Theor Appl Genet (2005) 111: 1385–1395
DOI 10.1007/s00122-005-0070-8

ORIGINAL PAPER

Jinfa Zhang · Yingzhi Lu · Shuxun Yu

Cleaved AFLP (cAFLP), a modified amplified fragment length polymorphism analysis for cotton

Received: 18 April 2005 / Accepted: 1 August 2005 / Published online: 25 August 2005

Abstract In certain plant species including cotton (*Gossypium hirsutum* L. or *Gossypium barbadense* L.), the level of amplified fragment length polymorphism (AFLP) is relatively low, limiting its utilization in the development of genome-wide linkage maps. We propose the use of frequent restriction enzymes in combination with AFLP to cleave the AFLP fragments, called cleaved AFLP analysis (cAFLP). Using four Upland cotton genotypes (*G. hirsutum*) and three Pima cotton (*G. barbadense*), we demonstrated that cAFLP generated 67% and 132% more polymorphic markers than AFLP in Upland and Pima cotton, respectively. This resulted in 15.5 and 25.5 polymorphic cAFLP markers per AFLP primer combination, as compared to 9.1 and 11.0 polymorphic AFLP. The cAFLP-based genetic similarity (GS) is generally lower than the AFLP-based GS, even though both marker systems are overall congruent. In some cases, cAFLP can better resolve genetic relationships between genotypes, rendering a higher discriminatory power. Given the high-resolution power of capillary-based DNA sequencing system, we further propose that AFLP and cAFLP amplicons from the same primer combination can be pooled as one sample before electrophoresis. The combination produced an average of 18.5 and 31.0 polymorphic markers per primer pair in Upland and Pima cotton, respectively. Using several restriction enzyme combinations before pre-selective amplification in combination with various frequent 4 bp-cutters or 6 bp-cutters after selective amplification, the pooled AFLP and cAFLP will provide unlimited number of polymorphic markers for genome-wide mapping and fingerprinting.

Keywords *Gossypium* spp · Cultivated tetraploid · AFLP · Cleaved AFLP · Genetic similarity

Communicated by R. Hagemann

J. Zhang (✉) · Y. Lu
Department of Agronomy and Horticulture,
New Mexico State University
Las Cruces, NM 88003, USA
E-mail: jinzhang@nmsu.edu
Tel.: +1-505-6463438
Fax: +1-505-6466041

S. Yu
China Cotton Research Institute
455112, Anyang, Henan, China

Introduction

The cotton genus *Gossypium* (Malvaceae) comprises approximately 50 species distributed in various continents except Europe: North, Central, and South Americas (18 species), north-east Africa and Arabia (14 species) and Australia (17 species) (Wendel and Cronn 2003). It has four cultivated species: two New World tetraploid species, *Gossypium hirsutum* L. and *Gossypium barbadense* L. and two Old World diploid species, *Gossypium arboreum* L., and *Gossypium herbaceum* L. Tetraploid Upland cotton (*G. hirsutum*, AD_1; $2n = 4x = 52$), is the predominant cultivated cotton with high yield and wide adaptation, accounting for more than 90% of the world cotton production, while its closely related species, American Pima cotton or Egyptian cotton (*G. barbadense*, AD_2; $2n = 4x = 52$) is grown for its extra long, strong, and fine fiber in Egypt and limited area in a few other countries (e.g., southwestern states of U.S., northwest China, Uzbekistan, Sudan, India, and Pakistan). Two cultivated diploid species, *G. arboreum* L. (A_1) and *G. herbaceum* L. (A_2) are only cultivated in very small acreage in South Asia (China, India, and Pakistan). The two tetraploid species arose about 1–2 million years ago through hybridization between A-genome related extant diploid species and D_5-genome (*Gossypium raimondii* Ulbrich) related species followed by chromosome doubling. The ancestral tetraploid evolved and diverged in the New World following a long distance separation, giving rise to five species including the cultivated *G. hirsutum* and *G. barbadense*, and three wild *Gossypium tomemtosum* Nutall ex Seemann, *Gossypium*

mustelinum Miers ex Watt, and *Gossypium darwinii* Watt.

During the most part of the past century, cotton breeding had made significant contributions to increase cotton yield, improve fiber quality and enhance biotic tolerance. Current and obsolete cultivars and strains in Upland cotton have been and still are the main sources in cotton breeding programs worldwide. However, the desirable and amenable genetic variations for breeders are limited or lacking or difficult to dissect. Due to the narrow genetic base of cotton germplasm that cotton breeders have been utilizing and low efficiency of traditional selection methods, cultivar improvement in cotton has slowed down in the past 10–15 years in the U.S. In fact, the past 10 years has seen cotton yield stagnant. Many factors have been discussed for the contributing causes: narrow germplasm base, shift in breeding method to backcrossing for transgenic introgression, nematodes, and weather changes, among others (May et al. 1995; Meredith 2000; Lewis 2001). A number of studies have suggested that cultivated Upland cotton germplasm possesses a low level of genetic diversity, when evaluated by isozymes, random amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), restricted fragment length polymorphism (RFLP), and simple sequence repeats (SSR) (Wendel et al. 1992; Multani and Lyon 1995; Tatineni et al. 1996; Pilley and Myers 1999; Zuo et al. 2000; Abdalla et al. 2001; Iqbal et al. 2001; Gutierrez et al. 2002; Lu and Myers 2002; Rahman et al. 2002).

The AFLP has been widely used to rapidly generate molecular markers among various organisms from bacteria to plants (Vos et al. 1995). The AFLP analysis combines the reliability of restriction enzyme digestion with the utility of the polymerase chain reaction. Genomic DNA or cDNA is first restricted and followed by ligation of the fragments with adaptors. The ligated fragments are subsequently amplified by PCR using selective AFLP primers with amplified products resolved by denaturing polyacrylamide gel electrophoresis. The AFLP fragments can be detected by radioactive-labeling or fluorescent labeling, or by silver staining the gel. The techniques are highly versatile and can be applied to studies of DNA of any origin and complexity, without prior sequence information. Typically, AFLP fragments are inherited in a Mendelian fashion as dominant or co-dominant markers, making the techniques amenable to tracking inheritance of genetic loci in a segregating population. As with RFLP markers, AFLP detects the presence of point mutations, insertions, deletions, and other genetic rearrangements and is very reproductive and reliable, but with higher multiplex ratio. The AFLP has been used in cotton for linkage map construction (Altaf et al. 1998; Lacape et al. 2003; Brubaker and Brown 2003; Lu et al. 2005), gene mapping (Lacape et al. 2005; Zhang et al. 2005), germplasm diversity assessment (Pilley and Myers 1999; Iqbal et al. 2001; Abdalla et al. 2001; Westengen et al. 2005), and evolutionary study (Liu et al. 2001).

