

TITLE: *Gossypium* Germplasm Resources for Cotton Improvement

DISCIPLINE: Genomics and Biotechnology

AUTHORS: Richard G. Percy (Corresponding author)
USDA, ARS
Arid Land Agricultural Research Center
21881 North Cardon Lane
Maricopa, AZ 85239
Phone: 520-316-6357
Fax: 520-316-6330
Email: richard.percy@ars.usda.gov

ACKNOWLEDGEMENT: The author would like to thank the members of the Cotton Crop Germplasm Committee, members of the Crop Germplasm Research Unit, USDA, ARS, and Mark Bohning of the Germplasm Research Information Network (GRIN) for their assistance.

ABBREVIATIONS: BIL, backcross introgression line; CIRAD, Centre de cooperation internationale en recherche agronomique pour le développement; NIIL, near isogenic introgression line; QTL, quantitative trait loci; RIL, recombinant inbred line; U.N. FAO, United Nations, Food and Agriculture Organization; USDA-ARS, United States Department of Agriculture – Agricultural Research Service

***Gossypium* Germplasm Resources for Cotton Improvement**

ABSTRACT

Only a very small fraction of the genetic diversity residing in the *Gossypium* genus is represented in improved, elite cotton germplasm. Although genetic diversity in elite germplasm is reported to be narrow, diversity on the farm is narrower due to preferential planting of successful cultivars and breeding techniques that tend to promote an over-reliance on a few genotypes. By contrast, the genetic variability present in *in situ* populations and *ex situ* collections represents a large, under-utilized resource for cotton improvement efforts. *In situ* resources are in many places being threatened by human development or are disappearing. *Ex situ* resources have been reported to include 49,000 accessions residing in over 25 institutional germplasm collections worldwide. In the United States, the Agricultural Research Service (ARS) of the USDA maintains a *Gossypium* collection of 9,303 accessions, representing 42 species. These accessions are divided into primary, secondary, and tertiary germplasm pools according to their relative genetic accessibility and utility. In 2006, 4,106 accessions of the USDA-ARS collection were distributed to 105 researchers and institutions, worldwide. Among problems challenging the ARS collection are limited resources for seed renewal, incomplete characterization and evaluation of accessions, and unrecognized gaps and redundancies in the collection. Efforts to find and direct resources toward these problems are ongoing. Outside the USDA-ARS germplasm collection, numerous genetic resources for the investigation and improvement of cotton have been created by public researchers, including recombinant inbred line (RIL) populations, back-cross introgression line (BIL) populations, near isogenic introgression line (NIIL) populations, chromosome substitution lines, and day-neutral converted race stocks, among others.

KEY WORDS:

cotton, germplasm, *in situ* conservation, *ex situ* collections

Historically, four species of *Gossypium* have been cultivated for fiber (Brubaker et al., 1999). Of these species, *Gossypium hirsutum* L. is by far the most widely grown worldwide. *Gossypium barbadense* L., grown primarily for its superior fiber attributes, is a distant second in world acreage. In recent years there has been a growing recognition of the limited genetic diversity of elite commercial germplasm of both species, and a growing concern about genetic vulnerability (Bowman et al., 1996). In both species the process of domestication is thought to have dramatically narrowed the germplasm base (Brubaker et al., 1999). Early and modern improvement efforts have further narrowed this base. May (2000) reported that intensive selection imposed to maximize yield and adaptation, along with selection for early maturity, has eliminated substantial variation from elite *G. hirsutum* germplasm pools. Although genetic diversity in elite germplasm is reported to be narrow, actual diversity on the land is narrower, due to preferential planting of successful cultivars and breeding techniques that tend to promote an over-reliance on a few genotypes (Wallace et al., 2007). Against the narrow genetic base of commercially elite germplasm is the diversity residing in wild, commensal, and landrace cottons of six tetraploid and forty-three diploid species of the *Gossypium* genus. The genetic variability present in *in situ* populations and in *ex situ* collections of these species represents a large, under-utilized resource for cotton improvement efforts.

