Theor Appl Genet (2000) 101:931–938 © Springer-Verlag 2000

S.A. Khan · D. Hussain · E. Askari · J.McD Stewart
K.A. Malik · Y. Zafar

Molecular phylogeny of *Gossypium* species by DNA fingerprinting

Received: 2 January 2000 / Accepted: 12 February 2000

Abstract Total genomic DNA from 31 available *Gossypium* species, three subspecies and one interspecific hybrid, were analysed to evaluate genetic diversity by RAPD, using 45 random decamer primers. A total of 579 amplified bands were observed, with 12.9 bands per primer, of which 99.8% were polymorphic. OPJ-17 produced the maximum number of fragments while the minimum number of fragments was produced with primer OPA-08. Cluster analysis by the unweighted paired group method of arithmetic means (UPGMA) showed six main clusters. Cluster 'A' consisted of two species and one subspecies of the A-genome, with a 0.78–0.92 Nei's similarity range. Cluster B, composed of all available tetraploid species and one interspecific hybrid, showed the same sister cluster. Nei's similarity ranged from 0.69 to 0.84. The B-genome formed the UPGMA sister cluster to the E-genome species. Cluster 'C' consisted of five *Gossypium* species of which three belong to the B-genome, with Nei's similarity values of 0.81 to 0.86. Although there was considerable disagreement at lower infra-generic ranks, particularly among the D-genome (diploid New World species) and C-genome (diploid Australian species) species. The sole F-genome species *Gossypium longicalyx* was resolved as a sister group to the D-genome species. *Gossypium herbaceum* and *G. herbaceum* Africanum showed the maximum Nei's similarity (0.93). Minimum similarity (0.29) was observed between *Gossypium trilobum* and *Gossypium nelsonii*. The average similarity among all studied species was 50%. The analysis revealed that the interspecific genetic relationship of several species is related to their centre of origin. As expected, most of the species have a wide genetic base range. The results also revealed the genetic relationships of the species *Gossypium hirsutum* to standard cultivated *Gossypium barbadense*, *G. herbaceum* and *Gossypium arboreum.* These results correspond well with previous reported results. The level of variation detected in closely related genotypes by RAPD analysis indicates that it may be a more efficient marker than morphological marker, isozyme and RFLP technology for the construction of genetic linkage maps.

Key words *Gossypium* species · RAPD · Phylogeny · Cluster · Diversity

Communicated by H.F. Linskens

S.A. Khan · D. Hussain · E. Askari · K.A. Malik · Y. Zafar (✉)
National Institute for Biotechnology and
Genetic Engineering (NIBGE), PO Box 577, Jhang Road,
Faisalabad, Pakistan
e-mail: y-zafar@yahoo.com
Fax: +92-41-651472

J. McD Stewart
Department of Agronomy, University of Arkansas, Arkansas,
Fayetteville AR 72701, USA

Introduction

The genus, *Gossypium* comprises about 50 species of trees, shrubs and herbs (Fryxell 1971). The study of the genetics and evaluation of *Gossypium* diversity is important for improvement in the existing gene pool. Cotton is the most important textile fibre crop and is the second most important oil seed crop in the world (Cherry and Leffler 1984). Eight cytogenetically defined diploid genome groups (A through G and K) have been recognized (Stewart 1995). Wild cotton species has proved to be a valuable source of the genetic material-wide spectrum of a biological and economical index that could be useful for the modelling of varieties corresponding to modern requirements and conditions.

Previously morphological markers, with their complex and undeciphered genetic control, were used for individual identification. Morphological markers may be affected by environmental effects and growth practices. In contrast to the morphological markers, proteins and DNA-based data are available from genetically based descriptors. Apart from identification they can also be used in tests of parentage, in genetic mapping and in the measurement of genetic diversity (Chapparo et al. 1994; Iqbal et al. 1997). The environment, however, can inter-

fere with the quality and quantity of proteins in seed (Higgins 1984). Proteins in the seed also show interaction with other compounds which are found in roots and leaves. Protein is also growth-stage and tissue specific. The roots, stem and leaves of cotton plants are very rich in polyphenolic products (Cherry and Leffler 1984). These phenolic compounds interact with proteins in various ways so that electrophoretic analysis of isozymes becomes difficult. As a result, genetic research based on isozymes has been relatively difficult in polyphenolic-rich plants such as cotton.

Within the last few years restriction fragment length polymorphism (RFLP) technology has been applied to several cotton species to study evolution, population genetics, phylogenetic relationships and genome mapping (Shappley et al. 1996; Yu et al. 1997), but it creates low variation in cotton compared to other plant taxa (Brubaker and Wendel 1994).

Random amplified polymorphic DNA (RAPD) markers (William et al. 1990) have been successfully used for cultivar analysis and species identification in most plants, due to the technical simplicity and speed of RAPD methodology (Gepts 1993). These include papaya (*Carica papaya* L*)* (Stiles et al. 1993), rye *(Secale cereale* L.) (Iqbal and Rayburn 1994), 14 Australian cotton cultivars (Multani and Lyon 1995), 54 genotypes of lentil (Sharma et al. 1995), 16 elite US cotton varieties (Tatineni et al. 1996), 23 elite commercial cotton varieties (Iqbal et al. 1997) and *Phaseolus vulgaris* (Yonghe Bai et al. 1998).