Great genetic diversity and many desirable or potentially desirable genes or traits from *G. barbadense* have encouraged cotton geneticists working on interspecific hybridization for the past century. But the success is limited except for fiber quality improvement in Acala cotton. Using several marker systems such as RFLP, AFLP, and SSR, high-density linkage maps were developed from interspecific hybrid populations between Upland and Pima cotton (Reinisch et al. 1994; Lacape et al. 2003; Mei et al. 2004; Zhang et al. 2003; Nguyen et al. 2004; Rong et al. 2004). Many genes and quantitative trait loci have been mapped (Wright et al. 1998; Saranga et al. 2002; Paterson et al. 2003; Chee et al. 2004; Han et al. 2004; Hinchliffe et al. 2005). However, the number of polymorphic markers has shown to be limited for intraspecific mapping populations in Upland cotton (Shappley et al. 1998; Ulloa et al. 2002, 2005). No linkage map is currently available for Pima cotton. Compared with other plant species, Upland cotton and Pima cotton has much lower level of within-species DNA sequence polymorphisms (Lu et al. 2005) that has impeded the progress in constructing genome-wise linkage map and gene mapping.

In order to take advantages of the AFLP marker system and generate more polymorphic markers for Upland cotton or Pima cotton, we have developed a modified AFLP technique, called cleaved AFLP (cAFLP) based on fluorescent labeling and capillary electrophoresis. In this paper, we will provide evidence that further restriction of AFLP products by a restriction enzyme will release many more polymorphic fragments.

Materials and methods

Plant materials

To compare the resolution power between AFLP and cAFLP, seven genotypes were used: three Pima cotton (Pima S-1, Pima 57-4, and Pima Phytogen 76), and four Upland cotton (ARK8518, TM-1, NM24016, and Acala 1517-99). Pima S-1 and Pima 57-4 is a pair of natural isogenic lines since 57-4 was a double haploid from a haploid mutant isolated from Pima S-1 (Zhang and Stewart 2004), while Phy 76 is a commercial Pima cultivar. TM-1 is a genetic standard for Upland cotton (Kohel et al. 1970); ARK8518 was a breeding line and later released as H1330 from the University of Arkansas (Bourland 1996). NM 24016 was an Acala breeding line with substantial germplasm introgression from *G. barbadense* (Cantrell and Davis 2000), while Acala 1517-99 was an Acala cotton cultivar released also from New Mexico State University (Cantrell et al. 2000).

AFLP and cAFLP analysis

Genomic DNA was extracted from the leaf tissue of each genotype following the mini-prep protocol (Zhang and

Stewart 2000). The quality of the extracted DNA was checked by electrophoresis on a 1.4% agarose gel, stained with ethidium brominde (EB), and visualized on a UV light. DNA quantity was measured by a Fluorometer.

The AFLP was done following the protocol of Vos et al. (1995) with minor modifications. The genomic DNA (500 ng) was incubated for 3 h at 37°C with the following reagents: 0.2 µl of T4 DNA Ligase (400 U/µl), 1 µl of 10 times Ligase buffer, 1 µl of NaCl (0.5 M), 0.5 µl BSA (1 mg/ml), 1 µl of *Mse*I adaptor (25 µM), 1 µl of *Eco*RI adaptor (5 µM), 0.5 of µl *Mse*I enzyme (10 U/µl), and 0.25 µl of *Eco*RI enzyme (20 U/µl). After ligation, the reaction was diluted ten times with TE buffer (10 mM Tris–HCl, pH 8.0, 0.1 mM EDTA) and stored at -20°C. To determine if the DNA templates were restricted completely, the ligation reaction was run on a 1% agarose gel and stained with EB.

Two pre-selective primers with a single selective nucleotide extension (forward *Eco*RI-PSA E: 5′GACTGCGTACCAATTCA3′; reverse *Mse*I-PSA M: 5′GATGAGTCCTGAGTA AC3′) were used to amplify fragments of the DNA template. The PCR reaction mix consisted of 1.0 µl of 10 times buffer, 0.8 µl dNTP (2.5 Mm each), 0.3 µl *Mse*I primer, 0.3 µl *Eco*RI primer, 0.5 U TaqGold Polymerase and 4.5 µl deionized water (ddH_2O) per reaction. The total volume of the reaction was 10 µl containing 2 µl of the diluted restricted-ligated DNA template. The PCR conditions were as follows: after the initial 72°C for 30 min and 95°C for 5 min, followed by 25 cycles of denaturing at 94°C for 20 s, annealing at 56°C for 30 s and extension at 72°C for 2 min, with a final extension at 72°C for 2 min and 60°C for 30 min after the last PCR cycle. The PCR products were diluted with 90 µl of TE buffer (10 mM Tris–HCl, pH 8.0, 0.1 mM EDTA) and run on a 1% agarose gel and stained with EB to determine if the DNA templates were amplified correctly. To survey for AFLP polymorphism within Upland cotton, all the 64 primer combinations that had two nucleotide extensions from PSA primers were used to amplify preselective PCR products from TM-1 and NM 24016. Then, a subset of four primer combinations (A1, B2, C3, and D7) was selected for selective PCR amplification of the seven preselective PCR products. Selective PCR reactions contained 4.1 µl of ddH_2O, 10 times PCR buffer, 0.2 mM dNTP′s each, 0.25 µM of *Mse*I primer, 0.20 µM of *Eco*RI primer with a fluorescent dye WellRed D4 (Beckman-Coulter Inc., Fullerton, CA, USA) and 2.0 µl of diluted pre-selective amplification DNA. The PCR program for the selective amplification consisted of an initial denaturation at 94°C for 5 min, ten cycles of 94°C for 20 s, 66°C for 30 s, and 72°C for 2 min, followed by 20 cycles of 94°C for 20 s, 56°C for 30 s, and 72°C for 2 min, each with 1°C lowering of annealing temperature, and a final extension at 60°C for 30 min after the last PCR cycle. After selective amplification, 2.0 µl diluted (ten times) DNA was added to 18.0 µl of deionized formamide and 0.12 µl of a 600 bp size standard (Beckman-Coulter Inc., Fullerton, CA, USA). The samples were analyzed and sequenced using the CEQ 8000 Fragment Analysis Software (Beckman-Coulter Inc., Fullerton, CA, USA). For cAFLP analysis, 5.0 µl of the selective amplified PCR products were used in restriction with enzyme *Taq*I followed the manufacturer's instructions and 1 µl of the restricted AFLP products was analyzed using the CEQ 8000 Sequencer as described above.

