GENETIC EVALUATION OF EXOTIC CHROMATINS FROM TWO OBSOLETE INTERSPECIFIC INTROGRESSION LINES OF UPLAND COTTON

by

PAWAN KUMAR

(Under the Direction of Peng Wah Chee)

ABSTRACT

Germplasm lines derived through interspecific introgressions are important resource for plant breeders targeting novel alleles from wild or unadapted relatives to improve elite germplasm. Among the four Gossypium species domesticated for their cotton fibers, cultivated form of *Gossypium hirsutum* L. commonly called Upland cotton accounts for 96% of global cotton fiber production. The other three cultivated species, including *Gossypium barbadense* L., *Gossypium arboreum* L., and *Gossypium herbaceum* L., have been replaced by Upland cotton due its high yield potential and broader environmental adaptation. In Upland cotton historical records suggested that several obsolete germplasm lines with excellent fiber quality may have been derived via interspecific introgression; however the effective utilization of such lines in marker assisted breeding rests on the ability to target and exploit segments harboring favorable alleles.

The objective of this study is to identify and evaluate the effects of alien chromosomal segments in two *G. barbadense* introgressed lines, Sealand 542 & Sealand 883. F2 mapping populations were developed by crossing the Sealand lines with Upland

cotton lines representing major types of genetic backgrounds of the US cotton. We used 1170 SSR markers to identify locations of the putative introgressed segments. Our results show that a total of 22 putative introgressions were detected including 12 introgressions on 7 chromosomes in the Sealand 542 genome and 10 introgressions on 5 chromosomes in the Sealand 883 genome. QTL analysis revealed that a number of the identified introgressed regions harbored positive alleles for fiber quality traits. Efficacy of these introgressed alleles were tested over two generations ($F_{2:3}$ and $F_{2:4}$) across multiple genetic backgrounds and our results show that several of the introgressed fiber quality alleles from *G. barbadense* were consistently detected over generations as well as across genetic backgrounds. However the percent phenotypic variation explained by the introgressed QTLs varied across genetic backgrounds suggesting epistatic interaction with the genetic background in which these were present. Prospects of utilizing Sealand lines in Upland cotton breeding programs and germplasm improvement are also discussed.

INDEX WORDS: Cotton, Sealand 542, Sealand 883, Interspecific introgression, Fiber quality, Upland cotton, Quantitative Trait Loci (QTL), Genetic backgrounds

GENETIC EVALUATION OF EXOTIC CHROMATINS FROM TWO OBSOLETE INTERSPECIFIC INTROGRESSION LINES OF UPLAND COTTON

by

PAWAN KUMAR

B.S., University of Agricultural Sciences, Dharwar, India, 2005

M.S., Punjab Agricultural University, Ludhiana, India, 2001

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

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2012

GENETIC EVALUATION OF EXOTIC CHROMATINS FROM TWO OBSOLETE INTERSPECIFIC INTROGRESSION LINES OF UPLAND COTTON

by

PAWAN KUMAR

Major Professor: Peng Wah Chee

Committee: Andrew Paterson
Peggy Ozias-Akins
Benjamin Todd Campbell
Guy Collins

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2012

DEDICATION

I dedicate this dissertation to my parents, to my loving and supporting wife Rippy Singh and to my loving son Aryan Sharma.

ACKNOWLEDGEMENTS

This dissertation could not be had been completed without the help and support of so many people. First of all, I would like to thank Dr. Peng Chee, my major professor, for his guidance, patience, encouragement and assistance in completion of my doctoral dissertation.

Also, I would like to thank my present and former committee members including Dr. Andrew Paterson, Dr. Peggy Ozias-Akins, Dr. Benjamin Todd Campbell, Dr. Guy Collins and Dr. E. Charles Brummer for serving on my committee and for their critical advices during the course of my research.

I am also grateful to the past and present members of Cotton Molecular Breeding Lab. Especially Jennifer McCurdy and Dr. Edward Lubbers for their unequivocal support during planning and execution of field research projects. I would like to thanks members of National Environmentally Sound Production Agriculture Laboratory (NESPAL) for their love and friendship during ups and downs of my personal life. I would also like to extend my thanks to members of Plant Genome Mapping Laboratory (PGML) for their support in making my stay in Athens comfortable.

I thank my loving and supporting wife Rippy who stood next to me in both difficult and fun time during my journey through graduate research. I also thank the Almighty to have blessed us with ever loving and smiling son Aryan.

Finally, I'd like to extend my thanks to all of those that their names are not mentioned here but have helped and encouraged me through my study.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1

INTRODUCTION

Upland cotton (*Gossypium hirsutum*) is the most important natural fiber crop of the world accounting for over 90% of world's total cotton production. Cotton has been used to make clothes since ancient times; it was spun, woven and dyed by people of ancient India, Egypt and China. Today, cotton is commercially cultivated in more than 50 countries in drier tropical and sub-tropical environmental conditions (Smith and Cothren, 1999). The major share of global cotton production comes from countries such as USA, China, India, Pakistan and Australia where the climatic conditions like periods of hot and dry weather, photoperiod and adequate soil moisture favor natural growth requirements of cotton (Khadi et al., 2010). The primary objective of cotton cultivation is lint production, which is harvested as seed-cotton and later seeds are ginned out leaving behind fiber mass referred to as lint. Along with being an important source of lint fiber, cotton is also an important source of food and feed. Cottonseeds are a rich source of oil and this cottonseed-oil is of industrial value. The oil content of cottonseed is about 20%. Cottonseed meal serves as protein rich feed for livestock. Cotton hulls are used as fertilizer and fuel. Pressed paper and cardboard is manufactured from stalk fibers.

The US is the third largest producer of Upland cotton accounting for 16% of global cotton production and is preceded only by China and India (Figure 1). Cotton lint is a major source of revenue in the US; over 18 million bales of Upland cotton were

produced in 2010-11 which is estimated to have added over $6.8 billion to the US economy (National Agricultural Statistics Service, 2011).

There are about 50 species of cotton (Fryxell et al., 1992) most of the species 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 an allotetraploid Upland cotton (*G. hirsutum*) and about 8% come 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 that 2% towards total cotton production.

Gossypium gene pools

Gossypium species are classified into primary, secondary and tertiary gene pools based on the ease of gene exchange between different species. The concept of gene pools was proposed by Harlan and De Wet (1971) to emphasize genetic compatibility over formal taxonomy. The primary gene pool corresponds to species between which hybrids are fertile with normal chromosome pairing, gene segregation and gene transfer is simple. Therefore, for Upland cotton everything within species *G. hirsutum* is classified under the primary *Gossypium* gene pool (Lubbers and Chee, 2009). The secondary gene pool includes species which, following hybridization, may participate to a limited extent of mutual exchange of genes. Hybrids of *G. hirsutum* with all other tetraploid *Gossypium* species (*G. barbadense, G. mustelinum, G. tomentosum* and *G. darwinii*) are vigorous and fertile but show genetic breakdown during segregation in subsequent generations

(Hutchinson, 1951). Hence, all of these tetraploid *Gossypium* species are grouped into the secondary *Gossypium* gene pool (Lubbers and Chee, 2009). The tertiary gene pool includes species from which gene transfer requires extreme methods such as embryo rescue, tissue culture, chromosome doubling or the use of bridging species. Diploid *Gossypium* species and any distant groups are categorized in the tertiary *Gossypium* gene pool.

Upland cotton

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. It is sometimes referred to as New World cotton or short-staple cotton and has light green fuzzy seed to which lint is firmly attached. This species has broad adaptation and is grown on 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 cytoplasmic background (Wendel and Albert, 1992).

Types of Upland cotton in the US.

Nine types of Upland cotton were described at the beginning of 20th century. These groups were- Eastern Big Boll type, Western Big Boll type, Semi-cluster type, Cluster type, Rio Grande type, Early type, Long Limb type, Upland Long Staple type and the Intermediate or Miscellaneous type (Duggar, 1907; Tyler, 1910). These nine types of Upland cotton illustrated the phenotypic diversity available in the middle to late 1800s

prior to the era of the boll weevil. Many of these types were lost during the boll weevil era and are no longer available, marking a great loss in the history of cotton improvement.

A new grouping of different types of Upland cotton emerged in the post boll weevil era (late 1970's). Upland cotton is now grouped into four distinct types- Acala, Plains, Delta and Eastern (Niles and Feaster, 1984). The Acala type of Upland cottons is derived from introductions made into the US after the advent of the boll weevil in 1892. Production of Acala type is primarily concentrated in California, New Mexico, high altitudes of Arizona and western Texas, occupying about 10% of the U.S. production area. Among several Acala cultivars released, Acala SJ4 occupies an important place in history; released in 1976, Acala SJ-4 is a high quality cultivar with increased fiber and yarn strength and high yield on verticillium-wilt infested soils. The Delta types are predominately grown in the rain-belt area from southern Texas to Alabama and selective regions of Arizona and the Imperial Valley of California, covering approximately one third of the U.S. cotton production area. These were derived either completely or partially from the germplasm of the older pre-boll weevil cultivars with broad adaptability. The Deltapine and Stoneville series are prominent examples of Delta type cultivars. The Plains type of Upland cotton is exclusively grown in Texas, Oklahoma and eastern New Mexico occupying almost one half of the total U.S. cotton production area. Most of the cultivars of this type are characterized by compact plant habit, relatively determinant fruiting habit and storm resistant bolls. A prominent example of this type is the Paymaster family which was developed in the High Plains of Texas from an introduction made from Guatemala. The Eastern type cultivars represent the widest diversity of

Upland cotton germplasm. These types predominate cotton production in North Carolina, South Carolina and Georgia and were developed by hybridization of post boll weevil cultivars with pre boll weevil era stocks and introgression from Sea Island (Niles and Feaster, 1984). The most prominent examples are the Pee Dee germplasm lines.

Sea Island cotton

Sea Island (*G. barbadense*) cottons were introduced from West Indies to the coastal areas of South Carolina and Georgia in the late 18th century (Fryxell, 1965; Stephens, 1976). Racial relationships based on various genetic characteristics indicate that Sea Island cottons are most closely related to the primitive forms of *G. barbadense* in northeastern South America (Stephens, 1974). Sea Island cottons were distinguished by their exceptionally long and fine fiber and were commercially grown from the late 19th century until the first quarter of the 20th century. However, lower yields, exceedingly late maturity and limited adaptability led to serious decline in their cultivation after the 1920's. Today, Sea Island cottons residing in germplasm collections are often identified by their island of origin (St. Vincent, Barbados, St. Lucia, St. Kitts, Nevis, etc.) or by their fiber traits (St. Vincent Superfine, OSI Ordinary, Superfine V46, etc.) (Percy, 2009).

Sealand cotton

The primary sources of superior quality cotton that existed in the US during the early 20th century were the Sea Island (*G. barbadense*) varieties grown in northern Florida and the coastal plains of Georgia and South Carolina, extra-long staple upland (*G. hirsutum*) varieties grown in the Mississippi Delta and some south-eastern areas and

extra-long staple varieties of Egyptian (*G. barbadense*) cotton. However, Egyptian cotton varieties lacked uniformity and Sea Island varieties were late maturing, low yielding and highly susceptible to boll weevil, thereby hindering their large scale commercial cultivation.

Extra-long staple varieties of upland cotton were high yielding and early maturing but the fiber quality was not on par with *G. barbadense* varieties. Therefore, to improve fiber quality of Upland cotton varieties a massive breeding program of interspecific crossing and backcrossing was initiated at the Pee Dee Experiment Station, Florence, South Carolina in the mid 1930's. This program developed several Upland cotton germplasm lines with improved fiber characteristics, presumably from introgressions from *G. barbadense* parents. One such germplasm line was the "Sealand line". In the mid 1940's, Sealand cultivars such as Sealand-542 became popular among cotton growers and were planted in over one thousand acres in coastal plains of the Carolinas and Georgia. Commercial cultivation of Sealand cultivars was abandoned in the early 1950's due to their increased susceptibility to boll weevil and arrival of several improved Pima cottons (Jenkins, 1948).

Fiber quality

The market-value of raw cotton is judged 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 which the textile industry utilizes to predict a fiber sample include fiber length, fiber strength, fiber elongation, fiber fineness and uniformity index.

Improving yielding potential, along with widening the adaptability of Upland cotton has remained the primary breeding goal of cotton breeders in the post boll-weevil era. However, in the last few years, there has been a shift in focus to include improving fiber quality parameters of Upland cotton. There are two driving forces behind this shift: Foremost is technological breakthroughs in textile industry. The older ring spinning machines have now been replaced by high speed open-end and air jet spinning machines (Felker, 2001) , which are up to eight times faster than their older counterparts, giving the textile mills a tremendous boost in productivity and thereby in profit margins. The spinning efficiency of these new machines can only be sustained with the use of high quality cotton fibers. The second reason for focusing on fiber quality is the transformation of the U.S. cotton industry, from a domestic market to an export oriented market. Until the late 1990's local cotton mills accounted for more than 70% of the total demand for raw fiber; however in recent years, local demand has sharply declined as more and more textile mills moved overseas due to labor costs. As a result, more than 80% of raw cotton produced in the U.S. is now being exported (Foreign Agricultural Services, 2011). International consumers more greatly emphasize fiber quality than the previously established standards for the domestic market and until recently a majority of the Upland cotton cultivars grown in the U.S. barely meet these quality requirements. Therefore, improving fiber quality traits of Upland cotton has become a priority for cotton breeders.

Fiber quality traits and their measurements

Since 1980, USDA-AMS classing offices have been using the High Volume Instrumentation (HVI) for measuring fiber properties (Moore, 1996). The standard HVI measurements currently consist of fiber length, length uniformity, strength, micronaire, and color, while subjective determinations of leaf grade, preparation, and extraneous matter are recorded by trained classing personnel.

Fiber length

HVI determines fiber length by scanning a test beard that is prepared by a comb or a brush attachment. These readings are reported in percentage of the total number of fibers present at each length value and into other length parameters, such as mean length and upper-half mean length (Behery, 1993). The most common reporting of fiber length in the literature is in units of Upper Half Mean (UHM) length, which is the average length of the longer one-half of the fibers reported in inches. Cotton Incorporated designates Upland cotton with an UHM of 1.11 to 1.26 inches as 'Long' and fibers with UHM of above 1.26 inches as 'Extra Long' (Cotton Incorporated and Textile World, 2003). Longer fibers tend to produce stronger yarn, resulting in improved overall appearance of the fabric.

Fiber strength

Fiber strength in HVI is measured by breaking the fiber bundle (beards) formed during length measurements and is reported in grams per tex. It is the amount of force in grams required to break one tex unit of fiber (one tex unit is the weight in grams of 1000 meters of fiber). The HVI strength measurements for Upland cotton range from below 20 grams per tex (very weak) to above 32 grams per tex (very strong) (Cotton Incorporated

and Textile World, 2003). Fiber strength describes the sturdiness of fiber; stronger fibers can effectively withstand the mechanical impacts in the course of yarn production, thereby drastically improving the efficiency of the whole fabric manufacturing process.

Fiber elongation

Fiber bundle elongation is measured directly from the displacement of the jaws during the bundle breaking process and is usually reported along with fiber strength (Bradow and Davidonis, 2000). It measures the elasticity of cotton fibers before the break occurs and is expressed in percentage. Fiber breakage causes inefficiency in yarn manufacturing performance and decreased end product quality (Chee and Campbell, 2008). Yarns spun out of fibers with higher elongation percentage tend to withstand the weaving process of fabric manufacture more efficiently (May, 2000).

Micronaire

Micronaire is a measure of fiber fineness and maturity, and is determined by the amount of resistance offered by a unit bundle of fibers to a constant airflow (Jhonson, 1951). The acceptable micronaire range for Upland cotton without any price penalty is 3.5 to 4.9, with a premium range of 3.7 to 4.2 (Bradow and Davidonis, 2000). High micronaire fibers are thick and mature and often do not spin efficiently resulting in coarser yarn. Conversely, fibers with low micronaire value are fine but immature and can cause neps and dye defects (Hake et al., 1996). Therefore, micronaire values are often used in combination with other fiber quality parameters such as fiber length and strength before fibers are processed.

Uniformity index

HVI fiber length uniformity is calculated from a ratio of the HVI mean length and upper half mean length of the fibers, and is expressed as a percentage. It is an important quality parameter for the textile industry because it indirectly indicates the short fiber content in the sample. If all fibers in the sample are of the same length, then the mean length and the upper half mean length will be the same, and the uniformity index therefore will be 100. Fiber sample with UI value of 85 represents a very high degree of uniformity whereas UI value less than 77 represents very low degree of uniformity or in other words high short fiber content (Cotton Incorporated and Textile World, 2003).

References

Bajaj, Y.P.S. (1998) Cotton Springer, Berlin ; New York.

Behery, H.M. 1993. Short fiber content and uniformity index in cotton. ICAC Review Articles on Cotton Production Research (United Kingdom).

Bradow, J.M., G.H. Davidonis. 2000. Quantitation of fiber quality and the cotton production-processing interface: a physiologist's perspective [electronic resource]. Journal of cotton science (Online) Journal of cotton science 4:34-64.

Chee, P., B. Campbell. 2008. Bridging Classical and Molecular Genetics of Cotton Fiber Quality and Development. Genetics and Genomics of Cotton:283.

Cotton Incorporated, Textile World. (2003) U.S. cotton fiber chart for 2001, Available at http://www.cottoninc.com/ClassificationofCotton, Cotton Incorporated, Cary, NC, and Textile World, Atlanta, GA.

Duggar, J. 1907. Description and classification of varieties of American Upland Cotton. Ala. Agr. Exp. Sta. Bull 140.

Felker, G.S. (2001) Fiber Quality and New Spinning Technology. pp. 5-7.

Fryxell, P.A. 1965. Need for Taxonomic Research in Gossypium. Am. J. Bot. 52:650-&.

Fryxell, P.A., L.A. Craven, J.M. Stewart. 1992. A Revision of Gossypium Sect Grandicalyx (Malvaceae), Including the Description of 6 New Species. Syst. Bot. 17:91-114.

Hake, K.D., D.M. Bassett, T.A. Kerby, W.D. Mayfield. 1996. Producing quality cotton. Cotton production manual. Publication 3352:134-149.

Harlan, J., J. De Wet. 1971. Toward a rational classification of cultivated plants. Taxon 20:509-517.

Hutchinson, J. 1951. Intra-specific differentiation in Gossypium hirsutum. Heredity 5:161-193.

Jenkins, J.G. 1948. Sea Island cotton breeding. 28th Ann. Rep. Ga Coastal Plains Exp Stn Bull 16:1.

Jhonson, B. 1951. Use and application of Fibre spinning Tests. Natl. Cotton Counc. Am., Memphis, TN 1:51.

Khadi, B., V. Santhy, M. Yadav. 2010. Cotton: An Introduction. Cotton:1-14.

Lubbers, E.L., P.W. Chee. 2009 The Worldwide Gene Pool of G. hirsutum and its Improvement, *In* A. H. Paterson (Ed.), Genetics and Genomics of Cotton. pp. 1-30.

May, O.L. 2000. Genetic variation in fiber quality. Cotton fibers–developmental biology, quality improvement, and textile processing. Food Products Press, New York:183–229.

Moore, J. 1996. Cotton classification and quality. The cotton industry in the United States. ASA, CSSA, and SSSA, Madison, WI:51-57.

Niles, G.A., C.V. Feaster. 1984 Breeding, *In* R. J. Kohel and C. F. Lewis (Eds.), Cotton, Madison, WI. pp. 202-229.

Percy, R.G. 2009 The Worldwide Gene Pool of Gossypium barbadense L. and Its Improvement, *In* A. H. Paterson (Ed.), Genetics and Genomics of Cotton. pp. 1-16.

Smith, C.W., J.T. Cothren. (1999) Cotton: origin, history, technology and production John Wiley & Sons.

Stephens, S.G. 1974. The use of two polymorphic systems, nectary fringe hairs and corky alleles, as indicators of phylogenetic relationships in New World cottons. Biotropica:194-201.

Stephens, S.G. 1976. The origin of Sea Island cotton. Agricultural history:391-399.

Tyler, F.J. (1910) Varieties of American upland cotton (USDA-ARS, BPI, Bulletin no. 163).

Wendel, J.F., 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. Syst. Bot. 17:115-143.

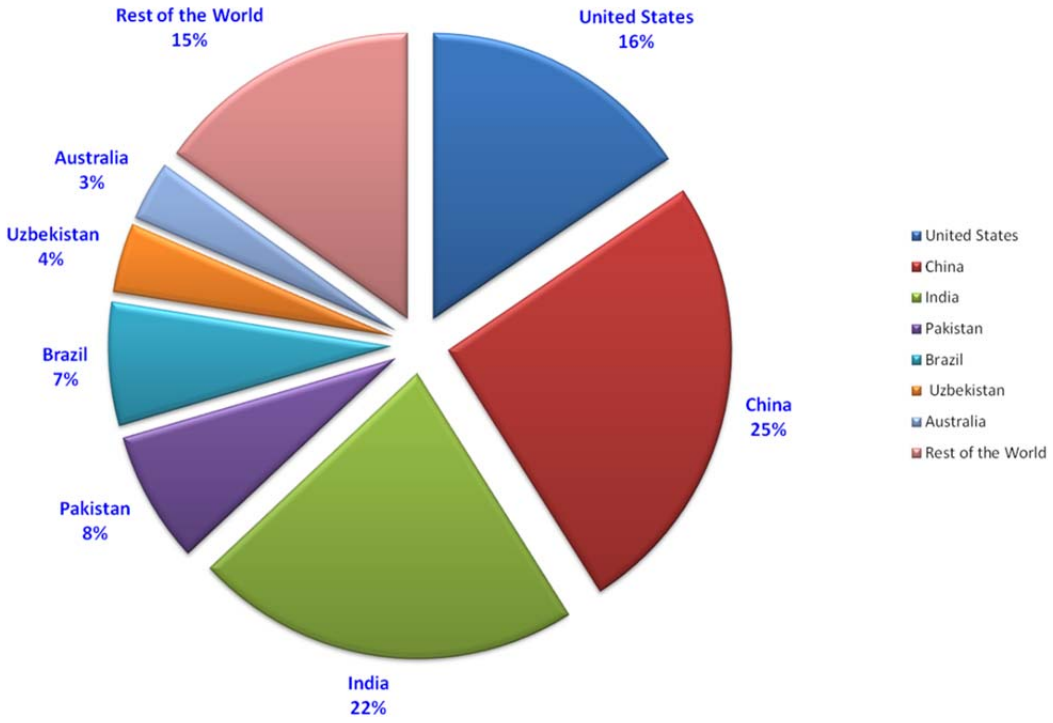


Figure 1.1 Upland cotton production by major countries as percentage of the total world cotton production in 2010-11.

CHAPTER 2

REVIEW OF LITERATURE

Origin of cotton

The genus *Gossypium* consists of more than 50 species (Fryxell et al., 1992) and is the largest and most widely distributed of the eight genera in tribe *Gossypieae*. Phylogenetic analysis indicates that *Gossypium* diverged from its sister group of *Gossypioides* and *Kokia* 12.5 million years ago (Cronn et al., 2003; Wendel and Cronn, 2003). A wide range of genomic diversity exists within the genus *Gossypium*. There are eight diploid genomic groups (A-G & K) within this genus (Endrizzi et al., 1985; Stewart, 1995). Genomes A, B, E, and F are considered to be of African or Asian origin, C and G genomes are of Australian origin and the D genome species originated in the western hemisphere. Diploid species of Asian or African origin are often referred to as Old World cotton species whereas diploid D genome and allotetraploid species with AADD genome are referred as New World cottons (Endrizzi et al., 1985). Allopolyploid *Gossypium* originated prior to the evolution of modern humans but relatively recently in geological terms, perhaps 1-2 mya (Wendel and Cronn, 2003) by the fusion of A and D genomes in an A genome cytoplasmic background.

Molecular markers and cotton linkage maps

Gossypium hirsutum is an allopolyploid with a genome size of 2.3 Gb (Arumuganathan and Earle, 1991) and to map all the recombinational events in such a large genome would require a large number of evenly spaced molecular markers. The first molecular map of cotton was published in 1994 (Reinisch et al., 1994) developed by RFLP markers on an interspecific mapping population derived from a cross between *G. hirsutum* race Palmeri and *G. barbadense* acc. K101. Markers were assorted in 41 linkage groups covering 4,675 cM of the cotton genome. Various linkage groups were associated with chromosomes using previously developed series of interspecific monosomic substitution stocks (Stelly, 1993). In an another study a genetic linkage joinmap was constructed from four different mapping populations (Ulloa et al., 2002). This map covered 1,502 cM or approximately 31% of the total recombinational length of the cotton genome and consist of 284 RFLP loci that sorted onto 47 linkage groups with average distance of 5.3 cM between markers. In order to better understand evolutionary differentiation of diploid and tetraploid genomes of cotton, comparative RFLP mapping was employed (Brubaker et al., 1999) to compare genetic maps for the allotetraploid AD genome (n=26) and diploid A&D genomes (n=13). This comparison facilitated synteny, chromosomal rearrangement and gene order studies among the thirteen homoeologous linkage groups. It was found that allotetraploid genomes and diploid genomes were recombinationally equivalent even though there exists a two-fold difference in their physical size (Brubaker et al., 1999). A detailed recombination map of sequence-tagged sites was developed using 2669 probes which mapped 3347 loci with average distance of less that 2cM between markers (Rong et al., 2004). This was the first detailed cotton map

that combined the markers into the expected 26 linkage groups. The map gave better insight into genome organization, transmission and evolution of the cotton genome. No major structural changes between tetraploid and diploid D-genomes were detected, however several reciprocal translocations and inversions were found in the tetraploid A-genome. This map provided a framework for an integrated high resolution map of cotton genome.

The RFLP marker system suffers from limitations such as requiring large amounts of high quality DNA, not being very cost effective, labor intensive and not amenable to automation. To overcome these limitations, Polymerase Chain Reaction (PCR) based marker systems for cotton were developed. Microsatellite markers, also known as Simple Sequence Repeats (SSR), are tandem repeat elements that are abundant in plants and animals. Polymorphism in SSRs often arises either due to unequal crossing over or DNA polymerase slippage during replication which results in either increase or decrease in copy numbers (Yu et al., 1999).

SSR markers are becoming increasingly popular as these are PCR based, follow co-dominant inheritance and are relatively abundant in plant genomes. It is estimated that on an average there is one microsatellite every 170kb DNA of the cotton genome (Zhao et al., 1994). A multi-institutional program to capture and characterize SSRs was undertaken (Reddy et al., 2001). Both inter- and intra- specific polymorphism was detected by 150 loci which were amplified by primers designed from fragments obtained using biotin capture method. This study not only revealed the abundance of microsatellites in cotton but also the possible exploitation of SSR markers in genome mapping studies. A molecular linkage map developed from an interspecific haploid

population using SSRs and RAPD markers revealed possible homoeologous association between several cotton chromosomes (Zhang et al., 2002). This map covered 3314.5 cM with 489 mapped loci arranged in 43 linkage groups. Several attempts were made to take advantage of well-established RFLP markers in conjunction with a newly developed PCR based marker system. A combined RFLP-SSR-AFLP linkage map from an interspecific cross was developed (Lacape et al., 2003) which consisted of 888 mapped loci ordered in 37 linkage groups covering 4400 cM of the cotton genome. A set of unique genomic SSRs predominately containing CA/CT repeats were developed and mapped at CIRAD, France (Nguyen et al., 2004). Other markers such as Sequence Related Amplified Polymorphism (SRAP) (Lin et al., 2005) and Retrotransposon-Microsatellite Amplified Polymorphism (REMAP) markers (He et al., 2007) were also employed along with SSR markers to construct dense molecular maps of cotton genome with evenly spaced markers.

EST derived markers are valuable tools in understanding structural and functional genomics of cotton. EST-SSRs are generally conserved across *G. barbadense* and *G hirsutum*, therefore can also be used as markers for comparative mapping and evolutionary studies of transcribed genes (Saha et al., 2003). PCR markers based on ESTs can reveal size or sequence polymorphisms which can then be mapped (Chee et al., 2004). Polymorphisms in EST are often due to genetic variations in introns, which accumulate four-fold more variation than exons (Kumar et al., 2006). EST derived SSR primers were able to detect polymorphism within *G. hirusutm* and between *G. hirsutum* and *G. barbadense* genotypes (Qureshi et al., 2004). Many microsatellite markers developed from ESTs transcribed during fiber elongation in the A-genome species *G.*

arboreum (Han et al., 2004; Han et al., 2006) were successfully mapped using interspecific *Gossypium* crosses. Markers targeting complex sequence repeats in ESTs (Park et al., 2005) and markers developed from BAC-end genomic sequences from *G. hirsutum* Acala 'Maxxa' (Frelichowski et al., 2006) were employed to construct an updated linkage map of cotton. SSR markers from ESTs involved in various pathways such as lignin metabolism, starch biosynthesis and stress response have also been developed (Taliercio et al., 2006). Mapping these SSRs will essentially map the gene represented by the ESTs and will be a valuable resource for cotton improvement.

In many plants, Single Nucleotide Polymorphism (SNP) markers are increasingly becoming the marker system of choice. SNPs are the most abundant source of polymorphism and are present in both coding and non-coding regions of the genome. In cotton, the frequency of SNPs was found to be different between the genomes of two cultivated species, *G. hirsutum* and *G. barbadense* with Sea Island cotton (*G. barbadense*) having greater occurrence of SNPs (Rahman et al., 2009). Moreover the type and distribution of SNPs varied between and within genes of both species, with insertions-deletions (indels) occurring more frequently than base substitutions (Rahman et al., 2009). Gene-specific SNP markers were used for phylogenomic characterization and assignment of chromosomal locations of six *EXPANSIN A* genes (An et al., 2007) and six *MYB* genes (An et al., 2008) and the authors suggested that SNP markers may be useful in exploring the roles of candidate genes in the expression of complex traits. SNP markers were successfully employed in characterization of two genome specific transcription factor genes *GhMyb8* and *GhMyb10* in Upland cotton (Hsu et al., 2008), however the exact chromosomal location of the two genes couldn't be determined using

the SNP deletion method due to the incomplete coverage of cytogenetic stocks. Nucleotide diversity in cotton was explored using over 1000 SNPs from 270 loci and 279 indels from 92 loci segregating in *G. hirsutum* and *G. barbadense* which were genotyped across a standard panel of elite cotton breeding lines and mapping parents of populations from six cotton species (Van Deynze et al., 2009). A highlight of this study is the finding that the A and D genomes in both diploid and tetraploid cotton remain distinct from each such that paralogs can be distinguished using SNP markers. Recently, a high density cotton genetic map was developed using SSR markers and the first public set of single nucleotide polymorphism (SNP) markers (John et al., 2012). This genetic map comprised 2072 loci (1825 SSRs and 247 SNPs) and covered 3380 cM of the allotetraploid cotton genome with an average marker interval of 1.63 cM.

Although the availability of densely saturated maps has greatly facilitated precise QTL mapping however genetic linkage maps developed using different bi-parental combinations may differ due to individual specific chromosomal aberrations and rearrangements which can obscure the interpretation of individual map. Combining various linkage maps available into a consensus map was needed. In an early attempt a consensus map was constructed by aligning At-subgenome, Dt-subgenome and D-genome with the help of common shared molecular markers (Rong et al., 2005). A 'pooled map' was created (Ulloa et al., 2005) by combining marker data from four intraspecific mapping populations using JoinMap software. A unified consensus linkage map was constructed (Lacape et al., 2009) from 1,745 unique and shared loci between previously developed BC1 and RIL maps. This consensus map spanned a total distance of 3,637 cM with an average distance of 2.1 cM between two loci. A comprehensive

reference map of tetraploid cotton genome is constructed (Yu et al., 2010) by combining mapping information from 28 individual maps and contained 7,424 markers. This comprehensive reference map emphasized on marker order rather than marker distances estimated in individual maps.

QTL mapping in cotton

Untill the last decade of the 20th century, geneticists relied mostly on morphological markers. About 145 mutant markers have been identified in or transferred to cultivated cotton (Percy and Kohel, 1999). Exploitation of these markers for cotton improvement is restricted by their phenotypic effects and difficulty in congregating multiple markers in single plant types (Percy and Kohel, 1999). Since many of the economically important traits such as yield, yield components and fiber characteristics are quantitatively inherited, there was a need to develop a dense molecular map which can guide breeders to identify chromosomal regions contributing towards the phenotype.

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 Trail 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.

The first published molecular map of cotton was from an interspecific cross between. *G. hirsutum* race palmeri and *G. barbadense* acc. K101 (Reinisch et al., 1994). No analysis for agronomic or fiber traits was performed on this population. 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 et al., 1998) and also found that the most QTLs influencing fiber quality and yield were located on the "D" subgenome, 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 et al., 1998). 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. RFLP markers were used to map and characterize QTLs determining cotton leaf morphology and several other traits (Jiang et al., 2000). Sixty-two QTLs were identified for 14 morphological traits in a F_2 mapping population developed from an interspecific cross between a *G. hirsutum* parent carrying four morphological mutants, and *G. barbadense*. A majority of the QTLs were mapped on the D-genome of cotton. Twenty-six QTLs for various agronomic and fiber quality traits were identified on nine linkage groups in an intraspecific cross of *G. hirsutum* (Ulloa and Meredith, 2000). The phenotypic variation explained by these QTLs ranged from 3.4-44.6%. The parent 'Prema" contributed toward long, strong, and fine fiber whereas parent 'MD5678ne' contributed for high yield but short, coarse, and weak fibers. A single major QTL for fiber strength originating from a *G. anomalum* introgression line 'Suyuan 7235' was detected on chromosome 10 (Zhang et al., 2003) which could explain more than 30% of the phenotypic variation. Pedigree records indicated that a possible origin of this major QTL might be from Acala 3080 cotton which was one of the parental lines in developing Suyuan 7235. Using an interspecific map developed from RFLP, SSR and AFLP

markers, seven QTLs for various fiber related traits could be identified (Mei et al., 2004). Interestingly five of these seven QTLs were contributed by the A-genome and were present in clusters, indicating that some chromosomes contribute more toward phenotypic variation in fiber traits than others.

