EFFECTS OF EXOTIC GENOTYPES AND GENETIC BACKGROUNDS ON FIBER QUALITY AND PLANT ARCHITECTURAL TRAITS IN UPLAND COTTON (*Gossypium hirsutum* L.)

By

Jeevan Adhikari

(Under the Direction of Andrew H. Paterson)

ABSTRACT

The genetic diversity of the world's cultivated cotton gene pool is poor due to a series of bottlenecks caused by migration, domestication and selection. Low genetic diversity of cultivated cotton is a major hindrance to genetic improvement of the crop. Utilization of wild and primitive accessions of cotton in breeding programs would allow the introduction of useful and favorable alleles to the cultivated gene pool and identification of significant associations between genetic markers and traits of interest. In our study, we used three primitive cotton genotypes collected from different parts Mexico and Guatemala and converted to day-neutral flowering to create experimental populations. We screened the segregating F_2 and $F_{2:3}$ generations of these populations with 85 polymorphic SSR markers selected from 18 "hotspot" regions in the cotton genome rich in fiber quality quantitative trait loci (QTLs). We investigated the association of these markers with six different fiber quality traits as well as with yield related and morphological/ plant architectural traits in those populations. Significant associations were identified many traits of interest in different population for different years.

Index words: *Gossypium hirsutum*, SSR, QTL, polymorphism, photoperiod, exotic genotypes

EFFECTS OF EXOTIC GENOTYPES AND GENETIC BACKGROUNDS ON FIBER QUALITY AND PLANT ARCHITECTURAL TRAITS IN UPLAND COTTON (*Gossypium hirsutum* L.)

By

JEEVAN ADHIKARI

B.S., Tribhuvan University, Nepal, 2010

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

MASTER OF SCIENCE

ATHENS, GEORGIA

2015

EFFECTS OF EXOTIC GENOTYPES AND GENETIC BACKGROUNDS ON FIBER QUALITY AND PLANT ARCHITECTURAL TRAITS IN UPLAND COTTON

(*Gossypium hirsutum* L.)

By

JEEVAN ADHIKARI

Major Professor: Andrew H. Paterson

Committee: Peng Chee
Cheolwoo Park

Electronic Version Approved:
Suzanne Barbour
Dean of the Graduate School
The University of Georgia
August 2015

DEDICATION

I would like to dedicate this thesis to my parents, my grandmother, my brother and my lovely wife, Roshani.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my advisor, Dr. Andrew Paterson, for his inspiration, unconditional support and mentoring. I would also like to thank my committee members Dr. Peng-Wah Chee and Dr. Cheolwoo Park, for their constant support, suggestions and precious time. I am also grateful to all the members of Plant Genome Mapping Laboratory, for their support and help throughout my research.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

SUPPLEMENTARY TABLES

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Purpose of the study

There has been stiff competition between synthetic fibers and naturally occurring cotton fibers in the world textile industry. Despite the immense competition posed by these man-made synthetic fibers, the preference for cotton products has aided its survival as one of the world's most widely cultivated crops (Saleem, Bilal et al. 2010). However, agronomic performance and fiber quality must continually be improved if cotton is to maintain its economic viability (McCarty, Wu et al. 2006). Improvement of any crop requires diverse genetic resources. The worldwide Upland cotton gene pool is genetically impoverished due to a series of bottlenecks imposed by polyploid formation, domestication and migration. In addition, overexploitation of a few genetic backgrounds by breeders may have contributed to a plateau of yield and fiber quality traits (Van Esbroeck and Bowman 1998). Such a narrow genetic base might result in a crop being highly vulnerable to stresses (McCarty, Jenkins et al. 1998). A decline in the genetic diversity of Upland cotton and the need to broaden the genetic base of cotton germplasm useful for the improvement of lint yield, fiber quality and biotic or abiotic stresses has been widely reported (Esbroeck, Bowman et al. 1999, Bowman 2000, Iqbal, Reddy et al. 2001, Gutierrez, Shoemaker et al. 2002).

One solution to the problem of cotton genetic vulnerability might be the exploration of exotic genotypes (Paterson, Boman et al. 2004). Incorporating favorable alleles, genes or gene complexes from wild relatives or accessions has been a high strategic priority for practical improvement of some crops (Hajjar and Hodgkin 2007). Research has shown that primitive accessions of cotton are highly diverse and have useful genetic variability (Percival 1987, Meredith 1991, McCarty and Jenkins 1992, McCarty, Jenkins et al. 1995, McCarty, Jenkins et al. 1998, McCarty, Jenkins et al. 1998, McCarty and Jenkins 2001, McCarty, Wu et al. 2003). Many of the accessions in the US national collection of cotton germplasm have been reported to have useful genetic variability (Meredith 1991, McCarty and Jenkins 1992, McCarty and Jenkins 2001). In practice, however, the utilization of genetic variability from primitive cotton accessions has been limited due to their photoperiod response (Stephens, Miller et al. 1967, Holley and Goodman 1988, McCarty and Jenkins 1992, Uga, Nonoue et al. 2007).

Conversion of primitive, usually short-day flowering, genotypes into day-neutral forms by repeated backcrossing has been undertaken to reduce the obstacles to using this germplasm. McCarty, et, al. (1995, 1998a, 1998b) evaluated F_5, BC_1F_5, BC_2F_5, BC_3F_5 and BC_4F_5 progenies of 16 converted day neutral germplasm accessions for several agronomic and fiber traits by crossing to converted germplasm. Exotic alleles are introduced into cultivated genetic backgrounds when they are crossed with exotic and/or converted germplasm. Many of these alleles and quantitative trait loci (QTLs) can have significant effects on plant morphology and fiber quality traits. Furthermore, introgression of exotic alleles increases the genetic diversity available in the elite gene pool of the crop and can provide a foundation for further crop improvement.

With the aim of identifying new alleles and allele combinations useful in cotton improvement, we designed an exploration of utilizing converted exotic and primitive accessions of cotton. The major aim of this exploration is to mine alleles and marker loci significantly associated with QTLs that could affect morphological and/or fiber quality traits of Upland cotton.

Economic importance of cotton

Cotton (*Gossypium spp.*) is an important cash crop grown in the United States and many parts of the globe for its spinnable fiber. Known to the agricultural world as "White Gold" for its soft fluffy staple fiber, cotton is the major source of natural fiber for the textile industry. Cotton is cultivated primarily for its lint; which is harvested as seed-cotton and later seeds are ginned out. Apart from being an important source of lint fiber, cottonseeds are a rich source of oil, which is of high industrial value and can also be consumed by human beings like other vegetable oils. Cottonseed meal serves as protein rich feed for ruminant livestock. Cotton hulls are used as animal feed, fertilizer and fuel.

Cotton has been associated with ancient civilizations and has greatly contributed to industrial and economic development of many countries. The worldwide economic impact of cotton is estimated at around $500 billion/year with an annual utilization of around 115 million bales of cotton fiber (Chen, Scheffler et al. 2007). Today, cotton is commercially cultivated in more than 50 countries in drier tropical and sub-tropical areas of the world (Smith and Cothren 1999). The major share of global cotton production comes from countries such as China, India, the United States, Pakistan and Australia where climatic conditions such as periods of hot and dry weather, photoperiod and adequate soil moisture favor natural growth requirements of cotton (Khadi, Santhy et al. 2010).

The United States is typically the third largest producer of cotton in the world, accounting for about 16% of the total world production and is usually the leading exporter, accounting for over one-third of the global trade in raw cotton. The United States cotton industry accounts for more than $25 billion in products and services annually, generating about 200,000 jobs in industry sectors from farm to textile mill (USDA 2013). The surviving United States textile industry faces a challenge to improve its competitiveness and increase domestic demand. Modernizing equipment will no doubt improve production efficiency and the competitiveness of United States textile products. Modern high efficiency equipment has also raised the requirements for fiber quality to maximize the efficiency of high-speed spinning equipment. The new technology requires stronger and longer fibers and fewer short fibers (Zeng, Meredith et al. 2007). Unfortunately, high yielding cultivars in the United States frequently do not possess fiber quality high enough to meet these standards (Zeng, Meredith et al. 2007).

Domestication of cotton

There are more than 50 cotton species in the genus *Gossypium* (Fryxell and Craven 1992) of which most are diploid (2n=26) and five are polyploid (2n=52). Only four species of cotton have been domesticated and are cultivated commercially for lint production. Approximately 90% of world's cotton production comes from allotetraploid *G. hirsutum* (Upland cotton) and about 8% comes from another allotetraploid, *Gossypium barbadense*, commonly referred to as Sea Island cotton, Pima cotton, or Egyptian cotton. Two diploid species *Gossypium arboreum* and *Gossypium herbaceum* contribute less than 2% towards total cotton production (Kumar 2012). Cotton was introduced into the eastern coastal areas of North America by European immigrants. *G. hirsutum* was grown in the upland area of the country hence it received its common name, Upland cotton. It is a perennial plant but cultivated as an annual crop. Sometimes referred to as

New World cotton or short-staple cotton, it has light green fuzzy seed to which lint is firmly attached. This species has broad adaptation and is grown over a wide geographical range. Upland cotton is an allotetraploid species (2n=52) formed by union of the A genome and D-genome in A-genome cytoplasm (Wendel 1989).

Genetic diversity of upland cotton

The level of genetic diversity of any crop species is an essential element of sustainable crop production and development. Cotton productivity and the future of cotton breeding efforts depend on the level of genetic diversity of cotton gene pools and its effective exploitation in breeding programs. In the past few decades, efforts in molecular marker technology have helped reveal the level of genetic diversity of many crops. These efforts reinforced a serious concern about the narrow genetic base of cultivated cotton germplasm, which has obviously been associated with a 'genetic bottleneck' during its history of domestication and improvement (Iqbal, Reddy et al. 2001, Chee, Lubbers et al. 2004, Lubbers, Chee et al. 2004).

There is rich genetic diversity in the *Gossypium* genus, including all its morphological, physiological and agronomic properties (Mauer 1954), with characteristics such as plant architecture, stem pubescence and color, leaf shape, flower color, pollen color, boll size and shape, fiber quality, yield potential, early maturity, photoperiod dependency, and resistance to multi-adversity environmental stresses (Abdurakhmonov, Abdullaev et al. 2005). This wide phenotypic diversity of cotton motivates the potential use of these diverse species in the breeding programs as the initial materials.

There are numerous examples of utilization of such genetic variations in solving many fundamental problems in cotton breeding and production (Abdurahmonov 2007). The exploration of genetic diversity for *Verticillium* wilt fungi from the exotic *G. hirsutum*

germplasm and its mobilization into elite cultivars solved the wilt epidemics in Uzbekistan (Abdurakhmonov, Abdullaev et al. 2005). One of the progenies from this *Verticillium* resistant accession turned out to also be salt tolerant. A number of other examples of the creation of natural defoliation, disease and pest resistance, tolerance to multi-adversity stresses, improved seed oil content and fiber quality parameters, utilizing natural genetic diversity have been documented (Abdurakhmonov, Abdullaev et al. 2005).

Fiber quality and exotic genotypes

The market value of raw cotton is determined in part by its fiber quality, which is a collective term given to a set of measurements that describe the physical properties of a sample of fibers extracted from a cotton bale (Bradow and Davidonis 2000). Cotton fiber with desirable quality not only helps in maintaining and enhancing yarn processing efficiencies but also influences the quality of the end product. The main fiber quality parameters that the textile industry utilizes to predict the properties of a fiber sample include fiber length, fiber strength, fiber elongation, fiber fineness and fiber uniformity index.

Fiber quality is of growing importance in cotton breeding programs. A lot of studies and experiments have been done in cotton while utilizing exotic cotton germplasm for fiber quality traits (Verhalen, Murray et al. 1969, McCarty, Jenkins et al. 1998, Uga, Nonoue et al. 2007, Zeng, Meredith et al. 2007). Approaches range from classical phenotype-based selection to the use of modern molecular tools and techniques. Pedigree selection of exotic germplasm collection has been advocated for improving traits in cotton such as micronaire and fiber length (Verhalen and Murray 1967), lint percent , seed cotton yield, lint yield, and earliness (Verhalen, Murray et al. 1969) .

Lines derived from primitive accessions of *G. hirsutum* have been reported to contain useful genetic resources for Upland cotton improvement (McCarty and Jenkins 1992, McCarty, Jenkins et al. 1995, McCarty, Jenkins et al. 1998, Mei, Syed et al. 2004, McCarty, Wu et al. 2005). In order to assess their utility in breeding programs, some derived day-neutral lines have been further investigated in their hybrid forms (Swindle 1993, McCarty, Jenkins et al. 1995, McCarty, Jenkins et al. 1996, McCarty and Jenkins 2001, McCarty, Wu et al. 2003, McCarty, Jenkins et al. 2004). More than 50 of 70 hybrids displayed improved fiber strength compared to Deltapine 50 and other commercial cultivars (McCarty, Wu et al. 2003, McCarty, Jenkins et al. 2004, McCarty, Wu et al. 2005).

Zeng et al., 2007, developed a species polycross population to exploit and evaluate the alleles from other tetraploid cotton genotypes. Association analysis using this population (Zeng, Meredith et al. 2009) revealed 59 significant marker-trait associations with six fiber quality traits. Fiber quality alleles from wild Hawaiian cotton (*G. tomentosum*) were evaluated (Zhang, Rong et al. 2011) using 17 interspecific backcross-self families. Alleles from *G. tomentosum* were found to be associated with multiple favorable effects on fiber quality traits. Potential exploitation of alleles from another wild cotton relative (*G. mustelinum*) is being evaluated (Dantas, Barroso et al. 2012).

QTL mapping for fiber quality traits

In recent years, molecular marker technology has facilitated the construction of detailed molecular maps of the cotton genome. These genetic linkage maps can now be employed in mapping Quantitative Trait Loci (QTLs) governing yield and fiber related traits. A significant marker-trait association would be beneficial for marker-assisted selection in cotton breeding as well as for cloning genes of interest. A pioneering study mapped 14 QTLs for agronomic and

fiber quality traits on a linkage map derived from an interspecific cross of *G. hirsutum* cv. 'CAMD-E' and *G. barbadense* cv. 'Sea Island Seaberry' (Jiang, Wright et al. 1998) and also found that the most QTLs influencing fiber quality and yield were located on the "D" sub-genome, derived from an ancestor that does not produce spinnable fibers. Similar results were observed in a study involving an interspecific cross between four different *Xanthomonas campestris* pv. *Malvacearum* (Xcm) resistant *G. hirsutum* parents and *G. barbadense* parent 'Pima S-7' (Wright, Thaxton et al. 1999). Together these studies suggested that the merger of A and D genomes with different evolutionary histories in a common nucleus offered unique avenues for phenotypic response to selection.

Twenty-six QTLs for various agronomic and fiber quality traits were identified on nine linkage groups in an intraspecific cross of *G. hirsutum* (Ulloa, Meredith Jr et al. 2002). The phenotypic variation explained by these QTLs ranged from 3.4- 44.6%. A single major QTL for fiber strength originating from a *G. anomalum* introgression line 'Suyuan 7235' was detected on chromosome 10 (Zhang, Guo et al. 2002). Using an interspecific map developed from RFLP, SSR and AFLP markers, seven QTLs for various fiber-related traits were identified (Mei, Syed et al. 2004). A comprehensive analysis of advanced generation backcross populations from a cross between *G. hirsutum* cv. Tamcot2111 and *G. barbadense* cv. using RFLP markers detected 22 QTLs for fiber elongation (Chee, Draye et al. 2005), 32 and 9 QTLs for fiber fineness and micronaire (Draye, Chee et al. 2005), 28, 9 and 8 QTLs for fiber length, uniformity and short fiber content, respectively (Chee et al., 2005b).

QTL mapping for plant architectural traits

Plant architecture is defined as the three-dimensional organization of a plant, which is dependent on the function and relative arrangement of each of its parts. Plant architecture is of

major agronomic importance because it strongly influences the suitability of a plant for cultivation, its overall yield and its economic coefficient (Reinhardt and Kuhlemeier 2002). One of the notable successes in plant architectural breeding was the mining and introduction of dwarfing genes in wheat and rice during the Green Revolution, leading to increased harvest index and grain yield (Peng, Richards et al. 1999). Total yield in cotton is highly dependent on different architectural traits such as plant height and branching as well as on flowering and maturity indices. Alterations in canopy architecture allow increased light penetration to the lower portion of the plant canopy, resulting in increased number of bolls and total yield (Pettigrew 1994, Reta-Sánchez and Fowler 2002). However, a more economical alternative would be to develop genotypes with suitable plant architecture (Song and Zhang 2009). Identification of molecular markers diagnostic of these architectural traits would accelerate breeding schemes to develop suitable plant genotypes for higher yield and quality.

Plant height is an important agronomic trait used in plant architecture breeding. Since this trait is associated with plant morphogenesis, lodging resistance and harvest efficiency, it has gained focus in breeding schemes to achieve yield potential in different crops. With the advance of molecular marker technologies, QTLs affecting plant height have been identified in many crops like rice (Li et al 2003, Zhang et al 2006b), maize (Zhang et al 2006c), wheat (Cui et al 2011) and sugarcane (Ming et al 2002). In cotton, this trait is closely related to canopy size and photosynthetic capacity of the plant (Liu, Ai et al. 2014). Plant height in cotton primarily involves the number and length of the mainstem nodes and is determined by cell expansion during the growing season. The primary stem of the cotton plant supports necessary fruiting and vegetative branches, and thus the total number of bolls produced by the plant depends on the appropriate height of the plant. Although different growth environments and planting patterns

allow for varying levels of individual plant height in cotton production (Reta-Sánchez and Fowler 2002), the global trend toward machine picking makes shorter plants a better alternative, since taller plants are often associated with excessive vegetative growth and later maturity and can present harvesting difficulties (Percy, Cantrell et al. 2006).

Plant height in cotton is inherited both quantitatively and qualitatively. There are many genes associated with plant height (Ellis, Rebetzke et al. 2005). Several cotton dwarfing mutants have been identified (Wu, Zhou et al.) and some phytohormone signaling pathway-related dwarfing genes have been characterized (Yang, Foo et al. 2006, Aleman, Kitamura et al. 2008, Liao, Ruan et al. 2009). Traditional quantitative genetic studies have revealed that plant height is also a complex trait, with additive effects (Wu, McCarty et al. 2009), both additive and dominant effects (Naveed, Abdul et al. 2006) and/or epistasis (Kalsy and Garg 1988, Khan and Khan 1993). Over the last decade, many QTL for cotton plant height have been identified (Shappley, Jenkins et al. 1998, Wang, Wu et al. 2006, Qin, Liu et al. 2009, Song and Zhang 2009, Muhammad, Wangzhen et al. 2011, Liu, Ai et al. 2014).

Plant architecture can significantly affect light penetration and distribution in the crop canopy and affect plant growth, biomass partitioning, boll distribution, boll weight and yield potential (Kaggwa-Asiimwe, Andrade-Sanchez et al. 2013). It can also play a significant role in reducing losses to diseases and insects. The number of nodes in the primary stem of the cotton plant determines both the height of the plant and the number and orientation of branches the plant bears, thus affecting the overall plant architecture. Usually, one secondary branch arises from a node of the primary stem, but more than one secondary branch can emanate from a single node. The number of branches arising from a single node determines the plant canopy density and also affects plant architecture, thus having significant effects on boll number, boll

distribution and boll weight. This will have a direct effect on the quality and quantity of fiber yield from the cotton plant. Identification of QTLs or molecular markers associated with the number of branches arising from a node will be helpful not only for breeding for cotton plant architecture but also for overall quantity and quality of fiber yield.

QTL mapping for yield and yield related traits

Cotton fiber development is a complex process that involves fiber initiation, elongation primary wall synthesis), wall thickening (secondary wall synthesis) and desiccation or maturation (Barsa and Malik, 1984). The seed coat of cotton is covered with lint (long) and fuzz (short) fibers. Lint is separated from fuzz and the seed during ginning. Lint yield is an important component of total fiber yield because it is that yield component that is the raw material for textile industries. Lint yield is a complex trait in cotton, which is controlled by a large number of QTLs. In the past two decades many QTLs related to lint yield have been identified in Upland cotton (Mei, Zhu et al. 2013).

Lint percentage is another important yield component in Upland cotton and has been found to have very high heritability (Wang, Li et al. 2014). It correlates with seed cotton yield, lint yield and other yield components to different degrees. A handful of research efforts relating to increasing lint percentage in cotton have been carried out in the last few decades. Zhang et al (2003a) conducted genetic analysis of historic cotton varieties in China and suggested that different yield components contributed to total seed yield differently in different time period. Their results suggested that screening cotton with excellent lint percentage and big bolls is an effective way not only for yield breeding but also for quality breeding. Correlation and path analysis on yield characteristics in different Upland cotton varieties revealed the largest contribution of bolls per plant to total yield followed by lint percentage (Li, Guo et al. 2008). So,

increasing lint percentage has significant effects on increasing lint yield. Exploration and identification of markers that are closely related to lint percentage QTLs is a useful way of breeding for higher lint percentage. Identification of markers related to lint percentage QTLs and mapping of these QTLs have been widely reported (Ulloa and Meredith 2000, Abdurakhmonov, Buriev et al. 2007, Li, Wang et al. 2012, Liu, Wang et al. 2012, Zhang, Qian et al. 2013, Wang, Li et al. 2014). More than 50 QTLs related to lint percentage have been detected in these studies.

Plant flowering time is an adaptive trait with biological and agricultural significance (Murfet 1977) and is directly related to quantity and quality of cotton yield. Primitive cottons are photoperiod sensitive and conventional genetic analysis of photoperiod sensitivity of cotton has been carried out in different intraspecific (*G. hirsutum*) hybrids (Guo, McCarty et al. 2009). Flowering time has been found to be under multigenic control and different segregation patterns have been observed (Waddle, Lewis et al. 1961). However, the number of loci controlling genetic variation of the trait and their map positions have not been well characterized (Guo, McCarty et al. 2009). Identification of molecular markers significantly associated with flowering is important to characterize and localize the loci controlling this trait. Several morphological indices have been reported to help estimate the cotton crop maturity and earliness such as: node of the first fruiting branch, number of vegetative branches and percentage of bolls on vegetative branches (Ray and Richmond, 1966); the time to first square or first flower (Joham, 1979). In cotton, the main stem node of the first fruiting branch (NFB) was positively related to flowering time and has been used as a practical measurement of earliness. Ray and Richmond (1966) studied various morphological measures of earliness in cotton and considered NFB as the most reliable and practical measurement of earliness. Since then, many studies focusing on NFB as a reliable measure of earliness in flowering have been conducted and many QTL for NFB has been

identified (Guo, McCarty et al. 2009). The proportion of flowers in a given time of a cotton crop season can also be taken as a measure of earliness or lateness in flowering. This method could provide a direct measure of the flowering behavior of different cotton genotypes.

