J. Cotton Res. Dev. **29** (2) 232-236 (July, 2015)

Molecular marker based genetic diversity analysis in cotton using RAPD and SSR markers

SALONI TYAGI*, S.K. VERMA , KIRAN NEHRA AND D.R. BAJYA
Central Institute for Cotton Research,Regional Station, Sirsa - 125 055
**E-mail : salonityagi10@gmail.com*

ABSTRACT : The relative efficiency of Randomly Amplified Polymorphic DNA (RAPD) and Simple Sequence Repeats (SSR) markers was assessed for their suitability and reliability for estimating genetic diversity in cotton. The genetic variability and relationships among 23 cultivars of *Gossypium hirsutum* were investigated using 10 RAPD and 10 SSR primers, which produced 34 and 22 amplicons, respectively. The number of polymorphic amplicons was found to be 18 and 5, resulting into a polymorphism equivalent to 52.9 and 22.7, respectively. The genetic relationship was examined using the Numerical System Package (NTSYS-PC), UPGMA analysis was performed and dendrograms were constructed. Based on the RAPD and SSR combined data, the genetic similarities ranged from 54 to 96 per cent, respectively. All dendrograms clustered the 23 genotypes into 2 major clusters (A and B) showing affinity at genus level; most of the genotypes (20 out of 23) falling into cluster A. Maximum similarity (83%) was found in the cultivars RS 2013, H1098, LD694, LD327, H1117, LH1556, HD123, CISA310, RS921, HS6, F1861 and F846, followed by the cultivars H1226, RST9, H777, F505 and F1378 showing 78 per cent similarity; and RS875, LH900 and RS810 showing 76 per cent similarity.

Keywords : Cotton, genetic diversity, molecular markers, RAPD, SSR

Cotton, *Gossypium* spp., is an economically important crop that is grown throughout the world as a source of fibre, food and feed. The cotton belonging to the genus *Gossypium* (Malvaceae) comprises of approximately 50 species, distributed in various continents except Europe. It has 4 cultivated species: *Gossypium hirsutum* L., *G. barbadense, G. arboreum* L. and *G. herbaceum* L. *G. hirsutum* is the predominant cultivated cotton with high yield and wide adaptation, while its closely related species, *G. barbadense* is grown for its extra long, strong, and fine fibre. The diploid species, *G. arboreum* L. and *G. herbaceum* L. are cultivated only in very small acreage.

Due to the narrow genetic base of cotton germplasm that cotton breeders have been utilizing low efficiency of traditional selection methods, cultivar improvement in cotton has slowed down in the past 10–15 years. Analysis of genetic diversity and relatedness between species and among genotypes is useful in plant breeding programs because it provides a tool for accurate organization of germplasm and efficient parental selection. During the past few decades, a number of molecular techniques have been recruited to complement traditional methods for the evaluation of biodiversity, estimation of relatedness and genotype identification (Rana *et al.*, 2005). DNA markers are, by far, the most powerful and the most widely used tool to uncover informative polymorphism and genome variability (Gostimsky *et al.*, 2005) With the development of the polymerase chain reaction (PCR) technology, the number of useful DNA based marker types has been expanded considerably. Different PCR based marker types are available now a days including Randomly Amplified Polymorphic DNA (RAPDs) Microsatellites (SSRs). Expressed Sequence Tags (ESTs) (Guo *et al.*, 2007) and Amplified Fragment Length Polymorphisms (AFLPs)

Keeping in view the above facts, the present investigation was conducted as an attempt to address the following objectives: (1) to estimate the genetic polymorphism among twenty three cotton genotypes using two types of PCR based DNA markers, *i.e.*, RAPDs and EST-

SSRs, (2) to assess the genetic relationship among these genotypes, and (3) to identify unique DNA markers and determine a typical fingerprint for each genotype.

Plant material and DNA Extraction: The cotton genotypes included in this study consisted of 23 genotypes of *Gossypium hirsutum* which includes LD694, RS2013, HD 123, LD 327, CISA 310, F 846, H 1098, RS 875, RS 810, RS 921, H 1236, LH 900, F 1861, RST 9, F 1378, H 777, F 505, H 1117, LH 1556, HS 6, RG 8, RST 9, PA255. Seed material was obtained from the collection of the Cotton Research Institute, Sirsa. Total genomic DNA was isolated from young leaf material . This was standardized further for cotton leaf tissue.

