Gel Based DNA Marker Technologies in Cotton

S. Saha[1], H. Tan[2], M. Karaca[2] and J.N. Jenkins[1]

[1]USDA/ARS, Genetics and Precision Agriculture, Crop Science Research Laboratory, P.O. Box 5367, Mississippi State, MS 39762, USA; [2]Plant and Soil Sciences, Mississippi State University Mississippi State, MS 39762, USA.

Introduction

Genetic mapping with DNA markers has become an essential tool for revealing the genetic basis of both qualitative and quantitative traits in crop plants. DNA markers have special advantages over morphological traits in genetic analysis because they are normally polymorphic, have no pleiotropic or epistatic effects, employ non-destructive methods requiring very small amounts of plant materials and there is no effect of environment on the phenotype. Many of the economically important traits of crop plants are controlled by complex quantitative traits (QTL). An accurate and detailed genetic map of DNA markers linked to QTLs provides an invaluable tool to the breeders to use molecular markers in marker-assisted selection by allowing selection for the markers rather than the complex QTL phenotype. Recently several laboratories in USA and abroad have launched special initiatives on molecular genome mapping programs in cotton to exploit the available gene pools and enhance the improved germplasm resources (Cantrell et al., 1999; Yu and Kohel 1999; Rungis et al. 1999; Cantrell et al., 1998, 1999; He et al. 1999; Wright et al. 1998; Luo et al. 1999; Akbari and Lyon, 1999).

Scientists have had to face three primary challenges in molecular mapping the cotton genome: 1) developing an extraction method to produce high quality DNA suitable for molecular analysis; 2) developing an appropriate method that can quickly score large numbers of polymorphic molecular markers and 3) creating a suitable mapping population that can be used for identifying molecular markers linked to the traits of interest. However, the method for creating the appropriate mapping population varies based on the research interest. The overall objectives of this paper are to address these concerns and also provide a summarized report on the current status of molecular mapping in cotton.

DNA Extraction

High quality DNA is an essential requirement for developing DNA markers in cotton. Cotton tissue is notoriously recalcitrant to many common DNA extraction methods due to high levels of polyphenolic compounds (e.g. gossypol). When cells are disrupted during sample grinding, phenolic compounds interact with proteins and nucleic acids causing problems in molecular analysis. Therefore, special precautions must be taken during DNA extraction from cotton tissues. Several methods have recently been developed for DNA extraction in cotton (Callahan and Mehta 1991; Paterson et al., 1993). However, these methods require either fresh tissue or tissue stored at very low temperatures (-50° to -70°C). Saha et al. (1997) reported a DNA extraction method using freeze-dried cotton tissue. The freeze-dried tissue can be ground as a dry powder and stored in a freezer (-20°C) for an indefinite period. Recently, the method has been modified to be more efficient and less time consuming for large scale DNA analyses using a slightly modified method of Dneasy Plant Mini Kit (Qiagen, Santa Clarita, CA). Freeze-dried crushed cotton leaf tissue powder (35 mg) was used instead of liquid N_2. The remainder of the method was followed as per the manufacturer's protocol. This method yielded high molecular weight DNA that was readily digestible with restriction enzymes and served as a template for PCR amplification.

Molecular Genome Methods

Four different types of molecular methods were used for developing DNA markers in cotton namely Restriction Fragment Length Polymorphisms (RFLPs), Random Amplified Polymorphic DNAs (RAPDs), Simple Sequence Repeats (SSRs) and Amplified Fragment Length Polymorphisms (AFLPs).

Restriction Fragment Length Polymorphisms

RFLP techniques involve restriction enzyme digestion of DNA into small fragments which are electrophoretically separated on an agarose gels and transferred to a solid support or filter (nitrocellulose membrane) and detected as discrete bands following hybridization to a radioisotope or non-radioisotope labeled probes. Alleles are recognized by differences, due to mutation at the restriction sites, in the size of the restriction fragments to which probes have hybridized. The advantages of the RFLP method are that it is normally polymorphic and markers are co-dominant distinguishing heterozygote from the homozygote (Figure 1A). However, the conventional detection of RFLPs by Southern blot analysis is laborious, time consuming, costly and it always requires sufficient quantity of good quality DNA and a probe.

