# Exploiting Genetic Diversity

**Conference Paper** · September 2007

CITATIONS
24

READS
708

**1 author:**

Ibrokhim Y Abdurakhmonov
Ministry of Innovational Development
**130** PUBLICATIONS **2,044** CITATIONS

SEE PROFILE

**TITLE**: **Exploiting Genetic Diversity**

**DISCIPLINE**: Genetics &Breeding, Molecular Biology &Physiology

**AUTHORS**: Ibrokhim Y. Abdurakhmonov

Laboratory of genetic engineering and biotechnology, Institute of Genetics and Plant Experimental Biology, Academy of Sciences of Uzbekistan. Yuqori Yuz, Qibray region Tashkent district, 702151 Uzbekistan. Phone: (99871) 1629965; Fax:(998712) 642230. Correspondence (IYA: *ibrokhim_a@yahoo.com*);

**ACKNOWLEDGMENTS**: I thank the Academy of Sciences of Uzbekistan and the Office of International Research Programs (OIRP) of USDA-ARS for continual funding support of cotton genomics research in Uzbekistan. I would like to thank Prof. Abdukarimov Abdukarimov, Prof. Abdumalyan Abdullaev, Dr. Sofiya M. Rizaeva, Dr. Sukumar Saha and Mr. Ilkhom Salakhutdinov for the useful suggestions during manuscript preparation. Author apologizes for not citing many other important papers of colleagues in this short overview.

**ABBREVIATIONS**: RAPD (Random Amplified Polymorphism); RLFLP (Restricted Fragment Length Polymorphism); AFLP (Amplified Fragment Length Polymorphism) SSR (Simple Sequence Repeat); LD (Linkage Disequilibrium); GD (Genetic distance); MLM (Mixed Linear Model); WGA (Whole Genome Association).

## ABSTRACT

The level of genetic diversity of crop species is an essential element of sustainable crop production. Cotton productivity and the future of cotton breeding efforts also depend on genetic diversity of cotton gene pools. The amplitude of genetic diversity of cotton (*Gossypium* species) is exclusively wide, but cotton researchers and producers are concerned with the narrow genetic base of cultivated cotton germplasm that caused recent cotton yield and quality declines. These declines, however, is largely due to challenges and the lack of innovative tools to effectively exploit genetic diversity of *Gossypium* species. The most effective utilization of genetic diversity of *Gossypium* species further requires modern genomics technologies that reveal the molecular basis of genetic variations of agronomic importance. Ongoing efforts toward sequencing cotton genome (s) are the pivotal step that will facilitate the fine-scale mapping and better utilization of functionally significant variations in cotton gene pools. Here, we briefly review the status of worldwide cotton germplasm collections, the genetic base of cotton species, the examples and challenges of exploiting genetic diversity, genomics resources &efforts, and future perspectives in effective exploiting genetic diversity in cotton.



**Key words**: *Gossypium* genus, cotton genetic diversity, molecular markers, cotton genomics resource

## INTRODUCTION

Genetic diversity generally refers to different ranges or variants of morphophysiological patterns and properties among individuals of a population – often referred as traits, which is imperative for both positive respond to the rapid environmental changes and consequent survival of biological species. The lack of genetic diversity or its narrowness in the crop species creates a potential threat to sustain crop productivity due to rapid vulnerability of genetically uniform cultivars by potentially new biotic and abiotic stresses. Therefore, wider genetic diversity of crop species ensures a potential to protect crops from massive new pathogen and pest epidemics and sudden global environment changes as well as provides an opportunity to further improve complex traits of interest by combining or pyramiding genetic variations within population(s) (Van Esbroeck et al., 1998; Van Esbroeck et al., 1999). The narrow genetic diversity is usually caused by extensive use of one or more closely related genotypes in breeding programs (Van Esbroeck et al. 1999) or 'genetic bottleneck' through historic crop domestication that selected out a few genotypes to spread (Iqbal et al., 2001). Cotton productivity and the future of cotton breeding efforts, as in many other agronomic crops, also depend on genetic diversity of cotton gene pools. Worldwide cotton researchers and producers are concerned with the narrow genetic basis of the cultivated cotton germplasm that caused recent cotton yield and quality declines (Cotton Incorporated, USA). The existing and predicted problems of worldwide cotton breeding programs associated with narrowness of the genetic basis of the cultivar germplasm, caused by genetic bottleneck through historic domestication events and selection (Iqbal et al., 2001), is largely due to challenges and the lack of innovative tools to mobilize the useful genetic variations from diverse exotic cotton species of *Gossypium* genus to the breeding cultivars.

Although wild cottons (*Gossypium* species) are perennial shrubs and trees, the domesticated cottons are tropically and sub-tropically annual crops cultivated since prehistoric times of the development of human civilization. The genus *Gossypium* includes approximately 46 diploid and 5

allotetraploid species (Percival et al., 1999) that are largely spread in tropical and subtropical regions of the world. Diploid cottons, referred as Old World cottons, are classified into eight (A-G to K) cytogenetically defined genome groups that have African/Asian, American, and Australian origin (Endrizzi et al., 1985). Two of these Old World cottons from Asian origin, *G. arboreum* and *G. herbaceum*, with a spinnable seed fiber, were originally cultivated in Asian continent. Hybridization between A-genome (Asian cottons) and D-genome (American cottons) diploids and subsequent polypoidization about 1.5 million years ago created the five AD allotetraploid lineages that are indigenous to America and Hawaii (Phillips, 1963; Wendel and Albert, 1992; Adams et al., 2004). These New World allotetraploid cottons include the commercially important species, *G. hirsutum* and *G. barbadense*, which are widely cultivated worldwide.

*G. hirsutum* is the most widely cultivated (90%) and industrial cotton among all *Gossypium* species. *G. hirsutum* includes the Upland cotton and other early maturing, annually grown herbal bushes. Guatemala is considered the center of origin for *G. hirsutum*, but it spread throughout Central America and Caribbean. According to archaeobotanical findings, *G hirsutum* probably was domesticated originally within the Southern end of Mesoamerican gene pool (Wendel, 1995; Brubaker et al., 1999). Consequently, two centers of genetic diversity exist within *G. hirsutum*: Southern Mexico-Guatemala and Caribbean (Brubaker et al., 1999); Mexico-Guatemala gene pool is considered site of original domestication and primary center of diversity. Within this range, *G. hirsutum* exhibits diverse types of morphological forms, including wild, primitive to domesticated accessions. According to Mauer (1954), there are four groups of sub-species of *G. hirsutum*: 1) *G. hirsutum* ssp. mexicanum, 2) *G. hirsutum* ssp. paniculatum, 3) *G. hirsutum* ssp. punctatum, and 4) *G. hirsutum* ssp. euhirsutum (domesticated cultivars). These four groups of sub-species include itself a number of wild landraces such as yucatanense, richmondi, punctatum, latifolium, palmeri, morilli, purpurascens and their accessions as well as a number of domesticated variety accessions from 80 different cotton growing countries worldwide.

The wide-distribution of *G. barbadense* included mostly South America, southern Mesoamerica and the Caribbean basin (Fryxell, 1979). *G. barbadense*, accounting for about 9% of world cotton production, was originally cultivated in costal islands and lowland of the USA that became known as Sea Island cotton. Sea Island cottons, then, were introduced into Nile Valley of Egypt and widely grown as Egyptian cotton to produce long staple fine fibers (Abdalla et al., 2001). The other three AD tetraploid species of cotton, *G. mustelinum* with specific distribution in the Northeast Brazil (Wendel et al., 1994), *G. darwinii* endemic to Galapagos Islands (Wendel and Percy, 1990), and *G. tomentosum* Nutall ex Seemann endemic to Hawaiian Islands (DeJoode and Wendel, 1992; Hawkins et al., 2005), are truly wild species (Westengen et al., 2005).

Thus, the *Gossypium* genus, encompassing wide geographic and ecological niches, have a large amplitude of morphobiological and genetic diversity that conserved *in situ* at centers for cotton origin (Ulloa et al. 2006) and preserved *ex situ* within worldwide cotton germplasm collections and materials of breeding programs. Once exploited effectively, these wide ranges of genetic diversity of the genus, in particular reservoir of potentially underutilized genetic diversity in exotic wild cotton germplasms, are the 'golden' resources to improve cotton cultivars and solve many fundamental problems associated with fiber quality, resistance to insects and pathogens and tolerance to abiotic stresses. Here, we briefly review the efforts on cotton germplasm collections worldwide, the genetic base of cotton species, describe some examples of the morphobiological and genetic diversity within the Upland cotton cultivars and landrace stock accessions, worldwide efforts of utilization of genetic diversity in solving of breeding problems and improvement of cotton production with its potential challenges, and perspectives of new innovative genomic technologies to better exploiting the natural variations preserved within cotton germplasm collections.

**Worldwide cotton germplasm collections**

There are the main cotton germplasm collections preserved *ex situ* in Australia, China, India, France, Mexico, USA and Uzbekistan. The history of collecting an initial cotton germplasm through

the specific expeditions of cotton scientists to the centers of *Gossypium* origins are well described by Ulloa et al. (2006) that were the basis, perhaps, for the majority of the current cotton germplasm collections worldwide. Consequently, cotton germplasm collections worldwide were enriched with a numerous cotton germplasm accessions and breeding materials/lines as source of the genetic diversity through continuous research efforts of specific cotton breeding programs and mutual germplasm exchange.

The brief descriptions for some of worldwide cotton germplasm collections were highlighted in 'White Paper' for cotton genome sequencing (available at http://algodon.tamu.edu/sequencing/, verified on June 28, 2007). Briefly, Australian cotton germplasm collection includes 2000 *G. hirsutum* accessions and 500 accessions of 17 Australian cotton species that are maintained by Australian Commonwealth Scientific and Research Organization (CSIRO). In India, the collection of cotton germplasm includes 5980 *G. hirsutum*, 1049 *G. barbadense* 1867 *G. arboreum*, 568 *G. herbaceum* and 173 inter-genomic hybrids, which are being preserved at several cotton research institutions (C. D. Mayee and V. V. Singh, unpublished). The Central Institute for Cotton Research (CICR) at Nagpur, the responsible organization for the national germplasm of cotton, preserves a total of 3,760 accessions of *G. hirsutum*, 289 of *G. barbadense*, 1,455 of *G. herbaceum*, 24 wild species and 21 other perennials that are being maintained and evaluated for utilization in cotton improvement programs *(http://www.bioversityinternational.org/publications/Web_version/174/ch11.htm,* verified on August 4, 2007). Besides, the wild species, perennials and cytogenetic stocks have been conserved vegitatively in species garden of CICR at Nagpur as well as at many other cotton research centers such as Akola, Gujrat Agricultural University, Surat and CICR, Tamil Nadu Agricultural University (C. D. Mayee and V. V. Singh, unpublished). In France, French Agricultural Research Center for International Development (CIRAD) preserves about 3200 *G. hirsutum* cotton accessions, including 1000 improved cultivars and 1000 perennial wild accessions.

