3, Y-4, QY-3, QY-4, Y2-1, Y2-2, QYA2-2 clusters provided the evidences of why most of the yield traits are highly positive correlated since all the favorable alleles of QTLs in these clusters are derived from same parent with additive effects. Similarly, Q-2, Q-7 explained a negative correlation (-0.62 and -0.38) between UHM and FE. Q-7, QYA2-1 explained a negative correlation (-0.79 and -0.24) between UHM and SFC. QY-1, QY-5, QYA-1, QYA-2, QY2-1, QY2-5, QY2-8, QYA2-1, QYA2-2 clusters contained the fiber quality traits and yield traits with the favorable alleles from different parents, partially explained why the fiber quality performance always against the yield performance. However. some of the clusters contained both fiber quality traits and yield traits QTLs with favorable alleles from same parents, which provided us an efficient way to improve the quality traits and yield traits at the same time. For example, QY-2 contained 3 QTLs from BW, UI and SFC, of which the favorable alleles were all derived from NC05AZ06. Therefore, the QTLs in QY-2 can help to improve the BW, UI and SFC at the same time. QY-3 were shared by 3 QTLs from LI, SI and MIC, of which the favorable alleles were all contributed by NC05AZ06. In this case, the QTL markers in QY-3 can help improving the MIC,

LI and SI concurrently. Similarly, the QTLs in QY-4 had the potential to improve the UHM, MIC, LI and LP simultaneously. The QTLs in QY2-2 can help to improve the FE and LI. QY2-4 were shared by 3 QTLs from LI, MIC and SFC with favorable alleles from NC05AZ06, which had the potential to improve the LI, MIC and SFC together. The QTLs in QY2-6 can help improving BW and UHM. The favorable alleles of QTLs for LP, FE and LP, MIC in clusters QY2-3 and QY2-7 were all derived from NC11-2100. These two clusters can help improving the LP, FE and MIC at the same time.

Table

Table 6. 1 The regions of shared SNPs identified as QTLs in both NC05AZ06 × NC11-2091 and NC05AZ06 × NC11-2100 RILs population

Shared markers	Chromosome	QTLs in pop.2091[a]	QTLs in pop.2100
i02388Gh	A1(Chr16)	qFE-16-CH16-1	q2BW-CH16-1
i04855Gh	A2(Chr23)	qUHM-17-CH23-1 qFE-17-CH23-1 qSFC-16-CH23-1	q2UI-CH23-1
i10683Gh-i10578Gh	A6(Chr18)	qSI-16-CH18-1	q2FS-CH18-1
i00531Gh-i02022Gh	A7(Chr7)	qUHM-16-CH7-1 qUI-17-CH7-1 qPH-17-CH7-1	q2LI-CH7-1
i04056Gh-i45346Gh	A8(Chr14)	qLP-17-CH14-1	q2LP-CH14-1 q2FG-CH14-1 q2SFC-CH14-1 q2UI-CH14-1 q2FS-CH14-1
i32404Gh-i04358Gh	A8(Chr14)	qSI-17-CH14-1	q2SFC-CH14-1 q2UI-CH14-1 q2FS-CH14-1 q2UHM-CH14-1
i00538Gh-i49272Gh	A10(Chr11)	qUI-CH11-1 qSFC-16-CH11-1	q2FE-CH11-1 q2SFC-CH11-1 q2UI-CH11-1
i24397Gh	A11(Chr4)	qFS-CH4-1	q2MIC-CH4-1
i08503Gh-i08424Gh	A12(Chr15)	qLP-17-CH15-1 qSI-CH15-1	q2MIC-CH15-1
i20990Gh-i22689Gh	A12(Chr15)	qSI-17-CH15-2	q2SI-CH15-1
i02249Gh-i02214Gh	D1(Chr19)	qPH-17-CH19-1 qFS-CH19-1 qFE-CH19-1	q2PH-CH19-1
i21903Gh	D2(Chr21)	qUI-17-CH21-1 qSFC-CH21-1	q2BW-CH21-1 q2UHM-CH21-1
i15482Gh	D2(Chr21)	qFS-16-CH21-1	q2PH-CH21-1
i14909Gh-i03364Gh	D3(Chr24)	qUHM-17-CH24-1 qLI-17-CH24-1 qLP-CH24-1	q2FS-CH24-1 q2UHM-CH24-1
i03177Gh-i03124Gh	D3(Chr24)	qFS-17-CH24-1 qUHM-16-CH24-1 qLP-16-CH24-2	q2LP-CH24-1 q2MIC-CH24-1

Table 6. 1 (continued).

Shared markers	Chromosome	QTLs in pop.2091[a]	QTLs in pop.2100
i12578Gh-i00502Gh	D4(Chr20)	qFE-17-CH20-1	q2PH-CH20-1
i12852Gh-i47995Gh	D4(Chr20)	qFE-16-CH20-1	q2FE-CH20-1 q2FS-CH20-1 q2SI-CH20-1
i11260Gh-i11229Gh	D6(Chr26)	qUI-17-CH26-1	q2LP-CH26-1 q2FG-CH26-1 q2MIC-CH26-1
i00999Gh-i03902Gh	D8(Chr17)	qBW-17-CH17-1	q2BW-CH17-1
i11908Gh	D10(Chr12)	qLI-16-CH12-1 qBW-17-CH12-1 qSI-CH12-1	q2PH-CH12-1
i21510Gh-i17616Gh	D10(Chr12)	qUI-16-CH12-1	q2MIC-CH12-1
i49210Gh-i21605Gh	D11(Chr5)	qUHM-17-CH5-1	q2PH-CH5-1 q2SFC-CH5-1 q2UI-CH5-1 q2FG-CH5-1 q2UHM-CH5-1
i58446Gb-i07030Gh	D11(Chr5)	qFE-CH5-1	q2LP-CH5-1
i00585Gh-i01227Gh	D12(Chr25)	qLP-CH25-1 qFG-CH25-1 qFS-CH25-1 qMIC-CH25-1	q2LP-CH25-1 q2LI-CH25-1 q2SFC-CH25-1
i13018Gh-i12977Gh	D13(Chr22)	qFG-CH22-1	q2FG-CH22-1

