Identification of a RAPD marker linked to a male fertility restoration gene in cotton

(Gossypium hirsutum L.)

Tien-Hung Lan¹, Charles G. Cook² & Andrew H. Paterson¹, *

¹Department of Soil and Crop Sciences, Texas A&M University, College Station, TX 77843-2474
²USDA-ARS, Weslaco, TX 78596

(*author for correspondence, e-mail ahp0918@acs.tamu.edu, fax: 409-845-0456)
KEY WORDS
bulked segregant analysis, cytoplasmic male sterility, near-isogenic line, linkage mapping

ABSTRACT
One RAPD marker, 6 cM away from a gene which restore male-fertility of a male-sterile cytoplasm was found in upland cotton (Gossypium hirsutum L.). This marker was discovered after screening 400 decamers to identify DNA polymorphism between near-isogenic lines and subsequently verified by bulked segregant analysis in an F2 population of 89 individuals derived from a cross between a cytoplasmic-male-sterile line and a restorer line. The RAPD marker was sub-cloned, sequenced, and mapped to a cotton high density RFLP map. The evaluation and utilization of this RAPD for tagging and ultimately cloning the fertility-restoring gene is discussed.
INTRODUCTION

Cytoplasmic male sterility (CMS) is a maternally inherited trait conferring the inability to produce functional pollen because of interaction between cytoplasmic and nuclear genes. Since CMS does not affect female fertility, male sterile plants are able to set seeds as long as viable pollen are provided. The presence of certain nuclear genes, Rf (Restoring fertility), can effectively suppress the male-sterile cytoplasm and restore pollen fertility. The application of CMS/Rf system has proved to be an effective means to produce commercial F1 hybrid seed for many crops (Williams 1992). Cotton is predominantly a self-pollinated crop, however, cotton breeders have long been trying to breed F1 hybrid cotton, to harness F1 heterosis for many desirable traits such as high seedling-vigor, earliness, superior fiber quality and yield (Davis 1978), but also because F1 hybrid seeds could generate huge revenue for the seed industry (Anonymous 1985). Therefore, attempts to produce F1 hybrid cotton on a commercial scale have never stopped (Anonymous 1987). Before the introduction of the CMS/Rf system into Upland cotton, the only way to generate commercial hybrid cotton was through hand emasculation and crossing, which was economical only in China and India where labor cost are low. In US where labor costs are high, hand-crossing made the price of hybrid cotton prohibitive.

The first CMS line of commercial cotton was introduced by crossing an upland cotton, *G. hirsutum*, as male parent, to a wild species, *G. harknessii* (Meyer 1973). A Rf line was also developed by transferring a nuclear restorer gene from *G. harknessii* into *G. hirsutum* simultaneously (Meyer 1975). The F1 hybrid population generated by crossing the CMS line to the Rf line showed a wide range of male fertility expression. Several CMS/Rf lines were then developed through the backcross method (Weaver and Weaver 1977). Later, a second Rf which expresses incomplete dominance, was also identified in *G.
*barbadense* (Sheetz and Weaver 1980). Identification of molecular markers closely linked to the nuclear Rf genes could help breeders to distinguish male-sterile and fertile plants prior to pollen shed. Here we report the identification of molecular markers closely-linked to the nuclear Rf gene under male-sterile cytoplasm in commercial cotton.

**MATERIAL AND METHOD**

A pair of cotton near-isogenic lines, HAF277 and DELCOT277 (kindly provided by R. Bridge, USDA-ARS, Stoneville, MS; Meyer 1975; Weaver and Weaver 1977; Sheetz and Weaver 1980), carrying the Rf gene and CMS respectively were used for initial RAPD-PCR screening. A single primer was used in each PCR reaction, and the PCR products were resolved in 1.6 % agarose gels using established techniques (Williams *et al*. 1990). An F<sub>2</sub> segregating population comprised of 89 individuals from a cross of a CMS line, A2, and a Rf line, B418, were used for bulked segregant analysis (Giovannoni *et al*. 1991, Michelmore *et al*. 1991).

Linkage analysis was performed using MapMaker (Lander *et al*. 1987), Macintosh version 2.0 (kindly provided by S. Tingey, DuPont), using the Kosambi centiMorgan function. A threshold of LOD 3.0 was used to test linkage.

**Result**

**RAPD screening**

After screening 400 random decamers obtained from the University of British Columbia, 15 positive RAPDs were identified, which represent 3.75 % of the tested primers. These positive RAPDs were verified twice, on fresh DNA extractions, to avoid false-
positives. We applied bulked segregant analysis (Giovannoni et al. 1991; Michelmore et al. 1991) for further exploration of the 15 positive RAPD markers identified in the near-isogenic lines. From the A2 x B418 F2 population, we chose 19 individuals that were clearly fertile, and 19 individuals that were clearly sterile to construct the DNA pools, and to avoid individuals that were of doubtful phenotype. To increase the stringency of our test, two pairs of synthetic DNA pools were constructed for each phenotype (fertile and sterile). One pair of pools contained 9 individuals each in the CMS and Rf pool, the other contained 10 individuals each. The pooled DNA was then used as template for RAPD-PCR analysis. Among the fifteen identified markers, two RAPD markers, R6952 (sequence CGGTTTCGTA) and R6861 (sequence CGTGACAGGA), putatively distinguished between male-sterile and male-fertile pools.

