Chromosome Substitution Lines: Concept, Development and Utilization in the Genetic Improvement of Upland Cotton

Sukumar Saha¹ et al.*

¹United States Department of Agriculture-Agriculture Research Service, Crop Science Research Laboratory, USA

1. Introduction

Cotton is the most important natural fiber source for the textile industry world-wide. It is also an alternative of the man-made petroleum-based “synthetic fibers” providing an advantage for a sustainable environment. Cotton is formed by developing seed of several Gossypium species, which are mainly grown as an important cash crop in more than 70 countries including USA, India, China and Uzbekistan (Smith and Coyle, 1997). Although cotton plants are best known as the renewable source of textile materials for clothing, the fiber, seed and plants have many other uses, including home insulation to save energy, protein-rich seed-derived feed for animals, cottonseed oil as a foodstuff for humans, and as a source of mulch and biomass (Cotton Incorporated, 2010, http://cottontoday.cottoninc.com/sustainability-about/responsible-economic-development, verified on October 14, 2011). This brings significant humanitarian and economic benefits. For example, scientists are exploring genetic means to better harness its highly nutritious seed for food and feed (Sunilkumar et al., 2006).

Cotton is facing some serious challenges in production and marketing, such as competition from synthetic fibers, large year-to-year variability in yield and fiber qualities, and recent changes in textile technologies, for which the optimization of cotton requires altered fiber quality characteristics. Genetic solutions to these and other challenges require adequate genetic variation to be present in the breeding germplasm. In cotton, however, the genetic diversity available in the breeding gene pool is narrow, and is recognized as a cause of yield stagnation, declining fiber quality, and increasing genetic vulnerability to biotic and abiotic stresses in worldwide cotton production.

* David M. Stelly², Dwaine A. Raska², Jixiang Wu³, Johnie N. Jenkins¹, Jack C. McCarty¹, Abdusalom Makamov⁴, V. Gotmare⁵, Ibrokhim Y. Abdurakhmonov⁴ and B.T. Campbell⁶

¹United States Department of Agriculture-Agriculture Research Service, Crop Science Research Laboratory, USA
²Department of Soil and Crop Sciences, Texas A&M University, USA
³Plant Science Department, South Dakota State University, USA
⁴Center of Genomic Technologies, Institute of Genetics and Plant Experimental Biology, Academy of Sciences of Uzbekistan, Uzbekistan
⁵Division of Crop Improvement, Central Institute for Cotton Research, India
⁶USDA-ARS Coastal Plains Soil, Water, and Plant Research Center, USA
Cotton breeding programs were extremely successful over several decades by increasing the frequency of beneficial alleles for important traits at many loci, which generated high yielding lines with superior agronomic qualities. As a consequence, breeders continued making crosses among closely related high-yielding cultivars. In fact, most Upland cotton breeding programs rely primarily on crosses among closely related elite domesticated genotypes with high yield and superior fiber qualities, and on reselection from existing cultivars (Esbroeck and Bowman, 1998).

The rate of change of USA cotton yield has steadily declined since 1985. The absolute cotton yield reached at a disturbing rate (3.3% annual rate) by 1998, demanding the immediate need for genetic improvement (Paterson et al., 2004). The history on the genetic improvement of Upland cotton suggested that some of the favorable alleles were fixed in the elite breeding gene pool and that significant additional improvements would require introgressing genes from outside of the pool (Bowman et al., 1996). Exotic unadapted species have contributed beneficial alleles for improving agronomically valuable traits in many crop species (Tanksley and McCough, 1997). Although it is well known that exotic germplasm contains potentially valuable genes, the exotic gene pools of *Gossypium hirsutum* L. have been under-characterized and under utilized.

There are several impediments in conventional methods of interspecific introgression in cotton: 1) complex antagonistic relationships among important traits; 2) cytogenetic differences among the species due to different ploidy levels, meiotic affinity and chromosomal structural differences including translocations and inversions; 3) “linkage drag effects” leading to poor agronomic qualities; 4) reduced recombination; 4) loss of alien genetic materials in early generations; 5) sterility in the hybrids; 6) complex genetic interactions such as Muller-Dobzhansky complexes and 7) distorted segregation (Endrizzi et al., 1985). These kinds of difficulties are most severe in crosses of *G. hirsutum* with diploid species, but most of them also apply significantly to the crosses of the "primary gene pool" which includes all of the other natural tetraploid (2n=52) cotton species *G. barbadense*, *G. tomentosum*, *G. mustelinum* and *G. darwinii* along with *G. hirsutum*.

As part of the primary gene pool, these four tetraploid species are especially accessible reservoirs of important genes for pest and disease resistance, and for improved agronomic and fiber traits. However, the effects of beneficial genes from wild unadapted germplasm are often obscured by other genes that affect the trait negatively in the wild species. Only upon appropriate genetic analysis, the presence of such valuable genes can be detected. Though single genes can significantly affect individual traits, they typically involve complex direct and indirect molecular interactions with other genes. Thus, genes often affect multiple traits, and traits are generally determined by multiple genes. Furthermore, the potential utility and value of specific alien genes is usually compromised by co-inheritance of closely linked genes that have deleterious agricultural effects on productivity. Rates of recombination with alien chromosomes can be reduced considerably, especially in certain segments, therein greatly raising the likelihood of undesirable linkages. To physically separate the beneficial and undesirable alien genes by breeding can thus be extremely difficult. Several generations of crosses with special breeding strategies are needed to eliminate such undesirable traits and fix the selected few desirable traits in interspecific introgression to improve cultivars.
The success of introgressing useful genetic variation from unadapted species into Upland cotton depends on several factors: 1) the breadth of diversity that can be accessed; 2) the speed and efficacy with which the useful alleles can be transferred, given various biological constraints; and 3) whether useful variation can be transferred without the deleterious traits. An alternative approach to conventional pedigreed or population-based interspecific introgression is to use alien chromosome substitution, which entails a modified pedigreed backcross breeding approach (Stelly et al., 2005). A similar and potentially complementary approach is marker-assisted chromosome segment substitution, as first demonstrated in tomato "introgression lines (ILs)". Categorically, these methods trace back to the classical study on quantitative trait analysis of wheat "backcross inbed lines (BILs)" by the famous quantitative geneticist/breeder Robert Allard (Wehrhahn and Allard, 1965).

