Theor Appl Genet (2002) 105:325–331
DOI 10.1007/s00122-002-0947-8

H. J. Lu · G. O. Myers

Genetic relationships and discrimination of ten influential Upland cotton varieties using RAPD markers

Received: 22 May 2001 / Accepted: 8 October 2001 / Published online: 19 June 2002

Abstract Influential Upland cotton (*Gossypium hirsutum* L.) varieties are those that have the higher genetic contributions to modern Upland cultivars than other germplasms. Our previous research has shown significant differences in general combining ability (GCA) effects for yield, yield components, and fiber properties among ten influential cotton varieties. In this study, we used random amplified polymorphic DNA (RAPD) data to evaluate DNA variation of these ten varieties. Of 86 random decamer primers screened for their capability of amplifying DNA via the polymerase chain reaction (PCR), 63 generated a total of 312 DNA fragments. Forty two bands were polymorphic, which showed a low percentage (13.5%) of DNA variation among these influential varieties. Genetic similarities among the ten varieties based on RAPD data were from 92.7% to 97.6%. All of the varieties were individually identified by variety specific markers in genetic fingerprinting. One primer, UBC-149, amplified a 1,430-bp DNA fragment that was absent in five varieties and present in the other five varieties. This RAPD marker had significant negative relationships with GCA-effect estimates for seed cotton yield, lint yield, number of bolls per plant and micronaire, and significant positive relationships with GCA effects for boll size and seed index. This finding, for the first time, identifies a DNA fragment in cotton that is a potential DNA marker linked to a yield gene(s) or a yield-related gene(s).

Communicated by P. Langridge

H.J. Lu (✉) · G.O. Myers
Department of Agronomy,
Louisiana Agricultural Experiment Station,
Louisiana State University Agricultural Center,
Baton Rouge, LA 70803, USA
e-mail: luh@missouri.edu
Fax: +1-573-882-7017

Present address:
H.J. Lu, Department of Agronomy, University of Missouri,
Columbia, MO 65211, USA

Keywords RAPD · Genetic diversity · Molecular marker · Influential varieties · Upland cotton (*Gossypium hirsutum* L.)

Introduction

The 1970 corn leaf blight epiphytotic has drawn plant breeders' attention to the genetic uniformity of cultivars in crop species. Information on genetic uniformity (or genetic diversity) of crop cultivars in USA has been available in corn (Smith 1988; Gerdes and Tracy 1993), soybean (Delannay et al. 1983; Specht and Williams 1983), wheat (Cox et al. 1986), spring oat (Rodgers et al. 1983) and peanut (Knauft and Gorbet 1989). Genetic interrelationships among cotton cultivars released between 1970 and 1990 have been recently investigated by several researchers (Bowman et al. 1995, 1996; May et al. 1995). They found that the genetic base was diverse among clusters of cotton varieties, but narrow within each cluster.

One approach to evaluating the genetic diversity of cultivars released during a certain period of time is through calculation of the coefficients of parentage (*r*). Some of the assumptions made to compute *r* values (Cox et al. 1986) include: (1) that a cultivar derived from a cross obtains one-half of its genes from each parent; (2) that all lines used in crossing are homozygous and homogeneous; (3) that all original germplasm lines are unrelated to each other; (4) that the *r* value between a selection from a cultivar and the cultivar equals 0.75; and (5) that the *r* value between two selections from the same cultivar is $r = 0.75^2 = 0.56$. When Bowman et al. (1996) estimated the average genetic contribution of each of the 49 breeding programs to the 260 cotton cultivars released between 1970 and 1990, they assumed that all influential lines were unrelated. But sometimes such assumptions are not met. For example, 'PD 2165' and 'Delcott 277' were developed from crosses containing Beasley's triple hybrid germplasm; 'McNair 235' was selected from the cross of 'Coker 201' × 'PD 2165';

'DES 56' was bred from the cross of 'Stoneville 213' × 'PD 2164'; and 'PD 2164' and 'PD 2165' are sister germplasm lines (Culp 1998). Therefore, an understanding of the genetic similarity or distance of germplasm lines will provide additional information on the genetic diversity of modern cultivars.

