Indian J. Genet., **62**(1): 11-14 (2002)

Genetic diversity analysis in Indian diploid cotton (*Gossypium* spp.) using RAPD markers

M. K. Rana and K. V. Bhat

National Research Centre on DNA Fingerprinting, N.B.P.G.R., Pusa Campus, New Delhi 110 012

(Received: June 2001; Accepted: December 2001)

Abstract

Random amplified polymorphic DNA (RAPD) analysis was employed in commercially released 18 cultivars of diploid cotton belonging to two species i.e. *Gossypium arboreum* and *Gossypium herbaceum*. The 20 primers, selected from 60 initially screened, generated 224 amplification products of which 127 amplicons were polymorphic. All primers produced polymorphic amplification products, however, the extent of polymorphism varied with each primer. Statistical analysis was carried out using NTSYS-pc software and a dendrogram was generated using Jaccard's similarity coefficients. The values for similarity coefficients ranged from 0.50 to 0.99. Cluster analysis showed clear-cut separation of the genotypes of the two species. *G. herbaceum* cultivar Jayadhar did not cluster with any other cultivar. *G. arboreum* cultivars formed three clusters (II, III and IV). Cluster II contained three out of four cultivars from the same place, whereas cluster IV contained cultivars from Punjab and Haryana.

Key words : *Gossypium* spp., cultivars, genetic diversity, random amplified polymorphic DNA (RAPD)

Introduction

Cotton (*Gossypium* spp.) is comprised of about 50 diploid and tetraploid species but world's cotton fiber is produced from four species, *G. arboreum* L. (n = 13, A genome), *G. herbaceum* L. (n = 13, A genome), *G. barbadense* L. (n = 26, AD genome), and *G. hirsutum* L. (n = 26, AD genome). Though *G. hirsutum* accounts for 70% of the total acreage in India, the present species composition of 30% under *G. arboreum* and *G. herbaceum* varieties should be stabilized at least at this level, by systematic improvement so that this genetic resource for tolerance to moisture stress, high harvest index, resistance to leaf curl irus, and suitability for some specific end uses, is not eroded further.

For any crop improvement programme, analysis of genetic diversity is the first and foremost step. To have a reliable estimate of genetic relationships and genetic diversity a large number of polymorphic markers are required. This limits the use of morphological characteristics which are few or which lack adequate levels of polymorphism in *Gossypium*. Most of genetic diversity analysis studies in cotton have been carried out using morphological markers only. Now-a-days Polymerase Chain Reaction (PCR)-based molecular markers [1-4] have developed into powerful tools to analyze genetic relationships and genetic diversity. Random amplified polymorphic DNA (RAPD) first reported by Williams *et al.* [1] and Welsh & Mc Clelland [2] is one of such tools. It is a powerful technique and its resolving power is several folds higher than morphological and isozyme markers and is much simpler and technically less demanding than other such techniques. Recently, RAPD analysis has been used for diversity analysis in a vast array of field crops [5-11] and results for genetic diversity analysis in diploid cotton are presented here for the first time. It has been used earlier for genetic diversity analysis of tetraploid species in cotton [12-14].

Materials and methods

Plant material and DNA extraction

The plants of 18 cultivars (Table 1) were grown in pots in the greenhouse. DNA was extracted from 5g of a bulked sample of leaves from 10 plants of each cultivar. The CTAB DNA extraction procedure of Sagahai-Maroof [15] was used with some modifications for DNA extraction. The concentration of DNA in the RNA-free samples was determined with a Hoefer DNA Fluorometer and the DNA samples were diluted in 10:1 TE to a working concentration of approximatey 10ng μl^{-1} and stored at 4°C until PCR amplification.

