COTTON IMPROVEMENT'S FUTURE: A COOPERATIVE RESEARCH PROPOSAL

Gerald O. Myers
Department of Agronomy
Louisiana State University
Baton Rouge, LA

Abstract

Cotton (primarily *Gossypium hirsutum* L. and *Gossypium barbadense* L.) is the leading textile fiber in the world and the second most important oilseed. Widely grown in the U.S., it is a significant contributor to the U.S. economy. It is grown in 17 states from Virginia to California and from Missouri to the lower Rio Grande Valley of Texas, covering an estimated 15.7 million acres in 2000. In the U.S., cotton is the fourth most important crop (raw product value 5 to 6 U.S.$ billion/yr; economic impact U.S.$15 to 25 billion/yr), the most important industrial crop (the underpinning of the U.S. textile industry, i.e., several hundred thousand jobs of about U.S.$80 billion/yr), and the second most important oilseed crop. Clearly, productivity and profits from cotton fiber, cottonseed, and their products are important overall to the U.S. economy, but especially to that of rural America. Technological advances have enabled U.S. cotton producers to achieve unparalleled production efficiency. Approximately half of that efficiency is attributable to genetic improvement through breeding. The increases followed rapid producer adoption of new technologies that emerged from cotton research and development at public institutions and in private industry. Production in the U.S. has, however, been recently characterized as increasingly variable and unstable. It is relevant, therefore, to determine the cause or causes of those problems and to propose appropriate solution(s). This proposal presents a coordinated plan of research to be undertaken during the 5 yr period 2001-2006 based upon a consensus of public and private cotton researchers.

Introduction

The importance of genetic improvement have long been undersold relative to other U.S. investments in cotton, particularly price support systems. Whereas price supports are a short-term fix, genetic improvements offer long-term base level improvements for U.S.-grown cottons. Breeding of economically important plants improves the economic return from those plants, e.g., more yield and/or quality for the same or reduced economic input. The biological prerequisites to breeding include genetic diversity, hybridization, genetic segregation, and phenotypic manifestation. Scientific requirements include a diversity of laboratory, greenhouse, field and specialized tests, and testing environments. Breeding programs necessarily integrate a very broad scope of supporting scientific and practical disciplines such as genetics, statistics, agronomy, pathology, entomology, physiology, reproductive biology, taxonomy, and so on. In the case of cotton, breeding results in improved productivity and quality of cotton and cottonseed. Because of its high value as a crop, foodstuff, and base for U.S. textile industries, economic gains from cotton breeding far exceed their cost.

As successful as cotton improvement efforts have been, however, the pace of innovation must accelerate to keep U.S. cotton production globally competitive and to meet the demands of increasing world population. The rate of genetic gain for cotton fiber yield from traditional breeding methods was estimated to be roughly 1.3%/yr (Lewis et al., 2000), i.e., an increase in the U.S. of about $50 million in raw product and $150 million to $200 million in economic impact per year. Unfortunately, investments in cotton breeding do not reflect the magnitude of returns. A recent publication (Report of the American Cotton Producers, 1999) highlights some disturbing trends in the improvement of cotton. Following a period of steady increases in cotton yield from 1970 through the mid 1980s, cotton yield increases leveled off and even become negative for a time. Primary causes of this decline most likely include reduced public staffing and funding for germplasm collection, introgression, genetics and breeding, and the diversion of commercial investments from conventional breeding to backcross breeding programs. Whereas the former elevate base performance, the latter merely add a trait or two by incorporating transgenes into existing elite genotypes.

Because genetic improvement is the most cost-efficient means of improving the profitability of economically important plants and because cotton is so valuable to the U.S., it is of paramount importance that the U.S. capitalizes on contemporary technologies for the improvement of U.S.-grown cottons. One of the rapidly developing, new technologies with potential to greatly accelerate cotton improvement is the broad area of genomics. Genes control all metabolic processes required to sustain life. Cotton can only be crossed with closely related plants. Wild relatives and more distant species are also sources of useful genes, but in many cases, the reproductive specificity of cotton prevent effective utilization of other plant genes. To do this, the U.S. must meld conventional and modern approaches to plant improvement by:

1. Developing public structural and functional genomics resources for cotton that are made fully available to the public researcher in a timely (i.e., immediate) manner:
2. Applying them to field-based and laboratory-based cotton breeding and genetic improvement efforts; and
3. Strengthening those components of its cotton program that address germplasm, its analysis (e.g., markers, traits), usage (e.g., introgression), genetics and breeding.

High-yielding cultivated upland cotton types have long been regarded as being rather uniform genetically and thus relatively vulnerable to potential pathogen or insect epidemics (Bowman et al., 1996). This has been demonstrated repeatedly with a number of technologies, most recently and perhaps most convincingly with isozymes and molecular DNA markers. Without genetic variation, genetic improvement is impossible. Empirically, we can deduce that the biggest constraint in cotton improvement has been the limited amount of genetic variation available for selection in the elite backgrounds present in breeder's nurseries. Increased genetic diversity must be made available to breeders in a usable form. To be incorporated into a varietal breeding program, the genetic diversity must be available in a form that allows expeditious incorporation of the trait, without inadvertent incorporation of undesirable genes. Otherwise, the elite genetic composition of the varietal breeding materials becomes "polluted" with agronomically unacceptable genes ("trash"). In essence, germplasm utilization thus requires extensive "pre-breeding" to generate genotypes that are "elite" or at least "advanced", i.e., free of large numbers of agronomically undesirable genes. The presence of such genes undermines the statistical probability of recovering genotypes of sufficient caliber to compete with existing varieties.