Tetraploid species of cotton *Gossypium hirsutum* L. occupies more than 90% of the cotton growning area in Pakistan. All available cotton cultivars in Pakistan are the intraspecific crosses of *G. hirsutum* L. evaluated with a narrow genetic base (Iqbal et al. 1997). Therefore, we were interested to explore the application of RAPD marker technology to fingerprinting potential for the genotypic discrimination of a number of *Gossypium* species. The objectives of this current work were, therefore: (1) to analyse and quantify the genetic diversity in cultivated and wild *Gossypium* species using RAPDs, and (2) to check the validity of already available data about the genus with those based on alternative markers.

Materials and methods

Plant materials

Thirty one species, three subspecies and one interspecific hybrid of cotton (*Gossypium*) were collected from different sources (Table 1).

Table 1 List of the *Gossypium* species used in the present study to determine their genetic phylogeny

No.	Species name	Genome	Genome size (2n)	Source
1	*G. herbaceum*	A_1	26	CCRI, Multan[a]
2	*G. herbaceum*	Africanum A1	26	CCRI, Multan[a]
3	*G. arboreum*	A_2	26	NIBGE[b]
4	*G. anomalum*	B_1	26	CCRI, Multan
5	*G. barbosanum*	B_3	26	CCRI, Multan
6	*G. capitis*-viridis	B_4	26	UOA, USA[c]
7	*G. sturtianum*	C_1	26	UOA, USA[c]
8	*G. nandewarence*	C_1	26	CCRI, Multan
9	*G. robinsonii*	C_2	26	CCRI, Multan
10	*G. australe*	C_3	26	UOA, USA
11	*G. thurberii*	D_1	26	CCRI, Multan
12	*G. herknesii*	D_{2-2}	26	CCRI, Multan
13	*G. davidsonii*	$D_{3\text{-}d}$	26	CCRI, Multan
14	*G. aridum*	D_4	26	CCRI, Multan
15	*G. raimondii*	D_5	26	CCRI, Multan
16	*G. lobatum*	D_7	26	CCRI, Multan
17	*G. trilobum*	D_8	26	CCRI, Multan
18	*G. laxum*	D_9	26	CCRI, Multan
19	*G. turnerii*	D_{10}	26	UOA, USA
20	*G. schwendimanii*	D_{11}	26	UOA, USA
21	*G. stocksii*	E_1	26	CCRI, Multan
22	*G. incanum*	E_4	26	CCRI, Multan
23	*G. longicalyx*	F_1	26	CCRI, Multan
24	*G. bickii*	G_1	26	CCRI, Multan
25	*G. nelsonii*	G_3	26	CCRI, Multan
26	*G. rotundifolium* 499788K		26	UOA, USA
27	*G. pilosum* 499786K		26	UOA, USA
28	*G. hirsutum*	(AD_1)	52	NIBGE
29	*G. hirsutum* Yucatenanse	(AD)	52	UOA, USA
30	*G. barbadense*	$(AD)_2$	52	UOA, USA
31	*G. barbadense* Mummy	(AD)	52	UOA, USA
32	*G. tomentosum*	$(AD)_3$	52	CCRI, Multan
33	*G. mustilinum*	(AD)	52	CCRI, Multan
34	*G. lanceolatum*	$(AD)_?$	52	CCRI, Multan
35	*G. arboreum*×*G. herknesii*	$A_2 \times D_{2-2}$	52	Breeding Line[d]

[a] Central Cotton Research Institute, Multan, Pakistan
[b] National Institute for Biotechnology and Genetic engineering, Faisalabad Pakistan
[c] University of Arkansas, USA
[d] Breeding line provided by Professor James McD. Stewart, USA

Total genomic DNA isolation

Total genomic DNA was isolated with the CTAB method described by Doyle and Doyle (1987) and treated with RNase to eliminate RNA. DNA concentration was measured by DyNAQuant 200 and the DNA was diluted in distilled water to a concentration of 12.5 ng/μl for PCR reactions.

Polymerase chain reaction and primers

PCR-amplification reactions were performed exactly as described by Iqbal et al. (1997). Out of the 45 primers used, 20 were of the M series, 17 of kit J series and seven of kit A series purchased from Operon Technologies Inc., Alameda, California, USA.

Analysis of amplification profiles

Amplification profiles of 35 Genotypes were compared with each other and bands of DNA fragments were scored as present (1) or absent (0). The data for all 45 primers was used to estimate the similarity on the basis of the number of shared amplification products (Nei and Li 1979). A dendrogram based on similarity coefficients was generated by using the unweighted pair group of arithmetic means (UPGMA).

Results and discussion

DNA of 34 *Gossypium* species and one interspecific hybrid of *Gossypium* was amplified with 45 different random primers purchased from Operon Technologies. All 35 genotypes with the 45 primers revealed a unique banding pattern and so can be used for species identification. This might be indicative of a wide genetic base of the cotton species studied. Different primers produced a different level of polymorphism among the different species (Fig. 1a, b, c).

A total of 579 DNA fragments were amplified, with an average of 12.9 RAPD markers per primer. Out of 579 DNA-amplified fragments, only one (0.2%) was found to be monomorphic (less than 0.5 kb). The remaining 578 (99.8%) were polymorphic. The approximate size of the largest fragment produced was 3.5 kb and the smallest easily recognisable fragment produced was approximately 0.4 kb. Out of 35 genotypes studied, *Gossypium herbaceum* (A) produced the maximum number of DNA-amplified fragments (289), while *Gossypium lobatum* (D7) produced 119 bands, which is the minimum number. Other species produced between 151 and 283 bands in general. *G. hirsutum* (A1) and *G. hirsutum* Africanum (A1) produced 264 bands which are the maximum common bands, while the minimum common bands (41) were produced between *Gossypium trilobum* (D8) and *Gossypium nelsonii* (G3). Other species produced between 51 to 230 common DNA-amplified fragments in general. The pair-wise genetic distance estimates of the 35 genotypes in this study were analysed and are given in Table 2.

