A comparative analysis of genetic distances among 24 upland cotton genotypes using RAPD markers and phenotypic characters

Terezinha Aparecida Teixeira-Cabral[1], Julio Cesar Viglioni Penna*[2] And Luiz Ricardo Goulart[3]
[1]Former Grad. Student, Institute of Genetics And Biochemistry, Univ. Fed. Uberlândia, P.O.Box 593, Uberlândia, MG, Brazil; [2]Professor, Institute of Agronomic Sciences, Univ. Fed. Uberlândia, P.O.Box 593, Uberlândia, MG, Brazil; [3]Professor, Institute Of Genetics And Biochemistry, Univ. Fed. Uberlândia, P.O. Box 593, Uberlândia, MG, Brazil. (* Corresponding Author. E-mail: penna@umuarama.ufu.br)

ABSTRACT

The knowledge of genetic distances among individuals or populations in plant breeding programs can be used to complement phenotypic information and to help the selection of individuals to breed towards segregating populations with high variability. The objectives of this study were to identify primers that produce polymorphisms among cotton (*Gossypium hirsutum* L.) genotypes, to compare the effect of primer length in the generation of polymorphisms, to determine genetic distances among 24 cotton genotypes from eight countries using the RAPD technique and compare them with results obtained through phenetic analysis of the genotypes based on 52 qualitative and quantitative characteristics using numerical taxonomy. The 24 genotypes were evaluated by the RAPD technique utilizing short primers and combinations of long primers. RAPDs were scored as presence (1) or absence (0) of prominent bands reproducible in duplicated PCR reactions. These data were subjected to analyses to produce a dissimilarity matrix based on Simple, Nei and Li's and Jaccard's Distances. UPGMA was utilized to create dendrograms based on the genetic distance matrices. Short primers produced polymorphic bands that were efficient to discriminate among genotypes, while long primers couldn't detect differences among the genotypes. Comparing the three methods used for estimation of genetic distances, the method that better dealt with distances obtained by using phenotypic character analyses was the Simple Distance Method. The correlation among the results of this method with Jaccard's and Nei and Li's Distances were 93 % and 91 %, respectively. The maximum simple distance found between pairs of genotypes based in the polymorphisms produced by short primers was 18%, which possibly indicates a narrower genetic basis than in the phenotypic analysis. The correlation among phenetic analysis and RAPD based on short primers was moderate and negative, tough significant (r=-0.43).

KEY WORDS: Cotton, genetic distances, *Gossypium hirsutum*, RAPD.

INTRODUCTION

The genus *Gossypium* comprises 50 diploid and tetraploid species indigenous to Australia, Central and South Americas, Africa, Asia and the Galapagos, and the Hawaiian Archipelagos (Fryxell, 1992). Upland cotton, *Gossypium hirsutum* L., accounts for about 90% of world cotton lint production and has a long history of improvement through breeding (Sparks, 1995). In plant breeding programs, the knowledge of the genetic distances among entries in germplasm collections - individuals or populations - can be used to complement phenotypic information for a wise choice of plants to breed in order to achieve genetic variable populations, as well as indicate more efficient breeding methods to be used (Thormann and Osborn, 1992; Skroch et al., 1992). For the determination of phenotypic distances, qualitative morphological characters and quantitative characters may be used, and through them, genotypic distances are inferred. However, this procedure is limited since quantitative traits are influenced by environmental effects and may also lack adequate generation of polymorphisms. Many of the complications of a phenotype-based comparison could be mitigated through direct genotype identification by using a DNA-based diagnostic assay (Beckman and Soller, 1983). In fact, molecular genetic markers are powerful tools to analyze genetic relationships and genetic diversity. Among them, random amplified polymorphic DNA's (RAPD's) have been employed in genetic research due to their speed and simplicity (Williams et al., 1990; Welsh and McClelland, 1990).

Several crops have been the object of genetic distance analyses by means of molecular markers. In *Arachis*, for example, a variability study based on RFLP and

RAPD markers demonstrated low variability levels among American cultivars of *A. hypogaea* L., but a great variability among species within the genus, which suggests that interspecific hybridizations could increase variability (Halward et al., 1991).

A high level of variation in fragment patterns was observed among and within species of *Hordeum* when 10-base primers were used for random DNA amplifications (González and Ferrer, 1993). Genetic similarities among eight rye cultivars estimated through RFLP and RAPD analyses showed good agreement with the known genealogy (Loarce et al., 1996). Elite cotton pure line genotypes derived from interspecific hybridization and, as well as typical *Gossypium hirsutum* and *G. barbadense* entries, were used in the generation of genetic distance dendrograms on RAPD and taxonomic distances (morphologic data). Classification of genotypes based on both methods yielded similar results, with a correlation coefficient of 0.63 between genetic and taxonomic distances (Tatineni et al., 1996).

This research aimed at identifying primers in order to generate polymorphism's among a group of 24 cotton genotypes from eight countries; determining genetic distances among them using RAPD markers; and comparing distances obtained by means of a RAPD analysis and by a phenetic analysis of qualitative and quantitative characters.

MATERIAL AND METHODS

Genotypes: Seeds of the 24 *Gossypium hirsutum* L. genotypes utilized in this study (Table 1) were donated by "Empresa de Pesquisa Agropecuária de Minas Gerais" (EPAMIG) and "Empresa Brasileira de Pesquisa Agropecuária" (EMBRAPA). The entries do not represent a random sample of cultivars grown in the eight countries of origin, nor such countries represent a random sample of cotton-growing nations. Sowing was performed in three-liter pots in a glasshouse located in the campus of the University of Uberlândia in 1996. The substrate used was a mixture of 1/3 peat moss and 2/3 soil. Fertilization was done according to the official recommendation (Comissão de Fertilidade..., 1989) after soil routine chemical analysis. Five seeds of each entry were sowed in each pot and three plants out of five were sampled for DNA extraction. This was done by harvesting three leaves (one per plant, in bulk) 10 days after emergence of the first foliar primordia.

DNA extraction: 300 mg of the harvested material were macerated in liquid nitrogen for immediate DNA extraction according to the protocol proposed by Doyle and Doyle (1990), modified by the addition of polyvinilpirrolidone (PVP-40). The macerate was transferred to 2 mL eppendorf tubes and added to 600 µl extraction buffer (2% CTAB, 1,4 M NaCl, 20 mM EDTA, 100 mM Tris-HCl pH 8.0, 1% PVP-40 and 0.2% ß mercaptoethanol). Tubes were then incubated at 65°C for 30 minutes. After incubation, an equal volume of chloroform was added: isoamyl alcohol (24:1) with five-min centrifugation at 12,000 rpm. Proteins and cellular debris were precipitated and supernatant DNA was transferred to another tube. This sequence was repeated once in order to increase DNA quality. DNA was precipitated with isopropanol at -20°C and the pellet recovered was rinsed with 70% ethanol. After drying, DNA was resuspended in TE buffer (10 mM Tris-HCl, 0.01 mM EDTA, pH 8.0) with 10 µg/ml of RNase and incubated at 37°C for 30 min. Nucleic acid concentration and purity were determined spectrophotometrically at 260 nm (DNA reading) and 280 nm (protein reading) and ran in 1.2% agarose gel for quality confirmation. For amplification, DNA was diluted in pure water to a final concentration of 10 ng/µl.