Similarly to NFB, the percentage of bolls on vegetative branches and the proportion of opened and closed bolls in a given time of a cotton crop season (boll maturity) is an indicative measure of earliness in cotton. As information related to earliness in cotton is scanty, identification of molecular markers associated strongly with boll maturity as well as flowering is very useful in breeding for early maturing genotypes.

Meta-analysis of QTLs in cotton

Hundreds of QTLs related to fiber, yield, plant morphology and architecture, biotic and abiotic stresses have been identified in cotton from studies using different parents and environments. Differences in mapping populations, genotypes and environments can yield heterogeneous results in QTL mapping (Rong, Feltus et al. 2007, Zhang and Percy 2007). Such data would also result in identification of partly or wholly non-overlapping sets of QTLs (Rong, Feltus et al. 2007). With an aim to merge QTL data from different sources and compile them into a consensus map, a few meta-analyses of QTLs in cotton have been carried out by Rong et al., (2007), Lacape et al., (2010), and Said et al., (2013).

Rong et al., (2007) performed a comprehensive meta-analysis of cotton QTLs using a high density reference map consisting of 3475 loci and reported the alignment of 432 QTLs involving cotton fiber quality, yield, leaf morphology, flower morphology, and other traits in 11 different populations. They revealed that more QTL were detected in the D sub-genome of tetraploid cotton than in the A subgenome. They also came to a conclusion that QTLs are clustered in certain locations in the cotton genome rather than being distributed randomly. A

total of 18 “hotspots” in the cotton genome that contained 104 of 432 QTLs in this study were mapped to a whole genome marker map of cotton based on the D-genome sequence of *Gossypium raimondii* (Wang, Zhang et al. 2013). Each hotspot consisted of 3 to 7 QTLs related to fiber quality traits such as fiber elongation, fiber length, fiber strength, fiber fineness, fiber color, fiber uniformity and short fiber content.

Genetic background and their effects

Numerous studies have shown interactions between QTL alleles and genetic backgrounds (Sebolt, Shoemaker et al. 2000, Lecomte, Duffé et al. 2004, Li, Li et al. 2009). Lecomte et al. (2004) introgressed five QTLs controlling fruit quality into three tomato lines and found that the phenotypic effect of each QTL varied according to the recipient parent. The effect of dent maize genetic background on grain yield QTLs was significant- among the total of 33 QTLs identified, only one was in common between the two mapping population developed by crossing a single high-oil maize line with two dent maize inbred lines (Li, Li et al. 2009). QTL analysis for rice panicle number in doubled haploid as well as recombinant inbred populations derived from crossing a japonica variety with two indica varieties showed that the effect of genetic background on QTL detection can be greater than the environment effect (Liao, Wu et al. 2001). Furthermore, the magnitude of epistatic interaction among loci is not only greatly influenced by the genetic background in which the QTLs were detected but can be greater than that of QTL main effects.

In Upland cotton, significant interaction between QTL and genetic backgrounds or family interactions has been observed (Chee, Draye et al. 2005, Draye, Chee et al. 2005). At least 11 loci affecting fiber fineness (Draye, Chee et al. 2005) and 19 loci affecting fiber length (Chee, Draye et al. 2005) showed significant interaction with family background, thereby altering the

genetic effects of the introgressed alleles. Additive x Additive, Additive x Dominance, Dominance x Additive and Dominance x Dominance epistatic interactions were observed among nine loci associated with six fiber quality traits (Wang, Wu et al. 2007).

Summary

Cotton has always been a crop of choice for producing natural fiber for the textile industry, despite stiff competition posed by synthetic fibers. A major challenge for the cotton industry is to improve the quality and quantity of the product, which in part has been hindered by bottlenecks of domestication, migration and selection, resulting in low diversity of cultivated cotton species. The improvement in fiber quality traits of US cotton cultivars has been slow owing to these bottlenecks and overexploitation of a few closely related genotypes in the already narrow gene pool of the crop. As wild and exotic accessions of cotton are diverse and might contain valuable alleles for improvement of cultivated US cotton genotypes, inclusion of these accessions in breeding programs may be very useful in increasing genetic diversity, introducing favorable alleles from wild donors. Here, we have used three exotic accessions of cotton introduced from different parts of Mexico and Guatemala and converted to day-neutral flowering, evaluating them in different cultivated backgrounds representing the US cotton gene pool to identify the effects of these wild genotypes on different fiber quality traits as well as on different plant agronomic and architectural traits.

Identification of molecular markers associated with different fiber quality and plant architectural traits may accelerate breeding and improvement of these traits. SSRs have become a marker class of choice in many crop species including cotton because of their co-dominant nature, high reproducibility and adaptability to high-throughput genotyping platforms (Powell, Morgante et al. 1996). In our study we have used SSR markers selected from genomic regions of

cotton that are reported to be “hotspots” for fiber quality traits (Rong, Feltus et al. 2007) and screened them in different populations to identify significant associations with fiber quality as well as plant architectural traits.

References

Abdurahmonov, I. 2007. Exploiting genetic diversity. World Cotton Research Conference-4, Lubbock, Texas, USA, 10-14 September 2007. International Cotton Advisory Committee (ICAC), Washington; USA. p. un-un.

Abdurakhmonov, I.Y., A.A. Abdullaev, S. Saha, Z.T. Buriev, D. Arslanov, Z. Kuryazov, et al. 2005. Simple sequence repeat marker associated with a natural leaf defoliation trait in tetraploid cotton. Journal of Heredity 96: 644-653.

Abdurakhmonov, I.Y., Z.T. Buriev, S. Saha, A.E. Pepper, J.A. Musaev, A. Almatov, et al. 2007. Microsatellite markers associated with lint percentage trait in cotton, Gossypium hirsutum. Euphytica 156: 141-156. doi:10.1007/s10681-007-9361-2.

Aleman, L., J. Kitamura, H. Abdel-mageed, J. Lee, Y. Sun, M. Nakajima, et al. 2008. Functional analysis of cotton orthologs of GA signal transduction factors GID1 and SLR1. Plant Molecular Biology 68: 1-16. doi:10.1007/s11103-008-9347-z.

Bowman, D.T. 2000. Attributes of public and private cotton breeding programs [electronic resource]. Journal of cotton science (Online) Journal of cotton science 4: 130-136. doi:http://www.jcotsci.org/toc.html.

Bradow, J.M. and G.H. Davidonis. 2000. Quantitation of fiber quality and the cotton production-processing interface: a physiologist's perspective. Journal of Cotton Science 4: 34-64.

Chee, P.W., X. Draye, C.X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, et al. 2005. Molecular dissection of phenotypic variation between Gossypium hirsutum and

Gossypium barbadense (cotton) by a backcross-self approach: III. Fiber length. TAG Theoretical and Applied Genetics 111: 772-781.

Chen, Z.J., B.E. Scheffler, E. Dennis, B.A. Triplett, T. Zhang, W. Guo, et al. 2007. Toward Sequencing Cotton (Gossypium) Genomes. Plant Physiology: 1303. doi:10.2307/40065773.

Dantas, A.C.A., P.A.V. Barroso, L.V. Hoffmann, M.F. Alves and F.P.d. Andrade. 2012. SSR markers to detect gene flow from upland to mocó cotton. Revista Ciência Agronômica 43: 163-169.

Draye, X., P. Chee, C.X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, et al. 2005. Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense (cotton) by a backcross-self approach: II. Fiber fineness. TAG Theoretical and Applied Genetics 111: 764-771.

Ellis, M.H., G.J. Rebetzke, F. Azanza, R.A. Richards and W. Spielmeyer. 2005. Molecular mapping of gibberellin-responsive dwarfing genes in bread wheat. Theoretical & Applied Genetics 111: 423-430. doi:10.1007/s00122-005-2008-6.

Esbroeck, G.A.v., D.T. Bowman, O.L. May and D.S. Calhoun. 1999. Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars. Crop Science 39: 323-328.

Fryxell, P.A. and L.A. Craven. 1992. A Revision of Gossypium Sect. Grandicalyx (Malvaceae), Including the Description of Six New Species. Systematic Botany: 91. doi:10.2307/2419068.

Guo, Y., J.C. McCarty, J.N. Jenkins, C. An and S. Saha. 2009. Genetic detection of node of first fruiting branch in crosses of a cultivar with two exotic accessions of upland cotton. Euphytica 166: 317-329. doi:10.1007/s10681-008-9809-z.

Gutierrez, O.A., D.B. Shoemaker, C.L. Cheatham, J.C. McCarty, Jr., S. Basu, S. Saha, et al. 2002. Genetic distance among selected cotton genotypes and its relationship with F2 performance. Crop science 42: 1841-1847.

Hajjar, R. and T. Hodgkin. 2007. The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica 156: 1-13. doi:10.1007/s10681-007-9363-0.

Holley, R.N. and M.M. Goodman. 1988. Yield potential of tropical hybrid maize derivatives. Crop Science 28: 213-218.

Iqbal, M.J., O.U.K. Reddy, K.M. El-Zik and A.E. Pepper. 2001. A genetic bottleneck in the 'evolution under domestication' of upland cotton Gossypium hirsutum L. examined using DNA fingerprinting. Theor Appl Genet 103: 547-554. doi:10.1007/PL00002908.

Jiang, C., R.J. Wright, K.M. El-Zik and A.H. Paterson. 1998. Polyploid formation created unique avenues for response to selection in Gossypium (cotton). Proceedings of the National Academy of Sciences of the United States of America 95: 4419-4424.

Kaggwa-Asiimwe, R., P. Andrade-Sanchez and G. Wang. 2013. Plant architecture influences growth and yield response of upland cotton to population density. Field Crops Research 145: 52-59. doi:http://dx.doi.org/10.1016/j.fcr.2013.02.005.

Khadi, B.M., V. Santhy and M.S. Yadav. 2010. Cotton: An Introduction. BIOTECHNOLOGY IN AGRICULTURE AND FORESTRY 65: 1-14.

Kumar, P. 2012. Genetic evaluation of exotic chromatins from two obsolete interspecific introgression lines of Upland cotton [electronic resource] / by Pawan Kumar2012.

Lecomte, L., P. Duffé, M. Buret, B. Servin, F. Hospital and M. Causse. 2004. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theoretical & Applied Genetics 109: 658-668.

Li, C., W. Guo, X. Ma and T. Zhang. 2008. Tagging and mapping of QTL for yield and its components in upland cotton population with varied lint percentage. Cotton Science 20: 163-169.

Li, C., C. Wang, N. Dong, X. Wang, H. Zhao, R. Converse, et al. 2012. QTL detection for node of first fruiting branch and its height in upland cotton (Gossypium hirsutum L.). Euphytica 188: 441-451. doi:10.1007/s10681-012-0720-2.

Li, Y.L., X.H. Li, J.Z. Li, J.F. Fu, Y.Z. Wang and M.G. Wei. 2009. Dent corn genetic background influences QTL detection for grain yield and yield components in high-oil maize. Euphytica 169: 273-284. doi:10.1007/s10681-009-9966-8.

Liao, C.Y., P. Wu, B. Hu and K.K. Yi. 2001. Effects of genetic background and environment on QTLs and epistasis for rice (Oryza sativa L.) panicle number. Theoretical & Applied Genetics 103: 104-111.

Liao, W., M. Ruan, B. Cui, N. Xu, J. Lu and M. Peng. 2009. Isolation and characterization of a GAI/RGA-like gene from Gossypium hirsutum. Plant Growth Regulation 58: 35-45.

Liu, R., N. Ai, X. Zhu, F. Liu, W. Guo and T. Zhang. 2014. Genetic analysis of plant height using two immortalized populations of "CRI12 × J8891" in Gossypium hirsutum L. Euphytica 196: 51-61.

Liu, R., B. Wang, W. Guo, Y. Qin, L. Wang, Y. Zhang, et al. 2012. Quantitative trait loci mapping for yield and its components by using two immortalized populations of a heterotic hybrid in Gossypium hirsutum L. Molecular Breeding 29: 297-311. doi:10.1007/s11032-011-9547-0.

Lubbers, E., C. Peng, J. Gannaway, R. Wright, K.M. El-Zik and A.H. Paterson. 2004. Levels and patterns of genetic diversity in upland cotton. Plant and Animal Genome XII Conference. San Diego, CA.

Mauer, F.M. 1954. Origin and systematics of cotton. Origin and systematics of cotton. Izdateljstvo Akad. Nauk Uzbeksk. SSR. Taskent.

McCarty, J., J. Wu and J. Jenkins. 2006. Genetic diversity for agronomic and fiber traits in day-neutral accessions derived from primitive cotton germplasm. Euphytica 148: 283-293. doi:10.1007/s10681-005-9027-x.

McCarty, J.C. and J.N. Jenkins. 2001. Primitive cotton germplasm : yield and fiber traits for 16 day-neutral accessions / Jack C. McCarty, Jr. and Johnie N. Jenkins. Research report / Mississippi Agricultural & Forestry Experiment Station ; vol. 22, no. 19. [Mississippi State, Miss.] : Mississippi Agricultural & Forestry Experiment Station, [2001].

McCarty, J.C., J.N. Jenkins, B. Tang and C.E. Watson. 1996. Genetic analysis of primitive cotton germplasm accessions. Crop Science 36: 581-585.

McCarty, J.C., J.N. Jenkins and J.X. Wu. 2004. Primitive accession derived germplasm by cultivar crosses as sources for cotton improvement: I. Phenotypic values and variance components. Crop Science 44: 1226-1230.

McCarty, J.C., J.N. Jenkins and J. Zhu. 1998. Introgression of Day-Neutral Genes in Primitive Cotton Accessions: I. Genetic Variances and Correlations. Crop Sci. 38: 1425-1428. doi:10.2135/cropsci1998.0011183X003800060002x.

McCarty, J.C., Jr. and J.N. Jenkins. 1992. Cotton germplasm. Characteristics of 79 day-neutral primitive race accessions. Technical Bulletin - Mississippi Agricultural and Forestry Experiment Station. p. 17-17.

McCarty, J.C., Jr., J.N. Jenkins and B. Tang. 1995. Primitive cotton germplasm: variability for yield and fiber traits. Technical Bulletin - Mississippi Agricultural and Forestry Experiment Station. p. 8-8.

McCarty, J.C., Jr., J.N. Jenkins and J. Zhu. 1998. Introgression of day-neutral genes in primitive cotton accessions: II. Predicted genetic effects. Crop Science 38: 1428-1431.

McCarty, J.C., J. Wu and J.N. Jenkins. 2003. Use of primitive accessions of cotton as sources of genes for improving yield components and fiber properties / Jack C. McCarty, Johnie N. Jenkins, and Jixiang Wu. Bulletin / Mississippi Agricultural and Forestry Experiment Station ; 1130. [Mississippi State, Miss.] : Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University, [2003].

McCarty, J.C., J. Wu and J.N. Jenkins. 2005. Potential of primitive accessions for cotton improvement / Jack C. McCarty, Jr., Johnie N. Jenkins, and Jixiang Wu. Bulletin / Mississippi Agricultural and Forestry Experiment Station ; 1141. [Mississippi State, Miss.] : Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University, [2005].

Mei, H., X. Zhu and T. Zhang. 2013. Favorable QTL Alleles for Yield and Its Components Identified by Association Mapping in Chinese Upland Cotton Cultivars. PLoS ONE 8: 1-12. doi:10.1371/journal.pone.0082193.

Mei, M., N.H. Syed, W. Gao, P.M. Thaxton, C.W. Smith, D.M. Stelly, et al. 2004. Genetic mapping and QTL analysis of fiber-related traits in cotton (Gossypium). TAG. Theoretical And Applied Genetics. Theoretische Und Angewandte Genetik 108: 280-291.

Meredith, W.R., Jr. 1991. Contributions of introductions to cotton improvement. In: H. L. Shands and L. E. Wiesner, editors, Use of plant introductions in cultivar development. Part 1. Proceedings of a symposium sponsored by Division C-1 of the Crop Science Society of America in Las Vegas, Nevada, USA, 19 October 1989. Crop Science Society of America, Inc., Madison, Wisconsin; USA. p. 127-146.

Muhammad, S., G. Wangzhen, U. Ihsan, T. Nabila, Z. Yusuf, R. Mehboob ur, et al. 2011. QTL mapping for physiology, yield and plant architecture traits in cotton (Gossypium hirsutum L.) grown under well-watered versus water-stress conditions. Electronic Journal of Biotechnology 14: 1-13. doi:10.2225/vol14-issue3-fulltext-3.

Murfet, I.C. 1977. Environmental Interaction and The Genetics of Flowering. Annual Review of Plant Physiology 28: 253-278. doi:doi:10.1146/annurev.pp.28.060177.001345.

Naveed, M., Q. Abdul, M. Waqas and N. Etrat. 2006. Genetic study of yield of seed cotton and plant height in cotton genotypes. International Journal of Agriculture and Biology 8: 630-635.

Paterson, A.H., R.K. Boman, S.M. Brown, P.W. Chee, J.R. Gannaway, A.R. Gingle, et al. 2004. Reducing the genetic vulnerability of cotton. Crop Science 44: 1900-1901.

Peng, C., E. Lubbers, L. May, J. Gannaway and A. Paterson. 2004. CHANGES IN GENETIC DIVERSITY OF U.S. UPLAND COTTON. Proceedings of the Beltwide Cotton Production Conference: 1054-1054.

Peng, J., D.E. Richards, N.M. Hartley, G.P. Murphy, K.M. Devos, J.E. Flintham, et al. 1999. 'Green revolution' genes encode mutant gibberellin response modulators. Nature 400: 256-261.

Percival, A.E. 1987. The National Collection of Gossypium GermplasmDepartment of Agricultural Communications, Texas A & M University.

Percy, R.G., R.G. Cantrell and J.F. Zhang. 2006. Genetic variation for agronomic and fiber properties in an introgressed recombinant inbred population of cotton. Crop Science 46: 1311-1317.

Pettigrew, W.T. 1994. Source-to-sink manipulation effects on cotton lint yield and yield components. Agronomy journal 86: 731-735.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey, et al. 1996. The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. Molecular Breeding 2: 225-238.

Qin, Y., R. Liu, H. Mei, T. Zhang and W. Guo. 2009. QTL mapping for yield traits in Upland cotton (Gossypium hirsutum L.). Acta Agronomica Sinica 35: 1812-1821.

Reinhardt, D. and C. Kuhlemeier. 2002. Plant architecture. EMBO reports 3: 846-851. doi:10.1093/embo-reports/kvf177.

Reta-Sánchez, D.G. and J.L. Fowler. 2002. Canopy light environment and yield of narrow-row cotton as affected by canopy architecture. Agronomy Journal 94: 1317-1323.

Rong, J., F.A. Feltus, V.N. Waghmare, G.J. Pierce, P.W. Chee, X. Draye, et al. 2007. Meta-analysis of Polyploid Cotton QTL Shows Unequal Contributions of Subgenomes to a Complex Network of Genes and Gene Clusters Implicated in Lint Fiber Development. Genetics 176: 2577-2588. doi:10.1534/genetics.107.074518.

Saleem, M.F., M.F. Bilal, M. Awais, M.Q. Shahid and S.A. Anjum. 2010. Effect of nitrogen on seed cotton yield and fiber qualities of cotton (Gossypium hirsutum L.) cultivars. JAPS, Journal of Animal and Plant Sciences 20: 23-27.

Samuel Yang, S., F. Cheung, J.J. Lee, M. Ha, N.E. Wei, S.-H. Sze, et al. 2006. Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton. The Plant Journal: For Cell And Molecular Biology 47: 761-775.

Sebolt, A.M., R.C. Shoemaker and B.W. Diers. 2000. Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean. Crop Science 40: 1438-1444.

Shappley, Z.W., J.N. Jenkins, J. Zhu and J.C. McCarty, Jr. 1998. Quantitative trait loci associated with agronomic and fiber traits of upland cotton. Journal of Cotton Science 2: 153-163.

Smith, C.W. and J.T. Cothren. 1999. Cotton : origin, history, technology, and production / editors, C. Wayne Smith, J. Tom CothrenNew York : Wiley, c1999.

Song, X. and T. Zhang. 2009. Quantitative trait loci controlling plant architectural traits in cotton. Plant Science 177: 317-323.

Stephens, J.C., F.R. Miller and D.T. Rosenow. 1967. Conversion of alien sorghums to early combine genotypes. Crop Science 7: 396-396.

Swindle, M.G. 1993. Performance of F2 Hybrids in Cotton from Crosses of Exotic Germplasm and CultivarsMississippi State University. Department of Agronomy.

Uga, Y., Y. Nonoue, Z.W. Liang, H.X. Lin, S. Yamamoto, U. Yamanouchi, et al. 2007. Accumulation of additive effects generates a strong photoperiod sensitivity in the extremely late-heading rice cultivar 'Nona Bokra'. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik 114: 1457-1466.

Ulloa, M., W.R. Meredith Jr, Z.W. Shappley and A.L. Kahler. 2002. RFLP genetic linkage maps from four F2.3 populations and a joinmap of Gossypium hirsutum L. Theoretical & Applied Genetics 104: 200-208.

Ulloa, M. and W.R. Meredith, Jr. 2000. Genetic linkage map and QTL analysis of agronomic and fiber quality traits in an intraspecific population. Journal of Cotton Science 4: 161-170.

USDA. 2013. Cotton and Wool. In: L. Meyer, editor Economic Research Service. USDA.

Van Esbroeck, G. and D.T. Bowman. 1998. Cotton germplasm diversity and its importance to cultivar development [electronic resource]. Journal of cotton science (Online) Journal of cotton science 2: 121-129. doi:http://www.jcotsci.org/toc.html.

Verhalen, L.M. and J.C. Murray. 1967. Diallel analysis of several fiber property traits in Upland cotton Gossypium hirsutum L. Crop science 7: 501-505.

Verhalen, L.M., J.C. Murray, E.S. Oswalt, J.W. Simmons, M. Simmons and K.C. Fun. 1969. Cotton variety tests in Oklahoma 1963-67. Bulletin, Agricultural Experiment Station, Oklahoma State University. p. 34-34.

Waddle, B.M., C.F. Lewis and T.R. Richmond. 1961. The genetics of flowering response in cotton. III. Fruiting behaviour of Gossypium hirsutum race lati-folium in a cross with a variety of cultivated American upland cotton. Genetics 46: 427-437.

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

Wang, B., Y. Wu, N. Huang, X. Zhu, W. Guo and T. Zhang. 2006. QTL mapping for plant architecture traits in upland cotton using RILs and SSR markers. Acta Genetica Sinica 33: 161-170.

Wang, M.I.N., C. Li and Q. Wang. 2014. Quantitative trait loci mapping and genetic dissection for lint percentage in upland cotton (Gossypium hirsutum). Journal of Genetics 93: 371-378. doi:10.1007/s12041-014-0385-9.

