EVALUATION OF CONVENTIONAL AND MARKER-ASSISTED BREEDING METHODS FOR THE IMPROVEMENT OF FIBER QUALITY IN *GOSSYPIUM* SPP.

A Dissertation

by

KARI LYNNE HUGIE

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	C. Wayne Smith
Co-Chair of Committee,	Hongbin Zhang
Committee Members,	Lori Hinze
	Steve Hague
	Mark Hood
Head of Department,	David Baltensperger

December 2015

Major Subject: Plant Breeding

ABSTRACT

Excess supply coupled with domestic reliance on export markets dictates that the U.S. cotton industry compete globally in terms of both price and fiber quality. Two cycles of divergent selection for fiber upper-half mean length (UHML) and bundle strength (Str) were conducted within five genetically diverse populations. Realized heritability estimates for UHML and Str were calculated for each cycle, and correlated responses among fiber properties and lint percent (LP) were measured as they responded to selection for UHML and Str. The results suggest that early generation selection for UHML and Str was an effective strategy for the genetic improvement of fiber quality within four of the five populations at College Station, TX. There were consistent negative correlations between fiber properties and LP. However, several strains with simultaneously improved fiber quality and LP were identified within each population, providing evidence of repulsion phase linkage.

Marker-assisted selection (MAS) may help mitigate some of the current challenges regarding the genetic improvement of fiber quality, such as low genetic diversity and the negative association between fiber quality and lint yield. A multitude of quantitative trait loci (QTL) for UHML and Str have been identified in the literature, but the use of MAS for the improvement of fiber quality is still rare in public cotton breeding programs. Validation studies are necessary to develop portable genetic markers and to identify QTL with stable effects on trait expression across environments and genetic backgrounds. The effects of previously reported microsatellite markers (SSRs)

linked to QTL for UHML and Str were evaluated in three genetic backgrounds, and the efficiency of MAS for fiber quality utilizing SSRs linked to stable QTL for UHML and Str was investigated. Using the results of 31 published QTL mapping studies, six SSRs associated with stable QTL for UHML and six SSRs associated with stable QTL for Str were identified. In all but one case, the genetic gain achieved through marker-based selection of individual plants having four-to-six beneficial alleles for UHML or Str was similar to that achieved by phenotypic selection of the top 20%.

DEDICATION

In dedication to my family, William, Nancy, and Josie Hugie, without whose inspiration and support this work would not have been possible.

ACKNOWLEDGEMENTS

I would like to acknowledge my graduate advisory committee, Dr. C. Wayne Smith, Dr. Hongbin Zhang, Dr. Steve Hague, and Dr. Lori Hinze, for their important contributions in the preparation of this dissertation and in the supporting research. This work would not have been possible without their patience, guidance, and support, for which I am sincerely grateful.

I am indebted to Dawn Deno and Richard Hermes, whom I have depended on for their knowledge and support throughout my research project. This work would not have been possible without their help with all aspects of planting, field maintenance, harvest, and post-harvest processing. I would also like to thank Chantel Scheuring, Yunhua Liu, and Yang Zhang for their patience and help with laboratory work. I owe a special thanks to all of the graduate students and undergraduate student workers at the TAMU Cotton Improvement Lab for their invaluable contributions to this research.

I would like to extend my sincere gratitude to Dr. David Fang and Dr. Ping Li with the USDA-ARS in New Orleans, LA for their help with genotyping. This research would not have been possible without their financial support and hard work. Lastly, I would like to express thanks to Dr. Don Jones and Cotton Incorporated for the fellowship and financial support of this research and my graduate education.

NOMENCLATURE

AFIS	Advanced fiber information system
Elon	Elongation
ELS	Extra-long staple
FBRI	Fiber and Biopolymer Research Institute
HVI	High volume instrument
IP	Individual plant
LD	Linkage disequilibrium
LP	Lint percent
MAS	Marker assisted selection
Mic	Micronaire
PR	Progeny row
QTL	Quantitative trait loci
RFLP	Restriction fragment length polymorphism
RIL	Recombinant inbred line
SSR	Simple sequence repeat/ Microsatellite
Str	Bundle strength
UHML	Upper-half mean length
UI	Length uniformity index

TABLE OF CONTENTS

Page

ABSTRACT ii

DEDICATION iv

ACKNOWLEDGEMENTS v

NOMENCLATURE vi

TABLE OF CONTENTS vii

LIST OF FIGURES ix

LIST OF TABLES x

1. INTRODUCTION 1

2. DIVERGENT SELECTION IN *GOSSYPIUM* SPP. FOR FIBER LENGTH AND BUNDLE STRENGTH 4

    2.1. Literature Review 4
    2.2. Objectives 10
    2.3. Materials and Methods 10
    2.4. Results and Discussion 15

3. IDENTIFICATION OF MICROSATELLITE MARKERS FOR STABLE FIBER LENGTH AND BUNDLE STRENGTH QTL IN *GOSSYPIUM* SPP. 35

    3.1. Literature Review 35
    3.2. Objectives 41
    3.3. Materials and Methods 41
    3.4. Results and Discussion 47

4. CONCLUSIONS 69

REFERENCES 72

Page

APPENDIX 85

LIST OF FIGURES

Page

Figure 2.1. Observed versus predicted response to selection for high and low UHML and College Station, TX 19

Figure 2.2. Distribution of UHML within five $F_{2:5}$ populations derived by divergent selection for UHML when evaluated at College Station, TX in 2014 22

Figure 2.3. Observed versus predicted response to selection for high and low Str at College Station, TX 25

Figure 2.4. Distribution of Str within five $F_{2:5}$ populations derived by divergent selection for Str when evaluated at College Station, TX in 2014 28

Figure 2.5. Scatterplot of fiber UHML, Str, and LP measured on $F_{2:5}$ strains derived from HS624 x ELS33 grown at College Station, TX in 2014 32

Figure 2.6. Scatterplot of fiber UHML, Str, and LP measured on $F_{2:5}$ strains derived from TAM22 x SID84 grown at College Station, TX in 2014 32

Figure 2.7. Scatterplot of fiber UHML, Elon, and LP measured on $F_{2:5}$ strains derived from TAM22 x ELS33 grown at College Station, TX in 2014 34

Figure 3.1. Frequency distribution of UHML in three F_3 populations grown at College Station, TX in 2013 49

Figure 3.2. Frequency distribution of Str in three F_3 populations grown at College Station, TX in 2013 58

LIST OF TABLES

Page

Table 2.1. Mean performance of parental lines grown at College Station, TX in 2014 15

Table 2.2. Fiber UHML of selected plants, their progeny, and realized heritability estimates for cycles one (F_2 to $F_{2:3}$) and two ($F_{2:3}$ to $F_{2:4}$) of divergent selection within five populations evaluated at College Station, TX in 2012 and 2013 16

Table 2.3. ANOVA and mean UHML of selected groups (i.e., high and low UHML) composed of $F_{2:5}$ strains derived from divergent selection and evaluated at College Station, TX in 2014 20

Table 2.4. Fiber Str of selected plants, their progeny, and realized heritability estimates for cycles one (F_2 to $F_{2:3}$) and two ($F_{2:3}$ and $F_{2:4}$) of divergent selection within five populations evaluated at College Station, TX in 2012 and 2013 23

Table 2.5. ANOVA and mean Str of selected groups (i.e., high and low Str) composed of $F_{2:5}$ strains derived from divergent selection and evaluated at College Station, TX in 2014 26

Table 2.6. Phenotypic correlations among HVI measured fiber properties and lint percent of $F_{2:5}$ strains grown at College Station, TX in 2014 29

Table 3.1. Publications from which the microsatellite markers (SSRs) used to identify stable QTL for fiber quality were selected 44

Table 3.2. Population-specific models derived from forward stepwise linear regression of fiber UHML of F_3 plants grown at College Station, TX in 2013 using a common set of selected SSR alleles 49

Table 3.3. Regression of fiber UHML of F_3 individual plants and the resulting $F_{3:4}$ progeny rows when grown at College Station, TX based on SSR alleles identified through single-marker analysis 52

Table 3.4. ANOVA of UHML based on the number of beneficial SSR alleles in $F_{3:4}$ progeny rows grown at College Station, TX in 2014 54

Page

Table 3.5. ANOVA of UHML among $F_{3:4}$ progeny rows based on selection method (i.e., marker-based versus phenotypic) applied to the F_3 IPs grown at College Station, TX in 2013 and 2014 56

Table 3.6. UHML based on marker-based and phenotypic selection among $F_{3:4}$ progeny rows within three cotton populations grown at College Station, TX in 2014 57

Table 3.7. Population-specific models derived from forward stepwise linear regression of fiber Str of F_3 plants grown at College Station, TX in 2013 using a common set of selected SSR alleles 58

Table 3.8. Regression of fiber Str of F_3 individual plants and the resulting $F_{3:4}$ progeny rows when grown at College Station, TX based on SSR alleles identified through single-marker analysis 60

Table 3.9. ANOVA of Str based on the number of beneficial SSR alleles in the $F_{3:4}$ progeny rows grown at College Station, TX in 2014 62

Table 3.10. ANOVA of Str among $F_{3:4}$ progeny rows based on selection method (i.e., marker-based versus phenotypic) applied to the F_3 IPs grown at College Station, TX in 2013 and 2014 64

Table 3.11. Str based on marker-based and phenotypic selection of $F_{3:4}$ progeny rows within three cotton populations grown at College Station, TX in 2014 64

Table 3.12. Correlation between fiber UHML and Str among F_3 progeny plants and $F_{3:4}$ progeny rows grown at College Station, TX in 2013 and 2014, respectively 65

Table 3.13. UHML and Str based on the presence or absence of $BNL1604_{98}$ in the $F_{3:4}$ progeny rows grown at College Station, TX in 2014 67

Table A.1. P-values derived from single-marker analysis (i.e., ANOVA) of fiber UHML and Str for 229 SSR alleles among three experimental F_3 populations grown at College Station, TX in 2013 85

1. INTRODUCTION

Cotton (*Gossypium* spp.) is the most widely planted row crop in Texas and provides approximately $2 billion in gross value of production to the state (USDA, 2015a). For the fifth consecutive year, world cotton production will exceed consumption (Meyer et al., 2015), and excess supply coupled with U.S. reliance on export markets dictates that U.S. cotton production competes globally in terms of price and quality (USDA, 2015b). Currently, the foreign textile market is characterized by a high demand for cotton fibers suitable for the manufacture of finer yarns produced on ring-spinning systems (Faulkner et al., 2012). Rotor (i.e., open-end) and air-jet spinning systems also occupy significant market share in the global textile industry.

Yarn quality parameters are expensive to measure directly, and a substantial amount of lint is required for yarn production. Therefore, plant breeders rely on fiber parameters as indicators of yarn quality. The most important fiber quality parameters affecting yarn quality are fiber length and length uniformity for ring-spun yarn, fiber length and fineness for air-jet-spun yarn, and fiber bundle strength on rotor-spun yarn due to higher processing speeds (Bhortakke et al., 1997; Joy et al., 2010; Smith and Zhu, 1999). Demand for longer, stronger, and more uniform fibers has led to increased emphasis on the genetic improvement of cotton fiber quality (Bourland and Jones, 2012; Cantrell et al., 2000; Meredith and Nokes, 2011; Smith et al., 2009).

There are four cultivated species within the *Gossypium* L. genus, two Old World diploid species, *G. arboreum* L. and *G. herbaceum* L. ($2n = 26$), and two New World

allotetraploid species, *G. hirsutum* L. and *G. barbadense* L. (2*n* = 52). *Gossypium hirsutum*, also known as upland cotton, cultivars are characterized by high yields and broad adaptation and account for the majority (≥ 90%) of global production. *Gossypium barbadense*, also known as Sea Island, Pima, and Egyptian cotton, which accounts for most of the remaining 10% of global production, is valued for superior fiber quality but is characterized by inferior yield and adaptation. Upland cotton and *G. barbadense* readily hybridize, but attempts to introgress fiber quality traits from *G. barbadense* into upland cotton have been largely unsuccessful due to skewed chromatin transmission and the elimination of donor alleles (Jiang et al., 2000; Stephens, 1949). Therefore, the majority of cotton breeding programs in the U.S. are primarily focused on genetic improvement within the upland cotton gene pool.

Genetic improvement of fiber quality traits, in regard to public cotton breeding programs, has been attributed primarily to the accumulation of beneficial additive genetic effects through hybridization between elite upland cotton genotypes, followed by inbreeding and phenotypic selection (Chee and Campbell, 2009). Yet, the improvement of fiber quality traits through conventional breeding within upland cotton germplasm has been hindered by the negative relationship between lint yield and fiber quality traits (Al-Jibouri et al., 1958; Hinze et al., 2011; McCall et al., 1986; Meredith, 1984; Miller et al., 1958; Miller and Rawlings, 1967; Smith and Coyle, 1997; Ulloa, 2006) and low genetic diversity (Fang et al. 2013; Hinze et al., 2012; Lacape et al., 2007; Van Esbroeck and Bowman, 1998).

Molecular genetic studies have provided additional insights into the genetic basis of quantitative traits leading to the development of novel molecular breeding methods. These methods, such as DNA marker-assisted selection (MAS), may help mitigate some of the challenges of genetic improvement of fiber quality traits. Knowledge of specific quantitative trait loci (QTL) controlling the phenotypic expression of fiber length and strength may be utilized within breeding programs to more efficiently select for the simultaneous improvement of yield and fiber quality. Molecular genetics approaches also may be utilized to identify novel genetic diversity associated with fiber quality traits.

There are numerous publications identifying and characterizing such QTL for fiber length and strength, but the use of molecular breeding methods, such as MAS, for fiber quality traits is rare in public cotton breeding programs. One major challenge in the use of MAS for fiber quality traits is inconsistency regarding the genomic location and effect of individual QTL. Additional research is needed to develop portable genetic markers, tightly linked to QTL and to identify QTL which have consistent effects on trait expression across environments and genetic backgrounds. Validation studies of published fiber quality QTL effects across genetic backgrounds and environments are needed to address current gaps in the literature, provide information on the efficiency of MAS for fiber quality traits, and ultimately aid in the development of effective genomics-assisted breeding methods for fiber quality traits.

2. DIVERGENT SELECTION IN *GOSSYPIUM* SPP. FOR FIBER LENGTH AND BUNDLE STRENGTH[*]

2.1. Literature Review

In order for a trait, such as fiber length, to be responsive to genetic improvement through selection there must be a reliable and accurate method of measurement. There are instruments available that measure multiple fiber parameters simultaneously, more specifically the High Volume Instrument (HVI) system and the Advanced Fiber Information System (AFIS) (Uster Technologies, Knoxville, TN). This study will focus on fiber quality measurements obtained from the HVI system. It is currently the most commonly used instrument to measure fiber quality parameters due to its high efficiency and low cost. Breeders can obtain a panel of HVI fiber property measures for approximately $2.50 per sample (E.F. Hequet, personal communication, 2015). Additionally, HVI measures serve as the standards upon which the U.S. cotton price support programs are based. Fiber length parameters obtained through the HVI system include upper-half mean length (UHML), which is the mean of the longest 50% of fibers, and the length uniformity index (UI), which is the ratio between the average fiber length and UHML. Fiber length parameters are estimated by providing the HVI system

[*]Part of the data reported in this section is reprinted with permission from Joy, K.S. 2014. Inheritance of cotton fiber length and strength. Ph.D. diss., Texas A&M University, College Station, TX. Copyright 2014, Kolbyn S. Joy.

with a random sub-sample of cotton fibers, which are aligned as a beard, and then the beard length is measured by optical sensing. HVI bundle strength (Str) is reported as the force, in grams, required to break one tex unit (i.e., grams per 1,000 meters) of fibers. The HVI system measures Str by clamping a bundle of fibers at a specified length (3.18 mm) and measuring the force required to break the bundle. Beard mass is estimated through optical sensing. Multiple studies have demonstrated a significant relationship between HVI fiber properties, particularly UHML and Str, and ring-spun yarn properties, such as tenacity (Meredith et al., 1991; Faulkner et al., 2012; Üreyen and Kadoğlu, 2006). The HVI system also measures fiber elongation at break (Elon), color, trash content, neps (i.e., a small entanglement of fibers), UV florescence to measure whiteness, and micronaire (Mic). Elon is based on the distance traveled by the clamps before the fiber bundle breaks, and Mic is an indirect measure of fiber maturity and fineness. HVI Mic is essentially a measure of the specific surface area of fibers, therefore it does not provide a particularly accurate estimate of the maturity or fineness of a sample.

A trait must be heritable to be amenable to selection. Studies have demonstrated that UHML and Str are inherited quantitatively and controlled by multiple genes (Meredith, 1984). Heritability estimates for UHML and Str commonly range from moderate-to-highly heritable, and the majority of genetic variation is additive in nature (Ali et al., 2008; Baker and Verhalen, 1973; Al-Rawi and Kohel, 1970; Hinze et al., 2011; Jenkins et al., 2009; Lee et al., 1967; May, 1999; Ulloa, 2006; Zeng et al., 2011). Moreover, UHML and Str are, in general, minimally influenced by genotype x

environment (G x E) interactions (Abou-El-Fittouh et al., 1969; Al-Jibouri et al., 1958; Chee and Campbell, 2009; Meredith and Bridge, 1973; Miller et al., 1958; Lacape et al., 2010). However, heritability estimates apply only to a specific population or genetic background and to the environments in which the populations are grown (Holland et al., 2003). Therefore, it is not surprising that a number of studies have reported contradictory findings. Several studies have reported a predominance of dominance genetic variance for UHML and Str in upland breeding populations (Tang et al., 1996; May and Green, 1994) and in populations derived from crosses with exotic upland cotton germplasm (Campbell et al., 2014; Cheatham et al., 2003; McCarty et al., 2004). Nevertheless, considering the predominantly additive nature of genetic variation in UHML and Str and relatively minor G x E interaction effects, selection for these traits typically is practiced in early generations within pedigree-based breeding programs.

The final criterion that must be satisfied for a trait to be responsive to selection is that there must exist genetic variation. There are numerous reports of extremely low levels of genetic diversity within cultivated upland cotton germplasm (Fang et al. 2013; Hinze et al., 2012; Lacape et al., 2007; Van Deynze et al., 2009; Van Esbroeck and Bowman, 1998). In relation to the genetic improvement of fiber quality, plant breeders often exploit interspecific crosses between upland cotton and *G. barbadense* as a source of beneficial alleles for fiber quality, but interspecific introgression of fiber quality traits through conventional breeding has been largely unsuccessful (Jiang et al., 2000; Stephens, 1949). Alternatively, some breeders suggest that sufficient genetic variation exists within the upland cultivated gene pool, despite evidence of low genetic diversity.

Bowman et al. (1996) estimated the coefficient of parentage of 260 upland cultivars released between 1970 and 1990 and concluded that sufficient genetic variation existed within the cultivated germplasm to sustain continued genetic improvement through selection. More recent studies on the inheritance of fiber quality traits have continued to report moderate-to-high heritability estimates, suggesting the presence of genetic variation among public breeding program germplasm (Campbell and Meyers, 2015).

Relationships between traits are an important consideration for selection schemes. Generally, there is a positive association between UHML and Str facilitating simultaneous improvement of both traits, but the reported correlations for UHML and Str with other fiber properties are inconsistent across genetic backgrounds (Campbell et al., 2012; Jenkins et al., 2009). Despite relatively abundant additive genetic variation and favorable correlations among fiber quality traits, the observed rate of genetic gain for fiber length and strength has been considerably lower than that achieved for lint yield (Campbell et al., 2011; Campbell and Meyers, 2015). One challenge in breeding elite upland cultivars with improved fiber quality is the negative relationship between fiber quality traits, specifically UHML and Str with yield components (Al-Jibouri et al., 1958; Hinze et al., 2011; McCall et al., 1986; Meredith, 1984; Miller et al., 1958; Miller and Rawlings, 1967; Smith and Coyle, 1997; Ulloa, 2006). Notably, many of the fiber quality traits of interest in upland cotton are neither maximized nor minimized during selection by breeders not only because of the aforementioned negative association but also because of market demand. For example, the current upland cotton market demands

what one may consider average quality that is required for the majority of final products, from sheeting to denim.

Research suggests that the negative relationship between fiber quality and lint yield is, at least in part, attributable to repulsion phase linkage. Meredith and Bridge (1967) were able to decrease in magnitude the negative correlation between lint yield and Str through intermating. There are also several reports of germplasm lines exhibiting high lint yield and Str within Pee Dee germplasm (Campbell et al., 2012; Culp and Green, 1992; Green and Culp, 1990). Zeng et al. (2011) identified germplasm lines derived from crosses between elite upland cultivars and exotic lines with positive general combining ability for both lint yield and Str, and they also identified progeny lines harboring favorable alleles for UHML and Str, with no compensatory reduction in yield. The magnitude of the negative association between UHML and lint yield varies widely depending on genetic background compared to Str (Constable and Bange, 2007). Research suggests that hybridization between germplasm lines with elite fiber quality and high yield and the use of breeding methods to increase recombination, such as intermating, can be utilized to combine high lint yield with improved UHML and Str.

Joy (2014) reported on the combining ability and inheritance of cotton fiber quality traits among several genetically diverse parental lines that he selected based on pedigree and fiber properties. The study found significant genetic variation for UHML and Str among the parental lines and identified several lines which should be beneficial for the genetic improvement of UHML and Str. The narrow-sense heritability estimates (i.e., the proportion of additive genetic variation to phenotypic variation) for UHML and

Str among the parental combinations were moderate-to-high, and the majority of the genetic variation in fiber properties was attributable to additive genetic effects. Specifically, the results suggest that 04 SID 84-2 (SID84), an unreleased, experimental interspecific hybrid (*G. hirsutum* x *G. barbadense*), would be useful as a parental line for the improvement of UHML, and TAM B182-33 ELS (ELS33) (Smith et al., 2009; PI 654362), an extra-long staple upland (ELSU) line, would be useful for the simultaneous improvement of UHML and Str. Additionally, the transgressive segregation in populations derived from crosses with SID84 indicated that the line may contain favorable alleles for Str not present in the upland parental lines included in the study.

Based on these previous findings, the current study presents the results of divergent selection for HVI measured fiber UHML and Str conducted within selected parental combinations from the research conducted by Joy (2014). Realized heritability was estimated based on how much of the selection differential imposed was observed as a response in the progeny (Hill, 1972; Holland et al., 2003) and was used also to test the predictions for selection made by Joy (2014). A similar experiment was conducted by McCall et al. (1986). They performed multi-directional selection for Str within an upland cotton population and used realized heritability estimates to evaluate effectiveness of multiple cycles of selection. McCall et al. (1986) found that selection for greater Str resulted in reduced lint percent, increased UHML and UI, as well as earliness. Similarly, Miller and Rawlings (1967) conducted three cycles of recurrent selection for lint yield and observed variable correlated responses for UHML, Str, fineness, and Elon among

cycles. The current study examined the correlated responses to selection for fiber UHML and Str within the parental combinations selected from the study by Joy (2014).

2.2. Objectives

- Estimate realized heritability for fiber UHML and Str in five genetically diverse populations subjected to divergent selection.
- Define the relationship between HVI measured fiber properties, including UHML, Str, UI, Mic, and Elon, as traits respond to selection for UHML and Str.

2.3. Materials and Methods

Plant material. Five Texas A&M AgriLife Research (AgriLife) experimental populations were selected for the study based on previously estimates of breeding potential for fiber length and strength (Joy 2014). Four genetically diverse parental lines were crossed to derive the five populations: 06 WE 62-4 (HS624), an unreleased, upland experimental line with exceptional fiber strength, 'Tamcot 22' (TAM22) (Thaxton et al., 2005; PI 635877), an upland commercial cultivar with moderate fiber qualities, SID84, and ELS33. The parental lines of HS624 were upland commercial cultivars, Delta and Pine Land (DPL) 491 (PI 618609 PVPO) and DPL 90 (PI 529529 PVPO), along with TAM 96WD-18 (Thaxton et al., 2005; PI 635879) and TAM 91C-95Ls (Smith, 2001; PI 612326). The experimental line, SID84, was derived from a cross between the upland commercial cultivar, TAM 94L-25 (Smith, 2003: PI 631440), and the *G. barbadense* cultivar, New Mexico Sea Island (NMSI) 1331 (Roberts et al., 1997). Five F_2 populations were derived from the following crosses (ignoring reciprocals): (1) HS624 x

ELS33, (2) TAM22 x ELS33, (3) ELS33 x SID84, (4) HS624 x TAM22, and (5) TAM22 x SID84.

