Theor Appl Genet (1996) 92:915–919

Z. W. Shappley · J. N. Jenkins · C. E. Watson Jr
A. L. Kahler · W. R. Meredith

Establishment of molecular markers and linkage groups in two F_2 populations of Upland cotton

Received: 11 September 1995 / Accepted: 10 October 1995

Abstract Two F_2 populations of cotton (*Gossypium hirsutum* L.) from the crosses of HS46 × MARCABUCAG8US-1-88 (MAR) and HS46 × Pee Dee 5363 (PD5363) were characterized for restriction fragment length polymorphisms (RFLPs) using DNA probes. Seventy-three probe/enzyme combinations were used in the HS46 × MAR population analysis, which resulted in 42 informative polymorphic fragments. These 42 moleclar markers represented 26 polymorphic loci, which consisted of 15 codominant and 11 dominant (+/−) genotypes. Chi-square analyses of these loci fit expected genotypic ratios of 1:2:1 and 3:1, respectively. An analysis of these loci with the MAPMAKER program resulted in the establishment of four linkage groups *A*, *B*, *C*, and *D* with 4, 2, 2, and 2 loci, respectively, as well as 16 unlinked loci. Six probe-enzyme combinations were assayed on the HS46 × PD5363 population, which resulted in 11 informative polymorphic fragments. These 11 fragments represented 6 polymorphic loci, 1 dominant (+/−) and 5 codominant genotypes. The MAPMAKER analysis of these loci yielded 2 linked loci. Thus, a total of 53 polymorphic fragments and 32 polymorphic loci, representing five linkage groups, were identified among the two families.

Key words *Gossypium hirsutum* L. · Restriction fragment length polymorphisms · Linkage map

Contribution of the USDA-ARS in cooperation with the Miss Agric For Exp Stn.

Communicated by G. Wenzel

Z. W. Shappley
Department of Plant and Soil Sciences, Mississippi State University, Mississippi State, MS 39762, USA

J. N. Jenkins (✉)
USDA, ARS, CSRL, P.O. Box 5367, Mississippi State, MS 39762, USA

C. E. Watson, Jr.
Department of Experimental Statistics, Mississippi State, MS 39762, USA

A. L. Kahler
Biogenetic Services, Inc., Brookings, SD 57006, USA

W. R. Meredith
USDA, ARS, Cotton Physiology and Genetics Research Unit, Stoneville, MS 38776, USA

Introduction

With the advent of molecular marker technology, great advances have been made in gathering genetic information about plant, animal, and prokaryotic genomes. Increased information concerning gene identification and map placement would facilitate enhancement of desired agronomic traits. Knowledge of the inheritance and linkage relationships of molecular markers and traits of economic importance provides scientists with useful molecular tools that may be used to improve the efficiency of crop improvement programs.

Restriction fragment length polymorphisms (RFLPs) of DNA have been found to be useful in several crops, including corn (*Zea mays* L.) (Melchinger et al. 1991) and soybean (*Glycine max* (L.) Merr.) (Skorupska et al. 1993). Within the last 3 years, an intense effort has been made to apply RFLP technology to several species of cotton (Meredith 1992; Wang et al. 1993; Stelly 1993; Cantrell and Davis 1993; Patterson 1993; Wing 1993; Reinisch et al. 1994; Shappley 1994). These projects have been fueled by an attempt to map the cotton genome and establish molecular markers for yield components and fiber quality genes.

Previous work has shown that RFLP technology is feasible for use in Upland cotton comparisons (Shappley et al. 1993). Parent selections were made based on this prior research, and F_2 populations were established. The objective of the study presented here was to determine the inheritance relationships of RFLPs and linkage relationships among the observed RFLP loci in two Upland F_2 populations of cotton.

Materials and methods

Greenhouse procedures

One hundred F_2 plants from each of the crosses HS46 × MARCABUCAG8US-1-88 (MAR) and HS46 × Pee Dee 5363 (PD5363) (Green et al. 1990) were grown in the greenhouse at the USDA-ARS research facility at Mississippi State, Mississippi. Twenty grams of wet leaf tissue was collected from each plant. RFLP analyses were conducted by Biogenetic Services, (BGS), Brookings, South Dakota.

