FIBER QUALITY IMPROVEMENT OF UPLAND COTTON (GOSSYPIUM HIRSUTUM, L.) FROM TARGETED INTERSPECIFIC SOURCES USING INTROGRESSION LINES AND EXOTIC UPLAND GERMPLASM WITH THE AID OF MARKER-ASSISTED INTROGRESSION.

by

ISMAEL NINO BROWN

(Under the Direction of Peng W. Chee and Andrew H. Paterson)

ABSTRACT

An apparent stagnation of fiber quality improvement in Upland cotton *(Gossypium hirsutum)* is likely due to a limited priority by breeders to improve fiber quality, exacerbated by limited genetic diversity for fiber quality within G. hirsutum. Increasing competition from man-made fibers such as polyester has put additional impetus for improvement of cotton fiber quality. Extensive genetic resources exist, however, from which valuable alleles for fiber quality can be mined. In addition to divergent, unadapted, or even wild G. hirsutum, valuable QTLs have been identified in other related species within the Gossypium genus. These interspecific sources of fiber quality represent an underutilized resource within the available gene pool of G. hirsutum. While numerous QTL have been mapped and identified in interspecific populations or introgression lines, few have been verified or have progressed to the point of introgression into adapted material. This dissertation seeks to address some of these topics by examining QTL identified from interspecific sources, their interaction with elite, divergent genetic backgrounds, their genotype by environment interactions, and their application in cotton improvement.

INDEX WORDS: Fiber quality, Gossypium, Interspecific introgression, QTL, Upland cotton

FIBER QUALITY IMPROVEMENT OF UPLAND COTTON (GOSSYPIUM HIRSUTUM, L.) FROM TARGETED INTERSPECIFIC SOURCES USING INTROGRESSION LINES AND EXOTIC UPLAND GERMPLASM WITH THE AID OF MARKER-ASSISTED INTROGRESSION.

by

ISMAEL NINO BROWN

B.S., Texas A&M University, 2006

M.S., Texas A&M University, 2011

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

2017

FIBER QUALITY IMPROVEMENT OF UPLAND COTTON (GOSSYPIUM HIRSUTUM, L.) FROM TARGETED INTERSPECIFIC SOURCES USING INTROGRESSION LINES AND EXOTIC UPLAND GERMPLASM WITH THE AID OF MARKER-ASSISTED INTROGRESSION.

by

ISMAEL NINO BROWN

Co-Major Professors:	Peng W. Chee
	Andrew H. Paterson
Committee:	Todd Campbell
	Brian Schwartz
	Jared Whitaker
	Dayton Wilde

Electronic Version Approved:

Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2017

DEDICATION

I would like to dedicate this document to the people I hold dearest in this world and who have sacrificed, supported, and encouraged me during my pursuit of this degree and in the completion of this dissertation.

To my dear Aiden, who has helped Mommy with his siblings while I spent the weekends typing rather than playing outside. To my sweet Chloe, whose pronunciation of the word "dissertation" was unintelligible when I began, but as the years have slipped by, is as clear as can be. To my dear Zachary, whose constant questioning of whether Daddy "will *ever* be done with his dissertation," encouraged me to work harder and faster. To my sweet little Etta Lowe, who has seldom seen weekends where Daddy didn't have to work. To my beautiful babies whom I adore with all my heart, I dedicate this dissertation to you, because it is for you that your Mommy and I have worked so hard, and it is to you that we dedicate ourselves.

To my sweet, adoring, beautiful, wife Emily, I dedicate this dissertation. The completion of this project has certainly been a team effort, and without you none of this would have been possible. You have been my unflappable and fervent supporter for these many years. You have given me the strength to keep going when down, and the encouragement to succeed when doubtful. I love you Emily, and I am so thankful to have you by my side in this life. You and I truly could not have made it to this point without your unwavering spirit. I dedicate this dissertation to you, as it is equally your own achievement. I love you Emily Brown!

To my beautiful little family, I dedicate this to you.

And guess what… Daddy gets to play this weekend!

ACKNOWLEDGEMENTS

There are many people that have been of tremendous help in the pursuit of my PhD, and I am grateful for the chance to thank them here. First of all, I would like to thank my co-major professors, Drs. Peng Chee and Andy Paterson. Dr. Chee, thank you for accepting me into your program and teaching me so much about applied molecular genetics in cotton breeding. The years spent in your lab were an incredible learning experience and time of growth. The fishing trips were pretty great too! Dr. Paterson, thank you for hosting me during my time in Athens; the short time I spent there was a valuable learning experience. To the members of my graduate committee, Drs. Todd Campbell, Brian Schwartz, Jared Whitaker, and Dayton Wilde, thank you for your guidance and collaboration.

Thank you to the people of the Molecular Cotton Breeding Lab at the University of Georgia for all of the help you have given me over the years. To Jennifer McBlanchett, thank you for all your help in the field and greenhouses, ginning and delinting. To Rippy Singh, thank you for your help and guidance in the lab, an environment in which I was completely lost. To Dr. Pawan Kumar, thank you for being a helpful and collaborative lab mate and colleague, whose advice was invaluable. To Dr. Ed Lubbers, thank you for your help with Agrobase matters and data curation. Thank you to all of the student workers who helped collect and process my samples over the years.

To my colleagues and managers at Bayer, thank you for supporting me and giving me my dream job before graduating. To Jody, Clint, Jeff, and Elliott, thank you for putting up with me during these years.

Thank you also to my sweet in-laws who helped Emily keep the kids occupied on the many weekends I spent working, and for the support they gave our family during this trying time. To Mrs. Herzog, who has been a great and supportive influence on our little transplanted family. Thank you also to the Schwartz's for their friendship and support over the years. Finally, I want to thank my parents for instilling in me a strong work ethic and abhorrence for the idea of quitting. I love you Mom and Dad, and appreciate everything you have taught me and I appreciate above all the great example you have been to me in how to live one's life.

Thank you everyone!

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

CHAPTER 1

INTRODUCTION

Cultivated and valued by humans for millennia, members of the genus *Gossypium*, collectively termed cotton, have become an important commodity crop that influences the daily lives of people across the globe. Grown for the valuable seed epidermal trichomes or hairs that are produced on the seed surface, two of the New World, tetraploid species, *G. hirsutum* and *G. barbadense* account for the vast majority of international cotton acreage. In 2015, *G. hirsutum*, referred to as Upland cotton, accounted for over 98% of the 3.5 million hectares of cotton planted in the United States, the remaining acreage planted to *G. barbadense*, commonly termed Pima, Egyptian, or extra-long staple (ELS) cotton (USDA-NASS, 2016). Over 101 million 218 kg cotton lint bales were produced globally in 2015 (USDA-FAS, 2016). India and China led world production of the crop with 27.8 and 23.8 million bales respectively, followed by the US, Pakistan, and Brazil with 12.9, 7.2, and 6.7 million bales, respectively (USDA-FAS, 2016). In the US, the 2014 crop was worth an estimated $5.1 billion (USDA-NASS, 2016), $964 million of which was produced by growers in the state of Georgia. Among the various industries in the state of Georgia, agriculture is the leader, with cotton the leading row crop for both acres planted and farm gate value (UGA-CAES, 2014).

Due to a combination of improved agronomics, genetic gain from breeding, and the introduction of GMO cotton into the marketplace, national average Upland cotton lint yield has seen a 70% improvement from 500 kg/ha in 1975 to 850 kg/ha in 2015. Improvements in fiber quality, however, have not maintained a similar pace of improvement. From 1980, fiber bundle

strength (FBS) increased from 235 kN m kg-1 to 294 kN m kg-1 in 2012, a 25.1% increase encouraged by the industry's need for stronger fibers to feed open-end and rotor spinning facilities, but an increase of much lesser proportion. Fiber upper half mean length (UHML) has increased only 5.6% in nearly 40 years, from 26.7 mm in 1975 to 28.2 mm in 2014 (Incorporated, 2015). The stagnation of fiber quality improvement is likely due to a limited priority by breeders to improve fiber quality, exacerbated by limited genetic diversity for fiber quality within *G. hirsutum*. Increasing competition from man-made fibers such as polyester has put additional impetus on improvement of cotton fiber quality. It is of some concern, however, that *G. hirsutum* germplasm shows little genetic diversity from which to make significant gains for fiber quality improvement (Bertini, et al., 2006; Fang, et al., 2013; Van Becelaere, et al., 2005).

Extensive genetic resources exist, however, from which valuable alleles for fiber quality can be mined. In addition to divergent, unadapted, or even wild *G. hirsutum*, valuable QTLs have been identified in other related species within the *Gossypium* genus. These interspecific sources of fiber quality represent an underutilized resource within the available gene pool of *G. hirsutum*. While numerous QTL have been mapped and identified in interspecific populations or introgression lines, few have been verified or have progressed to the point of introgression into adapted material. This dissertation seeks to address some of these topics.

LITERATURE REVIEW

Cotton Fiber Quality Traits and Methods of their Measure

In recent decades, man-made fibers such as polyester have taken significant textile market share from natural fibers such as cotton. In 1990, cotton and other natural fibers accounted for over 60% of the global textile production market and in 2000 it was down to 45%. In 2010, natural fibers accounted for 36% of world fiber consumption, with cotton accounting for 32.9% of the total consumption (FAO-ICAC, 2013). This rapid decline in cotton market share can be tied to recent advances in man-made fiber technologies as well as the historically high price of the raw cotton fiber and the processing costs associated with the natural product vs. man-made synthetic fibers. For cotton to remain competitive with man-made fibers, it is important that the cotton breeding community improve the fiber quality genetic potential of the cotton crop. Improving fiber length, strength, elongation, fineness, and uniformity are crucial to the continued viability of US cotton farming operations.

The seed trichomes, or cotton fibers, have characteristics intrinsic to them that determine their value for spinning into yarns from which to make textiles. All fibers, natural or man-made, have these characteristics. These include, but are not necessarily limited to, the length of the fiber, the fiber strength, how much those fibers will stretch before breaking, and fiber diameter. These parameters, as well as the chemical composition and the physical nature or three dimensional shape of the fibers, all contribute to the selection of a fiber as a suitable source for spinning into yarn. They also influence the characteristics of the yarn, the textile produced from the yarn, and the costs associated with processing the fiber into yarns.

The production of a cotton textile begins when the individual fibers of a cotton bale are spun into yarn. The fibers from many bales are combined, blended, cleaned and arranged so that

they are parallel to each other through a process called combing and carding. Once this process is complete, the fibers go through several series of drawing or stretching before being spun together to produce a yarn. Cotton, being a natural fiber requires multiple rounds of such processing, which can cause fiber breakage. Fiber length is important for producing thin and strong yarns, whereas strong fibers are required to withstand the rigors of mechanical processing and also helps to create a strong yarn. Uniformity of length is also important for the efficient production of yarn at an industrial scale. After being spun to create yarn, the yarns are woven to produce textiles or twisted together to produce string. Yarns produced from long strong fibers are used in the production of higher end products such as fine dress shirts or high thread-count sheets, whereas shorter fibers are used to produce coarse yarns used in towels and denim jeans.

Chemical and physical composition of fibers influences their spinning characteristics as well as their dyeing and finishing qualities. For instance, cotton fiber is 99.5% cellulose, a polysaccharide linear chain of hundreds to thousands of glucose molecules. Primarily composed of α-cellulose, the crystallinity of the cellulose as well as the non-cellulosic components affects dye uptake. The chains of cellulose are deposited to constitute the secondary cell wall. Immature fibers have a lesser degree of this secondary cell wall deposition, thereby making immature cotton fiber weaker than a fully mature fiber and poorer dye uptake.

Cotton fibers were spun by hand prior to the Industrial Revolution, which made fiber length of great importance. Today, the different technologies used in spinning yarns have different priorities and preference for fiber quality characteristics. Ring spinning mills, for instance, which produce some of the finest yarns, capable of producing very high thread count and therefore valuable fabrics, rank fiber length as the number one priority, followed by strength and then fiber fineness. However, rotor spinning mills require fibers with high strength

foremost, followed by fineness and lastly fiber length (Deussen, 1993). Rotor spinning produces a more coarse, and hairier yarn, but production speed is much faster.

By the turn of the 20th century, cotton fiber quality was judged subjectively by highly trained individuals known as "classers," who would assign the various grades of cotton to help textile mills in determining fiber quality. However, their method of judging fiber quality visually and by hand was highly variable and resulted in discrepancies between markets. In the 1920s, the United States Department of Agriculture (USDA) introduced a standard method of classification for cotton fiber quality and disseminated these standards to the various classers across the nation. In 1965, the micronaire (MIC) machine was introduced, and by the 1990s, the classing of all US cotton was mandated to be run on a high volume instrument (HVI) classing machine (Ramey Jr, 1999). The HVI machine has the capability of measuring fiber length (UHML), fiber bundle strength (FBS), elongation (ELONG), color, trash content, and micronaire. All of which are very important to the textile mills for determining the cotton bales they will purchase.

Fiber Length

The length of an individual cotton fiber is very important to the yarn and the textile it will be used to produce, and has long been considered one of the most important aspects of fiber quality for yarn strength and quality (Brown, 1938; Perkins, et al., 1984). Fiber length is positively correlated with yarn tenacity and yarn elongation (Benzina, et al., 2007; Ramey, et al., 1977). Additionally, the length of the fiber determines the fineness of the yarn capable of being produced. When fibers are spun into a yarn to produce textiles, they are arranged in a linear fashion so that they all face the same direction. These linearized fibers are then twisted together in the spinning stage. The applied twist creates an interlocking of fibers and frictional force

between individual fibers at their points of contact. Longer fibers, after twisting, have more points of contact along their length with their neighboring fibers, giving them more interlocking and frictional forces. This creates a yarn that requires more force to pull apart. So the longer the fibers used to make a yarn, the stronger that yarn will be. As the average length of a cotton fiber sample increases, it takes fewer and fewer fibers per cross-section of yarn to produce an equally strong yarn, and if shorter fibers are used it takes more fibers per cross section to produce a yarn with equal strength. Because it takes fewer long fibers per yarn cross section to produce an adequately strong yarn, yarns made with longer fibers generally have a smaller diameter. Yarns produced with long fibers are strong and usually quite fine, which is desirable in high-end textile products such as high thread count bed sheets and fine dress shirts. Yarn tenacity, or the amount of force required to break a given yarn, is highly correlated with fiber length (r=0.768) (Bedez Üte and Kadoğlu, 2014)

Fiber length is a very important trait to the textile industry for the preceding reasons. However, it has been shown to have a negative correlation with yield (Azhar, et al., 2004; Meredith and Bridge, 1971; Miller, et al., 1958; Scholl and Miller, 1976; Shen, et al., 2007; Smith and Coyle, 1997). Along with this negative correlation with yield, it is against a grower's best interest to grow a variety with good fiber quality if it will not yield competitively. Under the current US pricing model, it takes a very high premium paid for quality to make up for a small increase in cotton lint kg per hectare (kg/ha). While discounts are applied to cotton bales with low quality, the grower can make up for this situation by producing more bales per acre. Hence, market forces driven by growers are not always aligned with the market needs of the textile industry. Breeders have attempted to address the shortcomings in fiber quality, especially length without sacrificing yield. However, it is because of the negatively correlated relationship

between many fiber quality traits and yield that breeders in the public and private sectors have had such difficulty in improving cotton fiber quality without harming yield potential.

Fiber length is measured in several different ways. Cotton is a natural product, and as such it is important to think of a fiber sample as a population of fibers, and so the entire fiber distribution must be described to fully understand the nature of the sample. Within each sample even from the same plant there are fibers of greater length, fibers of shorter length, some that are strong, weak, etc. HVI provides what was previously described as the "fibrogram" by Hertel, which is the histogram of the second integral of the length frequency (Hertel, 1940)(1940). Three measurements are derived from this fibrogram, the mean length (ML) which is simply the average length of fibers in the sample, upper-half mean length (UHML), which is the mean length of the longest 50% of fibers in the sample, and the Uniformity Index (UI) which is the ratio of the ML compared to the UHML, which gives some idea of the fiber length distribution. Uniformity ratio (UR) can also be discerned from the fibrogram, which is the ratio of the 50% span length over the 2.5 span length. Because the industry primarily looks at UHML to describe fiber length, UR is lesser used than UI, though the two measurements are very similar. For ring-spinning, uniformity of fiber length is an important determinant of yarn quality and spinning performance (Deussen, 1993)

Short fiber content (SFC), defined as the percentage of fibers by weight with length less than 12.7mm (Behery, 1993), is important to the manufacturer as high levels of SFC can cause processing issues such as excess fiber waste, decreased yarn strength, increased ends-down, yarn defects and decreased productivity (Backe, 1986). SFC has increasingly become more important as a problem in cotton textile production as modern textile production speeds increase and the adverse effects of short fiber become more apparent. Sources of SFC include genotype, the

interaction of the genotype with environment, and mechanical breakage during harvesting, cleaning, ginning, and processing.

HVI measurement of fiber length and other traits has been the standard applied to all bales produced within the US, however, more precise and descriptive measures are now available with the Advanced Fiber Information System (AFIS) by Uster Technologies.

Along with measuring the average length of the total number of fibers scanned (Ln) and by weight (Lw), AFIS also measures the upper quartile length, length coefficient of variation (CV) by number and weight, SFC by weight and number, and the length upper percentiles such as length by number of the upper 2.5% and upper 5%. These measurements provide a much more thorough picture of the entire fiber length distribution. As fiber quality becomes increasingly important for textile manufacturers, AFIS measurement of cotton fiber quality will likely become more commonplace. Currently, the wide scale adoption of this technology is limited by its low speed and high cost, especially in breeding where large numbers of samples are processed and the testing is cost prohibitive.

Fiber Strength

The strength which is required to break a given sample of fibers has a large effect on the yarn that will be produced, the textile that it will be used in, but also how well the fiber handles the harsh mechanical rigors of processing. It has been noted that the ginning process creates a large amount of fiber breakage that leads to shorter fiber and/or high SFC (Krifa, 2006; Schenek, et al., 1998; Schneider, et al., 1994). Stronger fibers are better equipped to make the journey from the cotton boll to the shirt without being damaged or broken into shorter fibers. After ginning, the fibers are exposed to further mechanical rigors during the spinning and weaving processes, making fiber strength crucial to efficient textile manufacture (Faerber, 1995; Patil and Singh,

1995). Individual strength of a fiber also contributes to yarns with high strength (May and Taylor, 1998; Meredith, et al., 1991). Although fiber length and length uniformity are arguably the most important factors in yarn strength, individual fiber strength is very close in importance and has been given increasing importance along with fiber fineness (FINE) by textile manufacturers (Moore, 1996).

High strength fibers allow for faster production speeds (Perkins, et al., 1984), and while it is an important characteristic for all yarn production, it is especially important for open-ended/rotor spun yarn production (Deussen, 1993). Yarn tenacity has been shown to be highly correlated with fiber strength (r=0.734) (Bedez Üte and Kadoğlu, 2014). Fiber strength has a greater degree of importance in open-end yarn spinning vs. ring spun due to the yarn structural differences and the mechanical differences of the manufacturing processes.

Current estimation of fiber strength is conducted during HVI analysis using the same beard or bundle of fibers used in estimation of fiber length, which is of a standard weight. The bundle is grasped on each end by two pairs of metal clamps, which are then spread apart at a constant speed, and the force required to break the bundle is recorded. This breaking force is related back to the weight of the fiber bundle as kN m kg^{-1}. This fiber bundle strength (FBS) is a good measure of the strength of fibers, however, it can be biased toward fine fibers, since more fine fibers compose a bundle of equal weight, which leads to an over-estimate of the fiber strength due to the larger number of fibers in the bundle.

Fiber Elongation

Fiber elongation (ELO) refers to the amount a fiber stretches before they rupture or break. Measurement of this characteristic is achieved during the fiber bundle strength measurement during HVI analysis or by Stelometer. As the bundle of fibers is pulled apart to determine the

amount of force needed to break, the machine also measures how far the sample stretches prior to breakage. ELO is thought to confer better processing ability to fiber. The ability to stretch gives some protection during the rigorous processing forces incurred during ginning, bale opening, carding, drafting, spinning, weaving and knitting. Faulkner et al. (2012) found that fiber from cultivars with similar strength and length, but different ELO values had different work-to-break values. Described as the total force required to rupture a fiber bundle, work-to-break is a function of strength and ELO combined, and is a more accurate predictor of spinning performance (Benzina, et al., 2007). Indeed, cultivars with lower strength but better ELO values required more work-to-break. While fiber strength is an important aspect of spinning performance, fiber ELO is shown in this study to be of increased importance to yarn production and spinning efficiency (Faulkner, et al., 2012).

Unfortunately, ELO as measured by HVI has some problems. Although the measurement is repeatable with acceptably low coefficient of variation (CV), the data isn't comparable between machines because as of now there are no methods to calibrate and thereby standardize the measurement. So while one machine might produce consistent results, the results are impacted by daily and perhaps hourly fluctuations in ambient environmental conditions. It is because of this unreliability that ELO is not included in classing reports. Work is being done in an effort to standardize the measurement of this important fiber quality so that breeders can improve fiber ELO and textile producers will have a metric to help them select bales for purchase (Benzina, et al., 2007).

Fiber Fineness and Maturity

Intuitively, finer fibers allow for more fibers per cross-sectional area of a given yarn thickness, which confers that yarn greater strength. This idea is presented by Faerber and Deussen (1994)

as the spin limit which is the minimum yarn count, or yarn thickness, that can be produced from a sample of fibers and still produce a quality yarn with acceptable spinning efficiency. Finer fibers and more fibers per cross section of yarn allows for more frictional forces among yarn fibers, which increases strength of the yarn, reduces the required level of yarn twisting, and can greatly increase spinning efficiency and productivity during manufacture. Fiber maturity, which is the degree to which the secondary cell wall has thickened over the course of fiber development, is also a very important fiber trait. Immature fibers cause poor performance during dye uptake, decreased fiber strength and increased breakage during processing, defects in the fabric and waste during textile and yarn manufacture. Immature fiber also causes neps, which are little knots of entangled fibers that develop in yarns and can be seen as fabric defects.

There are two primary factors that influence the fineness of a fiber, the genetically determined fiber perimeter, and the developmental maturity of the fiber, which has both a genetic component as well as a very large environmental component. We use perimeter to describe the width of a fiber because after the boll opens and fiber cells begin to dry out, the fiber cell collapses on itself, transforming from a circular cross section, to one that resembles a half or quarter moon. After primary cell wall formation of the cotton fiber cell, successive layers of crystalline cellulose are deposited to the interior forming the secondary cell wall. The degree to which the secondary cell wall is thickened by these successive layers of cellulose determines what is generally referred to as fiber maturity. Assuming fiber perimeter is constant for several fiber samples, one with low maturity or low levels of secondary cell wall deposition will have weak fibers that will create a weak yarn. Fiber with very high maturity and overly thick secondary cell wall deposition will result in fibers that do not collapse in on themselves when the fiber cells dry out, thus creating a fiber that is circular in cross section that does not spin well and

creates a weak yarn. Fibers with optimum maturity have sufficiently thick secondary cell walls, have strong fibers that still collapse after drying, and are spun to create strong yarns.

Fiber fineness measurement is most commonly conducted during HVI and AFIS analyses though a few other techniques are used at lower frequency. HVI attempts to quantify fineness and maturity together through micronaire (MIC) reading. MIC is achieved by applying a standardized amount of pressurized air to a standard weight of fibers, the resistance of that air caused by the sample of fibers gives us the MIC value. Thinner fibers provide more resistance to air flow than coarse fibers. This measure is confounded however, by fiber maturity, which can also influence MIC. Immature fibers are improperly considered fine fibers, whereas overly mature fibers are considered to be coarse. Immature fiber, which generally has low MIC values less than 3.5 cause neps and weak spots in the yarn, as well as reduced dye-uptake due to the limited cellulose deposition. High MIC cotton, which is usually considered to be >5.0, generally does not spin well and cannot be used to produce high quality yarns (Hake, et al., 1996). If relative fiber or crop maturity is known, for instance, if one has a check variety with which to compare readings, then one can ascertain the relative level of maturity or fineness using MIC. Without any information regarding maturity or fineness, cotton with low MIC, can result from immature fibers or fibers that are inherently fine (May, 1999). Coarse fibers, or thin fibers with excessive secondary cell wall thickening are represented by high MIC readings (Hsieh, 1999). Therefore, HVI MIC is an imperfect measurement of fiber fineness.

Cross-sectional analysis of cotton fibers is the most accurate method to determine fineness and maturity and the international standard (Adedoyin, et al., 2011), however, it is labor intensive, slow, and very expensive (Paudel, et al., 2013). AFIS fineness (FINE) and maturity (MAT) readings are determined by measurements taken on individual fibers. The machine

analyzes light reflected off fibers flowing past optical sensors, which it uses to determine circularity of the fiber and cross-sectional area from which maturity or degree of secondary cell wall thickening is calculated (Bradow, et al., 1996). The cottonscope method, which has gained some level of notoriety in recent years, is an automated system that produces snippets of fibers from a sample and suspends these snippets in water. Several images of the suspension are taken and are analyzed to determine maturity and fineness. The cottonscope gives highly correlated results to cross-sectional image analysis, however, it too is time consuming and expensive (Paudel, et al., 2013).

Foreign Matter and Trash Content

Although foreign matter and trash content are not fiber quality parameters intrinsic to the cotton fiber themselves, it can be considered a fiber quality parameter of a cultivar, a region, or an individual bale. Foreign matter and trash content are very important to textile manufacturers, and as such should be treated as an important aspect of cotton fiber quality. Due to the highly automated nature of cotton processing, one of the biggest problems for spinners is cleanliness of the raw product and high foreign matter content (ITMF, 2014). Keeping trash and foreign matter levels to a minimum should be a goal of breeding programs, production systems, and ginners as well as spinners and weavers.

Common contaminants include leaf, bark and stem particles, burs, seeds, seed fragments, motes, grass, sand, oil and dust. The majority of these contaminants are removed during the ginning process, however, not all of it can be affectively or efficiently removed entirely by mechanical cleaning. Thus, spinners end up with contaminated materials, which cause significant problems at the mill. The recent past has seen some improvement in lint cleanliness,

however, the most recent survey conducted by the International Textile Manufacturers Federation (ITMF), indicates increases in cotton bale contamination.

Traits such as leaf shape and leaf pubescence can have a significant effect on leaf trash. Both of which are highly heritable traits controlled by few genes, and have been shown to significantly reduce motes, small-leaf trash, and improve grades (Novick, et al., 1991). These results suggest that selecting for smooth or semi-smooth leaf and okra leaf shapes should improve lint trash from leaf material.

Genetics of Fiber Quality

Each of the major fiber quality properties discussed previously are under some level of genetic control, however, they are quantitative traits that exhibit a great deal of variation in their phenotypic expression. Each of these traits are controlled or at least influenced by many genes, thus the environment in which they are grown plays a large role in their expression. Heritability is the fraction of this phenotypic variance that is controlled solely by genetic variance, and studies have shown that fiber quality traits tend have a higher heritability than such traits as yield.

Fiber length, which can be measured a number of different ways as discussed previously, has a generally high level of heritability. In May's extensive review (1999), it was pointed out that genetic variance is generally higher than non-genetic variance for fiber length (May, 1999). Most of the research done to date has shown that Additive genetic variance is usually higher than non-additive variance (Green and Culp, 1990; Miller and Marani, 1963; Miller and Rawlings, 1967; Tang, et al., 1993). Quisenberry (1975) and May and Green (1994) found non-Additive genetic variance to be larger than additive variance for length in the populations used in their experiments (May and Green, 1994; Quisenberry, 1975). However, as

in the case of May and Green (1994), May (1999) suggested that this could have been due to an exhaustion of additive genetic variance in the Pee Dee breeding program due to several decades of selection for fiber length. In measuring response from selection, both mass selection and pedigree-based selection were successful in increasing genetic gain (May, 1999). In an examination of multi-year, multi-site variety trials, Campbell and Jones (2005) found that the environment (E) explained 47% of the total variation for fiber length, but that genotype (G), and the interaction of genotype by environment (G x E) explained 45% and 8%, respectively. The high heritability of fiber length, especially UHML, would suggest that extensive environmental replication, like that which is typically done for yield testing, is probably unnecessary and gains can be made through breeding without extensive replication.

Genetic variation for fiber length uniformity, as measured by UI and UR does exist, however, the environment plays a large role in its expression, but not so much that it limits gains from selection (Meredith Jr, et al., 1996; Meredith, et al., 1991). Of the genetic variance reported in May and Green (1994), 80% was attributed to additive genetic variance, and 20% to additive x additive genetic variance, giving further evidence that selection should be effective in improving length uniformity. SFC, and the genetics underlying this complex trait is poorly understood. It has been suggested that breakage during ginning and textile processing due to weak fibers could be a large cause of SFC (Behery, 1993). FBS and SFC by weight (SFCw), indeed showed a negative correlation of -0.25 in an interspecific *G. hirsutum* x *G. barbadense* F2 population (Mei, et al., 2004). Jiang et al. (1998) identified a QTL for SFCw in an interspecific mapping population that explained 14.7% of the total variance for the trait (Jiang, et al., 1998). However, in a *G. hirsutum* intraspecific population, Meredith et al. (1996) showed little if any correlation between FBS and SFC (r-=-0.02). These conflicting results could be due to

population composition, i.e., inter- vs. intraspecific populations. Additionally, breeding studies on SFC are confounded by the lack of mechanical cleaning steps on laboratory gins, which likely causes us to underestimate SFC. UI, UR, and SFC, while components of fiber length, seem a little more difficult to modify than UHML, however, as mills are beginning to put more pressure on the industry to provide a more uniform product, these components of fiber quality will likely become more important.

FBS is largely under genetic control with low GxE interaction, mainly additive genetic effects, and high heritability (May, 1999). Estimates of the number of genes controlling FBS have ranged from 5 to 14 genes (Self and Henderson, 1954; Tipton, et al., 1964), evidence of the trait's quantitative nature. There are likely many more genes affecting the trait as mapping studies to date have identified several QTL in various populations that affect FBS. Interestingly, in a few examples, FBS has been shown to be controlled by few major genes, however, these examples all involve Beasley's Triple Hybrid (Beasley, 1940a) somewhere in their pedigree (May and Green, 1994; Meredith Jr, 1992; Meredith, 1977; Richmond, 1950). It is likely that this source of FBS derived from Beasley's Triple Hybrid, a three-species interspecific hybrid, is novel to *G. hirsutum*. If there occurs little or no variation in *G. hirsutum* germplasm, any allele of interspecific origin would produce a large effect on the phenotype. It also presents an interesting possibility of being a novel locus, heretofore unrepresented in *G. hirsutum* germplasm before the introgression from *G. thurberi* and *G. arboreum*. The negative correlation between FBS and yield (Culp and Green, 1992; Culp, et al., 1979) have made improvement of the commercial crop difficult for breeders whose primary customer are the growers, who require quantity over quality of fiber. The market is, however, moving toward improvement of fiber quality being a major concern, which will impact growers increasingly in the future.

Interestingly, HVI measurement of FBS is unable to differentiate small differences in FBS between experimental lines (Cooper, et al., 1988; Green and Culp, 1988). This is an important limitation to consider in breeding for improvement of FBS because these small differences are within the range of improvement made by breeders over the course of years of selection, and could be potentially limiting sustained genetic gain for this trait. Because of this limitation in the ability of HVI to discriminate minor differences, it might be important to replicate over large numbers of replications and environments in contrast to UHML determination.