Opportunities to enlarge the scope of breeding activities continues to be extended by in vitro culture and molecular genetic technologies which allow genes to be captured from increasingly distant sources. In general in vitro culture allows for the transfer of larger numbers of genes, with relatively little precision; whereas, molecular genetic technologies allow for the transfer of a relatively small numbers of genes, with greater precision. These new technologies complement each other and more conventional methods of improvement.

There are numerous benefits that would arise from the development of more detailed information regarding cotton genetics and cotton improvement procedures that would allow present and future cotton researchers to meet demands for improved cotton yield and quality. New techniques,

Reprinted from the *Proceedings of the Beltwide Cotton Conference*
Volume 1:393-400 (2001)
National Cotton Council, Memphis TN

particularly those involving molecular genetics, offer additional approaches to genetic dissection and manipulation for genetic improvement and complement more traditional methods of germplasm usage. These methods take advantage of molecular techniques for analysis of genomes (structural genomics), genome expression (functional genomics), indirect selection for desired traits [marker assisted selection (MAS)], as well as gene cloning and genetic engineering The objectives of this proposal will allow for the development of these new techniques in addition to the refinement of existing techniques and the development of a uniform database of information for the extraction and input of research results.

Beyond the benefits afforded to cotton, success in this research will aid in the development of new genetic research tools to expedite and expand the scope of genetic improvement possible in other plants and animals. For cotton, as for these other organisms, there are new opportunities to enhance economic yield, lower production costs, increase plant productivity and reduce pesticide and other chemical inputs.

Once implemented for cotton, the methods will increase the rates of genetic improvement and will expand the scope of targets amenable to breeding and genetic engineering. Of particular importance for breeding are molecular markers and linkage maps for genetic dissection and MAS. Existing markers and maps are grossly inadequate for serious breeding applications in terms of their numbers, kinds, accessibility, and quality. This must be remedied by rapid expansion of effort by researchers who desire to share their products in an immediate and cost-effective manner. For longer term applications in breeding and genetics, comprehensive structural genomic and functional genomic resources are essential.

The research outlined in this proposal also affords additional benefits that positively impact the sustainability of U.S. agricultural research. Chief among these is that researchers involved in this project are primarily responsible for the training of the next generation of public and private breeders. Broad societal benefits relate to how the results of cotton breeding and genetic research into plant resistances to insects, pathogens, and abiotic stresses can reduce the need for pesticides and other chemicals which are expensive, pollute the environment, and require energy for synthesis and distribution.Objectives

Objectives

The overall goal of this new project is to enable U.S. researchers to address the short- and long-term genetic needs of the U.S. cotton industry. To do so, we will augment the five main areas of investigation established in S-258 by developing and assimilating new technologies relevant to germplasm, genetics, genomics, and genetic improvement of cotton. The specific objectives of this proposal are:

1. To acquire, curate, characterize, and evaluate the species, races, and genetic types of Gossypium (cotton).
2. To develop, maintain, and distribute molecular genetic, genetic, and cytogenetic tools for the evaluation and enhancement of cotton germplasm.
3. To adapt and develop methodologies to evaluate, modify, and utilize cotton germplasm.
4. Germplasm enhancement for biotic and abiotic stress resistance and agronomic traits.
5. To refine and develop cotton breeding and variety testing methodologies and techniques.
6. To develop cotton bioinformatic systems.

Procedures

Objective 1: To acquire, curate, characterize, and evaluate the species, races, and genetic types of *Gossypium* (cotton).

The U.S. National Cotton Germplasm Collection is comprised of the germplasm collected as seed during various expeditions, along with materials obtained as donations and exchanges with other germplasm banks around the world. It is maintained by the USDA-ARS at College Station, TX. The Collection is a part of the National Plant Germplasm System (NPGS), and personnel in this ARS unit are responsible for the acquisition, maintenance (including multiplication and preservation), and distribution of those germplasm resources. Data on the materials maintained is accessible through the Germplasm Resources Information Network (GRIN) [http://www.ars-grin.gov/npgs]. Multiple accessions of most of the 49 recognized Gossypium species are maintained as seed.

The collection will be expanded and maintained under five subcollections representing different categories of germplasm. Duplicate seed lots will be stored in the working collection at the Crop Germplasm Research Unit, College Station, TX, and in the base collection at the National Seed Storage Laboratory, Ft. Collins, CO. The collection consists of five subcategories:

1. OBSOLETE CULTIVARS (VARIETIES). These are, with rare exceptions, released cultivars of *G. hirsutum* that are no longer commercially popular, as well as recent releases that are not protected under PVP. Entries in this collection carry an "SA" number designation.
2. LANDRACE COLLECTION. Accessions in this collection originated from wild, feral, and "dooryard" (few or single plants, usually grown for home use) seed of *G. hirsutum*. Of these, 25% have been classified according to geographic-morphological "races" palmeri, richmondi, punctatum, latifolium, marie-galante, morilli, and yucatanense. These are identified with a "TX-" designation.
3. *G. barbadense* COLLECTION. All germplasm of this tetraploid species is maintained with a "GB-" designation and includes cultivars (old and recent), dooryard, and wild accessions.
4. ASIATIC COLLECTION. The cultivated and wild A-genome cotton *G. herbaceum* and *G. arboreum* are maintained in this collection. The former species carries an "A1-" designation while the latter carries an "A2-"
5. WILD SPECIES COLLECTION. Currently, multiple accessions of three wild tetraploid species and 38 wild diploid species are maintained in the collection.