In order to study and explore the secondary gene pool such as that of *G. barbadense* a detailed analysis of interspecific phenotypic variation for fiber elongation (Chee et al., 2005a), fiber fineness (Draye et al., 2005), and fiber length (Chee et al., 2005b), was undertaken. For this study an advanced generation backcross population from a cross between *G. hirsutum* cv. Tamcot2111 and *G. barbadense* cv. Pima-S6 was developed. A comprehensive analysis of this population using RFLP markers detected 22 QTLs for fiber elongation (Chee et al., 2005a), 32 and 9 QTLs for fiber fineness and micronaire (Draye et al., 2005), 28, 9 and 8 QTLs for fiber length, uniformity and short fiber content, respectively (Chee et al., 2005b). The phenotypic variation explained by these QTLs ranged from 8 to 28%.

Knowledge of stable QTLs or QTL rich genomic regions consistently identifiable across mapping populations and over generation is highly valuable for cotton breeders targeting fiber quality improvement. Blocks of gene rich regions are observed in the genomes of allotetraploid cotton. For example, nineteen QTL rich regions across fifteen chromosomes were identified (Lacape et al., 2005) which were mapped on a RFLP-SSR-AFLP genetic map of *G. hirsutum* by *G. barbadense* (Lacape et al., 2003). Simultaneous occurrence or co-localization of QTLs for fiber quality in these QTL rich regions leads to the observed phenotypic correlations (Lacape et al., 2005; Saranga et al., 2002; Zhang et al., 2011). Three of the seventeen fiber quality QTLs showed stable expression and were

detected across three different mapping populations (Shen et al., 2005) suggesting the existence of elite fiber quality genes, probably of similar origin. Similarly, three stable QTLs for boll size, two for lint percentage and one for boll number per plant were detected above threshold on chromosome D8 in two populations at multiple locations over two years (Chen et al., 2010).

Availability of dense linkage maps facilitated identification of new QTLs affecting fiber quality and yield. For example QTL analysis in the F_2 generation using a dense interspecific genetic map, developed from 834 SSR, 437 SRAP, 107 RAPD, and 16 REMAP markers (He et al., 2007) covering 5472.3 cM of cotton genome, identified as many as fifty-two QTLs. These included QTLs for seed index, lint yield, seed cotton yield, number of seeds per boll, fiber strength, fiber length, and micronaire values. Genetic analysis of fiber quality traits using this map detected up to fifteen markers for fiber strength, five significant loci for fiber length, four for uniformity, two for micronaire and as many as ten loci for fiber elongation in an advance generation interspecific population (He et al., 2008).

Limited polymorphism within Upland cotton restricts the use of linkage maps developed from intraspecific crosses. Therefore to achieve greater level of intraspecific polymorphism, a four way cross (4WC) population was developed (Qin et al., 2008). QTL mapping using linkage map constructed from 4WC population detected 31 QTLs for yield and fiber quality traits that were able to account for 5.1-25.8% of the total phenotypic variation. An intraspecific map developed from TM-1 by Suyuan 7235 cross was instrumental in identifying major fiber strength QTL (Shen et al., 2007; Shen et al., 2006; Zhang et al., 2003) which accounted for up to 16% of the total phenotypic variation

and was consistently detected in multiple environments. This QTL region was further dissected (Chen et al., 2009) by using three overlapped RILs. QTL analysis of these 3 RILs over 2 environment and over 2 years revealed 5 clustered QTLs (2.5<LOD<29.8). This cluster accounted for 28.8-59.6% of the total phenotypic variation for fiber strength. The effect of this QTL was validated in two Upland cotton genetic backgrounds (Kumar et al., 2012) where the allele from Suyuan 7235 accounted for over 40% of the total phenotypic variation.

Alleles from wild cotton species

Broadening the genetic base is essential for genetic improvement of lint yield and fiber quality in upland cotton (*Gossypium hirsutum* L.). Novel alleles from exotic germplasm can be a major source for broadening the genetic base of Upland cotton. For evaluation and exploitation of alleles from other *Gossypium* species a species polycross (SP) population was developed (Zeng et al., 2007) by crossing cultivars and strains of Upland cotton (*G. hirsutum*) with the four other tetraploid species, *G. barbadense* L., *G. tomentosum* Nutt., *G. mustelinum* Watt., and *G. darwinii* Watt. and allowing 11 generations of random mating and selfing. Association analysis using this population (Zeng et al., 2009) revealed 59 significant marker trait associations with six fiber quality traits. Population structure and substructure are often observed in association analysis and can significantly influence result and interpretation of association analysis. In this study 39 of the 59 marker-trait associations remained significant after correction by accounting for population substructure and relatedness among individual lines. Boll weight was most

significantly influenced by population structure whereas micronaire and elongation registered least influence.

Fiber quality alleles from wild Hawaiian cotton (*G. tomentosum*) were evaluated (Zhang 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. These findings hold promise of rapid genetic gains by increasing diversity in Upland cotton by introgressing such alleles from wild cotton. The authors also suggested that alleles from *G. tomentosum* might be best exploited by use in F1 hybrid cottons in order to mitigate effects of linkage drag often associated with introgression breeding. Potential exploitation of alleles from another wild cotton (*Gossypium mustelinum*) is being evaluated (Dantas et al., 2012). Together the wild cottons offer a source of novel fiber quality and yield alleles that can be deployed to improve yield and quality of Upland cotton.

QTLs by genetic background interactions

Interaction with genetic backgrounds can significantly alter the expression of QTLs. Numerous studies have shown interactions between QTL and genetic backgrounds (Lecomte et al., 2004a; Li et al., 2009; Sebolt et al., 2000) and suggested the need to test the efficacy of a QTL in multiple genetic backgrounds before its utilization in MAS. For example, Lecomte et al. (2004) introgressed five QTL controlling fruit quality into three tomato lines and found that the breeding efficiency of each QTL varied according to the recipient parent. Sebolt et al. (2000) introgressed a QTL conferring high seed protein concentration into three soybean lines with varying levels of protein content and detected

the effects in only two of the three genetic backgrounds; the effect of this QTL was not detected in the line having the highest seed protein concentration. The effect of dent maize genetic background on grain yield QTLs was found significant where only one QTL of the total of 33 identified was found common between two mapping population developed by crossing a single high-oil maize line with two dent maize inbred lines (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 is greater than the environment effect (Liao 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 the main-effect QTLs.

In Upland cotton, significant interaction between QTL and genetic backgrounds or family interactions has been observed (Chee et al., 2005b; Draye et al., 2005). A genetic effect of an introgressed *G. barbadense* allele for fiber elongation at a locus pAR137a was altered in opposite directions when present in different families. Similarly at least 11 loci for fiber fineness (Draye et al., 2005) and 19 loci affecting fiber length components (Chee et al., 2005b) 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 (Zhang et al., 2007).

References

An, C., S. Saha, J.N. Jenkins, D.P. Ma, B.E. Scheffleer, R.J. Kohel, J.Z. Yu, D.M. Stelly. 2008. Cotton (Gossypium spp.) R2R3-MYB transcription factors SNP identification, phylogenomic characterization, chromosome localization, and linkage mapping. Theor. Appl. Genet. 116:1015-1026.

An, C.F., S. Saha, J.N. Jenkins, B.E. Scheffler, T.A. Wilkins, D.M. Stelly. 2007. Transcriptome profiling, sequence characterization, and SNP-based chromosomal assignment of the EXPANSIN genes in cotton. Mol. Genet. Genomics 278:539-553.

Arumuganathan, K., E. Earle. 1991. Nuclear DNA content of some important plant species. Plant molecular biology reporter 9:208-218.

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

Chee, P., X. Draye, C.X. Jiang, L. Decanini, T.A. Delmonte, R. Bredhauer, C.W. Smith, A.H. Paterson. 2005a. Molecular dissection of interspecific variation between Gossypium hirsutum and Gossypium barbadense (cotton) by a backcross-self approach. I. Fiber elongation. Theoretical and applied genetics 111:757-763.

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

Chee, P.W., J. Rong, D. Williams-Coplin, S.R. Schulze, A.H. Paterson. 2004. EST derived PCR-based markers for functional gene homologues in cotton. Genome 47:449-462.

Chen, H., N. Qian, W.Z. Guo, Q.P. Song, B.C. Li, F.J. Deng, C.G. Dong, T.Z. Zhang. 2009. Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro.D8 in Upland cotton. Theor. Appl. Genet. 119:605-612.

Chen, H., N. Qian, W.Z. Guo, Q.P. Song, B.C. Li, F.J. Deng, C.G. Dong, T.Z. Zhang. 2010. Using three selected overlapping RILs to fine-map the yield component QTL on Chro.D8 in Upland cotton. Euphytica 176:321-329.

Cronn, R., R.L. Small, T. Haselkorn, J.F. Wendel. 2003. Cryptic repeated genomic recombination during speciation in Gossypium gossypioides. Evolution 57:2475-2489.

Dantas, A.C.A., P.a.V. Barroso, L.V. Hoffmann, M.F. Alves, F.P. Andrade. 2012. SSR markers to detect gene flow from upland to mocó cotton. Rev Cienc Agron 43:163-169.

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

Endrizzi, J.E., E.L. Turcotte, R.J. Kohel. 1985. Genetics, cytology, and evolution of Gossypium. Adv. Genet.:271-375.

Frelichowski, J.E., M.B. Palmer, D. Main, J.P. Tomkins, R.G. Cantrell, D.M. Stelly, J. Yu, R.J. Kohel, M. Ulloa. 2006. Cotton genome mapping with new microsatellites from Acala 'Maxxa' BAC-ends. Mol. Genet. Genomics 275:479-491.

Fryxell, P.A., L.A. Craven, J.M. Stewart. 1992. A Revision of Gossypium Sect Grandicalyx (Malvaceae), Including the Description of 6 New Species. Syst. Bot. 17:91-114.

Han, Z., C. Wang, X. Song, W. Guo, J. Gou, C. Li, X. Chen, T. Zhang. 2006. Characteristics, development and mapping of Gossypium hirsutum derived EST-SSRs in allotetraploid cotton [electronic resource]. Theoretical and applied genetics 112:430-439.

Han, Z.G., W.Z. Guo, X.L. Song, T.Z. Zhang. 2004. Genetic mapping of EST-derived microsatellites from the diploid Gossypium arboreum in allotetraploid cotton. Molecular genetics and genomics MGG 272:308-327.

He, D.H., Z.X. Lin, X.L. Zhang, Y.C. Nie, X.P. Guo, Y.X. Zhang, W. Li. 2007. QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of Gossypium hirsutum x Gossypium barbadense. Euphytica 153:181-197.

He, D.H., Z.X. Lin, X.L. Zhang, Y.X. Zhang, W. L.I, Y.C. Nie, X.P. Guo. 2008. Dissection of genetic variance of fibre quality in advanced generations from an interspecific cross of Gossypium hirsutum and G. barbadense [electronic resource]. Plant Breeding 127:286-294.

Hsu, C.Y., C. An, S. Saha, D.P. Ma, J.N. Jenkins, B. Scheffler, D.M. Stelly. 2008. Molecular and SNP characterization of two genome specific transcription factor genes GhMyb8 and GhMyb10 in cotton species. Euphytica 159:259-273.

Jenkins, J.N., S. Saha. (2001) Genetic improvement of cotton : emerging technologies Science Publishers, Enfield, NH.

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

John, Z.Y., R.J. Kohel, D.D. Fang, J. Cho, A. Van Deynze, M. Ulloa, S.M. Hoffman, A.E. Pepper, D.M. Stelly, J.N. Jenkins. 2012. A High-Density Simple Sequence Repeat and Single Nucleotide Polymorphism Genetic Map of the Tetraploid Cotton Genome. G3: Genes| Genomes| Genetics 2:43-58.

Kumar, P., A.H. Paterson, P.W. Chee. 2006. Predicting Intron Sites by Aligning Cotton ESTs with Arabidopsis Genomic DNA [electronic resource]. Journal of cotton science 10.

Kumar, P., R. Singh, E.L. Lubbers, X. Shen, A.H. Paterson, B.T. Campbell, D.C. Jones, P.W. Chee. 2012. Mapping and Validation of Fiber Strength Quantitative Trait Loci on Chromosome 24 in Upland Cotton. Crop Sci. 52:1115-1122.

Lacape, J.M., J. Jacobs, T. Arioli, R. Derijcker, N. Forestier-Chiron, D. Llewellyn, J. Jean, E. Thomas, C. Viot. 2009. A new interspecific, Gossypium hirsutum x G. barbadense, RIL population: towards a unified consensus linkage map of tetraploid cotton. Theor. Appl. Genet. 119:281-292.

Lacape, J.M., T.B. Nguyen, B. Courtois, J.L. Belot, M. Giband, J.P. Gourlot, G. Gawryziak, S. Roques, B. Hau. 2005. QTL analysis of cotton fiber quality using multiple Gossypium hirsutum x Gossypium barbadense backcross generations. Crop Sci. 45:123-140.

Lacape, J.M., T.B. Nguyen, S. Thibivilliers, B. Bojinov, B. Courtois, R.G. Cantrell, B. Burr, B. Hau. 2003. A combined RFLP-SSR-AFLP map of tetraploid cotton based on a Gossypium hirsutum x Gossypium barbadense backcross population. Genome 46:612-626.

Lecomte, L., P. Duffe, M. Buret, B. Servin, M. Causse. 2004. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theor. Appl. Genet. 109:658-668.

Li, Y., X. Li, J. Li, J. Fu, Y. Wang, M. Wei. 2009. Dent corn genetic background influences QTL detection for grain yield and yield components in high-oil maize. Euphytica 169:273-284.

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

Lin, Z., D. He, X. Zhang, Y. Nie, X. Guo, C. Feng, J.M. Stewart. 2005. Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. Plant breeding = Zeitschrift fur Pflanzenzuchtung 124:180-187.

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

Nguyen, T.B., M. Giband, P. Brottier, A.M. Risterucci, J.M. Lacape. 2004. Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. Theoretical and applied genetics 109:167-175.

Park, Y.H., M.S. Alabady, M. Ulloa, B. Sickler, T.A. Wilkins, J. Yu, D.M. Stelly, R.J. Kohel, O.M. El-Shihy, R.G. Cantrell. 2005. Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. Molecular genetics and genomics MGG 274:428-441.

Percy, R., R. Kohel. 1999 Cotton Qualitative genetics *In* C. W. Smith and J. T. Cothren (Eds.), Cotton: Origin, history, technology, and production, John Wiley and Sons, New York. pp. 319-360.

Qin, H., T. Zhang, Y.-M. Zhang, W. Guo. 2008. QTL mapping of yield and fiber traits based on a four-way cross population in Gossypium hirsutum L [electronic resource]. Theoretical and applied genetics TAG 117:883-894.

Qureshi, S.N., J.N. Jenkins, R.V. Kantety, S. Saha. 2004. EST-SSR: a new class of genetic markers in cotton [electronic resource]. Journal of cotton science 8.

Rahman, M., Y. Zafar, A.H. Paterson. (2009) Gossypium DNA markers: types, numbers and uses, Genetics and Genomics of Cotton, New York, Springer. pp. 1-39.

Reddy, O.U.K., T. Brooks, Y. Bolek, K.M. El-Zik, J.N. Jenkins, A.E. Pepper, I. Abdurakhmonov, S. Saha. 2001. New dinucleotide and trinucleotide

microsatellite marker resources for cotton genome research [electronic resource]. Journal of cotton science (Online) Journal of cotton science 5:103-113.

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

Rong, J., C. Abbey, J.E. Bowers, C.L. Brubaker, C. Chang, P.W. Chee, T.A. Delmonte, X. Ding, J.J. Garza, B.S. Marler, C.H. Park, G.J. Pierce, K.M. Rainey, V.K. Rastogi, S.R. Schulze, N.L. Trolinder, J.F. Wendel, T.A. Wilkins, T.D. Williams-Coplin, R.A. Wing, R.J. Wright, X. Zhao, L. Zhu, A.H. Paterson. 2004. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton (Gossypium). Genetics 166:389-417.

Rong, J., J.E. Bowers, S.R. Schulze, V.N. Waghmare, C.J. Rogers, G.J. Pierce, H. Zhang, J.C. Estill, A.H. Paterson. 2005. Comparative genomics of Gossypium and Arabidopsis: Unraveling the consequences of both ancient and recent polyploidy. Genome Res. 15:1198-1210.

Saha, S., M. Karaca, J.N. Jenkins, A.E. Zipf, O.U.K. Reddy, R.V. Kantety. 2003. Simple sequence repeats as useful resources to study transcribed genes of cotton. Euphytica 130:355-364.

Saranga, Y., M. Menz, C. Jiang, R. Wright, D. Yakir, A. Paterson. 2002. Genetic mapping implicates osmotic potential as a major component of crop adaptation to arid conditions. Genome Res 11:1988-1995.

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

Shappley, Z.W., J.C. Mccarty, Jr., J. Zhu, J.N. Jenkins. 1998. Quantitative trait loci associated with agronomic and fiber traits of upland cotton [electronic resource]. Journal of cotton science (Online) Journal of cotton science 2:153-163.

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

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

Shen, X.L., T.Z. Zhang, W.Z. Guo, X.F. Zhu, X.Y. Zhang. 2006. Mapping fiber and yield QTLs with main, epistatic, and QTL X environment interaction effects in recombinant inbred lines of upland cotton. Crop Sci. 46:61-66.

Stelly, D.M. 1993. Interfacing cytogenetics with the cotton genome mapping effort. Beltwide Cotton Conferences. Proceedings:1545-1550.

Stewart, J.M. (1995) Potential for crop improvement with exotic germplasm and genetic engineering, in: F. N. Constable GA (Ed.), Challenging the future: Proceedings of the world cotton research conference, CSIRO, Melbourne. pp. 313–327.

Taliercio, E., A.C.M. Millena, P. Payton, M.L. Pierce, A.L. Phillips, M.A. Patil, M. Essenberg, R.D. Allen, H. Nguyen, N. Klueva. 2006. Analysis of ESTs from

multiple Gossypium hirsutum tissues and identification of SSRs [electronic resource]. Genome 49:306-319.

Ulloa, M., W.R. Meredith, Jr. 2000. Genetic linkage map and QTL analysis of agronomic and fiber quality traits in an intraspecific population [electronic resource]. Journal of cotton science (Online) Journal of cotton science 4:161-170.

Ulloa, M., W.R. Meredith, Z.W. Shappley, A.L. Kahler. 2002. RFLP genetic linkage maps from four F-2.3 populations and a joinmap of Gossypium hirsutum L. Theor. Appl. Genet. 104:200-208.

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

Van Deynze, A., K. Stoffel, M. Lee, T.A. Wilkins, A. Kozik, R.G. Cantrell, J.Z. Yu, R.J. Kohel, D.M. Stelly. 2009. Sampling nucleotide diversity in cotton. BMC Plant Biol. 9.

Wendel, J.F., R.C. Cronn. 2003. Polyploidy and the evolutionary history of cotton. Advances in Agronomy 78:139-186.

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

Yu, J., R.J. Kohel, C.W. Smith. 2010. The construction of a tetraploid cotton genome wide comprehensive reference map. Genomics 95:230-240.

Yu, K., S.J. Park, V. Poysa. 1999. Abundance and variation of microsatellite DNA sequences in beans (Phaseolus and Vigna). Genome 42:27-34.

Zeng, L., E. Taliercio, D.L. Boykin, W.R. Meredith. 2007. Evaluation of an Exotic Germplasm Population Derived from Multiple Crosses among Gossypium Tetraploid Species [electronic resource]. Journal of cotton science 11:118-127.

Zeng, L.H., W.R. Meredith, O.A. Gutierrez, 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. Theor. Appl. Genet. 119:93-103.

Zhang, J., W. Guo, T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L. x Gossypium barbadense L.) with a haploid population. Theor. Appl. Genet. 105:1166-1174.

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

Zhang, T.Z., Y.L. Yuan, J. Yu, W.Z. Guo, R.J. Kohel. 2003. Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. Theor. Appl. Genet. 106:262-268.

Zhang, Z., J. Rong, V.N. Waghmare, P.W. Chee, O.L. May, R.J. Wright, J.R. Gannaway, A.H. Paterson. 2011. QTL alleles for improved fiber quality from a wild Hawaiian cotton, Gossypium tomentosum. Theor. Appl. Genet.:1-14.

Zhao, X.P., Y.R. Lin, A.H. Paterson. (1994) Characterization and genetic mapping of DNA microsatellites from cotton., Plant genome II conference. Town & Country Conference Center, San Diego, January 1994.

CHAPTER 3

GENETIC ASSESSMENT OF TWO INTROGRESSION LINES WITH EXOTIC GENOME FOR FIBER QUALITY IMPROVEMENT OF UPLAND COTTON (*GOSSYPIUM HIRSUTUM* L.)[1]

[1] Pawan Kumar, Rippy Singh, Edward L. Lubbers, Xinlian Shen, Andrew H. Paterson, B. Todd Campbell, Donald C. Jones and Peng W. Chee. To be submitted to *Theoretical and Applied Genetics*

Abstract

Intraspecific mapping populations of *Gossypium hirsutum* from crosses between S-7235 and two introgression lines (Sealand 542 and Sealand 883) were used to identify the locations of introgressed *G. barbadense* chromatin and to identify quantitative trait loci (QTLs) controlling fiber quality traits. A total of 21 QTLs and 4 putative QTLs for fiber quality traits were identified, including five for fiber elongation, three for fiber fineness, three for short fiber content, four for fiber strength, six for fiber length, and four for fiber uniformity. At least 15 fiber quality QTLs were detected on the introgressed chromosomal segments; of these, favorable alleles for 9 QTLs were contributed by the Sealand lines. Stable introgression of *G. barbadense* fiber quality alleles in the Sealand lines and the absence of these alleles in predominant types of Upland cotton grown in the US make these introgression lines a valuable resource for Upland cotton improvement.

Introduction

Cotton fiber, elongated epidermal cells of the cotton ovule with a thickened secondary wall composed primarily of cellulose, is the foremost source of natural fiber used in the textile industry. Four *Gossypium* species have been domesticated for their cotton fiber of which the cultivated form of *Gossypium hirsutum* L. commonly called Upland cotton accounts for 96% of global cotton fiber production (Smith et al., 2008) The other three cultivated species, including *Gossypium barbadense* L., *Gossypium arboreum* L., and *Gossypium herbaceum* L., were once widely grown, but have been largely replaced by Upland cotton due its high yield potential and broader environmental adaptation.

Despite having greater productivity, genetic improvement of yield and fiber quality traits in Upland cotton is generally slow and faces many challenges. First is the negative association between fiber quality traits and yield components (Culp and Green, 1992). Since yield improvement has been the foremost objective of a breeding program, it is challenging for a breeder to target fiber quality without compromising lint yield. Second is the low genetic diversity in the *G. hirsutum* gene pool due to genetic bottlenecks imposed during polyploidization about 1-2 million years ago and its subsequent domestication from a small subset of its wild ancestors (Paterson et al., 2004). In addition, the traditional approach in Upland cotton improvement has relied on crossing lines from a few genetic backgrounds and reselection within the population for yield and fiber quality traits (Gingle et al., 2006; May et al., 1995; Van Esbroeck and Bowman, 1998). Together these events resulted in a cultivated Upland germplasm with a narrow genetic base which constrains future genetic improvement.

Large amount of genetic variation exist in the secondary and tertiary gene pools of *Gossypium* (Lacape et al., 2007; Lubbers and Chee, 2009; Percival and Kohel, 1990; Percy, 2009) and historical accounts indicated that the Upland germplasm has substantial exposure to both gene pools through intentional breeding efforts (Culp and Harrell, 1975; Lubbers and Chee, 2009; Ware, 1951). However, the extent of the genetic contribution of wild donors in Upland germplasm is largely unknown because evidence of putative interspecific introgression into Upland cotton can only be detected in loci that are fixed or nearly fixed in secondary gene pools but are absent or rare in Upland cotton germplasm (Lubbers and Chee, 2009). Thus, while pedigree analyses indicated that many cultivars were developed through interspecific hybridization, molecular analysis based on isozymes and DNA markers on the Upland cotton germplasm indicated that the genepool is largely homogeneous. Rare alleles, which have likely arisen through introgression, are restricted to only a few closely related cultivars within a germplasm group (Brubaker and Wendel, 1994; Iqbal et al., 2001; Wendel et al., 1992). These results suggest that the secondary and tertiary gene pools remains largely un-utilized in Upland cotton improvement.

Nonetheless, a number of successes in utilizing the secondary gene pool in the development of Upland cotton cultivars have been documented. For example, the cultivated form of *G. barbadense* (known as Pima, Sea Island, or Egyptian cotton) has superior fiber quality to that of Upland cotton and has been implicated in the development of a number of obsolete, extra-long staple cottons (ELS) (Ware, 1951; Smith 2004), including the series of "Sealand" germplasm lines released in the late 1930's by the USDA-ARS Pee Dee public cotton breeding programs at Florence, SC

(Jenkins, 1948). These Sealand lines have improved fiber length and fineness compared to Upland cultivars available during that time (Campbell et al., 2011), and the varieties 'Sealand 542' and 'Sealand 883' were commercially grown as Extra Long Staple (ELS) Upland varieties on approximately 1,000 acres in South Carolina, Georgia, and Florida until the 1950's (Culp and Harrell, 1974b).

The significantly longer and finer-fibers of Sealand germplasm lines as compared to their recurrent parents suggests that they contain stable introgressions of genes conferring these phenotypes from the *G. barbadense* donor parent. However, the extent of introgression, chromosome locations, and genetic effects of the introgressed chromatin are largely unknown. Our objectives in this study were to (i) identify the genomic position of the introgressed chromatin in Sealand 542 and Sealand 883 and (ii) to evaluate genetic effects of the introgressed alleles on fiber quality traits. Detailed genetic dissection of the introgressed regions offer the opportunity to better understand the genetic effects of backcross introgression involving interspecific hybridization and to provide tools for selecting rare introgressed progenies that carry desirable combinations of genes to further improve cotton fiber quality.

Materials and methods

Plant material and phenotypic measurements

Two F_2 populations designated as Pop-542 and Pop-883 were generated by crossing an Upland cotton line Suyuan 7235 as female parent with Sealand 542 (PI 528730) and Sealand 883 (PI 528875), respectively. The germplasm line Suyuan 7235 (referred to herein as S-7235) was developed by the Cotton Research Institute (CRI),

China by crossing *G. anomalum* with *G. hirsutum*, and then backcrossing the progeny to the *G. hirsutum* cultivar "Acala 3080" (PI 529543) (Qian et al., 1992). The germplasm lines, Sealand 542 and Sealand 883 (referred to herein as SL-542 and SL-883, respectively) were developed at the USDA-ARS Pee Dee Experiment Station, Florence, South Carolina from an interspecific cross between *G. barbadense* (Sea Island) line "Bleak Hall" (PI 608115) and an Upland cotton line "Coker Wilds" (Bowman et al., 2006a; Culp and Harrell, 1974b) followed by four backcrosses to the Upland cotton recurrent parent.

Identification of putatively introgressed chromatin is based on the assumption that the two Sealand lines, SL-542 and SL-883, are sister lines originating from the same pedigree (Bowman et al., 2006a; Culp and Harrell, 1974b), therefore the genomic differences between these two lines are hypothesized to be due to differential retention of the introgressed chromatin from the *G. barbadense* donor parent 'Bleak Hall'. Loci polymorphic between these Sealand lines indicate the putative introgressed chromatin. In order to detect introgression and rare alleles in the Sealand lines, a polymorphism survey panel was developed consisting of DNA from the three mapping parents along with the DNA from four historically important genotypes which represented the predominant germplasm groups of the US cotton.

The two F_2 populations (Pop-542 & Pop-883) comprised of 350 individuals each were planted at the William Gibbs Farm in Tifton, Georgia (31.27°N, 83.30°W) in the summer of 2005. Standard production practices were followed. Twenty-five mature cotton bolls from the individual F_2 plants were hand-picked and ginned on a table-top saw gin. Fiber quality of each ginned sample was measured using High Volume

Instruments (HVI) at the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, NC). The fiber quality measurements comprised fiber length or Upper Half Mean (UHM) in inches, fiber strength (STR) in g/tex, fiber fineness or micronaire (MIC), percent fiber elongation (ELO), percent short fiber content (SFC) and percent uniformity index (UI).

DNA preparation and molecular marker analysis

Genomic DNA of the parents and F_2 plants from both mapping populations was extracted following a published protocol (Paterson et al., 1993). Quantity and quality of the extracted DNA was determined by running the samples on a 0.8% agarose gel. The concentration of genomic DNA was estimated by comparing the size and intensity of each sample band with those of the standard λ uncut DNA (1μg/5μl) and commercial MassRuler (Thermo Fisher Scientific).

PCR reaction conditions were slightly modified from the previously described protocol (Chee et al., 2004). PCR amplification was performed in a PTC-100 or PTC-200 thermocycler (MJ Research Inc.). A 10μl reaction contained 10ng of template DNA, 0.5μM primer mix, 100μM dNTPs, 1.5mM $MgCl_2$, 3U of DNA polymerase, and 1X reaction buffer (100 mM Tris-HCl, pH 8.3, 500 mM KCl). The cycling conditions for PCR were 94 °C for 3 min, followed by 35 cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 1.2 min. After the last cycle, reactions were incubated at 72 °C for 6 min before cooling to 4 °C. For primer pairs that did not amplify in the first two attempts, a gradient reaction from 45 to 60 °C was performed to empirically determine the best annealing temperature (Chee et al., 2004). The amplified PCR products were

electrophoretically separated on a 10% non-denaturing polyacrylamide gel electrophoresis (PAGE) and were visualized by staining with silver nitrate following the published procedures (Zhang et al., 2002).

The polymorphism survey panel was genotyped with a total of 1170 Simple Sequence Repeats (SSR) markers, covering all of the 13 homeologs. Sequences of the SSR markers were downloaded from the Cotton Marker Database (Blenda et al., 2006) and were commercially synthesized by Eurofins MWG Operon (Huntsville, AL). Polymorphic markers were grouped in two groups: i) polymorphic between SL-542 and SL-883 as well as S-7235 and ii) polymorphic either between SL-542 and S-7235 or between SL-883 and S-7235, but not between SL-542 and SL-883.

Genetic mapping and QTL analysis

The phenotypic distributions and the correlations among the fiber quality traits in the Pop-542 and Pop-883 mapping populations were calculated with SAS version 9.1 (SAS Institute Inc, 1989). Linkage groups were constructed using the Mapmaker/EXP 3.0 (Lander et al., 1987) software. The assembly of the linkage groups was done using the 'group' command with a LOD score of 3.0 and a maximum recombination fraction of 30 cM. Recombination units were converted into genetic distances by using the Kosambi mapping function (Kosambi, 1944) with the "error detection" command on. Unlinked markers were added to framework using the 'try' and 'compare' commands. The final order of marker sequence on a linkage group was confirmed using the 'ripple' command. Assignment of linkage groups to the chromosome and subgenome is based on the published comprehensive reference map of tetraploid cotton (Yu et al., 2010).

Detection of quantitative trait loci and estimation of various genetic parameters were performed by Composite Interval Mapping (CIM) function implemented in the software WinQTL Cartographer version 2.5 (Wang et al., 2005). The likelihood ratio (LR) threshold value ($\alpha = 0.05$) for declaring the presence of QTL were estimated after 1,000 permutations (Doerge and Churchill, 1996). Peaks below this threshold but with a LOD > 2.5 were considered putative QTLs. Mapping was performed at 1 cM walk speed in a 10 cM window and five background cofactors where the cofactors were selected via forward-backward stepwise regression method. QTLs were defined by one-LOD confidence intervals on either sides of the peak position and were named following a method used in rice (McCouch et al., 1997). Briefly, the QTL is designated as '*q*' followed by an abbreviation of the trait name which is then followed by the chromosome name. Multiple QTLs on the same chromosome are distinguished by an alphabetical suffix.

Results

Putative introgressed regions

Mapping locations of the 85 loci that were polymorphic between the two Sealand parents detected 12 putative introgressions on 7 chromosomes (3, 5, 6, 9, 18, 21 and 23) of SL-542 and 10 putative introgressions on 5 chromosomes (5, 11, 15, 16 and 25) of SL-883. Putative introgressed regions of SL-542 covered 351 cM (6% of the cotton genome) and 235 cM for SL-883 (4% of the cotton genome). Location of the introgressed segments detected in SL-542 and SL-883 are shown on a genome-wide, comprehensive reference map of tetraploid cotton (Yu et al., 2010) (Figure 2).

Of the 12 putative introgressions detected in SL-542, three were present as a single segments on chromosomes 3 (20.9 cM), 9 (53.7 cM), and 23 (63.7 cM); two each on chromosomes 5, 18, and 21; and three were detected on Chromosome 6 (Figure 3). Of the 10 introgressed regions detected in SL-883, two were present as single segments on chromosomes 16 (38.8 cM) and 25 (7.6 cM), two on Chromosome 15, and three each on chromosomes 5 and 11 (Figure 3).