An important trait especially for open-end yarns, fiber ELO, or the amount a fiber stretches before breaking, improves yarn evenness, strength and hairiness, and improves the yarn's resilience to weaving and textile production (Backe, 1996). HVI ELO is highly correlated with yarn work-to-break, which is indicative of a yarn's performance during processing and textile manufacture (Faulkner, et al., 2012). Ring-spun yarn performance was found to benefit from higher ELO values, with significant though moderate to low correlations with yarn evenness, hairiness, and breaking elongation with r=0.162, 0.211, and 0.202, respectively (Bedez Üte and Kadoğlu, 2014). Interestingly, in the same study, yarn tenacity of ring-spun yarn was negatively correlated with ELO (r=-0.353). ELO, like length and FBS is under a high degree of genetic control, with genotypes generally exhibiting more control over ELO values than GxE or E (May, 1999). In the studies reviewed by May (1999), ELO exhibited primarily Additive genetic variance though some exceptions did occur. Thus, due to the highly heritable nature of this trait and Additive genetic variance, it is safe to conclude that phenotypic selection is sufficient for sustained improvement of this fiber quality parameter. Though standardization of the measurement is sorely needed for improved comparability.

Fiber fineness determines the spin limit of a yarn, that is, the finest yarn count that can be spun with an acceptable level of yarn quality and ends down (Faerber and Deussen, 1994). Finer fibers have greater fiber to fiber contact or "cooperation" in the yarn, which allows for a lesser yarn twist requirement, which in turn increases yarn production speeds at the factory (May, 1999). The improved yarn and textile performance imparted by fine fibers, makes fiber fineness a coveted trait among cotton breeders. Fineness, as measured by AFIS, more accurately depicts the genetic variance of the trait rather than non-genetic variance (Meredith Jr, et al., 1996), and is less impacted by environmental variance or other non-genetic factors which is the case for MIC (Meredith and Whitten, 1994). Roughly half of the experiments reviewed by May (1999) showed that non-genetic factors influenced MIC reading more than genetic factors. Draye et al. (2005) identified several QTL for FINE in an interspecific mapping population, but very few for MIC in the same population, giving further evidence of the highly heritable nature of FINE and its superiority for determining fiber coarseness over MIC. Because MIC is a measure of both fiber diameter as well as maturity, the reading can be confounded, so without knowing relative maturity of the fibers, it is not possible to discriminate low MIC cotton as being due to fine fibers, due to immature fibers, or a combination of both. Likewise, high MIC cotton can be due to coarse fibers, excessive secondary cell wall deposition, or a combination of both. This makes MIC readings difficult to interpret in studies attempting to determine the genetic underpinnings of fiber fineness or maturity. Considering these factors, it would likely be more helpful to consider the individual components affecting MIC readings, that is, fineness separately from maturity by using AFIS rather than HVI when these traits are considered. However, due to the ubiquity, speed and relatively inexpensive HVI analysis, MIC has largely been relied on for improving fiber fineness (May, 1999). Given its shortcomings, MIC is mostly influenced by

Additive genetic variation, and has reasonably high heritability estimates, and as such, has been effective in improving fiber quality through selection over the years (May, 1999).

Evolutionary History of the *Gossypium* Genus

Upland cotton is a member of the family Malvaceae, which contains the mallows. The *Gossypium* genus is one of great species diversity and range, consisting of 52 distinct species distributed globally (Fryxell, 1992), the beginnings of which occurred some 5-15 million years ago (mya) (Wendel and Cronn, 2003). Almost exclusively, except for those species grown under cultivation, members of this group occur in the tropics or sub-tropics, and exist as perennials. Only under cultivation can members of this group be considered annual, and that, only because they are grown in such a way that annual behavior is forced. If allowed to grow as they please without the risk of freeze, they would decidedly behave as a perennial. This unique genus has an interesting evolutionary history reflected in its diversity both in morphology as well as cytology, though these topics are only briefly discussed here.

Of the species comprising this diverse genus, most are diploid (2n=26), and have the genome group designations A-G, and K (Endrizzi, et al., 1985). Groups with differing genome-group letters indicate their lack of inter-fertility. A great amount of cytological work was conducted on this genus in the mid-20th century, and from it came understanding of the relatedness and structural compatibility of many of these diploid and polyploid genomes as well as the evolutionary history of the genus. Much of this work was done by Beasley (1942), Skovsted (1933, 1934, 1935, 1937) and Webber (1935, 1939) . Their work focused on making calculated crosses of the various combinations of species to determine the inter-compatibility of the species, which determined the members of the various genome groups and how the polyploids related to each other and to the diploid progenitor species.

Species from the A genome group include *G. arboreum* and *G. herbaceum*, both of which are native to the African and Asian continents, and were independently domesticated by early agriculturalists. These two diploid species have had considerable history of cultivation, although today, they are not grown on a large industrial scale. They are the only examples of cultivated diploid species of the *Gossypium* genus. The D genome group includes 13 New World diploid species, most notably *G. raimondii*, which is believed to be one of the precursors of our modern-day cultivated allotetraploid species. The other diploid groups vary a good deal and include several species like the D- and K-groups - containing 12 and 13 species respectively - to those with very few, as is the case for the F-group which contains only one species, *G. longycalyx* of East Africa (Wendel and Cronn, 2003).

The *Gossypium* genus is believed to have originated some 5-15 mya, and quickly diversified into the extant genome groups. At the base of the phylogenetic tree, occurred the split of the New (D-genome) and Old World (A-genome) cotton species. It is understood that these A- and D-genome groups represent the earliest divergence of the genus, after which subsequent divergence events occurred. The diploid groups can be sub-divided by the continents in which they inhabit, giving us an idea of their lineages. The Australian diploid *Gossypium* species include those from the C-, G-, and K-genome groups, the Americas include only the D-genome species, and the Africa/Arabia/Asia group which consists of genome groups A, B, E and F (Wendel and Cronn, 2003).

Between 1.3 to 1.7 mya in the New World, the hybridization of a diploid D-genome species occurred with an Asiatic A-genome cotton species that likely made a trans-oceanic trip to the land-masses of the New World (Beasley, 1940a; Senchina, et al., 2003; Skovsted, 1934b; Skovsted, 1937a). This chance hybridization initiated a polyploidy event that gave rise to what

is now our AD-genome, allotetraploid cotton species, including the widely cultivated *G. hirsutum* and *G. barbadense* most notably, but also the non-cultivated species *G. tomentosum, G. darwinii, G. mustelinum, G. ekmanianum,* and G. stephensii. The serendipitous event or series of events that led to the marriage and duplication of the A and D genomes are especially intriguing in that there does not exist an A-genome diploid in the New World today. There exists only D-genome wild diploid cottons and allotetraploid AD-genome cotton species. D-genome species include *G. gossypioides,* which is thought to be an early member of the American D-genome diploids, as well as *G. raimondii* of Peru and *G. klotzchianum* of the Galapagos Islands. Although there is some debate as to the exact circumstances, it is generally accepted that *G. raimondii* (Gerstel and Sarvella, 1956; Phillips, 1962; Phillips, 1963; Sarvella, 1958) or a closely related ancestor served as the D-genome donor and *G. herbaceum* the A-genome donor (Gerstel, 1953) in this fortuitous merger. It is not clear how a distant relative of *G. herbaceum* would have made it to the New World, however, it could have been accomplished by traveling across what is today the islands of Polynesia, then on to Central or South America. What is really interesting is that there exists no examples today of any relative of the *G. herbaceum* individuals that traveled the Pacific.

Consequences of Polyploidy

Polyploidy in plants

Polyploid formation is a common occurrence in the history of plant evolution and has likely been a driving force of diversification. Fossil evidence has suggested polyploidy existing in plants at rates of up to 70% of all angiosperm species (Masterson, 1994). More recently, genomic analyses and study have revealed that an ancient polyploidy event occurred prior to the global

radiation of angiosperms, suggesting that all flowering plants are essentially ancient polyploids (Bowers, et al., 2003; Jiao, et al., 2011).

The possible mechanisms by which a polyploid genome is created are varied. Several previous studies have shown that in some cases, when an interspecific hybridization occurs, triploid progeny are created in the F_1 or F_2 (Hollingshead, 1930; Jones and Bamford, 1942; Muntzing, 1930; Muntzing, 1932; Skalinska, 1945; Thompson, 1931). Self-fertilization, or backcrossing to one of the diploid parents can then lead to a tetraploid individual. This concept has been called the triploid bridge by Ramsey and Schemske (1998) (Ramsey and Schemske, 1998). Allotetraploids have also been documented coming directly from an interspecific cross of two diploids in the F_1 (Hahn, et al., 1990; Levan, 1937; Nagaharu, 1935) and/or F_2 generation (Buxton and Newton, 1928; Levan, 1941). Another mechanism lies in the production of unreduced gametes, which has been shown to occur at rates of 0.5% per gamete, which is quite surprising considering the number of gametes produced in a life cycle (Ramsey and Schemske, 1998). These unreduced gametes can then go on to hybridize with or form a closely related, but distinct species. This is especially likely between species where the diploid genomes are sufficiently different that the homeologous chromosomes will not pair during meiosis. The inability to pair during meiosis can then lead the hybridized plant to produce unreduced gametes, and thereby polyploid progeny at high rates (Ramsey and Schemske, 2002). Certainly it is possible that other avenues might exist for the genesis of a polyploid, however, these present the most likely explanations and are given credence by having been observed in synthetically produced polyploids.

Phenotypically, polyploid plants are not always larger structurally than their diploid relatives, but individual cell sizes are indeed usually larger (Otto and Whitton, 2000). The

formation of a new, polyploid organism often results in an unstable genome that undergoes major changes including reorganizations, structural changes, major deletions and duplications (Wendel, 2000). In synthetic polyploids of Brassica, wheat, and Arabadopsis, major chromosomal rearrangements and total fragment losses have been documented (Chen and Ni, 2006; Song, et al., 1995). Changes in gene expression and gene silencing have been shown to occur (Adams and Wendel, 2005; Chen and Ni, 2006; Comai, 2005; Lee and Chen, 2001; Liu, et al., 1998). Genomes under stress, such as those going through a polyploidy event, are prime candidates for the activation of latent transposable elements (McClintock, 1983), and other genes as well (Kashkush, et al., 2003; O'Neill, et al., 2002).

This instrument of genomic change can have a range of consequences on a newly formed polyploid, or neopolyploid. First of all, changing the genomic context in which genes occur has the potential to greatly increase the phenotypic variability within a neopolyploid population (Otto, 2007). It can also result in new combinations of phenotypic traits from two previously ecologically isolated species, allowing the neopolyploid to fit a previously unused or underutilized niche. These can include previously glaciated regions, or areas in which two closely related species have recently been introduced and where competition is limited (Brochmann, et al., 2004; Stebbins, 1984).

The additional copy of a genome inherent in polyploids allows for the masking or dilution of deleterious mutations that might accumulate over time (Stebbins, 1971). This allows the mutation to persist in a population, which can either be detrimental or it can allow for further mutations to accumulate. The redundant genes that exist in neopolyploids have been shown to offer some advantage to fitness (Kafri, et al., 2006), and selection seems to preserve these redundancies (Hughes, 1994; Kondrashov, et al., 2002). Polyploidization is a common form of

mutation, important in species diversification and evolution, the result of which leads to "ecologically distinct lineages that are able to persist"(Vamosi and Dickinson, 2006).

Polyploidy in *Gossypium*

The fortunate encounter of an Old World, A-genome diploid, similar to *G. herbaceum* with a D-genome diploid species resembling *G. raimondii* in the New World resulted in what are now the highly prized allotetraploid, AD-genome cotton species, *G. hirsutum* and *G. barbadense*. The doubling of the genome that occurred and the ensuing diversity that it would have created, looks to have provided some unique opportunities on which natural and artificial selection could act (Jiang, et al., 1998; Wendel and Cronn, 2003). At least one example of this is the modification in gene expression that has resulted in the unusually high cellulose content of cotton fibers in the allotetraploids (Paterson, et al., 2012).

The tetraploid cotton species of the New World were the subject of much investigation in the 1920s and 1930s. It was Skovsted (1934) who initially concluded that cultivated cotton was an amphidiploid, created by the merger of an Asiatic A-genome species with a D-genome New World species (Skovsted, 1934b). Other researchers had noticed that the large chromosomes of the Asiatic species, *G. arboreum* and *G. herbaceum* would form 13 bivalents with the 13 large chromosomes in New World tetraploids, however, it was Skovsted (1934) who first postulated that the tetraploids were likely the product of a merger between an A-genome and D-genome. He reached this conclusion by observing chromosome behavior during meiosis in hybrids of A-genome diploids with AD-genome tetraploids, and noticed that the larger A-genome chromosomes formed bivalents with the larger chromosomes of the AD-genome, and smaller chromosomes from the D-genome always paired with the smaller chromosomes of the AD-genome (Skovsted, 1934a; Skovsted, 1934b; Webber, 1934).

Shortly after Skovsted's observation, came the creation of the first synthetic tetraploid produced from crossing *G. arboreum* (A-genome) with *G. thurberi* (D-genome) followed by doubling the chromosome number with Colchicine. The production of a stable, true-breeding allotetraploid was achieved independently by two labs at about the same time (Beasley, 1940b; Harland, 1940) . Using a very targeted, inter-specific crossing scheme, Beasley was able to eliminate all other possible sub-genome contributors to the New World tetraploids (Beasley, 1942). Other researchers during this time demonstrated that the sub-genomes were of A- and D-genome ancestry by showing the homology of loci due to the common origins of the genomes (Harland, 1937; Harland and Atteck, 1941; Silow, 1946). The conclusion that *G. raimondii* was the most closely related D-genome diploid species to the D-subgenome of the tetraploids was reached by Stephens (1944a,1944b) and further supported by Hutchison (1945, 1947). Soon after, *G. herbaceum* was confirmed as the most closely related A-genome diploid as it's chromosomes formed bivalents at a higher frequency with the A-sub-genome of the tetraploids (Gerstel, 1953; Menzel and Brown, 1954).

Some interesting things occurred after polyploidization of the New World cottons. Gene conversion of homoeologues from A to D versions of genes was common in the neopolyploid, which might have contributed to the success of these new species (Guo, et al., 2014). The bias toward D allele conversion in tetraploid *G. hirsutum* could be due to the increased fitness of the D-alleles, or simply to higher frequency of outcrossing of the neopolyploid to the native D-genome diploids (Paterson, et al., 2012). Evidence of outcrossing of the neopolyploid could be contained in the New World, D-genome diploid, *G. gossypioides*, which has repetitive DNA specific to the A-genome Asiatic species (Paterson, et al., 2012), and cytoplasmic DNA evidence (Wendel et al., 1995). Alleles from the D-genome account for more genetic variation seen in

fiber traits for both *G. hirsutum* and *G. barbadense* (Jiang, et al., 1998; Reinisch, et al., 1994). This idea is especially interesting since *G. herbaceum*, the A-genome contributor produces spinnable fiber, whereas *G. raimondii* does not. A similar example of the D-genome contributing to fiber traits lies in "Beasley's Triple Hybrid," the result of crossing *G. thurberi* (D-genome) x *G. arboreum* (A-genome), doubling the chromosomes with Colchicine, and then crossing the artificial tetraploid with a *G. hirsutum* variety (Culp and Slafer, 1994). This triple hybrid is believed to be the source of improved fiber strength in a great deal if not all of American cotton germplasm, and the genes imparting that improved fiber strength, are thought to have come from *G. thurberi*, the D-genome species that lacks spinnable fibers (Niles and Feaster, 1984). Polyploidy indeed created unique and serendipitous avenues for selection in *G. hirsutum* and *G. barbadense*, both natural and artificial.

Domestication of *G. hirsutum* and *G. barbadense*

The *Gossypium* genus presents a truly remarkable example of plant domestication by humans in that four geographically isolated and unique species were domesticated by four isolated groups of people during nearly the same period for essentially the same purpose, the production of useful fiber. Archaeological evidence of cultivation for the Asiatic/African species, *G. herbaceum* and *G. arboreum*, can be traced back to at least 4,700 years ago (Chowdhury and Buth, 1971). *G. barbadense* was originally domesticated in and around the Peruvian Andes somewhere between 4000 and 5000 years ago (Westengen, et al., 2005), and *G. hirsutum* in the northern Yucatan peninsula around the same time (Brubaker and Wendel, 1994; d'Eeckenbrugge and Lacape, 2014; Wendel, et al., 1992). Some interesting archaeological evidence exists that could push the date back a little further. For instance, some *G. barbadense* remains have been found in house floors and hearths in northern Peru that are over 5500 to 6000 years old. Along

with other crop plant evidence like peanut and cucurbits dating back to 9000 years found in this same area, it would suggest that the *G. barbadense* was utilized in an agricultural capacity, perhaps suggesting an older domestication (Dillehay, et al., 2007). Archaeological remains of *G. hirsutum* dating back to 4000 to 5000 years ago, exist as already cultivated forms (Smith and Stephens, 1971) in areas where the crop was introduced (Wendel, et al., 1992), suggesting that this date could be pushed back further too, especially if the previously discussed evidence of *G. barbadense* is considered.

For many crop plant species, the center of their origin is not always their center of diversity, as espoused by Harlan (1992). Wendel et al. determined the original source of today's productive Upland cotton as southern Mexico and Guatemala (1992), which is counter to the previously accepted origin of Upland cotton which was the Mexican highlands. It is likely that today's cultivated Upland can be traced back to these Mexican highland types, however, their earlier origin can be placed in southern Mexico and Guatemala. The first examples of domestication were likely carried out by Pre-Columbian Americans of the Yucatan peninsula (Brubaker and Wendel, 1994). These Yucatan-derived plants would have then radiated out from the peninsula across the arid subtropics of Mesoamerica, the Caribbean, and northern South America (Brubaker and Wendel, 1994). Native or wild populations of *G. hirsutum* exist today as members of coastal ecosystems (Fryxell, 1979), but Hutchison (1951), and Stephens (1958), argue that these are feral, re-established forms of semi-domesticated types. However, according to genetic comparisons using restriction fragment length polymorphism (RFLP), Brubaker and Wendel propose that these littoral types are truly wild and that the inland forms are likely feral or early domesticated types (Brubaker and Wendel, 1994). These inland populations almost always

occur in association with the presence of humans, also suggesting their semi-domesticated status along with the RFLP analysis (Brubaker and Wendel, 1994; Hutchinson, 1951; Stephens, 1958).

G. barbadense has a wide range across the tropics of South America, and in many areas, overlaps with *G. hirsutum* in the Caribbean and northern parts of South America, occurring as "dooryard" or "commensal" plants in these regions (Brubaker, et al., 1999; Wendel and Cronn, 2003). Like its cousin *G. hirsutum*, it can be difficult to differentiate wild, feral or domesticated forms of *G. barbadense* within its naturally occupied range. Amplified fragment length polymorphism (AFLP) marker diversity analyses pinpoint the primitive domestication center of *G. barbadense* as northwestern Peru and southwestern Ecuador (Westengen, et al., 2005), fortifying the conclusions previously reached based on archaeological evidence (Piperno and Pearsall, 1998; Rossen, et al., 1996). Field studies have supported the center of diversity and the center of origin to be the same with a high number of primitive types in coastal, northwestern Peru (Fernandez, et al., 2003). *G. barbadense* would have radiated and been transported out into Colombia, Venezuela, the Caribbean, and the Pacific islands. A great deal of diversity still exists in *G. barbadense* wild populations along the northern Peruvian coast (Westengen, et al., 2005). Primitive forms of *G. barbadense* often include arborescent growth habit, small, undesirable boll types, and colored lint in varying shades of brown. As *G. barbadense* radiated out from its core area, it became more domesticated, the lint became more white, and the kidney-seeded and tufted traits were favored for their ease in hand-ginning (Brubaker, et al., 1999; Turcotte and Percy, 1990).

Selection by early Americans probably reflected the observations made by Cook (1905) of some post-Columbian natives in Guatemala shortly after the turn of the century. He observed people intercropping cotton and peppers. The perennial shrubs were harvested as soon as the

first bolls were open in the field, after which the plants were removed to provide more room for the peppers to grow (Cook, 1905a, 1905b). Selection such as this for earliness would have contributed significantly to the annualized habit of modern *G. hirsutum* (Cook, 1905b). Early maturing plant types would have been preferred by peoples living in the highlands of central Mexico, and thus, similar selection practices likely took place by these people, further domesticating and improving this plant species. This rigorous selection for earliness in the crop would have caused a severe restriction in the genetic diversity of these early progenitors, in what has been referred to as "catastrophic selection" pressure (Lewis, 1962).

The effects of domestication on this crop species have been significant at the phenotypic level, and accordingly, great change has occurred at the molecular level. Most important of course is the limited genetic diversity that domestication created in the germplasm pool, the topic of which will be discussed further in a later section. Transcriptome analyses using RNA-seq methods have shown significant shifts in expression patterns including a prolonged fiber elongation phase (Applequist, et al., 2001), as well as a radical redirection of plant resources from stress response pathways toward fiber production pathways and overall plant growth (Yoo and Wendel, 2014). Extensive changes to molecular networks within the plant have also been caused by domestication (Chaudhary, et al., 2008; Hovav, et al., 2008a; Rapp, et al., 2010; Yoo and Wendel, 2014).

Early Breeding History of *G. hirsutum*

In the late 18th and early 19th centuries, around the American Revolution, cotton cultivation in the southeastern United States consisted of two types of cotton, "Georgia Green Seed" and "Creole Black Seed" types (Moore, 1956). Green seeded cotton varieties, *G. hirsutum*, had short, coarse lint, high yield and good disease resistance as compared to the long and fine fiber of the smooth

black seed types, *G. barbadense*. The Creole Black seed cottons were not very productive and the Georgia Green Seed types were more difficult to gin by hand, which was the processing technology of the time (Affleck, 1851; Stephens, 1958; Wailes, 1854). Eli Whitney's invention of the cotton gin allowed the green seeded biotypes to be more efficiently mechanically processed, and as a result, these cultigens predominated until the Mexican highland stocks were introduced to the US (Brown and Ware, 1958; Moore, 1956; Niles and Feaster, 1984; Wailes, 1854; Ware, 1951).

Introductions of the Mexican highland types to the United States occurred in the early and mid- 1800s, especially after the Mexican American War in 1846 to 1848. These highland stocks contributed to a shorter growing season, higher yields, better disease resistance, longer, finer fiber, and improved ease of harvest (Moore, 1956; Wailes, 1854). Thus began the focused improvement of American Upland cotton germplasm. Selection and beneficial outcrossing among selections and cultivars of these new highland types led to the high-yielding American Upland cotton varieties of the early 1900s (Endrizzi, et al., 1985). The infestation of the boll weevil in the early 1900s led to further concerted introductions of Mexican germplasm as breeders and growers sought to develop earlier maturing varieties that could develop a mature crop before weevil populations could reach catastrophic levels each season (Niles and Feaster, 1984; Ware, 1951).

The turn of the century also saw a more targeted approach to cotton variety development. There was a more concerted effort to produce cultivars adapted to the various macro-environments of the US Cotton Belt such as the Acala SJ4, Plains, Delta, and Eastern types (Meredith Jr, 1991; Niles and Feaster, 1984). Acala SJ4 type cultivars are grown primarily west of Texas, such as New Mexico and California. They are largely based upon the Mexican stocks

from Acala SJ4 and Tuxtla, Mexico (Chiapas), with some degree of influence from Beasley's "Triple Hybrid," as well as some *G. barbadense* germplasm (Meredith Jr, 1991; Niles and Feaster, 1984; Ware, 1951). Acala SJ4 cotton varieties are well known for their excellent fiber characteristics. In North and West Texas the Plains type cultivars are grown, which are distinguished for their short stature, early season maturation, and their tight, storm and wind proof boll. These cultivars were derived from a variety called "Big Boll Stormproof" collected in Mexico in the 1850s (Niles and Feaster, 1984; Ware, 1951). The Paymaster group includes "Kekchi" germplasm collected in Guatemala by Cook (Cook, 1905a). The Delta cultivars, grown primarily in the Mississippi Delta as well as Beltwide, were developed from pre-1900 cultivars such as "Lone Star," which was derived from some 1860 and 1880 introductions from Mexico (Ramey, 1966; Ware, 1951). Delta and Pine Land, and Stoneville were early cotton seed companies that bred directly for the Mississippi Delta region. The fertile, uniform soils of the Delta were excellent grounds on which to breed, leading to many outstanding varieties that were grown across the belt and prized for their wide adaptability. The Eastern cultivars were largely made up of Coker germplasm, which includes many selections from the 19th and 20th century Mexican introductions, including some possible *G. barbadense* introgression (Niles and Feaster, 1984). It is interesting to note here that the cotton grown in these macro-environments show some phenotypic diversity, however, molecular dissection of these relationships tell us a different story. These 4 biotypes actually show a great deal of genetic similarity, and while modern breeding has improved cotton yields and quality substantially, the limited genetic diversity in the US germplasm pool could be limiting our genetic gain.

Early advances in breeding in the US consisted largely of mass selection within existing varieties. Fortunate outcrossing and co-mingling of introduced cultigens resulted in some lucky

advances in early variety development (Calhoun, et al., 1994). The first appreciable application of Mendelian principles in cotton breeding occurred with Balls' study of the inheritance of lint color in *G. barbadense* (1907). Similarly, Shoemaker examined the inheritance of the okra leaf trait in *G. hirsutum* (1909), and the inheritance of red leaf coloration by McLendon (1912).

Genetic Diversity Within *Gossypium*

Genetic diversity is of critical importance to improvement of a crop species, without which genetic gain is limited and not sustainable. Diversity can provide a buffer against unforeseen limitations to production such as disease epidemics, pest introduction, or severe weather limitations. For an example of great loss due to lack of diversity, we need look no further than the corn leaf blight epidemic in the United States of the 1970s. The disease wiped out an average of 20-30% of the national crop, resulting in a loss of nearly one billion dollars, which today would be worth over six billion dollars. This loss was due to the overwhelming use of a single cytoplasmic male-sterile line or a close relative as the seed parent for 85% of the nation's hybrid corn seed (Ullstrup, 1972). Following this crisis, breeders have been much more vigilant about maintaining appropriate levels of diversity.

The specter of significant loss to abiotic and biotic stressors is sufficient reason to foster diversity within the germplasm base, however, additional economic incentive exists; the potential for genetic gain and the improvement of the quality and yield of the crop. Around the turn of the 21st century, cotton breeders and agronomists began to notice a downward trend or "plateau" in yield and fiber quality (Helms, 2000; Lewis, 2001; May, et al., 1995; Meredith, et al., 1997). These authors all attributed this trend to the limited diversity within the crop. Some of the downward trend that they noticed was likely due to the introduction of transgenic crops in the mid-1990s. As transgenic traits were introgressed primarily through backcross methods,

genetic gain was limited. It took a significant amount of time for the linkage drag due to trait introgression to be overcome. This phenomenon, however, does point to the problems that we can incur when diversity is not utilized.

Because breeders usually cross only the most elite materials in the development of new varieties, breeding materials become genetically quite similar as desirable phenotypes are identified and retained for further interbreeding. In addition to the diversity reduction due to breeding activities, the crop's history of origin also plays a large role.

Among the most important contributing factors to the limited diversity we see today in Upland cotton is domestication by man (Iqbal, et al., 2001). Like other crop species, domestication had severe consequences to the diversity of the cultivated form of the crop. The rigorous selection protocol, whether direct or indirect, that was likely practiced by early agriculturalists would have improved the annual habit of early Upland, probably reduced seed dormancy, and improved crop earliness, similar to the practices of the Kekchi Indians observed by Cook (1905a,1905b). This same selection practice would have severely limited the genetic base of the crop as crossing schemes would not have been implemented, simply recurrent mass selection from the same material. The only outside source of variation would have been the sharing of seeds and chance outcrossing. These early maturing domesticated types gave rise to the "Yucatanense" Upland cottons, which would have been prized by the inhabitants of the Mexican Highlands in south central Mexico. All existing examples of cultivated Upland cotton can be traced back to documentation that points to their origin as the Mexican Highland stocks, however, molecular genetic examination has shown that the Highland stocks were derived from the Yucatanense types, and are therefore considered the progenitor form of today's cultivated

Upland cotton. This very limited origin suggests that inherent diversity within cultivated forms of the crop is extremely limited.

Since the birth of modern cotton breeding in the early 1900s, there has been a focus on producing cultivars that are regionally adapted, as the cotton belt in the US exhibits diverse growing environments. These regional cultivars show little diversity within their group, but do show some dissimilarity between regional groups, suggesting some "built-in" diversity (Wallace, et al., 2009). The US cotton belt can be split into 4 primary macro-environments. In southern California and the desert southwest, the Acala SJ4-type Upland cottons are preferred, in the Texas High Plains, stormproof types, in the Mississippi Delta they prefer the 'Delta' type varieties, and in the Southeastern US, the Coker and Delta types are preferred. Perhaps an unintended benefit of this regionality is that significant economic losses due to some biotic or abiotic stress event would impact a region forcefully perhaps, but not likely the whole cotton belt. Allozyme analysis, however, between examples of these seemingly disparate groups of cultivars have shown limited genetic diversity (Wendel, et al., 1992). This could be due to the limited number of allozymes available at the time of the study.

Calhoun et al., (1994,1997) in a thorough investigation of Upland cotton germplasm, compiled the pedigrees of a vast number of commercially important cultivars and germplasm lines. Using these compiled pedigrees, May et al. (1995) and Bowman et al. (1996) calculated coefficients of parentage (CP) for 126 Upland cotton cultivars released between 1980 and 1990, and CP for 260 cultivars released between 1970 and 1990, respectively. May et al. (1995) calculated the mean CP of the diverse panel of genotypes included in their study as 0.07, with 1 being the highest degree of shared parentage (full-sibs), and 0 indicating no shared parentage. Genotypes were grouped in clusters based on their shared parentage of at least one cultivar.

Within the 12 identified clusters, CP ranged from 0.17 to 0.34, suggesting diversity suffers significantly among lines of common origin. These low average values of CP would seem to indicate a high degree of diversity among the cultivars represented. Similar results were obtained by Bowman et al. (1996) when they examined 260 lines released between 1970 and 1990. Both studies suggest sufficient levels of diversity with the Upland cotton germplasm base, however, these results are misleading.

Van Esbroeck et al. (1999) observed that pedigree-based estimates of genetic similarity likely overestimate levels of diversity because they do not account for genetic similarity of the ancestral parentage, and also under-estimate genetic similarity of re-selections from "inbred" lines. Interestingly, out of the 668 public germplasm releases registered prior to 1998 in the journal Crop Science, or its affiliate, the Journal of Plant Registrations, only 4 of those germplasm lines have been used in the pedigree of a commercial variety (Esbroeck and Bowman, 1998). In a comparison of pedigree-based CP versus restriction fragment length polymorphism (RFLP) genetic similarity index (GSI), Van Becelaere et al. (2005) found GSI ranging from 0.914 to 0.975, differing wildly from the CP of 0.07 by May et al. (1995). While the determined values were quite different, there was a significant though moderate correlation of r=0.41 between CP and GSI (Van Becelaere, et al., 2005). Similar conclusions have been reached by researchers of other self-pollinated crops such as wheat (Barrett, et al., 1998; Kim and Ward, 1997), barley (Melchinger, et al., 1994) and oat (O'Donoughue, et al., 1994). The moderate level of congruence between pedigree-based CP and RFLP-based GSI, would seem to indicate that CP grossly overestimates levels of true genetic similarity, however, the moderate level of correlation could indicate that pedigree-based methods of determining genetic similarity are still useful, at least in determining "ball-park" levels of relatedness and general relationships.

Pedigree based determination of relatedness might not be precise or perfectly accurate, but it remains a good, cheap, fast method for getting a rough idea of the relatedness of genotypes.

Additional diversity studies conducted using modern molecular techniques have reinforced the idea of limited diversity within the *G. hirsutum* germplasm pool. Among the earlier papers on this topic, low levels of diversity were found on small, but representative samplings of the *G. hirsutum* germplasm pool using amplified fragment-length polymorphism (AFLP) markers and simple sequence repeat (SSR) markers (Pillay and Myers, 1999; Rungis, et al., 2005). Similarly, Bertini et al. (2006) found low levels of SSR allelic diversity among 53 cultivars using 31 polymorphic SSR markers, with dissimilarity ratios that ranged from 0 to 0.41 among members of the panel (Bertini, et al., 2006). In a fairly recent and comprehensive study examining 193 cultivars of diverse adaptation, Fang et al. (2013) screened the panel using 448 SSR markers selected based on representation of the *G. raimondii* sequence. The authors identified 1590 alleles at 732 loci, 523 of which were polymorphic. Genetic similarity index ranged from as low as 0.64 for Paymaster HS200 vs. Pak 4F CB 4025, and as high as 0.993 for Rowden vs. DES 716 (Fang, et al., 2013). For the most part, the phylogenic relationships and genetic similarity indexes were in agreement with the breeding history and pedigrees of the lines as elucidated by Calhoun et al. (1997) and May et al. (1995).