Seed increases will be made at three locations. Photoperiodic stocks will be increased at the Cotton Winter Nursery, Tecoman, Colima, Mexico. Day-neutral stocks will be increased at Rio Farms, Monte Alto, TX. Wild species will be increased in greenhouses at College Station, TX. Distribution of seed will be made from the working collection.

Objective 2: To develop, maintain, and distribute molecular genetic, genetic, and cytogenetic tools for the evaluation and enhancement of cotton germplasm.

Advances in genetic research methodology have made possible the dissection and analysis of plant genomes at the molecular level. An area of research that promises to be important to applied efforts in plant breeding in the near future is in the area of genetic/molecular markers. Molecular markers, like sign posts and road maps, aid plant breeders in incorporating new genes into improved cotton cultivars. Development of additional markers will accelerate this process. Efforts to develop such markers are underway in several research settings (public and private) in the U.S. and worldwide. The main goals of this objective are to develop a universally "agreed upon" and "adhered to" system of nomenclature for these markers and a mechanism to share them among researchers. To do this, the following detailed objectives are proposed:

1. Create and distribute on the web recommendations for a group-approved, standardized nomenclature for cotton molecular markers, ESTs, cDNAs. This nomenclature should include embedded information as to lab source and/or developer.
2. Create a "running list" (user-undatable) of publicly available markers, donor; summary list of donors and respective contributions; interface this with maps to facilitate their usage.
3. Develop a large number of portable markers. Marker types would include RFLPs, RAPDs, AFLPs, SSRs, SNPs, CAPs and any others developed during the period of this proposal.
4. Develop a large number of TM-1 BAC clones to build an integrative genetic and physical mapping infrastructure during the period of this proposal.
5. The identification of "Core" marker sets (highly polymorphic) so that a common interface can be generated among multiple mapping populations.
6. The development and maintenance of mapping populations. This would include:
    a. A highly polymorphic public reference population for mapping low-polymorphism markers.
    b. Other ad hoc public mapping/breeding populations.
7. The refinement and synthesis of existing and new linkage, physical, and cytogenetic maps into a consensus map. Such a map would make use of available marker technology as well as information on QTLs and ESTs. Of further interest is the establishment of a comparative map with *Arabidopsis*, especially with regard to conserved sequences, regions of macro- and micro-synteny, and relative rearrangement. This subobjective includes the maintenance of a web site containing this information.
8. Mutants can provide powerful tools for the investigation of the cotton genome. Research covered in this proposal request would undertake their characterization, maintenance, and distribution and to maximize their utility through isoline development.
9. Utilize wide crosses to access valuable traits in the secondary and tertiary gene pools.
10. Maintain, develop, and utilize cytogenetic stocks. Specific objectives include:
    a. Interspecific F1 hybrids and chromosome-doubled derivatives.
    b. Maintain Cotton Cytogenetic Collection.

Maintain and distribute the Cotton Cytogenetic Collection: Working plant collections will be grown and seed collections will be maintained at College Station in two storage sites, and for long-term storage at Ft. Collins, CO. Seed and plants will be distributed upon request.

1. Mark all *G. hirsutum* chromosomes by hypoaneuploidy.
2. The primary goal will be to find, create, or obtain some sort of deficiency for each chromosome (whole or part), so that all linkage groups can be identified as to chromosome of origin. The secondary goal will be to obtain complete coverage for each chromosome, e.g., through a primary monosome, both telosomes, or the complementing segmental duplication-deficiencies from a single translocation. Presently, some coverage on just 2 to 3 more chromosomes is needed to meet the first objective; coverage remains incomplete on 10 or fewer of the 26 chromosomes. Cytogenetic types providing the desired coverage are being induced by irradiation and generated by two methods of directed chromosome manipulation that yield translocation-derived segmental or whole-chromosome aneuploids), all of which have been effective for other chromosomes or chromosome segments.
3. Produce, maintain, and distribute chromosome and chromatin substitution lines (intact germplasm introgression), their tissue, and/or DNA: The monosomes, telosomes, and critical segmental aneuploids (those uniquely marking a chromosome) will be crossed to produce cytogenetically identified aneuploid F1 hybrids useful for mapping as well as subsequent introgression. After repeated backcrossing and aneuploid recovery, an advanced generation will be inbred for seed increase such that the germplasm can be tested.
4. Evaluate (traits, markers, genes, expression and linkage groups): Backcross derived seed of chromosome-substitution lines (see above) will be used for trait detection and small-scale testing, e.g., for chromosome-specific fiber traits.
5. Use chromatin substitution lines as point of targeted introgression: Lines with advantageous traits will be used as parents in breeding efforts to further localize and identify the genes, markers allowing indirect selection, and introgression of the gene(s) into elite breeding germplasm.

Objective 3: To adapt and develop methodologies to evaluate, modify, and utilize cotton germplasm.

The evaluation and utilization of methodologies for analyses of cotton germplasm has been limited. Previous regional projects (such as S-258) have only scratched the surface of the available developing technologies to characterize and implement the vast stores of genetic variability within the cotton germplasm collection. The evaluation and utilization of the cotton germplasm collection has had only limited progress to date due to the complexity of the genome and the limited amount of funding available. To date, molecular tools such as RFLPs (Reinisch et al., 1994; Shappley et al., 1998), RFLPs and RAPDs (Yu et al., 1999), SSRs (Xiao, unpublished) and AFLPs (Reddy et al., 1997 Xiao, unpublished) have been utilized to evaluate selected cotton germplasm. QTL identification has been demonstrated by Jiang et al. (1998), Yu et al. (1998), and Shappley et al.(1998) thereby establishing the potential to discover new genes to enhance cotton production.

