Identification of DNA polymorphism in cultivars of groundnut(*Arachis hypogaea L.*) using RAPD and AFLP

By
C Srinivas Reddy

School of Biotechnology
Institute of Post graduate studies and Research
Jawaharlal Nehru Technological University
Hyderabad

CERTIFICATE

This is to certify that the thesis entitled *"Identification of DNA polymorphism in cultivated groundnut* (**Arachis hypogaea** *L.) using RAPD and AFLP markers"* is submitted by Mr. C. Srinivas Reddy towards partial fulfillment for his M. Tech (Biotechnology) to Jawaharlal Nehru Technological University, Hyderabad. This work embodies a record of bonafied research carried out under my supervision at the Genetic Resources Enhancement Program of the International Crops Research Institute for the Semi-Arid Tropics. This work has not been submitted earlier either to this University or to any other institution for fulfillment of the requirement of a course of study.

Dr. S.N. Nigam
Principal Scientist
GREP
ICRISAT
Patancheru 502 324
Andhra Pradesh, India

Dr. Laxmi Narasu
Head
School of Biotechnology
JNTU, Hyderabad
Andhra Pradesh

Executive Summary

The molecular markers based on DNA sequence variation have significant advantages over the other protein markers for genotype identification, genome analysis and mapping. DNA markers are neutral to various environmental factors, highly sensitive and sufficiently reproducible. Therefore, the objective was to familiarize with various molecular marker techniques. During this project work, I learnt RAPD, AFLP, and DAF to reveal polymorphism at DNA level in Groundnut. Apart from this, I have also undertaken some additional activities such as computer applications using C, C++ and information search through *Internet*, because of the unique opportunities available at ICRISAT.

RAPD analysis of five different genotypes in groundnut with 48 different primers revealed 96 PCR amplified products of which 10 were polymorphic. Primers Gn39, V4, B11, B13, GN20 showed best polymorphism. AFLP analysis of 10 genotypes and the whole progeny with 4 AFLP primer pairs identified a total 101 fragments of, which 17 were polymorphic. Four primer pairs E-coR1 primer E-ACA with MseI primersM-CAC,M-CAG,M-CTG,M- CTT showed the best results. Cluster analysis of 10 & 5 Groundnut genotypes for RAPD and AFLP data was carried out using statistical software package GENSTAT and a dendrogram was constructed

Each of these above techniques studied has some advantages and disadvantages. Depending on application, a range of markers can be selected. RAPD is simple, fast, free from hazardous material, and needs only small amounts of DNA. Although it has advantages over protein markers and RFLP, the uncertainty of reproducibility of RAPD markers and their dominant nature limits its use.

AFLP combines the advantages of both RFLP and RAPD, it requires less amount of DNA and is faster than RFLP, reveals several polymorphic fragments in a single reaction. These markers are reliable and reproducible. However, AFLP analysis is expensive and requires highly skilled workers.

With the experience gained above, I can now attempt development of Groundnut linkage map and identify DNA markers of agronomic interest.

Acknowledgments

I express deep sense of gratitude and sincere thanks to Dr. S.N. Nigam, ***Principal Scientist****, Genetic Resources and Enhancement Program, ICRISAT, Patancheru for his guidance, supervision and valuable suggestions during the course of this study. I am very thankful to Dr S L Dwivedi for his encouragement and suggestions.*

I am grateful to Dr. Laxmi Narsu, Head and Coordinator, Department of Biotechnology, Jawaharlal Nehru Technological University, Hyderabad for her cooperation and encouragement during the course of the endeavor.

I am thankful to Mrs. Santosh Gurtu Sr. Research associate for supervision, teaching all biotechnological tests including molecular markers, RAPDs, AFLP etc.

I am thankful to Dr. B. Diwakar, Head, Training and Fellowship Program, ICRISAT for his help. I sincerely thank Dr. N. Seetharama, Senior Scientist, GREP for his suggestions, guidance, and permiting me to participate in Training Course on Genetic Transformation using Particle Inflow Gun at ICRISAT. I sincerely thank Mrs. S.S. Beena, Research associate, for her encouragement, cooperation, and constructive suggestions during my research work. I sincerely thank Dr.NagaPrasad, Mrs.Sailaja, Dr. P.K. Mythili, ,Mr.Abdul Gaffoor, Mr. Mohd Yousuf, Mr. G. Somaraju, Mr.Brian Moss, Ms. A. Mohita,, Ms.Varalaxmi for their technical contribution and continuous help during my course work. Thanks are due to Mr. Ajay joseph. I am grateful to K.D.V. Prasad for his help in cluster analysis, construction of dendrograms and valuable suggestions for finalization of this report.

C. Srinivas Reddy

LIST OF CONTENTS

Page No.