Reinisch et al. (1994) published the first report on a RFLP map using an interspecific tetraploid cultivated *Gossypium* spp. They reported 705 RFLP loci assembled into 41 linkage groups cov-

ering 4675 cM. They also located several RFLP markers to different chromosomes using aneuploid chromosome substitution lines. Shappley et al. (1996) reported 16 RFLP loci arranged in 6 linkage groups in an intraspecific F_2 population of upland cotton. Shappley et al. (1998) presented another report of 120 loci arranged into 31 linkage groups covering 865 cM in upland cotton. Shappley et al. (1998) also reported the first report of RFLP-QTL linkage map using an intraspecific population of upland cotton. They determined the location of 100 QTLs in 24 linkage groups. They also observed that several QTLs are located on the same position in some of the linkage groups. This might be due to several reasons including a master QTL gene controlling several traits, the same genetic effect (QTL) measured in different ways or the QTL loci are strongly linked. Jiang et al. (1998) reported 261 RFLP markers in 27 linkage groups in an RFLP-QTL linkage map using an inter-specific population. They observed that 14 QTLs affecting fiber-related traits were located within the D subgenome group. They mentioned that fiber-related genes had already been fixed in the A sub genome and that the discovery of genes from a non-cultivated genome (D) is a reflection of a genetic revolution during polyploidization in tetraploid cotton. We reported the chromosomal association of several RFLP markers linked to important QTLs in upland cotton (Saha et al., 1998). However, our report indicated more influence of the A sub genome in fiber-related QTLs in the upland cotton demonstrating the importance of further investigation and introgression of diploid A-genome species in improving fiber traits in upland cotton. Our results also showed the portability and strength of RFLP markers that could be used between two different intra and inter-specific cotton populations. The high density genetic map of RFLP markers and QTLs of defined chromosome regions will be very helpful in transferring useful orthologous QTL loci among the tetraploid cultivated species of cotton. Wright et al. (1998) identified RFLP markers linked to the resistance allele for the bacterial blight pathogen (*Xanthomonas campestris* pv. *malvacearum* (*Xcm*).

To complement conventional RFLP markers, molecular markers based upon Polymerase Chain Reaction technology (PCR) have recently been widely used in plants. PCR is done following an enzymatic amplification of genomic DNA that is flanked by two synthetic oligonucleotide primers (3' and 5' end of DNA) that hybridize to complementary strands of the target sequence in template DNA. Repeated cycles of the heat denaturation of the template DNA, annealing of the primers and primer extension by *Taq* polymerase in the PCR reaction result in amplification of the specific segment of the template genomic DNA defined by the synthetic 5' and 3' primers. Normally the banding patterns are scored directly under UV light after electrophoresing the amplified products of PCR, stained with ethidium bromide, in an agarose gel. Polymorphism in the genome is detected due to 1) nucleotide changes that prevent amplification due to mismatch of bases at any one of the priming sites; 2) deletion of a priming site; 3) insertion or deletion in the template DNA that changes the size of the amplified products; and 4) insertion that causes the priming site to be too distant to support amplification. The PCR method has become one of the most popular tools because of its ease, speed, sensitivity and versatility to study large populations. In recent years 3 different types of PCR-based DNA markers have been used in cotton namely: RAPD, SSR and AFLP.

Random Amplified Polymorphic DNAs

Most commonly used PCR-based DNA techniques that fall in this category in which about 10 to 15 base-size oligonucleotide arbitrary primers are used for amplification of DNA segments to develop molecular markers (Figure 1B). The advantage of this method is that it does not require any prior DNA sequence information. Our results using inter and intra-specific tetraploid cotton showed that normally about 8 to 12 DNA markers ranging from 120 to 1000 bp in size per primer combination, are generated using RAPD method (Figure 1B). We observed about 2% to 10% polymorphic markers/primer combination in upland cotton depending on accessions. The disadvantages of using RAPD markers are that they are dominant markers and sometimes produce non-specific DNA markers. Yu et al. (1998) identified several RFLP, RAPD and SSR markers linked to fiber length and other properties in an inter-specific population. Wajahatullah and Stewart (1997) evaluated genomic affinity among the selected *Gossypium* species using RAPD method.

