Utilization of exotic cotton germplasm resources to increase genetic diversity

J. McD. Stewart and C. Feng
Dept. Crop, Soil, & Environmental Sciences, University of Arkansas, Fayetteville AR USA
Correspondence author jstewart@uark.edu

ABSTRACT

Utilization of a germplasm pool can be compared to fishing in a lake. As a result of heavy usage, the benefits gained for the effort expended are minimal. New possibilities need to be explored in both cases to obtain significant results. Molecular measurements of genetic similarity of cotton cultivars from locations around the world indicate that genetic diversity among current breeding programs is very limited for continued cotton improvement. Future sources of genetic diversity for cotton improvement must come from the germplasm pools represented in the 1E, 2E, and 3E pools comprising the Gossypium species. The University of Arkansas has a strong program to evaluate and utilize exotic Gossypium germplasm to increase available genetic diversity for breeding. Development of strategies for hybridisation to increase the efficiency of recombination, and utilization of molecular techniques to aid in monitoring and introgression of traits make the potential for exotic germplasm to contribute to genetic improvement of cotton very high.

A fable

Once there was a pristine lake that held a vast bounty of nature. In the course of human enterprise the lake offered sustenance to the inhabitants around the lake. One day a visitor from afar caught a very large fish from the lake, and as he returned to his home he carried stories of the wonderful lake. More people came to fish in the lake and word of its virtues spread far and wide. Soon many people came to fish in the lake. At first the lake seemed limitless and provided many fish of all sizes. However, with time the fish that were caught became fewer and smaller, until only an occasional very small fish was caught. Each new generation of fishermen considered their catches to be normal and they were happy. So, they continued to fish in the lake, unaware that it was nearly depleted of fish or that other lakes might now provide better fishing.

Now think of the fabled lake as the germplasm pool that cotton breeders are using in their efforts to "fish" for new genes for crop improvement. In the present day situation they continue to use the same germplasm pool that many have used before them, and they are able to obtain only very small gains.

Genetic similarity

All breeding efforts carry the assumption that genetic diversity exists within the breeding germplasm pool for the traits of interest. However, the breeding pools being used by cotton breeders around the world apparently are providing very little genetic diversity. This is strongly indicated by the genetic similarity among common cotton cultivars (Table 1). Each of the countries from which data are available report high genetic similarities in their common or even selected cultivars. In those cases where similarity among specific *G. hirsutum* genotypes is reported at 60%, the data are biased toward selection for differences. In these cases only PCR primers detecting polymorphisms were used, while those giving monomorphic bands among the genotypes were not used in the calculations. This results in artificially low similarity indexes. The cotton industry in the USA has recognized that the genetic base of existing commercial cultivars has become very narrow. The lack of yield improvement in the last few years is attributed, in part, to the lack of diversity in breeder's germplasm pools. To rectify this situation, a research initiative sponsored through Cotton Incorporated (CI) is underway to fund research with the specific intent of increasing the genetic diversity available for cotton improvement. The major component of the program seeks to capture diversity from the primary germplasm pool, but some effort is also directed toward the secondary germplasm pool. As a component of the research, molecular markers linked to any traits of interest are also being developed.

Cotton germplasm pools

Aside from the few genes that can be introduced into cotton via genetic engineering, the only source of genetic variation for cotton improvement is from the exotic germplasm comprised in the primary, secondary, and tertiary germplasm pools of *Gossypium* species. The primary germplasm pool is the one most readily usable by breeders. The five species in this group are tetraploid with AD genomes and can be hybridised relatively easily with cultigens. In the F_1 and subsequent generations, recombination is high, although areas of the exotic genome have very low or no recombination due to various phenomena including gamete selection. Extensive collections have been made of feral and ruderal plants of *G. hirsutum* and *G. barbadense* from their centers of diversity so representatives of these species make up the major portion of the collection. The accessions are maintained by the US Department of Agriculture at College Station, TX and accessions are freely available to all bona fide breeders upon request. The secondary germplasm pool comprises the diploid *Gossypium* species that have reasonably high chromosome homoeology to either the A subgenome or the D subgenome of the tetraploid AD genome species. This includes all the species in the A, B, D and F genomic groups. Strategies for obtaining hybrids and promoting genetic recombination with tetraploid cotton were presented at the first world cotton conference (Stewart, 1995). The tertiary germplasm pool is composed of the species grouped in the C, E, G, and K genomes of Australia and NE Africa. Although several interesting traits reside among