Molecular analysis : RAPD amplification was performed as described by using twenty 10 mer random primers (Table 2). SSR assays were performed. Ten SSR primer pairs (Table 2) to be employed in SSR analysis were selected from the cotton database. The PCR amplification products for RAPD's and SSRs were resolved on 1.5 per cent agarose gels, and were detected by staining with ethidium bromide.

Data analysis : The banding patterns generated by RAPD and SSR were examined to determine the level of polymorphism and the genetic relatedness among the 23 cotton genotypes. The amplified fragments were scored as present (1) or absent (0). Cluster analysis was based on similarity matrix obtained with unweighted pair group method using arithmetic average (UPGMA), and the relationships between genotypes were displayed as dendrogram.

Polymorphism as detected by RAPD analysis : A total of 10 primers against 23 cotton genotypes were selected for their scorable results. The amplification profiles for 23 cotton genotypes produced by 10 RAPDs revealed a total of 18 polymorphic bands out of 34 reproducible products (Table 1), corresponding to a level of polymorphism of 52.9 per cent. Hussein *et al.*, 2006 assayed 21 cotton accessions using 28 RAPD primers. The total number of amplicons detected was 323, while, the number of polymorphic amplicons was 191. Thus, the level of polymorphism among the 21 accessions was 59.1 per cent. On the other hand, working on 31 *Gossypium* species, recorded a level of polymorphism of 99.8 per cent. The number of amplicons/ primers ranged from 1 (RAPD 2 and 4) to 5 (RAPD 5, 6, 7 and 10), whereas the number of polymorphic bands/primer ranged from 0 (RAPD 2 and 4) to 4 (RAPD 10). The dendrogram chart (Fig. 1) shows genetic similarity and diversity among the 23 cotton cultivars, the range of dendrogram being 0.57 to 1.00. Dendrogram revealed 2 distinct clusters A and B. Maximum number (20 out of 23) of cultivars fell in cluster A. Cluster A was divided into 2 sub clusters A1 and A2. A1 contained genotype RS810, LH900 and RS875. Cluster A2 was further divided into sub clusters A2 (a) and (b). A2 (a) contained genotype RST9 and A2 (b) was further divided into sub clusters A2(c) and (d). A2(c) contained genotype PA255 and A2 (d) was further subdivided into sub clusters A2 (e) and (f). A2 (e) contained genotypes RG8 and F846. A2 (f) was further subdivided into sub clusters A2 (g) and (h). A2(g) was further divided into sub clusters having genotypes H777, RST9, H1226, HD123, H1117, LH1556, F1861, HS6, RS921, F505, H1098, CISA310 and RS2013. A2 (h) contained the genotype LD694. Varieties (RS2013, H1117 and HD123) were found to be genetically similar and could not be distinguished from each other. Varieties (H 1226 and RST 9, H777) showed a greater genetic similarity. The varieties F846, RG8, HD, PA255 and RST9 showed slight genetic diversity among them and fell in nearby sub clusters. The cultivars LD 327 and F 1378 showed genetic similarity among them, but were found to be diverse from other cultivars of this sub cluster group. LD 694 was also genetically different from other cultivars of the sub cluster. Varieties F 846 and RG 8 showed genetic similarity among themselves. PA 255 showed genetic differences with other cultivars of cluster A1.

Table 1. Primer sequence, the total number of amplicons, monomorphic amplicons, polymorphic amplicons, and the percentage of polymorphism as revealed by RAPD analysis.

