DNA Fingerprinting for Identification of Varieties

Selection within the existing population is a recognized method for development of varieties. Spontaneous mutations, out-crossings or a mixture could be responsible for variations within a pure breeding variety. Identification of such varieties based on morphological differences is always a major issue. Even if a variety was developed through hybridization between cultivars having a narrow genetic base, identification becomes difficult unless the new genotype has a marker gene. Marker genes with prominent morphological differences are, first, not easily available and, second, they may carry undesirable effects. Breeding for induction of a prominent morphological character or to avoid the undesirable effect of a marker gene is not recommended

because of the time involved in this process. Nevertheless, for purposes of maintaining the purity of a new variety in the seed production process, prominent morphological differences among varieties have to be found. Morphological features are indicative of the genetic make up of the plant, and it is usually not difficult for breeders to identify their own varieties even with minor differences from other varieties. Seed production systems all over the world are still based on visual differences. Off types—deviations from the variety norm—are eliminated, and it is assumed that the rest of the population is pure.

Identification of Genetic Diversity

The genetic information coded in Deoxyribose Nucleic Acid or DNA—the complex chemical which encodes the genetic information of all living organisms—is comprised of four bases abbreviated as A (Adenine), C (Cytosine), G (Guanine) and T (Thiamine). The sequence of arrangement of these bases determines the internal as well as external performance of a genotype. There is no easy way to ascertain genetic purity within a population. Genetic diversity resulting from intra or interspecific introgression can be evaluated with morphological characters, seed proteins, isozymes and (DNA) markers. Isozyme analysis is limited by the small number of marker loci, a general lack of polymorphism in these loci and chances of variability in the banding pattern due to plant development (Tanksley et al 1989). A large number of polymorphic markers are required to measure genetic diversity in a reliable manner which limits the use of morphological characters and isozymes to be used in measuring genetic diversity. Restricted fragment length polymorphism (RFLP) can be used but it is time consuming and expensive. However, DNA fingerprinting provides a powerful tool to identify the extent of similarities and differences among genotypes. Using Random Amplified Polymorphic DNA (RAPD) provides an unlimited number of markers for various purposes. The RAPD technique is simple and fast enough to measure the relationship among genotypes with a fair degree of reliability. RAPD has been able to provide the quantity of variation or confirm the relationship among genotypes. Moreover, Tatineni et al (1996) compared the genetic differences obtained from RAPD markers of sixteen US varieties with visual morphological differentiations of the same varieties and concluded that RAPD marker analysis confirmed taxonomic differences among varieties.

Quantification of Genetic Diversity

Limited work has been done on quantification of genetic diversity among species and varieties and within varieties. However, some of the latest work done in Australia, Pakistan and the USA has shown that interspecific differences for polymorphism are not large. Genetic fingerprinting of *G. hirsutum*, *G. barbadense* and *G. arboreum* with RAPD markers showed that *G. barbadense* and *G. arboreum* species were at least 50% similar to *G. hirsutum*. Multani and Lyon (1995) used RAPD markers generated by thirty random decamer primers and silver staining to fingerprint thirteen upland cottons and a variety of *G. barbadense*. Thirteen *G. hirsutum* varieties included eleven locally developed Australian genotypes, the Deltapine variety DP 90 and the S 295 variety from France. Seeds were sterilized in 20% household bleach for twenty minutes and grown on an artificial medium under laboratory conditions. For DNA extraction, cotyledonary leaves from two to five plants were mixed to make a bulk of 1.5 grams. Thirty decamer primers of four different sets were employed for DNA amplification. Amplified DNA samples were analyzed by electrophoresis and an amplification by a given primer was marked as present or absent for all cultivars.

According to Multani and Lyon, fourteen cotton cultivars and thirty primers resulted in the formation of 453 amplified DNA fragments or RAPD markers. 15% of the total markers were specific to the *G. barbadense* variety S-7 and they were absent in any other variety included in the trial. 33% of the total markers were unique to *G. hirsutum* varieties. Within *hirsutum* types, DP 90 was 60% genetically different from the Australian varieties; however, it was about 80% similar to CS 50. The French variety S 295, which is supposed to have a diverse origin (may have been developed in some African country), was approximately 90% similar to Australian type CS 50.