Within this proposal a concerted effort will be presented to create a process to hone down the large collection into a workable set which will provide detailed characterization, evaluation, and the utilization of valuable diverse genotypes in cotton improvement. The processes will combine molecular tools, model system genotyping, functional genomics, phenotypic screens, and physiological analyses to identify specific genotypes for integration into breeding programs and thus the improvement of cotton varieties in the marketplace. In addition, investigations into the genetics and physiological basis of tissue culture and somatic embryogenesis will be undertaken. The major objectives are as follows:

6. Determine genetic relationships among selected cotton germplasm. Within the large assortment of genotypes within the cotton germplasm collection, it is necessary to pare down the number to a subset of genetic outliers that will provide the maximum potential for success within the program. Selections will be based on molecular techniques using a combination of AFLPs, RFLPs, SSRs, RAPDs, SNPs, and ESTs. These molecular analyses will determine the relationships within the collection and provide information on their relationships with historical and current cultivars. A subset of genotypes will be selected based on these data for further characterization.
7. Utilize a plant system approach to discover potential genes within the germplasm collection. This will involve the construction, ordering, and distribution of BAC and EST libraries in cotton. To this end the *Arabidopsis* genome will be utilized as a model system to discover complementary conserved DNA sequences within the selected subset to use as potential candidates for cotton improvement. *Arabidopsis* ESTs will be evaluated against ESTs from the cotton germplasm collection and the information from the cotton BAC library to look at cotton functional genomics.
8. Develop microarrays for comparative analyses of genotypes, physiological/ developmental stages, organs, tissues, and cell types.
9 Tissue culture and transformation technologies.

Objective 4: Germplasm enhancement for biotic and abiotic stress resistance and agronomic traits.

Quantitative trait loci have been identified using populations of both the interspecific and intraspecific hybrids of the allotetraploids *G. hirsutum* and *G. barbadense*. Greater variability exists between *G. hirsutum* and *G. barbadense* than within each species. Reinisch et al. (1994) made use of the interspecific hybrid's variability to prepare an RFLP map. This map has been employed by several groups of experimenters. Jiang et al. (1998) used RFLP markers from the interspecific map to find QTL for various traits of fiber quality (elongation, length, strength and thickness), yield components and earliness. Wright et al. (1998) identified one A genome and five D genome chromosomes which had genes that conferred resistance to *Xanthomonas campestris* pv. *malvacearum*. A complementary F2 population was derived from a cross between two homozygous lines *G. hirsutum* acc. TM-1 and extra-long-staple *G. barbadense* acc. 3-79 at the USDA-ARS, Crop Germplasm Research Unit (TX). Because neither of these two genetic standards is photoperiodic, all filial generations are amenable to QTL analysis of quantitative traits that are maturity-dependent, i.e., many economically important fiber and seed traits (Yu et al., 1998). Segregation data of 171 F2 individuals of the TM-1 x 3-79 cross were obtained for a collection of over 1,000 genetic markers (Yu et al., 1997; Reddy et al., 1997; Burr et al., unpubl.; Pepper et al., unpubl.). Thirteen QTLs for fiber quality properties (strength, length, and fineness) have been identified in 3-79 (Yu et al., 1998; Kohel et al., 2000). These QTLs explained the phenotypic variance ranging 30% to 60% for each fiber property in the F2 population. Ulloa and coworkers (2000) obtained variation by crossing a *G. hirsutum* cultivar with a line that contained a significant introgression of *G. hirsutum* and *G. barbadense* (Cantrell et al., 1998). From this population, QTL for stomatal conductance were associated with markers (Ulloa et al., 2000). Though marker discovery has been more arduous with intraspecific crosses, populations derived from intraspecific hybrids of *G. hirsutum* have revealed QTL for lint yield and fiber strength (Meredith, 1994). Kloth (1995) used a morphological marker, pilose (T1), to identify loci influencing fiber length, micronaire, perimeter, and wall thickness.

The use of markers to introgress genes into elite lines is pending. The major barrier is availability of worthwhile genes. However, the USDA, ARS group at Mississippi State University seems on the verge of finding alleles for root-knot nematode resistance. Germplasm accessions and commercial lines with resistance to root-knot nematode have been identified (Robinson et al., 1977), proteins have been correlated with resistance to root-knot nematodes (Callahan et al., 1997) and prepared an RFLP map (Shappley et al. 1996, 1998) is under development. Work is also underway on identifying genes and gene products that would lead to enhanced fiber properties and resistance to diseases such as aflatoxin produced by *Aspergillus flavus*.

Traditional applied plant breeding approaches have made great progress in improving cotton yield, quality, and resistance to an array of abiotic and biotic stresses. The foundation for this progress has been the identification of desirable genetic variation. Efforts over the next 5 year period will be made to characterize accessions within the cotton germplasm collection not previously investigated for traits of interest. In certain instances, desirable genetic variation is not readily accessible and research into methods to make it more readily usable will be undertaken. Both efforts are seen as crucial to maintaining a broad genetic base in cotton.