[a] pop.2091: RILs population NC05AZ06 × NC11-2091
pop 2100: RILs population NC05AZ06 × NC11-2100

REFERENCES

1. Fryxell PA. A revised taxonomic interpretation of *Gossypium* L. (Malvaceae). Rheedea 2. 1992; 108–165.
2. Grover CE, Zhu X, Grupp KK, Jareczek JJ, Gallagher JP, et al. Molecular confirmation of species status for the allopolyploid cotton species, *Gossypium ekmanianum* Wittmack. Genet. Resour. Crop Evol. 2014; 62: 103-114.
3. Shim J, Mangat PK, Angeles-Shim RB. Natural variation in wild *Gossypium* species as a tool to broaden the genetic base of cultivated cotton. J. Plant Sci. Curr. 2018; Res 2: 005.
4. U.S. Department of Agriculture. Cotton: World Markets and Trade. USDA Foreign Agricultural Service. 2018. https://apps.fas.usda.gov/psdonline/circulars/cotton.pdf
5. National Cotton Council of America. Overview of the U.S. cotton industry. 2011. https://www.cotton.org/pubs/cottoncounts/upload/Cotton-Industry-Overview_Jan-19-2011.pdf
6. OECD/Food and Agriculture Organization of the United Nations. "Cotton", in OECD-FAO Agricultural Outlook 2018-2027, OECD Publishing, Paris/Food and Agriculture Organization of the United Nations, Rome. 2018. https://doi.org/10.1787/agr_outlook-2018-13-en
7. Shang L, Liu F, Wang Y, Abduwell A, Cai S, et al. Dynamic QTL mapping for plant height in Upland cotton (*Gossypium hirsutum*). Plant Breeding. 2015; 134(6): 703-712.
8. Lei L, Zheng HL, Wang JG, Liu HL, et al. Genetic dissection of rice (*Oryza sativa* L.) tiller, plant height, and grain yield based on QTL mapping and metaanalysis. Euphytica. 2018; 214:109.

9. Su J, Li L, Zhang C, Wang C, Gu L, et al. Genome-wide association study identified genetic variations and candidate genes for plant architecture component traits in Chinese upland cotton. Theor Appl Genet. 2018; 131: 1299-1314.
10. Bechere E, Turley RB, Auld DL, Zeng L. A New Fuzzless Seed Locus in an Upland Cotton (*Gossypium hirsutum* L.) Mutant. American Journal of Plant Sciences. 2012; 3: 799-804.
11. Turley RB and Kloth RH. Identification of a Third Fuzzless Seed Locus in Upland Cotton (*Gossypium hirsutum* L.). The Journal of Heredity 2002; 93(5).
12. Fang DD. Cotton Fiber Genes and Stable Quantitative Trait Loci. Cotton Fiber: Physics, Chemistry and Biology. 2018; 151-178.
13. Yao Y, Zhang B, Dong CJ, Du Y, Jiang L, Liu JY. Comparative proteomic and biochemical analyses reveal different molecular events occurring in the process of fiber initiation between wild-type Allotetraploid cotton and its fuzzless-lintless mutant. Plos One. 2014;10(2): e0117049.
14. Said J, Lin Z, Zhang X, Song M and Zhang J. A comprehensive meta QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. BMC Genomics. 2013; 14(1): 776.
15. Said J, Song M, Wang H, Lin Z, Zhang X et al. A comparative meta-analysis of QTL between intraspecific *Gossypium hirsutum* and interspecific *G. hirsutum* × *G. barbadense* populations. Mol Genet Genomics. 2015; 290:1003–1025.
16. Lacape JM, Llewellyn D, Jacobs J, Arioli T, et al. Research article Meta-analysis of cotton fiber quality QTLs across diverse environments in a *Gossypium hirsutum* x *G. barbadense* RIL population. BMC Plant Biology. 2010; 10:132.

17. Smith CW, and Coyle GG. Association of Fiber Quality Parameters and Within-Boll Yield Components in Upland Cotton. Crop Sci. 1997; 37:1775-1779.

18. Bourland M. History of Cotton Breeding and Genetics at the University of Arkansas. The Journal of Cotton Science. 2018; 22:171-182.

19. Zhang J. History and Progress in Cotton Breeding, Genetics, and Genomics in New Mexico. The Journal of Cotton Science. 2018; 22:191-210.

20. Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, et al. Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding. 1997; 3:87-103.

21. Boopathi NM, Thiyagu K, Urbi B, Santhoshkumar M, et al. Marker-Assisted Breeding as Next-Generation Strategy for Genetic Improvement of Productivity and Quality: Can It Be Realized in Cotton? International Journal of Plant Genomics. 2011; 2011: 16.

22. Ribaut JM and Hoisington D. Marker-assisted selection: new tools and strategies. Trends in Plant Science. 1998; 3(6), 236-238.

