ABSTRACT

ANDRES, RYAN JAMES. Investigation of Cotton Leaf Shape: Its Effect on Boll Rot Severity and Underlying Genetic Mechanism. (Under the direction of Vasu Kuraparthy and Daryl Bowman).

Cotton boll rot is a disease of increasing importance in the southeastern U.S. Cotton Belt ranging from Louisiana to North Carolina. The mutant leaf shapes of cotton have been investigated as a potential source of boll rot control, but an ideal avoidance strategy has not yet been developed. Okra leaf cotton has been suggested as an effective way to control boll rot, but its reduced yield potential has prevented its acceptance by growers. Sub-okra cotton has minimal field resistance to boll rot but yields competitively with normal-leafed varieties. Here, an experimental open-pollination approach was used to develop hybrid cotton segregating for both the sub-okra and okra alleles. The ability of these hybrids to counter boll rot was investigated. Abnormally hot and dry weather failed to create conditions conducive to boll rot at any of the five test locations, all of which had previously shown nearly annual incidences of the disease. This prevented an appropriate analysis of the hybrid's boll rot avoidance, although results were encouraging from the one location with the highest incidence of boll rot. Attempts to identify the causal gene underlying leaf shape were also employed. Known leaf shape modification genes from model plant and crop species were used to query the Cotton Gene Index for potential orthologs. Expression patterns of these genes in different tissue types of both okra and normal leaf cultivars were analyzed for variation. Results from this analysis show some evidence that an ectopic expression of Class I *KNOX* genes is responsible for the okra leaf shape in cotton. However, a more robust and sensitive technique is recommended to confirm these findings and such approaches are briefly discussed. In a separate project, sterile, triploid interspecific hybrids were also developed by crossing *Gossypium hirsutum* with *G. raimondii* and *G. sturtianum* and the efforts to overcome their sterility are discussed. Additionally, a new and more efficient method of mitotic chromosome preparation compatible for fluorescent *in situ* hybridization was developed for cotton.

© Copyright 2012 by Ryan James Andres

All Rights Reserved

Towards Understanding the Genetic Architecture of Mutant Leaf Shapes and Their Potential Use in Cotton Breeding

by
Ryan James Andres

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

Crop Science

Raleigh, North Carolina

2012

APPROVED BY:

_____	_____
Vasu Kuraparthy Committee Co-Chair	Daryl Bowman Committee Co-Chair
_____	_____
Ramsey Lewis	Michael Gore

DEDICATION

To my wife Ellie for her love and support over the past nine and a half years of our lives together. Also, to my parents Tony and Karen for their encouragement and support of my academic and life goals.

BIOGRAPHY

Ryan Andres was born on October 29, 1987 in Poughkeepsie, NY to Anthony and Karen Andres. He relocated along with his family to Cary, NC six years later. After graduating from high school at St. David's School in Raleigh, NC in 2005, Ryan obtained his B.S. in Biology from Elon University in 2009. Following summer internships at NC State University with Dr. David A. Danehower and Dr. Kasi Azhakanandam and the Athenix Corporation during his undergraduate years, Ryan enrolled in graduate school at NC State University in 2009. Following graduation, Ryan plans to continue on for a PhD in the Kuraparthy Lab at NC State. He and his wife Ellie were married on April 14, 2012.

ACKNOWLEDGMENTS

First off, I would like to thank Dr. Vasu Kuraparthy for giving me an opportunity to come to NC State and continue my education in the field of crop science. I would also like to thank the Cotton Incorporated Fellowship Program, especially Dr. Don Jones, for providing funding to my assistantship. This work would not have been possible without the help of lab technician Blake Bowen, my fellow graduate students Priyanka Tyagi and Linglong Zhu, and long-term undergraduate workers Wess McKnight and Evan Jones. I would not be in this position today if not for the loving support from my family including my parents Tony and Karen and my sister Lisa. Finally, I would like to thank my wife Ellie, without whom, none of this would be possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Effect of Leaf Shape on Boll Rot Severity in Cotton

1.1 Review of Literature

1.1.1 Cotton Boll Rot

Cotton boll rot is a fungal and/or bacterial infection associated with high humidity and a lack of ventilation in the canopy during the flowering, boll development, and dehiscence of cotton (*Gossypium* spp.*)*. The severity of infestation can vary widely but the disease manifests itself through reduced yields and a deterioration of fiber quality. Boll rot refers to the collective infection by various fungi and bacteria that differ by location and environmental conditions. Cauquil (1975) identified over 40 fungal species and several bacteria that are capable of causing boll rot. Of these, fungal species belonging to the genera *Ascochyta, Aspergillus, Diplodia, Fusarium,* and *Glomerella,* along with the bacteria *Xanthomonas malvacearum* are considered the most common and serious threats (Cauquil 1975, Bell et al. 2010, Sadras and Felton 2010). This wide spectrum of causal agents negates the possibility of developing direct genetic resistance as any attempts to find and combine all of the necessary resistance mechanisms in one plant would be difficult, if not impossible. Furthermore, there appears to be little genetic resistance among the cultivars and species of *Gossypium* to the necrotrophic pathogens associated with boll rot (Bell et al. 2010). The disease appears to be spreading throughout the southeastern United States and new causal agents continue to be identified (Medrano and Bell 2007). One of these new species, *Pantoea agglomerans,* is a bacterium commonly thought of as a soil-born saprophyte that does not produce any external symptoms on infected bolls (Medrano and Bell 2007).

Potentially affected areas range from Louisiana to North Carolina and incidence appears more severe with increasing proximity to the coast. Projected increases in temperature, humidity, precipitation, and abnormal weather events brought about by global warming will likely only amplify the incidence of boll rot in this area. Severe boll rot can cause yield losses ranging from 50-80% in hard hit areas of the southeastern U.S. and was responsible for a 15-30% yield reduction in Brazil during the 2012 growing season (Darryl Bowman, Gerald Myers – personal communication).

An excellent and comprehensive source on the basics of boll rot is the work of Cauquil (1975), which is summarized here: The interior of a closed cotton boll is a dark, wet and contained environment that is ideal for microbial growth. The fibers provide an abundant source of sugar and sufficient lipids and proteins are provided by the seeds themselves. External symptoms appear first along sutures, the base, or the tip of the developing fruit. Bacteria cause oblong blemishes with decomposed tissue being soft, black, and exuding a fluid-like substance. Fungal attacks generally cause circular and brown necrotic spots. Inside, fiber is immature, moist, and often discolored, usually grey but also yellow. Bolls fail to swell and remain compact. Entry can occur actively through weak points such as dehiscence sutures, tips, and bases. Microbes also enter through previously induced lesions and can diffuse to the bolls after entering elsewhere.

Cotton plants seem to possess few natural mechanisms to prevent boll rot. External resistance is seen in lines that excel at preventing microbial penetration. This includes the development of a stronger pericarp, increasing capsular tightness, and decreasing attractiveness to pricking insects. Internal resistance measures focus around slowing microbial growth after infection has already occurred. Examples include changing locular conditions to deviate from ideal microbial growth conditions (e.g. reduced sugar content) and developing stronger septa to prevent the spread of infection from one locule to another. However, these internal mechanisms decrease yield and reduce fiber quality and are therefore not viable options.

Efforts to develop resistant cultivars and/or effective chemical control methods have not yielded the desired level of success against boll rot. Cauquil (1975) stated that the best way to counter boll rot was to alter the microclimate that spurs its growth by reducing planting density, limiting mineral fertilizer application, and moderating irrigation. Jones (1982) noted that the use of skip rows and bottom defoliation showed effectiveness in reducing boll rot. However, these practices may come at the expense of overall yield and would be undesirable if sufficient boll rot failed to develop.

1.1.2 Leaf Shape and Boll Rot Resistance

Jones (1970) compared leaf shape isolines in numerous varietal backgrounds and found that okra leaf shape can reduce the incidence of boll rot by 40%. Super-okra was found to be even more effective, showing a 53% reduction in boll rot (Jones 1970). Leaf shape mutants reduce the incidence of boll rot by providing a more open canopy structure that permits increased sunlight and air-flow to reach the interior of the cotton canopy. This alters the microclimate by reducing humidity, shade, and moisture and increasing the temperature, creating a less hospitable environment for microbial growth. Prolonged periods of warm, cloudy, and humid weather or extensive lodging can reduce the effectiveness of okra in combating boll rot (Jones 1982).

The effect of okra leaf on yield has produced conflicting results. Some researchers have recorded a slightly positive effect on yield. Jones (1982) reported that okra lines out-yielded their normal leaf isolines by an average of 5% across multiple tests in Louisiana (range -3% to +18%). Other studies showed that okra leaf had no effect on yield in Mississippi (Meredith 1985) or a slight decrease in yield in Arizona (Wilson 1986). An increase in harvested yield for okra leaf lines may be the result of reduced losses to boll rot rather than an increase in yield potential. On the other hand, the reduction in total leaf area seen in okra plants (60-65% of that produced by normal) may result in less than optimal light capture, reduced photosynthesis, and by extension, reduced yield under optimal conditions (Jones 1982, Wells and Meredith 1986). This helps to explain the conflicting results noted above. Boll rot is likely to be a significant factor affecting production in Louisiana and Mississippi, but not a concern in Arizona. Furthermore, super-okra, which produces approximately half as much leaf area as normal, yielded on average 8% less than normal leaf isolines in Louisiana (range -40% to +2%) (Jones 1982). This provides further evidence for the hypothesis above; since it appears that the gain in production through reduced boll rot losses seen in super-okra is not able to compensate for the reduced plant production even in environments where boll rot is a frequent threat. Boll and fiber quality parameters are not seriously enough affected by leaf shape to be of concern (Jones 1982).

It should also be noted that mutant leaf shapes show accelerated rates of flowering compared to normal types. Okra plants produce 50-100% more flowers, while super-okra flowers 85-150% more often than normal (Jones 1982). This increased rate of flowering does not contribute to any yield advantage as the mutants also show a significantly greater rate of boll shedding. However, this accelerated flowering contributes to early maturity and increased ability to compensate for losses induced by temporary biotic and abiotic stresses (Jones 1982). Okra leaf shape lines have also shown increased resistance to the whitefly and pink bollworm, presumably through the lethal effects of greater canopy temperatures on larvae (Jones 1982). On the other hand, okra lines appear more susceptible to plant bugs such as tarnished plant bugs and cotton flea hoppers, and are also subject to greater weed pressure due to increased light penetration (Jones 1982). Both of these disadvantages can easily be compensated for by chemical means.

There appears to be no significant difference in yield between sub-okra and normal leaf cotton: however, the increase in air movement and light penetration is not great enough to produce a significant reduction in boll rot in the sub-okra genotypes. The more extreme leaf shape mutants okra and super-okra display a significant reduction in boll rot but the decrease in leaf area is likely intense enough to result in a substantial loss of yield under ideal conditions. This limits the application of okra and super-okra lines to environments where boll rot is likely and severe enough to risk their reduction in agronomic potential.

The creation of a more open canopy structure as a means of disease control has been successfully deployed in other agricultural species. Blad et al. (1978) showed that planting cultivars and adopting cultural practices that resulted in a more open canopy reduce the rate of white mold (*Sclerotinia sclerotiorum*) in common bean *(Phaseolus vulgaris* L.) by increasing temperature and reducing moisture. Gubler et al. (1987) showed that the removal of the lower leaves of grape plants (*Vitis vinifera*) reduced the incidence of *Botrytis* bunch rot but could also result in reduced yields. In strawberry (*Fragaria x ananassa*), the open, thin canopy of cultivar "Sweet Charlie" significantly outperformed the industry standard "Camarosa," in resistance to anthracnose fruit rot (*Colletotrichum acutatum*) and other rots

(Shuman 2001). In contrast to "Sweet Charlie," "Camarosa" possesses a much thicker and dense canopy. "Sweet Charlie" showed reduced hours of leaf wetness and increased mean daily temperature, factors that were both implicated in the cultivar's resistance to the fungal infections (Shuman 2001). Similar to boll rot and leaf shape in cotton, the results of these studies showed that an open-canopy cultivar can out-yield a conventional type in environments characterized by high disease incidence. However, in years and locations where sufficient disease fails to develop, normal canopy cultivars have a significant yield advantage.

1.1.3 Leaf Shape

Nearly all tetraploid cotton cultivars possess the normal or broad leaf type. Along with normal (l), numerous mutant leaf types make up an allelic series located on chromosome 15 of the cotton D sub-genome including okra (L^o), sub-okra (L^u), Sea-Island (L^e), and super-okra (L^s) (Jones 1982). Okra leaf shape is characterized by reduced photosynthetic area per leaf, a more pronounced lobing pattern with deeper sinuses, and ectopic outgrowths/abnormal leaf margins (Figure 1). Sub-okra possesses a less pronounced lobing pattern with a more complete leaf margin exhibiting only the occasional abnormality (Figure 1). Lobes of sub-okra leaf also tend to have a slight upward curl to them. Sea-Island is a third leaf shape mutant found commonly in cultivars of *Gossypium barbadense*, and super-okra is a severe leaf shape mutant characterized by an almost compound leaf structure consisting of narrow strips of leaf emanating from the petiole. Another allelic series for leaf shape is found on chromosome 1 of the A sub-genome consisting of laciniate (L^L), narrow (L), intermediate (L^I) and two alleles for broad leaf (L^B and l) (Jones 1982). Laciniate leaf shape is similar in appearance to the okra leaf shape on chromosome 15 of the D sub-genome

The sub-okra allele was discovered around 1953 in a breeding nursery in South Carolina (Green 1953). It was found in a synthetic tetraploid derived from a cross between *G. arboreum* (AA) and *G. thurberi* (DD) originally developed by J.O. Beasley (Green 1953). Subsequent experiments showed that sub-okra belonged to the same allelic series as the leaf

shape alleles normal, okra, and super-okra and that this series was located in the D-sub-genome (Green 1953). Therefore, it was concluded that the sub-okra allele could have originated from *G. thurberi* (DD). The location of the allelic series was later corroborated by Jiang et al. (2000) via QTL analysis. In this study, 180 F_2 plants from a cross between wild-type *G. barbadense* and a quadruple mutant in the Deltapine 61 *G. hirsutum* background were used to identify QTL affecting leaf shape in cotton. The Deltapine 61 variety carried mutations conferring frego bract, okra leaf, nectariless, and glabrousness traits. Plants were phenotyped for 14 different traits related to leaf size and shape such as lobe lengths, widths, and angles (Jiang et al. 2000). In addition to the major leaf shape locus on chromosome 15, 14 QTL with a LOD≥3.0 (P≤0.001) and 41 with a LOD≥2.0 (P≤0.01) were found, 61.3% of which mapped to the D sub-genome (Jiang et al. 2000). These findings indicate that there appears to be a substantial amount of variation available in cotton leaf shape that could be exploited to develop novel leaf shapes (Jiang et al. 2000).

The sub-okra allele is also very similar to the Sea-Island allele (L^e) commonly found in *G. barbadense* (Meredith 1984). It has not been conclusively proven that these are different alleles. Differences between the two leaf shapes may be due to the increased size of *G. barbadense* leaves compared to *G. hirsutum*. The differences therefore, may result from loci other than the classic leaf shape locus.

The origin of the okra allele is less certain with its appearance in the literature dating to at least 1890 (Mell 1890) while the super-okra allele dates to at least 1930 (Kearney 1930). Okra leaf has been found as a mutation in both cultivated and wild cotton, while super-okra is thought to have originated as a second mutation in both wild and cultivated okra leaf cotton (Stephens 1945, Jiang et al. 2010). Phenotypically, both super-okra and okra leaves appear to be most similar to the leaves of *G. thuberi* or it close relative *G. trilobum* (Figure 2). Coupled with the evidence that the sub-okra allele is derived from *G. thurberi* (Green 1953), it is possible that all three of these mutant leaf shapes can trace their origin to *G. thurberi*. However, *G. thurberi* is not thought to be the D-genome donor to tetraploid cotton, a role normally reserved for *G. raimondii* (Endrizzi et al. 1985, Hanson et al 1996), although

some recent evidence implicates *G. gossypioides*, as the tetraploid D-genome donor (Zhao et al. 2012). Therefore, how alleles from *G. thurberi* could be present in *G. hirsutum* remains a mystery. The use of *G. thurberi* in cultivar improvement (Trolinder 2010), or a more complex polyploidization process (Zhao et al. 2012) might provide some explanations. Leaves of both *G. raimondii* and *G. gossypioides* are oval in appearance and provide no indication that they could be the source of the leaf shape alleles (Figure 2). Additionally, *G. herbaceum* and *G. arboreum*, the only two extant members of the A genome group, possess leaves more phenotypically similar to those of sub-okra (Figure 2).

The presence of an allelic series on chromosome 1 of the A sub-genome with a range of phenotypes similar to those seen on chromosome 15 in the D sub-genome indicate that perhaps the two loci are orthologous. Furthermore, chromosomes 1 (or A1) and 15 (or D1) are thought to be homeologs (Wang et al. 2006). This provides additional evidence that the loci may be orthologous genes with similar allelic variations. These mutations could predate the divergence from a common ancestor or could have arisen separately and become fixed in response to varying environmental conditions.

1.2 Introduction

The sub-okra leaf shape is phenotypically expressed at about the 10th node (Meredith 1984). Earlier nodes show normal leaves. Whatever mechanism (genetic promoter or otherwise) triggers the expression of sub-okra is not associated with any other leaf shapes. If the trigger mechanism is a promoter closely linked with the sub-okra allele, we hypothesize that it may be possible to create a transgressive segregate derived from a cross between sub-okra and okra. The resultant transgressive segregant would have normal-shaped leaves at the first ten nodes and then produce the highly-lobed leaves characteristic of the okra cultivars for the remainder of its life cycle. The advantage is to allow the plant to have maximum growth potential in the earlier stages of development and then provide a canopy that would allow more air movement and potentially reduce boll rot. With this in mind, the sub-okra line

NC05AZ06 was crossed to the okra-leaf line NC05AZ21 with the objective of looking for a possible transgressive segregant in the F_2 population.

In addition to the development of this novel phenotype, it would be interesting to determine the boll rot prevention ability of a line heterogeneous for both the sub-okra and the okra allele. F_1 hybrid plants of okra x sub-okra show intermediate leaf shape between their parents (Figure 1) Due to its self-pollinating nature, hybrid seed production in cotton is a difficult endeavor. Along with tedious hand pollinations, an experimental open-pollination nursery was used to develop two lines with various proportions of heterozygosity. The ability of these hybrid lines to combat boll rot was compared to both parents and a normal leaf check variety (NC10AZ10) in order to gauge the applicability of this approach. Even though isogenic lines are not being used, any differences in boll rot resistance can be attributed to leaf shape since no other mechanism has yet been described that confers resistance to this disease.

1.3 Materials and Methods

In order to carry out the open-pollination, ten plots each of the sub-okra line NC05AZ06 and the okra line NC05AZ21 were planted in alternating fashion in an isolation nursery in Clayton, NC during the summer of 2010. Pollination was allowed to occur naturally resulting in either self-pollination or cross-pollination by native insect pollinators. The two parents were harvested separately and the F_1 seed was bulked over plants belonging to the same line. This created two separate half-sib families that differed based on their female parent for use in the hybrid study. Hand-pollinations were carried out in a separate field in Clayton. A total of 1891 crosses were made between July 23 and August 16, 2010 with 963 using NC05AZ06 as the female parent and 928 using NC05AZ21 as the female parent. Hand-pollinations were harvested and processed independently but mixed thoroughly following acid de-linting and fungicide treatment.

In the summer of 2011, the hand-pollinated hybrid and both of the open-pollinated F_1 hybrids were tested against a normal leaf line (NC10AZ10) and both parental lines for boll

rot incidence. NC10AZ10 was included as a check to gauge the level of boll rot in a normal-leafed variety relative to that seen in the experimental entries. The test covered five locations with four replicates per location. The locations were: Tifton, GA, Quincy, FL, Alexandria, LA, and two tests at Fairhope, AL. All locations were chosen for their near annual incidence of boll rot. Rot at these locations is highly regular even without artificial promoters of boll rot such as excess irrigation or nitrogen fertilizer application. Each test was arranged in a 6x4 randomized complete block design (RCBD). Each plot was four rows wide and approximately 40 feet in length with 12 foot alleys. Quincy utilized six row plots in order to accommodate available planting equipment. Boll rot severity was measured by counting the ratio of rotted bolls to total bolls in a six foot subsection of each of the two interior rows of each plot. Data was analyzed using SAS Proc GLM and Proc ANOVA (SAS, Cary, NC).

An additional 500 F_1 seeds from the hand-pollinations between NC05AZ06 and NC05AZ21 were sent to a winter nursery in Mexico in late 2010 for selfing. Approximately 80,000 resulting F_2 seed was planted in bulk at the Upper Coastal Plains Research Station near Rocky Mount, NC in the summer of 2011. When the plants were approximately 1m tall (about mid-season), the entire field was examined to find an individual possessing the chance recombination of early normal leaf production switching to the okra leaf shape.

1.4 Results and Discussion

1.4.1 Weather Conditions during the 2011 Growing Season

The four locations chosen for this study were all chosen because of their near annual incidence of boll rot. Their locations on or near the Gulf of Mexico places them in a region characterized by high temperatures as well as some of the highest annual rainfall totals in the contiguous United States. These conditions create an especially hot and humid environment during the summer months that is ideal for the development of boll rot, limiting the productivity of the region. However, the summer of 2011 was characterized by slightly elevated temperatures and a pronounced lack of precipitation, especially during the month of August. This prolonged period of hot and dry weather led to one of the least productive

environments for the development of boll rot seen in some time (Table 1). In addition to the lack of boll rot, overall poor growth due to late planting and excessively hot and dry weather led to the abandoning of the trial in Alexandria. Additionally, damage from Tropical Storm Lee led to abundant lodging and rank growth at both tests in Fairhope. Best growth occurred in Quincy, the only site where yield data were taken. Across all locations, the simple ratio of rotted bolls to total bolls was 12%, well below the 20-30% that was hoped for and the 50-80% that can sometimes be seen in this region.

Table 1: Weather Data for 2011 Growing Season: Difference in the average maximum daily temperature and in cumulative precipitation for the period of June-August 2011 versus historical averages. In parentheses is the percentage of the average rainfall received in 2011. All weather data was gleaned from www.wunderground.com.

Location	**Difference in Avg. Max Temp.**	**Difference in Cumulative Precipitation.**
Tifton, GA	+3.0^{o}F	Unavailable
Quincy, FL	+4.0^{o}F	-11.53in (51.7%)
Fairhope, AL	+3.0^{o}F	-8.21in (50.4%)
Alexandria, LA	+5.0^{o}F	-4.59 (63.6%)

1.4.2 Results from Tifton, GA

Tifton experienced the smallest increase in temperature of all four locations and also a marked increase in cumulative precipitation (as measured ~42 miles away in Albany, GA) for the period June-August 2011. Nevertheless, Tifton had the lowest incidence of boll rot among the three sites where measurements were taken. Only 6.6% of the total bolls at this location were rotten. There were no significant pair wise differences between any of the varieties at the 0.05 significance level, although sub-okra did show significantly higher levels of boll rot when tested at $\alpha = 0.10$ (Table 2a and 2b). Results from a mean separation test

showed that block 3 had significantly more rot than block 4 but that no other block pairs differed significantly (Table 2c). Since blocks 3 and 4 were adjacent to each other and there were no other differences between blocks, there appears to be no easily observable explanation for this difference. Furthermore, while there were no significant differences among the varieties in terms of boll rot reduction, the order of the means behaved somewhat as expected with boll rot incidence decreasing along with total leaf area (Table 2b).

1.4.3 Results from Quincy, FL

Despite warmer than average temperatures and nearly half of the typical precipitation, Quincy showed the best overall growth of any of the four tests visited and is the only test for which yield data are available. Decent growth, along with a significantly elevated incidence of boll rot compared to nearby Tifton, may be the result of the irrigation supplied at this site. Surprisingly, the normal leaf check NC10AZ10 had significantly higher boll rot than the sub-okra variety but not any of the other four lines tested at the 5% significance level (Table 3a and 3c). When the significance threshold was increased to 10%, additional differences were detected. At this level, the hand-pollinated F_1 had a significantly lower incidence of rot than both the normal line and the okra entry. Sub-okra also showed significantly lower rot than okra (Table 3a). The poor performance of the okra line, especially relative to the sub-okra variety, may simply be an artifact of the study or may be attributed to the significantly higher total boll number okra displayed over sub-okra in this location (data not shown). Block 2 was shown to have a significantly higher incidence of boll rot compared to blocks 3 and 4 but not block 1. There were no other significant pair wise differences between blocks. In terms of the yield data collected from Quincy, neither variety nor block had any significant effect on yield either at the 0.05 or 0.10 significance levels (Table 4a-c).

1.4.4 Results from Fairhope, AL

The first test location in Fairhope had the highest incidence of boll rot, affecting almost 20% of the total bolls. Tropical Storm Lee, through its sustained high winds and brief flooding, led to pronounced lodging of the test, a development known to offset the boll rot

resistance of mutant leaf cultivars (Jones 1982). Additionally, the fourth block was thrown out after being overrun by weeds, which combined with the unfavorable weather to yield fewer than five total bolls in the two meter plots for some varieties. However, the okra line NC05AZ21, and its most phenotypically similar F_1 hybrid, the open-pollinated line NC05AZ21 x NC05AZ06, showed significantly less boll rot ($\alpha = 0.05$) than the normal leaf cultivar (Table 5a and 5b). No other pair wise combinations were significant, even at the 10% level, but the fact that okra leaf outperformed the normal leaf by almost 18% in terms of boll rot incidence is encouraging and agrees with previous reports in the literature (Jones 1970, 1982). The open-pollinated seed collected from NC05AZ21 also showed a highly significant 15% reduction in the incidence of boll rot compared to the normal leaf.

That this was observed in the environment showing the highest level of boll rot, and in spite of extensive lodging, indicates at least limited potential for the deployment of an open-pollinated cultivar as a means to combat boll rot. Differences in the incidence of boll rot may have been even stronger had environmental conditions been more conducive to the development of disease and/or had lodging not occurred. Since no measurements were taken on the amount of out-crossing that occurred in the open-pollination nursery in Clayton in the summer of 2010, no estimates are available on the frequency of the sub-okra allele in the seed harvested from the okra parent. Therefore, it is possible that a high degree of self-pollination occurred in the okra line, leaving the resulting F_1 seed with few copies of the sub-okra allele. This could explain the similarity of the performance of NC05AZ21 and Open-Pollinated (OP) NC05AZ21 x NC05AZ06 at the first Fairhope test location in the summer of 2011. However, there is no reason to believe that a substantial amount of out-crossing did not occur in the open-pollination nursery. The OP NC05AZ21 x NC05AZ06 is certain to contain at least some copies of the sub-okra allele and the data from Fairhope 1 indicates that whatever increase in leaf area caused by the presence of these sub-okra alleles did not significantly alter the resistance to boll rot. Any increase in the yield potential brought about by the increase in leaf area would be independent of boll rot resistance and may make the use of an open-pollinated hybrid a feasible alternative to normal leaf cultivars in regions

conducive to boll rot. Block had no effect on the incidence of boll rot at the Fairhope 1 test location (Table 5c).

Despite being located adjacent to Fairhope 1, the incidence of boll rot in Fairhope 2 was significantly lower (19.0% and 13.4%, respectively) but not significantly different than that of Quincy (13.4%). The difference between the two Fairhope tests was planting date, with Fairhope 2 being planted approximately two weeks after Fairhope 1 in the hopes of getting at least one test with sufficient boll rot. There were no significant differences detected among varieties at either the 0.05 or 0.10 significance levels (Tables 6a and 6b). No block effects were significant either (Table 6c).

1.4.5 Conclusions from Boll Rot Study

The strategy used here may present a more practical approach to hybrid seed production in cotton. One of the open-pollinated lines used the sub-okra line (NC05AZ06) as the female parent and was biased to some extent towards the sub-okra allele. The other utilized an okra female parent (NC05AZ21) and was somewhat biased towards the okra leaf shape. Fiber quality is unlikely to show significant differences in the F_1 generation and thus was not collected. Since the lines under comparison were not isogenic, differences in yield cannot be solely attributed to leaf shape and data were only taken where convenient. It was expected that boll rot resistance will be negatively correlated to the varieties total leaf area. Okra leaf was expected to show the best ability to fight boll rot followed by the three hybrids, which may not have been significantly different from each other or from okra. Ideally, one or both of the open-pollinated hybrids would have performed equally as well as the okra parent and/or the hand-pollinated hybrid in response to boll rot. Sub-okra and normal leaf were expected to show the lowest ability to cope with boll rot and were not expected to be significantly different from each other, although sub-okra may have had a slight advantage.

The inability to detect significant differences in the boll rot performance of the leaf shape varieties across this experiment is mostly explained by a lack of disease development. All locations selected in this study were chosen due to a near annual incidence of boll rot. Using four different locations, along with two different planting dates at one location, was

thought to be enough to ensure that adequate data would be collected. However, elevated temperatures and a lack of rainfall during the flowering and boll development periods throughout the region led to rotted boll ratios below the 20-30% minimum desired. Mean boll rot indices across all tests did fall into an order that fits somewhat with the hypothetical results, but differences were not nearly significant enough to rule out their occurrence by chance. A larger sample size may have been more effective at obtaining statistically significant results but an overall lack of growth due to the poor environmental conditions would likely have hindered the collection of additional data. In retrospect, an approach that favored the sampling of a fixed number of bolls per plot (400-600) instead of the number of bolls in a fixed length of row, as proposed by Cauquil (1975) may have been more appropriate. Such an approach would require more resources but may better protect the experiment against low productivity and insufficient sample size.

Nevertheless, the results from the one test that neared an appreciable level of boll rot (Fairhope 1 at 19.0%) show that the methods used are sufficient when boll rot incidence develops to an acceptable level. This held true even in the low yielding environment of Fairhope 1. That the OP NC05AZ21 x NC05AZ06 (biased towards okra) was more resistant than the check provides some incentive to further investigate the ability of an open-pollinated hybrid as a means to combat boll rot. This hybrid may be more resistant to boll rot than normal leafed lines but should not suffer as severe a yield potential penalty as a pure okra leaf cultivar.

1.4.6 Development of Transgressive Segregant

Vegetative phase change, also known as heteroblastic development, is a phenomenon that is fairly common among plant species (Dolan and Poethig 1991). Maize (*Zea mays*) possesses two basic leaf types, juvenile and adult, that differ in size, cell type, cuticle thickness, trichome presence, and waxy coating (Chandler et al. 2011). The first four nodes are generally occupied by juvenile leaves which transition into full adult leaves above node eight (Chandler et al. 2011). Nodes five through eight are occupied by a transitional zone

where leaves possess characteristics that are both juvenile and adult (Chandler et al. 2011). *Arabidopsis* undergoes a similar change in leaf shape (Willmann and Poethig 2011). Young *Arabidopsis* leaves are flat and round with a smooth margin, long petiole, and trichomes only on the top side of the leaf (Willmann and Poethig 2011). Mature leaves on the other hand are curled, pointed, and serrated with a short petiole and trichomes on both sides of the leaf (Willmann and Poethig 2011). Cotton is another species that displays pronounced heteroblasty with early heart-shaped leaves giving way gradually to a more pronounced lobing pattern that culminates in a climax leaf shape (Dolan and Poethig 1991).