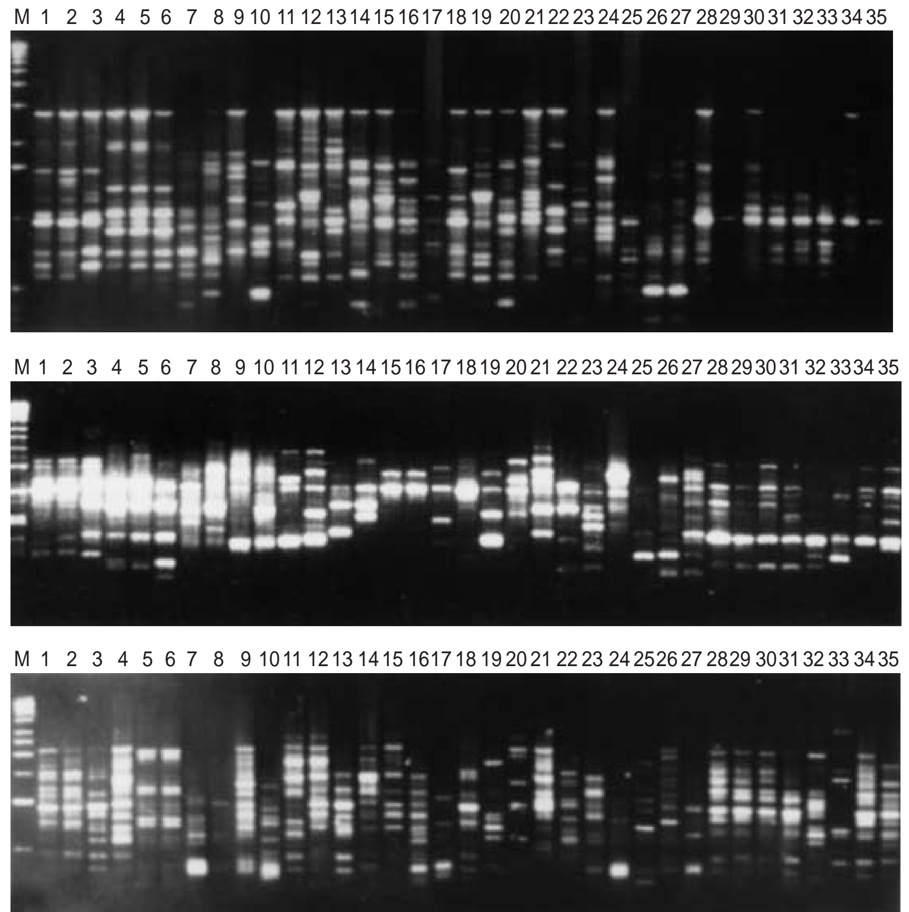

Fig. 1 Amplification profile of 35 *Gossypium* genotypes with primers (*a*) OPM-16 (*b*) OPM-18 and (*c*) OPA-13. *Lane M*=Size marker. *Lane 1=G. herbaceum. Lane 2=G. herbaceum* Africanum. *Lane 3=G. arboreum. Lane 4=G. anomalum. Lane 5=G. barbosanum. Lane 6=G. capitis* varidis. *Lane 7=G. sturtianum. Lane 8=G. nandewarence. Lane 9=G. robinsonii. Lane 10=G. australe. Lane 11=G. thurberii. Lane 12=G. herknesii. Lane 13=G. davidsonii. Lane 14=G. aridum. Lane 15=G. ramondii. Lane 16=G. lobatum. Lane 17=G. trilobum. Lane 18=G. laxum. Lane 19=G. turnerii. Lane 20=G. schwendimanii. Lane 21=G. stocksii. Lane 22=G. incanum. Lane 23=G. longicalyx. Lane 24=G. bickii. Lane 25=G. nelsonii. Lane 26=G. rotundifolium. Lane 27=G. pilosum. Lane 28=G. hirsutum. Lane 29=G. hirsutum* Yuca. *Lane 30=G. barbadense. Lane 31=G. barbadense* Mum. *Lane 32=G. tomentosum. Lane 33=G. mustilinum. Lane 34=G. lanceolatum. Lane 35*=interspecific hybrid

Table 2 Similarity matrix for Nei's and Li's coefficients of 35 *Gossypium* genotypes obtailned from RAPD markers