All five of the 52-chromosome allotetraploid *Gossypium* species, including the four species most amenable to interspecific introgression into Upland cotton (common name for the widely grown middle-staple *G. hirsutum* cultivars), are suggested to have arisen from a common ancient polyploidization event (Wendell, 1989). Cytological observations of hybrids and comparative molecular mapping indicate that synteny and colinearity are mostly conserved among the five tetraploid species, and that there are few major cytost ructural barriers to interspecific introgression among them. However, genetic limitations are manifested by the significant level of "$F_2$ breakdown" that is commonly observed after hybridization of Upland cottons with the other species due to extensive genetic incompatibility between these species at the whole-genome level (Beasley and Brown 1942; Reinisch et al. 1994). As Allard's work on wheat would suggest, backcross-inbreeding can be used to circumvent some of these issues. We have begun to exploit a similar approach by chromosome and chromosome segment substitution (CS) lines, using a modified form of backcross-inbred line development and subsequent quantitative genetic evaluations.

We released a set of 17 disomic alien chromosome substitution *G. barbadense* (CS-B) lines through hypoaneuploid-based backcrossing in a near-isogenic genetic background of Texas Marker-I (TM-1) line (Stelly et al., 2005). In a series of complementary quantitative analyses, the chromosomal effects on agronomic and fiber properties were documented using these CS-B lines in our previous studies (Saha et al., 2004, 2006, 2008, 2010, 2011a, Jenkins et al., 2006, 2007). We also showed that the chromosomal substitution lines constitute important breeding resources, increasing the genetic diversity available in Upland cotton (Jenkins et al., 2006; 2007). The objective of this paper is to provide a summarized report on the concept, development and utilization of CS lines from our previous studies in genetic analysis and germplasm improvement of Upland cotton.

### 2. Cytogenetic resources of cotton

Cotton was one of the first crop plants to which Mendelian principles of genetics were applied (Balls, 1906; Shoemaker, 1908). Cytogenetic analyses revealed that the cultivated species *G. hirsutum* and *G. barbadense* and several other New World allotetraploids contained 52 chromosome, whereas most other diploid species contained just 26 chromosomes. Furthermore, they indicated that all of 52-chromosome species were allotetraploids with two pairs of meiotically independent genomes, AADD. In that, the A and D components are very similar to the A and D diploid genomes of certain extant 26-chromosome species that occur naturally in Africa and the New World (Endrizzi, 1984, 1985; Wendel 1989).
Reciprocal translocations were developed for *G. hirsutum* in order to identify and meiotically detect each of the 26 different chromosomes in translocation heterozygotes, which allowed assignment of the individual chromosomes to the A and D subgenomes. The chromosomes of A genome and D genome of *G. hirsutum* were designated as chromosomes 1-13, and 14-26, respectively (Menzel and Brown, 1978; Brown, 1980; Brown et al., 1981).

Representative types of the original primary monosomic, monotelodisomic and tertiary monosomic reciprocal translocation (NTN) have been or are being backcrossed to TM-1, a highly inbred line from Deltapine 14 that serves as a standard reference for genetic and cytogenetic research (Fig. 1; Kohel et al., 1970). Stocks in the current Cotton Cytogenetics Collection are thus nearly isogenic to TM-1 and each other (Fig. 2). Monosomic stocks have

Fig. 1. Cytological identification of the aneuploid lines. (A) Metaphase showing 26II pairs of chromosomes in a normal plant (2n); (B) Metaphase stage from a monosomic plant (2n-1); (C) Metaphase stage from a monotelodisomic plant (2n-1/2) and (D) metaphase from a heterozygous translocation plant. The arrow shows the monosomic and telosomic chromosome and translocated chromosomes respectively.
been used extensively to locate and develop markers for specific chromosomes (Endrizzi, 1963; Kohel, 1978; White and Endrizzi, 1965; Guo et al., 2008; Gutierrez et al., 2009). Monosomic stocks have also been used to identify the chromosomes of translocations (Endrizzi, 1985). Monotelodisomic plants have been used to locate markers to a specific arm of a chromosome (Endrizzi and Kohel, 1966; Endrizzi and Taylor, 1968; Endrizzi and Bray, 1980) and to map the distance between a centromere and a marker locus (Endrizzi et al., 1985). The Cotton Cytogenetic Collection currently includes chromosome-deficient stocks of *G. hirsutum*, mainly primary monosomics (missing one chromosome) for 15 of the 26 chromosomes, and monotelodisomics (missing most of one chromosome arm) for 31 of the 52 chromosome arms, including at least four chromosomes not represented among the monosomic stocks. To achieve additional cytogenetic coverage of the genome, new monosomics and monotelodisomics have been identified, and tertiary monosomics have been synthesized from intercrosses between monosomics and related translocations, and repeated backcrossing to the tertiary monosomics. These chromosome-deficient stocks of *G. hirsutum* have been used as recurrent parents in the development of CS lines.

