Amplified fragment length polymorphism (AFLP) and genealogy analysis of the introgressed lines from *Gossypium hirsutum* × *Gossypium barbadense* varieties

Zhao-hui Liu¹,², Yong Zhang¹,³*, Jinhua Peng¹, Kejun Deng¹, Shunwen Dong² and Zhenglong Ren¹,³*

¹School of Life Sciences and Technology, University of Electronic Science and Technology of China, Chendu, 610054, P. R. China.
²Industrial Crop Breeding and Cultivation Institute, Sichuan Academy of Agriculture Sciences, Jianyang 641400, P. R. China.
³Key Laboratory of Plant Breeding and Genetics, Sichuan Agricultural University, Ya’an, 625014, P. R. China.

Accepted 21 May, 2010

In this study, amplified fragment length polymorphism (AFLP) technique was used to analyze the genealogical relationship of 7 introgressed cotton varieties from the hybridization of upland (*Gossypium hirsutum* L) and sea-island cottons (*Gossypium barbadense* L). Ten pairs of primer combinations with high polymorphism and resolution were selected from 64 primers. These 10 primer combinations resulted in a total of 480 bands, of which 374 bands (77.9%) were polymorphic and 51 bands (10.6%) were specific. Cluster analyses showed that 7 varieties of cottons were divided into two groups. Our results suggested that interspecific hybridization is feasible to broaden germplasm of upland cotton and AFLP technique could be applied for the identification of variety purity and genealogical relationship.

Key words: Cotton, introgressed line, AFLP, genealogical analysis.

INTRODUCTION

Besides the upland cotton cultivars in gossypium, there are multiple relative species including wild plant varieties and wild species. All these relative species are genetically diversified and contain many superior genetic resources for cotton breeding (Stalker, 1980; Altman, 1993). Recently, researchers from China and other countries explored distant hybridization between upland cotton cultivars and other cotton cultivars, wild and semi-wild species. A large amount of superior germplasm resources for the upland cotton cultivars with stable characters and normal fertilities were bred (Lacape et al., 2005; Liang et al., 2002; Hu et al., 1993; Qian et al., 1996). This kind of cotton germplasm was designated as introgressed lines from interspecific hybridization. Generation of introgressed lines has broadened the genetic basis of current cotton germplasm and overcame the bottleneck of upland cotton breeding, which has great application values.

Our laboratory has initiated the research project on the distant hybridization of cotton since 1977 and bred the first genetically stable, acarid-resistant hybrid lines (*Chuan 98*) from *Gossypium hirsutum* L. × *Gossypium barbadense* L. (*Wang et al., 1991*). *Chuan 98* was brought into the national gene bank and became a material with international sensitivity level. In addition, our laboratory has also bred approximately 10 aphid- or acarid-resistant lines including 9809, 169-6, etc. With these cotton lines, we obtained acarid-resistant cultivar *Chuan45* and aphid-resistant cultivar *Chuan109* (*Wang, 1993*). The harms caused by cotton acarid and aphid were partially resolved by popularizing the aphid- or acarid-resistant cotton varieties.

Abbreviations: AFLP, Amplified fragment length polymorphism; PCR, polymerase chain reaction; SSR, simple sequence repeats; RAPD, amplified polymorphic DNA.
Table 1. The cotton materials used in this experiment and their origins.

<table>
<thead>
<tr>
<th>The name of the materials</th>
<th>Origins</th>
<th>Genealogical information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chuan98</td>
<td>ICI-SAAS</td>
<td>(Dongting#1 × sea-island cotton)F12</td>
</tr>
<tr>
<td>9809</td>
<td>ICI-SAAS</td>
<td>(Chuan98 × Anti-diseasesae86-1) Bc1F13</td>
</tr>
<tr>
<td>9812</td>
<td>ICI-SAAS</td>
<td>(Chuan98 × Anti-diseasesae 86-1) Bc1F13</td>
</tr>
<tr>
<td>Chuan45</td>
<td>ICI-SAAS</td>
<td>[(Chuan98 × Jiangsu cotton #1)Bc2F1] × Zhongmiansuo#12</td>
</tr>
<tr>
<td>169-6</td>
<td>ICI-SAAS</td>
<td>(Kang77 × Jiangsu cotton #1)F8</td>
</tr>
<tr>
<td>296</td>
<td>ICI-SAAS</td>
<td>(Kang77 × Jiangsu cotton #1)F8</td>
</tr>
<tr>
<td>Chuan109</td>
<td>ICI-SAAS</td>
<td>(Kang77 × Jiangsu cotton #1)Bc1F5</td>
</tr>
</tbody>
</table>

Although these germplasm resources were relatively well applied, studies on the molecular mechanisms of these resistant cotton varieties are still lacking. In addition, these germplasm resources were obtained by hybridization between sea-island cotton and upland cotton lines, but characterization on the chromosome or DNA level has never been performed. Resistant cotton varieties were developed by transformation and systematic breeding through multiple generations. Some of these varieties were very similar and cannot be differentiated based on the morphology, which is inconvenient for the field selection and breeding of new resistant varieties.