Eighty markers that were not polymorphic between the Sealand lines but were polymorphic with S-7235 mapped to 9 genomic regions on 8 chromosomes. These genomic regions either represent random polymorphism between S-7235 and the Sealand line or may represent common introgressions residing in the genome of the two Sealand lines.

Linkage maps

Of the 1170 SSR markers surveyed, 165 loci were polymorphic between the mapping parents, including 125 loci polymorphic between the parents of Pop-542 and 120 loci polymorphic between the parents of Pop-883. In addition, 80 loci were common between the two Sealand parents but were polymorphic with S-7235 and 85 loci were polymorphic between the two Sealand parental lines.

Genetic mapping determined that the 125 polymorphic SSR loci in the population Pop-542 were grouped into 21 linkage groups on 15 chromosomes covering 530 cM recombinational length of the cotton genome. The number of markers per chromosome ranged from 2 (Chromosomes 6, 7, and 20) to 51 (Chromosome 24) with an average of 8

markers per chromosome. Similarly, in population Pop-883, the 120 polymorphic SSR markers were grouped into 19 linkage groups, which mapped to 13 chromosomes covering 411 cM recombinational length of the cotton genome. The number of markers per chromosome ranged from 2 (Chromosomes 6, 7, and 20) to 51 (Chromosome 24) with an average of 9 markers per chromosome.

Population biometric parameters

The fiber properties of the parental lines SL-542, SL-883, and S-7235 are shown in Table 1. Between the mapping parents of Pop-542, all fiber quality traits except for fiber fineness were significantly different with SL-542 having higher fiber elongation, short fiber content and S-7235 possessing greater fiber length, fiber uniformity and fiber strength. Between the mapping parents of Pop-883, S-7235 had significantly finer fiber, tighter uniformity, and stronger fiber while SL-883 had significantly longer fiber. Fiber elongation and short fiber content were not significantly different between SL-883 and S-7235 (Table 1). All six fiber quality traits were normally distributed in both Pop-542 and Pop-883 (Figure 1). Between the two mapping populations the mean fiber elongation, fineness, fiber strength, and uniformity index were higher for population Pop-542 while mean short fiber content and fiber length were greater for population Pop-883. Transgressive segregants were present in both mapping populations for all the six fiber quality traits.

Pearson correlation coefficients among the fiber quality traits are listed in Table 2. Fiber length, fiber strength, and uniformity index were positively correlated to each other and were negatively correlated with short fiber content. Fiber length was negatively

correlated with fiber elongation and fiber fineness. High negative correlation (R^2=0.80) between uniformity index and short fiber content was observed in both Pop-542 and Pop-883.

QTL analysis

A total of 25 QTLs affecting fiber quality were detected by composite interval mapping of both populations. Nine QTLs were unique to Pop-542 and twelve were unique to Pop-883 while four QTLs were common between the two populations. Of the total QTLs, 21 were at or above the threshold LOD score while 4 were putative. Nine QTLs were identified on 4 A subgenome chromosomes and 16 were on 9 D subgenome chromosomes. Details of the QTLs identified for different fiber quality traits is summarized below.

Fiber elongation (ELO)

Five QTLs were identified for fiber elongation (Table 3). Four QTLs (chr. 08, 23, and 24) in Pop-542 and one in Pop-883 (chr. 11). Of the QTLs identified in Pop-542, three above the threshold LOD score while one (chr. 24) was identified as putative. The percent phenotypic variation explained by the QTLs ranged from 2.81 to 9.46% with genetic effects ranging from 0.02 to 0.76%. The favorable alleles for four QTLs (chr. 08, 11, 23, & 24) originated from the Sealand parents while the favorable allele for *qELO-Chr08* originated from S-7235 (Table 3). Favorable alleles from SL-542 was expected as it had significantly greater fiber elongation than S-7235 parent.

Micronaire (MIC)

A total of three QTLs for micronaire were identified (Table 3). Two QTLs (chr. 11 & 17) were unique to Pop-883 while the QTL on chr. 24 was identified in both Pop-542 and Pop-883 mapping populations. The percent phenotypic variation (PV) explained ranged from 5.81 to 33.57% and the genetic effects ranged from 0.3 to 0.41. The favorable allele for *qMIC-Chr11* from SL-883 may be due to transgressive segregation while favorable alleles from S-7235 parent was expected (Table 3).

Short fiber content (SFC)

A total of three QTLs for short fiber content were identified (Table 3). Two (chr. 05 & 24) were detected in Pop-883 and one QTL on chr. 09 was uniquely identified in Pop-542 above the threshold LOD score. QTLs unique to Pop-883 accounted for 3.75 to 6.48% PV and the QTL in Pop-542 accounted for 9.4% PV. The genetic effects ranged from 0.19 to 1.18%. Alleles increasing short fiber content for QTLs *qSFC-Chr05*, *qSFC-Chr09,* and *qSFC-Chr24* originated from S-7235, SL-542, and SL-883 respectively (Table 3). Favorable alleles for short fiber content from Sealand lines was expected as both, SL-542 and SL-883, had greater short fiber content than S-7235 parent.

Fiber strength (STR)

A total of four QTLs affecting fiber strength was identified (Table 3). Two (chr. 16 & 26) were unique to Pop-883 while the other two (chr. 17 & 24) were identified in both Pop-542 and Pop-883 populations. The percent phenotypic variation accounted by these QTLs ranged from 5.96 to 30.92% and the genetic effects ranged from 0.53 to 2.32

g/tex. The favorable allele for *qSTR-Chr16* originated from SL-883 while S-7235 contributed the favorable alleles for the rest of the QTLs (Table 3). Favorable alleles from S-7235 parent was expected as it possessed fibers with significantly higher fiber strength whereas the QTL *qSTR-Chr16* for which SL-883 contributed favorable allele was unexpected and might be due to transgressive segregation.

Fiber length (UHM)

A total of four QTLs and two putative QTLs for fiber length were identified (Table 3). Three QTLs (chr. 05, 16, & 25) were identified in Pop-883 while a QTL on Chr. 09-b and two putative QTLs (chr. 9-a & 21) were identified in Pop-542. The phenotypic variation explained by these QTLs and putative QTLs ranged from 4.35 to 15.39% with genetic effects ranging from 0.02 to 0.03 inches. The favorable allele for *qUHM-Chr16* and *qUHM-Chr25* originated from SL-883 while S-7235 contributed the favorable alleles for remaining QTLs and putative QTLs (Table 3). Sealand parent SL-883 had significantly greater fiber length than S-7235 parent therefore favorable alleles from SL-883 for fiber length in Pop-883 were expected. Between the parents of Pop-542, S-7235 had significantly greater fiber length therefore favorable alleles from S-7235 parent in Pop-542 was also expected.

Fiber uniformity (UI)

Four QTLs were detected for fiber uniformity (Table 3). Two (chr. 08 & 15) were unique to Pop-883 while one on chr. 18 was unique to Pop-542. The QTL on chr. 24 met the significance threshold in both Pop-542 and Pop-883 however it was putative in Pop-

542. The phenotypic variation explained by these QTLs ranged from 2.01 to 9.67 % with genetic effects ranging from 0.36 to 0.96%. The favorable allele for *qUI-Chr15* originated from SL-883 while S-7235 contributed the favorable alleles for the remaining QTLs (Table 3). Favorable alleles form S-7235 parent was expected as it had significantly greater fiber uniformity than both Sealand parents; however the QTL *qUI-Chr15* for which SL-883 contributed favorable allele may be a transgressive segregant.

Discussion

Introgression from *G. barbadense* into the Sealand series was suspected, not only because of their pedigree (Culp and Harrell, 1974b) but also due to their unique fiber quality which was on par with Sea Island cotton. Genomic survey of the two Sealand lines confirmed twenty-two introgressed chromosome segments ranging in size from 1.2 to 63.7 cM. The introgression events were found to be equally distributed between A_t and D_t sub-genomes of tetraploid cotton and the introgression pattern were unique to each Sealand line. Only one (chr. 05) of the chromosomes with introgressed segments was common between the two Sealand lines and even then the introgressed regions on chr.05 of did not overlap.

The uniqueness in introgression pattern between SL-542 and SL-883 likely accounts for the differences observed in fiber quality between these two lines. The retention of *G. barbadense* chromatin in each Sealand line was unique, suggesting that these lines and perhaps others in the Sealand series may have been independently derived during backcrossing. Phenotypic selection pressure for fiber quality traits and yield components applied during backcrossing may have resulted in lines with different but

stable introgression of *G. barbadense* alleles. The Germplasm Resources Information Network (GRIN) (www.ars-grin.gov) indicates that at least eight Sealand lines, including SL-542 and SL-883 exist in the database. The six other 'Sealand lines' (SL-1, SL-3, SL-7 white flower, SL-7 yellow flower, SL-391 and SL-472) may also have been independently derived and thus the genomic composition of each of these lines may be unique. Therefore, the Sealand lines together as a group may serve as reservoir of novel fiber quality alleles of *G. barbadense* origin. Furthermore, the phenotypic observations at GRIN show that the remaining Sealand lines also have unique fiber quality and yield traits, therefore it may be worthwhile to explore these lines for yield and fiber quality alleles.

Sealand parents contributed favorable alleles for several fiber quality traits. Out of the 25 QTLs identified in this study, favorable alleles for 11 QTLs were contributed by the Sealand parents, including 3 QTLs (*qELO-Chr23-a*, *qELO-Chr23-b,* and *qSFC-Chr09*) for which SL-542 contributed favorable alleles were identified on the introgressed segments. Favorable alleles from SL-542 parent for fiber elongation QTLs were expected since the mean fiber elongation of SL-542 was the highest (6.01%) among the three mapping parents. Selection for fiber elongation is relatively new in cotton breeding programs (Zhang et al., 2011), therefore secondary and tertiary gene pool species may tend to harbor novel alleles that have not been fixed in the Upland cotton germplasm (Chee et al. 2005). Although it is doubtful that fiber elongation may have been a selection criteria during Sealand germplasm development (May, 2000), the introgression of these QTLs in SL-542 may have been the result of its correlation with

other traits such as fiber strength, which was an important fiber trait that was routinely tested during the 1930s.

Sealand parent SL-883 contributed favorable alleles for seven (43.75%) of the sixteen fiber quality QTLs detected in Pop-883. Of these seven QTLs, six (85.7%), *qELO-Chr11*, *qMIC-Chr11*, *qSTR-Chr16*, *qUHM-Chr16*, *qUHM-Chr25,* and *qUI-Chr15* were identified on four introgressed segments. A single introgressed segment on chr. 16 harbored favorable alleles for fiber length and fiber strength, and this segment alone could improve fiber strength by 0.9 g/tex and increase fiber length by 0.51 mm. A noteworthy observation is that all chromosomes registering an introgression event in SL-883 harbored one or more fiber quality QTLs. In contrast, only four of the seven chromosomes registering introgression event in SL-542 harbored fiber quality QTLs. A possible explanation for this observation is that perhaps the retention of these segments may have been due to other agronomic phenotypic effects such as yield components. The GRIN observation shows that SL-542 has better lint percentage than SL-883 and other Sealand lines. Alternatively, some of these *G. barbadense* segments could represent random introgression with no improvement value.

Many QTLs were identified in the same marker interval, for example, chr.16 and 24 carried multiple fiber quality QTLs. The correlation observed among fiber quality traits may be the direct result of co-localization of fiber quality QTLs. Co-localization of QTLs is observed in many crops (Lexer et al., 2003; Saliba-Colombani et al., 2001; Yan et al., 2009) including cotton (Lacape et al., 2010; Shen et al., 2011; Zhang et al., 2011) where QTLs with both positive and negative genetic effects were detected on same

chromosomal position. Another noteworthy observation is that, sixteen (64%) of the fiber quality QTLs were identified on D_t sub-genome. Although the D-subgenome progenitor did not produce spinnable fibers, it has loci that influence the quality of the fiber produced in allotetraploid cottons which indicates that the polyploidization of *Gossypium* might have given rise to novel variation for fiber quality traits (Jiang et al., 1998; Wendel, 2000). Similar results were previously reported in cotton (Chee et al., 2005b; Jiang et al., 1998; Mei et al., 2004; Paterson et al., 2003; Rong et al., 2007b) where the genetic control of fiber quality by the D_t sub-genome was significantly greater than that of A_t sub-genome.

Current efforts in genetic analysis and interspecific introgression to improve fiber quality involving *G. hirsutum* by *G. barbadense* crosses are conducted predominately with Pima cotton, Egyptian cotton, or with unimproved *G. barbadense* accessions (Chee and Campbell, 2008; Manikanda et al., 2011). Sea Island cottons are most closely related to the primitive forms of *G. barbadense* found in South America (Stephens, 1974) and were introduced in the US from the Caribbean Islands in the late 18th century (Fryxell, 1965; Stephens, 1976), whereas Pima cotton originated from the individual plant selection in Egyptian cotton cultivars introduced from the Nile Valley in the early 1900's (McGowan, 1961; Percy, 2009). The *G. barbadense* donor parent of the Sealand lines, a cultivar named Bleak Hall, is an improved Sea Island cotton variety with superior fiber quality including fiber length ranging from 51 to 63.5 mm (Smith et al., 2008) compared to 28 to 31 mm length of TM-1, a Upland cotton genetic standard (Zhang et al., 2003). Therefore, it is likely that the Sealand lines may harbor *G. barbadense* alleles of South

American origin that may be novel to those interspecific *G. hirsutum* by *G. barbadense* genetic populations currently developed from Pima cultivars and lines.

Because the foundation of the modern Pee Dee germplasm includes the Sealand lines, it is possible that one or more of the introgressed alleles detected in this study may be prevalent in modern Pee Dee Upland cotton. Historical records indicate that at least two Sealand lines, Sealand-542 and Sealand-7, were involved in Pee Dee germplasm development (Bowman et al., 2006a). Therefore the Extra Long Staple cotton developed from the Pee Dee germplasm lines may be carrying Sea Island alleles of South American origin. Work is now underway to test this hypothesis by genotyping the DNA markers identified in this study which detect *G. barbadense* introgression, on key Pee Dee germplasm lines developed since this breeding program was initiated in the 1940's (Culp and Harrell, 1974b).

References

Blenda, A., J. Scheffler, B. Scheffler, M. Palmer, J.M. Lacape, J.Z. Yu, C. Jesudurai, S. Jung, S. Muthukumar, P. Yellambalase, S. Ficklin, M. Staton, R. Eshelman, M. Ulloa, S. Saha, B. Burr, S.L. Liu, T.Z. Zhang, D.Q. Fang, A. Pepper, S. Kumpatla, J. Jacobs, J. Tomkins, R. Cantrell, D. Main. 2006. CMD: a Cotton Microsatellite Database resource for Gossypium genomics. Bmc Genomics 7:-.

Bowman, D.T., O.A. Gutierrez, R.G. Percy, D.S. Calhoun, O.L. May. 2006. Pedigrees of upland and pima cotton cultivars released between 1970 and 2005. Mississippi Agricultural and Forestry Experiment Station Technical Bulletin 1155:57.

Brubaker, C.L., J.F. Wendel. 1994. Reevaluating the Origin of Domesticated Cotton (Gossypium-Hirsutum Malvaceae) Using Nuclear Restriction-Fragment-Length-Polymorphisms (Rflps). Am. J. Bot. 81:1309-1326.

Campbell, B.T., P.W. Chee, E. Lubbers, D.T. Bowman, W.R. Meredith, J. Johnson, D.E. Fraser. 2011. Genetic Improvement of the Pee Dee Cotton Germplasm Collection following Seventy Years of Plant Breeding. Crop Sci. 51:955-968.

Chee, P., B. Campbell. 2008. Bridging Classical and Molecular Genetics of Cotton Fiber Quality and Development. Genetics and Genomics of Cotton:283.

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

Chee, P.W., J. Rong, D. Williams-Coplin, S.R. Schulze, A.H. Paterson. 2004. EST derived PCR-based markers for functional gene homologues in cotton. Genome 47:449-462.

Culp, T.W., C.C. Green. 1992. Performance of obsolete and current cultivars and Pee Dee germplasm lines of cotton. Crop Sci. 32:35-41.

Culp, T.W., D.C. Harrell. 1974. Breeding quality cotton at the Pee Dee Experiment Station Florence, SC. USDA-ARS. Pubi. ARS-S-30 12.

Culp, T.W., D.C. Harrell. 1975. Influence of Lint Percentage, Boll Size, and Seed Size on Lint Yield of Upland Cotton with High Fiber Strength. Crop Sci. 15:741-746.

Doerge, R., G. Churchill. 1996. Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285.

Fryxell, P.A. 1965. Need for Taxonomic Research in Gossypium. Am. J. Bot. 52:650-&.

Gingle, A.R.Y., H. Chee, P.W. May, O.L. Rong, J. Bowman, D.T. Lubbers, E.L. Day, J.L.D. Paterson, H. Andrew. 2006. An integrated web resource for cotton. Crop Sci. 46:1998.

Iqbal, M.J., O.U.K. Reddy, K.M. El-Zik, 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.

Jenkins, J.G. 1948. Sea Island cotton breeding. 28th Ann. Rep. Ga Coastal Plains Exp Stn Bull 16:1.

Jiang, C.X., R.J. Wright, K.M. El-Zik, A.H. Paterson. 1998. Polyploid formation created unique avenues for response to selection in Gossypium (cotton). P Natl Acad Sci USA 95:4419-4424.

Kosambi, D. 1944. The estimation of map distances from recombination values. Ann Eugen 12:172-175.

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

Lacape, J.M., D. Llewellyn, J. Jacobs, T. Arioli, D. Becker, S. Calhoun, Y. Al-Ghazi, S. Liu, O. Palaï, S. Georges. 2010. Meta-analysis of cotton fiber quality QTLs across

diverse environments in a Gossypium hirsutum x G. barbadense RIL population. BMC Plant Biol. 10:132.

Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln, L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174.

Lexer, C., M. Welch, J. Durphy, L. Rieseberg. 2003. Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: Implications for the origin of Helianthus paradoxus, a diploid hybrid species. Mol. Ecol. 12:1225-1235.

Lubbers, E.L., P.W. Chee. 2009 The Worldwide Gene Pool of G. hirsutum and its Improvement, *In* A. H. Paterson (Ed.), Genetics and Genomics of Cotton. pp. 1-30.

Manikanda, B., K. Thiyagu, B. Urbi, M. Santhoshkumar, A. Gopikrishnan, S. Aravind, S. Gat, R. Ravikesavan. 2011. Marker-assisted breeding as next-generation strategy for genetic improvement of productivity and quality: can it be realized in cotton? International journal of plant genomics 2011.

May, O.L. 2000. Genetic variation in fiber quality. Cotton fibers–developmental biology, quality improvement, and textile processing. Food Products Press, New York:183–229.

May, O.L., D.T. Bowman, D.S. Calhoun. 1995. Genetic diversity of U.S. upland cotton cultivars released between 1980 and 1990. Crop Sci. 35:1570-1574.

Mccouch, S., Y. Cho, M. Yano, E. Paul, M. Blinstrub, H. Morishima, T. Kinoshita. 1997. Report on QTL nomenclature. Rice Genet Newsl 14.

Mcgowan, J.C. (1961) History of extra-long staple cottons Hill Printing Co., El Paso, Tex.

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

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

Paterson, A.H., C.L. Brubaker, 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 - ISPMB 11:122-127.

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

Percival, A.E., R.J. Kohel. 1990. Distribution, collection, and evaluation of Gossypium. Advances in Agronomy 44:225-256.

Percy, R.G. 2009 The Worldwide Gene Pool of Gossypium barbadense L. and Its Improvement, *In* A. H. Paterson (Ed.), Genetics and Genomics of Cotton. pp. 1-16.

Qian, S., J. Huang, Y. Peng, B. Zhou, M. Ying, D. Shen, G. Liu, T. Hu, Y. Xu, L. Gu. 1992. Studies on the hybrid of G. hirsutum L. and G. anomalum Wawr. & Peyr. and application in breeding. Sci. Agric. Sinica 25:44-51.

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

Saliba-Colombani, V., M. Causse, D. Langlois, J. Philouze, M. Buret. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. Theor. Appl. Genet. 102:259-272.

Sas Institute Inc. (1989) SAS/STAT user's guide, SAS Institute Inc., Cary, NC.

Shen, X., Z. Cao, R. Singh, E.L. Lubbers, P. Xu, C.W. Smith, A.H. Paterson, P.W. Chee. 2011. Efficacy of qFL-chr1, a Quantitative Trait Locus for Fiber Length in Cotton (Gossypium spp.). Crop Sci. 51:2005.

Smith, C.W., P.S. Thaxton, I.N. Brown, S. Hague, E. Hequet. 2008. Development of Extra-Long Staple Upland Cotton. Crop Sci. 48:1823-1831.

Stephens, S.G. 1974. The use of two polymorphic systems, nectary fringe hairs and corky alleles, as indicators of phylogenetic relationships in New World cottons. Biotropica:194-201.

Stephens, S.G. 1976. The origin of Sea Island cotton. Agricultural history:391-399.

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

Wang, S., C. Basten, Z. Zeng. 2005. Windows QTL cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Ware, J.O. (1951) Origin, rise and development of American upland cotton varieties and their status at present University of Arkansas College of Agriculture, Agricultural Experiment Station, Fayetteville, AR

Wendel, J.F. 2000. Genome evolution in polyploids. Plant Mol. Biol. 42:225-249.

Wendel, J.F., C.L. Brubaker, A.E. Percival. 1992. Genetic Diversity in Gossypium-Hirsutum and the Origin of Upland Cotton. Am. J. Bot. 79:1291-1310.

Yan, X., J. Li, F. Fu, M. Jin, L. Chen, L. Liu. 2009. Co-location of seed oil content, seed hull content and seed coat color QTL in three different environments in Brassica napus L. Euphytica 170:355-364.

Yu, J., R.J. Kohel, C.W. Smith. 2010. The construction of a tetraploid cotton genome wide comprehensive reference map. Genomics 95:230-240.

Zhang, J., W. Guo, T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L. x Gossypium barbadense L.) with a haploid population. Theor. Appl. Genet. 105:1166-1174.

Zhang, T.Z., Y.L. Yuan, J. Yu, W.Z. Guo, R.J. Kohel. 2003. Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. Theor. Appl. Genet. 106:262-268.

Zhang, Z., J. Rong, V.N. Waghmare, P.W. Chee, O.L. May, R.J. Wright, J.R. Gannaway, A.H. Paterson. 2011. QTL alleles for improved fiber quality from a wild Hawaiian cotton, Gossypium tomentosum. Theor. Appl. Genet.:1-14.

Table 3.1. Fiber quality comparison between mapping parents

Mapping Parents	MIC		UHM		UI		STR		ELO		SFC	
	Mean	PMD[a]	Mean	PMD[a]	Mean	PMD[a]	Mean	PMD[a]	Mean	PMD[a]	Mean	PMD[a]
S-7235	4.039		32.070		84.833		36.311		4.117		6.250	
SL-542	3.992	0.048	28.100	3.97**	82.486	2.347**	27.623	8.688**	6.014	1.897**	7.633	0.086**
SL-883	3.604	0.435**	34.230	2.16**	84.118	0.715**	32.023	4.288**	4.027	0.089	6.767	0.516

** Significant at P= 0.001

a= Mapping parents mean difference

Table 3.2 Pearson Correlation Coefficients among fiber quality traits

Trait	Population	Micronaire	Short fiber content	Fiber strength	Uniformity index	Fiber length
Fiber elongation	Pop-542	0.05	0.09	-0.19*	-0.10	-0.18
	Pop-883	0.16	-0.15	-0.05	0.16*	-0.31**
Micronaire	Pop-542		-0.19*	0.30**	0.22*	-0.31**
	Pop-883		-0.34**	0.33**	0.50	-0.15*
Short fiber content	Pop-542			-0.42**	-0.80**	-0.47**
	Pop-883			-0.43**	-0.79**	-0.35**
Fiber strength	Pop-542				0.47**	0.31**
	Pop-883				0.54**	0.30**
Uniformity index	Pop-542					0.53**
	Pop-883					0.27**

** Significant at P= 0.001

* Significant at P= 0.01

Table 3.3. Summary of fiber quality QTLs identified in Pop-542 and Pop-883.

Trait	QTLs	Population	LOD	Add[a]	Dom[b]	d/a[c]	PV%[d]	+ve Allele
Fiber elongation								
	qELO-Chr11	Pop-883	4.0	-0.2	-0.18	0.89	2.81	SL-883
	qELO-Chr08	Pop-542	3.0	0.02	-0.46	-22.89	4.17	S-7235
	qELO-Chr23-a	Pop-542	6.8	-0.64	-0.38	0.60	9.46	SL-542
	qELO-Chr23-b	Pop-542	10.6	-0.76	-0.53	0.69	7.81	SL-542
	qELO-Chr24	Pop-542	2.5	-0.08	0.42	-5.26	4.19	SL-542
Micronaire								
	qMIC-Chr11	Pop-883	3.9	-0.38	-0.10	0.27	5.81	SL-883
	qMIC-Chr17	Pop-883	7.7	0.3	-0.07	-0.25	11.67	S-7235
	qMIC-Chr24	Pop-542	9.6	0.33	0.02	0.07	21.34	S-7235
		Pop-883	11.8	0.41	-0.07	-0.17	33.57	S-7235
Short fiber content								
	qSFC-Chr05	Pop-883	3.8	0.19	-0.53	-2.77	6.84	S-7235
	qSFC-Chr09	Pop-542	4.2	-1.18	-0.44	0.37	9.4	SL-542
	qSFC-Chr24	Pop-883	4.6	-0.75	-0.79	1.05	3.75	SL-883
Fiber strength								
	qSTR-Chr26	Pop-883	3.9	0.65	0.50	0.77	5.96	S-7235
	qSTR-Chr16	Pop-883	3.4	-0.9	0.40	-0.44	9.37	SL-883
	qSTR-Chr17	Pop-542	9.4	0.53	0.28	0.53	8.73	S-7235
		Pop-883	14.2	0.72	-0.02	-0.03	11.24	S-7235
	qSTR-Chr24	Pop-542	12.5	1.93	0.20	0.10	22.01	S-7235
		Pop-883	14.1	2.32	0.10	0.04	30.92	S-7235
Fiber length								
	qFL-Chr5	Pop-883	3.2	0.51	0.51	1.00	6.47	S-7235
	qFL-Chr09-a	Pop-542	2.7	0.51	-0.25	-0.50	10.45	S-7235
	qFL-Chr09-b	Pop-542	3.0	0.76	0.51	0.67	4.35	S-7235
	qFL-Chr16	Pop-883	4.4	-0.51	0.25	-0.50	8.22	SL-883
	qFL-Chr21	Pop-542	2.6	0.52	-0.25	-0.50	9.6	S-7235
	qFL-Chr25	Pop-883	5.0	-0.76	0.25	-0.33	15.39	SL-883
Fiber uniformity								
	qUI-Chr08	Pop-883	3.6	0.71	0.68	0.96	3.97	S-7235
	qUI-Chr15	Pop-883	3.4	-0.36	0.61	-1.70	5.1	SL-883
	qUI-Chr18	Pop-542	3.1	0.44	0.47	1.06	9.67	S-7235
	qUI-Chr24	Pop-542	2.6	0.95	0.33	0.35	2.01	S-7235
		Pop-883	7.9	0.96	0.55	0.58	7.81	S-7235

[a]-Additive effects, [b]-Dominance effect, [c]-Degree of dominance, [d]-The phenotypic variance explained by the QTL.
d/a value <1= additive gene action, >1=dominant gene action, >2= over dominance.

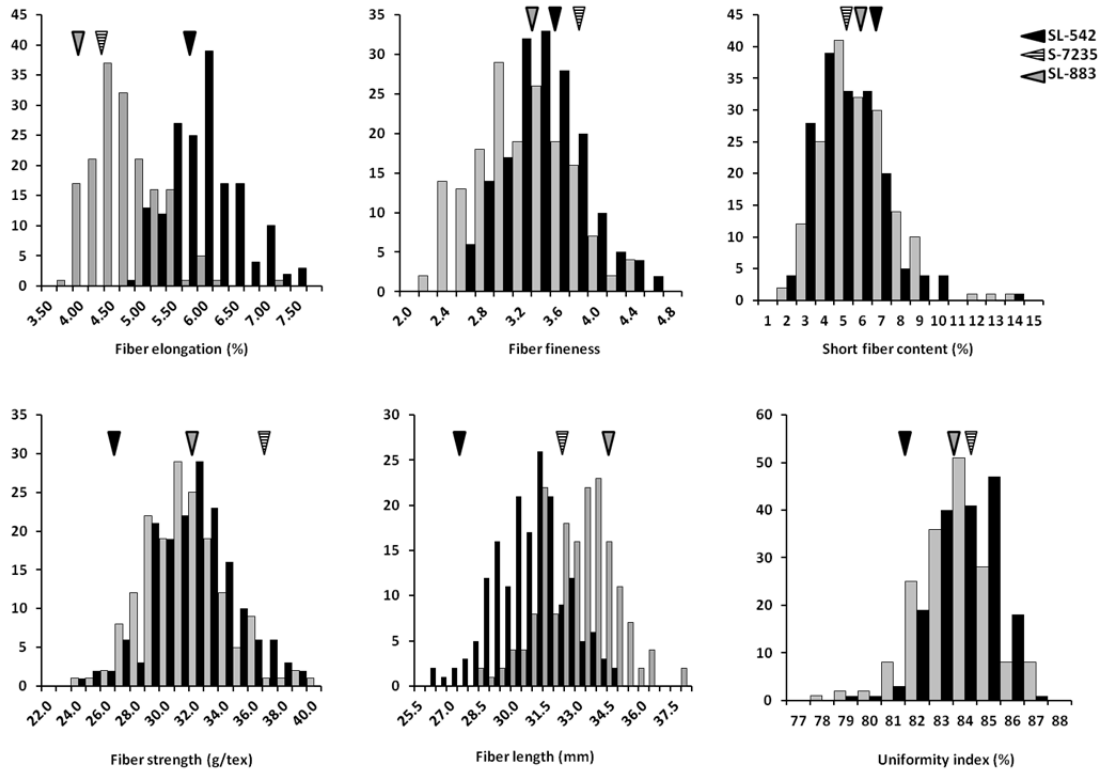


Figure 3.1. Distribution of fiber quality traits in Pop-542 (Black bars) and Pop-883 (Gray bars). Phenotypic values of the mapping parents are shown by the triangles.

