A Comparison Between RAPD and Phenotypic Analyses on 24 Cotton Genotypes

T.T. Cabral[1], J.C.V. Penna[2] and L.R. Goulart[3]
[1]Former graduate student, [2]. Professor, Agronomy Dept., [3]. Professor,
Genetics Dept., Univ. Fed. Uberlândia, P.O. Box 592, Uberlândia, MG, Brazil.

ABSTRACT

Knowledge of genetic distances among individuals or populations in plant breeding programmes can be used to complement phenotypic information in choosing individuals to use in breeding for populations with high genetic variability. Random Amplified Polymorphic DNA (RAPD) techniques may be advantageous in studies of genetic relationships. This paper discusses research to identify primers that produce polymorphism among cotton (G. hirsutum L.) genotypes, to compare the effects of primer length in the generation of polymorphism, to estimate genetic distances among 23 genotypes from eight countries using RAPD analysis and tro compare such estimates with those obtained through phenotypic analysis.

Introduction

Upland cotton, *Gossypium hirsutum* L., accounts for about 90% of the world's cotton production. A long history of breeding lies behind the gains in yield and fiber quality. In a breeding programme, knowledge of genetic distances among individuals or populations allows better germplasm utilization and more efficient selection of genotypes for crossing. In the determination of distances it is common to use phenotypic traits that may suffer interference from the environment or may not provide adequate levels of polymorphism. Molecular biology techniques utilise the structure of genetic material and have been used to determine genetic distances. Random Amplified Polymorphic DNA (RAPD) detects polymorphism of nucleotide sequences based on DNA, amplified using oligonucleotides with around 10 bases of arbitrary sequence as initiators (primers) (Welsh and McClelland, 1990; Williams *et al.* 1990).

The objective of this research was to identify primers to generate polymorphism among a group of cotton genotypes; evaluate the effect of primer length in detecting polymorphism; and estimate distances among 24 genotypes from eight countries using RAPD markers. The results were compared with those obtained with a phenotypic analysis of the same genotypes, evaluated for 52 qualitative and quantitative traits, utilising numerical taxonomy (Penna, 1980).

Techniques for polymorphism detection on DNA sequences, based on the chain reaction of DNA polymerase (PCR), find complementary sequences across the genome. If they recognize sequences located within a few hundred base pairs of each distant from one another and in opposite strand-strings, the included regions are amplified. In accordance with the amplified products, polymorphisms are generated that are interpreted as ultimate genetic variation. Several economically important crops have been analysed for genetic distances by means of molecular markers in both intraspecific and infraspecific levels. In cotton, near homozygous elite genotypes derived from interspecific hybridization as well as typical *G. hirsutum* and *G. barbadense* entries, were utilized in the generation of genetic distances dendrograms, built with both RAPD and morphologic data. Classifications obtained yielded similar results, with a correlation coefficient of 0.63 between genetic and taxonomic distances (Tatineni *et al.*, 1996). RAPD markers generated with 30 primers were used to

fingerprint 12 cultivars and one strain of *G. hirsutum* and one cultivar of *G. barbadense*, all of Australian origin. Out of the 453 markers identified, 15.2% were present only in the *G. barbadense* cultivar and 33.3% of the remaining markers were present in all 13 *G. hirsutum* cultivars. Cluster analysis of genetic distances among cultivars revealed a narrow agreement with known phylogenetic relations (Multani and Lyon, 1995).