Not all varieties were distinguishable with a number of markers. Some upland cultivars were distinguishable with only 10, 12, 15 and 19 RAPD markers while Pima S-7 varied from upland types by 104 markers (69 unique and 35 absent markers). Based on molecular markers, it was possible to group varieties into various groups in respect to their genetic similarity or differences. Multani and Lyon hypothesized that on the basis of polymorphic markers, it could be possible to understand the genetic relationship among genotypes of unknown origin.

Similar work has been done in Pakistan on 22 upland cultivars from Australia, Pakistan and the USA and the commercially grown Pakistani *G. arboreum* variety, Ravi. Iqbal et al (1996) also used the Williams et al (1990) RAPD technique to estimate genetic homology among cotton varieties of diverse origin. Fifty primers belonging to various groups of Operon were used for PCR amplification.

Iqbal and his group, in addition to genetic differences among varieties, also studied differences among plants within a variety. For this purpose, two varieties, S 12 and the candidate variety Krishma, were grown beyond seedling stage. Twenty plants within S 12 exhibited consistent amplification profiles for thirteen primers, thus showing a very high degree of similarity among plants. It was proved that the variety is true breeding and quite homozygous in its genetic make up. Fifteen plants within Krishma were studied for nine primers and polymorphism was found for only two primers. The polymorphic studies of Krishma showed that the variety is not homozygous and is still segregating for some characters. According to Iqbal et al (1996), morphological differences among plants within Krishma and their

differential response to leaf curl virus disease in the field also confirmed the heterogeneity of the population.

No single primer was capable of inducing polymorphic effects in all 23 varieties so that they could be differentiated among themselves. Forty out of fifty primers amplified a total of 349 DNA fragments in 22 varieties. One primer was not polymorphic for any of the 23 varieties. However, the inability of a single primer to produce polymorphic markers in all varieties under study indicated a closer genetic relationship among varieties. Like the Australian work, intervarietal differences for a number of amplified fragments were quite evident, and the size of amplified fragments with different primers also varied in different varieties.

Iqbal et al (1996) also confirmed interspecies differences in respect of polymorphic markers. They found that the *G. arboreum* variety had the least similarity with other varieties. Ravi showed a 55% similarity with the variety CIM 1100, more than for all other varieties. Diploid cottons are said to be tolerant to leaf curl virus disease and field experiments have proved that CIM 1100 (*G.hirsutum*) also has a high tolerance to this disease, one reason for the similarity between these two varieties. The RAPD data also showed that seventeen out of 22 varieties were similar in their DNA amplification from 82% to over 94%. There is an indication of a greater similarity among varieties developed at the same station. Iqbal et al (1996) have related this similarity to the narrow genetic base used in hybridization of new genotypes.

Tatineni et al (1996) studied the effect of eighty primers on amplification of DNA genomes of nineteen cotton genotypes including *G. hirsutum* and *G. barbadense* genotypes. The genotypes used had a very diversified origin and exhibited great variation in morphological data such as sympodia length, leaf length, depth of cut in leaf, petal color, pollen color, and boll shape and surface. Instead of cotyledonary leaves, Tatineni et al used fully expanded leaves from ten different plants and mixed them together before genomic DNA was extracted. Twenty-seven out of eighty primers yielded either monomorphic amplified products or no morphism was seen in any variety. Fifty-three primers were able to amplify 135 fragments that were clearly polymorphic thus showing an average of 1.7 RAPD per primer. As expected, the two barbadense genotypes had a low genetic distance of 0.42 from each other. On the basis of RAPD, it was possible to group all varieties into two major groups of *G. hirsutum* and *G. barbadense*.

This recent development in the study of DNA has created new opportunities to understand similarities and differences. It has also made it possible to establish relationships among genotypes of unknown origin. The above mentioned studies are among the first studies done in these countries to quantify genetic diversity in a broad array of genotypes belonging to three different cultivated species of cotton. The genotypes selected included distantly related genotypes and a variety of primers were tried.

Polymorphism Studies in Some Countries

Item	Australia	Pakistan	USA
No. of varieties	14	23	19
No. of primers	30	50	80
No polymorphism	0	1	27
Total No. of polymorphism	453	349	135
Average RAPD/primer	15.1	7.0	1.7

While details on primers can be obtained from the respective laboratories, results of the three studies are compared in the table above.