There is a large number of cotton germplasm collection preserved under the Southern Plains Agricultural research Center (SPARC), Crop germplasm Unit of USDA-ARS at College Station Texas, USA with its cotton winter nursery at Mexico for increasing of long seasoned, photoperiodic germplasm accessions. The collection contains a total of 9332 accessions representing 49 species from 74 countries. In that, the working cotton collection (responsible for germplasm characterization and documentation, enrichment, and coordination with breeding programs) maintains 3132 *G. hirsutum* accessions, 2120 landrace stock germplasm of *G. hirsutum*, 1585 *G. barbadense*, 1730 *G. herbaceum* and 194 *G. arboreum*, and germplasm of 581 wild cottons representing 40 species of wild diploids and allotetraploids. Besides, there are "Base and Ad Hoc" cotton collections in the USA (details can be found at http://www.ars-grin.gov/npgs/cgc_reports/cottonstatus2005.pdf, verified on August 4, 2007). Recently, the US cotton germplasm resources were enriched with newly collected 151 *Gossypium* accessions, comprising of the seven known and one not described taxa (Ulloa et al., 2006).

In Uzbekistan, there are over 17,000 cotton germplasm accessions including isogenic, inbred lines, recombinant inbred lines (RIL), elite AD allotetraploid varieties (*G. hirsutum* and *G. barbadense*), monosomic and translocation lines (A. Abdullaev, personal communication; Abdukarimov et al. 2003) along with wild, primitive and extant representatives of the A- to K-genome groups that have been developed in the Cotton Research Institutes of the Republic of Uzbekistan and collected over the world for the past century. The initiatives of collecting germplasm resources of cotton were made by Dr. N. I. Vavilov and F. M. Mauer in 1930, and the efforts successfully further continued by Dr. A. Abdullaev's group at the Academy of Sciences of Uzbekistan. During the past 50 years period, because of huge investment of Uzbekistan government, the efforts and multiple-time scientific expeditions to Australia, China, India, Mexico, Pakistan, Peru, and Sri-Lanka headed by Dr. Abdullaev, as well as through the collaborative efforts of germplasm exchange worldwide, the Uzbek cotton germplasm collection was effectively enriched

with wide range of worldwide genetic diversity of *Gossypium* genus. The majority of *G. hirsutum* wild and cultivar accessions (~12,000) are preserved at Cotton Breeding Institute of Agriculture Ministry of Uzbekistan. The remaining nearly 6,500 cotton accessions are preserved within the Institute of Genetics and Plant Experimental Biology at the Academy of Sciences of Uzbekistan. This 'exotic' sub-collection of cotton germplasm resources includes more than 40 A- to K-genome wild *Gossypium* species (also vegitatively conserved at germplasm Unit greenhouse), 4,500 unique and "exotic" *G. hirsutum* accessions, 900 *G. barbadense* accessions, 400 *G. arboreum* accessions (A. Abdullaev, personal communication), 200 photoperiod converted radiomutants of photoperiodic allotetraploid species (Djaniqulov, 2002; Abdurakhmonov et al., 2007), and a number of intra- and interspecific hybrids (Rizaeva, 1996). Additionally, collection of more than 500 'unique' genetic stocks of *G. hirsutum*, consisting of collection of fiber mutants and their RIL populations and isogenic lines segregating with many morphological traits (Musaev et al., 2000), and cytogenetic stocks (Musaev et al., 2000; Sanamyan and Rakhmatulina, 2003) are being maintained at the National University of Uzbekistan at Tashkent. These worldwide germplasm resources constitute a vast potential resource of genes for agronomically important traits, such as insect and pathogen resistance, tolerance to environmental stresses, fiber quality (length, strength and lint yield) and yield potential, and serve as fundamental base for both traditional and molecular cotton research programs.

**Genetic diversity revealed by molecular marker analyses**

A number of studies on the genetic diversity of *Gossypium* species revealed a low level of genetic diversity within Upland cultivars inferred from isozymes (Wendel and Percy, 1990; Wendel et al., 1992), random amplified polymorphisms – RAPDs (Multani and Lyyon, 1995; Tatineni et al., 1996; Iqbal et al., 1997), restricted fragment length polymorphisms – RFLPs (Wendel and Brubaker, 1993), amplified fragment length polymorphisms – AFLPs (Pilay and Myers, 1999; Abdalla et al., 2001; Iqbal et al., 2001; Rana et al., 2005) and Simple Sequence Repeats – SSRs (Liu et al., 2000a,

Gutierrez et al., 2002; Rungis et al., 2005; Zhang et al., 2005a). Although there is little variation in estimation of genetic diversity among Upland cultivars, in general, the genetic distance reported for Upland cultivars was in the range of 0.01-0.28. Recently, we analyzed a large number of *G. hirsutum* variety and exotic accessions from Uzbek cotton germplasm collection with SSR markers (Abdurakhmonov et al. 2007b; Abdurakhmonov et al., unpublished; see Fig.1 for an example of phylogenetic analysis with some ecotypes). Our results, obtained from phylogenetic analysis of a total number of 620 of *G. hirsutum* accessions from exotic germplasm and diverse ecotypes/breeding programs confirmed the narrow genetic base of Upland cotton breeding germplasm pool (with the genetic distance (GD) range of 0.005-0.26) and provided an evidence for the occurrence of a genetic 'bottleneck' during domestication events of the Upland cultivars in molecular level (Iqbal et al., 2001). We demonstrated existence of wide genetic diversity within the exotic germplasm (GD=0.02-0.50). Previous studies (Liu et al., 2000a, Lacape et al., 2006) also found wider genetic diversity in the land race stocks of *G. hirsutum*, suggesting the existence of sufficient genetic diversity in the exotic germplasm for future breeding programs. A wider genetic diversity (30-87%) within *G. hirsutum* breeding lines revealed by AFLP markers was also reported (Rana et al., 2005), suggesting the useful genetic diversity to broaden the genetic base of Upland cotton germplasm.

The molecular genetic diversity within *G. barbadense* germplasm accessions was also studied using molecular markers such as allozyme (Wendel and Percy, 1990), and AFLPs (Abdalla et al., 2001; Westengen et al., 2005) that revealed a narrow genetic base within *G. barbadense* accessions with the narrow genetic distance of 7-11% (Abdalla et al., 2001; Westengen et al., 2005) as was observed within the Upland cotton germplasm. The genetic diversity was also low within *G. tomentosum* germplasm with the genetic distance range of 2-11% (Hawkins et al., 2005).

The genomic diversity of the A- genome diploid cottons has also been studied using molecular marker technology (Liu et al., 2006; Kebede et al., 2007). The genetic distance within 39

*G. arboreum* L ($A_2A_2$-genome) accessions, analyzed with SSR markers, ranged from 0.13-0.42 (Liu et al., 2006) demonstrating the existence of wider genomic diversity in the A-genome diploids compared to the Upland cultivar germplasm. Kebede et al. (2007) reported, however, moderate level of genetic diversity within each $A_1$ and $A_2$- genome cottons that ranged from 0.03-0.20 with an average of 0.11 within *G. herbaceum* and 0.02-0.18 with an average of 0.11 for *G. arboreum* ($A_2$). The overall genetic distance between $A_1$ and $A_2$ genomes was up to 38 % (Kebede et al., 2007) demonstrating moderately wider genetic diversity than those reported for the AD cotton germplasm reviewed herein. A wider range of genetic diversity was observed among the D-genome diploid cottons with the genetic similarity of 0.08-0.94 (Guo et al. 2007a), suggesting usefulness of utilization of D-genome cottons in the breeding programs.

The genetic diversity among the AD allotetrapoids and A-and D-genomes also were characterized in many studies using various markers systems. AFLP marker analyses studies (Iqbal et al., 2001, Abdalla et al., 2001, Westengen et al., 2005) revealed that the genetic distance between *G. barbadense* and *G. hirsutum* was in a range of 21-33%. The other wild AD tetraploids (*G. mustelinum* and *G. tomentosum*) were close to the cultivated AD cottons sharing 75-84% similarity, where *G. tomentosum* was closer to *G. hirsutum* genome (GD=0.16) than the other allotetraploid species (Westengen et al., 2005). The genetic distance between the widely cultivated AD cottons (*G. barbadense* and *G. hirsutum*) and A-genome diploids varied from 45 to 69%, and the cultivated AD cottons and the D-genome varied from 55 to 71%. The GD between the wild AD tetraploids and the A-genome was in a range of 46-52%, and between the wild AD cottons and the D-genome was 58-59%. The genetic distance between the A- and D-genome cottons, in general, was in a range of 0.72-0.82 when analyzed with AFLPs (Iqbal et al., 2001, Abdalla et al., 2001, Westengen et al., 2005).

The use of SSR markers revealed that the GD between *G. hirsutum* and *G. barbadense* was in a range of 42-54% (Kebede et al., 2007; Abdurakhmonov et al., 2007b, unpublished). However, Lacape et al. (2006) reported higher genome dissimilarity values (D=0.89-0.91%) between *G.*

*hirsutum* and *G. barbadense* within their material. Also, high mean dissimilarity values were reported between *G. hirsutum* and *G. tomentosum* (D=0.71-0.75) and between *G. barbadense* and *G. tomentosum* (D=0.80) using highly polymorphic sets of SSRs (Lacape et al., 2006). The genetic distance among the AD tetraploids was also in a range of 0.80-0.88 (Liu et al., 2000a) with moderate closeness of *G. tomentosum* to the Upland cotton than *G. barbadense* cultivars that also was supported by other studies with different marker system (Dejoode and Wendel, 1992; Hawkins et al., 2005). The GD between the cultivated AD cottons and the A-genome was in a range of 31-43%, and GD between the cultivated AD cottons and the D-genome was in a range of 35-46% (Kebede et al., 2007). The GD between A-and D-genome cottons varied in a range of 29-42% (Kebede et al., 2007).

Thus, the results of genetic distance studies among cotton species reveal low level of genomic diversity within cultivated AD cottons; however, there are useful wide genetic diversity available within exotic land race stocks, wild AD cottons, and a putative A-and D-genome ancestors to AD cottons that have potential to search genetic gains and can be utilized in future improvement of cotton. Below we give some examples of biodiversity in cotton.

**The examples of morphobiological peculiarities of cotton**

The amplitude of genetic diversity of cotton (*Gossypium* ssp), including all its morphological, physiological and agronomic properties, is exclusively wide (Mauer, 1954). Knowing the details of genetic diversity is also very important to determine timeframe of cotton agronomy, develop a strategy for genetic gains in breeding, and conserve existing gene pools of cotton. There are a great deal of genetic diversity in *Gossypium* genus within characteristics such as plant architectures, stem pubescence and color, leaf plate shape, flower color, pollen color, boll shape, fiber quality, yield potential, early maturity, photoperiod dependency, and resistance to multi-adversity environmental stresses that are important for the applied breeding of cotton. However, it is difficult to exemplify the detailed description of morphobiological and agronomic diversities of

entire cotton genus in this short overview. Here, we give some examples of morphobiological diversity observed within the Upland cotton germplasm maintained in the Uzbek cotton germplasm resources to demonstrate existence of useful genetic diversity within the 'narrow-based' Upland cotton germplasm and its wild landrace stocks. This will give a glance off the richness of cotton with various useful biodiversities.