The field containing ~80,000 F_2 individuals resulting from the original crosses between NC05AZ21 and NC05AZ06 were examined at mid-season. Despite thorough observation, an okra leaf cultivar displaying the "delay mechanism" previously described in sub-okra was not found. This could indicate that the linkage is so tight that the probability of obtaining the desired recombination event is miniscule. However, it may be that the delay mechanism does not exist. It may be possible that the expression levels of the sub-okra allele are modulated in a growth stage specific manner through an unknown mechanism. The lower leaves of sub-okra, while not nearly as lobed as the upper leaves, are more lobed than and easily distinguishable from those of normal leaves (Figure 1b and 1d). Furthermore, the F_1 hybrid between an okra leaf line and a sub-okra line also show a fairly pronounced increase in lobing as the node number increases (Figure 1c). This increase in lobing is also visible to some extent in okra leaf (Figure 1a). Therefore, it may be that the increase in lobing of higher node sub-okra leaves is a result of the heteroblastic leaf development of cotton and is not related to the classical leaf shape locus. This phenomenon of heteroblasty has previously been reported in cultivars of *G. barbadense* whose Sea-Island leaf shape is similar to and may not be genetically different from that of sub-okra (Dolan and Poethig 1991, Meredith 1984). Furthermore, it has been shown that the okra leaf mutation enhances and intensifies the heteroblastic pattern of normal leaf cotton (Dolan and Poethig 1991). Since okra and sub-okra are mutant alleles of the same gene, it is likely that sub-okra also enhances the heteroblastic effect, although to a lesser degree than okra. The genetic architecture of the

heteroblastic effect is well established in the model plant *Arabidopsis*. Small RNAs and genes involved in flowering time have been implicated in the heteroblasty displayed by *Arabidopsis* (Poethig 2010, Willman and Poethig 2011). However, there is no information available on the genetic mechanism underlying heteroblasty in cotton.

Figure 1: Leaf Shape in Cotton (*Gossypium hirsutum*): A) NC05AZ21 – classical okra leaf shape, note increased lobing pattern with abundant ectopic outgrowths B) NC05AZ06 – sub-okra leaf shape with reduced lobing pattern relative to okra and rare ectopic outgrowths C) F_1 hybrid between NC05AZ06 and NC05AZ21 showing intermediate lobing but return of abundant ectopic outgrowths D) Texas Marker-1 (TM-1) – normal or broad leaf type characteristic of most cotton cultivars E) F_1 hybrid leaf compared to parental types F+G) – obsolete broad leaf cultivar Dixie Triumph vs. F_1 hybrid and parental leaf type of NC05AZ21 and NC05AZ06 respectively.

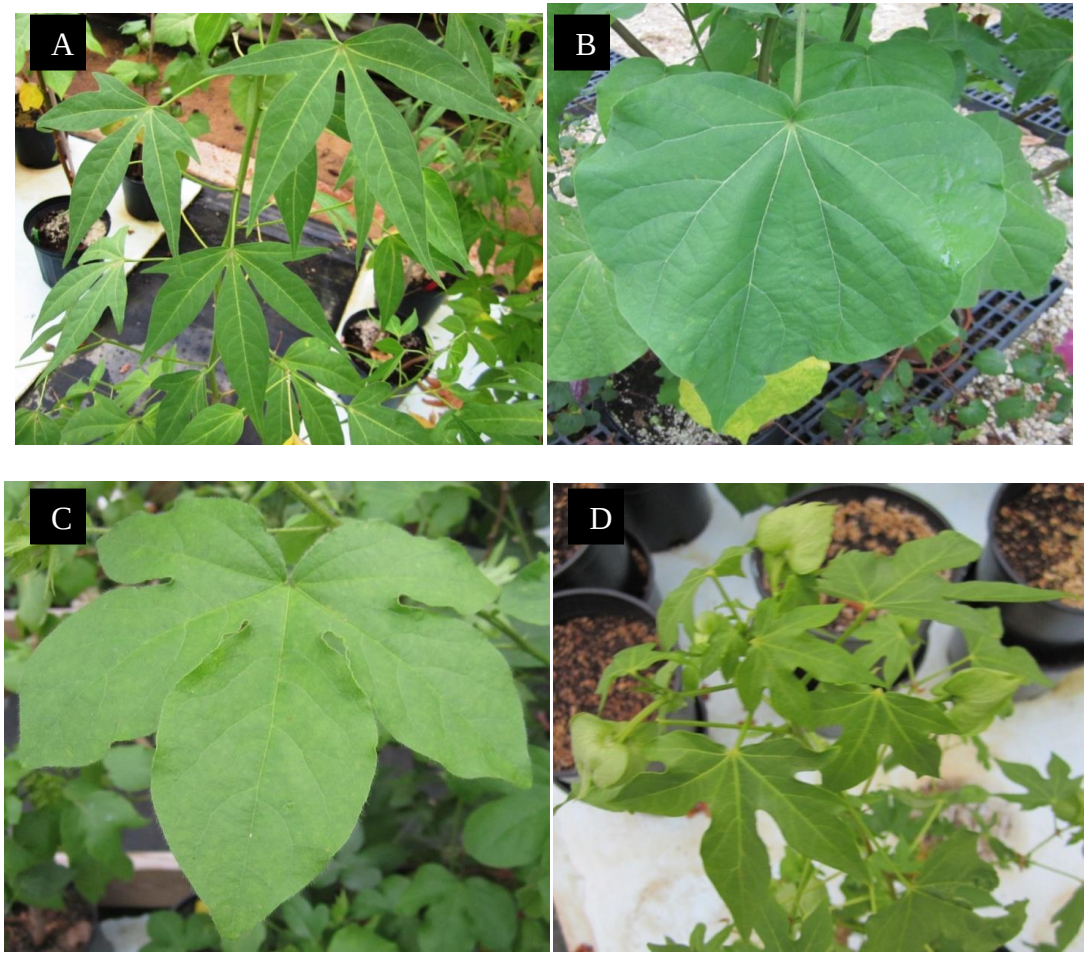


Figure 2: Leaf shapes of *Gossypium* diploids: A) *G. thurberi* accession NC0230, DI 37 (DD) B) *G. raimondii* accession NC0171, PI 530898 (DD) C) *G herbaceum* accession NC2086, PI 630017 (AA) D) *G. arboreum* accession NC0503, PI 529740 (AA).

Table 2a: Pairwise Orthogonal Comparisons for effect of variety on boll rot incidence – Tifton, GA. F-values followed by the probability of obtaining a greater F-value by chance (Pr > F) in parentheses. *Indicates significant at α = 0.10. **Indicates significant at α=0.05.

	NC05AZ06 (Sub-Okra)	NC05AZ21 (Okra)	Open-Pollinated NC05AZ21 x NC05AZ06	Open-Pollinated NC05AZ06 x NC05AZ21	Hand-Pollinated F_1	NC10AZ10 (Normal)
NC05AZ06 (Sub-Okra)						
NC05AZ21 (Okra)	3.58 (0.0780)*					
Open-Pollinated NC05AZ21 x NC05AZ06	1.26 (0.2802)	0.59 (0.4526)				
Open-Pollinated NC05AZ06 x NC05AZ21	0.96 (0.3423)	0.83 (0.3768)	0.02 (0.8908)			
Hand-Pollinated F_1	0.16 (0.6927)	2.22 (0.1573)	0.51 (0.4841)	0.33 (0.5719)		
NC10AZ10 (Normal)	0.59 (0.4539)	1.26 (0.2792)	0.12 (0.7301)	0.04 (0.8351)	0.13 (0.7195)	

Table 2b: Mean separation test for effect of variety on boll rot incidence in Tifton. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.000745
Critical Value of t	2.13145
Least Significant Difference	0.0411

t Grouping	Mean % Rot	Variety
A	0.08251	NC05AZ06 (Sub-Okra)
A	0.07473	Hand-Pollinated F_1
A	0.06767	NC10AZ10 (Normal)
A	0.06358	OP 06 x 21* F_1
A	0.06088	OP 21 x 06* F_1
A	0.04600	NC05AZ21 (Okra)

*OP – Open-pollinated, 06 – NC05AZ06, 21 – NC05AZ21

Table 2c: Mean separation test for effect of block on boll rot incidence in Tifton. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.000745
Critical Value of t	2.13145
Least Significant Difference	0.0336

t Grouping		Mean % Rot	Block
	A	0.07957	3
B	A	0.07347	1
B	A	0.06831	2
B		0.04224	4

Table 3a: Pairwise Orthogonal Comparisons for effect of variety on boll rot incidence – Quincy, FL. F-values followed by the probability of obtaining a greater F-value by chance (Pr > F) in parentheses. *Indicates significant at α = 0.10. **Indicates significant at α=0.05.

	NC05AZ06 (Sub-Okra)	NC05AZ21 (Okra)	Open-Pollinated NC05AZ21 x NC05AZ06	Open-Pollinated NC05AZ06 x NC05AZ21	Hand-Pollinated F_1	NC10AZ10 (Normal)
NC05AZ06 (Sub-Okra)						
NC05AZ21 (Okra)	4.52 (0.0506)*					
Open-Pollinated NC05AZ21 x NC05AZ06	0.24 (0.6318)	2.68 (0.1226)				
Open-Pollinated NC05AZ06 x NC05AZ21	0.91 (0.3540)	1.37 (0.2606)	0.22 (0.6470)			
Hand-Pollinated F_1	0.08 (0.7865)	3.42 (0.0841)*	0.05 (0.8338)	0.46 (0.5064)		
NC10AZ10 (Normal)	4.55 (0.0498)**	0.00 (0.9935)	2.70 (0.1208)	1.39 (0.2574)	3.45 (0.0829)*	

Table 3b: Mean separation test for effect of variety on boll rot incidence in Quincy. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.001926
Critical Value of t	2.13145
Least Significant Difference	0.0662

t Grouping		Mean % Rot	Variety
	A	0.16888	NC10AZ10 (Normal)
B	A	0.16862	NC05AZ21 (Okra)
B	A	0.13234	OP 06 x 21* F_1
B	A	0.11784	OP 21 x 06* F_1
B	A	0.11121	Hand-Pollinated F_1
B		0.10266	NC05AZ06 (Sub-Okra)

*OP – Open-pollinated, 06 – NC05AZ06, 21 – NC05AZ21

Table 3c: Mean separation test for effect of block on boll rot incidence in Quincy. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.001926
Critical Value of t	2.13145
Least Significant Difference	0.054

t Grouping		Mean % Rot	Block
	A	0.17433	2
B	A	0.14335	1
B		0.11323	3
B		0.10346	4

Table 4a: Pairwise Orthogonal Comparisons for effect of variety on yield (lbs of seed cotton per foot of row) in Quincy, FL. F-values followed by the probability of obtaining a greater F-value by chance (Pr > F) in parentheses. *Indicates significant at α = 0.10. **Indicates significant at α=0.05.

	NC05AZ06 (Sub-Okra)	NC05AZ21 (Okra)	Open-Pollinated NC05AZ21 x NC05AZ06	Open-Pollinated NC05AZ06 x NC05AZ21	Hand-Pollinated F_1	NC10AZ10 (Normal)
NC05AZ06 (Sub-Okra)						
NC05AZ21 (Okra)	0.20 (0.6622)					
Open-Pollinated NC05AZ21 x NC05AZ06	0.67 (0.4268)	1.59 (0.2260)				
Open-Pollinated NC05AZ06 x NC05AZ21	0.23 (0.6385)	0.86 (0.3695)	0.11 (0.7405)			
Hand-Pollinated F_1	1.80 (0.1998)	3.19 (0.0942)*	0.27 (0.6077)	0.74 (0.4024)		
NC10AZ10 (Normal)	0.12 (0.7387)	0.01 (.9171)	1.34 (0.2655)	0.67 (0.4254	2.83 (0.1134)	

Table 4b: Mean separation test for effect of variety on yield in Quincy. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.006124
Critical Value of t	2.13145
Least Significant Difference	0.1179

t Grouping		Mean (lbs./ft)	Variety
	A	0.64714	Hand-Pollinated F_1
	A	0.61813	OP 21 x 06* F_1
	A	0.59946	OP 06 x 21* F_1
	A	0.57293	NC05AZ06 (Sub-Okra)
	A	0.55412	NC10AZ10 (Normal)
	A	0.54826	NC05AZ21 (Okra)

*OP – Open-pollinated, 06 – NC05AZ06, 21 – NC05AZ21

Table 4c: Mean separation test for effect of block on yield in Quincy. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.006391
Critical Value of t	1.98969
Least Significant Difference	0.0448

t Grouping		Mean (lbs./ft)	Block
	A	0.62997	2
	A	0.61188	3
	A	0.58206	1
	A	0.53612	4

Table 5a: Pairwise Orthogonal Comparisons for effect of variety on boll rot incidence – Fairhope, AL 1. F-values followed by the probability of obtaining a greater F-value by chance (Pr > F) in parentheses. *Indicates significant at α = 0.10. **Indicates significant at α=0.05.

	NC05AZ06 (Sub-Okra)	NC05AZ21 (Okra)	Open-Pollinated NC05AZ21 x NC05AZ06	Open-Pollinated NC05AZ06 x NC05AZ21	Hand-Pollinated F_1	NC10AZ10 (Normal)
NC05AZ06 (Sub-Okra)						
NC05AZ21 (Okra)	1.57 (0.2388)					
Open-Pollinated NC05AZ21 x NC05AZ06	0.71 (0.4187)	0.17 (0.6910)				
Open-Pollinated NC05AZ06 x NC05AZ21	0.09 (0.7760)	2.39 (0.1534)	1.29 (0.2826)			
Hand-Pollinated F_1	0.11 (0.7450)	0.84 (0.3901)	0.26 (0.6218)	0..39 (0.5449)		
NC10AZ10 (Normal)	1.97 (0.1907)	7.06 (0.0240)**	5.05 (0.0484)**	1.24 (0.2924)	3.02 (0.1128)	

Table 5b: Mean separation test for effect of variety on boll rot incidence in Fairhope 1. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	10
Error Mean Square	0.006678
Critical Value of t	2.22814
Least Significant Difference	0.1487

t Grouping		Mean % Rot	Variety
	A	0.29202	NC10AZ10 (Normal)
B	A	0.21785	OP 06 x 21* F_1
B	A	0.19835	NC05AZ06 (Sub-Okra)
B	A	0.17604	Hand-Pollinated F_1
B		0.14208	OP 21 x 06* F_1
B		0.11477	NC05AZ21 (Okra)

*OP – Open-pollinated, 06 – NC05AZ06, 21 – NC05AZ21

Table 5c: Mean separation test for effect of block on boll rot incidence in Fairhope 1. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	10
Error Mean Square	0.006678
Critical Value of t	2.22814
Least Significant Difference	0.1051

t Grouping		Mean % Rot	Block
	A	0.24461	3
	A	0.16790	1
	A	0.15804	2

Table 6a: Pairwise Orthogonal Comparisons for effect of variety on boll rot incidence – Fairhope, AL 2. F-values followed by the probability of obtaining a greater F-value by chance (Pr > F) in parentheses. *Indicates significant at α = 0.10. **Indicates significant at α=0.05.

	NC05AZ06 (Sub-Okra)	NC05AZ21 (Okra)	Open-Pollinated NC05AZ21 x NC05AZ06	Open-Pollinated NC05AZ06 x NC05AZ21	Hand-Pollinated F_1	NC10AZ10 (Normal)
NC05AZ06 (Sub-Okra)						
NC05AZ21 (Okra)	0.12 (0.7321)					
Open-Pollinated NC05AZ21 x NC05AZ06	2.07 (0.1712)	1.18 (0.2936)				
Open-Pollinated NC05AZ06 x NC05AZ21	0.52 (0.4805)	0.14 (0.7131)	0.51 (0.4864)			
Hand-Pollinated F_1	0.56 (0.4644)	0.16 (0.6933)	0.47 (0.5030)	0.00 (0.9785)		
NC10AZ10 (Normal)	0.04 (0.8365)	0.02 (0.8915)	1.51 (0.2387)	0.26 (0.6151)	0.29 (0.5965)	

Table 6b: Mean separation test for effect of variety on boll rot incidence in Fairhope 2. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	10
Error Mean Square	0.006678
Critical Value of t	2.22814
Least Significant Difference	0.1487

t Grouping		Mean % Rot	Variety
	A	0.20904	OP 21 x 06* F_1
	A	0.14930	Hand-Pollinated F_1
	A	0.14691	OP 06 x 21* F_1
	A	0.11429	NC05AZ21 (Okra)
	A	0.10221	NC10AZ10 (Normal)
	A	0.08393	NC05AZ06 (Sub-Okra)

*OP – Open-pollinated, 06 – NC05AZ06, 21 – NC05AZ21

Table 6c: Mean separation test for effect of block on boll rot incidence in Fairhope 2. Means with the same letter are not significantly different.

Alpha	0.05
Error Degrees of Freedom	15
Error Mean Square	0.001926
Critical Value of t	2.13145
Least Significant Difference	0.054

t Grouping		Mean % Rot	Block
	A	0.18673	1
	A	0.14838	2
	A	0.10735	4
	A	0.09466	3

REFERENCES

Bell AA, Howell CR, Stipanovic RD. 2010. Chapter 18 – Cotton host-microbe interactions. In: Stewart MD, Oosterhuis DM, Heitholt JJ, Mauney JR, editors. *Physiology of Cotton* Springer Science + Business Media 187-205.

Blad BL, Steadman JR, Weiss A. 1978. Canopy structure and irrigation influence white mold disease and microclimate of dry edible beans. *Phytopathology* 68: 1431-1437.

Cauquil J. 1975. Cotton boll: rot laying out a trial of a method of control. New Delhi: Amerind Publishing Co. Pvt. Ltd.

Chandler MA, Riedeman ES, Tracy WF. 2011. Vegetative phase change in maize: biotic resistance and agronomic performance. *Plant Breeding Reviews* 34: 131-160.

Dolan L, Poethig RS. 1991. Genetic analysis of leaf development in cotton. *Development* 113: 39-46.

Endrizzi JE, Turcotte EL, Kohel RJ. 1985. Genetics, cytology, and evolution of *Gossypium*. *Advance in Genetics* 23: 271-375.

Green JM. 1953 Sub okra, a new leaf shape in Upland cotton. *The Journal of Heredity* 44: 229-232.

Gubler WD, Marois JJ, Bledsoe AM, Bettiga LJ. 1987. Control of *Botrytis* bunch rot of grape with canopy management. *Plant Disease* 71: 599-601.

Hanson RE, Islam-Faridi MN, Percival EA, Crane CF, Ji Y, McKnight TD, Stelly DM, Price HJ. 1996. Distribution of 5S and 18S-28S rDNA loci in a tetraploid cotton (*Gossypium hirsutum* L.) and its putative diploid ancestors. *Chromosoma* 105: 55-61.

Jiang C, Wright RJ, Woo SS, DelMonte TA, Paterson AH. 2000. QTL analysis of leaf morphology in tetraploid *Gossypium* (cotton). *Theoretical Applied Genetics* 100: 409-418.

Jones JE. 1970. The status of okra-leaf cotton. p. 55. Proceedings of the Beltwide Cotton Production Research Cotton Production Research Conference. National Cotton Council of America: Memphis, TN.

Jones JE. 1982. The present state of the art and science of cotton breeding for leaf-morphological types. p. 93-99. Proceedings of the Beltwide Cotton Production Research Conferences. National Cotton Council of America: Memphis, TN.

Kearney TH. 1930. Genetics of cotton: a survey of our present knowledge. *The Journal of Heredity* 21: 325-336.

Medrano EG, Bell AA. 2007. Role of *Pantoea agglomerans* in opportunistic bacterial seed and boll rot of cotton (*Gossypium hirsutum*) grown in the field. *The Journal of Applied Microbiology* 102: 134-143.

Mell PH. 1890. A microscopic study of the cotton plant. Bulletin No. 13 of the Agricultural Experiment Station of the Agricultural and Mechanical College, Auburn, AL.

Meredith WR. 1984. Influence of leaf morphology on lint yield of cotton - enhancement by the sub okra trait. *Crop Science* 24: 855-857.

Meredith WR. 1985. Lint yield genotype x environment interaction in Upland cotton as influenced by leaf canopy isolines. *Crop Science* 25: 509-512.

Poethig RS. 2010. The past, present, and future of vegetative phase change. *Plant Physiology* 154: 541-544.

Sadras VO, Felton GW. 2010. Chapter 20 – Mechanisms of cotton resistance to arthropod herbivory. In: Stewart MD, Oosterhuis DM, Heitholt JJ, Mauney JR, editors. *Physiology of Cotton* Springer Science + Business Media 213-228.

Shuman JL. 2001. *Anthracnose* fruit rot resistance in strawberry. PhD Dissertation, *North Carolina State University*

Stephens SG. 1945. A genetic survey of leaf shapes in new world cottons – a problem in critical identification of alleles. *Journal of Genetics* 46: 313-330.

Trolinder N. 2010. Chapter 33 – Cotton regeneration. In: Stewart MD, Oosterhuis DM, Heitholt JJ, Mauney JR, editors. *Physiology of Cotton* Springer Science + Business Media 379-382.

Wang K, Song X, Han Z, Guo W, Yu JZ, Sun J, Pan J, Kohel RJ, Zhang T. 2006. Complete assignment of the chromosomes of *Gossypium hirsutum* L. by translocation and fluorescence in situ hybridization mapping. *Theoretical Applied Genetics* 113: 73-80.

Wells R, Meredith Jr. WR. 1986. Normal vs. okra leaf yield interactions in cotton. II. Analysis of vegetative and reproductive growth. *Crop Science* 26: 223-228.

Willmann MR, Poethig SR. 2011. The effect of the floral repressor *FLC* on the timing and progression of vegetative phase change in *Arabidopsis. Development* 138: 677-685.

Wilson FD. 1986. Pink bollworm resistance, lint yield, and lint yield components of okra-leaf cotton in different genetic backgrounds. *Crop Science* 26: 1164-1167.

Zhao X, Si Y, Hanson RE, Crane CF, Price HJ, Stelly DM, Wendel JF, Paterson AH. 2012. Dispersed repetitive DNA has spread to new genomes since polyploidy formation in cotton. *Genome Research* 8: 479-492.

CHAPTER 2: Candidate Gene Analysis of Leaf Shape in Cotton

2.1 Review of Literature

Cultivated cottons (*Gossypium hirsutum* and *G. barbadense*) are allotetraploids (2n=4x=52, AADD) formed by the hybridization of two diploid progenitors, most closely related to *G. raimondii* (2n=2x=26, DD) and *G. herbaceum* (2n=2x=26, AA) or *G. arboreum* (2n=2x=26, AA) (Brubaker et al. 1999, Chen et al. 2007, Wendel et al. 2009). The genus consists of approximately 49 species with five allotetraploids (all AADD) and 43 diploids grouped into eight genomes labeled A through G plus K that are distributed evenly throughout the tropical and sub-tropical parts of the world (Wendel et al. 2009). Members of the A, B, E, and F genomes are found naturally in Africa and Asia while members of the D genome are native to the new world (Chen et al. 2007). The three remaining genomes, C, G, and K are indigenous to Australia (Chen et al. 2007). The hybridization event giving rise to all the allotetraploids occurred in the New World and all the polyploids are native to the Americas (Wendel et al. 2009). The current understanding is that an A genome diploid must have established in the New World, close to the D genome donor, and then subsequently died off following hybridization (Wendel et al. 2009). Four different species of cotton have been domesticated over the course of human history, all independently and for the same purpose (Smith and Cothren 1999). In addition to *G. hirsutum* and *G. barbadense*, the two remaining A genome diploids, *G. herbaceum* and *G. arboreum*, were domesticated in Asia, although their cultivation is almost non-existent today (Smith and Cothren 1999). All four exist naturally as perennial, photoperiodic sprawling bushes or small trees with minimal production of poor quality fiber (Wendel et al. 2009). Today, cotton is cultivated as an annual using photoperiod-insensitive, compact bushes with significantly enhanced yield and fiber quality (Wendel et al. 2009).

Despite its status as the world's leading natural textile source, as well as an important source of vegetable oil and seed meal, genomic resources are poor in cotton compared to other agronomic species like rice, wheat, and maize. The genomic sequence of cotton is not yet available and there are no good physical maps available for cotton research. However,

good expressed sequence tag (EST) resources are available for cotton (Udall et al. 2006). This enables the use of comparative genomic analysis as an effective way to utilize the genetic and genomic information from model plants and crop species to expedite the genetic and genomic analysis of cotton.

Identification and isolation of genes underlying desirable phenotypes can help utilize these genes more efficiently in crop improvement programs. The candidate gene approach is a promising strategy for discovering these causal genes, especially in species deficient in genomic resources. Utilizing information on gene function gathered from model organisms can provide clues as to which genes control a particular trait in an organism that is not as well-studied (Pflieger et al. 2001). The basic principle involves the use of the previously determined functional gene sequences to search for homologous genes in the species of interest. Variability in mRNA, protein, or enzyme activity levels between phenotypes lends strong credence to the notion that a putative candidate gene is in fact the causal gene (Pflieger et al. 2001). However, genetic transformation resulting in complementation of a deficient phenotype, over-expression, and/or silencing of a gene product is the only way to unequivocally validate a candidate gene (Pflieger et al. 2001). Nevertheless, candidate gene mapping approaches help establish the genomic location of the trait and lay the foundation to decipher molecular and biochemical processes involved in the trait under study.

Perhaps the most famous example of the candidate gene approach in plants involved the orthologous dwarfing genes *Reduced height-1* (*Rht-B1* and *Rht-D1*) in wheat and *dwarf-8* (*d8*) in maize (Peng et al. 1999). These genes are especially important because they underlie the dwarf phenotype in their respective crops that are responsible for the substantial yield increases of the Green Revolution (Peng et al. 1999). Another significant example involves the use of *FLOWERING LOCUS T* (*FT*) and *FLOWERING LOCUS C* (*FLC*) genes from *Arabidopsis* to identify similar loci influencing the photoperiod response in rice, barley, and wheat (Greenup et al. 2009). The candidate gene approach has also been used to identify genes involved in biochemical, quality, and disease resistance traits (Pflieger et al. 2001).

2.1.1 Class I KNOTTED1- LIKE HOMEOBOX (KNOX) Genes

KNOTTED1-LIKE HOMEOBOX (KNOX) genes are a homeodomain containing family of transcription factors that belong to the TALE super-family (Hake et al. 2004). The TALE is characterized by a three amino acid loop extension between the first and second helices of the homeodomain (Reiser et al. 2000). Homeodomains are encoded by a ~180bp long homeobox and are involved in the DNA-binding specificity of the protein (Hake et al. 2004). In addition to the homeodomain, KNOX proteins contain two other highly conserved domains, ELK and MEINOX, both of which are involved in protein-protein interactions (Reiser et al. 2000). KNOX proteins are closely related to the human myeloid ecotropic viral integration site (MEIS) proteins through a conserved N-terminus (Hake et al. 2004). This N-terminus is called the MEINOX domain after MEIS and KNOX (Hake et al. 2004). *KNOX* genes are divided into two classes (Class I and Class II) based on the conservation of the homeodomain, intron positioning, and expression pattern (Hake et al. 2004). Members of the Class I *KNOX (KNOX1)* subfamily are strongly expressed in meristem tissue but are generally excluded from developing organ primordia while Class II *KNOX (KNOX2)* genes display a more general expression pattern (Hake et al. 2004).

KNOX1 genes are required for the acquisition and maintenance of the shoot apical meristem (SAM) (Uchida et al. 2010). Severe mutations to certain *KNOX1* genes are effectively lethal, resulting in the termination of shoot growth shortly after germination and the occasional production of leaf tissue in place of the shoot (Uchida et al. 2010). The SAM is an indeterminate structure of totipotent stem cells which give rise to all of the above ground organs of a plant (Uchida et al. 2010). The indeterminacy of these stem cells requires the constant expression of *KNOX1* genes orthologous to the *Arabidopsis thaliana* gene *SHOOT MERISTEMLESS (STM)* (Uchida et al. 2010). Down-regulation of these *KNOX1* genes is one of the first steps in the development of leaves and other above ground organs (Uchida et al. 2010). The exclusion of *KNOX1* expression from the developing primordia is essential to the loss of indeterminacy and the development of the various determinate cell types (Uchida et al. 2010).

In simple leafed plants, this repression endures throughout the development of the leaf (Uchida et al. 2010). On the other hand, in compound leafed species, *KNOX1* gene expression is re-established leading to the development of more complex leaf architecture (Uchida et al. 2010). Numerous experiments have shown that over-expression of *KNOX1* genes in simple leafed species can cause increased leaf lobing, ectopic outgrowths, and ectopic meristem formation (Uchida et al. 2010). Over-expression in compound leafed species causes a dramatic increase in leaf compounding (Uchida et al. 2010). Based on these findings, it has been proposed that variation in *KNOX1* gene expression may be at least partially responsible for the wide variation of leaf shape seen in nature (Uchida et al. 2010). However, for this to be true, plants must have evolved a mechanism for altering the non-essential re-expression of *KNOX1* genes in developing leaf primordia while leaving the essential function of these genes in the SAM unchanged (Uchida et al. 2010).

Uchida et al. (2007) identified two conserved non-coding sequences (CNSs) in the 5' upstream region of the *STM* gene and its orthologs that persists across both simple and compound leafed species. This conservation extends across both monocots and dicots as well (Uchida et al. 2007). They illustrated that at least one of these elements, the K-box, is involved in whether *KNOX1* expression is re-established in leaf primordia but not in the initial down-regulation (Uchida et al. 2007). Furthermore, they showed that regulation of these *KNOX1* genes occurs in the same pathway as *ASYMMETRIC LEAVES1 (AS1)* and *ASYMMETRIC LEAVES2 (AS2)*, two genes previously indicated in leaf shape development (Uchida et al. 2007). *AS1* is a MYB transcription factor expressed in shoot lateral organs and with an expression pattern complementary to that of *KNOX1* genes (Uchida et al. 2010). *AS2* is a LATERAL ORGAN BOUNDARY (LOB)-domain transcription factor known to form a heterodimer with *AS1*. Together, these genes work together to suppress the expression of *KNOX1* genes (Uchida et al. 2010).

Briefly, the regulatory system as understood in *Arabidopsis* is: *STM* expression is controlled by some currently unknown factors that presumably act through the K-box. *STM* represses *AS1* which acts with *AS2* to down-regulate the expression of other *KNOX1* genes

including *KNAT1/BP* and presumably *KNAT2/ATK1* and *KNAT6* (Uchida et al. 2007). *KNAT3*, *KNAT4*, *KNAT5*, and *KNAT7* are all Class II *KNOX* genes and therefore are not implicated in leaf shape development. It is important to note that the K-box is conserved among simple and compound leafed plants so there is no evidence that a sequence alteration or mutation to the K-box itself alters leaf shape (Uchida et al. 2007). Instead it is possible that there is a mutation upstream of the K-box such as to an unidentified transcription factor or other protein that binds there (Uchida et al. 2007).