Materials and Methods

Seeds of 24 genotypes of *G. hirsutum* L. were donated by "Empresa de Pesquisa Agropecuária de Minas Gerais" and "Empresa Brasileira de Pesquisa Agropecuária". The entries represent neither a random sample of the cultivars grown in the eight countries, nor a random sample of cotton-growing nations (Fig. 1). Seeds were planted at a rate of five per pot in a glass house on the University of Uberlândia (UFU) campus in 1996. The substrate used was a mixture 1/3 peat moss and 2/3 soil. Three plants out of the five were sampled for DNA extraction and DNA amplification: From the harvested material, 300 mg were macerated in liquid Nitrogen following DNA extraction, according to the protocol by Doyle and Doyle (1990), modified by addition of polyvinilpirrolidone (PVP-40). Primers: There was a preliminary trial to evaluate the best programme in the thermocycle. This was 5 min at 94º C, 45 cycles [15 seconds at 94º C, 15 seconds at 36º C, one min at 72º C]. Twenty-four 10-base primers (Operon Techn.) were evaluated and eleven, A01- A15, L04, L11, L13, L15, O02, O05, O06, O08 and O18 were utilized. DNA's were extracted from cultivars Acala SJ-5 and IAC-RM4-SM5. Each reaction contained: 25 ng DNA, 1 U Taq polymerase, 2 mM $MgCl_2$, 0.05 mM of each dNTP, 0.32 µM primer, completed with pure water to 25 µl. The primers selected were utilized to amplify DNA's of the 24 genotypes. The reaction was performed twice for each cultivar. Three combinations of primers developed at the Molecular Genetics Laboratory of UFU were evaluated: GOU01/GOU07, GOU03/GOU10 and GOU8/GOU9. Visualisation: Amplification products were separated on 6% polyacrylamide gels at 100 V for 4 hours and stained with ethydium bromide (5 (g/ml). RAPD's were registered as present vs. absent of intense bands.

Data analyses: A matrix was generated according with presence-1 vs. absence-0 of bands, and was used to estimate the following pair-wise distances: Percent Disagreement (PD)=N'AB/NT (Puterka *et al.*,1993); Jacard Distance (JD) (%)=(1-[a/(n-d)](x100 (Sneath and Sokal,1973); Nei and Li Complement (NLC) (%)=[1-2 NAB/(NA + NB)]x100 (Nei and Li, 1979). UPGMA (un-weighted pair-group method using arithmetic averages) was used to group genotypes and obtain dendrograms.

Results and Discussion

Concentration of reaction components that revealed most consistent bands was: 0.32 µM primer; 2 mM MgCl2; 0.05 mM of each dNTP; 1 U Taq DNA and 25 ng DNA. Eleven of the 24 short primers tested generated 68 bands, averaging 6.18 bands per primer. Band sizes were estimated to range from 300 to 2600 bp and 30 were polymorphic (44.12% polymorphism). Genetic distances based on band presence and absence and calculated by Percent Disagreement 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 varied between 1% (Coker 5110 and Lankart LX 571) and 12% (2421 and IAC-13-1, Minas D. Beja and 4959, DPL 16 and 153-F). A high correlation (r = 0.95) was found between the two matrices. Cluster analysis based on PD utilising UPGMA defined two groups at an arbitrary level of 12% (Fig. 1): 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 DPL 16; A5 with Westburn and Tashkent 2 and A6 with Tashkent 3. 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. It is evident that the genotypes are relatively close, indicating a possible narrow genetic variation.

The three combinations of long primers generated 14 bands ranging from 150 to 520 bp long and five (35.71%) were polymorphic. Dissimilarities ranged from 0 to 29% and the cluster analysis distinguished 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 an 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, DPL 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: B1 with the cultivar 2421

short primers. Although significant at the 1% probability level, the correlation found was negative (r = - 0.43). Genotypes such as Del Cerro and SU 0450/8909 that were distant in the phenotypic analysis, were close in the analysis based on molecular markers. As expected, the correlation between EDs and GDs based on Nei and Li Complements, was also moderate and negative (r = - 0.39). When EDs based on phenotypic markers were correlated with GDs based on markers generated with both short and long primers, it was found values of r= - 0.4 for percent disagreement, r = - 0.37 for Jacard's coefficient and r = - 0.36 for Nei and Li Complement., all statistically significant. In the phenotypic comparison involving South American entries, the opposite was found in relation to the RAPD analysis. SU

and B2 with C-4727. These primers could separate groups but did not differentiate among genotypes as efficiently as short primers.

The PD dendrogram based on data calculated from 68 bands generated through short primers on the five South American genotypes, demonstrates that the most divergent genotype was IAC-13-1 and the most closely related were SU 0450/8909 and Del Cerro (Figure 2). When only 19 genotypes were considered by excluding those from South American, there was no separation of groups according to origin at the 14 % level (Fig 3).