DNA amplification and gel electrophoresis

Protocol for PCR was optimized by varying the concentration of template DNA, *Taq* DNA polymerase and magnesium salt. Twenty 10-base oligonucleotide primers (Operon Technologies, Inc., U.S.A.) were selected for final RAPD-PCR amplification. Each reaction

Table 1. List of cultivars used and their pedigrees

S. No.	Cultivar	Pedigree	Place of release
I) *Gossypium herbaceum*			
1.	Sujay	Derivative of B.C.1-6 × S.q-1-4	GAU, Surat
2.	G. Cot 21	1502E × (G. Cot 13 × 4011 D)	GAU, Viragram
3.	G. Cot 11	(3200 × EP2) F10	GAU, Viragram
4.	G. Cot 13	Kalyan × 1802	GAU, Viragram
5.	V797	Kalyan × (Kalyan × Vijay)	GAU, Viragram
6.	Jayadhar	Derivative from A-56-347 × Acala 44-2	UAS, Dharwad
7.	4011	3943/5021 × (Digvijay × Raniben) F.22	GAu, Surat
8.	G. CotDH9	4011 × 824	GAU, Surat
II) *Gossypium arboreum*			
9.	824	Reselection from long staple Gaorani cotton	GAU, Surat
10.	Sanjay	Derivative of C-520 × Jarilla	GAU, Amreli
11.	G. Cot 19	Jyoti × *arboreum* N.K.	GAU, Amreli
12.	G. Cot 15	Sanjay × Chandrolla	GAU, Amreli
13.	G27	Selection from local collection of Sanguineum cotton	GAU, Surat
14.	DDCC-1	-	UAS, Dharwad
15.	RG-8	Developed from G1 × *Cernum*	RAU, Sri Ganganagar
16.	LD 327	-	PAU, Ludhiana
17.	HD 107	Selection from local germplasm	HAU, Ludhiana
18.	402	-	-

- Indicates not available

mixture (25 µl) for PCR amplification consisted of 1x reaction buffer (10 mM Tris-HC1, pH 8.3 and 50 mM K C1), 3.0 mM Mg Cl_2, 2U of *Taq* DNA polymerase; 200 µM each of dATP, dTTP, dCTP and dGTP (all reagents from Perkin Elmer), 0.5 µM of 10 nt primer (Operon Technologies, Alameda, USA) and approximately 50ng of genomic DNA template. The PCR amplification conditions were as follows: initial extended step of denaturation at 94°C for 3 min followed by 40 cycles of denaturation at 94°C for 1 min, primer annealing at 36°C for 1 min and elongation at 72°C for 1 min. The 40th cycle was followed by an extended primer extension step at 72°C for 4 min and then held at 4°C until electrophoresis. The amplification products were electrophoresed on 1.6% agarose gel at 100 volts in 1x TAE buffer. A 1kb DNA ladder was used as a molecular standard. The gel was stained in the presence of ethidium bromide and gels were photographed under UV light.

Table 2. Primer, its sequence and level of polymorphism detected

Primer	Sequence (5'-3')	Total no. of bands	No. of polymorphic bands	Percent polymorphism
OPA-01	CAGGCCCTTC	13	6	46.2
OPA-02	TGCCGAGCTG	6	5	83.3
OPA-05	AGGGGTCTTG	16	7	43.8
OPA-06	GGTCCCTGAC	14	7	50.0
OPA-07	GAAACGGGTG	18	13	72.2
OPA-09	GGGTAACGCC	7	4	57.1
OPA-11	CAATCGCCGT	9	8	54.5
OPA-14	TCTGTGCTGG	9	3	33.3
OPA-15	TTCCGAACCC	11	8	72.7
OPA-16	AGCCAGCGAA	16	9	56.3
OPA-17	GACCGCTTGT	10	7	70.0
OPA-19	CAAACGTCGG	15	11	73.3
OPA-20	GTTGCGATCC	13	10	76.9
OPB-01	GTTTCGCTCC	8	2	25.0
OPB-02	TGATCCCTGG	11	9	81.8
OPB-03	CATCCCCCTG	8	3	37.5
OPB-04	GGACTGGAGT	11	4	36.4
OPB-05	TGCGCCCTTC	11	3	27.3
OPB-06	TGCTCTGCCC	7	2	28.6
Total		224	127	