DISCUSSION

***In Situ* conservation of genetic resources.** Most of the genetic diversity to be found in *Gossypium* remains *in situ*. Although the concept of *in situ* conservation of genetic resources has gained recognition and support in recent years, its practice presents unique challenges. Habitat loss due to conversion of wild lands to agricultural use is probably the greatest challenge and is most problematic in developing nations (Rubenstein et al., 2005). Although habitat loss affects

whole plant and animal communities as well as individual species, it is often easier to justify and gain support for community preservation than it is to justify conservation of a single wild species of little apparent economic value. Another source of genetic diversity and therefore a candidate for *in situ* conservation are landraces (also, in the case of cotton, dooryard or commensal cottons). Since landrace or dooryard cottons frequently are grown in subsistence or basic economies, the adoption of a modern cotton industry has often led to their abandonment in favor of improved cultivars (Singh et al., 2003). Unlike wild species that are often the beneficiaries of benign neglect in a wild environment, there is little chance of landrace preservation *in situ* on agricultural land without some form of economic incentive being offered (Rubenstein et al., 2005). Wild species also can be adversely affected by the adoption of modern cotton production practices in adjacent territories. Wild and feral populations of cotton often harbor the same insect and disease pests as adjoining cultivated cotton, are perceived as pest sources, and become the target of eradication programs. In such situations, *in situ* management of wild species may become necessary, as well as some forms of compensation and economic incentive for *in situ* preservation to be successful. Few examples of the implementation of active *in situ* preservation efforts currently exist for cotton. One example is an effort to protect native colonies of the species *G. mustelinum* Meirs ex Watt in Brazil (Barroso et al., 2006). This one conservation effort demonstrates that, with local support, conservation can be as simple as erecting a fence to exclude cattle.

With only one native cotton species, the relevancy of *in situ* conservation within the continental United States is low. However, this one species, *G. thurberi* Todaro, has faced depredations in the past. In Arizona in the 1930's an eradication program was instituted against *G. thurberi* in a misguided effort to control the cotton boll weevil (Kearney and Peebles, 1942).

This eradication program effectively removed *G. thurberi* from part of its range, but at present the species appears to be in little danger and does not require active *in situ* conservation measures. Although the U.S. has no *in situ* cotton conservation efforts, it has participated in a related activity – the establishment of a “living collection” of *Gossypium* species in Mexico (Ulloa, 2006). In a recent cooperative collecting effort with Mexico, the U.S. aided in the establishment a nursery containing diploid and tetraploid species in or near their native ranges.

***Ex Situ* collections and the USDA-ARS collection.** Currently the vast majority of cotton germplasm resources are held in *ex situ* collections. In 1996 the U.N. FAO reported approximately 49,000 germplasm accessions residing in collections worldwide (U.N. FAO, 1996b). Using FAO data, 73 percent of these accessions were held by the six largest national collections. There appear to be significant discrepancies between the 1996 FAO report and self-reporting by national collections (Table 1). The number of accessions held by India appears to be grossly over-estimated by the FAO report, while holdings of the United States (USDA-ARS, NGRP, GRIN, 2007) and China (Weidong et al., 2000) appear to be grossly under-estimated. It may be that the status of collections has changed significantly in the intervening ten years since the FAO report. Regardless of whether one uses the FAO report or self-reporting by individual public collections, data suggests that between 32,000 and 36,000 accessions of all species now reside in the world’s six largest collections. Data on the distribution of accessions among species are readily available for the USDA-ARS collection of the United States (USDA-ARS, NGRP, GRIN, 2007) and the CIRAD collection of France (Dessauw and Hau, 2007). These collections are probably typical in that the tetraploid, primary germplasm pool comprises 75.8 and 89.7 percent, respectively, of the two collections (Table 2). The proportion of tertiary germplasm, the most difficult source of genetic variation to access, is quite small in both collections –