Primers	Primer sequences	Total amplicons	Monomorphic amplicons	Polymorphic amplicons	Polymorphism (%)
RAPD 2	ACGTAGCGTC	1	1	0	0
RAPD 4	AAGTCCGCTC	1	1	0	0
RAPD 5	CCCAGTCACT	5	3	2	40
RAPD 6	CCACGGGAAG	5	2	3	60
RAPD 7	CAGGACTGAC	5	2	3	60
RAPD 9	TCCCACGCAA	3	1	2	66.6
RAPD 10	TCAGAGCGCC	5	1	4	80
RAPD 13	GTCAGAGTCC	4	3	1	25
RAPD 15	TGGCGTCCTT	3	1	2	66.6
RAPD 16	TCGGCGGTTC	2	1	1‘	50
	Total	**34**	**16**	**18**	
	Average	3.4	1.6	1.8	52.9

Polymorphism as detected by SSR analysis : Ten SSR primers were employed to investigate the genetic polymorphism as shown in Table 2, the number of alleles/locus ranged from 1 to 5, while the number of polymorphic alleles varied from 0 to 4 and the average level of polymorphism was 22.7 per cent. Liu *et al.*, (2005) investigated the genetic diversity of the Asian cotton *Gossypium arboreum L.*, widely cultivated in China, by microsatellite analysis. They found that of the 358 microsatellite markers analyzed, 74 primer pairs detected 165 polymorphic DNA fragments among 39 *G. arboreum* accessions examined. Chen and Du (2006) studied the genetic diversity among 43 sources of upland cotton germplasm with