1. Detection and mapping of qualitative and quantitative trait loci.
   a. Marker-assisted selection is a prized goal. For this to be practical, QTL must be assigned to intervals bracketed by markers. Quantitative trait loci are assigned to markers by the use of a segregating population which is produced from parental plants with ample variation in both markers and quantitative traits. Parental plants are selected on the assumption that divergence of phenotype reflects divergence of genotype for QTL. This assumption is extended to the anonymous phenotypes of the DNA markers and marker variation between candidate parental lines is not generally assessed beforehand. However, if the candidate lines are carefully chosen, prior analysis of marker genotypes may not be strictly necessary: U.S. cotton varieties show marker variation based on region of origin (Meredith 1994). Abundant variation can always be found by creating interspecific crosses, but segregation distortion in these combinations complicate analyses. Once two cotton varieties or lines are crossed, the hybrid is self-pollinated; and an F2 generation is produced. Frequently, but not necessarily, the F2 plants are used to screen markers. Most quantitative traits are not reliably evaluated on a single plant basis and, thus, require a replicated experimental design employing the progeny of F2 plants to estimate the phenotypic value of the quantitative trait. To produce sufficient seed for replicated designs several schemes are used, all based on individual harvest of F2 plants. Choices include the use of F3 progeny rows, inbreeding to near-homozygosity to develop single or multiple inbred lines (recombinant inbreds) from each plant in the F2, and populations created by intermating the plants in an F3 progeny row.

b. Data will be analyzed with a variety of publicly available computer programs. New markers will be checked for linkage and, if prior information is available, oriented with respect to previously utilized and discovered markers. For each quantitative trait, the data is tested to insure that significant variation exists between F2-derived progeny. Lastly, the marker and quantitative trait data are combined and tested to determine if the QTL lies between linked markers or are adjacent to markers that are not yet associated with a linkage group.

2. Characterize the germplasm subsets by phenotypic and physiological analyses to identify specific traits of interest.
   a. A series of experiments will be conducted to determine the potential of selected genotypes against biotic and abiotic stresses. Biotic stresses will include resistance to diseases and insect pests, and measurements will be taken to quantify the specific responses. Abiotic stresses will include extremes of temperature, salinity, and nutrients. Measurements for evaluation will include growth and development and physiological processes such as photosynthetic rates and nutrient accumulation in specific plant parts.
   b. Additional characterization to discover genetic variants in fiber traits and cottonseed components will be undertaken. Fiber characteristics such as length, strength, elongation, micronaire, and uniformity will be investigated from the germplasm pool. In addition, seed quality traits such as amino acid and fatty acid content and their respective specific chemical components will be measured.
3. Development and utilization of techniques to broaden the genetic base of cotton.
   a. Day-neutral conversion program.
   b. Interspecific hybridization.
   c. Methods of gene introgression.

The analysis of quantitative trait loci and their association with molecular markers is expected to produce substantial benefits to cotton breeding. The association of markers with quantitative traits provides an opportunity to understand the organization of the cotton genome and the interaction of the two genomes in the allotetraploid. Such understanding will allow improved results with transformation and traditional breeding. In the near term, knowing which QTLs are tethered to markers will bring a new level of efficiency to such projects as breeding cotton for improved fiber traits, disease resistance, and other traits of agronomic importance. Efficiencies are expected to come in the form of improved means of selecting parental plants, early generation testing for desirable alleles, or the transfer of genes to agronomically elite lines from variation-rich, but agronomically poor, germplasm lines.

Objective 5: To refine and develop cotton breeding and variety testing methodologies and techniques.

Many questions addressed by the research proposed in this section may seem to have been answered by previous research. However, plant breeding and genetics are dynamic activities that must respond to the challenges of cotton production. Two major factors make it prudent to revisit some of these questions. First, a "yield plateau" or even decline over the last decade has been noted by several leading scientists and commodity groups (Report of the American Cotton Producers, 1999). This suggests some fundamental change in the germplasm pool in active use by plant breeders since genetic gains in yield by traditional means do not seem to be occurring. One approach to solving this yield plateau problem is to broaden the germplasm base as addressed in other sections of this proposal. Another approach is to re-examine how we go about achieving genetic gains, both in terms of methodology used to develop and evaluate new genotypes and in terms of the values of critical genetic parameters such as heritability. Genetic parameters are a function of the genetic populations they describe. When genetic populations change, as seems to be the case recently, the genetic parameters of the population change and the approach to further manipulate the population may also require change.

The second factor that justifies re-visiting these questions is the advent of new technology. Utilizing new molecular genetic technology is addressed in another section of the proposal. Advances in information technology offer potential for progress in traditional field evaluation of germplasm. In addition to facilitating data acquisition, the vast computing power now widely available makes it feasible to utilize more sophisticated experimental design and data analysis than in the past. Design and analysis beyond the traditional randomized complete block (RCBD) may allow breeders to identify superior genotypes more accurately and dependably.

The first objective, updating pedigree information, builds on previous work and keeps breeders up to date. This document also provides historic information to future generations of cotton breeders.

Cotton breeders currently use a variety of breeding strategies, but the choice of method is seldom based on solid comparison with other methods. The second objective of the proposed research is to provide breeders with quantitative information on contrasting breeding strategies and on methods that are not in wide use in cotton. Comparisons among methods in breeding cotton for disease resistance have been made (Baker and Sappenfield, 1985), but few such comparisons are available in breeding for yield.

Early-generation, bulk testing has proven useful in some crops and situations (Cregan and Busch, 1977; Singh et al., 1990), but not in others (e.g., Fowler and Heyne, 1955). This approach or modifications of it have been used by some commercial and public breeders, but its utility has not been empirically measured to any great extent in cotton (Green and Culp, 1989). A preliminary study at the Delta Research and Extension Center indicated a moderate positive correlation between early generation performance and performance descendent pure lines (Barut, 1998), but this finding should be corroborated.