Amplification: A preliminary trial was ran in order to evaluate five different programs in a thermocycler PTC-100-96 MJ Research: **(1)** 5 min at 94° C, 45 cycles [1 min at 94° C, 2 min at 37° C, 2 min at 72° C] and termination for 7 min at 72° C; **(2)** 5 min at 94° C, 10 cycles [40 seconds at 94° C, 90 seconds at 30° C, 90 seconds at 72° C] and 30 cycles [40 seconds at 94° C, 60 seconds at 55° C, 70 seconds at 72° C]; **(3)** 5 min at 94° C, 45 cycles [1 min at 94° C, 2 min at 45° C, 2 min at 72° C] and termination 7 min at 72° C; **(4)** 5 min at 94° C, 45 cycles [15 seconds at 94° C, 15 seconds at 36° C, 30 seconds at 72° C] and termination 2 min at 72°C) and **(5)** similar to the previous condition with an additional minute in the extension time (72°C). Reagent concentrations were as follows: 30 ng DNA of cultivar Acala SJ-5, 1 U Taq polymerase, 0.200 mM of each dNTP, 2 mM $MgCl_2$, 0.2 µM of primer (OP L12), with final volume of 25 µl completed with pure water. These conditions comply with the ones used by Tatineni et al. (1996), except for $MgCl_2$ concentration.

Another trial was ran to optimize the reaction, by using the primer L12 (Operon Technologies) as follows: three primer concentrations (0.24; 0.32, and 0.4 µM), four Mg^{+2} concentrations (1.125; 1.500; 2.000, and 2.250 mM), three Taq DNA polymerase concentrations (0.75; 1.00 and 1.50 U), five dNTP concentrations (0.05; 0.10; 0.20; 0.40, and 0.80 mM of each), and two DNA concentrations (25 and 50 ng

per reaction). The final volume was 25 µl completed with pure water. Program (5) described under Amplification was utilized in the thermocycler.

Primers: Twenty-four 10-base primers (Operon Techn.) were evaluated: A01 through A15, L04, L11, L13, L15, O02, O05, O06, O08, and O18). DNA's were extracted from cultivars Acala SJ-5 and IAC-RM4-SM5. Each reaction contained: 25 ng DNA, 1 U Taq DNA polymerase, 2 mM $MgCl_2$, 0.05 mM of each dNTP, 0.32 µM primer, completed with pure water to 25 µl. Program 5 was utilized in the thermocycler. Eleven out of the 25 primers evaluated (Table 2) were utilized to amplify DNA's of the 24 genotypes, following the same procedures described. The reaction was performed twice for each cultivar. Three combinations of primers developed at the Molecular Genetics Lab. of the Universidade Federal de Uberlândia were also evaluated: GOU01/GOU07, GOU03/GOU10 and GOU8/GOU9 (Table 3). Reaction conditions were the same as the ones used for the short primers, using the first thermocycler program (1) described under Amplification.

Visualization: The amplification products were separated electrophoretically on 6% polyacrylamide gels at 100 V for four hours, stained with ethidium bromide (5 µg/ml), and viewed and photographed under UV light (VDS ImageMaster – Pharmacia Biotech). RAPD's were registered as presence vs. absence of bands. Only intense polymorphisms were considered.

Table 1. List of the 24 cotton genotypes utilized and their respective countries of origin.

Identification Number	Genotypes	Country of Origin (Abbreviation)
01	IAC-13-1	BRAZIL (BR)
02	IAC-RM4-SM5	BRAZIL (BR)
03	Minas Dona Beja	BRAZIL (BR)
04	SU 0450/8909	BRAZIL (BR)
05	4959	BULGARIA (BU)
06	6396	BULGARIA (BU)
07	4F	PAKISTAN (PK)
08	Del Cerro	PERU (PE)
09	SK 14	THAILAND (TH)
10	CA(68)41	UGANDA (UG)
11	Acala SJ-5	USA (US)
12	Coker 5110	USA (US)
13	Deltapine Land 16	USA (US)
14	Lankart LX 571	USA (US)
15	Paymaster 303	USA (US)
16	Stoneville 213	USA (US)
17	Westburn M	USA(US)
18	149-F	ex-USSR (UR)
19	153-F	ex-USSR (UR)
20	2421	ex-USSR (UR)
21	C-4727	ex-USSR (UR)
22	Tashkent-1	ex-USSR (UR)
23	Tashkent-2	ex-USSR (UR)
24	Tashkent-3	ex-USSR (UR)

Table 2. List of the eleven short primers utilized in the amplification of DNA's of the 24 genotypes and their respective nucleotide sequences.

Primer	Sequence
OPA-01	5'-CAGGCCCTTC-3'
OPA-02	5'-TGCCGAGCTG-3'
OPA-03	5'-AGTCAGCCAC-3'
OPA-09	5'-GGGTAACGCC-3'
OPA-10	5'-GTGATCGCAG-3'
OPA-15	5'-TTCCGAACCC-3'
OPL-04	5'-GACTGCACAC-3'
OPL-11	5'-ACGATGAGCC-3'
OPL-15	5'-AAGAGAGGGG-3'
OPO-02	5'-ACGTAGCGTC-3'
OPO-06	5'-CCACGGGAAG-3'

All lab work was undertaken at the Molecular Genetics Laboratory of the Universidade Federal de Uberlândia.

Data analyses: A matrix was generated according to the presence (1) and absence (0) of bands, and was used to estimate the following coefficient of similarity: 1) Simple matching (Sokal and Michener, 1958) using the equation $GS_{ij} = (a+d)/(a+b+c+d)$; 2) Jaccard's coefficient of similarity (Jaccard, 1901) using the equation $GS_{ij} = a/(a+b+c)$ and 3) Nei and Li's coefficient of similarity (Nei and Li, 1979) using the equation $GS_{ij} = 2a/(2a+b+c)$, where $\mathbf{GS_{ij}}$ is the genetic similarity between the *i* and *j* genotypes, **a** is the number of bands present in both *i* and *j*; **b** is the number of bands present in *i* and absent in *j*; **c** is the number of bands present in *j* and absent in *i*, **d** is the number of bands that are in both genotypes. Conversion to genetic distance (GD) was done using the equation $GD_{ij} = 1 - GS_{ij}$. The UPGMA (unweighted pair-group method using arithmetic averages) was used to group genotypes and construct dendrograms. Simple correlation coefficients (r) between GD were also estimated.