Data analysis

All of the seven DNA samples were used in the AFLP and cAFLP analysis. Each sample was scored for present (1) and absent (0) by the CEQ Cluster Fragment Analysis (Beckman-Coulter Inc., Fullerton, CA). Only those amplified fragments with high reproducibility were scored by manually checking the traces generated with each primer combination in that the loci with low reproducibility were deleted. To estimate the genetic similarities among genotypes a genetic distance matrix based on Jaccard coefficient was used in the Numerical Taxonomy System (NTSYSpc), Version 2.1 (Exeter Software, Setauket, New York, USA). A phylogenetic tree was constructed using the Neighbor-Joining (NJ) method. This program was used to group genotypes that are genetically related to each other based on the genetic similarity (GS) matrix.

Results

Survey of AFLP polymorphism between TM-1 and NM 24016

About 64 AFLP primer pairs were screened for AFLP using TM-1, the genetic standard of Upland cotton and NM 24016. NM 24016 is the most diverse germplasm in Upland cotton since it contained substantial genetic introgression from *G. barbadense*. The polymorphism between the two should represent the highest level of polymorphism within Upland cotton. Of 4679 AFLP fragments amplified from the 64 primer pairs, 211 AFLP fragments (4.5%) were polymorphic. This number of polymorphic AFLP is not sufficient to construct a genome-wide linkage map, considering that cotton genome has a total of more than 5000 cM in genetic distance. Therefore, developing new and more markers within Upland cotton is necessary.

Number and percentage of AFLP fragments cut by an enzyme

When restricted by Taq I, 30% of AFLP the fragments were cut (Table 1). In both Upland and Pima cotton, the most frequent AFLP fragments ranged from 100 to 300 bp in size. Among AFLP fragments of <100 bp,

Table 1 Distribution of AFLP fragments that are cut by *Taq*I

Fragment size (bp)	Upland cotton (AD_1)		Pima cotton (AD_2)		AD_1/AD_2	
	No. total	No. cut (%)	No. total	No. cut (%)	No. total	No. cut (%)
< 100	54	6 (11.1)	50	8 (16.0)	59	14 (23.7)
100–200	113	21 (18.6)	97	13 (13.4)	120	34 (28.3)
200–300	87	33(37.9)	57	22 (38.6)	81	37 (45.9)
300–400	52	28 (53.8)	33	18 (54.5)	36	17 (47.2)
400–500	17	14 (82.4)	12	7 (58.3)	8	4 (50.0)
> 500	1	1 (100.0)	6	6 (85.7)	5	6 (100.0)
Sum	324	103 (31.8)	255	74 (29.0)	310	111 (35.8)

Table 2 Distribution of number of polymorphic fragments

Fragment size (bp)	Upland cotton (AD_1)		Pima cotton (AD_2)		AD_1/AD_2	
	AFLP	cAFLP	AFLP	cAFLP	AFLP	cAFLP
< 100	4	18	10	21	12	31
100–200	9	21	7	43	42	50
200–300	10	14	8	9	53	35
300–400	7	5	12	17	20	17
400–500	7	4	2	7	12	4
> 500	0	0	5	5	5	5
Sum	37/324 (11.4)	62/283 (18.5)	44/255 (17.3)	102/291 (35.1)	144/370 (38.9)	142/374 (38.0)

11%–16% were cut, while 13%–19% AFLP fragments in size 100 bp–200 bp were cut. The restricted AFLP (cAFLP) fragments were increased to 37% for 200 bp–300 bp fragments. For AFLP fragments of 300 bp–500 bp in size, more than 50%–80% were cut. For AFLP fragments larger than 500 bp, more than 80% were restricted. Figure 1a and b show a comparison between AFLP and cAFLP fragments resolved on CEQ 8000 DNA Sequencer. The results indicated that AFLP fragments can be further restricted using a frequent restriction enzyme *Taq*I and readily resolved using the capillary-based CEQ 8000 DNA Sequencer.

Number and percentage of polymorphic cAFLP

Within Upland cotton, 11.4% of AFLP fragments were polymorphic, while 17.3% AFLP markers were polymorphic in Pima cotton (Table 2). The number of polymorphic AFLP markers appeared to be not associated with their fragment sizes, even though most of them ranged from below 100 to 400 bp. Due to the limitation of the sequencer system, fragments larger than 500 bp were not well detected. However, with cAFLP, many more polymorphic fragments (18.5% in Upland cotton and 35.1% in Pima cotton) were detected (Table 2). Some of the polymorphic markers came from the restriction of large AFLP fragments that were undetected by the CEQ system and others were derived from the smaller fragments. Most of the polymorphic cAFLP markers were below 300 bp in size. Compared with AFLP analysis that produced 37 and 44 polymorphic markers in Upland and Pima cotton, respectively, from four primer pairs, cAFLP generated 62 and 102 polymorphic fragments, respectively, an increase of 67%–132%. However, the number of polymorphic cAFLP markers (142) between Upland and Pima cotton did not increase as expected, as compared with 144 polymorphic AFLP markers between the two species. This indicates that cAFLP would not increase its resolution power at the interspecies level in cotton when high level of genetic diversity exists.

We further examined the distribution of AFLP and cAFLP markers (Table 3). Of the monomorphic markers that were produced by the four primer combinations, 218 and 181 are common between AFLP and cAFLP in Upland and Pima cotton, respectively, indicating that these AFLP fragments were not restricted; 102 (in Upland) and 73 (in Pima) fragments were AFLP specific, indicating that these monomorphic AFLP fragments were also not cut, while 65 (in Upland) and 110 (in Pima) monmorphic fragments were cAFLP specific, indicating that these were generated from restriction of other monomorphic AFLP fragments. However, the restriction of these monomorphic AFLP did not produce polymorphic cAFLP.

Of the polymorphic fragments, 17 (23%) in Upland and 26 (21%) in Pima were in common between AFLP and cAFLP, indicating that the enzyme *Taq*I did not cut these polymorphic AFLP markers; 22 (30%) in Upland and 16 (13%) in Pima were AFLP specific, indicating that these polymorphic AFLP were cut by *Taq*I and lost in cAFLP analysis. However, the number of cAFLP specific polymorphic markers were increased to 35 (47%) in Upland and 82 (66%) in Pima, including most, if not all of these polymorphic AFLP markers that were cut. However, the unique polymorphic cAFLP between