The heritability of lint percentage and fiber quality has been determined in numerous genetic populations. However, these studies were done some time ago, and values may not apply to genetic populations currently in use.

At early stages of variety development, when seed are too scarce for replicated testing, breeders traditionally evaluate a large number of genotypes by visual examination. Although this practice is often hallowed as the "art" of breeding, greater progress might be achieved by the application of science (i.e., objective yield determination) at this point. Field variability often masks genetic differences in the absence of replication. Various techniques have been proposed to account and/or adjust for field variability. Calhoun (1997) demonstrated that modified augmented designs (MAD) could be used to reduce the influence of field variability when such variability was fairly large. This finding needs to be corroborated and other methods need to be evaluated.

Among cotton breeders, the RCBD and simple analysis of variance (ANOVA) are used almost exclusively in replicated yield trials. Bowman (1989) demonstrated that more sophisticated designs and analyses can reduce and/or account for field variability that can mask genetic differences, even in replicated trials. These findings need to be corroborated in other regions by calculating the relative efficiency (Cochran and Cox, 1957) of the various designs. Furthermore, the techniques and computer code for field layout and data analysis of these more sophisticated designs needs to be made more accessible to the general community of cotton breeders.

The concept of increasing yield by selecting for increases in specific yield components was examined in some depth in the 1970s (e.g., Worley et al. 1976). The general consensus was that compensatory shifts among yield components precluded progress in final yield. Nevertheless, selection for yield per se has resulted in shifts among yield components (Bridge et al., 1971). Lewis et al. (2000) has argued fairly convincingly that (indirectly) measuring fiber number (or fiber weight) per seed, rather than yield per se, may be more effective in identifying genetically high yielding varieties because management and environment have a large impact on seed number per hectare, but a relatively small impact on fiber per seed. This finding needs to be corroborated in other regions and methods to economically determine fiber per seed must be developed.

The advent of genetically engineered cotton varieties has raised several questions about how to measure or use them for both producers and plant breeders. One difficulty posed is how to evaluate the new technology in official variety trials. To date, genetically engineered cotton varieties have been tested in trials that may mask the benefits of the value added traits. Alternative variety testing methods may need to be developed that can provide producers with information they can use in their variety selection efforts. Additionally, cotton breeders are faced with a new challenge in their variety development efforts due to the wide spread adoption of transgenic cotton in the marketplace and in the improved germplasm base: the merits and demerits of forward crossing of transgenic genotypes. The objectives of this research are:

1. Update information on the genetic relationships among potential parental germplasm to aid in identifying divergent parents for use in artificial hybridization: Information on the pedigrees of commercial and non-commercial germplasm released since 1995 will be collected from the Plant Variety Protection Office, official release notices, and other sources as necessary. This information will be used to update the MAFES bulletin, "Pedigrees of Upland and Pima Cotton Cultivars Released between 1975 and 1995" (Calhoun et al., 1997), and the USDA technical bulletin, "Coefficients of parentage for 260 cotton culitvars released between 1970 and 1990" (Bowman et al. 1997).
2. Quantify advantages and disadvantages of divergent variety development strategies.
3. Compare genetic gain from selection in pedigree breeding schemes (plant-to-row selection each generation) with that of a bulk breeding scheme (plant selection delayed until F4 or F5)
    a. Determine the value of measuring performance of unselected early generation (F2 to F4) bulk populations in identifying those with a high probability of producing pure lines with superior yield and/or fiber quality.
    b. Determine relationship between values for lint yield, lint percentage, and fiber quality parameters measured in single plants to those values measured in descendant progeny rows.
4. Develop guidelines for efficient and discriminating genotype evaluation.
5. Evaluate various methods of nonreplicated evaluation:
    a. Visual evaluation vs. yield measurement: Plots in replicated trials of new, fairly unfamiliar strains will be evaluated visually and assigned a numeric value for yield potential by one or more persons. Plots will then be harvested for yield determination. Visual evaluations and objective yield measurements will then be correlated on an individual plot basis and on an entry-mean basis. In addition, comparisons will be made of entries that would have been discarded on visual vs. measured evaluation.
    b. Compare various methods to remove or adjust for spatial field variability in small-plot yield trials in cotton (comparison of experimental designs): Several yield trials will be set up in rectangular lattice designs. Results will be analyzed using the lattice model and traditional analysis of variance as if the design had been an RCBD. Control of experimental error will be compared between the two methods.
    c. Comparison of post-mortem data analyses: Existing variety test data will be analyzed using traditional analysis of variance, nearest neighbor analysis, and trend analysis. Differences among analyses in terms of relative efficiency, experimental error, and measured entry performance will be examined.
6. Determine the value of yield component analysis in place of, or in addition to, traditional yield estimation.
7. Evaluate the need to incorporate systems testing in variety trials for transgenic technologies.
8. Investigate the genetic consequences of forward crossing of transgenic parents.

Updated information on pedigrees and genetic relationships among potential parental material will allow breeders to make more genetically divergent crosses thereby increasing the probability of generating new, unique genetic combinations. Quantifying costs and benefits of different breeding strategies will allow breeders to make informed decisions regarding the choice of strategies, rather than relying solely on habit or training. This information should result in more efficient utilization of resources and therefore greater benefit (to breeding entities, producers, and ultimately to consumers) at a given level of input. A document will be developed that outlines advantages, disadvantages, and appropriate applications of alternative experimental designs and data analyses relative to RCBD/traditional ANOVA with specific reference to cotton and to issues encountered in various regions. The document will include software available (including SAS and required SAS code) to facilitate these designs and analyses. Language and explanations in the document will be geared toward non-statisticians. Developing a user friendly guide to experimental design and analysis would allow non-statistically oriented breeders to get the most (in terms of identifying superior genotypes) out of the data they generate. Yield component analysis offers the promise of a more discriminating and stable method of evaluating genotypes. In combination, these outcomes could help overcome the current stagnation in yield improvement.