Wang, Z., D. Zhang, X. Wang, X. Tan, H. Guo and A.H. Paterson. 2013. A whole-genome DNA marker map for cotton based on the D-genome sequence of Gossypium raimondii L. G3 (Bethesda, Md.) 3: 1759-1767. doi:10.1534/g3.113.006890.

Wendel, J.F. 1989. New World Tetraploid Cottons Contain Old World Cytoplasm. 11): 4132.

Wendel, J.F. and V.A. Albert. 1992. Phylogenetics of the Cotton Genus (Gossypium): Character-State Weighted Parsimony Analysis of Chloroplast-DNA Restriction Site Data and Its Systematic and Biogeographic Implications. Systematic Botany: 115. doi:10.2307/2419069.

Wright, R.J., P.M. Thaxton, K.M. El-Zik and A.H. Paterson. 1999. Molecular mapping of genes affecting pubescence of cotton. Journal of Heredity 90: 215-219.

Wu, C.T., B.L. Zhou and T.Z. Zhang. Isolation and characterization of a sterile-dwarf mutant in Asian cotton (Gossypium arboreum L.).

Wu, J.X., J.C. McCarty, S. Saha, J.N. Jenkins and R. Hayes. 2009. Genetic changes in plant growth and their associations with chromosomes from Gossypium barbadense L. in G. hirsutum L. Genetica 137: 57-66.

Zeng, L.H., W.R. Meredith, D.L. Boykin and E. Taliercio. 2007. Evaluation of an exotic germplasm population derived from multiple crosses among Gossypium tetraploid species. Journal of Cotton Science 11: 118-127.

Zeng, L.H., W.R. Meredith, Jr., O.A. Gutiérrez and D.L. Boykin. 2009. Identification of associations between SSR markers and fiber traits in an exotic germplasm derived from multiple crosses among Gossypium tetraploid species. TAG Theoretical and Applied Genetics 119: 93-103.

Zhang, J., W. Guo and T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L. x Gossypium barbadense L.) with a haploid population. TAG. Theoretical And Applied Genetics. Theoretische Und Angewandte Genetik 105: 1166-1174.

Zhang, J.F. and R. Percy. 2007. Improving upland cotton by introducing desirable genes from pima cotton. World Cotton Research Conference-4, Lubbock, Texas, USA, 10-14 September 2007. International Cotton Advisory Committee (ICAC), Washington; USA. p. un-un.

Zhang, T., N. Qian, X. Zhu, H. Chen, S. Wang, H. Mei, et al. 2013. Variations and transmission of QTL alleles for yield and fiber qualities in upland cotton cultivars developed in China. PLoS ONE 8: e57220-e57220.

Zhang, Z., J. Rong, V.N. Waghmare, P.W. Chee, O.L. May, R.J. Wright, et al. 2011. QTL alleles for improved fiber quality from a wild Hawaiian cotton, Gossypium tomentosum. TAG Theoretical and Applied Genetics 123: 1075-1088.

CHAPTER 2

TARGETED IDENTIFICATION OF ASSOCIATION BETWEEN FIBER QUALITY TRAITS AND MICROSATELLITE MARKERS

[1]Adhikari J., Z. Wang, S. Khanal, R. Chandnani, J.D. Patel, L. Goff, D. Jones and A.H. Paterson. To be submitted to *Theoretical and Applied Genetics*.

Abstract

Primitive and wild accessions of cotton are potential sources of favorable alleles for genetic improvement- enriching genetic variation in the narrow elite gene pool of this important crop. In this study, three exotic accessions of cotton converted to day-neutral flowering were used to create experimental populations in four elite US cultivar backgrounds. These populations were screened with microsatellite markers selected from "hotspots" for fiber quality quantitative trait loci (QTLs) in the cotton genome and single marker analysis was performed to identify significant association of the markers with six fiber quality traits. A total of 134 nominal marker-trait associations were identified, out of which 15 associations were significant after Bonferroni correction for multiple comparisons. In 67 out of 134 putative associations identified, the exotic parents contributed favorable alleles to multiple backgrounds and for multiple traits, in addition to the traits for which they were selected. These results indicate that utilization of exotic and wild accessions of cotton is useful in introducing favorable alleles for cotton improvement.

Introduction

Cotton (*Gossypium spp*) is an important cash crop grown in the United States and many parts of the globe for its spinnable fiber. Known to the agricultural world as "White Gold" for its soft fluffy staple fiber, cotton is the major source of natural fiber for the textile industry. With the advent of highly sophisticated and efficient spinning technologies, there is a need to improve the quality of cotton fiber. Improvement of any crop requires diverse genetic resources. However, the worldwide Upland cotton gene pool is genetically impoverished due to a series of bottlenecks imposed by polyploid formation, domestication and migration.

The narrow genetic base of cotton has been considered as one of the major obstacles in cotton improvement. In addition, overexploitation of a few genetic backgrounds by breeders may

have contributed to a plateau of yield and fiber quality traits (Esbroeck and Bowman 1998). Such a narrow genetic base might result in a crop being highly vulnerable to stresses (McCarty, Jenkins et al. 1998). A decline in the genetic diversity of upland cotton and the need to broaden the genetic base of cotton germplasm useful for the improvement of lint yield, fiber quality and biotic or abiotic stresses has been widely reported (Esbroeck, Bowman et al. 1999, Bowman 2000, Iqbal, Reddy et al. 2001, Gutierrez, Shoemaker et al. 2002).

One solution to the problem of cotton genetic vulnerability might be the exploration of exotic genotypes (Paterson, Boman et al. 2004). Incorporating favorable alleles, genes or gene complexes from wild relatives or accessions has been a high strategic priority for practical improvement of some crops (Hajjar and Hodgkin 2007). Research has shown that primitive accessions of cotton are highly diverse and have useful genetic variability (Percival 1987, Meredith 1991, McCarty and Jenkins 1992, McCarty, Jenkins et al. 1995, McCarty, Jenkins et al. 1998, McCarty, Jenkins et al. 1998, McCarty and Jenkins 2001, McCarty, Wu et al. 2003). In practice, however, the utilization of genetic variability from primitive cotton accessions has been limited due to their photoperiod response (Stephens, Miller et al. 1967, Holley and Goodman 1988, McCarty and Jenkins 1992, Uga, Nonoue et al. 2007). Conversion of primitive usually short-day flowering genotypes into day-neutral forms by repeated backcrossing has been undertaken to reduce the obstacles to use this germplasm.

The development of molecular markers has accelerated the process of selection and improvement of traits of interest. Conventional breeding has played an important role in the improvement of yield and fiber quality in Upland cotton. However, the progress made by convention breeding schemes is very slow. Since the first restriction fragment length polymorphism (RFLP) linkage map of cotton has been reported (Reinisch, Dong et al. 1994),

many genetic maps have been developed for Gossypium intra-specific and inter-specific populations (Jiang, Wright et al. 1998, Wright, Thaxton et al. 1999, Jiang, Wright et al. 2000, Ulloa and Meredith 2000, Ulloa, Meredith Jr et al. 2002, Zhang, Guo et al. 2002, Shen, Guo et al. 2005, Wang, Wu et al. 2006, Shen, Guo et al. 2007, Qin, Liu et al. 2009, Muhammad, Wangzhen et al. 2011) and QTLs linked to different fiber quality, yield and yield related, and other agronomic and economic traits have been mapped. Identification of molecular markers linked to these QTLs or traits of interest is essential in breeding or selection of these traits.

In this study, we have selected SSR markers from "hotspots" for fiber quality traits on the basis of meta-analysis of QTLs in cotton (Rong, Feltus et al. 2007) and a whole genome marker map of cotton based on the D-genome sequence of *G. raimondii* (Wang, Zhang et al. 2013). Association of these markers with six different fiber quality traits was studied using different populations and significant associations are reported.

Materials and methods

Plant materials

We selected three converted accessions, MDN063, MDN101 and MDN257 based upon their phenotypes for yield, lint percentage, boll size, micronaire (fiber fineness), 2.5% span length (fiber length), elongation, and strength (Table 2.1). MDN063 was converted from race T0063, which was collected in the state of Chiapas, Mexico. It contributed large and positive additive genetic variation to progeny of crosses with Deltapine 16 (DPL 16) for increased boll size, reduced micronaire (indicative of finer fiber) and greater fiber elongation (Table 2.2). It ranked first among 79 lines tested for lowest micronaire and its progeny also ranked first among 79 progeny populations tested. It was also highly-ranked for boll size and fiber elongation. MDN101 was converted from race T101, which was collected in the state of Jutiapa, Guatemala.

It contributed large and positive additive genetic variation to progeny of crosses with DPL16 for increased fiber length, higher lint percentage, increased fiber strength, and reduced plant height (Table 2.2). Its progeny ranked first of the 79 populations tested for a "selection index" reflecting yield, fiber strength and lint percentage (McCarty, Jenkins et al. 1996). MDN257 was converted from race T257, which was collected in the state of Oaxaca, Mexico. It ranked first among 79 progeny tested for fiber length and its progeny also ranked first among 79 progeny population tested. Although this genotype is below average for some key attributes (Table 2.2), it brings additional botanical and geographic diversity to the sampling of genotypes. Four elite US genotypes, PD94042, DES56, PMHS200 and Acala Maxxa, respectively sampling the Eastern, Delta, Plains and Acala production regions of the US cotton genetic pool were evaluated for effects of QTLs from the exotic lines.

Development of experimental populations

Crosses were made in 2010-11 in the green house, between the day neutral exotic lines (MDN063, MDN101 and MDN257) and four elite *Gossypium hirsutum* genotypes (DES56, Acala Maxxa, PD94042 and PMHS200) in different combinations (Table S1). Bolls were hand harvested and collected separately for individual crosses. All samples were hand-ginned and delinted.

The resulting F_1 and the parents were planted at The University of Georgia Plant Sciences Farm, Watkinsville, Georgia. Standard cultivation practices were applied including irrigation, fertilization and pesticide application. As many flowers as possible from each individual F_1 plant were self-pollinated, and DNA of F_1 plants was checked with several SSR markers to verify hybridity. Selfed bolls from verified F_1 plants were hand collected separately for F_2 seed. Selfed bolls from the F_2 plants were also collected individually to obtain $F_{2:3}$ seed. $F_{2:3}$ progeny plants

for 2010 populations were grown in 2012 and those for 2011 populations were grown in 2013 in plots of 10 seeds. There were two planting dates for F_2 plants derived from populations created in 2010.

Sample collection and data analysis

Fiber samples from F_2 plants comprised of up to 25 open pollinated bolls from individual plants. For the $F_{2:3}$ progeny plots, 25 bolls were collected to sample all the plants in each plot. After ginning, samples were sent to Cotton Inc. for testing of six fiber quality traits [Micronaire or fiber fineness (MIC), Upper half mean length or fiber length (UHM) , Fiber uniformity index (UI), Fiber strength (STR), Fiber elongation (ELO) and Short fiber content (SFC)]. Data were analyzed using R statistical software. Single marker analyses (SMA) for fiber quality traits were done using Windows QTL Cartographer 2.5 (Wang, Basten et al. 2012). Bonferroni correction was used to correct for multiple comparisons.

DNA extraction and SSR assays

Genomic DNA was isolated from young unopened leaves from each F_2 plant from all the populations and from the parental samples using a modified CTAB method (Paterson, Brubaker et al. 1993). A total of 1-2 g of fresh tissue was ruptured in a 1:1 mixture of cotton lysis buffer and extraction buffer. After leaving the tubes in a 65°C water bath for about an hour, the extraction materials were purified twice with 800 µl of chloroform iso-amyl alcohol. DNA was then precipitated with 500 µl of isopropanol, cleaned with two washes of 75% ethanol (500µl each) and centrifuged. The clean dried DNA was dissolved in 200 µl of TE buffer. The extracted DNA was checked for quality and quantity and stored at -20˚C.

The SSR technique was used to identify polymorphic markers. We selected eighteen "hotspot" regions from the cotton genome based on their richness in fiber quality QTLs (Table

S6) to select SSR markers for genotyping of the populations. 720 SSR primer pairs from these regions were used to screen polymorphism among the parents of the mapping populations. SSR amplifications used the Touchdown PCR reaction program: 1. Seven cycles of four steps (94°C for 4 minutes, 94°C for 40 seconds, 58°C for 1 min and 72°C for 1 min); 2. Thirty six cycles of three steps (94°C for 45 seconds, 54°C for 45 seconds and 72°C for 1 minute); 3. One cycle of 72°C for 10 minutes, and 4. Hold at 4°C until the PCR plates are removed. PCR products were separated using acrylamide gel electrophoresis and visualized using silver staining (Bassam and Gresshoff 2007).

Results

Performance of parents, F_2 and $F_{2:3}$ progenies

The mean parental values for six fiber quality traits evaluated in this study are shown in Supplemental Table S1. Parents displayed wide variation for these traits over three years (Figures 2.1 and 2.2, Table S2). MDN101, which was selected for fiber length (UHM), showed better fiber lengths than DES56 and PMHS200 in 2011. Likewise, MDN063, selected for fiber fineness (MIC) and fiber elongation (ELO), showed better fiber elongation than DES56 both in 2011 and 2012. It also showed better performance than Acala Maxxa for fiber fineness in 2011. The mean F_2 and $F_{2:3}$ values for different fiber quality traits in the thirteen populations are presented in supplemental tables S3 and S4 respectively. The absolute values of skewness for all six fiber quality traits in all the populations were less than 1 (not shown), indicating approximately normal distributions of the traits.

The genetics and correlation among fiber quality traits

We revealed strong positive correlation between fiber length (UHM) and fiber uniformity index (UI); UHM and fiber strength (STR); and UI and STR, for all individual populations (data

not shown) as well as when data from all populations were pooled (Table 2.3). Fiber fineness (MIC) and fiber length (UHM) were negatively correlated, a good outcome as reduced micronaire is indicative of higher fiber quality. Short fiber content (SFC %) was negatively correlated with all other fiber quality traits for all the populations. A strong negative correlation between UI and SFC was revealed in all the populations. These results suggest that synchronous improvement of different fiber quality traits is possible in these populations.

We also calculated the heritability of these fiber quality traits on the basis of regression of trait values of $F_{2:3}$ progenies on trait values on corresponding F_2 progenies. Medium to highly significant positive regression values were found for all traits except SFC% in all the populations (Table 2.4). This indicates very high heritability of these traits in the respective populations and the suitability of early generation evaluation of these populations for diagnostic DNA markers.

Genomic distribution of polymorphic SSR markers

We screened 720 SSR primer pairs selected from eighteen "hotspot" genomic regions based on the whole genome marker map of cotton based on *G. raimondii* (Wang, Zhang et al. 2013). Among the 720 SSR primer pairs tested for polymorphism, 82 (11.38%) were polymorphic between different parental combinations, listed in Table S4. Polymorphic markers were identified for all hotspots except "hotspot VIII". For two QTL hotspots, only one polymorphic SSR marker was identified. For all other hotspots, multiple polymorphic markers were identified permitting us to sample significant proportions of most of the hotspots.

Association of polymorphic markers with fiber quality traits

The number of marker loci tested for association with phenotypes was different in different populations. For each population, there were the same numbers of statistical tests as the number of amplified loci. In such a condition, the experiment-wise Type I error rate would be

much higher than the nominal significance rate of any single test (0.05). To overcome this problem, the Bonferroni correction ($P \leq 0.05/N$, where N is the number of statistical tests/amplified loci used in each population) was used to obtain an appropriate significance threshold. In total, there were 134 nominal associations between different traits and the marker loci tested in different populations (Table S7), 15 of which remained after Bonferroni correction (Table 2.5).

Twenty five nominal associations were found for micronaire in total, with two remaining significant after Bonferroni correction. One of these two significant associations was identified in the MDN101 × PMHS200 population with the DPL0156_230 marker loci in the year 2011, explaining 16 percent of the phenotypic variation. The other significant association was identified in the MDN101 × Acala Maxxa with the CIR141_250 SSR, explaining 12 percent of the phenotypic variation. Among nominal associations that did not survive the Bonferroni correction, some gained support from multiple independent discoveries. BNL3359_270 was significant in the MDN101 × Acala Maxxa population in both 2011 and 2012. NAU2152_240 was significant in two populations, MDN101 × PD94042 and MDN063 × DES56, in 2012. The phenotypic variation explained by the nominal associations ranged from 9 to 27 percent.

Nineteen nominal associations were identified for fiber length, with three remaining significant after Bonferroni correction. DPL0270_150 was significantly associated with fiber length in the MDN101 × DES56 population in 2011, explaining 33 percent of the phenotypic variation. DPL0378_510 and JESPR37_1000 were associated with fiber length in the MDN257 × DES56 population in 2012, explaining 18 and 19 percent of phenotypic variation respectively. Among nominal associations that did not survive the Bonferroni correction, some gained support from multiple independent discoveries. NAU1042_270 was associated with fiber length in two

populations, MDN101 × PD94042 and MDN101 × PMHS200, in 2011. The phenotypic variation explained by the nominal associations ranged from 2 to 33 percent.

Fourteen nominal associations were identified for fiber uniformity index, with four remaining significant after Bonferroni correction. JESPR101_280 and NAU5120_170 were significant in the MDN101 × DES56 population in 2011, explaining 34 and 46 percent of the phenotypic variation respectively. NAU5120_450 was significant in the MDN101 × Acala Maxxa population in 2012, explaining 13 percent of the phenotypic variation; and in the MDN063 × DES56 population in 2011, explaining 14 percent of the phenotypic variation. Among nominal associations that did not survive the Bonferroni correction, some gained support from multiple independent discoveries. NAU5120_450 was associated with fiber uniformity ratio in the MDN063 × DES56 population for both the planting dates in 2011. The phenotypic variation explained by the nominal associations ranged from 6 to 46 percent.

Twenty two significant associations were identified for fiber strength, with only one remaining significant after Bonferroni correction. NAU3820_120 was associated with fiber strength in the MDN101 × DES56 population in 2012, explaining 10 percent of the phenotypic variation. Among nominal associations that did not survive the Bonferroni correction, some gained support from multiple independent discoveries. NAU3820_120 was significant in two populations, MDN101 × DES56 in 2012 and MDN101 × PMHS200 in 2011. MUSB1020_650 was also significant in two populations, MDN063 × DES56 and MDN063 × Acala Maxxa in 2012. The phenotypic variation explained by the nominal associations ranged from 3 to 23 percent.

Twenty one significant associations were identified for fiber elongation, three of which remained significant after Bonferroni correction. JESPR101_280, NAU5120_170 and

NAU5465_220 were associated with fiber elongation in the MDN101 × DES56 population in 2011, explaining 43, 47 and 40 percent of phenotypic variation respectively. Among nominal associations that did not survive the Bonferroni correction, some gained support from multiple independent discoveries. NAU1366_450 was significant for both planting dates in the MDN101 × DES56 population in 2011 and also in 2012 for the same population, which indicates particularly strong association of this marker locus with fiber elongation. The phenotypic variation explained by the nominal associations ranged from 4 to 45 percent.

Twenty six significant associations were identified for short fiber content, two of which remained significant even after Bonferroni correction. NAU5120_170 was associated with short fiber content in the MDN101 × DES56 population in 2011, explaining 37 percent of the phenotypic variation. DPL0378_180 was associated with short fiber content in the MDN101 × Acala Maxxa population in 2011, explaining 19 percent of the phenotypic variation. Among nominal associations that did not survive the Bonferroni correction, some gained support from multiple independent discoveries. NAU5120_450 was significant in the MDN063 × DES56 population in both the planting dates in 2011. The same locus was significant in the MDN101 × Acala Maxxa population in 2012, indicating stability of the association in multiple populations and over multiple planting dates/environments. The phenotypic variation explained by the nominal associations ranged from 4 to 37 percent.

Associations that were significant over multiple years indicate stability of the potential QTLs (Table 2.6). Marker locus DPL0279_190 was associated with MIC for the population MDN101 × Acala Maxxa in both 2011 and 2012, albeit both of these associations were only nominal. Likewise, NAU1366_450 was strongly linked with ELO for MDN101 × DES56 in both 2011 and 2012. These associations were also nominal. Similarly, some associations were

significant over different planting dates (Table 2.6). Marker locus NAU5120_450 was strongly linked with UI and SFC in the population MDN063 × DES56 for two planting dates in 2011. The association of NAU5120_450 with UI survived the Bonferroni correction for the second planting date.

Alleles from the exotic parent MDN101, the only one tested in all four cultivar backgrounds, were significant for different traits in different backgrounds (Table 2.6). Marker locus NAU3820_120 from the exotic parent was significant in two genetic backgrounds PMHS200 (nominally) and DES56 (even after Bonferroni correction) for STR. Similarly, NAU1042_220 from the same exotic parent was nominally associated with UHM in PMHS200 and PD94042 backgrounds. Likewise, BNL1317_230 from the same exotic parent was nominally associated with STR in two different genetic backgrounds DES56 and PMHS200. Since the different elite cultivars represent different cotton production regions of the United States, this indicates that alleles from the exotic donor may have different value in different elite cotton backgrounds and production regions.

Some marker loci were significantly associated with multiple traits in the same population (Table 2.6). For example, BNL1317_220 was significantly associated with UHM, STR and SFC for MDN101 × PMHS200 in the year 2012. Similarly, BNL3359_390 was significantly associated with UHM, UI, STR and SFC for MDN257 × DES56 in the year 2012. Likewise, DPL0270_150 was significantly associated with ELO, MIC, UHM and UI for MDN101 × DES56 in the year 2011. Among these associations, only the association of DPL0270_150 with UHM in the MDN101 × DES56 population survived Bonferroni correction, the rest being only nominally significant. Nonetheless, this suggests potential usefulness of

exotic donors for simultaneous improvement of different fiber quality traits following the introduction of these alleles into the cultivated genetic background.

Contribution of exotic parents

The three exotic parents (MDN063, MDN101 and MDN257) used in our study were selected based on the additive genetic variance contributed by them in crosses with DPL16 for various traits (Table 2.2, summarized from McCarty, Jenkins, et al., 1996). Our results show that these parents contributed positive additive effects across different genetic backgrounds not only for the fiber quality traits they were selected for, but also for other fiber quality traits (Table 2.5 and Table S7).

MDN063, selected for its positive effect for reduced fiber fineness and increased fiber elongation in cross with DPL16, contributed favorable alleles (albeit only nominal) for these two traits in crosses with DES56 and for fiber fineness in cross with Acala Maxxa, confirming the selection of this parent for these traits. In addition, it also contributed favorable alleles (nominal) for other traits like fiber strength, fiber uniformity ratio and short fiber content (in crosses with DES56 and Acala Maxxa) and for fiber length in crosses with DES56.