It is clear that the lack of diversity inherent within cultivated Upland cotton is due in large part to mankind's selective improvement. This is exemplified by the extreme selection pressure brought on by domestication events, the severe bottlenecking of the germplasm base that occurred shortly after, and in the practices of modern breeding where relatively few genotypes are used in the development of new commercial varieties. However, while the

diversity of cultivated types is quite low, it might be argued that sufficient diversity exists among accessions within our germplasm repositories. These resources, unfortunately, go under-utilized.

Introgression and *Gossypium* Genetic Resources

When the question arises of how we are to increase diversity within cultivated Upland cotton, one must first consider the potential sources of genetic diversity available to breeders of the crop. Harlan and de Wet proposed a method of classifying sources from which breeders can extract useful genetic variation. Foremost is the primary gene pool, which consists of members of the species and closely related species that readily intermate with few abnormalities. Gene transfer is straightforward with fertile F_1 and F_2 hybrids, and segregation in subsequent generations follows typical patterns. It is from the primary gene pool that genes are most readily accessible for crop improvement. Consequently, it is within the primary gene pool that most of the attention and resources are typically spent. The secondary gene pool includes related species that can be crossed with the cultivated crop of interest, but the resulting hybrids exhibit hybrid breakdown in the F_1 or later generations. Chromosome pairing is limited or abnormal, and so sterile progeny are often obtained. Their usage is difficult but with a few backcrosses, gene transfer is achievable. Tertiary gene pools represent an interesting but difficult to utilize resource. This pool represents plants that are quite different from the cultivated species, but can produce a hybrid using some very specialized techniques such as embryo rescue, chromosome doubling, or grafting. Hybrids resulting from tertiary gene pool introgression projects face large levels of sterility in both the hybrid and the progeny.

Within the primary gene pool of cultivated Upland cotton, as we have seen in studies like (Bertini, et al., 2006; Fang, et al., 2013), limited diversity exists. Although the boundaries of the gene pools are open to some interpretation, the primary gene pool of *G. hirsutum* can be

generally considered as being restricted to members of the *G. hirsutum* species. The cultivated form of Upland has limited genetic diversity, however, there do exist sources of great diversity within the species. An example of this are lines such as TX2093, a "wild," or at least minimally improved example of the Yucatanense race previously discussed, which likely descends from some of the very first cultivated Uplands. In an SSR-based diversity survey, TX2093 shared only ~30% of its marker alleles with TM-1, generally considered the cultivated Upland cotton genetic standard. To put this into perspective, within the same study, TM-1 shared ~20% of its SSR alleles with *G. tomentosum*, a tetraploid cotton species of Hawaiian origin (Liu, et al., 2000b). Other sources of diversity exist, although perhaps less extreme than the early progenitor races such as Yucatanense within the *G. hirsutum* species and within its primary gene pool, but they remain under-utilized. White and Richmond (1963) noticed significant heterosis for yield among lines of diverse *G. hirsutum* origin.

The primary gene pool is likely sufficient for small, incremental genetic gains, however, for sustained improvement, novel phenotypes, or drastic genetic gain for a particular trait, it is likely necessary that we branch out to the secondary gene pool. Thought it should be noted that examples of beneficial, intra-specific wide-crosses with exotic, unadapted, unimproved, or wild germplasm from the primary gene pool exist for other crops. Wheat germplasm has been improved by use of dwarfing genes from Japanese germplasm (Krull, et al., 1970), disease resistance from landraces (Hoisington, et al., 1999), and tomatoes have seen a great deal of valuable allelic introgression coming from wild tomato (Tanksley and McCouch, 1997). While the primary gene pool is an important source for crop improvement, there are some limitations. For instance, the use of landraces and wild specimens depends on exchange of germplasm across international borders for the benefit of both parties involved.

Interspecific matings of *G. hirsutum* by the other tetraploid species such as *G. barbadense* have been successful, however, they are generally fraught with hybrid breakdown, some amount of sterility, abnormal segregation, and loss of desirable genes during backcross generations and population truncation. For these reasons, we can consider the other tetraploid species as members of *G. hirsutum*'s secondary gene pool, since gene transfer is tractable, but can be complicated and time-consuming. The use of interspecific introgression has seen some interest and has likewise seen some success, especially where *G. barbadense* is concerned. *G. barbadense* has seen a good deal of introgression into *G. hirsutum*, and *G. hirsutum* into *G. barbadense*, both naturally and artificially. Some authors have considered *G. barbadense* to be a member of the primary gene pool with *G. hirsutum*, however, there are no examples of a direct *G. hirsutum* x *G. barbadense* single cross hybrid resulting in an inbred commercial variety or otherwise commercially acceptable germplasm line, aside from the interspecific F_1 hybrids cultivated in India. Examples of distant introgressions making an impact on commercial germplasm do indeed exist, and are discussed in greater detail below. These 2 species have seen 1 to 2 million years of divergent evolution and independent domestication events, which would suggest that a great deal of genetic variability between these species likely exists. Usually, these interspecific introgressions require several generations of backcrossing prior to being acceptable for use in breeding programs, especially commercial programs where the stakes are too high to take steps backward. Hence, members of the tetraploid cotton species are here considered as part of the secondary gene pool, which includes *G. barbadense*, *G. darwinii*, *G. tomentosum*, *G. mustelinum, G. ekmanianum, and G. stephensii*.

The use of *G. barbadense* for introgression has been very helpful for increasing selectable variation within *G. hirsutum*. Synthetic interspecific hybrid populations often exhibit

hybrid breakdown (Stephens, 1949), lower numbers of recombination in early generations (Reinisch, et al., 1994), preferential recombination, infertility, and unexpected segregation ratios in late generations (Jiang, et al., 2000) as well as agronomically unsuitable offspring. Extensive backcrossing and introgression efforts have been made to improve many traits, but especially fiber quality in *G. hirsutum* (Campbell, et al., 2011; Campbell, et al., 2010; Cantrell and Davis, 1993; Tatineni, et al., 1996). Successful cases of introgression usually involve the incorporation of the interspecific segments of DNA by way of introgression lines rather than direct or single cross hybrids (Shen, et al., 2007; Sun, et al., 2012; Zhang, et al., 2012; Zhang, et al., 2009). Percy et al. (2006), using stabilized introgression lines, created a group of recombinant inbred lines (RILs) that had significant portions of *G. barbadense* introgression in a *G. hirsutum* background, and identified several RILs that exhibited positive transgressive segregation for lint percent (LP) and 2.5% span length (SL), as well as no significant correlation between lint yield and fiber quality traits. In another example of using stable *G. hirsutum* lines introgressed with *G. barbadense* chromatin to improve fiber traits, Zeng and Meredith (2009) found significant negative correlations between lint yield and FBS, 50% SL, 2.5% SL, ELO, but positive correlation between lint yield and FINE and MAT, indicating a positive improvement in both yield and fiber quality.

Molecular mapping of agronomically important quantitative trait loci (QTL) has benefited greatly from using molecularly diverse interspecific crosses due to marker polymorphism and easier tracking of extreme phenotypes. There have been several fiber quality QTL identified in *G. hirsutum* coming from germplasm introgressed with *G. barbadense* using markers (Chen, et al., 2009; Draye, et al., 2005b; Lacape, et al., 2003; Paterson, et al., 2003; Shen, et al., 2007; Wang, et al., 2011; Zhang, et al., 2014; Zhang, et al., 2009). Zhang (2011)

identified several QTL for fiber quality traits in a *G. hirsutum* x *G. tomentosum* mapping population. Similarly, introgression of *G. hirsutum* into *G. barbadense* has been very important to the commercial viability of *G. barbadense* varieties. Wang et al. (1995) surveyed *G. barbadense* cultivars and determined there to be 7.3% *G. hirsutum* chromatin in commercial Pima varieties, 9% in Sea Island cultivars, and 9.6% in so-called Egyptian varieties of *G. barbadense*. Recall that *G. barbadense* did not see the same effort in domestication and selection for agricultural productivity that *G. hirsutum* saw by the native Central Americans. Hence, the *G. hirsutum* contribution has likely aided in the commercial productivity of *G. barbadense*.

Interspecific introgression is a long-term effort that requires significant allocation of resources and an integrated approach of research due to the high degree of epistasis and the opposing effects of QTL in different backgrounds, linkage drag and hybrid breakdown (Paterson, et al., 2004). It is largely because of these difficulties that many researchers feel that it is important for the public sector to work with feral, wild, diverse germplasm introgression, including related species such as *G. barbadense*, to widen the genetic base of Upland cotton and incorporate novel alleles (Wallace, et al., 2009).

The tertiary gene pool resources available to *G. hirsutum* include the diploid cotton species such *G. arboreum*, *G. herbaceum*, and the 50 or so others that exist. Related genera might also be included here, however, these are less likely to be utilized to a large degree, as there is already a large amount of diversity existent in *Gossypium* spp. The most important example of a successful tertiary gene pool introgression is found in Beasley's Triple Hybrid. Beasley (1940) created an artificial tetraploid by crossing *G. arboreum*, an Asiatic, A-genome diploid, with *G. thurberi*, a wild American, D-genome diploid spp., followed by doubling the

chromosomes of the resulting progeny. This synthetic autotetraploid was then crossed to Deltapine 14, a *G. hirsutum* cultivar, to produce a fertile offspring that would readily cross to *G. hirsutum* provided *G. hirsutum* was used as the male parent (Lewis, 1957). Early breeding efforts with the Triple Hybrid involved a great deal of backcrossing with *G. hirsutum* essentially, but was very important to improving the fiber strength of American Upland cotton germplasm. Beasley's Triple Hybrid can be found in the foundational pedigrees of the New Mexico Acala SJ4 and Pee Dee breeding programs (Smith, et al., 1999), both of which have been incredibly important to the improvement of Upland cotton. Additionally, there have been fiber quality alleles introgressed from *G. arboreum* (Sun, et al., 2012), *G. tomentosum* (Zhang, et al., 2011), and *G. klotzchianum* (Xu, et al., 2012).

Given the limited genetic diversity in the primary gene pool, sustained improvement of the Upland cotton crop will require better utilization of the available genetic resources found in related species, both diploid and tetraploid in order to successfully enhance the economically important traits (Lubbers and Chee, 2009). Molecular tools will greatly help us to introgress alleles from interspecific sources into elite germplasm. Work done by Rong et al. (2007) and Wang et al. (2013) have indicated that it is common to see QTLs for traits that map to non-homeologous regions of the A and D subgenomes, and that variation in contribution of subgenomes to important traits likely depends on the ploidy lineages, which has also been noted in other polyploid species such as peanut (Fonceka, et al., 2012). Large differences in expression between A and D subgenome homeologues, which is supported by a rather large body of work in cotton (Adams and Wendel, 2005; Chaudhary, et al., 2009; Flagel and Wendel, 2010; Hovav, et al., 2008b; Rapp, et al., 2010; Yoo and Wendel, 2014), could conceivably allow us to pyramid

these novel QTL together by introgressing them into *G. hirsutum* in non-homeologous locations within the genome (Wang, et al., 2013).

Status of Cotton Genetic Mapping Research

As discussed previously, interspecific introgression of traits comes with a significant amount of baggage, most of which is negative. QTL are generally carried on small portions of genome from the donor parent, but sourced from other species, these small segments can have significant agronomic impact to the recurrent parent from the myriad of other genes carried on the segment, the size of which is limited by the number of backcrosses or other recombination events to chip away at the exotic portion of DNA. For instance, introgression of traits from *G. barbadense* into *G. hirsutum*, often creates progeny that resemble *G. barbadense*, however, this is undesirable, as *G. hirsutum* is the cotton species of greater economic impact. In the past, the only way to tell if an individual plant was carrying the gene or QTL of interest was by growing large populations, then screening the population phenotypically. With fiber quality traits, this can be cost prohibitive, but it has been done to great success, many of which were discussed previously. The problem with this method is that it requires significant cost as well as time and space. The backcrossing and subsequent phenotyping of generations to track the desired trait takes years and years if not decades. The use of molecular marker technologies has changed this paradigm significantly over the past few decades, making interspecific trait introgression a more tractable and efficient endeavor.

Linkage Maps and QTL

The most notable requirement for successful mapping of traits, is a densely populated genetic linkage map which gives the location of molecular markers on their respective chromosomes, allowing researchers to co-locate them with important phenotypic traits. Initial

genetic map construction of crop genomes was conducted using restriction fragment length polymorphism (RFLP) based markers. These markers, some of the first DNA-based markers used in molecular research, are very labor intensive, requiring large amounts of DNA and several time-consuming steps in order to resolve the polymorphisms for DNA-fragment size between individuals in populations

Many of the initial linkage maps for cotton were constructed using interspecific populations to maximize the opportunity for molecular polymorphism as well as phenotypic variance. The first RFLP map of tetraploid cotton identified 705 loci across 41 linkage groups and 4675 cM (Reinisch, et al., 1994). They were able to infer the A and D subgenomes and assign chromosomal identities to 14 of the identified linkage groups using an F_2 population derived from a cross between a *G. hirsutum* and *G. barbadense* parent . In what was the first example of QTL mapping in cotton, Jiang et al. (1998) identified 14 QTL for fiber quality, 4 QTL for yield components, and one QTL for plant maturity or "earliness". Ulloa et al. (2002) used a combination of four mapping populations to create a composite genetic linkage map covering 1502 cM, and 284 loci, on 47 linkage groups. Wright et al. (1999) identified several QTL for plant pubescence, a trait that affects insect resistance as well as harvest and ginning efficiency. Many more mapping studies have been undertaken to identify QTL in cotton.

The development of polymerase chain reaction (PCR)-based DNA markers made a significant impact in the genomics field across organisms, including plants. Because PCR-based markers can sample vast numbers of loci across a genome very rapidly with only small quantities of DNA, these methods are much faster, less expensive, and more user-friendly than RFLPs. Some common PCR-based markers include amplified fragment length polymorphism (AFLP), sequence related amplified polymorphism (SRAP), random amplified polymorphic DNA

(RAPD) and simple sequence repeats (SSR). SSR markers, which are repetitive motifs of a few base pairs up to complex patterns, also referred to as microsatellite markers, have a high incidence of polymorphism, which has been very helpful in cotton as there is little intraspecific polymorphism with some of the other marker systems. These markers are usually codominant, occur widely across the genome, can be easily shared between institutions, can be used with varied populations, and in cotton are numerous within the genome. For these reasons, SSRs have enjoyed greater popularity than some of the other marker types. It has been proposed that SSRs are sufficiently ubiquitous within the cotton genome, that extensive genome mapping and marker assisted selection (MAS) can be conducted to great effect (Reddy, et al., 2001).

The work that has been conducted to make SSRs useful in cotton has been extensive. To provide the research community with anchored SSR locations, Liu et al. (2000c) used aneuploid cytogenetic stocks to localize about 30 SSR markers on specific chromosomes. A similar strategy was employed in wheat (*Triticum aestivum* L.) to localize about 200 SSR markers on the wheat genome (Röder, et al., 1998a,1998b). SSRs have been used extensively in genetic diversity studies of cultivated cotton (Bertini, et al., 2006; Fang, et al., 2013; Guang and Xiong-Ming, 2006; Gutierrez, et al., 2002; Kalivas, et al., 2011; Lacape, et al., 2007; Liu, et al., 2000a; Rungis, et al., 2005; Zhang, et al., 2005a; Zhang, et al., 2005b), and have re-confirmed that the genetic diversity available in cultivated cotton is low.

Linkage maps and QTL mapping have also benefited from the use of SSR markers. In the following studies, SSRs have been used exclusively to map QTL for plant architecture (Bao-Hua, et al., 2006), yield and yield components (Wang, et al., 2007), resistance to pathogens such as nematode and Verticillium wilt (Shen, et al., 2006; Wang, et al., 2008), and extensively in mapping of fiber quality QTLs (Abdurakhmonov, et al., 2008; Lacape, et al., 2005; Mei, et al.,

2004; Shen, et al., 2005; Wu, et al., 2009; Zhang, et al., 2005c). Using three different intraspecific *G. hirsutum* F2 mapping populations, and between 127 and 77 polymorphic SSR markers per population, Shen et al. mapped 39 QTLs for fiber quality traits (2005).

Today, the use of SSRs as well as a myriad of other types of molecular markers has been instrumental in the identification of QTL across the cotton genome. He et al. (2007) used 834 SSRs, 437 SRAPs, 107 RAPDs, and 16 retrotransposon-microsatellite amplified polymorphism (REMAP) markers to reveal 1,029 loci mapped to 26 linkage groups that extend to 5,472 cM, with an average inter-locus distance of 5.3 cM. They identified 52 economically important QTL for lint index (4 QTL), seed index (8 QTL), lint yield (11 QTL), seed cotton yield (4 QTL), seed per boll (9 QTL), fiber strength (3 QTL), fiber length (5 QTL), and micronaire (8 QTL) in a cross between a *G. hirsutum* and *G. barbadense* parent.

QTL Validation and Genetic Background

An important aspect of QTL research is validation of their effect. The explosion of mapped QTLs has been a boon for cotton genomics research, however, little validation has occurred to test them in different genetic backgrounds, environments or years. The interaction of a QTL with the genetic background and/or environment in which it is deployed can have a significant influence on its expression (Lecomte, et al., 2004; Sebolt, et al., 2000; Tanksley and Hewitt, 1988), thus making validation an important step after mapping an economically important QTL.

Interactions between QTL and genetic backgrounds may be particularly prominent when the mapping population is of interspecific origin. Often the parents used in mapping populations are chosen both for their molecular diversity as well as their phenotypic diversity. When QTL of interspecific origin are introgressed into adapted germplasm, there are unexpected interactions with genetic background. These can include varied or total lack of effect in given backgrounds,

unintended impacts on other traits, or lack of expression in different environments. Chee et al. (2005a, 2005b) and Draye et al. (2005), studying a backcross-self population, each paper focusing on one of the following traits ELO, FINE, and UHML, noticed the QTLs of interest having varied or unintended effects depending on genetic background. For this reason, before extensive introgression efforts are undertaken, it is critical the QTL be evaluated and verified to have effects across target populations.

WORKS CITED

Abdurakhmonov, I., R. Kohel, J. Yu, A. Pepper, A. Abdullaev, F. Kushanov, et al. 2008. Molecular diversity and association mapping of fiber quality traits in exotic G. hirsutum L. germplasm. Genomics 92: 478-487.

Adams, K. and J. Wendel. 2005. Novel patterns of gene expression in polyploid plants. Trends in Genetics 21: 539-543.

Adedoyin, A.A., C. Li and M.D. Toews. 2011. Characterization of single cotton fibers using a laser diffraction system. Textile Research Journal 81: 355-367.

Affleck, T. 1851. The early days of cotton growing in the South-West. De Bow's Review 10: 668-669.

Applequist, W.L., R. Cronn and J.F. Wendel. 2001. Comparative development of fiber in wild and cultivated cotton. Evolution & Development 3: 3-17.

Azhar, F., M. Naveed and A. Ali. 2004. Correlation of Sed Yield with Fiber Characteristics of Cotton. International Journal of Agriculture and Biology 6: 656-658.

1996. The importance of cotton fiber elongation on yarn quality and weaving performance. Proceedings of the 9th annual Engineered Fiber Selection System Conference. Cotton Incorporated, Raleigh, NC.

Backe, E.E. 1986. Effect of Short Fiber Content in Cotton on Plant Performance and Quality'. Textile Research Journal 56: 112-115.

Balls, W.L. 1907. Note on Mendelian Heredity in cotton. Journal of Agricultural Science 2: 216-220.

Barrett, B., K. Kidwell and P. Fox. 1998. Comparison of AFLP and pedigree-based genetic diversity assessment methods using wheat cultivars from the Pacific Northwest. Crop Science 38: 1271-1278.

Beasley, J. 1940a. The origin of American tetraploid Gossypium species. American Naturalist 74: 285-286.

Beasley, J. 1940b. The production of polyploids in Gossypium. Journal of Heredity 31: 39-48.

Beasley, J. 1942. Meiotic chromosome behavior in species, species hybrids, haploids, and induced polyploids of Gossypium. Genetics 27: 25.

Bedez Üte, T. and H. Kadoğlu. 2014. Regressional Estimation of Cotton Sirospun Yarn Properties from Fibre Properties. Autex Research Journal 14: 161-167.

Behery, H. 1993. Short fiber content and uniformity index in cotton. International Cotton Advisory Committee and Center for Agriculture and Biosciences Rewiew Article 4.

Benzina, H., E. Hequet, N. Abidi, J. Gannaway, J. Drean and O. Harzallah. 2007. Using Fiber Elongation to Improve Genetic Screening in Cotton Breeding Programs. Textile Research Journal 77: 770-778.

Bertini, C.H., I. Schuster, T. Sediyama, E.G.d. Barros and M.A. Moreira. 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genetics and Molecular Biology 29: 321-329.

Bowers, J., B. Chapman, J. Rong and A. Paterson. 2003. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. Nature 422: 433-438.

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

Bradow, J., O. Hinojosa, L. Wartelle, G. Davidonis, G. Sassenrath-Cole and P. Bauer. 1996. Applications of AFIS Fineness and Maturity Module and X-Ray Fluorescence Spectroscopy in Fiber Maturity Evaluation. Textile Research Journal 66: 545-554.

Brochmann, C., A. Brysting, I. Alsos, L. Borgen, H. Grundt, A. Scheen, et al. 2004. Polyploidy in arctic plants. Biological Journal of the Linnean Society 82: 521-536.

Brown, H.B. 1938. Cotton. History, species, varieties, morphology, breeding, culture, diseases, marketing, and uses. Cotton. History, species, varieties, morphology, breeding, culture, diseases, marketing, and uses.

Brown, H.B. and J.O. Ware. 1958. Cotton. New York, Mc Grow-Hill Book Company. Inc.

Brubaker, C.L., F. Bourland and J.F. Wendel. 1999. The origin and domestication of cotton. Cotton: Origin, history, technology, and production. John Wiley & Sons, New York: 3-31.

Brubaker, C.L. and J.F. Wendel. 1994. Reevaluating the origin of domesticated cotton (Gossypium hirsutum; Malvaceae) using nuclear restriction fragment length polymorphisms (RFLPs). American journal of botany: 1309-1326.

Buxton, B. and W. Newton. 1928. Hybrids of *Digitalis ambigua* and *Digitalis purpurea*, their fertility and cytology. Journal of Genetics 19: 1269-1279.

Calhoun, D., D. Bowman and O. May. 1994. Pedigrees of upland and pima cultivars released between 1970 and 1990. Miss. Agric. Forestry Exp. Stn. Bul 1017.

Calhoun, D.S., D. Bowman and O.L. May. 1997. Pedigrees of upland and pima cotton cultivars released between 1970 and 1995Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University.

Campbell, B., P. Chee, E. Lubbers, D. Bowman, W. Meredith, J. Johnson, et al. 2011. Genetic improvement of the Pee Dee cotton germplasm collection following seventy years of plant breeding. Crop Science 51: 955-968.

Campbell, B. and M. Jones. 2005. Assessment of genotype× environment interactions for yield and fiber quality in cotton performance trials. Euphytica 144: 69-78.

Campbell, B., S. Saha, R. Percy, J. Frelichowski, J. Jenkins, W. Park, et al. 2010. Status of the global cotton germplasm resources. Crop science 50: 1161-1179.

1993. Characterization of Upland× Pima breeding lines using molecular markers. Proc. Beltwide Cotton Conf., New Orleans, LA.

Chaudhary, B., L. Flagel, R.M. Stupar, J.A. Udall, N. Verma, N.M. Springer, et al. 2009. Reciprocal silencing, transcriptional bias and functional divergence of homeologs in polyploid cotton (Gossypium). Genetics 182: 503-517.

Chaudhary, B., R. Hovav, R. Rapp, N. Verma, J.A. Udall and J.F. Wendel. 2008. Global analysis of gene expression in cotton fibers from wild and domesticated Gossypium barbadense. Evolution & development 10: 567-582.

Chee, P., R. Bredhauer, C.W. Smith, A.H. Paterson, T.A. Delmonte, X. Draye, et al. 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., R. Bredhauer, C.W. Smith, A.H. Paterson, T.A. Delmonte, X. Draye, et al. 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.

Chen, H., N. Qian, W. Guo, Q. Song, B. Li, F. Deng, et al. 2009. Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro. D8 in Upland cotton. Theoretical and applied genetics 119: 605-612.

Chen, Z. and Z. Ni. 2006. Mechanisms of genomic rearrangements and gene expression changes in plant polyploids. Bioessays 28: 240-252.

Chowdhury, K. and G. Buth. 1971. Cotton seeds from the Neolithic in Egyptian Nubia and the origin of Old World cotton. Biological Journal of the Linnean Society 3: 303-312.

Comai, L. 2005. The advantages and disadvantages of being polyploid. Nature Reviews Genetics 6: 836-846.

Cook, O. 1905a. Cotton culture in Guatemala. Yearbook of the United States Department of Agriculture: 1904. Government Printing Office, Washington, DC.

Cook, O.F. 1905b. Social organization and breeding habits of the cotton-protecting kelep of Guatemala.

1988. Fiber strength by different test methods. Proceedings-Beltwide Cotton Production Research Conferences (USA).

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

Culp, T., D. Harrell and T. Kerr. 1979. Some genetic implications in the transfer of high fiber strength genes to upland cotton. Crop Science 19: 481-484.

Culp, T. and G. Slafer. 1994. Genetic contributions to yield in cotton. Genetic improvement of field crops.: 321-360.

d'Eeckenbrugge, G.C. and J.-M. Lacape. 2014. Distribution and differentiation of wild, feral, and cultivated populations of perennial Upland cotton (Gossypium hirsutum L.) in mesoamerica and the Caribbean.

De Vicente, M. and S. Tanksley. 1993. QTL analysis of transgressive segregation in an interspecific tomato cross. Genetics 134: 585-596.

1993. Improved cotton fiber properties: The textile industry's key to success in global competition. Proceedings-Beltwide Cotton Conferences (USA).

Dillehay, T.D., J. Rossen, T.C. Andres and D.E. Williams. 2007. Preceramic adoption of peanut, squash, and cotton in northern Peru. Science 316: 1890-1893.

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

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

Endrizzi, J., E. Turcotte and R. Kohel. 1985. Genetics, cytology and evolution of Gossypium. Advances in Genetics 23: 271-375.

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

1995. Future demands on cotton fiber quality in the textile industry. Beltwide Cotton Conferences (USA).

D. Herber and D. Richter. 1994. Improved cotton fiber quality and improved spinning technology - a profitable marriage. Part I. Progress in rotor spinning and progress in the quality profile of U.S. upland cotton. Part II. The contributions of improved cotton quality and those of rotor spinning developments to higher profits in cotton production and in spinning. Beltwide Cotton Conference, Memphis, TN. National Cotton Council.

Fang, D.D., L.L. Hinze, R.G. Percy, P. Li, D. Deng and G. Thyssen. 2013. A microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton (Gossypium hirsutum L.) cultivars from major cotton-growing countries. Euphytica 191: 391-401.

FAO-ICAC. 2013. World Apparel Fibre Consumption. In: S. Shui and A. Plastina, editors, Food and Agriculture Organization of the United Nations - International Cotton Advisory Comittee.

Faulkner, W., E. Hequet, J. Wanjura and R. Boman. 2012. Relationships of cotton fiber properties to ring-spun yarn quality on selected High Plains cottons. Textile Research Journal 82: 400-414.

Fernandez, H., R. Rodriguez and O. Westengen. 2003. Biologıa y Etnobotanica del Algodon Nativo Peruano (Gossypium barbadense L., Malvaceae). Arnaldoa 10: 93-108.

Flagel, L.E. and J.F. Wendel. 2010. Evolutionary rate variation, genomic dominance and duplicate gene expression evolution during allotetraploid cotton speciation. New Phytologist 186: 184-193.

Fonceka, D., H.-A. Tossim, R. Rivallan, H. Vignes, I. Faye, O. Ndoye, et al. 2012. Fostered and left behind alleles in peanut: interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. BMC plant biology 12: 26.

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

Fryxell, P. 1992. A revised taxonomic interpretation of Gossypium L (Malvaceae). Rheedea 2: 108-165.

Gerstel, D. 1953. Chromosomal translocations in interspecific hybrids of the genus Gossypium. Evolution 7: 234-244.

Gerstel, D. and P. Sarvella. 1956. Additional observations on chromosomal translocations in cotton hybrids. Evolution 10: 408-414.

1988. The inheritance and relationships of yield and fiber properties and the implications to a cotton breeding program. Proceedings-Beltwide Cotton Production Research Conferences (USA).

Green, C. and T. Culp. 1990. Simultaneous improvement of yield, fiber quality, and yarn stregth in Upland cotton. Crop science 30: 66-69.

Guang, C. and D. Xiong-Ming. 2006. Genetic diversity of source germplasm of upland cotton in China as determined by SSR marker analysis. Acta Genetica Sinica 33: 733-745.

Guo, H., X. Wang, H. Gundlach, K.F. Mayer, D.G. Peterson, B.E. Scheffler, et al. 2014. Extensive and biased intergenomic nonreciprocal DNA exchanges shaped a nascent polyploid genome, Gossypium (Cotton). Genetics 197: 1153-1163.

Gutierrez, O., S. Basu, S. Saha, J. Jenkins, D. Shoemaker, C. Cheatham, et al. 2002. Genetic Distance among Selected Cotton Genotypes and Its Relationship with F Performance. Crop Science 42: 1841-1847.

Hahn, S., K. Bai and R. Asiedu. 1990. Tetraploids, triploids, and 2*n* pollen from diploid interspecific crosses with cassava. Theoretical and Apllied Genetics 79.

Hake, K., D. Bassett, T. Kerby and W. Mayfield. 1996. Producing quality cotton. In: S. Johnson Hake, T. Kerby and K. Hake, editors, Cotton Production Manual. University of California, Divison of Agriculture and Natural Resources, Oakland, CA.

Harlan, J. 1992. Crops and Man. American Society of Agronomy. Crop Science Society of America, Madison, Wisconsin: 63–262.

Harland, S. 1937. Homologous loci in wild and cultivated American cottons. Nature 140: 467-468.

Harland, S. 1940. New polyploids in cotton by the use of colchicine. Tropical Agriculture, Trinidad and Tobago 17: 53-54.

Harland, S.C. and O.M. Atteck. 1941. The genetics of cotton. Journal of Genetics 42: 21-47.

He, D.-H., Z.-X. Lin, X.-L. Zhang, Y.-C. Nie, X.-P. Guo, Y.-X. Zhang, et al. 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× Gossypium barbadense. Euphytica 153: 181-197.

2000. Yield study report. Proc Beltwide Cotton Production Conf. National Cotton Council of America, San Antonio, Tex.

Hertel, K. 1940. A Method of Fiber-Length Analysis Using the Fibrograph. Textile Research Journal 10: 510-525.

Hoisington, D., M. Khairallah, T. Reeves, J.-M. Ribaut, B. Skovmand, S. Taba, et al. 1999. Plant genetic resources: What can they contribute toward increased crop productivity? Proceedings of the National Academy of Sciences 96: 5937-5943.

Hollingshead, L. 1930. Cytological investigations of hybrids and hybrid derivatives of Crepis capillaris and Crepis tectorumUniversity of California Press.