Through a Cooperative (Co-op) Agreement between TAMU and USDA-ARS, Mississippi (MS), all core monosomic and monotelodisomic cotton plants are developed at Texas A&M University (College Station, Texas), where they become part of the Cotton Cytogenetics Collection; the Collection and most developmental stocks involving *G. barbadense* 3-79 chromatin substitution are duplicated in the greenhouse at USDA/ARS (Mississippi State, MS). This Co-op Agreement enables a safety backup to the Collection, provides some resources for maintenance and development of the Collection, places important research resources at USDA/MS, and fosters multi-dimensional collaborations. These increase capability in use of the cytogenetic stocks for basic genetics, genomics and applied breeding studies.

### 3. Development of chromosome substitution lines

The detailed method of chromosome substitution line development was discussed by Stelly et al. (2005). Each alien species chromosome substitution line development involves four specific stages: (1) development of the respective TM-1-like hypoaneuploid stock, (2) use of the cytogenetic stock as a recurrent seed parent in a recurrent backcrossing program to create a monosomic or monotelodisomic F$_1$ substitution stock, followed by (3) inbreeding with the TM-1-like hypoaneuploid stock to recover a euploid disomic substitution line and (4) confirmation of the cytogenetic and genetic constitution of the disomic lines by cytological analysis and chromosome-specific SSR markers (Figs. 1, 2, 3, and 4).

The strategy in developing these lines follows principles of cytogenetic behavior, transmission and inheritance as reported by Endrizzi et al. (1985). The strategy is based on the principle of differential transmission rate between mega- versus micro-gametophytes, i.e., between the “ovule” (“seed” or “female” parent) versus the pollen (“male”) parent. It is believed that transmission of hypoaneuploidy through the ovule parent is common (up to 50%) for most chromosomes or chromosome arms in monosomic and monotelodisomic conditions, but that transmission through cotton pollen is rare to nil for all whole-chromosomes and most large-segment deletions (e.g., telosomes).
In each hypoaneuploid F₁ stock, the targeted alien chromosome or segment is hemizygous and thus lacks a homologous partner to undergo meiotic pairing and recombination. By avoiding recombination, the alien chromosomes or segments remain intact when they are transmitted to the gametes and next generation (Endrizzi et al. 1963, 1984, 1985). Even marker-assisted selection based on large number of markers could not guarantee transfer of all alien loci, especially in high recombination regions, which tend to be gene rich (Zhang et al., 2008; Lacape et al., 2009). This is the clear difference between backcross derived materials in Upland cotton from G. barbadense versus interspecific chromosome substitution lines. Previous studies showed that the actual number of recombinants obtained were significantly fewer than the expected number in an interspecific backcross program of Upland cotton (Rhyne, 1958) and also limited genetic transmission in advanced generation of the interspecific hybrids (Lacape et al., 2009).

Chromosome substitution lines were developed as follows. The initial cross was made between a chromosomally normal (euploid) G. barbadense 3-79 line, as male, and a cytogenetic derivative of G. hirsutum TM-1 that was deficient for one copy of a specific chromosome (“monosomic” plant) or chromosome arm (“monotelodisomic” plant), as female. Due to their chromosomal constitution, primary monosomic, monotelodisomic, tertiary monosomic and euploid plants were isolated from diagnostic metaphase I meiotic configurations (25 II + I; 25 II + Ii; 24 II + III; and 26 II, respectively) in microsporocytes (“pollen mother cells”, Fig. 1). Interspecific F₁ progeny were screened phenotypically, cytogenetically (metaphase-I chromosome analysis of microsporocytes) and in some
instances by a molecular marker analysis [loss of heterozygosity (LOH)], to identify a plant deficient for the chromosome or chromosome arm. The identified interspecific aneuploid (BC₀F₁) plant was subsequently used as male parent in backcrosses onto the recurrent aneuploid TM-1 plant. As in the preceding (BC₀F₁) generation, an aneuploid BC₁F₁ progeny was selected based on metaphase I analysis, as in the BC₀F₁ generation. This procedure was repeated until the fifth backcross. The selected BC₅F₁ hypyoaneuploid was selfed to recover a euploid plant (BC₅F₁S₁ 26II). Euploids were again selected phenotypically, cytogenetically and, in some cases, by a molecular marker analysis. Finally, one euploid BC₅F₁ plant was selfed to establish each euploid BC₅S₁ substitution line. In each CS-B line, a pair of chromosomes (or chromosome arm segments) of *G. hirsutum* inbred TM-1 was replaced by the respective pair from *G. barbadense* doubled-haploid line 3-79 lines.

<table>
<thead>
<tr>
<th>Normal <em>G. hirsutum</em></th>
<th>Monosomic <em>G. hirsutum</em> (e.g. H₄)</th>
<th>Donor species (e.g. <em>G. barbadense</em>)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 2 3 4</td>
<td>AA</td>
<td>AA</td>
</tr>
<tr>
<td>5 6 7 8 9 10 11 12 13</td>
<td>14 15 16 17 18 19 20 21 22 23 24 25 26</td>
<td>AA</td>
</tr>
</tbody>
</table>

Fig. 3. The overall strategy on the development of a chromosome substitution line. The diagrammatic picture showed chromosome of two species in different colors.