MDN101 was selected for fiber length and fiber strength as it contributed positive additive genetic variance for these traits in crosses with DPL16. Our results show that it contributed positive additive effect for fiber length in crosses with PMHS200 and for fiber strength in crosses with PD94942, DES56 and PMHS200, confirming the selection of this parent for the selected traits. Among the favorable alleles contributed by this parent, the association of NAU3820_120 with STR, NAU5465_220 with ELO and NAU5120_170 with SFC met the Bonferroni standard for statistical significance in DES56 background. Similarly, the association of DPL0378_180 with SFC met the Bonferroni standard in PMHS200 background. In addition, it

showed nominal positive additive effects for fiber elongation in all four genetic backgrounds and for fiber fineness (in crosses with PD94042, DES56 and Acala Maxxa), fiber uniformity ratio (in cross with DES56) and for short fiber content across all four backgrounds.

MDN257 was selected for fiber length for its positive additive effect in crosses with DPL16. Two favorable alleles from this parent, DPL0378_510 and JESPR37_1000, which showed significant association with UHM (surviving the Bonferroni correction), contributed positive additive effects for this trait. In addition, it showed nominal positive additive effects for fiber elongation and short fiber content (in crosses with DES56 and Acala Maxxa) and for fiber fineness, fiber strength and fiber uniformity ration in cross with DES56.

Discussion

Knowledge of the genetic diversity within a crop gene pool is of great importance. The probability of recovering superior genotypes in segregating generations is greater when combinations of parents having different complementary alleles are taken as a basis. The use of parent plants with low genetic diversity in the formation of populations reduces genetic variability, making the selection of superior genotypes difficult (Rotili, Cancellier et al. 2012). Moreover, inclusion of diverse parents in any experimental population facilitates the identification of greater numbers of polymorphic markers and thus increases the probability of getting larger numbers of significant marker-trait associations.

In our study, diverse parental lines were used to create experimental populations. Parents not only showed diverse phenotypes but also differed by more SSR marker polymorphisms than most elite cottons. Among the 15 associations meeting the Bonferroni threshold for significance, four involved favorable alleles from exotic parents for different fiber quality traits. Exotic parents generally showed positive additive effects for traits for which they were selected based

on prior information (McCarty, Jenkins, et al., 1996). For example, MDN101 showed positive additive effects for STR and MDN257 for UHM. Furthermore, positive additive effects (Bonferroni significant) were observed not only for the phenotypes for which the parents were selected, but also for other phenotypes scored in our study. MDN101, which was selected for UHM and STR, contributed favorable alleles for ELO. Exotic parents also contributed favorable alleles to multiple genetic backgrounds, some of which were nominal and some met the Bonferroni standard. Marker locus NAU3820_120, which was contributed by MDN101 for STR, met the Bonferroni standard in DES56 background while it was nominal in the PMHS200 background. Although not all the favorable alleles met the Bonferroni standard, evidence that some exerted favorable effects in multiple backgrounds reduces the likelihood that they are false positives, and indicates the usefulness of the parents in those backgrounds for different fiber quality traits.

Use of diverse parents allowed us to identify large numbers of nominal associations between polymorphic SSR markers and the six fiber quality traits. Of the 134 associations identified in the study, there were 67 (50%) in which favorable alleles were from the exotic parent. While only four of the 67 favorable alleles survived the Bonferroni correction, these results support the hypothesis that exotic *G. hirsutum* lines confer favorable alleles to elite cottons that can contribute to fiber quality improvement. Moreover, one would postulate that exotic alleles may also improve other traits of agronomic interest, like disease resistance, insect tolerance or drought resistance, which might have eroded from the cultivated gene pool of cotton owing to decades of selection for high yield and fiber quality.

Among the 15 marker-trait associations that met the Bonferroni significance threshold, 12 included the exotic parent MDN101, wherein the parent contributed favorable alleles in 2 cases.

Two significant associations included the exotic parent MDN257, with the parent contributing favorable alleles in both cases; and one significant association included the exotic parent MDN063 where the parent did not contribute any favorable alleles. For 10 of the 15 significant associations, the associated marker loci were located in the D sub-genome and 5 were located in the A sub-genome. Similarly, 3 of the 4 favorable marker alleles were located in the D sub-genome and only one in the A sub-genome. While the sample sizes are small, these patterns continue to support the hypothesis that the D sub-genome, from an ancestor that did not produce spinnable fiber, nonetheless makes an important contribution to the fiber quality of elite tetraploid cottons (Jiang, Wright, et al., 1998; Rong, Feltus, et al., 2007).

Even with small numbers of markers from the fiber QTL hotspots, we identified large numbers of associations for different fiber quality traits for selected region in the cotton genome. Increasing the number of DNA markers within these regions as well as selecting markers throughout the genome could result in identification of larger numbers of significant associations. Comparing associations identified through markers selected by targeting certain regions with associations identified through markers selected evenly from the whole genome would also allow us to test the hypothesis that the non-random concentrations of QTLs found in interspecific cotton crosses (Rong, Feltus, et al., 2007) translate well to these intra-specific crosses. Progress toward low-cost discovery of SNPs (Kim, Hui, et al., 2015) may make this goal realistic in the near future.

Conclusion

In this study, three exotic *G. hirsutum* accessions converted to day-neutral flowering were used as parents in crosses with four elite US cultivars to develop experimental populations in two different years in different combinations. We studied F_2 and $F_{2:3}$ generations of these populations

for six different fiber quality traits and used 85 different SSR markers selected from 18 QTL "hotspot" regions in the cotton genome to identify significant associations of the primer pairs with fiber quality traits. Our results showed useful contribution of the exotic parents for different fiber quality traits for which they were selected across different genetic backgrounds used in this study as well as for other fiber quality traits that were scored in the study. Some significant and some nominal associations were identified in different populations in different years and across different backgrounds for same exotic parents. These associations could be potential QTLs for the selected traits in selected cotton production regions. Increasing the marker density and using interval mapping in these regions, potential QTLs with significant effects could be identified.

References

Bassam, B.J. and P.M. Gresshoff. 2007. Silver staining DNA in polyacrylamide gels. Nature protocols 2: 2649-2654. doi:10.1038/nprot.2007.330.

Bowman, D.T. 2000. Attributes of public and private cotton breeding programs [electronic resource]. Journal of cotton science (Online) Journal of cotton science 4: 130-136. doi:http://www.jcotsci.org/toc.html.

Esbroeck, G.A.v., D.T. Bowman, O.L. May and D.S. Calhoun. 1999. Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars. Crop Science 39: 323-328.

Esbroeck, G.v. and D.T. Bowman. 1998. Cotton germplasm diversity and its importance to cultivar development. Journal of Cotton Science 2: 121-129.

Gutierrez, O.A., D.B. Shoemaker, C.L. Cheatham, J.C. McCarty, Jr., S. Basu, S. Saha, et al. 2002. Genetic distance among selected cotton genotypes and its relationship with F2 performance. Crop science 42: 1841-1847.

Hajjar, R. and T. Hodgkin. 2007. The use of wild relatives in crop improvement: a survey of developments over the last 20 years. Euphytica 156: 1-13. doi:10.1007/s10681-007-9363-0.

Holley, R.N. and M.M. Goodman. 1988. Yield potential of tropical hybrid maize derivatives. Crop Science 28: 213-218.

Iqbal, M.J., O.U.K. Reddy, K.M. El-Zik and A.E. Pepper. 2001. A genetic bottleneck in the 'evolution under domestication' of upland cotton *Gossypium hirsutum* L. examined using DNA fingerprinting. Theor Appl Genet 103: 547-554. doi:10.1007/PL00002908.

Jiang, C., R.J. Wright, K.M. El-Zik and A.H. Paterson. 1998. Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). Proceedings of the National Academy of Sciences of the United States of America 95: 4419-4424.

Jiang, C., R.J. Wright, S.S. Woo, T.A. DelMonte and A.H. Paterson. 2000. QTL analysis of leaf morphology in tetraploid *Gossypium* (cotton). Theoretical & Applied Genetics 100: 409-418.

Kim C, Guo H, Kong W, Chandnani R, Shuang L, Paterson A.H. Review: Application of genotyping by sequencing technology to a variety of crop breeding programs. *Plant Science* [serial online]. January 1, 2015; Available from: ScienceDirect, Ipswich, MA. Accessed June 7, 2015.

McCarty, J.C. and J.N. Jenkins. 2001. Primitive cotton germplasm : yield and fiber traits for 16 day-neutral accessions / Jack C. McCarty, Jr. and Johnie N. Jenkins. Research report / Mississippi Agricultural & Forestry Experiment Station ; vol. 22, no. 19. [Mississippi State, Miss.] : Mississippi Agricultural & Forestry Experiment Station, [2001].

McCarty, J.C., J.N. Jenkins, B. Tang and C.E. Watson. 1996. Genetic analysis of primitive cotton germplasm accessions. Crop Science 36: 581-585.

McCarty, J.C., J.N. Jenkins and J. Zhu. 1998. Introgression of Day-Neutral Genes in Primitive Cotton Accessions: I. Genetic Variances and Correlations. Crop Sci. 38: 1425-1428. doi:10.2135/cropsci1998.0011183X003800060002x.

McCarty, J.C., Jr. and J.N. Jenkins. 1992. Cotton germplasm. Characteristics of 79 day-neutral primitive race accessions. Technical Bulletin - Mississippi Agricultural and Forestry Experiment Station. p. 17-17.

McCarty, J.C., Jr., J.N. Jenkins and B. Tang. 1995. Primitive cotton germplasm: variability for yield and fiber traits. Technical Bulletin - Mississippi Agricultural and Forestry Experiment Station. p. 8-8.

McCarty, J.C., Jr., J.N. Jenkins and J. Zhu. 1998. Introgression of day-neutral genes in primitive cotton accessions: II. Predicted genetic effects. Crop Science 38: 1428-1431.

McCarty, J.C., J. Wu and J.N. Jenkins. 2003. Use of primitive accessions of cotton as sources of genes for improving yield components and fiber properties / Jack C. McCarty, Johnie N. Jenkins, and Jixiang Wu. Bulletin / Mississippi Agricultural and Forestry Experiment Station ; 1130. [Mississippi State, Miss.] : Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University, [2003].

Meredith, W.R., Jr. 2000. Cotton yield progress - why has it reached a plateau? Better Crops with Plant Food 84: 6-9.

Meredith, W.R., Jr. 1991. Contributions of introductions to cotton improvement. In: H. L. Shands and L. E. Wiesner, editors, Use of plant introductions in cultivar development.

Part 1. Proceedings of a symposium sponsored by Division C-1 of the Crop Science Society of America in Las Vegas, Nevada, USA, 19 October 1989. Crop Science Society of America, Inc., Madison, Wisconsin; USA. p. 127-146.

Muhammad, S., G. Wangzhen, U. Ihsan, T. Nabila, Z. Yusuf, R. Mehboob ur, et al. 2011. QTL mapping for physiology, yield and plant architecture traits in cotton (*Gossypium hirsutum* L.) grown under well-watered versus water-stress conditions. Electronic Journal of Biotechnology 14: 1-13. doi:10.2225/vol14-issue3-fulltext-3.

Paterson, A.H., R.K. Boman, S.M. Brown, P.W. Chee, J.R. Gannaway, A.R. Gingle, et al. 2004. Reducing the genetic vulnerability of cotton. Crop Science 44: 1900-1901.

Paterson, A.H., C.L. Brubaker and J.F. Wendel. 1993. A rapid method for extraction of cotton (*Gossypium spp.*) genomic DNA suitable for RFLP or PCR analysis. Plant Molecular Biology Reporter 11: 122-127.

Percival, A.E. 1987. The National Collection of *Gossypium* GermplasmDepartment of Agricultural Communications, Texas A & M University.

Qin, Y., R. Liu, H. Mei, T. Zhang and W. Guo. 2009. QTL mapping for yield traits in Upland cotton (*Gossypium hirsutum* L.). Acta Agronomica Sinica 35: 1812-1821.

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

Rong, J., F.A. Feltus, V.N. Waghmare, G.J. Pierce, P.W. Chee, X. Draye, et al. 2007. Meta-analysis of Polyploid Cotton QTL Shows Unequal Contributions of Subgenomes to a Complex Network of Genes and Gene Clusters Implicated in Lint Fiber Development. Genetics 176: 2577-2588. doi:10.1534/genetics.107.074518.

Rotili, E.A., L.L. Cancellier, M.A. Dotto, J.M. Peluzio and E.V.d. Carvalho. 2012. Genetic divergence in populations of maize, in Tocantins State, Brazil. / Divergência genética em genótipos de milho, no Estado do Tocantins. Revista Ciência Agronômica 43: 516-521.

Shappley, Z.W., J.N. Jenkins, J. Zhu and J.C. McCarty, Jr. 1998. Quantitative trait loci associated with agronomic and fiber traits of upland cotton. Journal of Cotton Science 2: 153-163.

Shen, X., W. Guo, Q. Lu, X. Zhu, Y. Yuan and T. Zhang. 2007. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica 155: 371-380. doi:10.1007/s10681-006-9338-6.

Shen, X., W. Guo, X. Zhu, Y. Yuan, J.Z. Yu, R.J. Kohel, et al. 2005. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. Molecular Breeding 15: 169-181. doi:10.1007/s11032-004-4731-0.

Stephens, J.C., F.R. Miller and D.T. Rosenow. 1967. Conversion of alien sorghums to early combine genotypes. Crop Science 7: 396-396.

Team, R.D.C. 2010. R: A language and environment for statistical computing. Vienna, Austria: R foundation for Statistical Computing. doi:http://www.R-project.org.

Uga, Y., Y. Nonoue, Z.W. Liang, H.X. Lin, S. Yamamoto, U. Yamanouchi, et al. 2007. Accumulation of additive effects generates a strong photoperiod sensitivity in the extremely late-heading rice cultivar 'Nona Bokra'. TAG. Theoretical and applied genetics. Theoretische und angewandte Genetik 114: 1457-1466.

Ulloa, M., W.R. Meredith Jr, Z.W. Shappley and A.L. Kahler. 2002. RFLP genetic linkage maps from four F2.3 populations and a joinmap of *Gossypium hirsutum* L. Theoretical & Applied Genetics 104: 200-208.

Ulloa, M. and W.R. Meredith, Jr. 2000. Genetic linkage map and QTL analysis of agronomic and fiber quality traits in an intraspecific population. Journal of Cotton Science 4: 161-170.

Wang, B., Y. Wu, N. Huang, X. Zhu, W. Guo and T. Zhang. 2006. QTL mapping for plant architecture traits in upland cotton using RILs and SSR markers. Acta Genetica Sinica 33: 161-170.

Wang, S., C.J. Basten and Z.-B. Zeng. 2012. Windows QTL Cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC. doi:http://statgen.ncsu.edu/qtlcart/WQTLCart.htm.

Wang, Z., D. Zhang, X. Wang, X. Tan, H. Guo and A.H. Paterson. 2013. A whole-genome DNA marker map for cotton based on the D-genome sequence of *Gossypium raimondii* L. G3 (Bethesda, Md.) 3: 1759-1767. doi:10.1534/g3.113.006890.

Wright, R.J., P.M. Thaxton, K.M. El-Zik and A.H. Paterson. 1999. Molecular mapping of genes affecting pubescence of cotton. Journal of Heredity 90: 215-219.

Wu, J., J.N. Jenkins, J.C. McCarty, M. Zhong and M. Swindle. 2007. AFLP marker associations with agronomic and fiber traits in cotton. Euphytica 153: 153-163. doi:10.1007/s10681-006-9250-0.

Zhang, J., W. Guo and T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum* L. x *Gossypium barbadense* L.) with a haploid population. TAG. Theoretical And Applied Genetics. Theoretische Und Angewandte Genetik 105: 1166-1174.

Table 2.1: Origin and characteristics of converted exotic *G. hirsutum* genotypes.

Description							
Genotype	**Race**	**Country**	**State**	**Site**	**Trait**	**Line rank (of 79)**	**Progeny rank (of 79)**
MDN063	Latifolium	Mexico	Chiapas	Flores Magon	Micronaire	1	1
MDN101	Latifolium	Guatemala	Jutiapa	Jutiapa	Yield Fiber Strength Lint %	NA	1
MDN257	Morrillii	Mexico	Oaxaca	Mitla	2.5% span length	1	1

Table 2.2: Additive genetic variance of exotic parents in crosses with Delta-Pine 16.

Additive genetic variance in crosses with Delta-Pine 16								
Genotype	**Yield**	**Lint %**	**Boll size**	**Micronaire**	**2.5% Span length**	**Elongation**	**Strength**	**Height**
MDN063	-1.9	-0.57**	0.34**	-0.30**	0.55**	-0.14**	-1.89	-0.40
MDN101	84.7**	0.42**	-0.06	0.06	0.38**	-0.09	1.94**	-9.08**
MDN257	15.5	-0.76**	-0.23**	-0.08*	0.79**	-0.29**	1.81	11.09*

*, **Significant from zero at 0.05 and 0.01 respectively. Source:McCarty, Jenkins, et al., 1996.

Table 2.3: Correlation among six fiber quality traits (pooled).

Trait		UHM	UI	STR	ELO	SFC
MIC	F2	-0.32*	0.00	-0.06	0.18**	-0.17
	F3	-0.45**	0.05	-0.10	0.11	-0.43**
UHM	F2		0.59***	0.67***	0.26**	-0.52**
	F3		0.62***	0.57***	-0.32**	-0.14
UI	F2			0.69***	0.41**	-0.82***
	F3			0.55***	-0.08	-0.70***
STR	F2				0.38**	-0.69***
	F3				-0.32**	-0.44**
ELO	F2					-0.50**
	F3					-0.16

***Significant at P=0.001 ** Significant at P= 0.01 * Significant at P= 0.05

Table 2.4: Heritability of different fiber quality traits based on regression of $F_{2:3}$ progenies on F_2 progenies.

Population	MIC	UHM	UI	STR	ELO	SFC
MDN101 × PD94042	0.46**	0.38*	0.18	0.29*	0.44**	0.21
MDN101 × DES56	0.23	0.61**	0.26*	0.47**	0.45**	0.44**
MDN101 × PHMS200	0.43**	0.57**	0.12	0.48**	0.50**	0.06
MDN101 × Acala Maxxa	0.17	0.51**	0.21	0.41**	0.57**	0.31*
MDN063 × DES56	0.41**	0.50**	0.12	0.38*	0.47**	0.46**
MDN063 × Acala Maxxa	0.23	0.36*	-0.06	0.25*	0.62**	0.04*
MDN257 × DES56	0.42**	0.80***	0.20	0.35*	0.61**	0.20
MDN257 × Acala Maxxa	0.62**	0.62**	0.35*	0.34*	0.55**	0.34*

***Significant at P=0.001 ** Significant at P= 0.01 * Significant at P= 0.05

Table 2.5: Markers significantly associated with different fiber quality traits and their contribution to phenotypic variation.

Population	Year (F_2)	Trait	Associated Marker	P Value*	R^2	Additive Effect [a]
MDN101 × DES56	2011	ELO	JESPR101_280	0.0006	0.43	-0.45
	2011	ELO	NAU5120_170	0.0003	0.47	-0.48
	2011	ELO	NAU5465_220	0.0020	0.40	0.64
	2011	SFC	NAU5120_170	0.0020	0.37	0.53
	2011	UHM	DPL0270_150	0.0040	0.33	-0.03
	2011	UI	JESPR101_280	0.0030	0.34	-0.68
	2011	UI	NAU5120_170	0.0003	0.46	-0.88
	2012	STR	NAU3820_120	0.0017	0.10	0.34
MDN101 × PMHS200	2011	MIC	DPL0156_230	0.0020	0.16	-0.17
	2011	SFC	DPL0378_180	0.0020	0.19	0.58
MDN101 × Acala Maxxa	2011	MIC	CIR141_250	0.0003	0.12	-0.19
	2012	UI	NAU5120_450	0.0014	0.13	-0.69
MDN063 × DES56	2011	UI	NAU5120_450	0.0004	0.14	-0.19
MDN257 × DES56	2012	UHM	DPL0378_510	0.0003	0.18	0.48
	2012	UHM	JESPR37_1000	0.0002	0.19	0.03

*Significant after standard Bonferroni correction for multiple comparisons.

[a] Positive values indicate favorable alleles coming from exotic parents (MDN101, MDN063 or MDN257) and negative values indicate favorable alleles coming from elite parents(PD94042, DES56, PMHS200 or Acala Maxxa) in respective crosses. For SFC, negative values indicate favorable alleles coming from the exotic parents and positive values indicate favorable alleles coming from the elite parents.

Table 2.6: Significance of marker-trait associations by year, genetic background, planting dates and fiber quality traits.

Associated Marker	Trait	Population	Year (F_2)	Planting Date	P Value	R^2
By year						
DPL0279_190	MIC	MDN101 × Acala Maxxa	2011	1	0.03	0.05
	MIC	MDN101 × Acala Maxxa	2012	1	0.008	0.09
NAU1366_450	ELO	MDN101 × DES56	2011	1	0.009	0.37
	ELO	MDN101 × DES56	2012	1	0.02	0.05
By planting dates						
NAU5120_450	UI	MDN063 × DES56	2011	1	0.006	0.13
	UI	MDN063 × DES56	2011	2	0.0004*	0.14
NAU5120_450	SFC	MDN063 × DES56	2011	1	0.04	0.07
	SFC	MDN063 × DES56	2011	2	0.007	0.08
By cultivated genetic background						
BNL1317_230	ELO	MDN101 × PD94042	2012	1	0.03	0.04
	ELO	MDN101 × PMHS200	2011	1	0.007	0.13
NAU1042_270	UHM	MDN101 × PD94042	2011	2	0.03	0.25
	UHM	MDN101 × PMHS200	2011	1	0.02	0.09
NAU3820_120	STR	MDN101 × DES56	2012	1	0.0017*	0.10
	STR	MDN101 × PMHS200	2011	1	0.03	0.09
By fiber quality trait						
BNL1317_220	SFC	MDN101 × PMHS200	2012	1	0.03	0.13
	STR	MDN101 × PMHS200	2012	1	0.02	0.14
	UHM	MDN101 × PMHS200	2012	1	0.04	0.11
BNL3359_390	SFC	MDN257 × DES56	2012	1	0.02	0.09
	STR	MDN257 × DES56	2012	1	0.005	0.11
	UHM	MDN257 × DES56	2012	1	0.004	0.12
	UI	MDN257 × DES56	2012	1	0.004	0.12
DPL0270_150	ELO	MDN101 × DES56	2011	2	0.02	0.24
	MIC	MDN101 × DES56	2011	2	0.02	0.27
	UHM	MDN101 × DES56	2011	2	0.004*	0.33
	UI	MDN101 × DES56	2011	2	0.01	0.28

*Significant after standard Bonferroni correction for multiple comparisons.