Two RAPD markers were then subjected to linkage analysis using the 38 F2 plants used in constructing the synthetic DNA pools. Four replicas were used, and each amplified band was scored as a dominant locus. Linkage analysis show R6861 was 24.6 cM from the Rf gene, and R6592 was 2.3 cM away.

**Sequencing**

The marker R6592 was eluted from the agarose gel, and cloned into the EcoRV site of dTTP-end-filled pBluescript(KS), then sequenced from both ends using the T3/T7 primers on a Applied Biosystem 373 DNA Sequencing System. From the end sequences of R6592, the correct RAPD primer sequence was identified (Fig. 1), and the sequence of R6952 was deposited into Genbank with accession number AF094829 and AF094830. A blast search of R6592 in Genbank did not found any corresponding sequence at 95% confidence level.
Southern Analysis

Genomic DNA of A4 and B418 was digested by BamH1, CfoI, EcoR1, EcoRV, HindIII and XbaI, gel-separated, and blotted to Hybond N+ membrane as described (Reinisch et al. 1994). Gel-isolated R6592 and R6861 were P32 labeled and applied to the blots. However, a smear pattern was observed and no restriction-enzyme-polymorphism could be identified between A4 and B418, suggesting that both R6592 and R6861 contain repetitive sequence. Thus, R6592 was further digested with a mixture of AccI, CfoI, HinfI, HindIII and EcoR1, and sub-cloned into pBluescript to try to remove the interfering repetitive sequence. One clone, R6592a14, identified a HindIII polymorphism between A4 and B418, though the background on the p32-exposed film was still pretty high. R6592a14 was then used to genotype the 89 individuals of A4 x B418 F2 population for the HindIII polymorphism. Linkage analysis revealed the distance between Rf gene and R6592 to be 6 cM (Fig 2a), not significantly different from the earlier estimate based on 38 individuals. Further, we tried to map R6592a14 in a cotton high-density linkage map based on a cross of Gossypium hirsutum and Gossypium barbadense (Reinisch et al. 1994). R6592a14 did not detect HindIII polymorphism in this population, but one of six genomic restriction fragments did detect EcoRV polymorphism. R6592a14 mapped to a linkage group that was tentatively identified as chromosome 20 (Reinisch et al. 1994), between markers pAR959 and pAR3-41 (Fig. 2b).

Discussion

Since the near-isogenic lines we used have been backcrossed for at least 8 generations, the introgressed region should be less than 14.1 cM (Hanson 1959; using the average cotton chromosome length of 200 cM from Reinisch et al. 1994). This is consistent with the distance of R6592a14 to the Rf locus. The larger distance between R6861 and Rf
was unexpected, but R6861 was not extensively verified, so the estimated map distance is based on a small number of individuals.

For the purpose of marker-assisted selection, the best scenario was to find tightly linked markers on both side of the target gene to reduce the risk of mis-genotyping due to single recombination events between the marker and the target gene. More markers linked to the Rf gene would be desirable and could be found either by targeted RAPD or AFLP screens, or in the course of further enrichment of the cotton map. The CMS/Rf is a potentially cost-effective way to produce F1 hybrid seeds. Rf genes have been successfully mapped in rice and common bean by RAPD/bulk segregant analysis (He et al. 1995; Zhang et al. 1997), and we provide a marker diagnostic of this important phenotype in cotton. In addition to its utilization in marker-assisted selection, this marker may serve as a starting point for positional cloning of the Rf gene.

ACKNOWLEDGMENT

We thank Mark D. Burow for technical assistance. Aspects of the work described here were supported by USDA 91-37300-6570, to AHP, and Texas Higher Education Coordinating Board Award 999902-148 to AHP and Rod A. Wing.
Fig. 1
Partial DNA sequence of clone R6592. Primer sequences are underlined.

Fig. 2
a. Linkage of Rf gene and marker R6592 in *G. hirsutum* A2 x B418 F2 population.
b. Marker R6592 mapped to cotton chromosome 20 in *G. hirsutum* (race palmeri) x *G. barbadense* “k101” primary mapping population (Reinisch et al., 1994).
Reference

Hanson, W., 1959  Early generation analysis of lengths of heterozygous chromosome segments around a locus held heterozygous with backcrossing or selfing. Genetics 44: 833-837.


