Status of Genome Mapping and Use in Cotton Improvement

Yüksel BÖLEK
KSÜ, Ziraat Fakültesi, Tarla Bitkileri Bölümü, Kahramanmaraş

ABSTRACT

Cotton (*Gossypium* sp.) is the world's leading natural fiber crop. The use of molecular markers to enhance plant-breeding efforts is being widely studied. A major area of research is the use of molecular markers to identify and manipulate chromosome segments QTL (quantitative trait loci) controlling quantitative traits. The importance of genetic maps composed of molecular markers well documented for several economically important crops. The objective of this paper is to review current marker technology and genome mapping efforts and use of marker-assisted selection in cotton improvement.

Key Words: Cotton, molecular markers, plant breeding, marker-assisted selection, QTL.

Pamukta Genome Haritalamasının Durumu ve Islahta Kullanımı

ÖZET

Pamuk (*Gossypium* sp.) dünyanın önde gelen lif bitkilerinden birisidir. Moleküler markörlerden yararlanarak, bitki ıslahı çalışmalarının geliştirilmesi, üzerinde geniş bir şekilde çalışılan bir konudur. Çalışmaların büyük bir çoğunluğu moleküler markörler kullanılarak, kantitatif karakterleri kontrol eden ve QTL olarak adlandırılan kromozom parçalarını tesbit etmek ve düzenlemek üzerine yoğunlaşmıştır. Moleküler markörlerden oluşan genetik haritaların önemi, ekonomik açıdan önemli birçok bitki için ortaya konulmuştur. Bu derlemenin amacı; pamuk ıslahında, günümüzde popülaritesi giderek artan markör teknolojisinin ve gen haritalama çalışmalarının durumunu ve markör yardımıyla seleksiyonda kullanımı konusunda yapılan çalışmaları incelemektir.

Anahtar Kelimeler: Pamuk, Moleküler Markörler, Bitki Islahı, Markörler Yardımıyla Seleksiyon, QTL.

INTRODUCTION

Cotton is the world's leading natural fiber and second most important oilseed crop, has long been a focus of genetic, systematic and breeding research. Cotton includes about 50 diploid (2n=2x=26) and allotetraploid (2n=4x=52) species (Fryxell, 1992). Seven cytogenomic groups, designated A-G, are found in the genus *Gossypium* (Endrizzi et al, 1985). The allotetraploid species are made up of two genomic groups, the A and the D taxa (Beasley 1940, 1942; Wendell et al., 1992). Two tetraploid species, *G. hirsutum* L. and *G. barbadense* L., accounts for 90 and 5%, respectively, of the world's cotton production (Wendel et al., 1992). World cotton commerce of about $20 billion annually is dominated by "AD" tetraploid species.

Currently, the cotton industry faces many challenges including competition from synthetic fibers, demand for pest and disease resistant varieties, lack of genetic diversity within the improved germplasm, and improvement in textile technologies that require higher quality fibers. However, the promise and potential of new technologies to meet the challenges are immense. One of these technologies is to use plant genome mapping to identify and manipulate QTLs/genes that are linked to economically important traits. Detailed genetic maps of DNA markers have been constructed for most major crop plants, and mapping of many agriculturally significant genes has contributed to molecular cloning.

Genome mapping is a synthesis of concepts from classical genetics with the new tools from molecular biology. The priority of plant genome mapping for crop productivity is the identification of genes associated with important traits, and use of this information to further improve crops to meet the needs of a hungry world. In the past decade genome mapping has emerged as a powerful approach to research in agriculture.

Understanding the level of inheritance of agriculturally important traits at the DNA level creates new opportunities to streamline plant breeding, the process of altering plant genotypes to better fit the needs of human being. Further, this understanding provides a channel for communication between the farmer's field and the molecular biology laboratory, in principle enabling the scientist to identify the specific DNA elements responsible for particular plant characteristics. There has been a lot of work in cotton improvement by using genome-mapping approach and there is a need to review what has been accomplished. This paper presents current developments in cotton genome mapping and use of molecular markers in cotton improvement.

Why Is Genome Mapping?