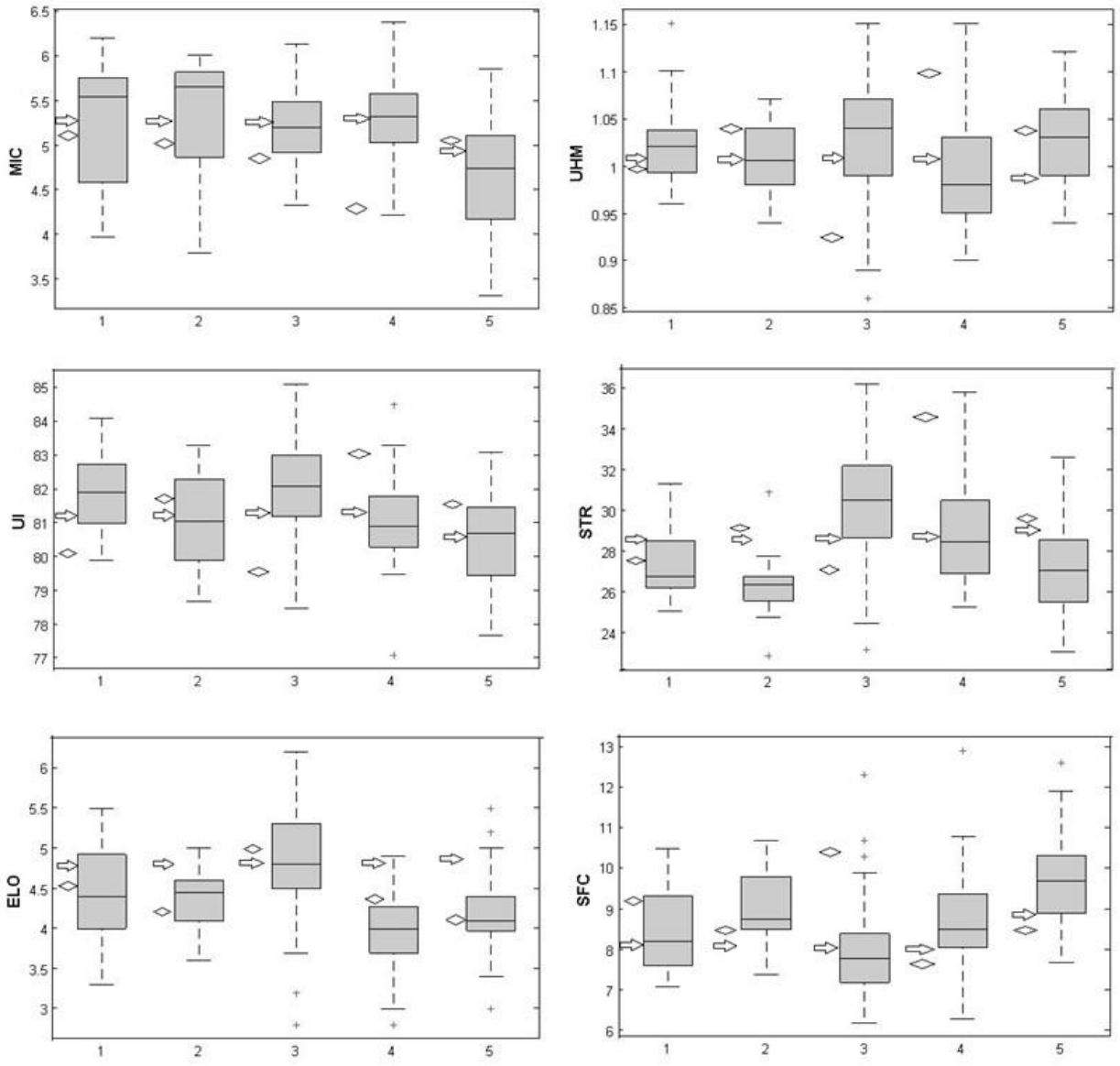

Figure 2.1: Distribution of populations (F_2 generation) created in 2010 for six different fiber quality traits. Y axis shows the range for each trait and x-axis shows different populations: (1) MDN101 × PD94042, (2) MDN101 × DES56, (3) MDN101 × PMHS200, (4) MDN101 × Acala Maxxa, (5) MDN063 × DES56. Average phenotypic values are shown by arrows for exotic parents (MDN101 or MDN063) and wedges for elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

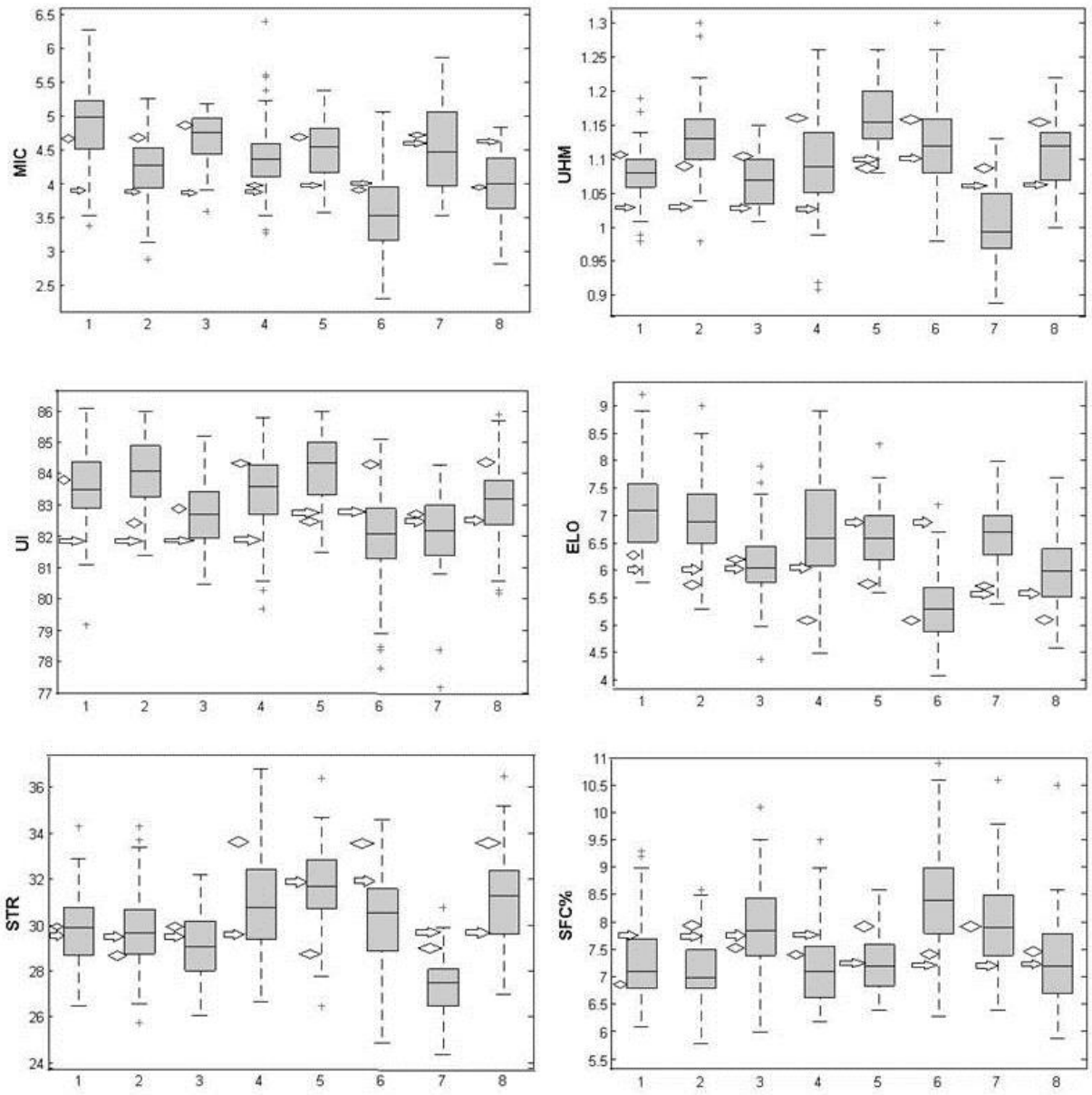

Figure 2.2: Distribution of populations (F_2 generation) created in 2011 for six different fiber quality traits. Y axis shows the range for each trait and x-axis shows different populations: (1) MDN101 × PD94042, (2) MDN101 × DES56, (3) MDN101 × PMHS200, (4) MDN101 × Acala Maxxa, (5) MDN063 × DES56, (6) MDN063 × Acala Maxxa, (7) MDN257 × DES56, and (8) MDN257 × Acala Maxxa. Average phenotypic values are shown by arrows for exotic parents

(MDN101, MDN063 or MDN257) and wedges for elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

CHAPTER 3

ASSOCIATION OF MICROSATELLITE MARKERS WITH YIELD RELATED AND PLANT ARCHITECTURAL TRAITS IN UPLAND COTTON

Adhikari J.[1], Z. Wang, S. Khanal, R. Chandnani, J.D. Patel, L. Goff, D. Jones and A.H. Paterson.

To be submitted to *Euphytica*.

Abstract

Cotton is the world's leading fiber crop and genetic improvement of fiber yield is the primary aim of cotton breeding programs. Primitive cotton genotypes offer diversity for genetic improvement of cultivated cotton species as they might introduce useful and favorable alleles into the cultivated gene pool and help in crop improvement. In this study, we used three exotic *Gossypium hirsutum* accessions converted to day-neutral flowering as parents with four elite cotton cultivars to create experimental populations. We selected 85 polymorphic microsatellite markers selected from 18 "hotspot" regions for fiber quality quantitative trait loci (QTLs) and studied the association of these markers with different yield related and morphological traits. We identified 126 nominal marker-trait associations for different yield related and plant morphological traits out of which 8 remained significant after Bonferroni correction for multiple comparisons. Four of these significant associations were for boll maturity, two for flowering, one for plant height and one for number of nodes with at least one branch. Utilization of these markers may be useful in improvement of plant architectural and morphological traits. The exotic parents contributed favorable alleles for plant height, flowering and boll maturity indicating their potential for introducing favorable alleles related to plant architecture into elite cottons.

Introduction

Cotton (*Gossypium sp*) is the world's most important natural textile fiber. The genus *Gossypium* consists of more than 50 species (Fryxell and Craven 1992) of which most are diploid (2n=26) and five are polyploid (2n=52). Only four species of cotton have been domesticated and are cultivated commercially for lint production. Approximately 90% of world's cotton comes from allotetraploid *G. hirsutum* (Upland cotton) and about 8% comes from another

allotetraploid, *Gossypium barbadense*, commonly referred to as Sea Island cotton, Pima cotton, or Egyptian cotton. Two diploid species *Gossypium arboreum* and *Gossypium herbaceum* contribute less than 2% towards total cotton production (Kumar 2012). Because of higher yields and long spinnable fibers, Upland cotton has always been the center of cotton research and breeding programs.

The worldwide economic impact of cotton is estimated at around $500 billion/year with an annual utilization of around 115 million bales of cotton fiber (Chen, Scheffler et al. 2007). Today, cotton is commercially cultivated in more than 50 countries in drier tropical and sub-tropical areas of the world (Smith and Cothren 1999). The major share of global cotton production comes from countries such as China, India, United States, Pakistan and Australia where climatic conditions such as periods of hot and dry weather, photoperiod and adequate soil moisture favor natural growth requirements of cotton (Khadi, Santhy et al. 2010). The United States is the third largest producer of cotton in the world, typically accounting for 16% of the total world production and is usually the leading exporter, accounting over one-third of the global trade in raw cotton.

In recent years, demand for cotton fiber in the world market has dramatically increased. But the acreage for cotton production has been declining due to competition by food and staple crops (Mei, Zhu et al. 2013). Thus improving yield and yield related traits has been the major challenge faced by cotton industry in the recent years. Improvement in fiber yield and quality has been in stalemate in the United States mainly due to low genetic diversity of cultivated cotton species. Since horizontal expansion of yield is not possible due to limited acreage of cultivation, there is a need to increase yield per unit area to support the increasing demands of cotton fibers. Improvement of current cultivars by breeding for higher yield is the only way to fulfill the

increasing demands of world cotton industry. Inclusion of exotic lines in breeding programs is essential to increase the genetic diversity of cultivated cotton pool as it introduces favorable alleles to the genetically impoverished pool (Paterson, Boman et al. 2004, Hajjar and Hodgkin 2007).

Plant architecture is defined as the three-dimensional organization of a plant, which is dependent on the function and relative arrangement of each of its parts. Plant architecture is of major agronomic importance because it strongly influences the suitability of a plant for cultivation, its overall yield and its economic coefficient (Reinhardt and Kuhlemeier 2002). Plant architecture can significantly affect light penetration and distribution in the crop canopy and affect plant growth, biomass portioning, boll distribution, boll weight and yield potential of the crop (Kaggwa-Asiimwe, Andrade-Sanchez et al. 2013). Apart from yield related traits, plant architectural traits like height of the plant, number and orientation of branches, earliness of flowering and maturity also have direct effects on the total cotton yield. Thus simultaneous improvement of all these traits needs to be taken into consideration while breeding for higher yielding cultivars.

Identification of associations of molecular markers with quantitative traits is valuable for breeding and improvement of different traits. Many studies relating to yield related and plant architectural traits have been conducted in cotton and lots of QTL or markers associated with these traits have been identified (Jiang, Wright et al. 2000, Naveed, Abdul et al. 2006, Percy, Cantrell et al. 2006, Abdurakhmonov, Buriev et al. 2007, Li, Guo et al. 2008, Guo, McCarty et al. 2009, Qin, Liu et al. 2009, Li, Wang et al. 2012, Liu, Wang et al. 2012, Mei, Zhu et al. 2013, Said, Lin et al. 2013, Liu, Ai et al. 2014). In this study, we have used genetically diverse parents including elite cotton cultivars and exotic accessions converted to day-neutral flowering,

studying F_2 and $F_{2:3}$ generations of different populations, to identify associations of selected polymorphic SSR markers with different yield related and plant architectural traits.

Materials and Methods

Plant materials

We selected three exotic *G. hirsutum* genotypes converted to day-neutral flowering, MDN063, MDN101 and MDN257, based on different agronomic characteristics. Four elite US genotypes, PD94042, DES56, PMHS200 and Acala Maxxa, representing the US cotton genetic pool were selected as the recipient parents in the study.

Crosses were made in 2010-11 in the green-house between the day-neutral exotic lines and elite genotypes in different combinations (Table S1). Bolls were hand harvested and collected separately for individual crosses. All samples were hand-ginned and delinted.

The resulting F1 and the parents were planted at The University of Georgia Plant Sciences Farm, Watkinsville, Georgia. Standard cultivation practices were applied including irrigation, fertilization and pesticide application. As many flowers as possible from each individual F_1 plant were self-pollinated, and DNA of F_1 plants was checked with several SSR markers to verify hybridity. Selfed bolls from verified F_1 plants were hand collected separately for F_2 seed. Selfed bolls from the F_2 plants were also collected individually to obtain $F_{2:3}$ seeds. $F_{2:3}$ progeny plants for 2010 populations were grown in 2012 and those for 2011 populations were grown in 2013 in plots of 10 seeds.

Sample collection and data analysis

Sample were collected from F_2 plants and corresponding $F_{2:3}$ progeny plots. Total seed yield was taken from 25 bolls to sample each plant in the plot and was partitioned into fiber yield and seed yield. Utilizing this partitioning, lint percentage for each sample was calculated. Data

on plant height, branching, flowering and boll maturity were taken shortly before harvesting. Plant heights for single plants were taken for F_2 plants while the average height of all the plants from a plot was taken for $F_{2:3}$ progeny. Data for flowering, boll maturity and branching used measuring system shown in Table 3.1. Since data on flowering were taken shortly before harvest, lateness in flowering was analyzed using the scored data. Data were analyzed using R statistical software. Single marker analyses (SMA) for plant height, total seed cotton yield and lint percentage were done using Win QTL cartographer 2.5 (Wang, Basten et al. 2012). Poisson regression was used to identify relationships between the markers and categorical traits like flowering, maturity and number of nodes with at least one branch. Bonferroni correction was used to correct for multiple comparisons.

DNA extraction and SSR assays

Genomic DNA was isolated from young unopened leaves from each F_2 plant from all the populations and from the parental samples using a modified CTAB method (Paterson, Brubaker et al. 1993). A total of 1-2 g of fresh tissue was ruptured in a 1:1 mixture of cotton lysis buffer and extraction buffer. After leaving the tubes in a 65˚C water bath for about an hour, the extraction materials were purified twice with 800 µl of chloroform iso-amyl alcohol. DNA was then precipitated with 500 µl of isopropanol, cleaned with two washes of 75% ethanol (500µl each) and centrifuged. The clean dried DNA was dissolved in 200 µl of TE buffer. The extracted DNA was checked for quality and quantity and stored at -20˚C.

The SSR technique was used to identify polymorphic markers. 720 SSR primer pairs were selected from 18 hotspots for fiber quality QTLs from the cotton genome and were used to screen for polymorphism among the parents of the mapping populations. SSR amplifications used a Touchdown PCR reaction program: 1. Seven cycles of four steps (94°C for 4 minutes,

94°C for 40 seconds, 58°C for 1 min and 72°C for 1 min); 2. Thirty six cycles of three steps (94°C for 45 seconds, 54°C for 45 seconds and 72°C for 1 minute); 3. One cycle of 72°C for 10 minutes, and 4. Hold at 4°C until the PCR plates are removed. PCR products were separated using acrylamide gel electrophoresis and visualized using silver staining (Bassam and Gresshoff 2007).

In a separate study, 85 polymorphic SSR markers were selected from 18 different regions of the cotton genome known as "hotspots" for fiber quality QTLs. These sets of markers were used to identify associations with fiber quality traits. Same set of markers were deployed in this study to identify their associations with yield related and plant architectural traits in cotton.

Results

Performance of parents, F_2 and $F_{2:3}$ progenies

The average phenotypic values of the parents for various yield, yield related and plant morphological traits are listed in Table S8. Parents showed diverse phenotypic values for the traits scored. None of the exotic parents had higher average values for lint percentage than the elite parents. MDN101 showed shorter average heights than the other exotic parents and three of the four elite cultivars. Elite lines showed lower average values for flowering indicating lateness in flowering than the exotic lines. However, the earliness in flowering shown by the exotic lines could be due to alleles coming from DPL16 since these exotic parents were repeatedly backcrossed with DPL16 to convert them to day-neutral flowering. On the other hand, elite genotypes had earlier boll maturity than the exotic lines. The distribution of the populations for plant height and lint percentage is shown in Figure 3.1.

Markers associated with yield, yield related and morphological traits

A total of 14 associations were identified for plant height (Table 11). Only the association of the marker locus DPL0270_150 with plant height met Bonferroni correction for multiple comparisons (Table 3.4). It explained 9 percent of the total phenotypic variation. Among the nominal associations that did not survive the Bonferroni correction, one gained support from multiple independent discoveries. DPL0279_200 was significant in MDN101 × DES56 as well as MDN101 × Acala Maxxa in 2012. The phenotypic variation explained by the nominal associations ranged from 2 to 12 percent.

A total of twenty four associations were identified for flowering (Table S12), of which two were still significant after Bonferroni correction for multiple comparisons (Table 3.4). One of these two associations was between the trait and marker locus MONDPL0696_180 in MDN101 × PD94042 population explaining 9 percent of the phenotypic variation, while the other was between the trait and marker locus NAU3074_180 in MDN063 × DES56 population explaining 21 percent of the phenotypic variation. Among nominal associations that did not survive the Bonferroni correction, some gained support from multiple independent discoveries. NAU1221_510, NAU5120_550 and DPL0378_590 were each associated with the trait in two different populations in the year 2012. The phenotypic variation explained by the nominal associations ranged from 1 to 19 percent.

A total of thirty two associations were identified for boll maturity (Table S13), with four (BNL3989_170, MUCS616_160 and NAU3820_120 in the population MDN101 × DES56; and NAU5120_550 in the populations MDN101 × Acala Maxxa) remaining significant after Bonferroni correction for multiple comparisons (Table 3.4). The phenotypic variations explained by these four associations were 16, 12, 10 and 15 percent respectively. Among nominal associations that did not survive the Bonferroni correction, two gained support from multiple

independent discoveries. NAU4042_180 was significant in two different populations and also in both the years. Similarly, NAU5120_510 was significant in three different populations and in both years. The phenotypic variation explained by nominal associations ranged from 1 to 39 percent.

A total of seventeen associations were identified for nodes with one or more branches (Table S14) with one (BNL3977_130 in MDN063 × DES56, explaining 6 percent of phenotypic variation) remaining significant after Bonferroni correction (Table 3.4). The phenotypic variation explained by the nominal associations ranged from 1 to 15 percent.

For several additional traits, we found nominal associations but which did not meet the threshold of Bonferroni correction for multiple comparisons. A total of nineteen associations were identified for total seed cotton yield (Table S9) with phenotypic variation explained by these associations ranging from 4 to 22 percent. Twenty associations were identified for lint percentage (Table S10) with phenotypic variation explained by these associations ranging from 2 to 21 percent. Finally, we found no associations with number of nodes with two or more branches.

Contribution of the exotic parents

All three exotic parents contributed to shorter plant height (Table S11). Of the fourteen associations identified for plant height, only the association of DPL0270_150 with the trait in MDN063 × Acala Maxxa survived the standard Bonferroni correction. The exotic parent MDN063 contributed the favorable allele (shorter plant height) for this association. For the remaining 13 nominal associations, the three exotic parents contributed favorable alleles in nine cases. MDN101, which showed positive additive effects for plant height in crosses with DPL16

(Table 2.2), contributed favorable alleles in six cases in three different backgrounds, DES56, PMHS200 and Acala Maxxa. MDN257 contributed favorable alleles in three cases.

Exotic parents also contributed nominal favorable alleles for lint percentage (Table S10), none of which met the Bonferroni threshold. MDN101, which showed positive additive effects in crosses with DPL16 (Table 2.2), also showed positive additive effects in crosses with PD94042 and Acala Maxxa. Similarly, MDN063 contributed favorable alleles to DES56 background and MDN257 to Acala Maxxa background.

Positive effects of the exotic parents were also observed for total seed cotton yield, but none of which met the Bonferroni threshold. The exotic parents contributed favorable alleles in 11 of 19 nominal associations identified for total seed cotton yield. MDN101 contributed favorable alleles to three genetic backgrounds (PD94042, PMHS200 and Acala Maxxa), MDN257 to two backgrounds (DES56 and Acala Maxxa), and MDN063 to only one background (DES56).