Scoring and data analysis

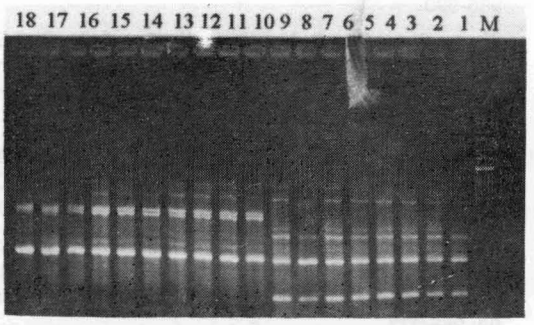
For all the genotypes, bands on RAPD gels were scored as present (1) or absent (0). Missing and doubtful cases were scored as '9'. Jaccard's similarity coefficient values [16] were calculated for each pair-wise comparison between cultivars and a similarity coefficient matrix was constructed. This matrix was subjected to unweighted pair-group method for arithmetic averages analysis (UPGMA) to generate a dendrogram using average linkage procedure. All computing were carried out using NTSYS-pc software [17].

Results and discussion

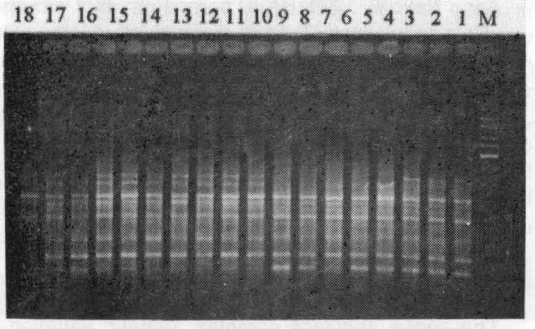
Eighteen cultivars (Table 1), representing two species of cotton i.e. *G. herbaceum* and *G. arboreum*, were used for genetic diversity analysis using RAPD markers. Out of these 18, seven belonged to *G. herbaceum* and 10 to *G. arboreum* species. One genotype G. Cot DH9 was an inter- specific hybrid between these two species.

A total of 20 10-mer primers (Operon Technologies, Inc., U.S.A.) were selected for final RAPD-PCR analysis from 60 initially screened. Selection of these 20 primers was done on the basis of their multibanded and easily scorable amplification products. A typical example with primers OPA-11 and OPA-14 is shown in Fig. 1. Total 224 amplicons were obtained with these 20 primers across the two species with 11.2 bands per primer. The number of bands produced

ranged from 6 (OPA-02) to 18 (OPA-07). Out of these 224 bands, 127 were found to be polymorphic and the level of polymorphism was 56.7% across the two species of Old World. Species-wise polymorphism was found to be 6.3% and 23.2% in *G. herbaceum* and *G. arboreum*, respectively. Of the total 224 bands 27.5% were specific to *G. herbaceum* whereas, 17.9% were specific to *G. arboreum*, indicating their importance in species identification.



a) Primer OPA-11



b) Primer OPA-14

Fig. 1. RAPD amplification of diploid cotton cultivars with primer OPA-11 (a) and OPA-14(b). The lane numbers written on the top correspond to cultivars as listed in Table 1. The lane labeled M is Gene Ruler 1 kb DNA ladder (MBI Fermentas, UK)

Jaccard's pair-wise similarity coefficient values for 18 cultivars were calculated and are presented in Table 3. The range of genetic similarities was found to be 0.50-0.99. Average genetic similarity among these 18 genotypes was found to be 0.72 ± 0.02. However, average genetic similarities individually among *G. herbaceum* and *G. arboreum* cultivars were recorded to be 0.98 ± 0.002 and 0.90 ± 0.006, respectively. A comparison of these values shows that *G. herbaceum* possesses less genetic diversity than *G. arboreum*. Average genetic similarity between the two species was 0.52 ± 0.001. Multani and Lyon [12] also found high genetic similarity (92.1-98.9%) among nine Australian cotton cultivars belonging to *G. hirsutum*. Similarly, Iqbal *et al.* [13] found 81.5-93.5% genetic similarity among 17 *G. hirsutum* cultivars. We observed 28.0% genetic diversity in these two species, which is corroborated by the findings of Wendel *et al.* [17] who also reported 28.0% genetic diversity in *G. arboreum* cultivars.