A genetic map could be thought as a road map, reflecting the relative proximity of different landmarks to one another, and molecular markers at defined places along each linear chromosome enable the geneticist to determine a particular gene of interest (Paterson, 1996). Genetic linkage maps are useful tools for studying genome structure, evolution, identifying introgression and for marker-assisted selection in breeding programs because they are closely associated with important agronomic traits. Modern *G. hirsutum* and *G. barbadense* cultivars show significant variation for important traits including fiber production, pest resistance, and tolerance to environmental adversities such as heat, cold and drought (El-Zik and Thaxton, 1989). Molecular markers are playing a critical and increasing role in the development of superior cultivars that combine the favorable traits. In addition, wild *Gossypium* germplasm also harbors many valuable traits including disease and insect resistance, stress tolerance and fiber quality attributes. However, the transfer of genes from wild species is time consuming and not always successful. DNA markers and genetic maps that assess the introgression of alien genes into cultivated cotton will greatly accelerate breeding efforts.

Morphological features are indicative of the phenotype but are represented by only a few loci because there is not a large enough number of a character, and they can also be affected by environmental factors and growth practices. Improving yield

and quality of cotton has been mostly attributed to improve its genetic potential. Currently, there are about 145 morphological markers identified in cultivated cotton but their utility in breeding programs has remain limited because of their deleterious effect and the difficulties in combination of multiple markers in a single genotype (Percy and Kohel, 1999). Besides not producing a complete linkage map, most of these markers have major effects on other important quantitative traits and have very limited usefulness in selection and in many cases the heterozygosity condition is not identifiable (Meredith, Jr., 1994). To have an accurate and reliable estimate of genetic relationship and genetic diversity, a large number of polymorphic markers are essentially required.

Biochemical markers such as isozymes were used to identify QTLs in maize, tomato, oats and soybean (Stuber, 1992; Stuber and Edwards, 1986). Although isozyme markers likely have no phenotypic effects, their analysis has certain limitations due to the availability of a limited number of marker loci (insufficient for broad applications), a general lack of polymorphisms for these loci in elite breeding materials, and the chance of variability in breeding patterns being due to plant development (Tanksley et al., 1989). Therefore, many more genetic markers that are polymorphic among a large collection of cotton germplasm will be needed. Shanti et al. (2001) showed that molecular markers are more reliable indicators of fertility restorer than morphological markers in several lines of cotton, thus using highly saturated genetic maps made up of molecular markers will be very powerful approach in the improvement of crop species. Furthermore, cloning of agriculturally important genes are challenging. Since many agriculturally important traits are influenced by multiple genes, and because the effects of multiple genes can be obscured by the nongenetic factors such as environmental variation, and plant genes are large and complex, requiring a search through tens of thousands of protein-encoding sequences together with a much larger quantity of non-coding DNA to find the single element that directs a particular biochemical step, map based cloning is amenable to isolation of genes known only from phenotype (Meyer et al., 1996).

Linkage Maps of Cotton

Molecular linkage maps based upon DNA markers are widely recognized as essential for genetic research in many species (Dudley, 1993). A detailed molecular map of the cotton genome has been published by Reinisch et al. (1994) who used an interspecific cross of *G. hirsutum* and *G. barbadense* to assemble 705 RFLP loci into 41 linkage groups and 4675 cM, with average spacing between markers of about 7 cM. Since each gamete contain 26 chromosomes, at least 15 gaps exist in the map to bring overall genetic size of cotton genome to 5125cM. Fourteen of the 26 chromosomes have been associated with linkage groups by using a series of monosomic interspecific substitution stocks developed by Stelly (1993). The physical size of the cotton genome is relatively large, in a range of 2702 Mb to 2246 Mb (Arumugunathan and Earle, 1991; Michaelson et al., 1991). With a genome size of 2246 Mb, the average physical size of a cM in cotton is about 400 kb. The genetic map 5000 cM of cotton genome (Yu et al., 1998) will require 3000 DNA probes to map at average 1 cM density, and the physical genome of 2246 Mb will

require 75.000 YACs/BACs of average size 150 kb for 5x coverage (Arumugunathan and Earle, 1991).