Probe construction

The initial cDNA library was developed by BGS using leaf material from six different cotton varieties including Acala, Coker, Deltapine, Pee Dee, and Texas types. Total cellular RNA was isolated from leaf tissue of each variety and prepared according to the methods of MacDonald et al. (1987). Poly(A)-RNA was purified and isolated according to the methods of Aviv and Leder (1972). The preparations were pooled and oligo-(dT) chromatography was performed to isolate polyadenylated RNA. This poly(A)-RNA-isolated mRNA was used as a template for double-stranded cDNA (ds-cDNA) synthesis (Gubler and Hoffman 1983). A dT-tailed *Not*I primer/adaptor oligonucleotide was used to prime a reverse transcriptase enzyme reaction. *Eco*RI adaptors were then joined to the 5′ ends of the ds-cDNA. Ligation and transformation techniques were followed according to Hanahan (1983). Following *Not*I digestion and size fractionation steps, the cDNA was inserted into the *p*GEM-11zf(−) vector (Promega). Multiple copies of the vector were generated via the polymerase chain reaction (PCR) (Saiki et al. 1988). Selections of transformants were based on Amp^R and Lac^-. Radiolabelling was done by random priming (Feinberg and Vogelstein 1983). The methodology was slightly modified or adapted from that given in the literature for use by Biogenetic Services. (unpublished).

RFLP analysis

Freeze-dried leaf material from individual samples was ground into a fine powder. DNA was isolated from each plant sample and purified by sedimentation until equilibrium in CsCl density gradients. In addition to individual F_2 plants, the analysis included bulk samples of each parental line. Techniques for DNA isolation and other RFLP methods were similar to those reported in "Current protocols in molecular biology" (Ausbel 1987) with slight modifications by BGS (Biogenetic Services, unpublished). Either *Eco*RI or *Eco*RV was used as the restriction enzyme in the digest of the sample DNA, depending on the probe used. Gel electrophoresis was conducted with a TAE buffer. Southern blots were made on nitrocellulose membranes, and the resulting complex of fragments was subjected to hybridization techniques (Budowle and Baechtel 1990). After washing the membranes to remove the unhybridized probe solution, the membranes were placed on Kodak AR Safety Film and incubated for 7–14 days depending on optimal exposure time. The radiolabelled DNA fragments on the resulting autoradiograph were digitized using the DIGIGEL software package with phage lambda DNA as a standard. The autoradiographs were scored for the presence or absence of polymorphic fragments. Polymorphic fragments were analyzed with the MAPMAKER program (Lander et al. 1987) to establish molecular marker linkage groups.

Mapmaker analysis

The MAPMAKER (1993) program was acquired from the Whitehead Institute (Cambridge, Mass.) and run on a Unix system at Mississippi State, Mississippi. The MAPMAKER program (Lander et al. 1989) estimates a maximum likelihood distance (Haldane and Smith 1947; Morton 1955; Mather 1957) and an LOD score (Morton 1955) between two loci for a given number of F_2 individuals based on the presence or absence of RFLP fragments in relation to the parental lines. Codominant fragment patterns and dominant (+/−) fragment patterns identified by a specific probe/enzyme combination were designated as individual loci.

The maximum likelihood estimate used in the program is derived from an algorithm by Lander and Green (1987). LOD score indicates the odds of linkage between two loci when they are separated by their maximum likelihood distance. According to Morton (1955) LOD is more specifically defined as the $\log_{10}$ of the ratio of the likelihoods when the loci are taken to be at their maximum likelihood recombination fraction and when the loci are considered unlinked.

The LOD score parameter was set at the program default of 3.00 and the distance parameter was set at 50 cM. This level gives linkage values at a minimum confidence level of 10^3:1. This significantly lessens the chances of the program reporting false linkage. Multipoint and three-point analyses were used in determining the order of the loci in a particular linkage group (Morton et al. 1986).