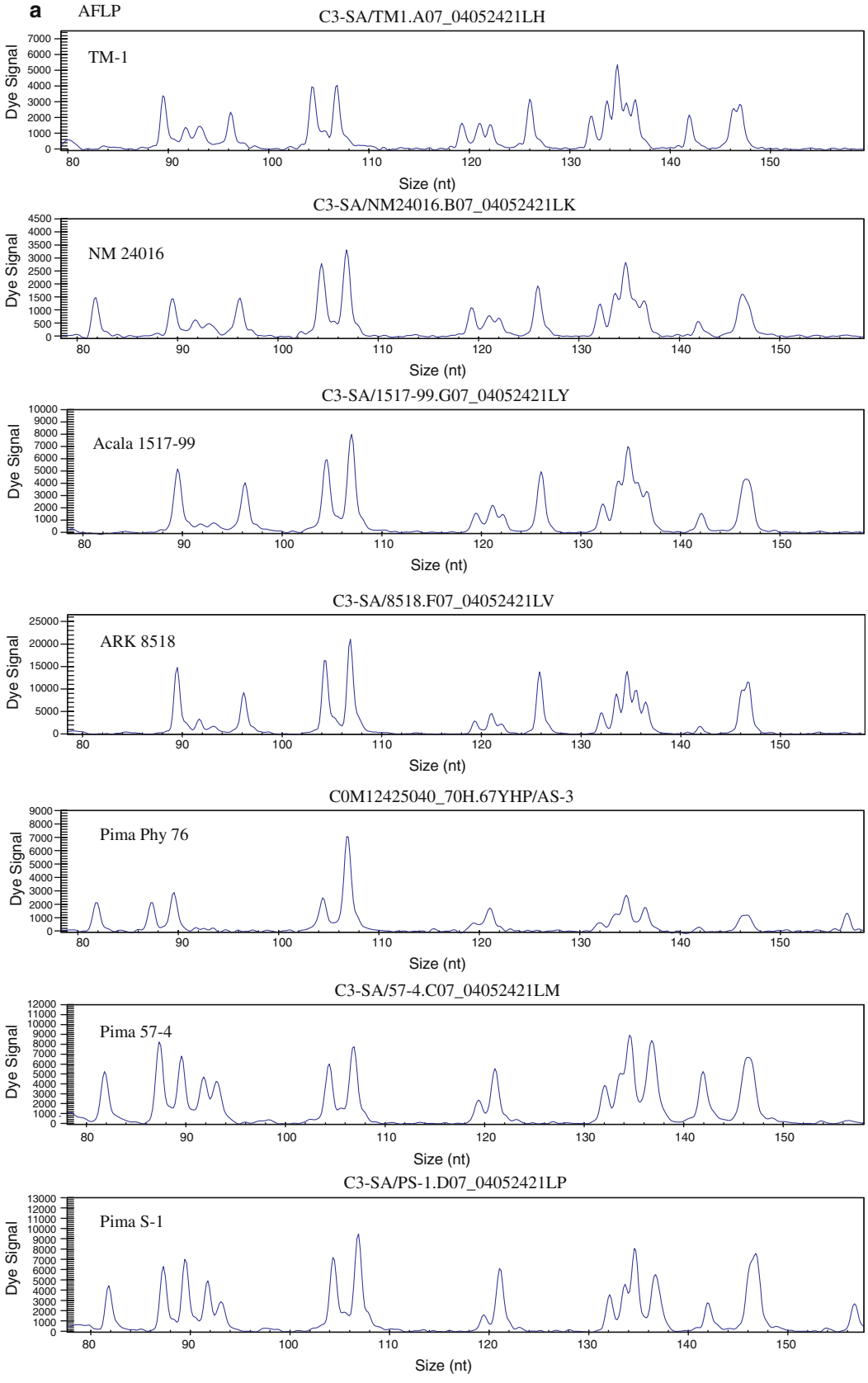

Fig. 1 Comparison between AFLP (**a**) and cAFLP (**b**) amplified by primer combination C3. *Arrows* indicate new fragments after AFLP products were restricted with *Taq*I

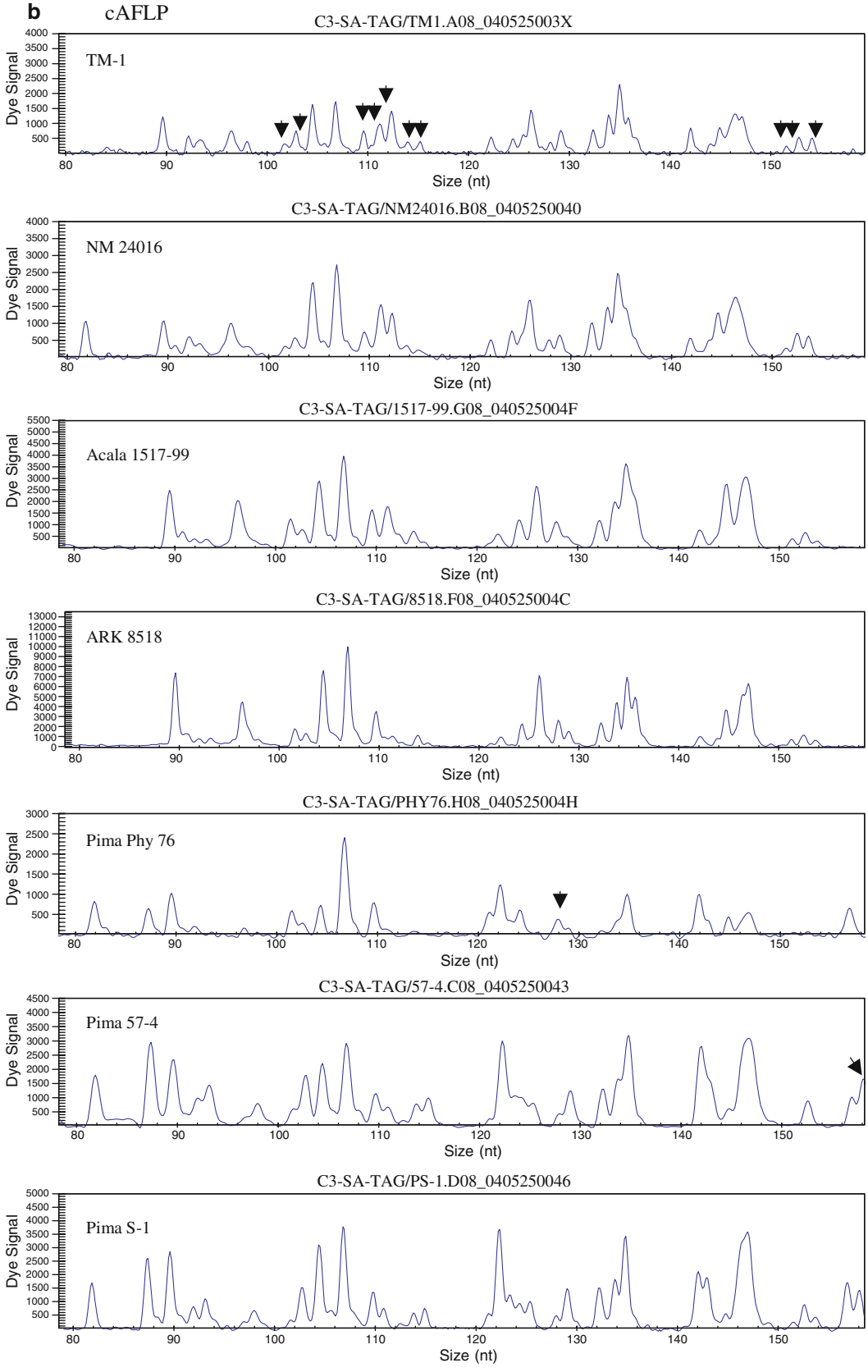
b
cAFLP
C3-SA-TAG/TM1.A08_040525003X
TM-1
C3-SA-TAG/NM24016.B08_0405250040
NM 24016
C3-SA-TAG/1517-99.G08_040525004F
Acala 1517-99
C3-SA-TAG/8518.F08_040525004C
ARK 8518
C3-SA-TAG/PHY76.H08_040525004H
Pima Phy 76
C3-SA-TAG/57-4.C08_0405250043
Pima 57-4
C3-SA-TAG/PS-1.D08_0405250046
Pima S-1
Dye Signal
Size (nt)