these species, their chromosomes have limited homoeology with either A or D genomic chromosomes. Hence, transfer of traits from these species to cultivated diploid or tetraploid cottons is usually very difficult. The number of known species (formally described and undescribed) in *Gossypium* is currently 50. A vast amount of genetic diversity resides in these species, but as all breeders know, linkage drag introduces many genes from the exotic genetic material that have deleterious effects on yield and quality. The available genetic resource for cotton improvement can be equated to a large pile of oysters. It is the job of those geneticists involved in pre-breeding to discover which ones within that pile contain the pearls, and then to separate these from all the shells without pearls. Molecular techniques to measure genetic similarity and to provide a measure of chromatin introgression in recipient populations has the potential to greatly improved the efficiency of utilization of exotic germplasm. An incomplete listing of some of the uses for which molecular markers have been employed include 1) identification and protection of varieties and/or patented clones; 2) certification of genetic purity of lines and hybrids; 3) monitoring of outcrossing and self pollination; 4) evaluation of germplasm and populations (variability, diversity, classification, genetic distance and phylogeny); 5) construction of genetic linkage maps; 6) genetic mapping of QTLs; and 7) MAS - Marker Assisted Selection.

The remainder of this presentation is devoted to a summary of the work with exotic *Gossypium* germplasm that has been done, or is in progress, at the Cotton Biotechnology and Germplasm Development Laboratory at the University of Arkansas. These activities range from germplasm acquisition and taxonomy, to cytogenetics, to evaluation and introgression.

Germplasm acquisition, evaluation and utilization at the University of Arkansas

The cotton research program at the University of Arkansas has a strong component devoted to evaluation and utilization of exotic *Gossypium* germplasm to increase the genetic diversity of breeding pools. When considering germplasm, diversity may be described at several levels. A fundamental recognition of distinctiveness occurs at the level of species. The criteria that distinguish species are in the realm of taxonomy and distinction traditionally is based on morphological traits, with primary emphasis on the flower, and secondarily on the foliage. The circumscription of a species may be very broad or very narrow, based on the specimens available for study and the opinion and philosophy of the taxonomic authority. As a result of germplasm expeditions by JMS in the last few years, two undescribed, and previously uncollected, species have been discovered. In addition, extensive *in situ* observations of known Mexican *Gossypium* species suggest that the current circumscription of at least one taxon encompasses a wide range of morphological and geographic diversity that justify re-examination of the species delineation. A classical taxonomic study of this species, as well as of two taxa collected in Australia and Mexico deemed to be undescribed, is in progress. Another measure of germplasm diversity lies in the degree of homoeology of chromosomes among related species. The *Gossypium* species have been placed into genomic groups designated A, B, C, D, E (Beasley, 1942) F (Phillips and Strickland, 1966), and G (Edwards and Mirza, 1979). Based on unpublished data, Stewart (1995) proposed an eighth diploid genomic group, designated K, that encompasses the *Gossypium* species taxonomically classified in Section *Grandicalyx*. We are currently collecting data on meiotic figures of microspore mother cells from numerous interspecific hybrids involving the Australian *Gossypium* species classified in the Sections *Grandicalyx*, *Hibiscoidea*, and *Sturtia*. These data should provide insight into the correct genomic affinities of all the Australian species. Within the primary pool, we are developing recombinant inbred lines (RIL) with a common commercial cultivar and four wild species including *G. hirsutum* race *yucatanense, G. tomentosum, G. mustelinum* and *G. darwinii*. We consider the first taxon to be the progenitor, or wild ancestor, of cultivated *G. hirsutum*. At least 500 RILs from each hybrid combination are being developed by single seed descent. Two of the lines are in BC_2, one is at BC_1 with partial development of BC_2, and an F_2 population has been developed for the fourth. At BC_4 progeny rows will be developed for evaluation. The large number of lines being developed for each is an attempt to thoroughly dissect the exotic genome while capturing as much of it as possible. Any beneficial alleles or QTLs identified should already be sufficiently introgressed in the BC_4 that a breeder could use them directly in a breeding program. The secondary germplasm pool has been of interest to us for several years, but with the advent of the new CI program we have renewed efforts in this area. Several years ago a (*G. herbaceum* x *G. armourianum*) x *G. hirsutum* hybrid was made and allowed to self, intercross and outcross for 3-4 generations. Dr. Fred Bourland at the University of Arkansas has made selections from this hybrid swarm based on yield and fiber quality parameters. Four of those selections have made it to the advanced stains tests. A number of morphological characters were also selected from this material and have been introgressed into the cotton AD genome. These include red anther and thick leaf from *G. armourianum*, red calyx and boll from *G. herbaceum* and short fruiting branch as a transgressive character. This latter trait has been selected for yield and will be evaluated for its potential in narrow-row cotton production. Our current objective is to develop a series of synthetic allotetraploid hybrids involving the A, F and B species with as many D genome species as possible. Theoretically this provides the most direct approach to obtain recombination of these species with upland cotton. In past work a *G. davidsonii x G. anomalum* tetraploid was used to introgress the D_3 cytoplasm into the cultivated tetrap-