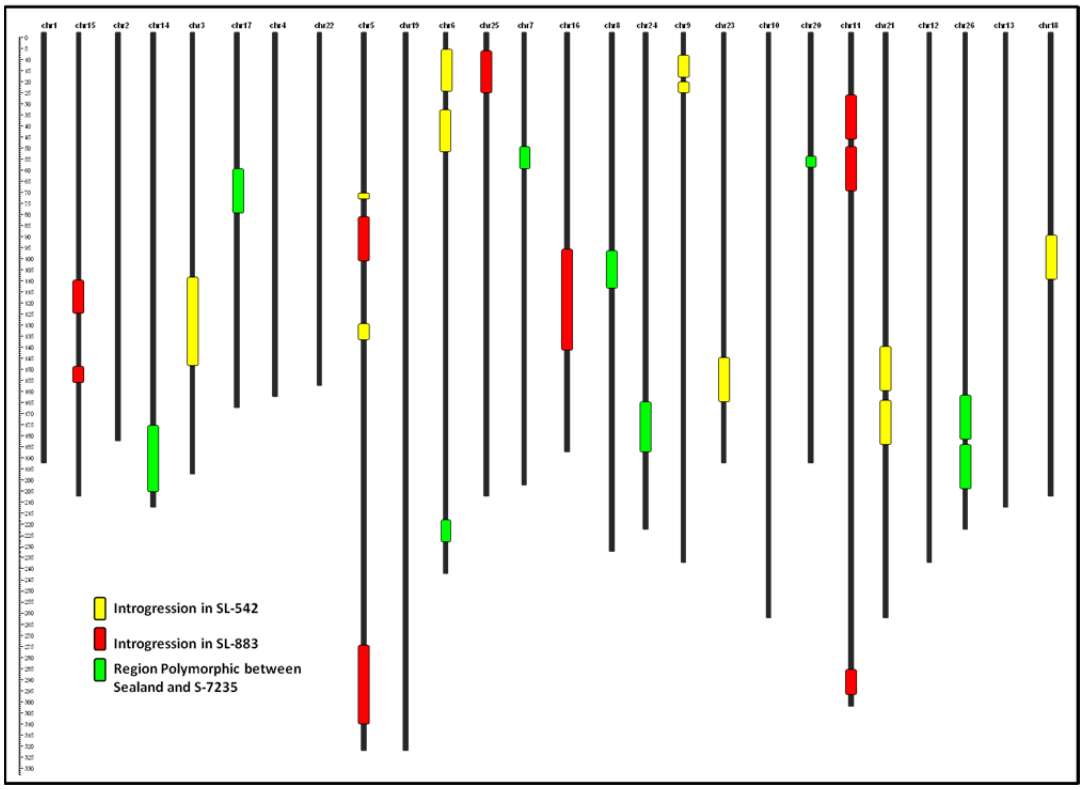
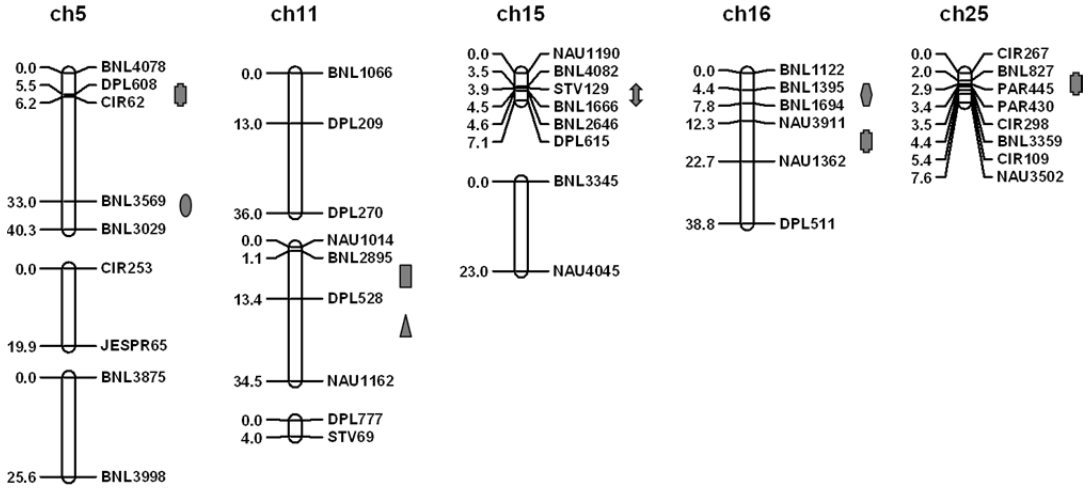
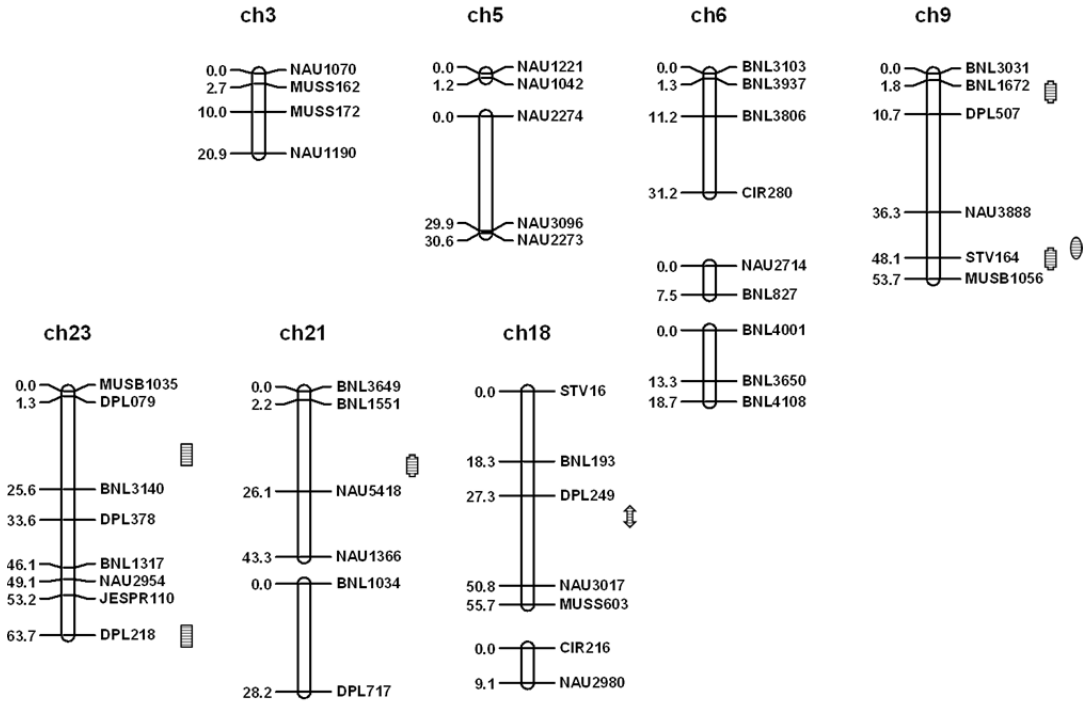


Figure 3.2. Tentative locations of the introgressed segments plotted on the genome wide comprehensive reference map of tetraploid cotton by Yu et al. (2010).

A


ch5
0.0 BNL4078
5.5 DPL608
6.2 CIR62
33.0 BNL3569
40.3 BNL3029
0.0 CIR253
19.9 JESPR65
0.0 BNL3875
25.6 BNL3998
ch11
0.0 BNL1066
13.0 DPL209
36.0 DPL270
0.0 NAU1014
1.1 BNL2895
13.4 DPL528
34.5 NAU1162
0.0 DPL777
4.0 STV69
ch15
0.0 NAU1190
3.5 BNL4082
3.9 STV129
4.5 BNL1666
4.6 BNL2646
7.1 DPL615
0.0 BNL3345
23.0 NAU4045
ch16
0.0 BNL1122
4.4 BNL1395
7.8 BNL1694
12.3 NAU3911
22.7 NAU1362
38.8 DPL511
ch25
0.0 CIR267
2.0 BNL827
2.9 PAR445
3.4 PAR430
3.5 CIR298
4.4 BNL3359
5.4 CIR109
7.6 NAU3502

B


ch3
0.0 NAU1070
2.7 MUSS162
10.0 MUSS172
20.9 NAU1190
ch5
0.0 NAU1221
1.2 NAU1042
0.0 NAU2274
29.9 NAU3096
30.6 NAU2273
ch6
0.0 BNL3103
1.3 BNL3937
11.2 BNL3806
31.2 CIR280
0.0 NAU2714
7.5 BNL827
0.0 BNL4001
13.3 BNL3650
18.7 BNL4108
ch9
0.0 BNL3031
1.8 BNL1672
10.7 DPL507
36.3 NAU3888
48.1 STV164
53.7 MUSB1056
ch23
0.0 MUSB1035
1.3 DPL079
25.6 BNL3140
33.6 DPL378
46.1 BNL1317
49.1 NAU2954
53.2 JESPR110
63.7 DPL218
ch21
0.0 BNL3649
2.2 BNL1551
26.1 NAU5418
43.3 NAU1366
0.0 BNL1034
28.2 DPL717
ch18
0.0 STV16
18.3 BNL193
27.3 DPL249
50.8 NAU3017
55.7 MUSS603
0.0 CIR216
9.1 NAU2980

C

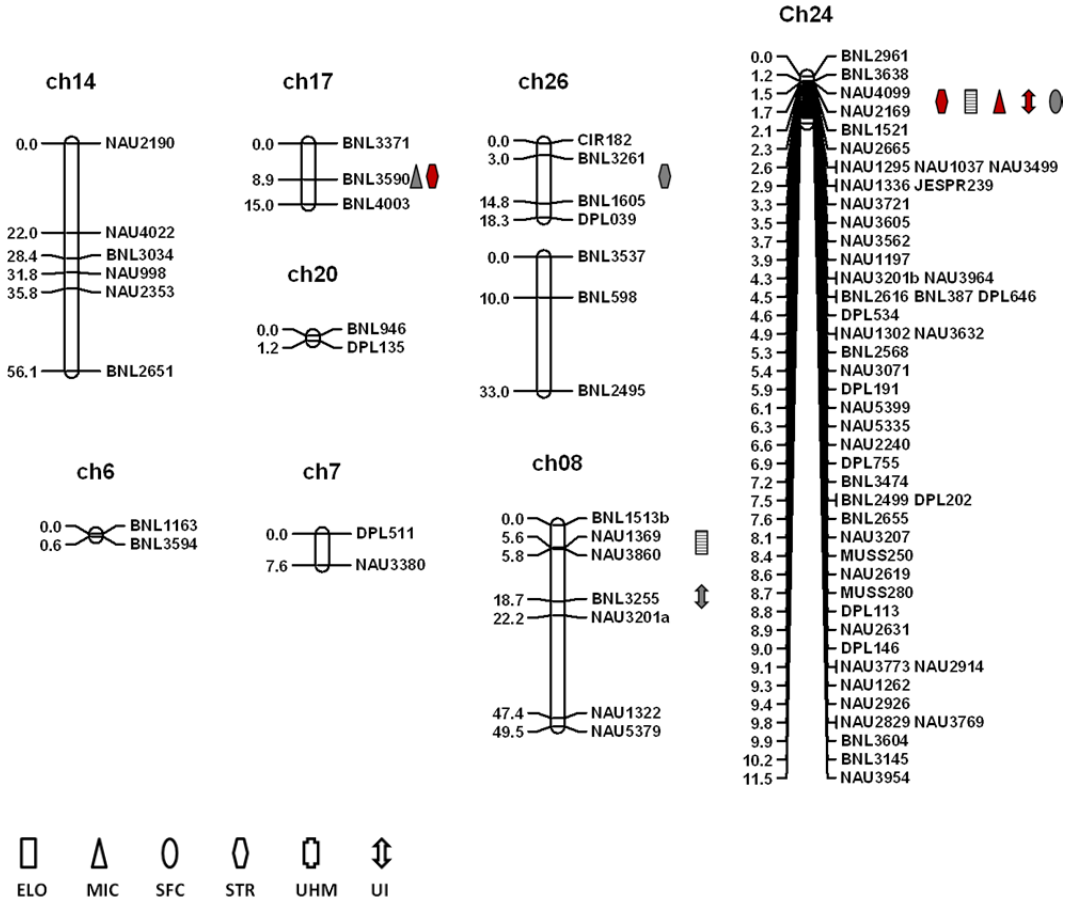


Figure 3.3. QTLs for fiber quality traits identified in Pop542 and Pop883.

Panel A- Linkage map of introgressed segments in SL883
Panel B- Linkage map of introgressed segments in SL542
Panel C- Linkage map of polymorphic regions between Sealand lines and S7235
Legends filled with solid grey are detected in Pop883, with hash are detected in Pop542 and those filled with solid red are common QTLs detected in both Pop542 and Pop883.

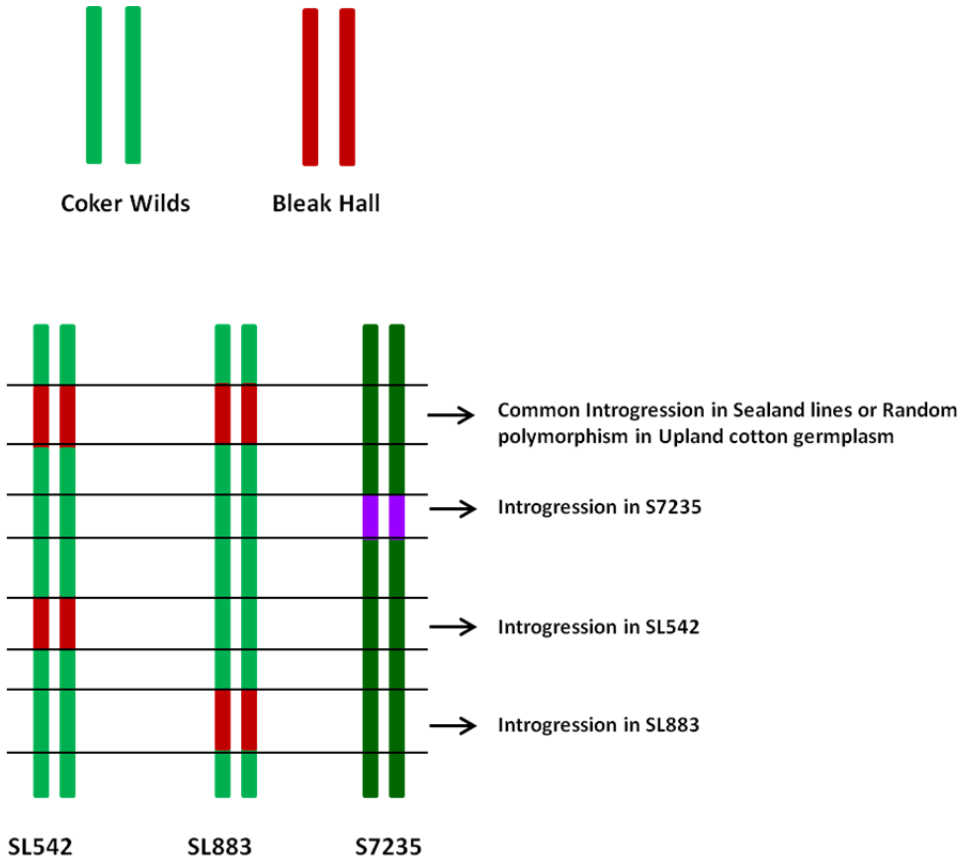


Figure 3.4. Identification of the introgressed segments.

CHAPTER 4

EVALUATING THE EFFICACY OF INTERSPECIFIC INTROGRESSION ON FIBER QUALITY COMPONENTS OF UPLAND COTTON (*GOSSYPIUM HIRSUTUM* L.)[2]

[2] Pawan Kumar, Rippy Singh, Edward L. Lubbers, Xinlian Shen, Andrew H. Paterson, B. Todd Campbell, Donald C. Jones and Peng W. Chee. To be submitted to *Crop Science.*

Abstract

Genetic exploration of obsolete cotton germplasm lines may lead to discovery of novel fiber quality alleles that may not be prevalent in today germplasm. In an earlier study involving F_2 mapping populations, several introgressions from *G. barbadense* were identified in two Sealand germplasm lines which carried alleles that significantly affect fiber quality traits. Efficacy of these introgressed alleles were tested over two generations ($F_{2:3}$ and $F_{2:4}$) and our results show that out of 24 fiber quality QTLs identified, 10 QTLs were consistently detected in both generations and of these 10 QTLs, 8 were also detected in the previous study. Our results suggest that the Sealand lines can be utilized as a source of useful novel genetic variation in Upland cotton and since a majority of the genomic background of Sealand lines is derived from the *G. hirsutum*, it is unlikely that there will be significant segregation distortion when these lines are incorporated in a breeding program. Efficacy of several of the introgressed alleles could be validated over generations, making these alleles as ideal candidate for marker assisted selection.

Introduction

Over the past decade, a fundamental shift has taken place in the market for Upland cotton (*Gossypium hirsutum* L.) produced in the United States, from supplying raw fiber to primarily domestic yarn and textile mills to exporting nearly two-thirds of the lint fiber. While the Upland cotton grown in the U.S. has average uniformity and strength with medium staple length, which is well suited for the domestic open-end rotor based spinning mills, international textile mills are shifting to more advanced ring or air-jet/vortex spinning technologies which require stronger and longer staple cotton fiber with high fiber uniformity index (Felker, 2001). To remain competitive in the international market, both public and private breeding programs would need to emphasize improving the genetic potential of yield and fiber quality of U.S. cotton.

Historically, cotton breeders have utilized only a limited number of elite parental lines for population development and imposed intensive selection to maximize genetic gain in yield, fiber quality, and adaptation to a broad environment (Bowman et al., 1996; May, 2000). This approach to breed improved cotton cultivars has led to a narrow genetic base of the Upland cotton germplasm (Wendel et al., 1992; Brubaker and Wendel, 1994; Iqbal et al., 2001). Analyses of cotton yield data for cultivars released in the last 2 decades indicated that genetic gain in lint yield has declined over the past several years (Meredith, 2006) While there is not a consensus as to the underlying cause of this decline, many researchers recognized that retaining sufficient genetic diversity in the Upland cotton gene pool is essential for its continual genetic improvement (Campbell et al., 2010; Zhang et al., 2011), and have raised concerns regarding erosion in the genetic

base leading to genetic vulnerability of Upland cotton (Campbell et al., 2010; Paterson et al., 2004; Van Esbroeck et al., 1998; Wallace et al., 2009).

Genetic and phenotypic variation in fiber quality is abundant in the other four species that form the allotetraploid *Gossypium* clade (Lacape et al., 2007; Percival and Kohel, 1990) and can be exploited for Upland cotton improvement. In particular, *G. barbadense* (also known as Pima, Sea Island, or Egyptian cotton) is the second most widely cultivated species of cotton and has superior fiber quality traits such as longer, stronger, and finer fibers than Upland cottons. Transferring the superior fiber quality traits of *G. barbadense* into Upland germplasm is attractive but has been an elusive objective due to complications. Typical obstacles faced in introgressive breeding are hybrid breakdown (Stephens, 1949), distorted genetic segregation in early generations (Reinisch et al., 1994) and in advanced generations (Jiang et al., 2000), linkage drag, and a high level of epistasis between fiber quality QTLs and genetic backgrounds harboring different unlinked introgressed alleles (Chee et al., 2005a; Chee et al., 2005b). Therefore, although germplasm lines with stable introgression from *G. barbadense* have been developed after many cycles of crossing, backcrossing and selection (Chee et al., 2005a; Chee et al., 2005b; Draye et al., 2005), the practical utility of these introgression lines in mainstream breeding will depend on the successful 'genetic cleaning' of chromatin with undesirable effects.

A number of obsolete cultivars developed during the boll weevil era displayed fiber quality on par with that of *G. barbadense*, suggesting that these lines may have been developed via introgressive breeding. For example, two sets of Upland cotton lines (Sealand and Earlistaple) with improved fiber quality were developed by the USDA-ARS

Pee Dee program in the mid-1940's and were commercially grown on limited acreage until the 1950's (Culp and Harrell, 1977). During that period, the Pee Dee program focused primarily to improve yield of Sea Island cottons and to improve fiber quality of Upland cotton *via* interspecific hybridization with Sea Island cotton (Culp and Harrell, 1973). The Sealand lines are Upland cottons with remarkable improvement in fiber length and strength but with low yield potential and limited adaptability (Campbell et al., 2011), traits that resemble the *G. barbadense* parent. Therefore, thorough evaluation of the introgressed chromatin in the introgression lines is warranted before these lines are utilized in breeding programs.

Recent molecular analysis confirmed that the Sealand lines contain introgression of *G. barbadense* chromatin (Kumar et al. submitted). In addition, preliminary QTL analysis by using a single year unreplicated phenotypic data suggests that a number of the introgressed chromosomal segments may harbor genes for improved fiber quality. Validation of the marker-trait association between the introgressed regions and fiber quality traits is critical before employing these lines in a MAS program to improve fiber quality. The objectives of this study are to validate the presence of QTLs for fiber quality on the introgressed chromosomal segments of Sealand 542 and Sealand 883 and to evaluate the efficacy of the *G. barbadense* alleles on fiber quality in the Upland genetic background.

Materials and methods

Population development and phenotypic evaluation

Two mapping populations were developed for this study. The population 'Pop542' was derived from the cross of Suyuan 7235 by Sealand 542 (PI 528730) and 'Pop883' was developed by crossing Suyuan 7235 with Sealand 883 (PI 528875). Suyuan 7235 (herein designated as S-7235) is a high fiber strength germplasm line released by the Institute of Industrial Crops, Jangsu Academy of Agricultural Sciences, China, developed through interspecific hybridization of *G. hirsutum* with *G. anomalum* followed by backcrossing to the *G. hirsutum* cultivar "Acala 3080" (PI 529543). Sealand 542 and Sealand 883 (herein designated as SL-542 and SL-883 respectively) are two cultivars developed at the Pee Dee Experiment Station, Florence, South Carolina. Both SL-542 and SL-883 were developed from crossing the Upland cotton line 'Coker Wilds' with the Sea Island cotton (*G. barbadense*) 'Bleak Hall' (PI 608115) followed by backcrossing to Coker Wilds (Bowman et al., 2006a; Culp and Harrell, 1974b).

F_1s of the two crosses were grown in the greenhouse and F_2 seeds were collected from a single plant in both crosses. Over 350 F_2 seeds from each of the two crosses were planted at the William Gibbs Farm, University of Georgia-Tifton Campus in Tifton, Georgia in the summer of 2005. One hundred and seventy-five individuals were randomly tagged in each of the two F_2 populations and seeds were harvested from these tagged individuals. In 2006, 175 $F_{2:3}$ families along with the three parents were planted as progeny rows in a completely randomized design (CRD) with two replications in Tifton, Georgia. All the 175 $F_{2:3}$ families were advanced to the $F_{2:4}$ generation in 2007 where they were again planted together with the three parents in a CRD design with two

replications in Tifton, Georgia. The plots were single row plots, 9 m by 1 m, planted at four seeds per row foot in early May and harvested in early October. Standard production practices were followed in each test. The Pop542 and Pop883 populations are considered genetically distinct backgrounds because, despite having a common pedigree, the two Sealand parents, SL-542 and SL-883, inherited different segments of the donor parent genome (Kumar et al., submitted).

A sample of 25 bolls was harvested from individual F_2 plants, and from each $F_{2:3}$ and $F_{2:4}$ progeny row, ginned on a table-top saw gin, and tested for fiber quality. Fiber quality was measured using the High Volume Instruments (HVI) at the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, NC). The fiber quality measurements comprised Upper Half Mean (UHM) in inches, fiber strength (STR) in kilonewton meter per kilogram (kN m kg−1), where one newton equals 9.81 kg-force, Fiber fineness or micronaire (MIC), percent fiber elongation (ELO), percent short fiber content (SFC), and percent uniformity index (UI).

Genotyping and data analysis

Genomic DNA from each of the F_2 plants and the three parents was extracted following an established procedure (Paterson et al., 1993). A total of 1170 SSR markers, covering all of the 26 homeologous chromosomes, were screened for polymorphism between the mapping parents. Primer sequences of the SSR markers were obtained from the Cotton Marker Database (www.cottonmarker.org) (Blenda et al., 2006) and were commercially synthesized by Eurofins MWG Operon (Huntsville, AL). PCR amplification was performed as described (Chee et al., 2004) and the PCR products were

electrophoretically separated using 10% non-denaturing polyacrylamide gel electrophoresis. The DNA fragments were visualized by staining with silver nitrate (Zhang et al., 2002). The SSR primer pairs were first screened for polymorphism between the parents and then the polymorphic primers were tested on the mapping populations.

Independent linkage maps were constructed using the F_2 data for both Pop542 and Pop883 mapping populations (Kumar et al. submitted). Briefly, linkage groups were constructed using Mapmaker/EXP (Lander et al., 1987) using a LOD score 3.0 and a maximum recombination fraction of 30 cM as grouping thresholds. Recombination units were converted into genetic distances by using the Kosambi mapping function (Kosambi, 1944).

The phenotypic distributions of fiber quality traits and the Pearson correlation coefficients among the traits in $F_{2:3}$ and $_{F2:4}$ of both populations were calculated using PROC UNIVARIATE and PROC CORR procedures of SAS version 9.1 (SAS, 1999), respectively. Heritability of the traits was estimated by parent-offspring regression of 175 $F_{2:3}$-$F_{2:4}$ in each population. QTLs affecting fiber quality traits were analyzed using the Composite Interval Mapping (CIM) function of Windows QTL Cartographer version 2.5 (Wang et al., 2005). For all four phenotypic data sets, LSMEANS were used for the CIM analysis, which was performed using Model 6 (standard model) at 1 cM walk speed on a sliding window of 10 cM with five cofactor markers. Forward and backward stepwise regressions were performed for selecting markers as cofactors and the likelihood ratio statistics (Haley and Knott, 1992) was used to test the significance of locus-trait association. Likelihood ratio (LR) threshold values ($\alpha = 0.05$) for declaring the presence of QTLs were estimated after 1,000 permutations for each trait (Churchill and Doerge,

1994). QTLs were termed putative when the peaks were detected below threshold but above LOD 2.5. Percent phenotypic variation (R^2) was calculated at the peak LR score and the QTL likelihood interval was estimated by marking a one-LOD score drop on either side of the peak. QTL names start with 'q' followed by an abbreviation of the trait name and then the chromosome number. Multiple QTLs on the same chromosome are distinguished by an alphabetical suffix. For brevity, we classified QTLs as 'common' when a QTL is detected in both populations, 'consistent' if a QTL is detected in more than one generation (either F_2, $F_{2:3}$ and $F_{2:4}$) and 'unique' if detected in only one generation of any population.

Results

Phenotypic distribution and correlations

Based on the means of replicated data from 2006 ($F_{2:3}$) and 2007 ($F_{2:4}$), both SL-542 and SL-883 possess fiber quality traits that were different from S-7235. Between the three parents, SL-542 had the greatest fiber elongation and fineness while SL-883 had the greatest fiber length. SL-542 and SL-883 had lower fiber strength and fiber uniformity but greater short fiber content than the S-7235 parent (Table 1). Pairwise comparisons show that except for fiber fineness in $F_{2:4}$ all fiber quality traits were statistically different (LSD, $P>0.05$) between SL-542 and S-7235 in $F_{2:3}$ and $F_{2:4}$. Similarly, statistical differences were observed for all traits between SL-883 and S-7235 in the $F_{2:3}$, however in $F_{2:4}$ only fiber strength and fiber length were statistically different (LSD, $P>0.05$) between SL-883 and S-7235 (Table 1).

Continuous variation for all fiber traits was observed in both $F_{2:3}$ and $F_{2:4}$ generations of Pop-542 and Pop-883 confirming polygenic inheritance of these traits. The population mean was near to the mid-parent value for all traits except for fineness in Pop-542 and for fiber length in Pop-883 where they were skewed towards the lower value parent (S-7235) (Figure 1.) Transgressive segregation was observed for most of the quality traits analyzed.

Fiber quality traits had moderate to high heritability estimates ranging from 0.16 for fiber uniformity to 0.49 for fiber fineness (Table 2). The range of h^2 values estimated from the parent-offspring regressions were in congruence with the earlier estimates of narrow sense heritability (May, 1999). Pearson correlation coefficients between fiber quality traits, calculated for all four data sets, showed strong positive correlation between fiber length, fiber strength, and uniformity index (Table 3) in all three of the pairwise comparisons. Fiber length showed significant negative correlation with fiber elongation, fiber fineness, and short fiber content. Very strong negative correlations (up to r=0.88) were detected between short fiber content and uniformity index.

QTL mapping

The details of linkage maps developed and the chromosomal regions of introgressions in both SL542 and SL883 have been presented elsewhere (Kumar et al. submitted). Briefly, a total of 125 and 120 markers detected polymorphism between the mapping parents of Pop-542 and Pop-883, respectively. Linkage maps of Pop-542 and Pop-883 spanned 530 and 411 cM respectively (Lacape et al., 2009). Twelve

chromosome regions were determined to contain *G. barbadense* introgressions in SL-542 and 10 *G. barbadense* introgressed segments were detected in SL-883.

A total of twenty-four QTLs affecting six fiber quality traits were detected by composite interval mapping (Table 4). Six QTLs were specific to Pop-542, 14 were identified only in Pop-883 while 4 QTLs were identified in both Pop-542 and Pop-883. The QTLs detected for each fiber traits are described below.

Fiber elongation (ELO)

A total of five QTLs were identified (Table 4), two (chr. 21 and 23) in Pop-542 and three (chr. 5, 11, and 24) in Pop-883. The QTLs on chromosomes 11, 21, and 23 were consistent over generations. The percent phenotypic variation (PV%) explained ranged from 2.04 to 13.57% with the genotypic effects ranging from 0.07 to 0.15%. The favorable alleles for all QTLs were contributed by the Sealand parents except for the QTL *qELO-Chr24*, where the favorable allele originated from the S-7235 parent.

Micronaire (MIC)

Four QTLs were identified (Table 4), two (chr. 11 & 17) unique to Pop-883 while one on chromosome 9 was unique to Pop-542. The QTL on chromosome 24 was common to both Pop-542 and Pop-883. It was also consistently detected in both generations of Pop-542 while it was detected as putative QTL at 2.6 LOD in $F_{2:3}$ generation of Pop-883. The PV% explained by these QTLs ranged from 2.44 to 10.03% with genetic effects ranging from 0.1 to 0.22 units. SL-542 and SL-883 contributed favorable alleles for *qMIC-Chr09* and *qMIC-Chr11* respectively, while S-7235 contributed favorable alleles for *qMIC-Chr17* and *qMIC-Chr24*.

Short fiber content (SFC)

Three QTLs affecting short fiber content were detected (Table 4), two (chr. 8 and 17) unique to Pop-883 while a QTL on chromosome 9 was unique to Pop-542. The QTL *qSFC-Chr09* was consistently identified in both $F_{2:3}$ and $F_{2:4}$ datasets of Pop-542. The PV% explained by these QTLs ranged from 4.04 to 15.14% with additive effects ranging from 0.12 to 0.2%. Alleles reducing short fiber content for all QTLs were contributed by the S-7235 parent.

Fiber strength (STR)

A total of five QTLs affecting fiber strength were identified (Table 4). A QTL on chromosome 8 and two QTLs (chr. 16 & 17) were unique to Pop-542 and Pop-883 respectively. The QTL *qSTR-Chr24* was common among Pop-542 and Pop-883 and was consistently identified in $F_{2:3}$ & $F_{2:4}$ generations datasets of both populations. The QTL, *qSTR-Chr26* was identified above the threshold LOD score in Pop-883, however it was putative in Pop-542 (LOD=2.5). The PV% explained by these QTLs ranged from 6.8 to over 40% with additive effects ranging from 0.35 to 1.22 g/tex. Alleles from S-7235 increased fiber strength at all loci except *qSTR-Chr26* in Pop-883 where the alleles from SL-883 improved fiber strength.

Fiber length (UHM)

A total of three QTLs affecting fiber length was identified (Table 4). One on chromosome 9 and two (chr. 16 & 25) unique to Pop-542 and Pop-883 respectively. The QTL *qFL-Chr25* was consistently detected in $F_{2:3}$ & $F_{2:4}$ datasets of Pop-883. The PV% explained by these QTLs ranged from 4.23 to 9.38% with additive effects ranging from

0.014 to 0.036 inches. Favorable alleles for *qFL-Chr09* originated from S-7235 and for *qFL-Chr16* and *qFL-Chr25* these originated from the SL-883 parent.

Uniformity index (UI)

A total of four QTLs were identified (Table 4). Three (chr. 8, 12, & 17) unique to Pop-883 while one (chr. 24) common between the two populations. The PV% explained by the Pop-883 specific QTLs ranged from 3.3 to 7.4% with genetic effects ranging from 0.32 to 0.41% in $F_{2:3}$ generation, however none of these regions showed significant association in $F_{2:4}$ generation. The common QTL, *qUI-Chr24*, was consistently identified in both $F_{2:3}$ and $F_{2:4}$ generation of both mapping populations. QTL *qUI-Chr24* accounted for 2.87 and 3.87% PV in $F_{2:3}$ and $F_{2:4}$ generation, respectively in Pop-542 and 8.51 and 7.53% in $F_{2:3}$ and $F_{2:4}$ generation respectively in Pop-883. The favorable allele was contributed by the S-7235 parent for all the QTLs detected.

Discussion

The use of introgressive breeding in cotton improvement has been an attractive approach since the beginning of modern Upland germplasm development, particularly with the focus on transferring genes from *G. barbadense* to improve the fiber quality (Ware, 1951). Examples of interspecific introgressive breeding in Upland cotton include two well-known successes in improving fiber quality: fiber strength in Acala cotton (Smith and Cothren, 1999; Zhang et al., 2005) and fiber length in Extra Long Staple Cottons including Sealand and Earlistaple lines (Culp and Harrell, 1977). Until recently, the extent of introgression in SL-542 and SL-883 were largely unknown, however the fact that these lines displayed fiber quality on par with that of *G. barbadense* indicated

possible stable introgression of *G. barbadense* chromatin. Kumar et al. (submitted) identified multiple introgressed segments in both SL-542 and SL-883. Interestingly, few of the introgressed *G. barbadense* chromosome segments were shared between the two Sealand lines, reaffirming that the members of the Sealand series may have been independently derived via backcrossing. Herein, we evaluated the introgressed *G. barbadense* chromosome segments found in SL-542 and SL-883 for their potential to provide alleles for improving fiber quality.

A total of 24 fiber quality QTLs were identified in this study, including ten QTLs and one putative QTL that were consistently detected in both $F_{2:3}$ and $F_{2:4}$ generations. This is similar in number to the 25 putative QTLs detected in the previous study using unreplicated F_2 data (Kumar et al. submitted). Interestingly, only 60% (or 15) QTLs were likely to be the same loci detected between the two studies. However, eight of the ten consistent QTLs identified in present study, those that were detected in both $F_{2:3}$ and $F_{2:4}$ generations, were also identified in the F_2 generation. The confirmation of these fiber quality QTLs is an essential step toward utilizing these loci for MAS to improve fiber quality (Chee and Campbell, 2008). However, the two consistent QTLs, *qELO-Chr21* and *qMIC-Chr09,* which were mapped to the introgressed segment in the SL-542 parent were not detected in the F_2 generation. Fiber fineness is greatly influenced by the environment (Chee and Campbell, 2008; May, 2000), which may also account for their missing condition in our previous study. This reinforces the need for QTL validation based on phenotypic data obtained from replicated trials over generations.

An obstacle in implementing MAS to improve fiber quality is due to incomplete understanding of the QTL position as well as its predictive phenotypic effect in different

genetic backgrounds (Chee and Campbell, 2008). Since the two Sealand parents are genetically distinct, Pop-542 and Pop-883 provide not only information on detection but also validation for common QTLs contributed by the S-7235 parent. Therefore, the four common QTLs, *qMIC-Chr24*, *qSTR-Chr26*, *qSTR-Chr24* and *qUI-Chr24*, which were contributed by the S-7235 parent are likely to be authentic. In addition, because of the population design, we detected significant interactions between these common QTLs and the genetic backgrounds. For example, the S-7235 allele for *qMIC-Chr24* accounted for greater genetic variation when present in SL-542 background whereas alleles for the QTLs *qUI-Chr24* and *qSTR-Chr24* explained larger phenotypic variation when present in SL-883 genetic background (Kumar et al., 2012). The QTL *qSTR-Chr26* had variable effects in different genetic backgrounds. For this QTL the SL-542 alleles had increasing effect when detected in Pop542 but when present in Pop883, allele from S-7235 tends to increase fiber strength. Similar epistatic interactions between QTLs and genetic background have been observed in other interspecific populations with alien chromosome segments including cotton (Chee et al., 2005a,b; Zhang et al., 2011) and other crop species (Blanc et al., 2006; Lecomte et al., 2004b; Li et al., 2009; Liao et al., 2001; Sebolt et al., 2000).