Results and discussion

HS46 × MAR

Ninety-eight individual F_2 plants from the HS46 × MAR cross were analyzed with 73 probe/enzyme combinations, which resulted in 53 polymorphic fragments. Forty-two of these fragments were considered informative for the linkage analysis based on the fact that either codominant or single fragment patterns were identified. These two types of fragment patterns were easily characterized for use in the MAPMAKER analysis. The 42 polymorphic fragments, or molecular markers, constituted 26 polymorphic loci. All of the loci fit the expected genotypic ratios of either 1:2:1 or 3:1 as determined by chi-square analyses (Table 1). Fifteen loci showed codominance of their restriction fragments. The remaining 11 loci were termed dominant (+/−) loci because they showed a single fragment versus no fragment pattern.

Pairwise linkage was calculated for all possible two-locus combinations of the 26 loci using the MAPMAKER program (Lander et al. 1987). The pairwise analyses suggested four linkage groups *A*, *B*, *C*, and *D* with 4, 2, 2, and 2 loci, respectively (Table 2). CentiMorgan (cM) distances were determined for each locus pair.

Three-point and multi-point analyses were conducted to determine the map order of the four linked loci in group A (Table 3). According to the log-likelihood estimates in the multi-point analysis, the most likely order calculated was only slightly better than the second-best order (0.00 and − 0.03, respectively). This minute difference was due to the tight linkage of loci C26D5RV and C34F5RV, suggested to be 0.51 cM. Thus, either order of *A* could be possible (Fig. 1). The three other linkage groups (*B*, *C*, and *D*) (Fig. 1.) can also be constructed. The relative distances between loci of each linkage group also were calculated in centiMorgans using the maximum likelihood method (Fig. 1).

HS46 × PD5363

An F_2 population of 99 individuals of this cross was analyzed with 6 different probe/enzyme combinations

Table 1 Genotypic distribution, expected ratios, chi-square (χ^2) values, and genotypes among 26 loci for 98 F_2 plants of HS46 × MARCABUCAG8US-1-88 and 6 loci for 99 F_2 plants of HS46 × PD5363

Locus	Fragment size (kb)	Observed ratio	Expected ratio	χ^2	Probability	Genotype[a]
		HS46 × MARCABUCAG8US-1-88				
C05E5RV	8400 4400	24:45:29	1:2:1	1.16	0.75–0.50	Codom.
C05F4RI	6400 3500	24:53:21	1:2:1	0.84	0.75–0.50	Codom.
C13B1RV	8000 7000	25:51:22	1:2:1	0.35	0.90–0.75	Codom.
C14E1RI	8000	79:19	3:1	1.65	0.50–0.25	+/−
C16A1RI	3500	74:24	3:1	0.01	0.95–0.90	+/−
C16F4RV	3200 2200	27:48:23	1:2:1	0.37	0.90–0.75	Codom.
C17A6RI	6400	66:22	3:1	1.02	0.50–0.25	+/−
C18A4RI	6000	74:24	3:1	0.01	0.95–0.90	+/−
C18A4RV	14000 8000	18:55:25	1:2:1	2.47	0.50–0.25	Codom.
C24A4RI	5000 4700	19:52:27	1:2:1	1.67	0.50–0.25	Codom.
C26B5RV	8000	72:26	3:1	0.12	0.75–0.50	+/−
C26D5RV	9000 7000	26:53:19	1:2:1	1.65	0.50–0.25	Codom.
C27B6RI	9400 8000	15:56:27	1:2:1	4.94	0.10–0.050	Codom.
C33A6RV	6200	66:32	3:1	3.06	0.10–0.050	+/−
C33A6RI	3500	78:20	3:1	1.10	0.50–0.25	+/−
C34F5RV	8500 4500	19:54:25	1:2:1	1.76	0.50–0.25	Codom.
C38E2RI	5800 1400	19:51:28	1:2:1	1.82	0.50–0.25	Codom.
C41E4RI	8200 3600	25:54:19	1:2:1	1.76	0.50–0.25	Codom.
C41F5RI	5000 4500	19:51:28	1:2:1	1.82	0.50–0.25	Codom.
C42E4RI	13000 9600	28:51:19	1:2:1	1.82	0.50–0.25	Codom.
C44D2RI	14000 12000	22:51:25	1:2:1	0.35	0.90–0.75	Codom.
C47B1RV	2800	72:26	3:1	0.12	0.75–0.50	+/−
C50C1RI	4400	74:24	3:1	0.01	0.95–0.90	+/−
C51D3RV	12000	72:26	3:1	0.12	0.75–0.50	+/−
C53A6RV	7900	76:22	3:1	0.34	0.75–0.50	+/−
C44B4RV	13000 11000 2500	25:47:26	1:2:1	0.18	0.75–0.50	Codom.
		HS46 × PD5363				
C14E1RI	8000	80:19	3:1	1.78	0.25–0.10	+/−
C20A5RV	9000	18:42:39	1:2:1	11.18	$P < 0.005$	Codom.
C25D5RI	4000 3000	17:53:29	1:2:1	3.40	0.25–0.10	Codom.
C36A5RI	11000 9000	21:56:22	1:2:1	1.73	0.50–0.25	Codom.
C38C2RI	9000 7000	22:55:22	1:2:1	1.22	0.75–0.50	Codom.
C41E4RI	8000 7000	24:44:31	1:2:1	2.21	0.50–0.25	Codom.