2.1.2 BEL1-like HOMEODOMAIN (BLH) Genes

The TALE super-family of proteins consists of two major groups; the KNOX proteins and the BELL or BEL1-like homeodomain proteins, of which there are 13 members in *Arabidopsis* (Kumar et al. 2007). All BELL members possess the same basic structure consisting of a DNA-binding domain and the BELL and SKY domains involved in interactions with the other proteins, especially KNOX (Kumar et al. 2007). These heterodimer formations are evolutionarily conserved and required for site-specific DNA binding (Kumar et al. 2007). Some BELL proteins like BELLRINGER (BLR) are positive regulators of *KNOX1* activity (Kumar et al. 2007). BLR forms heterodimers with STM and KNAT1 in the shoot to maintain the SAM and lead to normal plant development (Kumar et al. 2007). Two other BEL1-like HOMEODOMAIN (BLH) proteins are expressed in the leaves and negatively regulate *KNOX1* genes; SAWTOOTH1 (SAW1)/BLH2 and SAW2/BLH4 (Kumar et al. 2007). Saw1 and saw2 single mutants have no obvious phenotypic differences compared to the wild-type but a saw1 saw2 double mutant line showed increased leaf lobing and ectopic expression of *KNAT1* in these serrations (Kumar et al. 2007). These findings implicate that the two BLH proteins act redundantly to suppress *KNAT1* since it appears one can compensate for a mutation in the other (Kumar et al. 2007). Furthermore, over-expression of *SAW1* can rescue the *as1* mutant phenotype and reduce the ectopic expression of *KNAT1* in the leaf (Kumar et al. 2007). Therefore, it appears that

SAW1/BLH2 and SAW2/BLH4 work to ensure normal leaf development by repressing multiple *KNOX* genes in the developing leaf (Kumar et al 2007).

SAW1 and *SAW2* have 87% identity in their sequences, which combined with extensive sequence similarity in their surrounding regions and their positions on different chromosomes, indicates that one arose as a gene duplication event of the other (Kumar et al. 2007). This sequence similarity also implies that the two genes have similar expression patterns and regulatory elements which would support the finding of their redundancy (Kumar et al. 2007). A third BELL protein with significant sequence similarity to SAW1 and SAW2 is BEL1 (73% and 72% respectively) (Kumar et al. 2007). In *in vitro* experiments BEL1 was capable of binding STM, KNAT1, KNAT2 and KNAT5 while the more distantly related BLH1 was only capable of binding KNAT2 and KNAT5 (Kumar et al. 2007). These findings indicate that multiple BELL proteins are capable of binding multiple KNOX proteins but the exact binding partners may vary (Kumar et al. 2007). Since *BEL1* has a partially over-lapping expression domain with *SAW1/SAW2*, including in the leaves, it was postulated that BEL1 may also have a role in leaf shape development (Kumar et al. 2007). However, the *saw1 saw2 bel1* triple mutant failed to show any increase in leaf lobing indicating that *BEL1* and *SAW* genes maintain different functions in corresponding tissues or that mutations to *BEL1* can be compensated for by some unknown mechanism (Kumar et al. 2007).

The normal expression patterns of *SAW1/SAW2* and *KNOX1* genes are essentially mutually exclusive with *KNOX1* expression restricted to the stems, meristems, and reproductive organs in *Arabidopsis* (Kumar et al. 2007). *SAW* gene expression is concentrated in the leaves, sepals, and petals meaning that the two gene families are either never co-expressed, overlap only in the border regions, or are coincident only very early before being resolved (Kumar et al. 2007). Furthermore, the *saw1 saw2 as1* triple mutant showed increased leaf lobing compared to both the *saw1 saw2* double mutant and the *as1* single mutant (Kumar et al. 2007). This indicates that SAW1/SAW2 and AS1 negatively regulate *KNOX1* expression in independent but parallel pathways that may be separated

spatially and/or temporally (Kumar et al. 2007). It is likely that the *saw1 saw2 as2* triple mutant would mimic the *saw1 saw2 as1* triple mutant phenotypically (Kumar et al. 2007).

2.1.3 Modified KNOX Genes

The *PETROSELINUM (PTS)* gene of tomato (*Solanum lycopersicum)* encodes a novel type of *KNOX* gene whose protein consists of only a MEINOX domain and lacks the typical homeodomain required for DNA binding (Uchida et al. 2010). The MEINOX domain facilitates the homodimerization of *KNOX* gene products and also interacts with BELL (BEL-like homeodomain) transcriptional regulators (Uchida et al. 2010). Over-expression of *PTS* due to a single nucleotide deletion in the promoter region leads to greatly increased leaf complexity and compounding compared to wild-type tomato (Uchida et al. 2010). The gene has been found in other dicot species as well, including the *Arabidopsis* ortholog *KNATM* (Uchida et al. 2010).

The tomato STM ortholog LeT6 is known to interact with BIP (a BELL protein orthologous to SAW1 and SAW2 in *Arabidopsis*) (Uchida et al. 2010). The heterodimerization between BIP and LeT6 drives the normal leaf shape development of tomato by either directly repressing *KNOX1* gene expression or sequestering LeT6 (Uchida et al. 2010). PTS also has the ability to bind BIP, and by doing so prevents BIP from interacting with LeT6 (Uchida et al. 2010). Thus over-expression of *PTS* leads to an excess of PTS protein which sequesters a sufficient amount of BIP to prevent it from blocking LeT6 (Uchida et al. 2010). Unmolested LeT6 in turn suppresses *AS1* and *AS2* orthologs, leading to the activation of *KNAT1* and other *KNOX1* orthologs and a more complicated leaf shape (Uchida et al. 2010).

2.1.4 Cup-Shaped Cotyledon (CUC) Genes and Associated miRNAs

Leaves are initiated from the flanks of the self-maintaining and totipotent SAM. *CUP-SHAPED COTYLEDON1 (CUC1), CUC2,* and *CUC3* are *Arabidopsis* genes involved in the definition of organ boundaries originating from the SAM (Nikovics et al. 2006). *CUC1*

and *CUC2*, but not *CUC3* are targeted by a miRNA (*miR164*) that is transcribed from the *MIR164A, B,* and *C* genes (Nikovics et al. 2006). Mutations to the *MIR164A* gene deepen the serration of *Arabidopsis* leaf margins, while over-expression leads to leaves with smooth margins (Nikovics et al. 2006). Furthermore, modification of *CUC2* to disrupt *miR164* binding results in increased leaf lobing, but the modification of *CUC1* does not alter leaf shape (Nikovics et al. 2006). Additionally, *cuc2* knockout mutants have smooth leaves in both a wild-type and *mir164a* mutant background while the *cuc1* knockout mutant does not (Nikovics et al. 2006). These findings indicate that *CUC2* acts specifically to control the development of leaf margins (Nikovics et al. 2006).

CUC1, CUC2, and *CUC3* are all members of the NAC family of transcription factors and are required for SAM formation and the identification of organ margins (Nikovics et al. 2006). *MIR164A, MIR164B,* and *MIR164C* target six different members of the NAC family. *MIR164C* has been implicated in controlling the petal number in Arabidopsis while *MIR164B* does not yet have a determined role in the development of any above ground organ (Nikovics et al. 2006). The 3' ends of all *miR164s* are highly variable, resulting in precursor-miRNAs of varying length (Nikovics et al. 2006). However, all still have the ability to form stem-loop structures and be processed into mature miRNAs (Nikovics et al. 2006). Full activity of *miR164* is dependent on the proper functioning of a large promoter region (Nikovics et al. 2006). Insertions into this promoter increase leaf serration but not as severely as a knockout mutation (Nikovics et al. 2006).

CUC2 and *MIR164* are co-expressed in the margins of young primordia (Nikovics et al. 2006). The pattern of lobing is first determined independently and then the ratio of CUC2 versus *MIR164* determines the extent of the serrations (Nikovics et al. 2006). Transgenic over-expression of *MIR164* genes can result in smooth leaf margins, even in mutant backgrounds displaying highly lobed leaves (Nikovics et al. 2006). However, not even constitutive expression of *MIR164* genes can overcome the mutant leaf shape seen in lines in which CUC2 has been modified to evade *miR164* (Nikovics et al. 2006).

Further tests showed that the observed increase in leaf serration occurred independently of *KNOX1* expression (Nikovics et al. 2006). Elevated levels of KNAT1 or KNAT2 were not found in the leaves of plants displaying a defective *CUC2* regulatory system (Nikovics et al. 2006). This evidence implicates *CUC2* as a possible determinant of leaf shape independent of *KNOX1* throughout plant species (Nikovics et al. 2006). Other studies have indicated that CUC genes are positive regulators of *KNOX* gene expression and the exact relationship between CUC and KNOX is not yet fully understood (Koyama et al. 2010).

Classical miRNA theory states that target mRNAs are expressed at consistent levels and it is the variability in the expression of the miRNA that dictates repression. In this model it appears that both MIR gene and the target are expressed at the same time throughout development (Nikovics et al. 2006). Repression then depends on the copy number of the target relative to the miRNA and the efficiency with which the miRNA can bind the target (Nikovics et al. 2006). Currently, *Arabidopsis* is the only species studied to date that possesses both a *CUC1* and a *CUC2*, indicating that a gene duplication event specific to *Arabidopsis* could have given rise to this duality. Therefore, the role of CUC2 could have evolved only in *Arabidopsis*, making it possible that this mechanism is not found elsewhere in plants (Nikovics et al. 2006).

2.1.5 TEOSINTE BRANCHED1, CYCLOIDEA, AND PCF (TCP) Transcription Factors

CINCINNATA-like (CIN-like) TEOSINTE BRANCHED1, CYCLOIDEA, and PCF (TCP) transcription factors are negative regulators of CUC genes, indicating them in leaf shape development (Koyama et al. 2010). A model TCP from *Arabidopsis*, *TCP3*, directly activates the expression of two previously discussed leaf shape genes; *MIR164A* and *AS1* (Koyama et al. 2010). *TCP3* also activates *INDOLE-3-ACETIC ACID3/SHORT HYPOCOTYL2* (*IAA3/SHY2*), and *SMALL AUXIN UP RNA* (*SAUR*) (Koyama et al. 2010). *IAA3/SHY2* represses the auxin response and *SAUR* has been shown in rice to prevent auxin

accumulation and transport (Koyama et al. 2010). Since *TCP3* induces the expression of these genes, it is likely that *TCP3* acts to repress the auxin response (Koyama et al. 2010).

Over-expression of TCPs generates fused cotyledons and suppresses the development of shoot meristems while loss-of-function leads to ectopic expression of *CUC* genes (Koyama et al. 2010). *Arabidopsis* has 24 TCP genes and this redundancy inhibits the study of their function (Koyama et al. 2010). Mutations to individual TCP genes have little to no effect on phenotype while double, triple, quadruple, and quintuple mutants show a dose-dependent increase in the waviness of leaf margins, fusion of cotyledons, and excess leaf growth (Koyama et al. 2010). The miRNA *miR319A* is a negative regulator of TCP gene products and its over-expression exacerbates the mutant phenotype (Koyama et al. 2010).

The CIN-like TCP transcription factors seem to influence leaf shape through three separate but parallel pathways (Koyama et al. 2010). First, CIN-like TCPs activate the expression of *AS1* as well as some of its binding partners including *AGO7/ZIPPY*, *SAW*, and *GAI* which act cooperatively to promote cell differentiation mainly through the negative regulation of the *KNOX1* genes (Koyama et al. 2010). Second, these TCPs activate the expression of *MIR164*, transcribing miRNAs that directly target the *CUC* genes (Koyama et al. 2010). Finally, CIN-like TCPs activate repressors of the auxin response, a finding that initially seemed at odds with the previously known antagonistic relationship between auxin levels and *KNOX1* gene expression (Koyama et al. 2010). It has been hypothesized that TCPs maintain leaf cell auxin concentrations within a narrow range, maintaining these cells sensitivity to auxin signaling (Koyama et al. 2010).

2.1.6 FLORICAULA/LEAFY (FLO/LFY) Orthologs

Across vascular plants, *KNOX1* gene expression in developing leaf primordia is required for the development of compound leaves. A noticeable exception occurs in the inverted repeat-lacking clade (IRLC) of the Fabaceae, where orthologs of the pea (*Pisum sativum*) gene *UNIFOLIATA (UNI)* have evolved to assume the role of the *KNOX1* genes (Champagne et al. 2007).

UNI is orthologous to floral regulators *FLORICAULA (FLO)* and *LEAFY (LFY)* from snapdragon (*Antirrhinum majus*) and *Arabidopsis* respectively (Champagne et al. 2007). *FLO/LFY* genes have previously been implicated in the transition from vegetative to reproductive growth and are required for the development of flowers (Champagne et al. 2007). Reducing the expression of *FLO/LFY* orthologs outside the IRLC (in soybean, *Glycine max*) resulted in only slightly less compound leaves while over-expression of the tomato *STM* ortholog *LeT6* in the IRLC member alfalfa (*Medicago sativa*) increased leaflet number (Champagne et al. 2007). This indicates that *KNOX1* targets are still present in alfalfa, and presumably the rest of the IRLC, and that they remain sensitive to *KNOX1* genes (Champagne et al. 2007). It is therefore possible that *FLO/LFY* orthologs and *KNOX1* could have had at least partially over-lapping roles in compound leaf development in ancestral Fabaceae but that *KNOX1* gene expression was lost in the developing primordia of the IRLC (Champagne et al. 2007). The role was then assumed completely by *FLO/LFY* orthologs (Champagne et al. 2007). It is not currently known how much over-lap exists between *KNOX1* and *FLO/LFY* orthologs in the developing compound leaves of the IRLC and other species (Champagne et al. 2007). The genes could have the same targets, creating a dosage effect whereby any increase in expression causes an increase in leaf complexity (Champagne et al. 2007). They could also function in separate but complementary pathways that both promote leaf complexity or have a partial overlap existing somewhere in between these two extremes (Champagne et al. 2007).

2.2 Introduction

Two loci are known to affect leaf shape in cotton; one is the classical leaf shape locus on chromosome 15 of the D sub-genome and the other is on the homeologous chromosome 1 of the A sub-genome (Jones 1982, Wang et al. 2006). Both loci are multi-allelic and carry alleles that conferring altered leaf types with varying levels of leaf lobing. The leaf shape mutants at both loci are characterized by a more complex leaf architecture featuring increased lobing and more variable margins compared to normal or broad-leaf types. This lends credence to the notion that the two loci are orthologous genes with phenotypically

similar mutations at homeologous chromosomes. These mutations may or may not predate a common ancestor or the polyploidization event.

The Class I *KNOTTED1- LIKE HOMEOBOX (KNOX)* genes are a family of transcription factors known to influence leaf shape in numerous species including tomato, *Arabidopsis*, tobacco, and potato (Lincoln et al. 1994, Janssen et al. 1998, Tanaka-Ueguchi et al. 1998, Rosin et al. 2003, Uchida et al. 2010). Their main role is to maintain the indeterminacy of the shoot apical meristem (SAM) and they are down-regulated at the sites of cell differentiation and organ development (Uchida et al. 2010). Increased and/or ectopic expression of *KNOX1* genes are associated with increased leaf complexity, ectopic outgrowths, leaf lobing, and leaf compounding, presumably through a brief re-establishment of cell indeterminacy (Uchida et al. 2010). Furthermore, alterations in the expression of numerous positive and negative regulators of *KNOX1* genes have been shown to alter the levels of members of *KNOX1* and in turn, leaf shape.

In addition to normal leaf, the multi-allelic series on chromosome 15 contains alleles for sub-okra, okra, super-okra, and Sea-Island. The increased lobing of these mutants may be the result of a prolonged period of indeterminacy in the developing leaf primordia. Furthermore, okra leaves frequently display ectopic outgrowths and abnormal leaf margins, another hallmark of an inability to properly regulate cell differentiation (Figure 1). Sub-okra leaves display a lesser degree of leaf lobing and very few ectopic outgrowths (Figure 1). However, the F_1 progeny of a sub-okra by okra cross shows lobing that appears to be additive in nature as well as the return of abundant ectopic outgrowths (Figure 1).

Since the control of leaf shape is known to be under the control of a single gene (Jones 1982), the re-establishment of expression of a *KNOX1* gene early and briefly in developing cotton leaf primordia is a plausible explanation for the increased leaf lobing seen in cotton mutants. Ectopic outgrowths and abnormal leaf margins could be caused by additional *KNOX1* expression at later points in development. Mutations to *KNOX1* genes or any of its positive or negative regulators appear to be good candidate genes influencing cotton leaf shape. In the present study, expression levels of the candidate genes implicated in

leaf shape and development in model plants and crop species were investigated in normal and mutant leafed cotton plants. Relatively minor changes in DNA sequence at a single locus could drastically alter phenotype in a model befitting a single gene effect. Mutations resulting in varying levels of expressional changes could be responsible for the varying degree of leaf shape alterations.

2.3 Materials and Methods

mRNA coding sequences of candidate genes identified in the literature were obtained from the NCBI databases (http://www.ncbi.nlm.nih.gov/) in the FASTA format. Sequences were then BLAST searched against the DFCI Cotton Gene Index (http://compbio.dfci.harvard.edu/tgi/cgi-bin/tgi/gimain.pl?gudb=cotton) to search for Expressed Sequence Tags (ESTs) and tentative contigs (TCs) that showed a significant amount of similarity to the candidate genes (Tables 7-13). The significance threshold used in this study required that the Smallest Sum Probability be less than or equal to 1.0e-50. In cases where no tentative contigs (TCs) met this significance level, but a candidate gene was desired anyway, the highest scoring contig was used. EST sequence was then used along with the Primer 3 software (http://frodo.wi.mit.edu/) to design primers specific to each TC (Table 9). Primers were ordered from Integrated DNA Technologies (Coralville, IA, USA).

RNA samples were collected from cotton plants approximately 1m tall grown under greenhouse conditions. Shoot or SAM-enriched tissue consisted of the top 1-2cm of the main stem with any developing leaves removed. Primordial tissue samples were made up of developing leaves no greater than 2cm in maximum length harvested only off of the main stem. Mature leaf tissue was taken from a central section of fully expanded leaves. Samples were collected in 1.7mL micro-centrifuge tubes, frozen in liquid nitrogen, and ground using a micro-pestle. RNA was isolated from 100mg samples using the Spectrum ™ Plant Total RNA Kit (Sigma-Aldrich, St. Louis, MO, USA) and quantified on a NanoDrop 2000 UV-Vis spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). Briefly, tissue samples were lysed, cellular debris was pelleted, and the lysate was filtered out. Then RNA was

bound to the provided column, treated with DNase, and washed three times with the provided solutions. Following the wash steps, the column was dried and eluted prior to spectrophotometry. Quantification results were used to dilute the RNA concentration to approximately 125ng/µL and dilutions were confirmed using spectrophotometry.

RNA samples were then converted to cDNA using the ImProm-II™ Reverse Transcription System (Promega Corporation, Fitchburg, WI, USA). Experimental RNA (250ng at 125ng/µL) were combined with the supplied Oligo$(dT)_{15}$ primer (0.5µg/reaction) and treated at 70^{o}C for five minutes in an Eppendorf Mastercycler EP Gradient thermo-cycler (Eppendorf International, Hamburg, Germany). Following five minutes at 4^{o}C in the thermo-cycler, samples were spun briefly in a micro-centrifuge to collect any condensate. The reverse transcription reaction mixture consisting of ImProm-II^{TM} 5X reaction buffer, $MgCl_2$ (3mM), dNTP mix (0.5mM each dNTP), recombinant RNasin ribonuclease inhibitor (20 units), and ImProm-II^{TM} reverse transcriptase were added to each sample. Tubes were treated at 25^{o}C for five minutes in the thermo-cycler in order to allow primers to anneal. Extension was carried out at 42^{o}C for one hour and then reverse transcriptase was thermally inactivated by 15 minutes of 70^{o}C treatment. Inactivation of the reverse transcriptase is a necessary step prior to PCR amplification of the cDNA.

cDNA samples were then used as templates in PCR reactions in a semi-quantitative approach to measure mRNA copy number and, by extension, transcription levels of the previously identified genes. A volume of 1.0µL of the reverse transcription reaction was combined with 10X thermophilic standard polymerase reaction buffer (with 15mM $MgCl_2$) (Apex^{TM} Bioresearch Products, San Diego, CA, USA), dNTP mix (0.2mM), forward and reverse primers (1µM each), and *Taq* DNA polymerase (Apex^{TM} Bioresearch Products, San Diego, CA, USA) in 30uL reactions. The PCR protocol started with two minutes at 94^{o}C and continued with cycles consisting of one minute at 94^{o}C, one minute at 60^{o}C, and two minutes at 72^{o}C. Following the 40th cycle, the samples were treated for five minutes at 72^{o}C. PCR products were mixed with 25% of their volume in 1X loading dye and nine µL of product was loaded and separated on 2% agarose gels along with five µL of Quanti-Marker 100bp

ladder (Bioline GeneMate London, UK). Gels were imaged using an InGenius Manual Gel Documentation System (Syngene Bio Imaging, Bangalore, India).

Glyceraldehyde-3-phosphate dehydrogenase (GAPDH) was used as a positive control in all expression analyses. This enzyme is involved in glycolysis and thought to be expressed constantly in identical tissue types across cotton cultivars. It is also expressed relatively constantly across tissue types both within and among cotton varieties justifying its use as a positive control in these expression studies (Rich Tuttle, personal communication).

2.4 Results

2.4.1 Class I KNOX Genes

There are four Class I *KNOX* genes in *Arabidopsis*: *STM*, *KNAT1*, *KNAT2*, and *KNAT6* (*KNAT3-5* and *7* are Class II *KNOX* genes, a separate but related family that are not implicated in leaf shape) (Uchida et al. 2010). All other *KNOX1* genes identified in other species are orthologous to one of these four. Of the four *Arabidopsis KNOX1* genes, *STM* and *KNAT1* both yielded a pair of highest scoring contigs when blast searched in the cotton DFCI gene index (Table 7). Both *KNAT2* and *KNAT6* yielded a quartet of highest scoring contigs, but the quartets of both genes consisted of the same four members in a slightly different order (Table 7). Blasts conducted with *STM* orthologs from snapdragon (*AmSTM*), tobacco (*NTH15*), tomato (*Tkn2/LeT6*), and soybean (*SBH1)* yielded the same pair of highest scoring contigs as *STM* (data not shown). Queries with other Class I *KNOX* genes from apple (*KNAP1*-2) and potato (*POTH1*) failed to detect any additional contigs either (data not shown).

Figure 3 shows the expression pattern of a handful of selected Class I *KNOX* orthologs across four different tissue types (SAM-enriched stem, developing primordia, young leaf, and mature leaf) and the two cultivars used: the okra-leafed NC05AZ21 and the normal-leafed TM-1. The GAPDH positive control shows uniform expression over all four tissue types and both cultivars as expected (Figure 3). For TC203839, the highest scoring TC of *KNAT1*, both the okra and TM-1 variety show strong expression in the stem tissue (Figure

3). However, the okra variety shows fairly strong expression in the primordia and faint expression in the mature leaf while expression is absent from all other TM-1 tissues (Figure 3). TC204345 was highest scoring candidate for *STM*. It showed the strongest expression in the stem of both varieties and was fairly similar between the two (Figure 3). Surprisingly, TC204345 showed fainter expression in all three of the remaining tissue types although expression appears marginally stronger in the okra-leafed mutant (Figure 3). TC222827 was the second best match with *STM*. Again, its expression is strongest in the stem and appears equally strong in both of the varieties tested (Figure 3). The okra-leaf variety NC05AZ21 shows expression in both the primordial tissue and the mature leaf, but not in the young leaf (Figure 3). On the other hand, the normal leaf variety TM-1 does not show any expression outside of the stem, although there appears to be a good level of non-specific amplification (Figure 3).

TC203687 was the third highest TC after the blast with *KNAT1*. This placed it outside of the top pair of TC's. Nevertheless, it was still chosen for this expression analysis because it exceeded the high score for similarity set for this study of 1.0e-50 (Table 7). It also showed strong expression in the stem that was fairly uniform between the two varieties and gave a very faint band in the primordia of the okra-leaf mutant (Figure 3). Expression was not detected in any other tissues for either the mutant or the wild-type (Figure 3). The final TC chosen for the *KNOX1* expression analysis was DW236843. This TC was flagged in blasts with both *KNAT2* and *KNAT6* and showed similar scores with two other members of the quartets, ES810183 and ES843392 (Table 7). Like all the other *KNOX1* putative orthologs tested, this TC was expressed strongest in the stem and was equal in both the mutant and the normal leaf (Figure 3). Expression was also fairly strong in the primordia of NC05AZ21, and this expression continued at slightly weaker levels into both the young and mature leaves of the okra variety (Figure 3). This was the only time that the expression of a *KNOX1* putative ortholog was detected in the young leaf of either variety (Figure 3). Surprisingly, expression of DW236843 was also detected in the normal leaf primordia

(Figure 3). However, expression was absent from the mature leaf and appeared as a very faint, barely visible band in the young leaf (Figure 3).

2.4.2 AS1/AS2 Orthologs

Sequences were obtained for five different MYB transcription factors including *RS2* from maize, *AS1* from *Arabidopsis*, *PHANa* and *PHANb* from soybean, and *PHAN* from snapdragon. These sequences indentified two potentially orthologous sequences in the Cotton DFCI database and primers were designed for both (Table 7). Additionally, the sequence of the *AS2* LOB Domain Factor from *Arabidopsis* was queried against the database. No contigs met the desired significance level but a primer was designed for the lowest scoring contig ES819678 (4.9e-19).

Expression patterns for all three of these primers along with the GAPDH positive control are shown in Figure 4. There appears to be no difference in the expression of one of the two putative *AS1* orthologs, TC248918. Expression of this TC was strong and constant across the three tissue types tested and over both cultivars used (Figure 4). It even rivaled the GAPDH positive control for consistency across all tissue types. For the other putative *AS1* ortholog, expression was found in all three tissue types in both cultivars (Figure 4). Expression was fairly even across the three tissue types within each cultivar (Figure 4). However, the okra-leafed variety showed slightly stronger expression in all three tissues (Figure 4). For the weakly orthologous ES819678, expression again appears slightly stronger in the okra-leafed NC05AZ21, especially in the primordia tissue when compared to the TM-1 normal leaf (Figure 4).

2.4.3 BLH Genes

The sequences of two proposed cotton *BLH* transcription factors were obtained based on their similarity to one tomato (*BIP)* and two *Arabidopsis BLH* transcription factors (*BLH2/SAW1* and *BLH4/SAW2*) (Table 7). Primers were designed for both and the results of their expression patterns can be seen in Figure 5. Candidate TC820944 showed fairly

consistent expression across all tissue types and cultivars tested. Minor differences in expression were observed but varied with repetition indicating that it was a result of the experimental procedure used. The other candidate contig ES838507 showed variable expression over both tissue type and cultivar but did so in a non-repeatable way that varied with replication.

2.4.4 Modified Class I KNOX Genes

Neither the *Arabidopsis KNATM* nor the tomato *PTS* modified *KNOX1* gene had any significantly scoring contigs within the DFCI Cotton Gene Index (Table 7).

2.4.5 CUC Genes

The sequences of three different *CUC* genes, all from *Arabidopsis*, discovered two putative *CUC* orthologs in cotton (Table 7). Primers were made for both of the genes, but one (TC195788) failed to amplify. The other candidate (TC221650) did not show any variability in expression. miRNAs associated with CUC genes were not used in blasts because they are not long enough to generate significant results.

2.4.6 TCP Transcription Factors

Four different *TCP* genes from *Arabidopsis* (*TCP3, 4, 5,* and *10)* along with one from tomato (*Lanceolate*) were used to identify three different putative orthologs in the cotton genome (Table 7). All three primers showed bands at the expected product size. There was no significant or repeatable variation between any of the tissue types or cultivars.

2.4.7 FLO/LFY Floral Regulators

FLO/LFY floral regulators from pea (*UNI*), *Arabidopsis* (*LFY* and *UFO*), and snapdragon (*FLO*) were blasted against the DFCI Cotton Gene Index. No significant matches were obtained (Table 7).

Table 7: Known modifiers of leaf shape and their highest scoring ESTs/TCs/RFLPs in the DFCI Cotton Gene Index.