RESULTS AND DISCUSSION

The addition of PVP-40 was of great importance to obtain high-quality DNA. Several attempts to extract DNA without it had failed, resulting in low amounts of poor-quality DNA. Cotton cells contain high rates

Table 3. List of long primers utilized in the amplification of DNA's from the 24 genotypes and their respective nucleotide sequences.

Primer	Sequence
GOU-01	5'-CTTTAATTCCATATGCCTAAGCGGG-3'
GOU-03	5'-GCATCTGCAGGCCCACATYGTCTTYCCNGT-3'
GOU-07	5'-AACGCGGCACACATAGTTTT-3'
GOU-08	5'-CGAAAGTCTTAGTAACATAT-3'
GOU-09	5'-TAAATAAAGGTTACACTGGT-3'
GOU-10	5'-CTCGAACCTTCCCATTTCAA-3'

OBS: Y = C,T; N = A,C,G,T

of phenolic therpenoids and tannins, which bound to the DNA after wall lysis and can not be removed by conventional methods. Soluble PVP complexes with polyphenols avoiding such binding (John, 1992).

When short primers were used, both slow (**1**) and fast (**5**) programs were efficient in amplifying DNA. Two hours were saved per reaction with the fast program. However, when long primers were used, short programs did not work well, specially during the annealing stage, which required slower stages (**1**).

As for the reaction components, concentrations which most revealed consistent bands were: 0.32 µM primer; 2 mM $MgCl_2$; 0.05 mM of each dNTP; 1 U Taq DNA and 25 ng DNA. Reactions with 0.4 and 0.8 mM of dNTP's did not show amplified fragments, but concentrations of 0.05 mM and 0.10 mM did not show different results. Eleven out of the 24 ten-base long primers tested generated a great number of highly defined bands and so were chosen for DNA amplification. They generated 68 bands, averaging 6.18 bands per primer. Band sizes were estimated to be in the range of 300 to 2600 bp and 30 of them were polymorphic (44.12% polymorphism) (Table 4). In Figure 1 a typical band pattern obtained with primer OPO-06 for 12 genotypes is presented as an example.

Table 4. List of short primers utilized, number and estimated size of bands obtained and number of polymorphic bands.

Primers	Bands		
	Band number	Size (bp)	Number of Polymorphic bands
OPA-01	3	1550-2000	1
OPA-02	5	800-1600	2
OPA-03	5	520-1400	3
OPA-09	5	600-1100	1
OPA-10	8	480-1070	3
OPA-15	5	580-1500	2
OPL-04	8	450-930	4
OPL-11	8	620-1950	4
OPL-15	3	300-630	1
OPO-02	5	600-1150	3
OPO-06	13	450-2600	6

The genetic distance matrix, based on band presence versus absence, and calculated by using Simple Distances (SD), is not shown in this paper. Distances ranged from 1% (between Coker 5110 and Lankart LX 571) to 18% (between 153-F and IAC-13-1 or 2421 and IAC-13-1, among other pairs). Nei and Li Distances (matrix also not shown) varied from 1% (Coker 5110 and Lankart LX 571) to 12% (2421 and IAC-13-1, Minas Dona Beja and 4959, Deltapine 16 and 153-F). A high correlation (r = 0.95) was found between those two matrices.

The cluster analysis based on SD utilizing UPGMA defined two groups at an arbitrary level of 12% (Figure 2): group A with 23 genotypes and group B with one (C-4727). At a level of 10%, group A was subdivided into six subgroups: A1 with IAC-13-1, 4959 and Tashkent 1; A2 with 6396, Paymaster 303, SK 14 and Stoneville 213; A3 with IAC-RM4-SM5, 149-F, 153-F, Minas D. Beja, 2421, SU 0450/8909, Del Cerro, CA(68)41, Coker 5110, Lankart LX 571, Acala-SJ-5 and 4F; A4 with Deltapine 16; A5 with Westburn and Tashkent 2 and A6 with Tashkent 3 alone. C-4727 was the most divergent genotype. Pairs with distances smaller than 4% were: Coker 5110 and Lankart LX 571; 6396 and Paymaster 303; SU 0450/8909 and Del Cerro. Such results indicate a narrow

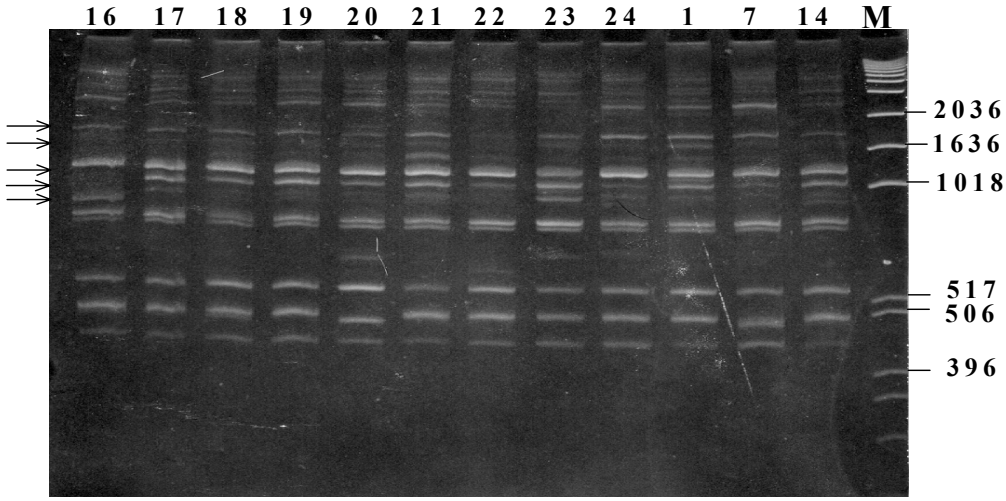

Figure 1. **RAPD patterns obtained with primer OPO-06 for 12 genotypes.** Sequence of genotypes from left to right, follows the identification numbers on Table1; M: DNA marker MW of 1 kb "Ladder". Arrows indicate polymorphisms.