[a] Codom., Codominant fragment patterns

resulting in 11 polymorphic fragments representing 6 polymorphic loci. Of these 6 loci, 5 had codominant alleles (Table 1). All but one set of the fragments segregated in the F_2 in an expected 1:2:1 genotypic ratio according to a chi-square analysis (Table 1). In this cross, MAPMAKER suggested that C25D5RI and C38C2RI were linked (Fig. 1) with a map distance of 24.5 cM and an LOD score of 8.97.

A total of 53 molecular markers were established from the analyses of the two populations. Five linkage groups, *A*, *B*, *C*, *D*, and *E*, were constructed from the MAPMAKER analysis. One of these groups (*A*) contained 4 loci, while the remaining four had 2 each. According to three-point and multi-point data, two orders of the loci in group *A* are possible. Additional data will be needed to distinguish the correct order.

Estimates of relative distances were also calculated between the loci within these five groups. The map distances for *A*, with its two possible orders, were 11.7 and 11.3 cM. Group *B* and group *C* gave small distances

Table 2 Pairwise linkage comparisons of loci and centiMorgan (cM) distances for four linkage groups of HS46 × MARCABUCAG8US-1- 88 and one linkage group of HS46 × PD5363

Loci X	Loci Y	LOD score	Distance in cM
HS46 × MARCABUCAG8US-1-88			
Linkage group A			
C05F4RI	C26D5RV	27.46	5.55
C05F4RI	C34F5RV	28.44	4.95
C05F4RI	C50C1RI	12.77	8.89
C26D5RV	C34F5RV	40.32	0.51
C26D5RV	C50C1RI	14.96	6.30
C34F5RV	C50C1RI	15.92	5.23
Linkage group B			
C24A4RI	C41F5RI	40.92	0.51
Linkage group C			
C18A4RV	C18A4RI	21.54	1.08
Linkage group D			
C33A6RV	C33A6RI	4.11	5.11
HS46 × PD5363			
Linkage group E			
C25D5RI	C38C2RI	8.97	24.5