Gene(Reference)	Plant	Description	Highest Scoring	BLASTn (e-value)
STM(NM104916)	*Arabidopsis*	Class I KNOX	TC254864(TC204345)	2.40e-93
			TC273462(TC222827)	1.5e-87
			TC254362(TC203839)	7.9e-53
KNAT1/BP(NM116884)	*Arabidopsis*	Class I KNOX	TC254362(TC203839)	5.0e-79
			TC255316(TC204796)	7.6e-79
			TC254202(TC203687)	2.5e-57
KNAT2/ATK1(NM105719)	*Arabidopsis*	Class I KNOX	TC254507(TC203990)	3.2e-60
			ES843392	1.1e-50
			ES810183	3.8e-50
			DW236483	5.9e-50
			TC273538(TC222904)	3.0e-43
KNAT6(NM102187)	*Arabidopsis*	Class I KNOX	TC254507(TC203990)	5.8e-71
			DW236483	8.0e-59
			ES843392	8.2e-59
			ES810183	7.8e-49
			TC255316 (TC204796)	1.5e-32
RS2(NM001112039)	Maize	MYB TF	TC247240(TC196751)	3.3e-37
AS1(NM129319)	*Arabidopsis*	MYB TF	TC248918(TC198414)	1.1e-41
PHAN(AJ005586)	Snapdragon	MYB TF	TC248918(TC198414)	6.7e-102
		MYB TF	TC247240(TC196751)	3.3e-66
PHANa(AY790252)	Soybean	MYB TF	TC247240(TC196751)	8.4e-66
			TC248918(TC198414)	9.4e-61

Table 7 Continued:

PHANb(AY790253)	Soybean	MYB TF	TC248918(TC198414)	1.0e-109
AS2(NM001124081)	Arabidopsis	LOB Domain Factor	ES819678	4.9e-19
BIP(EU352654)	Tomato	BLH TF	ES820944	4.6e-86
			ES838507	7.8e-85
BLH2/SAW1(NM001036720)	Arabidopsis	BLH TF	ES820944	5.0e-63
			ES838507	1.5e-62
BLH4/SAW2(NM179713)	Arabidopsis	BLH TF	ES820944	3.8e-66
			ES838507	3.9e-66
KNATM(NM001160868)	Arabidopsis	Modified KNOX Gene	AI054970	0.31
PTS(EU352653)	Tomato	Modified KNOX Gene	TC261486(TC210962)	0.021
CUC1(NM112380)	Arabidopsis	CUC Gene	TC272289(TC221650)	5.1e-49
CUC2(NM124774)	Arabidopsis	CUC Gene	TC272289(TC221650)	5.4e-53
			TC246269(TC195788)	7.8e-50
CUC3(AF543194)	Arabidopsis	CUC Gene	CO102970	3.7e-31
TCP3(NM104201)	Arabidopsis	TCP TF	TC248503(TC197998)	5.8e-47
TCP4(NM112365)	Arabidopsis	TCP TF	TC248503(TC197998)	2.4e-63
TCP5(NM125490)	Arabidopsis	TCP TF	TC236882(TC186604)	7.6e-33
TCP10(NM128662)	Arabidopsis	TCP TF	TC248503(TC197998)	3.1e-27
Lanceolate(EF091573)	Tomato	TCP TF	TC248503(TC197988)	1.2e-74
			TC237216(TC186920)	7.7e-74
			TC245172(TC194705)	1.9e-66
UNI(AF035163)	Pea	FLO/LFY	ES799148	2.5e-05
LFY(NM125579)	Arabidopsis	FLO/LFY	None	N/A
UFO(NM102834)	Arabidopsis	FLO/LFY	None	N/A
FLO(M55525)	Snapdragon	FLO/LFY	TC245095(TC194609)	0.038

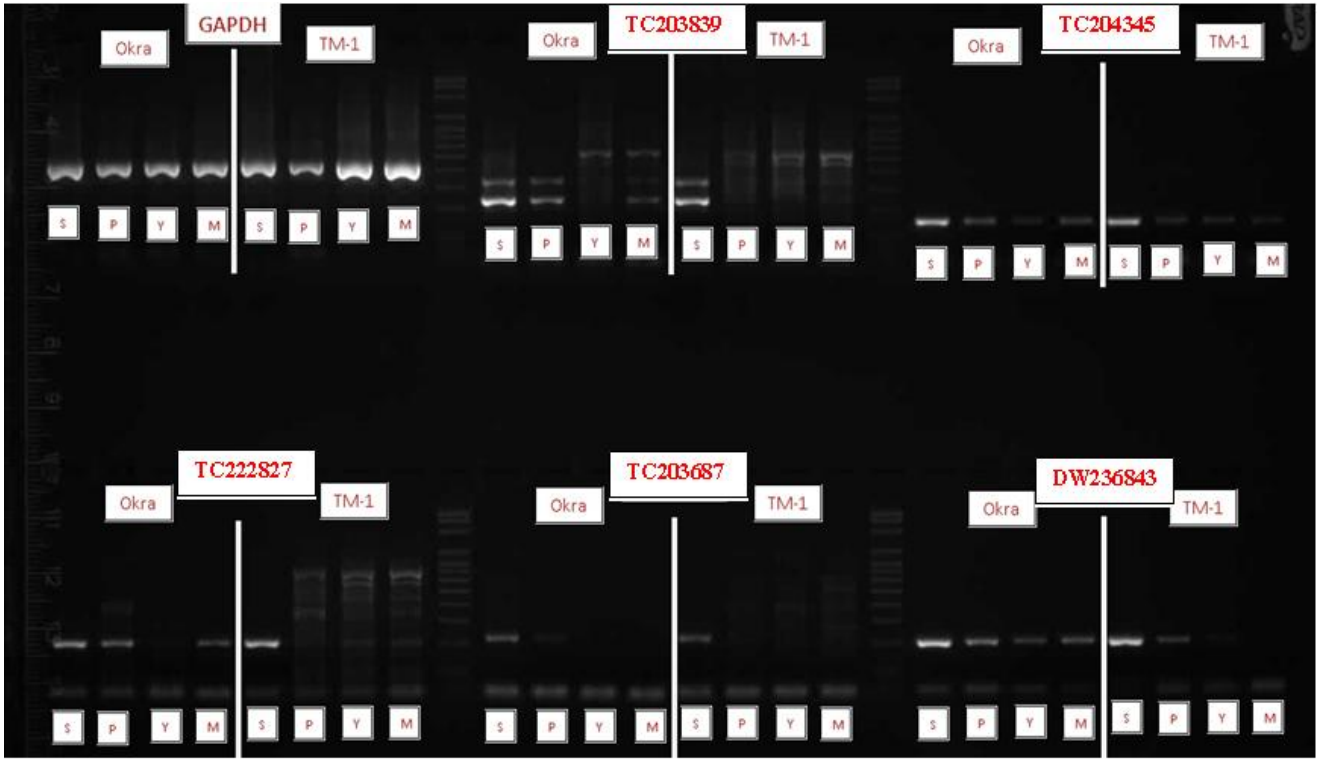


Figure 3: Semi-quantitative cDNA expression levels of *KNOX1* orthologs. Comparisons were made between an okra leaf variety (NC05AZ21) and a normal leaf line (TM-1). From each variety four tissue types were analyzed, SAM-enriched stem (S), primordia (P), young leaf (Y), and mature leaf (L). Expected product sizes are GAPDH – 451bp, TC203839 – 210bp, TC204345 – 163bp, TC222827 – 208bp, TC203687 – 227bp, DW236843 – 217bp. Reference for DNA ladder appears in Appendix A.

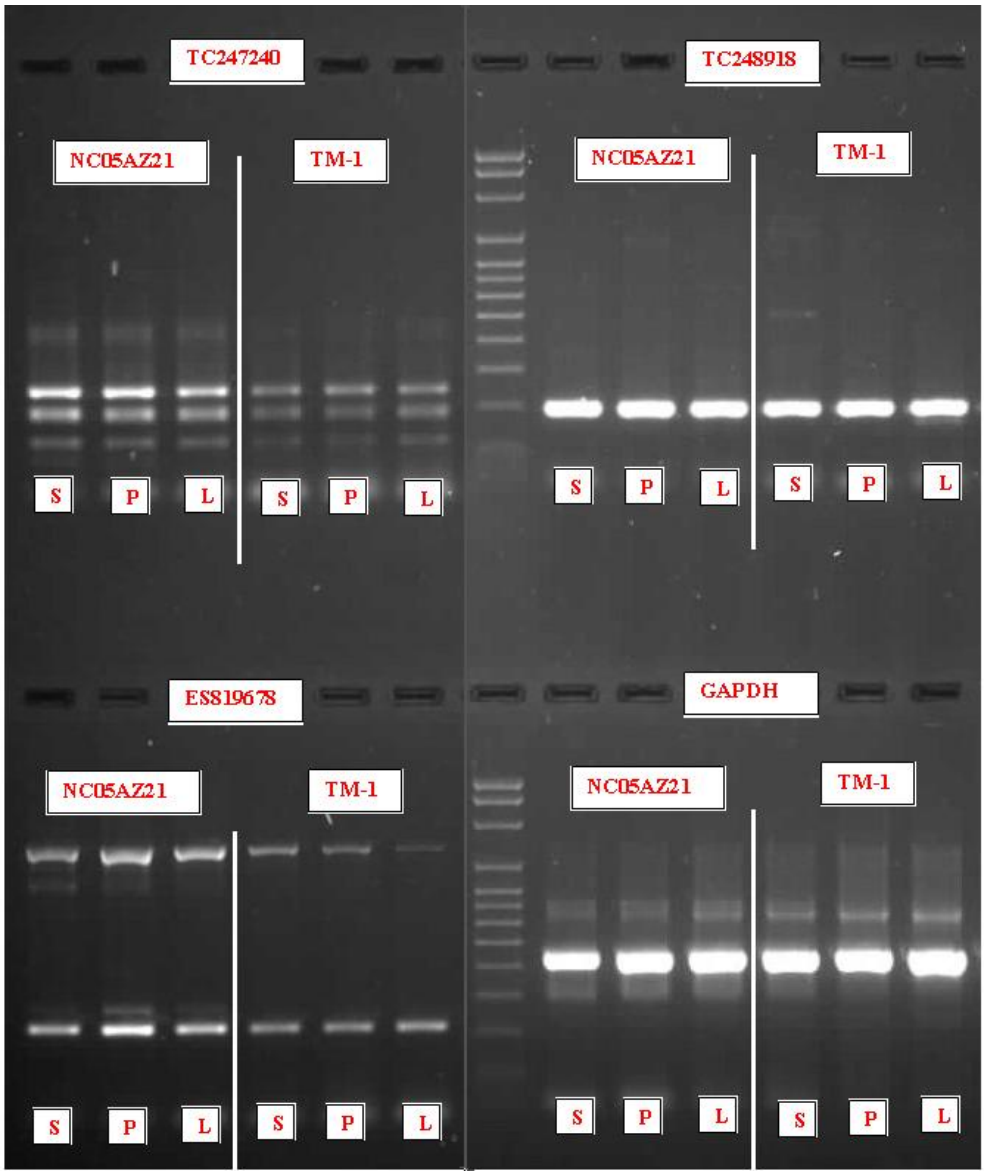


Figure 4: Semi-quantitative cDNA expression levels of *AS1/AS2* orthologs. Comparisons were made between an okra leaf variety (NC05AZ21) and a normal leaf line (TM-1). From each variety three tissue types were analyzed, SAM-enriched stem (S), primordia (P), and mature leaf (L). Expected product sizes are GAPDH – 451bp, TC247240 – 194bp, TC248918 – 201bp, ES819678 – 230bp. Reference for DNA ladder appears in Appendix A.

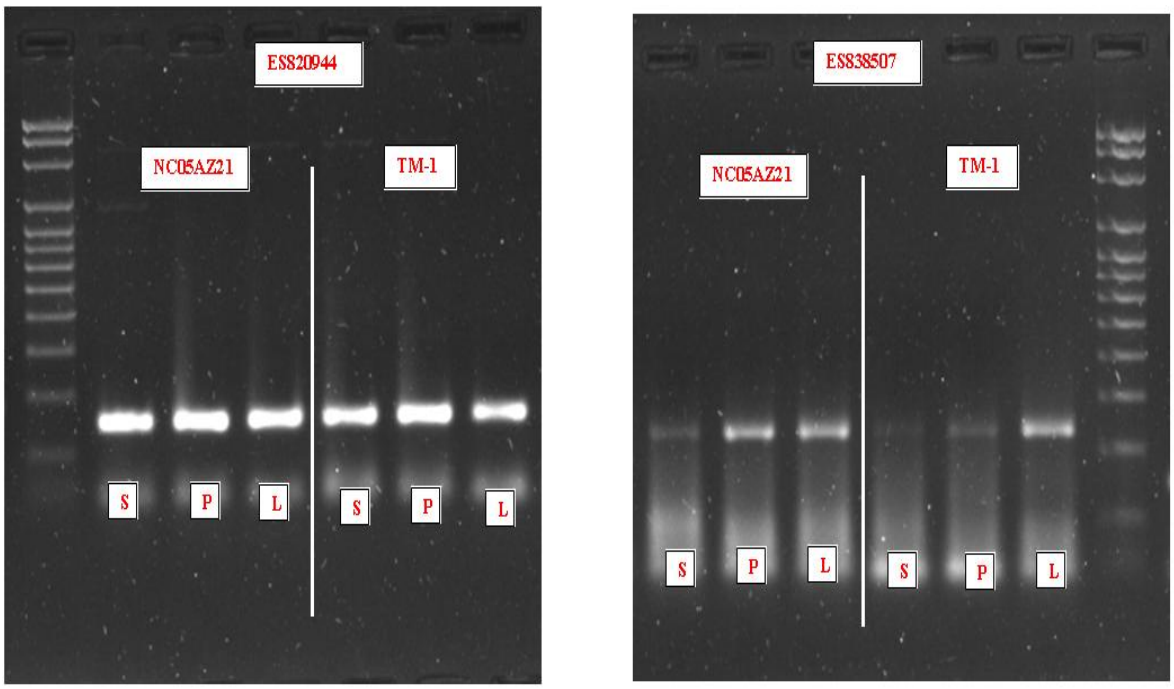


Figure 5: Semi-quantitative cDNA expression levels of *BLH* transcription factor orthologs. Comparisons were made between an okra leaf variety (NC05AZ21) and a normal leaf line (TM-1). From each variety three tissue types were analyzed, SAM-enriched stem (S), primordia (P), and mature leaf (L). GAPDH not shown but all samples were equal. Expected product sizes are ES820944 – 153bp, ES838507 – 243bp. Reference for DNA ladder appears in Appendix A.

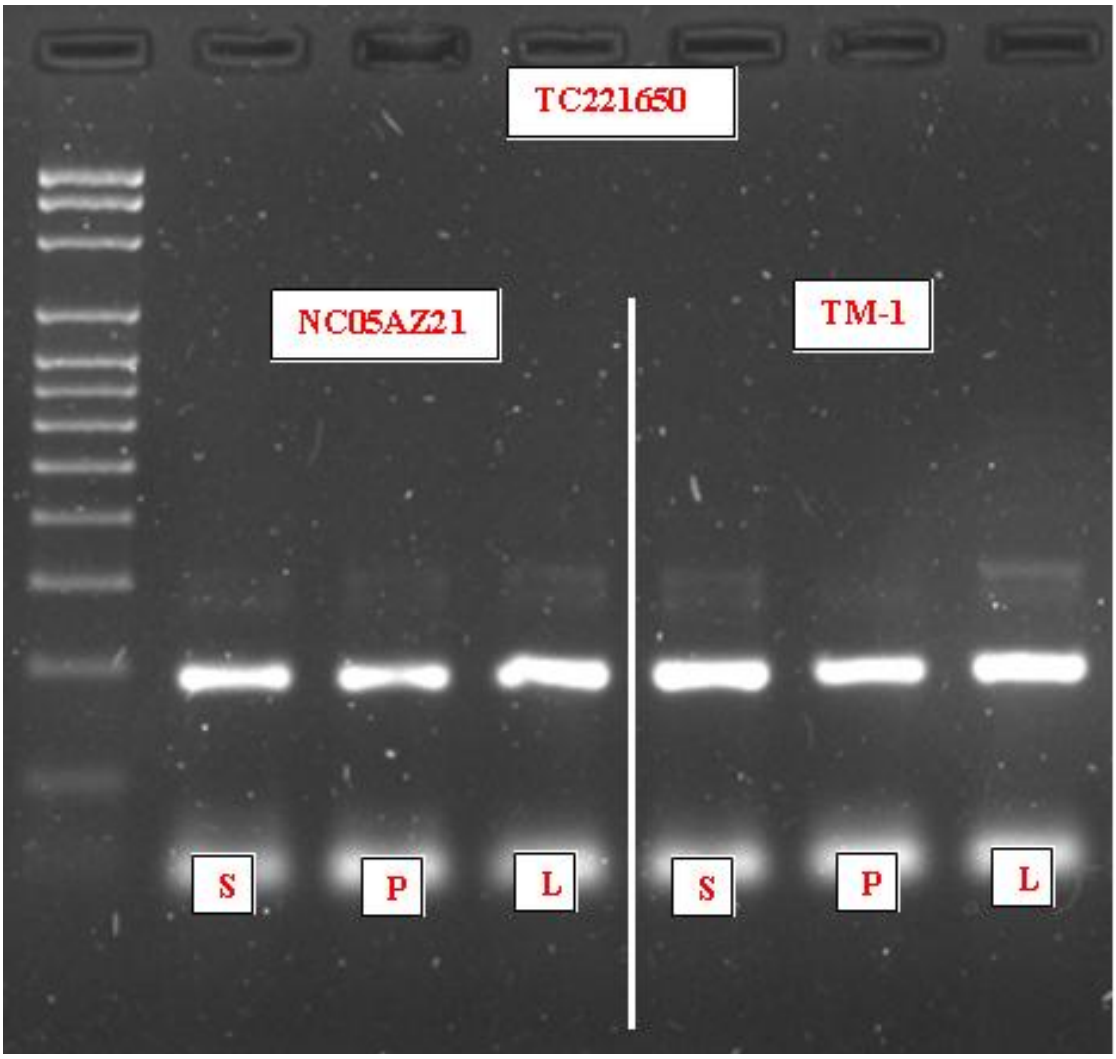


Figure 6: Semi-quantitative cDNA expression levels of *CUC* gene ortholog. Comparisons were made between an okra leaf variety (NC05AZ21) and a normal leaf line (TM-1). From each variety three tissue types were analyzed, SAM-enriched stem (S), primordia (P), and mature leaf (L). GAPDH not shown but all samples were equal. Expected product size is TC221650 – 192bp. Reference for DNA ladder appears in Appendix A.

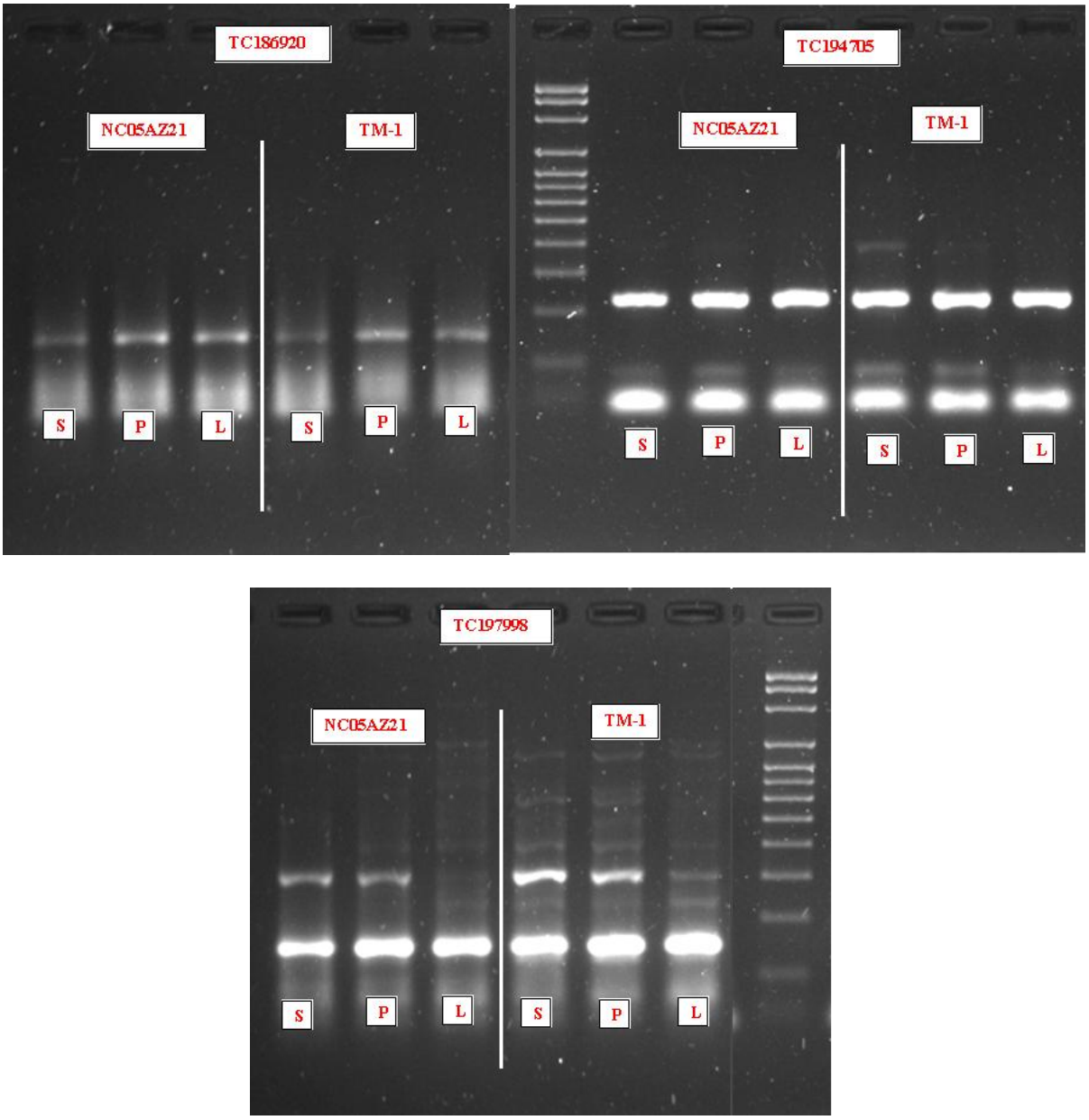


Figure 7: Semi-quantitative cDNA expression levels of *TCP* genes. Comparisons were made between an okra leaf variety (NC05AZ21) and a normal leaf line (TM-1). From each variety three tissue types were analyzed, SAM-enriched stem (S), primordia (P), and mature leaf (L). GAPDH not shown but all samples were equal. Expected product sizes are TC186920 – 160bp, TC194705 – 230bp, TC197998 – 150bp. Reference for DNA ladder appears in Appendix A.

2.5 Discussion

In theory, the protocol used provides a rough estimate of the mRNA copy number of a particular candidate gene and compares that to the expression level displayed in other tissues and cultivars (Chang et al. 1993). The approach is semi-quantitative in the sense that only major differences in expression can be detected (Chang et al. 1993). Slight variations in the copy number of a gene can be caused by a variety of reasons including: growth stage, stress, time of day, and experimental error. All of these can cause one band to appear more intense than another and lead to a false positive conclusion if the results are not viewed skeptically. Furthermore, most of the genes studied are transcription factors, so very slight and temporally short changes in their expression could cause dramatic phenotypic changes that are hard to detect with the method used.

2.5.1 Class I KNOX Genes

All five of the putative *KNOX1* orthologs tested showed the characteristic expression pattern of the gene family. Expression was strongest in the SAM-enriched stem tissue, followed by a pronounced down-regulation/exclusion from the developing primordia and leaves. The high degree of sequence similarity coupled with the characteristic expression profile indicates that the method was likely successful in identifying cotton *KNOX1* orthologs. The five *KNOX1* putative orthologs tested showed strong and equal expression in the SAM-enriched stems of both varieties. However, all five TCs tested showed expression levels that were stronger and/or persisted later in leaf development in the okra leaf than in the normal leaf (Figure 3). These results indicate that okra leaf has ectopic and/or over-expression of known leaf shape modification genes in a manner that is consistent with increased leaf complexity and lobing observed in tomato, *Arabidopsis*, tobacco, and potato (Lincoln et al. 1994, Janssen et al. 1998, Tanaka-Ueguchi et al. 1998, Rosin et al. 2003, Uchida et al. 2010). The abnormal expression of these genes could delay the normal process of cell differentiation temporarily, before being resolved. Resolution would need to occur in order for the leaf to develop at all. Persistent expression of these genes would maintain stem

cell indeterminacy indefinitely, preventing the differentiation necessary to develop the leaf and other above ground organs.

Therefore, the increased lobing and abnormal outgrowths characteristic of okra leaf cotton may be caused by the ectopic expression of cotton *KNOX1* genes. Since *STM* orthologs appear to be abnormally expressed, it may in turn cause the temporary down-regulation of negative regulators of *KNOX1* genes and therefore activate the expression of the other *KNOX1* genes. This is consistent with the role of *STM* as a master switch for the expression of other *KNOX1* genes in *Arabidopsis* proposed by Uchida et al. (2010). The point that there appears to be two putative cotton orthologs for every one *Arabidopsis KNOX1* gene is significant. Since cotton is a tetraploid, it should possess twice as many members of each gene family as *Arabidopsis* in the absence of any subsequent addition or deletion of members. This was the case for *STM* and *KNAT1* while blasts with *KNAT2* and *KNAT6* resulted in identical groups of four. This may be because the sequences of *KNAT2* and *KNAT6* are similar enough to each other that there are not major differences between their four putative cotton orthologs. The blast results show that cotton appears to have at least doubled the complement of *KNOX1* genes as *Arabidopsis*. This would be consistent with cotton's status as a tetraploid but it may have more based on the results of TC203687. This contig was outside the pair of high scoring *KNAT1* orthologs yet is very similar in sequence to other putative *KNOX1* orthologs. Its expression pattern was also typical of the gene family and the other putative orthologs indicating that it might be an extra member specific to cotton. In any case, due to the appearance of duality in cotton *KNOX1* gene number, mis-expression of the D sub-genome *STM* ortholog could be involved with the leaf shape mutants at the chromosome 15 locus. Meanwhile mis-expression of the A sub-genome *STM* copy could be responsible for the allelic series on chromosome 1. Exact copy number of these genes in cotton can be determined by using RFLP analysis.

2.5.2 AS1/AS2 Orthologs

AS1 and *AS2* expression is normally excluded from the SAM. Its appearance in the SAM-enriched stem tissue likely comes from the fact that pure SAM was not isolated for these tissue samples. Isolation of the pure SAM is technically difficult and its small size makes extraction of RNA more difficult. Therefore, the tissue meant to embody the SAM in this experiment is likely contaminated by flanking stem tissue leading to the detection of *AS1* and *AS2* expression in the SAM-enriched stem. Nevertheless, the slight over-expression of TC247240 and ES819678 in the okra line might be in response to the mis-expression of *KNOX1* genes discussed earlier. Since these genes are known negative regulators of *KNOX1* expression, over-expression of these genes may be part of the plant's response. By increasing the expression of *AS1* and *AS2* the plants may be attempting to reduce and eventually silence the expression of *KNOX1* genes in order to get the cells back on the necessary path towards cell differentiation. However, the difference in the expression is slight, and since the method used was semi-quantitative in nature, a more robust and repeatable technique is needed to confirm our results.

2.5.3 Other Candidate Genes

Neither of the two *BLH* candidate genes identified showed any kind of repeatable variation. The high amount of repeatability was likely due to the model used and based on these findings, there is no reason to believe that alterations in the expression of *BLH* transcription factors are at the root of the cotton leaf shape mutants. Additionally, the cotton putative *CUC* ortholog showed no variation of any kind between tissue types or cultivars. Therefore, it does not appear that variation in *CUC* gene expression is responsible for the okra leaf phenotype. Like the *CUC* ortholog, none of the three identified putative *TCP* orthologs in cotton showed any significant variation. Therefore, there is no evidence that differential expression of TCP transcription factor, either temporally or spatially, is responsible for the mutant leaf shape seen in okra cotton.

Although putative orthologs were not found for either *KNATM* or *PTS*, this does not necessarily mean cotton lacks a modified *KNOX1* gene that may play a similar role. It may simply be that the EST sequences from these genes have not yet been added to the database. Furthermore, a blast search with the *FLO/LFY* floral regulators shown to play a role in leaf shape in the IRLC of the bean family failed to generate any matches. This was surprising considering cotton almost assuredly possesses a *FLO/LFY* ortholog given its essential role in flowering. However since cotton lies well outside the IRLC, it is not likely to have the same function regarding leaf shape as those found in pea and other closely related legumes.

2.5.4 Conclusions from the Candidate Gene Expression Study

In retrospect, a more appropriate approach would have manipulated the gene expression in a more controlled manner with easily observable and correlated downstream phenotypes. The method used here is semi-quantitative and repeatability is poor. The method was simply not robust enough to capture a change in expression that is likely brief and occurs in a tiny cluster of cells that are difficult to identify. Additionally, difficulties arise in isolating primordia, ensuring uniformity of primordia, and isolating RNA from primordia samples.

Based on the findings of this study, it appears that the change in gene expression responsible for mutant leaf shape in cotton happens well before the primordial stage tested here. Dolan and Poethig (1991) showed that the okra mutation acts very early in leaf development, causing an increase in the early relative growth rate up until the leaf reaches 3-5mm in length. After this point, mutant and normal leaves have the same relative growth rate (Dolan and Poethig 1991). Primordial samples tested here were between 10-20mm in length, which means that samples were most likely already in phases of equal growth at the time of isolation. Collection of younger primordial samples is even more technically demanding and would require a significant investment of time and resources. As such, other methods, such as transgenic complementation, exist that are significantly more economical and efficient than the development of such a protocol. Dolan and Poethig (1991) also indicate that the

gene responsible for okra leaf shape in cotton may also be responsible for heteroblasty. More importantly, the gene behind both okra and heteroblasty is likely expressed within the primordia (Dolan and Poethig 1991). Even if the genes behind the two are different, the finding that okra leaf shape results from a change in gene expression very early and within primordia provides further evidence for our hypothesis that a re-establishment of *KNOX1* in the developing primordia is responsible for mutant leaf shape in cotton. If heteroblasty is also a result of a variation of *KNOX1* expression in the primordia, then growth stage has a much bigger effect on *KNOX1* expression levels than previously thought, which should be considered in future studies.

In summary, the results show that okra leaf mutants display an increase in the expression of a family of known leaf shape modification genes (*KNOX1*) compared to a normal leaf control (Figure 3). Okra leaf shape mutants also display an increase in lobing pattern and ectopic outgrowths that are characteristic of the mis-expression of these genes (Figure 1). Furthermore, this increase in expression occurs in the primordia, the region of cells that have been implicated as the origin of mutant leaf shape development (Dolan and Poethig 1991).

2.5.5 Future Research on Leaf Shape Genes in Cotton

In addition to the candidate gene, reverse genetics approach discussed above, a forward genetics approach to map the location of the leaf shape gene is currently underway. Both an okra line (NC05AZ21) and a sub-okra line (NC05AZ06) have been crossed to the normal leaf landrace accessions NC11-2100 and NC11-2091. F_1 hybrids are currently being grown in a summer nursery in Clayton, NC for selfing and advancement to the F_2 generation. After an additional round of selfing, phenotypic data and DNA samples will be collected from $F_{2:3}$ plants. Approximately 650 cotton SSR markers will be run on the four parents and polymorphic markers will be advanced to a bulk segregant analysis (BSA) on pooled members of the $F_{2:3}$ population sharing the same phenotype. SSRs still polymorphic after the BSA will be run on all F_2 plants individually in an effort to map the location of the leaf shape

locus. Candidate gene mapping can then be performed to see if any of the cloned genes map to the classical leaf shape locus in cotton. RFLP markers will be used to estimate the copy number of the leaf shape genes in cotton. The final step will involve cloning the leaf shape gene and identifying the causal nucleotide variation that results in the mutant leaf shapes seen in cotton.

The re-establishment of Class I *KNOTTED1-LIKE HOMEOBOX* (*KNOX1*) genes in developing leaf primordia has been implicated in increased leaf lobing and the appearance of ectopic outgrowths in otherwise simple-leafed species (Uchida et al. 2007, 2010). Both of these characteristics are hallmarks of the okra leaf shape mutation seen in cotton. Furthermore, the okra mutation is known to act very early in leaf development, before the primordia reach 3-5mm in length (Dolan and Poethig 1991). After this point, leaf shape growth and development appears to be equal between mutant and normal leaf types and this pattern is consistent with classical *KNOX1* expression (Dolan and Poethig 1991, Uchida et al. 2007, 2010). Our preliminary findings show that an okra leaf shape variety (NC05AZ21) displays elevated RNA levels of a known family of leaf shape modification genes (*KNOX1*) in the tissue where the mutation is thought to originate.

Transgenic complementation is the only way to unequivocally validate the findings of a candidate gene study. Transgenic over-expression of Class I *KNOX* genes has been shown to increase leaf complexity in numerous cultivated and model species including lettuce, potato, tobacco, tomato, dandelion, *Arabidopsis*, and the ornamental *Kohleria* (Frugis et al. 2001, Rosin et al. 2003, Tanaka-Ueguchi et al. 1998, Janssen et al. 1998, Muller et al. 2006, Lincoln et al. 1994, Barth et al. 2009). The over-expression of Class I *KNOX* genes in the normal leafed transgenic standard Coker 312 should push the leaf shape of this variety towards a more mutant shape. Placement of *Arabidopsis KNOX1* genes or their putative cotton orthologs in an appropriate plant transformation vector should be suitable for this purpose. It is likely that strong, constitutive over-expression of Class I *KNOX* genes would lead to highly irregular leaf development due to the continued re-expression of these genes

after the primordial stage. Therefore, the goal will be to develop weakly expressing lines capable of developing mutant leaf shapes without any off target effects.