Plant architectures and branching patterns vary within *G. hirsutum*. There are creepy and vertical stem types having compact and branchy types of bushes (Fig. 2). Plant stem also varies with the antocyanin color (Fig. 3) and the pubescence (Fig. 4). Leaf plate shape diversity is also wide. A various types of leaf plate architecture can be observed in *G. hirsutum* accessions including variations of different three-lobbed, five-lobbed, palmate, digitate, semi-digitate, and okra leave types (Fig. 5). Different flower color variations can be observed such as white, light yellow, yellow, cream-color, lavender, pink and bicolor (Fig.6). Likewise, there is wide diversity in pollen color (Fig. 7) and boll shape (Fig. 8). In the analyses of ~1000 *G. hirsutum* exotic and cultivated accessions in the two different environments, Mexico and Uzbekistan, we found a wide range of useful agronomic diversities (Abdurakhmonov et al., unpublished). In one or two environments, the cotton boll mass varies in a range of 1-9 grams /per boll, 1000 seed mass varies in a range of 50-170 grams, the lint percentage varies in a range of 0-45%, micronaire varies in a range of 3-7 mic, the fiber length varies in a range of 1-1.28 inch, and fiber strength varies in a range of 26-36 g/tex. There were also a wide range of variation on photoperiodic flowering (day neutral, week to strong photoperiodic dependency) and maturity (data not shown, Abdurakhmonov et al., unpublished). This wide phenotypic diversity of cotton shows the exclusive plasticity of cotton plants and potential of their wide utilization in the breeding programs as an initial material.

**Exploiting of genetic diversity through traditional approaches**

In general, cotton breeding programs worldwide are built on the utilization of above-mentioned genetic diversity of cotton through genetic gains of combining variations and further

application of mutational breeding tools to enhance useful diversity or remove undesired variations from the breeding material(s). In the consequences of utilization of the existing natural variations, a large collection of superior quality cotton cultivars, resistant to various multi-adversity stresses, were developed by worldwide breeding programs to make the cotton profitable in the specific growing environments. All elite commercialized cultivars are the examples of effective exploiting the genetic diversity, which are difficult to cover all within this overview. Hence, below we give a couple of examples of introduction of the useful diversities from the wild cotton germplasm using traditional breeding approaches that had a great impact in cotton production, at least, in some cotton growing countries.

In Uzbekistan (a second cotton fiber exporter), cotton is the major economy source that nowadays produces annually ~ 3.5-4 million tons of seed cotton fibers and exports~$900 million worth cotton fiber. In 1960's *Verticillium* wilt fungi began widely attacking the cotton plantations and fungi epidemics caused ~16-20% of crop yield lost at that time. In other words, cotton production (estimated production at that time was 5-6 million tons/year) lost up to 1 million tons of seed cotton fibers annually (Dr. A. Abdulalev, a curator of cotton germplasm in Uzbekistan, personal communication). Therefore, the development of wilt resistant varieties became a priority for Uzbekistan, and all efforts in cotton breeding programs were directed to solve that problem. The search for genetic diversity from the exotic *G. hirsutum* germplasm, and on-time mobilization of new wilt resistant genes of the exotic germplasm into the elite cultivars saved country's cotton production, so and economy of the country. Because of the tremendous efforts of group of scientist headed by academician S. M. Mirakhmedov, the wilt resistant variety series named as "Tashkent" were developed by genetic hybridization of an early – maturing superior quality variety C-4727 with the wilt resistant wild accessions of *G. hirsutum* (*G. hirsutum* ssp mexicanum var nervosum) and backcrossing it with C-4727. Developed hybrids were extensively selected in the background of wilt fungi for multiple-years and highly wilt resistant Tashkent –1, Tashkent-2, and Tashkent-3 varieties

were developed, and immediately released for the commercial use in 1971. These varieties were highly wilt disease resistant, early maturing (less than 120 days) and productive. In 1988, using extensive individual selection among hybrids from the cross of C-4727 and *G. hirsutum* spp mexicanum, the early-maturing Tashkent-6 variety with improved fiber quality was developed and commercialized within the Republic. As a continuation of a 'genetic diversity imprint' introgressed from the wild landrace stock of *G. hirsutum* to the improved Tashkent cultivar series, in 1980's, a new highly salt tolerant variety, AN-Boyovut-2, was developed from the multiple-years of selection of Tashkent variety biotypes in the saline lands of Uzbekistan (Abdullaev, personal communication, Abdukarimov et al., 2003). Tashkent variety series and it's many 'sister-varieties', like AN-Boyovut-2, are the basis of today's cotton breeding programs and production in Uzbekistan, and one of the success story in the utilization of genetic diversity and its impact from the single landrace stock germplasm, *G. hirsutum* ssp. mexicanum.

Another interesting example of the utilization of wild germplasm is the development of naturally early leaf defoliating Upland lines. Scientists have developed a naturally early leaf defoliating Upland cotton line named "Listopad Beliy" (white leaf defoliant, AADD, 2n=2x=52), from tri-species crosses between *G. thurberi* ($D_1D_1$, 2n=2x=26, natural defoliant) x *G. harknessii* ($D_2D_2$, 2n=2x=26, wilt resistant) x *G. hirsutum* (variety Tashkent-1, AADD, 2n = 2x = 52), using an integrated approach of chromosome doubling via colchicine and conventional breeding methods at the Institute of Genetics and Plants Experimental Biology (IG&PEB), Uzbekistan (Rizaeva et al., 2001). This line, with earliness, drought- and wilt-resistance and a natural leaf defoliation trait, has been used as parental donor material in breeding programs to introgress the natural leaf defoliation trait into many commercial varieties in Uzbekistan (Abdurakhmonov et al., 2005). Many such an example of utilization of genetic diversity in solving disease and pest resistance, tolerance to multi-adversity stresses, seed oil content and fiber quality in Uzbekistan were reviewed by Abdukarimov et al. (2003). In India, there are several examples of the development of cotton cultivars using exotic

germplasm resources that solved problems with Jassid resistance (B 1007, Khandwa-1, and Khandwa-2 varieties) introgressed from *G. tomentosum*, drought resistance (Deviraj variety) introgressed from *G. arboreum*, *Verticillium* tolerance (VRS-7 variety) introgressed from *G.hirsutum* ssp. mexicanum, and bacterial blight resistance (Arogya variety) introgressed from *G. anomalum* (Mayee and Singh, unpublished). There are many such examples in the other cotton breeding programs worldwide (for more examples, refer to Dr. Stelly et al., 2007 in this conference).

**Challenges with utilization of diversity of wild germplasm**

The utilization of useful genetic diversity of the wild germplasm using traditional breeding efforts is challenging. For instance, even within *G. hirsutum* cotton germplasm, the majority of wild and primitive accessions with useful genetic diversity is photoperiod-sensitive, short-day plants that never flower in long-day conditions of summer cultivations, making exotic cotton germplasm largely under-utilized in crossing programs. Introgression of day-neutral genes into wild cotton germplasm or conversion of photoperiodic wild and primitive cottons to day-neutral types is, therefore, of particular interest of cotton breeders for effective utilization of 'exotic' germplasm in introgression of potential genetic diversity into the breeding gene pool. There are some traditional breeding approaches available to overcome this challenge. Successful photoperiodic conversion programs in cotton have been developed to mobilize day-neutral genes into the primitive accessions of *G. hirsutum*, where day-neutral genes have been introgressed into 97 primitive cotton accessions by a large backcrossing effort (McCarty et al., 1979; McCarty and Jenkins, 1993, Liu et al., 2000a). This converted cotton germplasm is an important reservoir for potential genetic diversity and can be used as a source to introgress genes into breeding germplasm. The day-neutral versions of the other allotetraploid cottons are not available (Liu et al., 2000a). One of the effective alternative approaches in directly converting photoperiodic wild cotton races to day-neutral versions is the using of induced mutation. This has produced photoperiod converted mutant cotton germplasm in Uzbekistan by converting wild cottons directly into day-neutral plants and a number of commercial

cotton varieties such as AN-401, AN-402 and Kupaysin were released for the cotton farmers that are day-neutral cultivars with superior fiber and other agronomic quality in Uzbekistan (Abdurakhmonov et al., 2007c). Although these suggest the existence of useful approaches to partially overcome problems with photoperiodic flowering, those approaches and tools require tremendous efforts and investment.

Additionally, there are many other challenges due to 1) hybridization issues between various cotton genomes, 2) sterility issues of interspecific multi-genome hybrids, 3) photoperiodic flowering of wild cottons, as mentioned above examples, and 3) long timescale (10-12 years of efforts) required for successful introgression and recovering superior quality homozygous genotypes using traditional breeding approaches. This underlies necessity for development of new innovative genomics approaches to support and accelerate the traditional efforts of exploiting the genetic diversity in cotton breeding. The most effective utilization of the genetic diversity of *Gossypium* species further requires 1) characterization of candidate gene(s) underlying the phenotypic and agronomic diversities based on genomic information in other species, 2) estimation of genetic distances, geneology and phylogeny of gene pools, 3) acceleration of linkage mapping and markers assisted selection, 4) development cotton transgenomics, and 5) sequencing cotton genome(s). Furthermore, it is very important to characterize and describe the exiting cotton germplasm collections of Australia, China, India, France, Mexico, USA, and Uzbekistan both phenotypic and genomic level. Consequently, incorporation of information into electronic web-based cotton databases such as cotton DB (http://cottondb.tamu.edu/), Cotton Portal (http://gossypium.info), and the Cotton Diversity Database (http://cotton.agtec.uga.edu; Gingle et al., 2006) as well as further improvement of data management tools are pivotal to facilitate an effective exploiting the genetic diversity of cotton in the future. Cotton germplasm exchange among collections and research groups is also imperative part toward this goal. Since efforts of cotton research community on above-highlighted points will be discussed in other plenary talks in detail (refer the program – T. Zhang,

2007; T. Wilkins, 2007; A. Patterson, 2007; in this conference), below we briefly describe overall efforts on cotton genomics with some examples of investigation from our laboratory.

**Cotton genomics efforts in the utilization of genetic diversity in cotton**

During past years, international cotton research community has developed extensive genomic resources, the best reviewed in 'White Paper' for cotton genome sequencing (available at http://algodon.tamu.edu/sequencing/, verified on June 28, 2007), which are imperative for future utilization of the genetic diversity in cotton. Although there are many marker systems such as isozyme, RAPD, RFLPs, AFLPs (extensively referenced herein), and their various modifications (Zhang et al. 2005b; Zhang et al., 2007; in this conference) successfully used in cotton, the development of a large collection of a robust, portable, and PCR-based molecular marker resources, in particular, one of the widely used genetic markers, Simple Sequence Repeats (SSRs) and Single Nucleotide Polymorphisms (SNPs) for cotton during past decade, were one of tremendous accomplishments of cotton research community that accelerated studies of genetic diversity in cotton in genomic level. These resources were made available for cotton research community through cotton marker database (CMD) (Blenda et al., 2006) that are being extensively used to create cotton genetic linkage maps (Zhang et al., 2002; Lacape et al., 2003; Han et al., 2004; Rong et al., 2004; Gao et al., 2004; Nguyen et al., 2004; Lacape et al., 2005; Han et al., 2006; Song et al., 2005; Zang et al., 2005; Frelichowski et al., 2006; Guo et al., 2007b) and map important agronomic QTLs (Kohel et al., 2001; Karaca et al., 2002; Paterson et al., 2003; Mei et al., 2004; Shen et al., 2005; Abdurakhmonov et al., 2005; Bolek et al., 2005; Park et al 2005; Zhang et al., 2005a; Wang et al., 2006; Shen et al., 2006; Abdurakhmonov et al., 2007a; He et al., 2007; Shen et al., 2007). These studies are very important to mobilize useful genes of agronomic importance to the elite cultivars through marker-assisted breeding programs. Another very important effort in exploiting the genetic diversity of crops is the estimation of genetic distances, geneology and phylogeny of gene pools that some aspects was briefly reviewed above (Wendel et al., 1992; Brubaker and Wendel, 1994;

Tatineni et al., 1996; Iqbal et al., 1997; Lui et al., 2000a; Abdalla et al., 2001; Iqbal et al., 2001, Gutierrez et al., 2002; Wendel and Cronn, 2003; Rungis et al., 2005; Lacape et al., 2007). Furthermore, researchers have reported several potential candidate genes of many agronomic traits in cotton. Tremendous efforts were made to study molecular bases of one of the most complex, but important traits – cotton fiber development (Nolte et al., 1995; Pear et al. 1996; Orford and Timmis 1998; Wilkins and Jernstedt, 1999; Ruan et al. 2001; Wu et al., 2001; Ruan et al., 2003; Suo et al., 2003; Ji et al., 2003; Giband et al., 2003; Arpat et al., 2004; Li et al., 2005; Sun and Allen, 2005; Lee et al., 2006; Jacob-Wilk et al. 2006). These efforts, including many more recent reports from above-referenced researchers laboratories in dissection of candidate genes that are specifically expressed in developing fibers are undoubtedly imperative for future exploiting of genetic diversity in cotton fiber traits using transgenomics approaches (refer to T. Wilkins, 2007; in this conference).