Field trials and selection scheme. Field trials evaluating the populations were conducted at the AgriLife Research Farm near College Station, TX on a Weswood silt loam, a fine-silty, mixed, thermic Fluventic Ustochrept, integrated with Ships clay, a very fine, mixed, thermic Udic Chromustert. Standard cultural practices for cotton production in central Texas were conducted, including pesticide and herbicide applications and furrow irrigation. Joy (2014) obtained HVI measured UHML and Str on approximately 350 F_2 progeny for each of the five populations grown in 2010 and 2011. The first cycle of divergent selection was conducted within each population by selecting the top and bottom five percent of F_2 individual plants (IPs) for both UHML and Str. The F_2 IP selections were planted on April 12, 2012 as single $F_{2:3}$ progeny rows measuring 13.1 m x 1 m, and plants within each row were thinned to an approximate density of one plant per 0.4 m. Eight IPs were selected at random within each $F_{2:3}$ progeny row. Fifteen-boll samples were hand harvested from each of the eight IPs within each progeny row in late September 2012, and seed cotton samples were ginned on a 10-saw laboratory gin without lint cleaners. Fiber properties as measured by HVI were determined at Texas Tech University's Fiber and Biopolymer Research Institute (FBRI) in Lubbock, TX. The top and bottom ten percent of $F_{2:3}$ IPs for UHML and Str were selected within each population for the second cycle of divergent selection. The selected $F_{2:3}$ IPs were then planted as single $F_{2:4}$ progeny rows (13.1 m x 1 m) on April 24, 2013, and rows were thinned to an approximate density of one plant per 0.4 m. Four IPs were selected at

random within each of the $F_{2:4}$ progeny rows, and fifteen-boll samples were hand harvested from each selected IP in mid-to-late October, 2013. Seed cotton samples were ginned on a laboratory saw-gin without lint cleaners, and HVI measured fiber properties were determined at Texas Tech University's FBRI in Lubbock, TX.

The $F_{2:4}$ IPs with the highest/lowest UHML or Str within each of the progeny rows, corresponding with the direction of divergent selection, were selected and planted as replicated $F_{2:5}$ strains along with the parental lines on May 6, 2014. The trial was planted as a randomized complete block design (RCBD) with two replications. Thirty-boll samples were hand harvested from each $F_{2:5}$ progeny row in late October, preferentially picking first- and second-position bolls from the middle of the fruiting zone to minimize variation due to environmental factors. Seed cotton samples were processed as described above. The data collected on the $F_{2:5}$ strains were used to evaluate the effectiveness of divergent selection for UHML and Str.

Statistical analysis. The results of divergent selection were used to estimate the realized heritability of a quantitative trait (Falconer and Mackay 1996; Hill, 1972). The following formula described by Fehr (1987) was used to estimate realized heritability for UHML and Str on a single-plant basis for selection cycle one,

$$h^2 = \frac{\bar{x}_{high,F3} - \bar{x}_{low,F3}}{\bar{x}_{high,F2} - \bar{x}_{low,F2}}$$

where $\bar{x}_{high,F3}$ and $\bar{x}_{low,F3}$ represent the mean performance of the $F_{2:3}$ progeny of the F_2 IPs selected for the high and low groups, respectively, and $\bar{x}_{high,F2}$ and $\bar{x}_{low,F2}$ represent the mean performance of the F_2 IPs in the high and low groups, respectively. The same

formula was used to obtain realized heritability estimates for cycle two using the mean performance of the selected $F_{2:3}$ IPs and the resulting $F_{2:4}$ progeny. One major limitation of this study was that the resulting realized heritability estimates for UHML and Str were biased due to the evaluation of each generation in a different environment. However, this bias is likely minor considering that variation in fiber properties due to environment and G x E interaction effects is generally much smaller than the variation due to genotypic effects (Chee and Campbell, 2009; Meredith and Bridge, 1973; Miller et al., 1958).

Predicted and observed responses to selection for each cycle were calculated for each population. The observed response to selection for cycle one was calculated by taking the difference between the mean performance of the F_2 IPs and the mean performance of the $F_{2:3}$ progeny resulting from the selected F_2 IPs. The same formula was used to calculate the observed response to selection for cycle two using the mean performance of the $F_{2:3}$ IPs and the mean performance of the $F_{2:4}$ progeny derived from the selected $F_{2:3}$ IPs. The following formula described by Falconer and Mackay (1996) was used to calculate the predicted response to selection for each cycle,

$$R = h^2S$$

Where R represents the predicted response to selection, h^2 represents the heritability, and S represents the selection differential, which is the mean difference between the base population and the selected parents. The narrow-sense heritability estimates for each population obtained by Joy (2014) were used in predicting the response to the first cycle of selection for UHML and Str, and the realized heritability estimates from cycle one were used in predicting the response to the second cycle of selection. The selection

intensity (i.e., standardized selection differential), *i*, for each population and cycle and was calculated by dividing the selection differential by the phenotypic standard deviation (Falconer and Mackay, 1996). Figures illustrating response to selection were generated in JMP (SAS Institute Inc., 2013).

Analyses of variance (ANOVA) were conducted within each population on UHML and Str among the $F_{2:5}$ strains resulting from divergent selection. Data were analyzed as a RCBD with subsampling using the General Linear Models procedure (PROC GLM) in SAS (SAS Enterprise Guide, Version 6.1, SAS Institute Inc., Cary, NC, 2013). The $F_{2:5}$ strains were grouped according to the corresponding trait and direction of divergent selection, resulting in a high and low group for both UHML and Str. Groups (i.e., high and low UHML/Str) were analyzed as fixed effects, and replication and group x replication terms were treated as random effects. Analyses of UHML and Str conformed to the assumptions of ordinary least squares estimation, including the normality and homogeneity of residuals. The group x replication mean square error was used to test the significance of replication and group. Mean comparisons between groups were conducted using the MEANS statement in SAS to specify Fisher's protected least significant difference based on the mean square error of the group x replication interaction. Pearson's correlation coefficients among fiber properties and lint percent within each population (computed on the $F_{2:5}$ strain means) were calculated using the Correlation procedure in SAS (PROC CORR), and histograms and scatterplots were generated in R (R Development Core Team, 2013).

2.4. Results and Discussion

Selection for fiber UHML. The mean fiber UHML of the parental lines when grown at College Station in 2014 are shown in Table 2.1. The realized heritability estimates for fiber UHML among the five populations were moderate-to-high, ranging from 0.89 for TAM22 x ELS33 to 0.32 for HS624 x ELS33 (Table 2.2). As noted previously, a major limitation to this study and to the interpretation of the results was the evaluation of each generation in a separate environment. Consequently, the realized heritability estimates are biased due to environmental effects and G x E interaction effects (Holland et al., 2003). Nevertheless, the results support the findings of Joy (2014) that there is sufficient additive genetic variation among these parental combinations to enable selection for improved UHML. The realized heritability estimates from the second cycle of divergent selection ($F_{2:3}$ to $F_{2:4}$) were either similar or slightly higher than the estimates from cycle one (F_2 to $F_{2:3}$). They were also higher than the narrow-sense heritability estimates derived from the F_2 populations, with the exception of HS624 x ELS33 in cycle one.

Table 2.1. Mean performance of parental lines grown at College Station, TX in 2014.

	HVI fiber properties[‡]					
	UHML	**Str**	**Mic**	**UI**	**Elon**	**LP[§]**
Genotype[†]	**- mm -**	**- kN m kg^{-1} -**		**- % -**		**- % -**
ELS33	33.5	338.5	3.7	84.5	5.2	34.6
HS624	30.2	377.2	4.5	85.2	5.8	37.2
SID84	33.7	360.5	3.5	85.6	7.0	30.7
TAM22	27.9	272.7	4.0	83.5	6.7	36.8

[†]ELS33 = TAM B182-33 ELS; HS624 = 06 WE 62-4; SID84 = 04 SID 84-2; and TAM22 = Tamcot 22.
[‡]UHML = Upper-half mean length; Str = bundle strength; Mic = micronaire; UI = uniformity index; Elon = elongation.
[§]LP = Lint percent.

Table 2.2. Fiber UHML of selected plants, their progeny, and realized heritability estimates for cycles one (F_2 to $F_{2:3}$) and two ($F_{2:3}$ to $F_{2:4}$) of divergent selection within five populations evaluated at College Station, TX in 2012 and 2013. Selection intensity (*i*) is shown for each group in each cycle.

	HS624 x ELS33	**TAM22 x ELS33**	**ELS33 x SID84**	**HS624 x TAM22**	**TAM22 x SID84**
Cycle 1[†]			- mm -		
$\mathbf{F_2}$	32.0	30.6	32.0	29.1	30.5
$\mathbf{F_{2,\ high}}$	34.5	34.0	36.1	32.0	33.5
$\mathbf{F_{2:3,\ high}}$	34.0	34.0	35.6	31.8	32.8
$\boldsymbol{i_{high}}$[‡]	*1.8*	*1.9*	*2.1*	*2.0*	*1.7*
$\mathbf{F_{2,\ low}}$	28.2	26.2	27.4	26.2	27.2
$\mathbf{F_{2:3,\ low}}$	32.0	28.4	28.2	28.2	27.4
$\boldsymbol{i_{low}}$	*-2.6*	*-2.5*	*-2.4*	*-2.0*	*-1.8*
Cycle 2 [§]			- mm -		
$\mathbf{F_{2:3,\ high}}$	35.3	35.3	37.1	33.5	34.5
$\mathbf{F_{2:4,\ high}}$	35.3	34.8	37.3	32.3	34.8
$\boldsymbol{i_{high}}$	*0.7*	*0.3*	*0.4*	*0.7*	*0.5*
$\mathbf{F_{2:3,\ low}}$	29.5	26.4	26.9	26.9	25.4
$\mathbf{F_{2:4,\ low}}$	32.3	26.9	28.7	27.7	27.2
$\boldsymbol{i_{low}}$	*-1.5*	*-0.6*	*-0.3*	*-0.5*	*-0.5*
Heritability					
Narrow-sense[¶]	0.47	0.67	0.61	0.56	0.42
Realized					
Cycle 1	0.32	0.71	0.85	0.61	0.84
Cycle2	0.52	0.89	0.85	0.69	0.83

[†]Cycle one of divergent selection, where F_2 represents the overall mean UHML of the F_2 populations, F_2, *high* and F_2, *low* represent the mean UHML of the high and low selected F_2 plants, and $F_{2:3}$, *high* and $F_{2:3}$, *low* represent the mean UHML of the $F_{2:3}$ plants derived from the high and low selected F_2 plants.
[‡]Selection intensity for high and low UHML. Selection intensity is inversely proportional to the percent of plants selected.
[§]Cycle two of divergent selection, where $F_{2:3}$, *high* and $F_{2:3}$, *low* represent the mean UHML of the high and low selected $F_{2:3}$ plants, and $F_{2:4}$, *high* and $F_{2:4}$, *low* represent the mean UHML of the $F_{2:4}$ plants derived from the high and low selected $F_{2:3}$ plants.
[¶]Narrow-sense heritability estimates for each population derived from the generation means analysis conducted by Joy (2014).

Joy (2014) reported relatively high dominance effects for fiber UHML among the five parental combinations. It is possible that further inbreeding led to more effective selection in later generations resulting in higher realized heritability estimates. Alternatively, it may be that 2012 was a more favorable environment for the expression

of UHML compared to 2011 leading to higher realized heritability in cycle two. The highest realized heritability estimates were observed among ELS33 x SID84, TAM22 x SID84, and TAM22 x ELS33. Bearing in mind that realized heritability provides a measure of how much of the selection differential imposed was observed as a response in the progeny, the results suggest that early generation selection UHML within the three aforementioned populations is highly effective at College Station. The realized heritability estimates for HS624 x ELS33 were substantially lower compared to the other four populations. This may be attributable to reduced genetic variance between the two parental lines, which are both products of a long-term selection program for fiber quality, in combination with environmental and G x E interaction effects.

The selection differential is the difference between the mean performance of selected genotypes and the overall population mean, but it is not suitable for comparing the strength of selection across different populations or traits. Therefore, selection intensity was used instead which provides a measure of the selection differential expressed as a proportion of the phenotypic standard deviation. Consequently, selection intensity is inversely proportional to the percent of selected plants. Selection intensities differed slightly between populations within selection cycles for high and low UHML, yet these differences were generally minimal (~0.5) allowing for the comparison of the effectiveness of selection across populations (Table 2.2). Selection intensity was lowered for the second cycle of selection, resulting in smaller responses to selection for UHML across all populations. The observed response to selection was similar to the predicted response for each population, and selection for high UHML was generally more

effective than selection for low UHML (Figure 2.1). The minor differences between observed and predicted responses to selection may be attributable to non-additive genetic effects, environmental effects, G x E interactions, or a combination thereof. The percent gain or loss in UHML can be calculated from Table 2.2 by subtracting the mean of the progeny derived from the selected genotypes by the overall mean of the generation in which selection was performed, dividing by the same overall mean, and multiplying by 100. Consistent with the realized heritability estimates, the greatest cumulative percent change (i.e., gain plus loss for cycles one and two) in UHML was observed for ELS33 x SID84 (high: 16.3%; low: -10.1%), TAM22 x ELS33 (high: 13.5%; low: -12.5%), and TAM22 x SID84 (high: 12.7%; low: -10.4%). The smallest cumulative percent change in UHML was observed for HS624 x ELS33. Two cycles of selection resulted in a 10.1% increase in the mean UHML for HS624 x ELS33, but no net decrease in mean UHML was observed, despite a relatively high selection intensity. The failure to effectively decrease the mean UHML of progeny derived from HS624 x ELS33 may be attributable to environmental effects masking relatively low levels of additive genetic variation.

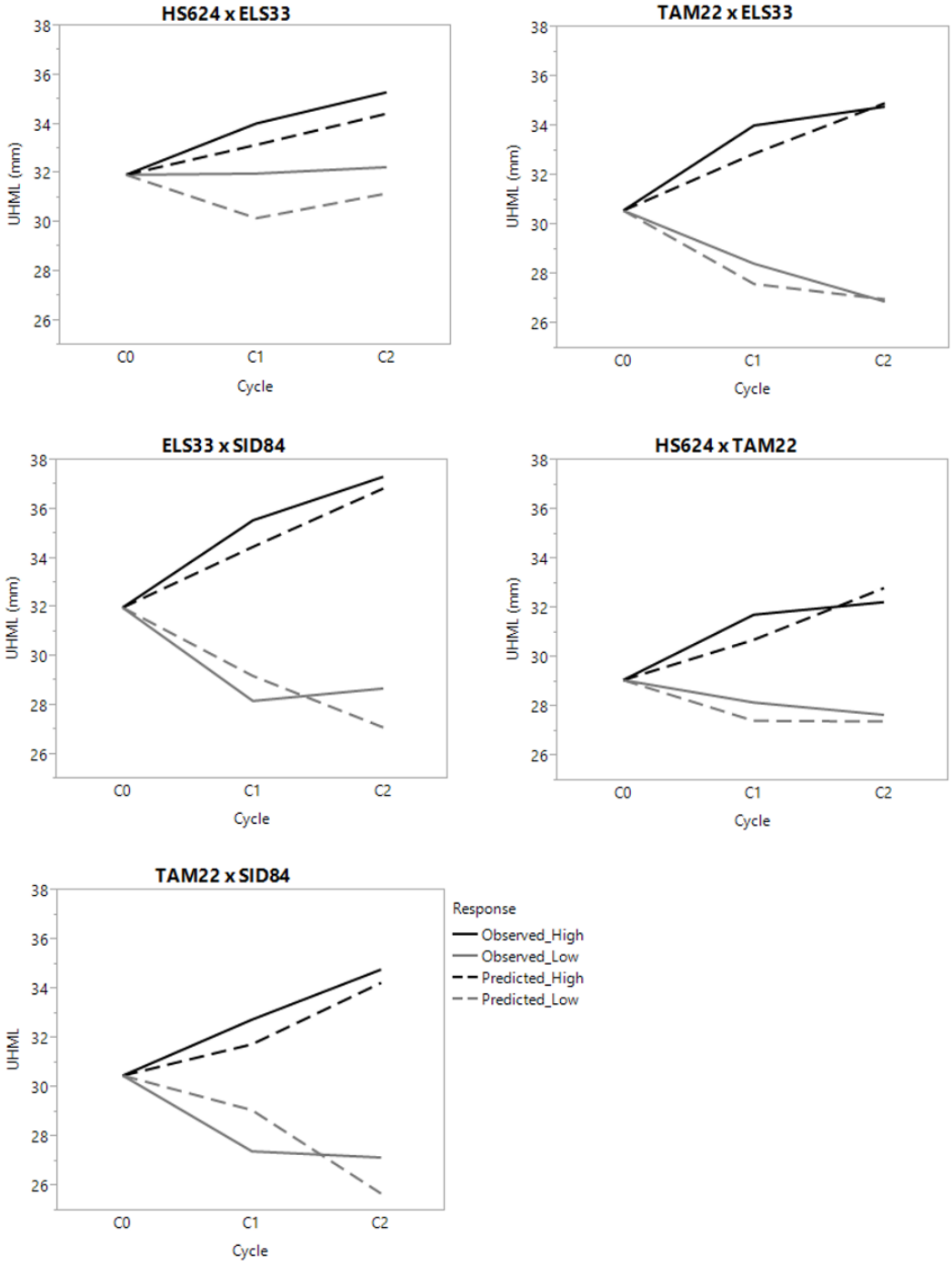

Figure 2.1. Observed versus predicted response to selection for high and low UHML at College Station, TX. Cycle zero (C0) refers to the mean UHML among all F_2 IPs; cycle one (C1) represents the mean UHML of the $F_{2:3}$ progeny derived from high and low selected F_2 IPs; and cycle two (C2) represents the mean UHML of the $F_{2:4}$ progeny derived from the high and low selected $F_{2:3}$ IPs.

ANOVA of the two UHML groups, i.e., high and low, and considering progeny within each group within each population as subsamples (Compton, 1994) indicated that early generation divergent selection for UHML, was effective in four of the five populations, resulting in a significant difference in mean UHML (Table 2.3).

The exception was divergent selection within HS624 x TAM22. The mean difference in UHML between the high and low length groups derived from HS624 x TAM22 was 4.29 mm, but the difference was not statistically significant due to relatively large plot error (i.e., replication x group interaction) in comparison to genetic (i.e., group) effects.

Table 2.3. ANOVA and mean UHML of selected groups (i.e., high and low UHML) composed of $F_{2:5}$ strains derived from divergent selection and evaluated at College Station, TX in 2014.

	HS624 x ELS33		**TAM22 x ELS33**		**ELS33 x SID84**		**HS624 x TAM22**		**TAM22 x SID84**	
Source[‡]	***df***	**MS**[†]	***df***	**MS**	*df*	**MS**	***df***	**MS**	***df***	**MS**
Rep	1	0.98	1	23.56	1	0.17	1	50.03	1	13.22
Group	1	251.01*	1	1270.8*	1	1306.97*	1	454.34	1	489.86*
Rep*Group	1	0.48	1	5.75	1	1.32	1	4.59	1	2.54
Sampling Error	101	1.46	119	1.61	96	2.22	98	0.96	86	1.87
LS Means[§]										
High	33.46 a		33.31 a		36.05 a		31.72 a		33.48 a	
Low	30.35 b		26.53 b		28.78 b		27.43 a		28.77 b	

* significant at the 0.05 probability level.

[†]Mean squares.

[‡] $F_{2:5}$ strains were categorized into two groups for analysis, based on selection for high or low UHML.

[§]Means within the same column (i.e., within population) with the same letter are not significantly different according to Fisher's protected least significant difference (LSD) at the 0.05 probability level.

The greatest difference in mean UHML between high and low length groups was observed for ELS33 x SID84 (7.27 mm), and the smallest difference in mean UHML

was observed for HS624 x ELS33 (3.11 mm). The mean UHML of $F_{2:5}$ strains derived from HS624 x ELS33 selected for high UHML was similar to that of $F_{2:5}$ strains derived from TAM22 x ELS33 and TAM22 x SID84, even though progeny derived from HS624 x ELS33 exhibited the smallest response to selection (Table 2.3). Positive and negative transgressive segregants for UHML were observed in all five populations (Figure 2.2). The most extreme transgressive segregation was observed within $F_{2:5}$ strains derived from ELS33 x SID84 suggesting a high level of allele dispersion between the two parents.

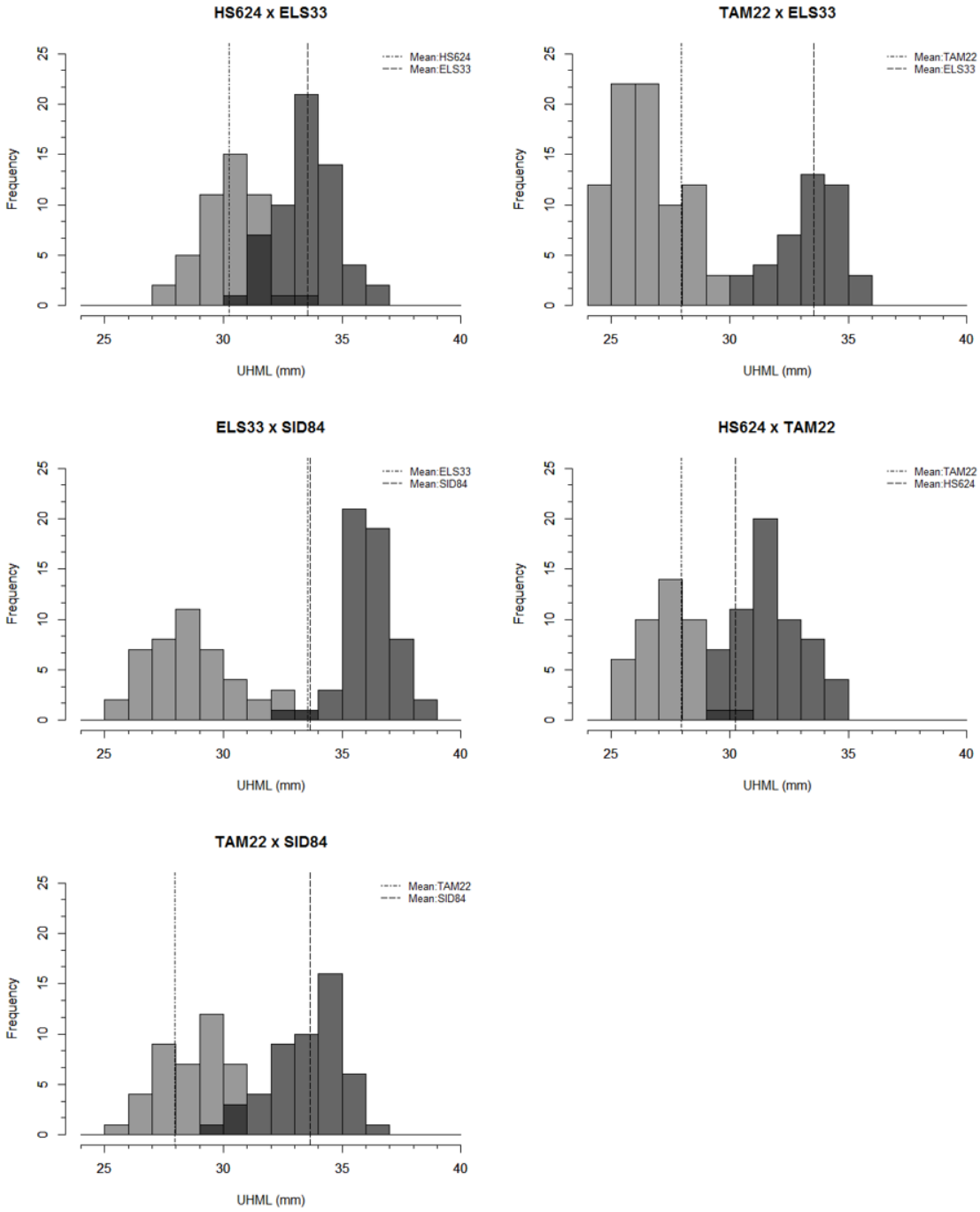

Figure 2.2. Distribution of UHML within five $F_{2:5}$ populations derived by divergent selection for UHML when evaluated at College Station, TX in 2014. Vertical lines represent the mean UHML of the parental lines also evaluated at College Station, TX in 2014.