If the transgenic over-expression of *KNOX1* genes successfully results in a move towards the desired leaf shape, the next step in the process would be the placement of the *KNOX1* transgenes under the control of an inducible promoter. By delaying the expression of the transgene in a controlled manner, the cotton plant would produce normal leaves up until the point of transgene induction. Following induction, plants should produce okra leaves, generating the desired transforming phenotype described in Section 1.2. This inducible promoter may either be growth stage specific or chemically induced. A potential growth stage specific promoter is the one driving the expression of the *LFY* gene in *Arabidopsis*. This promoter is induced only after the plant has initiated flowering and is restricted to primordial tissues (Amasino et al. 2010). This typically occurs around the 10^{th} node, allowing plants to produce normal leaves up until this point. An added benefit is that this system would not require the application of an inducer. An example of a suitable chemically inducible promoter is the *cis*-Jasmone promoter offered by PBL Technology. This system offers the advantage of leaving the transgene turned off if desired. If an inducible transgenic line can be successfully developed, the system will need to be field tested in order to determine its commercial applicability. If those tests are successful, the line can be tested against Coker 312 and a near-isogenic line (NIL) of Coker 312 possessing the okra leaf. These three lines would be genetically identical except for variations in leaf shape modification genes. The lines can then be tested at several locations throughout the southeastern U.S. Cotton Belt over multiple years and be tested for boll rot severity, yield, and fiber quality characteristics. Superior performance of the transformable leaf shape line would justify its advancement to the next stage of cultivar development. This process outlined here is likely a long and arduous one that is likely to take the better part of a decade. However, this provides a brief outlook at how knowledge of the genetic mechanism underlying leaf shape can be used to improve cotton production.

A second method that could be used to clarify the role of *KNOX1* genes in cotton leaf shape mutants could be Viral Induced Gene Silencing (VIGS). This system could be used to target *KNOX1* genes in the meristematic regions of an okra-leafed cultivar. If there was a deviation from the mutant leaf shape towards normal, this would provide evidence that *KNOX1* genes are at least somehow involved in the process. However, the multitude of *KNOX1* genes present in cotton could hinder this approach. Alternatively, the *AS1* and/or *AS2* orthologs could be targeted in normal leaf cultivars to test for a switch to the okra leaf phenotype. A third approach could utilize mRNA or protein tagging along with microscopy to detect and monitor the expression and movement of these molecules. However, without an established system, such an approach is likely to require a significant amount of time and resources.

REFERENCES

Amasino R. 2010. Seasonal and developmental timing of flowering. *The Plant Journal* 61: 1001-1013.

Barth S, Geier T, Eimert K, Watillon B, Sangwan RS, Gleissberg S. 2009. *KNOX* overexpression in transgenic *Kohleria* (Gesneriaceae) prolongs the activity of proximal leaf blastozones and drastically alters segment fate. *Planta* 230: 1081-1091.

Brubaker CL, Paterson AH, Wendel JF. 1999. Comparative genetic mapping of an allotetraploid cotton and its diploid progenitors. *Genome* 42: 184-203.

Champagne CEM, Goliber TE, Wojciechowski MF, Mei RW, Townsley BT, Wang K, Paz MM, Geeta R, Sinha NR. 2007. Compound leaf development and evolution in legumes. *The Plant Cell* 19: 3369-3378.

Chang PF, Narasimhan ML, Hasegawa PM, Bressan RA. 1993. Quantitative mRNA-PCR for expression analysis of low-abundance transcripts. *Plant Molecular Biology Reporter* 11: 237-248.

Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang T, Guo W, Chen X, Stelly DM, Rabinowicz PD, Town CD, Arioli T, Brubaker C, Cantrell RG, Lacape JM, Ulloa M, Chee P, Gingle AR, Haigler CH, Percy R, Saha S, Wilkins T, Wright RJ, Deynze AV, Zhu Y, Yu S, Abdurakhmonov I, Katageri I, Kumar PA, ur-Rhaman M, Zafar Y, Yu JZ, Kohel RJ, Wendel JF, Paterson AH. 2007. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiology* 145: 1303-1310.

Dolan L, Poethig RS. 1991. Genetic analysis of leaf development in cotton. *Development* 113: 39-46.

Frugis G, Giannino D, Mele G, Nicolodi C, Chiappetta A, Bitonti MB, Innocenti AM, Dewitte W, Onckelen HV, Mariotti D. 2001. Overexpression of *KNAT1* in lettuce shifts leaf determinate growth to a shoot-like indeterminate growth associated with an accumulation of isopentenyl-type cytokinins. *Plant Physiology* 126: 1370-1380.

Greenup A, Peacock WJ, Dennis ES, Trevaskis B. 2009. The molecular biology of seasonal flowering-responses in *Arabidopsis* and the cereals. *Annals of Botany* 103: 1165-1172.

Hake S, Smith HMS, Holtan H, Magnani E, Mele G, Ramirez J. 2004. The role of *KNOX* genes in plant development. *Annual Reviews of Cell and Developmental Biology* 20:125-151.

Janssen BJ, Lund L, Sinha N. 1998. Overexpression of a homeobox gene, *LeT6*, reveals indeterminate features in the tomato compound leaf. *Plant Physiology* 117: 771-786.

Jones JE. 1982. The present state of the art and science of cotton breeding for leaf-morphological types. p. 93-99. Proceedings of the Beltwide Cotton Production Research Conferences. National Cotton Council of America: Memphis, TN.

Koyama T, Mitsuda N, Seki M, Shinozaki K, Ohme-Takagi M. 2010. TCP transcription factors regulate the activities of *ASYMMETRIC LEAVES1* and miR164 as well as the auxin response, during differentiation of leaves in *Arabidopsis. The Plant Cell* 22: 3574-3588.

Kumar R, Kushalappa K, Godt D, Pidkowich MS, Pastorelli S, Hepworth SR, Haughn GW. 2007. The *Arabidopsis* BEL1-LIKE HOMEODOMAIN proteins SAW1 and SAW2 act redundantly to regulate *KNOX* expression spatially in leaf margins. *The Plant Cell* 19: 2719-2735.

Lincoln C, Long J, Yamaguchi J, Serikawa K, Hake S. 1994. A knotted1-like homeobox gene in Arabidopsis is expressed in the vegetative meristem and dramatically alters leaf morphology when overexpressed in transgenic plants. *The Plant Cell* 6: 1859-1876.

Muller KJ, He X, Fischer R, Prufer D. 2006. Constitutive *knox1* gene expression in dandelion (*Taraxacum officinale*, Web.) changes leaf morphology from simple to compound. *Planta* 224: 1023-1027.

Nikovics K, Blein T, Peaucelle A, Ishida T, Morin H, Aida M, Laufs P. 2006. The balance between the *MIR164A* and *CUC2* genes controls leaf margin serration in *Arabidopsis. The Plant Cell* 18: 2929-2945.

Peng J, Richards DE, Hartley NM, Murphy GP, Devos KM, Flintham JE, Beales J, Fish LJ, Worland AJ, Pelica F, Sudhakar D, Christou P, Snape JW, Gale MD, Harberd NP. 1999. Green revolution genes encode mutant gibberellins response modulators. *Nature* 400: 256-261.

Pflieger S, Lefebvre V, Causse M. 2001. The candidate gene approach in plant genetics: a review. *Molecular Breeding* 7: 275-291.

Reiser L, Sanchez-Baracaldo P, Hake S. 2000. Knots in the family tree: evolutionary relationships and functions of *knox* homeobox genes. *Plant Molecular Biology* 42: 151-166.

Rosin FM, Hark JK, Horner HT, Davies PJ, Hannapel DJ. 2003. Overexpression of a *Knotted*-like homeobox gene of potato alters vegetative development by decreasing gibberellin accumulation. *Plant Physiology* 132: 1106-1117.

Smith CW, Cothren JT. 1999. Cotton: Origin, history, technology, and production. John Wiley & Sons, Inc. New York, NY.

Tanaka-Ueguchi M, Itoh H, Oyama N, Koshioka M, Matsuoka M. 1998. Over-expression of a tobacco homeobox gene, *NTH15*, decreases the expression of a gibberellin biosynthetic gene encoding GA-20 oxidase. *The Plant Journal* 15: 391-400.

Uchida N, Townsley B, Chung K, Sinha N. 2007. Regulation of *SHOOT MERISTEMLESS* genes via an upstream-conserved noncoding sequence coordinates leaf development. *PNAS* 104(40): 15953-15958.

Uchida N, Kimura S, Koenig D, Sinha N. 2010. Coordination of leaf development via regulation of *KNOX1* genes. *Journal of Plant Research* 123:7-14.

Udall JA, Swanson JM, Haller K, Rapp RA, Sparks ME, Hatfield J, Yu Y, Wu Y, Dowd C, Arpat AB, Sickler BA, Wilkins TA, Guo JY, Chen XY, Scheffler J, Taliercio E, Turley R, McFadden H, Payton P, Klueva N, Allen R, Zhang D, Haigler C, Wilkerson C, Suo J, Schulze ST, Pierce ML, Essenberg M, Kim H, Llewellyn DJ, Dennis ES, Kudrna D, Wing R, Paterson AH, Soderlund C, Wendel JF. 2006. A global assembly of cotton ESTs *Genome Research* 16: 441-450.

Wang K, Song X, Han Z, Guo W, Yu JZ, Sun J, Pan J, Kohel RJ, Zhang T. 2006. Complete assignment of the chromosomes of *Gossypium hirsutum* L. by translocation and fluorescence in situ hybridization mapping. *Theoretical Applied Genetics* 113: 73-80.

Wang H, Chen J, Wen J, Tadege M, Li G, Liu Y, Mysore KS, Ratet P, Chen R. 2008. Control of compound leaf development by *FLORICAULA/LEAFY* ortholog *SINGLE LEAFLET1* in *Medicago truncatula*. *Plant Physiology* 146: 1759-1772.

Wendel JF, Brubaker CL, Seelanan T. 2009. The origin and evolution of *Gossypium*. In: Physiology of Cotton. Editors: Stewart JM, Oosterhuis DM, Heitholt JJ, Mauney JR. Springer New York, NY.

CHAPTER 3: Development of Interspecific Hybrids in *Gossypium*

3.1 Review of Literature and Introduction

Genetic uniformity in a cultivated crop species across a wide geographic area can become catastrophic if this widely adopted cultivar suddenly becomes vulnerable to a particular biotic or abiotic stress (Wallace et al. 2009). Genetic diversity, both in cultivation and in breeding populations, can help mitigate these threats by providing additional alleles conferring resistance to these stresses (Wallace et al. 2009). The *Gossypium* genus presently consists of ~50 species whose natural habitats are spread throughout much of the tropical and sub-tropical regions of the globe (Wallace et al. 2009). Due to this wide distribution and the accompanying range of environmental conditions *Gossypium* species have evolved under, there is a considerable amount of genetic diversity to exploit for the enhancement of cotton cultivars (Wallace et al. 2009).

However, despite this wealth of genetic variability, only a few germplasm lines have been involved in the majority of the recently released commercial varieties (Wallace et al. 2009). It is estimated that only 1% of the available genetic diversity is present in the current upland cultivars (Wallace et al. 2009). This bottleneck is even more severe in transgenics, which occupy an increasingly high percentage of the acreage (Wallace et al. 2009). Nearly all GMO lines are derived from the Coker 312 background due to its ease of regeneration, inserting a common linkage block into nearly all current cotton cultivars (Wallace et al. 2009).

Wild species germplasm is an excellent source of genes for biotic and abiotic stress tolerance as well as unique biochemical properties (Wallace et al. 2009). Genes from wild crop relatives have been successfully incorporated into released cultivars in many species including: rice, wheat, barley, maize, sorghum, millet, cassava, potato, chickpea, banana, groundnut, tomato, sunflower, and lettuce (Hajjar and Hodgkin 2007). For these 14 crops, over 60 different wild relatives have been used to transfer over 100 different traits (Hajjar and Hodgkin 2007). Of these traits, the vast majority (80%) have been for pest and disease resistance, although traits for abiotic stress tolerance, yield, cytoplasmic male sterility,

fertility restoration, and quality have all been incorporated as well (Hajjar and Hodgkin 2007).

Preservation and collection of the wild germplasm is needed before adequate characterization can be carried out to identify desirable traits (Wallace et al. 2009). Other issues limiting the use of exotic germplasm in traditional breeding include long breeding cycles, infrequent flowering, linkage drag, incompatibility, and the sterility often associated with interspecific (especially intergenomic and across ploidy level) hybrids (Hajjar ad Hodgkin 2007). Another significant hurdle to increasing genetic diversity in cotton is photoperiodism, which requires the use of resources such as greenhouses and growth chambers in order for breeders in temperate environments to carry out crosses between these species.

Nevertheless, there have been examples of wild trait introgression in cotton. J.O. Beasley developed a triple hybrid between *G. thurberi*, *G. arboreum*, and *G. hirsutum* that was backcrossed for several generations to *G. hirsutum* which was instrumental in breaking the negative correlation between yield and fiber quality (Percival et al. 1999, Schwartz and Smith 2008). Traits for resistance to the reniform nematode (*Rotylenchulus reniformis*) were transferred into *G. hirsutum* from *G. arboreum* and *G. longicalyx* using *G. aridum* and *G. armourianum* or *G. herbaceum* as bridging species, respectively, in a tri-species hybrid (Robinson et al. 2007, Sacks and Robinson 2009). In these instances, a tetraploid *G. hirsutum* was crossed to the bridging diploid to generate a sterile triploid (Robinson et al. 2007, Sacks and Robinson 2009). Triploids were subsequently doubled to hexaploids and crossed to the donor diploid (Robinson et al. 2007, Sacks and Robinson 2009). Resulting hybrids were backcrossed to cultivated *G. hirsutum* in order to regain the cultivated phenotype (Robinson et al. 2007, Sacks and Robinson 2009). However, there appears to be no publicly available synthetic polyploid, alien addition, monosomic, nullisomic, or chromosome substitution lines available at any of the cotton germplasm collections except for 32 “synthetic introgressed derivatives” at the Indian Central Institute for Cotton Research (CICR) (Stewart 2010, Campbell et al. 2010).

One of the most desirable alterations to the biochemistry of cotton is the elimination of gossypol production from the seeds. A significant amount of cottonseed is produced as a by-product of fiber production. This seed is very high in protein making it an ideal food for livestock and potentially protein-starved human populations throughout the world (Cai et al. 2010). However, the presence of gossypol, a toxic terpenoid aldehyde produced in the roots and stored in above-ground pigment glands, limits the feeding of cottonseed to animals with a ruminant digestive system. Cai et al. (2010) stated that gossypol-free cottonseed has the potential to feed 500 million people a year. Additionally, cottonseed is the sixth largest source of vegetable oil in the world, but the presence of gossypol stains this oil and other protein products with undesirable colors (Eamens et al. 2008).

There have been four different approaches used to develop an acceptable variety possessing gossypol-free seeds. First, lines have been developed in which gossypol production is knocked out throughout the plant using mutations to critical genes. This is easily done, but since gossypol is an important component of plant defense, these plants are extremely susceptible to damage from insects and diseases and are therefore not a viable option (Cai et al. 2010). A second approach was to cross a glandless variety with a cultivar known to produce low levels of gossypol in the hopes of developing a desirable phenotype. One line that came out of these efforts in China, Xiangmian 18, possessed glandless seeds with very low gland number and gossypol content (Cai et al. 2010). This line has had limited commercial success to date, presumably due to an insufficient level of gossypol production (Cai et al. 2010). A third approach is to use genetic engineering to develop this trait. Sunilkumar et al. (2006) has successfully used RNAi to silence gossypol genes only in the seed. Other GMO approaches include the activation of gland forming genes only in above-ground tissues in plants with a glandless background and the production of a less toxic form of gossypol that is more appropriate for humans and non-ruminant livestock (Cai et al. 2010). However, all of these GE approaches suffer from the traditional obstacles to GMO development including high developmental and regulatory costs and rejection in certain foreign markets.

The fourth and final source of gossypol-free seeds lies in wild species. Certain species contain traits that have been referred to as a delay in pigment gland morphogenesis or the delayed development of gossypol glands (Cai et al. 2010). This trait has been observed in five wild species native to Australia: *Gossypium sturtianum*, *G. nandewarense* Dereda, *G. australe* Mueller, *G. nelsonii* Fryxell, and *G. bickii* (Cai et al. 2010). These species show little to no gossypol content or pigment glands in dormant seeds and cotyledons one day prior to germination (Cai et al. 2010). However, within two days of germination gossypol has accumulated to low levels in the cotyledons and production will increase rapidly over the course of the next nine days (Cai et al. 2010). Some other species, including the African species *G. somalense,* possess gossypol glands in their seeds but gossypol fails to accumulate to a detectable level (Cai et al. 2010). Efforts to breed the delayed gland development trait into elite cotton cultivars or breeding lines have largely been fruitless. The Australian C- and G-genome species are located in the tertiary gene pool to the cultivated cotton tetraploids (AADD) and direct fertilization has proved elusive. Utilizing a diploid bridging species, such as the D-genome *G. raimondii,* has allowed the successful development of a tri-species hybrid (Diouf et al. 2010). However, these hybrids still suffer from serious problems relating to abnormal chromosome pairing and linkage drag (Diouf et al. 2010). In addition to the glandless-seed, glanded plant phenotype, *G. sturtianum* may be a good source of alleles for other desirable traits including cold and drought tolerance, fiber quality, and resistance to biotic stresses (Ahoton et al. 2003)

The development of new polyploid synthetics, alien addition, and/or chromosome substitution lines would greatly expand the diversity of the cotton germplasm. Efforts were undertaken to cross *G. hirsutum* with numerous wild, diploid relatives for this purpose. After the establishment of resulting sterile triploids, efforts were made to overcome their infertility in the next step towards the successful integration of desirable alleles from these wild relatives.

3.2 Materials and Methods

In an effort to increase the genetic diversity of cotton, crosses were carried out between the *Gossypium hirsutum* inbred line Texas Marker-1 (TM-1, 2n=4x=52, AADD) and numerous wild relatives. Crosses were also made among the wild relatives. Wild species included in this crossing regimen were *G. anomalum* (2n=2x=26, BB), *G. arboreum* (2n=2x=26, AA), *G.darwinii* (2n=4x52, AADD), *G. gossypioides* (2n=2x=26, DD), *G. herbaceum* (2n=2x=26, AA), *G. longicalyx* (2n=2x=26, FF), *G. raimondii* (2n=2x=26, DD), *G. sturtianum* (2n=2x=26, CC), and *G. thurberi* (2n=2x=26, DD). Additionally, *G. somalense* (2n=26=26, EE) and *G. capitis-viridis* (2n=2x=26, BB) were originally included but failed to germinate. Seeds were planted in two adjacent growth chambers in the N.C. State University Phytotron in the fall of 2009. Growth chambers were used in order to induce flowering in the photoperiod sensitive wild species. Growth chambers were set to eleven hour days with 28^{o}C days and 22^{o}C nights or thirteen hour days with 30^{o}C days and 24^{o}C nights. Crosses were made throughout the fall of 2009 and into the early spring of 2010. A small piece of KimWipe was soaked briefly in 50ppm giberellic acid and applied to the ovary of the female parent the day the cross was made in an effort to promote fertilization. On any given day, priority was given to cross combinations that had not previously been attempted. Due to the infrequent availability of some flowers, not all cross combinations could be attempted.

3.2.1 Backcrossing

In the first attempts to generate progeny from the sterile triploids, both the TM-1 x *G. sturtianum* F_1 triploid and the TM-1 x *G. raimondii* F_1 triploid were backcrossed repeatedly to TM-1. Crosses were conducted in a reciprocal manner so that each line had a chance as both the male and female parent. Attempts to facilitate fertilization were again made by applying a KimWipe soaked in 50ppm GA_3 to the ovary of the female parent on the day of the pollination.

3.2.2 Attempts to Double the Chromosome Number of Sterile Triploids

After no success with backcrossing, efforts were made to double the chromosome number of the mature *G. hirsutum* TM-1 x *G. sturtianum* F_1 triploid progeny. The protocol outlined by Brubaker et al. (1999) was utilized. This report stated that only 15 out of 234 sterile triploids were doubled by this procedure, 10 of which were ADC triploids resulting from a cross of *G. hirsutum* by *G. sturtianum*. To carry out this approach, mature plants approximately one meter tall were trimmed of all side branches and leaves leaving only the main stem. The top of the main stem was then trimmed to remove the shoot apical meristem and the plants were held without water for one day. The following day, the stem was re-trimmed and a standard plastic transfer pipette with the top of the bulb removed was attached to the main stem. Either 0.5% or 1% colchicine solution was applied to the pipette for 24 hours. Water was withheld again on this day in order to facilitate the movement of the colchicine solution through the stem. The next day, the transfer pipette and any excess colchicine was removed and the plants were returned to a normal watering regimen.

In another attempt to induce self-pollination or doubling of the interspecific hybrid, injection of 50ppm GA_3, 0.5-1.0% colchicine solution, or a 1:1 mixture of the two was injected into hybrid buds. In addition to being allowed to self-pollinate, flowers that had undergone these treatment were also backcrossed to TM-1 in a reciprocal manner. In order to promote fertilization, small pieces of KimWipe were soaked in 50ppm GA_3 and applied to the ovary the day of the pollination. A third approach involved the application of a small piece of KimWipe soaked in 1.0% colchicine solution to buds ranging in size from barely visible to a few days before anthesis. This approach was also applied to a plant derived from an interspecific cross between TM-1 and the presumed diploid donor of the cotton D genome, *Gossypium raimondii*.

3.2.3 Cotton Embryo Culture

After continued failure to obtain progeny from these triploids, the following embryo culture technique was employed. One day prior to anthesis, buds were injected with a 50:50

mixture of 1.0% colchicine and 50ppm gibberellic acid (GA_3). After 10-15 days, the developing boll was removed from the plant and transferred from the greenhouse to the laboratory. Bolls were dipped in 70% ethanol solution and flame sterilized. Then bolls were soaked in autoclaved double distilled water for 20 minutes in order to soften them. After soaking, a sterilized scalpel was used to remove 10-30 ovaries from each boll. The ovules were then floated on 50mL of a liquid medium in a 125mL Erlenmeyer flask. Liquid medium was prepared by adding 4.4g of MS basal media powder to 900 mL of Nanopure ddH_2O, followed by 40g of sucrose (final concentration 115mM) and 1.2g of ammonium nitrate (final concentration 15mM). Final volume was brought to one liter using Nanopure ddH_2O and liquid was partitioned in 50mL aliquots into 125mL Erlenmeyer flasks. Flasks were then capped with aluminum foil and autoclaved at 121^oC and 15psi for 15 minutes. Once flasks were cool to the touch, 174µL of 50ppm GA_3 (final concentration 0.5µM), 5µL of 50mM 3-indolacetic acid (IAA) (final concentration 5µM), and 2.5µL of 1mM kinetin (final concentration 0.05µM) were added to each flask. Flasks were then incubated at 30^oC for 60 days in the dark.

3.3 Results and Discussion

The *Gossypium sturtianum* accession NC0213 (PI 464863) was successfully crossed to the *Gossypium hirsutum* inbred line Texas Marker-1 (TM-1) yielding six vigorous F_1 plants that are presumed ADC triploids. Hybrid nature of the plants was confirmed by their novel flower color and morphology, intermediate leaf shape, and complete sterility (Figure 8). Fertilization was enhanced by the application of 50ppm gibberellic acid (GA_3) to the ovary of the TM-1 female parent. The hybrid plants flowered aggressively, with each producing up to eight flowers in a single day throughout the year. From all of these flowers not a single boll was produced without manipulation.

Attempts to backcross using TM-1 as the female parent along with GA_3 ovary treatment resulted in the development of smaller and abnormally shaped bolls (Figure 10). Approximately 50% of these bolls were shed before opening. Upon examination, the shed

bolls showed no signs of fertilization and seed development. From the opened bolls, lint was coarse and yield was minimal (Figure 10). In most, there was no evidence of fertilization and no hybrid plants were obtained from any of the crosses. Backcrosses with the TM-1 x *G. sturtianum* as the female parent failed to produce any seed although the same phenomenon of boll development and opening was seen with the application of exogenous GA_3 to the ovary on the day of pollination.

Efforts to double the chromosome number of these hybrids via the application of colchicine to floral buds or exposed stem all proved unsuccessful. Attempts to backcross pre-treated buds to TM-1 failed to produce any seed, even with the GA_3 ovary treatment. Again, crosses treated with GA_3 developed bolls, some of which opened yielding small amounts of lint, but no seed. All attempts at embryo culture were unsuccessful.

The *Gossypium raimondii* accession NC0171 (PI530898) was also successfully crossed with TM-1 resulting in four healthy F_1 plants. Complete sterility and unique leaf shape confirmed the hybrid nature of these lines (Figure 9). Flowers appear very similar to those of *G. raimondii* but the overall plant architecture and flowering rates match closely with those of TM-1 (Figure 9). *G. raimondii* resembles a tree, growing to the tops of the greenhouse and flowering only a handful of times each year. Efforts to double the chromosome number of these lines and to backcross them to TM-1 as both a male and a female have proven unsuccessful.

In addition to the sterile triploids discussed above, successful crosses were also made between TM-1 and two other allotetraploid members of *Gossypium, G. darwinii* (NC0061, AD5 3.02 SD) and *G. mustelinum* (NC0135, AD4 17.01 SD). The TM-1 x *G. darwinii* cross is currently in the F_2 stage, while the TM-1 x *G. mustelinum* cross is in the $F_{2:3}$ stage after F_2's were selfed in a winter nursery in Mexico. Both of these populations are currently planted in Clayton, NC.

A successful hybridization was also made between the diploid species *G. herbaceum* (NC2086, PI630017, AA) and *G. anomalum* (NC0003, PI530883, BB). Leaf morphology of the hybrid is similar to both of the parents, which very closely resemble each other.

However, flower morphology appears intermediate between the two parental types (Figure 11 A-C) and bract shape is variable as well. Furthermore, bolls are misshapen and seed set and fiber development is poor, neither of which are the case with the parental species (Figure 11 D - F)

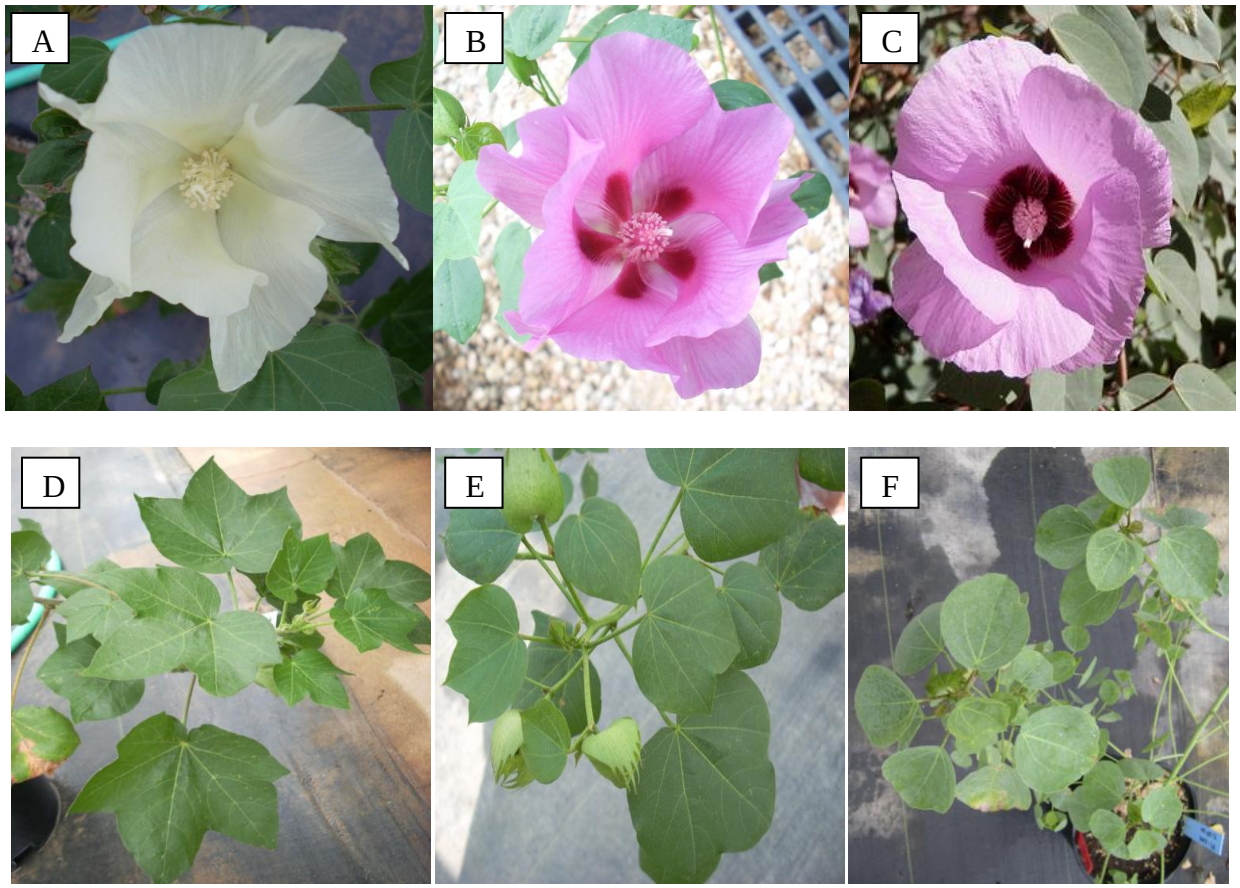
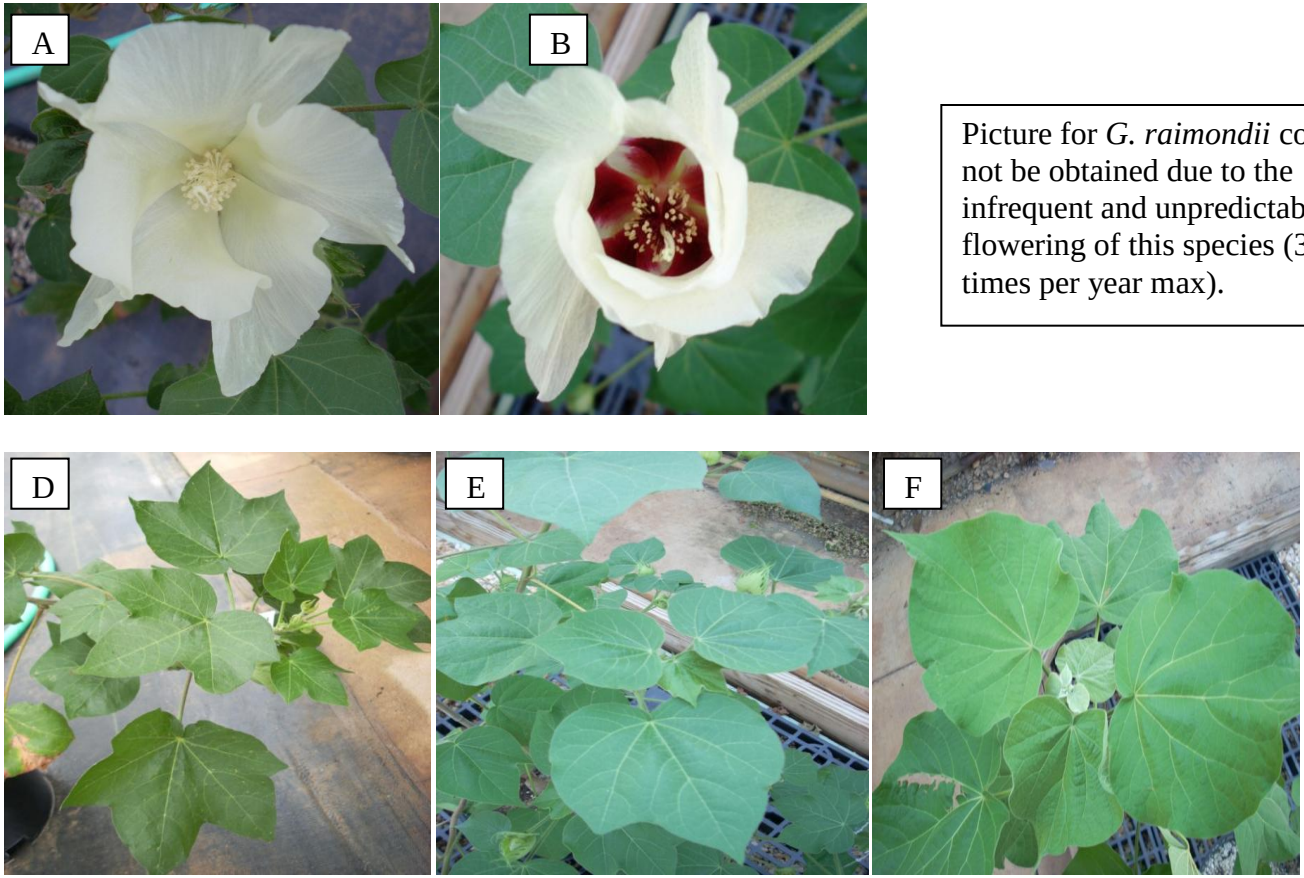


Figure 8: TM-1 x NC0213 (*Gossypium sturtianum*) F_1 Hybrid and Parents. A-C. TM-1, TM-1 x NC0213 F_1 Hybrid, and *G. sturtianum* flower*, respectively. D-F. TM-1, TM-1 x NC0213 F_1 Hybrid, and *G. sturtianum* leaf shape, respectively. *Photo *G. sturtianum flower taken from*http://www.flickr.com/photos/40325561@N04/4092363972/.