As mentioned above, seventeen CS-B lines were developed using the above procedure and released for use in Upland cotton improvement (Stelly et al., 2005). We designated the line with the number of the specific substituted chromosome or chromosome segment (sh=short arm, Lo=long arm). We also designated NTN to the lines derived from the cross with translocation stocks carrying two segments of the substituted chromosomes from the donor alien species. We dubbed the 3-79 chromosome substitution lines as "CS-B" lines, e.g. CS-B01 for chromosome-1, to reflect the species of origin, i.e., *G. barbadense*. Future series involving other species will be named analogously, e.g., "CS-T" for *G. tomentosum* chromosomes or...
chromosome segment(s), and CS-M for *G. mustelinum*. Additional CS-B, -T and -M lines are in development; several of the most advanced are now being investigated with regard to effects on fiber and agronomic traits.

We could not confirm the genetic identity of all of the CS lines based on molecular markers at the release time because very few chromosome specific molecular markers were available at that time in the public domain. Recently, in addition to the cytological analysis, we undertook an assignment to confirm the genetic identity of the CS lines using chromosome specific SSR markers (Fig. 4). We used a slightly modified protocol of our previous studies (Guo et al., 2008; Gutierrez et al., 2009). It is expected that a CS line will have the allele of the donor alien species specific to the substituted chromosome or chromosome segment and will miss the TM-1, the recurrent parent, allele specific to the locus of the substituted chromosome or chromosome segment in molecular results.

Most marker results were concordant, but several were inconclusive or discordant (Fig. 4). For the latter, we are investigating further to discern the underlying basis, which is likely an infrequent type(s) of event, and could be a procedural and/or biological in nature. It is interesting that most of these CS lines had different agronomic and fiber properties than TM-1 suggesting that these lines carried genetic materials from some other sources than TM-1, possibly the donor parent, *G. barbadense* (Saha et al., 2004, 2006, 2008, 2010 and 2011a).

Fig. 4. Electrophoregram results by ABI3130 xl capillary electrophoresis showing CS-B08Lo line has only the same allele of 3-79 for BNL3792 SSR marker specific to the long arm of chromosome eight, but missing the allele of TM-1. Accordingly the molecular results confirmed that CS-B08Lo line has the substituted the long arm of chromosome eight from 3-79.
4. Development of different genetic resources using the CS lines

4.1 Chromosome substitution lines

In each CS-B line, a pair of chromosomes (or chromosome arms) of *G. hirsutum* inbred TM-1 was replaced by the respective pair from *G. barbadense* doubled-haploid line 3-79 (Fig. 5). We also developed several chromosome substitution lines from *G. tomentosum* species (CS-T) and for some new chromosomes of *G. barbadense*. Currently, we are evaluating these lines for fiber and agronomic traits. These substitution lines are nearly isogenic to the common parent TM-1 for 25 chromosome pairs, as well as to each other, for 24 chromosome pairs. The comparative analysis of such unique genetic materials greatly empowers the detection of genetic effects of novel alleles by specific alien chromosomes associated with quantitative traits.

![Chromosome substitution line for chromosome 25, e.g. CS-B25](image)

Fig. 5. A diagrammatic picture showing different chromosomes in a CS-B line genome.

4.2 Chromosome specific recombinant inbred line development (CS-RILs)

The CS-RILs were developed from crosses of a CS-B line (e.g. CS-B17) of *G. barbadense* chromosome substitution lines with their common recurrent parent, inbred *G. hirsutum* TM-1 (Fig. 6, Saha et al., 2011b). An individual plant of the F$_2$ population from this cross was maintained by selfing via single seed decent method until F$_6$ generation. It is expected that the CS-RIL population gets stabilized in a near homozygous condition by the F$_6$ generation. The CS-RIL populations are currently being evaluated in field trials for fiber and agronomic traits. This CS-RIL population will be very useful for high-resolution mapping of QTLs. Using molecular markers in such a RIL will facilitate detection and cloning of QTLs.

4.3 Crosses among the CS lines to create different chromosome specific genetic resources

The near-isogenic nature of the substitution lines to the common parent TM-1 provides a unique opportunity to use specific mating designs among the CS lines of interest to create different chromosome specific combinations of genetic resources (Fig. 7). Using a mating design within the same alien species of two CS lines of interest or between two alien species for the same substituted chromosome or different substituted chromosome, we will be able to create a unique source of genetic resources useful as genomic tools for germplasm improvement. For example, a cross between the same chromosome substitution lines of two
different species will provide the opportunity to create a tri-species F\textsubscript{1} plant where the substituted chromosomes will be a heterozygous chromosomes from two donor alien species, while the other 25 chromosome pairs will be from the recurrent species of \textit{G. hirsutum}. Selfing such a F\textsubscript{1} plant will create the F\textsubscript{2} progenies segregating primarily for the genes located on the substituted heterozygous chromosomes from two donor species. Such a RIL will help to unveil epistasis, a difficult gene interaction to measure, due to interactions of the alleles from the two donor alien species with the alleles from other chromosomes of Upland cotton. Also, the crosses between two different substitution lines from the same donor alien species can provide unique opportunity to combine beneficial alleles from two alien substituted chromosomes of our interest (Fig. 7). Such genetic materials can be used as valuable tools for high-resolution genetic mapping and targeted chromosome specific introgression of valuable traits from wild and unadapted species.

5. Discovery of alien chromosomal effects on important traits using the CS lines

We have used several types of family structure, experimental design and statistical analysis to discover the alien chromosomal effects on important traits using the CS lines (Saha et al., 2004, 2006, 2008, 2010, 2011a; Jenkins et al., 2006, 2007; McCarty et al., 2006; Wu et al., 2006). Studies have enabled the discovery of some important chromosomal effects on agronomic and fiber quality traits, and point toward additional opportunities, as more CS lines are developed.