The PD dendogram (Fig. 4) at 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 divergence smaller than 2%.

Penna (1980) estimated Euclidean Distances (ED) between the same 24 genotypes, based on phenotypic markers (52 qualitative and quantitative traits), using numerical taxonomy. A correlation was calculated between EDs and the distances obtained by PDs determined through RAPD markers produced by

0450/8909 was the most divergent, whereas Minas D. Beja and IAC-13-1 were the more closely related (Penna, 1980). In the phenotypic analysis, the 19 genotypes from the New World and Old World were clearly separated into two groups: New World and Old World. This did not happen in RAPD analysis. Distances obtained with phenotypic data, revealed a greater genetic variability than those with RAPD markers. That could be due to the fact that the former may be affected by genotype x environment interactions that may inflate estimates.

Conclusions

1) Short primers: OPA-01, OPA-02, OPA-03, OPA-09, OPA-10, OPA-15, OPL-04, OPL-11, OPL-15, OPO-02, OPO-06 and the combinations of long primers GOU-

01/GOU-07, GOU-03/GOU-10 and GOU-08/GOU-09 generated at least one polymorphic band;

2) Short primers were more efficient than long primers in differentiating genotypes;

3) Maximum divergence between genotypes was 18% PD, according to RAPD markers (short primers), possibly indicating a narrow genetic variability;

4) PD was highly correlated with Nei and Li Complements (r = 0.95) and with Jacard Coefficients (r = 0.96); and 5) Genetic distances (PD's) obtained with RAPD markers (short primers) correlated negatively with Euclidean Distances based on phenotypic traits (r = - 0.43).

References

Doyle, J.J. and J.L. Doyle. (1990): Isolation of Plant DNA From Fresh Tissue. Focus: v.12:13-15.

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

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

Penna, J.C.V. (1980): Comparisons among selected upland cotton cultivars and strains utilizing the methods of numerical taxonomy., Oklahoma State Univ., Oklahoma 1980. 57p. (Ph.D. Thesis).

Puterka, G.J., W.C. Biack IV, W.M. Steiner and R.L. Burton. (1993): Genetic variation and phylogenetic relationships among world-wide collections of the Russian wheat aphid, *Diuraphis noxia* (Mordvilko), inferred from allozyme and RAPD-PCR markers. Heredity, v.70:604-618.

Sneath, P.H.A. and R.R. Sokal. (1973): Numerical Taxonomy. The Principles and Practice of Numerical Classification. W.H. Freeman and Co., San Francisco.

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

Welsh, J., and M. McClelland. (1990): Fingerprinting genomes using PCR with arbitrary primers. Nucleic Acid Research, Oxford, v.18:7213-7218.

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

Figure 1. Dendrogram representing genetic distances estimated among 24 cotton genotypes from eight countries, based on 68 RAPD markers generated by short primers.