comprising approximately 0.6% of each. The USDA-ARS and CIRAD collections differ in the proportion of the respective collections devoted to the secondary germplasm pool – specifically to the two A genome species *Gossypium arboreum* L. and *G. herbaceum* L. Due to acquisitions in recent years, 20% of the USDA-ARS collection is comprised of *G. arboreum* and *G. herbaceum* accessions - in comparison to 4% of the CIRAD collection. Otherwise, the proportions of the two collections devoted to the secondary germplasm pool are quite similar. Although it is reasonable to expect that the major portion of germplasm collections would be devoted to the most accessible primary germplasm pool, the secondary and tertiary germplasm pools are probably under-represented in world collections and potential genetic variability is being overlooked. Information sharing and comparison among collections reveal opportunities for germplasm exchange, broadening and balancing of genetic variability within collections, and added security of genetic resources by means of inter-collection redundancy.

Uses, problems, and challenges facing the USDA-ARS germplasm collection are probably typical of many collections worldwide, although other collections may face challenges not experienced by the USDA-ARS collection. Over the past twenty years the USDA-ARS collection has been fairly successful in acquiring materials through collecting trips and germplasm exchange programs (Wallace, 2007). In this time period, there have been eight collecting trips to the Galapagos Islands, Brazil, Mexico, and Australia; and germplasm exchange activities with India, China, Russia, and Uzbekistan. However, characterization and evaluation of germplasm lags seriously behind acquisition, and seed renewal in the collection has become an issue. In the past twenty years the seed renewal for the working collection has been extended from a five year schedule to a seven year schedule, and then to an “as needed” schedule prioritized by the curator. Although it is thought that the current seed renewal schedule

maintained by the curator is adequate, it probably is not optimal. Characterization of the USDA-ARS collection is primarily the responsibility of the collection curator, but these efforts have been supplemented by the other public researchers within the USDA, at universities, and at state experiment stations. Supplemental characterization efforts have generally been conducted on agronomically improved subsets of the collection, leaving the often more difficult characterization of the wild accessions of the primary germplasm pool, the secondary germplasm pool, and the tertiary germplasm pool to the curator. A standardized set of descriptors exists for the collection that appears with the catalog of accessions in the Germplasm Resource Information Network (USDA-ARS, NGRP, GRIN, 2007). A limitation of the descriptors used to characterize the germplasm is that they are primarily botanical or taxonomic in nature, whereas many of the descriptors that are of agronomic interest are lacking. Characterization of significant portions of the collection is incomplete; which sometimes limits the utility of the germplasm catalog. Despite the above deficiencies, over 4160 accessions were distributed to 105 researchers and institutions worldwide in 2006 (Frelichowski, 2007). There are ongoing efforts to fill the gaps in descriptor information in the collection.

Resources for evaluation of germplasm within the USDA-ARS collection have been limited, preventing large scale, systematic evaluation efforts. As a result, most evaluations have been upon smaller subsets of the collection, have been highly specific in their goals, and often have been conducted by individual scientists not affiliated with the collection. In the past ten years, evaluation efforts have included evaluation for: resistance to tobacco budworm, tolerance/resistance to cotton fleahopper, resistance to reniform nematode, resistance to root-knot nematode, etc (Table 3). Previous evaluations of the U.S and other collections have resulted in the discovery of numerous useful agronomic, pest resistance, and fiber quality traits, as can be

seen in Table 4. While individual, non-systematic evaluation efforts have resulted in the identification of numerous useful traits, the system results in duplication of efforts and poor reporting and recording of results. Specifically, no system exists for reporting results back to the central germplasm collection, and failure to report negative results perpetuates duplication of effort.