In our laboratory, we are developing genomics approaches to utilize the genetic diversity. We characterized phytochrome gene family and other light signal transduction factor genes to understand the molecular basis the photoperiodic flowering of cotton based on Arabidopsis genome information (Abdurakhmonov et al., 2006a, Abdurakhmonov and Pepper, 2006). To further understand day-neutral flowering associated genomic modification, we recently characterized a number of directly photoperiod converted induced mutants in molecular level (Abdurakhmonov et al., 2007c). The extensive research efforts on understanding the molecular basis of photoperiodic flowering of cotton are in progress. Molecular tagging of fiber quality traits in a natural fiber mutant derived RIL lines and natural leaf defoliation trait introgressed from wild germplasm in trigenomic hybridization were also conducted in our laboratory that accelerates utilization of useful diversity in the current cotton breeding programs of Uzbekistan (Abdurakhmonov et al., 2005; Abdurakhmonov et al., 2007a). We are also extensively characterizing ~1000 *G. hirsutum* germplasm resources in both phenotypic and molecular levels to revel potential new genetic diversity. We developed electronic database for characterized cotton germplasm resources using pcGRIN (pcGRIN Data

Management version 1.21) and Microsoft Access software that contains 72 characteristics for each accession, including passport data, collection data, site data, plant data and inflorescence and fruit data including all main fiber quality traits recorded from two very diverse environments, Uzbekistan and Mexico (Abdurakhmonov et al. 2006b). Further efforts to improve our electronic database are in progress.

**Association mapping to better exploit the genetic diversity**

In spite of extensive cotton genomics efforts, cotton lags behind other major crops for marker-assisted breeding due to limited polymorphism in the cultivated germplasm. QTL-mapping is a now-classical approach to identify molecular markers linked to complex traits that are segregating in specific populations. This approach requires that specific mapping populations, usually consisting of several hundred $F_2$ or recombinant inbred (RI) progeny, be developed from each germplasm accession to be examined for important genes effecting fiber quality and yield. Each population must be genotyped using hundreds, perhaps thousands of molecular markers. Population development and marker screening is extremely time-consuming, high-risk and expensive work — prohibitively expensive if dozens, let alone hundreds or thousands, of cotton germplasm accessions are to be examined. However, use of LD-based association mapping circumvents the need for large $F_2$ or RI mapping populations by making use of information contained within the genetic recombinations that have occurred in natural populations during the course of recent evolution (Abdurakhmonov et al., unpublished). Although novel to cotton research, the association genetics strategy is, in fact, highly applicable to the identification of markers linked to fiber quality and yield through the examination of linkage disequilibrium (LD) of DNA-based markers with fiber quality and yield traits in a large, diverse germplasm collection (Abdurakhmonov et al., 2004).

LD-based association mapping is a powerful molecular-genetic tool for high-resolution mapping of complex quantitative traits that facilitates broad and comprehensive utilization of natural genetic

diversity within populations or germplasm collections As a starting point, it is important to gain knowledge of the LD patterns for genomic regions of the 'target' organisms and the specificity of LD extent among different populations or groups to design and conduct association mapping (Nordborg et al., 2002; Nordborg at al., 2005). Although association mapping in plants is less advanced than in humans, the available plant germplasm collections are likely to be suitable materials to be exploited for association mapping when the extent of genome-wide LD is known with the further opportunity of creating mapping populations with required amount of LD and diversity. These gives more power for association mapping and results in even better utilization of natural diversity in plants (Rafalski and Morgante, 2004), including cotton as well.

In the frame of Former soviet Union (FSU) and USDA-ARS collaborative programs, we explored the Uzbek cotton germplasm collection, one of the largest collections with broad geographic and genetic diversity coverage that provided us an opportunity to identify the genetic diversity, population structure, kinship and an average extent of a genome-wide LD in cotton using large number of SSR markers. Our results revealed that pairwise LD level between SSR marker loci in cultivar and exotic *G. hirsutum* germplasm (in a range of 4-12%, $p \leq 0.005$) with genome-wide averages of LD decay within ~10-30 cM at $r^2=0.1$. We successfully applied 'association-mapping' strategy in cotton to tag potentially new fiber QTLs in the two diverse environments, Uzbekistan and Mexico and identified number of SSR markers associated with the microniare, fiber length and strength, uniformity, elongation and reflectance (Abdurakhmonov et al., 2007b; Abdurakhmonov et al., unpublished). In contrast to the human genome, where a very high density of molecular markers is needed for association mapping in the majority of cases (Kruglyak 1999), the cotton genome may require significantly fewer numbers of markers for effective LD–mapping of complex traits, which is also the case reported for other crops (Kraakman et al., 2004; Barnaud et al., 2006). Considering the tetraploid cotton genome with a total recombinational length of about 5200 cM, and an average 400kb per cM (Paterson and Smith, 1999), LD block sizes of 10-30 cM distance in cotton is large

enough to conduct an association mapping of complex traits that would require a maximum of ~200-500 polymorphic markers for successful association mapping. The mapping resolution may be limited, in particular with breeding germplasm. However, high levels of LD between SSR loci and to a lesser extent of the LD blocks in exotic *G. hirsutum* cotton germplasm suggest an opportunity to develop a set of mapping populations with the required amount of LD and diversity for high resolution mapping through directed crossing between selected germplasm pools (Rafalski and Morgante, 2004) The first insights of LD-based association mapping using MLM approach (Yu et al., 2006) highlight the promise of this approach in more effective exploiting the natural diversity of cotton preserved *ex situ* in worldwide cotton collections .

It should be noted that our estimate of genome-wide averages for the extent of LD in cotton might not adequately reflect LD patterns of specific regions or specific population groups. Each of these specific regions or population groups should additionally be explored for the extent of LD in order to conduct successful association mapping of variants within regions or populations of interest. This requires further efforts and investments in the development of fine-scale association mapping studies in cotton. Sequencing the cotton genome(s) will greatly facilitate high resolution and cost-effective linkage disequilibrium (LD)-based mapping and future whole genome-association (WGA) strategies in cotton, providing the evolutionary signatures of 'selective sweeps' in the genome and epigenetic details of the traits of interest. This will accelerate the fine-scale mapping and better utilization of functionally significant natural variations of gene pools preserved *ex situ* within worldwide cotton germplasm collections.

**Conclusions**

In conclusion, the *Gossypium* genus encompasses wide geographic and ecological niches and represents large amplitude of morphobiological and genetic diversity. The existed and predicted problems of worldwide cotton breeding programs associated with the narrowness of genetic base of the cultivar germplasm is due to challenges and the lack of innovative tools to exploit the useful

genetic variations of *Gossypium* genus. Although the utilization of genetic diversity of *Gossypium* genus using traditional breeding approaches was successful with a number of 'large impact' examples in worldwide cotton production, it is somewhat challenging. The most effective utilization of genetic diversity of *Gossypium* species further requires characterization of candidate gene(s) underlying the phenotypic and agronomic diversities, acceleration of linkage mapping, map-based cloning and markers assisted selection that underlie development of more innovative genomics technologies such as high-resolution, cost effective LD-based association mapping for cotton. The development of cotton transgenomics and complete sequencing of cotton genome(s) further will accelerate exploiting genetic diversity in highly specific manner and with clear vision. Future application of whole genome-association strategy with epigenomics perspectives, which currently is widely being applied in human and the other model plants such as Arabidopsis, will have a significant impact to identify true functions of genes controlling available genetic diversity, and consequently, its effective utilization.

## REFERENCES


Abdalla, A.M., O.U.K. Reddy, K.M. El-Zik, and A.E. Pepper. 2001. Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theor Appl Genet. 102:222-229.

Abdukarimov, A., S. Djataev, and I.Y. Abdurakhmonov. 2003. Cotton Research in Uzbekistan: Elite Varieties and Future of Cotton Breeding. The proceedings of WCRC-3, South Africa, P. 5-16.

Abdurakhmonov, I.Y., A. A. Abdullaev, S. Saha, Z.T. Buriev, D. Arslanov, Z. Kuryazov, G.T. Mavlonov, S.M. Rizaeva, U.K. Reddy, J.N. Jenkins, A. Abdullaev, and A. Abdukarimov. 2005. Simple sequence repeat marker associated with a natural leaf defoliation trait in tetraploid cotton. Journal of Heredity 96(6):644-653.

Abdurakhmonov, I.Y., A. Abdullaev, S. Rizaeva, Z. Buriev, A. Adylova, A. Abdukarimov, S. Saha, R. Kohel, J. Yu, and A.E. Pepper. 2004. Evaluation of *G. hirsutum* exotic accessions from Uzbek cotton germplasm collection for further molecular mapping purposes. National Cotton Council of America. The Proc. Beltwide Cotton Conf., San Antonio, Texas, USA, p.1133-1142.

Abdurakhmonov, I.Y., Z.T. Buriev, S. Saha, A.E. Pepper, J.A. Musaev, A. Almatov, S.E. Shermatov, F.N. Kushanov, G.T. Mavlonov, U.K. Reddy, J.Z. Yu, J.N. Jenkins, R.J. Kohel, and A. Abdukarimov. 2007a. Microsatellite markers associated with lint percentage trait in cotton, *Gossypium hirsutum*. Euphytica DOI 10.1007/s10681-007-9361-2.

Abdurakhmonov, I.Y., Z.T. Buriev, A.A. Abdukarimov, and A.E. Pepper. 2006a. Molecular Cloning And Characterization Of Phytochrome Gene Family In Cotton (*Gossypium spp*.). Plant & Animal Genome XIV Conference. Jan. 14-19, 2006. San Diego, CA, W159.

Abdurakhmonov, I. Y., Z. T. Buriev, I. B. Salakhuddinov, S. M. Rizaeva, A. T. Adylova, S. E. Shermatov, A. Adukarimov, R. J. Kohel, J. Z. Yu, A. E . Pepper, S. Saha, and J.N. Jenkins.