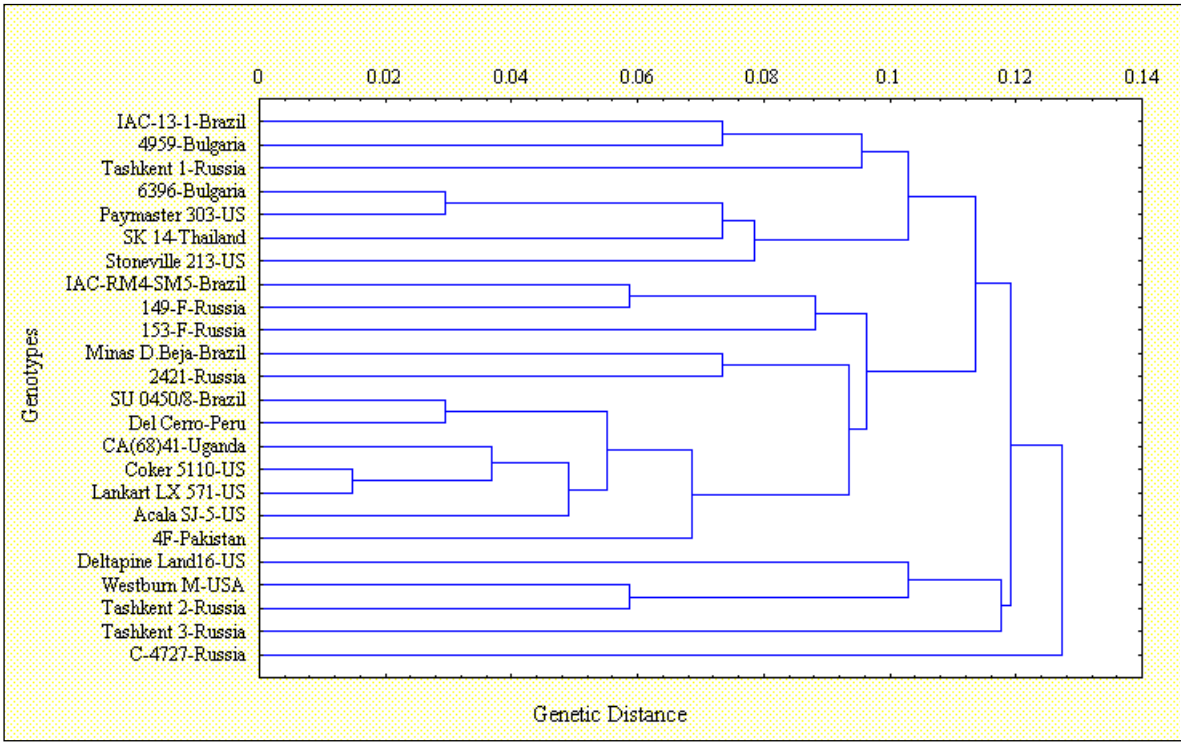

Figure 2. Dendrogram representing genetic distances between five South American genotypes based on 68 RAPD markers generated by short primers.