Fiber quality alleles from Sealand lines

Both SL-542 and SL-883 contributed positive QTL alleles for important quality traits. These results support previous speculation that many of the *G. barbadense* chromosomal segments in SL-542 and SL-883 were retained as a result of phenotypic selection pressure imposed during backcrossing. For example, of the five fiber elongation

QTLs identified, favorable alleles for four originated from the introgressed *G. barbadense* chromatin. The Sealand parent SL-542 has significantly higher fiber elongation, which could be attributed to the consistent QTLs present in chromosomes 21 and 23. Interestingly, fiber elongation was not a selection criteria during the development of the two Sealand parents (May, 1999). However, it is possible that the improvement in fiber elongation of SL-542 resulted from its correlation with other fiber quality traits such as fiber strength and fineness, which were emphasized in cotton breeding during that time (May, 1999). Other examples of favorable alleles contributed by the Sealand parents include two QTLs for Micronaire, which measures fiber fineness and three QTLs for short fiber content QTLs. Micronaire values are indicators of both fiber maturity and fiber fineness with higher values (>4.5) indicating more mature cotton fibers while lower values (<3.5) indicating immature fibers. Lower micronaire fiber are finer in texture and are sought by the textile mills, however when lower fineness value predominate due to immature fiber, this can cause neps and dye defects (Draye et al., 2005). Finally, two fiber length QTLs were contributed by the SL-883 parent, and the consistent QTL *qFL-Chr25* was detected at very high confidence level (13 LOD) in $F_{2:4}$ generation. Longer fibers require less twist in the roving process during ring spinning and thus are required for the production of finer yarn (May, 1999).

Correlations among fiber quality traits could occur either due to linkage or pleiotropy. In the present study, we found QTLs for fiber length and fiber strength on the introgression segment on chromosome 16 of SL-883 and QTLs for fiber strength and uniformity index on chromosome 24 of S-7235, suggesting the correlation between these traits could be in part due to co-localization of QTLs. Co-localization of QTLs have been

demonstrated in several crops including Brassica (Yan et al., 2009), sunflower (Lexer et al., 2003), and tomato (Saliba-Colombani et al., 2001). In cotton, QTLs for fiber quality have been observed to be co-localized and confined in a QTL rich region across the cotton genome (Chee et al., 2005b; Lacape et al., 2010; Lacape et al., 2005; Rong et al., 2007a; Saranga et al., 2002; Shen et al., 2011). Co-localization of QTLs with desirable effects could simplify their manipulation but becomes problematic when the QTLs are of opposite effects or linked to agronomic traits such as lint yield.

Historical significance of Sealand lines

The Sealand germplasm was within the foundation pool of the modern day Pee Dee germplasm (Bowman et al., 2006a). Pedigree records indicate that Sealand lines SL-542 and SL-7 were intercrossed with Earlistaple, AHA 6-1-4, and Beasley's Triple Hybrid 108 and 171 (*G. arboretum* L. X *G. thurberi* Tod. X *G. hirsutum* L.) to develop a germplasm pool with improved fiber properties (Culp and Harrel, 1980). Thus, SL-542 was instrumental in the development of Pee Dee germplasm. However, there are no documented records that show SL-883 was used at some point during the course of development of Pee Dee germplasm (Bowman et al., 2006a; Culp and Harrell, 1974b). The fiber length and strength QTLs in SL-883 could therefore contribute alleles that are not found in modern Pee Dee germplasm, and the SSRs markers linked to these QTLs could be utilized to assist in transferring these loci to improve fiber quality of elite Upland cotton germplasm.

References

Blanc, G., A. Charcosset, B. Mangin, A. Gallais, L. Moreau. 2006. Connected populations for detecting quantitative trait loci and testing for epistasis: an application in maize. Theor Appl Genet 113:206.

Blenda, A., J. Scheffler, B. Scheffler, M. Palmer, J. Lacape, J. Yu, C. Jesudurai, S. Jung, S. Muthukumar, P. Yellambalase. 2006. CMD: a cotton microsatellite database resource for Gossypium genomics. BMC Genomics 7:132.

Bowman, D.T., O.A. Gutierrez, R.G. Percy, D.S. Calhoun, O.L. May. 2006. Pedigrees of upland and pima cotton cultivars released between 1970 and 2005. Mississippi Agricultural and Forestry Experiment Station Technical Bulletin 1155:57.

Bowman, D.T., O.L. May, D.S. Calhoun. 1996. Genetic base of upland cotton cultivars released between 1970 and 1990. Crop Sci. 36:577-581.

Campbell, B.T., P.W. Chee, E. Lubbers, D.T. Bowman, W.R. Meredith, J. Johnson, D.E. Fraser. 2011. Genetic Improvement of the Pee Dee Cotton Germplasm Collection following Seventy Years of Plant Breeding. Crop Sci. 51:955-968.

Campbell, B.T., S. Saha, R. Percy, J. Frelichowski, J.N. Jenkins, W. Park, C.D. Mayee, V. Gotmare, D. Dessauw, M. Giband, X. Du, Y. Jia, G. Constable, S. Dillon, I.Y. Abdurakhmonov, A. Abdukarimov, S.M. Rizaeva, A. Adullaev, P.a.V. Barroso, J.G. Padua, L.V. Hoffmann, L. Podolnaya. 2010. Status of the Global Cotton Germplasm Resources. Crop Sci. 50:1161-1179.

Chee, P., B. Campbell. 2008. Bridging Classical and Molecular Genetics of Cotton Fiber Quality and Development. Genetics and Genomics of Cotton:283.

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

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

Chee, P.W., J. Rong, D. Williams-Coplin, S.R. Schulze, A.H. Paterson. 2004. EST derived PCR-based markers for functional gene homologues in cotton. Genome 47:449-462.

Churchill, G.A., R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963.

Culp, T., D. Harrel. 1980. Registration of extra-long staple cotton germplasm. Crop Sci. 20:290.

Culp, T.W., D.C. Harrell. 1973. Breeding Methods for Improving Yield and Fiber Quality of Upland Cotton (Gossypium-Hirsutum L). Crop Sci. 13:686-689.

Culp, T.W., D.C. Harrell. 1974. Breeding quality cotton at the Pee Dee Experiment Station Florence, SC. USDA-ARS. Pubi. ARS-S-30 12.

Culp, T.W., D.C. Harrell. 1977. Yield and fiber quality improvements in Upland cotton (Gossypium hirsutum L.). Technical Bulletin-South Carolina Agricultural Experiment Station (USA).

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

Felker, G.S. (2001) Fiber Quality and New Spinning Technology, in: P. D. a. D. Richter. (Ed.), Beltwide Cotton Conferences, Anaheim, CA. 9–13 Jan. 2001, National Cotton Council of America, Memphis, TN., . pp. 5-7.

Haley, C.S., S.A. Knott. 1992. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. Heredity 69:315-324.

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

Kosambi, D. 1944. The estimation of map distances from recombination values. Ann Eugen 12:172-175.

Kumar, P., R. Singh, E.L. Lubbers, X. Shen, A.H. Paterson, B.T. Campbell, D.C. Jones, P.W. Chee. 2012. Mapping and Validation of Fiber Strength Quantitative Trait Loci on Chromosome 24 in Upland Cotton. Crop Sci. 52:1115-1122.

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

Lacape, J.M., J. Jacobs, T. Arioli, R. Derijcker, N. Forestier-Chiron, D. Llewellyn, J. Jean, E. Thomas, C. Viot. 2009. A new interspecific, Gossypium hirsutum x G.

barbadense, RIL population: towards a unified consensus linkage map of tetraploid cotton. Theor. Appl. Genet. 119:281-292.

Lacape, J.M., D. Llewellyn, J. Jacobs, T. Arioli, D. Becker, S. Calhoun, Y. Al-Ghazi, S. Liu, O. Palaï, S. Georges. 2010. Meta-analysis of cotton fiber quality QTLs across diverse environments in a Gossypium hirsutum x G. barbadense RIL population. BMC Plant Biol. 10:132.

Lacape, J.M., T.B. Nguyen, B. Courtois, J.L. Belot, M. Giband, J.P. Gourlot, G. Gawryziak, S. Roques, B. Hau. 2005. QTL analysis of cotton fiber quality using multiple Gossypium hirsutum x Gossypium barbadense backcross generations. Crop Sci. 45:123-140.

Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln, L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174.

Lecomte, L., P. Duffe, M. Buret, B. Servin, F. Hospital, M. Causse. 2004. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theor. Appl. Genet. 109:658–668.

Lexer, C., M. Welch, J. Durphy, L. Rieseberg. 2003. Natural selection for salt tolerance quantitative trait loci (QTLs) in wild sunflower hybrids: Implications for the origin of Helianthus paradoxus, a diploid hybrid species. Mol. Ecol. 12:1225-1235.

Li, Y., X. Li, J. Li, J. Fu, Y. Wang, M. Wei. 2009. Dent corn genetic background influences QTL detection for grain yield and yield components in high-oil maize. Euphytica 169:273-284.

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

May, O. 1999 Genetic variation in fiber quality, Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing, Food Products Press, New York, NY. pp. 183-229.

May, O.L. 2000. Genetic variation in fiber quality. Cotton fibers–developmental biology, quality improvement, and textile processing. Food Products Press, New York:183–229.

Meredith, W.R., Jr. 2006. Registration of MD 15 Upland Cotton Germplasm. Crop Sci. 46:2722-2723.

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

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

Percival, A.E., R.J. Kohel. 1990. Distribution, collection, and evaluation of Gossypium. Advances in Agronomy 44:225-256.

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

Saliba-Colombani, V., M. Causse, D. Langlois, J. Philouze, M. Buret. 2001. Genetic analysis of organoleptic quality in fresh market tomato. 1. Mapping QTLs for physical and chemical traits. Theor. Appl. Genet. 102:259-272.

Saranga, Y., M. Menz, C. Jiang, R. Wright, D. Yakir, A. Paterson. 2002. Genetic mapping implicates osmotic potential as a major component of crop adaptation to arid conditions. Genome Res 11:1988-1995.

Sas. (1999) The SAS system for Windows: Version 8, SAS Institute, Inc. Cary, NC.

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

Shen, X., Z. Cao, R. Singh, E.L. Lubbers, P. Xu, C.W. Smith, A.H. Paterson, P.W. Chee. 2011. Efficacy of qFL-chr1, a Quantitative Trait Locus for Fiber Length in Cotton (Gossypium spp.). Crop Sci. 51:2005.

Smith, C.W., J.T. Cothren. (1999) Cotton: origin, history, technology and production John Wiley & Sons, New York.

Stephens, S.G. 1949. The Cytogenetics of Speciation in Gossypium .1. Selective Elimination of the Donor Parent Genotype in Interspecific Backcrosses. Genetics 34:627-637.

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

Wallace, T., D. Bowman, B. Campbell, P. Chee, O. Gutierrez, R. Kohel, J. Mccarty, G. Myers, R. Percy, F. Robinson. 2009. Status of the USA cotton germplasm collection and crop vulnerability. Genet. Resour. Crop Evol. 56:507-532.

Wang, S., C. Basten, Z. Zeng. 2005. Windows QTL cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Ware, J.O. (1951) Origin, rise and development of American upland cotton varieties and their status at present University of Arkansas College of Agriculture, Agricultural Experiment Station, Fayetteville, AR

Yan, X., J. Li, F. Fu, M. Jin, L. Chen, L. Liu. 2009. Co-location of seed oil content, seed hull content and seed coat color QTL in three different environments in Brassica napus L. Euphytica 170:355-364.

Zhang, J., W. Guo, T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L. x Gossypium barbadense L.) with a haploid population. Theor. Appl. Genet. 105:1166-1174.

Zhang, J.F., Y. Lu, H. Adragna, E. Hughs. 2005. Genetic improvement of New Mexico Acala cotton germplasm and their genetic diversity. Crop Sci. 45:2363-2373.

Zhang, Z., J. Rong, V.N. Waghmare, P.W. Chee, O.L. May, R.J. Wright, J.R. Gannaway, A.H. Paterson. 2011. QTL alleles for improved fiber quality from a wild Hawaiian cotton, Gossypium tomentosum. Theor. Appl. Genet.:1-14.

Table 4.1. Mean fiber quality of mapping parents planted in 2006 and 2007

Trait	Year	P1 (S7235)	P2 (SL542)	P3 (SL883)	P1-P2[a]	P1-P3[b]
Fiber elongation	2006	4.17	5.62	3.78	-1.45**	0.38**
	2007	4.04	5.44	4.08	-1.4**	-0.04
Micronaire	2006	4.22	4.79	3.27	-0.57**	0.95**
	2007	3.95	4.13	3.78	-0.17	0.18
Short fiber content	2006	6.78	8.03	8.28	-1.26**	-1.51**
	2007	6.77	7.60	7.24	-0.83**	0.47
Fiber strength	2006	37.68	27.68	30.78	9.99**	6.89**
	2007	35.63	28.18	33.44	7.45**	2.19**
Fiber length	2006	1.28	1.10	1.34	0.18**	-0.06**
	2007	1.26	1.11	1.33	0.14**	-0.07**
Fiber uniformity	2006	85.71	82.83	81.48	2.88**	4.23**
	2007	84.35	82.60	83.76	1.75**	0.59

[a] Difference in fiber quality traits between S7235 and SL542
[b] Difference in fiber quality traits between S7235 and SL883

Table 4.2. Parent-offspring heritability estimates of fiber quality traits for SL542 and SL883

Trait	Population	h^2 estimate
Fiber elongation	SL542	0.38
	SL883	0.45
Micronaire	SL542	0.50
	SL883	0.50
Short fiber content	SL542	0.38
	SL883	0.35
Fiber strength	SL542	0.41
	SL883	0.49
Fiber length	SL542	0.50
	SL883	0.45
Fiber uniformity	SL542	0.16
	SL883	0.30

Table 4.3. Pearson correlation coefficients between fiber quality traits in two generations of Pop542 and Pop883 (shown in red)

		Micronaire	Short fiber content	Fiber strength	Fiber uniformity	Fiber Length
Fiber elongation	$F_{2:3}$	0.36**\0.38**	-0.09\0.04	-0.08\0.13	0.06\0.23*	-0.29**\-0.30**
	$F_{2:4}$	0.15\0.40**	0.09\-0.32**	-0.18\0.30**	-0.03\0.24**	-0.30**\-0.13
Micronaire	$F_{2:3}$		0.03\0.20	0.04\0.18	0.01\0.20*	-0.41**\-0.37**
	$F_{2:4}$		-0.10\-0.47**	0.03\0.42**	0.20**\0.48**	-0.29**\-0.01
Short fiber content	$F_{2:3}$			-0.54**\-0.47**	-0.87**\-0.46**	-0.57**\-0.85**
	$F_{2:4}$			-0.46**\-0.58**	-0.78**\-0.88**	-0.71**\-0.41**
Fiber strength	$F_{2:3}$				0.57**\0.45**	0.52**\0.35**
	$F_{2:4}$				0.48**\0.56**	0.41**\0.33**
Fiber uniformity	$F_{2:3}$					0.56**\0.23**
	$F_{2:4}$					0.44**\0.29**

* and ** represent significance with *P*-values of 0.01 and 0.001, respectively

Table 4.4. Summary of fiber quality QTLs identified in F2:3 and F2:4 generation of the two mapping populations

Trait	QTL Name	Population	LOD		Add		R^2 (PV%)		+ve Allele
			$F_{2:3}$	$F_{2:4}$	$F_{2:3}$	$F_{2:4}$	$F_{2:3}$	$F_{2:4}$	
Elongation (%)									
	qELO-Chr05	Pop883	3.8		-0.09		2.04		SL883
	qELO-Chr11	Pop883	3.7	4.3	-0.10	-0.13	7.94	13.57	SL883
	qELO-Chr21	Pop542	3.8	3.1	-0.15	-0.13	5.78	9.10	SL542
	qELO-Chr23	Pop542	3.2	3.7	-0.12	-0.07	7.33	5.27	SL542
	qELO-Chr24	Pop883	3.1		0.11		6.59		S7235
Micronaire									
	qMIC-Chr09	Pop542	3.5	2.6	-0.12	-0.11	10.03	4.73	SL542
	qMIC-Chr11	Pop883		3.0		-0.12		4.22	SL883
	qMIC-Chr17	Pop883	9.8	1.7	0.22	0.10	7.23	2.44	S7235
	qMIC-Chr24	Pop542	3.4	7.4	0.15	0.17	9.12	9.30	S7235
		Pop883	2.6	3.5	0.14	0.17	6.43	5.41	S7235
Short fiber content (%)									
	qSFC-Chr08	Pop883	3.2		-0.12		4.04		SL883
	qSFC-Chr09	Pop542	4.4	3.6	-0.26	-0.16	8.11	14.05	SL542
	qSFC-Chr17	Pop883	6.5		-0.20		15.14		SL883
Fiber strength (kNm/Kg)									
	qSTR-Chr08	Pop542	3.3		-0.73		13.50		SL542
	qSTR-Chr26	Pop542	2.5	2.1	-0.45	-0.46	7.21	8.30	SL542
		Pop883	3.7		0.45		7.63		S7235

qSTR-Chr16	Pop883	3.8		-0.35		6.77		SL883
qSTR-Chr17	Pop883	5.3	3.7	0.61	0.60	10.3	6.44	S7235
qSTR-Chr24	Pop542	15.0	15.1	10.63	9.35	24.85	18.15	S7235
	Pop883	15.3	17.4	11.97	11.99	40.1	36.36	S7235
Fiber length (In)								
qFL-Chr09	Pop542	3.4		0.014		5.14		S7235
qFL-Chr16	Pop883		3.5		-0.025		6.50	SL883
qFL-Chr25	Pop883	5.7	12.6	-0.036	-0.030	4.23	9.38	SL883
Fiber uniformity (%)								
qUI-Chr08	Pop883	2.6		0.32		3.27		S7235
qUI-Chr12	Pop883	3.7		0.41		7.40		S7235
qUI-Chr17	Pop883	3.2		0.40		6.58		S7235
qUI-Chr24	Pop542	3.0	3.2	0.31	0.11	2.87	3.87	S7235
	Pop883	3.1	3.5	0.43	0.23	8.51	7.51	S7235

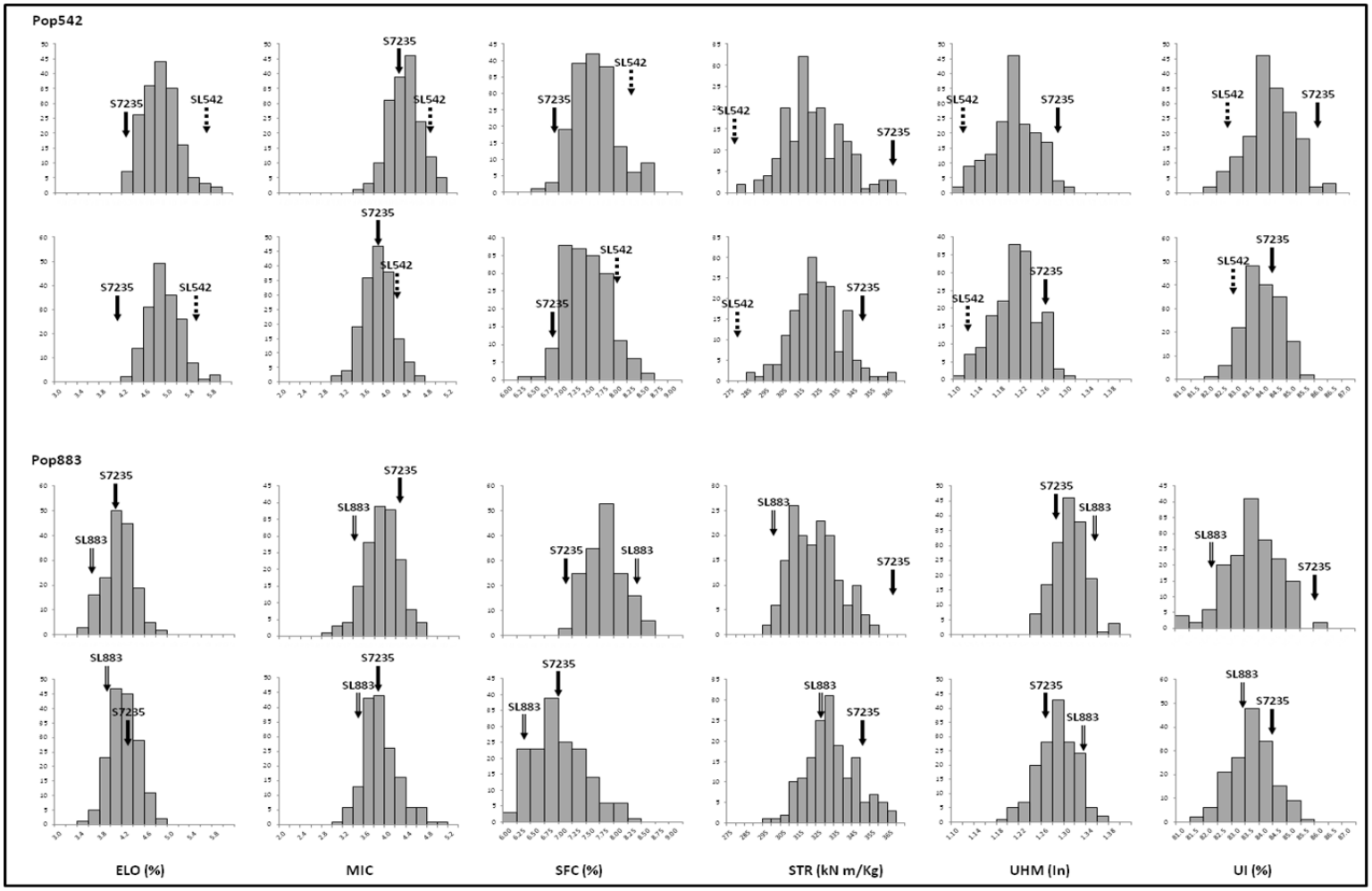


Figure 4.1. Distribution of fiber quality $F_{2:3}$ and $F_{2:4}$ generation of the two mapping populations

CHAPTER 5

MAPPING AND VALIDATION OF FIBER STRENGTH QTLS ON CHROMOSOME 24 IN UPLAND COTTON (*GOSSYPIUM HIRSUTUM* L.)[3]

[3] Pawan Kumar, Rippy Singh, Edward L. Lubbers, Xinlian Shen, Andrew H. Paterson, B. Todd Campbell, Donald C. Jones, and Peng W. Chee. Submitted to *Crop Sci* 06/26/12

Abstract

A major fiber strength QTL has been identified on chromosome 24 in the Chinese germplasm line "Suyuan 7235", however the effects of this QTL have not been tested in different genetic backgrounds. In this study, we confirmed the effects of this QTL by crossing Suyuan 7235 with two U.S. germplasm lines with different fiber strength. This QTL was consistently expressed in both populations, and over generations and years. The Suyuan 7235 allele explained up to 40% of the total phenotypic variation and accounted for an increase of up to 11.7 kN m kg^{-1}. The effects on fiber strength appear to be greater in Pop-883 than in Pop-542 despite the SL-883 parent having stronger fiber than the SL-542 parent. DNA fingerprinting on a collection of elite cotton lines indicated that this QTL is not present in a survey of the elite U.S. public germplasm. These results indicate that this fiber strength QTL could significantly improve the economic value of Upland cottons in the U.S. The identification of 27 novel markers tightly-linked in this region adds additional tools to allow this QTL to be more efficiently deployed in breeding cultivars with improved fiber strength.

Introduction

Historically, the U.S. has produced medium grade fiber Upland cotton (*Gossypium hirsutum* L.) primarily for the consumption of domestic mills. Early in the twenty-first century, technological advances in spinning technology (Felker, 2001) created demand for cotton with greater fiber quality than the standards established for the domestic market. Yet, studies have shown that fiber quality of Upland cotton in the U.S. has declined after 2000 (Bowman and Gutiérrez, 2003) as the U.S. cotton industry shifted from a domestic based market to an export oriented market with nearly two-thirds of the cotton fiber produced in the U.S. now sold on the world market (National Cotton Council 2011). Since the international textile mills impose a more stringent demand for fiber quality than domestic mills, further improvements in fiber quality are needed in order for U.S. cotton to remain competitive with other cotton producing countries.

Among several physical properties that collectively describe cotton fiber quality, fiber strength is one of the main quality traits that can greatly influence the yarn manufacturing process. Stronger fibers can effectively withstand mechanical impacts of the yarn spinning process better than weaker fiber (Deussen, 1992; Meredith et al., 1991), therefore it can be spun at a greater speed. In addition, yarn strength is directly influenced by fiber strength, thus, raw cotton with higher tenacity generally produces more durable fabrics and also maintains cotton's natural qualities after chemical processing of the fabric (Deussen, 1992; May, 1999).

Classical genetic studies have shown the polygenic nature of fiber strength (summarized by May, 1999), therefore improving fiber strength will involve stacking multiple favorable alleles into one genetic background. Although high heritability and

additive gene action make phenotypic selection effective for fiber strength improvement (May, 1999), the general inverse correlation between yield and fiber strength hinders simultaneous improvement of both traits (Culp and Green, 1992; Culp and Harrell, 1979). As a result, elite high yielding varieties often have average fiber strength even though some germplasm accessions display excellent fiber strength (Bowman and Gutiérrez, 2003).

More than 80 QTLs for fiber strength have been identified from 17 different QTL mapping studies (Summarized by Chee and Campbell, 2008). The number of QTLs that were detected in each study ranged from one (Zhang et al., 2003) to 21 (Paterson et al., 2003) and explained from 2.4% to 53.8% of the total phenotypic variations. Therefore, it may now be possible to stack multiple favorable alleles conferring improved fiber strength into a single genotype. However, a majority of the QTLs were identified in early generation interspecific hybrid populations, with the favorable alleles originating from *Gossypium barbadense* L. (also known as Pima cotton, Sea Island cotton or Egyptian cotton) complicating their manipulation in Upland cotton genetic backgrounds. An exception to this is a major QTL for fiber strength located on Chromosome 24 in the germplasm line Suyuan 7235 (Qian et al., 1992). This QTL was identified using F_2, $F_{2:3}$, backcross and recombinant inbred mapping populations derived from the cross of line Suyuan 7235 by TM-1 (Shen et al., 2007; Shen et al., 2005; Shen et al., 2006). This fiber strength QTL from Suyuan 7235 may be a good candidate for improving the fiber strength of the U.S. Upland cotton germplasm.

There are several challenges that need to be addressed if a breeding program desires to integrate MAS to improve fiber quality traits. For example, it is important to

determine if a QTL will produce similar phenotypic effects in multiple genetic backgrounds and/or environments. Although previous studies have shown that the fiber strength QTL in Suyuan 7235 can be detected over different generations and in different environments in China, this QTL has only been tested in the TM-1 genetic background (Chen et al., 2009; Guo et al., 2003; Shen et al., 2007; Shen et al., 2005; Shen et al., 2006; Zhang et al., 2003). It has been shown in rice (Liao et al., 2001; Steele et al., 2006), maize (Li et al., 2009), tomato (Lecomte et al., 2004), and cotton (Chee et al., 2005a; Chee et al., 2005b) that QTLs detected in genetic populations chosen to maximize phenotypic differences may be less effective in other genetic backgrounds due to interaction with other loci or epistasis (Holland, 2007). The objectives of this study are to validate the presence of a QTL for fiber strength on chromosome 24 of Suyuan 7235 and to determine the efficacy of this QTL in two cotton cultivars with different fiber strength.

Materials and methods

Population development and phenotyping

We developed two genetic populations by crossing "Suyuan 7235" to "Sealand 542" (PI 528730) and "Sealand 883" (PI 528875), designating as Pop-542 and Pop-883, respectively. The germplasm line Suyuan 7235 (designated from here on as S-7235) was released by the Institute of industrial crops, Jangsu academy of agricultural sciences, China, because it possessed superior fiber strength (Qian et al., 1992). S-7235 was developed by crossing *Gossypium anomalum* with *G. hirsutum*, and then backcrossing the progeny to the *G. hirsutum* cultivar "Acala 3080" (PI 529543) (Qian et al., 1992). The cultivars Sealand 542 and Sealand 883 (designated from here on as SL-542 and SL-

883, respectively) were developed at the Pee Dee experiment station, Florence, South Carolina, and were selected as parents in this study because they are genetically similar but had different fiber strength (Figure 1). The Sealand cultivars were developed from crossing the Upland cotton line 'Coker Wilds' with the Sea Island cotton (*G. barbadense*) 'Bleak Hall' (PI 608115), followed by backcrossing to Coker Wilds (Bowman et al., 2006b; Culp and Harrell, 1974b).

The F_1 hybrids for Pop-542 and Pop-883 were grown in the greenhouse and F_2 seeds were collected from a single F_1 plant for each cross combination. The F_2 populations, Pop-542 and Pop-883, comprising 175 individuals each, were planted at William Gibbs Farm, in Tifton, Georgia in the summer of 2005. Seed cotton from the individual F_2 plants was hand-picked, ginned on a table-top saw gin and tested for fiber quality. In 2006, $F_{2:3}$ families along with the three parents were randomly selected from each of the F_2 populations and were planted as progeny rows in a completely randomized design (CRD) with two replications in Tifton, Georgia. All 175 $F_{2;3}$ families were advanced to the $F_{2:4}$ generation in 2007 where they were again planted together with the three parents in a CRD with two replications in Tifton, Georgia. The plots were single row plots, 9 meters by 1 meter, planted at 4 seeds/row foot in early May and harvested in early October. Standard production practices were followed in each test. Twenty-five boll samples were harvested from each $F_{2:3}$ and $F_{2:4}$ progeny row, ginned on a table-top saw gin and tested for fiber quality. Fiber strength was measured using the High Volume Instruments (HVI) at the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, NC). HVI measures of fiber strength are reported as kilonewton meter per kilogram (kN m kg^{-1}), where one Newton equals 9.81 kg-force.

Molecular marker analysis

Genomic DNA from each of the F_2 plants and parents was extracted following an established procedure (Paterson et al., 1993). Quantity and quality of extracted DNA was checked on a 0.8% agarose gel before diluting for PCR amplification. Sequences of Simple Sequence Repeats (SSR) primers were downloaded from the Cotton Marker Database (http://www.cottonmarker.org) (Blenda et al., 2006) and were commercially synthesized by Operon (Eurofins MWG Operon, Huntsville, AL). A total of 68 SSR markers mapping either to chromosome 24 or its homeolog chromosome 8, were selected from the genetic maps available at Cotton Marker Database (Blenda et al., 2006). PCR amplification was performed as described (Chee et al., 2004) and the PCR products were electrophoretically separated using 10% non-denaturing Polyacrylamide gel electrophoresis. The DNA fragments were visualized by staining with silver nitrate as described (Zhang et al., 2002). The SSR primer pairs were first screened for polymorphism between the parents and then the polymorphic primers were tested on the mapping populations.

Data analysis

The phenotypic distribution of fiber strength in both mapping populations was calculated with SAS version 9.1 (SAS Institute Inc, 1989). Linkage maps of both populations were constructed using the Mapmaker/EXP (Lander et al., 1987) software. Logarithms of odds (LOD) score of 5.0 and maximum recombination fraction of 30 cM were set as grouping thresholds. Recombination units were converted into genetic distances by using the Kosambi mapping function (Kosambi, 1944). Detection of QTL

and estimation of genetic parameters were performed with Composite Interval Mapping (CIM) function of the software WinQTL Cartographer version 2.5 (Wang et al., 2005). The phenotypic data from the F_2, $F_{2:3}$ and $F_{2:4}$ generations in each population were analyzed separately. However, since the error mean squares for fiber strength from the $F_{2:3}$ and $F_{2:4}$ were homogeneous in the Pop-883 dataset (Levene, 1960), marker-trait association was also analyzed using the pooled data in this population. One thousand permutations at the 0.05 significance level was used to calculate the appropriate likelihood ratio (LR) threshold values for each phenotypic dataset. Forward regression method with the walk speed of 0.5 cM was used for scanning the region for QTLs.

Results

Population biometrical parameters

The three parents had significantly different fiber strength in 2006 and 2007 (Figure 1). S-7235 had mean fiber strength of 363.8 kN m kg^{-1} in 2006 and 350.1 kN m kg^{-1} in 2007. Among the Sealand parents, SL-883 had higher mean fiber strength of 301.1 and 324.6 kN m kg^{-1} in 2006 and 2007 respectively, than SL-542, which had fiber strength of 271.5 kN m kg^{-1} in 2006 and 277.5 kN m kg^{-1} in 2007. The fiber strength of both Pop-542 and Pop-883 populations were normally distributed in F_2, $F_{2:3}$ and $F_{2:4}$ generation (Figure 1). In both populations, transgressive segregation was observed in all generations; however, a greater number of transgressive segregants were detected in Pop-883. The range of phenotypes and population mean differed among populations, with Pop-883 displaying a higher population mean in all generations (Figure 1).