Fig. 1 (Contd.)

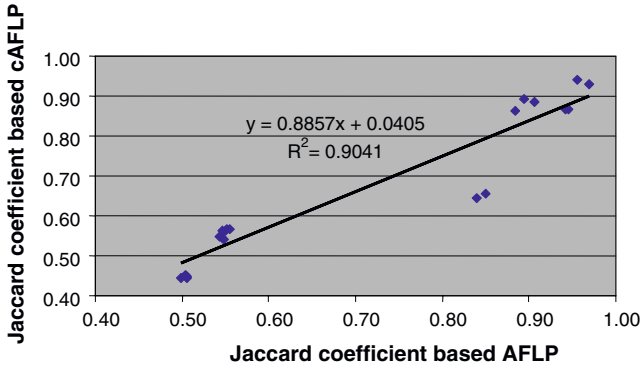

Fig. 2 Relationship between AFLP and cAFLP

Upland and Pima were only increased to 72, not as great as expected, as compared with 58 unique polymorphic AFLP markers.

Based on the fragment size, the restriction of most AFLP fragments should only produce two fragments, of which only the fragment with the fluorescent-labeled *Eco*RI primers was detected by the sequencer system. Therefore, the restriction of AFLP fragments within the range of detection does not necessarily increase the number of cAFLP fragments, unless the fragments to be restricted are larger than the sequencer can resolve.

Genetic similarity between genotypes as evaluated by AFLP and cAFLP

Jaccard similarity coefficients among the seven genotypes are listed in Table 4. Overall, the GS coefficients were lower based on cAFLP markers, indicating a higher discriminatory power of cAFLP in genotype differentiation than AFLP. However, the power in genotype discrimination between the two species (Upland and Pima cotton) did not significantly increase since they already have tremendous genetic diversity. For comparison between genotypes within species, the discriminatory power was significantly enhanced by cAFLP. For example, the AFLP-based GS between Pima 57-4 and Pima Phy 76, and between Pima Phy 76 and Pima S-1 was 0.84–0.85, while their similarity based on cAFLP was decreased to 0.65–0.66. The GS between Acala 1517-99 and ARK 8518, and between Acala 1517-99 and TM-1 was 0.94–0.95 based on AFLP, while their cAFLP-based GS was lower (0.87). However, the two marker systems were overall congruent in estimating the genetic diversity among the seven genotypes tested (Fig. 2). Therefore, cAFLP is a robust DNA fingerprinting technique with higher resolution power than AFLP.

Cluster analysis

Based on AFLP analysis, as expected, Pima S-1 and 57-4 are grouped together first before they joined with Pima Phy 76 to form the Pima cotton group, while the other four Upland cotton formed a separate group (Fig. 3). Within this Upland cotton group, TM-1 and ARK8518 grouped together as expected, since they are both Delta type Upland cotton. Unexpectedly, the two Acala cotton genotypes, Acala 1517-99 and NM 24016, developed from the same breeding program at New Mexico State University did not group together. However, cAFLP-based analysis correctly grouped these two together in a sub-group before they joined with another sub-group (TM-1 and ARK8518) to form the Upland cotton group (Fig. 4).