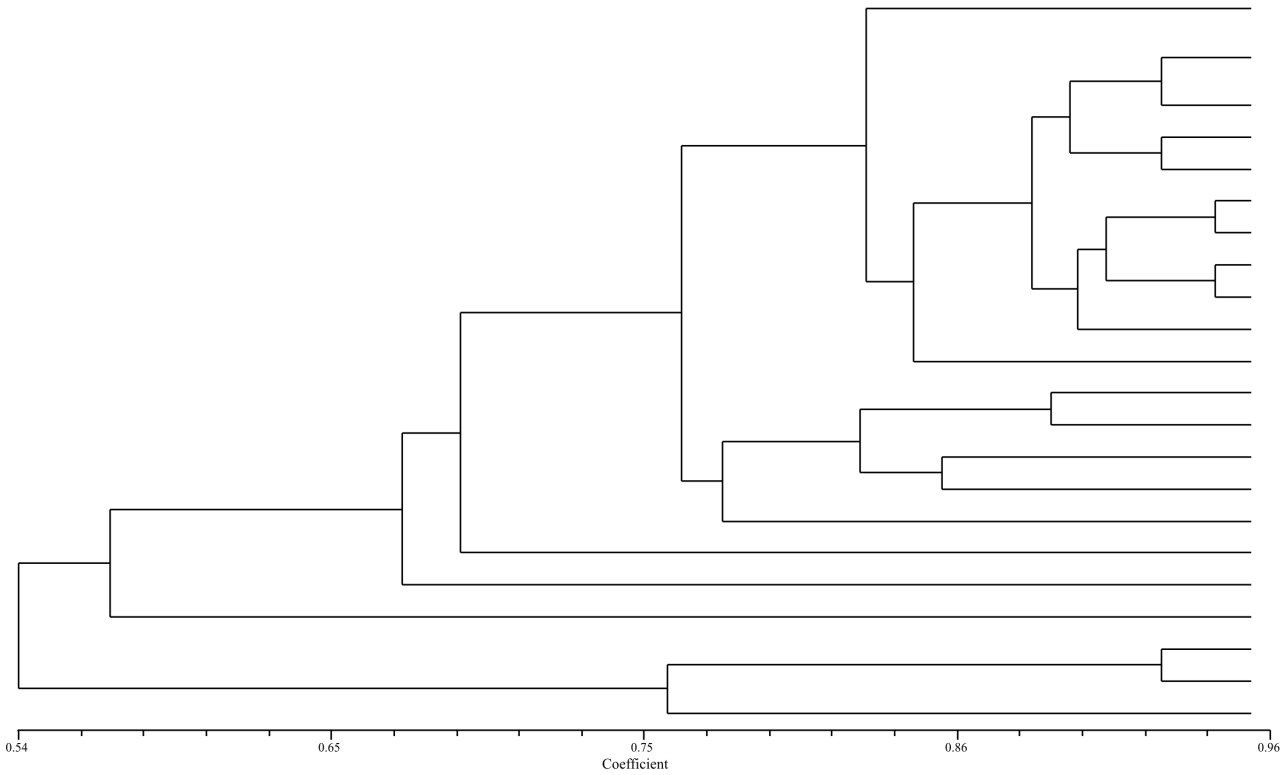

Fig. 1. Dendrogram of 23 cotton cultivars based on 10 RAPD and 10 SSR primers

different parental origins, breeding periods, and ecological growing areas in China on the basis of simple sequence repeat (SSR) markers. A total of 130 alleles with 80 per cent polymorphism were detected from 36 SSR primers. The number of alleles per primer ranged from two to eight with an average of 3.6. Fig. 1 shows the dendrogram chart which indicates the genetic similarity among 23 cotton cultivars, the range of dendrogram being 0.57 to 1.00. All cultivars fell into two major clusters (A1 and A2) which were further divided into sub clusters. Cluster A1 comprised of two major sub clusters (A1 (a) and (b). In the sub cluster A1 (a) contained the genotypes RST 9, H777, F1378 and sub cluster A1 (b) contained the genotypes RS180, H1226, LH900 and RS875. Cluster A2 comprised of two major sub clusters A2 (a) and (b). A2 (a) contained genotypes F505 and RST9. A2 (b) was further divided into sub clusters A2(c) and (d). A2(c) contained genotypes PA255 and RG8. A2 (d) was further divided into sub clusters A2 (e) and (f). A2(e) contained genotype F1861 and sub cluster A2(f) was further divided into sub clusters having genotypes LH1556, H1117, HS6, H1098, LD327, RS2013, RS921, F846, CISA310, HD123 and LD694. Cultivar (H1098 and LD327) and (F846, CISA 310, HD123) were found to be genetically similar and could not be distinguished from each other. The genotypes RS921, RS2013, HS6, RG8 and PA255 showed greater genetic similarity. RST9 and F505 showed genetic differences with another cultivar of cluster A2. In sub cluster A1 the varieties (RST9, H777 and F1378), (RS180 and H1226) showed greater genetic similarity, while other cultivars (LH900 and RS875). Cultivar LD694 showed greater genetic differences with other cultivars.

Polymorphism as detected by RAPD and SSR markers : Fig. 1 showed dendrogram chart which showed the similarity coefficient between 0.57 and 1.00. Dendrogram revealed two distinct clusters A and B. 20 varieties out of total 23 fell in cluster A. Cluster A could further be divided into sub clusters (A1 and A2). A1 contained genotype RST9. Cluster A2 was further divided into subclusters A2 (a) and (b). A2 (b) contained genotype RG 8 and A2 (a) was further divided into sub clusters A2(c) and (d). Genotypes PA255 fell in A2(c). Sub clusters A2 (d) was further bifurcated into A2 (e) and (f). A2 (f) contained genotypes F1378, F505, H777, RST9 and H1226. A2 (e) was further divided in to sub clusters A2

Table 2. Primer sequences, the total number of amplicons, monomorphic amplicons, polymorphic amplicons and percentage of polymorphism as revealed by SSR analysis

Primer	Primer sequence	Total amplicons	Mono amplicons	Poly amplicons	Polymorphism (%)
SSR-BNL252	F–TGAAGAGCTCGTTGTTGCAC				
	R-CGAAAGAGACAAGCAATGCA	**2**	2	0	0
SSR-BNL1053	F-AGGGTCTGTCATGGTTGGAG				
	R- CATGCATGCGTACGTGTGTA	**2**	1	1	50
SSR-BNL1721	F-TGTCGGAATCGGAAGACCGG				
	R- GCGCAGATCCTCTTACCAAA	**2**	2	0	0
SSR-BNL2572	F-GTCCTATTACTAAAATTGTTAAC				
	R- CGATGTTAAATCAATCAGGTCA	**2**	2	0	0
SSR-MGHES6	F-TCGCTTGACTTTCCATTTCC				
	R-AACCCTCGGGATTATCGTCT	**1**	1	0	0
SSR-MGHES13	F-CAGGGGAGCCATTGTTAGAA				
	R-CAGGGGTCCTGTGTTTCAGT	**2**	2	0	0
SSR-MGHES32	F-CGTCGCTTCCTTTGCTTAAC				
	R-GTCGGGTTAATTGCAAATCG	**1**	1	0	0
SSR-MGHES14	F-GAGGAGGCTGTGGTTGAAGA				
	R-ATGGTGACCCTGCTTACACC	**4**	4	0	0
SSR-MGHES16	F-ACCCCAATACAACCCCATTT				
	R-GCAGAGAAAAGGGACAGAGG	**1**	1	0	0
SSR-MGHES35	F-TCGAACGGCTCGTTAAATCT				
	R- CAGCAAAGAGTGGTTCTCTGG	**5**	1	4	80