Picture for *G. raimondii* could not be obtained due to the infrequent and unpredictable flowering of this species (3-4 times per year max).

Figure 9: TM-1 x NC0171 (*Gossypium raimondii*) F_1 Hybrid and Parents. A-C. TM-1, TM-1 x NC0171 F_1 Hybrid, and *G. raimondii* flower, respectively. D-F. TM-1, 1 x NC0171 F_1 Hybrid, and *G. raimondii* leaf shape, respectively.

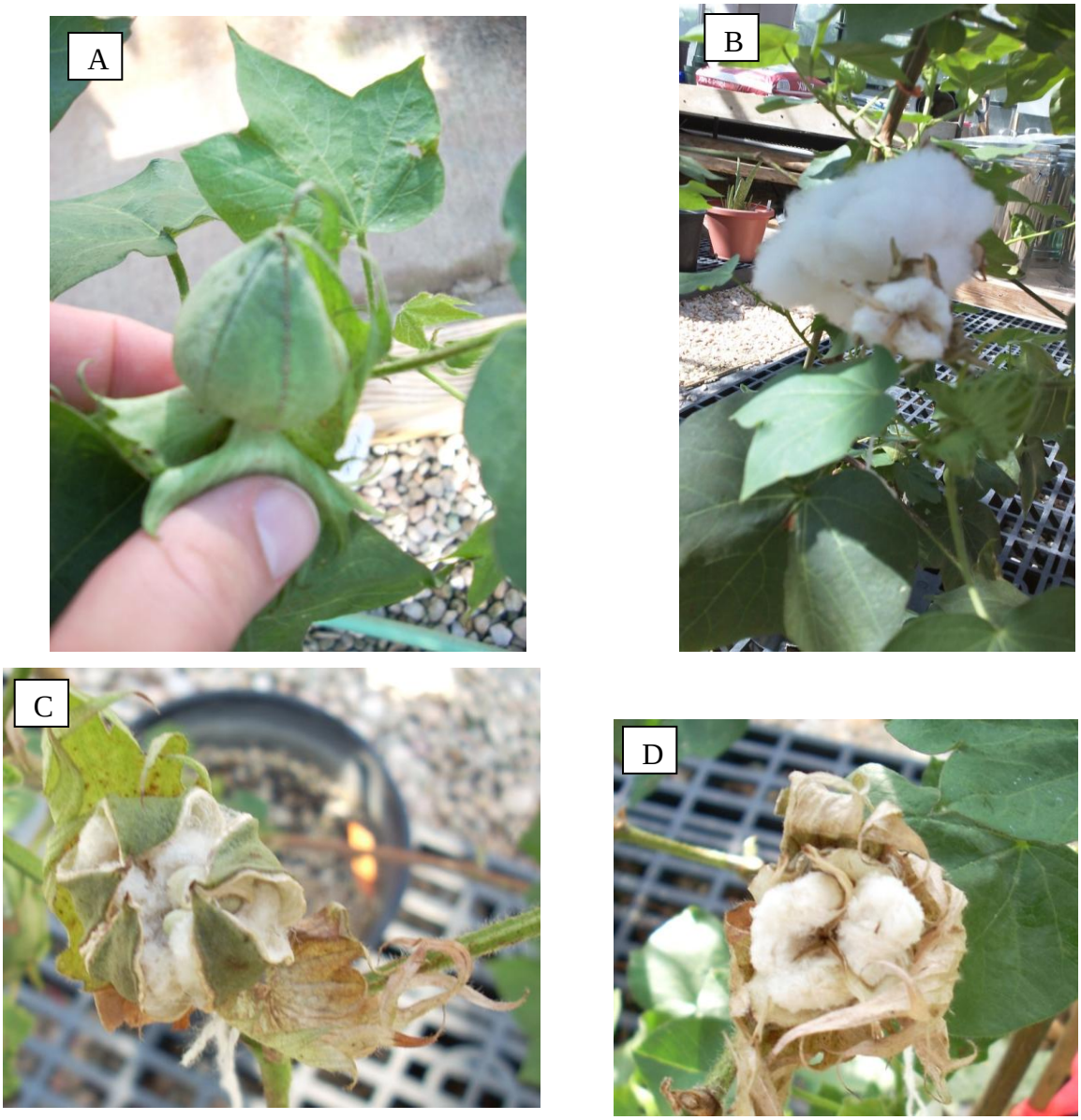


Figure 10: Results of Backcrossing TM-1 x NC0213 to TM-1: A. Abnormal and misshapen boll on TM-1 plant following fertilization with TM-1 x *G. sturtianum* and GA_3 treatment. B. In foreground is an opened TM-1 boll after fertilization with TM-1 x *G. sturtianum* and GA_3 treatment, in background is a normal, opened TM-1 boll that was allowed to self-fertilize. C+D. Opened, abnormal TM-1 bolls crossed with TM-1 x *G. sturtianum* and GA_3 treatment. Bolls of this type produced lint, but no seeds.

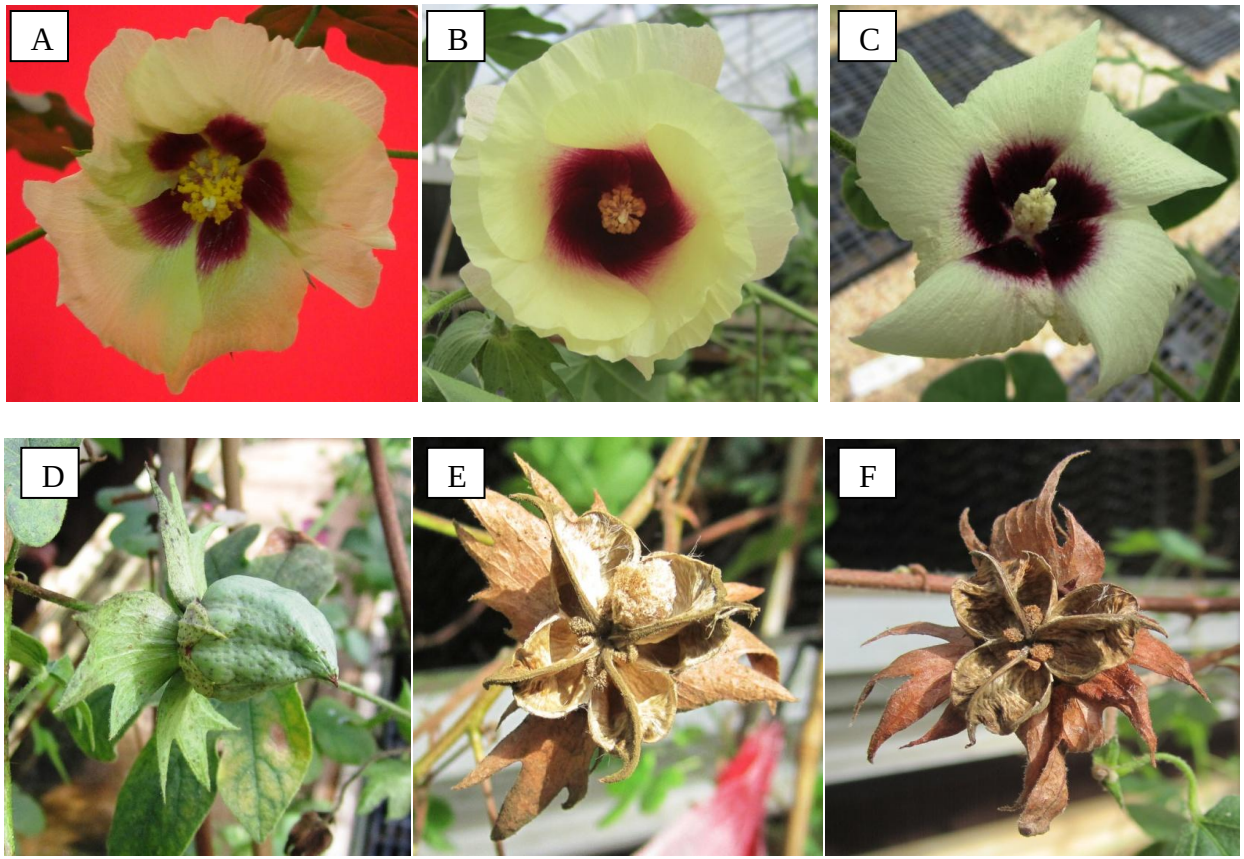


Figure 11: NC2086 (*Gossypium herbaceum*) x NC0005 (*G. anomalum*) F_1 Hybrid and Parents. A-C. NC2086, NC2086 x NC0005 F_1 Hybrid, and NC0005 flower, respectively. D. Misshapen boll typical of F_1 Hybrid. E-F. Opened bolls of F_1 hybrid showing poor seed set and lint production.

REFERENCES

Ahoton L, Lacape JM, Baudoin JP, Mergeai G. 2003. Introduction of Australian diploid cotton genetic variation into Upland cotton. *Crop Science* 43: 1999-2005.

Brubaker CL, Brown AHD, Stewart JM, Kilby MJ, Grace JP. 1999. Production of fertile hybrid germplasm with diploid Australian *Gossypium* species for cotton improvement. *Euphytica* 108: 199-213.

Cai Y, Xie Y, Liu J. 2010. Glandless seed and glanded plant research in cotton. A review. *Agronomy for Sustainable Development* 30: 181-190.

Campbell BT, Saha S, Percy R, Frelichowski J, Jenkins JN, Park W, Mayee CD, Gotmare V, Dessauw D, Giband M, Du X, Jia Y, Constable G, Dillon S, Abdurakhmonov IY, Abdukarimov A, Rizaeva SM, Adullaev A, Barroso PAV, Padua JG, Hoffman LV, Podolnaya L. 2010. Status of the global cotton germplasm resources. *Crop Science* 50: 1161-1179.

Diouf FBH, Benbouza H, Konan NO, Sarr D, Ndir KN, Baudoin JP, Mergeai G. 2010. Effect of gammaray in the progeny of trispecific hybrid [(*Gossypium hirsutum x G. raimondii*)^2 x *G. sturtianum*]. *Notulae Botanicae Horti Agrobotanici Cluj-Napoca* 38: 78-83.

Eamens A, Wang MB, Smith NA, Waterhouse PM. 2008 RNA silencing in plants: Yesterday, today, and tomorrow. *Plant Physiology* 147: 456-468.

Hajjar R, Hodgkin T. 2007. The use of wild relatives in crop improvement: A survey of developments over the last 20 years. *Euphytica* 156: 1-13.

Percival AE, Wendel JF, Stewart JM. 1999. Taxonomy and germplasm resources. In: Smith CW, Cothren JT, Editors: *Cotton origin, history, technology, and production*. John Wiley & Sons, NY, NY.

Robinson AF, Bell AA, Dighe ND, Menz MA, Nichols RL, Stelly DM. 2007. Introgression of resistance to nematode *Rotylenchulus reniformis* into Upland cotton (*Gossypium hirsutum*) from *Gossypium longicalyx*. *Crop Science* 47: 1865-1877.

Sacks EJ, Robinson AF. 2009. Introgression of resistance to reniform nematode (*Rotylenchulus reniformis*) into upland cotton (*Gossypium hirsutum*) from *Gossypium arboreum* and a *G. hirsutum* / *Gossypium aridum* bridging line. *Field Crops Research* 112: 1-6.

Schwartz BM, Smith CW. 2008. Genetic gain in fiber properties of upland cotton under varying plant densities. *Crop Science* 48: 1321-1327.

Stewart JM. 2010. Chapter 2 – Germplasm resources for physiological research and development. In: Stewart JM, Oosterhuis DM, Heitholt JJ, Mauney JR, editors. *Physiology of Cotton.* New York Springer Science + Business Media p. 19-23.

Sunilkumar G, Campbell LM, Puckhaber L, Stipanovic RD, Rathore KS. 2006. Engineering cottonseed for use in human nutrition by tissue-specific reduction of toxic gossypol. *Proceedings of the National Academy of Sciences* 103: 18054-18059.

Wallace TP, Bowman D, Campbell BT, Chee P, Gutierrez OA, Kohel RJ, McCarty J, Myers G, Percy R, Robinson F, Smith W, Stelly DM, Stewart JM, Thaxton P, Ulloa M, Weaver DB. 2009. Status of the USA cotton germplasm collection and crop vulnerability. *Genetic Resources and Crop Evolution* 56: 507-532.

CHAPTER 4: Development of an Improved Method of Mitotic Metaphase Chromosome Preparation in Cotton (*Gossypium hirsutum)*

4.1 Review of Literature

Fluorescent *in situ* hybridization (FISH) is a cytogenetic technique that utilizes site-specific nucleic acid hybridization between denatured chromosomal DNA fixed to a slide and selected, complementary probes (Bauman et al. 1980, Birchler and Danilova 2009). It holds great potential for the physical localization and mapping of unique, single-copy gene sequences, transgene insertion sites, and introgressed segments from wild and exotic sources (Figueroa et al. 2009, Peng et al. 2012). Additionally it can be used to identify and karyotype chromosomes, estimate the gap size between bacterial artificial chromosomes (BACs), and study chromosome structure and orientation (Peng et al. 2012). Even with the glut of sequencing and recombination data being generated today, FISH remains the only proven tool to physically map genes, linkage groups, and large segments of DNA to their respective chromosomes (Figueroa et al. 2009). These physical maps can reveal a clear picture of the order, spacing, and distribution of genes along the chromosomes and throughout the genome. (Figueroa et al. 2009). FISH probes, when integrated into genetic maps can reveal the physical distribution of gene spacing and recombination (Jiang and Gill 2006). FISH can also be used to detect and estimate alien gene introgression in wide hybridization programs (Jiang et al. 1993, Gill and Friebe 1998).

4.1.1 Cell Type used to Prepare Chromosomes

The first major decision to be made before embarking on any FISH experiment is which type of chromosome to use. The type of chromosome is influenced by the particular growth stage a cell happens to be in when the chromosomes are isolated. Mitotic prophase, prometaphase, and metaphase, meiotic prophase, and interphase chromosomes have all been used (Figueroa et al. 2009). The type of chromosome chosen depends largely on the objective of the experiment, but can also be influenced by the species under investigation.

Meiotic prophase chromosomes are usually harvested from the pachytene stage of prophase I; the point at which crossing over occurs between non-sister chromatids of homologous chromosomes (Kato et al. 2004). Pachytene chromosomes excel in the identification of major chromosomal landmarks, high-resolution mapping of sequences of interest, and in chromosome pairing studies (Figueroa et al. 2009). The major advantage of these chromosomes is that they are less condensed and can be stretched out to greater lengths than mitotic chromosomes (Figueroa et al. 2009). This can enhance cytogenetic work by improving the sensitivity and resolution of bound probes (Figueroa et al. 2009, Wang et al. 2009). In cytogenetics, the term resolution refers to the ability to differentiate between adjacent, bound probes (Wang et al. 2009). It is a measure of how far apart two bound probes need to be in order to be reliably perceived as different. Sensitivity, on the other hand is the smallest DNA sequence (i.e. the shortest length of the probe) that can be detected (Wang et al. 2009).

Pachytene chromosomes are isolated from the pollen mother cells (PMCs) located in the anthers (Figueroa et al. 2009). The major drawback to this approach is that PMCs can only be harvested after the plant has reached sexual maturity and begun to flower (Figueroa et al. 2009). This drastically increases the amount of time and resources necessary to collect these cell types compared to other growth stages (Figueroa et al. 2009). Furthermore, certain plant species flower only briefly, limiting the collection window to a few days (Figueroa et al. 2009). In other species, such as *Arabidopsis*, the anthers can be so small that the isolation of the PMCs is quite difficult (Koornneef et al. 2003, Figueroa et al. 2009). Neither of these last two issues is likely to be a problem in cotton since flowers appear almost daily over a two month span and are relatively large in size.

FISH on meiotic pachytene chromosomes is an important cytogenetic technique, but it lags behind in cotton due to difficulties in chromosome preparation (Peng et al. 2012). With a high chromosome count (2n=4x=52) and length (40X the length of metaphase), it is easy for the pachytene chromosomes of tetraploid cotton to become mixed and intertwined with each other (Peng et al. 2012). Failure to get a good chromosomal spread can prevent the

application of FISH, since the chromosomes cannot be distinguished from one another (Peng et al. 2012). Also, a hard cell wall and thick organic constituent of secondary metabolites in the cytoplasm can hinder FISH on cotton pachytene chromosomes (Wang et al. 2009, Peng et al. 2012).

After two unsuccessful attempts, Wang et al. (2009) documented the first successful FISH on tetraploid cotton (*Gossypium hirsutum*) pachytene chromosomes. Previous attempts to directly squash the PMCs without an enzyme solution or to digest them with a weak (3%) enzyme solution were unable to obtain acceptable spreads (Wang et al. 2009). Both of these approaches used ice-cold ethanol: acetic acid (3:1) to spread the chromosomes (Wang et al. 2009). However, a long digestion with a strong enzyme solution (5-6 hours at 37^{o}C with 2% cellulase and 2% pectolyase) followed by squashing the PMCs over a hot plate produced satisfactory spreads for FISH.

A second common source of chromosomes for FISH is the mitotic interphase nuclei (Jiang and Gill 2006, Peng et al. 2012). These chromosomes can provide the same or better resolution as meiotic pachytene chromosomes by serving as a source of extended DNA fibers (Jiang and Gill 2006, Peng et al. 2012). Additionally, interphase nuclei can be harvested from cotyledons, reducing the time and resources necessary for collection (Peng et al. 2012). Etiolated seedlings are often used in the isolation of interphase nuclei as the absence of light results in cotyledons with fewer chloroplasts and other debris that can hinder extraction (Peng et al. 2012). However, it can be difficult to obtain interphase nuclei and chromosomes with sufficient quality for FISH (Peng et al. 2012).

The most common target for FISH is the mitotic metaphase chromosome which are ideal for chromosome counting, rough mapping, determination of gene copy number and distribution, and GISH. These chromosomes can be easily harvested from root meristematic cells a few days after germination (Figueroa et al. 2009). Seeds are usually germinated in the lab and can be harvested in 3 to 5 days when the roots are about 1-2cm in length (Singh 2003). Somatic metaphase chromosomes are highly compact and small which makes them ideal for chromosome counting experiments (Figueroa et al. 2009). FISH on metaphase

chromosomes can be used readily to identify regions of repetitive and multi-copy DNA, but had lacked the sensitivity and resolution to localize unique and single-copy sequences in the genome (Jiang and Gill 1994). However, with the improvements in mitotic metaphase chromosome spreads, FISH using single copy gene is becoming a reality in corn, wheat, soybean, and *Arabidopsis* (Danilova 2012). In order to obtain mitotic metaphase chromosome spreads, various agents can be used to help synchronize the cell cycle and arrest a large proportion of the cells at metaphase.

4.1.2 Synchronizing the Cell Cycle

The root meristem is a highly active region of dividing cells that are at various stages of mitosis. Synchronization is performed in order to accumulate more mitotic cells at the metaphase stage. The metaphase stage is desirable because chromosomes reach maximum condensation at this stage, which in turn results in the appearance of distinct morphological features. Synchronization takes places before metaphase arrest which is the actual stopping of mitosis at the metaphase stage. Coordinating the cell cycle prior to the application of a metaphase arrestor can help increase the proportion of the cells arrested, providing more candidates for FISH and observation.

Halfmann et al. (2007) examined the treatment of cotton (*G. hirsutum*) root tip cells with hydroxyurea (HU) prior to the application of a metaphase arrestor. HU is a reversible nucleotide synthesis inhibitor that freezes cells in either the late G_1 or early S phase (Halfmann et al. 2007). It inhibits the activity of the enzyme ribonucleotide reductase, which in turn blocks the production of deoxyribonucleotides and ribonucleotides, thus preventing DNA synthesis (Halfmann et al. 2007). Upon removal of HU, cells should start dividing again and reach metaphase at roughly the same time, increasing the effectiveness of the subsequent arrestor treatment (Halfmann et al. 2007). Cells were treated with HU for ~18 hours and the proportion of cells at metaphase (i.e. the metaphase index) was maximized roughly 4.5 hours after removal of HU (Halfmann et al. 2007). Treatment with 3.5mM HU can increase the metaphase index by almost four-fold (without the subsequent addition of a

metaphase arrestor) (Halfmann et al. 2007). However, the increase was only from ~2% of cells to just under 8%, which is far less than the 50% observed in other species (Dolezel et al. 1999, Halfmann et al. 2007). In other studies, the concentration of HU did not exceed 2.5mM (Halfmann et al. 2007). High concentrations of HU can slow recovery and prevent synchrony (Halfmann et al. 2007) The increased concentration needed in cotton to obtain optimal results could be the result of its tetraploid nature (Halfmann et al. 2007). Multiple or paralogous copies of the ribonucleotide reductase gene may increase resistance to HU and dictate the use of higher concentrations in order to obtain acceptable results (Halfmann et al. 2007).

In wheat, rice and barley, 24 hour ice water treatment is given after HU and amiprophos-methyl (APM) treatment to synchronize the cell cycle and accumulate more cells at the metaphase stage (Pan et al. 1992). Later, colchicine treatment is given to prevent the growth of spindle fibers, thereby arresting the cells at metaphase. Other treatments to promote synchronization such as nutrient starvation, mimosine, or aphidicolin have yet to be tested in cotton (Arumuganathan et al. 1991, Halfmann et al. 2007).

4.1.3 Pre-treatments for Arresting at Metaphase

One of the earliest examples of the manipulation of the cell cycle for cytological purposes appears in Hsu and Pomerat (1953). Here, a mild hypotonic salt solution was used to bathe the cells prior to observation (Hsu and Pomerat 1953). This caused the cells to take up water and swell, which in turn negatively affected the formation of the spindle fibers (Hsu and Pomerat 1953). More recently a hypotonic 0.075M KCl solution was used as a pre-treatment in rice and potato (Kurata and Omura 1978, Pijnacker and Ferwerda 1984, Singh 2003).

Other documented pre-treatments for somatic metaphase chromosomes include: 8-hydroxyquinoline for species with small chromosomes (Tjio and Levan 1950, Kato 1999, Singh 2003), the classical spindle poison colchicine (Kato 1999, Singh 2003, Halfmann et al. 2007), bromonapthalenes in wheat and barley (Kato 1999, Singh 2003), and

paradichlorobenzene (PDB) for plants with small chromosomes (Palmer and Heer 1973, Singh 2003). Problems associated with some of these chemicals include poor water solubility, reduced effectiveness with larger chromosomes, poor affinity in plants, and toxicity to humans (Singh 2003). Longer treatment times with most of these chemicals are known to increase the metaphase index, but can also kill cells and damage the chromosomes (Halfmann et al. 2007). In addition there are numerous other chemicals known to target plant spindle fibers, including members of the dinitroanilines and phosphorothioamidates, although most of these have not been tested for their effectiveness in chromosome preparations (Kato 1999, Halfmann et al. 2007).

Kato (1999) recommended the use of nitrous oxide (N_2O) to obtain the highest numbers of countable, non-overlapping chromosomes for FISH. Nitrous oxide was compared to colchicine, trifluralin, amiprophos-methyl (APM), 8-hydroxyquinoline, and two temperature treatments (cold and cold-hot-cold) in maize (Kato 1999). Trifluralin is a member of the dinitroanilines and APM is a member of the phosphorothioamidates (Halfmann et al. 2007). The superiority of the N_2O treatment is not only due to a higher metaphase index, but also to the superior dispersing effect of high pressure N_2O on metaphase chromosomes (Kato 1999). Excluding N_2O, treatment with 0.04% 8-hydroxyquinoline for three hours at room temperature provided the best results in maize (Kato 1999).

Similarly, Halfmann et al. (2007) compared nitrous oxide to colchicine, APM and two novel chemicals (the benzamide RH-4032 and the phenylcyclohexene colchicine mimic RH-9472) in both synchronized and unsynchronized cotton (*G. hirsutum*) root tip cells (Halfmann et al. 2007). Colchicine was the least effective of the chemicals tested, which was not surprising given its low affinity for plant microtubules (Halfmann et al. 2007). RH-9472 showed an exceptional ability to increase the metaphase index but, similar to APM, colchicine, and RH-4032, caused a severe clumping of the metaphase chromosomes making most preparations unsuitable for FISH (Halfmann et al. 2007). Nitrous oxide, administered at pressures of 800 or 1000 kPa (~8-10atm) for 1, 3, 4, or 5 hours showed little ability to

increase the metaphase index compared to the other chemicals (Halfmann et al. 2007). However, the N_2O treatment still displayed the highest quality preparations due to its spreading effect on the chromosomes (Halfmann et al. 2007). N_2O treatment was also tested in combination with the other four metaphase arrestors in an attempt to combine the favorable aspects of each treatment (Halfmann et al. 2007). None of the combinations showed any improvement over the N_2O treatment alone (Halfmann et al. 2007). The cells remained well spread with no increase in the metaphase index (Halfmann et al. 2007). Halfmann et al. (2007) suggested that cotton may be overly sensitive to N_2O when compared to other crops resulting in excessive arrest before metaphase, or cell death.

4.1.4 Fixatives

Fixation of cells at the chosen cell stage is a necessary step in the preparation of chromosomes (Singh 2003). A good fixative will not cause any distortion, swelling, or contraction of the chromosomes (Singh 2003). The most commonly used fixative in chromosome preparation is Carnoy's Solution 1 which consists of a 3:1 mix of 95-100% ethanol and glacial acetic acid, respectively (Singh 2003). It is prepared fresh right before use and is widely used in meiotic, mitotic, and interphase preparations (Singh 2003, Peng 2012). Kato (1999) and Kato et al. (2004) used 90% acetic acid as a fixative following pre-treatment with nitrous oxide. A 2:1 solution of acetone: acetic acid has also been used in cotton (Crane et al. 1993, Ji et al. 1996, Ji et al. 2007). Other fixatives include Carnoy's Solution 2 (1:3:6 of glacial acetic acid: chloroform: 95-100% ethanol) and propionic acid alcohol solution (1:3 of propionic acid: 95-100% ethanol with 1g/100mL of ferric chloride) (Singh 2003). The propionic acid alcohol solution works well with small chromosomes and provides clear cytoplasm and good staining (Swaminathan et al. 1954, Singh 2003). Fixation times range anywhere from 10 minutes to several months (Kato 1999, Singh 2003). Following fixation, meiotic and mitotic preparations are commonly washed with water (Crane 1993, Hanson 1996, Kato 1999, 2004, Peng 2012) although a more recent approach is

to use citrate buffer before digestion (Kato 2006). Interphase nuclei are commonly baked for 30 minutes on a hot plate at 60^{o}C following fixation (Li et al. 2005, Peng et al. 2012).

4.1.5 Enzyme Digestions

One of the reasons animal cells are significantly more amenable to FISH than plants is that the thick cell walls of plants inhibit the preparation of chromosomes (Singh 2003). This is especially true in cotton, which has a relatively thick cell wall (Peng et al. 2012). The thick cell wall components of plants makes chromosome squashing more challenging and also requires technical skill to apply optimal and even pressure to the cells to generate good spreads of the entire chromosome complement. Besides the cell wall, cellular debris reduces the efficiency of FISH by interfering with the probe hybridization and increasing background noise. Enzyme digestions consist of varying concentrations of cellulase and pectinase that weaken and break down these cell walls, facilitating chromosome preparations (Singh 2003). This approach is also commonly referred to as the enzyme maceration method. Concentrations for pectinase range from 0.5% to 40% and for cellulase from 1.5% to 6% (Kurata and Omura 1978, Pijnacker and Ferwerda 1984, Geber and Schweizer 1988, Singh 2003, Peng et al. 2012). The length of the digestion can be anywhere from 15 minutes to 4 hours and generally take place at the enzyme's optimal temperature of 35-37^{o}C (Kurata and Omura 1978, Singh 2003, Peng et al. 2012). Following digestion, cells are generally washed in order to remove all remaining digestive enzymes (Kurata and Omura 1978, Gerber and Schweizer 1988). Samples are then re-suspended in fixative before being applied to the slide (Kurata and Omura 1978, Gerber and Schweizer 1988).

4.1.6 Preparing the Slide

The classical way to prepare slides of chromosome preparations is the squash technique. After being placed on the slide, previously digested or weakened cells are flattened and spread using physical pressure created by the placement of a cover slip over the solution (Gerber and Schweizer 1988). This helps to break apart the macerated cell wall,

open up the cell, clear the cytoplasm, and spread the chromosomes (Rayburn and Gold 1982). However, in order to be used for FISH, GISH, and certain kinds of staining such as Giemsa C- or N- banding, the cover slip must be removed (Singh 2003). This can be done using either liquid nitrogen (Pryor et al. 1980) or dry ice (Gill and Kimber 1974, Stack and Comings 1979, Chow and Larter 1981, Rayburn and Gold 1982). However, this approach can often lead to damaged chromosomes and poor quality (Rayburn and Gold 1982). Furthermore, equal and even pressure across the cover slip is hard to obtain and chromosomes can easily be fragmented or lost, especially near the edges of the cover slip (Fox and Zeiss 1961). Cell wall fragments also contribute to the loss of chromosomes or their damage as these fragments unevenly adhere to the cover slip.

Air-drying is an alternative method to the squash technique that was developed by Rothfels and Siminovitch (1958) (Fox and Zeiss 1961). It can be used in plants and animals to obtain preparations that are of higher quality than those resulting from the squash method (Gerber and Schweizer 1988). The change in surface tension during the drying period is effective at spreading the chromosomes to optimal distances (Fox and Zeiss 1961). A third technique, dropping, involves some technical skill and is performed by dropping a cell suspension from a pre-determined height onto either frozen or heated slides (Pradeep 2011). However, mixing the protoplasts in acetic acid and methanol and dropping them onto the slide resulted in the efficient spreading of chromosomes in corn (Danilova and Birchler 2008).