23. Meredith WR Jr. Contributions of introductions to cotton improvement, in "Use of plant introductions in cultivar development, Part 1," Shands HL and Wiesner LE eds. Crop Science Society of America. Madison, WI. 1991; 127–146.

24. Reinisch AJ, Dong J, Brubaker CL, Stelly DM, Wendel JF and Paterson AH. A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics. 1994; 138(3):829–847.

25. Multani DS, Lyon BR. Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome. 1995; 38(5):1005–1008.

26. Lu H, Myers G. Genetic relationships and discrimination of ten influential Upland cotton varieties using RAPD markers. Theor Appl Genet. 2002; 105(2-3): 325-331.

27. Lin Z, He D, Zhang X, Nie Y, Guo X, et al. Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. Plant Breeding. 2005; 124: 180-187.

28. Noormohammadi Z, Hasheminejad-Ahangarani FY, Sheidai M, Ghasemzadeh-Baraki S, and Alishah O. Genetic diversity analysis in Opal cotton hybrids based on SSR, ISSR, and RAPD markers. Genetics and Molecular Research. 2013; 12(1): 256-269.

29. Abdalla AM, Reddy OUK, El-Zik KM and Pepper AE. Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theor Appl Genet. 2001;102(2-3): 222-229.

30. Yu J, Yu S, Lu C, Wang W, Fan S, et al. High-density Linkage Map of Cultivated Allotetraploid Cotton Based on SSR, TRAP, SRAP and AFLP Markers. Journal of Integrative Plant Biology. 2007; 49(5): 716-724.

31. Yu J, Zhang K, Li S, Yu S, Zhai H, et al. Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* × *Gossypium barbadense* backcross inbred line population. THEOR APPL GENET. 2013; 126(1): 275-287.

32. Yu JZ, Ulloa M, Hoffman SM, Kohel RJ, Pepper AE, et al. Mapping genomic loci for cotton plant architecture, yield components, and fiber properties in an interspecific (*Gossypium hirsutum* L. × *G. barbadense* L.) RIL population. Mol Genet Genomics. 2014; 289(6): 1347-1367.

33. Wang H, Huang C, Guo H, Li X, Zhao W, Dai B, et al. QTL Mapping for Fiber and Yield Traits in Upland Cotton under Multiple Environments. PLOS ONE. 2015;10(6): e0130742.

34. Liu B, and Wendel FJ. Intersimple sequence repeat (ISSR) polymorphisms as a genetic marker system in cotton. Molecular Ecology Notes. 2001; 1: 205–208.

35. Su J, Pang C, Wei H, Li L, Liang B, et al. Identification of favorable SNP alleles and candidate genes for traits related to early maturity via GWAS in upland cotton. BMC Genomics. 2016; 17:687.

36. Kumar NVM, Katageri IS, Gowda SA, Adiger S, Yadava SK et al. 63K SNP chip based linkage mapping and QTL analysis for fiber quality and yield component traits in *Gossypium barbadense* L. cotton. Euphytica. 2019; 215: 6.

37. Kumar S, Banks TW and Cloutier S. SNP Discovery through Next-Generation Sequencing and Its Applications. Int J Plant Genomics. 2012; 2012: 831460.

38. Hulse-Kemp, AM, Lemm J, Plieske J, Ashrafi H, Buyyarapu R, et al. Development of a 63K SNP array for cotton and high-density mapping of intraspecific and interspecific populations of *Gossypium spp.* G3 (Bethesda). 2015; 5(6): 1187–1209.

39. Zhang T, Hu Y, Jiang W, Fang L, Guan X, Chen J, et al. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nat. Biotechnol. 2015; 33(5): 531–537.

40. Said JI, Knapka JA, Song M, and Zhang J. Cotton QTLdb: a cotton QTL database for QTL analysis, visualization, and comparison between *Gossypium hirsutum* and *G. hirsutum*× *G. barbadense* populations. Mol. Genet. Genomics. 2015; 290(4):1615-1625.

41. Iqbal MJ, Reddy OUK, El-Zik KM, Pepper AE. A genetic bottleneck in the evolution under domestication of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. Theor Appl Genet. 2001; 103: 547–554.

42. Wang B, Guo w, Zhu X, Wu Y, Huang N, Zhang T. QTL mapping of fiber quality in an elite hybrid derived-RIL population of upland cotton. Euphytica. 2006; 152: 367-378.

43. Yang X, Zhou X, Wang X, Li Z, Zhang Y, et al. Mapping QTL for cotton fiber quality traits using simple sequence repeat markers, conserved intron-scanning primers, and transcript-derived fragments. Euphytica. 2015; 201: 215-230.

44. Tang S, Teng Z, Zhai T, Fang X, Liu F, et al. Construction of genetic map and QTL analysis of fiber quality traits for Upland cotton (*Gossypium hirsutum* L.). Euphytica. 2015; 201: 195-213.

45. Huang C, Shen C, Wen T, Gao B, Zhu D, et al. SSR-based association mapping of fiber quality in upland cotton using an eight-way MAGIC population. Mol Genet Genomics. 2018; 293(4): 793-805.

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

47. Liu R, Gong J, Xiao X, Zhang Z, Li J, et al. GWAS Analysis and QTL Identification of Fiber Quality Traits and Yield Components in Upland Cotton Using Enriched High-Density SNP Markers. Front. Plant Sci. 2018; 9:1067.