In this study, we used highly efficient and reliable amplified fragment length polymorphism (AFLP) technique (Pieter et al., 1995) to screen the primer combinations for the studies on the genetic relationship between the introgressed cotton varieties. We expect to reveal the genetic diversity of the introgressed germplasm, define their genealogical relationship and provide foreseeable guidelines for the parental selection and genetic variation of later generations. These results will be instrumental to establishing new strategies for the collection and preservation of cotton variety sources.

MATERIALS AND METHODS

Cotton subjects

Seven natural cotton varieties hybridized between G. barbadense and G. hirsutum were provided by Sichuan Academy of Agriculture Science, Institute of Industrial Crop Breeding and Culture (Sichuan, China). The detailed origins and genealogical relationship of the 7 varieties are shown in Table 1.

DNA extraction

Young leaves (the age of the leave was less than 1 week) from 10 cotton seedlings were randomly collected from each variety. DNA was extracted according to the methods described previously with slight modifications (Guo et al., 2005). The purity of DNA was determined by electrophoresis at 0.8% agarose gel and the concentration of DNA was measured using fluorimetry (BioSpec-mini, Shimadzu, Japan).

AFLP analysis

AFLP analysis was performed according to the method described by Vos et al. (1995) using the adapter and polymerase chain reaction (PCR) primer sequences and PCR cycles with minor modifications as previously described (Zhang et al., 2008). The amplified products were separated by electrophoresis in 6% denatured polyacrylamide gel at 30 V for 1.5 h. After silvering staining, the gel was naturally dried and the images were taken accordingly. Selective amplification with each primer combination was repeated twice to obtain replicable results (Table 2).

Data analysis

Presence of a band obtained by AFLP at a specific position was recorded as “1” and absence of a band in this position was recorded as “0”. In addition, only those bands that ranged from 100 to 500 bp and were replicable and discernable in different experiments were recorded. Jacard’s similarity coefficient was calculated using the NTSYS-pc2.10 software. Dendrogram analysis was performed using unweighted pair group method with arithmetic mean (UPGMA) method.

RESULTS

Fingerprint evaluation for the 7 cotton varieties

A total of 480 amplification sites were obtained from the genomes of 7 different cotton varieties by using 10 primer pairs. The size of these amplified products ranged from 100 to 500 bp and were evenly distributed. The number of common amplification sites among these 7 cotton varieties was 106 and the number of polymorphic sites was 374, which accounted for 77.9% (374/480) of the total amplification sites. Representative gel images for the AFLP using primers CAC/M-CTA and E-ACC/ M-CAT are shown in Figure 1. These results suggested that abundant polymorphism is present among these 7 cotton varieties. These polymorphic band patterns are stable for a specific cotton variety and thus can be used to differentiate different varieties. A total of 51 (10.6%, 51/480) specific amplified bands were detected in the 7 cotton varieties using 10 AFLP primer combinations. These specific bands are owned by some varieties, but not by other
Table 2. Specific bands of 7 varieties (lines) amplified by different primer combinations.

<table>
<thead>
<tr>
<th>Primer pairs</th>
<th>Chuan98</th>
<th>9809</th>
<th>9812</th>
<th>Chuan45</th>
<th>169-6</th>
<th>169-296</th>
<th>Chuan109</th>
<th>Total no. of unique bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>E-AGG/M-CTT</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>E-ACT/M-CTG</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>E-AAG/M-CTG</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>E-ACT/M-CTC</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>E-ACT/M-CAC</td>
<td>10</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>13</td>
</tr>
<tr>
<td>E-ACT/M-CAT</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>E-CAC/M-CTA</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>11</td>
</tr>
<tr>
<td>E-AAC/M-CAG</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>E-ACG/M-CAC</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td>Total</td>
<td>29</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>9</td>
<td>7</td>
<td>51</td>
</tr>
</tbody>
</table>

Therefore, these bands can be used as specific markers to differentiate different varieties. In this experiment, 99% of the bands were replicable for the DNA extracted at 2 different times. Only very weak or very bright bands had relatively low repeatability, suggesting that the results obtained in this study are highly repeatable.