Genetic relationships and genetic diversity can also be measured using polymorphic markers. Tatineni et al. (1996) used morphological characters to investigate the genetic diversity of 16 near-homozygous elite cotton genotypes derived from interspecific hybridization. They measured 19 morphological traits traditionally used to distinguish *Gossypium hirsutum* from *Gossypium barbadense* on 16 genotypes and Pima S-6 (*G. barbadense*). Using the morphological data, they were able to classify the 16 genotypes into two clusters, one resembling *G. hirsutum* and the other *G. barbadense*. Isozyme analysis is another method to study the genetic diversity of cultivars. To assess levels and patterns of genetic variation in *G. hirsutum*, Wendel et al. (1992) evaluated 538 accessions representing the full spectrum of morphological and geographical diversity for allozyme variation at 50 loci (14 allozymes). Of the 538 accessions, 50 were modern Upland cultivars that represented the four categories of cotton commercially grown in the USA (Eastern, Plains, Delta, and Acala). They found modest levels of variation in all accessions but low levels of variation in modern cultivars. The use of morphological characters and isozymes for evaluation of genetic variation is, however, limited due to a lack of adequate levels of polymorphism. Molecular markers generated by different techniques developed in the 1980s and 1990s have become powerful tools to examine genetic variation because a large number of polymorphic markers are available. Due to its simplicity and speed, RAPD analysis is one of the most-commonly used techniques because it produces numerous molecular markers (Demeke et al. 1996; Cao and Oard 1997; Castagna et al. 1997; Heun and Helentjaris 1993). In cotton, RAPD markers have been used to study the genetic relationships and genetic variation of cultivars or germplasm lines (Multani and Lyon 1995; Tatineni et al. 1996; Iqbal et al. 1997).

The most influential Upland cotton lines from 16 cotton breeding programs were identified based on the analysis of the pedigrees of 260 cultivars released between 1970 and 1990 (Bowman et al. 1996). The general combining ability (GCA) and specific combining ability (SCA) for yield, yield components and fiber properties of the ten influential cotton varieties have been estimated (Lu and Myers 1998; Myers and Lu 1998). Genetic variation in these traits among the ten influential varieties was found to be significant for all the traits except for seed cotton yield and length uniformity. The objectives of this research were to: (1) evaluate the level of genetic diversity of ten influential cotton varieties using RAPD markers, (2) identify variety specific RAPD markers to fingerprint the ten varieties, and (3) exploit possible relationships between results of the diallel analysis and the DNA sequence polymorphism.

Materials and methods

Plant materials

The ten influential cotton varieties used in this research were 'McNair 235' from McNair/Northrup King, 'Coker 100W' from Coker Pedigreed Seed Co., 'Paymaster 54' from Paymaster/Cargill Seeds, 'Stoneville 2' from Stoneville Pedigreed Seed Co., 'Lankart 57' from Lankart Seed Farms, 'PD 2165' from USDA-ARS and South Carolina AES, 'Deltapine 15' from Delta and Pine Land Co., 'CA 614' from Texas AES-Lubbock, 'DES 56' from Delta Experiment Station, Miss., and 'Delcot 277' from Missouri AES. The ten varieties were used as parents to produce 45 F_1s in a half-diallel crossing.

DNA extraction

About ten plants were grown for each variety in pots in a greenhouse. The F_1s were grown at Winnsboro, Louisiana, in 1997. Young fully expanded leaves (<1 week old) were harvested and bulked for DNA isolation from plants of each of the ten varieties and the 45 F_1s. Total genomic DNA was extracted using a procedure slightly modified from that of Altaf et al. (1997). About 3 g of leaf tissue was put into a 50-ml centrifuge tube and agitated into small pieces after liquid nitrogen was added to the sample. The leaf sample was then mixed with 10 ml of DNA extraction buffer [0.1 M Tris–HCI (pH 8.0), 0.02 M EDTA (pH 8.0), 1 M of NaCl, 2% (w/v) CTAB (hexadecyl triammonium bromide), 2% (w/v) PVP 40 (polyvinylpyrolidone), 1 mM of phenanthroline, and 0.2% (w/v) β-mercaptoethanol] and was homogenized at 10,000 rpm for approximately 15 s. The sample was incubated at 70 °C for 1 h in a water bath. After the sample was cooled for 10 min, an equal volume of chloroform-isoamyl alcohol (24:1 v/v) was added and mixed gently with the sample to remove proteins. The mixture was centrifugated at 12,000 rpm for 10 min. The aqueous phase was transferred to a clean tube and extracted with chloroform-isoamyl alcohol (24:1 v/v) and centrifugated again. The supernatant was transferred to a new tube. Approximately 2 vol of ice-cold ethanol was added and the sample was kept at –20 °C for 30 min to precipitate DNA. DNA was spooled out with a small glass hook and washed two times, first with a washing solution of 80% ethanol and 15 mM of ammonium acetate for 20 min with gentle shaking and secondly with 100% ethanol for 20 min with gentle shaking. The DNA pellet was air-dried, dissolved in 5 ml of high salt TE [10 mM of Tris (pH 8.0), 1 mM of EDTA (pH 8.0), and 1 M of NaCl], and incubated at 60 °C for 2 h. A 5-ml of extraction buffer without β-mercaptoethanol was then added and the sample was placed on a shaker for 2 h to dissolve the remaining pellet. Again, DNA was precipitated with ice-cold ethanol and washed with ethanol as described above. The DNA pellet was dried and resuspended in 200 µl of 1 × TE buffer. The DNA sample was treated with 2 µl of RNase A (10 mg ml^{-1}) to destroy RNA overnight at room temperature. The concentration of the genomic DNA preparations was determined by spectrophotometry at 260 nm using a Gene Quant (Pharmacia Biotech) and diluted to 10 ng μl^{-1} with sterile water. DNA was then stored in a freezer (–20 °C) until used in PCR amplification.