Selection for fiber Str. The realized heritability estimates for fiber Str were lower compared to fiber UHML, ranging from 0.72 for TAM22 x SID84 to 0.14 for HS624 x ELS33 (Table 2.4). This may indicate that environmental effects and G x E interactions had larger effects on the expression of fiber Str compared to UHML at College Station. Nonetheless, the realized heritability estimates were still moderate-to-high suggesting that early generation selection for Str is effective at College Station, especially within the parental combinations TAM22 x SID84, ELS33 x SID84, and HS624 x TAM22. The realized heritability estimate for Str in cycle one within HS624 x ELS33 was substantially lower than the narrow-sense heritability and cycle two realized heritability estimates. These results suggest the importance of non-additive genetic effects, environmental effects, and/or G x E interactions for Str within this parental combination.

Table 2.4. Fiber Str of selected plants, their progeny, and realized heritability estimates for cycles one (F_2 to $F_{2:3}$) and two ($F_{2:3}$ to $F_{2:4}$) of divergent selection within five populations evaluated at College Station, TX in 2012 and 2013. Selection intensity (*i*) is shown for each group in each cycle.

	HS624 x ELS33	**TAM22 x ELS33**	**ELS33 x SID84**	**HS624 x TAM22**	**TAM22 x SID84**
Cycle 1[†]			- kN m kg^{-1} -		
F_2	367.0	304.6	339.2	318.4	338.0
$F_{2,\ high}$	424.3	359.1	412.0	383.1	424.6
$F_{2:3,\ high}$	359.9	339.0	368.8	350.9	360.5
i $_{high}$[‡]	*2.1*	*2.1*	*2.4*	*2.1*	*1.9*
$F_{2,\ low}$	297.8	246.2	263.1	252.8	259.4
$F_{2:3,\ low}$	342.1	287.5	280.1	293.9	278.4
i $_{low}$	*-2.5*	*-2.2*	*-2.5*	*-2.1*	*-1.7*

Table 2.4. *Continued.*

Cycle 2 [§]	- kN m kg^{-1} -				
$\mathbf{F}_{2:3,\ high}$	379.6	363.0	396.3	385.5	397.3
$\mathbf{F}_{2:4,\ high}$	380.6	338.4	367.9	367.9	398.3
i_{high}	*0.9*	*1.0*	*0.7*	*0.9*	*0.7*
$\mathbf{F}_{2:3,\ low}$	312.0	259.0	257.0	264.9	248.2
$\mathbf{F}_{2:4,\ low}$	344.3	296.3	284.5	289.4	290.4
i_{low}	*-1.3*	*-1.1*	*-0.6*	*-0.8*	*-0.6*
Heritability					
Narrow-sense[¶]	0.55	0.46	0.46	0.58	0.61
Realized					
Cycle 1	0.14	0.46	0.60	0.44	0.50
Cycle2	0.54	0.41	0.60	0.65	0.72

[†]Cycle one of divergent selection, where F_2 represents the overall mean Str of the F_2 populations, F_2, *high* and F_2, *low* represent the mean Str of the high and low selected F_2 plants, and $F_{2:3}$, *high* and $F_{2:3}$, *low* represent the mean Str of the $F_{2:3}$ plants derived from the high and low selected F_2 plants.
[‡]Selection intensity for high and low Str. Selection intensity is inversely proportional to the percent of plants selected.
[§]Cycle two of divergent selection, where $F_{2:3}$, *high* and $F_{2:3}$, *low* represent the mean Str of the high and low selected $F_{2:3}$ plants, and $F_{2:4}$, *high* and $F_{2:4}$, *low* represent the mean Str of the $F_{2:4}$ plants derived from the high and low selected $F_{2:3}$ plants.
[¶]Narrow-sense heritability estimates for each population derived from the generation means analysis conducted by Joy (2014).

Selection intensity, which is inversely proportional to the percent of selected plants, was similar between populations within selection cycles for high and low Str, and the selection intensity applied for the second cycle was substantially less than that applied for cycle one (Table 2.4). The observed responses to selection for Str were similar to the predicted responses, with the exception of HS624 x ELS33 (Figure 2.3). Effective selection for fiber Str within this parental combination may require additional generations of inbreeding and evaluation across different environments. The greatest cumulative percent change in Str was observed within TAM22 x SID84 (high: 17.1%;

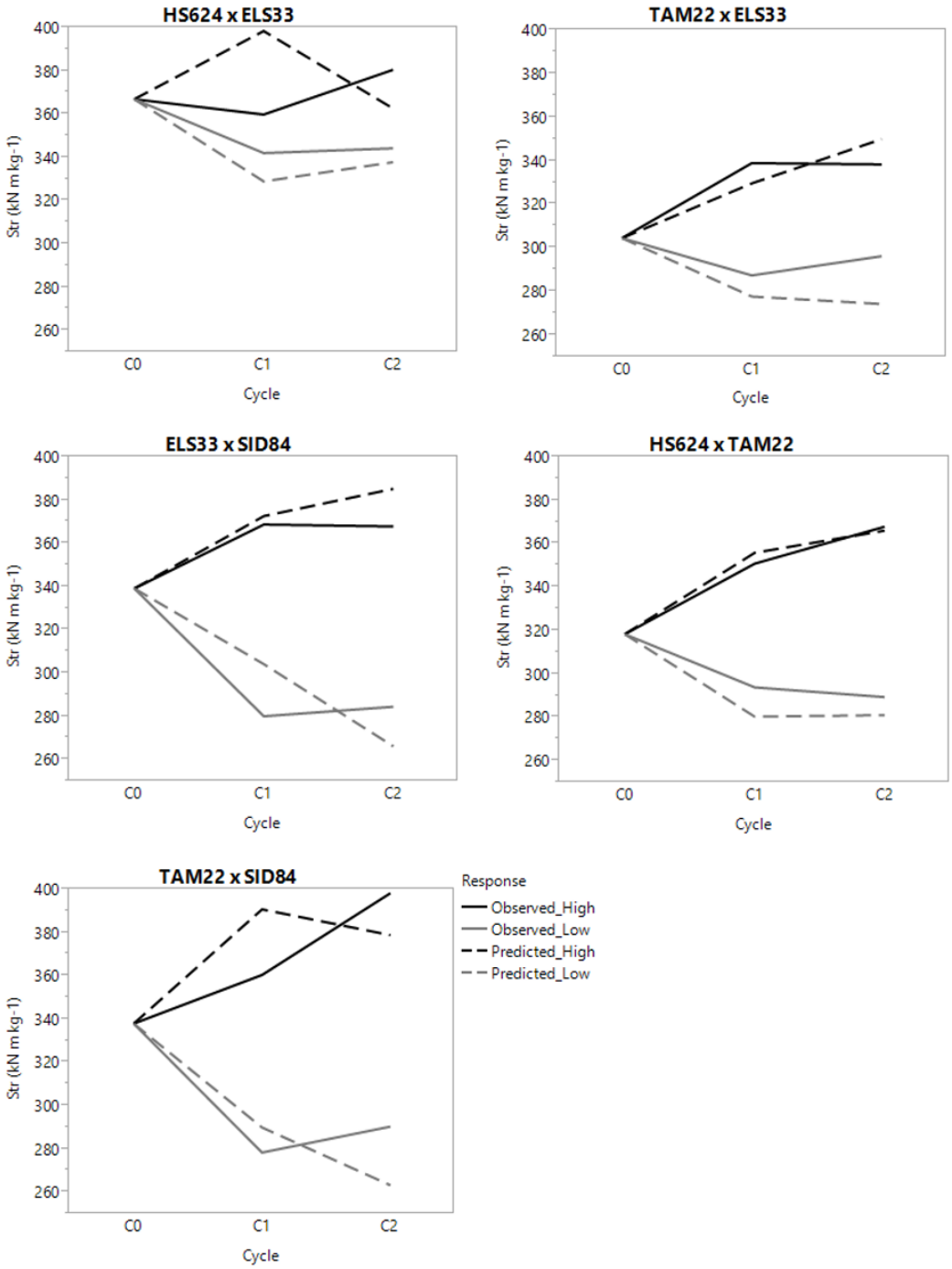

Figure 2.3. Observed versus predicted response to selection for high and low Str at College Station, TX. Cycle zero (C0) refers to the mean Str among all F_2 IPs; cycle one (C1) represents the mean UH Str ML of the $F_{2:3}$ progeny derived from high and low selected F_2 IPs; and cycle two (C2) represents the mean Str of the $F_{2:4}$ progeny derived from the high and low selected $F_{2:3}$ IPs.

low: -13.3%) followed by HS624 x TAM22 (high: 15.0%; low: -9.2%) and ELS33 x SID84 (high: 8.5%; low: -15.8%).

ANOVA of the two Str groups, i.e., high and low, within each population of $F_{2:5}$ strains indicated that divergent selection was effective in four of five populations, resulting in a significant difference in mean Str between strains selected for high versus low Str. The exception was the TAM22 x ELS33 population (Table 2.5).

Table 2.5. ANOVA and mean Str of selected groups (i.e., high and low Str) composed of $F_{2:5}$ strains derived from divergent selection and evaluated at College Station, TX in 2014.

	HS624 x ELS33		**TAM22 x ELS33**		**ELS33 x SID84**		**HS624 x TAM22**		**TAM22 x SID84**	
Source	***df***	**MS**[†]	***df***	**MS**	***df***	**MS**	***df***	**MS**	***df***	**MS**
Rep	1	44.3	1	75.8	1	49.6	1	475.5	1	1093.4
Group[‡]	1	32856.3*	1	77555.2	1	123119.2*	1	163928.1*	1	282418.2*
Rep*Group	1	187.0	1	1124.3	1	342.1	1	746.1	1	840.5
Sampling Error	86	593.6	98	185.4	93	609.7	96	327.4	114	625.3
LS Means[§]										
High		371.2 a		333.2 a		377.5 a		365.7 a		391.5 a
Low		330.7 b		273.8 a		296.0 b		284.8 b		293.3 b

* significant at the 0.05 probability level.
[†]Mean squares.
[‡] $F_{2:5}$ strains were categorized into two groups for analysis, based on selection for high or low Str.
[§]Means within the same column (i.e., within population) with the same letter are not significantly different according to Fisher's protected least significant difference (LSD) at the 0.05 probability level.

The mean difference between high and low Str groups within TAM22 x ELS33 was 59.4 kN m kg^{-1}, but the mean difference was not significant due to the large portion of variation in Str attributable to plot error. The greatest difference in mean Str between high and low strength strains was 98.2 kN m kg^{-1} within TAM22 x SID84. The smallest difference in mean Str of selected $F_{2:5}$ strains was within HS624 x ELS33 (40.5 kN m kg^{-}

[1]); however, the mean of the $F_{2:5}$ strains selected for high Str was comparable to the mean among $F_{2:5}$ strains from ELS33 x SID84 and HS624 x TAM22 selected for high Str. Positive and negative transgressive segregants for fiber Str were observed within all populations (Figure 2.4). The most extreme transgressive segregation was observed within $F_{2:5}$ strains derived from ELS33 x SID84, again suggesting a high level of allele dispersion between the two parents.

Correlated responses to selection. Several phenotypic correlations were detected between fiber properties among the $F_{2:5}$ strains derived from divergent selection for UHML and Str. The positive relationship between UHML and Str has been well established based on previous studies (Jenkins et al., 2009; Percy et al., 2006; Tang et al., 1996; Ulloa, 2006). In this study, a positive correlation between fiber UHML and Str was detected for each of the populations except HS624 x ELS33 (Table 2.6). While the results might indicate that the linkage between UHML and Str has been broken within HS624 x ELS33, it may also be a function of lower allele dispersion between the two parental lines for UHML and Str resulting in little variability in fiber quality among the $F_{2:5}$ progeny. Both fiber UHML and Str were highly positively correlated to UI within each of the populations. These results suggest that the simultaneous improvement of UHML, Str, and UI should be readily achieved within these parental combinations.

Correlations of UHML, Str, and UI with Elon were relatively inconsistent across genetic backgrounds, which corresponds to the literature (Campbell et al., 2012; Jenkins et al., 2009). Fiber UHML, Str, and UI were negatively correlated with Elon among $F_{2:5}$ strains derived from ELS33 x SID84 and TAM22 x ELS33.

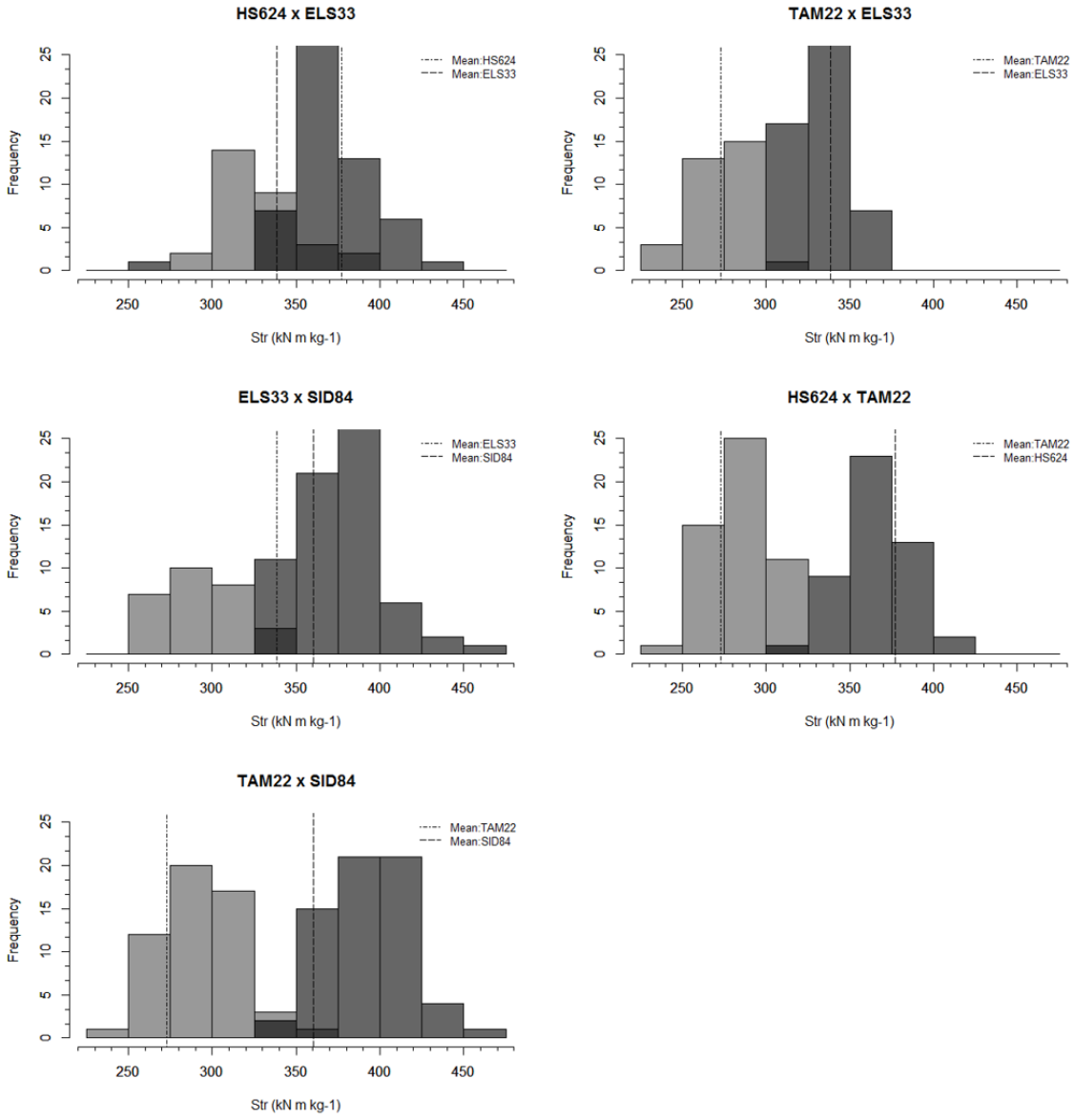

Figure 2.4. Distribution of Str within five $F_{2:5}$ populations derived by divergent selection for Str when evaluated at College Station, TX in 2014. Vertical lines represent the mean Str of the parental lines also evaluated at College Station, TX in 2014.

Table 2.6. Phenotypic correlations[†] among HVI measured fiber properties and lint percent of $F_{2:5}$ strains grown at College Station, TX in 2014.

	Str[‡]	Mic	UI	Elon	LP
HS624 x ELS33 (*N* = 88)					
UHML	0.13	-0.59***	0.43***	-0.07	-0.39***
Str		0.00	0.61***	0.43***	-0.07
Mic			-0.05	-0.03	0.58***
UI				0.23*	-0.17
Elon					0.13
TAM22 x ELS33 (*N* = 102)					
UHML	0.65***	-0.92***	0.78***	-0.79***	-0.72***
Str		-0.50***	0.84***	-0.22*	-0.63***
Mic			-0.61***	0.81***	0.67***
UI				-0.41***	-0.68***
Elon					0.54***
ELS33 x SID84 (*N* = 95)					
UHML	0.72***	-0.40***	0.79***	-0.76***	-0.70***
Str		0.02	0.85***	-0.47***	-0.47***
Mic			-0.01	0.17	0.55***
UI				-0.52***	-0.51***
Elon					0.47***
HS624 x TAM22 (*N* = 101)					
UHML	0.71***	-0.72***	0.76***	-0.39***	-0.54***
Str		-0.46***	0.81***	-0.07	-0.57***
Mic			-0.57***	0.14	0.52***
UI				-0.17	-0.56***
Elon					-0.03
TAM22 x SID84 (*N* = 90)					
UHML	0.67***	-0.27*	0.83***	0.03	-0.62***
Str		0.18	0.85***	0.13	-0.42***
Mic			0.02	-0.22*	0.46***
UI				0.09	-0.49***
Elon					-0.25*

*, **, *** significant at 0.05, 0.01, and 0.001 probability levels, respectively.
[†]Pearson's correlation coefficients.
[‡]UHML = upper-half mean length; Str = bundle strength; Mic = micronaire; UI = uniformity index; Elon = elongation; LP = lint percent.

Fiber UHML and UI were negatively correlated with Elon among $F_{2:5}$ strains derived from HS624 x TAM22, but there was no significant correlation detected between fiber Str and Elon. Ng et al. (2014) reported similar observations regarding correlations between UHML, Str, and Elon in five different genetic backgrounds grown at College

Station. There was no detectable significant correlation of UHML, Str, and UI with Elon among $F_{2:5}$ strains derived from TAM22 x SID84. Positive correlations between Str, UI, and Elon were detected within HS624 x ELS33, while there was no significant correlation between UHML and Elon. These two parental combinations may provide unique opportunities for the simultaneous improvement of Str, UHML, and Elon.

None of the parental lines had Mic indices outside of the non-discount range (3.5 to 4.9) as defined in the 2015 CCC Loan Premium and Discount Schedule for upland cotton (http://www.cotton.org/econ/govprograms/cccloan/). Few, to none, of the $F_{2:5}$ strains derived from HS624 x ELS33 and TAM22 x HS624 had Mic indices outside of the non-discount range. As expected, there was a consistent negative correlation between Mic and UHML, which was particularly strong for TAM22 x ELS33 (Table 2.6). While the majority of the $F_{2:5}$ strains within TAM22 x ELS33 selected for low UHML had Mic indices greater than 4.9, only two $F_{2:5}$ strains selected for high UHML had Mic indices less than 3.5. Therefore, the strong negative correlation between UHML and Mic should not deter the improvement of fiber UHML within this parental combination. Conversely, several of the $F_{2:5}$ strains derived from TAM22 x SID84 and the majority of $F_{2:5}$ strains derived from ELS33 x SID84 had Mic indices below 3.5, likely resulting from the interspecific nature of the SID84 parent. Thus, the negative correlation between UHML and Mic may impede selection for high UHML within interspecific populations having SID84 as a parent. The associations of the remaining fiber properties (i.e., Str, UI, and Elon) with Mic were inconsistent across genetic backgrounds.

The negative association between lint percent (LP), a primary lint yield component, and fiber quality traits is a major constraint to cultivar development in upland cotton (Hinze et al., 2011; Meredith, 1984; Miller et al., 1958; Smith and Coyle, 1997). Fiber UHML and UI were negatively correlated with LP across all genetic backgrounds in this study; Str was negatively correlated with LP across all genetic backgrounds, except for HS624 x ELS33; and the association between Elon and LP was inconsistent across genetic backgrounds. Despite the consistent negative relationship, there was not complete dependence between any of the fiber quality traits and LP. Increasing the number of progeny evaluated should enable the simultaneous improvement of fiber quality and LP. In this study, several $F_{2:5}$ strains were identified that combined beneficial alleles for fiber quality and LP. The parental combination, HS624 x ELS33, provided a unique opportunity to combine alleles for LP and fiber Str from HS624 with alleles for UHML from ELS33 (Figure 2.5). Three $F_{2:5}$ strains were identified with improved fiber UHML relative to HS624 with no compensatory reduction in LP or Str. Within the same parental combination, eight $F_{2:5}$ strains were identified with improved Str and LP in relation to ELS33 with UHML $\geq$ 33.5 mm. Thirteen $F_{2:5}$ strains derived from TAM22 x SID84 had improved UHML and Str relative to TAM22 with no compensatory loss in LP (Figure 2.6), and six strains derived from TAM22 x ELS33 had improved Elon and LP in relation to ELS33 with UHML $\geq$ 33.5 (Figure 2.7). Lastly, one $F_{2:5}$ strain derived from ELS33 x SID84 had UHML and Str greater than either parental line with LP equivalent to ELS33.

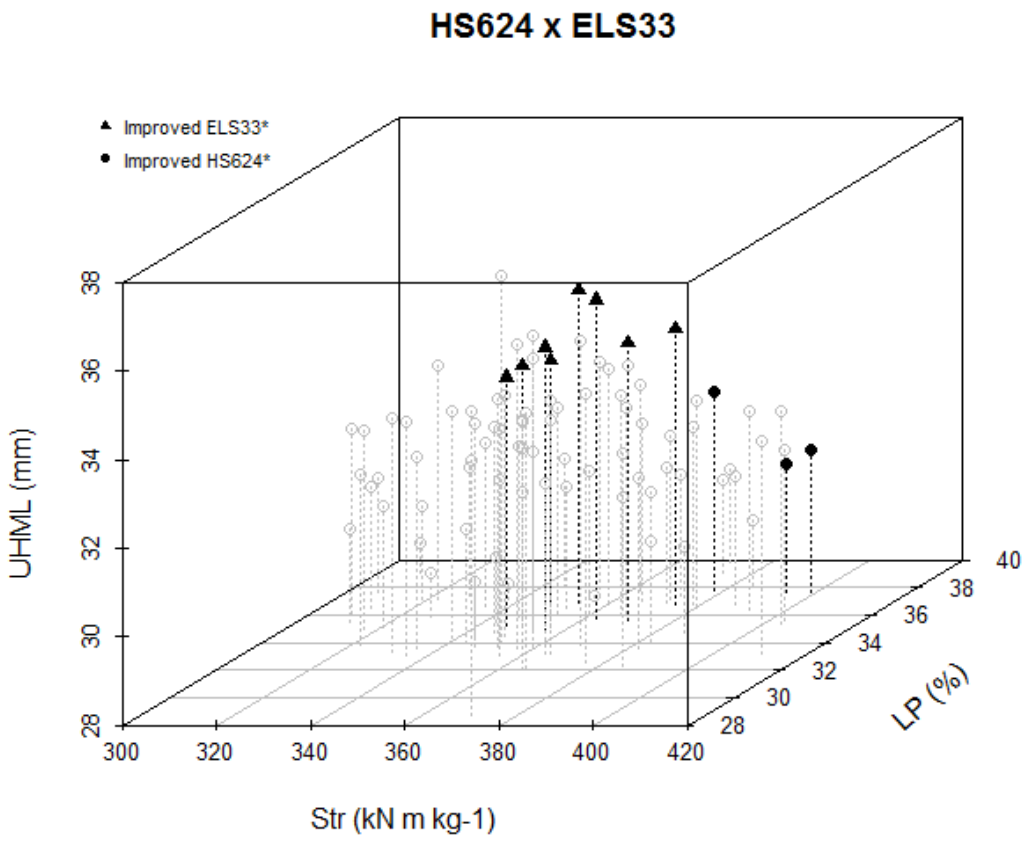

Figure 2.5. Scatter plot of fiber UHML, Str, and LP measured on $F_{2:5}$ strains derived from HS624 x ELS33 grown at College Station, TX in 2014. Improved ELS33 refers to $F_{2:5}$ strains with improved Str and LP in relation to ELS33 with UHML ≥ 33.5 mm. Improved HS624 refers to $F_{2:5}$ strains with improved fiber UHML in relation to HS624 with no compensatory reduction in fiber Str or LP.