(g) and (h). A2 (h) contained only one genotype LD694. A2(g) was further divided into sub clusters having genotypes F846, F1861, HS6, RS921, CISA310, HD123, LH1556, H1117, LD327, H1098 and RS2013. Varieties RS 2013 and H1098 were found to be genetically similar and could not be distinguished from each other. Varieties (H1117 and LH1556), (HD123 and CISA310), (RS921 and HS6), (H1226 and RST 9) and (H777 and F505) showed greater genetic similarity, the varieties LD 327, F1861,F846 and F1378 showed slight genetic diversity among themselves. LD694 was genetically different from other cultivars of the sub cluster A2 (g) and (h). RG8 showed genetic differences with other cultivars of cluster A. Cluster B comprised of 2 sub clusters (B1 and B2). Two cultivars RS 875 and LH900 of sub cluster B1 exhibited genetic similarity with each other and were genetically different from cultivars RS 810 which fell in sub cluster B2.

CONCLUSION

As a result of the analysis of the banding profile generated by different primers, it was seen that amplification products generally ranged from 200-400 bp in SSR primers and 500-1000 bp in RAPD primers. It was observed that a large number of high intensity bands were generally polymorphic in nature. The banding profile obtained in 23 cotton cultivars suggested that the SSR primer MGHES 14 identified maximum number (23) of alleles, and among the RAPD primers, the primer RAPD 16 identified a maximum number of 24 alleles. The polymorphism percentage generated by these primers was 100 per cent. The cotton fingerprint generated has further revealed a variety of specific primers. The RAPD primer 6 was specific for PA 255, RAPD 7 specific for RST 9, RAPD 9 specific for RST 9, RAPD 10 for LH 900, and RAPD 15 for HD 123. Similarly, BNL primer 1053 & MGHES 14 were specific for the variety RST 9, and MGHES 35 was specific for RS 810.

REFERENCES

Chen, G. and Du, X.M. 2006. Genetic diversity of source germplasm of Upland cotton in China as determined by SSR marker analysis. *Yi Chuan Xue Bao,* **33** : 733-45

Gostimsky, S.A., Kokaeva, Z.G. and Konovalov, F.A. 2005. Studying plant genome variation using molecular markers. *Russian J. Genet.* **41** : 378-88

Guo, W.Z., Sang, Z.Q., Zhou, B.L. and Zhang, T.Z. 2007. Genetic relationships of D genome species based on two types of EST-SSR markers derived from *G. arboretum* and *G. raimondii* in *Gossypium. J. Plant Sci.*, **172** : 808-14

Hussein, Ebtissam, H.A., Amina, A.M., Soltan A.M. and Adawy. 2006. Molecular characterization and genetic relationship among cotton genotypes 1- RAPD, ISSR, and SSR analysis. *Arab J. Biotech.,* **9**: 222-29.

Liu, D. Guo, Z., Lin, Y. and Nie, X. 2005. Genetic diversity of Asian cotton (*Gossypium arboreum L.*) in China evaluated by microsatellite analysis. *Gene. Resou. Crop Evol.,* ***53*** : 1145-52.

Rana, M.K., Singh, V.P. and Bhat, K.V. 2005. Assessment of genetic diversity in Upland Cotton breeding lines by using amplified fragment length polymorphism (AFLP) markers and morphological characterstics. *Gene. Resou. Crop Evol.,* **52** : 989-97

Recieved for publication : December 3, 2013
Accepted for publication : July 17, 2014