Other public genetic resources within the United States. Aside from the national germplasm collection, numerous genetic resources for the investigation and improvement of cotton have been created by public researchers in the U.S. These resources include recombinant inbred line (RIL) populations, back-cross introgression line (BIL) populations, near isogenic introgression line (NIIL) populations, chromosome substitution lines, day-neutral converted race stocks, etc (Wallace, 2007). A number of these populations have been developed for mapping, QTL investigations, and marker development and localization. Quite a few of the populations have been developed with specific goals in mind – such as fiber quality improvement. It is not surprising therefore that numerous of the RIL, BIL, and NIIL populations are *G. barbadense* introgressed *G. hirsutum* populations. Efforts are now underway to create *G. hirsutum* NIIL populations introgressed with *G. mustelinum*, and *G. hirsutum* alien chromosome substituted lines with *G. mustelinum* and *G. tomentosum* Nuttall ex Seemann as the donor species. A few of these resources, such as the *G. barbadense* chromosome substitution lines (Stelly et al., 2005), the *G. hirsutum* day-neutral converted race stocks (McCarty et al., 1993; McCarty et al., 2005), and a couple of the RIL populations (Percy et al., 2007) have been publicly released and are in the USDA-ARS national germplasm collection. However, the majority of these resources presently are held by their developers where they are undergoing evaluation.

CONCLUSIONS

With the increased recognition of the limited genetic diversity and the genetic vulnerability of elite cotton germplasm, there has come an intensified interest in conserving and evaluating the diversity of wild, commensal, and landrace cottons of the six tetraploid and forty-three diploid species of the *Gossypium* genus. However, much of the genetic variability to be found in *ex situ* cotton collections and *in situ* populations remains under-characterized and under-utilized in cotton improvement efforts. *In situ* preservation is a relatively new concept for germplasm preservation and faces many challenges in implementation. Genetic resources residing in germplasm collections of the USDA-ARS and other public and private research institutions or in public research programs represent a wealth of variability still to be tapped. Information sharing and comparison among collections reveal opportunities for germplasm exchange, broadening and balancing of genetic variability within collections, and added security of genetic resources by means of inter-collection redundancy. The utility of RIL, BIL, and NIL populations, chromosome substitution lines, day-neutral converted race stocks and other populations created for research purposes can be enhanced by coordinated collaborative research efforts.

REFERENCES

- 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.
- Brubaker, C.L., F.M. Bourland, and J.F. Wendel. 1999. The origin and domestication of cotton. p. 3-31. *In* C.W. Smith and J.T. Cothren (eds.) *Cotton: origin, history, technology, and production*. John Wiley & Sons, Inc., New York, NY.

- Barroso, P.A., C.E. Batisita, L.V. Hoffman, and A.Y. Ciampi. 2006. Genetic structure and *in situ* conservation of natural populations of *Gossypium mustelinum*. International Cotton Genome Initiative Res. Conf. Brasilia, Brazil.
- Dessauw, D., and B. Hau. 2007. Inventory and history of the CIRAD cotton (*Gossypium* spp.) germplasm collection. Plant Gen. Res. Newsletter 147:52-58 [Online]. Available at <http://www.ipgri.cgiar.org/pgnewsletter/archive.asp> (verified 23 May 2007).
- Frelichowski, J.E. 2007. U.S. National Germplasm Collection Update. Report to Cotton Germplasm Committee. Proc. Beltwide Cotton Conf., New Orleans, LA.
- Kearney, T.H., and R.H. Peebles. 1942. Flowering Plants and Ferns of Arizona. U.S. Print. Office, Washington, DC.
- May, O.L. 2000. Genetic variation for fiber quality. p. 183-229. In A.S. Basra (ed.) Cotton fibers – developmental biology, quality improvement, and textile processing. Food Products Press, New York.
- McCarty, J.C. Jr., and J.N. Jenkins. 1993. Seventy-nine Day-Neutral Primitive Cotton Germplasms. Crop Sci 33(2): 351.
- McCarty, J., and J. Jenkins. 2005. Twenty-one Day-Neutral Primitive Cotton Germplasms. Crop Sci 45(5): 2134.
- Percy, R.G., J. Zhang, and R. Cantrell. 2007. Notice of release of 96 introgressed recombinant inbred upland cotton germplasm lines. (ARS Release)
- Rubenstein, K.D., P. Heisey, R. Shoemaker, J. Sullivan, and G. Frisvold. 2005. Crop Genetic Resources. An Economic Appraisal. USDA-ERS. Economic Information Bulletin No. 2 [Online]. Available at <http://www.ers.usda.gov/Publications/EIB2/> (verified 23 May 2007).