An F_2 population was derived from a cross between homozygous lines *G. hirsutum* cv. TM-1 and *G. barbadense* cv. 3-79 at the USDA-ARS in Texas and segregation data of 171 F_2 individuals of this cross were obtained for 868 genetic markers (Yu et al., 1997; Reddy et al., 1997). These markers have been mapped into 50 linkage groups, and spanning nearly 5000cM of the cotton genome. By using of diploid and aneuploid cottons, 21 linkage groups have been derived from the A sub-genome, and 19 from the D sub-genome. Eighteen of the 26 coton gametic chromosomes were identified with linkage groups.

A trispecific F_2 population was developed from three different cultivars to study inheritance patterns of segregating loci and to establish linkage groups among three genome species by Altaf et al. (1997). 11 linkage groups were identified that spanned 521.7 cM with an average distance of 16.8 cM between markers with a large number of markers showed distorted segregation. Brubaker et al. (1999) compared both allotetraploid cotton and its diploid progenitors with a set of RFLP markers. Besides interspecific linkage maps, intraspecific maps are also constructed by several researchers (Shapley et al., 1998;Ulloa and Meredith, 1999; Ulloa et al., 2001) to investigate cotton genome and identify molecular markers linked to agriculturally important genes/QTLs.

Mapping Genes in Cotton

Numerous introgression events in cotton may be detectable using existing molecular markers. More than 50% of the introgressed chromatin is found in five specific chromosomal regions, which are largely common to genotypes from each of the leading *G. barbadense* breeding programs throughout the world (Wang et al., 1995). Molecular markers hold promise for mapping many traits in cotton which have been introgressed from exotic germplasm, such as Verticillium wilt resistance, and bacterial blight resistance (Staten, 1971), restoration of cytoplasmic male-sterility (Weaver and Weaver, 1977), improved fiber quality (Culp et al., 1979). In view of most measures of cotton quality and productivity are polygenic, QTL mapping is in a high priority of many research programs.

Genetic mapping of 145 monogenic traits, characters and mutants has assembled 65 suchmarker genes into 18 linkage groups, with 13 being assigned into specific chromosomes (Percy and Kohel, 1999). Association of DNA markers with these genes and others of agronomic importance would provide cotton breeders new opportunities during marker-assisted selection (MAS). Currently, several cotton genes have been tagged with DNA markers, including leaf shapes (Yu et al., 1997), plant trichomes, photoperiodism (Yu and Kohel, 1999), disease resistance (Wright et al., 1998; He et al., 1999; Bolek, 2002), fertility restorer (Lan et al., 1992; Guo et al., 1998; Zhang et al., 1999), phylogenomic relationships (Multani and Lyon, 1995; Iqbal et al., 1997; Pillay and Myers, 1999). Pendse et al. (2001), and Meredith and Brown (1998) used molecular markers to identify hybrids, Ulloa et al. (2000) found QTLs for stomatal conductance in a segregating population.

Liu et al (2000) determined the molecular variation for desirable alleles and accurate characterization of the variability within and among germplasm accession

in converted race stock collection, and they postulated that the recovery of the primitive recurrent parent could be improved by marker-assisted backcrossing. Most cotton QTLs identified so far influence fiber quality and yield, among other agronomic traits (Jiang et al., 1998; Yu et al., 1998; Kohel et al., 2001; Khan et al., 1998; Shapley et al., 1998; Ulloa and Cantrell, 1998). In *G. hirsutum* two RFLP genetic maps have been developed and used to identify QTLs for agronomic and fiber quality traits (Shappley et al., 1998; Ulloa and Meredith, 2000; Ulloa et al., 2000). DNA markers identified for fiber quality genes are being used to monitor *G. barbadense* introgression in the segregation progeny (Jiang et al., 2000).

CONCLUSION

Molecular markers are a valuable new tool with much promise for uses in future cotton genetic-breeding investigations. They offer a relatively simple method of tracing genetic sources of useful variability, thus specific chromosome regions with important QTLs can be identified and appropriate selection strategies developed. A detailed map holds the opportunity for identification of DNA markers for rapid assay of segregants, even using nondestructive assay of ungerminated seed. In view of the economic importance of cotton, future research will focus on agricultural productivity and quality. Characterization of gene pool, DNA fingerprint, molecular dissection of complex traits i.e. fiber attributes, characterization of genome organization and map-based cloning of mutations unique to cotton are not far away. Cotton may also make a unique contribution to genome evolution and distribution over different continents to study genome divergence.