Confirmation of fiber strength QTL on chromosome 24

Fifty-two of the 68 SSR markers were polymorphic between S-7235 and the Sealand parents, and were genotyped on both F_2 populations. We were able to establish linkage for 49 loci in Pop-542 and 50 loci in Pop-883. The SSR marker NAU3954 was polymorphic only in Pop-883 and therefore mapped only in this population. In both F_2 populations, with the exception of markers BNL3860 and NAU1369, all loci were mapped to a single linkage group at a LOD score of 7.0. The two unlinked markers were not included in further analysis. The linkage map of Pop-542 had an overall length of 8.7 cM with an average distance between markers of 0.18 cM. Similarly, the genetic linkage map of Pop-883 covered a distance of 11.0 cM with an average distance between markers of 0.22 cM. The maps developed from the two populations were nearly identical. However, the map order was not resolved for number of markers due to the population size of only 175 individuals. Therefore, the two maps were compared to prior published maps (Shen et al., 2007) and a consensus map was developed using individuals from both populations (Figure 2). In comparison to the complete linkage maps of chromosome 24 available at the Cotton Marker Database, the most probable localization of this region is on the second quartile of chromosome 24, covering approximately 11cM or 9.8% of the total recombinational length of the chromosome (Chen et al., 2009; Guo et al., 2007).

Composite interval mapping detected a major fiber strength QTL between markers BNL2961 and BNL3145 (LOD>12.5) in both populations (Figure 2). The percent of phenotypic variance (PV%) explained by this QTL differed across populations and between generations within a population, ranging from 18.2% to 22% in Pop-542 and from 30.9% to 40.1% in Pop-883 (Table 1). In Pop-883, association analysis performed

on the pooled dataset across $F_{2:3}$ and $F_{2:4}$ generations detected this QTL at a higher LOD of 16.6 explaining 44.5% PV. In all generations of both populations, the allele from S-7235 conferred positive additive effects, increasing fiber strength from 9.35 to 19.02 kN m kg^{-1} in Pop-542 and from 11.97 to 22.82 kN m kg^{-1}in Pop-883 (Table 1). It is interesting to note that while the efficacy of each QTL appears to be consistent across generation, the effects on fiber strength appear to be greater in Pop-883 than in Pop-542 despite the SL-883 parent having stronger fiber than the SL-542 parent.

Based on all the markers mapped within the BNL2621 and BNL3145 region, which flanked the fiber strength QTL, we identified 20 lines homozygous for the S-7235 allele, 74 lines heterozygous, and 31 homozygous for the SL-542 allele in Pop-542, and 17 lines homozygous for the S-7235 allele, 79 lines heterozygous, and 30 homozygous for the SL-883 allele in Pop-883. Figure 3 shows the mean fiber strength of the three genotypic classes. In both populations, a significant difference in fiber strength (LSD at $P<0.05$) was observed between the three genotypic classes. In population Pop-542, the difference in mean fiber strength between the two homozygous genotypic classes ranged from 18.33 to 36.28 kN m kg^{-1}. Again, the difference in mean fiber strength was consistently larger in Pop-883, ranging from 21.96 to 42.95 kN m kg^{-1}. These results strongly support the QTL analysis that the substitution of an S-7235 allele in this QTL region will result in significant improvement in fiber strength.

Discussion

Previous QTL mapping studies have reported the presence of a major QTL for fiber strength on chromosome 24 (previously referred as LGD03) in the high fiber

strength germplasm line S-7235 (Guo et al., 2003; Shen et al., 2007; Shen et al., 2005; Yuan et al., 2001; Zhang et al., 2003). In the present study, we confirmed the association of chromosome 24 on fiber strength in the germplasm line S-7235. In both mapping populations, the markers associated with fiber strength indicated that the QTL we identified was the same locus as previously reported by Shen et al. (2007). By utilizing three overlapping recombinant inbred lines develop from the same S-7235 by TM-1 cross, Chen et al. (2009) showed that this fiber strength QTL region may harbor as many as five distinct QTLs clustered within a 14 cM region. A close examination of our association analysis revealed three sub regions separated by 'dips' of more than 2 LOD value from the likelihood peaks in both populations (Figure 2), suggesting the presence of more than one QTL within this interval of chromosome 24 (Lynch and Walsh, 1998). We note that the populations studied here are too few and too small to provide the genetic resolution necessary to rigorously test for multiple QTLs in such a small genetic interval. However, the current data lend credence to prior results that this region of chromosome 24 may contain more than one QTL for fiber strength (Chen et al. 2009).

Numerous studies have shown interactions between QTLs and genetic backgrounds, and suggested the need to test the efficacy of a QTL in multiple genetic backgrounds before its utilization in MAS (Lecomte et al., 2004a; Li et al., 2009; Sebolt et al., 2000) For example, Lecomte et al. (2004) introgressed five QTLs controlling fruit quality into three tomato lines and found that the breeding efficiency of each QTL varied according to the recipient parent. Sebolt et al. (2000) introgressed a QTL conferring high seed protein concentration into three soybean lines with varying levels of protein content and detected the effects in only two of the three genetic backgrounds; the effect of this

QTL was not detected in the line having the highest seed protein concentration. In this study, we selected SL-883 and SL-542 as parents because they are genetically distinct from TM-1, the sole genetic background in which this fiber strength QTL has been tested (Guo et al., 2003; Shen et al., 2007; Shen et al., 2005; Yuan et al., 2001; Zhang et al., 2003). In addition, SL-883 and SL-542 differ significantly in fiber strength, which indicates that they contain different sets of alleles for this phenotype, offering the opportunity to validate both the efficacy and marker association of the chromosome 24 region from S-7235 with fiber strength. Both SL-883 and SL-542 cultivars were marketed commercially as extra-long staple (ELS) Upland varieties in late 1940s (Culp and Harrell, 1974b). SL-542 was more widely planted due to higher yield potential, with approximately 1000 acres grown in South Carolina, Georgia and Florida in 1948 (Jenkins, 1948).

In an earlier study, Shen et al. (2007) indicated that the genetic effects of this QTL in TM-1 background improved fiber strength by 6.08-12.16 kN m kg^{-1}. Using RILs derived from the same TM-1 genetic background, Chen et al., (2009) later reported that they observed a similar effect for this QTL region, accounting for an additive increase in fiber strength from 7.35 to 13.63 kN m kg^{-1}. Our results indicated that the effects of this QTL region in both the SL-883 and SL-542 cross combinations were similar to that observed in the TM-1 background as reported by Shen et al. (2007) and Chen et al. (2009). However, the efficacy of this QTL region appears to be slightly greater in Pop-883 than in Pop-542 in all tested generations despite the SL-883 parent having stronger fiber than the SL-542 parent. For example, the genetic effects based on replicated data in the $F_{2:3}$ and $F_{2:4}$ generations indicated that the S-7235 alleles increased fiber strength by

9.35-10.63 kN m kg^{-1} in Pop-542 compared to 11.98 kN m kg^{-1} in Pop-883 (Table 1). Further, the mean fiber strength for progenies homozygous for the S-7235 alleles at the QTL region ranged from 18.33 to 36.28 kN m kg^{-1} in Pop-542, but the range was relatively higher from 21.96 to 42.95 kN m kg^{-1} in Pop-883 (Figure 3). Because the differences we observed in the two genetic backgrounds were small, further evaluation of the QTL region is needed to confirm the interactions with genetic backgrounds.

The potential value of this QTL region for improving U.S. Upland cotton is significant because cotton fiber with strength above 304 kN m kg^{-1} receives a premium price (www.ams.usda.gov/cotton), but Upland cotton fiber produced in the U.S. seldom exceeds this threshold. For example, none of the high yielding varieties from 2009 state-wide multi-location cotton varietal trials in Georgia (www.swvt.uga.edu) or Texas (http://cottonimprovementlab.tamu.edu), the two top cotton producing states in the USA, had fiber strength higher than 294 kN m kg^{-1}. Introgression of this QTL cluster into a suitable Upland cotton genetic background could potentially improve fiber strength by up to 20 kN m kg^{-1} when present in a suitable background and therefore may result in an improved cultivar with sufficient fiber strength to receive a premium price in both domestic and international markets.

The possibility that this region of chromosome 24 from S-7235 may harbor multiple QTLs for fiber strength suggests that the most feasible approach for marker assisted backcrossing would be to introgress the QTL region as a single unit. The mapping of 27 additional markers in this QTL region, expanding it from 35 loci (Shen et al., 2007) to a total of 50 loci, has provided new SSRs for tagging this QTL region. Our data suggests that BNL2621 and BNL3145 are the best candidate flanking markers within

this interval to facilitate foreground selection for this QTL region as a single unit. While linkage drag is always a concern given that this QTL region quite large, it does not affect other important fiber quality traits such as length, elongation, and micronaire (data not shown). Therefore, the use of this QTL region for improving fiber strength is not expected to penalize other components of fiber quality.

To further evaluate the utility of this QTL region for improving the U.S. cotton germplasm, we tested ten markers within the BNL2621 and BNL3145 flanking region on a panel of elite germplasm lines that were submitted to the 2010 Regional Breeders Testing Network (www.cottonrbtn.com). These lines represent the most elite breeding materials in the U.S. public cotton breeding programs, and include five recently released cultivars (Figure 4). The result shows that the alleles linking the fiber strength QTLs from S-7235 are absent in all the lines in the test panel. Therefore, the markers could be employed for MAS in segregating populations involving S-7235 and any of the elite germplasm lines tested.

References

Blenda, A., J. Scheffler, B. Scheffler, M. Palmer, J.M. Lacape, J.Z. Yu, C. Jesudurai, S. Jung, S. Muthukumar, P. Yellambalase, S. Ficklin, M. Staton, R. Eshelman, M. Ulloa, S. Saha, B. Burr, S.L. Liu, T.Z. Zhang, D.Q. Fang, A. Pepper, S. Kumpatla, J. Jacobs, J. Tomkins, R. Cantrell, D. Main. 2006. CMD: a Cotton Microsatellite Database resource for Gossypium genomics. Bmc Genomics 7:132.

Bowman, D., O. Gutiérrez. 2003. Sources of fiber strength in the US upland cotton crop from 1980 to 2000. J. Cotton Sci. 7:164-169.

Bowman, D.T., O.A. Gutierrez, R.G. Percy, D.S. Calhoun, O.L. May. 2006. Pedigrees of upland and pima cotton cultivars released between 1970 and 2005. Bull. 1017. Miss. Agric. and For. Exp. Stn., Mississippi State.

Chee, P., B. Campbell. 2008. Bridging Classical and Molecular Genetics of Cotton Fiber Quality and Development. Genetics and Genomics of Cotton:283.

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

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

Chee, P.W., J. Rong, D. Williams-Coplin, S.R. Schulze, A.H. Paterson. 2004. EST derived PCR-based markers for functional gene homologues in cotton. Genome 47:449-462.

Chen, H., N. Qian, W.Z. Guo, Q.P. Song, B.C. Li, F.J. Deng, C.G. Dong, T.Z. Zhang. 2009. Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro.D8 in Upland cotton. Theor. Appl. Genet. 119:605-612.

Culp, T., D. Harrell. 1974. Breeding Quality Cotton at the Pee Dee Experiment Station, Florence, SC. USDA Res. Pubt. ARS-S-30:1-12.

Culp, T.W., C.C. Green. 1992. Performance of Obsolete and Current Cultivars and Pee Dee Germplasm Lines of Cotton. Crop Sci. 32:35-41.

Culp, T.W., D.C. Harrell. 1979. Registration of five germplasm lines of cotton (Reg. No. GP44 to GP48). Crop Sci. 19:751-752.

Deussen, H. 1992. Improved Cotton Fiber Properties: The Textile Industry's Key to Success in Global Competition. p. 43-63. In Proceedings from Cotton Fiber Cellulose: Structure, Function and Utilization Conference. National Cotton Council of America, Memphis, TN.

Felker, G.S. 2001 Fiber quality and new spinning technologies. p. 5-7, *In* P Dugger and D Richter (Eds.), 2001 Beltwide Cotton Conferences, Anaheim, CA. 9-13 Jan. 2001. National Cotton Council of America, Memphis, TN

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

Guo, W.Z., T.Z. Zhang, X.L. Shen, J.Z. Yu, R.J. Kohel. 2003. Development of SCAR marker linked to a major QTL for high fiber strength and its usage in molecular-marker assisted selection in upland cotton. Crop Sci. 43:2252-2256.

Holland, J. 2007. Genetic architecture of complex traits in plants. Curr. opin. plant biol. 10:156-161.

Jenkins, J.G. 1948. Sea Island cotton breeding. 28th Ann. Rep. Ga Coastal Plains Exp Stn Bull 16:1.

Kosambi, D. 1944. The estimation of map distances from recombination values. Ann Eugen 12:172-175.

Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln, L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174.

Lecomte, L., P. Duffe, M. Buret, B. Servin, M. Causse. 2004. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theor. Appl. Genet. 109:658-668.

Levene, H. 1960 Robust tests for equality of variances, Contributions to probability and statistics, Stanford University Press. pp. 278–292.

Li, Y., X. Li, J. Li, J. Fu, Y. Wang, M. Wei. 2009. Dent corn genetic background influences QTL detection for grain yield and yield components in high-oil maize. Euphytica 169:273-284.

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

Lynch, M., B. Walsh. (1998) Genetics and analysis of quantitative traits Sinauer Associates INC.

May, O. 1999 Genetic variation in fiber quality, Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing, Food Products Press, New York, NY. pp. 183-229.

Meredith, W., T. Culp, K. Robert, G. Ruppenicker, W. Anthony, J. Williford. 1991. Determining future cotton variety fiber quality objectives. Text Res J 61:715.

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

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

Qian, S., J. Huang, Y. Peng, B. Zhou, M. Ying, D. Shen, G. Liu, T. Hu, Y. Xu, L. Gu. 1992. Studies on the hybrid of G. hirsutum L. and G. anomalum Wawr. & Peyr. and application in breeding. Sci. Agric. Sinica 25:44-51.

Sas Institute Inc. (1989) SAS/STAT user's guide, SAS Institute Inc., Cary, NC.

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

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

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

Shen, X.L., T.Z. Zhang, W.Z. Guo, X.F. Zhu, X.Y. Zhang. 2006. Mapping fiber and yield QTLs with main, epistatic, and QTL X environment interaction effects in recombinant inbred lines of upland cotton. Crop Sci. 46:61-66.

Steele, K., A. Price, H. Shashidhar, J. Witcombe. 2006. Marker-assisted selection to introgress rice QTLs controlling root traits into an Indian upland rice variety. Theor. Appl. Genet. 112:208-221.

Wang, S., C. Basten, Z. Zeng. 2005. Windows QTL cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Yuan, Y., T. Zhang, W. Guo, X. Shen, J. Yu, R. Kohel. 2001. Molecular tagging and mapping of QTLs for super quality fiber properties in upland cotton. Acta Genet. Scinica 28:1-11.

Zhang, J., W. Guo, T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L. x Gossypium barbadense L.) with a haploid population. Theor. Appl. Genet. 105:1166-1174.

Zhang, T.Z., Y.L. Yuan, J. Yu, W.Z. Guo, R.J. Kohel. 2003. Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. Theor. Appl. Genet. 106:262-268.

Table 5.1. Biometrical parameters of QTLs for fiber strength on chromosome 24.

Population	Generation (Year)	LOD	Additive	Dominance	R^2
Pop542	F_2 (2005)	12.5	19.02	-1.34	22.01
	$F_{2:3}$ (2006)	15	10.63	-2.89	24.85
	$F_{2:4}$ (2007)	15.1	9.35	1.37	18.15
Pop883	F_2 (2005)	14.1	22.82	-1.24	30.92
	$F_{2:3}$ (2006)	15.3	11.97	-2.58	40.1
	$F_{2:4}$ (2007)	17.4	11.99	-3.32	36.36
Combined across generations	$F_{2:3}$, $F_{2:4}$	16.6	11.17	-3.39	44.48

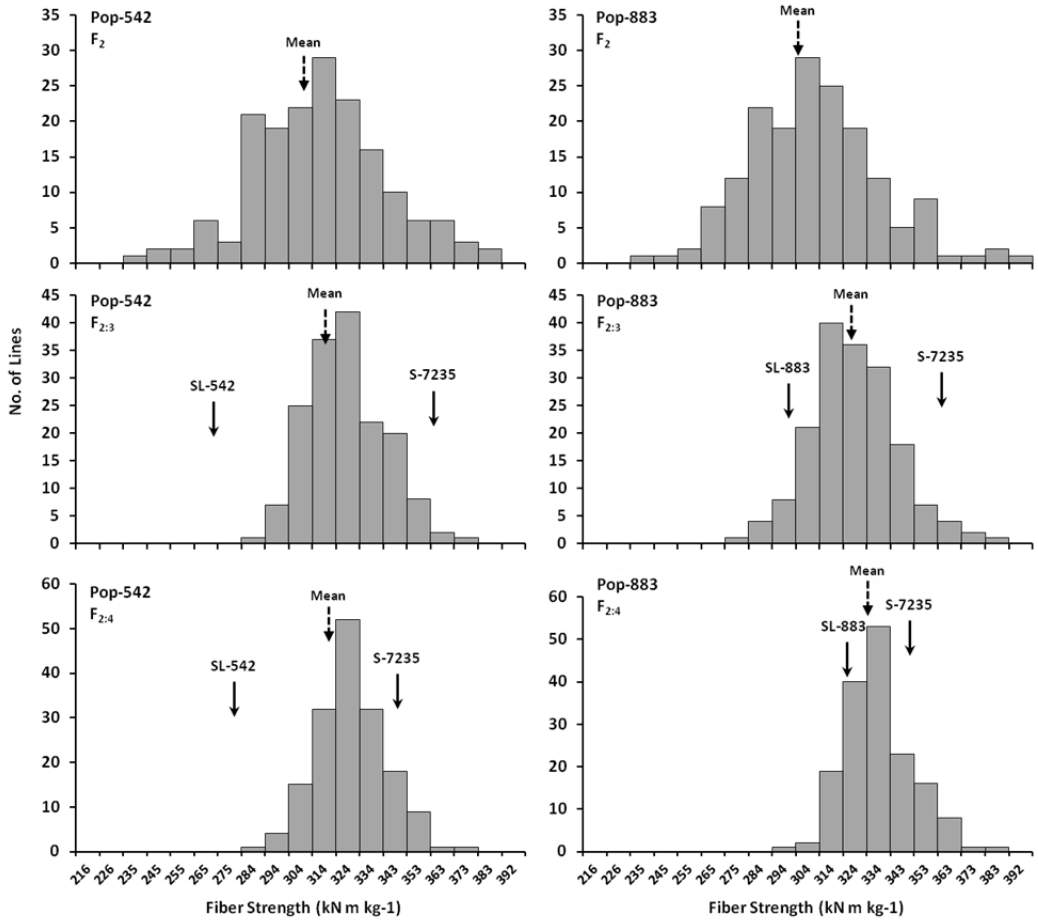


Figure 5.1. Distribution of fiber strength in F_2, $F_{2:3}$ and $F_{2:4}$ generations of Pop-542 and Pop-883

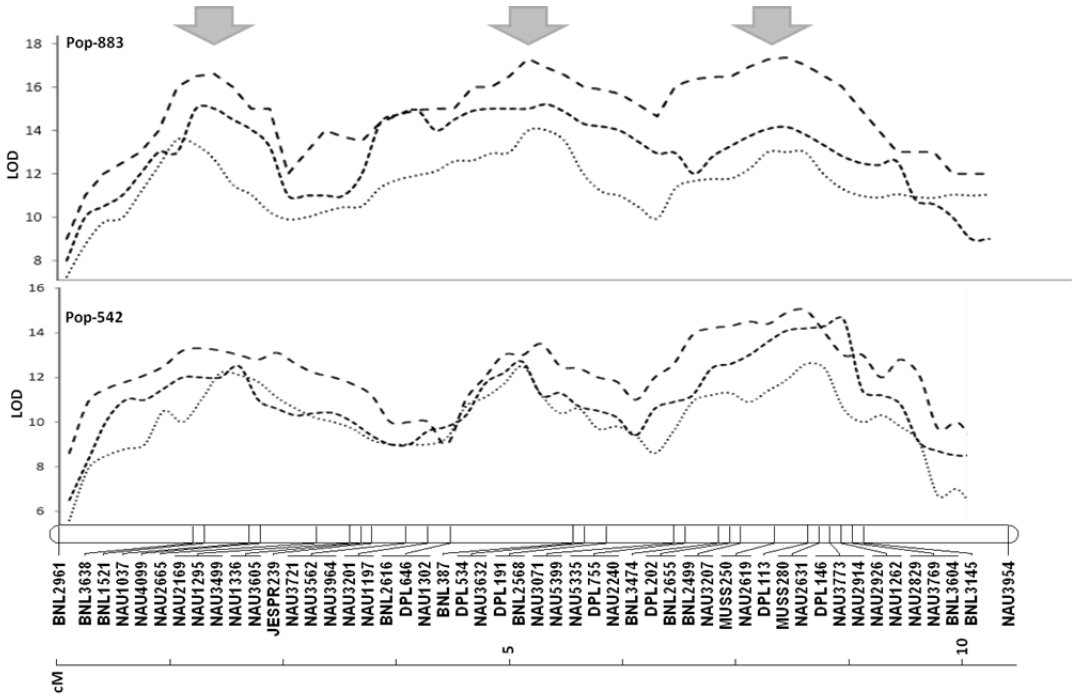


Figure 5.2. Consensus genetic linkage map of Chromosome 24 with LOD score profile for fiber strength in Pop-542 and Pop-883. Arrows indicate the likely peak of QTLs.

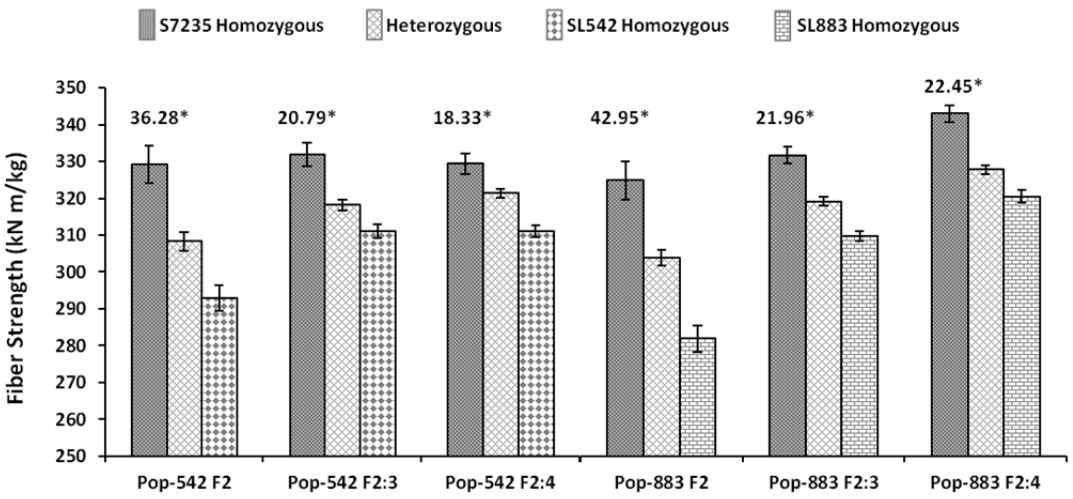
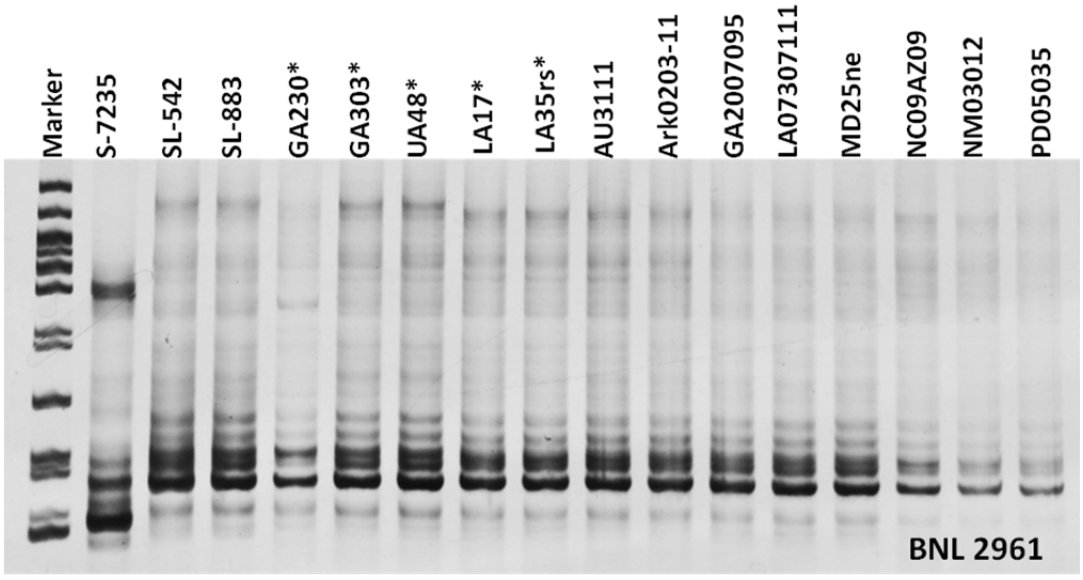


Figure 5.3. Mean fiber strength of different genotypic classes over three generations of Pop-542 and Pop-883. Error bars are standard error of the means.

* Difference in fiber strength (kN m kg^{-1}) between the homozygous genotypic classes



* Recently released cultivars

Figure 5.4. Polymorphism analysis of marker BNL2961 on a panel of elite cultivars and germplasm lines.

CHAPTER 6

SIMULTANEOUS MAPPING AND VALIDATION OF FIBER QUALITY QTLS IN MAJOR GENETIC BACKGROUNDS OF US UPLAND COTTON (*GOSSYPIUM HIRSUTUM* L.)[4]

[4] Pawan Kumar, Rippy Singh, Edward L. Lubbers, Andrew H. Paterson, B. Todd Campbell, Donald C. Jones and Peng W. Chee. To be submitted to *Theoretical and Applied Genetics*

Abstract

Genetic backgrounds can have profound effects on the identification and performance of introgressed QTLs, therefore validation of the effects of these introgressed QTLs in multiple genetic backgrounds is essential before they are deployed in a marker assisted breeding program. In order to validate the effectiveness of fiber quality QTLs introgressed into Sealand 883 across genetic backgrounds, we developed four mapping populations by crossing Sealand 883 with Acala SJ-4, Deltapine 50, Paymaster HS 26, and GA 2004089 respectively, each representing a major genetic background of Upland cotton grown in the US. QTL mapping in F_2 and $F_{2:3}$ generations of each mapping population resulted in the identification of 31 fiber quality QTLs including 10 QTLs that were consistently detected over generations in all the four genetic backgrounds. Effects of genetic background on detection and performance of fiber quality QTLs was observed at multiple loci. Thirteen (42%) QTLs identified in this study were also reported in our previous studies where the favorable allele for six QTLs was contributed by Sealand 883. Our results validate Sealand 883 as a repository of fiber quality QTL that can be targeted via marker assisted selection to improve elite Upland germplasm.

Introduction

The use of interspecific introgression from *Gossypium barbadense* L. into Upland cotton (*G. hirsutum* L.) has been an attractive approach to improve fiber quality (Saha et al., 2006; Saha et al., 2004; Stelly et al., 2005). However, stable introgression of desirable alleles between the two species is often hindered by segregation distortion (Jiang et al., 2000), hybrid breakdown (Stephens, 1947), repressed recombination (Reinisch et al., 1994), and linkage drag (Dighe et al., 2009). While a number of stable introgressions of fiber quality alleles have been reported, these studies also found significant interaction between introgressed alleles and genetic background (Chee et al., 2005a; Chee et al., 2005b; Draye et al., 2005) indicating that the use of interspecific introgression to improve fiber quality may not be a straightforward.

One of the most effective Upland cotton breeding programs in the US that successfully utilized interspecific introgression for fiber quality improvement is the USDA-ARS Pee Dee breeding program at the Pee Dee Experiment Station, Florence, SC. The Pee Dee breeding program was established in 1935 with the objective of improving fiber attributes of Upland cotton (Culp and Harrell, 1973) and has produced several Extra Long Staple (ELS) Upland cotton cultivars including Sealand and Earlistaple with fiber quality on par with Sea Island (*G. barbadense*) cotton. At present, the germplasm lines developed at the Pee Dee breeding program along with those developed at the New Mexico Acala breeding program accounts for over 50% of the fiber quality improvements present in commercial cultivars (Bowman and Gutierrez, 2003; Campbell et al., 2011). The high fiber quality of Pee Dee germplasm can be attributed to its unique breeding history involving germplasm lines with *G. barbadense* introgressions and Beasley's

Triple Hybrid (*G. arboretum* L. X *G. thurberi* Tod. X *G. hirsutum* L.) (Bowman et al., 2006). The pedigree records also highlight that only a subset of the high fiber quality germplasm lines developed in mid-1940's were included in the pool that constituted the base for the modern Pee Dee germplasm.

Historically, Upland cotton grown in the U.S. cotton belt can be broadly divided into four types based on their adaptation to specific production regions as well as differences in fiber quality or production practices; Acala, Delta, Plains, and Eastern (Hague et al., 2009; Smith and Cothren, 1999; Ware, 1951). The Acala type is recognized by its high fiber quality derived from their introductions after the advent of the boll weevil in early 1900s and is adapted to the Western US cotton belt. The Delta type has high yield potential with broad adaptation and is grown in the rain-belt area from southern Texas to Alabama. The Plains type is characterized by its compact plant architecture and stormproof bolls to minimize losses due to high winds in the Texas High Plains. Finally, the Eastern type is adapted to the coastal plains from the Carolinas to Georgia and Florida, has bigger bolls, and is developed by hybrid breeding of post boll weevil cultivars with pre boll weevil era stocks. Although the pedigrees of modern Upland cultivars have become intertwined (Bowman, 1999; Bowman et al. 2006) and the strict distinction of these germplasm types have blurred, the genetic background of each of these germplasm types may still possesses unique genetic attributes.

Genetic backgrounds can have profound effects on the identification and performance of introgressed QTLs in several crops such as rice (Liao et al., 2001; Steele et al., 2006), maize (Li et al., 2009), tomato (Lecomte et al., 2004b), and soybean (Sebolt et al., 2000) where the QTLs detected in one genetic background were less effective in

other genetic backgrounds due to interaction with other loci or epistasis (Holland, 2007). For example in rice, Liao et al. (2001) mapped QTLs for panicle number in populations derived from crossing a japonica variety with two indica varieties and found that the effect of genetic background on QTL detection is greater than the environment effect. Furthermore, the magnitude of epistatic interaction among loci is not only greatly influenced by the genetic background in which the QTLs were detected, but the interaction effect can be greater than the main-effect of the QTLs.

In a previous study, we identified introgressed regions from *G. barbadense* in the Upland cotton cultivar "Sealand 883" and genetic analysis indicated that a number of these introgressed chromosomal segments contain alleles which enhance the fiber quality of Upland cotton (Chapters 3 & 4). Validation of the effects of these introgressed QTLs in multiple genetic backgrounds is essential before they are deployed in a marker assisted breeding program. The objective of the present study was to evaluate the effects of genetic background on the efficacy of fiber quality QTLs introgressed from *G. barbadense* into Upland cotton.

Materials and methods

Mapping populations and phenotyping

Four F_2 populations were created by crossing Sealand 883 with Acala SJ-4, Deltapine 50, Paymaster HS 26 and GA 2004089, representing Acala type, Delta type, Plains type, and Eastern type genetic backgrounds of Upland cotton, respectively.

The germplasm line Sealand 883 (PI 528875), henceforth referred to as SL-883, was developed in 1945 at the Pee Dee Experiment Station, Florence, SC. It was a

progeny row selection from 'Bleak Hall' (*G. barbadense*, Sea Island) (PI 608115) X 'Coker Wilds' hybrid backcrossed four times to the 'Wilds' parent (Culp and Harrell, 1974b; Culp and Harrell, 1980) Acala SJ-4 or AcalaSJ4 (PI 529538) (Pedigree C6TE/NMB3080) was released in 1976 by USDA-ARS, Shafter, CA., is a high fiber quality cultivar with increased fiber and yarn strength and better yielding on verticillium-wilt infested soils. Deltapine 50 or DP50 (PI 529566) (Pedigree DP 16//DP Smoothleaf/DP 45/3/DES 56) was developed by Delta & Pine Land Co., Scott, MS and was released for commercial cultivation in 1984. Paymaster HS 26 or PMHS26 (Pedigree Acala SJ-4/5B9-184) was developed in 1983 by Paymaster Technologies, Inc., Aiken, TX and is characterized by compact plant habit, determinant fruiting habit, and storm resistant bolls. GA 2004089 or GA089 (Pedigree PD 94042/AP 7126) is a high yielding germplasm line developed at the Tifton campus, The University of Georgia, USA.