2006b. Characterization of *G. hirsutum* Wild and Variety Accessions from Uzbek Cotton Germplasm Collection for Morphological and Fiber Quality Traits and Database Development. Cotton Beltwide Conference. Jan. 3-6.San, 2006. Antonio, Texas, USA. P5306

Abdurakhmonov, I.Y., R.J. Kohel, S. Saha, A.E. Pepper, J. Yu, Z.T. Buriev, S.E. Shermatov, A.A. Abdullaev, F.N. Kushanov, J.N. Jenkins, B.E. Scheffler, and A. Abdukarimov. 2007b. Genome-Wide Linkage Disequilibrium Revealed By Microsatellite Markers And Association Study Of Fiber Quality Traits In Cotton. Plant & Animal Genome XV Conference. Jan. 13-17, 2007. San Diego, CA, W199.

Abdurakhmonov, I.Y., F.N. Kushanov, F. Djaniqulov, Z.T. Buriev, A.E. Pepper, N. Fayzieva, G.T. Mavlonov, S. Saha, J.N. Jenkins, and A. Abdukarimov. 2007c. The Role of Induced Mutation in Conversion of Photoperiod Dependence in Cotton. Journal of Heredity 98:258 – 266.

Abdurakhmonov, I.Y. and A.E. Pepper. 2006. Mining of A. thaliana transcription regulator HY5 from Cotton (*Gossypium spp.*). Plant & Animal Genomes XIV Conference. Jan. 14-19, 2006. San Diego, CA. P64.

Adams, K.L., R. Percifield, and J.F. Wendel. 2004. Organ-specific silencing of duplicated genes in a newly synthesized cotton allotetraploid. Genetics 168(4):2217-26.

Arpat, A. B., M. Waugh, J. P. Sullivan, M. Gonzales, D. Frisch, D. Main, T. Wood, A. Leslie, R. A. Wing , and T. A. Wilkins. 2004. Functional genomics of cell elongation in developing cotton fibers. Plant Mol Biol 54: 911–929.

Barnaud, A.T., Lacombe, and A. Doligez. 2006. Linkage disequilibrium in cultivated grapevine, *Vitis vinifera L*. Theor Appl Genet 112:708-716.

Blenda, A., J. Scheffler, B. Scheffler, M. Palmer, J.M. Lacape, J.Z. Yu, C. Jesudurai, S. Jung, S. Muthukumar, P. Yellambalase, S. Ficklin, M. Staton, R. Eshelman, M. Ulloa, S. Saha, B.

Burr, S. Liu, T. Zhang, D. Fang, A. Pepper, S. Kumpatla, J. Jacobs, J. Tomkins, R. Cantrell, and D. Main. 2006. CMD: A Cotton Microsatellite Database Resource for Gossypium Genomics. BMC Genomics 7: 132

Bolek, Y., K.M. El-Zik, A.E. Pepper, A.A. Bell, C.W. Magill, P.M. Thaxton, and O.U. Reddy. 2005. Mapping of verticillium wilt resistance genes in cotton. Plant Sci. 168:1581-1590.

Brubaker, C.L., A.H. Paterson, and J.F. Wendel. 1999. Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. Genome 42:184-203.

Brubaker, C.L. and J.F. Wendel. 1994. Reevaluating the origin of domesticated cotton (*Gossypium hirsutum, Malvaceae*) using nuclear restriction fragment length polymorphism (RFLP). Am J Bot. 81:1309-1326.

Dejoode, D.R. and J. F. Wendel. 1992. Genetic diversity and origin of the Hawaiian-Islands cotton, *Gossypium tomentosum*. Am J Bot. 79:1311-1319.

Djaniqulov, F. 2002. About relation between radio- sensitivity and mutability of wild and tropical cultivated cotton forms. Proceedings of Russian Academy of Agricultural Sciences. 2:19-22. (in Russian)

Endrizzi, J.E., E.L. Turcotte, and R.J. Kohel. 1985. Genetics, cytology, and evolution of *Gossypium*. Adv Genet 23:272–375.

Frelichowski, J.E. Jr., M.B. Palmer, D. Main, J.P. Tomkins, R.G. Cantrell, D.M. Stelly, J. Yu, R.J. Kohel, and M. Ulloa. 2006. Cotton genome mapping with new microsatellites from Acala 'Maxxa' BAC-ends. Mol Gent Genomics 275(5):479-491.

Fryxell, P.A. 1979. Natural history of the cotton tribe. Texas A&M University Press, College Station, Texas, USA.

Gao, W., Z.J. Chen, J.Z. Yu, D. Raska, R.J. Kohel, J.E. Womack, and D.M. Stelly. 2004. Wide-cross whole-genome radiation hybrid mapping of cotton (*Gossypium hirsutum L.*). Genetics 167(3): 1317-29.

Garris, A.J., S.R. McCouch, and S. Kresovich. 2003. Population structure and its effect on haplotype diversity and linkage disequilibrium surrounding the xa5 locus of rice (*Oryza saliva L.*). Genetics 165:759-769.

Gingle, A. R., H. Yang, P. W. Chee, O. L. May, J. Rong, D. T. Bowman, E. L. Lubbers, J. L. Day, and A. H. Paterson. 2006. An Integrated Web Resource for Cotton. Crop Sci 46:1998-2007

Giband, M., S. Pagant, C. Pannetier, and H. Hofte. 2003. *Arabidopsis thaliana* as a source of candidate genes for cotton fiber quality. In: The proceedings of the 3rd World Cotton Research Conference. Cape Town, South Africa. P.58–63

Guo, W., C. Cai, C. Wang, Z. Han, X. Song, K. Wang, X. Niu, C. Wang, K. Lu, B. Shi, and T. Zhang. 2007b. A microsatellite-based, gene-rich linkage map reveals genome structure, function, and evolution in *Gossypium*. Genetics DOI: 10.1534/genetics.107.070375.

Guo, W., Z. Q. Sang, B. L. Zhou, and T. Zhang. 2007a.Genetic relationships of D-genome species based on two types of EST-SSR markers derived from *G. arboreum* and *G. raimondii* in *Gossypium*. Plant Science 172:808-814.

Gutierrez, O.A., S. Basu, S. Saha, J.N. Jenkins, D.B. Shoemaker, C.L. Cheatham, and J.C. McCarty. 2002. Genetic distance among selected cotton genotypes and its relationship with F2 performance. Crop Sci. 42:1841-1847.

Han, Z.G., W. Guo, X.L. Song, and T. Zhang. 2004. Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. Mol Genet Genomics 272: 308-327.

Han, Z.G., C. Wang, X.L. Song, W. Guo, J.Y. Gou, C. Li, X. Chen, and T. Zhang. 2006. Characteristics, development and mapping of *Gossypium hirsutum* derived EST-SSRs in allotetraploid cotton. Theor Appl Genet. 112:430-439.

Hawkins, J.S., J. Pleassants, and J.F. Wendel. 2005. Identification of AFLP markers that discriminate between cultivated cotton and the Hawaiian islan endimic, *Gossypium tomentosum* nuttall ex seeman. Genet Resour Crop Evol. 52:1069-1078.

He, D.H., Z.X. Lin, X.L. Zhang, Y.C. Nie, X.P. Guo, Y.X. Zhang, and W. Li. 2007. QTL mapping for economic traits based on a dense genetic map of cotton with PCR-based markers using the interspecific cross of *Gossypium hirsutum* x *Gossypium barbadense*. Euphytica 153 (1-2):181-197.

He, D.H., Z.X. Lin, X.L. Zhang, Y.C. Nie, X.P. Guo, C.D. Feng, and J.Mc.D. Stewart. 2005. Mapping QTLs of traits contributing to yield and analysis of genetic effects in tetraploid cotton. Euphytica 144:141-149.

Iqbal, M.J., N. Aziz, N.A. Saeed, Y. Zafar, and K.A. Malik. 1997. Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. Theor Appl Genet 94:139-144.

Iqbal, J., O.U.K. Reddy, K.M. El-Zik, and A.E. Pepper. 2001. A genetic bottleneck in the 'evolution under domestication' of upland cotton *Gossypium hirsutum L*. examined using DNA fingerprinting. Theor Appl Genet. 103:547-554.

Jacob-Wilk, D., I Kurek, P Hogan, and D. P. Delmer.2006.The cotton fiber zinc-binding domain of cellulose synthase A1 from Gossypium hirsutum displays rapid turnover in vitro and in vivo. Proc Natl Acad Sci U S A 103(32):12191-6.

Ji, S.S., Y.C. Lu, J.X. Feng, G. Wei, S. Li, Y.H. Shi, Q. Fu, D. Liu, J.C Luo, and Y.X. Zhu. 2003. Isolation and analyses of genes preferentially expressed during early cotton fiber development by subtractive PCR and c DNA array. Nucleic Acid Res 31: 2534–2543.

Karaca, M.S., J.N. Saha, A. Jenkins, A. Zipf, R.J. Kohel, and M.D. Stelly. 2002. Simple sequence Repeat (SSR) markers Linked to the ligon lintless (Li) mutant in cotton. Journal of Heredity 93:221-224.

Kebede, H., G. Burow, R. G. Dani, and R. D. Allen. 2007. A-genome cotton as a source of genetic variability for Upland cotton (*Gossypium hirsutum*). Genet Resour Crop Evol. DOI10.1007/s10722-006-9157-6.

Kohel, R.J., J.Z. Yu, Y.H. Park, and G.R. Lazo. 2001. Molecular mapping and characterization of genes controlling fiber quality in cotton. Euphytica 121:163-172.

Kraakman, A.T.W., R.E. Niks, P. M. M. M Van den Berg, P. Stam, and F.A. Van Eeuwijk. 2004. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. Genetics 168:435- 446.

Kruglyak, L. 1999. Prospects for whole-genome linkage disequilibrium mapping of common disease genes. Nat Genet. 22:139-144.

Lacape, J.M., D. Dessauw, M. Rajab, J.L. Noyer, and B. Hau. 2007. Microsatellite diversity in tetraploid *Gossypium* germplasm: assembling a highly informative genotyping set of cotton SSRs. Mol Breeding 19 (1):45-58.

Lacape, J.M., T.B. Nguyen, B. Courtois, J.L. Belot, M. Giband, J.P. Gourlot, G. Gawryziak, S. Roques, and B. Hau. 2005. QTL analysis of cotton fiber quality using multiple *Gossypium hirsutum* x *Gossypium barbadense* backcross generations. Crop Sci. 45:123-140.

Lacape, J.M., T.B.Nguyen, S. Thibivilliers, B. Bojinov, B. Courtois, R.G. Cantrell, B. Burr, and B. Hau. 2003. A combined RFLP-SSR-AFLP map of tetraploid cotton based on a *Gossypium hirsutum* x *Gossypium barbadense* backcross population. Genome 46 (4):612-26.

Lee, J.J., O.S. Hassan, W. Gao, N.E. Wei, R.J. Kohel, X.Y. Chen, P. Payton, S.H. Sze, D.M. Stelly, and Z.J. Chen. 2006. Developmental and gene expression analyses of a cotton naked seed mutant. Planta 223:418–432.

Li, X.B., X.P. Fan, X.L. Wang, L. Cai, and W.C. Yang.2005. The cotton ACTIN1 gene is functionally expressed in fibers and participates in fiber elongation. Plant Cell 17:859–875.