Each substitution line is nearly isogenic to the common parent TM-1 for 25 chromosome pairs; pairs of non-homologous CS lines are nearly isogenic to each other for 24 chromosome pairs. Given \( n=26 \), each CS contains about 4\% of the \textit{G. barbadense} genome in a common \textit{G. hirsutum} background (~96\%). The near-isogenic nature of CS lines with TM-1 and each other renders them extremely amenable to dissection of quantitative traits of interest because they
largely eliminate genetic "noise" from 96% of the genome. This greatly facilitates the detection of genetic effects of novel alleles by specific alien chromosomes on quantitative traits. We demonstrated that interspecific chromosome substitution is among the most powerful means of introgression and steps toward identification of chromosomal association with a QTL in cotton.
As a matter of thoroughness, let us note that many of the *G. hirsutum* aneuploids used to develop CS lines were first discovered in non-TM-1 lines; they were backcrossed multiple times to TM-1 to create TM-1-like hypoaneuploids, which served as recurrent parents for CS line development. Through repeatedly backcrossing, it is nonetheless likely that small residual pieces of the original aneuploid remain in the recurrent parents with TM-1 background used to create the CS lines. These could conceivably affect chromosomal associations of some traits. Similarly, note that residual genetic contributions from the donor, e.g., *G. barbadense* '3-79' could also affect CS lines, because 5-6 backcrosses is unlikely to eliminate 100% of alien germplasm in the non-substituted portion (~96%) of the genome. Thus, there is also a possibility that the observed genetic effects could have been due to some unlinked residual effect of the donor (3-79) genome, i.e., independent of the substituted chromosome or chromosome arm (Saha et al., 2006). For statistical quantitative genetic calculations, however, we have considered that the differences in any trait among the lines exclusively due to the contribution of the alien species substituted chromosome on the assumption of isogenic nature of non-substituted chromosomes.

5.1 Direct comparison of the CS lines

The comparative analysis of CS-B lines in a uniform genetic background has provided an opportunity to detect net genetic effects of agronomic and fiber quality traits from all genes on the specific substituted chromosome or chromosome arms (Saha et al., 2004). Empirically, one can predict that single cotton chromosome contain an average of 1000-4000 genes. The mode of development of these CS-B lines uses hemizygosity to preclude recombination during introgression, so all genes within the alien chromosome or chromosome segment are transferred into Upland cotton. The CS-B line evaluation for QTL localization does not require segregating populations, so it differs from traditional QTL mapping, and offers certain statistical advantages.

A mixed linear model of genotype with genotype by environment (G × E) interaction model was used in the analysis of results (Saha et al., 2004). The results showed that the genotypic effects greatly exceeded G × E interaction effects for all quantitative traits. The residual variance accounted for less than 20% contribution to the phenotypic variance for all traits except elongation, suggesting that the genotypic effects or G × E interaction effects were readily detected. CS-B25 had significantly lower micronaire than TM-1 (Fig 8). The detailed results of this research were presented in Saha et al. (2004). Seven CS-B lines (2, 6, 16, 18, 5sh, 22Lo, 22sh) had greater lint percentage than 3-79 or TM-1. Micronaire for substituted chromosomes 17 and 25 was lower than TM-1 (Fig. 8). Substituted chromosome 17 line had a significantly greater fiber elongation than either TM-1 or 3-79. Fibers of substituted chromosomes 2 and 25 were significantly stronger than TM-1. The results showed that backcrossed chromosome substitution lines had both positive and negative net effects on various fiber traits. Moreover, they clearly showed the potential of some of the backcrossed chromosome substitution lines from comparative analysis for improving Upland cotton germplasm.

5.2 Comparison of CS lines and their F2 hybrids

The quantitative genetic analysis of CS lines can be extended by collective analysis of the lines and various hybrid generations, which allows partitioning of source effects. We initiated
such analyses by characterizing the chromosomal association of important agronomic and fiber traits by comparative analysis of CS-B lines in a TM-1 background along with their parental lines and their F2 hybrids (Saha et al., 2006; Saha et al., 2008). Given the isogenic nature of the whole-chromosome CS-B lines, the expected relative genetic complexity for single-locus effects would be approximately inversely proportional to the haploid chromosome number, 1/n. For multilocus interactions, the reductions would be geometric and thus much more extreme. In CS-B F2’s, segregation is largely to completely eliminated for 25 of the 26 chromosomes (about 96% of genome), rendering CS-B F2 analyses relatively free of the extensive numbers and kinds of interchromosomal interactions that arise in a conventional interspecific F2 population. The CS-B F2’s thus provided an opportunity to discern effects of alleles in homozygous versus heterozygous conditions, on a chromosome-by-chromosome or arm-by-arm basis. Multiple comparisons such as this can be used to determine if a substituted chromosome(s) or a chromosome arm(s) is associated with a quantitative trait of interest.