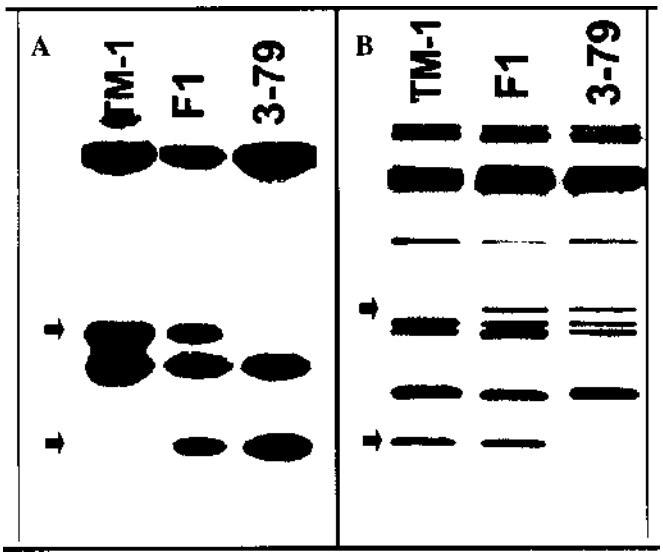

Figure 1A. Arrows indicated the presence of co-dominant RFLP markers in samples of TM-1, 3-79 and a F1. TM1 is an inbred line of *G. hirsutum*, 3-79 is an double haploid line of *G. barbadense* and F1 is a hybrid between TM-1 and 3-79.

Figure 1B. Agarose gel picture of RAPD analysis in TM-1, 3-79 and a F1. Note arrows indicated the presence of polymorphic dominant bands.

Simple Sequence Repeats

SSRs are tandemly repeated di-, tri- or tetra nucleotide loci in a genome. These repeats vary in number and occur in many sites within the genome. The analysis of short tandem repeat (SSR) loci by PCR methods has proven to be very useful because of the high degree of repeat number polymorphisms at many loci within the genome. The development of SSR markers is time consuming and requires sequencing of the genome to identify flanking primers for SSR markers. Dr. Ben Burr's lab at Brookhaven National Laboratory, New York, has developed several hundred SSR primer pairs from the cotton genome which are currently available commercially through Research Genetics, Huntsville, Alabama. SSR primers were also developed from a cotton genomic library enriched in $(GA)_n$ repeats (Reddy et al., 1998). The primary limitations in the use of SSR markers are their cost of development and the presence of non-specific stutter or shadow bands due to the slippage of the *Taq* polymerase during PCR.

Agrawal et al. (1999) reported that cross-species SSR primers could be used to detect polymorphism in tetraploid cotton. We observed that SSR markers can be dominant, similar to RAPD (unpublished data) or co-dominant (Figures 2). Cantrell et al. (1998) identified the chromosomal location of over 50 SSR markers in cotton. They also used this method to develop linkage map of QTL and SSR markers in tetraploid cotton.

Amplified Fragment Length Polymorphisms

AFLP is a PCR-based method of DNA fingerprinting. This method is based on the selective amplification of restriction enzyme digested fragments from genomic DNA using different primer combinations. In comparison to other methods, the AFLP method generates virtually unlimited numbers of DNA fragments from nanogram quantities of genomic DNA. AFLP techniques use normally very stringent PCR conditions, which provide a better reproducibility of the results. In this method, genomic DNAs are digested by two restriction enzymes; subsequently the oligonucleotide adapters, specific to the restriction site, are ligated to the digested DNA fragments to generate template DNA and these template DNAs are amplified using primers complementary to the adapter sequence. Our results indicated that AFLP method can be used to develop both dominant or co-dominant markers in cotton depending on the appropriate primer combination (unpublished data). Typically, detection of AFLP requires radioactive-labeling (Feng et al., 1997) or fluorescence labeling of the primers or silver staining of the gel (Feng et al., 1997). However, in silver staining both strands of the PCR products are detected, resulting in the formation of multiple double-band patterns that may complicate the interpretation of AFLP results. Moreover, we observed that low sensitivity of silver staining sometime limits the power of detected AFLP, especially with DNA fragments in the lower molecular weight range.

Khan et al. (1998) used AFLP and RAPD methods to identify DNA markers linked to three morphological traits. They constructed a linkage map comprising 51 linkage groups covering 6663 cM with 332 AFLPs, 91 RAPDs and 3 morphological traits. Feng et al. (1997) identified the chromosomal location of several AFLP markers. Currently, we are using AFLP markers to identify root-knot nematode resistance genes in cotton (unpublished data).