Liu, S., R.G. Cantrell, J.C.J. McCarty, and J. M. Stewart. 2000a. Simple sequence repeat-based assessment of genetic diversity in cotton race stock accessions. Crop Sci. 40:1459-1469.

Liu, D., X. Guo, Z. Lin, Y. Nie, and X. Zhang. 2006. Genetic diversity of Asian cotton (*Gossypium arboreum L.*) in China evaluated by microsatellite analysis. Genetic Resour Crop Evol. 53(6): 1145-1152.

Liu, S., S. Saha, D. Stelly, B. Burr, and R.G. Cantrell. 2000b. Chromosomal assignment of microsatellite loci in cotton. Journal of Heredity 91 (4):326-32.

Mauer, F. M. 1954. Origin and taxonomy of cotton. In Cotton. Academy of Sciences of USSR, Tashkent, Uzbekistan pp. 383 (In Russian).

McCarty, J.C.Jr. and Jenkins JN, 1993. Registration of 79 day-neutral primitive cotton germplasm lines. Crop Sci. 33:351.

McCarty, JC Jr., J.N. Jenkins, W.L. Parrott, and R.G. Greech. 1979. The conversion of photoperiodic primitive race stocks of cotton to day-neutral stocks. Mss Agric and Forestry Exp Stn Res Rep. 4(19):1-4.

Mei, M., N.H. Syed, W. Gao, P.M. Thaxton, C.W. Smith, D.M. Stelly, and Z.J. Chen. 2004. Genetic mapping and QTL analysis of fiber-related traits in cotton (*Gossypium*). Theor Appl Genet.108: 280-291.

Multani, D.S., and B.R. Lyon.1995. Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38:1005-1008.

Musaev, J.A., M.F. Abzalov, A. Almatov, M.F. Sanamyan, N. Gubanova, and U. Nadjimov, 2000. Cotton Genetics and Genetic Collection of Isogenic, Monosomic and Translocation Lines. Bulletins SCST of the republic of Uzbekistan p.28–39 (in Russian).

Nguyen, T.B., M. Giband, P. Brottier, A.M. Risterucci, and J.M. Lacape. 2004. Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. Theor Appl Genet. 109:167-175.

Nolte, K.D., D.L. Hendrix, J.W. Radin, and K.E. Koch. 1995. Sucrose synthase localization during initiation of seed development and trichome differentiation in cotton ovules. Plant Physiol 109: 1285–1293.

Nordborg, M., J.O. Borevitz, J. Bergelson, C.C. Berry, J. Chory, et al. 2002. The extent of linkage disequilibrium in *Arabidopsis thaliana*. Nat Genet. 30:190-193.

Nordborg, M., T.T. Hu, Y. Ishino, J. Jhaveri, C. Toomajian, et al. 2005. The pattern of polymorphism in *Arabidopsis thaliana*. PLoS Biol. 3 7: el96.

Orford, S.J. and J.N. Timmis.1998. Specific expression of an expansin gene during elongation of cotton fibers. Biochim Biophys Acta 1398:342–346.

Park, Y.H., M.S. Alabady, M. Ulloa, B. Sickler, T.A. Wilkins, J. Yu, D.M. Stelly, R.J. Kohel, O.M. El-Shihy, and R.G. Cantrell. 2005. Genetic mapping of new cotton fiber loci using EST-derived microsatellites in an interspecific recombinant inbred line cotton population. Mol Genet Genomics 274: 428-441.

Paterson, A.H., Y. Saranga, M. Menz, and C.X. Ziang. 2003. QTL analysis of genotype X environment interactions affecting cotton fiber quality. Theor Appl Genet 106:384-396.

Paterson, A.H., and R.H. Smith. 1999. Future horizons: biotechnology of cotton improvement. In: Cotton: Origin, History, Technology, and Production. John Wiely&Sons, Inc., New York.

Pear, J.R., Y. Kawagoe, W.E. Schreckengost, D.P. Delmer, and D.M. Stalker. 1996. Higher plants contain homologs of the bacterial celA genes encoding the catalytic subunit of cellulose synthase. Proc Natl Acad Sci USA 93: 12637–12642.

Percival, A.E., J.M. Stewart, and J.F. Wendel. 1999. Taxonomy and germplasm resources. In: Cotton: origin, history, technology and production, (Smith CW and Cothren JT, eds). New York: John Wiley; 33-63.

Phillips, L.L. 1964. Segregation in new allopoliploids of *Gossypium*. V. Multivalent formation in New world x Asiatic and New world x wild American Hexaploid. Am J Bot 51:324-329.

Pillay, M. and G.O. Myers. 1999. Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. Crop Sci. 39:1881-1886.

Rafalski, A. and M. Morgante. 2004. Corn and humans: recombination and linkage disequilibrium in two genomes of similar size. Trends Genet. 20:103-111.

Rana, M. K., V. P. Singh, and K. V. Bhat. 2005. Assessment of genetic diversity in upland cotton (*Gossypium hirsutum L.)* breeding lines by using amplified fragment length polymorphism (AFLP) markers and morphological characteristics. Cenet Resour Crop Evol. 52:989-997.

Rizaeva, S. M. 1996. Interspecific hybridization of cotton and development of new cotton donor accessions (in the example of New World cottons) (Doctor of Science dissertation). Tashkent, Uzbekistan: Institute of Plant Experimental Biology of the Academy of Sciences of Uzbekistan (in Russian).

Rizaeva ,SM, Abdullaev AA, Klyat VP, Arslonov DM, and Kuryazov ZB, 2001. Creation of donors with naturally early leaf defoliation. Uzbek Biol. Journal 4:65-70 (in Russian).

Rong, J., C. Abbey, J.E. Bowers, C.L. Brubaker, C. Chang, P. Chee, T.A. Delmonte, X. Ding, J.J. Garza, B.S. Marler, C-h. Park, G.J. Pierce, K.M. Rainey, V.K. Rastogi, S.R. Schulze, N. Troiinder, J.F. Wendel, T.A. Wilkins, T.D. WilHams-Coplin, R.A. Wing, R.J. Wright, X. Zhao, L. Zhu, and A.H. Paterson. 2004. A 3347 locus genetic recombination map of sequence-tagged sites reveals features of genome organization, transmission and evolution of cotton *(Gossypium).* Genetics 166:389-417.

Ruan, Y.L., D.J. Llewellyn, and R.T. Furbank. 2001. The control of single-celled cotton fiber elongation by developmentally reversible gating of plasmodesmata and coordinated expression of sucrose and K+ transporters and expansin. Plant Cell 13: 47–63.

Ruan, Y.L., D.J. Llewellyn, and R.T. Furbank. 2003. Suppression of sucrose synthase gene expression represses cotton fiber cell initiation, elongation, and seed development. Plant Cell 15: 952-64

Rungis, D., D. Llewellyn, E.S. Dennis, and B.R. Lyon. 2005. Simple sequence repeat (SSR) markers reveal low levels of polymorphism between cotton (*Gossypium hirsutum L.*) cultivars. Aus J Agr Res. 56:301-307.

Sanamyan, M. F. E. M. Rakhmatullina. 2003. Cytogenetic analysis of translocations in cotton. Plant Breeding 122 (6): 511–516.

Shen, X., W. Guo, Q. Lu, X. Zhu, Y. Yuan, and T. Zhang. 2007. Genetic mapping of quantitative trait loci for fiber quality and yield trait by RIL approach in Upland cotton. Euphytica DOI: 10.1007/s10681-006-9338-6.

Shen, X., W. Guo, X. Zhu, Y. Yuan, J.Z. Yu, R.J. Kohel, and T. Zhang. 2005. Molecular mapping of QTLs for fiber qualities in three diverse lines in Upland cotton using SSR markers. Mol Breeding 15(2):169-181.

Shen, X., G. Van Becelaere, P. Kumar, R.F. Davis, O.L. May, and P. Chee. 2006. QTL mapping for resistance to root-knot nematodes in the M-120 RNR Upland cotton line (*Gossypium hirsutum L.*) of the Auburn 623 RNR source. Theor Appl Genet. 113(8):1539-1549.

Song, X., K. Wang, W. Guo, J. Zhang, T. Zhang, and G.J. Scoles. 2005. A comparison of genetic maps constructed from haploid and BC1 mapping populations from the same crossing between *Gossypium hirsutum L.* and *Gossypium barbadense L*. Genome 48 (3):378-390.

Sun Y. and R. D. Allen. 2005. Functional analysis of the BIN 2 genes of cotton. Mol Genet Genomics. 274:51-9.

Suo, J., X. Liang, L. Pu, Y. Zhang, and Y. Xue. 2003. Identification of GhMYB109 encoding a R2R3 MYB transcription factor that expressed specifically in fiber initials and elongating fibers of cotton (*Gossypium hirsutum* L.). Biochim Biophys Acta. 1630(1):25–34.

Tatineni, V., R.G. Canlrell, and D.D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPD. Crop Sci. 36:186-192.

Ulloa, M., J. M. Stewart, E. A. Garcia-C, S. Goday-A, A. Gaytan-M, and S.Acosta-N. 2006. Cotton genetic resources in the western states of Mexico: *in situ* conservation status and germplasm collection for *ex situ* preservation. Genet resour Crop Evol. 53:653-668.

Van Esbroeck, G. A., D.T. Bowman, O. L. May and D. S. Calhoun. 1999. Genetic similarity indices for ancestral cotton cultivars and their impact on genetic diversity estimates of modern cultivars. Crop Sci. 39:323-328.

Van Esbroeck, G. A., D.T. Bowman, D. S. Calhoun, and O. L. May. 1998. Changes in the genetic diversity of cotton in the USA from 1970 to 1995. Crop Sci. 38:33-37.

Wang, B., W. Guo, X. Zhu, Y. Wu, N. Huang, and T. Zhang. 2006. QTL mapping of fiber quality in an elite hybrid derived-RIL population of upland cotton. Euphytica 152 3:367-378.

Wendel, J.F. 1995. Cotton. *In*: SimmondsS, SmarttJ (eds). Evolution of crop plants, 1st ed. Longman, London, pp. 358-366.

Wendel, J. F. and V. A. Albert. 1992. Phylogenetics of the cotton genus (*Gossypium*) - character state weighted Parsimony analysis of chloroplast-DNA restriction site data and its systematic and biogeographic implications. Sys Bot. 17:115-143.

Wendel, J.F. and C. L. Brubaker.1993. RFLP diversity in *Gossypium hirsutum L*. and new insights into the domestication of cotton. Am J Bot. 80 (Suppl 6):71.

Wendel, J.F., C.L. Brubaker, and A.E. Percival. 1992. Genetic diversity in *Gossypium hirsutum* and the origin of Upland cotton. Am J Bot. 79:1291-1310.

Wendel, J.F. and R.C. Cronn. 2003. Polyploidy and the evolutionary history of cotton. Adv Agronomy 78:140-186.

Wendel, J.F. and R. G. Percy.1990. Allozyme diversity and introgression in the Galapagos-Islands endemic *Gossypium darwinii* and its relationship to continental Gossypium *barbadense*. Bioch Syst Ecol. 18:517-528.

Wendel, J.F., R. Rowley, and J.M. Stewart. 1994. Genetic diversity in and phylogenetic-relationships of Brazilian endemic cotton, *Gossypium mustelinum* (Malvaceae). Plant Syst Evol.192:49-50

Westengen, O.T., Z. Huaman, and M. Heum. 2005. Genetic diversity and geographic pattern in early South American cotton domestication. Theor Appl Genet. 110:392-402.