RAPD analysis

The 25 µl PCR amplification mixture contained 2.5 µl of 10 × PCR-II buffer [50 mM of Tris (pH 8.3), 500 mM of KCl]; 1.5 mM of $MgCl_2$; 0.2 mM each of dATP, dCTP, dGTP and dTTP (Pharmacia Biotech); 0.3 mM of a decamer primer (University of British Columbia, Vancover); 0.5 units of Ampli*Taq* DNA polymerase (Perkin Elmer); and 25 ng of genomic DNA as a template. DNA was amplified in a Perkin Elmer GeneAmp PCR System 9600, programmed for a first denaturation step of 2 min at 94 °C followed by 45 cycles of 94 °C for 1 min, 35 °C for 1 min and 72 °C for 2 min. After the completion of 45 cycles, a final exten-

sion at 72 °C was carried out for 5 min. The completed reactions were then held at 4 °C until electrophoresis.

PCR products were separated by loading 12 μl of each sample and 2 μl of loading buffer type II [0.25% bromophenol blue, 0.25% xylene cyanol FF, and 15% Ficoll (Type 400; Pharmacia Biotech) in water] on a 1.2% agarose gel prepared with 0.5 × TBE buffer. The samples were subject to electrophoresis at 90–100 V for 4 h in 0.5 × TBE buffer. The gel was stained with ethidium bromide and visualized under UV light. The amplification patterns were photographed on a Mitsubishi thermal paper attached to a Gel Doc 1000 (Bio-RAD) and Mitsubishi Video Copy Processor system controlled by the software Molecular Analyst. A total of 86 primers (from UBC-115 to UBC-200) were screened for their ability to amplify DNA fragments. The amplified DNA fragments were scored as present or absent for all the varieties.

Data analysis

The RAPD data were first checked for variety specific DNA markers which could be used to genetically fingerprint the ten influential cotton varieties. The coefficients of genetic similarity between all pairwise comparisons were computed using Jaccard's formula $J_{ij} = a/(n–d)$ (Jaccard 1908). In this formula, a is the number of DNA fragments present in both variety *i* and variety *j*, d represents the number of DNA fragments absent in both varieties, and n is the total number of DNA bands. The obtained similarity coefficients were transformed into genetic distance (D) according to Swofford and Olson (1990) as: D = –ln J. Genetic distances were utilized to produce a dendrogram by using the UPGMA method.

GCA effects of the ten varieties for yield, yield components and fiber properties have been evaluated (Lu and Myers 1998; Myers and Lu 1998). To exploit possible relationships between GCA and an amplified DNA fragment showing polymorphism among the varieties, linear regression analysis was conducted using the following model

$$Y_i = b_0 + b_1 X_i + e_i,$$

where Y_i is GCA value, X_i is presence (1) or absence (0) of the band in question, b_0 is the intercept, b_1 is the slope, and e_i is the residual (i = 1, 2,…,10). b_1 was statistically tested for Ho: $b_1 = 0$ using a *t*-test. Each of the polymorphic DNA fragments and each of the traits were analyzed for correlation in F_1s to confirm the relationships detected in the ten varieties in the linear regression analysis.

Results and discussion

Primer screening and variety fingerprinting

Eighty six primers were screened for their ability to amplify the genomic DNA of the ten influential cotton varieties. Twenty three amplified no fragments or yielded smeared bands that could not be clearly identified. A total of 312 fragments was generated by the remaining 63 primers, with an average of about five bands per primer. The number of DNA fragments amplified with each primer was a little higher than that reported by Tatineni et al. (1996) but much lower than that reported by Multani and Lyon (1995). Bands that a primer yielded in this study ranged from one (UBC-116) to ten (UBC-141 and UBC-173). Of these 312 DNA bands amplified by 63 primers, 42 fragments showed polymorphism among the ten varieties, which is an approximately 13.5% polymorphism. The rest of the bands (86.5%) were monomorphic in all the ten varieties. These results indicate that the level of DNA variation was low among the ten varieties.