The analyses were facilitated by use of an additive dominance (AD) genetic model with G × E interaction for genetic analysis of the parental lines and their F2 hybrids (Zhu 1994, Tang et al. 1996). Whereas additive and additive epistatic effects were confounded in the comparison of homozygous CS-B and parental lines in Saha et al. (2004), application of the AD model to F2 generations enabled dissection of genetic effects into additive and dominance effects. In this model, additive genetic effect reveal an estimate of general combining ability due to the presence of specific alien chromosome or chromosome arm of 3-79 in TM-1 background. Thus, evaluating the additive effects may help cotton researchers choose an appropriate CS-B line as a source of good general combining ability in cultivar development. Heterozygous dominance effects can be considered as an estimate of specific combining ability (SCA) of parents in specific hybrid combinations. The dominance effects include homozygous and heterozygous effects which are related to inbreeding depression and heterosis (Jenkins et al., 2006).
The genetic background of the CS lines and the comparative analysis among the lines and their hybrids provided an opportunity to dissect the effect of each chromosome or chromosome segment under four different conditions: (i) the homozygous condition in the TM-1 genetic background, totally lacking the alien 3-79 chromosome or chromosome segment, ii) the homozygous condition of euploid CS lines in the TM-1 genetic background, iii) the average heterozygous condition in an F2 generation segregating for the TM-1 and 3-79 alleles in the specific substituted chromosome or chromosome arm and iv) the homozygous condition in the 3-79 genetic background. Additive and dominance genetic effects were detected when F2 hybrids and parents were evaluated (Saha et al., 2006, 2008). Additive effects were predicted based on the genetic effects of CS lines in TM-1 genetic background, TM-1 and 3-79 lines. On the other hand, dominance effects were predicted based on the results of the homozygous condition (euploid CS lines) in the TM-1 genetic background, TM-1, and 3-79 and heterozygous condition in the hybrids between CS lines and TM-1 (Saha et al., 2006; Saha et al., 2008). For example, results revealed that the alien chromosome 2 and 25 had significant net positive additive effects on fiber strength compared to TM-1, suggesting that these chromosomes carry genes for improving fiber strength and the alien chromosomes 17 and 25, respectively, had negative additive and homozygous dominance effects on micronaire, indicating that these chromosomes carry genes potentially useful for manipulation of micronaire (Saha et al., 2006, 2008).

5.3 Comparison of parental lines with their F2 and F3 hybrids

Six CS-B lines (CS-B14sh, CS-B16, CS-B17, CS-B22sh, CS-B22Lo, and CS-B25) and TM-1 (the recurrent parent) were crossed in a half diallele mating design and agronomic and fiber properties of the parental lines were compared with their F2 and F3 hybrids (Saha et al., 2010, 2011a). By applying the additive-dominance model, we were able to partition genetic effects into additive, dominance, additive and additive interaction effects for each of the substituted chromosome or chromosome arm (Saha et al., 2010; Saha et al., 2011a). An extended AD model including additive-by-additive epistatic effects (ADAA model) was used to estimate variance components and to predict genetic effects (Wu et al. 2006). The additive-dominance model provided means to dissect further the genetic effects into additive, dominance, additive and additive interaction effects for each of the substituted chromosome or chromosome arm in these studies (Saha et al., 2010, 2011a). We also used the method of Patterson (1939) to test the significance of the difference between the genetic effects for two lines based on the standard error of the difference between two effects. We used a one-tailed t-test to estimate the significance of variance components, and a two-tailed t-test for estimating the genetic effects. A significant difference in genetic effects between a specific CS-B line and TM-1 was considered a genetic effect attributable to the specific substituted chromosome or chromosome arm. Assuming uniform genetic background of the CS-B lines, the comparative analysis of the double-heterozygote combinations (CS-B × CS-B F1) versus their respective single heterozygotes from the cross of the respective CS-B line and TM-1 with their F1, F2 and F3 progenies revealed that epistatic effects between the genes in the chromosomes strongly affected most of the fiber quality and agronomic traits.

A summary of important genetic effects is presented in this section from these studies (Saha et al., 2010 and 2011a). CS-B14sh and CS-B25 had the highest Upper half means (UHM), and both were significantly longer than TM-1. The average UHM was higher in the hybrids of
CS-B17 × CS-B22sh and CS-B22sh × CS-B22Lo than their respective parents, which seems to indicate dominance effects of the alien allele(s) caused the hybrid vigor. Although fiber strength of the donor, 3-79, greatly exceeds that of the recurrent parent, TM-1, we observed that only one of the six substitution lines, CS-B14sh, had fiber significantly stronger than TM-1. The micronaire values of CS-B17 and CS-B25 were significantly lower than all other CS-B lines and TM-1. The F₁, F₂ and F₃ hybrids between these two lines, CS-B17 × CS-B25, also had exceptionally low micronaire relative to the hybrids of other CS-B lines, revealing the respective 3-79 chromosomes carry genes that can reduce micronaire, a positive breeding value. We observed positive additive effects greater than TM-1 for CS-B14sh and CS-B25 for fiber length, strength, and uniformity. Almost all (≥98%) CS-B lines and their crosses had additive-by-additive epistatic effects on fiber quality traits. Additive-by-additive interaction effects for uniformity ratio were higher for most of the CS-B17 hybrids and lower for CS-B22sh hybrids than their respective CS-B parents, suggesting opposite epistatic interaction between the alleles in homozygous versus heterozygous condition for this trait. CS-B16 and CS-B17 had negative additive-by-additive interaction effects for fiber strength when carrying the only alien chromosome in homozygous condition or heterozygous condition for the respective single alien chromosome, however, when the hybrid between CS-B16 and CS-B17 carrying both of the alien chromosomes in heterozygous condition showed high positive additive-by-additive interaction effect for fiber strength implying that the epistatic interaction between the alleles in heterozygous condition of the two substituted chromosomes caused the increased genetic effect in fiber strength.