48. Li C, Dong Y, Zhao T, Li L, Li C, et al. Genome-Wide SNP Linkage Mapping and QTL Analysis for Fiber Quality and Yield Traits in the Upland Cotton Recombinant Inbred Lines Population. Front. Plant Sci. 2016; 7: 1356.

49. Ma L, Zhao Y, Wang Y, Shang L and Hua J. QTLs Analysis and Validation for Fiber Quality Traits Using Maternal Backcross Population in Upland Cotton. Front. Plant Sci. 2017; 8:2168.

50. Tan Z, Zhang Z, Sun X, Li Q, Sun Y, et al. Genetic Map Construction and Fiber Quality QTL Mapping Using the CottonSNP80K Array in Upland Cotton. Front. Plant Sci. 2018; 9:225.

51. Li C, Zhao T, Yu H, Li C, Deng X, et al. Genetic basis of heterosis for yield and yield components explored by QTL mapping across four genetic populations in upland cotton. BMC genomics. 2018; 19:910.

52. Rong JK, Abbey C, Bowers JE, Brubaker CL, Chang C, et al. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (*Gossypium*). Genetics. 2004; 166: 389–417.

53. Lacape JM, Nguyen TB, Thibivilliers S, Bojinov B, et al. A combined RFLP–SSR–AFLP map of tetraploid cotton based on a *Gossypium hirsutum* × *Gossypium barbadense* backcross population. 2003; 46(4): 612-626.

54. Nasu S, Suzuki J, Ohta R, Hasegawa K, et al. Search for and Analysis of Single Nucleotide Polymorphisms (SNPs) in Rice (Oryza sativa, Oryza rufipogon) and Establishment of SNP Markers. DNA Research. 2002; 9:163-171.

55. Zhang Z, Hu M, Zhang J, Liu D, Zheng J, et al. Construction of a comprehensive PCR-based marker linkage map and QTL mapping for fiber quality traits in upland cotton (*Gossypium hirsutum* L.). Mol Breeding. 2009;24(1):49–61.

56. Shen X, Guo W, Lu Q, Zhu X, Yuan Y, Zhang T. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in upland cotton. Euphytica. 2007; 155: 371–380.

57. Kumar S, Gill BS, Faris JD. Identification and characterization of segregation distortion loci along chromosome 5B in tetraploid wheat. Mol Genet Genomics. 2007; 278(2):187-196.

58. Zhu C, Wang F, Wang J, Li G, Zhang H, Zhang Y. Reconstruction of linkage maps in the distorted segregation populations of backcross, doubled haploid and recombinant inbred lines. Chinese Science Bulletin. 2007; 52(12): 1648-1653.

59. Zhang L, Wang S, Li H, Deng Q, Zheng A, Li S, Li P, Li Z, Wang J. Effects of missing marker and segregation distortion on QTL mapping in F2 populations. Theor Appl Genet. 2010;121(6):1071–1082.

60. Zhang Z, Rong J, Waghmare VN, Chee PW, et al. QTL alleles for improved fiber quality from a wild Hawaiian cotton, *Gossypium tomentosum*. Theor Appl Genet. 2011; 123:1075–1088.

61. Baytar A, Peynircioglu C, Sezener V, Basal H, Frary A, et al. Identification of stable QTLs for fiber quality and plant structure in Upland cotton (*G. hirsutum* L.) under drought stress. IND CROP PROD. 2018; 124: 776-786.

62. Song X, Wang K, Guo W, Zhang J, Zhang T. A comparison of genetic maps constructed from haploid and BC1 mapping populations from the same crossing between *Gossypium hirsutum* L. and *Gossypium barbadense* L. Genome. 2005; 48(3):378-390

63. Paterson AH, Saranga Y, Menz M, Jiang C, Wright RJ. QTL analysis of genotype × environment interactions affecting cotton fiber quality. Theor Appl Genet. 2003; 106: 384–396.

64. Shao Q, Zhang F, Tang S, Liu Y, Fang X, et al. Identifying QTL for fiber quality traits with three upland cotton (*Gossypium hirsutum* L.) populations. Euphytica.2014; 198:43–58.

65. Andres RJ, Bowman DT, Lawrence KS, Myers G, Chee PW, et al. Effect of Leaf Shape on Boll Rot Incidence in Upland Cotton (*Gossypium hirsutum*). International Journal of Plant Breeding and Genetics. 2013; 7: 132-138.

66. Ooijen JW. JoinMap 4: Software for the Calculation of Genetic Linkage Maps in Experimental Populations. Kyazma BV, Wageningen. 2006.

67. Kosambi DD. The estimation of map distances from recombination values. Ann. Hum. Genet. 1943; 12: 172–175.

68. Paillard S, Schnurbusch T, Winzeler M, Messmer M, Sourdille P, Abderhalden O, et al. An integrative genetic linkage map of winter wheat (*Triticum aestivum* L.). Theor Appl Genet. 2003; 107: 1235–1242.

69. Zeng ZB. Precision mapping of quantitative trait loci. Genetics. 1994; 136(4): 1457-1468.

70. Wang S, Basten CJ, and Zeng ZB. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. 2012. http://statgen.ncsu.edu/qtlcart/WQTLCart.htm

71. Voorrips RE. MapChart: Software for the graphical presentation of linkage maps and QTLs. The Journal of Heredity. 2002; 93(1): 77-78.