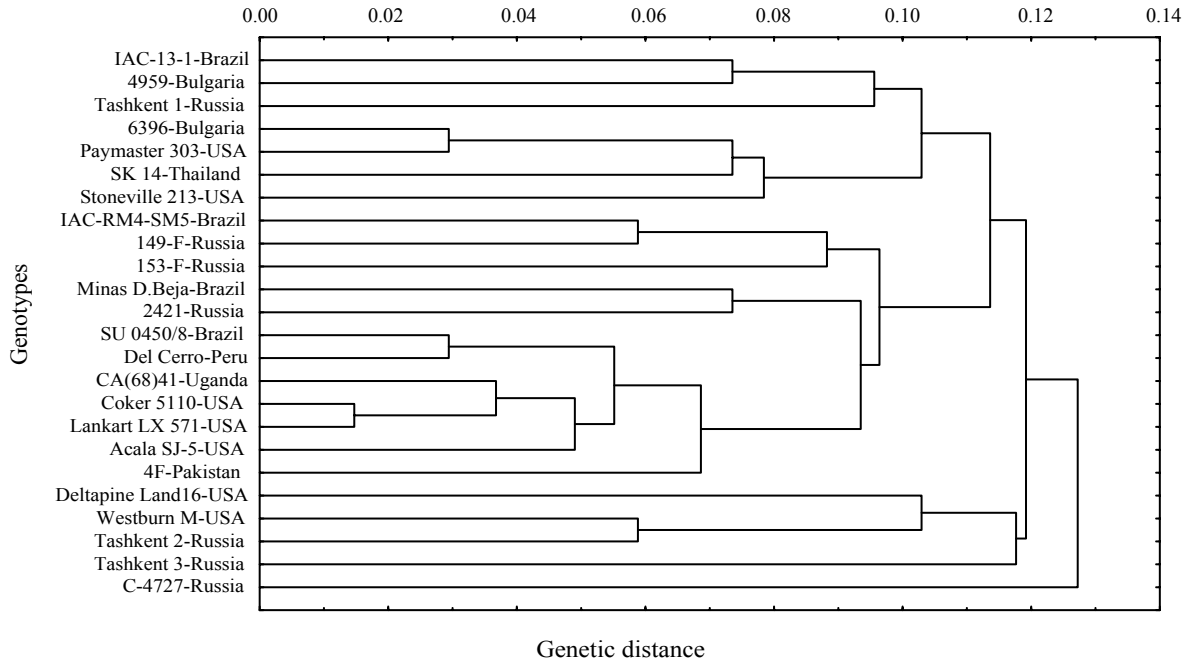

Figure 2. Dendrogram representing genetic distances estimated among 24 cotton genotypes from eight countries, based on 68 RAPD markers generated by short primers.

genetic variation amongst such entries.

The three combinations of long primers tested (Table 5) generated 14 bands ranging from 150 to 520-bp long of which, five (35.71%) were polymorphic. Figure 3 demonstrates the band pattern obtained with the combination of primers GOU01/GOU07 for eight genotypes as example.

Dissimilarities estimated with the long primers presented values ranging from 0 to 29% (matrix not shown). The cluster analysis (Figure 4) showed a few groups of genotypes. At a 19% genetic distance level, cultivars were divided into two groups: A with 22 genotypes and B with two (2421 and C-4727). At the 11 % level, group A was subdivided into two components: A1 with five cultivars (IAC-13-1, Lankart LX 571, 4F, Paymaster 303 and Tashkent 2) and A2 with 17 (IAC-RM4-SM5, Minas D. Beja, Del Cerro, CA(68)41, Acala SJ-5, Coker 5110, Deltapine 16, Stoneville 213, Westburn M, 153-F, 4959, SK 14, SU 0450/8909, 6396, 149-F, Tashkent 1 and Tashkent 3). Group B was also divided into two subgroups:

Table 5. List of long primers, number and sizes of bands obtained, and number of polymorphic bands.

Combination of Primers	Bands		Number of Polymorphic bands
	Total number	Size (bp)	
GOU 01/07 [1/]	6	160-520	2
GOU 03/10	4	150-360	2
GOU 08/09	4	160-480	1

[1/] GOU 01 and GOU 07.

B1 with the cultivar 2421 and B2 with C-4727. Although these long primers could also separate clusters, they did not differentiate among genotypes as efficiently as did the short primers.

When data obtained with short and long primers were pooled together, a total of 82 bands were obtained (an average of 5.86 bands per primer, with 42.7% polymorphism). Simple Distances between genotypes (matrix not shown) varied from 2% (Acala SJ-5 and Coker 5110; Coker 5110 and Lankart LX 571) to 20% (IAC-13-1 and 2421). Nei and Li Distances (matrix also not shown) ranged from 2% (Del Cerro and SU 0450/8909, CA(68)41 and Coker 5110, Acala SJ-5 and Coker 5110, among other pairs) to 13% (IAC-13-1 and 2421). Jaccard Distances resulted in genetic distances (matrix not shown) varying from 3% (Acala SJ-5 and Coker 5110; Coker 5110 and Lankart LX 571) to 22% (IAC-13-1 and 2421). The results obtained with SD's were closely related to Jaccard Distances and to Nei and Li Distances, as indicated by the correlation coefficients of 0.96 and 0.95, respectively, among SD´s and the other two. The dendrogram based on SD's (Figure 5), at the 14% genetic distance divided cultivars into two groups: A with 23 genotypes and B, with one (C-4727). Considering a level of 11%, group A was subdivided into four groups: A1 with three cultivars (IAC-13-1, 4959 and SK 14); A2 with 17 (SU 0450/8909, Del Cerro, CA(68)41, Lankart LX 571, Acala SJ-5, Coker 5110, 4F, 149-F, 153-F, 2421, 6396, Paymaster 303, Stoneville 213, Tashkent 1, Deltapine 16, Westburn M and Tashkent 2); A3 with Tashkent 3; and A4 with two (IAC-RM4-SM5 and Minas D. Beja). The pairs Acala SJ-5 and Coker 5110; SU 0450/8909 and Del Cerro; 6396 and Paymaster 303 had dissimilarity values smaller than 4%.

The SD dendrogram based on data calculated from 68 bands generated through short primers on the five South American genotypes demonstrated that the most divergent genotype was IAC-13-1 and the most closely related were SU 0450/8909 and Del Cerro (Figure 6). In the case of South American entries, for example, there was a fairly good ship relation between the RAPD analysis and genealogy: Brazilian cultivar IAC-13-1 is an Acala-derived genotype; IAC-RM4-SM5 is a selection on the cultivar Auburn 56; Minas Dona Beja originated from the cross Auburn 56 vs. DPL 11. The sequence in which those genotypes clustered (Figure 6) seemed related with their genealogy.