A dendrogram was generated by UPGMA cluster analysis based on Jaccard's similarity coefficients (Fig. 2). Cluster analysis revealed that at about 50% genetic similarity, the genotypes of the two species clustered in their respective groups. Among. *G. herbaceum* cultivars, all except Jayadhar clustered together. All the cultivars in this group are from Gujarat Agricultural University and Jayadhar only is from different breeding station i.e. University of Agricultural Sciences, Dharwad. Distinctness of Jayadhar can also be attributed to the alien blood (Acala 44-2) in its ancestry. *G. arboreum*

Table 3. Jaccard's similarity coefficient values among diploid cotton cultivars using RAPD analysis

S.No.	Cultivars	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	Sujay	1.00																	
2	G.Cot 21	0.99	1.00																
3	G.Cot 11	0.98	0.97	1.00															
4	G.Cot 13	0.98	0.98	0.99	1.00														
5	V 797	0.99	0.99	0.98	0.98	1.00													
6	Jayadhar	0.96	0.95	0.96	0.95	0.97	1.00												
7	4011	0.99	0.98	0.98	0.98	0.99	0.96	1.00											
8	G.CotDH9	0.98	0.98	0.97	0.97	0.99	0.97	0.99	1.00										
9	824	0.52	0.51	0.52	0.52	0.52	0.53	0.52	0.52	1.00									
10	Sanjay	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.53	0.95	1.00								
11	G.Cot 19	0.53	0.52	0.53	0.53	0.53	0.53	0.53	0.53	0.95	0.96	1.00							
12	G. Cot 15	0.52	0.52	0.53	0.52	0.53	0.53	0.53	0.53	0.95	0.96	0.96	1.00						
13	G27	0.53	0.52	0.52	0.52	0.53	0.53	0.52	0.53	0.9	0.94	0.9	0.93	1.00					
14	DDCC-1	0.54	0.54	0.54	0.54	0.55	0.55	0.54	0.55	0.91	0.91	0.91	0.92	0.94	1.00				
15	RG8	0.54	0.54	0.55	0.54	0.55	0.55	0.55	0.55	0.92	0.94	0.93	0.94	0.94	0.94	1.00			
16	HD327	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.50	0.83	0.88	0.86	0.86	0.88	1.00		
17	HD107	0.52	0.51	0.51	0.51	0.51	0.51	0.51	0.51	0.84	0.88	0.87	0.85	0.89	0.88	0.88	0.91	1.00	
18	402	0.50	0.50	0.51	0.50	0.51	0.51	0.51	0.51	0.85	0.86	0.86	0.86	0.87	0.87	0.89	0.89	0.93	1.00

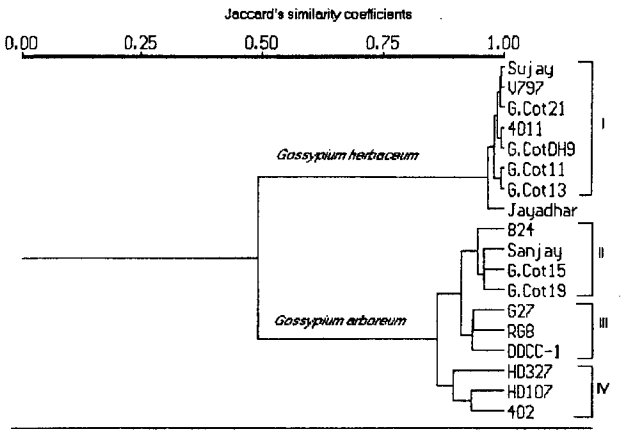


Fig. 2. Dendrogram generated using UPGMA analysis showing relationships between 18 cultivars of cotton using RAPD markers

cultivars can be grouped in three clusters (clusters II, III and IV), which is also in agreement with their pedigrees. In cluster II, three cultivars (Sanjay, G. Cot 19, and G. Cot 15) are from the same breeding station and they have commonness in their pedigrees. However, in cluster III all three cultivars are from three different places. Cultivars from Punjab and Haryana were grouped in cluster IV. Earlier cluster analysis using RAPD markers has been carried out in tetraploid cotton and clear-cut separation of *G. hirsutum* and *G. barbadense* cultivars was observed [14].