Table 1 Variety identification by variety specific RAPD markers

Variety	Primer	DNA fragment (bp)
McNair 235	UBC-148	850 (+)
	UBC-185	590 (–)
Coker 100W	UBC-132	1,710 (+)
Paymaster 54	UBC-191	420 (+)
Stoneville 2	UBC-148	850 (+)
	UBC-185	590 (+)
Lankart 57	UBC-166	940 (+)
PD 2165	UBC-192	430 (+)
Deltapine 15	UBC-180	1,100 (–)
CA 614	UBC-163	1,500 (+)
DES 56	UBC-141	850 (–)
	UBC-148	950 (+)
Delcot 277	UBC-141	850 (+)

All of the ten varieties could be identified individually by using one variety specific RAPD marker or a combination of two or more markers (Table 1). Primer UBC-132 yielded a 1,700-bp fragment (UBC-132-1710) that distinguished 'Coker 100W' from the other nine varieties. 'Paymaster 54' could be recognized by a single marker (UBC-191-420). 'Lankart 57' could be identified by the RAPD marker UBC-166-940. The marker UBC-192-430 can be used to identify 'PD 2165'. 'CA 614' was distinguishable by the marker UBC-163-1500. The above five varieties were identified by the presence of a special fragment each, while 'Deltapine 15' could be recognized by the absence of the band UBC-180-1100. The remaining four varieties could only be identified by a combination of two markers. For example, 'McNair 235' could be identified by a combination of the presence of UBC-148-850 and the absence of UBC-185-590. There was no single primer that could identify more than one variety without the aid of other primers. All ten varieties could be differentiated by using nine primers (Table 1). Therefore, RAPD markers can be used for identification of cotton varieties.

Cluster analysis

The genetic similarity matrix of RAPD data for the ten varieties were constructed based on Jaccard's (1908) coefficients of similarity and is shown in Table 2. The genetic similarities of the ten varieties were high, ranging from 92.7% to 97.6%. 'Coker 100W' and 'PD 2165' had the greatest similarity (97.6%). The number of amplified fragments from both 'Coker 100W' and 'PD 2165' with 63 primers was 290. These two varieties differed from each other only in seven bands. The genetic similarity between 'McNair 235' and 'Stoneville 2' was the second highest (97.3%). They had 286 amplified fragments in

Table 2 Similarity coefficients between pairs of ten influential cotton varieties based on RAPDs

	1	2	3	4	5	6	7	8	9	
1	McNair 235									
2	Coker 100W	0.966								
3	Paymaster 54	0.935	0.933							
4	Stoneville 2	0.973	0.953	0.933						
5	Lankart 57	0.946	0.946	0.945	0.959					
6	PD 2165	0.956	0.976	0.929	0.963	0.956				
7	Deltapine 15	0.956	0.956	0.955	0.943	0.949	0.939			
8	CA 614	0.940	0.940	0.946	0.960	0.953	0.950	0.943		
9	DES 56	0.943	0.949	0.956	0.963	0.956	0.953	0.959	0.956	
10	Delcott 277	0.930	0.937	0.932	0.937	0.943	0.946	0.927	0.934	0.946

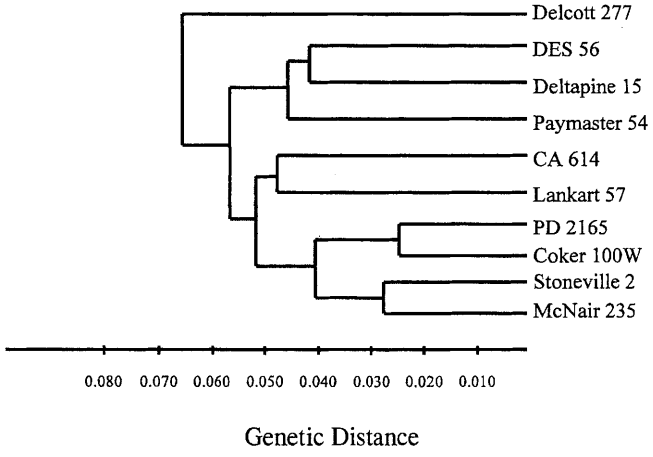

Fig. 1 Cluster analysis of the ten influential cotton cultivars based on RAPD data

common and eight fragments different. The similarity coefficient of 'Delcott 277' with 'Deltapine 15' was the lowest (92.7%). The primers generated 291 and 287 fragments from 'Decott 277' and 'Deltapine 15', respectively. The number of fragments that could be amplified from both varieties was 278 with the other 22 bands being polymorphic in these two varieties. 'PD 2165' and 'Paymaster 54' were 92.9% similar and the second least-similar variety pair. Among the ten varieties, 'Delcott 277' had the least genetic similarity with all other varieties. Its similarity was from 92.7% (with 'Deltapine 15') to 94.6% (with 'PD 2165' and 'DES 56').