loid nuclear background. Also, 4X hybrids were produced that contained the genomes of *G. arboreum x G. trilobum*, *G. herbaceum* x *G. armourianum* and *G. longicalyx* x *G. armourianum*. Recently we have obtained 2x hybrids of several *G. arboreum* and/or *G. herbaceum* accessions with *G. trilobum*, *G. aridium* and *G. raimondii* that will be subjected to colchicine treatment to double their chromosome number. When specific traits are needed, directed evaluation of unimproved germplasm is often more efficient than arbitrary hybridization. For example, resistance to reniform nematodes has not been found in any *G. hirsutum* cultivar or landrace accession evaluated. However, evaluation of approximately 300 accessions of *G. arboreum* and *G. herbaceum* identified good resistance in 8-10 accessions. The specific accessions are now being hybridised with D genome species as the first step in transferring the resistance to tetraploid cottons. Arbitrary hybridisation of an Asiatic cotton accessions at random would not have been fruitful in this case. With the advent of molecular techniques, our ability to measure genetic diversity has gained a powerful tool that allows evaluation of differences at the DNA level. In addition, molecular markers associated with specific traits can be used to efficiently introgress the trait into a novel genetic background. We employ molecular techniques both to gain an estimate of relatedness among species and to develop markers that we hope will serve in marker assisted selection of traits of interest in the exotic germplasm. As an example of the use of markers to measure species relatedness, a series of more than 900 AFLP and RAPD fragments were generated from among the arborescent *Gossypium* species of Mexico. Analysis of the data with phylogenetic software revealed that greater diversity exists within the species *G. aridum* than among the well recognized species *G. laxum, G. lobatum* and *G. schwendimanii*. The uniqueness of a new taxon collected in 2002 was also demonstrated. We have used RAPDs to identify fragments of DNA associated with the restorer genes for cytoplasmically conditioned male sterility. Two markers have been identified that appear to co-segregate with the Rf_1 gene and one that is approximately 2.5 cM from the Rf_2 gene. Because RAPDs have a reputation of being variable among laboratories or different plant populations, we have converted the co-segregating RAPD markers to STS markers by designing specific primers of approximately 20 bp each that yield a single fragment associated with the restorer gene. In our work with nematode resistance, the current techniques to measure nematode reproduction on test plants are laborious and require more than 60 days to complete. In this instance one or more markers associated with resistance genes would have a very significant impact on time and cost for progeny plant selection during introgression of the resistance. We developed an F_2 population from a hybrid between a reniform nematode-susceptible *G. arboreum* line and a resistant *G. arboreum* line to use in bulked segregant analysis. Two DNA bulked samples were prepared from the 10 most resistant progeny and the 10 most susceptible progeny, and the bulked samples were screened for polymorphisms in polymerase chain reactions with 600 random 10-mer primers. Three primers revealed polymorphisms that appear to distinguish the two samples. These are being tested on individual plants of the population to confirm association with the trait.