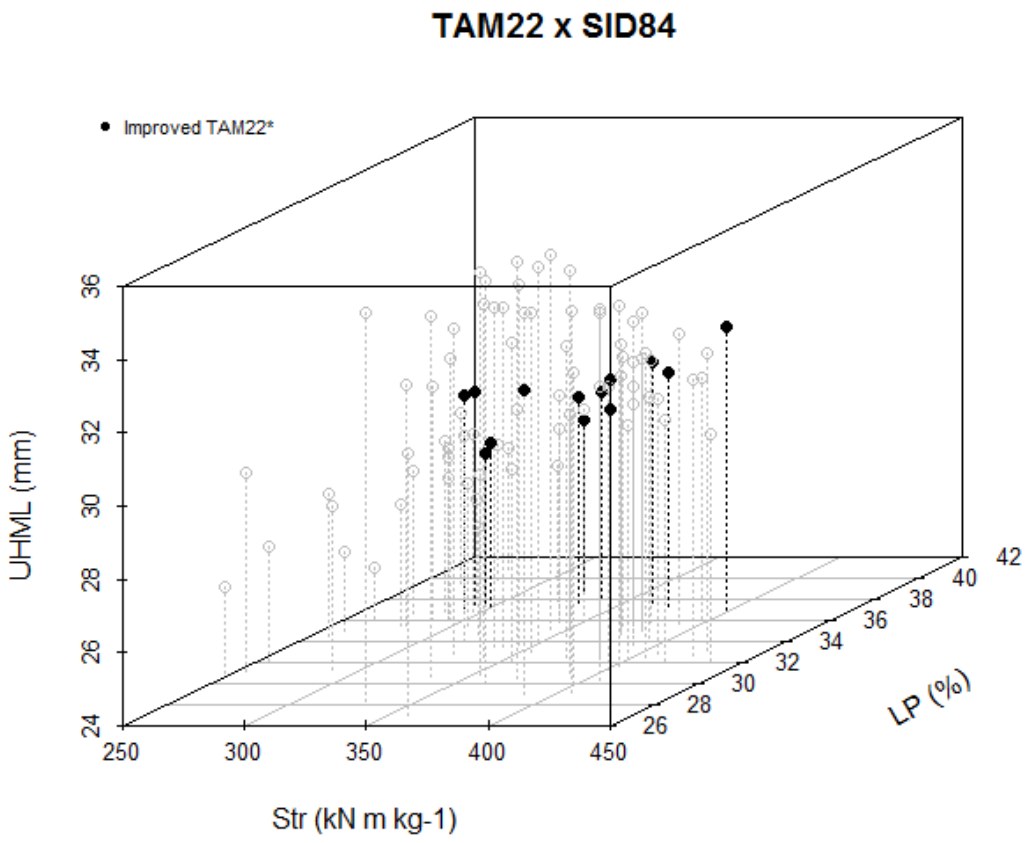

Figure 2.6. Scatterplot of fiber UHML, Str, and LP measured on $F_{2:5}$ strains derived from TAM22 x SID84 grown at College Station, TX in 2014. Improved TAM22 refers to $F_{2:5}$ strains with improved fiber UHML and Str in relation to TAM22 with no compensatory reduction in LP.

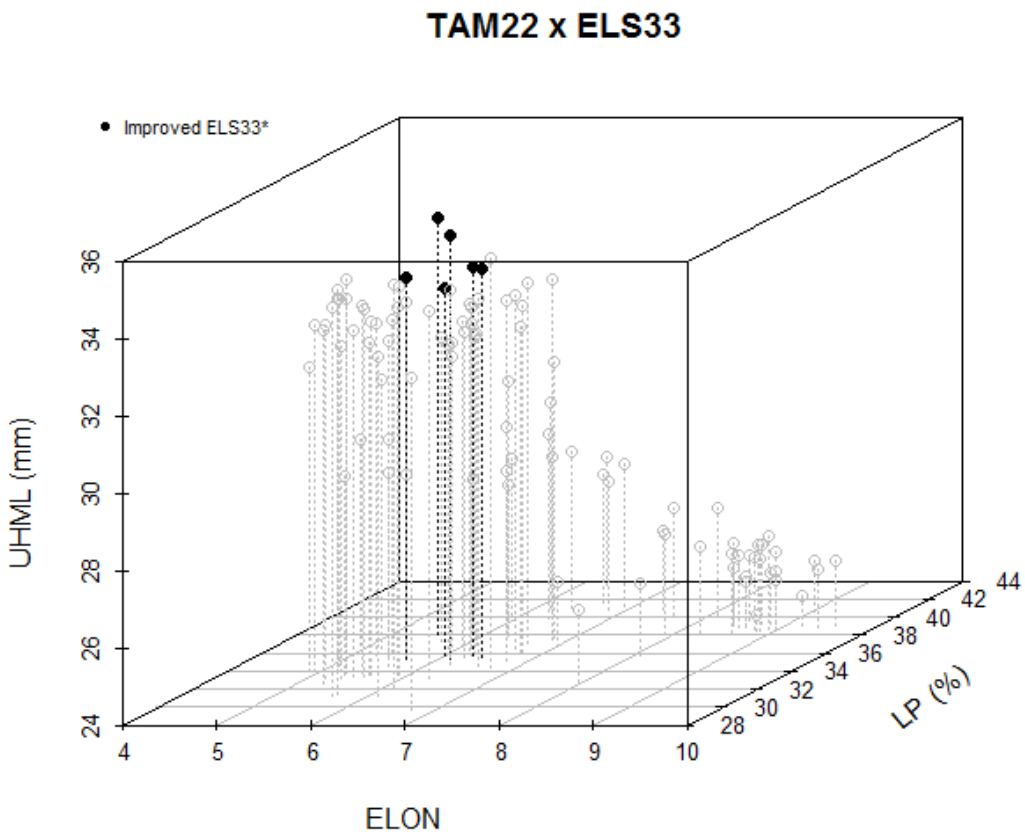

Figure 2.7. Scatterplot of fiber UHML, Elon, and LP measured on $F_{2:5}$ strains derived from TAM22 x ELS33 grown at College Station, TX in 2014. Improved ELS33 refers to $F_{2:5}$ strains with improved Elon and LP in relation to ELS33 with UHML ≥ 33.5 mm.

3. IDENTIFICATION OF MICROSATELLITE MARKERS FOR STABLE FIBER LENGTH AND BUNDLE STRENGTH QTL IN *GOSSYPIUM* SPP.

3.1. Literature Review

Advances in genetic marker and sequencing technologies have enabled more in-depth study into the genetic basis of quantitative traits and led to the development and application of novel molecular genetic approaches for the improvement of complex traits. Genetic mapping of QTL is a frequently used method to investigate quantitative trait variation at the molecular level. A QTL is a chromosomal region contributing to the expression of a quantitative trait that can be detected through statistical association between genetic marker loci and phenotypic variation. QTL mapping relies on linkage disequilibrium (LD), or the non-random association of alleles at different loci. Physical linkage of loci is the predominant cause of LD in segregating populations derived from controlled matings (Tanksley, 1993). Thus, statistically detectable LD between alleles at a genetic marker locus and a nearby QTL serves as the basis for mapping QTL to specific genomic locations. QTL mapping not only enables the location of loci contributing to quantitative trait variation, but also the estimation of the number of genes controlling a quantitative trait and their individual contribution to variation in phenotypic expression.

Law (1967) used morphological marker loci among inter-varietal chromosome substitution lines in wheat to map QTL for grain weight, grain number, height, and tiller number. The development and use of genetic markers, such as restriction fragment

length polymorphisms (RFLPs), provided increased resolution for QTL mapping and enabled the identification of discrete genetic factors (Paterson et al., 1988). Shappley et al. (1998) demonstrated that RFLPs could be used to map QTL for agronomic and fiber traits of interest in upland cotton (*Gossypium hirsutum* L.). More recent developments in genetic marker and next generation sequencing technologies have resulted in reduced genotyping costs and improved genome coverage for genetic mapping studies. Specifically, the development of microsatellite markers (SSRs) and single nucleotide polymorphism markers (SNPs) has enabled the construction of high-density genetic maps in tetraploid cotton (Blenda et al., 2012; Hulse-Kemp et al., 2015; Yu et al. 2012). The use of genetic markers coupled with the genomic sequence of the D sub-genome diploid progenitor, *G. raimondii* (Paterson et al., 2012; Wang et al., 2012), and the recently published draft sequence of tetraploid, upland cotton (Li et al., 2015; Zhang et al., 2015) will improve our ability to dissect and understand quantitative traits such as fiber length and bundle strength at the molecular level and ultimately help facilitate germplasm improvement in cotton.

Another goal of QTL mapping is to provide useful genetic markers to assist breeders in selection. MAS integrates traditional phenotypic selection with molecular genetics. Tanksley et al. (1981) first proposed using measurable isozyme differences in tomato (*Solanum* spp.) to improve the efficiency of introgressing monogenic traits from exotic germplasm into elite cultivars. Since then, MAS has proved highly effective for introgressing monogenic traits, specifically resistance to biotic stresses, through marker-assisted backcross selection in a variety of crop species (Buerstmayr et al., 2009; Chen et

al., 2001; Colton et al., 2006; Toojinda et al., 1998; Yang and Francis, 2005). MAS is useful when the trait of interest is difficult or expensive to phenotype or is not expressed in the environment or developmental stage in which selection is performed, and for pyramiding multiple monogenic traits or multiple QTL for a single quantitative trait (Xu and Crouch, 2008). Polyploidization, domestication, and the prevalent use of elite-by-elite crosses in upland cotton has reduced the levels of genetic diversity within the cultivated gene pool (Fang et al., 2013; Hinze et al., 2012; Lacape et al., 2007; Van Esbroeck and Bowman, 1998). Additionally, a negative genetic relationship between fiber quality traits and yield components attributed to repulsion linkage has been suggested by several authors (Culp et al., 1979; Green and Culp, 1990; Meredith, 1984; Miller et al., 1958; Smith and Coyle, 1997; Ulloa, 2006). QTL mapping and MAS in cotton may provide an opportunity to identify unexploited genetic diversity and increase the efficiency of simultaneous selection for yield and fiber quality traits.

Numerous QTL mapping studies have been conducted over the past few decades to investigate the genetic basis of fiber quality traits in cotton and develop genetic markers for use in MAS. Said et al. (2013) published a meta-analysis of QTL mapping studies, nearly half of which were conducted on fiber quality traits. Among the 20 QTL mapping publications for fiber quality reviewed by Said et al., fiber length and fiber bundle strength were the most comprehensively studied traits. The meta-analysis included 151 and 132 reported QTL for fiber length and strength, respectively. Chee and Campbell (2009) also performed an extensive review of QTL mapping studies for fiber quality traits, and the majority of individual QTL reported for fiber length and strength

accounted for less than 20% of the total phenotypic variation. QTL analyses indicate that fiber length and strength are controlled by multiple genes of relatively minor effect, congruent with the findings of most traditional quantitative genetics studies (Meredith, 1984). Despite the multitude of QTL described in the literature, there are very few reports of public breeding programs utilizing MAS for the improvement of fiber quality traits. The few published accounts of QTL being utilized for MAS involve the improvement of fiber strength. Zhang et al. (2003) identified a major QTL for fiber strength on chromosome 10 originating from the *G. hirsutum* x *G. anomalum* introgression line 7235. Using MAS to introgress the QTL into three different genetic backgrounds, they observed significant increases in mean fiber bundle strength ranging from 9.12 to 19.13 kN m kg^{-1}. Guo et al. (2005) also identified two QTL for fiber strength originating from 7235. The two QTL were introgressed into the Chinese cultivar, Simian 3, and they observed a mean difference of 43.46 kN m kg^{-1} between plants homozygous for both beneficial alleles versus plants homozygous for the alternative alleles.

A major challenge concerning the use of MAS for the improvement of fiber quality traits is inconsistency across different studies regarding the number, location, and effect of QTL for fiber quality traits (Chee and Campbell, 2009). This inconsistency can be attributed to multiple factors, such as the use of different experimental populations, growing environments, and genetic markers. Differences in population size also can introduce a substantial amount of variation in QTL mapping. Beavis (1998) demonstrated that minor effect QTL fail to be detected up to 97 % of the time in small

populations (N = 100), and the effects of the detected QTL are most often overestimated. As population size increases, this effect (i.e., the Beavis effect) lessens. Differences in experimental design and statistical analysis also introduce variation across studies. Studies suggest that QTL x environment interactions play an important role in the phenotypic expression of fiber length and strength (Lacape et al., 2010; Paterson et al., 2003; Shen et al., 2006). QTL x genetic background interactions (e.g., interlocus interactions or epistatic effects) also play a large role in the variation observed across mapping studies (Holland et al., 1997; Liao et al., 2001; Shen et al., 2006). Due to the low levels of genetic diversity among upland cotton germplasm, a majority of mapping studies have been conducted using interspecific populations, most commonly derived from crosses between *G. hirsutum* x *G. barbadense*. In the recent review by Said et al. (2013), 14 of the 20 mapping studies for fiber quality QTL were conducted among interspecific populations. QTL discovered in interspecific populations may have little utility within upland cotton breeding programs. The beneficial allele at a particular QTL may not be present within upland germplasm, and even if the alleles are present, they may not be segregating. Moreover, attempts to improve upland cotton through the introgression of fiber quality traits from *G. barbadense* have proved challenging due to skewed chromatin transmission and the elimination of donor alleles (Jiang et al., 2000; Stephens, 1949).

Early QTL mapping studies in cotton generally utilized small, interspecific, bi-parental populations consisting of 100-200 F_2 or $F_{2:3}$ progeny (Kohel et al., 2001; Mei et al., 2004; Ulloa et al., 2000). Many of the recent QTL mapping studies for fiber quality

traits have utilized larger upland intraspecific populations, replicated experimental designs, multiple testing locations, and more complex mating designs to improve the accuracy and reliability of QTL estimates (Cai et al., 2014; Fang et al., 2014; Lacape et al., 2010; Qin et al., 2008). Statistical approaches to QTL detection have progressed from using single-marker analysis, which analyzes the effect of each genetic marker separately, to methods such as composite interval mapping, which combine interval mapping with multiple regression to model the combined effects of genetic markers and more accurately locate QTL (Doerge, 2002). Shen et al. (2006) evaluated a recombinant inbred line (RIL) population at two locations for two years. Utilizing a mixed-model QTL mapping approach they were able to identify environment-specific QTL for fiber quality traits, as well as stable QTL expressed across locations and years (Shen et al., 2006). QTL mapping methods have also advanced to achieve broader inference across genetic backgrounds. For example, association mapping was designed to take advantage of historical recombination to obtain a greater number of segregating alleles, capture wider genetic diversity, and achieve higher mapping resolution (Yu and Buckler, 2006). Cai et al. (2014) conducted an association study of fiber quality traits using a panel of 99 diverse upland cotton (*G. hirsutum*) cultivars and accessions. They identified 70 significant SSR – trait associations, 36 of which coincided with previously reported QTL and explained between 4.60 and 22.52 % of the phenotypic variation in fiber length and between 5.64 and 18.75 % of the variation in fiber strength.

The use of MAS for the improvement of fiber quality traits in public breeding programs is still uncommon even despite advances in QTL detection methods, and still

the number of QTL mapping studies greatly outweighs the number of studies evaluating MAS (Xu and Crouch, 2008). The identification of QTL is only the first of many steps necessary for the incorporation of MAS into a breeding program. Due to the quantitative nature of fiber length and bundle strength, QTL must be independently validated in different genetic backgrounds and environments, and QTL effects must be re-estimated within the target germplasm (Heffner et al., 2009). There are a number of ways to validate a QTL, including evaluating QTL effects in different mapping populations (Castro et al., 2003; Knoll and Ejeta, 2008) or by creating near-isogenic lines by introgressing the QTL into different genetic backgrounds (Thabuis et al., 2004; Zhang et al., 2009b). Validation studies and the identification of tightly linked, portable markers associated with stable QTL are critical in the development of and evaluation of breeding strategies to incorporate MAS for fiber quality traits.

3.2. Objectives

- Evaluate the effects of previously reported SSRs for fiber UHML and Str QTL in three diverse genetic backgrounds.
- Investigate the efficiency of MAS for fiber quality utilizing SSR markers for stable UHML and Str QTL.

3.3. Materials and Methods

Plant material. Three AgriLife experimental populations were selected for the study based on pedigree and fiber quality characteristics, including two upland intraspecific populations and one interspecific population. The two upland intraspecific (*G. hirsutum* x *G. hirsutum*) populations were derived from crosses between AgriLife

experimental lines, 04WL-19 x 09207 and 09 PP-03-02 x 09917, all internal, unreleased breeding lines. The parental lines of 04WL-19 were 'Acala 1517-99' (Cantrell et al., 2000; PI 612326), TAM 96WD-18 (Thaxton et al., 2005; PI 635879), TAM 91C-95Ls (Smith et al., 2001; PI 614952), and TAM 94L-25 (Smith, 2003; PI 631440). The experimental line, 09207, was derived from a cross between HAR U 585-12 (PI 529381), an accession collected from Cote D'Ivoire, and TAM B182-33 ELS (Smith et al., 2009; PI 654362). The experimental line, 09 PP-03-02, was derived from a cross between 03 HIL B147-23 and 03 HIL B182-34, both unreleased AgriLife experimental lines with exceptional fiber length. The experimental line, 09917, was derived from a cross between TAM 96WD-18 and 'Tamcot 73' (Smith et al., 2011; PI 662044). The interspecific population was derived from a cross between TMC-9-2, an unreleased experimental *G. tomentosum* x *G. mustelinum* introgression line, and the upland cultivar LA 887 (Jones et al., 1991; PI 547084).

Field study. The three populations were grown at the AgriLife Research Farm near College Station, TX on a Weswood silt loam, a fine-silty, mixed, thermic Fluventic Ustochrept, integrated with Ships clay, a very fine, mixed, thermic Udic Chromustert. Standard cultural practices for cotton production in central Texas were conducted, including pesticide and herbicide applications and furrow irrigation. Nineteen rows (13.1 m x 1 m) of each F_3 population were planted on April 22, 2013, and plants within each row were thinned to a density of approximately one plant per 0.4 m. A total of 731 F_3 IPs were selected randomly in 2013 and tagged in the field, specifically 269 plants within 04WL-19 x 0927, 243 plants within 09PP-03-02 x 09917, and 219 plants within TMC-9-

2 x LA 887. Fifteen fully developed open bolls were harvested from each tagged plant in late October 2013 and ginned on a 10-saw laboratory gin without lint cleaners. Fiber properties, including UHML and Str, were determined at the FBRI in Lubbock, TX using the HVI system.

Approximately half of the harvested F_3 IPs were selected randomly and planted as $F_{3:4}$ progeny rows (9 m x 1 m) on May 6, 2014. One hundred and thirty-one plants were selected within 04WL-19 x 0927, 120 plants within 09PP-03-02 x 09917, and 108 plants within TMC-9-2 x LA 887. Thirty-boll samples were hand harvested from each $F_{3:4}$ progeny row in late October 2014. First- and second-position bolls from the middle of the fruiting zone were preferentially harvested to minimize variation due to environmental factors. Seed cotton samples were ginned on a laboratory saw-gin without lint cleaners, and fiber properties determined at the FBRI in Lubbock, TX using the HVI system.

Genotyping. Young leaf tissue was collected during the summer of 2013 from each of the 731 tagged F_3 IPs, and nuclear DNA was extracted using a modified CTAB (cetytrimethylammonium bromide) method described by Zhang et al. (2010). All genotyping was conducted at the USDA-ARS Cotton Fiber Bioscience Research Unit at the Southern Regional Research Center in New Orleans, LA. Seeds of five of the six parental lines were not available. Therefore, DNA was extracted from the most recent parental lines of 04WL-19, 09207, 09 PP-03-02, and 09917 to screen for polymorphisms. No germplasm was available to represent the parental line, TMC-9-2.

The DNA samples from the parental lines and LA 887 were screened with 536 SSR primer pairs selected from 31 published fiber quality QTL mapping studies (Table 3.1).

Table 3.1. Publications from which the microsatellite markers (SSRs) used to identify stable QTL for fiber quality were selected.

Publication	Type of population	Publication	Type of population
Cai *et al.*, 2014	Intraspecific[†]	Shen *et al.*, 2006	Intraspecific
Chen *et al.*, 2009	Intraspecific	Shen *et al.*, 2007	Intraspecific
Fang *et al.*, 2014	Intraspecific	Su *et al.*, 2013	Interspecific
Frelichowski *et al.*, 2006	Interspecific	Tan *et al.*, 2015	Intraspecific
Gore *et al.*, 2014	Interspecific	Wang *et al.*, 2006	Intraspecific
He *et al.*, 2007	Interspecific	Wang *et al.*, 2013	Interspecific
Islam *et al.*, 2014	Intraspecific	Yu *et al.*, 2013a	Interspecific
Lacape *et al.*, 2005	Interspecific	Yu *et al.*, 2013b	Interspecific
Lacape *et al.*, 2010	Interspecific	Zeng *et al.*, 2009	Interspecific
Li *et al.*, 2013	Interspecific	Zhang *et al.*, 2003	Intraspecific
Lin *et al.*, 2005	Interspecific	Zhang *et al.*, 2005	Intraspecific
Park *et al.*, 2005	Interspecific	Zhang *et al.*, 2009a	Intraspecific
Qin *et al.*, 2008	Intraspecific	Zhang *et al.*, 2012	Intraspecific
Rong *et al.*, 2004	Interspecific	Zhang *et al.*, 2013	Intraspecific
Said *et al.*, 2013	Both	Zhiyuan *et al.*, 2014	Intraspecific
Shen *et al.*, 2005	Intraspecific		

[†]Intraspecific refers to populations derived from *G. hirsutum* cultivars/accessions, and interspecific refers to populations derived from hybridization between *G. hirsutum* and another tetraploid *Gossypium* species.

Two-hundred and twenty-three of the SSRs showed polymorphism between the parental lines and were used to genotype the F_3 IPs. Primer sequences for the selected SSR markers were obtained from CottonGen (https://www.cottongen.org). Multiplex PCR was performed when screening the F_3 samples with the primer pairs. Forward primers were fluorescent-labeled at the 5' end with 6-FAM (6-carboxyfluorescein), HEX (4, 7, 2', 4', 5, 7-hexachloro-carboxyfluorescein), or NED (7', 8'-benzo-5-fluoro 2', 4, 7,-trichloro-5-carboxyfluorescein). Polymerase chain reaction (PCR) conditions used for

the SSR primer pairs were described by Fang et al. (2010). Amplified fragments with fluorescent labels were separated and sized by an automated capillary electrophoresis system ABI 3730XL (Applied Biosystems Inc., Foster City, CA, USA) using GeneScan-400^{TM} ROX$^{®}$ as the internal DNA standard. Allele calling was performed using GeneMapper 4.0 software (Applied Biosystems Inc., Foster City, CA, USA).

Statistical analysis. Pearson's correlation coefficient was used to measure the phenotypic correlations between fiber UHML and Str across and within the three populations. Frequently, multiple loci (i.e., more than two alleles) were amplified by a single SSR primer pair which could be the result of the primers annealing to homologous loci within sub-genomes or to homeologous loci across sub-genomes. The SSR markers could not be mapped to specific genomic locations because the parental genotypes were not available. Therefore, each SSR allele was analyzed separately, and genomic positions were estimated according to the map position(s) specified in CottonGen (https://www.cottongen.org). The majority of SSR primer pairs had been mapped to multiple chromosomes in the CottonGen database. Since the parental genotypes were not available to construct genetic maps for each population, multiple chromosome positions were listed for many of the SSRs included in this study. Alleles were scored as either "present" or "absent". Alleles having a score of $\leq 5\%$ in either the present or absent categories (i.e., rare alleles) within one or more of the three populations were excluded from the analyses to prevent spurious associations due to population structure (Fang et al., 2013). Statistical analyses were conducted using JMP (SAS Institute Inc., 2013). All

models conformed to the assumptions of ordinary least squares estimation, including the normality and homogeneity of residuals.