Next Generation of PCR-based DNA Markers Labeled with Fluorophores

Currently we are using a Capillallary Electrophoretic system (CE) to separate fluorescently-labeled amplified DNA markers that are visualized as peaks on electropherograms using the automated PE-Applied Biosystems ABI PRISM 310 Genetic Analyzer equipped with Genotyping and GeneScan analysis software (PE Applied Biosystem, Foster City, CA; Figure 2A, 2B and 2C). The multiple fluorescent detection system in CE permitted correction of lane to lane variation by co-electrophoresis of an internal size standard DNA labeled with a dye different from that of the sample. Automatic sample loading, digitized output of peak position, automated data collection and analysis and no need for gel preparation makes CE attractive to process large numbers of samples within a very short time for analysis. This method is sensitive enough to detect one/two base pair DNA fragment size differences between alleles of a DNA marker. This feature makes this method more attractive to distinguish a large number of polymorphic markers that can not be detected in a regular agarose gel.

The SSR markers were received from Research Genetics, Huntsville, AL and the PCR method was used as per manufacturer's protocol in the presence of fluorescently labeled base or primer. The AFLP analysis was done using AFLP Selective Amplification Module for large plant genome from PE Applied Biosystem (Foster City, CA 94404).The overall AFLP method was performed according to the manufacturer's protocol.

It was observed that SSR markers in CE can produce about 3-4 times more DNA markers of cotton in comparison to detection on agarose gel. This non-radioactive method requires very small, nanogram amounts of amplified product, very short time for sample analysis and accordingly very fast and cost-effective. This method also provides the capacity to visualize different amplified products of the same sample in one sample from multiplexing different primers labeled with multicolor fluorescence labeling techniques. DNA marker system in cotton using CE is an effective tool replacing gel electrophoresis. It provides high speed, sensitivity and reproducibility because it requires less human intervention and the automation of the system. The advantage of automated computer data storage and analysis of the sample for DNA marker makes it easy to analyze a large number of samples. Currently we are using CE routinely to screen a large number of germplasms for SSR and AFLP analyses (Figure 2). However, this CE system is only