The exotic parents also showed positive additive effects for flowering (Table S12) and boll maturity (Table S13). Among the two (of 24) associations identified for flowering which met the Bonferroni threshold, the exotic parents (MDN101 and MDN063) contributed favorable alleles in both cases. However, since the exotic parents were converted to day-neutral flowering by repeated backcrossing with DPL16, there might be a possibility for these alleles to be associated with DPL16. Of the remaining 22 nominal associations, the exotic parents contributed favorable alleles in nine cases. The exotic parents also contributed favorable alleles in two of the four significant associations identified for boll maturity, both from in crosses with DES56 and Acala Maxxa. Among the 28 nominal associations, the exotic parents contributed favorable alleles in 15 cases. MDN257 did not contribute any favorable alleles to boll maturity. One

significant association was identified for nodes with at least one branch (Nodes 1) in which the exotic parent did not contribute favorable alleles to the elite background (Table S14). Of the sixteen nominal associations for this trait, the exotic parents contributed favorable alleles in eleven cases.

Discussion

Inclusion of wild and exotic genotypes in breeding programs is one way of increasing the low genetic diversity in cultivated cotton and thus increasing the scope for improvement in different fiber and plant morphology related traits. In our study we used three exotic *G. hirsutum* genotypes converted to day-neutral flowering and four elite *G. hirsutum* cultivars representing different US cotton production regions as parents. Exotic parents not only showed diverse phenotypes but also contributed to SSR marker polymorphisms. Apart from the fiber quality traits for which the parents were actually selected for, useful additive effects were also observed for yield related and plant architecture related traits from these exotic parents. The exotic parents contributed favorable alleles meeting the Bonferroni significance threshold for three of the plant architectural and yield related traits tested in our study. Favorable alleles (coming from significant associations) were contributed by the exotic parents for five of 126 associations identified. Among the nominal associations, the exotic parents contributed favorable alleles in 57 of 121 cases. In some instances, alleles with only nominal effects in one population or year gained some support from the discovery of nominal effects in additional populations or years, suggesting that some of these may warrant future testing. These results generally indicate the usefulness of diverse genotypes in breeding programs for the introgression of useful alleles from the wild to the cultivated cotton pool.

Plant height is important to a crop as it determines the crop canopy. In most cotton-growing regions, moderately shorter plants are preferred as they are frequently associated with an improved canopy structure and hence, yield (Liu, Ai et al. 2014). Moreover, shorter plant height than wild cottons is a prerequisite for machine harvest. Understanding of genetics of plant height and identification of DNA markers strongly associated with this trait is useful for breeding for plant height in cotton. We identified 14 marker-trait associations for plant height, out of which one remained significant after Bonferroni correction for multiple comparisons. The exotic parent MDN063 contributed favorable allele to the Acala Maxxa background for shorter plant height.

Similarly, early flowering and boll maturity contribute to the yield and quality of cotton fibers. In many studies, node of first fruiting branch (NFB) has been used as an indication of flowering and the time for the first fruiting branch to emerge has been taken as an indication of earliness in flowering (Guo, McCarty et al. 2008, Li, Li et al. 2010, Li, Wang et al. 2012). In the present study, we have used the proportion of open and closed flowers on the plant near the time of harvesting as an index of lateness in flowering and used this index to identify association of SSR markers with this trait/index. Using this approach we identified 24 marker-trait associations for flowering and 32 associations for boll maturity. In most cases, elite parents contributed alleles related to lateness in flowering. The exotic parents contributed favorable alleles for earlier flowering (opposite of lateness) in both (Bonferroni-corrected) significant associations identified for this trait, and two of the four significant associations identified for boll maturity. These associations could be potentially useful in breeding for earliness of flowering and boll maturity in cotton. No significant associations were identified for total seed cotton yield, lint percentage

and number of nodes with at least one branch, although exotic parents contributed some nominal favorable alleles that may indicate the potential for useful contributions to the elite pool.

The exotic parents used in our study were selected based on the additive genetic variance shown by them for different fiber quality traits (McCarty, Jenkins, et al., 1996). Moreover, SSR markers used to screen the progeny populations were also selected from fiber quality QTL hotspots. Here, we found (Bonferroni-corrected) significant associations for traits other than those for which the parents were selected, indicating the usefulness of diverse exotic genotypes, in improvement of not only limited sets of traits but a broader range of traits. These parents could also be potential source of favorable alleles for biotic and abiotic stresses and other traits of interest in cotton. Utilization of these diverse resources in cotton breeding programs promises to improve the performance of elite cultivars and to broaden the narrow genetic base of the world's cultivated cotton gene pool.

Conclusion

In this study, three exotic *G. hirsutum* accessions converted to day-neutral flowering were used as parents in crosses with four elite US cultivars to develop experimental populations in two different years in different combinations. We studied F_2 and $F_{2:3}$ generations of these populations for different yield related and plant architecture related traits; and used 85 different SSR markers selected from 18 QTL "hotspot" regions in the cotton genome to identify significant associations with these traits. Eight significant associations (meeting Bonferroni thresholds) were identified in different populations, among which the exotic parents contributed favorable alleles in five cases. These associations show that utilization of the exotic parents in cotton breeding may help to improve on the performance of currently elite cultivars and increase the genetic diversity of the cultivated cotton gene pool.

References

Abdurahmonov, I. (2007). Exploiting genetic diversity. World Cotton Research Conference-4, Lubbock, Texas, USA, 10-14 September 2007. Washington; USA, International Cotton Advisory Committee (ICAC)**:** un-un.

Abdurakhmonov, I. Y., A. A. Abdullaev, S. Saha, Z. T. Buriev, D. Arslanov, Z. Kuryazov, G. T. Mavlonov, S. M. Rizaeva, U. K. Reddy, J. N. Jenkins, A. Abdullaev and A. Abdukarimov (2005). "Simple sequence repeat marker associated with a natural leaf defoliation trait in tetraploid cotton." Journal of Heredity **96**(6): 644-653.

Abdurakhmonov, I. Y., Z. T. Buriev, S. Saha, A. E. Pepper, J. A. Musaev, A. Almatov, S. E. Shermatov, F. N. Kushanov, G. T. Mavlonov, U. K. Reddy, J. Z. Yu, J. N. Jenkins, R. J. Kohel and A. Abdukarimov (2007). "Microsatellite markers associated with lint percentage trait in cotton, Gossypium hirsutum." Euphytica **156**(1/2): 141-156.

Aleman, L., J. Kitamura, H. Abdel-mageed, J. Lee, Y. Sun, M. Nakajima, M. Ueguchi-Tanaka, M. Matsuoka and R. D. Allen (2008). "Functional analysis of cotton orthologs of GA signal transduction factors GID1 and SLR1." Plant Molecular Biology **68**(1-2): 1-16.

Bassam, B. J. and P. M. Gresshoff (2007). "Silver staining DNA in polyacrylamide gels." Nat Protoc **2**(11): 2649-2654.

Bowman, D. T. (2000). "Attributes of public and private cotton breeding programs [electronic resource]." Journal of cotton science (Online) Journal of cotton science **4**(2): 130-136.

Bradow, J. M. and G. H. Davidonis (2000). "Quantitation of fiber quality and the cotton production-processing interface: a physiologist's perspective." Journal of Cotton Science **4**(1): 34-64.

Chee, P., E. Lubbers, L. May, J. Gannaway and A. Paterson (2004). "CHANGES IN GENETIC DIVERSITY OF U.S. UPLAND COTTON." Proceedings of the Beltwide Cotton Production Conference: 1054-1054.

Chee, P. W., X. Draye, C. X. Jiang, L. Decanini, T. A. Delmonte, R. Bredhauer, C. W. Smith and A. H. Paterson (2005). "Molecular dissection of phenotypic variation between Gossypium hirsutum and Gossypium barbadense (cotton) by a backcross-self approach: III. Fiber length." TAG Theoretical and Applied Genetics **111**(4): 772-781.

Chen, Z. J., B. E. Scheffler, E. Dennis, B. A. Triplett, T. Zhang, W. Guo, X. Chen, D. M. Stelly, P. D. Rabinowicz, C. D. Town, T. Arioli, C. Brubaker, R. G. Cantrell, J.-M. Lacape, M. Ulloa, P. Chee, A. R. Gingle, C. H. Haigler, R. Percy, S. Saha, T. Wilkins, R. J. Wright, A. Van Deynze, Y. Zhu, S. Yu, I. Abdurakhmonov, I. Katageri, P. A. Kumar, Y. Zafar, J. Z. Yu, R. J. Kohel, J. F. Wendel and A. H. Paterson (2007). "Toward Sequencing Cotton (Gossypium) Genomes." Plant Physiology(4): 1303.

Dantas, A. C. A., P. A. V. Barroso, L. V. Hoffmann, M. F. Alves and F. P. d. Andrade (2012). "SSR markers to detect gene flow from upland to mocó cotton." Revista Ciência Agronômica **43**(1): 163-169.

Draye, X., P. Chee, C. X. Jiang, L. Decanini, T. A. Delmonte, R. Bredhauer, C. W. Smith and A. H. Paterson (2005). "Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense (cotton) by a backcross-self approach: II. Fiber fineness." TAG Theoretical and Applied Genetics **111**(4): 764-771.

Ellis, M. H., G. J. Rebetzke, F. Azanza, R. A. Richards and W. Spielmeyer (2005). "Molecular mapping of gibberellin-responsive dwarfing genes in bread wheat." Theoretical & Applied Genetics **111**(3): 423-430.

Esbroeck, G. A. v., D. T. Bowman, O. L. May and D. S. Calhoun (1999). "Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars." Crop Science **39**(2): 323-328.

Esbroeck, G. v. and D. T. Bowman (1998). "Cotton germplasm diversity and its importance to cultivar development." Journal of Cotton Science **2**(3): 121-129.

Fryxell, P. A. and L. A. Craven (1992). "A Revision of Gossypium Sect. Grandicalyx (Malvaceae), Including the Description of Six New Species." Systematic Botany(1): 91.

Guo, Y., J. C. McCarty, J. N. Jenkins, C. An and S. Saha (2009). "Genetic detection of node of first fruiting branch in crosses of a cultivar with two exotic accessions of upland cotton." Euphytica **166**(3): 317-329.

Guo, Y., J. C. McCarty, J. N. Jenkins and S. Saha (2008). "QTLs for node of first fruiting branch in a cross of an upland cotton, Gossypium hirsutum L., cultivar with primitive accession Texas 701." Euphytica **163**(1): 113-122.

Gutierrez, O. A., D. B. Shoemaker, C. L. Cheatham, J. C. McCarty, Jr., S. Basu, S. Saha and J. N. Jenkins (2002). "Genetic distance among selected cotton genotypes and its relationship with F2 performance." Crop science **42**(6): 1841-1847.

Hajjar, R. and T. Hodgkin (2007). "The use of wild relatives in crop improvement: a survey of developments over the last 20 years." Euphytica **156**(1/2): 1-13.

Holley, R. N. and M. M. Goodman (1988). "Yield potential of tropical hybrid maize derivatives." Crop Science **28**(2): 213-218.

Iqbal, M. J., O. U. K. Reddy, K. M. El-Zik and A. E. Pepper (2001). "A genetic bottleneck in the 'evolution under domestication' of upland cotton Gossypium hirsutum L. examined using DNA fingerprinting." Theoretical and Applied Genetics **103**(4): 547-554.

Jiang, C., R. J. Wright, K. M. El-Zik and A. H. Paterson (1998). "Polyploid formation created unique avenues for response to selection in Gossypium (cotton)." Proceedings of the National Academy of Sciences of the United States of America **95**(8): 4419-4424.

Jiang, C., R. J. Wright, S. S. Woo, T. A. DelMonte and A. H. Paterson (2000). "QTL analysis of leaf morphology in tetraploid Gossypium (cotton)." Theoretical & Applied Genetics **100**(3/4): 409-418.

Kaggwa-Asiimwe, R., P. Andrade-Sanchez and G. Wang (2013). "Plant architecture influences growth and yield response of upland cotton to population density." Field Crops Research **145**(0): 52-59.

Khadi, B. M., V. Santhy and M. S. Yadav (2010). "Cotton: An Introduction." BIOTECHNOLOGY IN AGRICULTURE AND FORESTRY **65**: 1-14.

Kumar, P. (2012). Genetic evaluation of exotic chromatins from two obsolete interspecific introgression lines of Upland cotton [electronic resource] / by Pawan Kumar, 2012.

Lecomte, L., P. Duffé, M. Buret, B. Servin, F. Hospital and M. Causse (2004). "Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds." Theoretical & Applied Genetics **109**(3): 658-668.

Li, C., W. Guo, X. Ma and T. Zhang (2008). "Tagging and mapping of QTL for yield and its components in upland cotton population with varied lint percentage." Cotton Science **20**(3): 163-169.

Li, C., C. Wang, N. Dong, X. Wang, H. Zhao, R. Converse, Z. Xia, R. Wang and Q. Wang (2012). "QTL detection for node of first fruiting branch and its height in upland cotton (Gossypium hirsutum L.)." Euphytica **188**(3): 441-451.

Li, K., Z. Li and Y. Wei (2010). "Correlation of early maturing traits and QTL mapping in cotton." Xinjiang Agricultural Sciences **47**(1): 78-81.

Li, Y. L., X. H. Li, J. Z. Li, J. F. Fu, Y. Z. Wang and M. G. Wei (2009). "Dent corn genetic background influences QTL detection for grain yield and yield components in high-oil maize." Euphytica **169**(2): 273-284.

Liao, C. Y., P. Wu, B. Hu and K. K. Yi (2001). "Effects of genetic background and environment on QTLs and epistasis for rice (Oryza sativa L.) panicle number." Theoretical & Applied Genetics **103**(1): 104-111.

Liao, W., M. Ruan, B. Cui, N. Xu, J. Lu and M. Peng (2009). "Isolation and characterization of a GAI/RGA-like gene from Gossypium hirsutum." Plant Growth Regulation **58**(1): 35-45.

Liu, R., N. Ai, X. Zhu, F. Liu, W. Guo and T. Zhang (2014). "Genetic analysis of plant height using two immortalized populations of 'CRI12 × J8891' in Gossypium hirsutum L." Euphytica **196**(1): 51-61.

Liu, R., N. Ai, X. Zhu, F. Liu, W. Guo and T. Zhang (2014). "Genetic analysis of plant height using two immortalized populations of "CRI12 × J8891" in Gossypium hirsutum L." Euphytica **196**(1): 51-61.

Liu, R., B. Wang, W. Guo, Y. Qin, L. Wang, Y. Zhang and T. Zhang (2012). "Quantitative trait loci mapping for yield and its components by using two immortalized populations of a heterotic hybrid in Gossypium hirsutum L." Molecular Breeding **29**(2): 297-311.

Lubbers, E., P. Chee, J. Gannaway, R. Wright, K. M. El-Zik and A. H. Paterson (2004). Levels and patterns of genetic diversity in upland cotton. Plant and Animal Genome XII Conference. San Diego, CA.

Mauer, F. M. (1954). Origin and systematics of cotton. Origin and systematics of cotton., Izdateljstvo Akad. Nauk Uzbeksk. SSR. Taskent.

McCarty, J., J. Wu and J. Jenkins (2006). "Genetic diversity for agronomic and fiber traits in day-neutral accessions derived from primitive cotton germplasm." Euphytica **148**(3): 283-293.

McCarty, J. C. and J. N. Jenkins (2001). Primitive cotton germplasm : yield and fiber traits for 16 day-neutral accessions / Jack C. McCarty, Jr. and Johnie N. Jenkins. Research report / Mississippi Agricultural & Forestry Experiment Station ; vol. 22, no. 19, [Mississippi State, Miss.] : Mississippi Agricultural & Forestry Experiment Station, [2001].

McCarty, J. C., J. N. Jenkins, B. Tang and C. E. Watson (1996). "Genetic analysis of primitive cotton germplasm accessions." Crop Science **36**(3): 581-585.

McCarty, J. C., J. N. Jenkins and J. X. Wu (2004). "Primitive accession derived germplasm by cultivar crosses as sources for cotton improvement: I. Phenotypic values and variance components." Crop Science **44**(4): 1226-1230.

McCarty, J. C., J. N. Jenkins and J. Zhu (1998). "Introgression of Day-Neutral Genes in Primitive Cotton Accessions: I. Genetic Variances and Correlations." Crop Sci. **38**(6): 1425-1428.

McCarty, J. C., Jr. and J. N. Jenkins (1992). Cotton germplasm. Characteristics of 79 day-neutral primitive race accessions. Technical Bulletin - Mississippi Agricultural and Forestry Experiment Station**:** 17-17.

McCarty, J. C., Jr., J. N. Jenkins and B. Tang (1995). Primitive cotton germplasm: variability for yield and fiber traits. Technical Bulletin - Mississippi Agricultural and Forestry Experiment Station**:** 8-8.

McCarty, J. C., Jr., J. N. Jenkins and J. Zhu (1998). "Introgression of day-neutral genes in primitive cotton accessions: II. Predicted genetic effects." Crop Science **38**(6): 1428-1431.

McCarty, J. C., J. Wu and J. N. Jenkins (2003). Use of primitive accessions of cotton as sources of genes for improving yield components and fiber properties / Jack C. McCarty, Johnie N. Jenkins, and Jixiang Wu. Bulletin / Mississippi Agricultural and Forestry Experiment Station ; 1130, [Mississippi State, Miss.] : Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University, [2003].

McCarty, J. C., J. Wu and J. N. Jenkins (2005). Potential of primitive accessions for cotton improvement / Jack C. McCarty, Jr., Johnie N. Jenkins, and Jixiang Wu. Bulletin / Mississippi Agricultural and Forestry Experiment Station ; 1141, [Mississippi State, Miss.] : Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University, [2005].

Mei, H., X. Zhu and T. Zhang (2013). "Favorable QTL Alleles for Yield and Its Components Identified by Association Mapping in Chinese Upland Cotton Cultivars." PLoS ONE **8**(12): 1-12.

Mei, M., N. H. Syed, W. Gao, P. M. Thaxton, C. W. Smith, D. M. Stelly and Z. J. Chen (2004). "Genetic mapping and QTL analysis of fiber-related traits in cotton (Gossypium)." TAG. Theoretical And Applied Genetics. Theoretische Und Angewandte Genetik **108**(2): 280-291.

Meredith, W. R., Jr. (1991). Contributions of introductions to cotton improvement. Use of plant introductions in cultivar development. Part 1. Proceedings of a symposium sponsored by

Division C-1 of the Crop Science Society of America in Las Vegas, Nevada, USA, 19 October 1989. H. L. Shands and L. E. Wiesner. Madison, Wisconsin; USA, Crop Science Society of America, Inc.**:** 127-146.

Meredith, W. R., Jr. (2000). "Cotton yield progress - why has it reached a plateau?" Better Crops with Plant Food **84**(4): 6-9.

Muhammad, S., G. Wangzhen, U. Ihsan, T. Nabila, Z. Yusuf, R. Mehboob ur and Z. Tianzhen (2011). "QTL mapping for physiology, yield and plant architecture traits in cotton (Gossypium hirsutum L.) grown under well-watered versus water-stress conditions." Electronic Journal of Biotechnology **14**(3): 1-13.

Murfet, I. C. (1977). "Environmental Interaction and The Genetics of Flowering." Annual Review of Plant Physiology **28**(1): 253-278.

Naveed, M., Q. Abdul, M. Waqas and N. Etrat (2006). "Genetic study of yield of seed cotton and plant height in cotton genotypes." International Journal of Agriculture and Biology **8**(5): 630-635.

Paterson, A. H., R. K. Boman, S. M. Brown, P. W. Chee, J. R. Gannaway, A. R. Gingle, O. L. May and C. W. Smith (2004). "Reducing the genetic vulnerability of cotton." Crop Science **44**(6): 1900-1901.

Paterson, A. H., C. L. Brubaker and J. F. Wendel (1993). "A rapid method for extraction of cotton (Gossypium spp.) genomic DNA suitable for RFLP or PCR analysis." Plant Molecular Biology Reporter **11**(2): 122-127.

Peng, J., D. E. Richards, N. M. Hartley, G. P. Murphy, K. M. Devos, J. E. Flintham, J. Beales, L. J. Fish, A. J. Worland, F. Pelica, D. Sudhakar, P. Christou, J. W. Snape, M. D. Gale and

N. P. Harberd (1999). "'Green revolution' genes encode mutant gibberellin response modulators." Nature **400**(6741): 256-261.

Percival, A. E. (1987). The National Collection of Gossypium Germplasm, Department of Agricultural Communications, Texas A & M University.

Percy, R. G., R. G. Cantrell and J. F. Zhang (2006). "Genetic variation for agronomic and fiber properties in an introgressed recombinant inbred population of cotton." Crop Science **46**(3): 1311-1317.

Pettigrew, W. T. (1994). "Source-to-sink manipulation effects on cotton lint yield and yield components." Agronomy journal **86**(4): 731-735.

Powell, W., M. Morgante, C. Andre, M. Hanafey, J. Vogel, S. Tingey and A. Rafalski (1996). "The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis." Molecular Breeding **2**(3): 225-238.

Qin, Y., R. Liu, H. Mei, T. Zhang and W. Guo (2009). "QTL mapping for yield traits in Upland cotton (Gossypium hirsutum L.)." Acta Agronomica Sinica **35**(10): 1812-1821.

Reinhardt, D. and C. Kuhlemeier (2002). "Plant architecture." EMBO Reports **3**(9): 846-851.

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

Reta-Sánchez, D. G. and J. L. Fowler (2002). "Canopy light environment and yield of narrow-row cotton as affected by canopy architecture." Agronomy Journal **94**(6): 1317-1323.

Rong, J., F. A. Feltus, V. N. Waghmare, G. J. Pierce, P. W. Chee, X. Draye, Y. Saranga, R. J. Wright, T. A. Wilkins, O. L. May, C. W. Smith, J. R. Gannaway, J. F. Wendel and A. H.

Paterson (2007). "Meta-analysis of Polyploid Cotton QTL Shows Unequal Contributions of Subgenomes to a Complex Network of Genes and Gene Clusters Implicated in Lint Fiber Development." Genetics **176**(4): 2577-2588.

Rotili, E. A., L. L. Cancellier, M. A. Dotto, J. M. Peluzio and E. V. d. Carvalho (2012). "Genetic divergence in populations of maize, in Tocantins State, Brazil. / Divergência genética em genótipos de milho, no Estado do Tocantins." Revista Ciência Agronômica **43**(3): 516-521.

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

Saleem, M. F., M. F. Bilal, M. Awais, M. Q. Shahid and S. A. Anjum (2010). "Effect of nitrogen on seed cotton yield and fiber qualities of cotton (Gossypium hirsutum L.) cultivars." JAPS, Journal of Animal and Plant Sciences **20**(1): 23-27.

Sebolt, A. M., R. C. Shoemaker and B. W. Diers (2000). "Analysis of a quantitative trait locus allele from wild soybean that increases seed protein concentration in soybean." Crop Science **40**(5): 1438-1444.

Shappley, Z. W., J. N. Jenkins, J. Zhu and J. C. McCarty, Jr. (1998). "Quantitative trait loci associated with agronomic and fiber traits of upland cotton." Journal of Cotton Science **2**(4): 153-163.