The conventional characteristics or parameters that distinguish varieties or genotypes are highly influenced by the environment and growing conditions. The gene loci codings for isozymes and proteins are conserved in nature and hence are not polymorphic enough to facilitate differentiation of varieties having narrow genetic differences. Identification of alternate marker systems capable of distinguishing closely related materials is necessary. With the discovery of DNA fingerprinting in 1985, attention has been focused on DNA as a source of informative polymorphism. The DNA sequence of each individual is unique and is not influenced by environment or growing conditions. The DNA sequence is stable and does not change with the increase in growth thus providing an excellent means to study genotype identification in cotton and other living organisms. In addition to studying the genetic diversity in a given population, DNA fingerprinting also provides for tagging of genes of economic importance and their efficient utilization in the breeding programs. Gene tagging will provide not only a sense of direction but much more precision in breeding for specified objectives.

DNA fingerprinting could be a very useful tool for implementing intellectual property rights in various countries. It could help resolve controversial issues related to utilization of patent specific genes and the introduction and multiplication of seed for commercial planting. In India, the National Research Center on DNA Fingerprinting has recently been established to conduct research in the areas of development of probes for different species, techniques and technologies for utilization of information coded on DNA.

Contributions from Parents in F_1 Hybrids

A genetic system which could directly measure the extent of heterozygosity in commercial cotton hybrids or even ordinary hybrids is very important. Measurement of the extent of hybridity can reflect performance to a great extent. Gwyn et al (1995) using RFLP analysis studied ten parental lines of known combining ability, their eight F_1 combinations and F_2 population from two parental combinations. From the two, F_1 combinations consisting of 100% and 50% hybrid subpopulations and the two F_2 populations produced from the 100% pure F_1 plants were analyzed for their genetic integrity. DNA banding patterns from

the F_1 population precisely reflected the 100% or 50% contribution of the male parent. The F_2 population also clearly demonstrated the expected genetic ratio. Gwyn et al (1995) proved that RFLP can successfully be employed to determine the genetic contribution of parents and possibly be used to maximize heterozygosity and thus heterosis in cotton hybrids.

Progress has already been made to find chromosomal homology and to locate the existence of characters on various chromosomes using mapping of DNA probes and RFLP analysis. It seems that through DNA fingerprinting shortly it will be possible to identify the extent of expression of a particular character before field tests. Better understanding of the location of specific genes will facilitate forecasting hybrid performance.

References

Gwyn, J. J., R. W. Whitmore, G. R. McPherson, J. C. Vasek, and B. A. Greenley. 1995. The application of RFLP's in hybrid cotton breeding and seed production. *Proceedings of the Beltwide Cotton Conferences*, National Cotton Council of America, P. O. Box 12285, Memphis, TN 38182, USA.

Iqbal, M. J., N. Aziz, N. A. Saeed, Y. Zafar and K. A. Malik. 1996. Genetic diversity evaluation of some elite cotton cultivars by RAPD analysis. *Theoretical and Applied Genetics* (in press).

Multani, D. S. and B. R. Lyon. 1995. Genetic fingerprinting of Australian cotton cultivars with RAPD markers. *Genome*, 38:1005-1008

Saha, S., D. M. Stelly and A. E. Percival. 1988. Isozyme polymorphism among diploid and tetraploid species of cotton. *Proceedings of the Beltwide Cotton Conferences*, National Cotton Council of America, P. O. Box 12285, Memphis, TN 38182, USA.

Tanksley, S. D., N. D. Young, A. H. Paterson and A. H. Bonierbal. 1989. RFLP mapping in plant breeding: new tolls for old science. *Biotechnology*, 7:257-264

Tatineni, V., R. G. Cantrell and D. D. Davis. 1996. Genetic diversity in elite cotton germplasm determined by morphological characteristics and RAPD. *Crop Science*, 36:186-192

Williams, J. G. K., A. R. Kubelik, K. J. Levak, J. A. Rafalski, and S. V. Tingey. 1990. DNA polymorphism amplification by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, 18:6531-6535