Genetic similarity coefficients for the ten varieties in Table 2 were converted to genetic distance estimates using the formula D = -ln J (Swofford and Olson 1990). Genetic distance estimates were then utilized to produce a cluster by the UPGMA method (Fig. 1). Since DNA similarities among the ten varieties were high as revealed by RAPD analysis, the genetic diversity of these varieties was low. All of the genetic distance estimates were less than 0.10. 'Coker 100W' and 'PD 2165' grouped first because the genetic distance between them was the least (0.024). The genetic distance between 'McNair 235' and 'Stoneville 2' was 0.027 and these two varieties grouped together. In the dendrogram, 'Lankart 57' clustered with 'CA 614', and 'Paymaster 54' grouped with 'Deltapine 15' and 'DES 56'. 'Delcott 277' did not cluster with any other varieties, reflecting a relatively higher genetic distance from other varieties. However, the results of the cluster analysis should be interpreted with caution as the genetic constitutions of these ten influential lines appear to be highly similar.

Information on genetic diversity or genetic similarity in cotton is gradually being accumulated. Multani and Lyon (1995) found between 92.1 and 98.9% genetic similarity in nine closely related Australian cultivars of *G. hirsutum* L. based on their RAPD data. However, two other cultivars ('DP 90' and 'CS 50') and one breeding line ('001–1223') were clearly distinct and showed an approximately 30% genetic distance from those nine cultivars. The *G. barbadense* cultivar 'Pima S-7' was only about 57% similar to the *G. hirsutum* cultivars. In a similar study conducted by Iqbal et al. (1997), the genetic similarity of 17 *G. hirsutum* varieties ranged from 81.51 to 93.41%. Four other *G. hirsutum* cultivars and one *Gossypium arboreum* cultivar showed lower similarity (in most cases < 80%) with the 17 cultivars and with each other. Tatineni et al. (1996) reported high genetic distance estimates (0.25 – 2.05) for 19 genotypes including 16 near-homozygous elite genotypes derived from interspecific hybridization between *G. hirsutum* and other *Gossypium* spp. ['TM-1' (*G. hirsutum*), 'Pima 3-79' and 'Pima S-6' (both *G. barbadense*)]. It is difficult, however, to compare their results with others because the experimental materials were progenies of interspecific hybrids plus two species, and those primers that yielded monomorphic amplified products in RAPD analysis were excluded from calculation of genetic similarity or genetic distance. In contrast, Wendel et al. (1992) applied allozyme analysis to 50 U.S. Upland cultivars to measure genetic diversity and found that 14 (28%) of 50 allozyme loci were polymorphic. Later, they (Brubaker and Wendel 1994) reported that only 14 (7%) of 205 RFLP loci were polymorphic among the 23 Upland cultivars used in their study. In the present study, only 13.5% of 312 DNA fragments generated by 63 primers showed polymorphism. This result is more similar to those of Wendel et al. (1992) and Brubaker and Wendel (1994) than those of Multani and Lyon (1995) and Iqbal et al. (1997). According to Wendel et al. (1992) and Brubaker and Wendel (1994), modern Upland cotton cultivars are genetically depauperate. The results of this study indicate that the most important germplasm represented by

highly influential cotton lines are destitute of variation at the DNA level as well.

Relationships between RAPD marker UBC-149-1430 and GCA effects

As shown above, 270 of 312 DNA fragments produced by 63 primers were present in all the ten parents, which means that they are monomorphic. The remaining 42 DNA fragments showed polymorphism in the ten varieties and were used as molecular markers to exploit possible relationships between certain markers and GCA-effect estimates for yield, yield components and fiber quality traits. Relatedness was detected between one DNA marker, UBC-149-1430, and several GCA effects. The nucleotide sequence of primer UBC-149 that amplified this DNA fragment (1430 bp) is 5′AGCAGCGTGG 3′. UBC-149-1430 was present in 'Stoneville 2', 'Lankart 57', 'PD 2165', 'CA614' and 'Delcott 277', but absent in 'McNair 235', 'Coker 100W', 'Paymaster 54', 'Deltapine 15' and 'DES 56' (Fig. 2). Relationships between this marker and GCA-effect estimates from the ten varieties are shown in Table 3. Five varieties in which the DNA fragment was missing had positive GCA-effect estimates for seed cotton yield, while the other five varieties from which the DNA fragment was amplified showed negative or low positive GCA effects, suggesting that there existed a negative relationship between the presence of this DNA band and GCA for seed cotton yield. Similarly, GCA effects of five varieties that did not have the band were positive (in four varieties) or slightly negative (in 'Coker 100W') for lint yield, and GCA estimates of the other five varieties containing the DNA band were negative (in four varieties) or low positive (in 'Delcott 277'). Correlation is more apparent between GCA effects for the number of bolls per plant and the banding patterns because all the five varieties that did not contain the DNA band had positive GCA effects for the trait and all the other five varieties having the band expressed negative GCA effects. GCA effects for boll weight appear to be positively related to the DNA band as indicated by four negative GCA estimates and one positive estimate (in 'Coker 100W') in five varieties without the DNA band, and by four positive GCA-effect estimates and one negative estimate (in 'PD 2165') in the varieties with the band. Of five varieties in which the band was absent, four expressed positive GCA effects and one ('Coker 100W') had a negative GCA estimate for lint percentage, whereas only three of the other five varieties with the band showed negative GCA effects and the remaining two had positive GCA estimates. Thus, the relationship between the band and GCA effects for lint percentage was weaker compared with those described above. Correlation between the band and the GCA-effect estimates for lint index was positive in that GCA effects were negative in four of the five varieties in which the DNA fragment was absent and positive in four of the other five varieties in which the fragment was