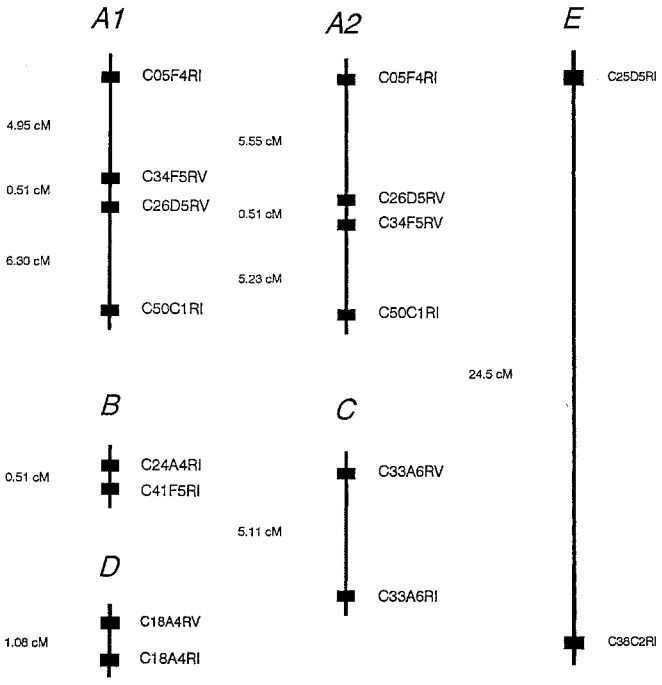

Fig. 1 *G. hirsutum* RFLP linkage groups, including centiMorgan distances, between loci of two possible orders of *A* (*A*1 and *A*2) and individual orders of *B*, *C*, *D*, and *E*. Groups *A*, *B*, *C*, *D* were detected in HS46 × MAR F_2s, and group *E* was detected in HS46 × PD5363 F_2s. There are 26 cotton chromosomes. RFLP marker loci are represented by ■.

between their two individual markers of 0.5 and 1.0 cM, respectively, while group *D* gave a distance of 5.2 cM between its loci. The largest distance was found in group *E* from the HS46 × PD5363 population, whch had a distance of 24.5 cM. All of the map distances cover a relatively small part of the cotton genome, which is estimated to span a minimum of 5125 cM (Reinisch et al. 1994).

This research demonstrates the availability of ample restriction fragment length polymorphisms of cDNA in the Upland cotton genome. While this is a preliminary map, further investigations of this type with additional RFLPs are increasing the working knowledge of the genetics of Upland cotton. Ongoing research with these

Table 3 Multi-point and three-point analyses of linkage group *A*, identified from HS46 × MAR F2's, as calculated by MAPMAKER

Multi-point	Order of loci				Log-likelihood
	C05F4RI	C34F5RV	C26D5RV	C50C1RI	0.00
	C05F4RI	C26D5RV	C34F5RV	C50C1RI	−0.03
	C26D5RV	C34F5RV	C05F4RI	C50C1RI	−2.19
	C34F5RV	C26D5RV	C05F4RI	C50C1RI	−3.17
	C26D5RV	C34F5RV	C50C1RI	C05F4RI	−4.74
	C34F5RV	C26D5RV	C50C1RI	C05F4RI	−5.78
	C05F4RI	C34F5RV	C50C1RI	C26D5RV	−8.71
	C34F5RV	C50C1RI	C26D5RV	C05F4RV	−9.69
	C26D5RV	C05F4RI	C34F5RV	C50C1RI	−11.91
	C34F5RV	C05F4RI	C26D5RV	C50C1RI	−12.86
	C34F5RV	C50C1RI	C05F4RI	C26D5RV	−17.60
	C34F5RV	C05F4RI	C50C1RI	C26D5RV	−17.66
Three-point					
	C05F4RI	C34F5RV	C26D5RV		0.00
	C05F4RI	C26D5RV	C34F5RV		−0.98
	C26D5RV	C05F4RI	C34F5RV		−12.86
	C05F4RI	C26D5RV	C50C1RI		0.00
	C26D5RV	C05F4RI	C50C1RI		−2.19
	C05F4RI	C50C1RI	C26D5RV		−4.80
	C05F4RI	C34F5RV	C50C1RI		0.00
	C34F5RV	C05F4RI	C50C1RI		−3.14
	C05F4RI	C50C1RI	C34F5RV		−5.69
	C26D5RV	C34F5RV	C50C1RI		0.00
	C34F5RV	C26D5RV	C50C1RI		−0.95
	C26D5RV	C50C1RI	C34F5RV		−9.66

and other RFLPs involves the association of RFLPs with quantitative trait loci that are agronomically important to the cotton industry. We are currently seeking molecular markers for genes that control fiber characteristics, earliness, plant physiology, and yield using data from the HS46 × MAR population in the F_2 through F_5 generations.

Acknowledgements We would like to thank Dr. E. Butler III for developing the cDNA library and Dee French for conducting the RFLP assays at Biogenetic Services.