- Singh, V.V., P. Mohan, V.N. Kulkarni, S.J.Baitule, and B.R. Pathak. 2003. Explorations within India for collection of cotton species germplasm. *Plant Gen. Res. Newsletter* 136: 40-46 [Online]. Available at <http://www.ipgri.cgiar.org/pgrnewsletter/archive.asp> (verified 23 May 2007).
- Stelly, D., S. Saha, D. Raska, J. Jenkins, J. McCarty, Jr., and O. Gutierrez. 2005. Registration of 17 upland (*Gossypium hirsutum*) cotton lines disomic for different *G. barbadense* chromosome or arm substitutions. *Crop Sci* 45(6): 2663.
- U.N. FAO. 1996a. Russia: Country report to the FAO international technical conference on plant genetic resources. FAO, Leipzig, Germany [Online]. Available at <http://www.fao.org/ag/AGP/AGPS/Pgrfa/pdf/urss.pdf> (verified 23 May 2007).
- U.N. FAO. 1996b. Report on the state of the world's plant genetic resources. FAO, Rome, Italy [Online]. Available at <http://www.fao.org/ag/AGP/AGPS/Pgrfa/pdf/SWRFULL2.PDF> (verified 23 May 2007).
- Ulloa, M., J.M. Stewart, E. Garcia, S. Godoy, A. Gaytan, and S. Acosta. 2006. Cotton genetic resources in the western states of Mexico: *in situ* conservation status and germplasm collection for *ex situ* preservation. *Genetic Res. and Crop Evol.* 53: 653-668.
- USDA-ARS-National Genetic Resources Program. 2007. Germplasm Resources Information Network (GRIN) database. Available at <http://www.ars-grin.gov/cgi-bin/npgs/html/taxfam.pl> (verified 23 May 2007).
- Wallace, T. (ed.) 2007. Cotton Germplasm Status Report and Crop Vulnerability Status. Germplasm Resources Information Network (GRIN) database. Available at http://www.ars-grin.gov/npgs/cgc_reports/cotton97.htm (verified 23 May 2007).

Weidong, G., J. Fang, D. Zheng, Y. Li, X. Lu, R. Rao, T. Hodgkin, and Z. Zongwen. 2000.

Utilization of germplasm conserved in Chinese national genebanks – a survey. *Plant*

Gen. Res. Newsletter 123:1-8 [Online]. Available at

<http://www.ipgri.cgiar.org/pgnewsletter/archive.asp> (verified 23 May 2007).

Table 1. Size of the seven largest ex situ cotton germplasm collections of the world, by percentage and accession number.

Germplasm accessions of all species			
Country	FAO Report, 1996		Reported in the literature
	%	(est. no)^z	no.
India	34	(16,660)	4,500^y
France	13	(6,370)	3,070^y
Russian Federation	12	(5,880)	6,300^x
United States	6	(2,940)	9,303^w
Pakistan	5	(2,450)	NA
China	3	(1,470)	6,724^v
Brazil	NA	NA	2,835^y
total	73	(35,770)	32,732
world total	100	49,000	?

^z Estimated by calculating the number of accessions equivalent to the FAO reported percentage when 49,000 is the total number.

^y Dessau and Hau, 2007

^x U.N. FAO, 1996a

^w USDA-ARS, NGRP, GRIN, 2007

^v Weidong et al., 2000

Table 2. *Gossypium* germplasm held in the USDA-ARS collection of the United States and the CIRAD collection of France, 2007.