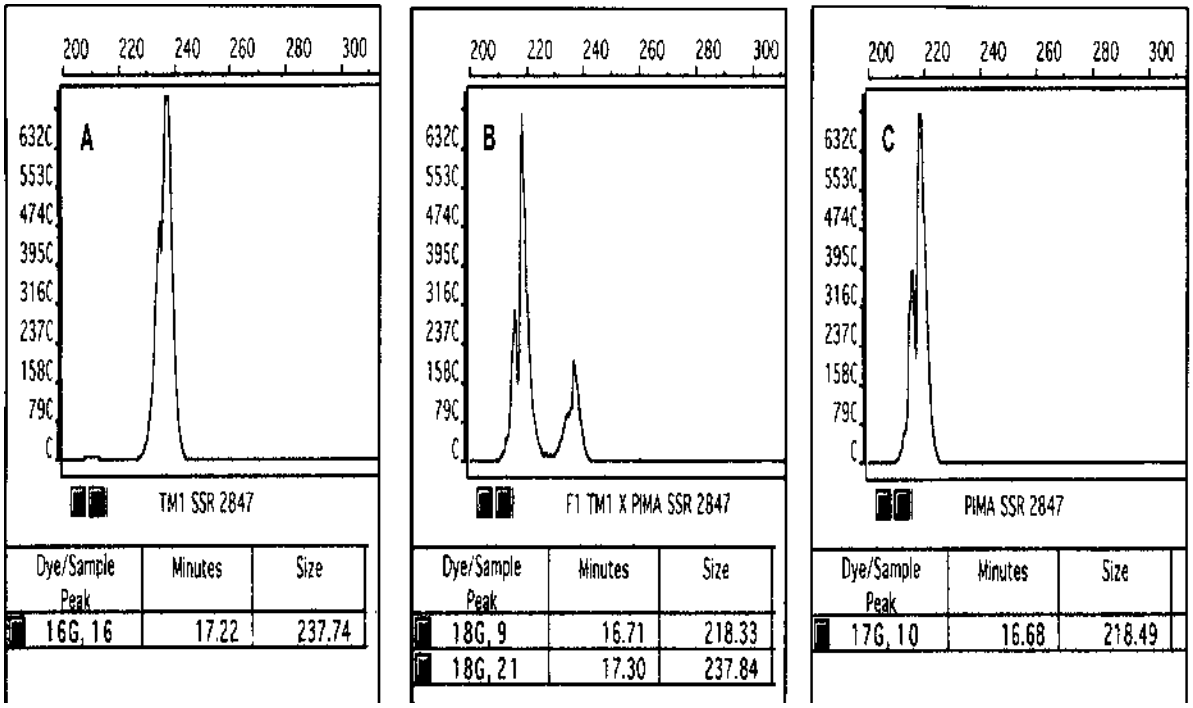

Figure 2. Automated Capillary Electrophoretic System (CE) showing polymorphic fluorescently-labeled amplified SSR markers which are visualized as peaks on electropherograms. X axis represented the size of the DNA fragment and the y-axis showed the amount of the amplified products. Note the presence of co-dominant DNA markers in F1 (B) showing the presence of same size allelic DNA fragments of TM-1 (A) and Pima (C), the two parents of the F1. The automated system provides also the table showing the size of different DNA fragments.

efficient at detecting DNA fragments ranging from 40 to 500 bp in size.

Preliminary results from CE detected about 25 polymorphic markers/AFLP primer combination at the inter-specific level of the tetraploid cotton (Table 1). Seven polymorphic SSR markers at the inter-specific level in comparison to 5 polymorphic SSR markers at the intra-specific level of the tetraploid cotton (Table 1) were also observed.

In conclusion, results demonstrated that AFLP method detected more polymorphic markers in comparison to the SSR technique (Table 1). Accordingly, it is suggested that the AFLP method can be used as a very useful tool for germplasm screening in cotton. Automated fluorescently-labeled amplified DNA markers using PE Applied Biosystem ABI 310 may be an alternative replacement for gel based DNA marker for screening large number of samples in cotton.

Acknowledgements

A word of appreciation to Mr. D. Dollar for his help in DNA extractions and also to Drs. Allan Zipf and Jack C. McCarty for their helpful suggestions on the manuscript.

Table 1. **Polymorphic DNA Markers in Cotton**

	Number of primer combinations		Number of DNA markers		Number of polymorphic DNA markers		Number of DNA markers /primer combination		Polymorphic DNA markers /primer combination		% Polymorphism	
Marker type	inter-specific	intra-specific	inter-specific	intra-specific	inter-specific	intra-specific	inter-specific	intra-specific	inter-specific	intra-specific	inter-specific	intra-specific
AFLP	10	10	607	631	256	166	60.7	63.1	25.6	16.6	42.17	26.30
SSR	10	10	150	130	79	53	15	13	7.9	5.3	52.66	40.77

This preliminary survey was conducted with the lines of TM-1, HS-46, and MARCABUCAG8US-1-88 (*G. hirsutum*) and 3-79 (*G. barbadense).*

References

Agrawal, D., Saha, S., Jenkins, J. N., Zipf, A. and May, L. 1998. DNA fingerprinting in cotton using SSR primers from diverse plant species. *J. of New Seeds* (in press).

Akbari, M., and Lyon, B. R. 1999. The use of AFLP technology for the identification of molecular markers in cotton (*Gossypium hirsutum* L). Plant and Animal Genome VII Conference. San Diego. January 17-21.

Callahan, F.E. and Mehta, A.M. 1991. Alternative approach for consistent yields of total genomic DNA from cotton (*Gossypium hirsutum* L.). *Plant Mol Biol. Rep.* 9(3):252-261.

Cantrell, R.G., Ullola, M., Liu, S. and Pederson, J. 1998. Concept of graphical genotypes for analyzing introgressed cotton germplasm. *Abstracts of the World Cotton Research Conference-2*, September 6-12, Athens, Greece. p27.

Cantrell, R. G., Pederson, J., and Liu, S. 1999. Mapping of introgeressed cotton population with DNA Markers. Plant and Animal Genome VII Conference. San Diego. January 17-21. W56.

Feng, X., Saha, S., and Soliman, K. M. 1997. AFLP markers in cotton. *Proc. Beltwide Cotton Conf.*, New Orleans, LA, 6-10 Jan. 1998. Natl. Cotton Council Am., Memphis, TN. p 483.

