Genetic Diversity Analysis and Fingerprinting of Diploid Cotton Cultivars Using Automated Amplified Fragment Length Polymorphism (AFLP)

MK Rana and KV Bhat

National Research Centre on DNA Fingerprinting, National Bureau of Plant Genetic Resources, Pusa Campus, New Delhi-110012

Key Words: Amplified Fragment Length Polymorphism (AFLP), Genetic diversity, *Gossypium* spp., Molecular markers

Cotton is an important cash crop of India. Nearly one third of India's earnings are from textile sector of which cotton alone contributes about 70% of raw material. All the four cultivated species of cotton are grown in the country of which 30% area is under diploid cottons.

The presence of genetic diversity in any species is important for its improvement. A large number of polymorphic markers are required for establishing genetic relationships among the cultivars and to assess the genetic diversity. Morphological markers are generally used for this purpose, but today we have various molecular marker systems available. Amplified fragment length polymorphism (AFLP) is a very powerful marker system that includes several-folds increase in the number of informative markers per analysis, highly reproducible banding pattern and no *a priori* sequence information of the DNA. It has been used to estimate genetic relationships in many studies including lentil (*Lens culinaris* Medikus) (Sharma *et al.* 1996), soybean (*Glycine max* L.) (Maughan *et al.* 1996), hops (*Humulus lupulus* L.) (Hartl and Seefelder, 1998), coconut (*Cocos nucifera* L.) (Perera *et al.* 1998), wheat (*Triticum aestivum* L.) (Barrett and Kidwell, 1998), neem (*Azadirachta indica*) (Singh *et al.* 1999), cotton (*Gossypium* spp.) (*Pillay and Myers*, 1999) *etc.* In the present study AFLP analysis was performed in diploid cotton cultivars to find the suitability of the technique for genetic diversity analysis and fingerprinting.

A total of 16 diploid cotton cultivars belonging to two species (Table 1) were used in the present study. These cultivars were developed and released under the Indian Agricultural Research Systems for cultivation in the specified regions of the country. Out of these 16 cultivars, seven belonged to *G. herbaceum* and eight to *G. arboreum*. One genotype was inter-specific hybrid between these two species.

DNA was extracted from 5 g of a bulked sample of leaves from 10 plants of each cultivar. The CTAB

Table 1. List of cotton cultivars used in the study

Cultivar	Species	Cultivar	Species
Sujay	*G. herbaceum*	Sanjay	*G. arboreum*
Gcot 21	*G. herbaceum*	Gcot 19	*G. arboreum*
Gcot 11	*G. herbaceum*	Gcot 15	*G. arboreum*
Gcot 13	*G. herbaceum*	G-27	*G. arboreum*
V797	*G. herbaceum*	DDCC-1	*G. arboreum*
Jayadhar	*G. herbaceum*	RG8	*G. arboreum*
4011	*G. herbaceum*	HD327	*G. arboreum*
824	*G. arboreum*	G.CotDH9	*G. herbaceum* x *G. arboreum* (hybrid)

DNA extraction procedure of Sagahai-Maroof *et al.* (1984) was used with some modifications for DNA extraction. The concentration of DNA in the RNA-free samples was determined with a Hoefer DNA Fluorometer Model DQ 200 (Hoefer Pharmacia Biotech Inc., San Francisco, CA) using Hoechst 33258 as the dye and calf thymus DNA as the standard (Brunk *et al.* 1979).

AFLP analysis was performed according to the protocol of Vos *et al.* (1995) that is supplied with AFLP Plant Mapping protocol from Perkin Elmer Applied Biosystems (USA) with slight modifications to suit to the material under investigation. DNA restriction, ligation, and pre-selective and selective amplifications of the samples were done as per the protocol. All amplifications were carried out in Perkin Elmer thermocycler (GeneAmp PCR System 9600). Six primer combinations of EcoRI+3 (E-plus three nucleotides) and MseI+3 (M-plus 3 nucleotides) were selected from initially 64 screened primer combinations (Table 2).

Samples containing 3 µl of selective PCR products, 0.5 µl of Gene Scan ROX 500 internal lane standard and Formamide were heated at 96°C for 5 min and quick chilled by placing on ice. DNA samples were capillary electrophoresed at constant power on an automated DNA sequencer (Perkin Elmer/Applied Biosystems model ABI Prism 310) equipped with GeneScan software (Perkin Elmer Applied Biosystems).