Summary

The exotic cotton germplasm pool is equivalent to a very large lake that is still relatively pristine. Great genetic diversity is available in this pool, but because of the perceived difficulties involved in hybridisation and introgression of useful traits, it has been utilized very little to improve our cultivated species. Strategies for hybridisation to increase the efficiency of recombination, and the availability of molecular techniques to aid in monitoring and selection of introgressed traits make the potential for exotic germplasm to contribute to the genetic improvement of cotton very high. Because most of the cultivated genotypes of the world, hence active genetic breeding pools, contain very little genetic diversity, the exotic germplasm pools must be used if we expect to have genetic improvement over the current performance of our cotton genotypes.

References

- Abdalla, A.M., Reddy, O.U.K., El-Zik K.M. and Pepper, A.E. (2001). Genetic diversity and relationships of diploid and tetraploid cottons revealed by AFLP. *Theor. Appl. Genet.*, **102**: 222-229.
- Beasley, J.O. (1942). Meiotic chromosome behavior in species, species hybrids, haploids and induced polyploids of *Gossypium*. *Genetics*, **27**:25-54.
- Edwards, G.A. and Mirza, M.A. (1979). Genomes of the Australian wild species of cotton. II. The designation of a new G genome for *Gossypium bickii*. *Can. J. Genet. Cytol.*, **21**: 367-372.
- Iqbal, M.J., Aziz, N., Saeed, N.A., Zafar, Y. and Malik, K.A. (1997). Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. *Theor. Appl. Genet.*, **94**: 139-144.
- Linos, A.A., Bebeli, P.J. and Kaltsikes, P.J. (2002). Cultivar identification in upland cotton using RAPD markers. *Aust. J. Agric. Res.*, **53**: 637-642.
- Lu, H.J. and Myers, G.O. (2002). Genetic relationships and discrimination of ten influential Upland cotton varieties using RAPD markers. *Theor. Appl. Genet.*, **105**: 325-331.
- Multani, D.S. and Lyon, B.R. (1995). Genetic fingerprinting of Australian cotton cultivars with RAPD markers. *Genome,* **38**: 1005-1008.
- Phillips, L.L. and Strickland, M.A. (1966). The cytology of a hybrid between *Gossypium hirsutum* and *G. longicalyx*. *Can. J. Genet. Cytol.*, **8**: 91-95.
- Rahman, M., Hussain, D. and Zafar, Y. (2002). Estimation of genetic divergence among elite cotton cultivars-genotypes by DNA fingerprinting technology. *Crop Sci.,* **42**: 2137-2144.

• Stewart, J. McD. (1995). Potential for crop improvement through germplasm enhancement and genetic engineering. p. 313-327. *In* G. Constable and N. Forrester (eds.). Challenging the future. *Proceedings of 1st World Cotton Research Conference*, CSIRO, Melbourne.

• Xu,Q-H., Zhang, X.L. and Nie, Y.C. (2001). Genetic diversity evaluation of cultivars (*G. hirsutum* L.) from the Yangzi River valley and the Yellow River valley by RAPD markers. *Acta Genet. Sinica,* **28**: 683-690.

Table 1. *Genetic similarity of cotton genotypes in various cotton producing countries based on molecular markers.*

Country	No. Access.	Method	Range GS	Reference
Australia	19	RAPD	92-99	Multani *et al.* (1995)
China	91	RAPD	Ave. 63*	Xu *et al.* (2002)
Greece	28	RAPD	61*-92	Linos *et al.* (2002)
Pakistan	17(1)	RAPD	82-93(57)	Iqbal *et al.* (1997)
	26	RAPD	82-91	Rahman *et al.* (2002)
USA	43	AFLP	83-99	Iqbal *et al.* (2001)
	8	AFLP	Ave. 86	Abdalla *et al.* (2001)
	10	RAPD	93-98	Lu *et al.* (2002)

* PCR primers that did not give polymorphisms among genotypes were not used in these studies, hence similarity values are biased to the low side.

Table 2. *Genomes and number of species comprised in the* Gossypium *germplasm pools.*

Pool	Genomes	No. Species	Comments
Primary	AD	5	Landraces of G. hirsutum and G. barbadense and 3 wild species.
Secondary	A,B,F,D	21	The races of G. arboreum and G. herbaceum and 19 wild diploid species of Africa and the Americas.
Tertiary	C,E,G,K	24	Wild diploid species of Australia and NE Africa to Pakistan.