Number[a]	Species name	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1	*G. herbaceum*	1													
2	*G. herbaceum* Africanum	0.923	1												
3	*G. arboreum*	0.709	0.708	1											
4	*G. anomalum*	0.609	0.606	0.565	1										
5	*G. barbosanum*	0.558	0.555	0.514	0.864	1									
6	*G. capitis*-viridis	0.548	0.542	0.549	0.818	0.816	1								
7	*G. sturtianum*	0.446	0.453	0.431	0.465	0.487	0.553	1							
8	*G. nandewarence*	0.510	0.492	0.475	0.516	0.488	0.543	0.690	1						
9	*G. robinsonii*	0.560	0.553	0.523	0.600	0.572	0.581	0.548	0.589	1					
10	*G. australe*	0.397	0.400	0.474	0.500	0.515	0.531	0.532	0.512	0.530	1				
11	*G. thurberii*	0.500	0.522	0.508	0.498	0.464	0.484	0.413	0.451	0.540	0.465	1			
12	*G. harknesii*	0.527	0.536	0.529	0.552	0.524	0.530	0.437	0.455	0.555	0.471	0.659	1		
13	*G. davidsonii*	0.474	0.486	0.481	0.446	0.443	0.427	0.432	0.463	0.491	0.456	0.546	0.562	1	
14	*G. aridum*	0.505	0.499	0.474	0.499	0.470	0.472	0.408	0.457	0.519	0.450	0.610	0.598	0.586	1
15	*G. raimondii*	0.460	0.462	0.413	0.483	0.484	0.477	0.364	0.395	0.544	0.434	0.555	0.593	0.518	0.587
16	*G. lobatum*	0.387	0.395	0.405	0.376	0.397	0.430	0.395	0.378	0.430	0.456	0.505	0.510	0.514	0.574
17	*G. trilobum*	0.330	0.323	0.301	0.316	0.333	0.327	0.423	0.347	0.389	0.387	0.310	0.308	0.383	0.321
18	*G. laxum*	0.560	0.545	0.524	0.473	0.455	0.453	0.421	0.450	0.473	0.428	0.448	0.461	0.477	0.494
19	*G. turnerii*	0.450	0.457	0.454	0.489	0.517	0.511	0.488	0.500	0.527	0.494	0.531	0.639	0.504	0.533
20	*G. schwendimanii*	0.447	0.449	0.450	0.478	0.501	0.512	0.453	0.471	0.468	0.459	0.513	0.508	0.467	0.555
21	*G. stocksii*	0.528	0.529	0.492	0.592	0.572	0.531	0.400	0.473	0.509	0.387	0.485	0.520	0.515	0.541
22	*G. incanum*	0.525	0.534	0.481	0.714	0.683	0.633	0.443	0.451	0.549	0.454	0.458	0.560	0.531	0.496
23	*G. longicalyx*	0.428	0.417	0.435	0.407	0.436	0.474	0.400	0.405	0.434	0.408	0.404	0.426	0.460	0.398
24	*G. bickii*	0.493	0.466	0.460	0.487	0.509	0.503	0.505	0.501	0.487	0.470	0.474	0.476	0.483	0.470
25	*G. nelsonii*	0.381	0.380	0.373	0.375	0.390	0.413	0.387	0.393	0.346	0.423	0.370	0.365	0.338	0.354
26	*G. rotundifolium* 499788	0.478	0.460	0.453	0.480	0.489	0.483	0.523	0.473	0.475	0.488	0.442	0.436	0.371	0.447
27	*G. pilosum* 499786	0.473	0.467	0.452	0.500	0.493	0.491	0.496	0.467	0.460	0.533	0.406	0.413	0.386	0.446
28	*G. hirsutum*	0.603	0.622	0.604	0.579	0.540	0.534	0.452	0.472	0.541	0.458	0.516	0.555	0.519	0.501
29	*G. hirsutum* Yucatenanse	0.605	0.613	0.590	0.561	0.519	0.545	0.460	0.489	0.503	0.457	0.523	0.542	0.502	0.488
30	*G. barbadense*	0.604	0.616	0.606	0.569	0.524	0.529	0.467	0.521	0.516	0.459	0.506	0.518	0.503	0.485
31	*G. barbadense* Mummy	0.603	0.604	0.604	0.564	0.514	0.548	0.484	0.542	0.482	0.480	0.483	0.488	0.492	0.475
32	*G. tomentosum*	0.598	0.606	0.611	0.542	0.522	0.540	0.496	0.502	0.506	0.497	0.491	0.508	0.501	0.483
33	*G. mustilinum*	0.489	0.483	0.502	0.470	0.442	0.477	0.432	0.426	0.446	0.471	0.445	0.429	0.456	0.425
34	*G. lanceolatum*	0.468	0.645	0.597	0.589	0.554	0.532	0.473	0.493	0.547	0.474	0.521	0.533	0.496	0.514
35	*G. arboreum*×*G. harknesii*	0.623	0.627	0.601	0.572	0.548	0.517	0.448	0.482	0.549	0.472	0.502	0.526	0.494	0.507

[a] Numbers on the top of the table represent the same *Gossypium* species as listed in the first vertical column

The similarity matrix is based upon Nei and Li's similarity coefficient. The genetic distances (Nei's similarity) ranged from 0.29 to 0.92. Maximum similarity was observed between *G. herbaceum* and *G. herbaceum* Africanum (0.92). *G. trilobum* (D8), a diploid cotton, in general showed a minimum degree of similarity with all other species ranging between 0.29 to 0.42.

Cluster analysis using RAPD resulted in six main cluster groups. The results in the present study are congruent with both the cytogenetic and geographic groupings as well as maternal cladistics relationships (Wendel and Albert 1992). In cluster 'A' out of three Genotypes *G. herbaceum* and *G. herbaceum* Africanum clustered together indicating that they are more closely related as compared to any other species. Their banding patterns were also more similar to each other as compared to other profiles. They have a high estimate of Nei's (1978) genetic identity (0.923). *Gossypium arboreum* showed 0.788 Nei's identity to *G. herbaceum*. Wendel et al. (1989) reported that *G. arboreum* and *G. herbaceum* are fixed, or nearly fixed, for alternate alleles at eight loci, leading to a low estimate of Nei's genetic identity (0.74). These two species comprise the A genome, based on chromosome pairing relationships (Beasley 1942). In this study they consist of one cluster, showing same results as Beasley (1942) and Endrizzi et al. (1985) and the same, or a near, Nei's similarity between *G. herbaceum* and *G. arboreum* also verify the conclusion of Wendel (1989).