4.1.7 Visualization and Hybridization

Initial cytogenetic maps relied on easily observable, physical characteristics such as total length, arm length, arm ratio, and centromere position in order to identify chromosomes (Figueroa et al. 2009). Additionally, simple stains like silver nitrate and Carmine-based stains were used to detect Nucleolar Organizing Regions (NORs) and heterochromatin vs. euchromatin, respectively (Wang et al. 2008, Figueroa et al. 2009). Silver nitrate easily incorporates into the proteins found at the NOR and Carmine stains chromatic regions and

centromeres dark red, while leaving euchromatic regions lighter red or unstained (Wang et al. 2008, Figueroa et al. 2009).

More advanced staining techniques followed, namely quinacrine and Giemsa (Figueroa et al. 2009). Quinacrine is a fluorescent dye that binds preferentially to AT-rich regions of the genome (Figueroa et al. 2009). Therefore, gene-poor, heterochromatic regions incorporate the most dye and stain darkest, while gene-rich euchromatic regions stain lightly or not at all (Figueroa et al. 2009). This approach, called Q-banding, results in a gradient of stained regions on each chromosome based on the base pair composition of each chromosome. (Vosa and Marchi 1972, Figueroa et al. 2009). Giemsa staining involves the de-naturation of DNA with trypsin or acid (Lavania 1978, Figueroa et al. 2009). It takes advantage of the fact that heterochromatin re-anneals faster than euchromatin due to its short, repetitive nature (Gill and Kimber 1974, Figueroa et al. 2009). This technique stains heterochromatin dark and euchromatin light, but without the gradient seen with Q-banding (Figueroa et al. 2009). There are numerous modifications to Giemsa staining. The classical approach described above is referred to as C-banding and specializes in the identification of centromeres. G-banding, which is most similar to Q-banding, involves the pre-treatment of chromosomes with a saline solution to cause stronger staining of AT-rich regions. Reverse- or R-banding involves denaturing the chromosomes with heat, followed by cooling them to 70^{o}C before staining. As its name implies, R-banding is the reverse of the other Giemsa stains, turning euchromatin dark and leaving heterochromatin light or unstained.

Karyotyping based on physical features and simple stains present numerous problems, especially in plants. Many important crop species, including cotton, are polyploids giving them large numbers of chromosomes that must be identified (Hanson et al. 1995, Figueroa et al. 2009). With a large number of chromosomes it is likely that some will have similar banding patterns that make them indistinguishable. Furthermore, almost all of cotton's chromosomes are small and either metacentric or sub-metacentric, making them physically indistinguishable from one another (Hanson et al. 1995). Meta- and sub-metacentric chromosomes contain centromeres that are at or near the center of the

chromosome, yielding arms that appear to be equal in length. Acrocentric chromosomes, where the centromere is biased towards one end resulting in arms of noticeably different length, are highly preferred for efficient physical karyotyping. This lack of physically identifiable features and differential stains necessitated the need for a more efficient tool for karyotyping that resulted in in situ hybridization (ISH).

Gall and Pardue (1969) and John et al. (1969) are credited with the development of in situ hybridization (ISH), also known as hybridization histochemistry (Jiang and Gill 1994, Figueroa et al. 2009). These approaches utilized radioactively labeled RNA probes and eventually developed the capacity to detect single-copy DNA sequences (Gall and Pardue 1969, John et al. 1969, McNeill et al. 1991, Jiang and Gill 1994, Figueroa et al. 2009). Despite these successes, the long exposure times, poor resolution, dangers of isotopes, and high work requirement led to the development of non-isotopic, enzymatic-based ISH in 1982 (Langer-Safer et al. 1982, Manuelidis et al. 1982, Rayburn and Gill 1985, Jiang and Gill 1994).

In this approach, biotinylated dUTP is incorporated into the probe (Jiang and Gill 1994, Figueroa et al. 2009). The DNA is denatured and the probe is allowed to bind to complementary site(s) in the genome. Excess probe is then washed away. Streptavidin/avidin, a molecule known to form very tight, borderline irreversible bonds to biotin, is then added (Jiang and Gill 1994). Streptavidin will only bind where it finds biotin and any excess molecule is washed away. An enzyme, such as horseradish peroxidase or alkaline phosphatase, is conjugated to the streptavidin before its addition (Jiang and Gill 1994). Therefore, when the substrate is added, signals are detected only at the probe binding sites, which should correlate to the desired DNA sequences. In an alternative method, the plant steroid dioxigenin is incorporated into the probe and is recognized by enzymatically-tagged antibodies specific to the steroid (Jiang and Gill 1994). The major advantage of enzyme-based ISH over related techniques is that the strength and duration of the signal are easily manipulated by the amount of substrate added (Jiang and Gill 1994). Furthermore, the signal will not fade over time (Jiang and Gill 1994).

Fluorescent in situ hybridization (FISH) uses fluorophores to replace the enzyme-substrate reaction as the reporter system in ISH (Langer-Safer et al. 1992, Pinkel et al. 1986, Jiang and Gill 1994, Kato et al. 2006, Figueroa et al. 2009). Fluorophores work by absorbing light at one wavelength and re-emitting it at another, longer wavelength. The wavelengths, amount of, and time before emission are all characteristics of the specific fluorophore. There are multiple advantages to using FISH over the other ISH techniques. First, fluorophores exist in a variety of colors, allowing multiple sequences to be probed and differentiated simultaneously (Jiang and Gill 1994, Figueroa et al. 2009). In addition, multiple colors can be incorporated into a single probe in order to produce new colors (i.e. green and red fluorophores to make a yellow probe) (Birchler and Danilova 2009). Second, better imaging via sophisticated microscopes attached to cameras and computers means more resolution and more powerful data processing applications (Jiang and Gill 1994, Figueroa et al. 2009). A third advantage of FISH is that the amount of probe, exposure time, and amount of fluorophores in the probe can all be easily adjusted in order to optimize the system (Birchler and Danilova 2009).

4.1.8 Preparation of the Probe

Probes for FISH come in two general categories: those from multi-copy/highly repetitive DNA sequences which are relatively easy to work with and those from low-copy/unique regions which are harder to detect, but generally more valuable (Singh 2003). The multi-copy sequences are usually repeats like the satellite sequences that make up centromeres, telomeres, and heterochromatin and gene clusters like the rDNA genes that exist as tandem clusters (Kato et al. 2006). These sequences are relatively large and usually at least five times the size of the probe so that binding happens readily (Kato et al. 2006). Probes for single-copy or unique regions are much smaller (usually 1-3kb) (Jiang and Gill 2006). They should be long enough to bind a unique sequence but short enough to avoid any repetitive sequences flanking the unique region that can cause background noise (Kato et al.

2006). Also, if the probe is too short, it will not contain enough fluorophores to generate a strong enough signal to distinguish it from background noise (Kato et al. 2006).

Probes can be labeled either directly or indirectly (Kato et al. 2006). In indirect labeling, the probe itself does not contain fluorophores, but instead carries either biotin or dioxigenin (Kato et al. 2006). The bound probe is recognized by streptavidin or the appropriate antibody which is conjugated to a fluorophore (Kato et al. 2006). The major problem with indirect labeling is that the antibodies and streptavidin can erroneously bind to non-specified parts of the chromosome or other debris on the slide (Kato et al. 2006). This can generate a background signal which can drown out the signal of a correctly hybridized probe (Kato et al. 2006). Furthermore, both of these are rather large molecules that can have difficulty reaching the probes, especially on condensed chromosomes and slides with lots of debris (Kato et al. 2006). These problems can be avoided by direct labeling in which dNTP's with fluorophores attached to them are incorporated into the probe (Kato et al. 2006). This approach greatly reduces background noise but can suffer from weaker signals (Kato et al. 2006). Solutions to this problem are to incorporate more fluorophore into the probe by using more DNA-Polymerase I during nick translation or to use a signal amplification system (Jiang and Gill 2006, Kato et al. 2006).

Tagging of the probe is done by nick translation in which three of the nucleotides are normal and one has been specially modified to carry a fluorophore group in a non-intrusive manner (Li et al. 2005, Kato et al. 2006). Which nucleotide carries the fluorophore varies on the fluorophore used (Kato et al. 2006). In nick translation, dsDNA to be labeled is treated with a DNase that induces single-stranded breaks in the DNA. DNA-Polymerase I is added, which recognizes the single-stranded nicks and binds to the exposed 3'-OH group (Lewin 2008). It uses its 5'→3' exonuclease activity to remove the nucleotides ahead of it and replaces them with the provided free-floating dNTPs (Li et al. 2005, Kato et al. 2006, Lewin 2008). When DNA-Polymerase I detaches from the DNA strand, it leaves another nick in the phosphate backbone (Lewin 2008). This nick is either sealed by DNA-Ligase or used again

by a DNA-Polymerase I as a starting point. Labeled probes can also be made by using PCR and supplying a tagged nucleotide.

4.2 Introduction

Upland cotton (*Gossypium hirsutum*) is an allotetraploid (2n=4x=52, AADD) with a genome size of approximately 2.5 Gb (Hendrix and Stewart 2005, Chen et al. 2007). Polyploid Upland cotton was formed by a chance hybridization between *G. raimondii* (2n=2x=26, DD) and either *G. herbaceum* (2n=2x=26, AA) or *G. arboreum* (2n=2x=26, AA) that occurred 1-2 million years ago (Wendel and Cronn 2003, Chen et al. 2007). Cotton is the most important and abundant natural fiber source in the world. It has an estimated economic impact of $500 billion per year worldwide and provides 300,000 jobs across both its production and associated textile industry in the United States alone (Wendel and Cronn 2003, Chen et al. 2007). Despite its importance, cotton lags behind many other important agronomic species, such as maize and wheat, in all manner of genetic progress and resources. Nowhere is this gap wider than in the field of cytogenetics.

The first step in the development of cytogenetic resources for any organism is the development of an effective and efficient way to number and identify the chromosomes (Jiang and Gill 2006). Once such a karyotype has been constructed, physical mapping, or the correlation of DNA sequence data to the chromosome, can be performed (Jiang and Gill 2006). Physical mapping remains the only way to determine the actual location of a gene within the genome (Li et al. 2005). Furthermore, it can help estimate the physical distance between two gene sequences, determine gene copy number, measure the length of sequences, and detect re-arrangements and other abnormalities (Li et al. 2005).

Fluorescent in situ hybridization (FISH) is a cytogenetic technique that can be used to map the physical location of genes and transgenes, determine aneuploids, detect chromosomal aberrations, study mitosis and meiosis, and analyze other variations in chromosome structure and orientation (Kato et al. 2004, Peng et al. 2012). However, none of this can be done without an effective way to karyotype. Additionally, genomic in situ

hybridization (GISH), BAC-FISH, and single copy probe detection are important variations of FISH that also require an efficient way to identify the chromosomes.

The development of a karyotype depends upon an efficient way to develop chromosome preparations, another resource that is currently unavailable to cotton cytogeneticists. Meiotic pachytene chromosomes are ideal for some techniques, such as determining gene order and fine physical mapping, due to their excellent length, resolution, and specificity (Kato et al. 2004). However, these chromosomes are only available from a certain type of cell in the anthers (Kato et al. 2004). This limits their availability as plants must be grown to sexual maturity and can only be harvested during flowering (Kato et al. 2004). Additionally, the large size of these chromosomes prevents their use in other cytogenetic applications such as chromosome counting, aneuploid detection, and rough physical mapping. Conversely, mitotic root cells can be harvested from the meristems of actively dividing root tips within a few days of germination. The downside to these chromosomes is that they are highly condensed, reducing the resolution and sensitivity of FISH (Kato et al. 2004). However, their small size and compact nature makes them ideal for many of the applications not appropriate for meiotic pachytene chromosomes. Furthermore, Wang et al. (2009) showed that pachytene chromosomes offered no significant advantage in terms of FISH sensitivity over metaphase chromosomes in cotton. This was despite an 18 fold increase in the resolution and despite numerous findings in other studies that pachytene chromosomes significantly improved sensitivity (Wang 2009). The failure of cotton pachytene chromosomes to increase resolution was attributed to cotton's high levels of secondary metabolites and excessively small and compact chromosomes. Nevertheless, this provides further support for the development of an effective somatic preparation (Wang et al. 2009).

Higher mitotic indexes and well spread or dispersed chromosomes whose morphology is not significantly affected during the preparation process are an essential requirement for mitotic chromosome spreads. This ensures their efficient utilization for karyotyping and in situ hybridization experiments such as FISH and GISH. Efficient

chromosome spreads with less contamination and/or hindrance from cell wall and cytoplasmic debris is even more important for performing FISH with single copy probes in plants. This is especially true in cotton due to the copious and diminutive disposition of the chromosomes, which plagues the identification of physical landmarks such as centromeric constrictions and the discrimination of the sister chromatids in mitotic metaphase cells. Furthermore, due to its dicotyledonous nature, cotton has only the main tap root available for mitotic chromosomal studies. Therefore, there is an urgent need to develop a protocol that could increase the mitotic index and improve the efficiency of chromosomal spreads while limiting contamination from cell wall debris.

Several methods have been used to increase the mitotic index in plants such as ice water, hydroxyurea (HU), mimosine, aphidicolin, and nutrient deprivation (Arumuganathan et al. 1991, Pan et al. 1992, Halfmann et al. 2007). Historically, compounds such as colchicine and amiprophos-methyl (APM) are then used to arrest root cells at metaphase (Hanson et al. 1996, Halfman et al. 2007). In animals, cells have long been treated with a hypotonic solution to improve the spread of chromosomes through swelling and bursting of the cells (Hsu and Pomerat 1953). However, this approach is not as effective in plants since the presence of the cell wall inhibits lysis.

The use of nitrous oxide (N_2O) to produce large quantities of ideal chromosome preparations was first shown in maize (Kato 1999). The superiority of the N_2O treatment comes not from an increase in the number of cells arrested, but rather from a pronounced positive effect on the spread of the chromosomes (Kato 1999, Halfmann et al. 2007). This is in contrast to most other chemicals with metaphase arresting characteristics such as colchicine and APM, which tend to cause a cytologically undesirable clumping of the chromosomes (Kato 1999, Halfmann et al. 2007). Since nitrous oxide is applied as a high pressure gas, exposed cells bugle or swell in a manner similar to that of a hypotonic solution. This helps to spread the chromosomes after the cell walls are macerated by enzyme digestion. The drop method, as opposed to the squash or smear method, also facilitates chromosomal spread and helps to remove cell components and other debris left over from

digestion. Good spreads are essential to unambiguously differentiate chromosomes from one another and facilitate efficient karyotyping and physical mapping.

The development of a protocol for the efficient preparation of cotton metaphase chromosomes, and a subsequent karyotype based on these metaphase chromosomes would be considered a major boon to cotton cytogenetics research. In an attempt to develop a fast, reliable, and effective way to prepare cotton chromosomes for FISH, the protocol outlined in Kato (1999) was performed with minor modifications. This new method of preparing chromosome spreads was tested for its compatibility to in situ hybridization by using rDNA as a probe in fluorescent in situ hybridization (FISH). Additionally, it has been noted that the longer the treatment with nitrous oxide, the more condensed the chromosomes become (Kato 1999). Therefore, varying treatment times of nitrous oxide were investigated in order to find the optimal duration of nitrous oxide exposure in cotton.

4.3 Materials and Methods

4.3.1 Preparation of Chromosomes from Root Tips

A modified version of the nitrous oxide and air drying method demonstrated in maize by Kato (1999) was used to prepare cotton mitotic metaphase chromosomes for fluorescent in situ hybridization. Seeds of the *Gossypium hirsutum* variety TM-1 (NC2630, PI 607172 originally obtained from Josh Lee/ Daryl Bowman) were germinated in moist vermiculite within plastic germination boxes (Figure 12). The seeds were kept in the dark at 30^{o}C in an incubator (VWR Digital Laboratory Incubator, VWR International, Radnor, PA) for two to three days.

Prior to root tip removal, standard 1.5mL Eppendorf micro-centrifuge tubes were prepared for sample collection. Four holes were placed in the top of each tube with a dissecting needle. From healthy and vigorous seedlings, 2-2.5cm long root tips were removed and placed in the micro-centrifuge tubes. One to two roots were placed in each tube and the tubes were misted lightly with ddH_2O using a standard spray bottle.

The 1.5mL tubes were placed in a custom-built gas-pressure chamber (Figure 13a-b) manufactured by the Murr Instrument Shop located at the Research Park of the University of Missouri-Columbia. Nitrous oxide was applied at a pressure of 160psi (~10.9 atm) for 1 hour and 45 minutes using a custom-built apparatus (Figure 13). To test the effect of length of N_2O exposure time on chromosome condensation, times of 80, 95, and 105 minutes were also used. When the exposure time was over, the gas was released slowly into the fume hood.

Cells in the root tips were then fixed by filling each tube with 90% ice-cold acetic acid and incubating on ice for 15 minutes. Following fixation, root tips were transferred to fresh tubes containing ice-cold 70% ethanol. The solution was gently mixed using a glass pipette; ethanol was removed, and then replaced with fresh ice-cold 70% ethanol. This was repeated 2-3 times. Following the final removal of ethanol, ice-cold 1x citric buffer was added. Samples were mixed gently using the glass pipette and citrate buffer was constantly removed and replaced with fresh buffer for 10 minutes.

Following the last removal of citrate buffer, the root tips were cleaned on filter paper. A 1-2mm long section of the milky white region was removed and transferred to a standard 0.7mL Eppendorf tube containing 20µL of ice-cold enzyme solution (1% pectolyase Y-23 (MP Biochemicals, Solon, OH), 4% cellulase Onozuka R-10 in 1x citric buffer (Yakult Pharmaceuticals, Tokyo, Japan)). The sample tubes were then placed in a Big Shot III Hybridization Oven (Boekel Scientific, Feasterville, PA) for incubation at 37^{o}C for 55 minutes.

Immediately following digestion, tubes were plunged into ice. The digested root tips were then washed with ice-cold 1x TE buffer, followed by 2-3 washes in 100% ethanol using a standard glass pipette. After thoroughly removing the ethanol, 30µL of freshly prepared 75% acetic acid – 25% methanol was added. The remains of the root tips, sufficiently weakened by the enzymatic degradation, were then ground gently using a rounded dissecting needle.

4.3.2 Dropping and Cross-Linking of Cells

Slides were labeled and kept on a Plexiglas scaffold. Using a standard pipette, 8µL of the cell suspension was dropped onto the marked spot on the slide. The slides were then allowed to dry by placing them inside a homemade humid chamber for five minutes. After drying, the 20X lens of a Leica DME compound microscope (Leica Microsystems, Wetzlar, Germany) was used to view slides in order to ensure selection of slides with good chromosomal spread for hybridization. Chosen slides where then placed on a thin piece of cardboard and cross-linked on a SpectroLinker XL-1000 UV Crosslinker (Spectronics Corporation, Lincoln, NE). Slides were then stored overnight at -20^{o}C in a standard microscope slide box.

4.3.3 FISH Analysis

The DNA probe PTA71 was used as a probe in the FISH experiment. Clone PTA71, the 18S-28S rDNA, is a 9-kb *Eco*RI fragment from common wheat (*Triticum aestivum*) (Gerlach and Bedbrook 1979). It contains the coding sequence for the 18S, 5.8S, and 25S rDNA genes and the intergenic spacer sequences (Gerlach and Bedbrook 1979). Plasmid PTA71-containing *E. coli* bacteria were grown in 3 mL of LB broth containing ampicillin in an incubating shaker. Plasmid DNA was isolated using a QIAGEN plasmid mini kit (Qiagen, Valencia, CA) according to the manufacturer's protocol. One microgram of plasmid DNA was labeled with fluorescein-12-dUTP (Enzo Life Sciences Inc., Farmingdale, NY) using nick translation according to the manufacturer's protocol (Roche Applied Sciences, Indianapolis, IN).

Five µL of denatured salmon sperm DNA (140ng/µL in 2x SSC) was dropped onto the center of each cell spread and covered with a 22x22mm cover slip. Root tip DNA (on slide) and probe DNA (in 200µL thin-wall PCR tubes) were denatured by floating in the boiling water baths for five minutes at 100^{o}C. The water bath consisted of a small aluminum tray covered with a wet (DI water) layer of Kimwipes. Slides and tubes were pressed firmly

onto the Kimwipes and covered with a plastic pipette tip box lid. The entire assembly floated in a larger tray containing boiling water on a hot plate.

After denaturation, the probes were plunged directly into ice. The denatured slides were placed on a thin metal sheet that had been placed upon the ice. The tube containing the probe was spun briefly to collect the total volume. The plastic cover slip was removed and 5µL of the probe solution (40ng of labeled probe/µL) was placed onto the center of the slide. After addition of the probe, the slides were incubated in a humid storage container lined with Kimwipes soaked with 2x SSC and placed in the Big Shot III Hybridization Oven at 37^{o}C overnight.

The slides were washed in 2x SSC at room temperature for five minutes in a Coplin jar. They were then washed at 42^{o}C in 2x SSC for 20 minutes in an incubating shaker (Precision Reciprocal Shaking Bath, Thermo Fisher Scientific, Waltham, MA). After removing the slides from the bath, the bottoms and sides were briefly dried to remove excess SSC without allowing the tops to dry. One drop of ice-cold propidium iodide (PI) pre-mixed with Vectashield mounting medium (Vector Laboratories, Burlingame, CA) was added to each slide in order to stain the chromosomes red. Slides were covered with 24x50mm cover slips prior to visualization. Cells were examined under an Olympus BX53 Dark Fluorescence microscope (Olympus Corporation, Tokyo, Japan) connected to a Prior Lumen 200 light source (Prior Scientific, Cambridge, UK). Images were captured using a Hamamatsu ORCA-03 camera (Hamamatsu Photonics, Hamamatsu, Japan

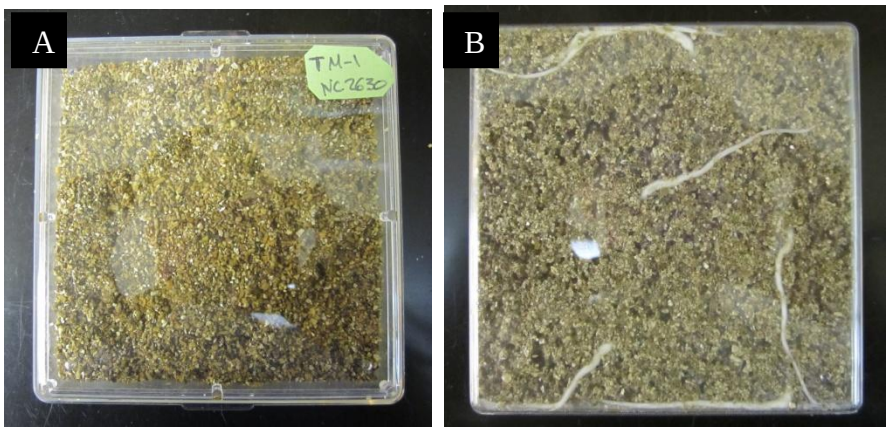


Figure 12: Germination conditions for mitotic metaphase chromosome preparation in cotton: A) Image of the plastic rectangular box containing moist vermiculite used to germinate seeds. B) Image of the bottom of the plastic box two days after planting showing the size of roots that were harvested for chromosome preparation.

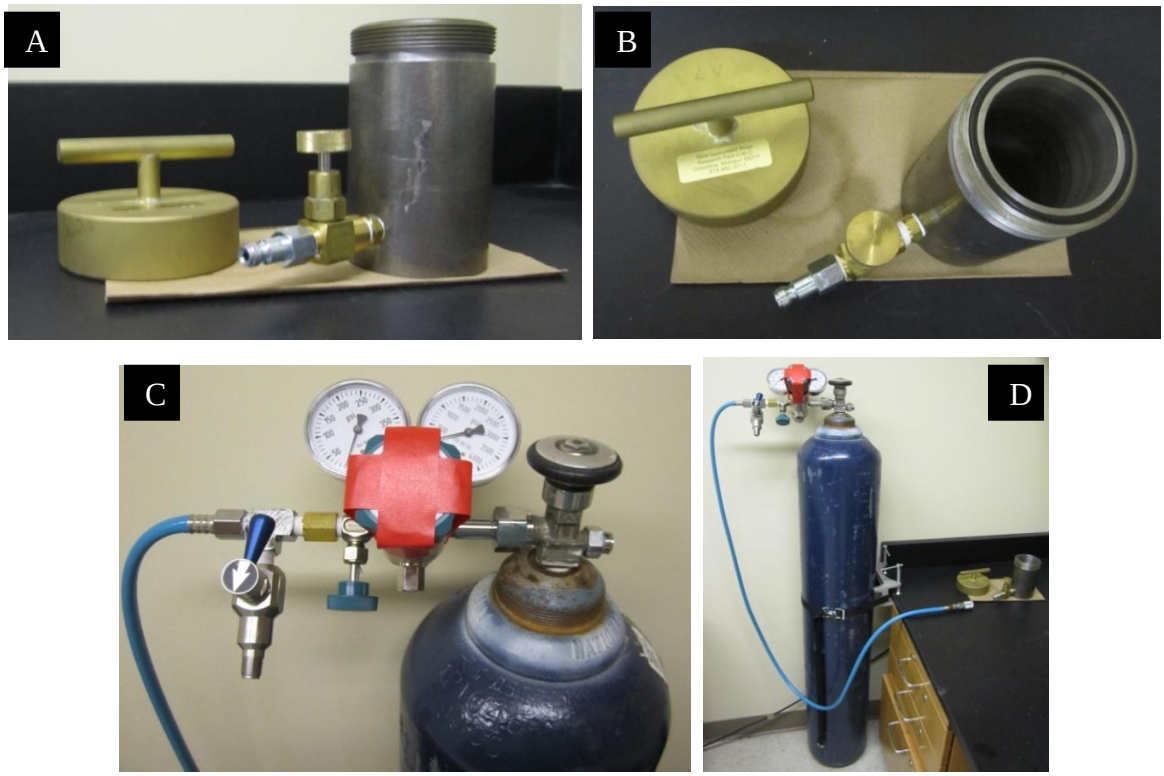


Figure 13: Custom-built gas-pressure chamber for exposing root tips to nitrous oxide. A+B) Side and overhead views of the pressure chamber in which root tips are treated. C) Close-up of the regulator built to regulate the N_2O air-flow. D) The entire N_2O delivery system.

4.4 Results

Root tips were used as source of mitotic metaphase chromosomes and pre-treated with nitrous oxide. Slides were examined briefly to ensure that an adequate chromosomal spread and metaphase index could be achieved through this method. Probes specific to the 18S-5S-26S rDNA tandem repeat from wheat (*Triticum aestivum*) were used to examine the performance of these slides in FISH.

The mitotic metaphase chromosome preparation used here generated high-quality chromosomal spreads and proved highly adept at arresting cells in metaphase. In each slide prepared, hundreds of cells could be seen arrested in metaphase. In addition, many of them had sufficient spreads to allow the chromosomes to be easily counted and differentiated from one another. In Figure 14, all 52 chromosomes of the tetraploid cotton variety TM-1 are easily visible and distinguishable. Furthermore, FISH with the 18S-5S-26S rDNA tandem repeat probe was successful. Six different bands were detected, representing three different genomic loci, each belonging to a different pair of homologous chromosomes (Figure 14). Additionally, very little background noise was detected, indicating that the hybridization protocol and probe worked efficiently in the present study.

With a two hour exposure to nitrous oxide, cotton chromosomes condense considerably. Cotton chromosomes are already small, numerous, and highly condensed when compared to other plant species. On top of this, nitrous oxide is known to produce more condensed chromosomes than other arresting agents. In an attempt to offset this limitation, shorter N_2O treatment times of 80, 95, and 105 minutes were investigated. The hope was to maintain the level of metaphase arrest and good chromosomal spread seen in the 110 minute treatment but generate longer, less condensed chromosomes to improve FISH sensitivity. The shorter N_2O treatments did result in chromosomes that were slightly longer and less condensed than the original treatment time (Figure 14 and 15). The sister chromatids and the narrowing at the centromere are more pronounced with the shorter treatments than with the longer treatment (Figure 15). In the longer treatment, sister chromatids cannot be distinguished on any chromosome and neither can the centromere (Figure 14). In contrast, with the shorter exposures, the chromosomes appear more defined

and elongated with indentations more clearly identifying the location of the centromeres and sister chromatids (Figure 14).

However, chromosome clumping and a failure to generate good spreads became a problem with the shorter exposures. This was especially true of the 80 minute treatment in which it was difficult to find cells where all 52 chromosomes could be unambiguously differentiated. Clumping was less of a concern with the 95 minute treatment times as more cells with ideal spreads were found than with the 80 minute treatment. However, their frequency was lower than that seen with the standard 110 minute treatment.

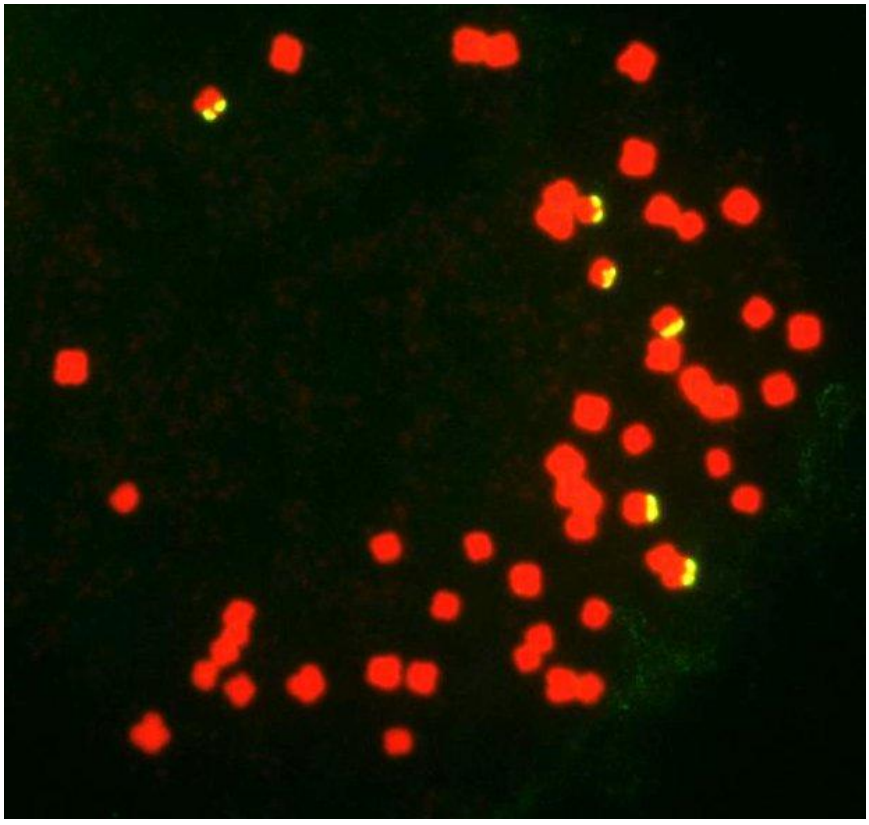


Figure 14: FISH on cotton mitotic metaphase chromosomes. All 52 chromosomes of *G. hirsutum* are visible and appear red in color. Green color represents binding of the 18S-5S-26S rDNA tandem repeat probe from wheat. Six signals (two each on three different pairs of homologous chromosomes) are visible corresponding to the three major 18S-5S-26S loci in cotton.