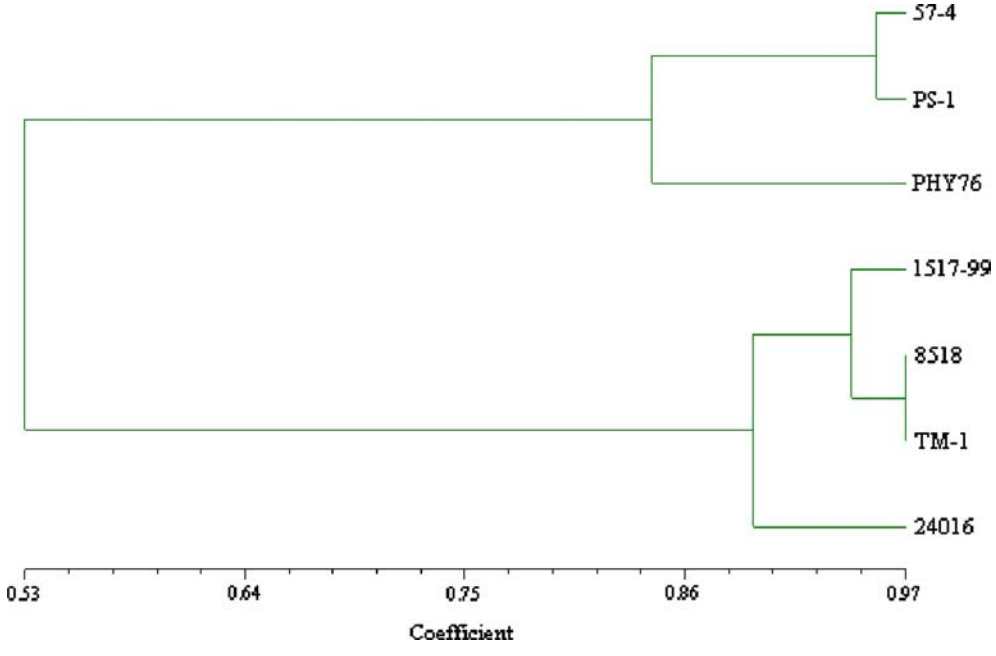

Fig. 3 An UPGMA dendrogram based on AFLP data from four primer combinations

Table 3 Comparison between AFLP and cAFLP fragments

Primer	Species	Common fragment	AFLP unique	cAFLP unique	Common poly	AFLP unique poly	cAFLP unique poly
A_1	AD_1[a]	70	49	19	1	5	10
	AD_2[b]	53	18	40	4	1	15
	AD_1/AD_2[c]	64	26	44	12	28	16
B_2	AD_1	58	22	14	7	11	11
	AD_2	46	19	19	7	8	26
	AD_1/AD_2	40	35	15	15	9	35
C_3	AD_1	47	16	27	3	1	8
	AD_2	43	24	40	9	3	35
	AD_1/AD_2	47	27	62	13	12	17
D_7	AD_1	43	15	5	6	5	6
	AD_2	39	12	11	6	4	6
	AD_1/AD_2	21	49	45	22	9	4
Sum	AD_1	218	102	65	17 (23.0)	22 (29.8)	35 (47.3
	AD_2	181	73	110	26 (21.0)	16 (12.9)	82 (66.1)
	AD_1/AD_2	172	137	166	62 (32.3)	58 (30.2)	72 (37.5)

[a]1517-99, 24016, TM-1, and A-RK8518
[b]57-4, Pima S-1, and Pima PH-Y76
[c]TM1, 57-4, and Pima S-1

Table 4 Similarity coefficient matrix based on AFLP (above diagonal) and cAFLP (below diagonal)

	1	2	3	4	5	6	7
Pima 57-4 (1)		0.8392	0.9556	0.5434	0.5549	0.5465	0.5501
Pima Phy 76 (2)	0.6452		0.8498	0.5042	0.5058	0.4986	0.5057
Pima S-1 (3)	0.9389	0.6545		0.5449	0.5520	0.5480	0.5473
Acala 1517-99 (4)	0.5493	0.4517	0.5465		0.9452	0.8951	0.9423
ARK 8518 (5)	0.5657	0.4434	0.5679	0.8669		0.8840	0.9701
NM 24016 (6)	0.5428	0.4462	0.5401	0.8929	0.8612		0.9054
TM-1 (7)	0.5646	0.4489	0.5619	0.8657	0.9296	0.8865	

Discussion

G. hirsutum was divided into seven races by Hutchinson et al. (1947) and was thought to be first domesticated in Yucatan peninsula of Mexico as the wild variety called ''yucatanense'', which could have given rise to another primitive variety ''punctatum'' (Brubaker and Wendel 1994). Their dispersion to the rest of Mesoamerica, northern South America, and the Caribbeans gave rise to other widespread commensal forms (i.e., ''marie'galante'' and ''palmeri'', and ''latifolium''). Intentional or unintentional selection in the early days created latifolium genotypes with reduced seed dormancy, compact growth habit, and photoperiod-neutral flowering, known as ''Mexican Highlands'' varieties. Introduction

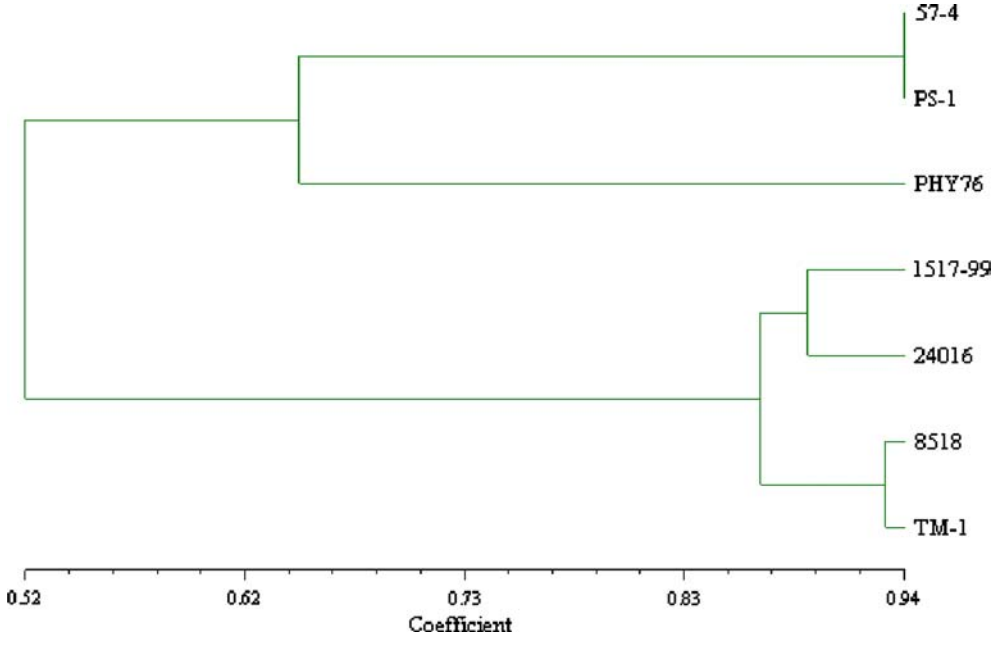

Fig. 4 An UPGMA dendrogram based on cAFLP data from four primer combinations

of several "Mexican" varieties into the U.S. during the late 1700 s and early 1800 s formed the germplasm foundation for modern American Upland cotton that was dispersed worldwide (Smith and Cothern 1999).

South America is the center of origin for *G. barbadense*, whose fiber was medium long and coarse, as typified by the current Tanguis cottons of Peru. The introgression of fiber length genes from outside the species, possibly from *G. hirsutum*, resulted in the development of extra long staple Sea Island cotton in Caribbeans. The Sea Island cotton was brought into Egypt in 1825 and crossed with tree cotton named Jumel, leading to the development of Ashmouni in about 1860 and several Egyptian cultivars between 1910 and 1940. The introduction of Mitafifi, an Egyptian cultivar and re-selection resulted in the release of the first extra long staple cultivar, Yuma in the U.S. in 1908. Between 1908 and 1949, four additional Pima cotton cultivars, Pima, SXP, Amsak, and Pima 32, were developed from the Egyptian germplasm base (Smith and Cothern 1999).