Hovav, R., J.A. Udall, B. Chaudhary, E. Hovav, L. Flagel, G. Hu, et al. 2008a. The evolution of spinnable cotton fiber entailed prolonged development and a novel metabolism. PLoS Genet 4: e25.

Hovav, R., J.A. Udall, E. Hovav, R. Rapp, L. Flagel and J.F. Wendel. 2008b. A majority of cotton genes are expressed in single-celled fiber. Planta 227: 319-329.

Hsieh, Y.-L. 1999. Structural development of cotton fibers and linkages to fiber quality. Cotton Fibers Developmental Biology, Quality Improvement, and Textile Processing. Food Product Press, Binghamton, nY: 167-183.

Hughes, A. 1994. The evolution of functionally novel proteins after gene duplication. Proceedings of the Royal Society of London B: Biological Sciences 256: 119-124.

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

Hutchinson, J., R. Silow and S. Stephens. 1947. The evolution of Gossypium and the differentiation of the cultivated cottonsOxford University Press, London.

Hutchinson, J., S. Stephens and K. Dodds. 1945. The seed hairs of Gossypium. Annals of Botany: 361-367.

Incorporated, C. 2015. Cotton Fiber Competition. Cotton Incorporated, Raleigh.

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

ITMF. 2014. Cotton Contamination Survey - 2013. International Textile Manufacturers Federation.

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

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

Jiao, Y., N.J. Wickett, S. Ayyampalayam, A.S. Chanderbali, L. Landherr, P.E. Ralph, et al. 2011. Ancestral polyploidy in seed plants and angiosperms. Nature 473: 97-100.

Jones, R. and R. Bamford. 1942. Chromosome number in the progeny of triploid *Gladiolis* with special reference to the contribution of the triploid. American Journal of Botany 29: 807-813.

Kafri, R., M. Levy and Y. Pilpel. 2006. The regulatory utilization of genetic redundancy through responsive backup circuits. Proceedings of the National Academy of Sciences 103: 11653-11658.

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

Kashkush, K., M. Feldman and A.A. Levy. 2003. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nature genetics 33: 102-106.

Kim, H. and R. Ward. 1997. Genetic diversity in Eastern US soft winter wheat (Triticum aestivum L. em. Thell.) based on RFLPs and coefficients of parentage. Theoretical and Applied Genetics 94: 472-479.

Kondrashov, F., I. Rogozin, Y. Wolf and E. Koonin. 2002. Selection in the evolution of gene duplications. Genome Biology 3: 1-8.

Krifa, M. 2006. Fiber Length Distribution in Cotton Processing: Dominant Features and Interaction Effects. Textile Research Journal 76: 426-435.

Krull, C., N. Borlaug, O. FRANKEL and E. BENNETT. 1970. The utilization of collections in plant breeding and production. Genetic resources in plants-their exploration and conservation.: 427-439.

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

Lacape, J.-M., T.-B. Nguyen, B. Courtois, J.-L. Belot, M. Giband, J.-P. Gourlot, et al. 2005. QTL analysis of cotton fiber quality using multiple× backcross generations. Crop Science 45: 123-140.

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

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

Lee, H.-S. and Z.J. Chen. 2001. Protein-coding genes are epigenetically regulated in Arabidopsis polyploids. Proceedings of the National Academy of Sciences 98: 6753-6758.

Levan, A. 1937. Cytological studies in the *Allium paniculatum* group. Hereditas 23.

Levan, A. 1941. The cytology of the species hybrid *Allium cepa* x *fistulosum* and its polyploid derivatives. Hereditas 27.

Lewis, C.F. 1957. Genetic Recombinations in a Hybrid Involving Three Species of Gossypium. Agronomy Journal 49: 455-460.

Lewis, H. 1962. Catastrophic selection as a factor in speciation. Evolution: 257-271.

2001. A review of yield and quality trends and components in American upland cotton. Proc. beltwide cotton conf., Anaheim, CA.

Liu, B., J. Vega and M. Feldman. 1998. Rapid genomic changes in newly synthesized amphiploids of Triticum and Aegilops. II. Changes in low-copy coding DNA sequences. Genome 41: 535-542.

Liu, S., R. Cantrell, J. McCarty and J.M. Stewart. 2000a. Simple sequence repeat–based assessment of genetic diversity in cotton race stock accessions. Crop Science 40: 1459-1469.

Liu, S., R. McCarty, J. Stewart and J. McD. 2000b. Simple Sequence Repeat–Based Assessment of Genetic Diversity in Cotton Race Stock Accessions. Crop Sci 40: 1459-1469.

Liu, S., S. Saha, D. Stelly, B. Burr and R. Cantrell. 2000c. Chromosomal assignment of microsatellite loci in cotton. Journal of Heredity 91: 326-332.

Lubbers, E.L. and P.W. Chee. 2009. The worldwide gene pool of G. hirsutum and its improvement. Genetics and genomics of cotton. Springer. p. 23-52.

Masterson, J. 1994. Stomatal size in fossil plants: Evidence for polyploidy in majority of angiosperms. Science 264: 421-424.

May, O. 1999. Genetic variation in fiber quality. In: A. Basra, editor Cotton fibers-developmental biology, quality improvement, and textile processing. Food Products Press, New York. p. 183-229.

May, O., D. Bowman and D. Calhoun. 1995. Genetic diversity of US upland cotton cultivars released between 1980 and 1990.

May, O.L. and C.C. Green. 1994. Genetic variation for fiber properties in elite Pee Dee cotton populations. Crop science 34: 684-690.

May, O.L. and R.A. Taylor. 1998. Breeding cottons with higher yarn tenacity. Textile research journal 68: 302-307.

McClintock, B. 1983. Nobel Lecture: The Significance of Responses of the Genome to Challenge, Nobel Foundation, Stockholm.

McLendon, C.A. 1912. Mendelian inheritance in cotton hybridsGeorgia Experiment Station.

Mei, M., N. Syed, W. Gao, P. Thaxton, C. Smith, D. Stelly, et al. 2004. Genetic mapping and QTL analysis of fiber-related traits in cotton (Gossypium). Theoretical and applied genetics 108: 280-291.

Melchinger, A.E., A. Graner, M. Singh and M.M. Messmer. 1994. Relationships among European barley germplasm: I. Genetic diversity among winter and spring cultivars revealed by RFLPs. Crop Science 34: 1191-1199.

Menzel, M.Y. and M.S. Brown. 1954. The significance of multivalent formation in three-species Gossypium hybrids. Genetics 39: 546.

1992. Improving fiber strength through genetics and breeding. Cotton Fiber Cellulose: Structures, Functions, and Utilization Conference.

Meredith Jr, W. 1991. Contributions of introductions to cotton improvement. In: H. Shands and L. Wiesner, editors, Use of plant introductions in cultivar development, Part I. ASA, CSSA, SSSA, Madison, WI. p. 127-146.

1996. Regional high quality fiber properties as measured by conventional and AFIS methods. Proceedings.

Meredith, W. 1977. Backcross breeding to increase fiber strength of cotton. Crop Science 17: 172-175.

Meredith, W. and R. Bridge. 1971. Breakup of Linkage Blocks in Cotton, Gossypium hirsutum L. Crop Science 11: 695-698.

1994. Genetics and management factors influencing textile fiber quality. Proc 7th Ann Cotton Incorporated Engineered Fiber Selection System Res Forum. Cotton Incorporated, Raleigh, NC.

Meredith, W.R., T. Culp, K. Robert, G. Ruppenicker, W. Anthony and J. Williford. 1991. Determining future cotton variety fiber quality objectives. Textile research journal 61: 715-720.

Meredith, W.R., J. Heitholt, W. Pettigrew and S. Rayburn. 1997. Comparison of obsolete and modern cotton cultivars at two nitrogen levels. Crop science 37: 1453-1457.

Miller, P. and A. Marani. 1963. Heterosis and combining ability in diallel crosses of upland cotton, Gossypium hirsutum. L. Crop Sci 3: 441-444.

Miller, P. and J. Rawlings. 1967. Selection for increased lint yield and correlated responses in upland cotton, Gossypium hirsutum L. Crop Science 7: 637-640.

Miller, P., J. Williams, H. Robinson and R. Comstock. 1958. Estimates of Genotypic and Environmental Variances and Covariances in Upland Cotton and Their Implications in Selection. Agronomy Journal 50: 126-131.

Moore, J. 1996. Cotton Classification and Quality. In: E. Glade, L. Meyer and H. Stults, editors, The Cotton Industry in the United States. USDA-ERS Agriculture Economy Report No. 739, Washington, DC. p. 51-57.

Moore, J.H. 1956. Cotton breeding in the Old South. Agricultural History: 95-104.

Muntzing, A. 1930. Outlines to a genetic monograph of the genus Galeopsis with special reference to the nature and inheritance of partial sterility. Hereditas 16: 185-341.

Muntzing, A. 1932. Cyto-genetic investigations on synthetic Galeopsis tetrahit. Hereditas 16: 105-154.

Nagaharu, U. 1935. Genome analysis in Brassica with special reference to the experimental formation of B. napus and peculiar mode of fertilization. Japanese Journal of Botany 7: 389-452.

Niles, G. and C. Feaster, editor. 1984. Breeding. Cotton Agronomy Monograph. ASA, Madison, WI.

Novick, R.G., J. Jones, W. Anthony, W. Aguillard and J. Dickson. 1991. Genetic trait effects on nonlint trash of cotton. Crop science 31: 1029-1034.

O'Donoughue, L., E. Souza, S. Tanksley and M. Sorrells. 1994. Relationships among North American oat cultivars based on restriction fragment length polymorphisms. Crop science 34: 1251-1258.

O'Neill, R., M. O'Neill and J. Graves. 2002. Undermethylation associated with retroelement activation and chromosome remodelling in an interspecific mammalian hybrid (vol 393, pg 68, 1998). Nature 420: 106-106.

Otto, S. 2007. The evolutionary consequences of polyploidy. Cell 131: 452-462.

Otto, S. and J. Whitton. 2000. Polyploid incidence and evolution. Annual Review of Genetics 34: 451-483.

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

Paterson, A., Y. Saranga, M. Menz, C.-X. Jiang and R. Wright. 2003. QTL analysis of genotype× environment interactions affecting cotton fiber quality. Theoretical and Applied Genetics 106: 384-396.

Paterson, A.H., J.F. Wendel, H. Gundlach, H. Guo, J. Jenkins, D. Jin, et al. 2012. Repeated polyploidization of Gossypium genomes and the evolution of spinnable cotton fibres. Nature 492: 423-427.

G. Constable and N. Forrester. 1995. Development of Medium Staple High-Strength Cotton Suitable for Rotor Spinning Systems. World Cotton Conference, Brisbane, Australia. February 14-17, 1994. CSIRO, Melbourne, Australia.

Paudel, D.R., E.F. Hequet and N. Abidi. 2013. Evaluation of cotton fiber maturity measurements. Industrial Crops and Products 45: 435-441.

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

Perkins, H., D. Ethridge and C. Bragg. 1984. Fiber. In: R. Kohel, editor Cotton. ASA, Madison, WI. p. 437-509.

Phillips, L. 1962. Segregation in new allopolyploids of Gossypium. IV. Segregation in New World x Asiatic and New World x wild American hexaploids. American Journal of Botany 49: 51-57.

Phillips, L. 1963. The cytogenetics of Gossypium and the origin of New World cottons. Evolution 17: 460-469.

Pillay, M. and G. Myers. 1999. Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. Crop Science 39: 1881-1886.

Piperno, D.R. and D.M. Pearsall. 1998. Origins of agriculture in the lowland neotropicsAcademic Press.

Quisenberry, J. 1975. Inheritance of fiber properties among crosses of Acala and high plains cultivars of upland cotton. Crop Science 15: 202-204.

1966. Historical review of cotton variety development. Proc. 18th Cotton Improvement Conf., Memphis, TN.

Ramey, H., R. Lawson and S. Worley. 1977. Relationship of cotton fiber properties to yarn tenacity. Textile Research Journal 47: 685-691.

Ramey Jr, H. 1999. Classing of fiber. Cotton: origin, history, technology, and production. John Wiley & Sons, new York: 709-727.

Ramsey, J. and D. Schemske. 1998. Pathways, mechanisms, and rates of polyploid formation in flowering plants. Annual Review of Ecology and Systematics 29: 467-501.

Ramsey, J. and D. Schemske. 2002. Neopolyploidy in flowering plants. Annual Review of Ecology and Systematics 33: 589-639.

Rapp, R.A., C.H. Haigler, L. Flagel, R.H. Hovav, J.A. Udall and J.F. Wendel. 2010. Gene expression in developing fibres of Upland cotton (Gossypium hirsutum L.) was massively altered by domestication. BMC biology 8: 139.

Reddy, O.U.K., A.E. Pepper, I. Abdurakhmonov, S. Saha, J.N. Jenkins, T. Brooks, et al. 2001. New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. J Cotton Sci 5: 103-113.

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

Richmond, T. 1950. Procedures and methods of cotton breeding with special reference to American cultivated species. Advances in genetics 4: 213-245.

Röder, M.S., V. Korzun, B.S. Gill and M.W. Ganal. 1998a. The physical mapping of microsatellite markers in wheat. Genome 41: 278-283.

Röder, M.S., V. Korzun, K. Wendehake, J. Plaschke, M.-H. Tixier, P. Leroy, et al. 1998b. A microsatellite map of wheat. Genetics 149: 2007-2023.

Rong, J., F.A. Feltus, V.N. Waghmare, G.J. Pierce, P.W. Chee, X. Draye, et al. 2007. Meta-analysis of polyploid cotton QTL shows unequal contributions of subgenomes to a complex network of genes and gene clusters implicated in lint fiber development. Genetics 176: 2577-2588.

Rossen, J., T.D. Dillehay and D. Ugent. 1996. Ancient cultigens or modern intrusions?: Evaluating plant remains in an Andean case study. Journal of Archaeological Science 23: 391-407.

Rungis, D., D. Llewellyn, E. Dennis and B. Lyon. 2005. Simple sequence repeat (SSR) markers reveal low levels of polymorphism between cotton (Gossypium hirsutum L.) cultivars. Crop and Pasture Science 56: 301-307.

Sarvella, P. 1958. Cytomixis and the loss of chromosomes in meiotic and somatic cells of Gossypium. Cytologia 23: 14-24.

1998. Short Fiber Content Determination by Means of AFIS. International Textile Committee on Cotton Testing Methods - Meeting of the ITMF Working Group on Fiber Length, Bremen, Germany. March 10-11. International Textile Manufacturers Federation, Zurich, Switzerland.

1994. Single Fiber Based Determination of Short Fiber Content. Beltwide Cotton Conferences, San Diego, CA. January 5-8, 1994. National Cotton Council of America, TN, USA.

Scholl, R. and P. Miller. 1976. Genetic Association between Yield and Fiber Strength in Upland Cotton. Crop Science 16: 780-783.

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

Self, F. and M. Henderson. 1954. Inheritance of fiber strength in a cross between the upland cotton varieties. AHA 50 and Half and Half. Agron. J 46: 151-154.

Senchina, D.S., I. Alvarez, R.C. Cronn, B. Liu, J. Rong, R.D. Noyes, et al. 2003. Rate variation among nuclear genes and the age of polyploidy in Gossypium. Molecular biology and evolution 20: 633-643.

Shen, X., W. Guo, Q. Lu, X. Zhu, Y. Yuan and T. Zhang. 2007. Genetic Mapping of Quantitative Triat Loci for FIber Quality and Yield Trait by RIL Approach in Upland Cotton. Euphytica 155: 371-380.

Shen, X., W. Guo, X. Zhu, Y. Yuan, Z.Y. John, R.J. Kohel, et al. 2005. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. Molecular breeding 15: 169-181.

Shen, X., G. Van Becelaere, P. Kumar, R.F. Davis, O.L. May and P. Chee. 2006. QTL mapping for resistance to root-knot nematodes in the M-120 RNR Upland cotton line (Gossypium hirsutum L.) of the Auburn 623 RNR source. Theoretical and Applied Genetics 113: 1539-1549.

Shoemaker, D. 1909. A study of leaf characters in cotton hybrids. Journal of Heredity: 116-118.

Silow, R. 1946. Evidence on chromosome homology and gene homology in the amphidiploid New World cottons. Journal of genetics 47: 213-221.

Skalinska, M. 1945. Cytogenetic studies in triploid hybrids of Aquilegia. Journal of Genetics 47: 87-111.

Skovsted, A. 1933. Cytological studies in cotton. I. The mitosis and the meiosis in diploid and triploid Asiatic cotton. Annals of Botany 47: 227-251.

1934a. Cytogenetics in relation to plant breeding in cotton. . Proc. 2nd Conf. Cotton Grow. Probl., Emp. Cotton Grow. Corp.

Skovsted, A. 1934b. Cytological studies in cotton. II. Two interspecific hybrids between Asiatic and New World Cottons. Journal of Genetics 28: 407-424.

Skovsted, A. 1935a. Cytological Studies in Cotton: III. A Hybrid Between Gossypium Davidsonii Kell and G. Sturtii F. Muell.

Skovsted, A. 1935b. Some new interspecific hybrids in the genus Gossypium L. J. Genet 30: 447-463.

Skovsted, A. 1937a. Cytological studies in cotton. Journal of Genetics 34: 97-134.

Skovsted, A. 1937b. Cytological Studies in Cotton. IV. Chromosome Conjugation in Interspecific Hybrids, by A. SkovstedUniversity Press.

Smith, C. and G. Coyle. 1997. Association of fiber quality parameters and within-boll yield components in upland cotton. Crop science 37: 1775-1779.

Smith, C.E. and S. Stephens. 1971. Critical identification of Mexican archaeological cotton remains. Economic botany 25: 160-168.

Smith, C.W., R. Cantrell, H. Moser and S. Oakley. 1999. History of cultivar development in the United States. Cotton: Origin, history, technology, and production. John Wiley & Sons, New York: 99-171.

Song, K., P. Lu, K. Tank and C. Osborn. 1995. Rapid genome change in synthetic polyploids of Brassica and its implications for polyploid evolution. Proceedings of the National Academy of Sciences 92: 7719-7723.

Stebbins, G. 1971. Processes of organic evolution. 2 ed. Prentice-Hall, Englewood Cliffs, N.J.

Stebbins, G. 1984. Polyploidy and the distribution of the arctic-alpine flora: new evidence and a new approach. Botanica Helvetica 94: 1-13.

Stephens, S. 1944a. The genetic organization of leaf-shape development in the genusGossypium. Journal of Genetics 46: 28-51.

Stephens, S. 1944b. Phenogenetic evidence for the amphidiploid origin of New World cottons. Nature 153: 53-54.

Stephens, S. 1949. The cytogenetics of speciation in Gossypium. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 34: 627.

Stephens, S.G. 1958. Factors affecting seed dispersal in Gossypium and their possible evolutionary significance.

Sun, F.-D., J.-H. Zhang, S.-F. Wang, W.-K. Gong, Y.-Z. Shi, A.-Y. Liu, et al. 2012. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Molecular breeding 30: 569-582.

Tang, B., J.N. Jenkins, J. McCarty and C. Watson. 1993. F2 hybrids of host plant germplasm and cotton cultivars: II. Heterosis and combining ability for fiber properties. Crop science 33: 706-710.

Tanksley, S., M. Ganal, J. Prince, M. De Vicente, M. Bonierbale, P. Broun, et al. 1992. High density molecular linkage maps of the tomato and potato genomes. Genetics 132: 1141-1160.

Tanksley, S. and J. Hewitt. 1988. Use of molecular markers in breeding for soluble solids content in tomato—a re-examination. Theoretical and Applied Genetics 75: 811-823.

Tanksley, S. and S. McCouch. 1997. Seed banks and molecular maps: unlocking genetic potential from the wild. Science 277: 1063.

Tatineni, V., R. Cantrell and D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Science 36: 186-192.

Thompson, W. 1931. Cytology and genetics of crosses between fourteen- and seven-chromosome species of wheat. Genetics 16: 309-324.

1964. Inheritance of fiber strength in two separate crosses of upland cotton having a common parent. Proc Sixteenth Annual Cotton Improvement Conference. Memphis, Tenn. National Cotton Council.

Turcotte, E. and R. Percy. 1990. Genetics of kidney seed in Gossypium barbadense L. Crop science 30: 384-386.

UGA-CAES. 2014. 2014 Georgia Farm Gate Value Report. University of Georgia - Center for Agribusiness and Economic Development.

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

Ullstrup, A. 1972. The impacts of the southern corn leaf blight epidemics of 1970-1971. Annual review of phytopathology 10: 37-50.

USDA-FAS. 2016. Cotton: World Markets and Trade. US Dept. of Agriculture - Foreign Agricultural Service, Washington, DC.

USDA-NASS. 2016. National statistics for cotton. In: USDA-NASS, editor Washington, DC.

Vamosi, J. and T. Dickinson. 2006. Polyploidy and diversification: A phylogenetic investigation in Rosaceae. International Journal of Plant Science 167: 349-358.

Van Becelaere, G., E.L. Lubbers, A.H. Paterson and P.W. Chee. 2005. Pedigree-vs. DNA marker-based genetic similarity estimates in cotton. Crop science 45: 2281-2287.

Van Esbroeck, G., D. Bowman, O. May and D. Calhoun. 1999. Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars.

Wailes, B.L.C. 1854. Report on the agriculture and geology of Mississippi.

Wallace, T., D. Bowman, B. Campbell, P. Chee, O. Gutierrez, R. Kohel, et al. 2009. Status of the USA cotton germplasm collection and crop vulnerability. Genetic Resources and Crop Evolution 56: 507-532.

Wang, B., W. Guo, X. Zhu, Y. Wu, N. Huang and T. Zhang. 2007. QTL mapping of yield and yield components for elite hybrid derived-RILs in upland cotton. Journal of Genetics and Genomics 34: 35-45.

Wang, B.,W. Yao-Ting, N.-T. Huang, Z. Xie-Fei, G. Wang-Zhen and T.-Z. Zhang. 2006. QTL mapping for plant architecture traits in upland cotton using RILs and SSR markers. Acta Genetica Sinica 33: 161-170.

Wang, F., Y. Gong, C. Zhang, G. Liu, L. Wang, Z. Xu, et al. 2011. Genetic effects of introgression genomic components from Sea Island cotton (Gossypium barbadense L.) on fiber related traits in upland cotton (G. hirsutum L.). Euphytica 181: 41-53.

Wang, F., Z. Xu, R. Sun, Y. Gong, G. Liu, J. Zhang, et al. 2013. Genetic dissection of the introgressive genomic components from Gossypium barbadense L. that contribute to improved fiber quality in Gossypium hirsutum L. Molecular breeding 32: 547-562.

Wang, G., J. Dong and A. Paterson. 1995. The distribution of Gossypium hirsutum chromatin in G. barbadense germ plasm: molecular analysis of introgressive plant breeding. Theoretical and Applied Genetics 91: 1153-1161.

Wang, H.M., Z.X. Lin, X.L. Zhang, W. Chen, X.P. Guo, Y.C. Nie, et al. 2008. Mapping and quantitative trait loci analysis of verticillium wilt resistance genes in cotton. Journal of Integrative Plant Biology 50: 174-182.

Ware, J.O. 1951. Origin, rise and development of American Upland cotton varieties and their status at presentUniversity of Arkansas College of Agriculture, Agricultural Experiment Station.

Webber, J. 1934. Cytogenetic notes on cotton and cotton relatives. Science 80: 268-269.

Webber, J.M. 1935. Interspecific Hybridization in Gossypium and the Meidtic Behavior of F1 Plants.

Webber, J.M. 1939. Relationships in the genus Gossypium as indicated by cytological data. J. agric. Res 58: 237-261.

Wendel, J. 2000. Genome evolution in polyploids. Plant Molecular Biology 42: 225-249.

Wendel, J., C. Brubaker and A. Percival. 1992. Genetic diversity in Gossypium hirsutum and the origin of upland cotton. American Journal of Botany 79: 1291-1310.

Wendel, J., A. Schnabel and T. Seelanan. 1995. An unusual ribosomal DNA sequence from Gossyipium gossypioides reveals ancien, cryptic, intergenomic introgression. Molecular Phylogenetics and Evolution 4: 298-313.

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

Westengen, O.T., Z. Huaman and M. Heun. 2005. Genetic diversity and geographic pattern in early South American cotton domestication. Theoretical and Applied Genetics 110: 392-402.

White, T. and T. Richmond. 1963. Heterosis and combining ability in top and diallel crosses among primitive, foreign, and cultivated American upland cottons. Crop Sci 3: 58-62.

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

Wu, J., O.A. Gutierrez, J.N. Jenkins, J.C. McCarty and J. Zhu. 2009. Quantitative analysis and QTL mapping for agronomic and fiber traits in an RI population of upland cotton. Euphytica 165: 231-245.

Xu, P., J. Zhu, X. Zhang, W. Ni, Y. Xu and X. Shen. 2012. Molecular mapping and identification of QTLs for fiber micronaire on chromosome 7 from Gossypium klotzschianum. Acta Agron. Sinica 38: 447-453.

Yoo, M.-J. and J.F. Wendel. 2014. Comparative evolutionary and developmental dynamics of the cotton (Gossypium hirsutum) fiber transcriptome.

Zeng, L. and W.R. Meredith. 2009. Associations among lint yield, yield components, and fiber properties in an introgressed population of cotton. Crop science 49: 1647-1654.

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

Zhang, J., Y. Lu, R. Cantrell and E. Hughs. 2005b. Molecular marker diversity and field performance in commercial cotton cultivars evaluated in the southwestern USA. Crop Science 45: 1483-1490.

Zhang, J., R.G. Percy and J.C. McCarty Jr. 2014. Introgression genetics and breeding between Upland and Pima cotton: a review. Euphytica 198: 1-12.

Zhang, K., J. Zhang, J. Ma, S. Tang, D. Liu, Z. Teng, et al. 2012. Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (Gossypium hirsutum L.). Molecular breeding 29: 335-348.

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

Zhang, Z.-S., Y.-H. Xiao, M. Luo, X.-B. Li, X.-Y. Luo, L. Hou, et al. 2005c. Construction of a genetic linkage map and QTL analysis of fiber-related traits in upland cotton (Gossypium hirsutum L.). Euphytica 144: 91-99.

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

CHAPTER 2

EVALUATION OF A CHROMOSOME SEGMENT FROM *GOSSYPIUM BARBADENSE* HARBORING THE FIBER LENGTH QTL *QUHM-CHR.25* IN FOUR DIVERSE UPLAND COTTON GENETIC BACKGROUNDS (*GOSSYPIUM HIRSUTUM*).

Nino Brown,* Pawan Kumar, Rippy Singh, Ed Lubbers, Todd Campbell, Gerald Myers, Robert Wright, Jay Subramani, Don Jones, Andrew H. Paterson, Peng W. Chee.

To be submitted to *Crop Science.*

EVALUATION OF A CHROMOSOME SEGMENT FROM *GOSSYPIUM BARBADENSE* HARBORING THE FIBER LENGTH QTL *QUHM-CHR.25* IN FOUR DIVERSE UPLAND COTTON GENETIC BACKGROUNDS (*GOSSYPIUM HIRSUTUM*).

Keywords

Cotton; fiber quality; interspecific; introgression; *Gossypium barbadense*.

Abbreviations

DP50, Deltapine 50

ELO, elongation

FBS, fiber bundle strength

GA089, GA2004089

HVI, high volume instrument

LPCT, lint percent

MIC, micronaire

NIL, nearly isogenic line

PM HS26, Paymaster HS26

QTL, quantitative trait locus

QTL+, lines with SL883 background, and with the SL883 allele at *qUHM-Chr.25*

QTL-, lines with SL883 background, but without the SL883 allele at *qUHM-Chr.25*

SFC, short fiber content

SL883, Sealand 883

SSR, single sequence repeat

UHML, upper half mean length

UI, uniformity index

INTRODUCTION

The quality of fibers produced by the Upland cotton plant, *Gossypium hirsutum*, has become a major focus of the cotton industry. Market pressures from textile mills, seeking more uniform, stronger, and longer cotton fiber, have encouraged cotton researchers to develop cultivars with improved fiber quality and production practices to maximize their genetic potential. A critical fiber quality to spinners today is fiber length, especially upper half mean length (UHML), which is the average length of the longest half of the fibers in a sample. The majority of textile production occurs in Asia today, making the cotton production market a major component of international trade. USDA-NASS (2016) estimates that global cotton production reached 96.9 million bales in 2015, 34.9 million of which were exported from the country producing them. The United States exported 10.2 million bales in 2015, most of the 12.8 million bales produced in the country that same year. Increased competition from man-made fibers to satisfy market trends from the mills as well as the fashion industry, are making it ever-more important to improve cotton fiber quality in order to maintain market competitiveness for US cotton producers.

While incremental gains have been achieved for UHML, those gains have been small and have shown signs of periodic stagnation (Helms, 2000; Lewis, 2001; May, et al., 1995; Meredith, et al., 1997). It has been assumed that stagnation in fiber quality improvement is due to lack of genetic diversity within the crop (Esbroeck and Bowman, 1998). To increase the genetic diversity within the available Upland cotton germplasm, it is an appealing strategy to introgress fiber traits from related species such as *Gossypium barbadense*, also known as Egyptian, Sea Island, and Pima cotton. However, interspecific introgression of *G. barbadense* into the more productive species, *G. hirsutum*, often results in deleterious effects such as hybrid breakdown

(Stephens, 1949), sterility, and abnormal segregation in the progeny (Jiang, et al., 2000; Reinisch, et al., 1994). These factors have made introgression difficult, however some successful cases involving extensive backcrossing and introgression efforts have been documented (Campbell, et al., 2011; Campbell, et al., 2010; Cantrell and Davis, 1993; Tatineni, et al., 1996). A useful strategy to incorporate some of these positive alleles from *G. barbadense* has been to use introgression lines with portions of *G. barbadense* genome in a primarily *G. hirsutum* genetic background (Percy, et al., 2006; Shen, et al., 2007; Sun, et al., 2012; Zeng and Meredith, 2009; Zhang, et al., 2012; Zhang, et al., 2009).

While numerous fiber quality QTLs have been mapped in interspecific populations or introgression lines derived from *G. barbadense* (Jiang, et al., 2000; Chen, et al., 2009; Draye, et al., 2005; Lacape, et al., 2003; Paterson, et al., 2003; Shen, et al., 2007; Wang, et al., 2011; Zhang, et al., 2014; Zhang, et al., 2009), few have been verified or have progressed to the point of being transferred into elite material. When a QTL of interspecific origin is introgressed into adapted germplasm, unexpected interactions with genetic background have been observed in soybean and tomato (Bernacchi, et al., 1998; Reyna and Sneller, 2001), including varied or total lack of effect in given backgrounds, unintended impacts on other traits, or lack of expression in different environments. In Upland cotton, Chee et al. (2005a, 2005b) and Draye et al. (2005), reported that genetic backgrounds can have profound interactions with effects of QTLs for fiber length, fineness, and elongation (ELO). For this reason, before extensive marker-assisted breeding efforts are undertaken, it is critical that the introgressed QTL allele be evaluated and verified to have effects across germplasm and target environments.

One of the longest-lived public US cotton breeding programs, USDA-ARS's Pee Dee breeding program in Florence, SC, used interspecific hybridization a great deal to influence the

improvement of *G. hirsutum* fiber quality in the US. An example of their use of interspecific introgression to improve fiber quality is Beasley's Triple Hybrid, a germplasm line derived from a cross between *G. arboreum*, *G. thurberi*, and *G. hirsutum*. This germplasm line has widely been credited for positively influencing much of the US cotton germplasm for fiber bundle strength (FBS). For example, Beasley's Triple Hybrid can be found in the foundational pedigrees of the New Mexico Acala SJ4 and Pee Dee breeding programs (Smith, et al., 1999), both of which have been incredibly important to the improvement of Upland cotton. More recently there have been fiber quality alleles introgressed from *G. arboreum* (Sun, et al., 2012), *G. tomentosum* (Zhang, et al., 2011), *G. mustelinum* (Wang, et al., 2016), and *G. klotzchianum* (Xu, et al., 2012) with significant effects.