DNA marker-assisted selection (MAS) would facilitate a breakage of the linkage drags and separate superior genes from negative factors in the breeding pedigrees. MAS also facilitate early detection of the targeted genes, especially for those that are difficult to score for phenotypes. Up to know, MAS has been limited to a few experimental programs and the genetic maps developed through interspecific hybridization of cotton currently have little use in conventional breeding programs. Most cotton breeding programs that employ MAS involve the DNA marker for major genes, as monogenic traits are easier to follow than complex traits. Limitations due to molecular markers use are cost, availability of quality molecular markers, intraspecific linkage maps, appropriate software for QTL analysis and the polyploid nature of cotton. Ongoing programs are great potential to reduce these limitations in near future.

REFERENCES

Altaf, M.K., J.M. Stewart, M.K. Wajahatullah, J-F. Zhang, R.G. Cantrell. 1997. Molecular and Morphological Genetics of a Trispecies F_2 Population of Cotton. In: Proc. of Beltwide Cotton Conferences (Ed. D.J. Herber and D.A. Richter),January 6-10, New Orleans, LA. pp. 448-452.

Arumuganathan, K., E.D. Earle. 1991. Nuclear DNA Content of Some Important Plant Species. Pl.Mol.Biol.Rep., 9:208-218.

Beasley, J.O. 1940. The Origin of American Tetraploid *Gossypium* Species. Am. Nat., 74:285-286.

Beasley, J.O. 1942. Meiotic Chromosome Behavior in Species, Species Hybrids, Haploids and Induced Polyploids of *Gossypium*. Genetics, 27:25-54.

Bolek, Y. 2002. Mapping of Verticillium Wilt Resistance Genes in Cotton. PhD. Dissertation, Texas A&M University Library, 113p.

Brubaker, C.L., A.H. Paterson, J.F. Wendel. 1999. Comparative Genetic Mapping of Allotetraploid Cotton and its Diploid Progenitors. Genome, 42:184-203.

Culp, T.W., D.C. Harrel, T. Kerr. 1979. Some Genetic Implications in the Transfer of High-Fiber Strength Genes to Upland Cotton. Crop Sci., 19:481-484.

Dudley, J.W. 1993. Molecular Markers in Plant Improvement: Manipulation of Genes Affecting Quantitative Traits. Crop Sci., 33:660-668.

El-Zik, K.M., P.M. Thaxton 1989. Genetic Improvement for Resistance to Pests and Stresses in Cotton. In: Integrated Pest Management System and Cotton Production (Ed. R. E. Frisbie, K. M. El-Zik, and L. T. Wilson), John Wiley and Sons, Inc., New York, pp. 191-224.

Endrizzi, J.E., E.L. Turcotte, R.J. Kohel. 1985. Genetics, Cytology and Evolution of *Gossypium*. Adv. Genetics, 23:271-375.

Fryxell, P.A. 1992. A Revised Taxonomic Interpretation of *Gossypium* L. Rheedea 2:108-165.

Guo, W., T. Zhan, J. Pan, R.J. Kohel. 1998. Identification of RAPD Markers Linked with Fertility-Restoring Gene of Cytoplasmic Male Sterile Lines in Upland Cotton. Chinese Sci., 43(1):52-54.

He, L., A. Bridges, F. Robinson, C.C. Cook, H-B. Zhang. 1999. Toward Map-Based Cloning of Root-Knot Nematode Resistant Genes in Cotton. Proc. Plant & Animal Genome VII, January 17-21, San Diego, CA.

Iqbal, M.J., N. Aziz, N.A. Saeed, Y. Zafar. 1997. Genetic Diversity Evaluation of Some Elite Cotton Varieties by RAPD Analysis. Theor. Appl. Genet., 94:139-144.

Jiang, C-X., R.J. Wright, K.M. El-Zik, A.H. Paterson. 1998. Polyploid Formation Created Unique Avenues for Response to Selection in Gossypium (cotton). Proc. Natl.Acad. Sci. USA., 95:4419-4424.

Jiang, C-X, R.J. Wright, S.S. Woo, T.A. Delmonte, A.H. Paterson. 2000. QTL Analysis of Leaf Morphology in Tetraploid *Gossypium* (cotton). Theor. Appl.Genet., 100:409-418.