Initially, forward stepwise linear regression was conducted within each population to identify subsets of key alleles associated with fiber UHML and Str among F_3 IPs. This method failed to identify any common alleles having a consistent and significant association with UHML or Str across all three populations. Therefore, a two-step approach combining single-marker analysis followed by stepwise multiple linear regression was conducted to identify SSR alleles associated with stable QTL for UHML and Str (Dudley, 1993). First, single-marker analysis was performed to identify alleles associated with UHML and Str across all three genetic backgrounds. The effect of each allele on fiber UHML and Str was estimated with the following ANOVA model,

$$y_{ijk} = \mu + p_i + a_j + pa_{ij} + e_{ijk}$$

where y_{ijk} represents the observed UHML or Str of 2013 F_3 IPs, μ represents the overall mean, p_i represents the main effect of the *i*th population, a_j represents the main effect of the *j*th allele, pa_{ij} represents the interaction effect of the *i*th population and *j*th allele, and e_{ijk} represents the residual error. All terms were analyzed as fixed effects, and a less stringent probability level of $\alpha = 0.05$ was used due to the likelihood of committing Type II errors (i.e., false negative). Alleles which had a significant association with 2013 UHML and Str were selected. Selected alleles for which the population x allele interaction effect was significant were examined further, and alleles associated with opposite effects on UHML or Str within different genetic backgrounds were excluded. These steps were performed to select alleles having a significant association with 2013

UHML and Str in the same direction across all three populations. Single-marker analysis does not provide insight into the joint effects of alleles. Therefore, stepwise multiple linear regression was subsequently conducted to collectively analyze the effects of the selected alleles and to identify a core subset of alleles associated with UHML and Str across all three F_3 populations. Mixed (i.e., forward and backward) stepwise linear regression was conducted using a probability-to-enter and probability-to-leave of $\alpha = 0.05$.

The associations between SSR alleles and QTL for fiber UHML and Str identified through the two-step analysis were validated by regressing the 2014 $F_{3:4}$ progeny row UHML and Str on the core subset of SSR alleles. Alleles maintaining a significant association with UHML and Str across 2013 and 2014 were then used to group the $F_{3:4}$ progeny rows based on the number of beneficial alleles detected in the F_3 IP progenitor. ANOVA of 2014 UHML and Str based the number of beneficial alleles was conducted both across and within populations. The top 20% for UHML and Str based on phenotype of the F_3 IP progenitor were selected as control groups. ANOVA of 2014 UHML and Str based on selection method (i.e., marker-based versus phenotypic) was conducted within populations to evaluate the efficiency of marker-based selection for fiber UHML and Str using the identified subsets of alleles. All mean comparisons were conducted using Fisher's Protected Least Significant Difference (LSD).

3.4. Results and Discussion

Five-hundred and fourteen alleles were amplified by the 223 SSR primer pairs. Over half (55.45%) of the alleles had a score of ≤ 5% in the 'presence' or 'absence'

category within one or more of the populations. These rare alleles were excluded from the analysis, resulting in a total of 229 shared polymorphic alleles across the three populations. It is important to note that the results presented below are specific to the genetic diversity present across the selected populations. The effects of SSR alleles that were either monomorphic or rare among at least one of the populations were not estimated. Also please note that in all regression models, the absent allelic state was used as the base level for estimating allelic effects, and TMC-9-2 x LA887 was used as the base level for estimating population effects. Consequently, a negative estimate for an allele actually indicates that the presence of the allele was associated with a positive effect on either UHML or Str.

Fiber UHML. The distributions and summary statistics of fiber UHML within the three F_3 populations are shown in Figure 3.1. The upland intraspecific populations had greater fiber UHML than the interspecific population, with UHML ranging from 29.21 – 38.35 and 25.15 – 37.85 mm among the F_3 progeny derived from 04WL-19 x 09207 and 09 PP-03-02 x 09917, respectively, and 25.27 – 33.53 mm for F_3 progeny derived from TMC-9-2 x LA 887. Forward stepwise regression of 2013 UHML on the 229 SSR alleles within each F_3 population failed to identify a single allele having a significant association with UHML in the same direction across all three genetic backgrounds (Table 3.2). One SSR allele, $DPL1201_{275}$, was significantly associated with UHML across all three populations; however, the direction of the relationship was not consistent.

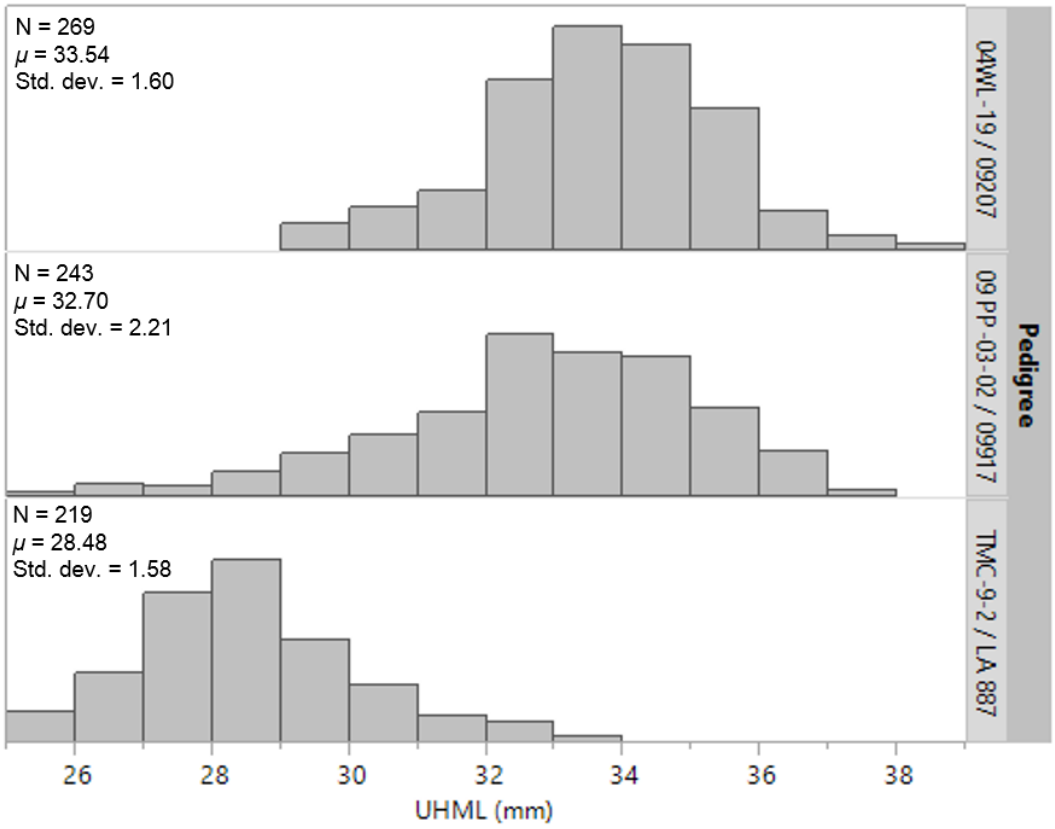

Figure 3.1. Frequency distribution of UHML in three F_3 populations grown at College Station, TX in 2013**.**

Table 3.2. Population-specific models† derived from forward stepwise linear regression of fiber UHML of F_3 plants grown at College Station, TX in 2013 using a common set of selected SSR alleles.

04WL-19 x 09207			**09 PP-03-02 x 09917**			**TMC-9-2 x LA 887**		
Term‡	**Estimate§**		**Term**	**Estimate**		**Term**	**Estimate**	
Intercept	33.814	***	Intercept	34.367	***	Intercept	29.128	***
BNL0830$_{104}$	0.554	***	CIR017$_{129}$	0.626	**	BNL1604$_{120}$	-0.368	***
BNL1604$_{98}$	0.493	***	CIR091$_{181}$	-0.791	***	BNL3545$_{115}$	-0.537	***
CGR5548$_{162}$	0.202	*	CIR246$_{168}$	1.787	***	BNL4017$_{222}$	0.567	***
CGR6383$_{223}$	-0.309	***	CIR253$_{187}$	1.036	**	CER0021$_{136}$	0.315	
CIR165$_{207}$	0.152		DPL0570$_{302}$	-0.422	***	CGR5106$_{191}$	0.256	**
CIR246$_{146}$	-0.385	***	**DPL1201$_{275}$**	**0.752**	***	CGR6170$_{208}$	0.227	**
CIR246$_{168}$	-0.354	**	JESPR050$_{218}$	0.346	*	CGR6329$_{230}$	-0.305	***
DPL0236$_{154}$	0.397	***	JESPR070$_{82}$	-1.218	*	CGR6902$_{145}$	-0.232	*
DPL0236$_{157}$	0.292	**	JESPR114$_{93}$	-1.146	**	CIR196$_{194}$	-0.375	***
DPL1201$_{275}$	**-0.242**	*	NAU1167$_{189}$	0.517	***	CIR213$_{236}$	-0.277	**

Table 3.2. *Continued.*

DPL1201$_{281}$	0.239	*	NAU2162$_{201}$	0.324	*	**DPL1201$_{275}$**	**0.427**	**
JESPR065$_{165}$	0.626	***	NAU2291$_{197}$	-0.433	*	HAU0087$_{181}$	-0.410	***
NAU1102$_{231}$	-0.402	***	NAU5046$_{226}$	-1.643	***	JESPR070$_{92}$	0.264	**
NAU2162$_{207}$	-0.201	**	SHIN1547$_{252}$	1.401	***	MUSS172$_{221}$	0.195	*
NAU2265$_{233}$	-0.162					MUSS422$_{200}$	0.281	***
TMB0382$_{179}$	0.278	***				NAU1302$_{221}$	-0.276	**
						NAU1369$_{247}$	-0.198	
						NAU2291$_{204}$	-1.030	***
						NAU2477$_{208}$	0.668	**
						NAU3308$_{220}$	-0.545	***
						SHIN1138$_{176}$	-0.223	**
						TMB0189$_{178}$	-0.558	***
						TMB1898$_{217}$	0.304	*
Adj. R^2 =	***0.48***			***0.27***			***0.61***	

*, **, *** significant at 0.05, 0.01, and 0.001 probability levels, respectively.
†Forward stepwise regression was conducted on fiber UHML within each population using a common set of 229 alleles using a p-value threshold of 0.05.
‡The 'absent' allelic state was used as the base level for regression estimates.

The presence of DPL1201$_{275}$ was positively associated with UHML in F_3 progeny derived from 04WL-19 x 09207, but negatively associated with UHML in F_3 progeny derived from 09 PP-03-02 x 09917 and TMC-9-2 x LA 887. The observed DPL1201$_{275}$ x genetic background interaction effect may be attributable to epistasis. Shen et al. (2006) reported multiple digenic epistatic interactions between fiber length QTL and unlinked background loci. Wang et al. (2006) also detected epistatic QTL for fiber length. Moreover, the additive effects and the epistatic effects of the QTL accounted for an equivalent proportion of the phenotypic variance in fiber length. Another potential explanation for the DPL1201$_{275}$ x genetic background effect may be that a recombination event occurred between the QTL for fiber UHML and the SSR during the development of one or more of the parental lines, causing DPL1201$_{275}$ to be in either repulsion or coupling linkage phase with the QTL depending on the genetic background.

The model resulting from stepwise regression of UHML among F_3 IPs derived from TMC-9-2 x LA 887 explained the greatest amount of variance (R^2 = 61%) in comparison to the derived models for the intraspecific populations with R^2 of 48% and 27% (Table 3.2).

The first approach using stepwise linear regression was unsuccessful at identifying portable SSR alleles for fiber length QTL. Therefore, a two-step process using single-marker analysis followed by stepwise regression was conducted to identify SSR alleles having a significant and consistent association with fiber UHML among the F_3 IPs. Single-marker analysis identified 98 SSR alleles that were significantly associated with 2013 UHML, 57 of which had significant allele x population effects (Table A.1). Fifty-five of the 57 alleles were associated with opposite effects on fiber UHML depending on the genetic background and were excluded from further analyses. Thus, the effects of the remaining 43 alleles associated with consistent effects on 2013 UHML were estimated using stepwise multiple linear regression. The resulting stepwise model retained 15 of the 43 SSR alleles and explained 65.44% of the total variation in UHML among the F_3 progeny (Table 3.3). The subset of 15 SSR markers originated from five different QTL mapping publications, the majority of which utilized populations derived from diverse germplasm. The publications included two mapping studies within upland intraspecific recombinant inbred lines (RILs) derived from a cross between strain 7235 and TM-1 that were evaluated in multiple environments (Shen et al., 2006; Shen et al., 2007), an association mapping study among lines derived from multiple crosses between upland and exotic cotton species, including *G. barbadense*, *G.*

tomentosum, *G. mustelinum*, and *G. darwinii* (Zeng et al., 2009), a meta-analysis of QTL mapping studies (Said et al., 2013), and a QTL mapping study using a random-mated RIL population derived from 11 diverse upland cotton cultivars (Fang et al., 2014).

Table 3.3. Regression of fiber UHML of F_3 individual plants and the resulting $F_{3:4}$ progeny rows when grown at College Station, TX based on SSR alleles identified through single-marker analysis.

Term[†]	2013 F_3 Estimate[‡]		2014 $F_{3:4}$ Estimate[‡]		Publication	Chr.[§]
Intercept	31.506	***	31.020	***		
Pedigree[04WL-19 x 09207]	1.677	***	1.140	***		
Pedigree[09 PP-03-02 x 09917]	1.288	***	0.942	***		
$BNL0830_{104}$	0.269	***	-0.022		Said *et al.* 2013	15
$BNL1604_{98}$	**0.318**	*******	**0.266**	******	**Said *et al.* 2013**	**7, 16, 17**
$BNL2986_{155}$	0.172	*	0.053		Shen *et al.* 2006; Zeng *et al.* 2009	16
$BNL4017_{234}$	**0.170**	******	**0.206**	*****	**Zeng *et al.* 2009**	**3, 14**
$CGR5548_{162}$	**0.212**	******	**0.215**	*****	**Fang *et al.* 2014**	**20**
$CIR196_{192}$	**0.184**	******	**0.203**	*****	**Zeng *et al.* 2009**	**11, 21**
$DOW067_{162}$	-0.224	**	-0.142		Fang *et al.* 2014	18
$DPL0270_{142}$	0.179	*	0.109		Fang *et al.* 2014	11
$DPL0570_{302}$	-0.131	*	-0.053		Fang *et al.* 2014	11
$JESPR050_{218}$	0.223	**	0.167		Shen *et al.* 2007	5, 19, 22, 25
$MUSS422_{207}$	-0.257	***	-0.160		Said *et al.* 2013	1, 15
$NAU0913_{206}$ (i.e., $NAU2291_{206}$)	0.154	*	0.112		Fang *et al.* 2014; Said *et al.* 2013	4, 22
$NAU1369_{247}$	**-0.170**	*****	**-0.182**	*****	**Shen *et al.* 2006; Shen *et al.* 2007**	**8, 24, 25**
$NAU2265_{233}$	-0.212	**	-0.066		Fang *et al.* 2014	2, 14
$NAU5046_{226}$	**-0.158**	*****	**-0.196**	*****	**Fang *et al.* 2014**	**5, 22**
Adj. R^2 =	***0.6544***		***0.5916***			

*, **, *** significant at 0.05, 0.01, and 0.001 probability levels, respectively.

[†]TMC-9-2 x LA 887 was used as the base level for regression estimates for population effects. The 'absent' allelic state was used as the base level for regression estimates for each allele.

[‡]The regression estimates from 2013 F_3 individual plant UHML were based on stepwise regression of alleles selected based on single-marker analysis. The regression estimates from 2014 $F_{3:4}$ progeny row UHML were based on multiple regression of the alleles identified through stepwise regression on 2013 F_3 UHML.

[§]Chromosomal assignments for each SSR correspond to the map position(s) specified in CottonGen (https://www.cottongen.org).

The associations between the 15 SSR alleles and 2013 UHML were validated by regressing the 2014 UHML data from the $F_{3:4}$ progeny rows onto the subset of 15 alleles. Less than half of the SSR alleles maintained a significant association with UHML among the $F_{3:4}$ progeny rows (Table 3.3). Loss of the statistically significant relationship between SSR alleles and 2014 UHML may be attributable to several factors, first and foremost is sample size. The population size of the $F_{3:4}$ progeny rows was approximately half of the sampled F_3 progeny. The smaller the population size, the less likely small effect QTL will be detected (Beavis, 1998). Second, QTL x environment interactions may also account for the lack of association with $F_{3:4}$ progeny row UHML. Although traditional quantitative genetics studies have shown that G x E interaction effects generally are minor regarding fiber length (Abou-El-Fittouh et al., 1969; Al-Jibouri et al., 1958; Chee and Campbell, 2009; Meredith and Bridge, 1973; Miller et al., 1958; Lacape et al., 2010), several mapping studies have reported significant QTL x environment interactions and environment-specific QTL for fiber length (Paterson et al., 1998; Shen et al., 2006; Sun et al., 2012). Third, recombination events between the SSR allele and QTL for UHML may have contributed to the loss of the significant association.

Six alleles maintained a significant association with UHML in both the F_3 and $F_{3:4}$ generations. The presence of $BNL1604_{98}$, $BNL4017_{234}$, $CGR5548_{162}$, and $CIR196_{192}$ was negatively associated with UHML, while the presence of $NAU1369_{247}$ and $NAU5046_{226}$ was positively associated with UHML. The parental genotypes were unavailable, therefore the SSR alleles could not be mapped to specific genomic

locations. The chromosome positions specified in CottonGen were used as estimates instead. Although the majority of the SSR markers had been mapped to two or more chromosomes, the results suggest that the SSR alleles were associated with six separate QTL for fiber UHML detected in all three populations as none of the six SSR markers had been mapped to the same chromosome nor to homeologous chromosomes.

The six SSR alleles were used to group the $F_{3:4}$ progeny rows into four groups, according to whether the F_3 IP progenitor had 0 – 1, 2, 3, or 4 – 6 beneficial alleles for analysis of MAS across populations. Progeny rows having 0 or 1 alleles and those having 4, 5 or 6 alleles were grouped together due to the small number of observations within those categories. ANOVA revealed that genetic background and the number of beneficial alleles had significant effects on UHML among $F_{3:4}$ progeny rows (Table 3.4).

Table 3.4. ANOVA of UHML based on the number of beneficial SSR alleles in $F_{3:4}$ progeny rows grown at College Station, TX in 2014.

Source	*df*	MS[†]		Adj. R^2
Combined				
Population	2	305.27	***	0.5900
No. alleles[‡]	3	27.97	***	
Population*No. alleles	6	3.22		
Error	347	1.98		
04WL-19 x 09207				
No. alleles[§]	4	5.51	**	0.0921
Error	126	1.28		
09 PP-03-02 x 09917				
No. alleles[§]	3	22.66	***	0.1429
Error	113	3.04		
TMC-9-2 x LA 887				
No. alleles[§]	4	6.87	**	0.1031
Error	103	1.69		

*, **, *** significant at the 0.05, 0.01, and 0.001 probability levels, respectively.

[†]Mean squares.

[‡]$F_{3:4}$ progeny rows were placed into four groups for the combined analysis according whether the F_3 progenitor had 0-1, 2, 3, or 4-6 beneficial alleles.

[§]$F_{3:4}$ progeny rows were placed into five groups for analysis within population according to whether the F_3 IP progenitor had 0-1, 2, 3, 4, or 5-6 beneficial alleles.

The interaction effect between genetic background and the number of beneficial alleles was not significant, suggesting that the two-step analysis was effective at identifying SSR alleles associated with stable QTL for UHML.

The $F_{3:4}$ progeny rows were grouped according to whether the F_3 IP progenitor had 0 – 1, 2, 3, 4, or 5 – 6 beneficial alleles to evaluate MAS within each population. One-way ANOVA of 2014 UHML demonstrated that the number of beneficial alleles had a significant effect on UHML within all three genetic backgrounds. The amount of variation in UHML among the $F_{3:4}$ progeny rows explained by the six alleles was relatively small, ranging from 9.21 – 14.29% (Table 3.4). These results suggest that the six SSR alleles are in LD with stable but minor effect QTL for fiber UHML. Another potential scenario is that the SSR alleles are not in complete LD with the QTL; therefore recombination events between the SSR marker and the QTL would reduce the accuracy of the estimated effects of the QTL but not occur frequently enough to prevent the detection of a statistical association.

The top 20% of $F_{3:4}$ progeny rows for UHML were selected based on the phenotype of the F_3 IP progenitor in order to compare the efficiency of marker-based and phenotypic selection. ANOVA of 2014 UHML revealed that selection method (i.e., marker-based versus phenotypic) had a significant effect on UHML among the $F_{3:4}$ progeny rows (Table 3.5).

Table 3.5. ANOVA of UHML among $F_{3:4}$ progeny rows based on selection method (i.e., marker-based versus phenotypic) applied to the F_3 IPs grown at College Station, TX in 2013 and 2014.

Source	*df*	MS[†]		Adj. R^2
04WL-19 x 09207				
Selection method[§]	5	9.60	***	0.1754
Error	151	1.26		
09 PP-03-02 x 09917				
Selection method[§]	4	32.33	***	0.2401
Error	137	2.66		
TMC-9-2 x LA 887				
Selection method[§]	5	14.24	***	0.2234
Error	127	1.66		

*** significant at the 0.001 probability level.

[†]Mean squares.

[§]$F_{3:4}$ progeny rows were placed into six groups for analysis within population according to whether the F_3 IP progenitor had 0-1, 2, 3, 4, or 5-6 beneficial alleles and whether the F_3 IP progenitor was in the top 20% for UHML.

The mean difference in UHML between $F_{3:4}$ progeny rows having the lowest number of beneficial alleles in the F_3 IP progenitor versus those having greatest ranged from 1.19 mm within 04WL-19 x 09207 to 2.53 mm within 09 PP-03-02 x 09917 (Table 3.6). The mean UHML of $F_{3:4}$ progeny rows having 5 or 6 beneficial alleles in the F_3 IP progenitor was not significantly different from the mean UHML of $F_{3:4}$ progeny rows derived from the F_3 IP progenitors in the top 20% for UHML for 04WL-19 x 09207 and TMC-9-2 x LA 887. Furthermore, selection of $F_{3:4}$ progeny rows with 4 beneficial alleles in the F_3 IP progenitor was equivalent to selecting the top 20% of F_3 IP progenitors for UHML within 09 PP-03-02 x 09917 and TMC-9-2 x LA 887.

Table 3.6. UHML based on marker-based and phenotypic selection among $F_{3:4}$ progeny rows within three cotton populations grown at College Station, TX in 2014.

	04WL-19 x 09207	09 PP-03-02 x 09917	TMC-9-2 x LA 887
Selection method[†]		**-- mm --**	
Marker-based			
0-1 alleles	31.85 cd[‡]	30.20 d	28.11 d
2 alleles	31.82 d	31.71 c	28.52 cd
3 alleles	32.27 cd	32.22 bc	28.80 bc
4 alleles	32.51 bc	32.81 ab	29.58 ab
5-6 alleles	33.04 ab	--	30.07 a
Phenotypic			
Top 20%	33.48 a	33.56 a	30.12 a

[†]$F_{3:4}$ progeny rows were selected based on genotype (i.e., the number of beneficial alleles in the F3 IP progenitor). $F_{3:4}$ progeny rows were also selected based on phenotype, selecting the top 20% for UHML based on the phenotype of the F_3 IP progenitors.
[‡]Means within columns followed by the same letter are not significantly different according to Fisher's protected least significant difference (LSD) at $\alpha = 0.05$.