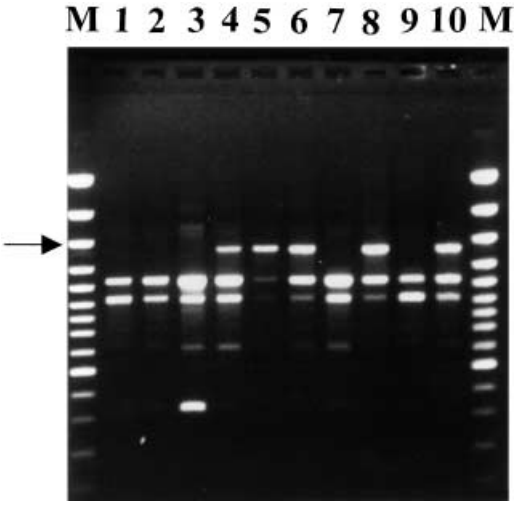

Fig. 2 The DNA banding pattern amplified by primer UBC 149 (RAPD) in ten influential cotton varieties. The *arrow* indicates the marker, UBC-149-1430. *Lane*1=McNair 235, *2*=Coker 100W, *3*=Paymaster 54, *4*=Stoneville 2, *5*=Lankart 57, *6*=PD 2165, *7*=Deltapine 15, *8*=CA 614, *9*=DES 56, *10*=Delcott 277. *M* is the molecular-weight marker (100-bp ladder plus)

Table 3 Relationships between Band UBC-149-1430 and GCA-effects for yield, yield components, and fiber properties of ten influential cotton cultivars

Cultivar	Band	SCY[a]	LY	B/P	BW	LP	SI	LI	UHM	E_1	Str.	UI	Mic.
McNair 235	0[b]	221.26	133.62	0.5928	–0.160	1.61	–0.711	0.071	–0.604	–0.076	–0.859	0.094	0.201
Coker 100W	0	70.95	–23.07	0.0006	0.101	–1.45	0.109	–0.367	0.052	–0.059	0.056	–0.067	–0.061
Paymaster 54	0	136.74	88.65	0.7023	–0.346	1.06	–1.150	–0.375	–0.272	0.043	–0.700	–0.337	0.048
Deltapine 15	0	117.90	89.39	0.5186	–0.249	1.34	–1.038	–0.224	–0.396	0.007	–0.905	–0.354	0.100
DES 56	0	341.09	139.84	1.0561	–0.335	0.57	–0.419	–0.063	0.281	0.049	0.167	0.134	0.203
Stoneville 2	1	–167.91	–127.50	–0.6150	0.250	–2.07	0.581	–0.287	–0.004	–0.061	0.309	–0.034	–0.117
Lankart 57	1	–375.60	–169.08	–1.2574	0.576	–0.97	1.716	0.694	0.465	0.329	–0.441	0.323	–0.145
PD 2165	1	–381.24	–154.43	–0.8253	–0.038	–0.49	0.751	0.301	0.684	–0.404	1.714	0.415	–0.079
CA 614	1	–26.09	–3.26	–0.1011	0.032	0.37	–0.133	0.058	–0.646	–0.257	–0.014	–0.399	0.023
Delcott 277	1	62.90	25.83	–0.0716	0.169	0.03	0.294	0.194	0.440	0.428	0.674	0.225	–0.174

[a] SCY = seed cotton yield, LY = lint yield, B/P = bolls/plant, BW = boll weight, LP = lint percentage, SI = seed index, LI = lint index, UHM = upper half mean fiber length (mm), E_1 = fiber elongation, Str. = fiber bundle strength (g/tex), UI = fiber length uniformity index, Mic. = micronaire

[b] 0 = absence of band; 1 = presence of band

Table 4 Regression analysis between Band UBC-149-1430 and GCA effects for each of traits in ten influential cotton varities

Trait[a]	*df*	Regression coefficient (b1)	Standard error	T for Ho: parameter = 0	*P*
SCY	1	–355.176	101.687	–3.493	0.0082
LY	1	–171.374	49.869	–3.436	0.0089
B/P	1	–1.148	0.282	–4.073	0.0036
BW	1	0.389	0.136	2.924	0.0192
LP	1	–0.013	0.007	–1.805	0.1087
SI	1	1.284	0.383	3.351	0.0101
LI	1	0.384	0.182	2.108	0.0681
UHM	1	0.376	0.285	1.320	0.2234
E_1	1	0.014	0.164	0.087	0.9331
Str.	1	0.896	0.433	2.070	0.0722
UI	1	0.214	0.179	1.195	0.2664
Mic.	1	–0.196	0.059	–3.319	0.0106