Among some of the germplasm developed through interspecific hybridization by the Pee Dee breeding program were the "Sealand" lines. These germplasm lines were the result of a cross between "Bleak Hall", a Sea Island cultivar (*G. barbadense*) with extra-long staple (ELS) fiber, and "Coker Wilds", a *G. hirsutum* cultivar (Bowman, 2006; Culp and Harrell, 1974). The Sealand lines were considered *G. hirsutum* Upland cultivars, but while not grown extensively on a commercial scale, they exhibit excellent UHML and FBS. Previous research indicated that many of the introgressed segments from *G. barbadense* within the Sealand lines harbored important quantitative trait loci (QTL) for FBS and UHML (Kumar, 2012; Kumar, et al., 2012). One of these QTL, *qUHM-Chr.25*, had a large effect explaining from 4.4 to 61.4% of the variation seen in UHML in the original mapping populations. Located near the single sequence repeat (SSR) markers CIR267 and BNL827, the region was previously determined to span 7cM (Kumar, 2012).

The turn of the 20th century saw a more concerted effort in cotton variety development to produce cultivars adapted to the various macro-environments of the US Cotton Belt such as the Acala SJ4, Plains, Delta, and Eastern types (Meredith Jr, 1991; Niles and Feaster, 1984). Acala SJ4 type cultivars are grown primarily west of Texas, such as New Mexico and California. They are largely based upon the Mexican stocks from Acala SJ4 and Tuxtla, Mexico (Chiapas), with some degree of influence from Beasley's "Triple Hybrid," as well as some significant *G. barbadense* germplasm introgression (Meredith Jr, 1991; Niles and Feaster, 1984; Ware, 1951). Acala SJ4 cotton varieties are well known for their excellent fiber characteristics. In North and West Texas, the Plains type cultivars were developed, which are distinguished for their short stature, early season maturation, and their tight, wind resistant boll. These cultivars were largely derived from a variety called "Big Boll Stormproof" collected in Mexico in the 1850s (Niles and Feaster, 1984; Ware, 1951). The Paymaster group, a member of the Plains type cultivars, includes "Kekchi" germplasm collected in Guatemala by Cook (1905). The Delta cultivars, grown primarily in the Mississippi Delta as well as Beltwide, were developed from pre-1900 cultivars such as "Lone Star," which was derived from some 1860 and 1880 introductions from Mexico (Ramey, 1966; Ware, 1951). Delta and Pine Land, and Stoneville were early cotton seed companies that bred directly for the Mississippi Delta region. The fertile, uniform soils of the Delta were excellent grounds on which to breed, leading to many outstanding varieties that were grown across the belt and prized for their wide adaptability. The Eastern cultivars were largely made up of Coker germplasm, which includes many selections from the 19th and 20th century Mexican introductions, including some possible *G. barbadense* introgression (Niles and Feaster, 1984). The cottons grown in these macro-environments show significant phenotypic diversity,

however, molecular dissection of these relationships tell us that they are actually quite similar genetically.

The purpose of this study was to, (1), test the effect of the *qUHM-Chr.25* QTL when deployed in several genetic backgrounds adapted to the 4 major US growing regions when grown in those regions, (2) demonstrate the effect of the QTL in a nearly-isogenic state to determine the agronomic effect of the QTL, and (3), to further fine-map the location of the QTL on Chromosome 25 previously identified as being within a 7cM region near the SSR markers CIR267 and BNL827 (Kumar, 2012).

MATERIALS AND METHODS

The lines developed for this study were derived from the cross of Sealand 883 (hereon referred to as SL883)(PI 528875) with four regionally adapted cultivar parents, Acala SJ4 (PI 529538), Paymaster HS26 (PM HS26) (PVP 8600087, PI 606814), Deltapine 50 (DP50) (PVP 8400154, PI 529566), and an unreleased breeding line from the University of Georgia Molecular Cotton Breeding Laboratory, GA2004089 (GA089). SL883 is an obsolete cultivar developed and released in 1945 by the USDA Pee Dee Cotton Breeding program in South Carolina resulting from the cross of a *G. hirsutum* parent, Coker Wilds, with Bleak Hall, a *G. barbadense* Sea Island cotton cultivar (Bowman, 2006; Culp and Harrell, 1974). The resulting progeny were backcrossed several times to Coker Wilds to create what is now the Sealand germplasm lines, including SL883 which is an Upland cotton germplasm line with excellent fiber quality, especially fiber length, due in part to significant *G. barbadense* introgression (Bowman, 2006; Culp and Harrell, 1974).

Statistical analyses were conducted using SAS/STAT Software, Version 9.4 (SASInstitute, 2014). The PROC GLM function was used for analysis of variance and mean

separations. Waller-Duncan's method for mean separation was chosen as a more conservative approach than Fisher's for separating means when F-values are low, and similarly conservative to Fisher's when F-values are high.

Nearly Isogenic Line Evaluation

Nearly isogenic lines (NIL) developed for this study were derived from the crosses used to develop the populations for the Genotype x Environment experiment described below. SL883 was crossed with four regionally adapted cultivar parents, Acala SJ4, PM HS26, DP50, and GA089. In 2008, individual F_3 plants from these 4 families were identified based on their heterozygosity for the genomic region containing *qUHM-Chr.25*. Seed from these heterozygous F_3 individuals were used to plant segregating F_4 nursery rows for seed increase in 2011 and the nursery rows were bulk harvested.

A randomly selected sample of about 22 segregating $F_{3:5}$ sub-families from each of the four SL883 x cultivar crosses were screened for the presence or absence of the genomic region spanning the QTL of interest. The screening populations from each segregating family were grown in the greenhouse and leaf samples from individual plants were collected and tested for the presence or absence of the QTL allele using SSR markers that span the ~7cM region of interest, BNL827, BNL3359, CIR109, CIR267, and CIR 298; all of which were identified during the earlier mapping study to be polymorphic and capable of distinguishing between SL883 and at least one of the 4 regionally adapted cultivars (Kumar, 2012). These 5 SSR markers were used during the screening to ensure the region containing the QTL would not be lost through recombination during selfing generations. In order to construct the NILs, 10 individual plants homozygous for the SL883 allele, and 10 plants homozygous for the cultivar allele were identified from each of the segregating lines. The 10 individuals for each QTL class were then

bulk-harvested together to serve as the seed source of the homozygous NILs. From the SL883 x Acala SJ4 populations there were a total of 10 NILs derived from 5 lines; 5 NILs with the SL883 allele and the 5 corresponding NILs without the SL883 allele. Similarly, the SL883 x PM HS26 lines consisted of 12 NILs; 6 with the QTL, 6 without. The SL883 x DP50 NILs were made up of 19 NILs; 9 lines with the QTL, 10 without. The SL883 x GA089 NILs consisted of 28 NILs, 14 with, and 14 without the QTL.

Genomic DNA extraction followed (Paterson, et al., 1993), and PCR amplification conditions were carried out in a modified form as described by Chee et al. (2004) , on a T100 Thermal Cycler (Bio-Rad Laboratories, 2011). The 10µl PCR reactions contained approximately 10ng of DNA, 0.5µM of forward and reverse SSR primers, 100µdNTPs, 1.5 mM $MgCl_2$, 3U of DNA taq polymerase, and 1X reaction buffer containing 100mM Tris-HCl, 500 mM KCl, at pH 8.3. The PCR thermal cycling conditions were 94ºC for 3 min, then 34 cycles of 94ºC for 45 sec, 55ºC for 45 sec, 72ºC for 1 min, then incubation at 72ºC for 7 min, prior to cooling down for storage at 12ºC. Viewing of the PCR products were accomplished using 10% polyacrylamide gel electrophoresis and staining with silver nitrate similarly to previously published procedures (Zhang, et al., 2002).

Genotype x Environment Interaction of *qUHM-Chr.25*

In 2014 and 2015, the NILs from each family were grown in a randomized complete block design along with the SL883 parent and the relevant cultivar parents as checks. In 2014, the plots were single rows spaced 1m apart and 3.3m long. In 2015, the plots were 2-rows wide, spaced 1m apart, and approximately 10m long. A 30-boll sample was harvested from the mid-fruiting zone of the plants to represent the genetic potential for fiber quality and provide uniformity within the sample. These 30-boll samples were ginned on laboratory 10-saw gins in

2014, and on 20-saw table top gins in 2015 to determine lint percent of the seedcotton sample (LPCT). Fiber was analyzed with a high volume instrument (HVI) by the Fiber Quality Laboratory at Cotton Incorporated in Cary, NC.

The lines used for this experiment were derived from individual F_3 plants from the single-cross hybridization of SL883 with the 4 diverse cultivar parent genetic backgrounds. A total of 50 individuals, 25 of which were homozygous for the QTL locus and 25 homozygous without the QTL locus, were selected from each of the four cultivar parent genetic background families, for a total of 200 lines. The lines were selected at random based on the presence or absence of the SL883 allele using the nearby SSR markers CIR 267 and BNL 827. After a progeny row seed increase in 2010, replicated yield trials were conducted across the cotton belt of the $F_{3:5}$ and $F_{3:6}$ lines in 2011 and 2012, respectively. Trials were conducted at the University of Arizona, Maricopa, AZ; Texas Tech University, Lubbock, TX; Louisiana State University, Baton Rouge, LA; and Clemson University's Pee Dee Research and Education Center, Florence, SC. The field trials were arranged in a randomized complete block design with two replications at each site. Trial plots were single-row plots approximately 6m in length, with rows spaced 1m apart. Best growing practices were followed in the regions in which the trials appeared. Field tests were conducted in 2011 and 2012, however, a problem in 2011 seed increases forced the discard of 2012 trial data. Boll samples were harvested from the mid-fruiting zone from each plot to represent the best bolls on the plant and to ensure a more uniform sample. Twenty five bolls were hand-harvested from each plot. After harvest, boll samples were ginned on a laboratory style 10-saw gin. Fiber samples were sent to Cotton Incorporated's Fiber Quality Testing Laboratory in Cary, NC for HVI analysis.

Fine Mapping of *qUHM-Chr.25*

Segregating F_4 populations from each genetic background were screened for recombination within the delineated region of *qUHM-Chr.25* spanning ~7cM flanked by SSR markers NAU2713 and CIR109 which are approximately 11cM apart (Kumar, 2012). Of the 1185 plants genotyped, two were positively identified and confirmed as recombinant within the QTL region, one from the PM HS26 background, and one from the GA089 background. The seed from these recombinant plants were planted in the greenhouse for an increase of seed supply. In 2012, the seed increase nursery rows were bulk harvested and advanced to the F_6 generation in 2013, from which individual plants were harvested again, and the heterozygous individuals retained for further testing.

New SSR markers were developed to further saturate the previously identified region containing *qUHM-Chr.25*. The new markers were developed using the reference *G. raimondii* genome sequence (Wang, et al., 2012). The previously identified SSR marker sequences were used to conduct BLAST searches to identify new nearby SSR markers within the QTL region . Sixteen PCR primers were developed (Operon, Huntsville, AL), with four (UGT2501, UGT2504, UGT2509, and UGT2516) showing reliable amplification and polymorphic between parents . PCR amplification was performed as described in (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 (Kumar, 2012).

Four individual plants, heterozygous for the recombined region spanning the markers UGT2501 to UGT2509, and homozygous for the cultivar parent for the rest of the region of interest (Figure 2.3.1), were selected from the PM HS26 background. The 2011 recombinant plant from the PM background was designated plant number 1066. The heterozygous plants

selected in 2013 were designated, 1066-73, 1066-77, 1066-78, and 1066-93. From the other recombinant families, derived from a single plant in 2011 from the GA089 background, designated 5064, 3 heterozygous plants were selected for further testing, 5064-418, 5064-422, and 5064-440. These plants served as the source of our segregating groups in 2014 and 2015. In 2014, individual plants grown in the field in Tifton, GA were tagged and DNA was extracted from each of them. The populations were screened using a panel of SSR markers spanning the region of interest. The SSR markers used included UGT2501, UGT2504, CIR298, UGT2509, CIR267, UGT2516, and BNL827. The two recombinant families, 1066 and 5064, unfortunately do not have staggered segments of the SL883 chromosome, but had crossovers in a similar region. Both families have cultivar parent background allelic forms of BNL827, UGT2516, and CIR267. Each had SL883 alleles for the markers UGT 2501, UGT2504, CIR298, and UGT2509. In 2014, individual plants from the 3 genetic classes within these families were ginned on laboratory gins, and fiber was sent to the Fiber Testing Laboratory at Cotton Incorporated in Cary, NC where they were analyzed using HVI. The populations derived from plants 1066-73, 1066-77, 1066-93, and 5064-418 were only grown in 2014. The populations derived from 1066-78, 5064-422, and 5064-440, however, were tested in 2014 and 2015. In 2015, seeds from the individual plants harvested in 2014 were used to plant a replicated trial of these lines. Plots were single rows, spaced 1m apart, and were approximately 3.3m long, the trial was replicated twice. A representative 25-boll sample was harvested from the mid-fruiting zone within the plots, ginned on laboratory gins, and sent to Cotton Incorporated for HVI fiber testing.

RESULTS AND DISCUSSION

Nearly Isogenic Line Trial

The genotypes used for this study were selected based on the status of the SSR loci flanking the *qUHM-Chr.25* QTL. This QTL was identified previously as having a significant effect on UHML, and as such, the UHML trait is the main focus of this discussion. However, as shown in the analysis of variance (Table 2.1), significant differences existed for some of the other measured fiber traits. The main effect of Year had a significant influence on average micronaire (MIC), UHML, fiber bundle strength (FBS), and ELO. The influence of the cultivar background main effect was significant for MIC, UHML, UI, FBS, ELO, and LPCT. This result is expected as the influence of the cultivar backgrounds should play a significant role in all the measured fiber traits, especially as the cultivar parents selected for this study were purposely chosen for their diversity and regional adaptation. The interaction of cultivar background with Year was significant for MIC and SFC, both of which are known to be traits heavily influenced by the environment in which the cotton is grown. The effect of the QTL genotypic classes, that is, NILs with the QTL (QTL+), NILs without the QTL (QTL-), the cultivar parents (Acala SJ4, DP50, GA089, and PM HS26) and the SL883 parent, exhibited significant differences not only for UHML, but also for MIC, UI, FBS, ELO, SFC, and LPCT as well. A QTL by cultivar background interaction existed for UHML, as well as LPCT. It is an expected result for a QTL to be influenced by the genetic background in which it is deployed, especially for quantitatively inherited traits such as UHML and LPCT. These interactions only accounted for 8.9% and 18.7% of the total genetic sources of variance for UHML and LPCT, respectively, compared to the QTL source of variance, which accounted for 85.6% and 74.8% of the totals, respectively. A significant effect was observed in the interaction of QTL with Year for SFC. No significant

differences were observed for the measured fiber traits due to the interaction of QTL by Background by Year.

The parents utilized for this study were selected for their representation of the four major cotton growing regions in the US, and the uniqueness of their fiber traits and agronomic characteristics that make them desirable for those regions. The fiber quality diversity of these cultivars is presented in Table 2.2 and Table 2.3. The interspecific introgression parent SL883 averaged UHML of 33.5mm over the 2 years that these lines were tested. The regionally adapted cultivar parents all had significantly shorter fiber than SL883, ranging from a low of 28.6mm for DP50, to a high of 29.9mm from GA089. The QTL- NILs averaged 31.5mm, significantly lower than SL833, but significantly higher than all 4 regionally adapted parents. The UHML of the QTL+ lines were numerically higher than the QTL- lines by 0.9mm, however the differences were not statistically significant. This result indicates a large influence of the SL883 background on UHML for the NILs. Again, this is to be expected from a single-cross hybrid. The numerically higher UHML for the QTL+ lines supports the idea that the QTL influences UHML over the cultivar background alone, although it appears that there is either too much environmental variance or experimental error to clearly link it to the QTL. Further testing in more environments and with more replication would likely give more power to detect the positive influence of the QTL.

MIC for the 4 regional cultivars, Acala SJ4, DP50, GA089, and PM HS26 ranged from 4.8 to 4.1, all of which are within the non-discount range for MIC. SL883, had a much finer or immature fiber with an average MIC value of 3.1, significantly lower than any of the groups in this study. Due to the relatively high FBS, it is unlikely that the low MIC readings are due to fiber immaturity, but rather to fiber fineness. It's also interesting to note that the influence of

SL883 on the lines with and without the QTL was significant, both classes had low MIC values. It is to be expected that progeny derived from a single cross hybridization resulting from the cross of one of the 4 regionally adapted cultivars with SL883 would have significant influence from SL883 on their fiber phenotypes. This is evident in the NILs, which exhibit MIC values that are mostly lower than the regionally adapted parents. The QTL+ lines have significantly lower MIC values than the regionally adapted cultivars, though higher than SL883. The QTL- NILs exhibit MIC values that are similar to the QTL+ lines, but are not significantly different than Acala SJ4 or PM HS26.

UI varied little among the NILs. SL883 had the lowest UI of 83.3%, and PM HS26 the highest with 85.4%. The other entries, including the NILs, were intermediate, ranging from 84.8% to 84.1%. FBS was lowest for DP50 with 270.9 kN m kg^{-1}. SL883 had FBS of 310.3 kN m kg^{-1}, which was not different than that of PM HS26 and GA089 which exhibited FBS of 312.6 and 313.8 kN m kg^{-1}, respectively. The NILs had similar FBS to each other, SL883, Acala SJ4, GA089, and PM HS26, but were improved over DP50. ELO was highest for PM HS26 at 7.1%, and lowest for SL883 with 4.4%. ELO of the NILs were intermediate to that of SL883 and PM HS26, and statistically no different than GA089 with 5.6% for QTL- lines, 5.3% for the QTL+ lines, and 5.4% for GA089. DP50 exhibited the highest SFC of 7.9 % compared to the rest of the lines, which ranged from 7.3 to 6.8%. GA089 had the highest LPCT with 40.9%, and SL883 the lowest LPCT of 27.8%. The QTL- and QTL+ NILs were not different than each other, with 32.9% and 32.3% respectively. Isolated effect of *qUHM-Chr.25* within each NIL cultivar parent genetic background

The effect of *qUHM-Chr.25* was limited when averaged across cultivar backgrounds, and while not statistically significant, it was a meaningful increase in fiber length of 0.9mm averaged

over 2 years (Table 2.2). When analyzed separately by cultivar backgrounds, the effect of the QTL on UHML becomes more evident (Table 2.4). In addition to QTL, Year also had a significant effect on UHML test average for 3 of the 4 cultivar backgrounds. For the PM HS26 background, a significant difference was observed in average across replication, indicating that some spatial variation might have been present in the trial, but was adequately blocked. Significant differences existed among the NILs and parent cultivars tested for UHML in each cultivar background trial. CVs were low, from 2.75 to 4.2% for the 4 trials when separated by background, indicating that the field trial was properly executed to discern phenotypic differences among treatments.

The effect of *qUHM-Chr.25* among the DP50 NILs provided a significant and marked increase in UHML for the QTL+ lines, with an increase of 1.4mm over the QTL- lines. DP50 exhibited the lowest UHML (28.6mm) in this test, compared to 33.6mm for the SL883 parent. With no obvious introgression from *G. barbadense* in the pedigree (Kumar 2012), DP50 exhibits a noticeable increase in UHML due to the presence of the SL883 background, as well as a large additional increase from the presence of the *qUHM-Chr.25* QTL. The DP50 NILs averaged 32.7mm for the QTL+ lines, and 31.3 for the QTL- lines, an improvement of 4.5% over the SL883 background alone.

GA089, which had the longest UHML (29.9mm) of the Upland cotton cultivar backgrounds used in this study with an UHML of 29.9mm, was derived from a cross of PD 94042/AP 7126, neither of which have any known *G. barbadense* introgression. The effect of the SL883 background in the QTL- lines was a significant improvement of 1.9mm over the GA089 parent. The effect of *qUHM-Chr.25* in addition to the SL883 background was also

significant, accounting for an additional 1mm in UHML over the QTL- lines, or a 3.1% increase due to the SL883 *qUHM-Chr.25* allele alone.

Deployed within the Acala SJ4 genetic background, the *qUHM-Chr.25* allele had no significant impact on UHML over the base-line effect of the SL883 background genome alone (Table 2.5). These NIL groups were very similar in UHML, 31.9mm vs. 31.8 mm for the QTL- and QTL+ NILs, respectively. It is conceivable that Acala SJ4 might already harbor the causal allele responsible for improved UHML in this QTL region. Alternatively, it could have an undetectable effect due to epistatic interactions or it could be masked by the beneficial fiber alleles Acala SJ4 already possesses. This would explain why the presence or absence of the region from SL883 has no effect on UHML in addition to the effect of the SL883 background alone.

Similarly, the PM HS26 NILs harboring *qUHM-Chr.25* performed statistically no different for UHML than did their near isogenic counterparts. Although not statistically significant, there was a numerical mean difference of 0.6mm between the QTL+ and QTL- lines, representing about a 2% increase. The Acala SJ4 cottons are thought to have introgression from *G. barbadense* in their pedigrees (Smith, et al., 1999) that have potentially lent them excellent fiber quality and resistance to diseases such as Verticilium wilt (Zhang, et al., 2005). PM HS26 is the result of a cross between Acala SJ4 and 5B9-184, a re-selection from PM266 (Calhoun, et al., 1997). Based on the pedigrees of these 2 cultivars, it seems plausible that Acala SJ4 could already harbor the allele responsible for improved UHML, likely coming from *G. barbadense* for both SL883 and Acala SJ4. This result is quite interesting and suggests an avenue for further study.

Another potential explanation for lack of effect in the Acala SJ4 and PM HS26 backgrounds is that these genotypes were not tested in their region of adaptation, therefore the effect of the QTL could be masked by their relatively unadapted phenotypic response to the growing environment in the lower southeastern US. In contrast, GA089 and DP50 lines are right at home in their preferred growing region. It is possible that testing these lines in the desert Southwestern US (Maricopa, AZ) and the Texas High Plains (Lubbock, TX) would result in better expression of the QTL.

Genotype x Environment Interaction of *qUHM-Chr.25*

An analysis of the experimental variance revealed significant differences due to the Location effect for UHML. That the effect of Location, in this case macro-environments within the US cotton Belt, had a significant effect on average UHML of the trial, indicates that the testing environments should have been good for the differential expression of the QTL based on genotype's area of adaptation(Table 2.6). The main effect of cultivar parent genetic background (Background) had a significant effect on UHML, and ELO. This is not surprising because the cultivars selected in this study vary considerably in UHML and ELO. However, it is surprising that the other fiber traits such as MIC, FBS, and UI were not also affected by the genetic background effect, likely due to little variation in these traits among this set of cultivars. The QTL effect can be compared via 4 possible genotypic categories in this study; (1) the QTL allele from SL883 present in a cultivar background genome, (2) without the QTL allele but still a hybrid progeny of a regionally adapted cultivar parent by SL883, (3) original cultivar parent, or (4) the SL883 parent. Among these categories, there were significant differences due to the QTL effect for MIC, UHML, and ELO. The QTL*Background interaction effect was significant for MIC. The QTL*Background*Location effect was significant for UHML, accounting for about

4.3% of the total variance involving the QTL. Although a small effect, the results support our speculation that the QTL would have variable expression depending on the genetic background in which it is deployed and the environment in which that genetic background is being grown vs. its adaptation region. It was expected that the genotypes with the regionally adapted backgrounds would express the QTL disproportionately in their region of adaptation, and was the primary reason for testing the QTL in these environments.

HVI fiber data for the SL883 x regional cultivar families are shown in Table 2.7, indicating family average HVI fiber performance of each background cultivar parent. Analysis of variance indicated that differences for UHML and ELO were the only statistically significant differences. SL883 had the lowest ELO value of 3.47, and the cultivar families ranged from 3.62% for the Acala SJ4 background to 4.34% for the PMHS26 background. UHML was significantly different between each group. The longest fibers, as measured by UHML, were produced by SL883 with 33.2mm, approaching the length of Pima cotton, or extra-long staple cotton (ELS) at about 35mm. UHML mean of cultivar parents was highest for GA089, with 32.1mm. The Acala SJ4 background lines were significantly shorter at 31.7mm, followed by the DP50 background with 31mm, and PM HS26 was shortest with UHML of 30.4mm.

When the data is partitioned by genetic background, the effect of the QTL is apparent (Table 2.9). An increase of UHML due to presence of the QTL region is significant for 3 out of the 4 families, representing an increase of 0.6mm, 0.5mm, and 0.4mm for the Acala SJ4, DP50, and PMHS26 lines, respectively. Although not significant, an increase of 0.2mm was apparent in the GA089 lines due to presence of the QTL. While the presence of SL883 background genome in the lines tested increases UHML overall, the presence of the *qUHM-Chr.25* has an additional positive effect over the SL883 background alone.

QTL groups are combined across backgrounds for HVI fiber properties in Table 2.8. SL883 exhibited the lowest MIC readings of 3.67. Combined across the cultivar parent genetic backgrounds, all lines derived from the cross with SL883 and harboring *qUHM-Chr.25* allele (QTL+), and those lines crossed with SL883 but not carrying the *qUHM-Chr.25* from SL883 (QTL-) had similar MIC values. The cultivar parents exhibited higher average MIC than the other categories. As these progenies are single-cross hybrids, meaning that on average they contain 50% SL883 genome and 50% cultivar background genome, it is intuitive that the MIC values would be intermediate to the parents. ELONG exhibited a similar result, where SL883 averaged ELO of 3.5% across the 4 testing locations, the cultivar parents, 4.6%, and the QTL+ and QTL- lines were intermediate with 3.9% and 4.0% respectively.

Similarly, UHML of the lines QTL+ and QTL- were intermediate to that of SL883 and the regionally adapted cultivars. SL883 exhibited UHML of 33.2mm across the testing sites, whereas the cultivar parents averaged far shorter UHML of 29.6mm. The hybrid progeny of the cultivar parents and SL883, both QTL+ and QTL- were significantly shorter than SL883, but longer than their cultivar parents on average. Lines QTL+ were no different statistically than those lines QTL-, however the UHML mean was numerically higher for lines QTL+ at 31.5mm versus the lines QTL- with an average UHML of 31.1mm. We did not see a significant increase overall based on the presence of the *qUHM-Chr.25* QTL in addition to that of the SL883 genome alone. This is likely due to the small effect of the QTL, and/or the lack of sufficiently high numbers of replications within and/or across environments.

Fine Mapping

Segregation of the recombined chromosome segment potentially harboring *qUHM-Chr.25*, had a significant effect on UHML of the individual plants (IPs) tested in 2014 with at least α=0.10

(Table 2.10). This was also evident in the group that was tested in 2014 as IPs and again as lines in a replicated trial in 2015 (Table 2.11). For the families that were only tested in 2014 as individual plants, families 1066-73 and -77, from the PM HS26 background all showed significantly longer UHML when the SL883 portion of the recombined region was present, vs. the cultivar portion (Table 2.12). The effect on UHML was an increase of 0.9mm, 1.5mm, and 1.3mm for families 1066-73, -77, and -93 respectively. For the family derived from the GA089 background, 5064-418, the effect was an increase of 1mm, not statistically significant at α=0.05 however it was at α=0.10. CV was quite low for these families within each of the genetic classes, ranging from 3.2 to 4.1%.

Similar results were seen in the replicated testing of these segregating lines, where the effect of the segregation of the recombinant chromosome region was statistically significant (Table 2.11). For the family 1066-78, which is a PM HS26 background genotype, UHML differed significantly from sister lines without the SL883 segment by 1.4mm overall (Table 2.12). For the two GA089 families, 5064-422 and -440, the effect of the SL883 portion of the chromosome segment improved UHML 0.9mm and 0.5mm, respectively. Again, CVs were quite low, ranging from 2.9 to 3.2%.

While these 2 recombinant lineages, 1066 from the PM HS26 background and 5064 from the GA089 background genotypes are unable to unequivocally resolve whether or not the other portion of the previously delineated QTL region harbors other QTL that influence UHML, the evidence suggests that the area between UGT2501 and UGT2509 harbors *qUHM-Chr.25* or part of it.

CONCLUSION

When the *qUHM-Chr.25* QTL was tested in 4 cultivars selected to represent genetic backgrounds adapted to the four macro-environments where Upland cotton is primarily grown in the United States cotton belt, the effect of this QTL was small, but significant in 3 out of the 4 cultivar backgrounds tested. When measured in a near-isogenic state comparing NILs with and without the QTL allele from SL883, the effect of *qUHM-Chr.25* on UHML was significant in 2 of the 4 cultivar backgrounds tested. The effect of the QTL, when averaged across the backgrounds tested in this trial, accounted for an additional 0.9mm in the QTL+ lines, though the difference was not significantly different than the QTL- lines. When separated by background, the effect of the QTL was not significantly greater than the SL883 background alone when deployed within the Acala SJ4 or PM HS26 background, however it was significant for the DP50 and GA089 backgrounds, with an UHML increase of 1.4mm (4.5%) and 1mm (3.1%), respectively. It is worth mentioning that the PM HS26 QTL+ lines did have a numerically higher average UHML of 0.6mm better than the QTL- lines, though the difference was not statistically significant. To give these numbers some perspective, in over 40 years of breeding the US cotton crop has gone from an average UHML of 26.7mm in 1975 to 28.2mm in 2014, an increase of only 1.5mm (5.6%) (Incorporated, 2015). The utilization of foreign alleles from closely related species by way of introgression lines, such as the SL883 *qUHM-Chr.25* allele, has the potential to make significant impacts on the cotton industry in relatively few years if deployed in commercial germplasm.

Testing of recombinant lines suggest that the QTL likely falls within the region flanked by UGT2501 and UGT2509. This region had a highly significant effect on lines from the PMHS26 as well as the GA089 background. Because the testing to narrow down the region of

interest was conducted on highly inbred recombinant sister lines, the data derived from the fiber samples was highly uniform, and likely made the effect of the QTL that much more apparent.

qUHM-Chr.25 represents a *G. barbadense*-sourced QTL for the improvement of Upland cotton germplasm. The effect of this QTL is significant, and along with the tightly linked marker sequence included here, should prove to be beneficial in improving Upland cotton germplasm. The portability of this QTL across diverse genetic backgrounds should make it valuable to breeders interested in improving fiber quality, a driving theme in cotton breeding today.

WORKS CITED

Bernacchi, D., T. Beck-Bunn, D. Emmatty, Y. Eshed, S. Inai, J. Lopez, et al. 1998. Advanced back-cross 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. Theoretical and Applied Genetics 97: 170-180.

Bio-Rad Laboratories, I. 2011. T100 Thermal CyclerBio-Rad Laboratories, Inc. - Life Science Group, Hercules, California.

Bowman, D. 2006. Pedigrees of upland and pima cotton cultivars released between 1970 and 2005.

Calhoun, D.S., D. Bowman and O.L. May. 1997. Pedigrees of upland and pima cotton cultivars released between 1970 and 1995Office of Agricultural Communications, Division of Agriculture, Forestry, and Veterinary Medicine, Mississippi State University.

Campbell, B., P. Chee, E. Lubbers, D. Bowman, W. Meredith, J. Johnson, et al. 2011. Genetic improvement of the Pee Dee cotton germplasm collection following seventy years of plant breeding. Crop Science 51: 955-968.

Campbell, B., S. Saha, R. Percy, J. Frelichowski, J. Jenkins, W. Park, et al. 2010. Status of the global cotton germplasm resources. Crop science 50: 1161-1179.

1993. Characterization of Upland× Pima breeding lines using molecular markers. Proc. Beltwide Cotton Conf., New Orleans, LA.

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

Chen, H., N. Qian, W. Guo, Q. Song, B. Li, F. Deng, et al. 2009. Using three overlapped RILs to dissect genetically clustered QTL for fiber strength on Chro. D8 in Upland cotton. Theoretical and applied genetics 119: 605-612.

Cook, O. 1905. Cotton culture in Guatemala. Yearbook of the United States Department of Agriculture: 1904. Government Printing Office, Washington, DC.