Germplasm pool	Genomes	USDA-ARS		CIRAD	
		Number of species	Number of accessions	Number of species	Number of accessions
1 ⁰ primary	AD	5	7050	5	2754
2 ⁰ secondary	A	2	1923	2	119
	B	3	10	2	23
	D	13	258	12	74
	F	1	4	1	2
total		19	2195	17	218
3 ⁰ tertiary	C	2	16	2	4
	G	3	20	3	6
	(K)	5	9	-	-
	E	4	13	4	9
total		14	58	9	19
Collection Total		38	9303	31	3070

Table 3. A partial list of evaluation efforts using USDA-ARS germplasm collection resources, conducted over the past ten years.

Evaluation objective	Institution	Species	Accession number and type
resistance to tobacco budworm	USDA-ARS, CSRL, Mississippi	<i>G. hirsutum</i>	168 day-neutral lines
tolerance/resistance to cotton fleahopper	TAES, Texas A&M University, Texas	<i>G. hirsutum</i>	110 day-neutral converted race stocks
resistance to reniform nematode	AAES, Auburn University, Alabama	<i>G. hirsutum</i>	2,500 accessions
resistance to root-knot nematode	Louisiana State University, Louisiana	<i>G. hirsutum</i>	150 accessions
resistance to <i>Rhizoctonia solani</i>	TAES, Texas A&M University, Texas	<i>G. hirsutum</i>	30 high tannin lines, 100 converted race stocks
resistance to root-knot nematode	USDA-ARS, CSRL, Mississippi	<i>G. hirsutum</i>	79 day-neutral lines
resistance to reniform nematode	USDA-ARS, CSRL, Mississippi	<i>G. arboreum</i>	8 accessions
		<i>G. hirsutum</i>	3 day-neutral converted race stocks
resistance to reniform nematode	USDA-ARS, CPRU, Texas	<i>G. hirsutum</i>	1419 primitive accessions
		<i>G. barbadense</i>	850 accessions
heat stress tolerance	AAES, Auburn University, Alabama	<i>G. hirsutum</i>	1380 accessions
seedling drought and salt tolerance	TAES, Texas A&M University, Texas	<i>G. hirsutum</i>	100 day-neutral converted race stocks and obsolete cultivars
agronomic traits and fiber quality	USDA-ARS, CSRL, Mississippi	<i>G. hirsutum</i>	151 day-neutral accessions
agronomic traits and fiber quality	TAES, Texas A&M University, Texas	<i>G. hirsutum</i>	80 day-neutral converted race stocks

Table 4. Partial list of useful traits identified in various germplasm pools.

Germplasm pool	Genome – species	Trait
1⁰	AD₁ – <i>G. hirsutum</i>	Resistance to:
		Bacterial blight
		Root-knot nematode
		Fusarium wilt
		Verticillium wilt
		Spidermite
		Boll/Bud worm
		Boll weevil
		Morphological traits:
		okra-leaf
	glandless	
	frego bract	
	AD₂ – <i>G. barbadense</i>	Bacterial blight resistance
		Thrips tolerance
		Fusarium wilt resistance
		Fiber quality traits
	AD₃ – <i>G. tomentosum</i>	Nectariless
2⁰	A₂ – <i>G. arboreum</i>	Southwestern cotton rust resistance
	B₁ – <i>G. anomalum</i>	Bacterial blight resistance
	D₁ – <i>G. thurberi</i>	Fiber strength
	D₂₋₁ – <i>G. armourianum</i>	Bacterial blight resistance
	D₂₋₂ – <i>G. harknessii</i>	Cytoplasmic male sterility – restorer system
	F₁ – <i>G. longicalyx</i>	Reniform nematode resistance
3⁰	C₁ – <i>G. stuartianum</i>	Terpenoid aldehydes absent in seed
	C₂ – <i>G. robinsonii</i>	
	G – <i>G. australe</i>	
	G – <i>nelsonii</i>	
	G₁ – <i>G. bickii</i>	