The F_1 hybrids for the four populations were grown in the greenhouse and F_2 seeds were collected from single F_1 plants for each cross combination. The four F_2 mapping populations comprised of 150 individuals each along with the five parents were planted at the William Gibbs Farm, in Tifton, GA (31.27°N, 83.30°W) in the summer of 2007 following standard production practices. Upon maturity, seed cotton from the individual F_2 plants was hand-picked and ginned on a table-top saw gin in cotton ginning facility at the University of Georgia-Tifton Campus, Tifton, GA. In 2008, $F_{2:3}$ families were planted as progeny rows in a completely randomized design with two replications along with the five parents as checks in Tifton, GA. The plots were single row plots, 9 m by 1 m, planted at four seeds per row foot in mid-May and harvested in early October. Standard production practices were followed in each test. Upon maturity, twenty-five

bolls from middle region of the plants were selected in each progeny row and ginned on a table-top saw gin. Fiber quality of each ginned sample was measured using the High Volume Instruments (HVI) at the Cotton Incorporated Textile Services Laboratory (Cotton Incorporated, Cary, NC). The fiber quality measurements comprised fiber length or Upper Half Mean (UHM) in inches, fiber strength (STR) in g/tex, fiber fineness or micronaire (MIC), percent fiber elongation (ELO), percent short fiber content (SFC), and percent uniformity index (UI).

Genotyping

A total of 1250 SSR markers, covering all of the 26 homeologous chromosomes, were screened for polymorphism between the mapping parents. Primer sequences of the SSR markers were obtained from the Cotton Marker Database (CMD) (www.cottonmarker.org) (Blenda et al., 2006) and were commercially synthesized by Eurofins MWG Operon (Huntsville, AL). PCR amplification was performed as described (Chee et al., 2004), and the PCR products were electrophoretically separated using 10% non-denaturing polyacrylamide gel electrophoresis. The DNA fragments were visualized by staining with silver nitrate as described by (Zhang et al., 2002). The SSR primer pairs were first screened for polymorphism between the SL-883 and the four genetic background parents, and the polymorphic primers were tested on the respective mapping populations.

Linkage map construction

Genetic maps for each of the four populations were constructed using Mapmaker/EXP 3.0 (Lander et al., 1987) software. Assembly of linkage groups was done with the 'group' command using a LOD score of 3.0 and the maximum recombination fraction of 30 cM was set as a grouping thresholds. Recombination units were converted into genetic distances by using the Kosambi mapping function (Kosambi, 1944) with the "error detection" command on. Unlinked groups or markers were added to framework using the 'try' and 'compare' commands. The final order of marker sequence on a linkage group was confirmed using the 'ripple' command. Assignment of linkage groups to the chromosome and subgenome is based on the published comprehensive reference map of tetraploid cotton (Yu et al., 2010).

Data analysis

The phenotypic distributions and correlations among the fiber quality traits in all four mapping populations were calculated with statistical analysis software SAS 9.2 (SAS Institute Inc, 1989). Detection of quantitative trait loci and estimation of various genetic parameters were performed by Composite Interval Mapping (CIM) implemented in the software WinQTL Cartographer version 2.5 (Wang et al., 2005). Likelihood ratio (LR) threshold values (α=0.05) for declaring a QTL were estimated after 1000 permutations for each phenotype (Doerge and Churchill, 1996). In addition, when a QTL was detected in one of the four populations above the threshold score, we scanned the LOD profile of remaining populations at that loci at lower threshold levels and peaks below the threshold, but LOD>2.5 were considered putative QTLs. Mapping was

performed at 1 cM walk speed in a 10 cM window and five background cofactors where the cofactors were selected *via* forward-backward stepwise regression method. QTL position was defined by one-LOD confidence interval on the either sides of the peak position. QTL are named following the protocol outlined by (McCouch et al., 1997), briefly, the QTL is designated as '*q*' followed by an abbreviation of the trait name which is then followed by the chromosome name. Multiple QTLs on a chromosome are distinguished by an alphabetical suffix. In this study we classified QTLs as 'common' when the QTL is detected in the same confidence interval of more than one population, 'consistent' if a QTL is detected in both F_2 and $F_{2:3}$ generations, and 'unique' if detected in only one generation of any population.

For markers segregating in two or more genetic backgrounds, two-way mixed model ANOVA was performed using the PROC MIXED procedure of SAS 9.2 (SAS Institute Inc, 1989). The model included genotype, background, and genotype x background interaction as fixed factors. Model parameters were estimated using the residual maximum likelihood (REML) method (SAS Institute Inc, 1989).

Results

Phenotypic performance

The fiber quality traits of the parents differed significantly when planted in 2007 and 2008; however the overall ranking remained the same (Figure 1). The SL-883 parent had significantly lower (LSD at $P<0.05$) fiber elongation, fiber fineness, and uniformity index but had significantly greater fiber length than the background parents in both the 2007 and 2008 planting years. AcalaSJ4 had greatest fiber strength in 2007 and 2008

followed by PMHS26. Short fiber content of all parents was significantly greater in 2007 than in 2008, and among them, SL-883 had the highest short fiber content (Figure 1).

Fiber quality traits were normally distributed in both F_2 and $F_{2:3}$ datasets of the four mapping populations (Figure 2). For fiber elongation, the trait means of all four populations were statistically different (LSD at $P<0.05$) in F_2 and $F_{2:3}$ datasets with the PMHS26 population having substantially high elongation in both datasets. Fiber fineness of the AcalaSJ4 population was significantly lower in both datasets. Short fiber content of all populations was significantly higher in the F_2 dataset, and the GA089 population had the lowest short fiber content in both datasets. There was no significant difference in mean fiber strength of the AcalaSJ4 and the PMHS26 populations however mean fiber strength of both, the AcalaSJ4 and the PMHS26 populations, was significantly higher than the mean fiber strength of the DP50 and the GA089 populations. Significant differences in fiber length were observed between the populations where the GA089 background population had substantially and significantly higher fiber length. The uniformity index significantly differed among populations in both F_2 and $F_{2:3}$ datasets of all populations except for the difference between the AcalaSJ4 and the PMHS26 populations. Pearson correlation coefficients among fiber quality traits show significant positive correlation between fiber length, fiber strength, and fiber uniformity traits. Fiber elongation and short fiber content were found to be negatively correlated with fiber length and fiber strength. Identical correlation trends were observed in F_2 and $F_{2:3}$ datasets. The Pearson correlation coefficients among fiber quality traits in $F_{2:3}$ dataset of the four mapping populations are presented in Table 2.

Linkage maps

Out of the 1250 SSR markers tested a total of 316 polymorphic SSR markers differentiated between the Upland background parents and SL-883 (84 for AcalaSJ4, 78 for DP50, 88 for PMHS26 and 66 for GA089). Independent linkage maps for each of the four mapping populations were developed. Out of the 316 polymorphic markers, 267 (84.4%) markers could be placed into linkage groups (LGs). The linkage maps consisted of 22 LGs with 79 mapped markers for the AcalaSJ4 population, 16 LGs with 66 markers for the DP50 population, 17 LGs with 70 markers for the PMHS26 population and 13 LGs with 52 markers for the GA089 population. Marker orders of the developed maps were confirmed by comparing with the consensus linkage map (Yu et al., 2010) available at cotton marker database.

QTL analysis

A total of 38 QTLs and 8 putative QTLs were detected by CIM independently performed in the four mapping populations. Eleven QTLs and three putative QTLs were mapped in the AcalaSJ4 background, twelve QTLs in the DP50 background, eight QTLs and four putative QTLs in the PMHS26 background and seven QTLs and one putative QTL in the GA089 background (Table 3). Pooling of results from these four independent mapping studies resulted in a total of 31 QTLs that included unique, common, and consistent QTLs. The description of QTLs identified for each fiber quality traits is summarized below.

Fiber elongation (ELO)

A total of five QTLs for fiber elongation were identified (Table 3). Two (chr. 12 & 24) were unique to the PMHS26 genetic background population and one on chromosome 5 to DP50 background population. A QTL on chromosome 25 was consistent in F_2 and $F_{2:3}$ generations of the PMHS26 population but was identified as putative QTL in both generations of the AcalaSJ4 population. The percent phenotypic variation (PV%) explained by these QTLs ranged from 2.35 to 48.62%. Favorable allele for *qELO-Chr12* was contributed by SL-883 while background parents contributed favorable alleles for other QTLs.

Micronaire (MIC)

A total of six QTLs were detected for micronaire (Table 3). Two (chr. 12 & 25) unique to the PMHS26 population while three (chr. 11, 15, & 23) unique to the AcalaSJ4, the DP50, and the GA089 populations respectively. *qMIC-Chr16* was common between the DP50 and the GA089 populations, while it was putative in the PMHS26 population. Phenotypic variance accounted by these six QTLs ranged from 2.63 to 36.29%. The favorable alleles for two QTLs, *qMIC-Chr11* and *qMIC-Chr23,* originated from SL-883, while for other QTLs, they were contributed by the respective background parents.

Short fiber content (SFC)

A total of two QTLs affecting short fiber content were identified (Table 3). Both (chr. 5 & 8) were uniquely identified in F_2 and $F_{2:3}$ generations of the AcalaSJ4 population. The genetic effects ranged from -0.12 to 0.18 explaining 9.54 to 12.3% of the total phenotypic variance. The favorable alleles decreasing short fiber content originated

from SL-883 for *qSFC-Chr5* while the AcalaSJ4 contributed the favorable allele for *qSFC-Chr8*.

Fiber strength (STR)

A total of 5 QTLs for fiber strength were detected (Table 3). Two (chr. 5 & 25) identified in the DP50 population and two (chr. 11 & 15) identified in the AcalaSJ4 population above threshold LOD score while these were putative in other populations (Table 3). The QTL on chromosome 16 was common between the DP50 and the GA089 population consistently detected in both F_2 and $F_{2:3}$ datasets. The percent phenotypic variation explained by these QTLs ranged from 3.22 to 20.54% with genetic effects ranging from -0.94 to 0.67 g/tex. The favorable alleles for four (chr. 5, 15, 16, and 25) were contributed by the SL-883 parent while the AcalaSJ4 parent contributed the favorable allele for *qSTR-Chr11*.

Fiber length (UHM)

A total of 8 fiber length QTLs were identified (Table 3). Two (chr. 12 & 23) unique to the PMHS26 population and one (chr. 11) unique to the AcalaSJ4 population. QTL on chromosome 5 and 25-c were identified above threshold LOD in $F_{2:3}$ datasets of the DP50 and the GA089 populations respectively, while these were putative in F_2 datasets of their respective populations (Table 3). The QTL on chromosome 16 was common between the DP50 and the GA089 populations while the QTL on chromosome 25-c was both common and consistent in all datasets of all populations; however it was putative in the GA089 F_2 population. The phenotypic effect of these QTLs ranged from 0.25 to 1.02 mm, explaining from 3.73 to 61.35% of the total phenotypic variance. The favorable alleles for 5 QTLs (62.5%) originated from SL-883 parent, while favorable

alleles for QTLs on chromosomes 11, 12 & 23 were contributed by their respective background parents.

Fiber uniformity (UI)

A total of 5 QTLs were identified (Table 3); two (chr. 18 and 24) were unique to the AcalaSJ4 population while one on chromosome 25 was unique to the DP50 population. The QTL on chromosome 8 were consistently detected in the F_2 and $F_{2:3}$ datasets of the AcalaSJ4 background, while the QTL on chromosome 15 was detected above threshold LOD score in the F_2 dataset of the AcalaSJ4 population but was putative in the $F_{2:3}$ dataset. The genetic effects of these QTLs ranged from -0.853 to 1.045, and the percent phenotypic variation explained ranged from 3.03 to 13.67 %. The favorable allele for *qUI-Chr15* originated from the SL-883 parent, while for other QTLs, background parents contributed the favorable allele.

Common and consistent QTLs

Four consistent QTLs each were identified in AcalaSJ4 and DP50 backgrounds while 3 and 6 consistent QTLs were identified in PMHS26 and GA089 backgrounds, respectively. Ten QTLs were common between two or more genetic background. Eight QTLs *qELO-Chr11*, *qUHM-Chr25-a*, *qELO-Chr25*, *qSTR-Chr25*, *qMIC-Chr16*, *qUHM-Chr25-b*, *qSTR-Chr16,* and *qUHM-Chr16* were both consistent and common QTLs identified over generations of multiple populations.

Genotype by background interactions

Genetic background main effects were significant for all fiber quality traits except for fiber fineness. A total of 47 markers were found segregating in more than one genetic background of which thirty-two markers were segregating in all four populations. Ten significant ($p<0.005$) genotype x background interactions were found by six markers. Three markers, BNL1694 for fiber strength and BNL3937 and NAU2714 for fiber length, also showed significant main effect.

Effect of genetic backgrounds on fiber length QTL qUHM-Chr25-a

The fiber length QTL on chromosome 25 was identified on introgressed region flanked by markers BNL2569 and NAU3502. This introgression region was found to be polymorphic and was segregating in all four populations. We identified 114 lines that were homozygous at all loci for background alleles, of these, 23, 38, 28 and 25 lines were from the AcalaSJ4, the DP50, the PMHS26, and the GA089 populations respectively. Similarly, 116 lines were identified that were homozygous for the SL-883 allele; 31, 31, 22, and 32 lines were from the AcalaSJ4, the DP50, the PMHS26, and the GA089 populations respectively. We also identified lines that were completely heterozygous for all loci in all the four populations. Significant differences in fiber length were observed between the two homozygous classes in all four populations (data not shown). Genetic effects of allelic substitution were calculated for the replicated $F_{2:3}$ dataset as the difference in deviation from population mean between homozygous (background allele) class and the heterozygous class (Figure 3). The largest effect of allele substitution was observed in the PMHS26 population, where substitution of one background allele by the SL-883 allele increased fiber length by 0.80 mm. Allelic substitution effects of 0.66,

0.46, and 0.16 mm were observed in the GA089, the DP50, and the AcalaSJ4 genetic backgrounds, respectively.

Discussion

Fiber quality alleles from SL-883

The SL-883 germplasm line may be a source of fiber quality alleles. Our results show that the majority of the *G. barbadense* introgressed segments in SL-883 are not prevalent in major genetic backgrounds of Upland cotton grown in the U.S. Comparison of our results with the two meta-analysis of fiber quality QTLs done in an interspecific F_2 mapping population (Rong et al., 2007a) and across diverse environments in an interspecific RIL population (Lacape et al., 2010) reveal that the majority of fiber quality QTLs detected on the introgressed segments of SL-883 in this study have not been previously reported. For example, we detected QTLs for fiber length and fiber strength on chromosome 25 and chromosome 16, and these chromosomes have not been reported to contain QTLs in the meta-analysis (Lacape et al., 2010).

Of the 31 QTLs detected in this study, the favorable alleles of 15 QTLs (or 48%) originated from the SL-883 parent. Interestingly, of those 15 QTLs, 11 (or 73%) were identified on the previously characterized introgressed segments. These results confirm our hypothesis that the high fiber quality of the SL-883 is primarily due to introgression from the *G. barbadense* parent. A favorable allele for only one of the five fiber elongation QTLs originated from the SL-883. *G. barbadense* has lower fiber elongation than Upland cotton (Chee et al., 2005a), therefore favorable alleles from the *G.*

barbadense introgressions in the SL-883 were less likely. In addition, fiber elongation might not have been the selection criteria for breeders during cultivar development (May, 1999). The SL-883 parent contributed favorable alleles for five of the eight fiber length QTLs predominantly with additive effects. The findings are consistent with other studies where the *G. barbadense* parent has contributed favorable alleles for the majority of fiber strength and fiber length QTLs (Chee et al., 2005b; Lacape et al., 2005).

Contribution of both At and Dt sub-genomes towards the control of fiber quality in the SL-883 was observed, however the Dt sub-genome had a slightly greater contribution with 18 (58%) QTLs identified on the Dt sub-genome chromosomes. A noteworthy observation is that a majority of QTLs for the two most important quality traits, fiber length and fiber strength, were located on the Dt sub-genome. Similar results were earlier reported in cotton (Chee et al., 2005b; Jiang et al., 1998; Mei et al., 2004; Paterson et al., 2003; Rong et al., 2007b) where the genetic control of fiber quality by the Dt sub-genome was significantly greater than that of At sub-genome. Although the D-genome progenitor does not produced spinnable fibers, it contains loci that influence the fiber quality in allotetraploid cottons. These results support the notion that polyploidization in *Gossypium* created novel variation for fiber quality (Jiang et al., 1998).

Effects of genetic background on fiber quality QTLs

The effects of genetic background on QTL detection and performance has been studied and documented in several crops (Bernacchi et al., 1998; Chaib et al., 2006; Lecomte et al., 2004a; Li et al., 2009). In cotton, interaction of QTLs with genetic

background has been previously studied only in biparental populations (Chee et al., 2005a; Chee et al., 2005b; Draye et al., 2005; Kumar et al., 2012; Zhang et al., 2011). Three of the four varieties, Acala SJ-4, Deltapine 50, and Paymaster HS 26, selected as background parents in this study, hold historical significance in Upland cotton breeding, representing the Acala, Delta, and Plains genetic backgrounds, respectively. The germplasm line GA089, which represent the Eastern type background, is a modern Pee Dee germplasm line. Significant interaction with background was detected for six markers of which three also had significant main effects. This result suggests that the different *G. hirsutum* parents may carry different sets of fiber quality alleles or different alleles at trans-acting loci that influence these QTLs (Zhang et al., 2011). For example, Sealand alleles contributed favorable alleles for two fiber strength QTLs, *qSTR-Chr05* and *qSTR-Chr11,* when present in Delta and Eastern type backgrounds. However, in the Acala background, the favorable allele for the same QTLs originated from AcalaSJ4. The Acala germplasm may harbor unique alleles for fiber strength, possibly introgressed from a three species triple hybrid (*G. arboreum* L. X *G. thurberi* Tod. X *G. hirsutum* L.) (Smith and Cothren, 1999; Zhang et al., 2005). Efficacy of the fiber length QTL *qUHM-Chr25-a* was greatest when present in the Plains type genetic background where it accounted for up to 61% of the phenotypic variation and the additive effect is up to 50% greater than in the other three backgrounds. Interestingly, the efficacy of this QTL was the least when present in the Acala type genetic background, which further suggests that different sets of fiber quality alleles may be present in the different germplasm types.

Stability of the fiber quality QTLs

The stability of QTLs is of great concern to breeders aiming to introgress favorable alleles from diverse sources into elite backgrounds. QTLs with minimal genetic background effects and that are consistent in expression over years are highly desirable for efficient marker assisted breeding. Therefore, robustness of the identified QTLs was djudged by comparing their association either in multiple backgrounds as well as in multiple generation datasets. Despite many fiber quality QTLs having been identified in cotton, a majority of the QTLs have not been validated in multiple backgrounds or evaluated in more than one generation (Chee and Campbell, 2008). In this study, eight (26%) of the QTLs reported here were detected in multiple generations and in more than one genetic backgrounds. In addition, two QTLs *qSTR-Chr5* and *qMIC-Chr11* were identified only in the $F_{2:3}$ dataset but were present in multiple background. The SL-883 parent contributed favorable alleles for five of the eight common-consistent QTLs. Although the majority of QTLs showed additive gene action, the expression of QTLs was greatly influenced by the genetic background in which these were expressing. For example, the fiber length QTL *qUHM-Chr25-a* on chromosome 25 accounted for 10 fold genetic variation when present in Plains type background. Therefore, the favorable allele for this QTL from SL-883 may not perform equally in cotton cultivars adapted to different regions of the US cotton belt, but may be particularly effective when introgressed in elite cultivars of the Plains type genetic background.

Thirteen (42%) QTLs identified in this study were also reported in our previous studies (Chapters 3 & 4) where favorable alleles for six QTLs was contributed by SL-883. The three consistent QTLs *qSTR-Chr16*, *qUHM-Chr16*, *qUHM-Chr25-a* from this

study also had stable expression in our previous studies using the parental line Suyuan-7235 (Chapters 3 & 4). It is interesting to note that Suyuan-7235 was developed via a cross involving the germplasm line Acala 3080 (Qian et al., 1992) and nine of these thirteen QTLs were identified in Acala type genetic background. For example, the QTL *qMIC-Chr11* which was detected in our previous studies had the favorable allele from SL-883 and in the present study, it was identified in both Acala and Plains type genetic background. However, when present in Acala type genetic background, the SL-883 allele had a positive effect. Another noteworthy observation is that a majority (80%) of the fiber uniformity QTLs was identified from the Acala type genetic background with AcalaSJ4 predominantly (75%) contributing favorable alleles for all except for QTL *qUI-Chr15* where the allele from SL-883 conferred a positive effect. Similar results were obtained in our previous study (Chapter 4) where the favorable alleles for all fiber uniformity QTLs originated from Suyuan-7235 except for the *qUI-Chr15* where the favorable allele originated from SL-883. However, this QTL was only detected in the F_2 and not in the $F_{2:3}$ or the $F_{2:4}$ generations (Chapter 4).

QTL identification is dependent on both the genetic background in which it is expressed and on the setting of the statistical threshold. Mapping experiments use stringent statistical thresholds to minimize experiment-wise error rates (Churchill and Doerge, 1994). If QTLs with large effects have been fixed in elite germplasm, lowering the threshold and tolerating the probability of committing a Type I error might be needed to detect small effect QTLs thereby reducing the chances of committing a Type II error or accepting a false negative (Anderson et al., 1993). Interaction within genetic backgrounds may render a QTL identified as a major locus in one population as a minor locus in

another population. We addressed this issue by simultaneous mapping of QTLs in different sets of genetic backgrounds that enabled us to compare mapping experiment results and draw conclusions regarding the stability of the QTLs identified. A major QTL identified in one background compelled us to scan LOD profiles of other mapping populations where the QTL had not been identified, using a lower threshold thereby controlling Type II error to some extent (Table 3 and Figure 4). The eight putative QTLs described in Table 3 were identified as major QTL in one or more mapping populations. Nonetheless, the QTLs with stable expression across genetic backgrounds and over generations are better targets for marker-assisted introgression without significant alteration in expression when pyramided in an elite cultivar.

Shared Introgression

Few of the *G. barbadense* introgressions in SL-883 may be prevalent in present day elite cultivars. For example, the introgressed segment flanked by markers NAU1042 and BNL3029 on chromosome 5 and the introgression flanked by BNL4082 and DPL615 on chromosome 15 (Figure 5) were non-polymorphic between SL-883 and Eastern type background parent suggesting the presence of some of the Sealand alleles in modern day Pee Dee germplasm.

Out of the nine Sealand lines listed in the USDA National Plant Germplasm Collection (USDA-ARS, 2012) two, Sealand 542 and Sealand 7, contributed alleles to develop modern day Pee Dee germplasm (Bowman et al., 2006; Culp and Harrell, 1974a). Results from our previous study (Chapter 3) show that Sealand 542 and SL-883

do not share the introgressions present on chromosome 5 and chromosome 15. It is likely that these introgressions were contributed by Sealand 7, however we did not test this hypothesis in the present study. The significance of this finding is that the Sealand germplasm lines have been instrumental in developing Extra Long Staple varieties from the Pee Dee germplasm and other Sealand lines including SL-883 may be sources of novel fiber quality alleles.

Unique QTLs from the background parents

Although the main focus of this study was to determine the effects of genetic background on the fiber quality QTLs of SL-883 introgressed from *G. barbadense,* a number of polymorphic regions unique to the background parents were also identified. These unique regions possess fiber quality QTLs. For example; *qSFC-Chr8* was identified uniquely in the Acala background where the favorable allele from AcalaSJ4 decreased short fiber content. Similarly, AcalaSJ4 contributed positive alleles for three unique fiber uniformity QTLs on chromosomes 8, 18, and 24 that increased fiber uniformity. Unique fiber quality QTLs were also identified on chromosomes 12, 23, and 24 in the Plains type genetic background (Table 3 and Figure 6). These results suggest that selective breeding over many generations in developing the different Upland types has resulted in the fixation of different sets of fiber quality alleles in different genetic backgrounds. Therefore, germplasm exchange between various breeding programs may help in tapping into yet unfixed fiber quality alleles.

References

Anderson, J.A., M.E. Sorrells, S.D. Tanksley. 1993. RFLP analysis of genomic regions associated with resistance to preharvest sprouting in wheat. Crop Sci. 33:453-459.

Bernacchi, D., T. Beck-Bunn, D. Emmatty, Y. Eshed, S. Inai, J. Lopez, V. Petiard, H. Sayama, J. Uhlig, D. Zamir. 1998. Advanced backcross QTL analysis of tomato. II. Evaluation of near-isogenic lines carrying single-donor introgressions for desirable wild QTL-alleles derived from Lycopersicon hirsutum and L. pimpinellifolium. Theor. Appl. Genet. 97:170-180.

Blenda, A., J. Scheffler, B. Scheffler, M. Palmer, J. Lacape, J. Yu, C. Jesudurai, S. Jung, S. Muthukumar, P. Yellambalase. 2006. CMD: a cotton microsatellite database resource for Gossypium genomics. BMC Genomics 7:132.

Bowman, D.T., O.A. Gutierrez. 2003. Sources of fiber strength in the U.S. upland cotton crop from 1980 to 2000 [electronic resource]. Journal of cotton science 7.

Bowman, D.T., O.A. Gutierrez, R.G. Percy, D.S. Calhoun, O.L. May. 2006. Pedigrees of upland and pima cotton cultivars released between 1970 and 2005. Mississippi Agricultural and Forestry Experiment Station Technical Bulletin 1155:57.

Campbell, B.T., P.W. Chee, E. Lubbers, D.T. Bowman, W.R. Meredith, J. Johnson, D.E. Fraser. 2011. Genetic Improvement of the Pee Dee Cotton Germplasm Collection following Seventy Years of Plant Breeding. Crop Sci. 51:955-968.

Chaib, J., L. Lecomte, M. Buret, M. Causse. 2006. Stability over genetic backgrounds, generations and years of quantitative trait locus (QTLs) for organoleptic quality in tomato. Theor. Appl. Genet. 112:934-944.

Chee, P., B. Campbell. 2008. Bridging Classical and Molecular Genetics of Cotton Fiber Quality and Development. Genetics and Genomics of Cotton:283.

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

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

Chee, P.W., J. Rong, D. Williams-Coplin, S.R. Schulze, A.H. Paterson. 2004. EST derived PCR-based markers for functional gene homologues in cotton. Genome 47:449-462.

Churchill, G.A., R.W. Doerge. 1994. Empirical threshold values for quantitative trait mapping. Genetics 138:963.

Culp, T., D. Harrell. 1974a. Breeding Quality Cotton at the Pee Dee Experiment Station, Florence, SC. Washington, DC: United States Department of Agriculture Publication S-30:1-12.

Culp, T.W., D.C. Harrell. 1973. Breeding Methods for Improving Yield and Fiber Quality of Upland Cotton (Gossypium-Hirsutum L). Crop Sci. 13:686-689.

Culp, T.W., D.C. Harrell. 1974b. Breeding quality cotton at the Pee Dee Experiment Station Florence, SC. USDA-ARS. Pubi. ARS-S-30 12.

Culp, T.W., D.C. Harrell. 1980. Registration of extra-long staple cotton germplasm (Reg. No. GP 150 to GP 154). Crop Sci. 20:291.

Dighe, N.D., A. Robinson, A.A. Bell, M.A. Menz, R.G. Cantrell, D.M. Stelly. 2009. Linkage Mapping of Resistance to Reniform Nematode in Cotton following Introgression from *Gossypium longicalyx* (Hutch. & Lee). Crop Sci. 49:1151-1164.

Doerge, R., G. Churchill. 1996. Permutation tests for multiple loci affecting a quantitative character. Genetics 142:285.

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

Hague, S., L. Hinze, J. Frelichowski. 2009 Cotton, *In* J. Vollmann and I. Rajcan (Eds.), Oil Crops. pp. 257-285.

Holland, J. 2007. Genetic architecture of complex traits in plants. Curr. opin. plant biol. 10:156-161.

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

Jiang, C.X., R.J. Wright, K.M. El-Zik, A.H. Paterson. 1998. Polyploid formation created unique avenues for response to selection in Gossypium (cotton). P Natl Acad Sci USA 95:4419-4424.

Kosambi, D. 1944. The estimation of map distances from recombination values. Ann Eugen 12:172-175.

Kumar, P., R. Singh, E.L. Lubbers, X. Shen, A.H. Paterson, B.T. Campbell, D.C. Jones, P.W. Chee. 2012. Mapping and Validation of Fiber Strength Quantitative Trait Loci on Chromosome 24 in Upland Cotton. Crop Sci. 52:1115-1122.

Lacape, J.M., D. Llewellyn, J. Jacobs, T. Arioli, D. Becker, S. Calhoun, Y. Al-Ghazi, S. Liu, O. Palaï, S. Georges. 2010. Meta-analysis of cotton fiber quality QTLs across diverse environments in a Gossypium hirsutum x G. barbadense RIL population. BMC Plant Biol. 10:132.

Lacape, J.M., T.B. Nguyen, B. Courtois, J.L. Belot, M. Giband, J.P. Gourlot, G. Gawryziak, S. Roques, B. Hau. 2005. QTL analysis of cotton fiber quality using multiple Gossypium hirsutum x Gossypium barbadense backcross generations. Crop Sci. 45:123-140.

Lander, E., P. Green, J. Abrahamson, A. Barlow, M. Daly, S. Lincoln, L. Newburg. 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1:174.

Lecomte, L., P. Duffe, M. Buret, B. Servin, M. Causse. 2004a. Marker-assisted introgression of five QTLs controlling fruit quality traits into three tomato lines revealed interactions between QTLs and genetic backgrounds. Theor. Appl. Genet. 109:658-668.

Lecomte, L., P. Duffe, M. Buret, B. Servin, F. Hospital, M. Causse. 2004b. Marker-assisted introgression of five QTLs controlling fruit quality traits into three

tomato lines revealed interactions between QTLs and genetic backgrounds. Theor. Appl. Genet. 109:658–668.

Li, Y., X. Li, J. Li, J. Fu, Y. Wang, M. Wei. 2009. Dent corn genetic background influences QTL detection for grain yield and yield components in high-oil maize. Euphytica 169:273-284.

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

May, O. 1999 Genetic variation in fiber quality, Cotton Fibers: Developmental Biology, Quality Improvement, and Textile Processing, Food Products Press, New York, NY. pp. 183-229.

Mccouch, S., Y. Cho, M. Yano, E. Paul, M. Blinstrub, H. Morishima, T. Kinoshita. 1997. Report on QTL nomenclature. Rice Genet Newsl 14.

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

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

Qian, S., J. Huang, Y. Peng, B. Zhou, M. Ying, D. Shen, G. Liu, T. Hu, Y. Xu, L. Gu. 1992. Studies on the hybrid of G. hirsutum L. and G. anomalum Wawr. & Peyr. and application in breeding. Sci. Agric. Sinica 25:44-51.

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

Rong, J., E.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.R. Wendel, A.H. Paterson. 2007a. 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.

Rong, J., R.J. Wright, Y. Saranga, O.L. May, T.A. Wilkins, X. Draye, V.N. Waghmare, F.A. Feltus, P.W. Chee, G.J. Pierce. 2007b. 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.

Saha, S., J.N. Jenkins, J. Wu, J.C. Mccarty, O.A. Gutierrez, R.G. Percy, R.G. Cantrell, D.M. Stelly. 2006. Effects of chromosome-specific introgression in upland cotton on fiber and agronomic traits. Genetics 172:1927-1938.

Saha, S., D.M. Stelly, R.G. Percy, D.A. Raska, O.A. Gutierrez, J. Wu, J.N. Jenkins, J.C. Mccarty, Jr. 2004. Effect of chromosome substitutions from Gossypium barbadense L. 3-79 into G. hirsutum L. TM-1 on agronomic and fiber traits. Journal of cotton science 8.

Sas Institute Inc. (1989) SAS/STAT user's guide, SAS Institute Inc., Cary, NC.

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

Smith, C.W., J.T. Cothren. (1999) Cotton: origin, history, technology and production John Wiley & Sons, New York.

Steele, K., A. Price, H. Shashidhar, J. Witcombe. 2006. Marker-assisted selection to introgress rice QTLs controlling root traits into an Indian upland rice variety. Theor. Appl. Genet. 112:208-221.

Stelly, D.M., S. Saha, D.A. Raska, J.N. Jenkins, J.C. Mccarty, Jr., O.A. Gutierrez. 2005. Registration of 17 upland (Gossypium hirsutum) cotton germplasm lines disomic for different G. barbadense chromosome or arm substitutions. Crop Sci. 45:2663-2665.

Stephens, S.G. 1947. Cytogenetics of Gossypium and the Problem of the Origin of New World Cottons. Advances in Genetics Incorporating Molecular Genetic Medicine 1:431-442.

Wang, S., C. Basten, Z. Zeng. 2005. Windows QTL cartographer 2.5. Department of Statistics, North Carolina State University, Raleigh, NC.

Ware, J.O. (1951) Origin, rise and development of American upland cotton varieties and their status at present University of Arkansas College of Agriculture, Agricultural Experiment Station, Fayetteville, AR

Yu, J., R.J. Kohel, C.W. Smith. 2010. The construction of a tetraploid cotton genome wide comprehensive reference map. Genomics 95:230-240.

Zhang, J., W. Guo, T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (Gossypium hirsutum L. x Gossypium barbadense L.) with a haploid population. Theor. Appl. Genet. 105:1166-1174.

Zhang, J.F., Y. Lu, H. Adragna, E. Hughs. 2005. Genetic improvement of New Mexico Acala cotton germplasm and their genetic diversity. Crop Sci. 45:2363-2373.

Zhang, Z., J. Rong, V.N. Waghmare, P.W. Chee, O.L. May, R.J. Wright, J.R. Gannaway, A.H. Paterson. 2011. QTL alleles for improved fiber quality from a wild Hawaiian cotton, Gossypium tomentosum. Theor. Appl. Genet.:1-14.