Wilkins, T.A. and J.A. Jernstedt. 1999. Molecular genetics of developing cotton fibers. Food product press 231–269.

Yu, J., G. Pressoir, W.H. Briggs, Bi. I. Vroh, M. Yamasaki, J.F. Doebley, M.D. McMullen, B.S. Gaut, D.M. Nielsen, J.B. Holland, S. Kresovich, and E.S. Buckler. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat Genet. 38(2): 203-208

Zhang, J., W. Guo, and T. Zhang. 2002. Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum L.* x *Gossypium barbadense L.*) with a haploid population. Theor Appl Genet.105 (8): 1166-1174.

Zhang, J., Y. Lu, R.G. Cantrell, and E. Hughs. 2005a. Molecular marker diversity and field performance in commercial cotton cultivars evaluated in the Southwestern USA. Crop Sci. 45: 1483-1490.

Zhang J.F., Y.Z. Lu, and S.X. Yu. 2005b. Cleaved AFLP (cAFLP), a modified amplified fragment length polymorphism analysis for cotton. Theor. Appl. Genet. 111:1385-1395.

Zhang, Z.S., Y.H. Xiao, M. Luo, X.B. Li, X.Y. Luo, L. Hou, D.M. Li, and Y. Pei. 2005. Construction of a genetic linkage map and QTL analysis of fiber-related traits in upland cotton (*Gossypium hirsutum L.*). Euphytica 144 (1): 91-99.

Zhao, X.P., Y.F. Ji, X.L. Ding, D.M. Stelly, and A.H. Paterson. 1998 Macromolecular organization and genetic mapping of a rapidly evolving chromosome-specific tandem repeat family (B77) in cotton (*Gossypium*). Plant Mol Biol. 38 (6): 1031-1042.

## FIGURE CAPTIONS

**Fig. 1.** Unrooted NJ-phylogenetic tree of G. hirsutum accessions from Uzbek cotton germplasm resources. Afghanistan (AFG), Azerbaijan (AZR), Bulgaria (BU), China (CN), Czechoslovakia (CZ), Hungarian (HU), India (IN), Iraq (IRQ), Korea (KR), Pakistan (PAK), Syria (SYR), Taiwan (TAI), Turkey (TUR), Ukraine (UA), United States (USA) and Uzbekistan (UZB). Branch lengths are shown.

**Fig. 2.** Diversity on plant bush architecture observed in *G. hirsutum* germplasm; A–creeping type; B to E–variations of vertical branchy (erect) type of bushes; F to I–variations of vertical compact type of bushes.

**Fig.3.** Diversity on stem color observed in *G. hirsutum* germplasm: A–no antocyanin color; B–weak antocyanin color; C–middle antocyanin color; and D–strong antocyanin color.

**Fig. 4.** Diversity on stem pubescence observed in *G. hirsutum* germplasm: A–naked; B–medium; C–sparse; D–hairy.

**Fig. 5.** Diversity on leaf plate shape observed in *G. hirsutum* germplasm: A to C–variations of 3-lobbed leaf plates; D– 5-lobbed leaf plate; E to H–variations of palmate type leaves; I to L–variations of semi-digitate and digitate type leaves; M to P–variations of okra and super okra type leaves (P)

**Fig. 6.** Diversity on flower color observed in *G. hirsutum* germplasm: A –white-yellow; B and C–light yellow; D and E–yellow; F–cream color; G–lavender; H and I–pink; J–bicolor.

**Fig.7.** Diversity on pollen color observed in *G.hirsutum* germplasm: A – light yellow; B– light cream color; C–cream color; D to G–variations of yellow pollen; H and I–tawny.

**Fig.8.** Diversity on boll shape observed in *G. hirsutum* germplasm: row A- cone-shaped; row B–oval-shaped; row C–round-shaped.

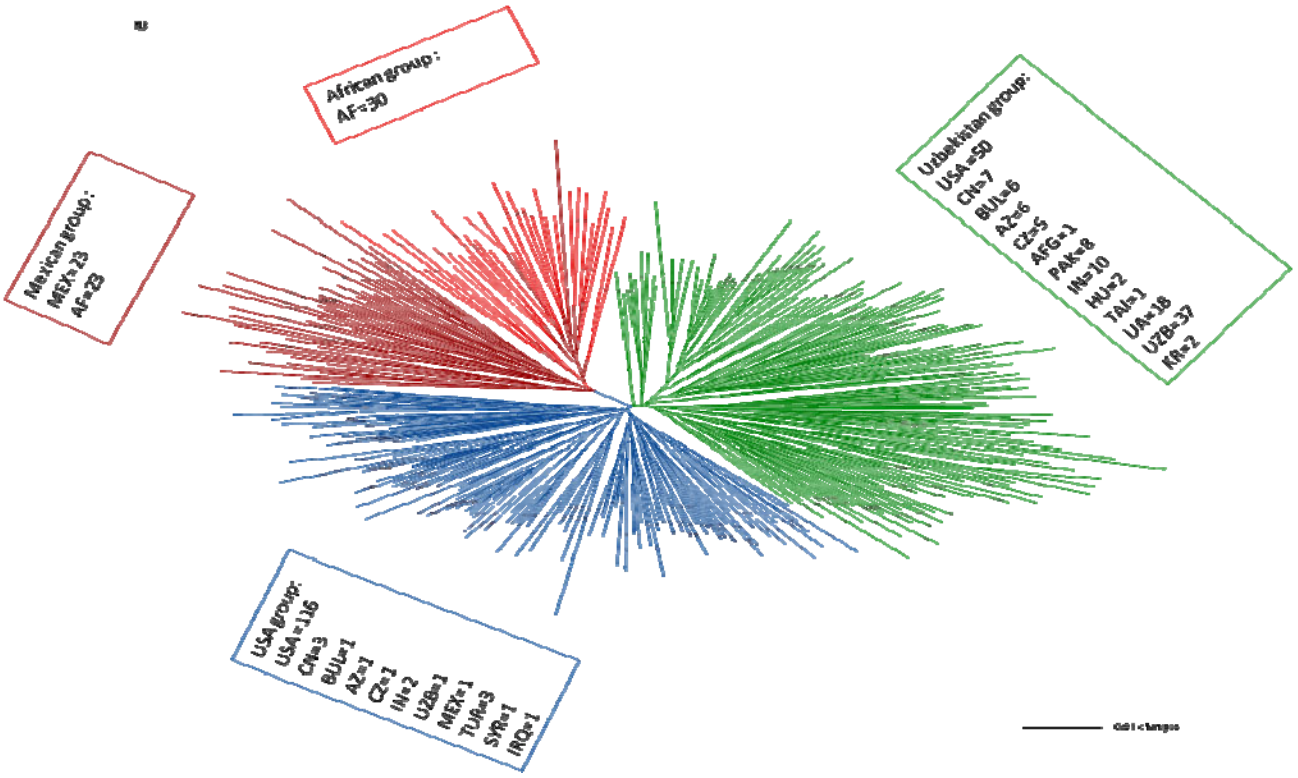


**Fig. 1. Unrooted NJ-phylogenetic tree of** G. hirsutum **accessions from** Uzbek cotton germplasm resources. Afghanistan (AFG), Azerbaijan (AZR), Bulgaria (BU), China (CN), Czechoslovakia (CZ), Hungarian (HU), India (IN), Iraq (IRQ), Korea (KR), Pakistan (PAK), Syria (SYR), Taiwan (TAI), Turkey (TUR), Ukraine (UA), United States (USA) and Uzbekistan (UZB). **Branch lengths are shown.**


**Fig. 1. Diversity on plant bush architecture observed in *G. hirsutum* germplasm; A–creeping type; B to E–variations of vertical branchy (erect) type of bushes; F to I–variations of vertical compact type of bushes.**


**Fig.2. Diversity on stem color observed in *G. hirsutum* germplasm: A–no antocyanin color; B–weak antocyanin color; C–middle antocyanin color; and D–strong antocyanin color.**


**Fig. 3. Diversity on stem pubescence observed in *G. hirsutum* germplasm: A–naked; B–medium; C–sparse; D–hairy.**

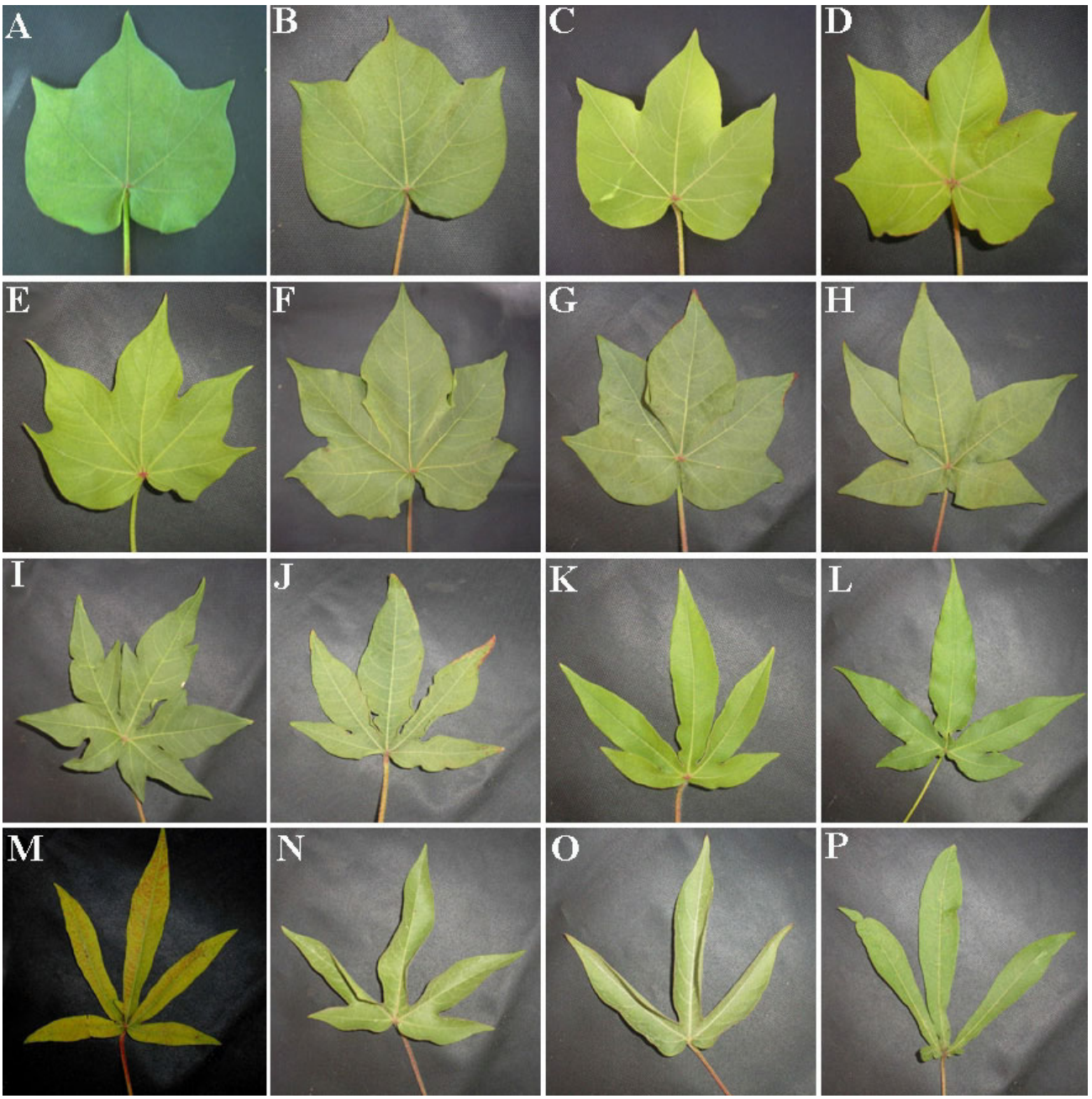


**Fig. 4.Diversity on leaf plate shape observed in *G. hirsutum* germplasm: A to C–variations of 3-lobbed leaf plates; D– 5-lobbed leaf plate; E to H–variations of palmate type leaves; I to L–variations of semi-digitate and digitate type leaves; M to P–variations of okra and super okra type leaves (P)**