One of the striking aspects of the cpDNA (chloroplast DNA) phylogeny is the sister grouping of the A-genome diploids and those of the allopolyploid species (Wendel and Albert 1992). In the present study, allopolyploids were grouped together in one cluster and showed more clustering similarity to species belonging to the A-genome. In cluster 'B' an the average Nei's similarity of 0.80 in allopolyploids was found by cluster analysis. Genetic similarity between allopolyploids indicates a very narrow genetic base. Despite the strong homology exhibited by many of the allopolyploids, this study demonstrated that it is now possible to differentiate very closely related species such as *Gossypium lanceolatum* and interspecific hybrid used in the present study.

In the present study the AD-genome (allopolyploid) clustered closer to the A-genome (diploid) as shown in the Fig. 2. The AD-genome combines two diploid A- and D-genomes. The A-genome has a larger chromosome size and more repetitive DNA sequences than the

15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1																				
0.485	1																			
0.316	0.422	1																		
0.428	0.431	0.398	1																	
0.535	0.528	0.400	0.498	1																
0.459	0.536	0.400	0.458	0.639	1															
0.463	0.407	0.313	0.471	0.503	0.537	1														
0.483	0.444	0.358	0.467	0.523	0.521	0.610	1													
0.384	0.508	0.398	0.348	0.430	0.459	0.458	0.464	1												
0.410	0.450	0.396	0.484	0.513	0.525	0.527	0.533	0.505	1											
0.355	0.375	0.290	0.394	0.428	0.445	0.344	0.389	0.431	0.435	1										
0.411	0.429	0.340	0.466	0.481	0.492	0.493	0.503	0.428	0.532	0.476	1									
0.358	0.394	0.383	0.414	0.485	0.531	0.473	0.497	0.404	0.521	0.447	0.745	1								
0.480	0.392	0.347	0.531	0.517	0.505	0.573	0.570	0.420	0.496	0.441	0.516	0.527	1							
0.452	0.376	0.349	0.511	0.492	0.484	0.514	0.528	0.436	0.506	0.429	0.518	0.539	0.848	1						
0.459	0.420	0.333	0.539	0.462	0.485	0.531	0.537	0.439	0.490	0.461	0.518	0.539	0.788	0.828	1					
0.431	0.434	0.353	0.542	0.501	0.488	0.514	0.523	0.413	0.510	0.474	0.517	0.529	0.725	0.778	0.796	1				
0.441	0.420	0.374	0.529	0.504	0.514	0.496	0.510	0.419	0.514	0.452	0.540	0.557	0.707	0.748	0.743	0.839	1			
0.392	0.403	0.354	0.433	0.464	0.459	0.404	0.454	0.390	0.455	0.446	0.462	0.461	0.567	0.591	0.554	0.660	0.697	1		
0.470	0.407	0.322	0.501	0.446	0.469	0.517	0.540	0.416	0.523	0.404	0.521	0.506	0.727	0.732	0.728	0.690	0.707	0.621	1	
0.450	0.381	0.317	0.476	0.449	0.481	0.523	0.533	0.419	0.512	0.399	0.501	0.486	0.726	0.723	0.719	0.678	0.711	0.617	0.884	1

D-genome (Geever 1980). Larger the chromosome size and more the repetitive sequences, greater the chances for 10-mer random primers to find homology, and to give more and different-sized DNA-amplified fragments. In this study we produced more and different sizes of RAPDs in the A-genome compared to other genomes. This cluster was followed by the AD-genome which has derived one set of chromosomess from A-diploid species, which is why the A-genome has a more UPGMA clustering similarity to the AD-genome.

The relationships portrayed by these allopolyploids were also in accordance with the available pedigree information. *G. hirsutum* and *G. hirsutum* Yucatenanse showed sister clustering, with a Nei's similarity of 0.848. *G. hirsutum* Yucatenase is a subspecies of *G. hirsutum*. Similarly, the *Gossypium barbadense* and *G. barbadense* Mummy gave sister clustering, showing a similarity of 0.839. *G. barbadense* Mummy is a subspecies of the cultivated *G. barbadense*; their differences being due to some spontaneous mutations. Spontaneous alternation of chromosomes are very common in the genome of *Gossypium* (Barrow and Dunford 1974), especially in the allopolyploids and have been occurring since the origin of the allotetraploids (Endrizzi et al. 1985). Wendel et al. (1994) used allozyme analysis to assess the levels and patterns of genetic diversity in *Gossypium mustilinum* and its relationships to other tetraploid species. Genetic variation was low and genetic-identity estimates between *G. mustilinum* and other polyploid species ranged from 0.471 to 0.591. These loosely equivalent and relatively low identity estimates quantify the isolated nature of the species. In the present RAPD analysis *G. mustilinum* occupied a basal phylogenetic position among the tetraploid *Gossypium*. The similarity matrix (Table 2) isolated *G. mustilinum* from the rest of the polyploids with Nei's similarity ranging from 0.554 to 0.697. Chloroplast DNA and ribosomal DNA restriction-site data also suggested a relatively basal position for *G. mustilinum* (Dejoode and Wendel 1992). Wendel et al. (1994) estimated the genetic diversity between *G. mustilinum* and other tetraploid species and found it to be lower than for other comparisons among the polyploids.

Gossypium tomentosum is morphologically distinct from other allopolyploids and its phylogenetic relationships with them is uncertain. *G. tomentosum* is distinguished from other amphiploid members of the genus on the basis of several morphological characters, including its sulphur-yellow corolla (Dejoode and Wendel 1992).