Given the narrow genetic bases of modern Upland and Pima cotton, low level of genetic polymorphisms at the intraspecific level was reported using isozymes and RFLP (Percy and Wendel 1990; Wendel et al. 1992). The results have been confirmed later using AFLP. Abdalla et al. (2001) reported an average GS of 0.86 and 0.89 in cultivated Upland and Pima cotton cultivars, respectively. Iqbal et al. (2001) included earliest Upland cotton varieties and cultivars developed in the U.S. and germplasm from several other countries and indicated that GS ranged from 0.83 to 0.99 in Upland and 0.91 to 0.99 in modern Pima cotton cultivars. More recently, Westengen et al. (2005) evaluated 94 *G. barbadense* germplasm accessions collected from South America and estimated that GS ranged from 0.83 to 0.98 with an average of 0.93. These AFLP-based results were consistent with our study using limited number of modern Upland (GS = 0.88–0.97) and Pima cotton cultivars (GS = 0.84–0.96). TM-1 was an inbred line derived from repeated selfing of Deltapine 14, a cultivar developed in the 1940 s; while the other Upland cotton tested were either commercial cultivars (ARK8518 and 1517-99) or breeding line (NM 24016) developed in the 1990 s. The inclusion of Acala 1517-99 and NM 24014 would represent the highest level of polymorphism in Upland cotton when compared with TM-1 and 8518. Pima S-1 was a cultivar developed in the 1950 s, while Phy 76 was a current commercial Pima cultivar released in the 1990 s. The polymorphism between the two would represent accumulation of genetic diversity gained due to breeding in the past 50 years.

Even though the interspecific polymorphism between Upland and Pima is very high (GS = 0.50–0.55 based on AFLP and 0.44–0.57 based on cAFLP in the present study), the low level of AFLP polymorphism within Upland and Pima would not allow the development of high-density intraspecific linkage maps. Based on Abdalla et al. (2001), on average, a primer combination (a total of 16 primer pairs) produced 73.8 AFLP fragments, of which 5.6 in Upland and 8.4 in Pima were polymorphic. Westengen et al. (2005) used eight primer combinations and found that polymorphic AFLP bands ranged from 7 (16%) to 16 (27%) with an average of 11.6 polymorphic AFLP markers among the diverse 94 *G. barbadense* accessions. Certainly, the level of AFLP polymorphism between any two intraspecific genotypes will be considerably lower. From four primer combinations, our data indicated 9.1 and 11.0 polymorphic AFLP markers per primer pair for Upland and Pima, respectively.

Since its invention of AFLP, various modifications have been proposed that have involved restriction enzyme combinations including using one enzyme, primer extensions (1, 2, or 3 bp) and combinations (with other primers based on SSR, retroposons, and disease resistance gene analogues), and fragment separation systems (e.g., Roy et al. 2002; Park et al. 2003; Soriano et al. 2005). The combination of six-base cutter *Eco*RI and four-base cutter *Mse*I was used in most cases of AFLP analysis. Some other six-base cutters, such as *Hin*dIII, *Pst*I, *Not*I, *Sac*I, and *Bgl*II, have also been combined with other four-base cutters, such as *Taq*I, *Hpa*II, *Msp*I, *Csp*6I, *Tru*1I, *TRu*9I, *Mfe*I, and *Hha*I. Wurff et al. (2000) proposed TE-AFLP using three enzymes before ligation, which was shown to reduce the number of AFLP bands and increase discriminatory power. The modified AFLP technique, cAFLP takes advantages of the convenience of traditional AFLP techniques (without any changes in AFLP protocols) and reliability of restriction digestion. Our work demonstrates that cAFLP analysis, based on further restriction of AFLP amplicons by a frequent restriction enzyme (here 4-bp cutter, *Taq*I), increases polymorphic markers by 67% in Upland and 132% in Pima. Given the high-resolution power of capillary sequencer, it is especially suited for cAFLP analysis. CEQ 8000 DNA Sequencer detected an average of 69–134 AFLP fragments per primer combination within 50–500-bp range. Since the number of cAFLP fragments was not unusually high and fragments with the same size are rare, to maximize the polymorphism that a primer combination can generate, we further propose that AFLP and cAFLP for the same primer combination can be combined before sample loading to the sequencer. The number of polymorphic markers per primer pair produced by combining AFLP and cAFLP was 18.5 (16%) in Upland and 31.0 (28%) in Pima. In doing this, except for purchasing inexpensive restriction enzymes and additional time for restriction, no extra reagents and time in using capillary sequencer are required, thereby highly cost-effective.

The main purpose of our present work was to prove the concepts using one of the most frequently used four-base cutter (e.g., *Taq*I) after AFLP amplification using *Eco*RI and *Mse*I. Though this cAFLP was initially developed for cotton, it is applicable to any AFLP analysis and its modifications. Using several restriction enzyme combinations before pre-selective amplification

(AFLP) in combination with various frequent four bp-cutters or six bp-cutters after selective amplification (cAFLP), the pooled AFLP and cAFLP will provide unlimited number of polymorphic markers for genome-wide mapping and fingerprinting.

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 103:547–554

Altaf MK, Zhang JF, Stewart JM, Cantrell RG (1998) Integrated molecular map based on a trispecific F2 population of cotton. In: Proceedings of the Beltwide Cotton Conference, San Diego, CA. 5–9 January 1998. National Cotton Council of America, Memphis, TN, USA, pp 491–492

Bourland FM (1996) Registration of 'H1330' cotton. Crop Sci 36:813

Brubaker CL, Wendel JF (1994) Reevaluating the origin of domesticated cotton (*Gossypium hirsutum* Malvaceae) using nuclear restriction fragment length polymorphisms (RFLP). Am J Bot 81:1309–1326

Brubaker CL, Brown AH (2003) The use of multiple alien chromosome addition aneuploids facilitates genetic linkage mapping of the *Gossypium* G genome. Genome 46:774–791

Cantrell RG, Davis DD (2000) Registration of NM24016, an interspecific-derived cotton genetic stock. Crop Sci 40:1208

Cantrell RG, Roberts CL, Waddell C (2000) Registration of 'Acala 1517–99' cotton. Crop Sci 40:1200–1201

Chee PW, Rong J, Williams-Coplin D, Schulze SR, Paterson AH (2004) EST derived PCR-based markers for functional gene homologues in cotton. Genome 47:449–462

Gutierrez OA, Basu S, Saha S, Jenkins JN, Shoemaker D, Cheatham CL, McCarty JC Jr (2002) Genetic distance among selected cotton genotypes and its relationship with F_2 performance. Crop Sci 42:1841–1847

Han ZG, Guo WZ, Song XL, Zhang TZ (2004) Genetic mapping of EST-derived microsatellites from the diploid *Gossypium arboreum* in allotetraploid cotton. Mol Genet Genomics 272:308–327

Hinchliffe DJ, Lu Y, Potenza C, Segupta-Gopalan C, Cantrell RG, Zhang JF (2005) Resistance gene analogue markers are mapped to homeologous chromosomes in cultivated tetraploid cotton. Theor Appl Genet 110:1074–1085

Hutchinson JB, Silow RA, Stephens SG (1947) The evolution of *Gossypium* and diiferentiation of the cultivated cotton, 1st edn. Oxford University Press, London

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 DNA fingerprinting. Theor Appl Genet 103:547–554