Khan, M. A., J. Zhang, J. McD. Stewart. 1998. Integrated Molecular Map Based on a Trispecific F_2 Population on Cotton. In: Proc. Beltwide Cotton Improvement Conference (Ed. D.J. Herber and D.A. Richter), San Diego, CA, pp.491-492.

Kohel, R.J., J. Yu, Y-H, Park, G.R. Lazo. 2001. Molecular Mapping and Characterization of Genes Controlling Fiber Quality in Cotton. Euphytica, 121(2):163-172.

Lan, T.H., C. Cook, A.H. Paterson. 1992. Molecular Mapping of a Gene for Restoration of Cytoplasmic Male-Sterility in Cotton. Proc. Plant Genome II, San Diego, CA.

Liu, S., R.G. Cantrell, J.C. McCarty, J.M. Stewart. 2000. Simple Sequence Repeat-Based Assessment of Genetic Diversity in Cotton Race Stock Accessions. Crop Sci., 40(5): 1459-1469.

Meredith (Jr), W.R. 1994. Use of Molecular Markers in Cotton Breeding. In: Challenging the Future: Proc. World Cotton Res. Conf. (Ed: G.A. Constable and N.W. Forrester), Feb 14-17. Brisbane, Australia, pp. 303-308.

Meredith (Jr), W.R., J.S. Brown. 1998. Heterosis and Combining Ability of Cottons Originating from Different Regions of the United States. J. Cotton Sci., 2(2):77-84.

Meyer, K.G., G. Benning, E. Grill. 1996. Cloning of Plant Genes Based on Genetic Map Location. In: Genome Mapping in Plants (Ed. A.H. Paterson), pp. 137-154.

Michaelson, M.J., H.J. Price, J.R. Ellison, J.S. Johnston. 1991. Comparison of Plant DNA Contents Determined by Feulgen Microspectrophotometry and Laser Flow Cytometry. Am. J. Bot., 78(2):183-188.

Multani, D.S, B.R. Lyon. 1995. Genetic Fingerprinting of Australian Cotton Cultivars with RAPD Markers. Genome, 38(5):1005-1008.

Paterson, A.H. 1996. Making Genetic Maps. In: Genome Mapping in Plants (Ed. A.H. Paterson), R.G. Landes Com.. and Academic Press, Inc., pp.23-39.

Pendse, R., S. Malhotra, S.E. Pawar, T.G. Krishna. 2001. Use of DNA Markers for Identifying Inbreds and Hybrid Seeds in Cotton (*Gossypium hirsutum* L.). Seed Science & Technology, 29(2):503-508.

Percy, R.G., R.G. Kohel. 1999. Cotton Qualitative Genetics. In: Cotton (Ed. C.W. Smith and J.T.Cothren), John Wiley & Sons NY, pp. 319-360.

Pillay, M., G.O. Myers. 1999. Genetic Diversity in Cotton Assessed by Variation in Ribosomal RNA Genes and AFLP Markers. Crop Sci., 39(6):1881-1886.

Reddy, A.S., R.M. Haisler, Z-H. Yu, R.J. Kohel. 1997. AFLP Mapping in Cotton. Proc. Plant & Animal Genome V, January 12-16, San Diego, CA.

Reinisch, A.J., C. Brubaker, D. Stelly, J. Wendel, A.H. Paterson. 1994. A Detailed RFLP Map of Cotton (Gossypium hirsutum x G. barbadense): Chromosome Organization and Evolution in a Disomic Polyploid Genome. Genetics, 138: 829-847.

Shanti, P., T.S. Raveendran, P.R. Kumar. 2001. Identification of Fertility Restorers in Cotton Germplasm Using Morphological and Molecular Markers. Journals of Soil & Crops, 11(1):66-73.

Shappley, Z.W., J.N. Jenkins, J. Zhu, J.C. McCarty. 1998. Quantitative Trait Loci Associated with Agronomic and Fiber Traits of Upland Cotton. J. Cotton Sci., 2:153-163.

Staten, G. 1971. Breeding Acala 1517 cottons, 1926-1970. New Mexico State University. College of Agric. and Home Econ., Memoir Series No:4, 48pp.