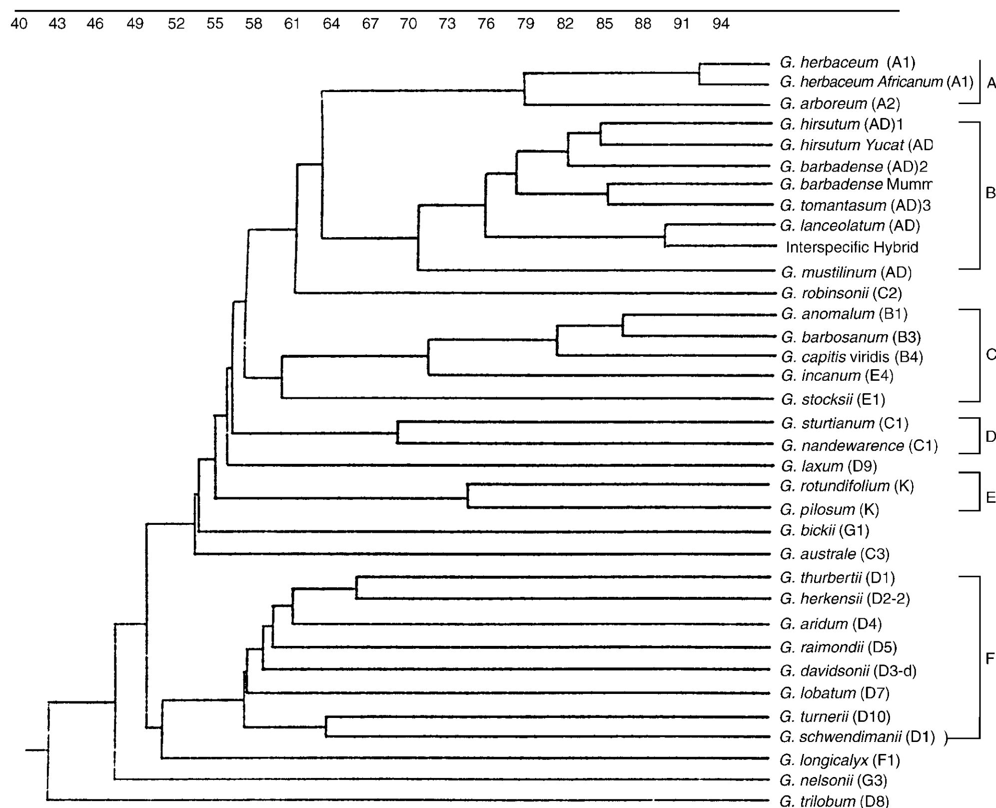

Fig. 2 Dendrogram of 31 *Gossypium* species, three subspecies and one interspecific hybrid of *Gossypium* generated by RAPD data using the UPGMA method

Fryxell (1979) noted a similarity in the leaf indentation of *G. tomentosum* and *G. mustilinum.* The flavonoid data of Parks et al. (1975) suggested that *G. tomentosum* is most similar to *G. barbadense* and to *Gossypium darwinii*. Similarly, Hansen-Kampf and Menzel (1980) reported little or no cytogenetic differentiation between *G. tomentosum* and *G. hirsutum*. Nei's (1978) estimated interspecific genetic identities are as follows: (*G. tomentosum–G. hirsutum* = 0.82); (*G. tomentosum–G. barbadense* = 0.65). In the present study, *G. tomentosum* showed Nei's similarity with other amphiploids ranging from 0.69 to 0.83. *G. tomentosum* forms a sister UPGMA cluster with *G. barbadense* Mummy. The highest interspecific genetic identity for *G. tomentosum* is with *G. barbadense* Mummy (0.83), whereas the genetic identity of *G. tomentosum* with *G. mustilinum* was considerably lower (0.69). *G. barbadense* formed a monophyletic assemblage with *G. hirsutum* and *G. hirsutum* Ycatenanse. *G. barbadense* has 0.78 Nei's similarity with *G. hirsutum*. Nei (1978) also estimated an interspecific genetic identity of 0.77 between *G. barbadense* and *G. hirsutum*. Each of these species appears to have been derived independently from different wild progenitors (Lee 1984; Wendel et al. 1989). *G. lanceolatum* is closely related to the interspecific hybrid ($A_2 \times D_{2-2}$).

The main UPGMA cluster 'C' comprised of five *Gossypium* species. Out of these three species belonged to the B-genome and the other two to the E-genome. B-genome species grouped in a sister cluster to the E-genome species. In the four Wagner trees (Wendel and Albert 1992), one resolved the B-genome clade as the sister group to the New World (D-genome) diploid species, a second showed it as a sister group to a clade consisting of the A-D-E and F genome groups, while a third placed it as part of an unresolved trichotomy. A fourth showed it as a sister cluster to a clade contained the D-genome species. Studies of flavonoids (Parks et al. 1975), seed proteins (Cherry et al. 1970; Johnson and Thein 1970) and phenetic relationships (Fryxell 1979; Valicek 1978) offer little or no convincing evidence for any of these alternative resolutions. Cytogenetic data are similarly equivocal (Endrizzi et al. 1985). Other evidence, however, strongly implicates a close alliance of the B-genome species, including foliar flavonoid data (Phillips 1966; Parks et al. 1975). The cpDNA data (Wendel and Albert 1992) demonstrated a remarkable similarity among the three B-genome species. In the present study the B-genome species (*Gossypium anomalum, Gossypium barbosanum* and *Gossypium capitis* viridis) demonstrated a remarkable Nei's similarity of 0.81 to 0.86 and showed a UPGMA sister group to a clade consisting of the A-AD *Gossypium robinsonii* and the E-genome group (Fig. 2).