References

Ausbel F (ed) (1987) Current protocols in molecular biology. Greene Publ, New York

Aviv H, Leder P (1972) Purification of biologically active globin messenger RNA by chromatography on oligothymidylic acid-cellulose. Proc Natl Acad Sci USA 69:1408–1412

Budowle B, Baechtel FS (1990) Modifications to improve the effectiveness of restriction fragment polymorphism typing. Appl Theor Electrophoresis 1:181–187

Cantrell RG, Davis DD (1993) Characterization of *hirsutum* × *barbadense* breeding lines using molecular markers. In: Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Natl Cotton Council Am, Memphis, Tenn., pp 1551–1553

Feinberg AP, Vogelstein B (1983) A technique for radiolabelling DNA restriction endonuclease fragments to high specific activity. Anal Biochem 137:266–267

Green CC, Culp TW, Kittrell BU (1990) Notice to plant breeders and geneticists relative to release of two cotton germplasm lines. SC Agric Exp Stn Germplasm Release

Gubler U, Hoffman BJ (1983) A simple and very efficient method for generating cDNA libraries. Gene 25:263–269

Haldane JBS, Smith CAB (1947) A new estimate of the linkage between the genes for color-blindness and hemophilia in man. Ann Eugenics 14:10–31

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Lander ES, Green P (1987) Construction of multi-locus linkage maps in humans. Proc Natl Acad Sci USA 84:2363–2367

Lander ES, Green P, Abrahamson J, Barlow A, Daly MJ, Lincoln SE, Newburg L (1989) MAPMAKER: an interactive computer package for constructing primary linkage maps of experimental and natural populations. Genomics 1:174–181

MacDonald RJ, Swift GH, Przybyla AE, Chirgwin JM (1987) Isolation of RNA using guanidinium salts. In: Berger SL, Kimmel AR (eds) Guide to molecular cloning techniques. Methods Enzymol 152:219–223

MAPMAKER (1993) MAPMAKER/EXP 3.0 Beta release "3.0b". January 1993. Whitehead Institute, Cambridge, Mass.

Mather K (1957) The measurement of linkage in heredity. John Wiley and Sons, New York

Melchinger AE, Messmer MM, Lee M, Woodman WL, Lamkey KR (1991) Diversity and relationships among U.S. maize inbreds revealed by restriction fragment length polymorphisms. Crop Sci 31:669–678

Meredith WR Jr (1992) RFLP association with varietal origin and heterosis. In: Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Natl Cotton Council Am, Memphis, Tenn., p 607

Morton NE (1955) Sequential tests for the detection of linkage. Am J Hum Genet 7:277–318

Morton NE, MacLean CJ, Lew R, Yee S (1986) Multipoint linkage analysis. Am J Hum Genet 38:868–883

Patterson AH (1993) Molecular markers in cotton improvement. In: Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Natl Cotton Council Am. Memphis, Tenn., p 1557

Reinisch MJ, Dong J, Bruaker CL, Stelly DM, Wendel JF, Patterson AH (1994) A detailed RFLP map of cotton, *Gossypium hirsutum* × *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. Genetics 138 :829–847

Saki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA (1988) Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science 239:487–491

Shappley ZW (1994) Restriction fragment length polymorphisms in cotton (*Gossypium hirsutum* L.): feasibility of use, diversity among plants within a line, and establishment of molecular markers and linkage groups among two F_2 populations. MSc thesis, Mississippi State University, Mississippi State, Miss.

Shappley ZW, Jenkins JN, McCarty JC Jr (1993) Potential for RFLP mapping in cotton. In: Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Natl Cotton Council Am, Memphis, Tenn., p 610

Skorupska HT, Shoemaker RC, Warner A, Shipe ER, Bridges WC (1993) Restriction fragment length polymorphism in soybean germplasm of the southern USA. Crop Sci 33:1169–1176

Stelly DM (1993) Interfacing cytogenetics with the cotton genome mapping effort. In: Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Natl Cotton Council Am, Memphis, Tenn., pp 1545–1550

Wang G, Kowalski S, Altman DB, El-Zik KM, Fryxell PA, Kohel RJ, Smith CW, Thaxton P, Wing RA, Patterson AH (1993) Introgression of agriculturally valuable QTLs from wide crosses in cotton using DNA markers. In: Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Natl Cotton Council Am, Memphis, Tenn., p 1558

Wing RA (1993) Prospects for physical mapping and map-based cloning of agriculturally important genes in cotton. In: Herber D (ed) Proc Beltwide Cotton Prod Res Conf. Natl Cotton Council Am, Memphis, Tenn., 1557