Kohel RJ, Richmond TR, Lewis CF (1970) Texas Marker-1. Description of a genetic standard for *Gossypium hirsutum* L. Crop Sci 10:670–671

Lacape JM, Nguyen TB, Thibivilliers S, Bojinov B, Courtois B, Cantrell RG, Burr B, Hau B (2003) A combined RFLP-SSR-AFLP map of tetraploid cotton based on a *Gossypium hirsutum* × *Gossypium barbadense* backcross population. Genome 46:612–626

Lacape JM, Lacape, Nguyen TB, Courtois B, Belot JL, Giband M, Gourlot JP, Gawryziak G, Roques S, Hau B (2005) QTL Analysis of cotton fiber quality using multiple *Gossypium hirsutum* × *Gossypium barbadense* backcross generations. Crop Sci 45:123–140

Lewis H (2001) A review of yield and fiber quality trends and components in American upland cotton. In: Proceedings of the Beltwide Cotton Conference, January 9–13, Anaheim, CA. National Cotton Council of America, Memphis, TN, USA, pp 1447–1453

Liu B, Brubaker CL, Mergeai G, Cronn RC, Wendel JF (2001) Polyploid formation in cotton is not accompanied by rapid genomic changes. Genome 44:321–330

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

Lu Y, Curtiss J, Percy RG, Cantrell RG, Zhang JF (2005) Discovery of single nucleotide polymorphism in selected fiber genes in cultivated tetraploid cotton. (2005). In: Proceedings of the Beltwide Cotton Conference, New Orleans, LA, USA, 5–9 January 2005. National Cotton Council of America, Memphis, TN, USA

May OL, Bowman DT, Calhoun DS (1995) Genetic diversity of U.S. Upland cotton cultivars released between 1980 and 1990. Crop Sci 35:1570–1574

Mei M, Syed NH, Gao W, Thaxton PM, Smith CW, Stelly DM, Chen ZJ (2004) Genetic mapping and QTL analysis of fiber-related traits in cotton (*Gossypium*). Theor Appl Genet 108:280–291

Meredith WR Jr (2000) Cotton yield progress—why has it reached a plateau. Better Crops 84:6–9

Multani DS, Lyon BR (1995) Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38:1005–1010

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

Park KC, Kim NH, Cho YS, Kang KH, Lee JK, Kim NS (2003) Genetic variations of AA genome *Oryza* species measured by MITE-AFLP. Theor Appl Genet 107:203–209

Paterson AH, Saranga Y, Menz M, Jiang CX, Wright RJ (2003) QTL Analysis of genotype × environment interactions affecting cotton fiber quality. Theor Appl Genet 106:384–396

Percy RG, Wendel JF (1990) Allozyme evidence for the origin and diversification of *Gossypium barbadense* L. Theor Appl Genet 79:529–542

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

Rahman M, Hussain D, Zafar Y (2002) Estimation of genetic divergence among elite cotton cultivars–genotypes by DNA fingerprinting technology. Crop Sci 42:2137–2144

Reinisch AJ, Dong JM, Brubaker CL, Stelly DM, Wendel JF, Paterson AH (1994) A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics 138:829–847

Rong J, Abbey C, Bowers JE, Brubaker CL, Chang C, Chee PW, Delmonte TA, Ding X, Garza JJ, Marler BS, Park CH, Pierce GJ, Rainey KM, Rastogi VK, Schulze SR, Trolinder NL, Wendel JF, Wilkins TA, Williams-Coplin TD, Wing RA, Wright RJ, Zhao X, Zhu L, Paterson AH (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

Roy JK, Balyan HS, Prasad M, Gupta PK (2002) Use of SAMPL for a study of DNA polymorphism, genetic diversity and possible gene tagging in bread wheat. Theor Appl Genet 104:465–472

Saranga Y, Menz M, Jiang C, Wright R, Yakir D, Paterson AH (2002) Genetic dissection of genotype × environment interactions conferring adaptation of cotton to arid conditions. Genome Res 11:1988–1995

Shappley ZW, Jenkins JN, Meredith Jr WR, McCarty JC Jr (1998) An RFLP linkage map of Upland Cotton, *Gossypium hirsutum* L. Theor Appl Genet 97:756–761

Smith CW, Cothern JT (1999) Cotton. Wiley, New York

Soriano JM, Vilanova S, Romero C, Llacer G, Badenes ML (2005) Characterization and mapping of NBS-LRR resistance gene analogs in apricot (*Prunus armeniaca* L.). Theor Appl Genet [Epub ahead of print]

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

Ulloa M, Meredith WR Jr, Shappley ZW, Kahler AL (2002) RFLP genetic linkage maps from four F(2.3) populations and a joinmap of *Gossypium hirsutum* L. Theor Appl Genet 104:200–208

Ulloa M, Saha S, Jenkins JN, Meredith WR Jr, McCarty JC Jr, Stelly DM (2005) Chromosomal assignment of RFLP linkage groups harboring important QTLs on an intraspecific cotton (*Gossypium hirsutum* L.) joinmap. J Hered 96:132–144

Van der Wurff AWG, Chan YL, van Straalen NM, Schouten J (2000) TE-AFLP: combining rapidity and robustness in DNA fingerprinting. Nucleic Acids Res 28:e105

Vos P, Hogers R, Bleeker M, Reijans M, van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, Zabeau M (1995) AFLP: a new technique for DNA fingerprinting. Nucleic Acids Res 23:4407–4414

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

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

Westengen OT, Huaman Z, Heun M (2005) Genetic diversity and geographic pattern in early South American cotton domestication. Theor Appl Genet 110:392–402

Wright RJ, Thaxton PM, El-Zik KM, Paterson AH (1998) D-subgenome bias of X cm resistance genes in tetraploid *Gossypium* (cotton) suggests that polyploid formation has created novel avenues for evolution. Genetics 149:1987–1996

Zhang J, Guo W, Zhang T (2003) Molecular linkage map of allotetraploid cotton (*Gossypium hirsutum* L. × *Gossypium barbadense* L.) with a haploid population. Theor Appl Genet 105:1166–1174

Zhang JF, Stewart JM (2000) Economical and rapid method for extracting cotton genomic DNA. J Cotton Sci 4:193–201

Zhang JF, Stewart JM (2004) Semigamy gene is associated with chlorophyll reduction in cotton. Crop Sci 44:2054–2062

Zhang JF, Lu Y, Cantrell RG, Hughs E (2005) Molecular marker diversity and field performance in commercial cotton cultivars evaluated in the Southwest U.S. Crop Sci 45:1483–1490

Zuo KJ, Sun JZ, Zhang JF, Nie YC, Liu JL (2000) Genetic diversity evaluation of some Chinese elite cotton varieties with RAPD markers. Acta Genetica Sinica 27:817–823