Table 6.1. Biometric parameters of four mapping populations (F_2 were planted in 2007 and $F_{2:3}$ in 2008)

Trait	Population	Mean	Min	Max	Std. Dev.
ELO	Acala	3.86	3.10	5.30	0.35
	Delta	4.79	3.70	6.90	0.54
	Plains	5.19	4.30	6.40	0.44
	Eastern	4.47	3.40	5.80	0.42
MIC	Acala	3.87	2.95	4.89	0.39
	Delta	4.00	2.83	5.12	0.38
	Plains	3.98	2.92	4.97	0.40
	Eastern	3.99	3.27	4.81	0.29
SFC	Acala	7.45	6.40	8.60	0.37
	Delta	7.30	6.40	8.80	0.35
	Plains	7.38	6.40	8.80	0.40
	Eastern	7.06	6.30	8.30	0.28
STR	Acala	32.51	27.50	37.80	1.74
	Delta	31.34	27.40	34.60	1.14
	Plains	32.45	29.10	36.30	1.47
	Eastern	31.75	28.60	36.00	1.15
UHM	Acala	1.30	1.10	1.44	0.05
	Delta	1.28	1.11	1.41	0.05
	Plains	1.22	1.12	1.33	0.04
	Eastern	1.34	1.24	1.45	0.04
UI	Acala	84.03	80.80	87.00	1.13
	Delta	84.52	81.30	86.40	0.93
	Plains	84.18	81.50	86.40	0.94
	Eastern	85.34	81.50	88.30	0.99

Table 6.2. Pearson correlation coefficient among fiber quality traits in four populations.

Trait	Population	ELO	MIC	SFC	STR	UHM
MIC	Acala	0.15*				
	Delta	0.24**				
	Plains	0.24**				
	Eastern	0.11				
SFC						
	Acala	-0.16*	-0.06			
	Delta	-0.29**	-0.06			
	Plains	-0.30**	-0.29**			
	Eastern	-0.03	-0.05			
STR						
	Acala	-0.02	0.02	-0.45**		
	Delta	-0.40**	-0.04	-0.02		
	Plains	-0.18*	0.23**	-0.17*		
	Eastern	-0.21**	0.00	-0.16*		
UHM						
	Acala	-0.32**	-0.55**	-0.301**	0.12*	
	Delta	-0.45**	-0.54**	-0.25**	0.32**	
	Plains	-0.44**	-0.46**	-0.09	0.06	
	Eastern	-0.23**	-0.32**	-0.34**	0.17*	
UI						
	Acala	0.10	-0.02	-0.80**	0.40**	0.37**
	Delta	0.14*	-0.04	-0.77**	0.07	0.37**
	Plains	0.22**	0.16**	-0.76**	0.11	0.19**
	Eastern	0.03	0.05	-0.81**	0.16*	0.41**

* and ** represent significance with *P*-values of 0.01 and 0.001, respectively

Table 6.3. Summary of fiber quality QTLs identified in four mapping populations.

Trait	QTL	Gen	Population	Nearest Marker	LOD	a	PV(%)
Fiber Elongation							
	qELO-Chr5	F_3	Delta	DPL608	5.2	0.29	19.95
	qELO-Chr11	F_2	Acala	DPL270	2.1	0.12	2.35
		F_2	Delta		2.3	0.30	15.01
		F_3	Delta	DPL270	3.1	0.20	6.57
	qELO-Chr12	F_2	Plains	NAU2868	3.6	-0.29	22.69
	qELO-Chr24	F_2	Plains	BNL3860	3.1	0.22	12.49
	qELO-Chr25	F_2	Acala	NAU2714	2.1	0.17	5.41
		F_2	Plains		5.9	0.35	31.94
		F_3	Acala	CIR267	2.1	0.11	2.99
		F_3	Plains		7.9	0.26	48.62
Micronaire							
	qMIC-Chr11	F_3	Acala	DPL270	3.5	-0.12	3.53
		F_3	Plain		2.4	0.18	18.42
	qMIC-Chr12	F_2	Plains	NAU2868	4.4	-0.29	36.29
	qMIC-Chr15	F_3	Delta	JESPR063	3.1	0.27	8.35
	qMIC-Chr16	F_2	Eastern	JESPR012	2.8	0.16	3.03
		F_3	Delta	BNL1395	2.7	0.13	6.48
		F_3	Plains		2.0	0.01	2.63
		F_3	Eastern		3.5	0.13	8.96
	qMIC-Chr23	F_2	Eastern	DPL507	3.1	-0.09	9.30
	qMIC-Chr25	F_3	Plains	BNL2569	3.5	0.27	20.46
Short Fiber Content							
	qSFC-Chr5	F_2	Acala	BNL4078	3.1	0.18	12.30
	qSFC-Chr8	F_3	Acala	JESPR308	3.1	-0.12	9.54
Fiber Strength							
	qSTR-Chr5	F_3	Acala	BNL4078	2.3	0.57	6.61
		F_3	Delta		3.1	-0.36	3.48
	qSTR-Chr11	F_2	Acala	DPL528	3.0	0.67	14.75
		F_2	Eastern		2.1	-0.69	7.45
	qSTR-Chr15	F_2	Acala	STV129	3.1	-0.63	3.22
	qSTR-Chr16	F_2	Delta	BNL1395	6.2	-0.94	20.54
		F_2	Eastern		2.3	-0.64	6.27
		F_3	Delta	BNL1395	3.7	-0.34	6.15
		F_3	Eastern		3.6	-0.56	11.51
	qSTR-Chr25	F_2	Delta	CIR298	5.2	-0.57	17.70
		F_2	Plains		2.2	-0.44	1.03
		F_3	Plains	NAU2714	2.2	-0.73	6.99
Fiber Length							
	qUHM-Chr5	F_2	Delta	DPL608	2.1	-0.37	8.22
		F_3	Delta	NAU1221	3.2	-0.37	7.68

qUHM-Chr11	F_3	Acala	DPL209	3.1	0.24	3.73
qUHM-Chr12	F_3	Plains	NAU2868	3.8	0.28	4.79
qUHM-Chr16	F_2	Eastern	NAU1362	2.3	-0.48	11.18
	F_3	Delta	NAU1362	8.3	-0.77	15.13
	F_3	Eastern		4.1	-0.46	14.41
qUHM-Chr23	F_3	Plains	NAU3414	4.5	0.17	3.76
qUHM-Chr25-a	F_2	Acala	CIR267	3.6	-0.75	4.52
	F_2	Delta		3.7	-0.63	15.55
	F_2	Plains		5.7	-1.03	46.18
	F_2	Eastern		2.1	-0.47	4.49
	F_3	Acala	BNL827	3.5	-0.26	4.44
	F_3	Delta		4.4	-0.54	10.59
	F_3	Plains		11.6	-1.02	61.35
	F_3	Eastern		6.2	-0.55	19.50
qUHM-Chr25-b	F_2	Plains	BNL3103	2.0	-0.40	6.97
	F_2	Eastern		3.1	-0.32	9.94
	F_3	Delta	BNL3103	2.6	-0.33	9.22
	F_3	Eastern		3.6	-0.30	12.84
qUHM-Chr25-c	F_2	Eastern	JESPR302	2.3	-0.29	9.02
	F_3	Eastern	JESPR302	4.7	-0.41	48.62
Fiber Uniformity						
qUI-Chr8	F_2	Acala	JESPR308	3.1	0.71	5.41
	F_3	Acala	JESPR308	2.7	0.36	5.57
qUI-Chr15	F_2	Acala	STV129	3.6	-0.85	4.91
	F_3	Acala	STV129	2.2	-0.16	3.03
qUI-Chr18	F_3	Acala	BNL193	3.0	0.43	10.29
qUI-Chr24	F_2	Acala	NAU4045	4.4	1.05	13.67
qUI-Chr25	F_2	Delta	BNL3359	3.0	0.15	3.45

-ve a value indicate favorable allele originating from SL-883

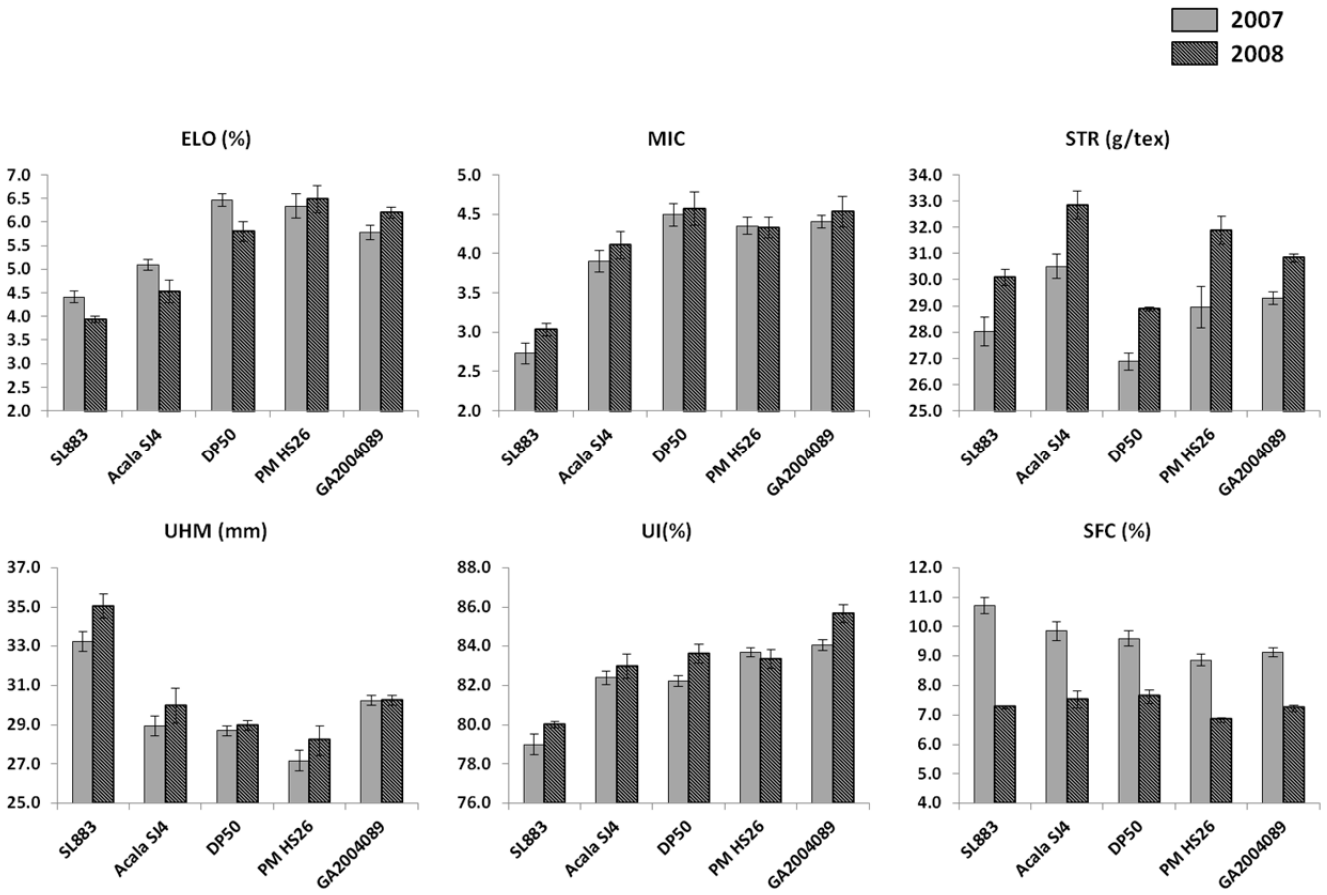


Figure 6.1. Mean fiber quality of mapping parents planted in 2007 and 2008.

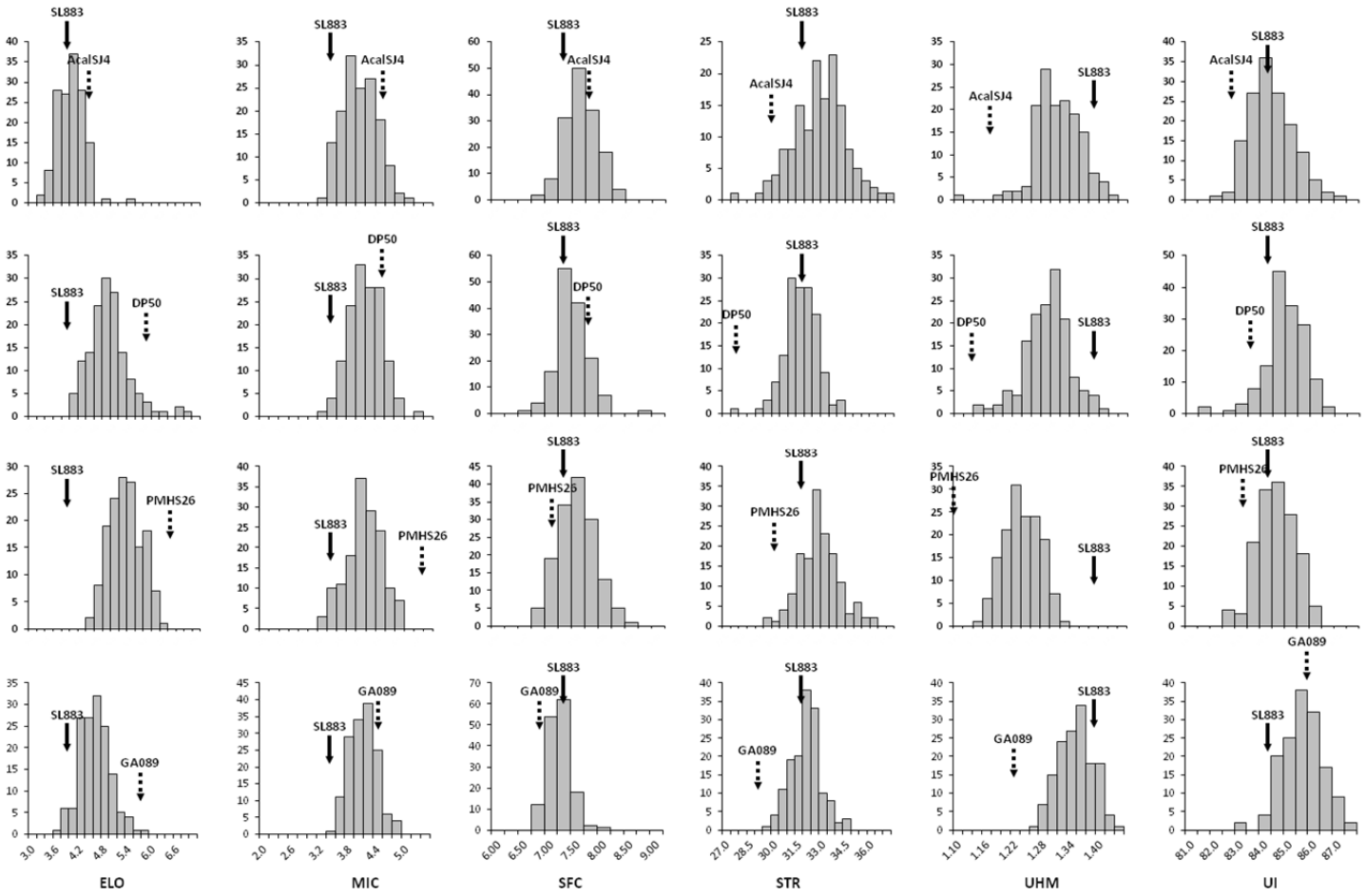


Figure 6.2 Distribution of fiber quality in $F_{2:3}$ generation of four mapping populations

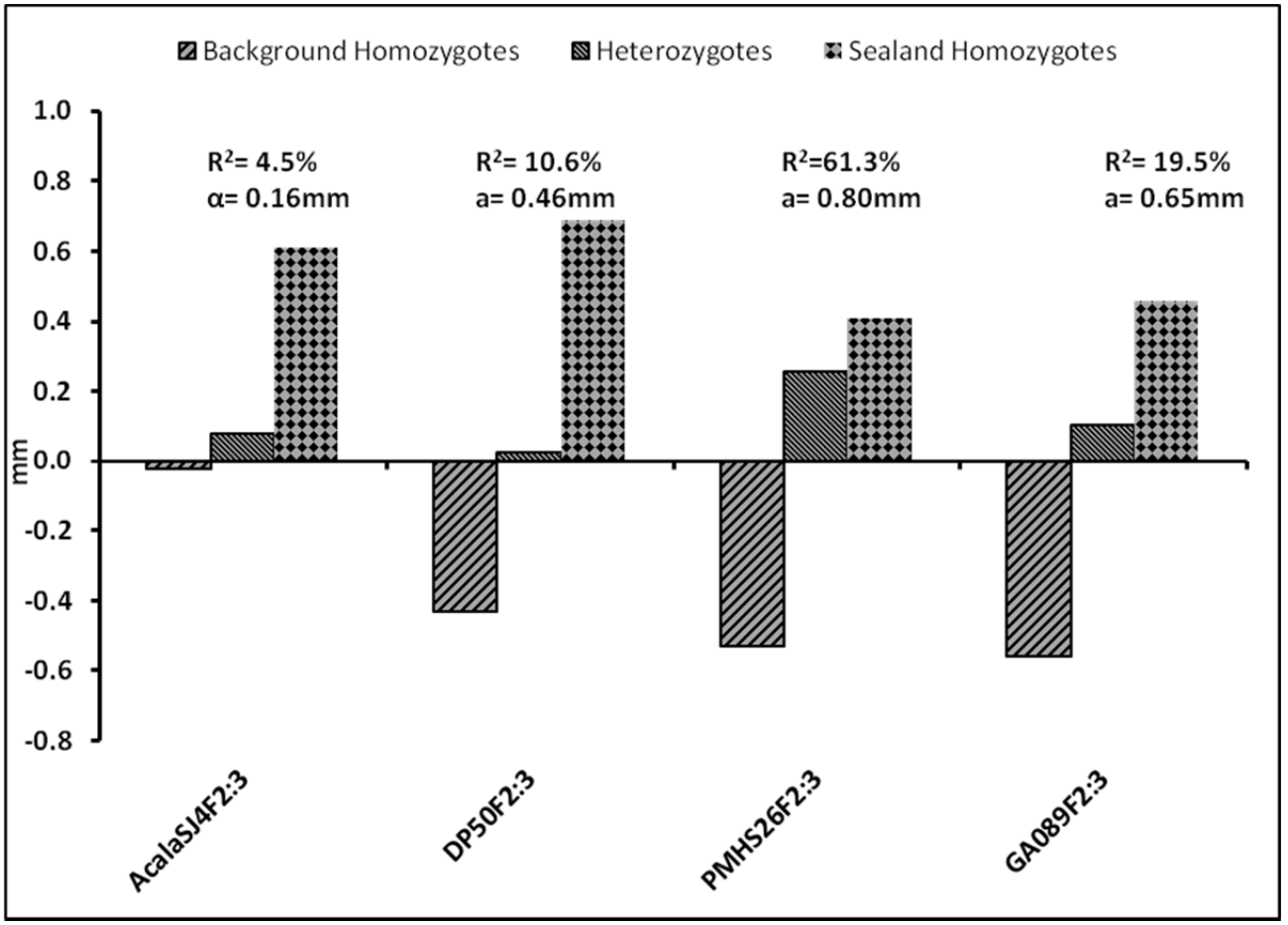


Figure 6.3. Effect of different background types on fiber length QTL *qUHM-Chr25-a*
R^2- Phenotypic variation explained
α- Effect of allele substitution

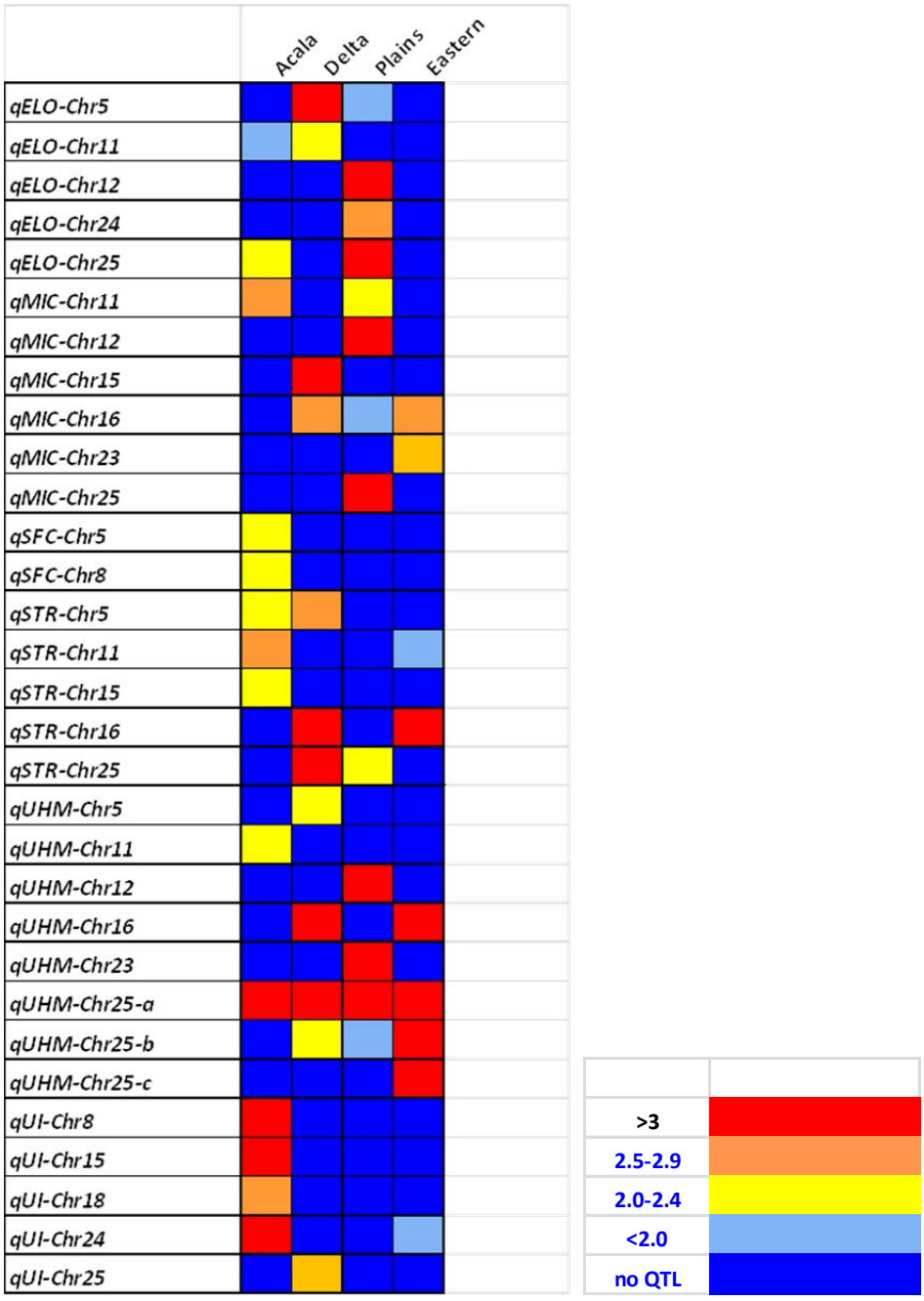


Figure 6.4. Heat map of the LOD score at which the QTLs were detected.

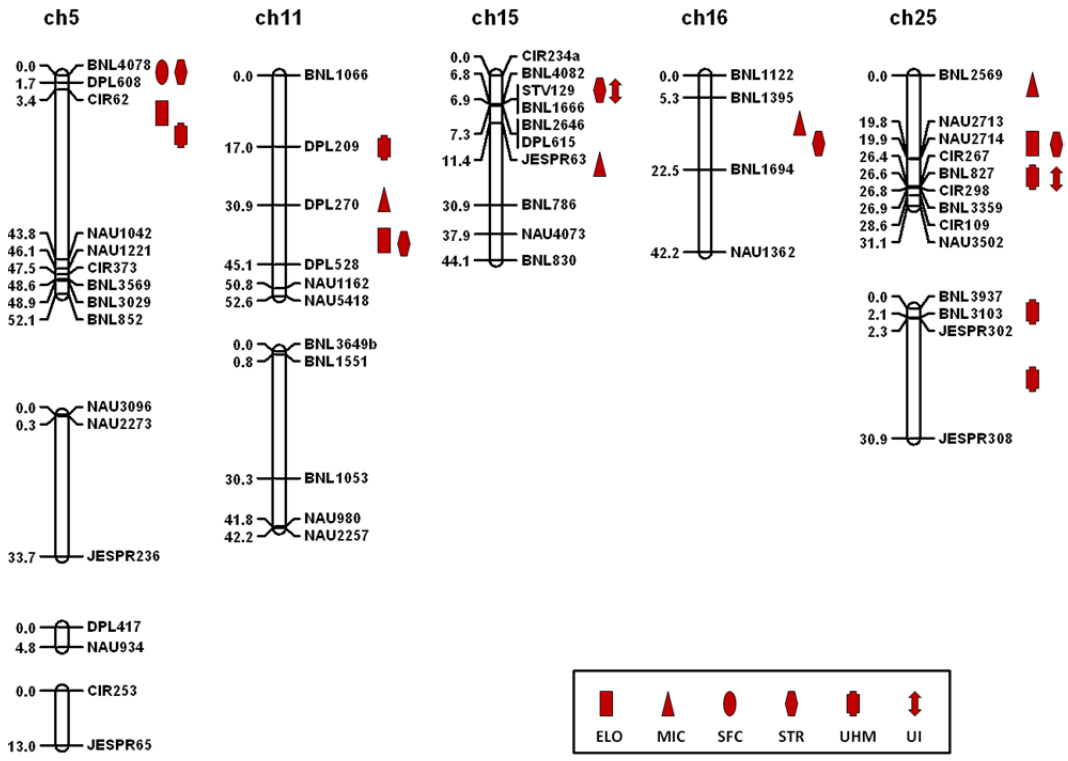


Figure 6.5. Mapping location of fiber quality QTLs on introgressed segments from Sealand 883

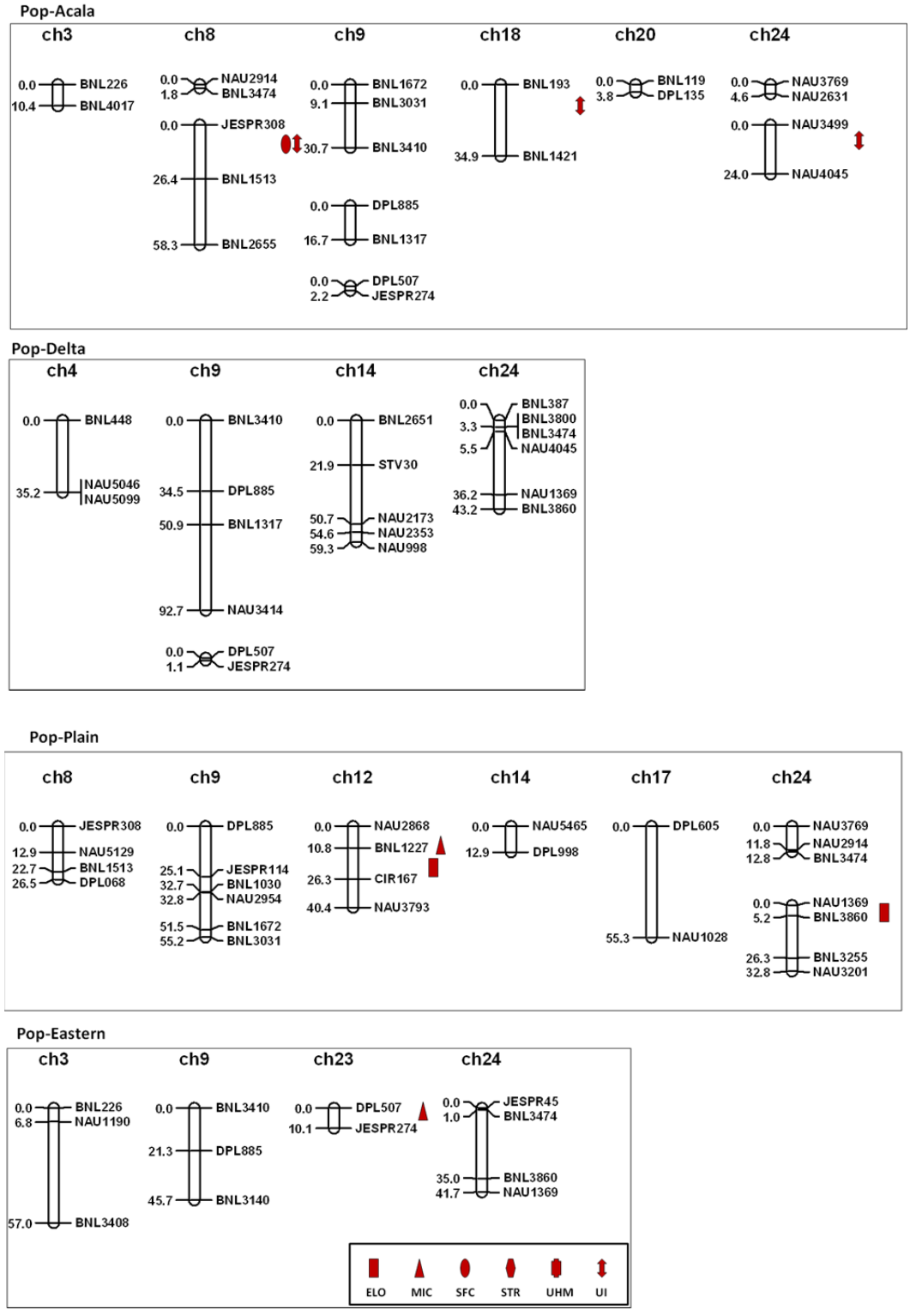


Figure 6.6. Mapping location of fiber quality QTLs on linkage maps developed for four mapping populations.

CHAPTER 7

CONCLUSION

The obsolete germplasm lines Sealand 542 and Sealand 883 of Upland cotton (*Gossypium hirsutum* L.) were evaluated for their potential to improve elite Upland germplasm of the US. These two lines were selected for genetic analysis in this study because of their pedigree records which indicated that these lines were developed at the USDA-ARS Pee Dee Experiment Station, Florence, South Carolina from an interspecific cross between *G. barbadense* (Sea Island) and Upland cotton. Phenotypically these lines possess significantly longer and finer-fiber compared to other Upland cotton lines suggesting that they contain stable introgressions of genes conferring these phenotypes from the *G. barbadense* parent. However, the extent of introgression, the chromosome locations, and the genetic effects of the introgressed chromatins are largely unknown.

In the first study (chapter 3), we identified a total of 22 putative introgressions including 12 introgressions on 7 chromosomes in the Sealand 542 genome and 10 introgressions on 5 chromosomes in the Sealand 883 genome. Putative introgressed regions of Sealand 542 covered 351cM (6% of the cotton genome) and 235cM for Sealand 883 (4% of the cotton genome). The introgression events did not show preferential sub-genomic affinity and were found to be equally distributed between At and Dt sub-genomes of tetraploid cotton, but the introgression pattern were unique to each Sealand line. Further QTL analysis revealed that the Sealand parents contributed favorable alleles for 11 of the 25 QTLs detected.

In the second study (chapter 4), we evaluated the efficacy of the fiber quality QTLs detected on the introgressed regions over two ($F_{2:3}$ and $F_{2:4}$) generations. A total of 24 fiber quality QTLs were detected, of which 10 QTLs were consistently detected in both $F_{2:3}$ and $F_{2:4}$ generations, including 8 QTLs which were also detected in first study. The robustness of the introgressed fiber quality QTLs suggests that they are ideal candidates for marker assisted selection.

In the third study (chapter 5), we mapped and validated a major fiber strength QTL on chromosome 24 with favorable alleles originating from Suyuan 7235, a high fiber strength germplasm line released by the Institute of Industrial Crops, Jangsu Academy of Agricultural Sciences, China. Interestingly, Suyan 7235 was developed through interspecific hybridization of *G. hirsutum* with *G. anomalum*. This QTL explained up to 40% of the total phenotypic variation and accounted for up to 22.8 kN m kg^{-1} increase in fiber strength. However, the efficacy of this QTL appeared to be greater when present in Sealand 883 background than when present in Sealand 542 background, reiterating the fact that the epistatic interaction with other loci in the background can affect the performance of the introgressed QTL.

In the final study (chapter 6), we investigated the effect of major different genetic backgrounds on the fiber quality QTLs from Sealand 883 before they are deployed in a marker assisted breeding program. We selected four cultivars, Acala SJ-4, Deltapine 50, Paymaster HS 26, and GA 2004089, which were of historical significance and represented Acala, Plains, Delta and Eastern types of Upland genetic background cultivated in the US. Of the 31 QTLs identified in this study, 8 QTLs were consistent over generations ($F_{2:3}$ and $F_{2:4}$) and were also common between two or more genetic

backgrounds. The effects of genetic background on detection and performance of the QTLs was observed at multiple loci. For example, the efficacy of the fiber length QTL present on the chromosome 25 of Sealand 883 was greatest when present in the Plains type genetic background where it accounted for up to 61% of the phenotypic variation and the additive effect was up to 50% greater than when present in Acala, Delta or Eastern type genetic backgrounds.

Overall, we were able to detect introgressed segments in the genomes of two Sealand lines and found that these introgression harbored favorable alleles for various fiber quality traits. Furthermore, we also validated the effects of these introgressed alleles over generations and across multiple genetic backgrounds. Together, these results determine that the Sealand germplasm lines are repository of fiber quality QTL alleles that can be targeted via marker assisted selection to improve elite Upland germplasm.