Objective 6: To develop cotton bioinformatic systems.

Advancements in information technology over the last decade have been impressive and now allow for the development of sophisticated, yet user friendly, programs for data collection, synthesis, and querying. Such bioinformatics tools are already being used to some degree, but are anticipated to become increasingly more important in future cotton improvement efforts. Over the next 5 yr period, existing information systems for cotton, namely CottonDB, should be enhanced to handle the large volume of classical and molecular genetic data being generated through the cotton research produced by this project and the cotton research community at large. The goals for this objective are:

1. Continued curation of CottonDB.
2. The development of user friendly data input and research collaboration linkage tools.
3. The establishment of linkages among plant genomic and cotton web sites.
4. The coordination and standardization of data and nomenclature for cotton.

5. Development of genomics analysis techniques (including graphical genotypes).
6. Methodologies for presenting and preserving variety testing data.
7. Development of internet information exchange sites devoted to cotton.
8. Information exchange (annual meetings, biennial Cotton Breeder's Tour; symposia, etc.)

Summary

The primary goal of this research is to improve the yield, fiber quality, and profitability of cotton production. The objectives of the research conducted under this Regional Research Project will address numerous important and/or unresolved issues relevant to increasing the agronomic potential of cotton. Success in achieving these objectives is expected to have a positive impact on cotton production and the work should be considered incomplete until it has done so. Several of the objectives will focus on germplasm enhancement and characterization, topics critical to improving cotton productivity. By the end of the proposed 5 yr period of coordinated work, it is anticipated that a significantly greater percentage of the cotton germplasm currently maintained in the U.S. will be characterized for standard Gossypium agronomic descriptors and additional traits such as insect and disease resistance. A more detailed and integrated genetic map of cotton will also be produced using both molecular genetic and cytogenetic techniques. Such information will help revolutionize the trait-specific improvement of cotton.

Marker-assisted selection has been a tantalizing promise of molecular biology for many years, but has not seen widespread utilization. With the development of a highly saturated cotton genetic map and the development of transportable, tightly linked markers (both of which are expected outcomes of the research outlined in this proposal), it is anticipated that such selection will be more commonly practiced in cotton within the next 5 yr period. This will allow cotton breeders and geneticists to tackle more effectively and efficiently those significant problems facing U.S. cotton producers and processors. Notable examples will likely include increased resistance to pests (e.g., root-knot and reniform nematodes), reduction of gossypol in seed, improved fiber strength and length, and higher yield.

Increasing the efficiency with which new cotton varieties are developed will allow plant breeders to more readily address the problems faced by cotton producers and consumers. A recent study (Bowman, 2000) found a great similarity in the procedures used by both public and private cotton breeders. This may represent the synthesis of over a century's worth of experience into a streamlined procedure, but it is deserving of critical evaluation. The availability and merits of increasingly sophisticated statistical analysis techniques should be examined. These two issues are a core component of this proposal.

New tools for the collection and dissemination of both historical and current data have been developed. This proposal contains several elements for data synthesis, curation, and for making the results of this research available for use by cotton producers primarily through the use of the internet. Future progress in cotton improvement is dependent on such activities.

Not explicitly addressed an objective in this research proposal, but critical to the future of U.S. cotton production and profitability, are the opportunities that regional project activities afford for the formal and informal interaction amongst cotton researchers. At one level, this includes professional information exchanges and collaborative research initiatives that are enhanced by personal contact. Just as important, but more nebulous, are the casual contacts established.

Finally, an expected outcome of the contacts and collaborations that arise from this regional research project is that greater mutual respect and appreciation will result between public and private cotton breeders and geneticists. In a break with the past, cotton researchers from the public and private sectors developed this proposal. An invitation to all U.S. cotton researchers to participate in the research objectives outlined above has been extended. Through our combined efforts we hope to keep U.S. cotton competitive, environmentally benevolent, and of the highest quality.

Acknowledgements

This proposal is the result of the efforts of members of two groups, SRDC-9801 and SRIEG-61, their advisors, and numerous reviewers. Their contributions are gratefully acknowledged.

References

Baker, I.A. and W.P. Sappenfield. 1985. Comparative efficiencies of the "Sequential Inoculation Selection" (SIS) and "Multi-Adversity Resistance" (MAR) systems of selection for multiple disease resistance in cotton. p. 63-69. In J.M. Brown and T.C. Nelson (ed.) Proc. Beltwide Cotton Prod. Res. Conf. Natl. Cotton Council Am., Memphis, TN.

Bowman, D.T. 1989. Use of trend analysis to improve precision in yield trials. p. 123-124. In J. Brown and D.A. Richter (ed.) Proc. Beltwide Cotton Prod. Res. Conf. Natl. Cotton Council Am., Memphis, TN.

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

Bowman, D.T., O.L. May, and D.S. Calhoun. 1997. Coefficients of parentage for 260 cotton cultivars released between 1970 and 1990. USDA-ARS Tech. Bull. 1852.

Bridge, R.R., W.R. Meredith, Jr., and J.F. Chism. 1971. Comparative performance of obsolete and current varieties of upland cotton. Crop Sci. 11:29-32.