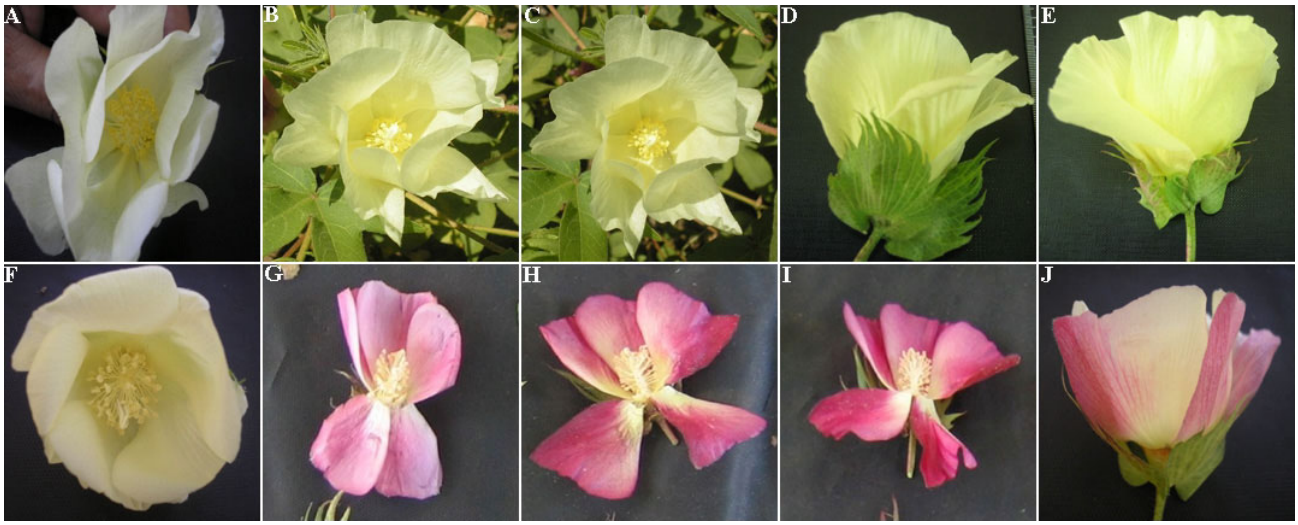


**Fig. 5. Diversity on flower color observed in *G. hirsutum* germplasm: A –white-yellow; B and C–light yellow; D and E–yellow; F–cream color; G–lavender; H and I–pink; J–bicolor.**

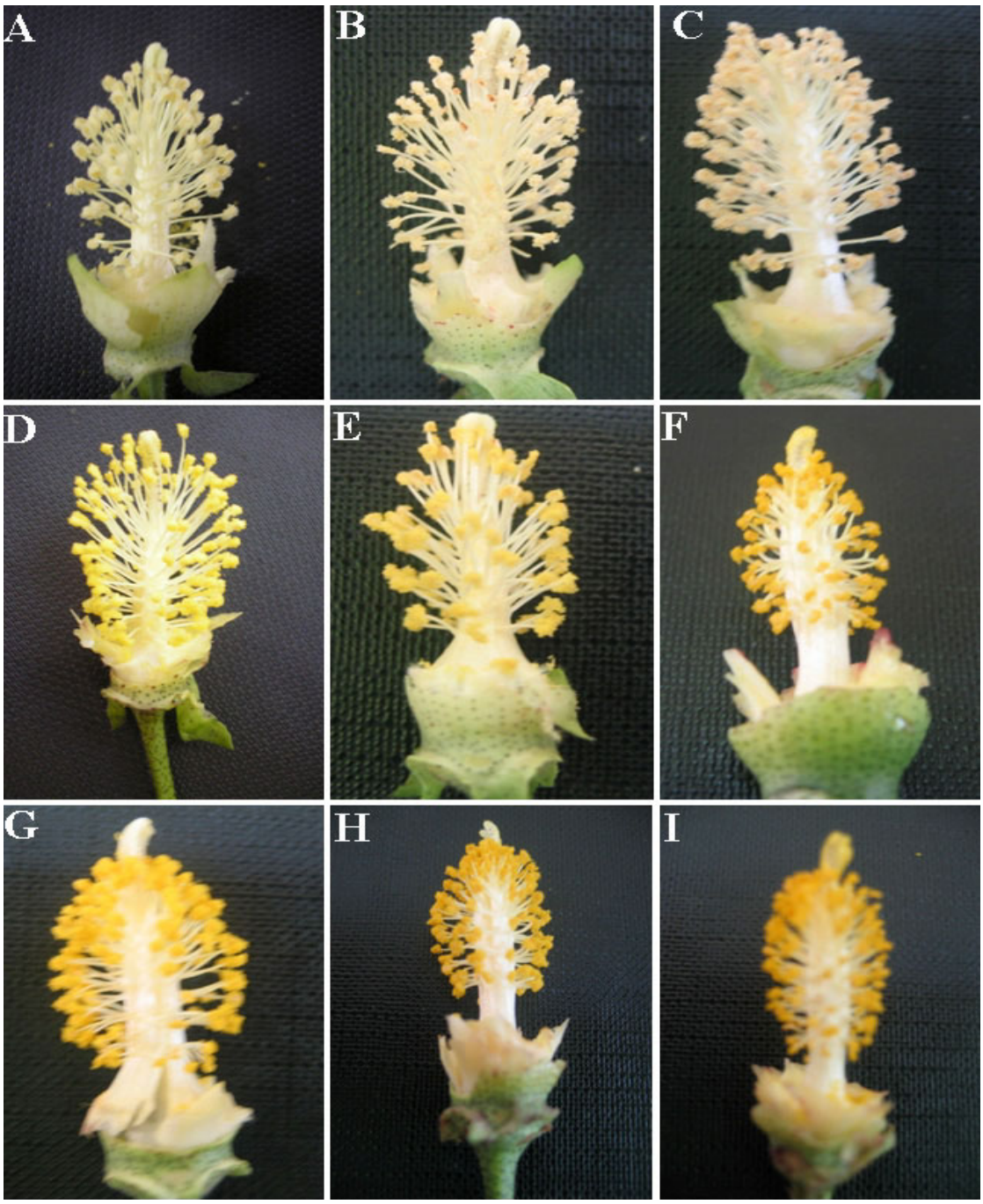


**Fig.6. Diversity on pollen color observed in *G.hirsutum* germplasm: A – light yellow; B– light cream color; C–cream color; D to G–variations of yellow pollen; H and I–tawny.**

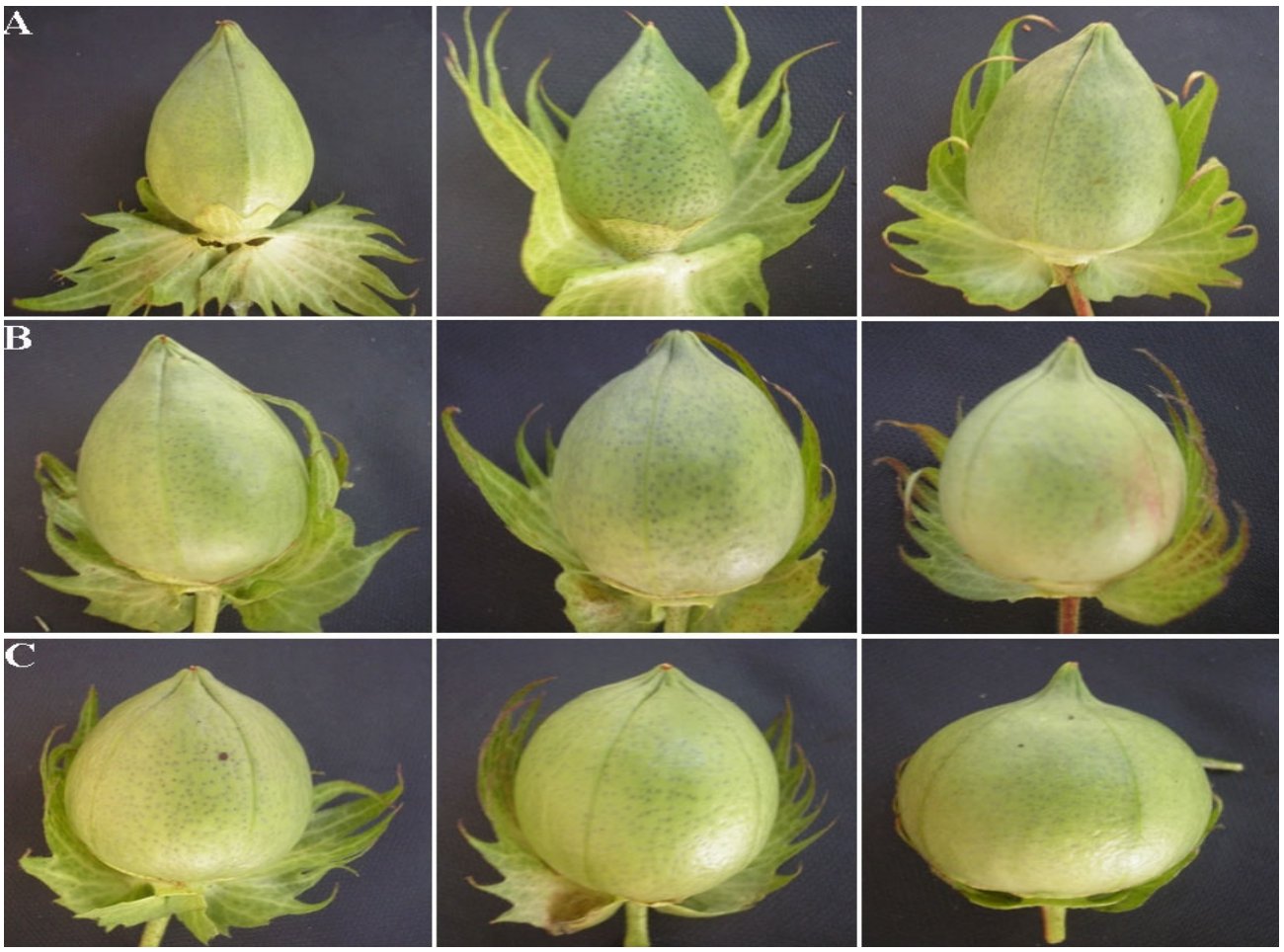


**Fig.7. Diversity on boll shape observed in *G. hirsutum* germplasm: row A- cone-shaped; row B–oval-shaped; row C–round-shaped**.