Table 2. AFLP primer combinations, cultivars distinguished and level of polymorphism detected

Primer Combination	Cultivars	Species-specific Distinguished	% Polymorphism Markers
E-AAG/ M-CAA	5	8	76.9
E-ACA/ M-CAA	3	10	50.0
E-ACT/ M-CTT	10	7	58.2
E-AAG/ M-CAG	7	8	83.5
E-ACA/ M-CAG	6	12	85.0
E-ACT/ M-CAG	8	7	64.6

After the ABI Prism 310 collected data, GeneScan analysis software was used to analyze and display the results. Results so obtained were transformed into binary data using software Genotyper version 2.5. Further computing for UPGMA analysis to generate a dendrogram and for calculating similarity coefficients (Jaccard, 1908) were carried out using NTSYS-software (Rohlf, 1988).

AFLP markers were assayed for their usefulness in analyzing molecular diversity and fingerprinting 16 cultivars of cotton belonging to two cultivated species *i.e. G. herbaceum* and *G. arboreum.* A total of 751 markers were obtained with six primer combinations across all the cultivars (Table 2). Out of these 751 markers, 523 were polymorphic indicating 70% polymorphism. On an average 87 markers per primer-combination were recorded.

Primer combination E-ACT/M-CTT discriminated the maximum number of cultivars (10) by producing cultivar specific markers, which was followed by primer combination E-ACT/ M-CAG that discriminated as many as 8 cultivars. All the cultivars but one (Gcot 19) could be distinguished from one another using this set of six primer combinations. Only two primer combinations E-ACT/M-CTT and E-AAG/M-CAG distinguished as many as 14 cultivars. An average of 8.6 markers per primer combination were found that were specific to either *G. herbaceum* or *G. arboreum* species (Table 2). Earlier cultivar identification in cotton has been reported using RAPD (Multani and Lyon, 1995; Iqbal *et al.* 1997) markers but no single RAPD primer could discriminate all the cultivars in their studies. So far we have studied only six primer combinations and further studies are underway to find a single primer combination that could discriminate all the cultivars.

AFLP markers obtained across all the cultivars were subjected to UPGMA analysis to determine genetic relationships among the cultivars (Table 3). A range of 0.50 to 0.88 for similarity coefficient values was observed across the two species. Gcot11 and Gcot13 had the maximum similarity value (0.85) among the *G. herbaceum* cultivars, while Jayadhar and Gcot11 had the least similarity coefficient value (0.68). Between *G. arboreum* cultivars Gcot15 with 824 and Gcot19 had the maximum similarity (0.88) while G-27 and Sanjay were the least similar (0.70). The close similarity between Gcot19 and Gcot15 and that between Gcot11 and Gcot13 may be explained by some commonness in their ancestry. Gcot15 and Gcot19 were procured from the same place. Average genetic similarity index across all 16-variety comparisons was found to be 66.0. It was 78.0 and 73.0 among the cultivars of *G. herbaceum* and *G. arboreum*, respectively. However,

Table 3. Similarity coefficient values among diploid cotton cultivars using AFLP analysis

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Sujay	1.00															
Gcot21	0.81	1.00														
Gcot11	0.81	0.80	1.00													
Gcot13	0.80	0.80	0.85	1.00												
V797	0.81	0.78	0.81	0.83	1.00											
Jayadhar	0.71	0.69	0.68	0.71	0.72	1.00										
4011	0.79	0.80	0.83	0.84	0.81	0.74	1.00									
GcotDH9	0.78	0.78	0.79	0.80	0.80	0.74	0.83	1.00								
824	0.66	0.52	0.56	0.56	0.54	0.50	0.56	0.55	1.00							
Sanjay	0.63	0.55	0.54	0.54	0.54	0.54	0.56	0.57	0.66	1.00						
Gcot19	0.64	0.53	0.57	0.55	0.55	0.54	0.57	0.57	0.78	0.76	1.00					
Gcot15	0.63	0.61	0.65	0.64	0.66	0.60	0.65	0.67	0.88	0.85	0.88	1.00				
G27	0.63	0.54	0.56	0.56	0.55	0.55	0.57	0.56	0.71	0.70	0.78	0.81	1.00			
DDCC1	0.63	0.55	0.57	0.55	0.54	0.53	0.59	0.56	0.77	0.73	0.81	0.81	0.77	1.00		
RG8	0.66	0.55	0.58	0.58	0.57	0.54	0.57	0.58	0.77	0.74	0.82	0.85	0.79	0.81	1.00	
HD327	0.64	0.53	0.56	0.54	0.53	0.50	0.56	0.54	0.74	0.67	0.75	0.75	0.73	0.78	0.76	1.00