The sole F-genome species *Gossypium longicalyx* was resolved as a sister group to the D-genome diploids. The Nei's similarity with other *Gossypium* species ranged from 0.34 to 0.50. *G. longicalyx* showed maximum similarity with *Gossypium bickii* and minimum similarity with *Gossypium laxum.* Wendel and Albert (1992) resolved the sole F-genome species *G. longicalyx* as a sister group to either the clade containing the

AD-genome or the clade comprising the E-genome species. Several other authors have commented previously on the relatively isolated position of *G. longicalyx* (Saunders 1961; Phillips and Strickland 1966; Valicek 1978; Fryxell 1979; Vollessen 1987), which is adapted to more mesic conditions than any other diploid *Gossypium* species. Phillips and Strickland (1966) showed that *G. longicalyx* merited its own genomic designation, and consider their cytogenetic data to imply that *G. longicalyx* is closer to the A-genome species than the E-genome species. The opposite conclusion was reached by Johnson and Thein (1970), based on comparative seed protein analysis, and by Valicek (1978) who conducted a phenetic analysis of morphology. Fryxell's (1971) phenetic analysis showed *G. longicalyx* to be about equidistant from both clades, as do the foliar flavonoid data of Parks et al. (1975). Our result suggested the placement of *G. longicalyx* as a sister cluster to the D-genome diploid species.

The cluster 'F' in Fig. 2 comprised of the D-genome species. Out of 35 available *Gossypium* genotypes, ten diploid species belonged to the D-genome, of which eight species are grouped together in the same sister UPGMA cluster. *G. laxum* showed more diversity and formed a sister cluster with the C-genome species. *G. trilobum* showed the minimum similarity (Nei's 0.29 to *G. nelsonii*) among all the species and formed a sister group to a clade consisting of the *G. nelsonii* and the D-genome groups. In the cpDNA data of Wendel and Albert (1992) cladistics relationships within the D-genome are incompletely resolved due to an absence of phylogenetic-informative cpDNA variation. Approximately half of the D-genome species (*Gossypium thurberii through Gossypium raimondii*) are included in an unresolved polychotomy, while species in the D-genome have been differentiated by previous authors (Brown and Menzel 1952; Phillips 1966; Johnson and Thein 1970; Parks et al. 1975; Valicek 1978; Fryxell 1979). In respect of cp DNA data, *G. thurberii* and *G. trilobum* showed a similarity. In the present RAPD analysis study, however, they showed a wide range of diversity from each other.

A similar lack of correspondence between the RAPD analysis and previous taxonomic concepts was apparent in the Australian cottons. The present study includes eight Australian species, designated as C-, G-, and K-genome species. From the section Grandicalyx two species were included in the present study, *Gossypium rotundifolium* and *Gossypium pilosum*. In this study these two species formed a sister group. From the section Sturtia of the four species *Gossypium sturtianum*, *Gossypium nandewarence Gossypium robinsonii* and *Gossypium australe, G. sturtianum* and *G. nandewarence* group together into a sister cluster, while *G. robinsonii* and *G. australe* showed more similarity with the AD-genome and G_1- and K-genome species. From the section Hibiscoidea two species *Gossypium bickii* and *Gossypium nelsonii* were analysed in the present study. *G. bickii* showed more similarity with other Australian species and forms a sister UPGMA cluster with them. *G. nelsonii*, showed itself as sister group to a clade consisting of D-, F- and C-group species. This means that *G. nelsonii* is more divergent than any of the other C-genome species. *G. nelsonii* also has a different morphology, ecology and geographic distribution (Fryxell 1965; Fryxell et al. 1992).

In this study, the average genetic similarity of all the *Gossypium* species analysed was found to be 50.0%. Similarly, the species studied indicated a wide range in their genetic base. Despite the strong homology exhibited by many of the cotton genotypes, this study has demonstrated that it is now possible to differentiate very closely related subspecies such as *G. herbaceum* and *G. herbaceum* Africanum by using molecular markers. Furthermore, most of the species can be individually characterized with species-specific RAPD markers which indicates the utility of this approach for fingerprinting purposes. The results obtained in the present study provide a good correspondence to the recognized taxonomic divisions within the genus *Gossypium.* These results suggest that the RAPD technique produces reliable results with which to construct the phylogenetic history of the genus *Gossypium.*

Acknowledgements This research was partly supported by the Cotton Biotechnology Project (Research and Extension) Government of the Punjab. The authors gratefully acknowledge the supply of the wild species from Dr. Zahoor Ahmad, Director, Central Cotton Research Institute, Multan.

References

Barrow JR, Dunford MP (1974) Somatic crossing as a cause of chromosome multivalents in cotton. J Hered 65:3–7

Beasley JO (1942) Meiotic chromosome behaviour in species, species hybrids, haploids and induced polyploids of *Gossypium*. Genetics 27:25–54

Brown ME, Menzel MY (1952) The cytology and crossing behaviour of *G. gossypioides*. Bull Torrey Bot Club 79:285–292

Brubaker CL, Wendel JF (1994) Re-evaluating the origin of domesticated cotton (*G. hirsutum* L., Malvaceae) using nuclear restriction fragment length polymorphism (RFLP). Am J Bot 81:1309–1326

Chapparo JX, Werner DJ, O'Malley D, Sederoff RR (1994) Targetted mapping and linkage analysis of morphological, isozyme and RAPD markers in peach. Theor Appl Genet 87: 805–815

Cherry JP, Leffler HR (1984) Seed. In.: Kohel RJ, Lewis CF (eds) Cotton. ASA, Madison, Wisconsin, pp 511–569

Cherry JP, Katterman FRH, Endrizzi JE (1970) Comparative studies of seed proteins of species of *Gossypium* by gel electrophoresis. Evolution 24:431–447