Culp, T. and D. Harrell. 1974. Breeding quality cotton at the Pee Dee experiment station Florence. SC USDA ARS-S-30, New Orleans, LA 71.

Draye, X., P. Chee, C. Jiang, L. Decanini, T. Delmonte, R. Bredhauer, et al. 2005. Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense (cotton) by

a backcross-self approach: II. Fiber fineness. TAG Theoretical and Applied Genetics 111: 764-771.

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

2000. Yield study report. Proc Beltwide Cotton Production Conf. National Cotton Council of America, San Antonio, Tex.

Incorporated, C. 2015. Cotton Fiber Competition. Cotton Incorporated, Raleigh.

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

Kumar, P. 2012. Genetic evaluation of exotic chromatins from two obsolete interspecific introgression lines of Upland cotton. University of Georgia, Athens, GA.

Kumar, P., R. Singh, E.L. Lubbers, X. Shen, A.H. Paterson, B.T. Campbell, et al. 2012. Mapping and validation of fiber strength quantitative trait loci on chromosome 24 in upland cotton. Crop science 52: 1115-1122.

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

2001. A review of yield and quality trends and components in American upland cotton. Proc. beltwide cotton conf., Anaheim, CA.

May, O., D. Bowman and D. Calhoun. 1995. Genetic diversity of US upland cotton cultivars released between 1980 and 1990.

Meredith Jr, W. 1991. Contributions of introductions to cotton improvement. In: H. Shands and L. Wiesner, editors, Use of plant introductions in cultivar development, Part I. ASA, CSSA, SSSA, Madison, WI. p. 127-146.

Meredith, W.R., J. Heitholt, W. Pettigrew and S. Rayburn. 1997. Comparison of obsolete and modern cotton cultivars at two nitrogen levels. Crop science 37: 1453-1457.

Niles, G. and C. Feaster, editor. 1984. Breeding. Cotton Agronomy Monograph. ASA, Madison, WI.

Paterson, A., Y. Saranga, M. Menz, C.-X. Jiang and R. Wright. 2003. QTL analysis of genotype× environment interactions affecting cotton fiber quality. Theoretical and Applied Genetics 106: 384-396.

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

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

1966. Historical review of cotton variety development. Proc. 18th Cotton Improvement Conf., Memphis, TN.

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

Reyna, N. and C. Sneller. 2001. Evaluation of marker-assisted introgression of yield QTL alleles into adapted soybean. Crop Science 41: 1317-1321.

SASInstitute. 2014. SAS/STAT Software. SAS Institute Inc., Cary.

Shen, X., W. Guo, Q. Lu, X. Zhu, Y. Yuan and T. Zhang. 2007. Genetic Mapping of Quantitative Triat Loci for FIber Quality and Yield Trait by RIL Approach in Upland Cotton. Euphytica 155: 371-380.

Smith, C.W., R. Cantrell, H. Moser and S. Oakley. 1999. History of cultivar development in the United States. Cotton: Origin, history, technology, and production. John Wiley & Sons, New York: 99-171.

Stephens, S. 1949. The cytogenetics of speciation in Gossypium. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 34: 627.

Sun, F.-D., J.-H. Zhang, S.-F. Wang, W.-K. Gong, Y.-Z. Shi, A.-Y. Liu, et al. 2012. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Molecular breeding 30: 569-582.

Tatineni, V., R. Cantrell and D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Science 36: 186-192.

USDA-NASS. 2016. National statistics for cotton. In: USDA-NASS, editor Washington, DC.

Wang, B., L. Liu, D. Zhang, Z. Zhuang, H. Guo, X. Qiao, et al. 2016. A Genetic Map between Gossypium hirsutum and the Brazilian Endemic *G. mustelinum* and its Application to QTL Mapping. G3: Genes| Genomes| Genetics 6: 1673-1685.

Wang, F., Y. Gong, C. Zhang, G. Liu, L. Wang, Z. Xu, et al. 2011. Genetic effects of introgression genomic components from Sea Island cotton (Gossypium barbadense L.) on fiber related traits in upland cotton (G. hirsutum L.). Euphytica 181: 41-53.

Wang, K., Z. Wang, F. Li, W. Ye, J. Wang, G. Song, et al. 2012. The draft genome of a diploid cotton Gossypium raimondii. Nature genetics 44: 1098-1103.

Ware, J.O. 1951. Origin, rise and development of American Upland cotton varieties and their status at presentUniversity of Arkansas College of Agriculture, Agricultural Experiment Station.

Xu, P., J. Zhu, X. Zhang, W. Ni, Y. Xu and X. Shen. 2012. Molecular mapping and identification of QTLs for fiber micronaire on chromosome 7 from Gossypium klotzschianum. Acta Agron. Sinica 38: 447-453.

Zeng, L. and W.R. Meredith. 2009. Associations among lint yield, yield components, and fiber properties in an introgressed population of cotton. Crop science 49: 1647-1654.

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

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

Zhang, J., R.G. Percy and J.C. McCarty Jr. 2014. Introgression genetics and breeding between Upland and Pima cotton: a review. Euphytica 198: 1-12.

Zhang, K., J. Zhang, J. Ma, S. Tang, D. Liu, Z. Teng, et al. 2012. Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (Gossypium hirsutum L.). Molecular breeding 29: 335-348.

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

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

Table 2.1. Analysis of variance mean squares for nearly isogenic lines harboring *qUHM-Chr.25* deployed within an Acala SJ4, Paymaster HS26, Deltapine 50, and Georgia 089 genetic background compared to the cultivar parents and Sealand 883 parent in Tifton, GA in 2014 and 2015.

Source		MIC†	UHML	UI	FBS	ELO	SFC	LPCT
	df‡							
Year	1	326.3*	91.7*	507.4	963.4*	21739.1	134.7*	6.4
Rep	1	2.8	0.6	0.2	28.2	0.1	1.9	0.5
Error A	1	0.1	0.5	30.1	5.3	1719.5	0.2	0.3
Background	3	22.7**	23.4**	50.7*	46.7*	61154.6**	6.7	28.3**
Background*Year	3	35.4**	0.7	2.2	14.8	1152.2	9.9*	0.3
Error B	10	3.2	0.5	10.8	7.7	1084.1	2.4	0.2
QTL	4	270.8**	39.3**	28.7*	199.0**	63783.4**	14.5**	8.0**
QTL* Background	3	10.1	4.1*	19.5	8.1	461.6	5.5	2.0*
QTL*Year	4	5.2	1.1	5.7	23.3	3046.7	8.8*	0.1
QTL* Background *Year	3	3.4	0.2	5.3	11.6	3104.5	2.3	0.1
Error C	274§	8.8	1.2	11.4	19.4	2685.2	2.9	0.5

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† MIC, micronaire; UHML, upper half mean length; UI, uniformity index; FBS, fiber bundle strength; ELO, elongation; SFC, short fiber content; LPCT, lint percent.

‡ df, degrees of freedom.

§ Error C df for LPCT are 275.

Table 2.2. Mean UHML values of nearly isogenic lines for *qUHM-Chr.25* deployed within several regionally adapted genetic backgrounds, Acala SJ4, Paymaster HS26 (PM HS26), Deltapine 50 (DP 50), and Georgia 2004089 (GA089) compared to the cultivar parents and Sealand 883 (SL883) parent in Tifton, GA in 2014 and 2015 field trials.

	UHML†
	mm
QTL (-)	31.5b‡
QTL (+)	32.4b
Acala SJ4	29.2cd
DP50	28.6d
GA089	29.9c
PM HS26	29.0cd
SL883	33.5a
CV	3.4

† UHML, upper half mean length.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 2.3. Agronomically important fiber trait mean values of nearly isogenic lines for *qUHM-Chr.25* deployed within several regionally adapted genetic backgrounds, Acala SJ4, Paymaster HS26 (PM HS26), Deltapine 50 (DP 50), and Georgia 2004089 (GA089) compared to the cultivar parents and Sealand 883 (SL883) parent in Tifton, GA in 2014 and 2015 field trials.

	MIC	UI	FBS	ELO	SFC	LPCT
	unit	%	kN m kg^{-1}	%	%	%
QTL (-)	4.0cd	84.4ab	303.6ab	5.6c	7.3b	32.9c
QTL (+)	3.7d	84.2bc	306.5ab	5.3c	7.1b	32.3cd
Acala SJ4	4.1bc	84.6ab	297.6b	6.2b	7.0b	32.7c
DP50	4.8a	84.1bc	270.9c	6.6b	7.9a	37.0b
GA089	4.3b	84.8ab	313.8a	5.4c	7.3b	40.9a
PM HS26	4.2bc	85.4a	312.6a	7.1a	6.8b	30.3d
SL883	3.1e	83.3c	310.3ab	4.4d	6.9b	27.8e
CV	7.7	1.3	4.5	9.5	7.8	7

† MIC, micronaire; UI, uniformity index; FBS, fiber bundle strength; ELO, elongation; SFC, short fiber content; LPCT, lint percent.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 2.4. Analysis of variance mean squares for the effect of *qUHM-Chr.25* QTL in a nearly-isogonic state on fiber upper half mean length (UHML) deployed within an Acala SJ4, Paymaster HS26 (PM HS26), Deltapine 50 (DP 50), and Georgia 2004089 (GA089) genetic background in Tifton, GA in 2014 and 2015.

	Acala SJ4		DP 50		GA089		PM HS26
	df‡		df		df		
Year	1	24.82	1	5574.96*	1	105.05**	1	22704.98**
Rep	1	1.95	1	0.02	1	0.35	1	155.58**
Error A	1	0.56	1	14.63	1	0.17	1	0.01
QTL	3	13.04**	3	4670.67**	3	25.87**	3	14777.59**
QTL*Year	3	0.42	3	78.78	3	1.67	3	351.38
Error B	37	1.80	81	138.87	117	0.79	40	1258.84
CV		4.2		3.7		2.8		3.6

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

‡ df, degrees of freedom.

Table 2.5. Effect of *qUHM-Chr.25* QTL on fiber upper half mean length (UHML) deployed in a nearly isogenic state within Acala SJ4, Paymaster HS26 (PM HS26), Deltapine 50 (DP 50), and Georgia 2004089 (GA089) genetic backgrounds compared to the cultivar parents (Bcknd Parent) and Sealand 883 (SL883) in Tifton, GA in 2014 and 2015.

	Acala SJ4	PM HS26	DP 50	GA089
	mm	mm	mm	mm
Bcknd Parent†	29.2c	29.0c	28.6c	29.9d
QTL (-)	31.9b	30.6b	31.3b	31.8c
QTL (+)	31.8b	31.2b	32.7a	32.8b
SL883	33.5a	33.3a	33.6a	33.6a
CV	4.2	3.6	3.7	2.8

† Bcknd Parent, Cultivar background parent; QTL+, Lines with *qUHM-Chr.25* from SL883; QTL- lines without *qUHM-Chr.25* from SL883.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 2.6. Analysis of variance mean squares for HVI fiber quality parameters in field trials of *qUHM-Chr.25* deployed within 4 regionally adapted cultivars, Acala SJ4, Paymaster HS26, Deltapine 50, and Georgia 089, compared to the Sealand 883 parent in Maricopa, AZ, Lubbock, TX, Baton Rouge, LA, and Florence, SC in 2011.

Source	df‡	MIC	UHML	FBS	ELO	UI
Rep	1	4.59	0.024	6.38	0.55	8.02
Location	3	3.46	0.146*	42.19	1.76	36.61
Error A	3	1.69	0.005	19.42	0.20	6.66
Background	3	1.14	0.057**	21.85	5.91**	3.68
Background*Location	12	0.43	0.007	3.57	0.45	1.91
Error B	15	0.54	0.007	7.56	0.51	2.18
QTL	2	5.50**	0.082**	1.47	4.76**	0.24
QTL*Background	6	0.42*	0.004	4.80	0.40	2.13
QTL*Background*Location	20	0.22	0.004*	2.08	0.26	1.00
Error C	1520	0.17	0.002	3.15	0.19	1.06

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† MIC, micronaire; UHML, upper half mean length; FBS, fiber bundle strength; ELO, elongation; UI, uniformity index; Background, cultivar parent genetic background.

‡ df, degrees of freedom.

Table 2.7. SL883 x regional cultivar family group means for HVI fiber quality parameters in field trials of *qUHM-Chr.25* deployed within 4 regionally adapted cultivars, Acala SJ4, Paymaster HS26 (PM HS26), Deltapine 50 (DP 50), and Georgia 089 (GA089) compared to the Sealand 883 (SL883) parent in Maricopa, AZ, Lubbock, TX, Baton Rouge, LA, and Florence, SC in 2011.

		MIC†	UHML	FBS	ELO	UI
	n	unit	mm	kN m kg^{-1}	%	%
Acala SJ4	393	4.1	31.7c‡	311.7	3.62d	82.6
DP50	391	4.3	31.0d	306.8	4.19b	82.9
GA089	390	4.2	32.1b	318.3	3.81c	83.3
PMHS26	387	4.2	30.4e	322.9	4.34a	83.2
SL883	26	3.7	33.2a	312.1	3.47e	82.8

† MIC, micronaire; UHML, upper-half mean length; FBS, fiber bundle strength; ELO, fiber bundle elongation; UI, uniformity index.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 2.8. QTL x Background means for HVI fiber quality parameters in field trials of *qUHM-Chr.25* deployed within 4 regionally adapted cultivars, Acala SJ4, Paymaster HS26 (PM HS26), Deltapine 50 (DP 50), and Georgia 2004089 (GA089) compared to the Sealand 883 (SL883) parent in Maricopa, AZ, Lubbock, TX, Baton Rouge, LA, and Florence, SC in 2011.

	Acala SJ4	DP50	GA089	PMHS26
QTL+†	32.1a‡	31.3a	32.2a	30.7a
QTL-	31.5b	30.8b	32.0a	30.3b
Cultivars	30.5b	n/a	n/a	28.5c

†QTL+, Lines with *qUHM-Chr.25* from SL883; QTL- lines without *qUHM-Chr.25* from SL883.

‡Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 2.9. Combined across-background QTL group means for HVI fiber quality parameters in field trials of *qUHM-Chr.25* deployed within 4 regionally adapted cultivars, Acala SJ4, Paymaster HS26, Deltapine 50, and Georgia 089 (Cultivars) compared to the Sealand 883 (SL883) parent in Maricopa, AZ, Lubbock, TX, Baton Rouge, LA, and Florence, SC in 2011.

		MIC†	UHML	FBS	ELO	UI
	n	units	mm	kN m kg-1	%	%
SL883	26‡	3.67c	33.2a	312.1	3.5c	82.8
QTL+	808	4.13b	31.5b	315.5	3.9b	83
QTL-	736	4.27b	31.1b	314.3	4.0b	83
Cultivars	17	4.63a	29.6c	311.6	4.6a	82.8

†MIC, micronaire; UHML upper-half mean length; FBS, fiber bundle strength; ELO, elongation; UI, uniformity index.

‡Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 2.10. Analysis of variance mean squares for UHML means from individual of several recombinant families segregating for the region containing *qUHM-Chr.25*, grown in Tifton, GA in 2014.

	PM HS26						GA089	
	1066-73†		1066-77		1066-93		5064-418	
	df‡	MS	df	MS	df	MS	df	MS
Source								
QTL	2	7.40*	2	8.92*	2	3.92	2	2.61
Error	57	0.92	52	1.04	40	1.38	68	1.63

* Significant at the 0.05 probability level.

† Families derived from individual heterozygous F6 plants designated 1066-73, 1066-77, 1066-93, 5064-418.

‡ df, degrees of freedom.

Table 2.11. Analysis of variance mean squares for UHML means from individual plants and replicated plots from recombinant families for the region containing *qUHM-Chr.25*, grown in Tifton, GA in 2014 and 2015.

	PM HS26		GA089			
	1066-78†		5064-422		5064-440	
	df‡	MS	df	MS	df	MS
Source						
QTL	2	20.03**	2	13.98**	2	4.20*
Rep	2	0.11	2	3.14*	2	1.45
Error	165	0.87	221	0.87	223	0.98

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† Families derived from individual heterozygous F6 plants designated 1066-78, 5064-422, and 5064-440.

‡ df, degrees of freedom.

Table 2.12. UHML means from individual plants from the 3 genetic classes of several recombinant families for the region containing *qUHM-Chr.25*, grown in Tifton, GA in 2014.

	PM HS- 26†						GA089	
	n	1066-73‡	n	1066-77	n	1066-93	n	5064-418
		mm		mm		mm		mm
Cultivar§	14	29.8b¶	16	28.6b	8	28.4	29	30.7
Heterozygous	28	29.6b	27	29.6a	25	29.1	28	31.2
SL883	18	30.7a	12	30.1a	10	29.7	15	31.7
CV		3.2		3.5		4.0		4.1

† Paymaster HS26 background families, PM HS26; Georgia 089 background families, GA089.

‡ Families derived from individual heterozygous F_6 plants designated 1066-73, 1066-77, 1066-93, 5064-418.

§ Identity of SSR-marker loci UGT2501 – UGT2509.

¶ Within columns, means followed by the same letter are not significantly different according to Waller-Duncan LSD, k=100.

Table 2.13. UHML means from individual plants and replicated plots from the 3 genetic classes of several recombinant families for the region containing *qUHM-Chr.25*, grown in Tifton, GA in 2014 and 2015.

			2014		2015				
Background	Family		n	IPS†	n	Rep 1	n	Rep 2	Overall Mean
				mm		mm		mm	mm
PM HS26	1066-78‡	Cultivar	15	28.7	14	28.8	14	28.8	28.8c
		Heterozygous	33	29.3	30	29.5	25	29.3	29.4b
		SL883	14	30.3	12	30.0	13	30.1	30.2a
		CV							3.2
				mm		mm		mm	mm
GA089	5064-422	Cultivar	28	31.0	13	31.2	13	31.0	31.1c
		Heterozygous	48	31.8	30	31.8	30	31.4	31.7b
		SL883	31	32.3	16	32.1	17	31.5	32.0a
		CV							2.9
				mm		mm		mm	mm
GA089	5064-440	Cultivar	27	31.5	22	31.0	23	30.7	31.1b
		Heterozygous	38	31.4	33	31.4	33	31.5	31.4ab
		SL883	25	31.5	15	31.7	12	31.6	31.6a
		CV							3.2

† IPS, individual plant data; Rep 1 and Rep 2, 30-boll sample from replicated plots.

‡ Families derived from individual heterozygous F6 plants designated 1066-78, 5064-422, and 5064-440.

CHAPTER 3

REGISTRATION OF R01-40-08, A *GOSSYPIUM HIRSUTUM*, UPLAND COTTON GERMPLASM LINE WITH *QFL-CHR.1* INTROGRESSED FROM *G. BARBADENSE* CONFERRING IMPROVED FIBER LENGTH

Nino Brown,* Xinlian Shen, Rippy Singh, Edward L. Lubbers, C. Wayne Smith, Pawan Kumar, Don Jones, Andrew H. Paterson, PengW. Chee.

To be submitted to *Journal of Plant Registrations.*

REGISTRATION OF R01-40-08, A *GOSSYPIUM HIRSUTUM*, UPLAND COTTON GERMPLASM LINE WITH *QFL-CHR.1* INTROGRESSED FROM *G. BARBADENSE* CONFERRING IMPROVED FIBER LENGTH.

Keywords

Cotton; fiber quality; interspecific; introgression; *Gossypium barbadense*.

Abbreviatons

ELO, elongation

FBS, fiber bundle strength

HVI, high volume instrument

LP, lint percent

MAS, marker assisted selection

MIC, micronaire

NIL, nearly isogenic line

QTL, quantitative trait locus

RFLP, restriction fragment length polymorphism

SFC, short fiber content

UHML, upper half mean length

UI, uniformity index

YLD, lint yield

INTRODUCTION

Upland cotton, *Gossypium hirsutum*, is used extensively across the globe for production of textiles. This unique fiber, feed and oil crop is especially valued for the ability of its fiber to produce a spinnable thread with which to make comfortable cotton clothes. The cotton industry has seen increased competition from man-made fibers in recent decades due to lower synthetic fiber prices and the high cost of cotton fiber production. Modern crop improvement in cotton has focused primarily on yield to bridge the price gap for the industry, however, modern textile production technologies are requiring a more uniform, longer, and stronger fiber that can be spun cheaply and efficiently into a high-quality textile yarn. A primary focus of improving fiber yield has likely contributed to stagnation of cotton fiber quality improvement. For example, from 1975 to 2015, Upland cotton yields have seen an increase of 70% from 500kg/ha to 850 kg/ha, whereas average cotton fiber upper half mean length (UHML) has increased only 5.6% from 26.7mm in 1975 to 28.2mm in 2014 (Incorporated, 2015). The evident stagnation in cotton fiber quality improvement is concerning given that the available *G. hirsutum* germplasm from which to draw shows little genetic diversity, especially among the commercial germplasm (Bertini, et al., 2006; Fang, et al., 2013; Van Becelaere, et al., 2005). The increasing competitive pressure of man-made fibers such as polyester provide further impetus for the improvement of Upland cotton fiber for the benefit of cotton producers in this country and abroad.

Gossypium barbadense, also known as Pima, Sea Island, or Egyptian cotton, exhibits fiber quality parameters that far exceed its cousin, Upland cotton. It has long been a goal of public and private breeders to introgress some of the excellent fiber traits of *G. barbadense* into the higher yielding and widely adapted *G. hirsutum*, but has had limited success. Hybrid breakdown (Stephens, 1949), sterility, and abnormal segregation (Jiang, et al., 2000; Reinisch, et

al., 1994) and other complications have hampered incorporation of positive alleles from *G. barbadense*. However, some important examples of successful incorporation of alleles from interspecific sources exist. Extensive backcrossing and introgression efforts have been made to improve many traits, but especially fiber quality in *G. hirsutum* (Campbell, et al., 2011; Campbell, et al., 2010; Cantrell and Davis, 1993; Tatineni, et al., 1996). Successful cases of introgression usually involve the incorporation of the interspecific segments of DNA by way of introgression lines rather than direct or single cross hybrids (Percy, et al., 2006; Shen, et al., 2007; Sun, et al., 2012; Zeng and Meredith, 2009; Zhang, et al., 2012; Zhang, et al., 2009).

The germplasm line reported here was derived from an inbred backcross mapping population of the recurrent parent, *G. hirsutum* cv. TAMCOT 2111 (Smith and Niles, 1991) (PVP 9100221, PI 554643), and *G. barbadense* cv., PIMA S-6 (Feaster and Turcotte, 1984) (PI 608346) as the donor of *qFL-Chr.1* for improved UHML (Chee, et al., 2005; Shen, et al., 2011; Xu, et al., 2017). The QTL of interest, *qFL-Chr1* was originally identified in a BC_3F_2 mapping population and explained 12-24% of the variation for UHML in the original population (Chee, et al., 2005). The line selected for release, R01-40-08, was derived from an individual plant selection in a BC_3F_3 line and harbors *qFL-Chr.1*, a QTL for improved fiber length from *G. barbadense* (Shen, et al., 2011). The development of R01-40-08 and the mapping population from whence it came is the result of a collaboration between the Plant Genome Mapping Lab at the University of Georgia in Athens, GA; Molecular Cotton Breeding Lab at the University of Georgia in Tifton, GA; and the Cotton Improvement Lab at Texas A&M University in College Station, TX. R01-40-08 represents a *G. hirsutum* line with improved UHML derived from the introgression of a portion of *G. barbadense* genome harboring a QTL for improved UHML, and

will be a useful resource for cotton breeders to improve their breeding programs using marker assisted selection (MAS) for more rapid fixation in their breeding pools.

MATERIALS AND METHODS

Originally mapped using an inbred backcross population from the cross of *G. hirsutum* cv. TAMCOT 2111 (recurrent) and *G. barbadense* cv. PIMA S-6 (donor), the *qFL-Chr1* QTL demonstrated a significant effect on UHML (Chee, et al., 2005). TAMCOT 2111 was the result of a cross between PD 6142 (Culp, et al., 1985) (CSR 87, GP 250) and an unreleased Texas Agrilife Research germplasm line. A series of 140 near-isogenic lines (NIL) were developed to further explore and demonstrate the effectiveness of the QTL. One of the BC_3F_2 parent plants used to develop the NILs, R01-40, was heterozygous for the QTL region of interest at restriction fragment length polymorphism (RFLP) marker pGH377 and p9-54, both of which flank the QTL region spanning about 50 cM (Shen, et al., 2011). R01-40 was fixed at the majority of other chromosomal regions for one parent or the other. BC_3F_3 families were derived from selfing R01-40, within which 24 plants were self-pollinated in the greenhouse and advanced to the BC_3F_4 to ensure genetic purity. Three sister lines (R01-40-03, R01-40-08, R01-40-20) from the R01-40 family were identified that exhibited both improved UHML and seemed to display beneficial agronomic potential.

To evaluate the agronomic performance of these lines, field trials were conducted in 2009, 2013, and 2014 at Tifton and Plains, GA. Soil types encountered at the Gibbs Research Farm in Tifton, GA include Tifton loamy sand, fine-loamy, kaolinitic, thermic Plinthic Kandiudult; and Clarendon loamy sands, fine-loamy, siliceous, semiactive, thermic Plinthaquic Paledult. Soils at the Southwest Georgia Research and Education Center in Plains, GA include Greenville sandy loam, a fine, kaolinitic, thermic Rhodic Kandiudult; Faceville sandy loam, fine,

kolinitic thermic Typic Kandiudults; and Tifton sandy loams. Field trials were laid out in a randomized complete block design with 2 row plots that were 12 m long and spaced 1 m apart. Standard regional production practices were employed each year for the field trials. Twenty-five bolls were harvested by hand from the mid-fruiting zone from each plot and ginned on a 10-saw laboratory gin. Fiber data was analyzed using the High volume instrument (HVI) by the Fiber Quality Laboratory at Cotton Incorporated in Cary, NC. Fiber yield was measured by harvesting the plots with a two-row cotton spindle picker modified for harvesting and weighing yield trial plots. Statistical analyses were conducted using the PROC GLM function in the SAS/STAT software, Version 9.4 of the SAS System for Windows (SASInstitute, 2014).

CHARACTERISTICS

Analysis of variance mean squares revealed that the main effect of Year had a significant effect on the experimental mean of UHML, UI, fiber bundle strength (FBS), and elongation (ELO) (Table 3.1). Location means were significantly different for lint yield (YLD) and micronaire (MIC), which is typical as these two traits are highly influenced by the environment in which they are grown; Tifton and Plains represent two very different growing regions in southern Georgia. Again, YLD and MIC means were significantly different for Year x Location, as well as short fiber content (SFC), likely due to the influence of the environment. Differences between genotypes were significant for each of the measured traits, YLD, lint percent (LP), MIC, UHML, UI, FBS, ELO, and SFC. LP and FBS showed an interaction with Genotype x Year, although they represented a small portion of the total Genotypic variance at 26.5% and 10.7%, respectively.

The R01-40 germplasm lines exhibited average MIC readings that were similar to both parents, PIMA S-6 and TAMCOT 2111, or intermediate to the parents (Table 3.2). R01-40-03 exhibited the highest MIC of 4.69, which was not different than TAMCOT 2111. R01-40-08 had an intermediate MIC of 4.53, and R01-40-20 had the lowest MIC of the germplasm lines. Low MIC can indicate either a fine fiber if the fibers are mature, or it can also be indicative of low fiber maturity, that is, poor secondary cell wall deposition within the fiber cells, which makes the fiber weak and the dye-uptake poor. The probability that R01-40-20 has poor fiber maturity is doubtful as its parent is PIMA S-6, which is known for having fine fibers when mature. Each of these three germplasm lines have MIC values well within normal or non-discount ranges, with R01-40-20 approaching the upper premium threshold of 4.3.

UHML improvement was significant both statistically as well as biologically. Introgression of this fiber length QTL from PIMA S-6 improved the UHML by at least 1mm in the case of R01-40-03, and as high as 1.5mm for R01-40-08, which is equivalent to nearly 2 staple units (1/32 inch). Although significantly shorter than PIMA S-6, the three germplasm lines had significantly longer UHML than TAMCOT 2111, representing a 3-5% increase in UHML due to the introgressed fiber length QTL. The PIMA S-6 parent used for comparison in these trials, showed uncharacteristically low mean UHML possibly because it is not adapted to the Southeast production region. Alternatively, the PIMA S-6 utilized in the test might have experienced some genetic drift from years of increasing seed in an area with a large population of pollinators.

R01-40-03, had the highest UI followed by R01-40-08 and TAMCOT 2111. Differences between these three were not significant. PIMA S-6 and R01-40-20 had lower UI, indicating that fiber length was slightly more variable in these genotypes than the others in the trial. It is

important to note, however, that all of the UI values seen here are within the premium range given to growers. So while the differences are significant, they are not agronomically meaningful since they are all at premium level for UI, which gives further evidence that these genotypes produce fibers that are desirable for spinners. SFC was highest for R01-40-20 with 7.61%, which makes it relatively more undesirable than the other lines. The other genotypes tested had significantly lower SFC than R01-40-20.

TAMCOT 2111 and R01-40-03 made up the significance group with the highest ELO values averaged across the 6 testing environments. R01-40-03 and R01-40-20 were similar in the second group, followed by R01-40-20, R01-40-08, and PIMA S-6 in the third significance group. It is interesting to note here that the lines with the shortest fibers had the highest ELO values such as TAMCOT 2111 and R01-40-03. Conversely, lines R01-40-08 and PIMAS S-6, which had the longest fibers had the lowest ELO values.

TAMCOT 2111 and R01-40-08 were the two highest yielding lines in the trial, and were not significantly different from each other for YLD. R01-40-20, PIMA S-6, and R01-40-03 were the lowest yielding lines. R01-40-08 was clearly the superior sister-line as it had the highest YLD and significantly improved fiber quality over the TAMCOT 2111 recurrent parent, especially for fiber length due in large part to the incorporation of the fiber length QTL *qFL-Chr1*. R01-40-08 is unique in that it demonstrates a significant improvement in fiber length of 5% over the recurrent parent, TAMCOT 2111, while simultaneously maintaining the desirable YLD potential of TAMCOT 2111. It could be argued that a 5% gain over several years is not groundbreaking. However, it is indeed unique to see a QTL of *G. barbadense* origin introgressed into *G. hirsutum* to make significant improvement in the fiber without sacrificing the YLD characteristics of the recurrent parent due to linkage drag. This QTL has a modest

effect, but using the molecular markers described previously, breeders can rapidly screen large populations prior to planting to eliminate those lines without the QTL, thereby focusing their limited resources toward lines with higher genetic potential for UHML. Germplasm resources such as these may prove beneficial to the cotton breeding community, as genetic gain becomes more stagnant for fiber quality traits. R01-40-08 is also unique in that introgression of fiber quality traits from *G. barbadense* rarely has resulted in genotypes that yield as well as the *G. hirsutum* recurrent parent.

LP and FBS values over the course of these trials showed interesting results, and a significant GxE interaction deserving of discussion here. Most of the genotypes, including the R01-40 sister lines and TAMCOT 2111, exhibited consistent LP as well as FBS values. PIMA S-6, however, did not show a great deal of consistency for either trait (Table 3.3). Across the three testing years, R01-40-08 exhibited LP values equal to or better than the recurrent parent, TAMCOT 2111, which was in the highest LP group. While R01-40-08 indeed has fiber length similar to PIMA S-6, it does not exhibit similar FBS. The FBS values of R01-40-08 were consistently lower than PIMA S-6, but equivalent to or better than the recurrent parent.

JUSTIFICATION FOR RELEASE

Before the advent of modern spinning technologies, the focus by breeders was on lint yield per hectare. In today's cotton industry, textile mills require a much higher quality fiber than previous varieties have supplied. The limited genetic diversity of Upland cotton germplasm does not lend itself to rapid genetic gain of the commercial Upland crop. It is for this reason that many public cotton breeding programs in the US have put increasing focus on fiber quality improvement.