He, L., Bridges, A., Robinson, F., Cook, C. C., and Zhang, H-B. 1999. Toward map-based cloning of root-knot nematode resistant genes in cotton. Plant and Animal Genome VII Conference. San Diego. January 17-21. p84

Jiang, Chuan-Xiao, Wright, R.J., El-Zik, K.M. and Patterson, A.H. 1998. Polyploid formation created unique avenues for response to selection in *Gossypium* (cotton). *Proc. Natl. Acad. Sci.* 95:4419-4424.

Khan, A.M., Zhang, J. and Stewart, J. McD. 1998. Integrated molecular map based on atrispecific F2 population of cotton. *Proc. Beltwide Cotton Conf.*, San Diego, CA, 5-9 Jan. 1998. Natl. Cotton Council Am., Memphis, TN. p 491.

Luo, J., Biswas, G., and Gould, J. 1999. Shoot apex transformation and regeneration of Texas cotton for disease resistance. Plant and Animal Genome VII Conference. San Diego. January 17-21. p533

Paterson, A.H., Brubaker, C.L., and Wendel, J.F. 1993. A rapid method for extraction of cotton (*Gossypium* spp.) genomic DNA suitable for RFLP or PCR analysis. *Plant Mol. Biol. Rep*. 11 (2):122-127.

Reddy, A. S.,, Connel, S., Pammi, S., and Iqbal, J. 1998. Genetic mapping of cotton: Isolation and polymorphism of microsatellites. *Proc. Beltwide Cotton Conf*., San Diego, CA, 5-9 Jan. 1998. Natl. Cotton Council Am., Memphis, TN. p 485.

Reinisch, A., Dong, J., Brubaker, C., Stelly, D., Wendel, J., and Paterson, A. 1994. A detailed RFLP map of cotton, *Gossypium hirsutum* x *Gossypium barbadense*: chromosome organization and evolution in a disomic polyploid genome. *Genetics* 138:829-847.

Rungis, D., Llewellyn, D., Dennis, E., and Lyon, B. R. 1999. Development of molecular markers for use in cotton. Plant and Animal Genome VII Conference. San Diego. January 17-21. P479.

Saha, S., Jenkins, J. N., McCarty, J.C. and Stelly, D.M. 1998. Chromosomal location of RFLP markers linked to QTL in cotton. *Abstracts of the World Cotton Research Conference-2*, September 6-12, Athens, Greece. p 52.

Saha, S., Callahan, F., Dollar, D., and Creech J. 1997. Effect of lyophilization of cotton tissue on quality of extractable DNA, RNA and protein. *J. Cotton Sci.* 1:11-14.

Shappley, Z. W., Jenkins, J. N., Matson, C. E., Kohler, A. L., Meredith, W. R. 1996. Establishment of molecular markers and linkage groups in two F2 populations of upland cotton. *Theor Appl. Genet.* 92: 915-919.

Shappley, Z. W., Jenkins, J. N., Meredith, W. R, and McCarty, J. C. 1998. An RFLP linkage map of upland cotton, Gossypium hirsutum. *Theor Appl. Genet*. 97: 756-761.

Wajahatullah, M.K. and Stewart, J.McD. 1997. Genomic affinity among *Gossypium* subgenus Sturtia species by RAPD analysis. *Proc. Beltwide Cotton Conf.*, New Orleans,6-10 Jan. 1997. Natl. Cotton Council Am., Memphis, TN. p 452.

Wright, R. J., Thaxton, P. M. El-Zik, K. M., and Paterson, A. H. 1998. D-subgenome bias of Xcm resistance genes in tetraploid *Gossypium* (cotton) suggests that polyploid formation has created novel avenues for evolution. *Genetics* 149(4):1987-1996.

Yu, J., and Kohel, R. J. 1999. Cotton genome mapping and applications. Plant and Animal Genome VII Conference. San Diego. January 17-21. W60.

Yu, J., Park Young-Ha, Lazo, G.R. and Kohel, R. J. 1998. Molecular mapping of the cotton genome:QTL analysis of fiber quality properties. *Proc. Beltwide Cotton Conf.*, San Diego, CA, 5-9 Jan. 1998. Natl. Cotton Council Am., Memphis, TN. p 485.