Dejoode DR, Wendel JF (1992) Genetic diversity and origin of Hawaiian Island cotton, *G. tomentosum*. Am J Bot 79:1311–1319

Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochem Bull 19:11–15

Endrizzi. JF, Turcotte EL, Kohell RJ (1985) Genetics, cytology and evolution of *Gossypium*. Adv Genet 23:271–375

Fryxell PA (1965) Stages in the evolution of *Gossypium.* Adv Frontiers Plant Sci 10:31–56

Fryxell PA (1971) Phenetic analysis and the phylogeny of the diploid species of *Gossypium* L. (Malvaceae). Evolution 25:554–562

Fryxell PA (1979) The natural history of the cotton tribe. College Station: Texas A & M University Press, USA
Fryxell PA, Craven LA, Stewart J McD (1992) A revision of *Gossypium* sect. Grandicalyx (Malvaceae), including the description of six new species. Systematic Bot 17:91–114
Geever RF (1980) The evolution of single-copy nucleotide sequences in the genome of *G. hirsutum*. PhD dissertation, University of Arizona, Tuscon, USA
Gepts P (1993) The use of molecular and biochemical markers in crop evaluation studies. In: Hecht MK (ed) Evolutionary biology, vol 27. Plenum Press, New York, pp 51–94
Hasen-Kampf CA, Menzel MY (1980) Incipient genome differentiation in *Gossypium.* Comparison of 12 chromosome in *G. hirsutum* and *G. tomentosum*, using different heterozygous translocation. Genetics 95:971–983
Higgins TJV (1984) Synthesis and regulation of major protein in seeds. Annu Rev Plant Physiol 35:191–221
Iqbal MJ, Rayburn AL (1994) Stability of RAPD markers for determining cultivar-specific DNA profiles in rye *(Secale cereale* L.). Euphytica 75:215–220
Iqbal MJ, Aziz N, Saeed NA, ZafarY, Malik KA (1997) Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. Theor Appl Genet 95:139–144
Johnson BL, Thein MM (1970) Assessment of evolutionary affinities in *Gossypium* by protein electrophoresis. Am J Bot 57: 1081–1092
Lee JA (1984) Cotton as a world crop. In: Kohel RJ, Lewis CF (eds) Cotton. American Society of Agronomy, Madison, pp 6–25
Multani DS, Lyon BR (1995) Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38:1005–1008
Nei M (1978) Estimation of average hetero zygosity and genetic distance from a small number of individuals. Genetics 89: 583–590
Nei M, Li W (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273
Parks CR, William De, Dreyer DL (1975) The application of flavonoid distribution to taxonomic problems in the genus *Gossypium.* Bull Torrey Bot Club 102:350–361
Phillips LL (1966) The cytology and phylogenetic of the diploid species of *Gossypium.* Am J Bot 53:328–335
Phillips LL, Strickland MA (1966) The cytology of a hybrid between *G. hirsutum* and *G. longicalyx*. Can J Genet Cytol 8:91–95
Saunders HE (1961) The wild species of *Gossypium* and their evolutionary history. Oxford University Press, UK
Sharma SK, Dawson IK, Waugh R (1995) Relationship among cultivated and wild lentils revealed by RAPD analysis. Theor Appl Genet 91:647–654
Shappley Zw, Jenkins JN, Watson Jr CE, Khaler AL, Meredith WR (1996) Establishment of molecular markers and linkage groups in two F_2 populations of upland cotton. Theor Appl Genet 92:915–919
Stewart J McD (1995) Potential for crop improvement with exotic germplasm and genetic engineering. In: Constable GA, Forrester NW (eds) Challenging the future. Proceedings of the World Cotton Research Conference-1, CSIRO, Melbourne, pp 313–327
Stiles JI, Lemme C, Sondur S, Morshidi MB, Maishardt R (1993) Using random amplified polymorphic DNA for evaluating genetic relationships among papaya cultivars. Theor Appl Genet 85:697–701
Tatineni V, Cantrell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characters and RAPDs. Crop Sci 36:186–192
Valicek P (1978) Wild and cultivated cottons. Coton et Fibres Tropicales (English edition) 33:363–387
Vollessen K (1987) The native species of *Gossypium* (Malvaceae) in Africa, Arabia and Pakistan. Kew Bull 42:337–349
Wendel J F (1989) New world cotton contains Old World cytoplasm. Proc Natl Acad Sci USA 86:4132–4136
Wendel JF, Albert VA (1992) Phylogenetic of the cotton genus (*Gossypium*): character-state weighted parsimony analysis of chloroplast DNA restriction site data and its systematic and biogeographic amplifications. Systematic Bot 17:115–143
Wendel JF, Olson PD, Stewart (1989) Genetic variation, introgression and independent domestication of Old World cultivated cottons. Am J Bot 76:1795–1806
Wendel JF, Rowley R, Stewart J McD (1994) Genetic diversity in and phylogenetic relationships of the Brazilian endemic cotton, *G. mustilinum*. Plant Syst Evol 192:49–59
William JGK, Kubenk AR, Livak KJ, Rafalski JA, Tingey SV (1990) DNA polymorphism amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Res 18:6531–6535
Yonghe Bai, Michaels TE, Pauls KP (1998) Determination of genetic relationships among *Phaseolus vulgaris* population in a conical cross from RAPD marker analysis. Mol Breed 4: 395–406
Yu ZH, Park YH, Lazo GR, Kohel RJ (1997) Molecular mapping of the cotton genome. Agron Abstracts, ASA, Madison, Wisconsin, p 147