APPENDICES

Appendix A

The summary of 2884 previously reported QTLs for 11 traits in 26 chromosomes

	MIC	UHM	UI	FS	FE	SFC	BW	LP	SI	LI	PH	Total
A1	18	27	3	9	7	1	1	7	10	3	3	89
A2	10	6	11	5	8	5	5	6	1	0	1	58
A3	12	23	6	10	4	2	3	25	8	1	4	98
A4	9	16	7	10	7	1	2	7	0	1	0	60
A5	35	28	14	21	19	0	11	11	7	3	11	160
A6	18	15	12	11	11	4	10	9	1	0	1	92
A7	13	26	12	34	12	4	5	25	23	1	1	156
A8	14	37	10	20	4	1	0	5	4	0	6	101
A9	22	12	11	17	12	0	2	13	6	2	2	99
A10	11	18	12	10	13	3	8	8	5	3	4	95
A11	6	15	6	20	6	2	5	6	4	2	7	79
A12	18	16	21	14	11	6	7	17	4	2	0	116
A13	11	9	10	10	7	0	5	39	4	2	5	102
A sub	197	248	135	191	121	29	64	178	77	20	45	1305
D1	17	23	22	26	18	4	6	9	8	4	14	151
D2	23	16	8	12	20	1	4	11	6	2	3	106
D3	24	15	9	23	7	2	5	36	1	0	2	124
D4	17	10	14	11	3	3	4	16	6	2	3	89
D5	13	21	11	25	9	0	15	4	4	0	1	103
D6	24	36	16	17	24	6	2	11	4	2	6	148
D7	14	12	10	7	6	1	0	2	1	1	1	55
D8	16	32	8	24	14	2	14	8	5	2	5	130
D9	6	5	3	5	8	0	18	13	19	2	11	90
D10	13	22	12	36	8	0	1	13	5	0	4	114
D11	28	26	21	42	22	5	3	13	6	0	6	172
D12	30	33	10	43	22	2	34	2	2	4	14	196
D13	20	25	10	8	5	3	6	11	3	3	7	101
D sub	245	276	154	279	166	29	112	149	70	22	77	1579
Total	442	524	289	470	287	58	176	327	147	42	122	2884

Appendix B

QTLs of fiber quality traits, yield traits and agronomic traits identified in the RILs population NC05AZ06 × NC11-2091

Trait[a]	QTL	Year	Range	Peak	LOD	PVE[b]	AE[c]
MIC	qMIC-CH10-1*	16	233.43-252.01	244.08	11.4	16.2	-0.25
		17	233.43-252.01	243.58	9.1	16.2	-0.207
	qMIC-CH24-1*	16	71.47-87.7	81.76	16.3	25.8	0.335
		17	68.4-87.7	78.81	12.5	23	0.256
	qMIC-CH25-1*	16	102.67-109.11	104.67	3.3	4.1	0.128
		17	92.32-112.58	102.67	6.1	10	0.167
	qMIC-16-CH3-1	16	78.61-95.32	88.82	11.7	17.2	0.344
	qMIC-16-CH6-1	16	0-15.09	3	7	19.3	0.274
	qMIC-16-CH7-1	16	118.91-132.42	126.63	5	6.4	0.158
	qMIC-17-CH25-2	17	58.93-59.91	59.91	3.4	5.2	0.117
UHM	qUHM-16-CH23-1	16	111.44-130.21	120.46	5	8.2	0.024
	qUHM-17-CH23-1	17	132.16-142.33	142.33	6.6	10.2	0.028
	qUHM-16-CH5-1	16	0-4.58	0	6.5	11.2	-0.028
	qUHM-17-CH5-1	17	74.41-95.71	87.3	5.8	8.8	0.026
	qUHM-16-CH7-1	16	43.66-65.19	55.42	6.9	12.1	0.029
	qUHM-17-CH7-1	17	17.16-35.23	22.56	6.1	10.1	0.027
	qUHM-16-CH8-1	16	143.62-150.29	150.29	4.3	7.8	0.023
	qUHM-17-CH21-1	17	158.54-171.02	171.02	3.8	5.5	0.02
	qUHM-17-CH24-1	17	46.58-68.4	60.96	4.8	7.5	0.024
	qUHM-16-CH24-1	16	123.73-142.8	134.44	6.7	11.9	0.028
UI	qUI-CH3-1*	16	46.82-68.03	58.14	10.4	21	-0.921
		17	46.82-68.03	55.53	3.5	6	-0.386
	qUI-CH11-1*	16	30.34-52.37	42.28	8.4	16.1	0.9
		17	46-50.81	46	2.8	4.9	0.362
	qUI-16-CH4-1	16	207.88-227.03	220.88	7.4	13	0.72
	qUI-17-CH7-1	17	47.29-65.19	61.65	5.4	8.9	0.459
	qUI-16-CH10-1	16	11.59-19.17	17.67	7.2	12.7	0.732
	qUI-16-CH11-2	16	9.49-18.49	14.65	4.5	7.5	-0.614
	qUI-16-CH12-1	16	152.39-163.74	156.51	4.1	7.2	0.548
	qUI-17-CH21-1	17	6.05-24.36	17.12	6.1	10	0.488
	qUI-17-CH22-1	17	67.73-79.76	77.25	5.7	9.5	0.483
	qUI-17-CH26-1	17	28.88-46.13	39.93	7	13.1	-0.596
FS	qFS-CH2-1*	16	107.73-118.93	112.84	4.9	7.3	0.639
		17	101.54-118.93	111.34	7.2	11.9	0.767
	qFS-CH4-1*	16	1.53-11.28	4.22	3.9	6.2	0.573
		17	0-8.64	8.64	3	4.1	0.425
	qFS-CH19-1*	16	121.23-124.72	124.72	3.6	5.2	-0.519
		17	112.85-133.39	121.23	7.3	11.8	-0.726
	qFS-CH25-1*	16	98.22-115.15	108.63	9.4	15.6	-0.936