Figure 1. AFLP profile generated for 7 cultivars of cotton using primer pairs (E-CAC/M-CTA and E-ACC/M-CAT). M, Marker; Lane 1, chuan98, Lane 2, 9809, Lane 3, 9812, Lane 4, Chuan45, Lane 5, 169-6, Lane 6, 169-296 and Lane 7, Chuan109.

Dendrogram analysis for the 7 cotton varieties

The fingerprint results obtained from AFLP were transformed into data matrix. The Jaccard's similarity coefficient was calculated using the NTSYS-pc software and the dendrogram analysis was performed using UPGMA method.

Dendrogram of the 7 cotton varieties is shown in
Figure 2. Dendrogram analysis for the 7 cotton varieties based on the AFLP fingerprint. AFLP analysis divided 7 cotton varieties into two groups. One group included Chuan98, Chuan109 and 169-296. The other group contained 9809, 9812, Chuan45 and 169-6.

Figure 3. Genealogical chart sketch map of 7 cotton varieties based on the parental components. Genealogical analysis divided 7 cotton varieties into two groups. One group contained Chuan98, Chuan45, 9809 and 9812. The other group contained 169-6, 169-296, Chuan109.

Figure 2. The results showed that 7 cotton varieties can be divided into two groups. Chuan109, Chuan169-296 and Chuan98 were clustered together, suggesting that these three cotton varieties are closely related. 9809, 9812, Chuan45 and aphid-resistant variety 169-6 were clustered together, suggesting that these four cotton varieties have close relationships. These results indicate that AFLP can be used to differentiate the genotypes of different cotton varieties and to reveal the relationship between these genotypes.

**Genealogical analysis of 7 cotton varieties**

Based on the genealogy in Table 1, the genealogical chart was generated as shown in Figure 3. Seven cotton varieties were divided into 2 groups: Chuan98, Chuan45, 9809 and 9812 were clustered as one group and chuan109, 169-296 and 169-6 as another group. Grouping of 169-6 and chuan98 were opposite in molecular and genealogical grouping methods. Even though the parental strain of 169-6, chuan109 and 169-296 was derived from Kang77 and Jiangsu cotton#1, each variety was obtained by self cross and selection and thus, the selection direction for each individual cotton variety was different. Therefore, different characters appeared in the offspring generations. For example, the morphology of 169-6 was more similar to that of other upland cotton varieties that were obtained by directional breeding, for example, chuan98. Due to the different
selective direction, the genotype of the offspring generations tended to the different breeding purposes, leading to the occurrence of different fingerprints for the varieties from the same parental varieties (Figures 2 and 3).

**DISCUSSION**

Molecular marking techniques are one of the important approaches for the investigation of the genetic diversity and relationship between different germplasm resources. Precise evaluation of the genetic diversity of the current germplasm can provide foreseeable guidelines for parental selection, prediction of heterosis and genetic variation in the offspring strains (Li and Ma, 2001; Luo et al., 2001). Since the first report on the application of AFLP in the detection of DNA sequence polymorphism, AFLP has been widely used in fingerprint and genetic diversity analysis of many plants including wheat and rice. Some studies also employed AFLP for the genetic analysis of cotton. For example, Wang and Li, (2002) obtained the studies also employed AFLP for the genetic analysis of cotton. Among the amplified products by 18 primer pairs, there were 52 differential bands among different varieties (Wang and Li, 2002). AFLP fingerprint analysis for 105 varieties showed that 46 varieties have specific markers (Wang et al., 2005). In addition, characterization and comparison of the cotton genetic resources were conducted using simple sequence repeats (SSR) and AFLP techniques (Wang et al., 2005). AFLP can detect many polymorphic sites and very slight genetic difference between cotton varieties with very close relationships can be identified by using different primer combinations. We used AFLP to investigate the introgressed lines from *G. hirsutum* L. × *G. barbadense* L varieties and amplified a total of 672 bands, which contained abundant genetic information. All the cotton varieties can be differentiated using 10 primer combinations. The highest efficiency for differentiation was primer E-CAC/M-CTA, which can differentiate 7 cotton varieties, with a differentiation rate of 91%. Our results confirmed that AFLP is an efficient and reliable tool for the studies of genetic relationship between cotton varieties. All the cotton varieties (except for 9812 and Chuan56) had specific AFLP bands. Since some of the introgressed lines have very similar morphology, we can use these specific bands to differentiate these varieties on the DNA level. Thus, we believe that specific band can be used as one of the molecular differentiation approaches to dynamically track the effectiveness of hybridization between different varieties. If we can further determine the link between these specific bands and some genes encoding superior characters, then these specific bands can be used for marker assisted selection, which can facilitate aggregation of the superior genes and breeding of superior varieties.