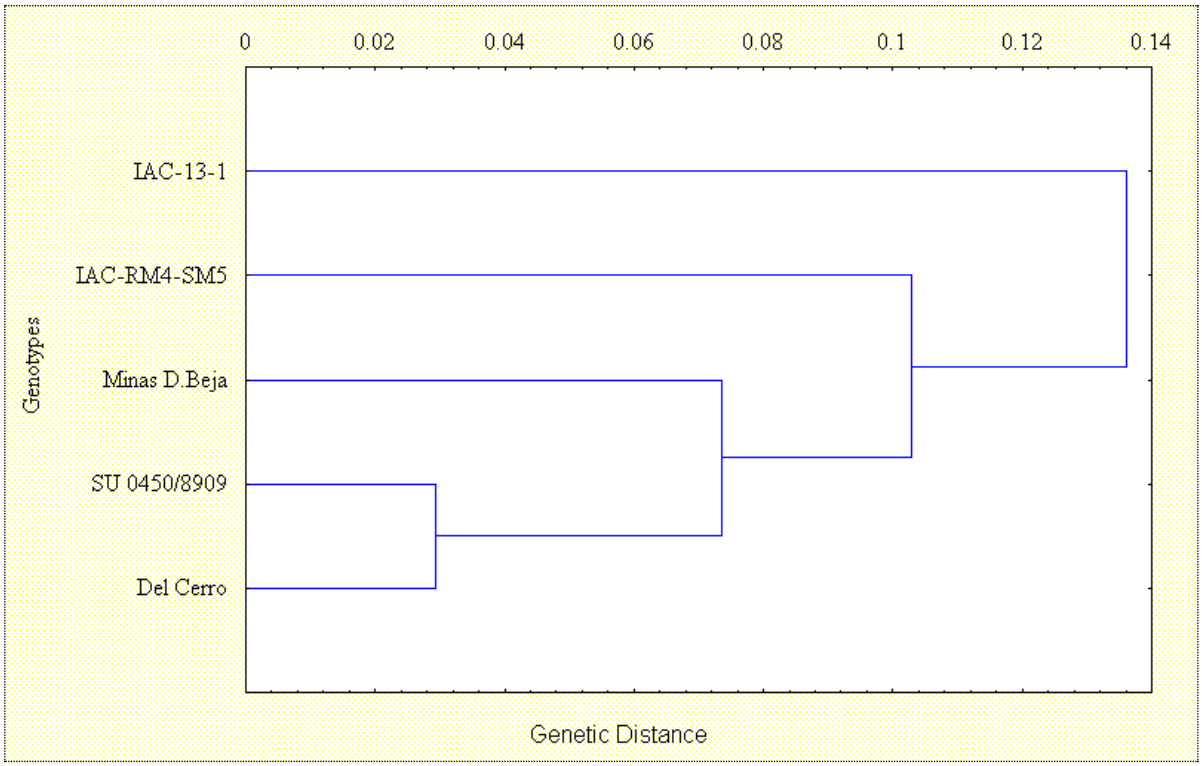

Figure 3. Dendrogram representing genetic distances between 19 genotypes from the United States and the Old World, based on 68 RAPD bands generated by short primers.

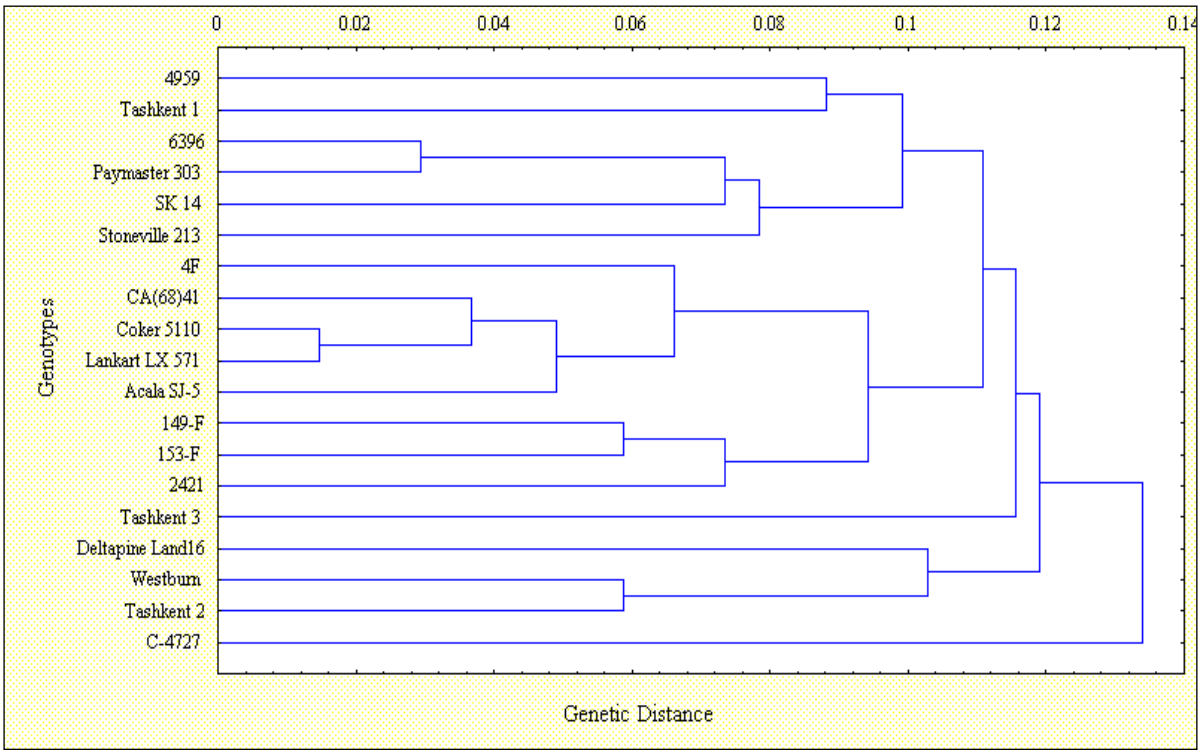

Figure 4. Dendrogram representing genetic distances between seven US cultivars, based on 68 RAPD bands produced by short primers.