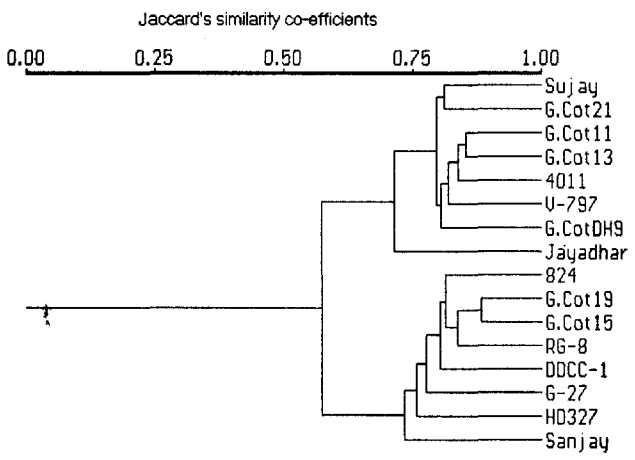

Fig. 1. Dendrogram generated through UPGMA analysis for 16 cultivars diploid cotton

genetic similarity index was 57.0 between the cultivars of the two species. Cluster analysis revealed separate clustering of cultivars of both the species (Fig. 1). GcotDH-9, an inter-specific hybrid, clustered with its female parent 4011. Jayadhar was the most distinct cultivar and did not cluster with any of the cultivars of its group of *G. herbaceum* cultivars. *G. arboreum* cultivars 824, Gcot19, Gcot15 and RG8 formed clear clustering into one group. Separate clustering of *G. hirsutum* and *G. barbadense* cultivars has been reported earlier using allozyme (Percy and Wendel, 1990), RAPD (Tatineni *et al.* 1996) and AFLP (Pillay and Myers, 1999) analyses.

In conclusion, this study shows that AFLP is able to discriminate closely related genotypes in cotton and provides sufficient numbers of polymorphic markers in a few experiments. With the automation of the AFLP technique and application of multiplex PCR a large number of samples can be analysed in a short time paving the way for effective characterization and conservation of cotton plant genetic resources.

References

Barrett BA and KK Kidwell (1998) AFLP-based genetic diversity assessment among wheat cultivars from the Pacific North-West. *Crop Sci.* **38:** 1261-1271.

Brunk CF, KC Jones and TW James (1979) Assay for nanogram quantities of DNA in cellular homogenates. *Anal. Biochem* **92:** 497-500.

Hartl L and S Seefelder (1998) Diversity of selected hop cultivars detected by fluorescent AFLPs. *Theor. Appl. Genet.* **96:** 112-116.

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

Jaccard P (1908) Nouvelles recherchés sur la distribution florale. *Bull. Soc. Vaud. Sci. Nat.* **44:** 223-270.

Maughan PJ, MA Saghai-Maroof, GR Buss and GM Huestis (1996) Amplified fragment length polymorphism (AFLP) in soybean: species diversity, inheritance, and near isogenic lines analysis. *Thoer. Appl. Genet.* **93**: 392-401.

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

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

Perera L, JR Russell, J Provan, JW Mc Nicol and W Powell (1998) Evaluating genetic relationships between indigenous coconut (*Cocos nucifera* L.) accessions from Sri Lanka by means of AFLP profiling. *Theor. Appl. Genet.* **96:** 545-550.

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

Rohlf FJ (1993). NTSYS-pc Numerical taxonomy and multivariate analysis system, version 1.80. Exeter Software, Setauket, New York.

Saghai-Maroof MA, KM Soliman, RA Jorgensen and RW Allard (1984) Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics, *Proc. Natl. Acad. Sci.* (USA) **81:** 8014-8018.

Sharma SK, MR Knox and THN Ellis (1996) AFLP analysis of the diversity and phylogeny of *Lens* and its comparison with RAPD anlaysis. *Theor. Appl. Genet.* **93:** 751-758.

Singh A, MS Negi, J Rajagopal, S Bhatia, UK Tomar, PS Srivastava, and M Lakshmikumaran (1999) Assessment of genetic diversity in *Azadirachta indica* using AFLP markers. *Theor. Appl. Genet.* **99:** 272-279.

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

Vos P, R Hogers, M Reijans, T Van de Lee, M Hornes, A Freijters, J Pot, J Pleman, M Kuiper and M. Zabeau (1995) AFLP: a new technique for DNA Fingerprinting. *Nucleic Acid Res.* **23:** 4407-4414.