Fiber Str. The distributions and summary statistics of fiber Str within each F_3 population are shown in Figure 3.2. The upland intraspecific populations had greater fiber Str compared to the interspecific population. The Str of F_3 progeny derived from 04WL-19 x 09207 and 09 PP-03-02 x 09917 ranged from 306.1 – 423.8 and 298.2 – 435.6 kN m kg^{-1}, respectively. While, the Str of the F_3 progeny derived from TMC-9-2 x LA 887 ranged from 271.7 – 372.8 kN m kg^{-1}. Forward stepwise linear regression within each F_3 population failed to identify any SSR alleles having a significant effect on Str in the same direction across all three genetic backgrounds (Table 3.7). Congruent with the results for fiber UHML, the set of 229 SSR alleles explained the least amount of variation in Str within the upland intraspecific population derived from 09 PP-03-02 x 09917 (18%) and the greatest amount of variation in Str among the interspecific population, TMC-9-2 x LA 887 (51%). The results suggest that published QTL for fiber quality traits may be most useful for MAS among interspecific populations.

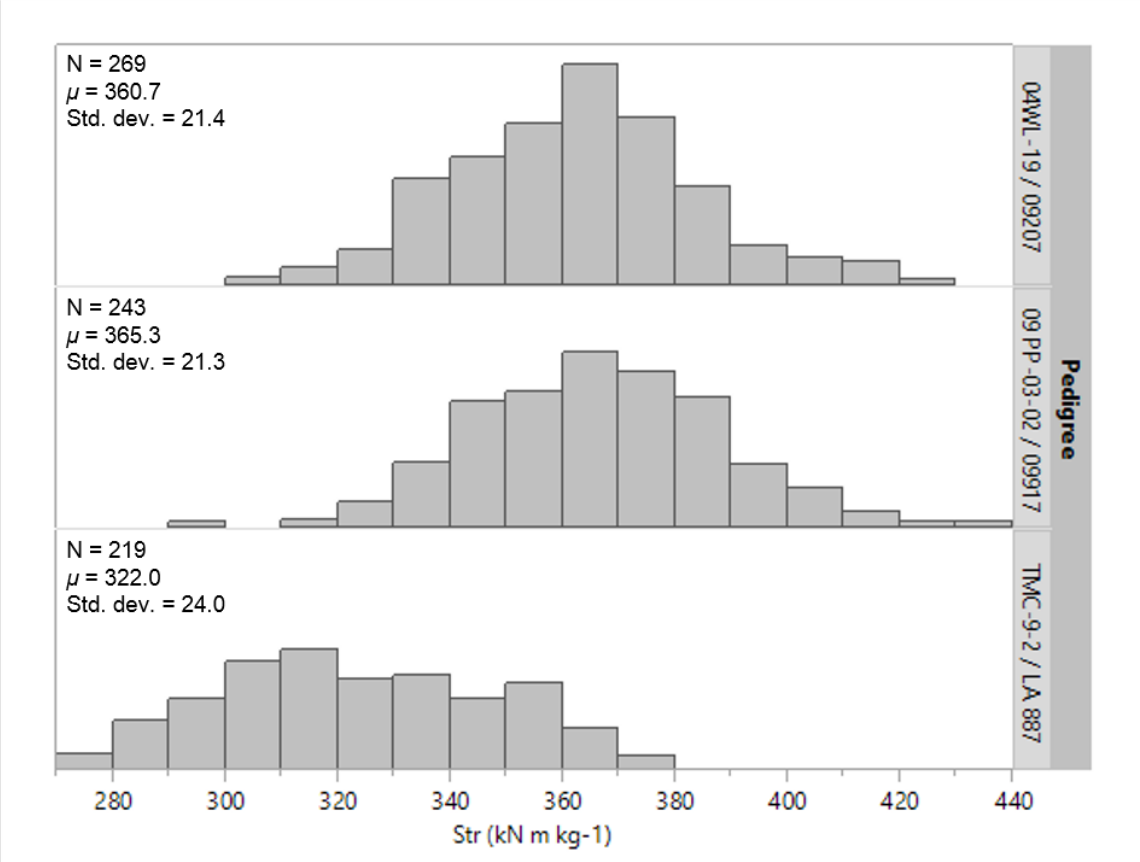

Figure 3.2. Frequency distribution of Str in three F_3 populations grown at College Station, TX in 2013.

Table 3.7. Population-specific models† derived from forward stepwise linear regression of fiber Str of F_3 plants grown at College Station, TX in 2013 using a common set of selected SSR alleles.

04WL-19 x 09207			09 PP-03-02 x 09917			TMC-9-2 x LA 887		
Term‡	**Estimate§**		**Term**	**Estimate**		**Term**	**Estimate**	
Intercept	38.023	***	Intercept	37.414	***	Intercept	32.729	***
$BNL1604_{98}$	0.940	***	$BNL1604_{98}$	0.298		$BNL1227_{186}$	-1.004	***
$BNL2495_{195}$	0.426	**	$BNL2599_{96}$	0.350	**	$BNL3031_{184}$	-0.662	***
$BNL2986_{155}$	-0.435	***	$CGR5139_{182}$	-0.469	***	$BNL3280_{212}$	-0.420	**
$BNL3031_{184}$	0.539	***	$DOW067_{162}$	-0.280	*	$BNL3410_{222}$	-0.373	*
$BNL3085_{106}$	-0.305	**	$DPL1379_{168}$	0.975	**	$BNL3545_{115}$	-0.370	
$BNL3452_{180}$	0.249	*	$JESPR295_{105}$	-0.486	**	$BNL4017_{222}$	1.032	***
$BNL3545_{183}$	-0.544	***	$NAU0913_{206}$	0.238		$CER0021_{136}$	0.727	*
$BNL3558_{210}$	0.371	*	$SHIN1138_{181}$	-0.376	*	$CGR6329_{232}$	0.438	*
$CGR5548_{171}$	-0.504	***				$CGR6383_{217}$	0.304	*

Table 3.7. *Continued.*

$DC30210_{148}$	-0.502	***	$CIR246_{157}$	0.586	*
$DOW067_{162}$	-0.347	*	$CIR249_{192}$	0.423	***
$JESPR065_{165}$	-0.370	**	$DPL0236_{154}$	-0.573	*
$NAU1102_{231}$	-0.764	***	$DPL1358_{205}$	0.740	***
$NAU1369_{247}$	0.385	***	$HAU0087_{188}$	-0.669	***
$TMB0382_{179}$	-0.335	**	$HAU2022_{163}$	-0.529	***
$TMB1898_{217}$	0.255	*	$JESPR114_{86}$	-0.459	***
			$MUSS422_{207}$	-0.122	
			$NAU3308_{220}$	-0.537	**
			$NAU3393_{196}$	-0.630	***
			$TMB0382_{182}$	0.566	**
Adj. R^2 =	***0.46***	***0.18***		***0.51***	

*, **, *** significant at 0.05, 0.01, and 0.001 probability level, respectively.

†Forward stepwise regression was conducted on fiber Str within each population using a common set of 229 alleles using a p-value threshold of 0.05.

‡The 'absent' allelic state was used as the base level for regression estimates.

This observation is somewhat surprising, considering that preference was given to QTL mapping studies within upland intraspecific populations in the selection of SSR markers for the study.

Single-marker analysis identified 104 SSR alleles that were significantly associated with Str among the F_3 plants, 68 of which had significant allele x population effects (Table A.1). Sixty-two of the 68 alleles were associated with opposite effects on Str depending on the genetic background and were excluded from further analyses. Stepwise linear regression of fiber Str among the F_3 plants was conducted based on the resulting subset of 42 SSR alleles associated with consistent effects across the three populations. The resulting model retained 17 alleles having a significant association with fiber Str among the F_3 plants and explained 53.48% of the total variation in Str (Table 3.8).

Table 3.8. Regression of fiber Str of F_3 individual plants and the resulting $F_{3:4}$ progeny rows when grown at College Station, TX based on SSR alleles identified through single-marker analysis.

Term[†]	2013 F_3 Estimate[‡]		2014 $F_{3:4}$ Estimate[‡]		Publication	Chr.[§]
Intercept	348.73	***	336.14	***		
Pedigree[04WL-19 x 09207]	13.23	***	6.61	**		
Pedigree[09 PP-03-02 x 09917]	15.37	***	8.88	***		
$BNL0830_{104}$	3.02	***	1.27		Said *et al.* 2013	15
$BNL1122_{166}$	-2.37	**	0.60		Shen *et al.* 2005; Shen *et al.* 2007; Zeng *et al.* 2009; Cai *et al.* 2014	7, 16
$BNL1604_{120}$	-2.76	*	-0.80		Said *et al.* 2013	7, 16, 17
$BNL1604_{98}$	**3.63**	****	**3.97**	*****	**Said *et al.* 2013**	**7, 16, 17**
$BNL2599_{96}$	2.83	***	0.97		Fang *et al.* 2014	1
$BNL3280_{213}$	2.05	*	-0.89		Zang *et al.* 2005	18, 20
$CGR6329_{232}$	**2.43**	****	**2.71**	*****	**Fang *et al.* 2014**	**26**
$CIR249_{192}$	2.74	**	1.96		Zeng et al. 2009	4
$DPL0236_{157}$	**1.91**	*****	**2.80**	****	**Fang *et al.* 2014**	**NA**
$DPL1358_{205}$	3.14	***	1.86		Fang *et al.* 2014	11
$JESPR050_{218}$	2.17	*	0.74		Shen *et al.* 2007	5, 19, 22, 25
$JESPR295_{105}$	-2.94	***	-1.87		Cai *et al.* 2014	12, 26
$NAU0913_{203}$ (i.e., $NAU2291_{203}$)	-3.56	***	-0.47		Fang *et al.* 2014; Said *et al.* 2013	4, 22
$NAU1102_{231}$	**-2.21**	*****	**-2.85**	*****	**Cai *et al.* 2014**	**19**
$NAU1369_{247}$	**2.71**	*******	**2.69**	****	**Shen *et al.* 2006; Shen *et al.* 2007**	**8, 24, 25**
$TMB0382_{179}$	**-1.96**	*****	**-2.48**	*****	**Tan *et al.* 2015**	**23**
$UCD120\_{266}$	-1.62	*	-0.73		Fang *et al.* 2014	22
Adj. R2 =	***0.5348***		***0.3747***			

*, **, *** significant at 0.05, 0.01, and 0.001 probability levels, respectively.

[†]TMC-9-2 x LA 887 was used as the base level for regression estimates for population effects. The 'absent' allelic state was used as the base level for regression estimates for each allele.

[‡]The regression estimates from 2013 F_3 individual plant Str were based on stepwise regression of alleles selected based on single-marker analysis. The regression estimates from 2014 $F_{3:4}$ progeny row Str were based on multiple regression of the alleles identified through stepwise regression on 2013 F_3 Str.

[§]Chromosomal assignments for each SSR correspond to the map position(s) specified in CottonGen (https://www.cottongen.org).

The set of 17 SSR markers originated from the five publications listed in the previous section on fiber UHML (Fang et al., 2014; Said et al., 2013, Shen et al., 2006; Shen et al., 2007; Zeng et al., 2009) plus three additional publications, including an association analysis of fiber quality traits among 99 upland cotton cultivars and accessions (Cai et al., 2013) and two QTL analyses within upland intraspecific bi-parental populations, Yumian 1 x T586 (Zhang et al., 2005) and CCRI 35 x Yumian 1 (Tan et al., 2015).

The associations between the 17 alleles and fiber Str were validated through multiple regression of fiber Str obtained on the $F_{3:4}$ progeny rows in 2014. Only six of the SSR alleles maintained a significant association to 2014 Str (Table 3.8). The presence of $BNL1604_{98}$, $CGR6329_{232}$, $DPL0236_{157}$, and $NAU1369_{247}$ was negatively associated with fiber Str, while the presence of $NAU1102_{231}$ and $TMB0382_{179}$ positively associated with Str. $DPL0236_{157}$ had not been mapped to a linkage group or chromosome position in the original publication (Fang et al., 2014) nor within CottonGen. Among the remaining SSR markers, no two had been mapped to the same chromosome nor to homeologous chromosomes; thus, it is plausible that the five markers were linked to separate QTL for fiber Str present in each of the three populations.

The six alleles maintaining significant associations with fiber Str across 2013 and 2014 were used to group the $F_{3:4}$ progeny rows based on the number of beneficial alleles in the F_3 IP progenitor. The $F_{3:4}$ progeny rows were grouped according to whether the F_3 IP progenitor had 0 – 3 or 4 – 6 beneficial alleles for the evaluation of MAS across populations. Two groups were specified instead of five due to missing data on $F_{3:4}$ rows

having 0, 1, 5, and 6 beneficial alleles within one or more of the populations. Genetic background and the number of beneficial alleles in the F_3 IP progenitor were significantly associated with fiber Str (Table 3.9). The interaction effect between population and the number of beneficial alleles was not significant, indicating that the two-step analysis was effective at identifying SSR alleles associated with stable QTL for fiber Str.

Table 3.9. ANOVA of Str based on the number of beneficial SSR alleles in $F_{3:4}$ progeny rows grown at College Station, TX in 2014.

Source	***df***	**MS‡**		**Adj. R^2**
Combined				
Population	2	14308.79	***	0.3251
No. alleles†	1	8241.76	***	
Population*No. alleles	2	289.87		
Error	353	325.80		
04WL-19 x 09207				
No. alleles†	4	2828.10	***	0.2137
Error	126	287.63		
09 PP-03-02 x 09917				
No. alleles†	3	616.57		0.0176
Error	116	360.42		
TMC-9-2 x LA 887				
No. alleles†	3	1036.16	*	0.0758
Error	102	267.82		

*, **, *** significant at 0.05, 0.01, and 0.001 probability levels, respectively.
†$F_{3:4}$ progeny rows were grouped according to whether the F_3 progenitor had 0-3 or 4-6 beneficial alleles for the combined analysis across populations. $F_{3:4}$ progeny rows were placed into five groups according whether the F_3 progenitor had 0-to1, 2, 3, 4, or 5-to-6 beneficial alleles for ANOVA within each population.
‡Mean squares.

The $F_{3:4}$ progeny rows were grouped according to whether the F_3 IP progenitor had 0 – 1, 2, 3, 4, or 5 – 6 beneficial alleles to evaluate MAS within each population. ANOVA of Str within each population revealed that the number of beneficial alleles was not significantly associated with Str among $F_{3:4}$ progeny rows derived from 09 PP-03-02 x 09917, explaining less than 2% of the total variation in Str (Table 3.9). The number of beneficial alleles was significantly associated with Str within the remaining two populations, explaining 7.58% and 21.37% of the variation in Str within TMC-9-2 x LA 887 and 04WL-19 x 09207, respectively. The results suggest that the six SSR alleles are associated with stable, but relatively minor effect QTL for fiber Str. It is also a possibility that the SSR markers are not in complete LD with the QTL, and as a consequence the reported QTL effects may be substantially underestimated.

The top 20% of $F_{3:4}$ progeny rows for Str were selected based on the phenotype of the F_3 IP progenitor to compare phenotypic and marker-based selection within each population. ANOVA of 2014 Str indicated that selection method had a significant effect on Str among the $F_{3:4}$ progeny rows (Table 3.10). The mean difference in Str between $F_{3:4}$ progeny rows having the lowest number of beneficial alleles in the F_3 IP progenitor versus those having greatest was 36.8 kN m kg^{-1} within 04WL-19 x 09207 and 13.8 kN m kg^{-1} within TMC-9-2 x LA 887 (Table 3.11). The mean Str of $F_{3:4}$ progeny rows having 5 or 6 beneficial alleles in the F_3 IP progenitor was not significantly different from the mean Str of $F_{3:4}$ progeny rows resulting from phenotypic selection of the top 20% of F_3 IPs for fiber Str for the upland intraspecific populations. Conversely, the mean Str of $F_{3:4}$ progeny rows resulting from phenotypic selection of the top 20% of F_3

IPs was significantly greater than the mean Str of $F_{3:4}$ progeny rows having 4 beneficial alleles in the F_3 IP progenitor.

Table 3.10. ANOVA of Str among $F_{3:4}$ progeny rows based on selection method (i.e., marker-based versus phenotypic) applied to the F_3 IPs grown at College Station, TX in 2013 and 2014.

Source	*df*	MS‡		Adj. R^2
04WL-19 x 09207				
Selection method†	5	4072.97	***	0.3009
Error	151	282.28		
09 PP-03-02 x 09917				
Selection method†	4	1886.75	***	0.1046
Error	135	372.94		
TMC-9-2 x LA 887				
Selection method†	4	2128.35	***	0.1850
Error	120	264.77		

*** significant at the 0.001 probability level.

‡Mean squares.

†$F_{3:4}$ progeny rows were placed into five groups within population according whether the F_3 progenitor had 0-to1, 2, 3, 4, or 5-to-6 beneficial alleles and whether the F_3 IP progenitor was in the top 20% for Str.

Table 3.11. Str based on marker-based and phenotypic selection of $F_{3:4}$ progeny rows within three cotton populations grown at College Station, TX in 2014.

Selection method†	04WL-19 x 09207	09 PP-03-02 x 09917	TMC-9-2 x LA 887
		--kN m kg^{-1}--	
Marker-based			
0-1 alleles	326.7 d‡	--	312.2 c
2 alleles	333.6 d	342.9 b	316.8 bc
3 alleles	345.9 c	344.6 b	324.7 b
4 alleles	349.7 bc	350.9 b	326.0 b
5-6 alleles	363.5 ab	353.2 ab	--
Phenotypic			
Top 20%	361.8 a	365.3 a	336.8 a

†$F_{3:4}$ progeny rows were selected based on genotype (i.e., the number of beneficial alleles in the F3 IP progenitor). $F_{3:4}$ progeny rows were also selected based on phenotype, selecting the top 20% for Str based on the phenotype of the F_3 IP progenitors.

‡Means within columns followed by the same letter are not significantly different according to Fisher's protected least significant difference (LSD) at $\alpha = 0.05$.

Fiber quality QTL clusters. Several SSR alleles identified through the two-step analysis were associated with both fiber UHML and Str. Numerous studies have reported the tendency for QTL of various fiber quality traits to co-localize, or form QTL 'clusters' (Lacape et al., 2010; Rong et al., 2007; Said et al., 2013). There was a significant positive phenotypic correlation between fiber UHML and Str across populations, although this association was weak to non-existent within the upland intraspecific populations (Table 3.12).

Table 3.12. Correlation between fiber UHML and Str among F_3 progeny plants and $F_{3:4}$ progeny rows grown at College Station, TX in 2013 and 2014, respectively.

	Correlation coefficient†			
Population	**F_3**		**$F_{3:4}$**	
Combined	0.6353	***	0.4939	***
04WL-19 x 09207	0.1757	**	0.0828	
09 PP-03-02 x 09917	0.2471	***	0.0886	
TMC-9-2 x LA 887	0.6035	***	0.6118	***

*, **, *** significant at 0.05, 0.01, and 0.001 probability levels, respectively.
†Pearson's correlation coefficient was calculated to estimate the phenotypic correlation between fiber UHML and Str.

Still, it is unclear whether the SSR alleles are associated with separate co-localizing QTL or single QTL with pleiotropic effects. BNL0830, JESPR050, and NAU0913 (i.e., NAU2291) were significantly associated with fiber UHML and Str among the F_3 populations but not in the $F_{3:4}$ progeny row populations, while BNL1604 and NAU1369 were significantly associated with UHML and Str across generations. The presence of $NAU1369_{247}$ was negatively associated with fiber Str but positively associated with UHML. Shen et al. (2006 and 2007) reported that NAU1369 was linked to a QTL for

fiber strength on chromosome 25 originating from upland germplasm line 7235, which contains alleles introgressed from *G. anomalum*. This QTL on chromosome 25 explained from 5.6 to 11.7% of the variation in fiber strength. Qin et al. (2008) mapped NAU1396 to chromosome 24, but did not detect any association between the SSR marker and fiber Str. Still, multiple studies suggest that chromosome 24 harbors multiple QTL for fiber strength (Chen et al., 2009; Kumar et al., 2012). In this study, the SSR markers were not used to create a genetic map due to the missing parental genotypes. Therefore, we were unable to resolve the chromosomal position NAU1396 or the associated QTL for fiber UHML and Str.

The absence of $BNL1604_{98}$ was positively associated with UHML and Str, making it a better candidate than NAU1396 for MAS concerning fiber quality. BNL1604 has been mapped to multiple linkage groups, including homeologous chromosomes 7 (Zhang et al., 2012) and 16 (Wang et al., 2011) as well as chromosome 17 (Song et al., 2005). Wang et al. (2011) reported that BNL1604 was linked to a QTL for fiber length and micronaire on chromosome 16 within an interspecific (*G. hirsutum* x *G. barbadense*) population. Mapping studies among upland intraspecific populations suggest that BNL1604 is also associated with a fiber quality QTL cluster on chromosome 7, which includes QTL for fiber length, strength, micronaire, length uniformity, and elongation (Said et al., 2013; Sun et al., 2012; Tan et al., 2015; Yu et al., 2013b; Zhang et al., 2012). Fang et al. (2014) identified two QTL clusters on chromosomes 7 and 16 that were associated with opposite effects on fiber length, strength, and fiber length uniformity. Based on these findings, they suggested that these

two clusters might represent duplicate loci with opposite effects. Selecting for favorable alleles at the two QTL clusters on chromosome 7 and 16, Fang et al. (2014) observed a significant increase in the mean UHML, Str, fiber length uniformity, and a significant decrease in the mean short fiber content. The SSR markers used for selection by Fang et al., C2-0114a (chr. 7) and CM0066a (chr. 16), were not segregating within the F_3 progeny derived from TMC-9-2 x LA 887, thus they were not included in our analyses. Yet, it is plausible that $BNL1604_{98}$ is in LD with one or both of the QTL described by Fang et al., considering BNL1604 has been mapped in close proximity to both C2-0114a and CM0066a (CottonGen.org).

In this study, selection against the deleterious allele $BNL1604_{98}$ was evaluated among the $F_{3:4}$ progeny rows across all three populations and resulted in a significant increase in UHML and Str (Table 3.13).

Table 3.13. UHML and Str based on the presence or absence of $BNL1604_{98}$ in the $F_{3:4}$ progeny rows grown at College Station, TX in 2014.

	UHML (mm)		**Str (kN m kg^{-1})**	
	$BNL1604_{98}$		**$BNL1604_{98}$**	
	absent	***present***	***absent***	***present***
Combined	31.27 a	30.73 b	339.8 a	331.2 b
04WL-19 x 09207	32.60 a	31.99 b	347.0 a	331.0 b
09 PP-03-02 x 09917	32.00 ns	31.76 ns	352.4 a	343.6 b
TMC-9-2 x LA 887	29.20 a	28.45 b	320.0 ns	318.8 ns

†Means within each row and within subheadings with the same letter are not significantly different according to Fisher's protected least significant difference (LSD) at $\alpha = 0.05$.

Selection against $BNL1604_{98}$ within 04WL-19 x 09207 significantly increased the mean UHML by 0.61 mm and Str by 16.0 kN m kg^{-1} and explained 5.38% and 15.38% of the

variation in UHML and Str, respectively. Selection against BNL1604$_{98}$ did not significantly increase UHML among $F_{3:4}$ progeny derived from 09 PP-03-02 x 09917, but it did have a significant effect on Str, increasing the mean Str by 8.8 kN m kg^{-1} (R^2 = 4.21%). Conversely, selection against BNL1604$_{98}$ within $F_{3:4}$ progeny derived from TMC-9-2 x LA 887 did not have a significant effect on fiber Str but did significantly increase the mean UHML by 0.75 mm (R^2 = 5.24%). BNL1604$_{98}$ appears to have the greatest utility in selection for fiber quality within 04WL-19 x 09207, explaining a relatively large portion of the total variation in fiber Str. The improvement observed through MAS for UHML and Str based on the single locus, BNL1604$_{98}$, was less than that reported by Fang et al. (2014).

4. CONCLUSIONS*

Low genetic diversity among elite upland cotton germplasm is a challenge concerning the improvement of fiber quality. The results obtained from divergent selection for UHML and Str correspond to those of Joy (2014) and provide further support that there is abundant genetic variation regarding UHML and Str among the following parental combinations: HS624 x ELS33, TAM22 x ELS33, ELS33 x SID84, HS624 x TAM22, and TAM22 x SID84. The realized heritability estimates suggest that the dominance effects reported by Joy for these parental combinations, with the exception of HS624 x ELS33, should not restrict early generation selection for UHML and Str at College Station. Evaluation and selection for UHML and Str should be conducted across multiple environments and generations to obtain stable fiber quality among lines derived from HS624 x ELS33.