Appendix B (continued)

		17	101.68-121.15	110.59	7.3	11.5	-0.701
	qFS-16-CH10-1	16	132.81-152.55	142.57	5.4	8.4	0.678
	qFS-17-CH13-1	17	9.58-15.12	15.12	3.7	5.6	0.502
	qFS-17-CH14-1	17	8.8-28.94	21.46	7.5	11.9	0.736
	qFS-17-CH21-1	17	31.07-45.91	39.35	5.4	8.2	0.609
	qFS-16-CH21-1	16	93-113.32	103.43	5.2	7.9	0.678
	qFS-16-CH22-1	16	52.69-63.22	56.26	3.7	5.8	-0.553
	qFS-17-CH24-1	17	119.17-125.91	119.17	3.5	5.8	-0.502
FE	qFE-CH4-1*	16	156.02-172.58	167.72	4.7	7.1	-0.385
		17	156.02-181.46	172.58	8.7	13.5	-0.562
	qFE-CH5-1*	16	129.33-143.65	129.33	5	9.4	-0.452
		17	138.14-152.07	148.56	5.1	7.3	-0.41
	qFE-CH8-1*	16	58.29-74.89	68.12	7.7	12.3	0.507
		17	65.15-85.78	71.03	5.3	7.5	0.437
	qFE-CH19-1*	16	124.72-137.87	136.4	7.7	12.4	0.499
		17	126.74-137.87	136.4	5.6	8.1	0.428
	qFE-16-CH7-1	16	17.16-27.41	27.41	4.5	6.7	-0.373
	qFE-16-CH16-1	16	113.05-131.94	119.54	4.6	7.4	0.384
	qFE-17-CH20-1	17	44.97-56.56	49.96	4.1	5.8	0.366
	qFE-16-CH20-1	16	111-121.92	121.92	3.8	5.7	-0.342
	qFE-17-CH23-1	17	132.16-142.33	142.33	7	11.2	-0.503
SFC	qSFC-CH3-1*	16	44.13-59.15	58.64	3.9	7.9	0.664
		17	46.82	46.82	2.5	18.4	0.807
	qSFC-CH21-1*	16	16.12-17.12	16.12	2.8	4.9	-0.507
		17	11.56-18.62	17.62	4.3	7	-0.493
	qSFC-17-CH2-1	17	36.99	36.99	3.6	20.6	-0.837
	qSFC-16-CH4-1	16	217.44-227.03	226.53	7.4	12.4	-0.811
	qSFC-17-CH5-1	17	176.37	176.37	3.1	20.1	-0.828
	qSFC-16-CH9-1	16	168.29-173.76	172.14	3.6	5.6	-0.562
	qSFC-16-CH10-1	16	12.09-19.17	17.67	3.8	6	-0.582
	qSFC-16-CH11-1	16	30.34-52.37	42.28	5.7	9.5	-0.729
	qSFC-17-CH18-1	17	57.29	57.29	3.2	20.4	-0.827
	qSFC-17-CH22-1	17	68.73-84.67	77.25	4.2	6.9	-0.507
	qSFC-16-CH23-1	16	133.16-142.33	140.29	4.1	6.4	-0.58
	qSFC-17-CH26-1	17	96.92-102.17	102.17	4.3	7	0.497
BW	qBW-16-CH4-1	16	204.27-222.9	214.51	9.3	14	0.217
	qBW-17-CH4-1	17	156.02-181.46	171.61	8.9	12.1	0.251
	qBW-17-CH4-2	17	75.25-92.29	89.69	4	5	0.159
	qBW-17-CH7-1	17	2.69-11.31	7	4.3	5.4	0.168
	qBW-16-CH11-1	16	78.54-92.09	88.65	4.8	6.4	0.149
	qBW-17-CH12-1	17	80.19-99.78	92.64	6.7	8.6	0.222
	qBW-16-CH14-1	16	33.01-46.41	41.29	5.1	6.9	0.153

Appendix B (continued)