CS-B16, CS-B22sh, and CS-B22Lo lines had higher additive genetic effects on lint percentage compared to TM-1 (Saha et al., 2010). All of the lines, except CS-B17, had opposite dominance effects on seedcotton yield in heterozygous versus homozygous conditions specific to the substituted chromosomes. CS-B14sh, CS-B22Lo and CS-B25 lines had higher dominance effects on both seed cotton and lint yield under homozygous condition compared to TM-1. The majority of the hybrids had positive additive and additive epistatic effects. Hybrids of CS-B16 x CS-B22sh had the highest additive and additive epistatic effect on boll weight among all others thus the epistatic interaction of the genes located on these two substituted chromosome likely responsible for this effect. We documented in this research that epistasis made a substantial contribution to each of the complex quantitative trait loci (QTLs) showing the effect of masking of alleles at one gene locus by an allele at another locus under homozygous and heterozygous condition for the same specific chromosome on complex agronomic traits. We detected for the first time in this research many cryptic alleles located on several 3-79 substituted chromosomes or chromosome arms of 3-79 (G. barbadense), commonly associated with poor agronomic qualities including yield, had the potential to improve agronomic traits including seed cotton and lint yield in TM-1 (G. hirsutum).

5.4 Evaluation of chromosome specific RILs to study genetic effects

Two different CS-RILs populations were developed from crosses of CS-B05sh and CS-B17 G. barbadense chromosome substitution lines with their common recurrent parent, inbred G. hirsutum TM-1 (Saha et al., 2011b). Each population included 50 CS-RILs, which were used in field trials with commercial varieties DP393 and PHY370 WR in four environments at Mississippi State University, MS. A randomized complete block design with four
replications was used in each environment. Four agronomic and five fiber traits were evaluated for these two CS-RIL populations. Mean comparisons among different lines were used subject to ANOVA analysis with least significance difference (LSD) at probability level of 0.05. The collective variation in each CS-RIL population was assessed by cluster analysis, using the Mahalanobis distance and Wald method, where all data were standardized with a variance of one and mean of zero. The detail results of this research will soon be presented in a separate paper.

Both CS-RIL populations showed significant genetic diversity for all of the traits being investigated. One RIL (CS-B05sh RIL5) had stronger fiber than DP393 and two RILs (RIL 5 and 18) had stronger fiber than PHY 370 WR. Sixty percent of CS-B17 RILs had lower micronaire than PHY 370 WR and 46% of CS-B17 RILs had lower micronaire than DP393. Boll weights were higher in 54% and 64% of CS-B17 RILs than DP393 and PHY 370WR, respectively. No RILs in the CS-B17 RIL population had higher lint percentage, seed cotton yield, or lint yield than each of two commercial cultivars. Cluster analysis results showed that the two commercial cultivars were in the same group and quiet distinct from CS-B17 RIL population. However, some of the CS-B05sh RILs were close to DP393 suggesting their potential to improve fiber traits. Results showed that some of these RILs have genetic potential to improve some agronomic and fiber traits in Upland cotton. This research provided a scope for additional resolution in genetic mapping and for the targeted use of exotic germplasm to improve fiber quality in a cotton breeding program.

5.5 Top crosses with improved cultivars

In commercial breeding programs, the use of unadapted germplasm has often been restricted to the introgression of traits that have already been identified, are simply inherited and offer a significantly advantageous near-term cost-benefit ratio. However, the breeding value of any new allele for a quantitative trait is dependent on the genetic backgrounds in which it is evaluated. Without introgression, it is difficult to identify the merits of alien genotypes or alleles in applied breeding programs (Jordan et al., 2011). We employed the classic AD genetic model to evaluate the merit of CS-B lines by crossing 13 CS-B lines with five elite cultivars from different cotton breeding companies (Jenkins et al., 2006, 2007). When these lines are crossed to commercial cultivars, these dominance effects included homozygous effects for cultivars and, heterozygous effects between CS lines and cultivars and TM-1 and cultivars respectively (Jenkins et al., 2006, 2007). The CS lines can be extensively utilized as “probes” in the detection of favorable alleles associated with traits of importance in the specific chromosome or chromosomes arm to improve inbred lines (Wu et al., 2010). Our results provided valuable genetic information to help breeders in the improvement of several traits of interest. Over the years different seed companies developed their improved lines by selecting desirable alleles located in different chromosomes. Discoveries of chromosomes with favorable alleles in different chromosomes of improved cultivars will provide the opportunity to combine beneficial alleles in a crossing program by selecting desirable cultivar lines. Thus we provided for the first time a genetic tool to uncover the desirable allele located in different chromosome of improved cultivars for the traits of interest (Jenkins et al., 2006, 2007). For example, chromosomes 2 and 25 from DP90, PSC355, and FM966 carried alleles for increased additive genetic effect in fiber strength. It is most likely that the use of these lines as parents in crossing program will influence fiber strength in cotton breeding program. We observed that both arms 22sh and 22Lo in FM966.
had additive effects associated with increased lint percentage; chromosome 25 and arms 5sh, 14sh, and 22Lo in FM 966 were associated with improved additive effects for boll weight. Chromosome 16 in crosses with SG747 and FM966 was associated with improved additive effects for lint yield. Chromosome 25, and arms 15sh, 22sh, and 22Lo in crosses with FM966 showed with improved additive effects for fiber length. Chromosomes 2 and 25 in FM966 were associated with improved additive effects for fiber strength. Thus, cultivar FM966 has many favorable genetic factors of agronomic and fiber traits associated with different chromosomes or chromosome arms of the CS-B lines. This is especially significant considering that neither TM-1 or 3-79, the parents of CS-B lines, is considered a line with high breeding values and thus not popular among cotton breeders.