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

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

Smith, C. W. and J. T. Cothren (1999). Cotton : origin, history, technology, and production / editors, C. Wayne Smith, J. Tom Cothren, New York : Wiley, c1999.

Song, X. and T. Zhang (2009). "Quantitative trait loci controlling plant architectural traits in cotton." Plant Science **177**(4): 317-323.

Stephens, J. C., F. R. Miller and D. T. Rosenow (1967). "Conversion of alien sorghums to early combine genotypes." Crop Science **7**: 396-396.

Swindle, M. G. (1993). Performance of F2 Hybrids in Cotton from Crosses of Exotic Germplasm and Cultivars, Mississippi State University. Department of Agronomy.

Team, R. C. (2015). " R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.".

Uga, Y., Y. Nonoue, Z. W. Liang, H. X. Lin, S. Yamamoto, U. Yamanouchi and M. Yano (2007). "Accumulation of additive effects generates a strong photoperiod sensitivity in the extremely late-heading rice cultivar 'Nona Bokra'." Theor Appl Genet **114**(8): 1457-1466.

Ulloa, M., W. R. Meredith Jr, Z. W. Shappley and A. L. Kahler (2002). "RFLP genetic linkage maps from four F2.3 populations and a joinmap of Gossypium hirsutum L." Theoretical & Applied Genetics **104**(2/3): 200-208.

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

USDA. (2013, March 5, 2013). "Cotton and Wool." Economic Research Service Retrieved August 12, 2013.

Van Esbroeck, G. and D. T. Bowman (1998). "Cotton germplasm diversity and its importance to cultivar development [electronic resource]." Journal of cotton science (Online) Journal of cotton science **2**(3): 121-129.

Verhalen, L. M. and J. C. Murray (1967). "Diallel analysis of several fiber property traits in Upland cotton Gossypium hirsutum L." Crop science **7**: 501-505.

Verhalen, L. M., J. C. Murray, E. S. Oswalt, J. W. Simmons, M. Simmons and K. C. Fun (1969). Cotton variety tests in Oklahoma 1963-67. Bulletin, Agricultural Experiment Station, Oklahoma State University.**:** 34-34.

Waddle, B. M., C. F. Lewis and T. R. Richmond (1961). "The genetics of flowering response in cotton. III. Fruiting behaviour of Gossypium hirsutum race lati-folium in a cross with a variety of cultivated American upland cotton." Genetics **46**: 427-437.

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

Wang, B., Y. Wu, N. Huang, X. Zhu, W. Guo and T. Zhang (2006). "QTL mapping for plant architecture traits in upland cotton using RILs and SSR markers." Acta Genetica Sinica **33**(2): 161-170.

Wang, M. I. N., C. Li and Q. Wang (2014). "Quantitative trait loci mapping and genetic dissection for lint percentage in upland cotton (Gossypium hirsutum)." Journal of Genetics **93**(2): 371-378.

Wang, S., C. J. Basten and Z.-B. Zeng (2012). "Windows QTL Cartographer 2.5." Department of Statistics, North Carolina State University, Raleigh, NC.

Wang, Z., D. Zhang, X. Wang, X. Tan, H. Guo and A. H. Paterson (2013). "A whole-genome DNA marker map for cotton based on the D-genome sequence of Gossypium raimondii L." G3 (Bethesda) **3**(10): 1759-1767.

Wendel, J. F. (1989). New World Tetraploid Cottons Contain Old World Cytoplasm, National Academy of Sciences of the United States of America**:** 4132.

Wendel, J. F. and V. A. Albert (1992). "Phylogenetics of the Cotton Genus (Gossypium): Character-State Weighted Parsimony Analysis of Chloroplast-DNA Restriction Site Data and Its Systematic and Biogeographic Implications." Systematic Botany(1): 115.

Wright, R. J., P. M. Thaxton, K. M. El-Zik and A. H. Paterson (1999). "Molecular mapping of genes affecting pubescence of cotton." Journal of Heredity **90**(1): 215-219.

Wu, C. T., B. L. Zhou and T. Z. Zhang Isolation and characterization of a sterile-dwarf mutant in Asian cotton (Gossypium arboreum L.).

Wu, J. X., J. C. McCarty, S. Saha, J. N. Jenkins and R. Hayes (2009). "Genetic changes in plant growth and their associations with chromosomes from Gossypium barbadense L. in G. hirsutum L." Genetica **137**(1): 57-66.

Yang, S., C. Foo, J. J. Lee, H. Misook, N. E. Wei, S. Sing-Hoi, D. M. Stelly, P. Thaxton, B. Triplett, C. D. Town and Z. J. Chen (2006). "Accumulation of genome-specific transcripts, transcription factors and phytohormonal regulators during early stages of fiber cell development in allotetraploid cotton." Plant Journal **47**(5): 761-775.

Zeng, L. H., W. R. Meredith, D. L. Boykin and E. Taliercio (2007). "Evaluation of an exotic germplasm population derived from multiple crosses among Gossypium tetraploid species." Journal of Cotton Science **11**(3): 118-127.

Zeng, L. H., W. R. Meredith, Jr., O. A. Gutiérrez and D. L. Boykin (2009). "Identification of associations between SSR markers and fiber traits in an exotic germplasm derived from multiple crosses among Gossypium tetraploid species." TAG Theoretical and Applied Genetics **119**(1): 93-103.

Zhang, J., W. Guo and T. Zhang (2002). "Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L. x Gossypium barbadense L.) with a haploid population." TAG. Theoretical And Applied Genetics. Theoretische Und Angewandte Genetik **105**(8): 1166-1174.

Zhang, J. F. and R. Percy (2007). Improving upland cotton by introducing desirable genes from pima cotton. World Cotton Research Conference-4, Lubbock, Texas, USA, 10-14 September 2007. Washington; USA, International Cotton Advisory Committee (ICAC)**:** un-un.

Zhang, T., N. Qian, X. Zhu, H. Chen, S. Wang, H. Mei and Y. Zhang (2013). "Variations and transmission of QTL alleles for yield and fiber qualities in upland cotton cultivars developed in China." PLoS ONE **8**(2): e57220-e57220.

Zhang, Z., J. Rong, V. N. Waghmare, P. W. Chee, O. L. May, R. J. Wright, J. R. Gannaway and A. H. Paterson (2011). "QTL alleles for improved fiber quality from a wild Hawaiian cotton, Gossypium tomentosum." TAG Theoretical and Applied Genetics **123**(7): 1075-1088.

Table 3.1: Scoring systems for morphological and plant architectural traits.

Phenotypes	Scoring
1. Plant height	Measured in cm
2. Flowering	0= Not yet flowered 1=Flowering, few if any bolls 2= Flowering, some bolls 3= Flowering, many bolls 4= Still flowering, near termini 5= Flowering completed, no flowers seen
3. Maturity of Bolls	0= No open bolls 1=at least 1 open boll, up to 25% 2=26-50% bolls open 3= 51-75% bolls open 4= more than 75% bolls open
4. Branching	- Number of nodes with at least one branch Nodes (1) - Number of nodes with two branches Nodes (2)

Table 3.2: Average performance of 2010 populations ($F_{2:3}$ progenies) for yield, morphology and plant architecture related traits.

Population	Total Yield (g)	Lint (%)	Height (cm)	Flowering	Maturity	Nodes (1)	Nodes (2)
MDN101 × PD94042	87.90	36.03	92.57	4	3	15.00	0
MDN101 × DES56	74.28	33.07	78.60	5	4	13.23	0
MDN101 × PMHS200	76.25	34.09	93.15	4	3	14.00	0
MDN101 × Acala Maxxa	95.30	35.29	87.37	4	4	14.33	1
MDN063 × DES56	114.55	34.45	87.89	4	4	13.34	0

Table 3.3: Average performance of 2011 populations (F_2 plants) for yield, morphology and plant architecture related traits.

Population	Total Yield	Lint %	Height	Flowering	Maturity	Nodes (≥1)	Nodes (≥2)
MDN101× PD94042	97.61	42.67	90.87	3.71	3.69	13.70	0.63
MDN101 × DES56	92.90	38.83	74.38	4.40	3.72	12.34	0.72
MDN101 × PMHS200	122.80	39.72	102.64	4.82	3.23	13.48	0.71
MDN101 × Acala Maxxa	77.26	37.24	91.14	3.77	3.47	14.01	0.52
MDN063 × DES56	97.98	40.21	73.57	3.82	3.91	12.80	0.40
MDN063 × Acala Maxxa	83.62	36.30	91.35	3.32	3.79	13.42	0.84
MDN257 × DES56	82.03	36.81	88.59	3.58	3.77	14.04	0.75
MDN257 × Acala Maxxa	90.98	36.31	91.32	3.64	3.63	14.6	0.95

Table 3.4: Markers significantly associated with yield related and plant architecture related traits.

Trait	Population	Year	Associated Marker	Pr
Plant Height	MDN063 × Acala Maxxa	2012	DPL0270_150	0.0009*
	MDN257 × DES56	2012	BNL3977_130	0.0024
	MDN101 × DES56	2012	MUCS616_180	0.009
	MDN101 × PMHS200	2012	NAU4047_390	0.005
Flowering	MDN101 × PD94042	2012	MONDPL0696_180	0.003*
	MDN063 × DES56	2012	NAU3074_180	0.002*
	MDN101 × DES56	2012	NAU1366_650	0.005
	MDN101 × PMHS200	2011	NAU1070_210	0.004
	MDN063 × DES56	2012	DPL0461_180	0.004
Maturity	MDN101 × DES56	2012	BNL3989_170	0.0002*
	MDN101 × DES56	2012	MUCS616_160	0.0001*
	MDN101 × DES56	2012	NAU3820_120	0.002*
	MDN101 × Acala Maxxa	2012	NAU5120_550	0.001*
Node (1)	MDN257 × DES56	2012	BNL3977_130	0.0006*
	MDN063 × DES56	2011	JESPR101_120	0.007
Lint Percentage	MDN101 × PD94042	2011	DPL0528_280	0.005
	MDN101 × PD94042	2012	DPL0270_150	0.005
	MDN101 × DES56	2012	NAU4042_450	0.009
	MDN063 × DES56	2012	BNL3359_280	0.008

*Significant after Bonferroni correction for multiple comparisons

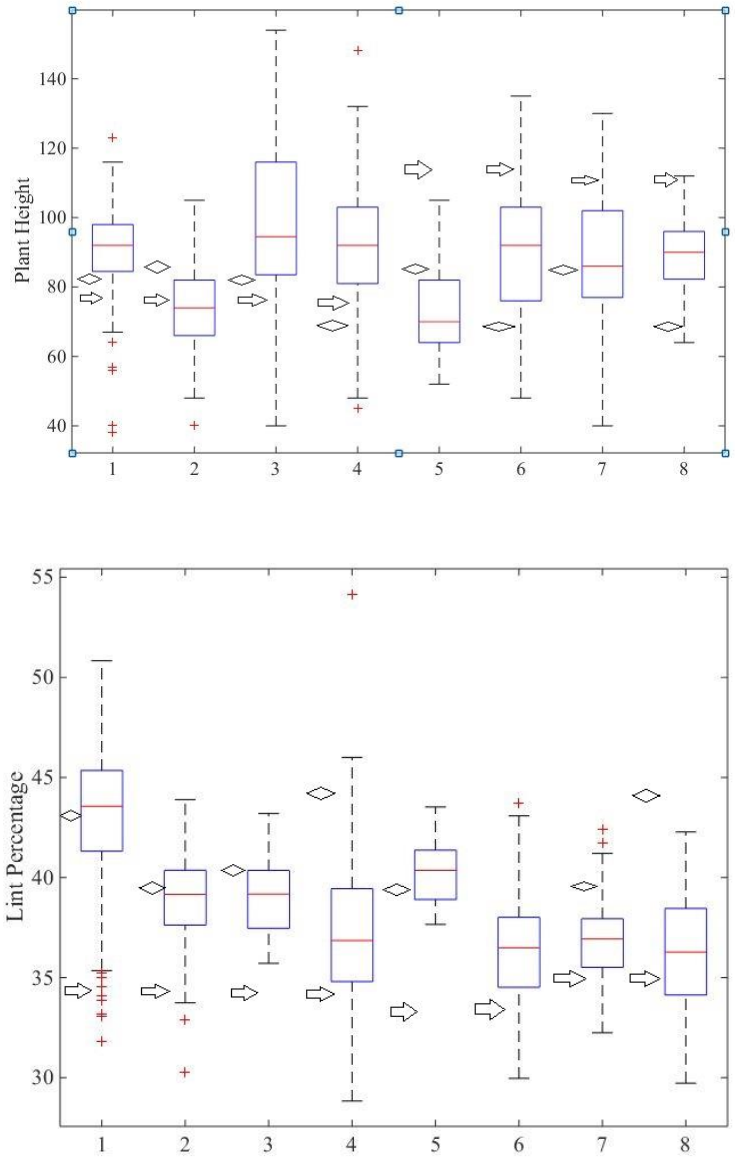

Figure 3.1: Distribution of populations for plant height and lint percentage. Y axis shows the range for each trait and x-axis shows different populations: (1) MDN101 × PD94042, (2) MDN101 × DES56, (3) MDN101 × PMHS200, (4) MDN101 × Acala Maxxa, (5) MDN063 × DES56, (6) MDN063 × Acala Maxxa, (7) MDN257 × DES56, and (8) MDN257 × Acala Maxxa. Average phenotypic values are shown by arrows for exotic parents (MDN101, MDN063 or MDN257) and wedges for elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

CHAPTER 4

SUMMARY

Despite stiff competition posed by synthetic fibers, the preference of the public for natural textile products has aided the survival of cotton as one of the major sources of natural fibers for the world textile industry. Modern spinning equipment requires long and strong cotton fibers for high efficiency. Moreover, the consumer is more interested in high quality textile products, requiring high quality cotton fibers. Much research efforts, both conventional and molecular, has been applied for the last few decades to improve both the quantity and quality of cotton fibers. Some progress has been achieved in increasing important parameters of cotton yield and fiber quality. Despite these advances, there is only slow improvement of these yield and fiber quality related traits. The narrow genetic base of cultivated cotton genotypes, mainly because of overexploitation of closely-related genotypes, has been a major obstacle in the improvement of cotton cultivars. Increasing the genetic diversity of the cultivated cotton gene pool, by increased utilization of widely available primitive and exotic accessions of cotton in breeding programs, is one way to improve the performance of cultivated cotton genotypes.

In our study, we made use of three exotic *G. hirsutum* accessions, collected from different parts of Mexico and Guatemala and converted to day neutral flowering, making experimental populations and studying the F_2 and $F_{2:3}$ generations of these populations for six different fiber quality traits and other yield, yield related and plant architecture related traits. Carefully selected SSR markers from 18 targeted regions of the cotton genome known to be rich

in fiber quality QTLs were deployed to identify significant association of these markers with the traits scored. We revealed significant associations between SSR markers and different fiber quality traits in all the populations. Some associations were identified in more than one population and some were identified in more than one year. This indicates the stability of these associations in different environments. Some associations were identified in more than one population. For many of these associations, the exotic parents were the source of useful and favorable alleles. This is indicative of the usefulness of the exotic parents not only for multiple fiber quality traits but also for their positive useful effects in multiple genetic backgrounds and in multiple environments.

The same set of markers used to identify associations with fiber quality traits were used to study their relation with yield related and plant architectural traits. Many marker loci that were significantly associated with fiber quality traits, also showed significant association with these traits, revealing the usefulness of these markers to track multiple traits of interest in the same or different populations. NAU3820_120, which showed significant association with STR in MDN101 × DES56 population, also showed significant association with boll maturity in the same population. Similarly, DPL0270_150, which showed significant association with UHM in MDN101 × DES56 populations, showed significant association with plant height in MDN063 × Acala Maxxa population. These associations represent potential QTLs for specific traits in specific region of the cotton genome and could be highly useful in breeding for the traits in question. Moreover, some of these associations were related to exotic parents contributing favorable alleles to different genetic backgrounds, indicating the usefulness of the exotic parents not only for fiber quality traits but also for other traits of agronomic importance. Further evaluation and validation of these associations for possible QTLs and their effects may be greatly

beneficial for breeding for improvement of fiber quality as well as useful plant morphology and architecture in Upland cotton, separately as well as simultaneously.

SUPPLEMENTARY TABLES

Table S1: List of populations made in 2010 and 2011 with the number of individuals in each population.

		Number of F_2 plants	
SN	**Population**	**2010 Population**	**2011 Population**
1	MDN101 × PD94042	28	110
2	MDN101 × DES56	15	106
3	MDN101 × PMHS200	66	39
4	MDN101 × Acala Maxxa	107	74
5	MDN063 × DES56	49	45
6	MDN063 × Acala Maxxa	-	125
7	MDN257 × DES56	-	72
8	MDN257 × Acala Maxxa	-	68

Table S2: Performance of parents for fiber quality traits in three years.

Parents	Year	MIC	UHM	UI	STR	ELO	SFC%
PD94042	2011	5.26	1.00	80.03	27.83	4.57	9.25
	2012	4.68	1.11	83.85	30.05	6.35	6.95
	2013	4.85	1.07	83.23	29.93	7.15	8.30
DES56	2011	5.04	1.04	81.50	28.84	4.36	8.58
	2012	4.68	1.09	82.28	28.98	5.00	8.07
	2013	4.69	1.06	82.60	25.56	5.39	8.58
PMHS200	2011	4.88	0.92	79.85	27.20	5.00	10.50
	2012	4.88	1.11	82.94	30.40	6.34	7.56
	2013	4.44	1.09	82.70	28.86	6.50	8.14
Acala Maxxa	2011	4.17	1.10	83.23	34.22	4.30	7.90
	2012	4.01	1.16	84.20	33.80	5.20	7.30
	2013	4.28	1.10	83.038	30.09	5.38	7.88
MDN063	2011	4.99	0.99	80.74	28.53	4.73	8.87
	2012	4.09	1.10	82.90	32.10	6.80	7.30
	2013	4.42	0.99	81.70	29.66	6.04	9.56
MDN101	2011	5.32	1.01	81.23	28.63	4.75	8.17
	2012	3.92	1.03	81.93	29.60	6.13	7.80
	2013	4.83	0.97	81.70	26.16	6.58	8.52
MDN257	2011	5.38	0.99	80.95	28.10	4.25	8.50
	2012	4.7	1.06	82.73	29.69	5.69	7.30
	2013	5.01	0.98	80.90	26.24	6.30	9.33

Table S3: Performance of populations created in 2010 for six different fiber quality traits.

Population		MIC	UHM	UI	STR	ELO	SFC%
MDN101× PD94042	F_2	5.25	1.02	81.84	27.39	4.37	8.44
	$F_{2:3}$	4.54	1.04	82.84	28.53	6.69	7.46
MDN101 × DES56	F_2	5.37	1.00	81.07	26.41	4.37	8.99
	$F_{2:3}$	3.87	1.09	82.28	28.92	5.54	8.27
MDN101 × PMHS200	F_2	5.20	0.99	81.08	28.81	3.97	8.72
	$F_{2:3}$	4.93	1.06	83.09	29.68	5.61	7.22
MDN101 × Acala Maxxa	F_2	5.31	1.03	82.1	30.42	4.84	7.92
	$F_{2:3}$	4.29	1.09	83.39	30.48	6.31	7.13
MDN063 × DES56	F_2	4.69	1.02	80.49	27.13	4.14	9.70
	$F_{2:3}$	3.75	1.11	82.70	29.17	5.82	8.03

Table S4: Performance of populations created in 2011 for six different fiber quality traits.

Population		MIC	UHM	UI	STR	ELO	SFC%
MDN101× PD94042	F_2	4.84	1.08	83.49	29.77	7.15	7.30
	$F_{2:3}$	4.94	1.04	82.68	27.70	6.88	8.22
MDN101 × DES56	F_2	4.22	1.13	83.99	29.69	6.95	7.13
	$F_{2:3}$	4.40	1.07	83.11	27.49	7.00	7.88
MDN101 × PMHS200	F_2	4.68	1.07	82.74	29.04	6.14	7.93
	$F_{2:3}$	4.76	1.04	82.53	27.20	6.32	8.23
MDN101 × Acala Maxxa	F_2	4.42	1.09	83.46	31.12	6.74	7.18
	$F_{2:3}$	4.68	1.05	83.12	29.38	6.64	7.67
MDN063 × DES56	F_2	4.47	1.16	84.04	31.74	6.61	7.28
	$F_{2:3}$	4.83	1.12	83.83	29.33	7.09	7.76
MDN063 × Acala Maxxa	F_2	3.55	1.12	81.97	30.18	5.35	8.52
	$F_{2:3}$	3.95	1.08	82.59	28.84	5.80	8.38
MDN257 × DES56	F_2	4.55	1.01	82.17	27.40	6.67	8.01
	$F_{2:3}$	5.02	0.96	81.95	25.51	7.66	8.66
MDN257 × Acala Maxxa	F_2	3.97	1.11	83.14	31.27	6.00	7.30
	$F_{2:3}$	3.94	1.05	82.71	28.39	6.30	8.20

Table S5: Genomic distribution of polymorphic SSR markers.

Hotspot	Dt	Markers Selected	Start Base Pair	End Base Pair	Span (Base Pairs)	Hotspot covered (%)
I	D02	5	5633399	17813958	12180559	63.07
II	D03	10	6090128	34035242	27945114	86.66
III	D05	7	4990385	35898072	30907687	68.35
IV	D07	3	4461008	5159719	698711	33.08
V	D07	3	14156087	18295298	4139211	74.51
VI	D07	5	35765492	51387770	15622278	55.25
VII	D08	9	42588287	54931420	12343133	87.07
VIII	D09	5	12852584	16027393	3174809	58.49
IX	D01	7	15639969	39803373	24163404	86.57
X	D03	1	42640991	42640991	-	-
XI	D04	2	55617129	56392452	775323	9.25
XII	D05	14	55367094	61974235	6607141	28.31
XIII	D06	11	38461976	47417863	8955887	84.56
XIV	D07	6	18295298	47659512	29364214	92.71
XV	D08	15	39438749	55198803	15760054	84.93
XVI	D09	1	7408472	7408472	-	-
XVII	D09	9	12852584	27281356	14428772	81.62
XVIII	D11	0	-	-	-	-

Table S6: Number and type of fiber quality QTLs reported in the "hotspots".