However, further investigation is needed regarding the relationship between the specific AFLP bands and the genes responsible for these traits.

Molecular marker techniques have been used to confirm the introgression of genetic information from parental varieties. For instance, Liu et al. (2000a, b) and Nie et al. (2000) used random amplified polymorphic DNA (RAPD) technique to analyze the hybridization between some varieties and found that unique bands specific for the wild parental varieties are present in the hybrids, suggesting that RAPD can detect the genetic component of the wild parental varieties in the hybrids between different varieties. Pang et al. (2006) investigated 155 introgressed cotton lines using SSR markers and found that exogenous genetic components in 8 alien cotton varieties were introgressed into the upland cotton varieties. Vroh et al. (1999) believe that a large amount of exogenous DNA fragments and mutated sites can be identified and the sources of these fragments can be tracked by AFLP. In this study, we used AFLP to track the specific fragment in the offspring strains of sea-island and upland cotton hybridized lines and found that there were 129 sites specific to the sea-island cotton lines. Among the 7 varieties, 8-97 bands were specific to the sea-island cottons. These results suggested that the detection efficiency of AFLP is much higher than that of RAPD and SSR. Pang et al. (2006) did not detect specific sites for Chuan109 with 44 SSR primers. However, we detected 97 specific sites in the sea-island cottons by using AFLP. Another type of site specific to the introgressed lines is not present in the parental strains of sea-island cotton and upland cotton lines. These sites occurred after the hybridization. The introgressed line containing these AFLP specific sites may also integrate DNA sequences from parental strains, which resulted in the molecular weight alterations in partial DNA sequences. Pang et al. (2006) detected 25 of this type of sites in 155 introgressed lines using SSR primers. However, only one site was detected in Chuan109. We detected 36 of this type of sites in Chuan109 using AFLP. In addition, most of the introgressed lines contained more than 10 specific sites. These sites might be closely related to the superior characters in the introgressed lines. We also identified a large amount of sea-island specific AFLP bands in the introgressed lines. These AFLP bands might be related to the resistance of the introgressed lines. Aphid- and acarid-resistant genes should be derived from sea-island cotton, which confirmed the correctness of using superior resistant genes from sea-island cottons through phenotypic observations. The relationship between the specific fragments generated through genetic recombination and the superior traits remained to be further investigated. With this information, we can facilitate the speed of superior genes introgressed from wild cottons into the upland cottons and the process of cotton breeding. Chuan98, 9809, 9812 and Chuan45 had the acarid-resistant character derived from sea-island cotton. Molecular markers for the acarid-resistant genes may be identified from the common sea-island cotton.
specific sites, which avoided generation of near isogenetic lines and bulked segregation analysis. However, those molecular markers for aphid- or acarid-resistant genes were not identified. One of the possible reasons was that the sea-island cotton used in this study cannot completely represent the original parent strains for the introgressed lines, which resulted in the inconsistency in the primary structure of DNA sequences in some genes. The other possible reason was that aphid- or acarid-resistant genes belonged to the main genes, which accounted for very small portion of the whole genome. We used very limited number of AFLP primers and thus these sites were not detected. In addition, every molecular marker technique has its own limitations. For example, AFLP cannot specifically detect the changes in the SSR alleles. Thus, these specific molecular markers can be identified by changing the restriction sites of AFLP or combining with other molecular markers.

The AFLP dendrogram reflected the relationship between each cotton variety on the genetic level, which can provide guidelines for the parental selection, prediction of the genetic variation in the offspring generations and the superiority of the hybrids. However, many RAPD and AFLP studies showed that genetic diversity between upland cotton strains is very limited, which is unfavorable for the breeding. Hybridization between different varieties can resolve this bottleneck. For example, previous studies have comprehensively analyzed the genetic composition of the main varieties obtained by the main Breeding Institutes in China using RAPD and performed genealogical analysis (Guo et al., 1997; Wang et al., 1997; Zhu et al., 2002). The results from these studies suggested that the genetic composition in the currently cultivated cottons is very limited. In addition, some of the varieties and derivatives were overcultivated. With the limited genetic basis, it became very difficult to have breakthroughs in the studies of stress resistance and some economic traits. This bottleneck could be resolved by long distance hybridization with wild varieties. Chuan98, Chuan109 and Chuan45 were obtained by hybridization between sea-island and upland cotton varieties. Many studies have shown that superior upland cotton varieties can be obtained by hybridization between sea-island and upland cotton varieties, but the offspring generations had the feature of severe segregations. Relatively stable introgressed lines, for example, Chuan98, Chuan109 and Chuan45 were successfully obtained by transformation and directional breeding in the late stage. Understanding the genetic relationship between these introgressed lines could provide guidelines for the further rational usage of cotton genetic materials. AFLP analysis showed that the 7 cotton varieties used in this study had a similarity coefficient ranging from 0.3778 to 0.9311. Abundant diversity was observed between different cotton varieties. For example, the smallest similarity coefficient (0.3778) was observed between the aphid-resistant variety, Chuan45 and acarid-resistant variety, Chuan109. Therefore, it is possible to breed cotton variety with the feature of multiple resistances through hybridization between aphid-resistant and acarid-resistant varieties. Our results also suggested that the property of the sea-island cotton can be introduced into the upland cotton variety. It is also feasible to obtain applicable varieties by artificial selection and domestication.