The germplasm line, R01-40-08, represents a milestone in Upland cotton germplasm improvement in that, as far as we know, it is the first publicly released germplasm carrying a purposefully introgressed *G. barbadense* fiber length QTL using MAS whose effect has been validated. R01-40-08 yields similarly to the recurrent, *G. hirsutum* parent, TAMCOT 2111, and has improved UHML from the introgressed QTL, *qFL-Chr.1* from *G. barbadense* donor, PIMA S-6. This introgression line is unique in that it is a predominantly *G. hirsutum* genetic background, yields well, and has introgression of interspecific origin. R01-40-08 may serve well the cotton industry in improving the fiber and spinning quality of American Upland cotton, and with the available tightly linked SSR markers published previously by this laboratory (Shen, et al., 2011; Xu, et al., 2017), this germplasm line will be an important addition to the Upland cotton breeding lexicon. The use of this introgression and the associated markers will help breeders to improve fiber length.

AVAILABILITY

Inquiries regarding availability of seed for research purposes or for commercial use should be directed to the corresponding author.

WORKS CITED

Bertini, C.H., I. Schuster, T. Sediyama, E.G.d. Barros and M.A. Moreira. 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genetics and Molecular Biology 29: 321-329.

Campbell, B., P. Chee, E. Lubbers, D. Bowman, W. Meredith, J. Johnson, et al. 2011. Genetic improvement of the Pee Dee cotton germplasm collection following seventy years of plant breeding. Crop Science 51: 955-968.

Campbell, B., S. Saha, R. Percy, J. Frelichowski, J. Jenkins, W. Park, et al. 2010. Status of the global cotton germplasm resources. Crop science 50: 1161-1179.

1993. Characterization of Upland× Pima breeding lines using molecular markers. Proc. Beltwide Cotton Conf., New Orleans, LA.

Chee, P.W., R. Bredhauer, C.W. Smith, A.H. Paterson, T.A. Delmonte, X. Draye, et al. 2005. 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.

Culp, T., R. Moore and J. Pitner. 1985. Registration of seven cotton germplasm lines. Crop Sci 25: 201-202.

Fang, D.D., L.L. Hinze, R.G. Percy, P. Li, D. Deng and G. Thyssen. 2013. A microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton (Gossypium hirsutum L.) cultivars from major cotton-growing countries. Euphytica 191: 391-401.

Feaster, C. and E. Turcotte. 1984. Registration of Pima S-6 cotton. Crop Sci 24: 382.

Incorporated, C. 2015. Cotton Fiber Competition. Cotton Incorporated, Raleigh.

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

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

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

SASInstitute. 2014. SAS/STAT Software. SAS Institute Inc., Cary.

Shen, X., Z. Cao, R. Singh, E.L. Lubbers, P. Xu, C.W. Smith, et al. 2011. Efficacy of qFL-chr1, a Quantitative Trait Locus for Fiber Length in Cotton (spp.). Crop science 51: 2005-2010.

Shen, X., W. Guo, Q. Lu, X. Zhu, Y. Yuan and T. Zhang. 2007. Genetic Mapping of Quantitative Triat Loci for FIber Quality and Yield Trait by RIL Approach in Upland Cotton. Euphytica 155: 371-380.

Smith, C. and G. Niles. 1991. Tamcot 2111: A high strength cotton cultivar for central and south Texas. TAES Leaflet 2465.

Stephens, S. 1949. The cytogenetics of speciation in Gossypium. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 34: 627.

Sun, F.-D., J.-H. Zhang, S.-F. Wang, W.-K. Gong, Y.-Z. Shi, A.-Y. Liu, et al. 2012. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Molecular breeding 30: 569-582.

Tatineni, V., R. Cantrell and D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Science 36: 186-192.

Van Becelaere, G., E.L. Lubbers, A.H. Paterson and P.W. Chee. 2005. Pedigree-vs. DNA marker-based genetic similarity estimates in cotton. Crop science 45: 2281-2287.

Xu, P., J. Gao, Z. Cao, P.W. Chee, Q. Guo, Z. Xu, et al. 2017. Fine mapping and candidate gene analysis of qFL-chr1, a fiber length QTL in cotton. Theoretical and Applied Genetics: 1-11.

Zeng, L. and W.R. Meredith. 2009. Associations among lint yield, yield components, and fiber properties in an introgressed population of cotton. Crop science 49: 1647-1654.

Zhang, K., J. Zhang, J. Ma, S. Tang, D. Liu, Z. Teng, et al. 2012. Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (Gossypium hirsutum L.). Molecular breeding 29: 335-348.

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

Table 3.1. Analysis of variance mean squares for R01-40 germplasm lines, and the parents TAMCOT 2111 and PIMA S-6 grown in 2009, 2013, and 2014 at Tifton and Plains, GA.

Source		YLD†		LP		MIC	UHML	UI	FBS	ELO	SFC
	df‡		df§		df¶						
Year	1	15	2	17	2	65	46*	61*	1695*	579*	49
Error A	2	48	4	6	4	12	5	3	155	74	12
Location	1	3440**	1	2	1	137**	19	5	1	36	4
Year x Location	1	326**	2	8	2	81*	189	19	452	110	124*
Error B	4	4	6	3	6	8	41	9	99	42	12
Genotype	4	96**	4	24**	4	59**	182**	34**	10208**	174**	135**
Genotype x Year	4	31	8	13**	8	11	27	7	1293**	49	21
Genotype x Location	4	44	4	5	4	2	6	3	47	4	11
Genotype x Year x Location	4	19	8	3	8	8	13	11	274	16	17
Error C	32	22	48	4	47	5	14	7	246	34	1

*Significant at $p \leq 0.05$ probability level.

**Significant at $p \leq 0.01$ probability level.

†YLD, lint yield; LP, lint percent; MIC, micronaire; UHML, upper-half mean length; UI, uniformity index; FBS,fiber bundle strength; ELO, fiber bundle elongation; SFC, short fiber content.

‡Degrees of freedom for YLD.

§Degrees of freedom for LP.

¶Degrees of freedom for MIC, UHML, UI, FBS,ELO, SFC.

Table 3.2. Genotype means for HVI fiber qualities and lint yield (YLD) in R01-40 germplasm lines, and the parents TAMCOT 2111 and PIMA S-6 grown in 2009, 2013, and 2014 at Tifton and Plains, GA.

Genotype	MIC†	UHML	UI	ELO	SFC	YLD
	units	mm	index	%	%	kg/h
PIMA S-6	4.37c‡	32.1a	84.4bc	5.4c	7.1b	869c
R01-40-03	4.69a	31.0b	85.2a	6.1ab	7.0b	901bc
R01-40-08	4.53b	31.5b	84.9ab	5.7c	7.2b	1033ab
R01-40-20	4.37c	31.3b	84.0c	5.8bc	7.6a	856c
TAMCOT 2111	4.73a	30.0c	84.6ab	6.3a	7.1b	1076a
CV%	5.1	3.0	1.0	10.0	3.9	17.6

†MIC, micronaire; UHML upper-half mean length; UI, uniformity index; ELO, elongation; SFC, short fiber content; YLD, lint yield.

‡Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 3.3. Genotype means by year for lint percent (LP) and HVI fiber bundle strength (FBS) for R01-40 germplasm lines, and the parents TAMCOT 2111 and PIMA S-6 grown in 2009, 2013, and 2014 at Tifton and Plains, GA.

	LP†			FBS		
Genotype	2009	2013	2014	2009	2013	2014
	%	%	%	g/tex	g/tex	g/tex
PIMA S-6	38.0ab‡	34.1b	35.7ab	39.8a	36.1a	34.2a
R01-40-03	36.7ab	35.5ab	35.9a	30.8b	32.8bc	31.7b
R01-40-08	40.3a	36.5a	35.8a	31.8b	33.6b	31.1bc
R01-40-20	34.2b	34.6b	34.2b	31.0b	31.9cd	30.7c
TAMCOT 2111	34.8b	36.9a	36.0a	30.1b	31.3d	30.8bc
CV%	8.5	4.0	3.1	7.0	3.6	2.7

†LP, lint percent; FBS,fiber bundle strength.

‡Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

CHAPTER 4

AGRONOMIC PERFORMANCE OF *QFL-CHR.1* FROM *GOSSYPIUM BARBADENSE* INTROGRESSED FROM GERMPLASM LINE R01-40-08 INTO FOUR UPLAND COTTON (*G. HIRSUTUM*, L.) GENETIC BACKGROUNDS

Nino Brown,* Xinlian Shen, Rippy Singh, Edward L. Lubbers, Pawan Kumar, Don Jones, Andrew H. Paterson, Peng W. Chee.

To be submitted to *Crop Science*.

AGRONOMIC PERFORMANCE OF *QFL-CHR.1* FROM *GOSSYPIUM BARBADENSE* INTROGRESSED FROM GERMPLASM LINE R01-40-08 INTO FOUR UPLAND COTTON (*G. HIRSUTUM*, L.) GENETIC BACKGROUNDS

Keywords

Cotton; fiber quality; interspecific; introgression; *Gossypium barbadense*.

Abbreviations

DP50, Deltapine 50

ELO, elongation

FBS, fiber bundle strength

GA089, GA2004089

HVI, high volume instrument

LPCT, lint percent

MIC, micronaire

PM HS26, Paymaster HS26

QTL, quantitative trait locus

RIL, recombinant inbred line

SFC, short fiber content

UHML, upper half mean length

UI, uniformity index

INTRODUCTION

Fiber quality of both natural and synthetic fibers continues to be of great importance to the textile industry. Competition from synthetic fibers and economic factors have made cotton (*Gossypium* spp.) fiber quality improvement more important for the continued success of cotton production in the US (Deussen, 1993; Faerber and Deussen, 1994). For the three major spinning technologies currently used in yarn production, ring, air-jet, and rotor spinning, fiber upper half mean length (UHML) is an important characteristic that gives the yarn strength and determines the fineness of the yarn produced (Bedez Üte and Kadoğlu, 2014; Ramey, et al., 1977). Fiber length, however, has been shown to be negatively correlated with yield (Azhar, et al., 2004; Meredith and Bridge, 1971; Miller, et al., 1958; Scholl and Miller, 1976; Shen, et al., 2007; Smith and Coyle, 1997), and as such requires focus from breeders and geneticists to try to improve quality without sacrificing yield.

Fiber quality improvement of cotton has likely been hampered by the limited genetic diversity within the crop caused by the genetic bottlenecks of domestication (Iqbal, et al., 2001), followed by selective breeding. Several studies have been conducted that support this idea, and have shown a great deal of genetic similarity within the modern Upland cotton (*Gossypium hirsutum*) germplasm base (Abdurakhmonov, et al., 2008; Bertini, et al., 2006; Fang, et al., 2013; Pillay and Myers, 1999; Rungis, et al., 2005). One of *G. hirsutum*'s closely related species, *G. barbadense*, or Pima cotton, is renowned for its exceptional fiber length and strength and is inter-fertile with *G. hirsutum* to some extent. Introgression of fiber traits from *G. barbadense* is possible but the early generations of hybrid populations often exhibit hybrid break down (Stephens, 1949), lower levels of recombination (Reinisch, et al., 1994), infertility, unexpected segregation ratios in later generations (Jiang, et al., 2000), and agronomically undesirable

offspring. Extensive backcrossing and introgression efforts have been made to improve many traits, but especially fiber quality in *G. hirsutum* (Campbell, et al., 2011; Campbell, et al., 2010; Cantrell and Davis, 1993; Tatineni, et al., 1996). Successful cases of introgression usually involve the incorporation of the interspecific segments of DNA by way of introgression lines rather than direct or single cross hybrids (Shen, et al., 2007; Sun, et al., 2012; Zhang, et al., 2012; Zhang, et al., 2009). Percy et al. (2006), using stabilized introgression lines, created a group of recombinant inbred lines (RILs) that had significant portions of *G. barbadense* introgression in a *G. hirsutum* background, and identified several RILs that exhibited positive transgressive segregation for lint percent (LPCT) and 2.5% span length, as well as no significant correlation between lint yield and fiber quality traits (Percy, et al., 2006).

One such introgression line, R01-40-08, developed by the University of Georgia Molecular Cotton Breeding Laboratory, was derived from an advanced backcross mapping population of the recurrent parent, *G. hirsutum* cv. TAMCOT 2111 (PVP 9100221, PI 554643), and the *G. barbadense* cultivar, Pima S-6 (Feaster and Turcotte, 1984) (PI 608346)(Shen, et al., 2011). R01-40-08 carries an introgressed quantitative trait locus (QTL) from Pima S-6, *qFL-Chr.1*, that explained 12-24% of the variation for UHML in the original mapping population, and conferred an average increase in UHML of 1.5mm, or 5% over the recurrent parent across 3 years of testing in Tifton and Plains, GA (Brown et al., in review, 2017). While this QTL had a decidedly significant effect in the original population, it was unclear if the QTL would have a similarly significant effect on UHML in other genetic backgrounds. Interactions between QTL, genetic background, and/or environment could influence the expression of a given QTL (Lecomte, et al., 2004; Sebolt, et al., 2000; Tanksley and Hewitt, 1988), making validation an important step after mapping an economically important QTL. Interactions between QTL and

genetic background are further expected when it is of interspecific origin (Chee, et al., 2005a; Chee, et al., 2005b; Draye, et al., 2005). The purpose of this study was to deploy *qFL-Chr.1*, which had a large effect on UHML in the original bi-parental population, in several genetic backgrounds adapted to the 4 major growing regions of the US, to determine its suitability as a candidate QTL for improving UHML in the US germplasm base.

MATERIALS AND METHODS

The lines used in this study were developed from the crossing of inbred introgression line, R01-40-08, which harbors *qFL-Chr.1*, with four regionally adapted cultivars chosen for their iconic representation of regional adaptation. These cultivar parents include Acala SJ4 (PI 529539), representing germplasm of the arid Southwest; Paymaster HS-26 (PM HS26, PI 606814) representing the stripper-harvested cultivars typical of the Texas High Plains; Deltapine 50 (DP50, PI 529566), a formerly popular cultivar in the Mississippi Delta production region; and GA2004089 (GA089), an unreleased breeding line from the University of Georgia Molecular Cotton Breeding Laboratory, typical of Southeastern type cotton cultivars.

Initial cross pollinations were made in 2011, with F_1 plants grown at the USDA Cotton Winter Nursery in Tecoman, Mexico. About 300 F_2 plants were individually harvested from each of the four R01-40-08 by cultivar background families. The $F_{2:3}$ progeny rows, grown in 2013, served as seed increase for the following year. Replicated trials were conducted in 2014 at the University of Georgia, Gibbs Research Farm in Tifton, GA and the Southwest Georgia Research and Education Center in Plains, GA. In 2015, replicated trials were planted only at the Gibbs Research Farm. Soil types encountered at the Gibbs Research Farm in Tifton, GA include Tifton loamy sand, fine-loamy, kaolinitic, thermic Plinthic Kandiudult; and Clarendon loamy sands, fine-loamy, siliceous, semiactive, thermic Plinthaquic Paledult . Soils at the University of

Georgia, Southwest Georgia Research and Education Center in Plains, GA include Greenville sandy loam, a fine, kaolinitic, thermic Rhodic Kandiudult; Faceville sandy loam, fine, kaolinitic thermic Typic Kandiudults; and Tifton sandy loams.

Field trials were laid out in a randomized complete block design, replicated twice at each location. In 2014 at Plains, plots were 2 rows wide spaced approximately 1m apart, and were 10m in length. In Tifton in 2014, plots were 2 rows wide, also spaced 1m apart, but were 5m long. In 2015, at Tifton, plots were approximately 7m long, with rows spaced 1m apart. A 30-boll sample was collected from each plot by hand from the mid-fruiting zone of the plants. After collection of boll samples, plots were harvested using a 2-row spindle type cotton picker modified for weighing plots, and plot weights were collected to estimate lint yield. Boll samples were ginned on laboratory gins to collect LPCT data and the fiber samples were sent for high volume instrument (HVI) testing at the Fiber Testing Laboratory of Cotton Inc. in Cary, NC. Statistical analyses were conducted using the PROC GLM function in the SAS/STAT software, Version 9.4 of the SAS System for Windows (SASInstitute, 2014). For the analysis of variance, the lines tested were separated into QTL classes which included (1) the cultivar parent checks; (2) the QTL donor parent R01-40-08; (3) lines resulting from the cross of a cultivar parent by R01-40-08 that are homozygous for the QTL (QTL+), (4) lines heterozygous for the (QTL(h)), and (5) lines homozygous without the Pima S-6 allele (QTL-), but crossed with R01-40-08.

RESULTS AND DISCUSSION

Analysis of variance revealed significant differences for all measured traits except for LPCT for the main effect of Environments (Table 4.1), indicating that testing environments were different. The effect of Background was significant for all of the measured agronomic and fiber traits. Significant interaction of Background*Environment was evident for LYLD, UHML, uniformity

index (UI), fiber bundle strength (FBS), elongation (ELO), and short fiber content (SFC). The QTL classes exhibited significant differences for LPCT, micronaire (MIC), UHML, UI, FBS, ELO, and SFC. Significant interaction of QTL*Background was apparent for MIC, UHML, and FBS. A significant difference in LPCT means was evident for the three-way interaction of QTL*Background*Environment.

Because of the significant interaction of the QTL*Background in the combined analysis (Table 4.1), it was necessary to run a separate analysis of variance in which the QTL effect could be analyzed separately for each of the 4 cultivar parent backgrounds in which the QTL was deployed (Table 4.2). The analysis of variance indicated that the effect of QTL class on UHML was significant for each genetic background, and that the DP50 background QTL lines exhibited a significant interaction of QTL*Environment. Mean separations for the genotypes tested are given in Table 4.3. For the lines in the Acala SJ4 data subset, the Acala SJ4 cultivar parent had the longest UHML of the group, with 31.2mm, which was significantly longer than R01-40-08 with 29.5mm. No significant difference was observed between those lines with Acala SJ4 background that were QTL+, QTL(h), or QTL-. When deployed within the DP50 and GA089 backgrounds, UHML was not affected by QTL status in the lines as no significant differences were observed between the QTL+, QTL(h) and QTL- lines. In the PM HS26 derived lines, the QTL(h) lines exhibited UHML that was significantly higher on average than the QTL+ lines, but were not significantly different than the QTL- lines.

Across the 4 testing environments, and averaged across genetic backgrounds, QTL+ lines showed no significant difference for UHML over the QTL- or QTL(h) lines, however, means were numerically different (Table 4.4). QTL(h) lines had the longest average UHML (29.0mm), followed by QTL+ lines with 28.7mm, and QTL- lines with 28.6mm. The donor parent, R01-40-

08, exhibited UHML of 29.5mm, which was significantly longer than the QTL+ and QTL- lines, but not significantly longer than the QTL(h) lines or any of the cultivar parents used in the trial, except Acala SJ4 which had an UHML of 31.2mm. Because the lines tested in this experiment were single cross hybrids, it is possible that the background effect of the TAMCOT 2111 within R01-40-08, could have brought the average UHML down, and segregation could be masking or confounding the effect of the QTL in these R01-40-08 x cultivar progenies. Recall that the interaction of QTL*Background was significant for UHML, and the means for each of the 4 cultivar backgrounds will be discussed later.

UI, which is a measure of length consistency, was highest for GA089, Acala SJ4, and PM HS26, ranging from 85.7% to 85.3%, indicating that their fiber length distribution is more uniform than R01-40-08 with 83.8%. The QTL+, QTL(h), and QTL- lines exhibited UI of 83.3%, 83.5%, and 83.7%, respectively, similar to R01-40-08 (83.8%) and DP50 (84.1%). Similarly, SFC was lowest for Acala SJ4, PM HS26, and GA089 with 6.9%, 7.0%, and 7.3%, respectively. R01-40-08, which had SFC content of 7.6%, was significantly higher than Acala SJ4 and PM HS26, and was not different than DP50 with 7.8%. The SFC of the QTL+, QTL(h), and QTL- lines were not significantly different from R01-40-08 with an average of 7.7%, 7.6% and 7.6%, respectively.

Although the trait of interest related to *qFL-Chr.1* is UHML, one of the goals of this experiment was to demonstrate the agronomic suitability of these lines as well as measure the effect of the QTL in the genetic backgrounds in which it was deployed. Table 4.5 shows the means for agronomic and fiber quality parameters measured during the course of this study. No significant difference for YLD was apparent in the analysis of variance, however GA089 followed by DP50 were the highest yielding lines numerically. The QTL(h) group was next

followed by QTL-, PM HS26, QTL+, and Acala SJ4. LPCT was highest for Acala SJ4 (42.3%) and GA089 (39.8%), and lowest for PM HS26 with 33.5%. LPCT of the QTL+, QTL(h), and QTL- lines were not significantly different, which is not surprising since the parents all had similar LPCT values except for PM HS26, which had low LPCT and Acala SJ4 had the highest in the group. MIC was highest for the hybrid lines and DP 50, ranging from 5.00 to 4.81 units, indicating that these fibers were either coarse or very mature. Acala SJ4 exhibited MIC that was significantly lower than any of the rest of the lines tested in this trial with 4.06 units, indicating a fine or immature fiber.

FBS was similar for R01-40-08, and the QTL+, QTL(h), and QTL- lines ranging from 296 kN m kg^{-1}, to 290 kN m kg^{-1}. Acala SJ4 had the highest FBS with 322 kN m kg^{-1}, followed by GA089 with 310 kN m kg^{-1}, and 306 kN m kg^{-1} for PM HS26, which were both significantly lower than Acala SJ4, but not significantly different than each other. DP50 exhibited the lowest FBS of the cultivar parents with 267 kN m kg^{-1}. ELO showed few differences, with only GA089 and Acala SJ4 exhibiting significantly lower ELO of 5.0% and 5.5% than the rest, which ranged from 7.1% to 6.6%.

CONCLUSION

The limited effect of this QTL in these 4 genetic backgrounds suggests that the effect of qFL-CHr.1 is either too small to be measured in this experiment, was confounded by a high degree of segregation still occurring in the lines, or that the QTL has no effect when deployed within these backgrounds. The first year of this experiment consists of individual plant fiber data, which is inherently not as accurate as boll samples harvested from an entire row. Because of the inclusion of all bolls in an individual plant sample, there is more variability in the data, whereas a 25-boll sample harvested from the mid-fruiting zone and picked from an entire plot of

a given line, is a much more uniform sample, and therefore better for detecting small differences between genotypes. The lines tested were single-cross hybrids, and as such, likely had a significant influence from background segregation occurring between R01-40-08, which is essentially TAMCOT 2111 plus the *qFL-Chr.1* QTL from Pima S-6, and the cultivar parents. The segregation of these divergent genetic backgrounds could very easily have overshadowed the effect of this QTL, which has been shown to have a modest effect (previous chapter). Lastly, there is a possibility that *qFL-Chr.1* from Pima S-6 by way of R01-40-08 has no effect in these 4 genetic backgrounds, further underscoring the necessity of QTL validation in genetically diverse backgrounds to verify a QTL's portability and practical utility in cotton improvement.

WORKS CITED

Abdurakhmonov, I., R. Kohel, J. Yu, A. Pepper, A. Abdullaev, F. Kushanov, et al. 2008. Molecular diversity and association mapping of fiber quality traits in exotic G. hirsutum L. germplasm. Genomics 92: 478-487.

Azhar, F., M. Naveed and A. Ali. 2004. Correlation of Sed Yield with Fiber Characteristics of Cotton. International Journal of Agriculture and Biology 6: 656-658.

Bedez Üte, T. and H. Kadoğlu. 2014. Regressional Estimation of Cotton Sirospun Yarn Properties from Fibre Properties. Autex Research Journal 14: 161-167.

Bertini, C.H., I. Schuster, T. Sediyama, E.G.d. Barros and M.A. Moreira. 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genetics and Molecular Biology 29: 321-329.

Campbell, B., P. Chee, E. Lubbers, D. Bowman, W. Meredith, J. Johnson, et al. 2011. Genetic improvement of the Pee Dee cotton germplasm collection following seventy years of plant breeding. Crop Science 51: 955-968.

Campbell, B., S. Saha, R. Percy, J. Frelichowski, J. Jenkins, W. Park, et al. 2010. Status of the global cotton germplasm resources. Crop science 50: 1161-1179.

1993. Characterization of Upland× Pima breeding lines using molecular markers. Proc. Beltwide Cotton Conf., New Orleans, LA.

Chee, P., R. Bredhauer, C.W. Smith, A.H. Paterson, T.A. Delmonte, X. Draye, et al. 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., R. Bredhauer, C.W. Smith, A.H. Paterson, T.A. Delmonte, X. Draye, et al. 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.

1993. Improved cotton fiber properties: The textile industry's key to success in global competition. Proceedings-Beltwide Cotton Conferences (USA).

Draye, X., R. Bredhauer, C.W. Smith, A.H. Paterson, T.A. Delmonte, P. Chee, et al. 2005. Molecular dissection of interspecific variation between Gossypium hirsutum and G. barbadense

(cotton) by a backcross-self approach. II. Fiber fineness. Theoretical and applied genetics 111: 764-771.

D. Herber and D. Richter. 1994. Improved cotton fiber quality and improved spinning technology - a profitable marriage. Part I. Progress in rotor spinning and progress in the quality profile of U.S. upland cotton. Part II. The contributions of improved cotton quality and those of rotor spinning developments to higher profits in cotton production and in spinning. Beltwide Cotton Conference, Memphis, TN. National Cotton Council.

Fang, D.D., L.L. Hinze, R.G. Percy, P. Li, D. Deng and G. Thyssen. 2013. A microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton (Gossypium hirsutum L.) cultivars from major cotton-growing countries. Euphytica 191: 391-401.

Feaster, C. and E. Turcotte. 1984. Registration of Pima S-6 cotton. Crop Sci 24: 382.

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

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

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

Meredith, W. and R. Bridge. 1971. Breakup of Linkage Blocks in Cotton, Gossypium hirsutum L. Crop Science 11: 695-698.

Miller, P., J. Williams, H. Robinson and R. Comstock. 1958. Estimates of Genotypic and Environmental Variances and Covariances in Upland Cotton and Their Implications in Selection. Agronomy Journal 50: 126-131.

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

Pillay, M. and G. Myers. 1999. Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. Crop Science 39: 1881-1886.

Ramey, H., R. Lawson and S. Worley. 1977. Relationship of cotton fiber properties to yarn tenacity. Textile Research Journal 47: 685-691.

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

Rungis, D., D. Llewellyn, E. Dennis and B. Lyon. 2005. Simple sequence repeat (SSR) markers reveal low levels of polymorphism between cotton (Gossypium hirsutum L.) cultivars. Crop and Pasture Science 56: 301-307.

SASInstitute. 2014. SAS/STAT Software. SAS Institute Inc., Cary.

Scholl, R. and P. Miller. 1976. Genetic Association between Yield and Fiber Strength in Upland Cotton. Crop Science 16: 780-783.

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

Shen, X., Z. Cao, R. Singh, E.L. Lubbers, P. Xu, C.W. Smith, et al. 2011. Efficacy of qFL-chr1, a Quantitative Trait Locus for Fiber Length in Cotton (spp.). Crop science 51: 2005-2010.

Shen, X., W. Guo, Q. Lu, X. Zhu, Y. Yuan and T. Zhang. 2007. Genetic Mapping of Quantitative Triat Loci for FIber Quality and Yield Trait by RIL Approach in Upland Cotton. Euphytica 155: 371-380.

Smith, C. and G. Coyle. 1997. Association of fiber quality parameters and within-boll yield components in upland cotton. Crop science 37: 1775-1779.

Stephens, S. 1949. The cytogenetics of speciation in Gossypium. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 34: 627.

Sun, F.-D., J.-H. Zhang, S.-F. Wang, W.-K. Gong, Y.-Z. Shi, A.-Y. Liu, et al. 2012. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Molecular breeding 30: 569-582.

Tanksley, S. and J. Hewitt. 1988. Use of molecular markers in breeding for soluble solids content in tomato—a re-examination. Theoretical and Applied Genetics 75: 811-823.

Tatineni, V., R. Cantrell and D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Science 36: 186-192.

Zhang, K., J. Zhang, J. Ma, S. Tang, D. Liu, Z. Teng, et al. 2012. Genetic mapping and quantitative trait locus analysis of fiber quality traits using a three-parent composite population in upland cotton (Gossypium hirsutum L.). Molecular breeding 29: 335-348.

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

Table 4.1. Analysis of variance mean squares for agronomic traits of *qFL-Chr.1* from *G. barbadense* introgressed into 4 genetic backgrounds of Upland cotton (*G. hirsutum*) grown in Tifton and Plains, GA in 2014 and Tifton, GA in 2015.

Source	df†	LYLD	df	LPCT	df	MIC	UHML	UI	FBS	ELO	SFC
Environment	2	1890.0*	3	0.94	3	3.53*	8.69*	27.53**	5839**	45.58**	2.07*
Rep	1	1.4	1	0.59	1	0.01	0.24	2.06	523	0.18	0.11
Error A	2	95.5	3	0.61	3	0.12	0.59	0.34	57	1.08	0.22
Background	3	327.9**	3	4.12**	3	2.02**	20.73**	3.44*	5013**	2.62**	1.55**
Background*Environment	12	185.5*	18	0.50	18	0.10	1.73*	2.56**	449**	0.87**	0.38*
Error B	18	61.3	24	0.47	24	0.06	0.83	0.74	123	0.30	0.17
QTL‡	6	87.4	6	0.69*	6	0.89**	6.19**	6.35**	2139**	4.78**	1.05**
QTL*Background	6	58.4	6	0.13	6	0.96**	3.83**	0.81	1630**	1.13	0.48
QTL*Background*Environment	21	32.1	31	0.45**	31	0.10	1.49	1.19	265	0.24	0.19
Error C	506	39.9	710	0.24	704	0.10	1.31	1.04	282	0.63	0.26

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† df, degrees of freedom; LYLD, lint yield; LPCT, lint percent; MIC, micronaire; UHML, upper half mean length; UI, uniformity index; FBS, fiber bundles strength; ELO, elongation; SFC, short fiber content.

‡ QTL, 7 QTL classes, genotypes derived from crosses with R01-40-08 and a regional cultivar w/ the *qFL-Chr.1* Pima S-6 allele and those without, Pima S-6, Acala SJ4, Paymaster HS26, Deltapine 50, GA2004089.

Table 4.2. Analysis of variance for Upland cotton (*Gossypium hirsutum*) lines from 4 genetic backgrounds crossed with R01-40-08, an introgression line carrying *qFL-Chr.1*, a previously mapped fiber length QTL from *G. barbadense*.

Source	df	PM HS26	df	Acala SJ4	df	GA089	df	DP50
Environment	3	26.3*	3	37.6	3	2.1	3	1.0
Rep	1	0.7	1	0.9	1	0.3	1	1.2
Error A	3	1.0	3	5.3	3	1.7	3	1.3
QTL	5	183.7**	5	180.0**	5	6.9**	5	4.1**
QTL*Environment	15	7.3	13	17.6	13	1.7	11	2.0*
Error B	227	13.8	193	13.0	187	1.4	92	1.0

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† df, degrees of freedom; QTL, genotypes grouped based on *qFL-Chr.1* QTL allele status, also includes parents.