When only 19 genotypes after exclusion of the South American lines were considered, there was no apparent separation of groups according to entry

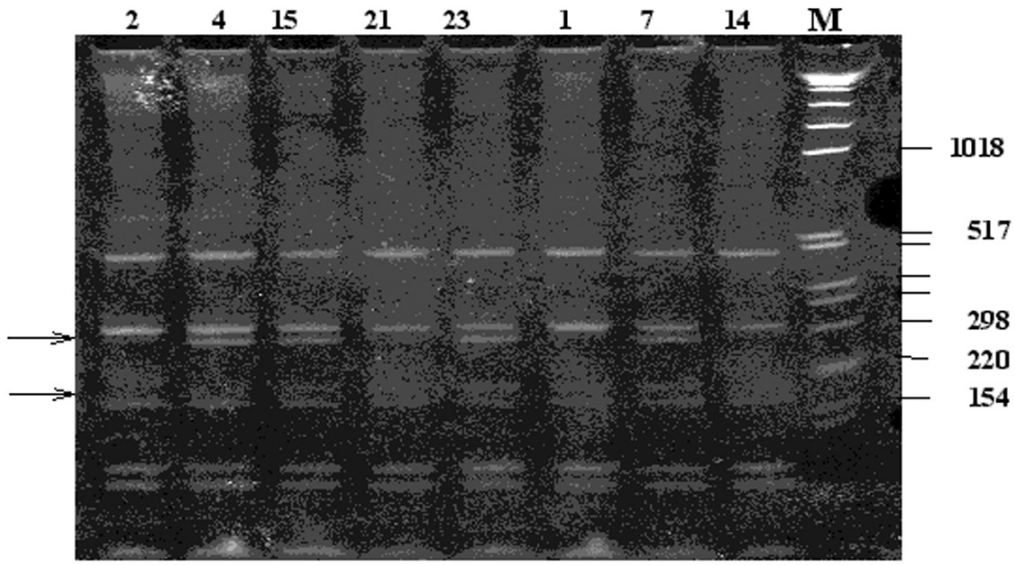

Figure 3. RAPD patterns obtained with combined primers GOU1/GOU7 for eight genotypes. sequence of genotypes from left to right: (a) iac-rm4-sm5, (b) su 0450/8909, (c) paymaster 303, (d) c-4727, (e) tashkent 2, (f) iac-13-1, (g) 4f, (h) lankart lx 571 and dna marker mw of 1 kb "Ladder". Arrows indicate polymorphisms.

origins at the 14 % level (Figure 7). Figure 8 shows grouping of 12 cultivars from Europe, Asia and Africa, based on SD's calculated on RAPD markers produced by short primers. Considering a level of 10%, genotypes 4F, SK 14, CA(68)41 and Tashkent 3 were confined to the same group. 4F and SK 14

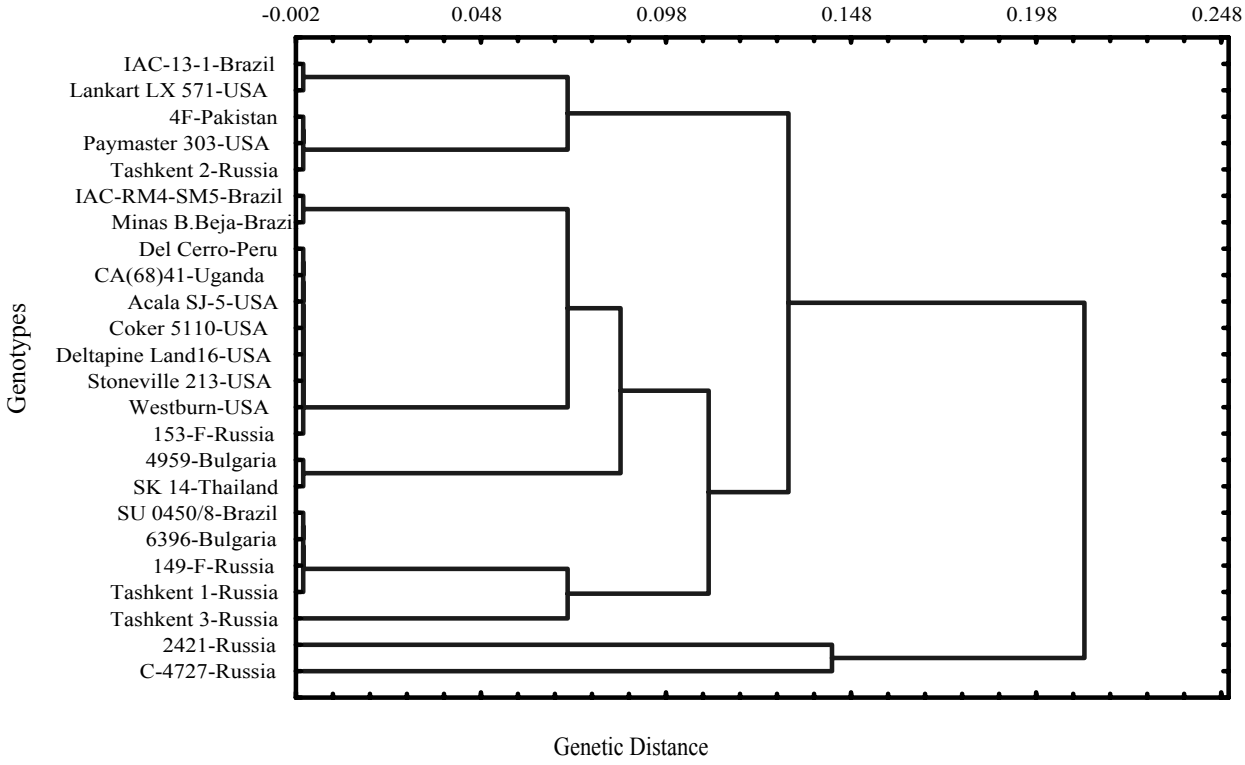

Figure 4. Dendrogram representing genetic distances estimated among 24 cotton genotypes from eight countries, based on 14 RAPD markers generated by long primers.