Dendrogram analysis showed that some of the results were unexpected. For example, acarid-resistant variety Chuan98 was clustered together with aphid-resistant varieties Chuan109, 296 and sea-island cottons. In contrast, 169-6 was clustered with other acarid-resistant varieties and Chuan56. Clustering results were not completely dependent on the cotton resistance. Based on the process of selective breeding of introgressed lines, the offspring generations of introgressed line Kang77 were segregated into two types: one was acarid-resistant and the other one was aphid-resistant. Acarid-resistant Chuan98 was bred by selecting an individual strain 98 from the Kang77 variety. Chuan45 was bred by hybridization of [(Chuan98 × Jiangsu cotton#1) Bc2F1×Zhongmiansuo12]. 9809 and 9812 were derived from (Chuan98 × Anti-disease86-1) Bc1F13. d115 was derived from [(Chuan98 × Jiangsu cotton#1) Bc2F1×Zhong381]. Aphid-resistant varieties Chuan109, 169-6 and 169-296 were derived from the segregation population of (Kang77 × Jiangsu cotton#1) Bc1. Directional breeding was based on the characteristics of the resistance and their own segregation. The characteristics of resistance only accounted for a very small portion of all the properties. Thus, Chuan98, Chuan109 and 169-296 were clustered together because these upland cotton varieties contained many features of sea-island. In contrast, although 169-6, 169-296 and Chuan109 were derived from the same segregation population, 169-6 was clustered together with Chuan56 which was more similar to the upland cotton 9809, 9812, Chuan45 and d115. 169-6 was clustered together with acarid-resistant varieties. We speculate that this cluster was not dependent on the resistance, but on the whole DNA sequence similarity. This also suggested that resistance gene only accounted for a very small portion of the whole DNA sequences. In addition, the specific sites between different varieties can also demonstrate the dendrogram results. There were 92 fragments that were introduced from sea-island cotton to Chuan98. The field characteristics of the Chuan98 were also tended to sea-island cottons. In contrast, there were fewer fragments that were introduced from sea-island cotton to 169-6. These results demonstrated the correctness of the AFLP technique.

Further dendrogram analysis showed that the clustering of 7 introgressed lines based on the molecular markers was not completely consistent with the genealogical clustering. Based on the breeding process of the introgressed lines, the offspring generation of Kang77 was segregated into two types: one was asphid-
resistant and the other one was acarid-resistant. Acarid-resistant variety Chuan98 was bred from the individual strain 98 of Kang77 variety. Chuan45 was derived from [(Chuan98 × Jiangsu cotton#1) Bc2F1×Zhongmiansuo12], 9809 and 9812 were derived from the segregation population of (Chuan98 × anti-disease86-1) Bc1F13. Aphid-resistant varieties Chuan109, 169-6 and 169-296 were derived from the segregation population of (Kang77 × Jiangsu cotton#1) Bc1. AFLP clustering results was generally consistent with the genealogical clustering. AFLP clustering of Chuan98 and 169-6 was different to their genealogical clustering. The possible reason was that the offspring generations of 169-6 were changed significantly after multiple self cross-over, even though the parental strain of 169-6, Chuan109 and 169-296 were derived from Kang77 and Jiangsu cotton#1. The phenotypic features of 169-6 were more similar to other upland cotton varieties after directional breeding example, Chuan98. Since the directional breeding was based on the resistance and the segregation characteristics during the process of breeding, resistance feature accounted for only a very small portion of the whole characteristics. Therefore, the difference between AFLP dendrogram and genealogical clustering occurred.

In summary, our results suggested that interspecific hybridization is feasible to broaden germplasm of upland cotton and AFLP technique could be applied for the identification of variety purity and genealogical relationship.

REFERENCES