1. OBJECTIVES 6

2. INTRODUCTION 7

3. REVIEW OF LITERATURE 15

Increased understanding and exploitation of gene 16
action

3.1 Foreign or exotic genes 19

3.2 Molecular of markers 20

3.3.1 Protein based markers 22

3.3.2 DNA based markers 23

3.3 Applications of DNA marker technologies 39

3.4.1 Molecular markers for estimating genetic diversity 40

3.4.2 Marker-assisted breeding (MAS) 41

3.4.3 Resistance breeding 42

3.4.4 Map based cloning of agronomically interesting genes 43

4. MATERIALS AND METHODS 44

4.1 RAPD analysis 45

4.2 AFLP analysis 47

5. RESULTS AND DISCUSSION 52

5.1 Detection of variability among Groundnut genotypes using Rapids

5.2 Detection of variability among Groundnut genotypes using AFLP 56

6. REFERENCES

1. OBJECTIVES

- To learn DNA fingerprinting techniques such as Amplified fragment Length Polymorphism (AFLP), Random Amplified Polymorph DNA (RAPD), use of radiolabeled biomolecules in AFLP .

- Review DNA-based molecular markers with particular reference to AFLP and PCR based methods, summarize the protocols used, providing an appreciation of technical difficulty and cost and highlighting the advantages and/or limitations of their use and finally outline the theoretical and practical considerations to be made when choosing a molecular marker or technique.

- Study diversity among potential mapping parents, segregation for markers in mapping populations and evaluate the applications of the above in plant breeding and construction of linkage maps.

INTRODUCTION

Groundnut is the major oilseed crop in India and accounts for 45% of the area and 55% of the production of total oilseeds in the country. In other countries of the region, it ranks either second or third among the annual oilseed crops grown. In India groundnut is grown in three seasons' i.e., rainy (85% area), postrainy (10% area) and summer (5% area). The rainy season groundnut, which is grown during the southwest monsoon period (June-November) is spread over the entire country and is generally rained. The cultivated groundnut (Arachis hypoogea L.) probably originated in Bolivia at the base of the Andes (Krapovickas 1968) extending into North Argentina. Groundnut was sown in 22.02 million ha spread 85 countries with the production of 22.59 million tones and the average productivity of 1.12 t ha. The cultivated groundnut belongs to the Arachis, series amphidiploides and the Fabaceae (Gregory et al. 1973). The species A.hypogea consists of the two subspecies, ssp hypogea and ssp fastigiata.

Knowledge of the groundnut genome is very limited and only in the recent years, have molecular techniques been used to interpret the genome organisation. Extensive variation for morphological and physiological variation has been observed in both wild and cultivated groundnut. Molecular tools such as DNA markers are increasingly becoming important and useful in crop breeding programs. This is necessitated by the presence of polymorphism at DNA level. Abundant polymorphism for wild Arachis sps has been observed, while there is little variation within American cultivars.Very little of low levels of polymorphism was detected using RFLP,RAPD and PCR-form-cutter analysis in

cultivated groundnut germplasm lines (Halward et al. 1991). Such a lack of DNA polymorphism has been reported in other self pollinated crops such as tomato (Helentjaris et al. 1985), melon (Shaattuck- Eidems et al. 1990), and Wheat (Joshi and Nguyen, 1993). Polyploidy coupled with large size of genome , highly self pollinating nature of groundnut with continuous inbreeding and a narrow genetic base would have contributed to low levels of DNA polymorphism in cultivated groundnut. Kochert et al. (1991) indicated that in only known wild tetraploid in this section Arachis of Arachis monticola , it is virtually identical in RFLP pattern to A. hypogaea. No variation in banding pattern was observed among the cultivars and germplasm lines of A. hypogaea using RAPDs techniques (Halward et al. 1992), whereas the wild arachis sps were uniquely identified with most primers tested. Paik- Ro et al. 1992 also reported that DNA polymorphism could not be detected within or between A.hypogaea , A.monticola and the lines of interspecific origin with the 32-endonuclease- probe combination of RFLP . Similarly restriction fragment diversity in self-pollinating species such as wheat, tomato, and soyabean has been much smaller (Sharp et al.1989 ; Helentjaris et al 1985 and Keim et al. 1989). However, cDNA probes were little effective in detecting polymorphism within tetraploid species, than Pst I-genomic clnes. This may be due to multigene families. Little isozyme variation was observed in cultivated in A.hypogaea (Lack and Stalker ,1993; Stalker et al 1994). Thus making a genetic map for cultivated groundnut will be difficult by the different techniques cited above . Lanham et al (1992) using an interspecific tetraploid progeny, of groundnut , TMV-2 (BxC), demonstrated the presence of 7 polymorphic loci using RAPD assay.