	qBW-16-CH16-1	16	35.22-54.81	44.41	6.7	9.9	0.192
	qBW-17-CH17-1	17	143.82-164.6	153.1	7.1	9.6	0.221
	qBW-16-CH22-1	16	91.66-106.73	91.66	7.8	12.4	0.212
	qBW-17-CH25-1	17	38.72-50.82	44.6	5.5	7.1	0.196
LP	qLP-CH24-1*	16	58-78.32	66.92	11.2	16.6	1.885
		17	58-76.37	65.94	7.3	8.8	1.208
	qLP-CH25-1*	16	86.34-105.64	97.24	4.5	5.9	1.077
		17	97.24-115.15	107.65	12.5	17.7	1.666
	qLP-16-CH4-1	16	106.65-115.64	111.13	5.4	7.4	1.245
	qLP-16-CH8-1	16	64.67-72.49	69.56	5.8	7.6	1.252
	qLP-17-CH14-1	17	61.88-75.77	68.84	11.1	15.2	1.561
	qLP-17-CH15-1	17	19.71-38.68	31.04	5.1	6.1	1.017
	qLP-16-CH24-2	16	128.96-149.29	141.3	5.2	6.7	1.185
	qLP-16-CH25-2	16	164.25-178.66	174.63	5	6.5	1.135
SI	qSI-CH12-1*	16	80.19-98.29	91.65	14.8	27.1	0.712
		17	80.19-100.77	90.17	14.8	30.4	0.669
	qSI-CH15-1*	16	19.71-38.68	33.54	10.7	17.5	-0.432
		17	19.71-47.61	38.68	6.7	10.6	-0.332
	qSI-17-CH7-1	17	126.63-141.29	129.79	4.4	7	0.271
	qSI-16-CH9-1	16	0-1.51	0	4.2	6	0.254
	qSI-17-CH14-1	17	88.04-100.05	98.08	3.8	6.9	-0.271
	qSI-17-CH15-2	17	141.83-159.08	152.59	4.7	7.4	-0.248
	qSI-16-CH16-1	16	47.06-61.27	53.76	4.2	6	0.256
	qSI-16-CH18-1	16	131.59-145.87	138.43	5	6.9	0.264
	qSI-17-CH26-1	17	102.17-114.07	109.53	3.9	5.8	0.239
LI	qLI-CH7-1*	16	129.79-141.29	139.78	3.1	4.2	0.176
		17	115.94-136.19	120.95	6.2	8.4	0.276
	qLI-16-CH2-1	16	4.21-17.19	11.03	7	10.8	0.297
	qLI-17-CH4-1	17	60.45-75.25	71.36	4.9	6.1	0.246
	qLI-16-CH9-1	16	183.45-203.28	193.74	5.9	8.9	0.266
	qLI-16-CH12-1	16	79.71-99.78	90.17	8.9	13.5	0.324
	qLI-17-CH14-1	17	153.61-158.4	155.68	5.5	7.1	0.244
	qLI-16-CH16-1	16	42.28-51.69	46.01	4.3	6.8	0.317
	qLI-17-CH17-1	17	111.31-126.18	121.67	5.6	7	0.251
	qLI-17-CH22-1	17	84.67-103.29	94.72	13.7	21.1	0.437
	qLI-17-CH24-1	17	55.56-76.86	66.92	9.2	12.6	0.337
PH	qPH-17-CH4-1	17	141.1-156.02	147.97	4.5	8.1	-2.025
	qPH-17-CH7-1	17	48.29-54.39	50.34	3.6	6.5	-1.705
	qPH-17-CH8-1	17	52.05-72.01	56.8	5.5	10.3	-2.557
	qPH-17-CH9-1	17	173.76-193.74	185.58	7.9	15.8	-2.625
	qPH-17-CH19-1	17	114.85-132.4	124.72	5.8	10.4	-2.204
FG	qFG-CH22-1*	16	124.58-134.55	134.55	23.5	39.2	24.361

Appendix B (continued)

	17	124.09-134.55	134.55	28.4	48	29.28
qFG-CH25-1*	16	88.36-108.63	98.71	14.8	19	-15.112
	17	98.71-104.67	101.19	4	3.6	-7.972
qFG-17-CH8-1	17	65.15-76.34	72.49	8.2	8.1	-12.182
qFG-16-CH25-2	16	11.52-24.44	22.89	6.3	6.6	10.494

*QTLs identified in both years
[a]MIC micronaire; UHM upper half mean; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzzless grade of seed
[b]PVE phenotypic variation explained [c]AE additive effect

Appendix C

QTLs of fiber quality traits, yield traits and agronomic traits identified in the RILs population NC05AZ06 × NC11-2100

Trait[a]	QTL	Range	Peak	LOD	PVE[b]	AE[c]
MIC	q2MIC-CH1-1	41.67-48.88	44.86	3.8	5.6	-0.117
	q2MIC-CH4-1	13.75-27.28	20.21	4.6	7.1	-0.131
	q2MIC-CH7-1	50.2-71.41	61.78	10.2	18.8	0.212
	q2MIC-CH12-1	121.07-131.65	121.07	4.6	7	0.129
	q2MIC-CH15-1	33.78-49.73	33.78	12.2	31.4	0.308
	q2MIC-CH24-1	114.44-117.55	115.69	4	6.5	-0.132
	q2MIC-CH26-1	39.59-55.23	48.13	7.7	13	0.176
UHM	q2UHM-CH1-1	91.64-111.46	99.33	8.8	13.3	-0.022
	q2UHM-CH5-1	80.53-97.46	91.64	7.5	11.3	0.02
	q2UHM-CH14-1	58.36-77.94	68.82	8.4	12.7	0.02
	q2UHM-CH21-1	2.53-21.12	10.34	8.9	13.5	0.021
	q2UHM-CH23-1	13.6-33.77	23.35	9	14	0.022
	q2UHM-CH24-1	56.35-66.82	56.35	4.5	6.8	0.016
UI	q2UI-CH1-1	19.92-30.66	23.83	3.2	5.4	-0.25
	q2UI-CH5-1	70.02-84.65	76.57	4.1	7.7	-0.298
	q2UI-CH5-2	20.48-30.93	20.48	3	5.1	0.249
	q2UI-CH7-1	73.29-81.66	81.66	3.1	5.4	0.245
	q2UI-CH11-1	68.95-73.61	72.33	4.1	7.4	0.292
	q2UI-CH14-1	52.88-72.3	64.1	8.7	17.5	0.465
	q2UI-CH23-1	102.21-102.85	102.21	3.2	5.5	0.251
	q2UI-CH25-1	182.76-191.94	183.99	6.1	12.1	0.396
FS	q2FS-CH2-1	2.6-20.28	8.23	6.8	10.5	0.662
	q2FS-CH14-1	54.96-77.31	68.16	12.2	20.7	-0.908
	q2FS-CH18-1	111.14-130.27	123.08	8.7	14	0.775
	q2FS-CH20-1	92.42-101.18	96.8	5	6.9	-0.555
	q2FS-CH24-1	46.05-56.35	46.05	5.4	7.9	0.566
	q2FS-CH26-1	46.85-62.92	56.49	6.9	10	-0.623
FE	q2FE-CH1-1	88.43-97.4	96.15	4.1	6.1	-0.24
	q2FE-CH3-1	90.4-107.43	96.25	6.2	10.5	0.302
	q2FE-CH6-1	70.48-80.33	70.48	3.2	16.5	-0.379
	q2FE-CH11-1	43.52-67.06	43.52	5	9.6	0.285
	q2FE-CH19-1	14.15-29.25	23.58	5.6	8.8	0.272
	q2FE-CH20-1	79.92-103.12	93.66	10.3	18.6	0.428
	q2FE-CH26-1	2.57-5.9	2.57	4.8	7.4	0.269
SFC	q2SFC-CH5-1	66.81-84.65	76.57	10.4	20.4	0.397
	q2SFC-CH7-1	50.2-66.93	65.63	3.9	8	-0.249
	q2SFC-CH11-1	60.67-73.61	73.61	4	18.9	-0.379
	q2SFC-CH13-1	85.12-120.78	120.78	4.2	6.9	0.234
	q2SFC-CH14-1	52.88-73.59	64.1	8.5	15.8	-0.346
	q2SFC-CH25-1	107.81-122.46	117.31	4.1	7.4	-0.233