Our studies suggested that CS-B lines could be used as tester stocks to reveal different sources of beneficial alleles, thereby, providing a tool to combine beneficial alleles for improving a fiber quality trait. Thus these CS-B lines open a new paradigm in cotton breeding programs to combine the favorable alleles located in different or same chromosomes by crossing the desirable improved cultivars.

5.6 CS lines in the study of developmental genetics

Genes act differentially at different stages of a plant life cycle, where their interactions in a network determine the phenotypes of complex traits. Therefore, discovering genetic changes in time-specific traits at different developmental stages during a growing season has recently become an important issue (McCarty et al., 2006; Wu et al., 2009). Cotton plants grow in an important time-specific manner to produce squares, flowers, bolls, and yields. However, most of the breeding research is based on the data at harvesting period. We used CS lines for time-specific genetic variance components determination using a mixed linear model approach, where phenotypic values observed at time $t$ were conditioned on the events occurred at time $t-1$, revealing the new genetic variations arising at several time intervals during cotton plant growth (McCarty et al., 2006; Wu et al., 2009). In our studies, we used CS-B lines to detect chromosomal associations with the flowering pattern (McCarty et al., 2006), and plant height, number of nodes, and internode length (Wu et al., 2009), major contributors to cotton yield, at different stages in the primary flowering season. We recorded the data every week in the primary growing season. We discovered that the short arm of chromosome 5 of 3-79 in TM-1 background exhibited a positive genetic association with flowering number during this primary flowering time (McCarty et al., 2006). There was no additive genetic effect detected for flower number increase at the initial stage, but additive genetic effects were detected during later stages for flower number increase. On the contrary, large dominance effects were detected for flower number increase for the first two weeks compared to the last two weeks in primary flowering period. We also discovered that the additive effects played a major role for plant height, number of nodes, and internode length at different developmental stages after initial flowering (Wu et al., 2009). We observed that CS-B16 and CS-B14sh had consistently lower additive effects for plant height across different growth stages implying that chromosome 16 and the short arm of chromosome 14 of 3-79 in TM-1 genetic background were associated with shorter plants. On the contrary CS-B26Lo consistently had greater additive effects across all of the growth stages among the CS-B lines suggesting the long arm of chromosome 26 in TM-1 genetic background was associated with taller plants in the primary flowering stages (Wu et al.,
Correlation analysis showed that plant height had significant additive correlations with number of nodes at six different developmental stages suggesting that plant height had common influences with number of nodes during growth after initial flowering in cotton.

6. Conclusion

We have developed, released and partially analyzed several backcrossed *G. barbadense* chromosome or chromosome arm substitution (CS-B) lines. More recently, we also have developed analogous chromosome substitution lines involving *G. tomentosum*, which we have dubbed "CS-T" lines. In current research, some of the CS-T lines are being characterized, while others are still being synthesized or increased. Moreover, we are similarly developing monosomic or monotelodisomic chromosome substitution lines from *G. mustelinum*, which are to be named "CS-M" lines.

The experimental results show that the CS lines are useful from at least four different perspectives: 1) to improve genetic diversity for important traits in Upland cotton, 2) to discover the untapped potential of the novel alleles from the other tetraploid species, 3) to understand the ramifications of epistasis in complex agronomic and fiber traits and 4) to identify chromosomal locations of important fiber and agronomic traits. Therefore, CS lines have been recommended for many cotton breeding programs worldwide and are being widely used in the cotton research community and by breeders. For instance, application of CS-B lines in improvement of commercialized cultivars of Uzbekistan is in progress within the frame of international collaborative projects.

We presented here a brief review on non-traditional method of interspecific germplasm introgression via the use of alien chromosome substitution lines in cotton. We showed that chromosome substitution line can be a useful tool for additional resolution in genetic mapping and for the targeted exploitation of exotic germplasm to improve fiber quality and agronomic traits in a cotton breeding program.

7. Acknowledgements

We thank Ms. Lillie Hendrix and Dr. Russell Hayes for helping in field and Greenhouse research. We acknowledge partial support from the following sources: Texas AgriLife Research, Cotton Inc., Texas State Support Committee, and Texas Dept. Agriculture Food & Fiber Research Grant Program. Joint publication of USDA/ARS, Mississippi Agricultural and Forestry Experiment Station, approved for publication as Journal Article No. J-12079 of the Mississippi Agricultural and Forestry Experiment Station. We thank the Office of International Research Programs (OIRP) of United States Department of Agriculture (USDA) for continual funding of our collaborative research on cotton germplasm characterization. We acknowledge Civilian Research Development Foundation (CRDF), USA for project coordination and Academy of Sciences of Uzbekistan for their continual in-house support of the research efforts.

Disclaimer: Mention of trademark or proprietary product does not constitute a guarantee or warranty of the product by the United States Department of Agriculture and does not imply its approval to the exclusion of other products that may also be suitable.
8. References


Modern plant breeding is considered a discipline originating from the science of genetics. It is a complex subject, involving the use of many interdisciplinary modern sciences and technologies that became art, science and business. Revolutionary developments in plant genetics and genomics and coupling plant "omics" achievements with advances on computer science and informatics, as well as laboratory robotics further resulted in unprecedented developments in modern plant breeding, enriching the traditional breeding practices with precise, fast, efficient and cost-effective breeding tools and approaches. The objective of this Plant Breeding book is to present some of the recent advances of 21st century plant breeding, exemplifying novel views, approaches, research efforts, achievements, challenges and perspectives in breeding of some crop species. The book chapters have presented the latest advances and comprehensive information on selected topics that will enhance the reader's knowledge of contemporary plant breeding.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following: