Pl Syst Evol 268: 199–208 (2007)
DOI 10.1007/s00606-007-0565-7
Printed in The Netherlands

Plant Systematics and Evolution

Phylogenetic diversity and relationship among *Gossypium* germplasm using SSRs markers

Y.-X. Wu[1,2], M. K. Daud[1], L. Chen[1], S.-J. Zhu[1]

[1]Department of Agronomy, College of Agriculture and Biotechnology, Zhejiang University, Hangzhou, P.R. China
[2]College of Agronomy, Shanxi Agricultural University, Taigu Shanxi, P.R. China

Received 29 January 2007; Accepted 9 May 2007; Published online 1 August 2007

Abstract. *Gossypium* species represent a vast resource of genetic multiplicity for the improvement of cultivated cotton. To determine genetic diversity and relationships within a diverse collection of *Gossypium*, we employed 120 SSR primers on 20 diploid species representing seven basic genome groups of the genus *Gossypium*, five AD allotetraploid cotton accessions while *T. populnea* served as an outgroup species. Out of 120 SSR primers, 49 pairs are polymorphic, which produced a total of 99 distinct alleles with an average of 2.0 alleles per primer pair. A total of 1139 major SSR bands were observed. Genetic similarities among all the diploid species ranged from 0.582 (between *G. herbaceum* and *G. trilobum*) up to 0.969 (between *G. arboreum* and *G. herbaceum*). Phylogenetic trees based on genetic similarities were consistent with known taxonomic relationships. The results also indicated that *G. raimondii* is the closest living relative of the ancestral D-genome donor of tetraploid species and the A-genome donor is much similar to the present-day *G. herbaceum* and *G. arboreum*. Ancient tetraploid cotton species were formed by hybridizing and chromosome doubling between them, then different tetraploid cotton species appeared by further geographical and genetic isolation and separating differentiation. The results showed that SSRs could be an ideal means for the identification of the genetic diversity and relationship of cotton resources at the genomic level.

Keywords: genetic diversity; tetraploid evolution; *Gossypium*; SSR

Introduction

The genus '*Gossypium*' comprises about 50 species including the world's most important fiber crop. It ranges from shrubs to small trees in arid regions spreading throughout the tropics and subtropics (Wendel 1995). Mostly it has diploid species ($2n = 26$), but the five species coming from the Western Hemisphere are classic genomic allopolyploids ('AD-genome', $2n = 52$), which is the result of hybridization between two diploid species ('A-genome' and 'D-genome'), possibly in the mid-Pleistocene (Wendel 1989, Seelanan et al. 1997, Cronn et al. 2002, Senchina et al. 2003). Since their parental genome groups exist in diploid

Correspondence: Zhu Shui-Jin, Department of Agronomy, College of Agriculture and Biotechnology, Zhjiang University, Hangzhou 310029, P.R. China
e-mail: shjzhu@zju.edu.cn

form in different hemispheres, the question "how and when allopolyploid cotton formed" has stimulated discussion for more than 50 years. According to some authors, *Gossypium* had an ancient, perhaps Cretaceous origin, due to its global distribution and high level of cytogenetic and morphological diversity. However, others suggest the origin of allopolyploids in agricultural times, thus presenting a scenario that involved human transfer of an African or Asiatic A-genome cultigen to the New World, followed by deliberate or accidental hybridization with a wild D-genome species (reviewed by Senchina et al. 2003).

Four different species have been independently domesticated at different times in human history, including the New World allopolyploids, *G. hirsutum* L. and *G. barbadense* L., and the Old World diploids, *G. herbaceum* L. and *G. arboreum* L. (Brubaker et al. 1999). These species provide most of the world's textile fiber and are important sources of oil and cotton seed meal (Pillay and Myers 1999). But currently, over 90% of the world's cotton derives from *G. hirsutum*, while *G. barbadense* produces most of the remaining commercial cotton.

Research on genetic diversity and phylogenetic relationships among cultivated and wild cotton species is necessary for better understanding of cotton evolution. In addition, molecular characterization of germplasm aids plant breeders in selecting appropriate materials for further genetic improvement of cultivars.

Molecular markers have been widely used in genetic analyses, breeding and investigations of genetic diversity and the relationship between cultivated and wild species because they have several advantages over other morphological markers, including high polymorphism and independence from effects related to environmental conditions and the physiological stage of the plant.

The utilities of comparative morphology (Fryxell 1979, 1992), intercross fertility and cytology (Endrizzi et al. 1985), allozymes (Wendel et al. 1992), RFLP (Wendel and Brubaker 1993), RAPD (Multani and Lyon 1995, Tatineni et al. 1996, Iqbal et al. 1997, Lu and Myers 2002), AFLP (Pillay and Myers 1999, Abdalla et al. 2001) and other molecular markers (Wendel and Albert 1992; Cronn et al. 1996, 2002) have all been used successfully in genetic diversity analyses in cotton species. In spite of the success of these methods, the level of polymorphism detectable is low, with allozymes and RFLP markers having particularly low intra- and inter-specific polymorphism, and these types of markers tend not to be efficient when applied to the genotyping of large germplasm collections (Liu et al. 2000).

To date, although SSRs are well established for human and mammalian genetics (Hanslik et al. 2000, Kanthaswamy et al. 2001), these markers have become available in plant species. They have been used to investigate genetic diversity in many plant species, including those of maize (Taramino and Tingey 1996), soybean (Morgante and Olivieri 1993), *Brassica* spp. (Poulsen et al. 1993), rice (Davierwala et al. 2000), barley (Macaulay et al. 2001), pearl millet (Chowdari et al. 1998), *Arabidopsis* (Depeige et al. 1995), tomato (Broun and Tanksley 1996), conifers (Tsumura et al. 1997), sorghum (Dean et al. 1999), wheat (Huang et al. 2002), spelt (Bertin et al. 2001), wild emmer wheat (Fahima et al. 2002), *Elymus caninus* (Sun et al. 2001) and ryegrass (Jones et al. 2001). However, SSRs have not been explored for the resolution of phylogenetic relationships in *Gossypium* spp.

The present phylogenetic relationships of the cotton genome groups was proposed by Wendel and Cronn (2003) according to recent molecular phylogenetic investigations, including largely cpDNA restriction site variation, and nucleotide sequence variation of a limited number of selected chloroplast genes, nuclear ribosomal DNA and low-copy nuclear genes (Wendel and Albert 1992; Cronn et al. 1996, 2002). Nevertheless, several significant questions and/or uncertainties about their phylogeny need to be further investigated. This work will provide evidence for the genetic diversity of *Gossypium* and valuable information for further conservation and cotton breeding programs. In addition, this work will elucidate the possible A- and D- genome donor of extant

tetraploid cotton species, and evaluate the feasibility of SSR markers in resolving phylogenetic relationships of cotton germplasm.

Materials and methods

Plant materials. Twenty-five cotton accessions including cultivated and wild (20 diploid and five tetraploid cotton species) were evaluated in the present study. One semi-wild *G. hirsutum* was also used. In order to serve as an outgroup species, *Thespesia populnea* was included as well. It was chosen because of its known distant relationship to the cotton species complex. The details of all accessions are presented in Table 1.

DNA extraction. All the accessions were grown in the greenhouse of Zhejiang University for DNA isolation. Total genomic DNA was extracted from fresh young leaves of each accession using the CTAB method (Paterson et al. 1993).

DNA quality was evaluated by photospectrometry using the 260/280 nm absorbance ratio method and by electrophoreses in 0.8% (w/v) agarose gel. Its concentration was estimated at 260 nm and quantified by means of comparison with DNA ladder of DL2000 (TaKaRa) ranging from 100 to 2000 bp. The stock DNA samples were stored at −20°C and working DNA samples (containing 50 ng μL^{-1}) at 4°C.

PCR amplification and SSR analysis. The sequences of SSR primers were obtained from Cotton DB database (http://algodon.tamu.edu/htdocs2cotton/cot2tondb.html). They were synthesized by the company of Shanghai biotechnology in China.

DNA amplification was carried out in a volume of 20μL containing 2μL 10 × buffer, 1.6μL $MgCl_2$ (25 mM), 0.2μL dNTPs (10 mM), 6μL template DNA (50 ng μL^{-1}), 2μL primers (2.5μM) (1μL of forward and reverse primer each), 0.2μL Taq polymerase (5 U/μL, Sangon), 8μL ddH_2O.

The following PCR profile was used in a DNA Mastercycler, Eppendorf, Germany. Initial denaturation at 94°C for 3 minutes followed by 35 cycles of denaturation at 94°C for 50 seconds, annealing at 58°C for 50 seconds, extension at 72°C for 2 minutes, followed by a final extension at 72°C for 10 minutes and a 4°C hold.

In all cases, PCR reactions were performed at least twice in order to ensure that absence was natural and was not due to failed reaction.

Alleles were scored on the basis of the presence or absence of amplified fragments. Each SSR fragment was considered a different locus as it corresponds to a unique position in the genome.

Electrophoresis. The PCR products were analyzed directly on 2.0% agarose gels in TBE buffer at 5 V/cm. The DNA fragments were visualised by staining with ethidium bromide and transillumination under UV light. A DNA ladder, labeled with DL2000 (TaKaRa), ranging from 100 to 2000 bp, was used to determine the size of the SSR fragments.

Data analysis and cluster analysis. The effectiveness of SSR markers in detecting polymorphism in the analyzed *Gossypium* genotypes was evaluated through trivial statistical parameters such as the

Table 1. Details of cotton accessions used in the present study

No	Materials	Genome	No	Materials	Genome
1	*G. arboreum*	A_2	14	*G. gossypioides*	D_6
2	*G. herbaceum*	A_1	15	*G. raimondii*	D_5
3	*G. africanum*	A_{1-2}	16	*G. aeysianum*	E_3
4	*G. capitis-viridis*	B_3	17	*G. somalense*	E_2
5	*G. anomalum*	B_1	18	*G. stocksii*	E_1
6	*G. sturtianum*	C_1	19	*G. longicalyx*	F
7	*G. australe*	G	20	*G. bickii*	G
8	*G. nelsonii*	G	21	*G. hirsutum*	$(AD)_1$
9	*G. aridum*	D_4	22	*G. barbadense*	$(AD)_2$
10	*G. thurberi*	D_1	23	*G. mustillinum*	$(AD)_4$
11	*G. klotzschianum*	D_{3-k}	24	*G. darwinii*	$(AD)_5$
12	*G. davidsonii*	D_{3-d}	25	*G. hirsutum* (semi-wild)	$(AD)_1$
13	*G. trilobum*	D_9	26	*T. populnea*	

number of polymorphism primer pairs and polymorphism bands, average polymorphism per primer pair, similarity coefficients and so on.

The SSR reproducible fragments were classified as present (1) or absent (0), and were typed into a computer file as a binary matrix, one for each molecular marker. The similarity coefficient was used to construct a dendrogram by the Unweighted Pair Group Method with Arithmetic Averages (UPGMA) according to Rohlf (1993). Clustering analyses were performed using NTSYS-pc (version 2.02h) to calculate the genetic similarity matrices.

Results

SSR analysis. For the purpose of genetic diversity studies, codominant SSR markers are coded as binary data so as to match the binary dominant data-set. In this study, the SSR markers provided a powerful assay for discriminating genetic diversity among cotton accessions.

120 pairs of SSR primers were used for amplication. From the profiles obtained, 49 pairs were polymorphic, which produced a total of 99 distinct alleles with an average of 2.0 alleles per primer pair. A total of 1139 major distinct SSR bands were observed. These reproducible polymorphic DNA fragments ranged from 50 to 500 base pairs (bp), which reflect a large difference in the number of repeats between the different alleles, and partial primer pairs amplification products are shown in Fig. 1.

Genetic relationships and diversity. Using these SSR polymorphisms, a dendrogram of the 26 accessions studied was constructed with an Unweighted Paired Group Method using Arithmetic Averages (UPGMA) clustering algorithm (Fig. 2). The dendrogram separated the 26 accessions into three main groups at a similarity coefficient value of 0.680. The first group consisted of 20 diploid cotton species. In this group, The A-genome cotton species and *G. nelsonii* differed considerably from the other and were clustered alone. The B-, C-, E-, D-, G- and F-genomes were clustered together, and within this group, *G. longicalyx* was distinctly differentiated from the others and then B-, C-, E-, D-, G-genome were divided into five groups respectively. The second group consisted of 5 tetraploid cotton species. Within the second group, *G. barbadense*, *G. mustillinum* and *G. darwinii* clustered together, while *G. hirsutum* and semi-wild *G. hirsutum* were more nearer. The *T. populnea* species formed an outgroup by itself (the third group).

Genetic relationships of *G. darwinii* with *G. barbadense* and *G. hirsutum* were studied using previously generated allozyme data, and significant introgression of *G. hirsutum*

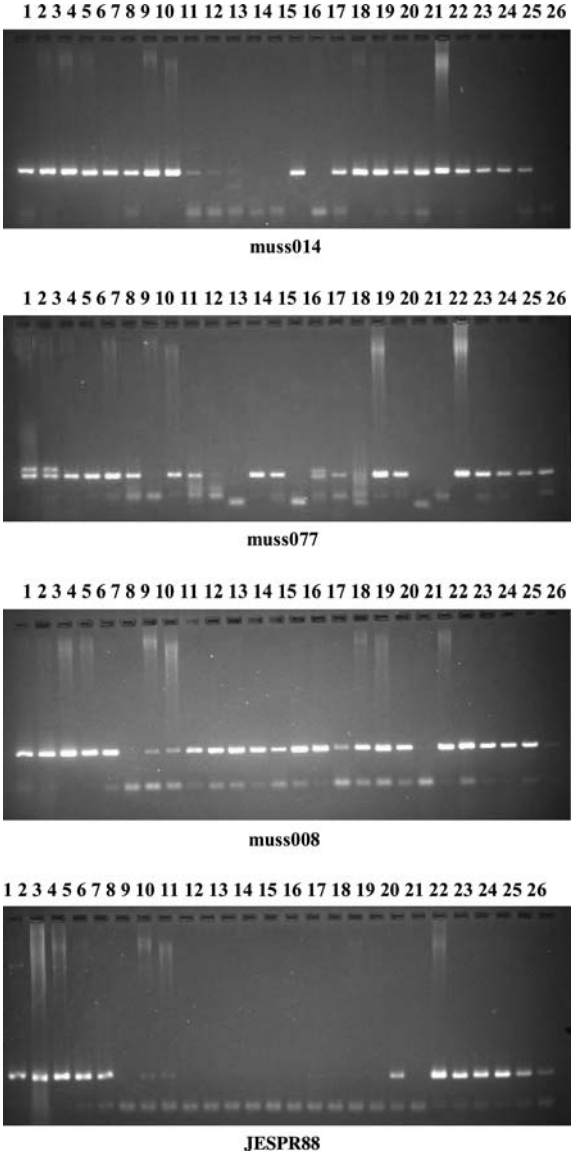

Fig. 1. SSR amplification products generated by partial primer pairs

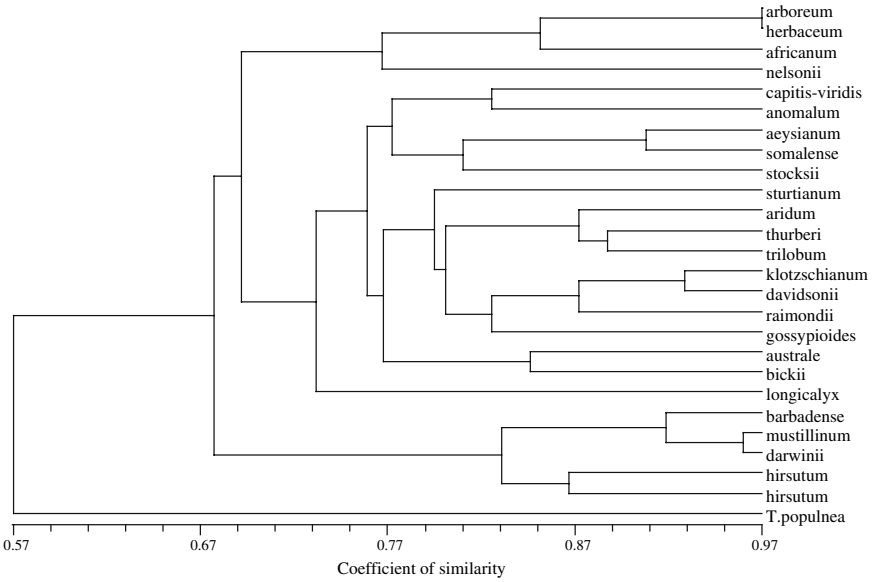

Fig. 2. The dendrogram of *Gossypium* spp. based on SSR molecular markers

alleles was detected. However, morphological considerations support the hypothesis that much of *G. darwinii*'s diversity stems from interspecific gene flow from *G. barbadense* (Wendel and Percy 1990). This is consistent with our result that *G. darwinii* and *G. barbadense* had a close relationship than *G. hirsutum*.

Genetic similarity coefficients of the diploid cotton species. Genetic similarity coefficients of 20 diploid cotton species based on UPGMA using the NTSYS-pc (Version 2.02h) by SSR were compared.

According to the similarity matrix of 20 diploid cotton species, *G. arboreum* and *G. herbaceum*, which belong to A- genome, had the greatest similarity (0.969), whereas *G. herbaceum* and *G. trilobum* had the least similarity (0.582). The average similarity coefficient of these source germplasms was 0.755. Thus we can conclude, that there are deep genetic backgrounds among diploid cotton germplasms for the genetic and breeding research. Within these diploid cotton species, there are many favorable characteristics such as fiber quality and yield traits, insect and pathogen resistance, tolerance to environmental stresses and so on, which may be available to our development of superior cultivars (Abdalla et al. 2001). So an understanding of the genetic and genomic relationships of cotton species is a critical step for further utilization of extant genetic diversity and genomic information.

Relationship between the species from A or D genome and tetraploid species. On the basis of obtained binary data, a matrix of similarity coefficients between A, D genome and tetraploid cotton species was calculated, and the results are shown in Table 2.

According to Table 2, the mean genetic similarity coefficients of nine diploid species comparing with 5 tetraploid species were from 0.618 to 0.731. The highest genetic similarity coefficient was between *G. raimondii* and five tetraploid species while the lowest was between *G. trilobum* and 5 tetraploid species.

By statistical analysis, we can find that the genetic similarity coefficients of *G. raimondii* and five tetraploid species were significant at 0.01 level when comparing with *G. aridum*, *G. thurberi*, *G. trilobum* and significant at 0.05 level comparing with *G. klotzschianum*, *G. davidsonii*, *G. gossypioides*. But there was no significant difference in their genetic simi-

Table 2. Similarity coefficients between tetraploid species and A or D genome species

Materials	Genetic similarity coefficients									
	G. arboreum	*G. herbaceum*	*G. aridum*	*G. thurberi*	*G. klotzschianum*	*G. davidsonii*	*G. trilobum*	*G. gossypioides*	*G. raimondii*
	A_2	A_1	D_4	D_1	D_{3-k}	D_{3-d}	D_9	D_6	D_5
G. hirsutum	0.765	0.755	0.643	0.694	0.724	0.694	0.643	0.714	0.735
G. barbadense	0.653	0.663	0.612	0.602	0.694	0.684	0.571	0.684	0.745
G. mustillinum	0.714	0.704	0.592	0.622	0.673	0.663	0.571	0.684	0.724
G. darwinii	0.714	0.704	0.612	0.663	0.673	0.663	0.612	0.684	0.704
G. hirsutum	0.673	0.663	0.714	0.684	0.755	0.745	0.694	0.724	0.745
Mean	0.704	0.698	0.635	0.653	0.704	0.690	0.618	0.698	0.731

larity coefficients between *G. arboreum* and *G. herbaceum* of A-genome species by comparing with 5 tetraploid species.

Overall, among D-Genome species, *G. raimondii* is the possible sole D genome donor of tetraploid cotton species. Among A-Genome species, *G. arboreum* and *G. herbaceum* are all the possible closest living relatives of the ancestral A-genome donor and the position of them is equal on the course of formation of tetraploids. These two A-genome diploid species are phylogenetically equidistant from extant tetraploid cotton species. Ancient tetraploid cotton species were formed possibly by hybridizing and chromosome doubling between the archaic species related to *G. arboreum* or *G. herbaceum* and *G. raimondii*, then different tetraploid species might have appeared by further geographical and genetic isolation and separating differentiation. The theory of one sole origin about tetraploid cotton (Beasley 1940, Gerstel 1953) was further testified by the present results.

According to the results, we can claim that (1) the results are in agreement with the genetic background of cotton germplasms which was proposed by Wendel and Cronn (2003). (2) *G. arboreum* and *G. herbaceum* are the possible closest living relatives of the ancestral A-genome donor and *G. raimondii* is the possible sole D-geome donor of tetraploid cotton species. (3) SSRs are the useful method for identification of genetic diversity and relationship of cotton germplasm.

Discussion

The diploid species have been divided into cytologically, genome-based groups (reviewed by Endrizzi et al. 1985, Stewart 1995), three primary centers of diversity, that is the African–Asian species (A-, B-, E-, and F-genomes), the Australian species (C-, G-, and K-genomes), and the New World species (D-genome) (Small and Wendel 2000). Each genome represents a group of morphologically similar species that can rarely form hybrids with species from other genomic groups. There are five allotetra-

ploid species with genome composition AD (2n = 4x = 52) including the commercially important *G. hirsutum* (''Upland cotton'') and *G. barbadense* (''Pima'' and ''Egyptian'' cotton), and they are traditionally considered to be allotetraploids (2n = 52), containing A- and D-subgenomes and being endemic to the New World (Fryxell 1992).

Ever since the discovery that tetraploid *Gossypium* species contain two different genomes, investigators have attempted to address the question of parentage, that is, which of the modern species of A-genome and D-genome diploids best serve as models of the progenitor genome donors. Over the decades, a diverse array of tools have addressed this question (reviewed in Wendel and Cronn 2003), collectively demonstrating that the best extant models of the ancestral genome donors are *G. arboreum* or *G. herbaceum* (A-genome) and *G. raimondii* (D-genome). These two genome groups diverged from each other early in the evolution of the genus, perhaps 7 to 11 million years ago (Seelanan et al. 1997, Cronn et al. 2002).

Earlier researchers suggested that the New World tetraploid species arose some 1-2 million years ago through the hybridization of an Old World taxon of the 'A genome' cytogenetic group, related to the present day species *G. herbaceum* and *G. arboreum* L. (2n = 2x = 26), with a taxon of the 'D-genome' group, related to the new world species *G. raimondii* and *G. gossipioides* L. (2n = 2x = 26) (Beasley 1940, 1942; Wendel et al. 1992). After polyploidization, the nascent 'AD' disomic tetraploid (2n = 2x = 52) gave rise to five extant tetraploid species, including *G. hirsutum* and *G. barbadense* which dominate world production of cotton fiber and seed (Iqbal et al. 2001).

Senchina et al. (2003) also suggested that the A-genome progenitor of polyploid cotton are thus interchangeable for the present purposes and the two diploid genomes, A and D, experienced approximately 5 million years of evolution in isolation from one another prior to their reunion at the time of polyploidy formation during the Pleistocene.

According to our present results, it is suggested that ancient tetraploid cotton species were formed possibly by hybridizing and chromosome doubling between the species of *G. arboreum* or *G. herbaceum* and *G. raimondii*, then different tetraploid species might have appeared by further geographical and genetic isolation and separating differentiation. This result is in agreement with the theory of sole origin about tetraploid cotton (Beasley 1940, Gerstel 1953).

In our current study, *G. nelsonii* has a close relationship with the A-genome, but a far distant correlation with other genomes. This result is not in agreement with Stewart (1995)'s finding who considered *G. nelsonii* to be G-genome, but a previous study suggested that it should belong to C-genome (Fryxell 1992). According to our study, perhaps it might belong to another new genome. Nevertheless, this is our own suggestion based on the present results. It needs more research work to approve our findings and to complete *Gossypium* classification.

Microsatellites, also called Simple Sequence Repeats (SSRs), are abundant throughout the eukaryotic genome (Tautz and Renz 1989, Kijas et al. 1995) and evolve rapidly (Levinson and Gutman 1987). They are of a variable number (up to 100) of tandem repeats of 1-6 base pair (Tautz 1989). They are useful molecular markers due to their abundance, uniform distribution, high polymorphism, codominance, rapid amplification by PCR, relatively simple to interpret and easily accessed by other laboratories via published primer sequences (Saghai-Maroof et al. 1994). In similar experiments carried out on other plant species, the superiority of the SSR marker system in comparison with the RAPD marker system is usually observed (Pejic et al. 1998, Palombi and Damiano 2002, Rajora and Rahman 2003).

The availability and abundance of microsatellite markers throughout the cotton genome make them particularly useful in genetic diversity studies of cotton (Reddy et al. 2001), with in excess of 1000 microsatellite primers having already been isolated from cotton DNA genome

libraries (Nguyen et al. 2004). The results of our work using SSR markers in cotton species suggest that they may provide an outstanding new tool for genetic diversity analysis. These results not only further confirm the previously phylogenetic analysis of cotton species and the utility of repetitive DNA sequences for phylogenetic analysis of the genus *Gossypium*, but also provide new insights into the phylogeny of the *Gossypium* genus. It is a very useful method in the identification of the genetic diversity and relationship for cotton germplasm at the genomic level.

Authors pay special thanks to Prof. Zhang xianyin for the technical assistance that made this research possible and providing valuable suggestions for the paper. This research is supported by National key basis research and development project. (973 project) (2004CB117305) and National Natural Science Foundation of China (No. 30471108).

References

Abdalla AM, Reddy OUK, El-Zik KM, Pepper AE (2001) Genetic diversity and relationships of diploid and tetraploid cottons revealed using AFLP. Theor Appl Genet 102: 222–229

Beasley JO (1940) The origin of American tetraploid *Gossypium* species. Amer Naturalist 74: 285–286

Beasley JO (1942) Meiotic chromosome behavior in species, species hybrids, haploids and induced polyploids of *Gossypium*. Genetics 27: 25–54

Bertin P, Gregoire D, Massart S (2001) Genetic diversity among European cultivated spelt revealed by microsatellites. Theor Appl Genet 102: 148–156

Broun P, Tanksley SD (1996) Characterization and genetic mapping of simple repeat sequences in the tomato genome. Molec Gen Genet 250: 39–49

Brubaker CL, Bourland FM, Wendel JF (1999) The origin and domestication of cotton. In: Smith WC (ed) Cotton: origin, history, technology and production. Wiley, New York

Brubaker CL, Paterson AH, Wendel JF (1999) Comparative genetic mapping of allotetraploid cotton and its diploid progenitors. Genome 42: 184–203

Chowdari KV, Davierwala AP, Gupta VS, Ranjekar PK, Govila OP (1998) Genotype identification and assessment of genetic relationships in pearl millet [*Pennisetum glaucum* (L.) R. Br.] using microsatellites and RAPDs. Theor Appl Genet 97: 154–162

Cronn RC, Small RL, Haselkorn T, Wendel JF (2002) Rapid diversification of the cotton genus (*Gossypium*: Malvaceae) revealed by analysis of sixteen nuclear and chloroplast genes. Amer J Bot 89: 707–725

Cronn RC, Zhao XP, Paterson AH, Wendel JF (1996) Polymorphism and concerted evolution in a tandemly repeated gene family: 5S ribosomal DNA in diploid and allopolyploid cottons. J Molec Evol 42: 685–705

Davierwala AP, Ramakrishna W, Ranjekar PK (2000) Sequence variations at a complex microsatellite locus in rice and its conservation in cereals. Theor Appl Genet 101: 1291–1298

Dean RE, Dahlberg JA, Hopkins MS, Mitchell SE, Kresovich S (1999) Genetic redundancy and diversity among 'Orange' accessions in the U.S. national sorghum collection as assessed with simple sequence repeat (SSR) markers. Crop Sci 39: 1215–1222

Depeige A, Goubely C, Lenior A, Cocherel S, Picard G, Raynal M, Grellet F, Delseny M (1995) Identification of the most represented repeated motifs in *Arabidopsis thaliana* microsatellite loci. Theor Appl Genet 91: 160–168

Endrizzi JE, Turcotte EL, Kohel RJ (1985) Genetics, cytology, and evolution of cotton. Adv Genet 23: 271–375

Fahima T, Roder MS, Wendehake K (2002) Microsatellite polymorphism in natural populations of wild emmer wheat, *Triticum dicoccoides*, in Isreal. Theor Appl Genet 104: 17–29

Fryxell PA (1979) The natural history of the cotton tribe. Texas A&M University Press, College Station

Fryxell PA (1992) A revised taxonomic interpretation of *Gossypium* L. (Malvaceae). Rheedea 2: 108–165

Gerstel DU (1953) Chromosomal translocation in interspecific hybrids of the genus *Gossypium*. Evolution 7: 235–244

Hanslik S, Harr B, Brem C, Schlotterer C (2000) Microsatellite analysis reveals substantial genetic differentiation between contemporary New

World and Old World Holstein Friesian populations. Anim Genet 31: 31–38

Huang XQ, Borner A, Roder MS (2002) Assessing genetic diversity of wheat (*Triticum aestivum* L.) germplasm using microsatellite markers. Theor Appl Genet 105: 699–707

Iqbal MJ, Aziz N, Saeed NA, Zafar Y (1997) Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. Theor Appl Genet 94: 139–144

Iqbal MJ, Reddy OUK, El-Zik KM, Pepper AE (2001) A genetic bottleneck in the 'evolution under domestication' of upland cotton *Gossypium hirsutum* L. examined using fingerprinting. Theor Appl Genet 103: 547–554

Jones ES, Dupal MP, Kolliker R (2001) Development and characterization of simple sequence repeat (SSR) markers for perennial rye grass (*Lolium perenne* L.). Theor Appl Genet 102: 405–415

Kanthaswamy S, Bininda-Emonds ORP, Warden C, Viray JL, Smith DG (2001) Use of SSR fragment length homozygotes for orangutan systematics. Primates 42: 35–45

Kijas JMH, Fowler JCS, Thomas MR (1995) An evaluation of sequence tagged microsatellite site markers for genetic analysis within *Citrus* and related species. Genome 38: 349–355

Levinson G, Gutman GA (1987) Slipped-strand mispairing: A major mechanism for DNA sequence evolution. Molec Biol Evol 4: 203–221

Liu S, Saha S, Stelly D, Burr B, Cantrell RG (2000) Chromosomal assignment of microsatellite loci in cotton. J Heredity 91: 326–332

Lu HJ, Myers GO (2002) Genetic relationships and discrimination of ten influential upland cotton varieties using RAPD markers. Theor Appl Genet 105: 325–331

Macaulay M, Ramsay L, Powell W (2001) A representative, highly informative 'genotyping set' of barley SSRs. Theor Appl Genet 102: 801–809

Morgante M, Olivieri AM (1993) PCR-amplified microsatellites as markers in plant genetics. Pl J 3: 175–182

Multani DS, Lyon BR (1995) Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38: 1005–1008

Nguyen TB, Giband M, Brottier P, Risterucci AM, Lacape JM (2004) Wide coverage of the tetraploid cotton genome using newly developed microsatellite markers. Theor Appl Genet 109: 167–175

Palombi M, Damiano C (2002) Comparison between RAPD and SSR molecular markers in detecting genetic variations in kiwifruit (*Actinida deliciosa* A. Chev.). Pl Cell Rep 20: 1061–1066

Paterson AH, Brubaker CL, Wendel JF (1993) A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. Pl Molec Biol Rep 11: 122–127

Pejic I, Ajmone-Marsan P, Morgante M, Kozumplick V, Castiglioni P, Taramino G, Motto M (1998) Comparative analysis of genetic similarity among maize inbred lines detected by RFLPs, RAPDs, SSRs, and AFLPs. Theor Appl Genet 97: 1248–1255

Pillay M, Myers GO (1999) Genetic diversity in cotton assessed by variation in ribosomal RNA genes and AFLP markers. Crop Sci 39: 1881–1886

Poulsen GB, Hahl G, Weising K (1993) Abundance and polymorphism of simple repetitive DNA sequences in *Brassica napus* L. Theor Appl Genet 85: 994–1000

Rajora OP, Rahman MH (2003) Microsatellite DNA and RAPD fingerprinting, identification and genetic relationships of hybrid poplar (*Populus canadiensis*) cultivars. Theor Appl Genet 106: 470–477

Reddy OU, Pepper AE, Abdurakhmonov I, Saha S, Enkins JN, Brooks T, Bolek Y, El-zik KM (2001) New dinucleotide and trinucleotide microsatellite marker resources for cotton genome research. Crop Sci 5: 103–113

Rohlf FJ (1993) NTSYS-pc, numerical taxonomy and multivariate analysis system. Version 2.02h, Applied Biostatistical Inc, New York

Saghai-Maroof MA, Biyashev RM, Yang GP (1994) Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal location and population dynamics. Proc Natl Acad Sci USA 91: 5466–5470

Seelanan T, Schnabel A, Wendel JF (1997) Congruence and consensus in the cotton tribe. Syst Bot 22: 259–290

Senchina DS, Richard IA, Cronn C, Liu B, Rong JK, Noyes DR, Paterson AH, Wing RA, Wilkins TA, Wendel JF (2003) Rate variation among nuclear genes and the age of polyploidy in *Gossypium*. Molec Biol Evol 20(4): 633–643

Small RL, Wendel JF (2000) Phylogeny, duplication, and intraspecific variation of *Adh* sequences in New World diploid cottons (*Gossypium* L., Malvaceae) Molec Phylogenet Evol 16(1): 73–84

Stewart JM (1995) Potential for crop improvement with exotic germplasm and genetic engineering. In: Constable GA, Forrester NW (eds) Challenging the future: proceedings of the world cotton research conference, CSIRO, Melbourne pp 313–327

Sun GL, Diaz O, Salomon B (2001) Genetic diversity and structure in a natural *Elymus caninus* population from Denmark based on microsatellite and isozyme analyses. Pl Syst Evol 227: 235–244

Taramino G, Tingey S (1996) Simple sequence repeats for germplasm analysis and mapping in maize. Genome 39: 277–287

Tatineni V, Cantrell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Sci 36: 186–192

Tautz D (1989) Hyper variability of simple sequences as a general source for polymorphic DNA markers. Nucl Acids Res 17: 6463–6471

Tautz D, Renz M (1989) Simple sequences are ubiquitous repetitive components of eukaryotic genomes. Nucleic Acids Res 12: 4127–4138

Tsumura Y, Suyama Y, Yoshimura Y, Mukai Y (1997) Sequence-Tagged-Sites (STSs) of cDNA clones in *Cryptomeria japonica* and their evaluation as molecular markers in conifers. Theor Appl Genet 94: 764–772

Wendel JF (1989) New World tetraploid cottons contain Old World cytoplasm. Proc Natl Acad Sci USA 86: 4132–4136

Wendel JF (1995) Cotton. In: Smartt J, Simmonds NW (eds) Evolution of crop plants, 2nd edn. Longman Scientific, Essex, pp 358–366

Wendel JF, Albert VA (1992) Phylogenetics of the cotton genus (*Gossypium*): Character-State Weighted Parsimony Analysis of chloroplast DNA restriction site data and its systematic and biogeographic implications. Syst Bot 17: 115–143

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

Wendel JF, Brubaker CL, Percival AE (1992) Genetic diversity in *Gossypium hirsutum* and the origin of upland cotton. Amer J Bot 79: 1291–1310

Wendel JF, Cronn RC (2003) Polyploidy and the evolutionary history of cotton. Adv Agron 78(2): 139–186

Wendel JF, Percy RG (1990) Allozyme diversity and introgression in the Galapagos Islands endemic *Gossypium darwinii* and its relationship to continental *G. barbadense*. Biochem Syst Ecol 18(7–8): 517–528