[a] SCY = seed cotton yield, LY = lint yield, B/P = bolls/plant, BW = boll weight, LP = lint percentage, SI = seed index, LI = lint index, UHM = upper half mean fiber length (mm), E_1 = fiber elongation, Str. = fiber bundle strength (g/tex), UI = fiber length uniformity index, Mic. = micronaire

present. Similarly, the presence of the DNA band appears to be associated with positive GCA effects for the seed index because four of the five varieties with the band had positive GCA estimates, while the absence of the DNA band was related with negative GCA effects for this trait as shown by negative GCA effects in four of five varieties without the band. When the data on GCA effects for fiber properties were analyzed, no clear relationships were observed between GCA effects for these traits and the DNA band, except for the relation of GCA effects for micronaire with the band. Among the five varieties that did not have band UBC-149-1430, four showed positive GCA effects for micronaire and one ('Coker 100W') had a negative GCA estimate. In contrast, four of the five varieties containing the DNA band expressed negative GCA effects for micronaire and the remaining one ('CA 614') had a positive GCA estimate. Therefore, a negative relationship occurred between the band and GCA effects for micronaire.

To determine the magnitude and significance of the relationships between the RAPD marker UBC-149-1430 and GCA effects for the traits investigated in this research, simple regression analysis was conducted with the banding patterns serving as the independent variable and GCA-effect estimates being the dependent variable. The regression coefficient (b1) was tested for its significance using a *t*-test (Table 4). The regression coefficients for seed cotton yield and lint yield were –355.176 and –171.374, respectively, which were highly significant (T = –3.493 and –3.436 and P = 0.0082 and 0.0089, respectively). Thus, there are strong negative relationships between the presence of the band and GCA effects for seed cotton yield, and between the band and GCA effects for lint yield. When a variety had the DNA fragment, it tended to produce F_1 hybrids with decreased seed cotton and lint yield; if the variety did not have this fragment, it tended to produce progenies with increased seed cotton and lint yield. The regression coefficient for the number of bolls per plant was also negative (–1.148) and highly significant (T = –4.073, P = 0.0036), indicating that the presence of the DNA fragment in a variety was associated with a decreased number of bolls per plant in the progenies of this variety. A positive relationship between GCA effects for boll weight and the band was confirmed by the regression coefficient of 0.389 which was significant (T = 2.924, P = 0.0192). The relationship between GCA effects for seed index and the DNA band was also positive (b_1 = 1.284) and significant (T = 3.351, P = 0.0101). Regression coefficients for two other yield components, lint percentage and lint index, were –0.013 and 0.384, respectively, and not significant at the P = 0.05 level. Regression analysis indicated that only one fiber trait, i.e., micronaire, had a significant association with the DNA marker. The regression coefficient for micronaire was –0.196, which was significant (T = –3.319, P = 0.0106), suggesting that a variety with the DNA fragment tended to produce progeny with a decreased micronaire. Simple regression analysis was also conducted to analyze relationships between each of the other 41 RAPD markers and GCA estimates for traits, but no relationships were detected at the 0.05 level.

Table 5 Correlation coefficients between Band UBC-149-1430 and each of the traits in 45 F_1s derived from a half diallel using the ten influential cotton varieties as parents

Trait[a]	Correlation coefficient (*r*)	*P*
SCY	–0.569	0.0028
LY	–0.657	0.0011
B/P	–0.743	0.0003
BW	0.498	0.0107
LP	–0.341	0.0851
SI	0.632	0.0015
LI	0.406	0.0537
UHM	0.234	0.3245
E_1	0.092	0.7924
Str.	0.326	0.1023
UI	0.243	0.3146
Mic.	–0.473	0.0142

[a] SCY = seed cotton yield, LY = lint yield, B/P = bolls/plant, BW = boll weight, LP = lint percentage, SI = seed index, LI = lint index, UHM = upper half mean fiber length (mm), E_1 = fiber elongation, Str. = fiber bundle strength (g/tex), UI = fiber length uniformity index, Mic. = micronaire

The correlation analysis conducted in the F_1s detected no relationships between each of the traits and each of the polymorphic DNA fragments except for the marker UBC-149-1430. The presence of the marker UBC-149-1430 was significantly negatively correlated with seed cotton

yield, lint yield, number of bolls per plant and micronaire, but was significantly positively correlated with boll weight and seed index (Table 5). These relationships were the same as those detected in the ten parents. The sample size in the correlation analysis was much larger (n = 45). The information obtained in both the parents and the F_1s indicate that the marker is associated with some of the economically important traits in cotton.