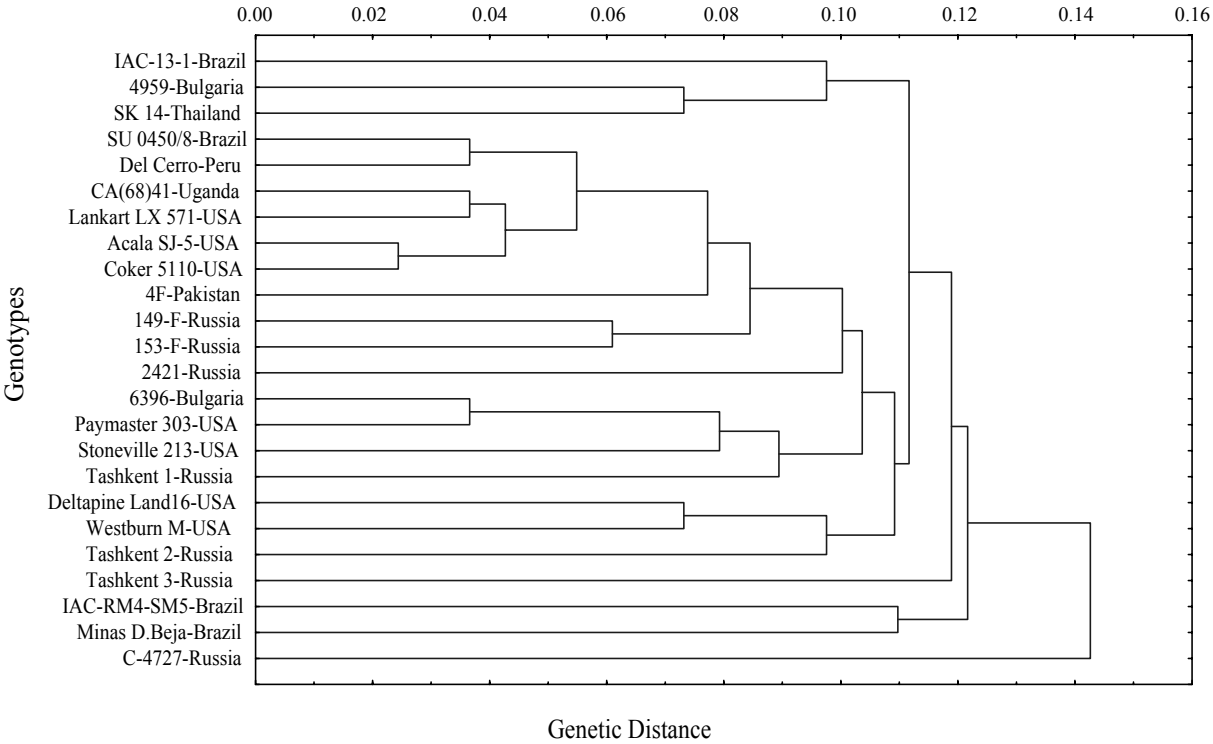

Figure 5. Dendrogram representing genetic distances estimated among 24 cotton genotypes from eight countries and based on 82 RAPD markers generated by short and long primers.

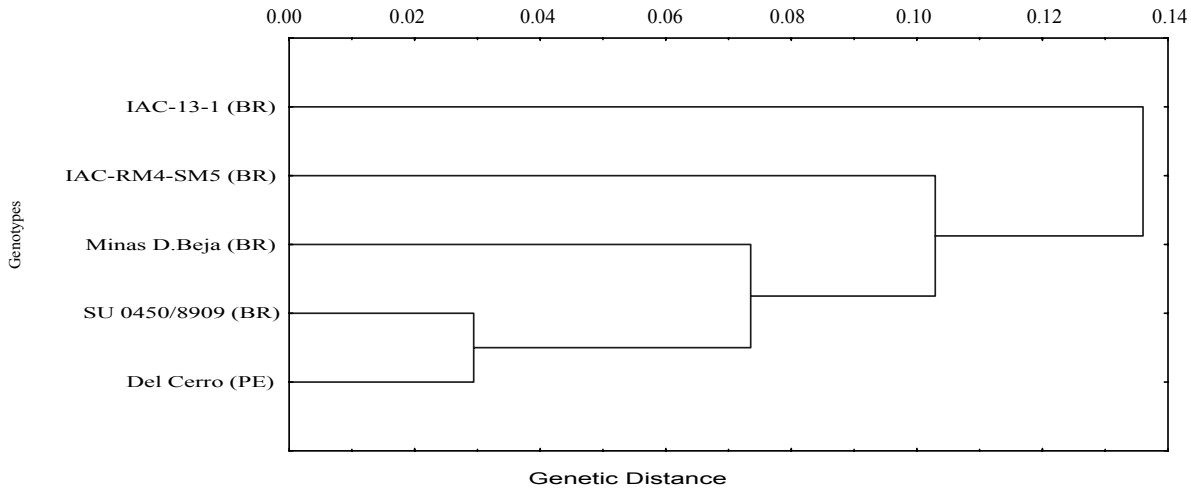

Figure 6. Dendrogram representing genetic distances between five South American genotypes based on 68 RAPD markers generated by short primers.

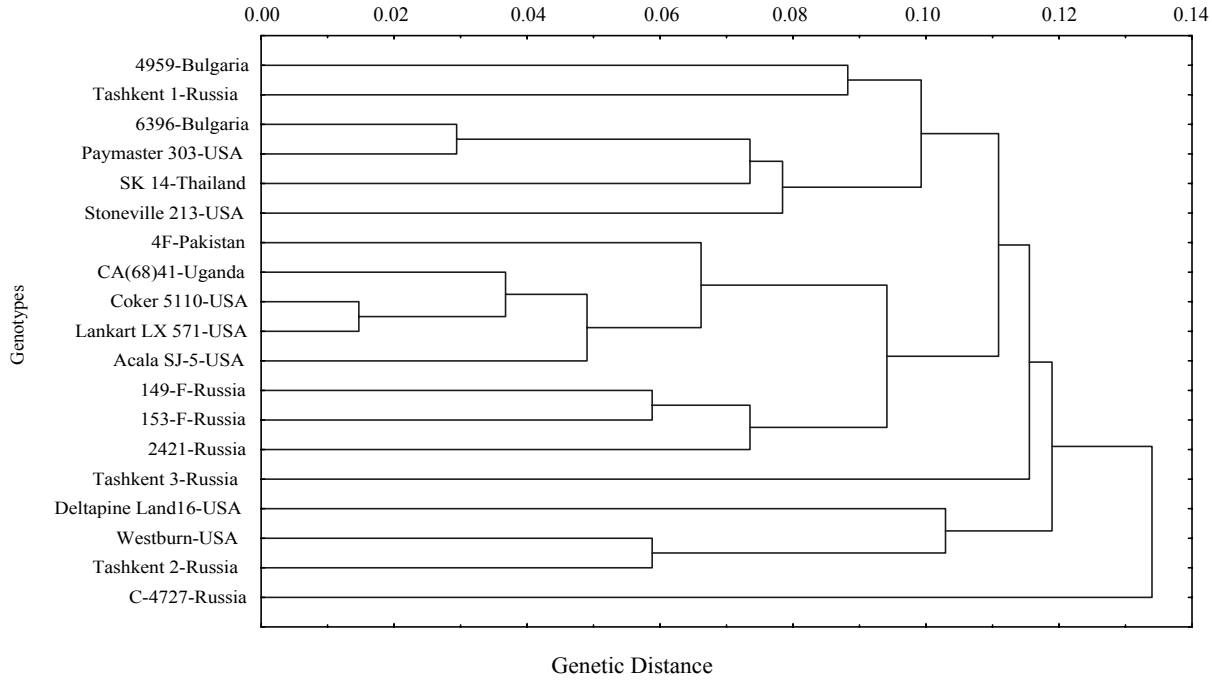

Figure 7. Dendrogram representing genetic distances between 19 genotypes from the United States and the Old World, based on 68 RAPD bands generated by short primers.

was the closest pair of entries (6% dissimilarity).