Hotspot	D Chr.	No of QTLs	At Chr.	At QTLs	Dt Chr.	Dt QTLs
I	D01	4	7	-	16	FU16.1, FL16.1, FF16.1, MIC16.1
II	D02	10	1	FL01.1, FL01.2, EL01.1, ELO1.2 FU01.1, FS01.1, FF01.1, FF01.2, MIC01.1,	15	FF15.1
III	D03	5	2/3	-	17	FL17.1, FU17.1, FS17.1,MIC17.1
IV	D03	5	2/3	FL02.1, FS02.2, FL03.1, FU03.1, SF03.1	17	-
V	D04	5	LGA02	-	LGD03	FFDO3.05, FFD03.1, FLD03.1, FUDO3.1, ELDO3.1
VI	D05	6	2,3	-	14	FF14.05, FF14.1, FU14.2, FL14.1, MIC14.2, FS14.1
VII	D05	8	2,3	EL02.1, FF02.1, FF02.2, FF03.1, FL02.1, FL02.2, FU02.1, MIC02.1	14	-
VIII	D06	7	9	-	23	EL23.03, EL23.05, EL23.1, FF23.1, FU23.1, FS23.1, FS23.2
IX	D07	5	LGA03	-	LGD02	FFD02.1, FLD02.2, FSD02.1, FSD02.2, ELD02.1
X	D07	6	LGA03	ELA03.1, FLA03.1,FFA03.2, FSA03.1, FLA03.2,FUA03.2	LGD02	-
XI	D07	5	LGA03	FUA03.1, ELA03.1, FLA03.1, FFA03.1, FFA03.2	LGD02	-
XII	D07	5	LGA03	FLA03.2,FUA03.2,ELA03.2, FUA03.3,FCA03.1	LGD02	-
XIII	D08	12	12	-	26	FL26.1,FL26.2,EL26.1, MIC26.1, FF26.1,MIC26.2, FS26.1,FU26.1, FS26.2, FS26.3,SF26.2,FU26.2
XIV	D08	6	12	FF12.1,FF12.15,FL12.1, FU12.2,EL12.1,FF12.2	26	-
XV	D09	9	4/5	-	LGD08	ELD08.1,FFD08.1,MICD08.1, ELD08.2,FLD08.1,FUD08.1, SFD08.1,FSD08.2,FFD08.2
XVI	D09	10	4/5	-	LGD08	FFD08.3,MICD08.2,ELD08.3, ELD08.4,FLD08.2,FUD08.2, SFD08.2,SFD08.15,FFD08.4, MICD08.3
XVII	D09	5	4/5	FF05.15,MIC05.1,FF05.2, EL05.05,FF05.3	LGD08	-
XVIII	D11	5	10	-	20	EL20.1,FL20.1,SF20.1, FL20.15,FF20.1

EL: fiber elongation, FF: fiber fineness, FL: fiber length, FS: fiber strength, FU: fiber uniformity, SF: short fiber content, QTLs: quantitative trait loci, At Chr.: A sub-genome chromosome, Dt Chr.: D sub-genome chromosome, D Chr.: D genome *Gossypium raimondii* chromosome. (Wang, Zhang et al. 2013)

Table S7: Markers associated with six different fiber quality traits.

Population	Year(F2)	Planting Date	Trait	Associated Marker	P Value	R^2
MDN101 × PD94042	2011	1	MIC	NAU3820_120	0.02	0.20
	2011	1	STR	TMB2933_550	0.04	0.15
	2011	2	ELO	NAU1042_270	0.02	0.31
	2011	2	MIC	HAU0536_270	0.03	0.25
	2011	2	STR	BNL2725_250	0.04	0.23
	2011	2	UHM	NAU1042_270	0.03	0.25
	2012	1	ELO	BNL1317_230	0.03	0.04
	2012	1	ELO	NAU3820_120	0.02	0.06
	2012	1	MIC	NAU2152_600	0.004	0.08
	2012	1	SFC	TMB1484_390	0.04	0.04
	2012	1	STR	NAU2152_600	0.02	0.05
	2012	1	STR	NAU1042_270	0.04	0.04
MDN101 × DES56	2011	1	ELO	NAU1366_450	0.009	0.37
	2011	1	ELO	NAU2190_510	0.007	0.45
	2011	1	ELO	NAU1221_390	0.02	0.33
	2011	1	UI	NAU4042_230	0.03	0.34
	2011	2	ELO	JESPR101_280	0.0006*	0.43
	2011	2	ELO	DPL0270_150	0.02	0.24
	2011	2	ELO	NAU2336_700	0.02	0.24
	2011	2	ELO	NAU5120_170	0.0003*	0.47
	2011	2	ELO	NAU5465_220	0.002*	0.40
	2011	2	MIC	DPL0270_150	0.02	0.27
	2011	2	MIC	NAU1221_390	0.02	0.22
	2011	2	SFC	JESPR101_280	0.007	0.29
	2011	2	SFC	NAU1366_450	0.02	0.10
	2011	2	SFC	NAU2336_700	0.03	0.22
	2011	2	SFC	NAU5120_170	0.002*	0.37
	2011	2	SFC	NAU5465_220	0.02	0.26
	2011	2	UHM	DPL0270_150	0.004*	0.33
	2011	2	UHM	NAU1221_390	0.009	0.28
	2011	2	UI	JESPR101_280	0.003*	0.34
	2011	2	UI	DPL0270_150	0.01	0.28
	2011	2	UI	NAU5120_170	0.0003*	0.46
	2011	2	UI	NAU5465_220	0.02	0.26
	2012	1	ELO	NAU1366_450	0.02	0.05
	2012	1	MIC	NAU5465_220	0.02	0.06
	2012	1	SFC	NAU4042_450	0.02	0.05
	2012	1	STR	BNL3989_170	0.02	0.05

	2012	1	STR	NAU3820_120	0.0017*	0.10
MDN101 × PMHS200	2011	1	ELO	BNL1317_230	0.007	0.13
	2011	1	MIC	NAU1042_180	0.005	0.15
	2011	1	MIC	NAU1070_210	0.04	0.07
	2011	1	MIC	DPL0156_230	0.002*	0.16
	2011	1	SFC	DPL0501_300	0.02	0.10
	2011	1	SFC	BNL3359_300	0.02	0.09
	2011	1	SFC	NAU5465_240	0.03	0.08
	2011	1	SFC	DPL0378_180	0.002*	0.19
	2011	1	STR	NAU3820_120	0.03	0.09
	2011	1	UHM	NAU1042_270	0.02	0.09
	2011	1	UHM	NAU1070_210	0.02	0.10
	2011	1	UHM	NAU3820_120	0.02	0.10
	2011	1	UHM	DPL0156_230	0.007	0.13
	2012	1	ELO	DPL0501_300	0.02	0.16
	2012	1	SFC	BNL1317_220	0.03	0.13
	2012	1	STR	BNL1317_220	0.02	0.14
	2012	1	STR	NAU2152_250	0.04	0.12
	2012	1	UHM	BNL1317_220	0.04	0.11
MDN101 × Acala Maxxa	2011	1	MIC	DPL0279_190	0.03	0.05
	2011	1	MIC	CIR141_250	0.0003*	0.12
	2011	1	STR	NAU5120_600	0.04	0.04
	2011	1	STR	NAU1221_220	0.02	0.05
	2011	1	STR	NAU1042_220	0.003	0.12
	2011	1	SFC	NAU1042_220	0.02	0.11
	2011	1	SFC	CIR141_240	0.04	0.03
	2011	1	UHM	NAU1042_220	0.03	0.06
	2011	1	UI	NAU1042_220	0.03	0.08
	2012	1	ELO	DPL0279_190	0.02	0.06
	2012	1	MIC	DPL0279_190	0.008	0.09
	2012	1	MIC	BNL3359_270	0.04	0.05
	2012	1	SFC	NAU5120_450	0.02	0.06
	2012	1	STR	DPL0279_190	0.006	0.09
	2012	1	UHM	DPL0279_190	0.02	0.08
	2012	1	UHM	BNL3359_270	0.03	0.06
	2012	1	UI	NAU5120_450	0.0014*	0.13
MDN063 × DES56	2011	1	SFC	NAU5120_450	0.04	0.07
	2011	1	UHM	DPL0665_180	0.04	0.09
	2011	1	UHM	JESPR37_350	0.03	0.10

	2011	1	UHM	JESPR37_270	0.02	0.11
	2011	1	UI	NAU5120_450	0.006	0.13
	2011	2	ELO	NAU5120_450	0.03	0.04
	2011	2	ELO	MUSB1020_650	0.02	0.08
	2011	2	ELO	MUCS616_240	0.007	0.10
	2011	2	ELO	NAU5465_270	0.02	0.09
	2011	2	MIC	DPL0528_350	0.02	0.08
	2011	2	MIC	NAU1188_750	0.02	0.07
	2011	2	MIC	NAU4047_390	0.005	0.10
	2011	2	MIC	BNL1673_240	0.02	0.08
	2011	2	SFC	NAU5120_450	0.007	0.08
	2011	2	SFC	MUSB1020_650	0.04	0.06
	2011	2	SFC	NAU4047_390	0.02	0.07
	2011	2	SFC	MUCS616_240	0.02	0.09
	2011	2	SFC	NAU5465_270	0.03	0.06
	2011	2	STR	DPL0528_350	0.004	0.11
	2011	2	STR	NAU4047_390	0.02	0.08
	2011	2	STR	MUCS616_240	0.02	0.07
	2011	2	UHM	DPL0528_350	0.003	0.12
	2011	2	UHM	NAU4047_390	0.02	0.08
	2011	2	UI	DPL0528_350	0.02	0.08
	2011	2	UI	NAU5120_450	0.0004*	0.14
	2011	2	UI	MUSB1020_650	0.04	0.06
	2012	1	MIC	NAU3074_180	0.008	0.16
	2012	1	MIC	NAU1042_130	0.04	0.09
	2012	1	MIC	NAU2152_240	0.004	0.17
	2012	1	STR	MUSB1020_650	0.04	0.09
	2012	1	UHM	NAU2190_510	0.04	0.02
MDN063 × Acala Maxxa	2012	1	MIC	JESPR183_150	0.03	0.04
	2012	1	MIC	MONDPL0696_250	0.006	0.06
	2012	1	SFC	HAU2653_750	0.006	0.06
	2012	1	STR	MUSB1020_700	0.04	0.03
	2012	1	UI	HAU2653_750	0.009	0.06
MDN257 × DES56	2012	1	ELO	BNL1059_200	0.03	0.07
	2012	1	MIC	NAU3016_250	0.02	0.09
	2012	1	MIC	NAU2152_240	0.03	0.07
	2012	1	MIC	JESPR37_1000	0.02	0.08
	2012	1	SFC	BNL3359_390	0.02	0.09
	2012	1	SFC	DPL0378_510	0.04	0.06
	2012	1	STR	BNL3359_390	0.005	0.11
	2012	1	STR	DPL0378_510	0.02	0.09

	2012	1	STR	JESPR37_1000	0.003	0.12
	2012	1	UHM	JESPR183_220	0.03	0.07
	2012	1	UHM	BNL3590_180	0.02	0.08
	2012	1	UHM	NAU3778_240	0.007	0.11
	2012	1	UHM	BNL3359_390	0.004	0.12
	2012	1	UHM	NAU3016_250	0.007	0.11
	2012	1	UHM	DPL0378_510	0.0003*	0.18
	2012	1	UHM	NAU5120_150	0.04	0.06
	2012	1	UHM	JESPR37_1000	0.0002*	0.19
	2012	1	UI	BNL3359_390	0.004	0.12
	2012	1	UI	NAU5120_150	0.02	0.08
MDN257 × Acala Maxxa	2012	1	ELO	NAU3016_270	0.02	0.09
	2012	1	MIC	NAU2715_220	0.02	0.09
	2012	1	SFC	MONDPL0696_270	0.04	0.07
	2012	1	SFC	BNL1317_300	0.03	0.08
	2012	1	STR	NAU4047_450	0.02	0.09

*Significant after Bonferroni correction for multiple comparisons.

Table S8: Performance of parents for yield, yield related and plant architectural traits.

Parents	Year	Total Yield (g)	Lint %	Height (cm)	Flowering	Maturity	Nodes (1)	Nodes (2)
PD94042	2012	142.70	43.00	85.00	5	4	17	1
	2013	137.43	45.09	108.00	4	4	13	1
DES56	2012	101.04	39.50	87.00	4	4	13	0
	2013	115.08	41.97	112.13	4	5	11	1
PMHS200	2012	145.76	41.30	86.00	4	4	13	1
	2013	144.39	41.89	112.20	4	4	13	0
Acala Maxxa	2012	125.01	44.00	73.00	3	4	11	1
	2013	107.30	42.34	91.13	4	3	10	1
MDN063	2012	179.63	32.40	116.00	5	4	18	0
	2013	142.96	33.45	104.00	4	3	14	0
MDN101	2012	137.51	34.90	78.00	5	3	15	1
	2013	129.83	37.28	109.70	4	3	13	0
MDN257	2012	98.61	35.00	115.00	5	3	18	1
	2013	88.23	38.78	125.91	4	3	15	1

Table S9: Markers associated with total seed cotton yield.

Population	Year	Associated Marker	P value	R^2	Additive Effect [a]
MDN101 × PD94042	2011	TMB1484_390	0.04	0.12	-22.51
	2011	NAU3820_260	0.02	0.22	13.49
MDN101 × DES56	2012	NAU5120_550	0.04	0.04	-2.66
MDN101 × PMHS200	2011	NAU3820_118	0.01	0.10	8.78
	2011	HAU0536_280	0.04	0.07	3.68
MDN101 × Acala Maxxa	2011	DPL0270_150	0.04	0.04	0.57
	2012	DPL0279_190	0.02	0.06	-15.31
	2012	DPL0378_610	0.01	0.06	-0.10
MDN063 × DES56	2011	JESPR37_350	0.02	0.11	0.08
	2011	NAU5465_270	0.04	0.08	15.58
	2012	NAU3074_200	0.03	0.09	8.76
	2012	DPL0270_220	0.04	0.09	-7.81
	2012	NAU5465_310	0.04	0.09	25.98
	2012	BNL3359_280	0.04	0.09	-25.98
MDN257 × DES56	2012	BNL3590_180	0.03	0.06	-0.10
	2012	NAU3778_240	0.01	0.09	2.24
	2012	NAU5120_150	0.01	0.08	0.47
	2012	JESPR37_1000	0.01	0.08	-0.22
MDN257 × Acala Maxxa	2012	BNL3359_280	0.03	0.07	1.29

[a] Positive values indicate favorable alleles coming from exotic parents (MDN063, MDN101 or MDN257) and negative values indicate favorable alleles coming from elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

Table S10: Markers associated with lint percentage.

Population	Year	Associated Marker	P value	R^2	Additive Effect [a]
MDN101 × PD94042	2011	TMB2933_550	0.04	0.14	-1.43
	2011	TMB1484_390	0.03	0.17	-1.60
	2011	DPL0279_200	0.02	0.19	3.31
	2011	DPL0528_280	0.005	0.19	1.68
	2011	HAU0536_270	0.02	0.21	-2.16
	2012	DPL0270_150	0.005	0.07	-0.20
	2012	NAU2152_170	0.03	0.02	-1.15
MDN101 × DES56	2012	NAU4042_450	0.009	0.06	-2.32
MDN101 × PMHS200	2012	HAU2525_250	0.02	0.14	-0.89
MDN101 × Acala Maxxa	2011	NAU2152_170	0.02	0.05	0.74
	2011	CIR141_250	0.02	0.04	0.44
	2011	DPL0605_230	0.02	0.03	-1.15
	2011	DPL0378_650	0.03	0.03	0.65
	2012	NAU1221_220	0.03	0.05	-1.09
	2012	NAU1042_200	0.04	0.04	1.33
MDN063 × DES56	2011	JESPR101_120	0.04	0.07	1.05
	2011	BNL3590_180	0.03	0.07	-1.14
	2011	BNL3359_280	0.008	0.14	-1.74
MDN063 × Acala Maxxa	2012	NAU2251_250	0.04	0.03	-0.71
MDN257 × Acala Maxxa	2012	NAU4047_450	0.02	0.09	1.12

*Significant after Bonferroni Correction

[a] Positive values indicate favorable alleles coming from exotic parents (MDN063, MDN101 or MDN257) and negative values indicate favorable alleles coming from elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

Table S11: Markers associated with plant height.

Population	Year	Associated Marker	P value	R^2	Additive Effect [a]
MDN101 × Acala Maxxa	2011	BNL3359_270	0.03	0.04	3.65
	2011	CIR141_240	0.04	0.03	0.81
MDN101 × DES56	2012	DPL0279_200	0.02	0.05	-2.38
	2012	MUCS616_180	0.009	0.06	-2.86
MDN101 × PHMS200	2012	DPL0378_180	0.005	0.19	-25.47
	2012	DPL0501_300	0.04	0.11	-10.40
	2012	NAU4042_180	0.05	0.15	-11.98
MDN101 × Acala Maxxa	2012	DPL0279_200	0.02	0.07	-5.12
	2012	NAU1221_220	0.04	0.02	5.67
MDN063 × Acala Maxxa	2012	DPL0270_150	0.0009*	0.09	-7.64
MDN257 × DES56	2012	BNL3977_130	0.003	0.12	-7.56
MDN257 × Acala Maxxa	2012	TMB1409_240	0.02	0.08	-3.10
	2012	MONDPL0696_270	0.04	0.07	2.73
	2012	MUCS426_350	0.03	0.07	-2.63

*Significant after Bonferroni Correction

[a] Negative values indicate favorable alleles for shorter plant height coming from exotic parents (MDN063, MDN101 or MDN257) and positive values indicate favorable alleles for shorter plant height coming from elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

Table S12: Markers associated with flowering.

Population	Year	Associated Marker	Pr	R^2	Additive [a] Effect
MDN101 × PD94042	2012	MONDPL0696_180	0.003*	0.09	0.24
	2012	NAU1221_510	0.06	0.04	-0.22
MDN101 × DES56	2011	NAU3820_120	0.06	0.43	0.04
	2012	BNL3359_280	0.02	0.06	0.19
	2012	DPL0279_200	0.07	0.03	0.13
	2012	MUCS616_160	0.07	0.03	0.14
	2012	NAU1221_510	0.02	0.04	0.23
	2012	NAU1366_650	0.005	0.07	-0.30
	2012	NAU2190_510	0.02	0.05	0.28
	2012	NAU3074_310	0.03	0.04	-0.16
	2012	NAU5120_550	0.07	0.03	0.15
MDN101 × PMHS200	2011	NAU1042_510	0.04	0.04	-0.11
	2011	NAU1070_210	0.004	0.05	-0.12
	2012	HAU2525_250	0.09	0.16	-0.09
MDN101 × Acala Maxxa	2012	DPL0378_590	0.04	0.02	-0.31
	2012	NAU5120_550	0.06	0.05	-0.30
MDN063 × DES56	2012	DPL0461_180	0.004	0.19	-0.01
	2012	NAU2152_240	0.07	0.07	0.27
	2012	NAU2190_450	0.02	0.14	-0.42
	2012	NAU3074_180	0.002*	0.21	0.51
MDN257 × DES56	2012	DPL0378_590	0.06	0.03	-0.18
MDN257 × Acala Maxxa	2012	BNL3977_450	0.04	0.06	0.25
	2012	NAU2671_310	0.04	0.01	-0.08
	2012	NAU5465_240	0.08	0.01	-0.02

*Significant after Bonferroni Correction

[a] Positive values indicate favorable alleles coming from exotic parents (MDN063, MDN101 or MDN257) and negative values indicate favorable alleles coming from elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

Table S13: Markers associated with boll maturity.

Population	Year	Associated Marker	P Value	R^2	Additive Effect [a]
MDN101 × PD94042	2012	BNL2725_550	0.04	0.04	0.09
	2012	TMB2933_550	0.03	0.05	-0.17
MDN101 × DES56	2011	NAU4042_650	0.06	0.39	-0.44
	2012	BNL3989_170	0.0002*	0.16	0.22
	2012	MUCS616_160	0.0001*	0.12	-0.18
	2012	NAU2640_230	0.07	0.03	0.10
	2012	NAU3074_310	0.02	0.05	0.12
	2012	NAU3820_120	0.002*	0.10	-0.00
	2012	NAU4042_230	0.06	0.04	0.13
	2012	NAU5120_550	0.04	0.05	-0.12
	2012	NAU5465_220	0.02	0.07	-0.15
MDN101 × PMHS200	2011	NAU4042_180	0.09	0.05	-0.07
	2012	DPL0378_180	0.08	0.11	0.83
	2012	DPL0501_300	0.03	0.13	0.51
MDN101 × Acala Maxxa	2011	BNL3359_270	0.02	0.04	0.19
	2011	DPL0378_590	0.09	0.01	0.07
	2011	DPL0605_230	0.06	0.01	-0.03
	2011	NAU1221_220	0.09	0.09	-0.24
	2011	NAU2152_170	0.02	0.04	0.13
	2011	NAU5120_510	0.049	0.03	0.11
	2012	DPL0378_590	0.02	0.01	-0.16
	2012	NAU1042_200	0.06	0.03	0.24
	2012	NAU5120_550	0.001*	0.15	0.61
MDN063 × DES56	2011	BNL1673_240	0.07	0.01	0.03
	2011	DPL0528_510	0.02	0.06	-0.08
	2011	DPL0665_180	0.06	0.03	-0.06
	2011	HAU2653_390	0.06	0.01	0.02
	2011	MUCS616_240	0.04	0.05	-0.08
	2011	MUSB1020_650	0.06	0.05	-0.04
	2011	NAU5465_270	0.02	0.07	0.13
	2012	DPL0461_180	0.08	0.06	0.12
MDN257 × Acala Maxxa	2012	NAU5120_450	0.08	0.01	-0.05

*Significant after Bonferroni Correction

[a] Positive values indicate favorable alleles coming from exotic parents (MDN063, MDN101 or MDN257) and negative values indicate favorable alleles coming from elite parents (PD94042, DES56, PMHS200 or Acala Maxxa).

Table S14: Markers associated with nodes with at least one branch (Nodes 1).

Population	Year	Associated Marker	P value	R^2
MDN101 × PD94042	2012	HAU0536_270	0.08	0.03
MDN101 × DES56	2012	NAU2190_510	0.003	0.01
MDN101 × PMHS200	2012	NAU2152_250	0.08	0.02
MDN101 × Acala Maxxa	2012	DPL0378_590	0.03	0.01
	2012	NAU5120_510	0.03	0.01
	2012	NAU5465_270	0.08	0.02
MDN063 × DES56	2011	BNL3435_220	0.09	0.01
	2011	DPL0528_390	0.06	0.01
	2011	JESPR101_120	0.007	0.02
	2011	MUSB1020_650	0.08	0.07
	2012	HAU2653_390	0.05	0.15
MDN257 × DES56	2012	BNL3977_130	0.0006*	0.06
	2012	NAU5120_150	0.01	0.04
	2012	NAU5461_510	0.03	0.03
MDN257 × Acala Maxxa	2012	BNL4035_150	0.03	0.02
	2012	NAU3016_450	0.03	0.05
	2012	NAU4047_450	0.02	0.04

*Significant after Bonferroni Correction