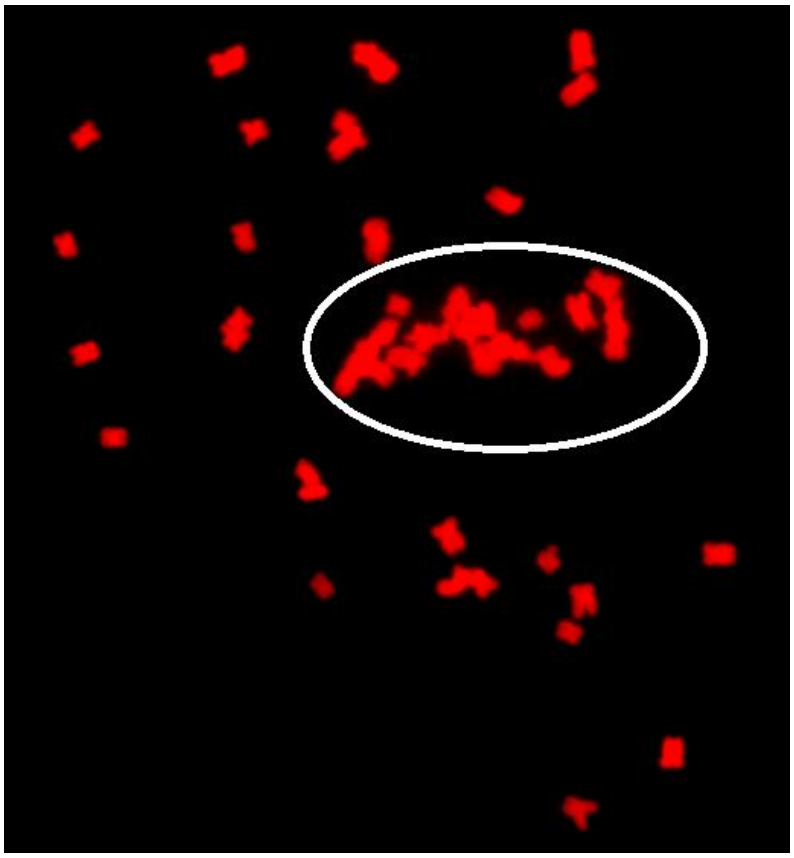


Figure 15a: Mitotic metaphase chromosomes of *G. hirsutum* inbred line TM-1 with an 80 minute exposure to nitrous oxide. Note the large clump of chromosomes circled in white.

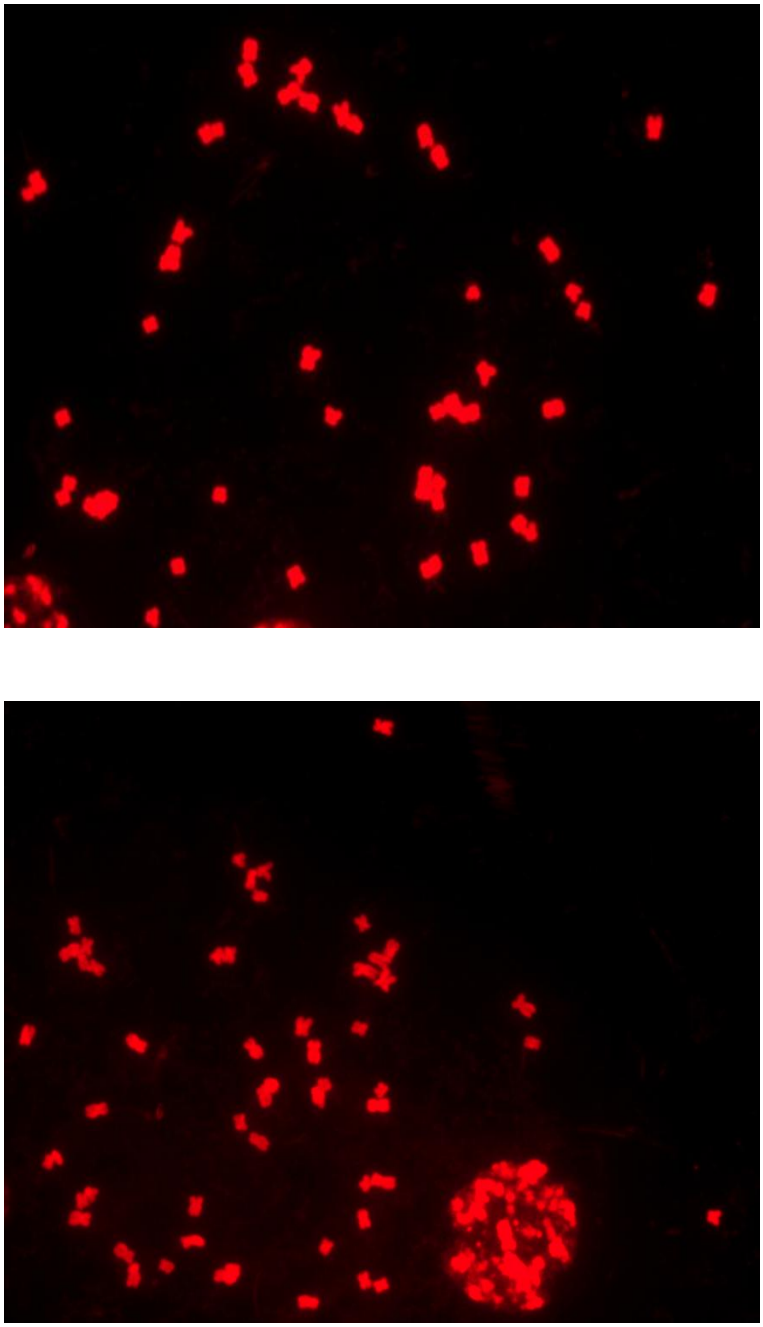


Figure 15b and c: Mitotic metaphase chromosomes of the *G. hirsutum* inbred line TM-1 with a 95 minute exposure to nitrous oxide.

4.5 Discussion

Previously, Peng et al. (2012) developed a method for isolating interphase nuclei from etiolated diploid cotton (*Gossypium arboreum*) cotyledons with an almost 100% success rate. Among other modifications, fresh cotyledons were chopped with a razor blade in ice-cold nucleus isolation buffer as opposed to grinding in liquid nitrogen (Peng et al. 2012). This idea was first demonstrated by Li et al. (2005) who showed that this approach yielded more intact nuclei in rice, wheat, and rye. The "Liquid Drainage Method" was also employed to spread the chromosomes (Peng et al. 2012). In this approach the edge of a cover slip was dragged just ahead of the solution and yielded better results than either tilting the slide or the squash method to spread the chromosomes (Peng et al. 2012).

Peng et al. (2012) also developed three modifications to the standard procedure found in the literature (Crane et al. 1993, Ji et al. 1997, Wang et al. 2009) to improve diploid cotton (*Gossypium arboreum*) pachytene chromosome preparation. First, the PMCs were isolated and digested directly, rather than applying the enzyme solution to the entire bud (Peng et al. 2012). This approach allows greater control and optimization of the digestion step (Peng et al. 2012). In order to improve the removal of cellular components, an additional incubation in 60% acetic acid plus enzyme mixture for three minutes at room temperature was added (Peng et al. 2012). Finally, an extra 50µL of ice-cold Carnoy's solution was used to facilitate the spread of the chromosomes; as opposed to the squash and cover slip removal methods which are prone to damaging and entangling the chromosomes (Peng et al. 2012).

Despite the development of excellent protocols for the preparation of pachytene and mitotic interphase chromosomes in cotton, an effective and efficient way to prepare mitotic metaphase chromosomes does not yet exist. Clumping, excessive condensation, high chromosome number and over- or insensitivity to many common arresting agents make cotton a troublesome species for developing mitotic metaphase spreads. The findings of the two studies mentioned above highlight many of the components necessary in a cotton mitotic metaphase preparation such as: a reliable method of isolating desired cell types, efficient digestion of cells, effective clearance of cellular debris, and obtaining an acceptable

chromosomal spread. Our method addresses these concerns by using an enzyme digestion directly onto isolated root tip meristematic cells while using the drop method to clear debris and spread the chromosomes along with a nitrous oxide pre-treatment.

Clumping of the metaphase chromosomes can make identifying individual chromosomes impossible and negate an otherwise ideal FISH. Obtaining an appropriate spread of the chromosomes can facilitate their counting, identification, and FISH. Many of the chemicals used as metaphase arrestors have a tendency to clump the chromosomes, decreasing their usefulness (Halfmann et al. 2007). Kato (1999), working with maize, was the first to show the dispersing effects of nitrous oxide on chromosome preparations, negating the issue of clumping seen with other arresting agents.

Halfmann et al. (2007) found that while generating spreads of great quality, nitrous oxide failed to arrest enough cells at metaphase in cotton to warrant its use in chromosome preparation. Nitrous oxide was deemed too toxic in cotton, resulting in arrest of the cell cycle before the metaphase stage (Halfmann et al. 2007). However, our results show that there are still more than enough metaphase cells for cytogenetic work such as chromosome counting and FISH. Metaphase indexes were not calculated and nitrous oxide was not compared to other chemicals in this study, but there is enough evidence to recommend the use of nitrous oxide in the preparation of cotton mitotic metaphase chromosomes. Furthermore, the great spreads provided by nitrous oxide compared to other chemicals appears to be more than enough to offset any potential decrease in the mitotic index. Even if fewer cells are arrested in metaphase by N_2O, more cells are available for cytogenetic work due to this pronounced spreading effect. The higher mitotic index observed in our study could be due to the fact that the enzyme digestion method (drop method) was used, which allows the available protoplasts to be mounted on the slide.

The exact reason why N_2O results in greater chromosomal spread is not definitively known. It may be due to the direct, physical effects of placing the chromosomes under a high pressure gas. Alternatively, it may be due to N_2O's mode of disrupting mitosis. Unlike most mitotic arrestors such as colchicine, N_2O has no effect on the formation of the spindle

fibers (Brinkley and Rao 1973). Instead, it prevents the chromosomes from properly aligning along the metaphase plate, which in turn prevents the spindle fibers from attaching to the chromosomes (Brinkley and Rao 1973). The clumping seen with other chemicals may simply be an artifact of the chromosomes aligning along the metaphase plate prior to division. The spreading effect of N_2O may then be at least partially attributed to the unnatural disorganization of the chromosomes, and this may or may not be enhanced by high pressure.

Three differences between the protocol used by Halfmann et al. (2007) and the one used here may explain the favorable results for nitrous oxide seen in this study. First, Halfmann et al. (2007) states that samples were brought to the treatment pressure with N_2O over a ten minute period. It may be this relatively long time necessary to get the samples up to the treatment pressure that resulted in poor rates of metaphase arrest. By contrast, our apparatus allowed cells to be brought to the treatment pressure almost immediately after the introduction of the nitrous oxide. The third difference is that Halfmann et al. (2007) used the protocol of Kim et al. (2002) in the preparation of slides. This approach takes digested root tips and places them directly on the slide, where they are macerated using forceps (Kim et al. 2002). By contrast, in our approach we used the drop technique that is known to produce an abundance of high-quality spreads in corn (Birchler and Danilova 2008). Digested root tips were ground using a rounded needle while still in the micro-centrifuge tube prior to dropping. The enhanced spreading effect of the drop method, while having no effect on the metaphase index, was likely to have further increased the proportion of analyzable cells.

To date, the majority of cotton cytogenetic work has utilized the squash method to spread the chromosomes (Crane et al. 1993, Hanson et al. 1996, Wang et al. 2006, Peng et al. 2012). The squash technique, which involves the placement of a cover slip over previously digested cells to flatten and spread the chromosomes, can cause damage and loss of chromosomes and requires a high degree of technical skill. The large numbers of analyzable cells seen with this approach indicate that the drop method is superior to both the squash method and the on-slide maceration approach for obtaining slides suitable for FISH in cotton.

The results of our FISH work are in agreement with the findings of Hanson et al. (1996) regarding the copy number of the 18S-5S-26S rDNA tandem repeat in the genome of *G. hirsutum*. Although Hanson et al. (1996) proposed that there at least 11 18S-5S-26S rDNA loci in cotton, only three of them were considered to be major loci. Of the remaining eight, one was considered intermediate and the other seven were deemed minor loci (Hanson et al. 1996). Our results showed six hybridization signals indicating three loci with pairs of signals coming from the same locus on homologous chromosomes. These three loci are likely the major loci detected by Hanson et al. (1996). Our procedure may have only detected the major loci due to the use of a probe from an unrelated species. Nevertheless, our results show a novel and highly efficient method for preparing cotton mitotic metaphase chromosomes that is appropriate for FISH and in some instances for chromosome counting.

By nature, cotton has relatively small and condensed chromosomes that are less than ideal for many cytogenetic applications. Furthermore, nitrous oxide is known to have a pronounced positive effect on chromosome condensation, further compounding this problem. The exact reason why N_2O causes condensation is not fully understood, but it may be due to the effect of high pressure on the chromosomes. Alternatively, as stated above, N_2O arrests the cell cycle later in metaphase than other arrestors. This may allow the chromosomes to condense further than with other chemicals. Here, a shortened exposure to N_2O leads to increased length and decreased condensation of mitotic metaphase chromosomes. This has the potential to improve the sensitivity and resolution of the N_2O pre-treatment for FISH in cotton. However, too short of an exposure to nitrous oxide results in increased clumping of chromosomes and poor spreading, negating one of the primary benefits of this approach. Therefore, a nitrous oxide treatment time of 95 minutes is recommended to produce the best mitotic metaphase chromosomes for FISH in cotton. This time provides a good balance of chromosome length and reduced condensation with an acceptable chromosomal spread.

In summary, our results show that the combination of a nitrous oxide pre-treatment with the drop method produces high-quality chromosomal spreads desirable for cytogenetic work in cotton. Furthermore, manipulation of the N_2O exposure time generates

chromosomes that are of ideal length and sufficient spread for FISH and other cytological techniques.

REFERENCES

Arumuganathan KJ, Slattery J, Tanksley S, Earle E. 1991. Preparation and flow cytometry analysis of metaphase chromosomes of tomato. *Theoretical Applied Genetics* 82: 101-111.

Bauman JG, Wiegant J, Borst P, van Duijn P. 1980. A new method for fluorescence microscopical localization of specific DNA sequences by in situ hybridization of fluorochromelabelled RNA. *Experimental Cell Research* 128: 485-490.

Birchler JA, Danilova TV. 2009. Chapter 13 – Fluorescence in situ hybridization and in situ PCR. In: Bass HW, Birchler JA, editors. *Plant Cytogenetics: Genome Structure and Chromosome Function.* Volume 4. New York: Springer 295-309.

Brinkley BR, Rao PN. 1973. Nitrous oxide: effects on the mitotic apparatus and chromosome movement in HeLa cells. *The Journal of Cell Biology* 58: 96-106.

Chen ZJ, Scheffler BE, Dennis E, Triplett BA, Zhang T, Guo W, Chen X, Stelly DM, Rabinowicz PD, Town CD, Arioli T, Brubaker C, Cantrell RG, Lacape JM, Ulloa M, Chee P, Gingle AR, Haigler CH, Percy R, Saha S, Wilkins T, Wright RJ, Deynze AV, Zhu Y, Yu S, Abdurakhmonov I, Katageri I, Kumar PA, ur-Rhaman M, Zafar Y, Yu JZ, Kohel RJ, Wendel JF, Paterson AH. 2007. Toward sequencing cotton (*Gossypium*) genomes. *Plant Physiology* 145: 1303-1310.

Chow C, Larter EN. 1981. Centromeric banding in maize. *Canadian Journal of Genetics and Cytology* 23: 255-258.

Crane CF, Price HJ, Stelly DM, Czeschin Jr. DG. 1993. Identification of a homeologous chromosome pair by in situ DNA hybridization to ribosomal RNA loci in meiotic chromosomes of cotton (*Gossypium hirsutum*). *Genome* 36: 1015-1022.

Danilova T. 2012. Physical mapping of the wheat and *Triticeae* genomes using single gene FISH. *Joint 22nd ITMI Summer Workshop and 4th U.S. National Wheat Genomics Workshop* Fargo, ND. June 25-29.

Danilova TV, Birchler JA. 2008. Integrated cytogenetic map of mitotic metaphase chromosome 9 of maize: resolution, sensitivity, and banding paint development. *Chromosoma* 117: 345-356.

Figueroa DM, Amarillo IE, Bass HW. 2009. Chapter 4 – Cytogenetic mapping in plants. In: Bass HW, Birchler JA, editors. *Plant Cytogenetics: Genome Structure and Chromosome Function.* Volume 4. New York: Springer 79-119.

Fox M, Zeiss IM. 1961. Chromosome preparation from fresh and cultured tissues using a modification of the drying technique. *Nature* 192: 1213-1214.

Gall JG, Pardue ML. 1969. Formation and detection of RNA-DNA hybrid molecules in cytological preparations. *Proceedings of the National Academy of Sciences* 63: 378-383.

Geber G, Schweizer D. 1988. Cytochemical heterochromatin differentiation in *Sinapis alba* (*Cruciferae*) using a simple air-drying technique for producing chromosome spreads. *Plant Systematics and Evolution* 158: 97-106.

Gerlach WL, Bedbrook JR. 1979. Cloning and characterization of ribosomal RNA genes from wheat and barley. *Nucleic Acids Research* 7: 1869-1885.

Gill BS, Kimber G. 1974. Giemsa C-banding and the evolution of wheat *Proceedings of the National Academy of Sciences* 71: 4086-4090.

Gill BS, Friebe B. 1998. Plant cytogenetics at the dawn of the 21st century *Current Opinion in Plant Biology* 1: 109-115.

Halfmann RA, Stelly DM, Young DH. 2007. Towards improved cell cycle synchronization and chromosome preparation methods in cotton. *The Journal of Cotton Science* 11: 60-67.

Hanson RE, Zwick MS, Chos S, Islam-Faridi MN, McKnight TD, Wing RA, Price HJ, Stelly DM. 1995. Fluorescent in situ hybridization of a bacterial artificial chromosome. *Genome* 38: 646-651.

Hanson RE, Islam-Faridi MN, Percival EA, Crane CF, Ji Y, McKnight TD, Stelly DM, Price HJ. 1996. Distribution of 5S and 18S-28S rDNA loci in a tetraploid cotton (*Gossypium hirsutum* L.) and its putative diploid ancestors. *Chromosoma* 105: 55-61.

Hendrix B, Stewart JM. 2005. Estimation of the nuclear DNA content of *Gossypium* species. *Annals of Botany* 95: 789-797.

Hsu TC, Pomerat CM. 1953. Mammalian chromosomes in vitro. II. A method for spreading the chromosomes of cells in tissue culture. *The Journal of Heredity* 44: 23-30.

Ji Y, Raska DA, McKnight TD, Islam-Faridid MN, Crane CF, Zwick MS, Hanson RE, Price HJ, Stelly DM. 1996. Use of meiotic FISH for identification of a new monosome in *Gossypium hirsutum* L. *Genome* 40: 34-40.

Ji Y, Zhao X, Paterson AH, Price HJ, Stelly DM. 2007. Integrative mapping of *Gossypium hirsutum* L. by meiotic fluorescent *in situ* hybridization of a tandemly repetitive sequence (B77). *Genetics* 176: 115-123.

Jiang J, Friebe B, Gill BS. 1993. Recent advances in alien gene transfer in wheat. *Euphytica* 73: 199-212.

Jiang J, Gill BS. 1994. Nonisotopic in situ hybridization and plant genome mapping: the first 10 years. *Genome* 37: 717-725.

Jiang J, Gill BS. 2006. Current status and the future of fluorescence in situ hybridization (FISH) in plant genome research. *Genome* 49: 1057-1068.

John HA, Birnstiel ML, Jones KW. 1969. RNA-DNA hybrids at the cytological level. *Nature* 223: 582-587.

Kato A. 1999. Air drying method using nitrous oxide for chromosome counting in maize. *Biotechnic & Histochemistry* 74: 160-166.

Kato A, Lamb JC, Birchler JA. 2004. Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *PNAS* 101: 13554-13559.

Kato A, Albert PS, Vega JM, Birchler JA. 2006. Sensitive fluorescence *in situ* hybridization signal detection in maize using directly labeled probes produced by high concentration DNA polymerase nick translation. *Biotechnic & Histochemistry* 81: 71-78.

Kim JS, Childs KL, Islam-Faridid M, Menz MA, Klein RR, Klein PE, Price J, Mullet JE, Stelly DM. 2002. Integrated karyotyping of sorghum by in situ hybridization of landed BACs. *Genome* 45: 402-412.

Koornneef M, Fransz P, de Jong H. 2003. Cytogenetic tools for *Arabidopsis thaliana*. *Chromosome Research* 11: 183-194.

Kurata N, Omura T. 1978. Karyotype analysis in rice. I. A new method for identifying all chromosome pairs. *Japanese Journal of Genetics* 53: 251-255.

Langer-Safer PR, Levine M, Ward DC. 1982. Immunological method for mapping genes on *Drosophila* polytene chromosomes. *PNAS* 79: 4381-4385.

Lavania UC. 1978. Differential staining and plant chromosomes – progress in cytogenetics. *Current Science* 47: 255-260.

Lewin B. 2008. Genes IX. Jones and Bartlett Publishers, Sudbury, MA. p.432

Li L, Yang J, Tong Q, Zhao L, Song Y. 2005. A novel approach to prepare extended DNA fibers in plants. *Cytometry Part A* 63A: 114-117.
Manuelidis L, Langer-Safer PR, Ward DC. 1982. High-resolution mapping of satellite DNA using biotin-labeled DNA probes. *Journal of Cell Biology* 95: 619-625.

McNeill JA, Johnson CV, Carter KC, Singer RH, Lawrence JB. 1991. Localizing DNA and RNA within nuclei and chromosomes by fluorescence *in situ* hybridization. *Genetic Analysis, Techniques, and Application* 8: 41-58.
Palmer RG, Heer H. 1973. A root tip squash technique for soybean chromosomes. *Crop Science* 13: 389-391.

Pan WH, Houben A, Schlegel R. 1992. Highly effective cell synchronization in plant roots by hydroxyurea and amiprophos-methyl or colchicine. *Genome* 36: 387-390.

Peng R, Zhang T, Liu F, Ling J, Wang C, Li S, Zhang X, Wang Y, Wang K. 2012. Preparations of meiotic pachytene chromosomes and extended DNA fibers from cotton suitable for fluorescence *in situ* hybridization *PLoS ONE* 7: 1-6.

Pijnacker LP, Ferwerda MA. 1984. Giemsa C-banding of potato chromosomes. *Canadian Journal of Genetics and Cytology* 26: 415-519.

Pinkel D, Straume T, Gray JW. 1986. Cytogenetic analysis using quantitative, high-sensitivity, fluorescence hybridization. *Proceedings of the National Academy of Sciences* 83: 2934-2938.

Pradeep PJ, Srijaya TC, Zain RBM, Papini A, Chatterji AK. 2011. A simple technique for chromosome preparation from embryonic tissues of teleosts for ploidy verification. *Caryologia* 64: 235-241.

Pryor A, Faulkner K, Rhoades MM, Peacock WJ. 1980. Asynchronous replication of heterochromatin in maize. *PNAS* 77: 6705-6709.

Rayburn AL, Gill BS. 1985. Use of biotin-labeled probes to map specific DNA sequences on wheat chromosomes. *Journal of Heredity* 76: 78-81.

Rayburn AL, Gold JR. 1982. A procedure for obtaining mitotic chromosomes from maize *Maydica* 27: 113-121.

Rothfels KH, Siminovitch L. 1958. An air-drying technique for flattening chromosomes in mammalian cells from *in vitro. Biotechnic & Histochemistry* 33: 73-77.

Singh RJ. 2003. Plant Cytogenetics. Second Edition. Boca Raton, FL: CRC Press.

Stack SM, Comings DE. 1979. The chromosomes and DNA of *Allium cepa*. *Chromosoma* 70: 161-181.

Swaminathan MS, Magoon ML, Mehra KL. 1954. A simple propiono-carmine PMC smear method for plants with small chromosomes *Indian Journal of Genetics and Plant Breeding* 14: 87-88.

Tjio JH, Levan A. 1950. The use of oxyquinoline in chromosome analysis. *Anales de la Estacion Experimental de Aula Dei* 2: 21-46.

Tsuchiya T. 1971. An improved aceto-carmine squash method, with special reference to the modified Rattenbury's method of making a preparation permanent. *Barley Genetics Newsletter* 1: 58-60

Vosa CG, Marchi P. 1972. Quinacrine fluorescence and Giemsa staining in plants. *Nature – New Biology* 237: 191-192.

Wang K, Song X, Han Z, Guo W, Yu JZ, Sun J, Pan J, Kohel RJ, Zhang T. 2006. Complete assignment of the chromosomes of *Gossypium hirsutum* L. by translocation and fluorescence in situ hybridization mapping. *Theoretical Applied Genetics* 113: 73-80.

Wang K, Guan B, Guo W, Baoliang Z, Hu Y, Zhu Y, Zhang T. 2008. Completely distinguishing individual A-genome chromosomes and their karyotyping analysis by multiple bacterial artificial chromosome-fluorescence *in situ* hybridization. *Genetics* 178: 1117-1122.

Wang K, Yang Z, Shu C, Hu J, Lin Q, Zhang W, Guo W, Zhang T. 2009. Higher axial-resolution and sensitivity pachytene fluorescence in situ hybridization protocol in tetraploid cotton. *Chromosome Research* 17: 1041-1050.

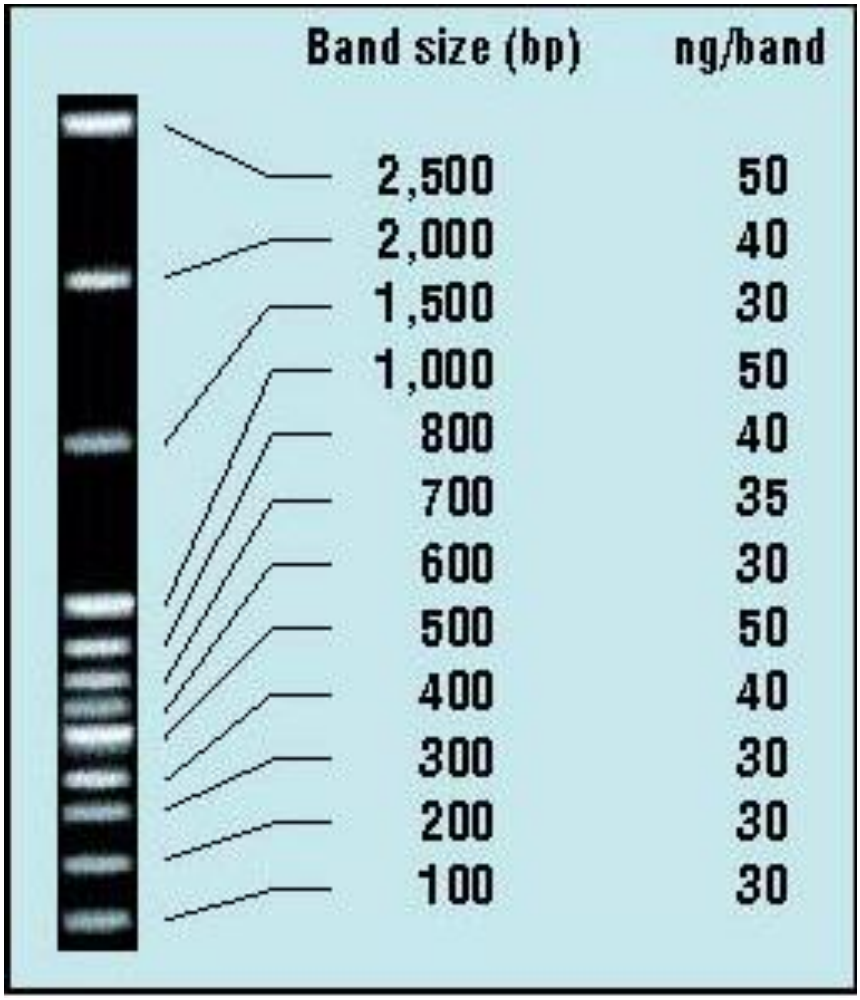
Wang K, Guo W, Yang Z, Hu Y, Zhang W, Zhou B, Stelly DM, Chen ZJ, Zhang T. 2010. Structure and size variations between 12A and 12D homeologous chromosomes based on high-resolution cytogenetic map in allotetraploid cotton. *Chromosoma* 119: 255-266.

Wendel JF, Cronn RC. 2003. Polyploidy and the evolutionary history of cotton. *Advances in Agronomy* 78: 139-186.

APPENDICES

APPENDIX A

Bioline GeneMate Quanti-Marker 100bp (5µL). Image obtained from http://www.bioexpress.com/divinity-cart/item/108100/BIOLINE-Quanti-Marker-100bp/1.html



APPENDIX B

Primer sequences for cotton leaf shape candidate genes.

EST/TC/RFLP	Putative Classification	Forward Primer Sequence 5'→3'	Reverse Primer Sequence 5'→3'	Expected Product Size (bp)
TC203839	Class I KNOX	GATCCTCTTCGGATGAACCA	GAGTTCGAAGCAAGGCAAAG	210
TC204345	Class I KNOX	ATCCATAGGGAAGGGATTGC	TGGCAGAGTCAACTGGTTTG	163
TC222827	Class I KNOX	CGCTTGTGGAGATGCTGTTA	GCAACTTCTGTCTGGCTTCC	208
TC203687	Class I KNOX	AGTGTGACGGTGTTGGTTCA	TGTAATGCAACTCCCACCAA	227
TC247240	MYB TF	GGCCTCCTTAATCCTTTGCT	AACACCACTTCCCATTGCTC	194
TC248918	MYB TF	GCAGCTGAAGTCCCTGGTAG	TAGAAGCAGCCATGGGAAAG	201
ES819678	LOB Domain Factor	GGGTCTTCTTCATCCCAACA	TGTGAGCCTTGGGATTTACC	230
ES820944	BLH TF	GTTCCTTACACAGCCCTTGC	ACCTCGGTGTCTCTCCCTTT	153
ES838507	BLH TF	GTAGTAACGGCGATGGTGGT	AATGCCTTGACATTGCCTTC	243
TC221650	CUC Gene	CGAACCACGTTCCACTTCTT	TCTCTAAGCCTTGCCTCAGC	192
TC186920	TCP TF	CCAGTTGACCGAACAATGTG	CCGTCGGCTTCTACTTCAAC	160
TC194705	TCP TF	CTGAAACCTCCCAATCTCCA	TTGTTGTCAAGGACGAGCAG	230
TC197998	TCP TF	GTAGAACTGGATGGCGGTGT	AAAGCATAGTGGAGGCGAGA	150
N/A	Glyceraldehyde-3-phosphate	ATCAAGGGCACCATGACTACCACT	ACCAGTTGAAGTCGGGACGATGTT	451