The SD dendogram (Figure 9) at the 10% level of GD divided US genotypes into three groups: A with Acala SJ-5, Coker 5110 and Lankart LX 571; B with Deltapine 16 and Westburn M; and C with Paymaster 303 and Stoneville 213. Coker 5110 and Lankart LX 571 showed a divergence smaller than 2%.

Penna (1980) estimated Euclidean Distances (ED) between the same 24 genotypes based on phenotypic markers (52 qualitative and quantitative traits) and utilizing numerical taxonomy. A correlation was calculated between these Euclidean Distances and the Genetic Distances obtained by SD's determined with RAPD markers produced by short primers. Although significant at the 1% probability level, the correlation found was moderate and negative (r = -0.43). Genotypes such as Del Cerro and SU 0450/8909 which were distant in the phenotypical analysis, were close in the analysis based on molecular markers. As expected, the correlation between ED's and GD's based on Nei and Li Distances, was also moderate and negative (r = -0.39).

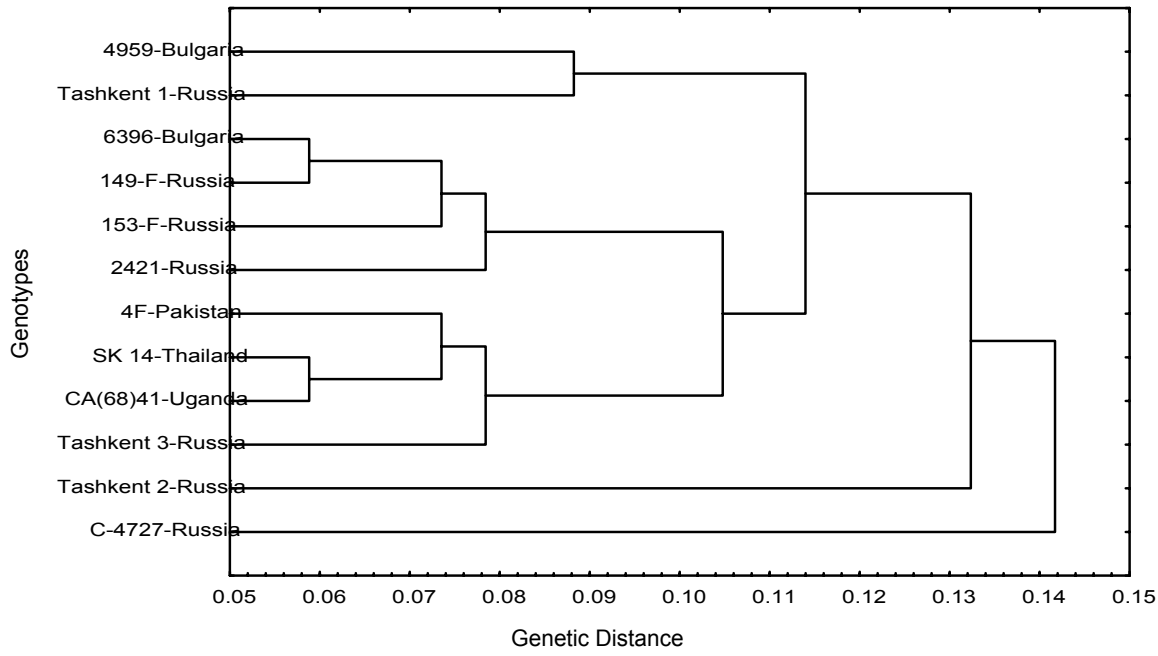

Figure 8. Dendrogram representing genetic distances between 12 genotypes from Europe, Asia and Africa based on 68 RAPD bands produced by short primers.

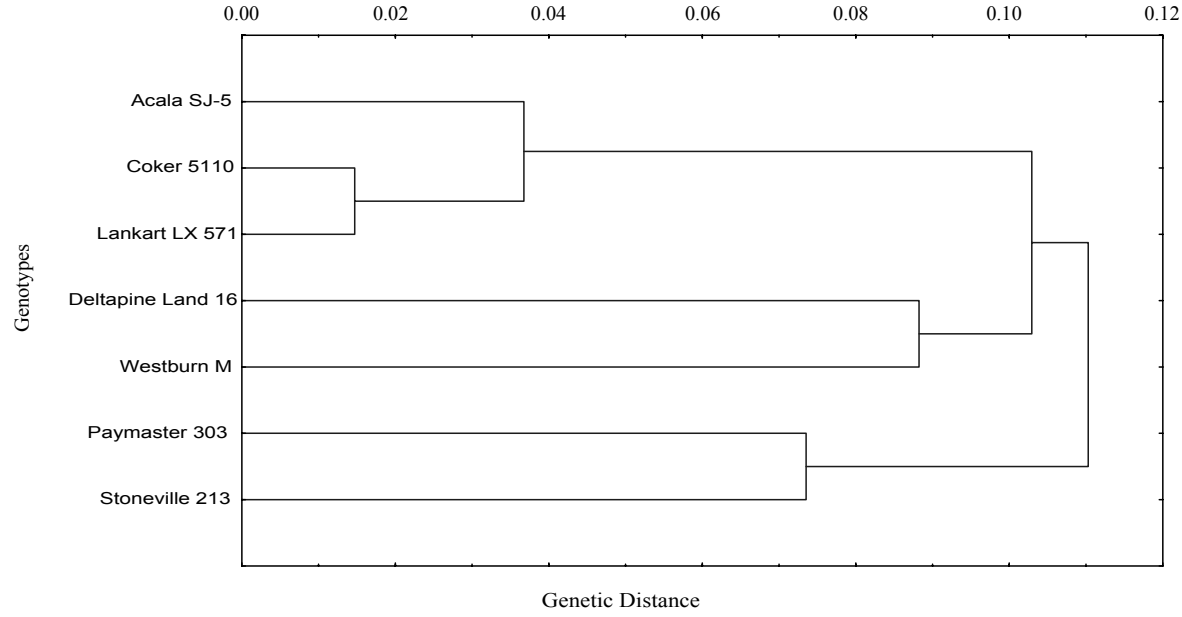

Figure 9. Dendrogram representing genetic distances between seven US cultivars, based on 68 RAPD bands produced by short primers.