Stelly, D.M. 1993. Interfacing Cytogenetics with the Cotton Genome Mapping Effort. In: Proc. Beltwide Cotton Improvement Conference (Ed. D.J. Herber and D.A. Richter), pp. 1545-1550.

Stuber, C.W., M.D. Edwards. 1986. Genotypic Selection for Improvement of Quantitative Traits in Corn Using Molecular Marker Loci. Proc. 41st Annual Corn and Sorghum Research Conf., Am. Seed Trade Assoc., 41:40-83.

Stuber, C.W. 1992. Biochemical and Molecular Markers in Plant Breeding. In: Plant Breeding Reviews (Ed. J. Jonick), John Wiley & Sons NY, pp. 37-61.

Tanksley, S.D, N-D. Young, A.H. Paterson, M.W. Bonierbale. 1989. RFLP Mapping in Plant Breeding: New Tools for an Old Science. Biotechnology, 7:257-264.

Ulloa, M., W.R. Meredith (Jr). 1999. A Partial RFLP Genetic Linkage Map from an Intraspecific Cross (*G. hirsutum* x *G. hirsutum* L.), Proc. Plant & Animal Genome VII,. January 17-21, San Diego, CA.

Ulloa, M., R.G. Cantrell. 1998. QTL Analysis in Cotton. Proc. Plant & Animal Genome VI, January 18-22, San Diego, CA.

Ulloa, M., W.R. Meredith (Jr). 2000. Genetic Linkage Map and QTL Analysis of Agronomic and Fiber Quality Traits in an Intraspcific Population. J. Cotton Sci., 4:161-170.

Ulloa, M., R.G. Cantrell, R.G. Percy, Z-M. Lu, E. Zeiger. 2000. QTL Analysis of Stomatal Conductance and Relationship to Lint Yield in an Interspecific Cotton. J. of Cotton Sci., 4(1): 10-18.

Ulloa, M., W.R. Meredith (Jr), Z.W. Shapley, A.L. Kahler. 2001. RFLP Genetic Linkage Maps from Four $F_{2:3}$ Populations and a Join Maps of Gossypium hirsutum L. Theor. Appl. Genet., 101:200-208.

Wang, G., J. Dong, A.H. Paterson. 1995. The Distribution of *Gossypium hirsutum* in *Gossypium barbadense* germplasm: Molecular Analysis of Introgressive Plant Breeding. Theor. Appl. Genet., 91(6/7):1153-1161.

Weaver, D.B., J.B. Weaver. 1977. Inheritance of Pollen Fertility Restoration in Cytoplasmic Male-Sterile Upland Cotton. Crop Sci., 17:497-499.

Wendel, J.F., C.L. Brubaker, A.E. Percival. 1992. Genetic Diversity in *Gossypium hirsutum* and the Origin of Upland Cotton. Am. J. Bot., 79:1291-1310.

Wright, R.J., P.M. Thaxton, K.M. El-Zik, A.H. Paterson. 1998. D-Subgenome Bias of *Xcm* Resistance Genes in Tetraploid Gossypium (Cotton) Suggests that Polyploid Formation Has Created Novel Avenues for Evolution. Genetics, 149:1987-1996.

Yu, Z-H, Y-H. Park, G.R. Lazo, R.J. Kohel. 1997. Molecular Mapping of the Cotton Genome. Prof of 5th International Congress of Plant Molecular Biology, Sept. 21-27, Singapore.

Yu, Z-H, Y-H. Park, G.R. Lazo, R.J. Kohel. 1998. Molecular Mapping of the Cotton Genome: QTL analysis of fiber quality characteristics. Proc. Plant & Animal Genome VI, January 18-22, San Diego, CA.

Yu, Z-H., R.J. Kohel. 1999. Cotton Genome Mapping and Applications. Proc. Plant & Animal Genome VII. San Diego, CA. January 17-21.

Zhang, J., J.McD. Stewart, R.B. Turley. 1999. Molecular Mapping of CMS-D_8 Restoration and Gene Cloning Specific to D_8 Restorer. In: Proc. Beltwide Cotton Improvement Conference (Ed. D.J. Herber and D.A. Richter), Orlando, FL. January 3-7, 1999, pp.448.