Based on similarity coefficients and cluster analysis genotypes Jayadhar, 402, HD 107 and LD 327 were found to be quite distinct and these can be used for their desirable characteristics in breeding programmes for cotton improvement. The genetic similarities obtained from the analysis can also be used for the selection of parents to generate mapping populations and for selecting parents for breeding purposes. The narrow genetic base reveled in the present study emphasizes the need to exploit large germplasm collections having diverse morphoagronomic traits in cultivar improvement programs.

Acknowledgements

The authors are grateful to the Indian Council of Agricultural Research, for the financial support for this work.

References

1. **Williams J. G. K., Kubelik A. R., Livak K. J., Rafalski J. A. and Tingey S. V.** 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res., **18**: 6531-6535.

2. **Welsh J. and Mc Clelland M.** 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Res., **18**: 7213-7218.

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

4. **Thomas M. R. and Scott N. S.** 1993. Microsattellite repeats in grapevine reveal DNA polymorphism when analysed as sequence-tagged sites (STSs). Theor. Appl. Genet., **86**: 985-990.

5. **Zhang C. L., Sun S. M., Jin D. M., Sun Z. L., Guo B. T. and Wang B.** 1998. Rapid identification of twelve elite maize inbred lines using RAPD markers. Acta Agronomica Sinica, **24**: 718-722.

6. **Ahmad F.** 1999. Random Amplified Polymorphic DNA analysis reveals genetic relationships among annual *Cicer* species. Theor. Appl. Genet., **98**: 657-663.

7. **Davila J. A., Loarce Y. and Ferrer E.** 1999. Molecular characterization and genetic mapping of random amplified microsatellite polymorphism in barley. Theor. Appl. Genet., **98**: 265-273.

8. **Fahima T., Sun G. L., Behrav A., Kuigama T., Beiles Aand Nevo E.** 1999. RAPD polymorphism of wild emmer wheat populations (*Triticum dicoccoides*) in Israel. Theor. Appl. Genet., **98**: 434-447.

9. **Rodrigue J. M., Berke T., Englé L. and Nienhuis J.** 1999. Variation among and within *Capsicum* species revealed by RAPD markers. Theor. Appl. Genet., **99**: 147-156.

10. **Lakhanpaul S., Chadha S. and Bhat K. V.** 2000. Random amplified polymorphic DNA (RAPD) analysis in Indian mung bean (*Vigna radiata* (L.) Wilczek) cultivars. Genetica, **109**: 227-234.

11. **Xu R. Q., Tomooka N., Vaughan D. A. and Doi K.** 2000. The *Vigna angularis* complex: genetic variation and relationships revealed by RAPD analysis and their implications for in situ conservation and domestication. Genet. Res. Crop Evol., **47**: 123-134.

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

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

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

15. **Saghai-Maroof M. A., Soliman K. M., Jorgensen R. A. and Allard R. W.** 1984. Ribosomal DNA spacer-length polymorphism in barley: Mendelian inheritance, chromosomal location, and population dynamics. Proc. Natl. Acad. Sci., (USA), **81**: 8014-8018

16. **Jaccard P.** 1908. Nouvelles recherches sur la distribution forale. Bull Soc Vaud Sci Nat., **44**: 223-270.

17. **Rohlf F. J.** 1993. NTSYS-pc Numerical taxonomy and multivariate analysis system, version 1. 80. Exeter Software, Setauket, New York.

18. **Wendel J. F., Olson P. D. and Stewart J. M.** 1989. Genetic diversity, introgression and independent domestication of old world cultivated cotton. Amer. J. Bot., **76**: 1795-1806.