Appendix C (continued)

BW	q2BW-CH11-1	109.87-127.67	120.42	5.9	10.5	0.158
	q2BW-CH13-1	0-12.93	3.89	4.1	7	-0.128
	q2BW-CH16-1	79.46-94.89	82.58	8.7	16.9	0.201
	q2BW-CH17-1	103.09-111.32	105.54	3.9	6.6	-0.125
	q2BW-CH21-1	2.53-21.12	3.77	7	12.9	0.174
	q2BW-CH23-1	47.38-64.11	55.8	7.4	13.6	0.181
	q2BW-CH26-1	53.26-71.21	62.92	6.3	11.4	0.165
LP	q2LP-CH5-1	118.9-140.93	128.11	10.5	20.6	1.281
	q2LP-CH6-1	47.48-69.87	61.05	5.1	8.7	-0.744
	q2LP-CH14-1	50.22-68.82	60.3	10.2	19.4	1.192
	q2LP-CH20-1	65.84-74.5	74.5	3.1	5	0.579
	q2LP-CH24-1	109.15-121.4	115.69	5.2	8.6	-0.77
	q2LP-CH25-1	96.82-109.09	109.09	7	17.7	-1.075
	q2LP-CH26-1	31.17-45.55	36.25	12.1	24.4	1.325
SI	q2SI-CH1-1	30.66-40.42	34.67	4.9	9.5	0.256
	q2SI-CH2-1	94.05-104.39	99.22	3.6	7	0.225
	q2SI-CH9-1	107.06-116.7	114.19	4.7	8.9	-0.273
	q2SI-CH14-1	95.2-99.17	95.2	4.5	8.6	-0.246
	q2SI-CH15-1	142.36-162.85	158.13	8.2	18.5	0.364
	q2SI-CH20-1	99.31-110.9	105.79	5.2	10.1	-0.277
LI	q2LI-CH3-1	93.71-112.87	98.09	4.7	9.9	0.225
	q2LI-CH7-1	35.03-50.2	46.49	6.3	13.4	0.275
	q2LI-CH9-1	7.84-22.71	22.71	3.8	7.6	-0.199
	q2LI-CH16-1	69.4-79.46	70.66	3.7	7.5	0.2
	q2LI-CH17-1	8.24-13.72	8.24	3.2	6.8	0.194
	q2LI-CH23-1	51.99-53.91	51.99	3.2	6.3	0.179
	q2LI-CH25-1	104.65-116.01	109.09	6.7	16.3	0.291
PH	q2PH-CH5-1	52.54-72.62	60.07	7.8	14.1	2.609
	q2PH-CH10-1	161.7-187.01	179.01	5.7	10.4	-2.401
	q2PH-CH12-1	69.23-73.56	69.23	3.8	6	1.816
	q2PH-CH19-1	121.88-133.55	129.76	8.6	15.4	-2.835
	q2PH-CH20-1	35.86-48.69	42.97	4.6	7.4	-2.032
	q2PH-CH21-1	73.79-79.22	74.41	3.6	5.6	1.656
FG	q2FG-CH5-1	80.53-97.46	86.62	7.5	9	13.024
	q2FG-CH14-1	52.88-64.76	52.88	3.2	14.6	15.823
	q2FG-CH22-1	105.95-118.66	118.04	19.8	34.1	25.519
	q2FG-CH26-1	32.42-53.26	45.55	8.2	10.4	-13.687

[a]MIC micronaire; UHM upper half mean; UI uniformity index; FS fiber strength; FE fiber elongation; SFC short fiber content; BW boll weight; LP lint percentage; SI seed index; LI lint index; PH plant height; FG fuzzless grade of seed

[b]PVE phenotypic variation explained [c]AE additive effect