The negative relationship between fiber quality and lint yield serves as another challenge regarding the genetic improvement of fiber quality. High levels of allele dispersion for UHML and Str were observed between SID84 and the upland parents, TAM22 and ELS33, but negative correlations between agronomic properties and fiber quality traits still limit the use of SID84 as a parental line. The best parental combination

*Part of the data reported in this section is reprinted with permission from Joy, K.S. 2014. Inheritance of cotton fiber length and strength. Ph.D. diss., Texas A&M University, College Station, TX. Copyright 2014, Kolbyn S. Joy.

for the simultaneous improvement of fiber quality and LP was HS624 x ELS33. Despite relatively low realized heritability and minor responses to selection for this parental combination, several $F_{2:5}$ strains with simultaneously improved LP, UHML and Str were identified, providing further evidence that the negative relationship between fiber quality and lint yield is attributable to repulsion phase linkage.

QTL identification and MAS may help identify unexploited genetic diversity and also provide a more efficient means for the simultaneous selection for yield and fiber quality traits. Hundreds of QTL for fiber quality traits have been described in the literature, but there is generally inconsistency regarding the genomic location and effect of individual QTL. Thus, QTL validation studies are necessary to enable the identification of fiber quality QTL that are stably expressed across genetic backgrounds and environments. Forward stepwise regression of fiber UHML and Str within three experimental populations, 04WL-19 x 09207, 09 PP-03-02 x 09917 and TMC-9-2 x LA 887, based on a common set of 229 segregating SSR alleles revealed the importance of genetic background effects in QTL mapping. The results demonstrate either that epistasis plays a substantial role in the phenotypic expression of fiber UHML and Str, or that the majority of SSR markers utilized in this study are not in tight enough LD with QTL for fiber UHML and Str to be portable across genetic backgrounds.

Despite the challenges in comparing QTL mapping results across different studies, six SSR alleles associated with stable QTL for fiber UHML and six alleles associated with stable QTL for fiber Str using the results of 31 published QTL mapping studies were identified. $BNL1604_{98}$ was identified as the best candidate for MAS, and

selection against the allele resulted in the simultaneous improvement of UHML and Str across three genetically diverse populations. The results also suggest that QTL studies which utilize diverse genetic backgrounds are more effective at identifying genetic markers linked to stable QTL for UHML and Str compared to QTL mapping within bi-parental populations (Cai et al., 2014; Fang et al., 2014; Said et al., 2013; Zeng et al., 2009). The SSR alleles associated with stable QTL for fiber UHML and Str explained only a minor proportion of the total phenotypic variance, yet marker-based selection for plants having five or six beneficial alleles for either UHML or Str was largely equivalent to phenotypic selection for the top 20%. There was only one exception, where phenotypic selection for the top 20% for Str within TMC-9-2 x LA 887 was more effective than marker-based selection. Lastly, these results provide further evidence that fiber UHML and fiber Str are quantitative traits controlled by a complex network of interacting genes of relatively small effect (Meredith, 1984).

REFERENCES

Abou-El-Fittouh, H.A., J.O. Rawlings, and P.A. Miller. 1969. Genotype by environment interactions in cotton—Their nature and related environmental variables. Crop Science 9: 377—381.

Al-Jibouri, H.A., P.A. Miller, and H.F. Robinson. 1958. Genotypic and environmental variances and covariances in an upland cotton cross of interspecific origin. Agronomy Journal 50: 633—636.

Al-Rawi, K.M. and R.J. Kohel. 1970. Gene action in the inheritance of fiber properties in intervarietal diallel crosses of upland cotton, *Gossypium hirsutum* L. Crop Science 10: 82—85.

Ali, M.A., I.A. Khan, S.I. Awan, S. Ali, and S. Niaz. 2008. Genetics of fibre quality traits in cotton (*Gossypium hirsutum* L.). Australian Journal of Crop Science 2: 10—17.

Baker, J.L. and L.M. Verhalen. 1973. The inheritance of several agronomic and fiber properties among selected lines of upland cotton, *Gossypium hirsutum* L. Crop Science 13: 444—450.

Beavis, W.D., 1998. QTL analyses: power, precision, and accuracy. *In* Paterson, A.H. (ed.) Molecular Dissection of Complex Traits. CRC Press, New York, pp. 145—162.

Bhortakke, M.K., T. Nishimura, and T. Matsuo. 1997. High speed yarn production with air-jet apinning: effects of some fiber parameters. Textile Research Journal 67: 101—108.

Blenda, A., D.D. Fang, J.F. Rami, O. Garsmeur, F. Luo, and J.M. Lacape. 2012. A high density consensus genetic map of tetraploid cotton that integrates multiple component maps through molecular marker redundancy check. PLoS One 7: e45739.

Bourland, F.M., and D.C. Jones. 2012. Registration of ‘UA48’ cotton cultivar. Journal of Plant Registrations 6: 15—18.

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

Buerstmayr, H., T. Ban, and J.A. Anderson. 2009. QTL mapping and marker-assisted selection for *Fusarium* head blight resistance in wheat: a review. Plant Breeding 128: 1—26.

Cai, C., W. Ye, T. Zhang, and W. Guo. 2014. Association analysis of fiber quality traits and exploration of elite alleles in Upland cotton cultivars/accessions (*Gossypium hirsutum* L.). Journal of Integrative Plant Biology 56: 51—62.

Campbell, B.T., P.W. Chee, E.L. Lubbers, D.T. Bowman, W.R. Meredith, J. Johnson, and D.E. Fraser. 2011. Genetic improvement of the Pee Dee cotton germplasm collection following seventy years of plant breeding. Crop Science 51: 955—968.

Campbell, B.T., P.W. Chee, E.L. Lubbers, D.T. Bowman, W.R. Meredith, J. Johnson, D. Fraser, W. Bridges, and D. Jones. 2012. Dissecting genotype x environment interactions and trait correlations present in the Pee Dee cotton germplasm collection following seventy years of plant breeding. Crop Science 52: 690—699.

Campbell, B.T., J.K. Greene, J. Wu, and D. Jones. 2014. Assessing the breeding potential of day-neutral converted racestock germplasm in the Pee Dee cotton germplasm enhancement program. Euphytica 195: 452—465.

Campbell, B.T., and G.O. Meyers. 2015. Quantitative Genetics. *In* Cotton, 2nd ed. D.D. Fang and R.G. Percy (eds.) ASA, CSSA, and SSSA, Inc. Madison, WI. doi:10.2134/agronmonogr57.2013.0024.

Cantrell, R.G., C.L. Roberts, and C. Waddell. 2000. Registration of 'Acala 1517-99' cotton. Crop Science 40: 1200—1201.

Castro, A.J., X. Chen, A. Corey, T. Filichkina, P.M. Hayes, C. Mundt, K. Richardson, S. Sandoval-Islas, and H. Vivar. 2003. Pyramiding and validation of quantitative trait locus (QTL) alleles determining resistance to barley stripe rust: effects on adult plant resistance. Crop Science 43: 2234—2239.

Cheatham, C.C., J.N. Jenkins, J.C. McCarty, Jr., C.E. Watson, and J. Wu. 2003. Genetic variances and combining ability of crosses of American cultivars, Australian cultivars, and wild cottons. Journal of Cotton Science 7: 16—22.

Chee, P.W. and B.T. Campbell. 2009. Bridging classical and molecular genetics of cotton fiber quality and development. *In* A.H. Paterson (ed.) Genetics and Genomics of Cotton. Springer, New York, NY. pp 283—31.

Chen, H., N. Qian, W. Guo, Q. Song, B. Li, F. Deng, D. Dong, and T. Zhang. 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, S., C.G. Xu, X.H. Lin, and Q. Zhang. 2001. Improving bacterial blight resistance of '6078', an elite restorer line of hybrid rice, by molecular marker-assisted selection. Plant Breeding 2: 133—137.

Colton, L.M., H.I. Groza, S.M Wielgus, and J. Jiang. 2006. Marker-assisted selection for the broad-spectrum potato late blight resistance conferred by gene *RB* derived from a wild potato species. Crop Science 46: 589—594.

Compton, M.E. 1994. Statistical methods suitable for the analysis of plant tissue culture data. Plant Cell, Tissue and Organ Culture 37: 217—242.

Constable, G.A., and M.P. Bange. 2007. Producing and preserving fiber quality: from the seed to the bale. *In*: Stephens, M. (ed.) Proceedings of the World Cotton Research Conference. Lubbock, TX, September. pp 10—14.

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

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

Doerge, R.W. 2002. Multifactorial genetics: Mapping and analysis of quantitative trait loci in experimental populations. Nature Reviews Genetics 3: 43—52.

Dudley, J.W. 1993. Molecular markers in plant improvement: Manipulation of genes affecting quantitative traits. Crop Science 33: 660—668.

Falconer, D.S. and T.F.C. Mackay. 1996. Introduction to quantitative genetics. Fourth edition. Longman, Essex, UK.

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.

Fang, D.D., J.N. Jenkins, D.D. Deng, J.C. McCarty, P. Li, and J. Wu. 2014. Quantitative trait loci analysis of fiber quality traits using a random-mated recombinant inbred population in Upland cotton (*Gossypium hirsutum* L.). BMC Genomics 15: 397—411.

Fang, D.D., J. Xiao, P.C. Canci, and R.G. Cantrell. 2010. A new SNP haplotype associated with blue disease resistance gene in cotton (*Gossypium hirsutum* L.). Theoretical and Applied Genetics 120: 943—953.

Faulkner, W.B., E.F. 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.

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

Gore, M.A., D.D. Fang, J.A. Poland, J. Zhang, R.G. Percy, R.G. Cantrell, G. Thyssen, and A.E Lipka. 2014. Linkage map construction and quantitative trait locus analysis of agronomic and fiber quality traits in cotton. The Plant Genome 7. doi:10.385/plantgenome2013.07.0023.

Green, C.C., and T.W. Culp. 1990. Simultaneous improvement of yield, fiber quality, and yarn strength in upland cotton. Crop Science 30: 66—69.

Guo, W.Z., T.Z. Zhang, Y.Z. Ding, Y.C. Zhu, X.L. Shen, and X.F. Zhu. 2005. Molecular marker assisted selection and pyramiding of two QLTs for fiber strength in upland cotton. Acta Genetica Sinica 32: 1275—1285.

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

Heffner, E.L., M.E. Sorrells, and J.L. Jannink. 2009. Genomic selection for crop improvement. Crop Science 49: 1—12.

Hill, W.G. 1972. Estimation of realised heritabilities from selection experiments I. Divergent selection. Biometrics 28: 747—765.

Hinze, L.L., B.T. Campbell, and R.J. Kohel. 2011. Performance and combining ability in cotton (*Gossypium hirsutum* L.) populations with diverse parents. Euphytica 181: 115—125.

Hinze, L.L., J.K. Dever, R.G. Percy. 2012. Molecular variation among and within improved cultivars in the U.S. cotton germplasm collection. Crop Science 52: 222—230.

Holland, J.B., H.S. Moser, L.S. O'Donoughue, and M. Lee. 1997. QTLs and epistasis associated with vernalization responses in oat. Crop Science 37: 1306—1316.

Holland, J.B., W.E. Nyquist, and C.T. Cervantes-Martinez. 2003. Estimating and interpreting heritability for plant breeding: an update. *In* J. Jannink (ed.) Plant Breeding Reviews, Volume 22. John Wiley & Sons, Inc. Oxford, UK. pp 9—112.

Hulse-Kemp, A.M., J. Lemm, J. Plieske, H. Ashrafi, R. Buyyarapu, D.D. Fang, J. Frelichowski et al. 2015. Development of a 63K SNP array for cotton and high-density mapping of intraspecific and interspecific populations of *Gossypium* spp. G3: Genes|Genomes|Genetics 5: 1187—1209.

Islam, M.S., L. Zeng, C.D. Delhom, X. Song, H.J. Kim, P. Li, and D.D. Fang. 2014. Identification of cotton fiber quality quantitative trait loci using intraspecific crosses derived from two near-isogenic lines differing in fiber bundle strength. Molecular Breeding 34: 373—384.

Jenkins, J.J., J.C. McCarty, Jr., J. Wu, and O.A. Gutierrez. 2009. Genetic variance components and genetic effects among eleven diverse upland cotton lines and their F2 hybrids. Euphytica 167: 397—408.

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

Jones, J.E., J.I. Dickson, W. Aguillard, W.D. Caldwell, S.H. Moore, R.L. Hutchinson, and R.L. Rogers. 1991. LA 887 cotton. Crop Science 31: 1701.

Joy, K.S. 2014. Inheritance of cotton fiber length and strength. Ph.D. diss., Texas A&M University, College Station, TX.

Joy, K.S., C.W. Smith, E. Hequet, S. Hague, P.S. Thaxton, and C. Souder. 2010. Fiber properties and mini-spun yarn performance of extra long staple upland cotton. Journal of Cotton Science 14: 82—90.

Kohel, R.J., J. Yu, Y.H. Park, and G.R. Lazo. 2001. Molecular mapping and characterization of traits controlling fiber quality in cotton. Euphytica 121: 163—171.

Knoll, J., and G. Ejeta. 2008. Marker-assisted selection for early-season cold tolerance in sorghum: QTL validation across populations and environments. Theoretical and Applied Genetics 116: 541—553.

Kumar, P., R. Singh, E. Lubbers, X. Shen, A.H. Paterson, B.T. Campbell, D.C. Jones, and P.W. Chee. 2012. Mapping and validation of fiber strength quantitative trait loci on chromosome 24 in upland cotton. Crop Science 52: 1115—1122.

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., D. Llewellyn, J. Jacobs, T. Arioli, D. Becker, S. Calhoun, Y. Al-Ghazi, et al. 2010. Meta-analysis of cotton fiber quality QTLs across diverse environments in *Gossypium hirsutum x G. barbadense* RIL population. BMC Plant Biology 10: 132—155.

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

Law, C.N. 1967. The location of genetic factors controlling a number of quantitative characters in wheat. Genetics 56: 445—461.

Lee, J.A., P.A. Miller, and J.O. Rawlings. 1967. Interaction of combining ability effects with environments in diallel crosses of upland cotton (*Gossypium hirsutum* L.). Crop Sci 7: 477—481.

Li, F. *et al.* 2015. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution. Nature Biotechnology 33: 524—530.

Li, X., D. Yuan, J. Zhang, Z. Lin, And X. Zhang. 2013. Genetic mapping and characteristics of genes specifically or preferentially expressed during fiber development in cotton. PLoS ONE 8: e54444. doi: 10.1371/journal/pone.0054444.

Lin, Z., D. He., Z. Zhang, Y. Nie, X Guo, C. Fend, and J.M. Stewart. 2005. Linkage map construction and mapping QTL for cotton fibre quality using SRAP, SSR and RAPD. Plant Breeding 124: 180—187.

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

May, O.L. 1999. Genetic variation in fiber quality. *In* Basra A.S. (ed). Cotton fibers: Developmental biology, quality improvement, and textile processing. Food Products Press, New York, NY. pp 183—229.

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

McCall, L.L., L.M. Verhalen, and R.W. McNew. 1986. Multidirectional selection for fiber strength in upland cotton. Crop Science 26: 744—750.

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

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

Meredith, W.R. 1984. Quantitative genetics. In Kohel, R.J. and Lewis, C.F. (eds.) Cotton, Agronomy Monographs 24. ASA, CSSA, and SSSA, Madison, WI. pp 131—150.

Meredith, W.R., and R.R. Bridge. 1973. Yield, yield component and fiber property variation of cotton (*Gossypium hirsutum* L.) within and among environments. Crop Science 13: 307—312.

Meredith, W.R., and W.S. Nokes. 2011. Registration of MD 9ne and MD 25 high-fiber-quality germplasm lines of cotton. Journal of Plant Registrations 5: 202—206.

Meredith, W.R. Jr., T.W. Culp, K.Q. Robert, G.F. Ruppenicker, W.S. Anthony, and J.R. Williford. 1991. Determining future cotton variety fiber quality objectives. Textile Research Journal 61: 715—720.

Meyer, L., S. MacDonald, and C. Valdes. 2015. Cotton and Wool Outlook. USDA Economic Research Service. http://www.ers.usda.gov/media/1820781/cws15d.pdf (accessed 4 June 2015).

Miller, P.A., J.C. Williams, Jr., H.R. Robinson, and R.E. Comstock. 1958. Estimates of genotypic and environmental variances and covariances in upland cotton and their implications in selection. Agronomy Journal 50: 126—131.

Ng, E.-H., C.W. Smith, E. Hequet, S. Hague, and J. Dever. 2014. Diallel analysis of fiber quality traits with an emphasis on elongation in upland cotton. Crop Science 54: 514—519.

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

Paterson, A.H. *et al.* 2012. Repeated polyploidization of *Gossypium* genomes and the evolution of spinnable cotton fibres. Nature 492: 423—427.

Paterson, A.H., E.S. Lander, J.D. Hewitt, S. Peterson, S.E. Lincoln, and S.D. Tanksley. 1988. Resolution of quantitative traits into Mendelian factors by using a complete linkage map of restriction fragment length polymorphisms. Nature 335: 721—726.

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

Qin, H., W. Guo, Y.M. Zhang, and T. Zhang. 2008. QTL mapping of yield and fiber traits based on a four-way cross population in *Gossypium hirsutum* L. Theoretical and Applied Genetics 117: 883—894.

R Development Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Url: http://www.R-project.org/.

Roberts, C., R.G. Cantrell, and S.T. Ball. 1997. Release of New Mexico Sea Island 1331. p. 473 *In* Proceedings of the National Beltwide Cotton Conference, New Orleans, LA. 7-10 Jan. 1997. National Cotton Council of America, Memphis, TN.

Rong, J. *et al*. 2004. A 3347-locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission, and evolution of cotton (*Gossypium*). Genetics 166: 389—417.

Rong, J., F.A. Feltus, V.N. Waghmare, G.J. Pierce, P.W. Chee, X. Draye, Y. Saranga, R.J. Wright, T.A. Wilkins, O.L. May, C.W. Smith, J.R. Gannaway, J.F. Wendel, and A.H. Paterson. 2007. Meta-analysis of polyploidy 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.

Said, J.I., Z. Lin, X. Zhang, M. Song, and J. Zhang. 2013. A comprehensive meta QTL analysis for fiber quality, yield, yield related and morphological traits, drought tolerance, and disease resistance in tetraploid cotton. BMC Genomics 14: 776—797.

SAS Institute. 2013. JMP Statistics and Graphics Guide. V. 11.2.0. SAS Inst., Cary, NC.

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

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

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

Shen, X., T. Zhang, W. Guo, X. Zhu, and X. Zhang. 2006. Mapping fiber and yield QTLs with main, epistatic, and QTL x environment interaction effects in recombinant inbred lines of upland cotton. Crop Science 46: 61—66.

Smith, C.W. 2001. Registration of three morphological variant upland cotton germplasm lines. Crop Science 41: 1371—1372.

Smith, C.W. 2003. Registration of TAM 94L-25 and TAM 94J-3 germplasm lines of upland cotton with improved fiber length. Crop Science 43: 742—743.

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

Smith, C.W., S. Hague, P.S. Thaxton, E. Hequet, and D. Jones. 2009. Registration of eight extra-long staple upland cotton germplasm lines. Journal of Plant Registrations 3: 81—85.

Smith, C.W., D. Jones, S. Hague. 2011. Registration of Tamcot 73 upland cotton cultivar. Journal of Plant Registrations 5: 273.

Smith, H., and R. Zhu. 1999. The spinning process. *In* C.W. Smith and J.T. Cothren (eds.) Cotton Origin, History, Technology, and Production. John Wiley & Sons, Inc., New York, NY. pp 729—749.

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

Stephens, S.G. 1949. The cytogenetics of speciation in *Gossypium*. I. Selective elimination of the donor parent genotype in interspecific backcrosses. Genetics 34: 627—637.

Su, C., W. Wang, X. Qiu, L. Yang, S. Li, M. Wang, and Q. Pan. 2013. Fine-mapping a fibre strength QTL *QFS-D11-1* on cotton chromosome 21 using introgressed lines. Plant Breeding 132: 725—730.

Sun, F.D., J.H. Zhang, S.F. Wang, W.K. Gong, Y.Z., Shi, A.Y. Liu, J.W. Li, J.W. Gong, H.H. Shang, and Y.L. Yuan. 2012. QTL mapping for fiber quality traits across multiple generations and environments in upland cotton. Molecular Breeding 30: 569—582.

Tan, Z. *et al.* 2015. Genetic map and QTL controlling fiber quality traits in upland cotton (*Gossypium hirsutum* L.) Euphytica 203: 615—628.

Tang, B., J.N. Jenkins, C.E. Watson, J.C. McCarty, and R.G. Creech. 1996. Evaluation of genetic variances, heritabilities, and correlations for yield and fiber traits among cotton F_2 hybrid populations. Euphytica 91: 315—322.

Tanksley, S.D. 1993. Mapping polygenes. Annual Review of Genetics 27: 205—233.

Tanksley, D., H. medina-Filho, and C.M. Rick. 1981. The effect of isozyme selection on metric characters in an interspecific backcross of tomato – Basis of an early screening procedure. Theoretical and Applied Genetics 60: 291—296.

Thabuis, A., A. Palloix, B. Servin, A.M. Daubeze, P. Signoret, F. Hospital, and V. Lefebvre. 2004. Marker-assisted introgression of 4 *Phytophthora capsici* resistance QLT alleles into a bell pepper line: validation of additive and epistatic effects. Molecular Breeding 14: 9—20.

Thaxton, P.M., C.W. Smith, and R. Cantrell. 2005. Registration of TAM 96WD-18 upland coton germplasm line with improved fiber length and strength. Crop Science 45: 1172.

Toojinda, T., E. Baird, A. Booth, L. Broers, P. Hayes, W. Powell, W. Thomas, H. Vivar, and G. Young. 1998. Introgression of quantitative trait loci (QTLs) determining stripe rust resistance in barley: an example of marker-assisted line development. Theoretical and Applied Genetics 96: 123—131.

Ulloa, M. 2006. Heritability and correlations of agronomic and fiber traits in an okra-leaf upland cotton population. Crop Science 46: 1508—1514.

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

Ůreyen, M.E., and H. Kadoğlu. 2006. Regressional estimation of ring cotton yarn properties from HVI fiber properties. Textile Research Journal 76: 360—366.

USDA. 2015a. 2014 State Agriculture Overview. USDA National Agricultural Statistics Service. http://www.nass.usda.gov/Statistics_by_State/Texas/ (accessed 5 June 2015).

USDA. 2015b. Baseline Related Historical Data. USDA Economic Research Service. http://ers.usda.gov/data-products/chart-gallery/detail.aspx?chartId=40082&ref=collection&embed=True (accessed 4 June 2015).

Van Deynze, A., K. Stoffel, M. lee, T.A. Wilkins, A. Kozik, R.G. Cantrell, J.Z. Yu, R.J. Kohel, and D.M. Stelly. 2009. Sampling nucleotide diversity in cotton. BMC Plant Biology 9: 125. doi:10.1186/1471-2229-9-125.

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

Wang, K. *et al.* 2012. The draft genome of a diploid cotton *Gossypium raimondii*. Nature Genetics 44: 1098—1103.

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

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

Wang, F. Z. Xu, R. Sun, Y. Gong, G. Liu, J. Zhang, L. Wang, C. Zhang, S. Fan, J. Zhang. 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.

Xu, Y., and J.H. Crouch. 2008. Marker-assisted selection in plant breeding: From publications to practice. Crop Science 48: 391—407.

Yang, W., and D.M. Francis. 2005. Marker-assisted selection for combining resistance to bacterial spot and bacterial speck in tomato. Journal of the American Society for Horticultural Science 130: 716—721.

Yu, J., and E.S. Buckler. 2006. Genetic association mapping and genome organization in maize. Current Opinion in Biotechnology 17: 1555—160.

Yu, J.Z., R.J. Kohel, D.D. Fang, J. Cho, A. Van Deynze, M. Ulloa, S.M. Hoffman, A.E. Pepper, D.M. Stelly, J.N. Jenkins, S. Saha, S.P. Kumpatala, M.R. Shah, W.V. Hugie, and R.G. Percy. 2012. A high-density simple sequence repeat and single nucleotide polymorphism genetic map of the tetraploid cotton genome. G3|Genes|Genomes|Genetics 2: 43—58.