Table 4.3. Mean separations of UHML for Upland cotton (*Gossypium hirsutum*) lines derived from crossing 4 genetic backgrounds with R01-40-08, an introgression line carrying *qFL-Chr.1*, a previously mapped fiber length QTL from *G. barbadense*.

	n†	Acala SJ4	n	DP50	n	GA089	n	PM HS26
QTL+	68	29.0bc‡	31	28.7a	29	28.6c	16	27.4d
QTL(h)	36	28.6c	27	29.0a	86	29.2abc	27	28.5bc
QTL-	95	28.9bc	42	29.2a	78	29.0bc	188	28.1cd
Cultivar parent	8	31.2a	6	28.7a	8	30.1a	8	29.1ab
R01-40-08	4	29.5b	4	29.2a	4	29.6ab	8	29.6a

† n, number of observations; QTL+, R01-40-08 by cultivar hybrids with the Pima S-6 allele; QTL(h), R01-40-08 by cultivar hybrids with the Pima S-6 allele in heterozygous or segregating condition; QTL-, R01-40-08 by cultivar hybrids without the Pima S-6 allele.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 4.4. HVI fiber length performance of lines from 4 genetic backgrounds of Upland cotton (*Gossypium hirsutum*), crossed with an introgression line, R01-40-08, some lines harboring *qFL-Chr.1*, a QTL from *G. barbadense*.

	n†	LYLD	n	LPCT	n	MIC	UHML	UI	FBS	ELO	SFC
		kg/ha		%		unit	mm	%	kN m kg^{-1}	%	%
QTL+	104	767	145	36.3cd‡	144	4.89ab	28.7d	83.3d	292d	6.9a	7.7ab
QTL(h)	123	858	176	37.1bcd	176	4.93a	29.0cd	83.5cd	293d	6.9a	7.6abc
QTL-	294	843	408	36.2cd	403	4.81ab	28.6d	83.7cd	290d	6.8a	7.6abc
Acala SJ4	6	725	8	42.3a	8	4.06e	31.2a	85.4a	322a	5.5b	6.9e
DP50	6	973	6	37.4bcd	6	5.00a	28.7cd	84.1c	267e	6.6a	7.8a
GA089	6	992	8	39.8abc	8	4.42d	30.1b	85.7a	310b	5.0b	7.3cd
PM HS26	6	779	8	33.5d	8	4.47cd	29.1cd	85.3a	306bc	7.1a	7.0de
R01-40-08	12	755	20	36.1cd	20	4.48cd	29.5bc	83.8cd	296cd	6.8a	7.6abc

† n, number of observations; LYLD, lint yield; LPCT, lint percent; MIC, micronaire; UHML, upper half mean length; UI, uniformity index; FBS, fiber bundles strength; ELO, elongation; SFC, short fiber content; QTL+, R01-40-08 by cultivar hybrids with the Pima S-6 allele; QTL(h), R01-40-08 by cultivar hybrids with the Pima S-6 allele in heterozygous or segregating condition; QTL-, R01-40-08 by cultivar hybrids without the Pima S-6 allele.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 4.5. Agronomic performance and other HVI fiber traits of lines from 4 genetic backgrounds of Upland cotton (*Gossypium hirsutum*), crossed with an introgression line, R01-40-08, some lines harboring *qFL-Chr.1*, a QTL from *G. barbadense*.

	n†	LYLD	n	LPCT	n	MIC	FBS	ELO
		kg/ha		%		unit	kN m kg^{-1}	%
QTL+	104	767	145	36.3cd‡	144	4.89ab	292d	6.9a
QTL(h)	123	858	176	37.1bcd	176	4.93a	293d	6.9a
QTL-	294	843	408	36.2cd	403	4.81ab	290d	6.8a
Acala SJ4	6	725	8	42.3a	8	4.06e	322a	5.5b
DP50	6	973	6	37.4bcd	6	5.00a	267e	6.6a
GA089	6	992	8	39.8abc	8	4.42d	310b	5.0b
PM HS26	6	779	8	33.5d	8	4.47cd	306bc	7.1a
R01-40-08	12	755	20	36.1cd	20	4.48cd	296cd	6.8a

† n, number of observations; LYLD, lint yield; LPCT, lint percent; MIC, micronaire; UHML, upper half mean length; UI, uniformity index; FBS, fiber bundles strength; ELO, elongation; SFC, short fiber content; QTL+, R01-40-08 by cultivar hybrids with the Pima S-6 allele; QTL(h), R01-40-08 by cultivar hybrids with the Pima S-6 allele in heterozygous or segregating condition; QTL-, R01-40-08 by cultivar hybrids without the Pima S-6 allele.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

CHAPTER 5

VALIDATION AND AGRONOMIC PERFORMANCE OF A FIBER STRENGTH QTL FROM *GOSSYPIUM MUSTELINUM, QSTR-CHR.19-1*, INTROGRESSED INTO 3 UPLAND COTTON (*G. HIRSUTUM*) GENETIC BACKGROUNDS.

Nino Brown,* Edward L. Lubbers, Rippy Singh, Baohua Wang, Pawan Kumar, Don Jones, Andrew H. Paterson, Peng W. Chee.

To be submitted to *Crop Science*.

VALIDATION AND AGRONOMIC PERFORMANCE OF A FIBER STRENGTH QTL FROM *GOSSYPIUM MUSTELINUM*, *QSTR-CHR.19-1*, INTROGRESSED INTO 3 UPLAND COTTON (*G. HIRSUTUM*, L) GENETIC BACKGROUNDS.

Keywords

Cotton; fiber quality; interspecific; introgression; *Gossypium mustelinum*.

Abbreviations

AB, advanced backcross

DP50, Deltapine 50

FBS, fiber bundle strength

GA089, GA2004089

HVI, high volume instrument

LPCT, lint percent

PM HS26, Paymaster HS26

QTL, quantitative trait locus

QTL+, lines crossed with *G. mustelinum* donor, with the *G. mustelinum* allele of interest

QTL-, lines crossed with *G. mustelinum* donor, but without the *G. mustelinum* allele of interest

UHML, upper half mean length

INTRODUCTION

Because of the low levels of genetic diversity within the Upland cotton gene pool, numerous researchers have suggested the need to explore the secondary gene pool as a source of novel alleles for fiber quality traits (Bowman and Gutierrez 2003; Campbell and Myers 2015; Campbell et al. 2009; Chee et al. 2016). *G. hirsutum* forms a monophyletic group with six other allotetraploid species, which include *G. barbadense*, *G. tomentosum*, *G. darwinii*, *G mustelinum*, *G. ekmanianum* and a recently discovered species, *G. stephensii*. While all allotetraploid *Gossypium* species are cross compatible, the use of interspecific sources as a means to incorporate novel alleles into the Upland germplasm base is problematic because progeny are often plagued with abnormal segregation, reduced fertility, photoperiodic flowering, and linkage drag (Stephen 1946; Jiang et al. 2000). To overcome many of these complications, advanced backcross populations have been employed to both map QTL and incorporate these newly discovered alleles into elite genetic backgrounds (Paterson et al. 2012). Advanced backcross mapping populations have been developed from the interspecific crosses of *G. hirsutum* by *G. barbadense* (Chee, et al., 2005; Chee, et al., 2005; Draye, et al., 2005; Wang, et al., 2012; Zhang, et al., 2015), *G. tomentosum* (Zhang, et al., 2011), *G. darwinii* (Wang, et al., 2012), and *G. mustelinum* (Wang, et al., 2016). Individual breeding lines derived from these populations have fewer genes from the donor parent, allowing for a more accurate evaluation of the effect of a putative QTL when introgressed. However, it is not uncommon for QTL identified in a mapping population to have no effect outside of that population, underscoring the necessity of QTL validation in genetically diverse backgrounds to verify a QTL's portability and practical utility in cotton improvement.

Gossypium hirsutum L. was first domesticated in the Yucatan peninsula of Central America (Brubaker, et al., 1999; Brubaker and Wendel, 1994; d'Eeckenbrugge and Lacape, 2014; Wendel, et al., 1992). The cultivated form of this species, commonly known as Upland cotton, has become widely grown globally for its valuable lint fiber, which is the primary source of natural fiber for the textile industry. This widely adaptable and high yielding species has been selectively bred by humans for nearly 5000 years, and has undergone significant genetic bottlenecks during the process of domestication (Iqbal, et al., 2001), dispersal from center of diversity, and modern breeding efforts (Bertini, et al., 2006; Fang, et al., 2013; Pillay and Myers, 1999; Rungis, et al., 2005; Van Becelaere, et al., 2005). Limited genetic diversity along with modern plant breeding focusing on fiber yield, physiological and agronomic traits have likely contributed to an impedence if not stagnation of fiber quality improvement in Upland cotton. Recent market pressures within the textile industry, however, have put the onus on fiber quality improvement. Higher quality cotton fibers are needed in the textile mills to satisfy the technological advances in yarn spinning, and necessary to compete with man-made synthetic fibers, which are becoming more preferred by textile producers due to their lower prices, higher product uniformity, and more reliable supply chain.

The standard method of measuring cotton fiber quality is conducted by using the High Volume Instrument (HVI). The HVI machine currently consists of determinations of fiber length, length uniformity, strength, micronaire, and color. Among these fiber properties, the Upper half mean fiber length (UHML) and fiber bundle strength (FBS) are the two most important fiber quality traits because they are most closely associated with the quality of the textiles produced through the spinning and manufacturing process. Traditionally, UHML is considered the most important character contributing to yarn quality and tenacity (Brown, 1938; Perkins, et al., 1984).

However, as spinning technology transitioned from the traditional ring spinning to the modern rotor and air jet spinning, fiber strength has become equally as important as UHML. Modern cotton mills are now seeking raw cotton with strong fibers because they can better withstand the rigors of mechanical processing such as during ginning, carding, spinning, and weaving processes, making FBS crucial to the high-speed textile manufacturing processing (Faerber, 1995; Patil and Singh, 1995; Perkins, et al., 1984). High quality yarns desirable in high-end textile products such as high thread count bed sheets and fine dress shirts are derived from fibers with both long fibers and high FBS (Deussen, 1993; May and Taylor, 1998; Meredith, et al., 1991).

A prior study using the advanced backcross (AB-QTL) approach has identified and introgressed a fiber strength QTL designated *qSTR-Chr.19-1-1* on chromosome 19 from *G. mustelinum*. This QTL was detected in both BC3F2 and BC3F2:3 generations, explained 12% of the phenotypic variation, and the *G. mustelinum* allele improved fiber strength by an average of 0.9 cN/tex (Wang et al. 2017). The current study was conducted to validate the efficacy and association of *qSTR-Chr.19-1* with fiber strength in three genetic backgrounds adapted to the four major US cotton production regions.

MATERIALS AND METHODS

The *G. mustelinum* introgression line used to develop the validation population was generated initially from an AB-QTL population consisting of the recurrent parent, a *G. hirsutum* germplasm line, PD94042 (PI 603219), and the donor parent, *G. mustelinum* (AD4-8, PS 85). The resulting F_1 generation was then backcrossed to the *G. hirsutum* parent, PD94042, for three cycles of backcrossing. The BC_3F_1s were allowed to self to generate the BC_3F_2, which was then individually harvested to form the basis of AB-QTL mapping experiments (Wang, et al., 2016).

Among the QTL mapped in that experiment was *qUHM-Chr.19* (Wang et al. 2017). From the BC_3F_2 populations, individual BC_3F_3 plants were harvested. These families were then self-pollinated and advanced to the $BC_3F_{3:6}$ generation in 2011. At this point they were individually harvested again, and progeny rows planted from individual plants. The resulting $BC_3F_{6:7}$ lines were genotyped and confirmed as carrying the QTL of interest.

Line B15-002 which carries the *qSTR-Chr.19-1* QTL was crossed to three *G. hirsutum* cotton cultivars or elite germplasm lines considered to be typical ideotypes of the growing regions they represent; Acala SJ4 (PI 529538), an Acala type cotton cultivar that was grown widely in the desert southwestern US; Paymaster HS26 (PM HS26) (PVP 8600087, PI 606814), a Texas High Plains variety that was grown in the western US; Deltapine 50 (DP50) (PVP 8400154, PI 529566), which was a widely grown, Delta-type variety for the Mid-South US around the Mississippi Delta growing region. Because this QTL was being donated from a line with significant amounts of PD94042, a line adapted for culture within the Southeastern US, we did not test it in an additional Southeastern background. The initial crosses were made in 2012, followed by individual harvest of the F_2 in 2014, and replicated trials of the $F_{2:3}$ in 2015.

Individual plants and the replicated trials in 2014 and 2015 respectively, were grown at the University of Georgia, Gibbs Research Farm in Tifton, GA. Soils encountered there include Tifton loamy sands, a fine-loamy, kaolinitic, thermic Plinthic Kandiudult; and Clarendon loamy sands, a fine-loamy, siliceous, semiactive, thermic Plinthaquic Paledults. In 2014, plants from populations developed from each of the 4 genetic backgrounds were individually genotyped using the SSR marker BNL3977, which previously was determined to be tightly linked to the *qSTR-Chr.19-1* locus. DNA was extracted from young leaf tissues in the field following the protocol discussed in Wang et al. (2017). PCR amplification conditions carried out in a modified

form as described in Chee et al. (2004) on a T100 Thermal Cycler (Bio-Rad Laboratories, 2011). The 10µl PCR reactions contained approximately 10ng of DNA, 0.5µM of forward and reverse SSR primers, 100µdNTPs, 1.5 mM $MgCl_2$, 3U of DNA taq polymerase, and 1X reaction buffer containing 100mM Tris-HCl, 500 mM KCl, at pH 8.3. The PCR thermal cycling conditions were 94ºC for 3 min, then 34 cycles of 94ºC for 45 sec, 55ºC for 45 sec, 72ºC for 1 min, then incubation at 72ºC for 7 min, prior to cooling down for storage at 12ºC. Viewing of the PCR products were accomplished using 10% polyacrylamide gel electrophoresis and staining with silver nitrate as previously described (Zhang, et al., 2002).

The genotyped plants were given unique identifiers and individually harvested by hand. Individual plant samples were ginned on a laboratory table-top 20-saw gin and the fiber samples sent to the Fiber Testing Laboratory at Cotton Inc., in Cary, NC for HVI testing.

In 2015, single-row plots were planted in a twice-replicated, randomized complete block design. Plots were 6.1m by 1m, planted at a seeding rate of approximately 60 seeds per plot. A twenty-five boll hand-harvested sample was collected from the mid-fruiting zone of the plants within each plot. The 25-boll samples were ginned on a table-top 20-saw gin to calculate lint percent (LPCT), and fibers were analyzed using high volume instrument (HVI) at the Fiber Testing Laboratory of Cotton Inc. in Cary, NC. Statistical analyses of the data were conducted using the PROC GLM function in the SAS/STAT software, Version 9.4 of the SAS System for Windows (SASInstitute, 2014).

RESULTS AND DISCUSSION

Analysis of variance revealed significant differences in replication means for FBS in the trial (Table 5.1). This result is not unexpected as the first replication was an individual plant sample from 2014, and the 2nd and 3rd replications come from 25-boll samples in 2015. The main

effect of the QTL classes, as well as the interaction of the QTL with the genetic background in which it was deployed (QTL*Background) was significant. Mean separations were conducted using Waller-Duncan's k-ratio=100, which approximates an α=0.05, but is more conservative than Fisher's LSD (Table 5.2). Analyzed across the genetic backgrounds of the cultivars used in this study, the mean FBS was significantly greater for the genotypes carrying the QTL region for qSTR-Chr.19-1, than the mean FBS for the genotypes that were not carrying the QTL region, 295kN m kg-1 versus 293kN m kg-1, respectively. Interestingly, the heterozygous genotypes had the lowest FBS with 289kn m kg-1. The donor parent, B15-002, exhibited FBS intermediate to that of the cultivar parents with 281 kN m kg-1. The cultivar parents Acala SJ4, DP50, and PM HS26 had FBS values of 289, 266, and 281 kN m kg-1, respectively.

Because the analysis of variance revealed a significant QTL*Background effect, it was necessary to separate the backgrounds for further analysis (Table 5.3). For all three genetic backgrounds used to test the effect of qSTR-Chr.19-1, the effect of replications on FBS means was significant. Also significant was the effect of the QTL in each of the genetic backgrounds, DP50, PM HS26 and Acala SJ4.

In the DP50 background population, genotypes with and without the QTL region had similar FBS (Table 5.4) while the heterozygous genotypes exhibited significantly lower FBS. Although statistically the means were not different, numerically the genotypes with the QTL had a higher mean with 299 kN m kg-1 versus 297 kN m kg-1 for the lines without the QTL region. The donor parent, for comparison exhibited FBS of In the PM HS26 background population , genotypes that were heterozygous for the QTL region again had the lowest FBS with 282 kN m kg-1. However, genotypes from the this background without the QTL region had FBS that averaged 288 kN m kg-1 , and genotypes with the QTL region averaged higher FBS with 294 kN

m kg-1, a statistically significant and biologically substantial difference. A similar positive effect was observed in the Acala SJ4 population. Genotyopes without the QTL region had a significantly lower 292kN m kg-1 FBS, whereas the heterozygtes and the genotypes with the QTL region exhibited significantly higher FBS of 299 kN m kg-1 and 300 kN m kg-1, respectively.

Validation and testing of mapped QTL are an integral step in deploying a putative QTL for use by breeders and other crop geneticists. In order for the full benefit of QTL mapping to be realized for the purposes of crop improvement, identified QTL must first be tested to ensure the effect is useful in genetic backgrounds outside of those used in mapping. To do this, diverse genetic backgrounds should be used as in this study, to determine if the QTL can be deployed across a diversity of breeding programs for many growing regions, or if it would be best utilized in specific genetic backgrounds or populations suited for specific growing regions.

The use of B15-002, harboring qSTR-Chr.19-1, represents an interesting case in that the line does not have FBS that stands apart on first glance. Indeed, in this experiment it was within the range of the cultivar parents; B15-002 having FBS of 281 kN m kg-1, versus Acala SJ4 with 289, DP50 with 266, and PM HS26 with a very similar 281 kN m kg-1. This lack of obvious phenotypic novelty might lead a breeder to discard a line such as this, however, this novel allele from G. mustelinum shows a positive effect on the adapted Upland germplasm backgrounds used in this study.

It is worth noting that from 1980, FBS of the US cotton crop increased from 235 kN m kg-1 to 294 kN m kg-1 in 2012, a 25.1% increase over more than 30 years, but only a 0.8% increase annually. The effect of qSTR-Chr.19-1 by comparison, across genetic backgrounds increased FBS by 2 kN m kg-1, or 0.7%. In the DP50 background, the realized FBS increase

was also 0.7%. The PM HS26 background population, however, exhibited an increase of 6 kN m kg-1, or 2.1%. Also registering a big improvement was the Acala SJ4 population which saw increases of 8 kN m kg-1, or 2.7%; more than 3 years of the accumulated genetic gain seen in the entire US cotton crop.

qSTR-Chr.19-1 therefore represents a QTL that could positively influence breeding efforts for FBS across the major growing regions and potentially across multiple genetic backgrounds favored in those regions. This QTL has the potential to make a significant impact on fiber quality across much of the cotton industry. An accumulation of alleles of seemingly small effect QTL such as qSTR-Chr.19-1, would benefit commercial Upland cotton germplasm and have significant and positive impacts on the cotton industry.

WORKS CITED

Bertini, C.H., I. Schuster, T. Sediyama, E.G.d. Barros and M.A. Moreira. 2006. Characterization and genetic diversity analysis of cotton cultivars using microsatellites. Genetics and Molecular Biology 29: 321-329.

Bio-Rad Laboratories, I. 2011. T100 Thermal CyclerBio-Rad Laboratories, Inc. - Life Science Group, Hercules, California.

Brubaker, C.L., F. Bourland and J.F. Wendel. 1999. The origin and domestication of cotton. Cotton: Origin, history, technology, and production. John Wiley & Sons, New York: 3-31.

Brubaker, C.L. and J.F. Wendel. 1994. Reevaluating the origin of domesticated cotton (Gossypium hirsutum; Malvaceae) using nuclear restriction fragment length polymorphisms (RFLPs). American journal of botany: 1309-1326.

Chee, P., X. Draye, C. Jiang, L. Decanini, T. Delmonte, R. Bredhauer, et al. 2005a. Molecular dissection of interspecific variation between Gossypium hirsutum and Gossypium barbadense (cotton) by a backcross-self approach: I. Fiber elongation. TAG Theoretical and Applied Genetics 111: 757-763.

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

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

d'Eeckenbrugge, G.C. and J.-M. Lacape. 2014. Distribution and differentiation of wild, feral, and cultivated populations of perennial Upland cotton (Gossypium hirsutum L.) in mesoamerica and the Caribbean.

1993. Improved cotton fiber properties: The textile industry's key to success in global competition. Proceedings-Beltwide Cotton Conferences (USA).

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

1995. Future demands on cotton fiber quality in the textile industry. Beltwide Cotton Conferences (USA).

Fang, D.D., L.L. Hinze, R.G. Percy, P. Li, D. Deng and G. Thyssen. 2013. A microsatellite-based genome-wide analysis of genetic diversity and linkage disequilibrium in Upland cotton (Gossypium hirsutum L.) cultivars from major cotton-growing countries. Euphytica 191: 391-401.

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

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

May, O.L. and R.A. Taylor. 1998. Breeding cottons with higher yarn tenacity. Textile research journal 68: 302-307.

Meredith, W.R., T. Culp, K. Robert, G. Ruppenicker, W. Anthony and J. Williford. 1991. Determining future cotton variety fiber quality objectives. Textile research journal 61: 715-720.

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

G. Constable and N. Forrester. 1995. Development of Medium Staple High-Strength Cotton Suitable for Rotor Spinning Systems. World Cotton Conference, Brisbane, Australia. February 14-17, 1994. CSIRO, Melbourne, Australia.

Perkins, H., D. Ethridge and C. Bragg. 1984. Fiber. In: R. Kohel, editor Cotton. ASA, Madison, WI. p. 437-509.

Pillay, M. and G. Myers. 1999. Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. Crop Science 39: 1881-1886.

Rungis, D., D. Llewellyn, E. Dennis and B. Lyon. 2005. Simple sequence repeat (SSR) markers reveal low levels of polymorphism between cotton (Gossypium hirsutum L.) cultivars. Crop and Pasture Science 56: 301-307.

SASInstitute. 2014. SAS/STAT Software. SAS Institute Inc., Cary.

Tanksley, S. and J. Nelson. 1996. Advanced backcross QTL analysis: a method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. Theoretical and Applied Genetics 92: 191-203.

Van Becelaere, G., E.L. Lubbers, A.H. Paterson and P.W. Chee. 2005. Pedigree-vs. DNA marker-based genetic similarity estimates in cotton. Crop science 45: 2281-2287.

Wang, B., X. Draye, Z. Zhang, Z. Zhuang, O.L. May, A.H. Paterson, et al. 2016. Advanced Backcross Quantitative Trait Locus Analysis of Fiber Elongation in a Cross between Gossypium hirsutum and *G. mustelinum*. Crop Science 56: 1760-1768.

Wang, B., Y. Nie, Z. Lin, X. Zhang, J. Liu and J. Bai. 2012a. Molecular diversity, genomic constitution, and QTL mapping of fiber quality by mapped SSRs in introgression lines derived from Gossypium hirsutum× *G. darwinii* Watt. Theoretical and applied genetics 125: 1263-1274.

Wang, P., Y. Zhu, X. Song, Z. Cao, Y. Ding, B. Liu, et al. 2012b. Inheritance of long staple fiber quality traits of Gossypium barbadense in G. hirsutum background using CSILs. Theoretical and Applied Genetics 124: 1415-1428.

Wendel, J., C. Brubaker and A. Percival. 1992. Genetic diversity in Gossypium hirsutum and the origin of upland cotton. American Journal of Botany 79: 1291-1310.

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

Zhang, J., J. Yu, W. Pei, X. Li, J. Said, M. Song, et al. 2015. Genetic analysis of Verticillium wilt resistance in a backcross inbred line population and a meta-analysis of quantitative trait loci for disease resistance in cotton. BMC genomics 16: 1.

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

Table 5.1. Analysis of variance mean squares for FBS comparison of lines from 3 Upland cotton (*Gossypium hirsutum*) genetic backgrounds crossed with a *G. mustelinum* introgression line harboring *qSTR-Chr.19-1*.

Source	df	MS
Background	2	10226
Rep	2	104967*
Error A	4	7401
QTL	2	3918*
QTL*Background	4	2626*
Error B	1463	382

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† df, degrees of freedom; MS, mean squares; QTL, allelic status of *qSTR-Chr.19-1*.

Table 5.2. Mean separations for FBS comparison of lines from 3 Upland cotton (*Gossypium hirsutum*) genetic backgrounds crossed with a *G. mustelinum* introgression line harboring *qSTR-Chr.19-1*.

	n	FBS
QTL-	477	293b
Hetz.	499	289c
QTL+	502	295a

† n, number of observations; FBS, fiber bundle strength; QTL-, progeny without the *G. mustelinum* allele at *qSTR-Chr.19-1* ; QTL(h), progeny with the *G. mustelinum* allele at *qSTR-Chr.19-1* in heterozygous or segregating condition; QTL+, progeny with the *G. mustelinum* allele at *qSTR-Chr.19-1*.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

Table 5.3. Analysis of variance mean squares for FBS comparison of lines from 3 Upland cotton (*Gossypium hirsutum*) genetic backgrounds crossed with a *G. mustelinum* introgression line harboring *qSTR-Chr.19-1*, analyzed separately by genetic background.

Source	df†	DP50	df	PM HS26	df	Acala SJ4
QTL	2	1677**	2	10625**	2	1932**
rep	2	24409**	2	132310**	2	12691**
Error	374	297	780	462	309	283

* Significant at the 0.05 probability level.

** Significant at the 0.01 probability level.

† df, degrees of freedom; DP50, Deltapine 50; PM HS26, Paymaster HS26; MS, mean squares; QTL, allelic status of *qSTR-Chr.19-1*.

Table 5.4. Mean separations of FBS for lines from 3 Upland cotton (*Gossypium hirsutum*) genetic backgrounds crossed with a *G. mustelinum* introgression line harboring *qSTR-Chr.19-1*, separated by genetic background.

	n†	DP50	n	PM HS26	n	Acala SJ4
QTL-	214	297a‡	179	288b	84	292b
Hetz.	71	290b	236	282c	192	299a
QTL+	94	299a	370	294a	38	300a

† n, number of observations; FBS, fiber bundle strength; QTL-, progeny without the *G. mustelinum* allele at *qSTR-Chr.19-1* ; QTL(h), progeny with the *G. mustelinum* allele at *qSTR-Chr.19-1* in heterozygous or segregating condition; QTL+, progeny with the *G. mustelinum* allele at *qSTR-Chr.19-1*.

‡ Values within columns followed by the same letter are not different at k=100 according to Waller-Duncan LSD.

CHAPTER 6

CONCLUSIONS

QTL mapping and DNA marker technologies have allowed researchers to track loci controlling or influencing complex traits such as cotton fiber length. This ability makes it desirable to break down some of these traits into simply inherited Mendelian factors. However, in many cases, it just is not that simple.

As described in Chapter 2, when the *qUHM-Chr.25* QTL was tested in 4 cultivars representative of the genetic backgrounds adapted to the four primary macro-environments of the cotton-growing United States, the effect of this QTL was small, but significant in 3 out of the 4 cultivar backgrounds tested in a Genotype x Environment experiment. However, when measured in a near-isogenic state comparing NILs with and without the QTL allele from SL883, the effect of *qUHM-Chr.25* on UHML was significant in 2 of the 4 cultivar backgrounds tested. The effect of the QTL, when averaged across the backgrounds tested in this trial accounted for an additional 0.9mm in the QTL+ lines, though the difference was not significantly different than the QTL- lines. When separated by background, the effect of the QTL was not found to be significantly greater than the SL883 background alone when deployed within the Acala SJ4 or PM HS26 background, however it was significant for the DP50 and GA089 backgrounds, with an UHML increase of 1.4mm (4.5%) and 1mm (3.1%), respectively. It is worth mentioning that the PM HS26 QTL+ lines did have a numerically higher average UHML of 0.6mm better than the QTL- lines, though the difference was not statistically significant. The utilization of foreign alleles from closely related species by way of introgression lines, such as the SL883 *qUHM-Chr.25*

allele, have the potential to make significant impacts on the cotton industry in relatively few years if deployed in commercial germplasm.

Testing of recombinant lines suggest that the QTL region likely falls within the region flanked by UGT2501 and UGT2509. This region had a highly significant effect on lines from the PMHS26 as well as the GA089 background. Because the testing to narrow down the region of interest was conducted on highly inbred recombinant sister lines, the data derived from the fiber samples was highly uniform, and likely made the effect of the QTL that much more apparent.

qUHM-Chr.25 represents a *G. barbadense*-sourced QTL for the improvement of Upland cotton germplasm. The effect of this QTL is significant, and along with the tightly linked marker sequence included here, should prove to be beneficial in improving Upland cotton germplasm. The portability of this QTL across diverse genetic backgrounds should make it valuable to breeders interested in improving fiber quality, a driving theme in cotton breeding today.

The germplasm line, R01-40-08, being submitted for public release here in Chapter 3, represents a milestone in Upland cotton germplasm improvement in that as far as we know, it is the first publicly released cotton germplasm line carrying a fiber length QTL whose effect has been validated. R01-40-08 yields similarly to the recurrent, *G. hirsutum* parent, TAMCOT 2111, and has improved UHML from the introgressed QTL, *qFL-Chr.1* from *G. barbadense* donor, PIMA S-6. This introgression line is unique in that it is a predominantly *G. hirsutum* genetic background, yields well, and has introgression of interspecific origin. R01-40-08 will serve well the cotton industry in improving the fiber and spinning quality of American Upland cotton, and with the available tightly linked SSR markers published previously by this laboratory, this

germplasm line will be an important addition to the Upland cotton breeding lexicon. The use of this introgression and the associated markers will help breeders to improve fiber length.

The limited effect of *qFL-Chr.1* in the 4 genetic backgrounds tested later in the experiments described in Chapter 4 suggests that the effect of the QTL is either too small to be measured in this experiment, was confounded by a high degree of segregation still occurring in the lines, or that the QTL has no effect when deployed within these backgrounds. The first year of this experiment consists of individual plant fiber data, which is inherently not as accurate as boll samples harvested from an entire row. Because of the inclusion of all bolls in an individual plant sample, there is more variability in the data, whereas a 25-boll sample harvested from the mid-fruiting zone and picked form an entire plot of a given line, is a much more uniform sample, and therefore better for detecting small differences between genotypes. The lines tested were single-cross hybrids, and as such, likely had a significant influence from background segregation occurring between R01-40-08 genome, which is essentially TAMCOT 2111 plus the qFL-Chr.1 QTL from Pima S-6, and the cultivar parents. The segregation of these divergent genetic backgrounds could very easily have overshadowed the effect of this QTL, which has been shown to have a modest effect (Chapter 3). Lastly, there is a possibility that qFL-Chr.1 from Pima S-6 by way of R01-40-08 has no effect in these 4 genetic backgrounds. It is not uncommon for QTL identified in a mapping population to have no effect outside of that population (REF), further underscoring the necessity of QTL validation in genetically diverse backgrounds to verify a QTL's portability and practical utility in cotton improvement.

Validation and testing of mapped QTL is an integral step in developing a putative QTL for use by breeders and other crop geneticists. In order for the full benefit of QTL mapping to be realized for the purposes of crop improvement, identified QTL must first be verified and tested to

ensure the effect is useful in genetic backgrounds outside of the ones used in mapping. To do this, diverse genetic backgrounds should be used as in this study, to determine if the QTL can be deployed across a diversity of breeding programs for many growing regions, or if it would be best utilized in specific genetic backgrounds or populations suited for specific growing regions.

In the study described in Chapter 5, we determined that the effect of a FBS QTL from *G. mustelinum*, *qSTR-Chr.19-1*, proved to have an effect in the 3 genetic backgrounds used in the study, although the effect was not statistically significant in one of those backgrounds. *qSTR-Chr.19-1* therefore represents a QTL that could positively influence breeding efforts for FBS across the major growing regions and potentially across multiple genetic backgrounds favored in those regions. This QTL has the potential to make a significant impact on fiber quality across much of the cotton industry.

Finally, it is evident from these results that introgression of positive alleles from interspecific sources is possible, valuable, and not entirely impractical. There exist some QTL with limited effect outside of the mapping population in which they were identified, but should not discourage breeders and geneticists to keep identifying positive novel alleles. Further utilization of closely related species has the potential to make significant, positive impacts on the cotton industry.