When ED's based on phenotypic markers were correlated with GD's based on RAPD markers generated with both short and long primers and the three methods tested, values of r = - 0.4 for simple distance, r = - 0.37 for Jaccard distances and r = - 0,36 for Nei and Li distances were found, all of them statistically significant.

In the comparison involving only South American cultivars, according to the phenotypical analysis, the opposite was found in relation to the RAPD analysis. For example, the genotype SU 0450/8909 was the most divergent, whereas Minas D. Beja and IAC-13-1 were the more closely related ones (Penna, 1980), which definitely did not happen in this study. In the phenotypical analysis of the 19 genotypes from the US and Old World, cultivars were clearly separated into two groups: the US and the Old World, which again did not happen in the RAPD analysis.

Although variabilities estimated through RAPD markers were quite similar among the studied groups, the grouping of European, Asian and African genotypes have shown the greatest variation, followed by North American and South American entries.

The same discrepancy between the two types of analyses was detected also when only the 12 genotypes from Europe, Asia and Africa were isolated. In the phenetic analysis, genotypes 4F, CA(68)41 and SK 14 were the ones with the greatest divergence (Penna, 1980).

The distances obtained from phenotypic data by the same author, revealed a greater genetic variability than distances inferred from molecular markers. That could be due to the fact that phenotypic relationships are normally affected by genotype x environment interactions, specially quantitative traits, which in turn may inflate the estimates of distances. On the other hand, groupings determined on the phenotypic analysis seemed more logical than on the RAPD´s if the geographical origins of the entries are considered.

ACKNOWLEDGEMENTS

We would like to thank the help and participation of our colleagues at Laboratório de Genética Molecular of the Universidade Federal de Uberlândia. We gratefully thank the financial aid for this work from the FUNDAP and CAPES.

RESUMO

Análise comparativa de distâncias genéticas entre 24 genótipos de algodoeiro estimadas por marcadores RAPD e caracteres fenotípicos

O conhecimento das distâncias genéticas entre indivíduos ou populações em programas de melhoramento de plantas pode complementar informações fenotípicas auxiliando na escolha de indivíduos para hibridações e obtenção de populações segregantes com maior variabilidade. Os objetivos deste trabalho foram: identificar *primers* que gerem polimorfismos entre genótipos de algodoeiro (*Gossypium hirsutum* L.), comparar o efeito do tamanho na determinação de polimorfismos, determinar distâncias genéticas entre 24 genótipos de algodoeiro oriundos de oito países utilizando a técnica de DNA polimórfico amplificado ao acaso (RAPD) e comparar com resultados obtidos pela análise fenética dos mesmos genótipos baseada na avaliação de 52 características qualitativas e quantitativas usando taxonomia numérica. Os 24 genótipos foram avaliados pela técnica RAPD utilizando *primers* curtos e combinações de *primers* longos. RAPDs foram rgistrados como presença (1) ou ausência (0) de bandas intensas e reprodutíveis em reações duplicadas de PCR. Esses dados foram analisados e geraram matrizes de dissimilaridades baseadas em Distâncias Simples, Distâncias de Nei e Li e Distâncias de Jaccard. O método UPGMA foi utilizado para criar dendrogramas baseados nas matrizes de distâncias genéticas. Comparando os três métodos usados para estimação das distâncias genéticas, o método que melhor correlacionou com distâncias obtidas através dos caracteres fenotípicos foi a Distância Simples. Este método concordou em 93% com as Distâncias de Jaccard e 91% com as Distâncias de Nei e Li. A máxima distância simples verificada entre pares de genótipos baseadas nos polimorfismos produzidos pelos *primers* curtos foi de 18%, indicando possivelmente uma base genética não tão ampla quanto revelada por caracteres fenotípicos. A correlação entre análise fenética e análise RAPD baseada em *primers* curtos, foi moderada e negatica (r=-0.43).

REFERENCES

Beckmann, J. S. and Soller, M. 1983. Restriction fragment lenght polymorphisms in genetic improvement: Methodologies, mapping and costs. Theor Appl Genet. 67:35-43.

Comissão de fertilidade do solo do estado de Minutesas Gerais. 1989. Recomendações para uso de corretivos e fertilizantes em Minas Gerais: p.106-107. IV Aproximação, Lavras.

Doyle, J. J. and Doyle, J.L. 1990. Isolation of plant DNA from fresh tissue. Focus. 12:13-15.

Fryxell, P.A. 1992. A revised taxonomic interpretation of *Gossypium* L. (Malvaceae). Rheedea 2:108-165.

González, J. M. and Ferrer, E. 1993. Random amplified polymorphic DNA analysis in *Hordeum* species. Genome. 36:1029-1031.

Halward, T. M.; Stalker, H. T.; Larue, E. A. and Kochert, G. 1991. Genetic variation detectable with molecular markers among unadapted germ-plasm resources of cultivated peanut and related wild species. Genome. 34:1013-1020.

Jaccard, P. 1901. Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull Soc Vaudoise Sci Nat. 37:547-579.

John, M. E. 1992. An efficient method for isolation of RNA and DNA from plants containing polyphenolics. Nucleic Acids Research. 20:2381.

Loarce, Y.; Gallego, R. and Ferrer, E. 1996. A comparative analysis of the genetic relationships between rye cultivars using RFLP and RAPD markers. Euphytica. 88:107-115.

Nei, M. and Li, W.H. 1979. Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci. 76:5269-5273.

Penna, J.C.V. 1980. Comparisons among selected upland cotton cultivars and strains utilizing the methods of numerical taxonomy. Ph.D. Thesis. Oklahoma State University.

Skroch, P.; Tivang, J. and Nienhuis, J. 1992. Analysis of genetic relationships using RAPD marker data. p. 26-30. In: Applications of RAPD Technology to Plant Breeding. Joint Plant Breeding Symposia Series. Ccsa, ASHS, Minneapolis.

Sokal, R. R. and Michener, C.D. 1958. A statistical method for evaluating systematic relationships. Univ Kans Sci Bull. 38:1409-1438.

Sparks, D.L. 1995. Malvaceae: Cotton (*Gossypium* spp.) Advances in Agronomy. 55:135-140.

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.

Thormann, C.E. and Osborn, T.C. 1992. Use of RAPD & RFLP markers for germplasm evaluation. p.9-11. In: Applications of RAPD Technology to Plant Breeding. Joint Plant Breeding Symposia Series. Ccsa, ASHS, Minneapolis.

Welsh, J. and McClelland, M. 1990. Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acids Research. 18:7213-7218.

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

Recepted: June 28, 2001;

Accepted: December 15, 2001.