Yu, J., S. Yu, M. Gore, M. Wu, H. Zhai, X. Li, S. Fan, M. Song, and J. Zhang. 2013a. Identification of quantitative trait loci across interspecific F_2, $F_{2:3}$ and testcross populations for agronomic and fiber traits in tetraploid cotton. Euphytica 191: 375—389.

Yu, J., K. Zhang, S. Li, S. Yu, H. Zhai, M. Wu, X. Li, S. Fan, M. Song, D. Yang, Y. Li, and J. Zhang. 2013b. Mapping quantitative trait loci for lint yield and fiber quality across environments in a *Gossypium hirsutum* x *Gossypium barbadense* backcross inbred line population. Theoretical and Applied Genetics 126: 275—287.

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

Zeng, L., W.R. Meredith Jr., and D.L. Boykin. 2011. Germplasm potential for continuing improvement of fiber quality in Upland cotton: Combining ability for lint yield and fiber quality. Crop Science 51: 60—68.

Zhang, T. *et al.* 2015. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement. Nature Biotechnology 33: 531—537.

Zhang, Z.S., M.C. Hu, J. Zhang, D.J. Liu, J. Zheng, K. Zhang, W. Wang, and Q. Wan. 2009a. 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, Y., L. Luo, T. Liu, C. Xu, and Y. Xing. 2009b. Four rice QTL controlling number of spikelets per panicle expressed the characteristics of single Mendelian gene in near isogenic backgrounds. Theoretical and Applied Genetics 118: 1035—1044.

Zhang, T., N. Qian, X. Zhu, H. Chen, S. Wang, H. Mei, and Y. Zhang. 2013. Variations and transmission of QTL alleles for yield and fiber qualities in upland cotton cultivars developed in China. PLoS ONE 8: e57220. doi: 10.1371/journal.pone.0057220.

Zhang, M., Y.H. Wu, M.K. Lee, Y.H. Liu, Y. Rong, T.S. Santos, C. Wu, F. Xie, R.L. Nelson, and H.B. Zhang. 2010. Numbers of genes in the NBS and RLK families vary by more than four-fold within a plant species and are regulated by multiple factors. Nucleic Acids Research 38: 6513—6525.

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

Zhang, T., Y. Yuan, J. Yu, W. Guo, and R.J. Kohel. 2003. Molecular tagging of a major QTL for fiber strength in Upland cotton and its marker-assisted selection. Theoretical and Applied Genetics 106: 262—268.

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

Zhiyuan, N., H. Chen, H. Mei, and T. Zhang. 2014. Molecular tagging of QTLs for fiber quality and yield in the upland cotton cultivar Acala-Prema. Euphytica 195: 143—156.

APPENDIX

Table A.1. P-values derived from single-marker analysis (i.e., ANOVA) of fiber UHML and Str for 229 SSR alleles among three experimental F_3 populations grown at College Station, TX in 2013.

SSR_Allele	UHML			Str		
	Population	Allele	Population* Allele	Population	Allele	Population* Allele
BNL0686_157	<.0001	0.4388	0.1992	<.0001	0.2135	0.0828
BNL0686_159	<.0001	0.1085	0.0030†	<.0001	0.4305	0.9334
BNL0830_104	<.0001	0.0017	0.1063	<.0001	0.0018	0.4640
BNL0830_99	<.0001	0.0272	0.0034†	<.0001	0.0132	0.2672
BNL1122_166	<.0001	0.7825	0.3157	<.0001	0.0028	0.1211
BNL1122_170	<.0001	0.5354	0.0791	<.0001	0.0194	0.0222†
BNL1227_186	<.0001	0.6600	0.5661	<.0001	0.0901	0.0840
BNL1395_154	<.0001	0.4601	0.3774	<.0001	0.0444	0.0463†
BNl1395_158	<.0001	0.3389	0.3359	<.0001	0.0258	0.0309†
BNL1604_120	<.0001	<.0001	0.0541	<.0001	<.0001	0.1040
BNl1604_98	<.0001	<.0001	0.2813	<.0001	<.0001	0.0784
BNL1672_134	<.0001	0.8634	0.0813	<.0001	0.1162	0.0492†
BNL2495_191	<.0001	0.9539	0.4725	<.0001	0.5294	0.9003
BNL2495_195	<.0001	0.6842	0.1446	<.0001	0.0089	0.0138†
BNL2599_88	<.0001	0.0298	0.1170	<.0001	0.1352	0.9708
BNL2599_96	<.0001	0.5349	0.4085	<.0001	0.0005	0.4720
BNL2634_197	<.0001	0.8273	0.5357	<.0001	0.0193	0.0465†
BNL2634_199	<.0001	0.4037	0.1617	<.0001	0.0008	0.0212†
BNL2733_99	<.0001	0.5787	0.1134	<.0001	0.0219	0.0327†
BNL2733_105	<.0001	0.7021	0.0482†	<.0001	0.0164	0.0181†
BNl2921_156	<.0001	0.8861	0.0072†	<.0001	0.0519	0.0612
BNL2921_158	<.0001	0.1646	0.0086†	<.0001	0.7878	0.5498
BNL2986_155	<.0001	<.0001	0.1486	<.0001	0.8779	0.0010†
BNL2986_157	<.0001	0.0027	0.1282	<.0001	0.8227	0.0163†
BNl3031_184	<.0001	0.5377	0.0189†	<.0001	0.6383	0.0056
BNL3085_104	<.0001	0.2097	0.9398	<.0001	0.8270	0.1197
BNL3085_106	<.0001	0.3951	0.8807	<.0001	0.0910	0.7969
BNl3257_236	<.0001	0.7510	0.5343	<.0001	0.2418	0.1253
BNL3261_210	<.0001	0.2254	0.1249	<.0001	0.9585	0.6167
BNL3280_212	<.0001	0.8476	0.0119†	<.0001	<.0001	0.0048†
BNL3280_213	<.0001	0.6362	0.3274	<.0001	0.0315	0.1120
BNl3410_222	<.0001	0.6002	0.8072	<.0001	0.0754	0.1365
BNl3410_224	<.0001	0.4784	0.2431	<.0001	0.3992	0.5645
BNL3442_113	<.0001	0.4770	0.5156	<.0001	0.9608	0.8688
BNl3452_178	<.0001	0.0037	0.9633	<.0001	0.8896	0.9465
BNl3452_180	<.0001	0.0351	0.8517	<.0001	0.3081	0.2100
BNl3463_232	<.0001	0.5257	0.4177	<.0001	0.0508	0.5290
BNl3463_233	<.0001	0.4119	0.0233†	<.0001	0.8545	0.3451
BNL3545_113	<.0001	0.9739	0.2002	<.0001	0.0048	0.0270†

Table A.1. *Continued*

BNl3545_115	<.0001	0.3519	0.1269	<.0001	<.0001	0.0010†
BNl3545_119	<.0001	0.2210	0.0164†	<.0001	0.4599	0.2531
BNl3545_128	<.0001	0.9079	0.1061	<.0001	0.0049	0.0141†
BNL3545_138	<.0001	0.0189	0.2399	<.0001	0.1839	0.1119
BNL3545_183	<.0001	0.2655	0.2659	<.0001	0.0729	0.1758
BNl3545_188*	<.0001	0.0126	0.7969	<.0001	0.1613	0.0328†
BNL3558_208	<.0001	0.9165	0.1115	<.0001	0.3235	0.2434
BNL3558_210	<.0001	0.8158	0.4662	<.0001	0.5476	0.8397
BNl3649_182	<.0001	0.7883	0.0005†	<.0001	0.3922	0.1091
BNL3649_192	<.0001	0.8277	0.0004†	<.0001	0.5515	0.1782
BNL4017_221	<.0001	0.5297	0.6428	<.0001	0.1749	0.0108†
BNL4017_222	<.0001	0.6657	0.3606	<.0001	0.5110	0.1164
BNL4017_226	<.0001	0.0548	0.2657	<.0001	0.0296	0.0155†
BNL4017_234	<.0001	0.0011	0.1318	<.0001	0.0255	0.3457
CER0021_136	<.0001	0.1765	0.1313	<.0001	0.7979	0.4428
CGR5106_178	<.0001	0.8405	0.7582	<.0001	0.2551	0.0464†
CGR5106_179	<.0001	0.2292	0.8739	<.0001	0.8117	0.0018†
CGR5106_190	<.0001	0.4672	0.2743	<.0001	0.0888	0.2814
CGR5106_191	<.0001	0.6609	0.2776	<.0001	0.0729	0.3966
CGR5139_180	<.0001	0.5869	0.4954	<.0001	0.4327	0.6813
CGR5139_182	<.0001	0.8754	0.5027	<.0001	0.0558	0.1681
CGR5510_170	<.0001	0.8357	0.0514	<.0001	0.0772	0.3381
CGR5510_172	<.0001	0.4106	0.1631	<.0001	0.0273	0.2246
CGR5548_162	<.0001	0.0342	0.4770	<.0001	0.7565	0.1598
CGR5548_171	<.0001	0.1299	0.5777	<.0001	0.0798	0.3801
CGR5565_155	<.0001	0.6590	0.7163	<.0001	0.2394	0.0288†
CGR5565_170	<.0001	0.1975	0.9571	<.0001	0.5521	0.1457
CGR5870_168	<.0001	0.1451	0.2812	<.0001	0.7354	0.2199
CGR6012_133	<.0001	0.7260	0.1681	<.0001	0.9029	0.3861
CGR6012_144	<.0001	0.0166	0.1018	<.0001	0.2751	0.0267†
CGR6170_208	<.0001	0.5098	0.1942	<.0001	0.3138	0.2393
CGR6329_230	<.0001	0.0008	0.0001	<.0001	0.0011	<.0001†
CGR6329_232	<.0001	0.0001	0.0009	<.0001	<.0001	0.0001
CGR6383_217	<.0001	0.0838	0.9876	<.0001	0.8983	<.0001†
CGR6383_223	<.0001	0.2258	0.2349	<.0001	0.9467	<.0001†
CGR6902_127	<.0001	0.6532	0.1543	<.0001	0.1566	0.3376
CGR6902_129	<.0001	0.0874	0.0098†	<.0001	0.9964	0.1722
CGR6902_145	<.0001	0.5015	0.1015	<.0001	0.2713	0.9877
CIR005_184	<.0001	0.0147	0.9133	<.0001	0.2513	0.4279
CIR005_186	<.0001	0.0128	0.9794	<.0001	0.3653	0.1633
CIR012_171	<.0001	0.0557	0.2664	<.0001	0.3157	0.7690
CIR017_129	<.0001	0.6752	0.9172	<.0001	0.4394	0.4254
CIR091_177	<.0001	0.6885	0.1568	<.0001	0.4857	0.8528
CIR091_178	<.0001	0.4324	0.9981	<.0001	0.2452	0.3399
CIR091_180	<.0001	0.4568	0.5449	<.0001	0.0033	0.0060†
CIR091_181	<.0001	0.4474	0.4996	<.0001	0.0739	0.0369†
CIR148_145	<.0001	0.1071	0.7141	<.0001	0.4696	0.6203
CIR165_207	<.0001	0.0160	0.3666	<.0001	0.4788	0.8294
CIR165_209	<.0001	0.0566	0.6494	<.0001	0.1294	0.1268
CIR167_206	<.0001	0.5806	0.0078†	<.0001	0.5053	0.6839
CIR167_207	<.0001	0.7577	<.0001†	<.0001	0.4672	0.6631

Table A.1. *Continued*

CIR170_158	<.0001	0.4793	0.3012	<.0001	0.4981	0.0048†
CIR170_162	<.0001	0.9633	0.2176	<.0001	0.2383	0.0775
CIR196_192	<.0001	0.0001	0.6306	<.0001	0.7755	0.2756
CIR196_194	<.0001	0.0025	0.0025†	<.0001	0.3035	0.3942
CIR213_236	<.0001	0.7281	0.1443	<.0001	0.8480	0.2368
CIR213_237	<.0001	0.3106	0.0668	<.0001	0.0006	0.0050
CIR246_146	<.0001	0.1095	0.0062†	<.0001	0.3439	0.1368
CIR246_157	<.0001	0.8916	0.1048	<.0001	0.0011	0.0048†
CIR246_168	<.0001	0.0071	0.1308	<.0001	0.2215	0.0737
CIR249_192	<.0001	0.3077	0.0189†	<.0001	0.0279	0.0921
CIR253_187	<.0001	0.0030	0.0039†	<.0001	0.0123	0.7490
CM0160_103	<.0001	0.7887	0.2463	<.0001	0.9605	0.9095
COT089_189	<.0001	0.0199	0.3983	<.0001	0.2895	0.0678
COT089_194	<.0001	0.1002	0.4202	<.0001	0.3588	0.0174†
DC30107_211	<.0001	0.7612	0.6998	<.0001	0.0243	0.0501
DC30107_213	<.0001	0.6138	0.4763	<.0001	0.0903	0.0171†
DC30210_148	<.0001	0.9003	<.0001†	<.0001	0.3553	0.0028†
DC40122_204	<.0001	0.9254	0.6141	<.0001	0.7886	0.2630
DOW067_156	<.0001	0.0001	0.1069	<.0001	0.1365	0.6635
DOW067_162	<.0001	0.0014	0.2681	<.0001	0.6521	0.2783
DPL0028_186	<.0001	0.5775	0.2145	<.0001	0.2977	0.5340
DPL0028_188	<.0001	0.6596	0.8922	<.0001	0.1874	0.1738
DPL0236_154	<.0001	0.0713	0.1295	<.0001	0.0224	0.0094†
DPL0236_157	<.0001	0.5118	0.0180†	<.0001	0.0379	0.0849
DPL0270_142	<.0001	0.0006	0.0975	<.0001	0.0002	0.1125
DPL0270_149	<.0001	0.0037	0.1058	<.0001	0.0015	0.0703
DPL0570_302	<.0001	0.0033	0.4113	<.0001	0.4605	0.8666
DPL0570_304	<.0001	0.3363	0.2888	<.0001	0.3851	0.0672
DPL1071_290	<.0001	0.8992	0.1699	<.0001	0.0064	0.3028
DPL1201_275	<.0001	0.0370	<.0001†	<.0001	0.3623	0.4342
DPL1201_281	<.0001	0.7357	0.0002†	<.0001	0.2079	0.1865
DPL1358_205	<.0001	0.0019	0.0251†	<.0001	<.0001	0.1298
DPL1358_212	<.0001	0.0118	0.0645	<.0001	0.0008	0.1010
DPL1362_285	<.0001	0.7740	0.6096	<.0001	0.0054	0.0482†
DPL1362_288	<.0001	0.5712	0.6012	<.0001	0.0001	0.0197†
DPL1379_168	<.0001	0.2577	<.0001†	<.0001	0.0354	0.0569
HAU006_205	<.0001	0.8100	0.9575	<.0001	0.5285	0.7061
HAU0086_199	<.0001	0.5673	0.8774	<.0001	0.7268	0.5404
HAU0086_202	<.0001	0.3708	0.3539	<.0001	0.0024	0.0031†
HAU0087_179	<.0001	0.3944	0.8505	<.0001	0.5339	0.1759
HAU0087_181	<.0001	0.9353	0.4522	<.0001	0.0039	0.0058†
HAU0087_188	<.0001	0.0430	0.0003†	<.0001	0.0328	0.0836
HAU0087_190	<.0001	0.0469	0.0081†	<.0001	0.1805	0.2612
HAU2022_163	<.0001	0.0370	0.8772	<.0001	0.4914	0.9739
HAU2022_166	<.0001	0.2552	0.3733	<.0001	0.5647	0.3256
HAU2065_318	<.0001	0.6324	0.1984	<.0001	0.0946	0.1374
HAU2065_325	<.0001	0.0244	0.0838	<.0001	0.3491	0.8631
HAU3233_261	<.0001	0.1764	0.2485	<.0001	0.7480	0.2728
HAU3233_263	<.0001	0.3387	0.0366†	<.0001	0.8712	0.0839
JESPR050_200	<.0001	0.2913	0.0024†	<.0001	0.1822	0.0002†
JESPR050_218	<.0001	0.0028	0.0699	<.0001	0.0400	0.1620

Table A.1. *Continued*

JESPR065_137	<.0001	0.1663	0.0122†	<.0001	0.4555	0.9973
JESPR065_165	<.0001	0.4244	<.0001†	<.0001	0.0542	0.2486
JESPR070_82	<.0001	0.1179	0.2380	<.0001	0.9570	0.5973
JESPR070_92	<.0001	0.5689	0.0160†	<.0001	0.0271	0.4380
JESPR114_86	<.0001	0.9839	0.4157	<.0001	0.9334	0.0074†
JESPR114_93	<.0001	0.9667	0.9810	<.0001	0.6620	0.3507
JESPR192_135	<.0001	0.1263	0.5595	<.0001	0.7140	0.0254
JESPR218_108	<.0001	0.4709	0.7783	<.0001	0.8099	0.1125
JESPR295_105	<.0001	0.9798	0.6093	<.0001	0.0050	0.0988
JESPR295_108	<.0001	0.4014	0.6158	<.0001	0.6725	0.0007†
MUSB0979_240	<.0001	0.7290	<.0001†	<.0001	0.3002	0.0986
MUSB0979_244	<.0001	0.3081	0.5931	<.0001	0.5482	0.0207†
MUSB0979_247	<.0001	0.0752	0.0133†	<.0001	0.0825	0.0502
MUSS172_200	<.0001	0.8703	0.0214†	<.0001	0.5684	0.0254†
MUSS172_221	<.0001	0.7483	0.0171†	<.0001	0.3836	0.0031†
MUSS422_200	<.0001	0.0045	0.0738	<.0001	0.0733	0.0354†
MUSS422_207	<.0001	0.0090	0.1004	<.0001	0.0283	0.0013†
NAU0895_192	<.0001	0.8403	0.0320†	<.0001	0.7244	0.6789
NAU0895_204	<.0001	0.1714	0.3397	<.0001	0.9654	0.2784
NAU0913_195	<.0001	0.8884	0.5778	<.0001	0.5642	0.3751
NAU0913_197	<.0001	0.3623	0.6995	<.0001	0.0213	0.0267†
NAU0913_203	<.0001	0.0366	0.0006†	<.0001	0.0200	0.1636
NAU0913_206	<.0001	0.0044	0.2052	<.0001	0.0723	0.1884
NAU0943_178	<.0001	0.5077	0.5777	<.0001	0.2730	0.0123†
NAU1037_190	<.0001	0.0264	0.4087	<.0001	0.2198	0.1806
NAU1037_193	<.0001	0.3379	0.4776	<.0001	0.1052	0.8209
NAU1042_219	<.0001	0.2899	0.9397	<.0001	0.1839	0.1094
NAU1042_242	<.0001	0.2787	0.8750	<.0001	0.0155	0.3339
NAU1102_231	<.0001	0.3715	0.0157†	<.0001	0.0411	0.1359
NAU1167_189	<.0001	0.6901	0.0829	<.0001	0.2025	0.0423†
NAU1167_195	<.0001	0.0756	0.0101†	<.0001	0.9122	0.3493
NAU1167_201	<.0001	0.2844	0.3218	<.0001	0.6378	0.5415
NAU1190_213	<.0001	0.3933	0.0308†	<.0001	0.4789	0.3651
NAU1190_233	<.0001	0.4303	0.0205†	<.0001	0.4705	0.0348†
NAU1217_157	<.0001	0.0278	0.5266	<.0001	0.1941	0.2119
NAU1217_158	<.0001	0.5043	0.4150	<.0001	0.0773	0.3217
NAU1221_220	<.0001	0.3498	0.5971	<.0001	0.0588	0.0106†
NAU1221_244	<.0001	0.6149	0.8535	<.0001	0.0323	0.3682
NAU1302_216	<.0001	0.0723	0.3623	<.0001	0.6846	0.3059
NAU1302_221	<.0001	0.0298	0.8168	<.0001	0.5562	0.0886
NAU1369_247	<.0001	0.0165	0.9984	<.0001	0.0049	0.1755
NAU1369_253	<.0001	0.1472	0.8138	<.0001	0.0182	0.1708
NAU2162_201	<.0001	0.1633	0.3410	<.0001	0.2223	0.4635
NAU2162_207	<.0001	0.7059	0.9506	<.0001	0.3521	0.9789
NAU2162_209	<.0001	0.8078	0.7107	<.0001	0.3097	0.8329
NAU2257_188	<.0001	0.7655	0.7386	<.0001	0.8633	0.7194
NAU2265_221	<.0001	0.9492	0.2062	<.0001	0.4368	0.1495
NAU2265_233	<.0001	0.0114	0.3528	<.0001	0.4502	0.9264
NAU2291_195	<.0001	0.1867	0.8742	<.0001	0.3099	0.2005
NAU2291_197	<.0001	0.6176	0.6276	<.0001	0.0445	0.0094†
NAU2291_204	<.0001	0.0535	0.0001†	<.0001	0.0380	0.1531

Table A.1. *Continued*

NAU2291_206	<.0001	0.0050	0.0443†	<.0001	0.0040	0.7985
NAU2477_202	<.0001	0.7565	0.8734	<.0001	0.0057	0.0082†
NAU2477_208	<.0001	0.0683	0.0031†	<.0001	0.0876	0.3495
NAU2477_211	<.0001	0.0036	0.0202†	<.0001	0.0059	0.6699
NAU3201_232	<.0001	0.6448	0.8494	<.0001	0.3107	0.2479
NAU3308_220	<.0001	0.0001	0.0003†	<.0001	0.2874	<.0001†
NAU3308_224	<.0001	<.0001	0.0325†	<.0001	0.2737	<.0001†
NAU3393_196	<.0001	0.5540	0.8136	<.0001	0.1856	0.0046†
NAU3393_205	<.0001	0.1828	0.0977	<.0001	0.2162	0.7412
NAU5037_282	<.0001	0.2366	0.0026†	<.0001	0.0009	0.0725
NAU5037_284	<.0001	0.1704	0.0010†	<.0001	0.0002	0.0226
NAU5046_219	<.0001	0.0570	0.0070†	<.0001	0.6231	0.0819
NAU5046_226	<.0001	0.0024	0.0609	<.0001	0.0988	0.3078
NAU5099_228	<.0001	0.0195	0.2849	<.0001	0.3065	0.3024
NAU5233_198	<.0001	0.2284	0.4460	<.0001	0.0201	0.0075†
NAU5233_204	<.0001	0.9805	0.4390	<.0001	0.1039	0.7288
SHIN0384_185	<.0001	<.0001	0.6914	<.0001	0.0587	0.6538
SHIN0384_187	<.0001	0.0141	0.0323†	<.0001	0.8721	0.2851
SHIN1138_176	<.0001	0.2417	0.2281	<.0001	0.6697	0.0479†
SHIN1138_177	<.0001	0.9568	0.9539	<.0001	0.8352	0.6069
SHIN1138_181	<.0001	0.0824	0.0531	<.0001	0.2583	0.0375†
SHIN1547_252	<.0001	0.9730	0.5295	<.0001	0.0993	0.5726
SHIN1635_241	<.0001	0.5986	0.0170†	<.0001	0.4026	0.0689
SHIN1635_243	<.0001	0.6633	0.5912	<.0001	0.7112	0.1282
TMB0189_178	<.0001	0.1397	0.0372†	<.0001	0.5781	0.0231†
TMB0189_184	<.0001	0.1559	0.3128	<.0001	0.0830	0.0420†
TMB0382_179	<.0001	0.2346	0.1376	<.0001	0.0343	0.8255
TMB0382_182	<.0001	0.7490	0.0739	<.0001	0.0889	0.4943
TMB0429_288	<.0001	0.0479	0.7841	<.0001	0.2958	0.2649
TMB0904_210	<.0001	0.6797	0.6846	<.0001	0.2390	0.5294
TMB1898_217	<.0001	0.2375	0.2458	<.0001	0.3077	0.5190
TMC005_178	<.0001	0.0837	0.1025	<.0001	0.3958	0.0019†
TMC005_183	<.0001	0.0481	0.2773	<.0001	0.2374	0.0267†
UCD120_264	<.0001	0.4574	0.9920	<.0001	0.3772	0.4818
UCD120_266	<.0001	0.2964	0.2389	<.0001	0.0442	0.0407
UCD120_273	<.0001	0.1235	0.0013†	<.0001	0.0347	0.1800
UCD120_275	<.0001	0.0213	0.0336†	<.0001	0.0325	0.4156

†Indicates alleles that were associated with opposite effects on UHML or Str depending on the population.