Calhoun, D.S. 1997. An evaluation of modified augmented designs for single-replication yield testing in cotton. p. 460-464. In P. Dugger and D. Richter (ed.) Proc. Beltwide Cotton Conf. Natl. Cotton Council Am., Memphis, TN.

Calhoun, D.S., D.T. Bowman, and O.L. May. 1997. Pedigrees of upland and Pima cotton cultivars released between 1970 and 1995. Miss. Agric. & Forestry Exp. Stn. Bull. 1069.

Callahan, F.E., J.N. Jenkins, R.G. Creech, and G.W. Lawerence. 1997. Changes in cotton root proteins correlated with resistance to root knot nematode development. J. Cotton Sci. 1:38-47.

Cantrell, R.G., M. Ulloa, R. Percy, E. Zeiger, and Z. Lu. 1998. Genetic variation for stomatal conductance in an interspecific cotton population. p. 485-486. In P. Dugger and D. Richter (ed.) Proc. Beltwide Cotton Conf. Natl. Cotton Council Am., Memphis, TN.

Cochran, W.G., and G.M. Cox. 1957. Experimental designs. 2nd ed. John Wiley & Sons, Inc. New York.

Cregan, P.B. and R.H. Busch. 1977. Early generation bulk hybrid yield testing of adapted hard red spring wheat crosses. Crop Sci. 17:887-891.

Fowler, W.L. and E.G. Heyne. 1955. Evaluation of bulk hybrid tests for predicting performance of pure line selections in hard red winter wheat. Agron. J. 47:430-434.

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

Kloth, R.H. 1995. Quantitative trait loci affecting cotton fiber are linked to the t_1 locus in uplant cotton. Theor. Appl. Genet. 91:762-768.

Kohel, R.J., J. Yu, Y.H. Park, and G. Lazo. 2000. Molecular mapping and characterization of genes controlling fiber quality in cotton. Euphytica (in press).

Green, C.C., and T.W. Culp. 1989. Comparison of selected lines, unselected and selected bulks for yield in cotton. p. 83. *In* 1989 Agronomy abstracts. ASA, Madison, WI.

Lewis, H., O.L. May, and F. Bourland. 2000. Cotton yield components and yield stability. *In* P. Dugger and D. Richter (ed.) Proc. Beltwide Cotton Conf. Natl. Cotton Council Am., Memphis, TN. (In press).

Meredith, W.R. Jr. 1994. Use of molecular markers in cotton breeding. p. 303-308. In G.A. Constable and N.W. Forrester (ed.) Challenging the Future: Proc. World Cotton Res. Conf.-1. Brisbane, Aust.

Reddy A.S., R.M. Haisler, Z.H. Yu, and R.J. Kohel.1997. AFLP mapping in cotton. Proc. Plant Genome V, The Int. Conf. on the Status of Plant Genome Research. San Diego, CA.

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

Report of the American Cotton Producers. 1999. Blue Ribbon Yield Committee. Natl. Cotton Council Am., Memphis, TN.

Robinson, M., J.N. Jenkins, and J.C. McCarty, Jr. 1997. Root knot nematode resistance of F_2 cotton hybrids from crosses of resistant germplasm and commercial cultivars. Crop Sci. 37:1041-1046.

Shappley, Z.W., J.N. Jenkins, W.R. Meredith, and J.C. McCarty, Jr. 1998. An RFLP linkage map of upland cotton, *Gossypium hirsutum* L. Theor. Appl. Genet. 97:756-761.

Shappley, Z.W., J.N. Jenkins, C.E. Watson, Jr., A.L. Kahler, and W.R.Meredith. 1996. Establishment of molecular markers and linkage groups in two F_2 populations of upland cotton. Theor. Appl. Genet. 92:915-919.

Singh, S.P., R. Lepiz, J.A. Gutierrez, C. Urrea, A. Molina, and H. Teran. 1990. Yield testing of early generation populations of common bean. Crop Sci. 30:874-878.

Worley, S., Jr., H.H. Ramey, Jr., D.C. Harrell, and T.W. Culp. 1976. Ontogenetic model of cotton yield. Crop Sci. 16:30-34.

Wright, R.J., P.M. Thaxton, K.M. El-Zik, and A.H. Paterson. 1998. D-Subgenome bias of *Xcm* resistance genes in tetraploid *Gossypium* (cotton) suggests that polyploid formation has created novel avenues for evolution. Genetics 149:1987-1996.

Ulloa, M., R.G. Cantrell, R.G. Percy, and E. Zeiger. 2000. QTL analysis of stomatal conductance and relationship to lint yield in an interspecific cotton. J. Cotton Sci. 40:10-18.

Yu, J., Y.H. Park, G.R. Lazo, and R.J. Kohel. 1997. Mapping cotton genome with molecular markers. p. 447. *In* P. Dugger and D. Richter (ed.) Proc. Beltwide Cotton Conf. Natl. Cotton Council Am., Memphis, TN.

Yu, J., R.J. Kohel, Y.H. Park, and G.R. Lazo. 1999. Localization and characterization of genes controlling fiber quality properties in cotton genomes (*G. barbadense* and *G. hirsutum*). Proc. 7th Int. Conf. on the Status of Plant and Animal Genome Res., San Diego, CA.

Yu, Z.H., R.J. Kohel, Y.H. Park, and G.R. Lazo. 1998. QTL analysis of fiber quality properties in cotton. Proc. Biochem. and Biotechnol. Cotton Workshop. Galveston, TX.