One purpose of developing molecular markers in crop species is to identify and map genes governing economically important traits. To locate genes, a mapping population such as a F_2 population, a backcross or testcross population, double-haploid lines and near-isogenic lines is necessary. Our study showed that the RAPD marker UBC-149-1430 was associated with several traits. The chromosome region containing the marker appears to be important in influencing the yield in cotton. However, some critical questions remain to be answered: what chromosome region is the marker located in, how is the marker close to QTLs (genes) for those traits, and how much variation does the marker explain for the traits? Research is underway to address these issues in the laboratory where the present study was carried out.

Acknowledgements We thank Dr. M.K. Altaf for help in extraction of the DNA. The experiments comply with the current law of the country in which the experiments were performed.

References

Altaf MK, Stewart MJ, Wajahatullah MK, Zhang J (1997) Molecular and morphological genetics of a trispecies F_2 population of cotton. Proc Beltwide Cotton Conf, pp 448–452

Bowman DT, May OL, Calhoun DS (1995) Contribution of various breeding programs to cultivars released between 1970 and 1990. Proc Beltwide Cotton Prod Res Conf, p 506

Bowman DT, May OL, Calhoun DS (1996) Genetic base of upland cotton cultivars released between 1970 and 1990. Crop Sci 36:577–581

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

Cao D, Oard JH (1997) Pedigree and RAPD-based DNA analysis of commercial U.S. rice cultivars. Crop Sci 37:1630–1635

Castagna R, Gnocchi S, Perenzin M, Heun M (1997) Genetic variability of the wild diploid wheat *Triticum urartu* revealed by RFLP and RAPD markers. Theor Appl Genet 94:424–430

Cox TS, Murphy JP, Rodgers DM (1986) Changes in genetic diversity in the red winter wheat regions of the United States. Proc Natl Acad Sci USA 83:5583–5586

Culp TW (1998) Public breeding in the southeast. Proc Beltawide Cotton Conf, pp 493–519

Delannay X, Rodgers DM, Palmer RG (1983) Relative genetic contributions among ancestral lines to North American soybean cultivars. Crop Sci 23:944–949

Demeke T, Lynch DR, Kawchuk LM, Kozub GC, Armstrong JD (1996) Genetic diversity of potato determined by random amplified polymorphic DNA analysis. Plant Cell Rep 15:662–667

Gerdes JT, Tracy WF (1993) Pedigree diversity within the Lancaster Surecrop Heterotic group of maize. Crop Sci 33:334–337

Heun M, Helentjaris T (1993) Inheritance of RAPDs in F_1 hybrids of corn. Theor Appl Genet 85:961–968

Iqbal MJ, Aziz N, Saeed NA, Zafar Y, Malik KA (1997) Genetic diversity evaluation of some elite cotton varieties by RAPD analysis. Theor Appl Genet 94:139–144

Jaccard P (1908) Nouvelles recherches sur la distribution florale. Bull Soc Vaud Sci Nat 44:223–270

Knauft DA, Gorbet DW (1989) Genetic diversity among peanut cultivars. Crop Sci 29:1417–1422

Lu H, Myers GO (1998) Combining ability for yield and yield components in influential cotton varieties. Proc Beltwide cotton conf. (Abstract). p 493

Myers GO, Lu H (1998) Combining ability for fiber properties in influential Upland cotton varieties. Proc World Cotton Research Conf Athens, Greece, pp 173–176

May OL, Bowman DT, Calhoun DS (1995) Genetic diversity of U.S. upland cotton cultivars released between 1980 and 1990. Crop Sci 35:1570–1574

Multani DS, Lyon BR (1995) Genetic fingerprinting of Australian cotton cultivars with RAPD markers. Genome 38:1005–1008

Rodgers DM, Murphy JP, Frey KJ (1983) Impact of plant breeding on the grain yield and genetic diversity of spring oats. Crop Sci 23:737–740

Smith JSC (1988) Diversity of United States hybrid maize germplasm: isozymic and chromatographic evidence. Crop Sci 28:63–69

Specht JE, Williams JE (1983) Genetic technology contributions to soybean productivity – retrospect and prospect. In: Fehr WR (ed) Genetic contributions to yield gains of major crop plants. CSSA special publ 7, AC 322.43. ASA and CSSA, Madison, Wisconsin

Swofford DL, Olson GJ (1990) Phylogeny reconstruction. In: Hillis DM, Moritz C (eds) Molecular systematics. Sinauer Associates, Sunderland, Massachusetts

Tatineni V, Cantrell RG, Davis DD (1996) Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPDs. Crop Sci 36:186–192

Wendel JF, Brubaker CL, Percival AE (1992) Genetic diversity in *Gossypium hirsutum* and the origin of Upland cotton. Am J Bot 79:1291–1310