Extensive studies have been made on wild, diploid Arachis species to identify DNA polymorphism (Kochert et al. 1992; Lacks and Stalker , 1993; Halward et al 1993). RFLP markers map has been developed for diploid species (Garcia et al. 1995). Although similar results were obtained using RAPD markers , these markers detecting introgressed fragments could not be placed on groundnut linkage map. Although there are difficulties in identification of molecular markers in cultivated groundnut will enhance the breeding capabilities for traits that are difficult to score, through conventional methods. Molecular markers may be used in four types of measurement needed for effective ex situ conservation all of which are useful in resolving the numerous operational, logistic and biological questions that face genebanks managers.these are:

* IDENTITY: the determination of whether an accession or individuals is catalogued correctly is true to maintained properly and whether genetic change or erosion has occurred in an accession or population.

- SIMILARTY: the degree of similarity among individuals in an accession or between accessions within a collection.
- STRUCTURE: the partioning of variation among individuals, accessions, populations, and species. Genetic structure is influenced by insitu demographic factors such as population size, reproduction biology and migration.
- DETECTION: the presence of particular allele or nucleotide sequence in a taxon,genebank accession, insitu population, individual, chromosome or cloned DNA segment

A whole range of different techniques can be used to detect polymorphisms at the DNA level. In fact the seemingly bewildering array of possible approaches is among the first problems faced by newcomers considering the application of these techniques to their own system, In reality, however, this wide array falls into three broad categories with respect to basic strategy: (A) Non-PCR based approaches; (B) PCR Arbitrary priming; and © Targeted-PCR and sequencing.

PCR arbitrary Priming techniques

With the advent of PCR, a number of techniques became available for the screening of genetic diversity. These require no prior sequence-specific information and can therefore be applied directly to any organism. The techniques are based on the use of a single arbitrary primer, which may be purchased from commercial companies in a PCR reaction on genomic DNA and result in the amplification of several discrete DNA products. Each of these products will be derived from a region of the genome that contains two short segments with some homology to the primer, which are on opposite stands, and sufficiently close together for the amplification to work. A number of closely related techniques based on this principle were developed at the same time and are collectively referred to as a multiple arbitrary amplicon profiling (MAAP)(Caetano-Annolles 1994). The most commonly used is RAPD analysis in which the primers are usually 10-mer or 20-mers and in which the amplification products are separated on agarose gels in the presence of ehidium bromide and visualized under ultraviolet light. AP-PCR (Arbitrary primed PCR) and DAF (DNA amplification Finger printing) (Caetano-Annolles et al. 1991) differ from RAPD's and detection of the fragments. In all cases, polymorphisms are detected in

one or both of the primer binding sites. For simplicity, only RAPD's will be referred to in this discussion.

The enormous attraction of RAPD's is that there is no requirement for DNA probes, nor for any sequence information for the design of specific primers. The procedure involves no blotting or hybridisation steps. The technoque is therefore, quick, simple and efficient and only requires the purchase of a thermocycling machine and agarose gel apparatus to set up in a laboratory for any new system under study. It requires small amounts of DNA(10ng per reaction)and sample throughput can be quite high throughput. Rapid's have also been proved to detect higher levels of polymorphism compared with RFLP in cases where the two techniques have been applied to the same material. They have been extensively used for screening diversity, particularly at intraspecific levels, including many population studies (Hadrys et al.1992). Unfortunately, the approach has serious limitations.
The first concerns the nature of the data generated RAPD's are dominant markers such that the homozygous conditions are the only genotypes discernible as presence or absence of the band .In addition the presence of a band of apparently identical molecular weight in RAPD gels of different individuals cannot be taken as evidence that the two individuals have the same band, although this assumption is commonly made. Further complications are that the single RAPD bands can be comprised of several co-migration amplification products and as in the case of DNA finger printing, there can be uncertainty in assigning markers to specific loci in the absence of preliminary pedigree analysis. Although completely unbiased estimators for RAPDs do not appear to be possible, they suggest several steps, which will ensure that the bias is negligible. In their article they derived estimators for gene and genotype frequencies ;within and between population

heterozygosities ;degree of inbreeding ; population subdivision and degree of individuals relatedness. One important conclusion from their study is that to achieve the same degree statistical power using RAPD's compared with co-dominant markers, two to ten times more individuals need to be sampled per locus and further to avoid bias in parameter estimation, the marker alleles for the most of these loci should be in relatively low frequency.

The use of RAPD's for determination the distribution and extent of variation is challenged even further when the second general problem of RAPD's is considered concerning the robustness of the data generated. RAPD's are notoriously prone to user error in that, unless the most consistent of conditions is strictly adhered to the RAPD's profiles obtained can vary considerably between different runs of the same sample. Even within a laboratory which may have different PCR machines or use different sources of polymerase and associated buffers. Even within a laboratory, the item saved by the direct application of RAPDs is often lost in achieving consistency and in confirming the reproducibility of the results obtained. As PCR machines are being improved all the time and new thermostable polymerases continue to appear on the market, it is predictable any particular data from RAPDs cannot be over-emphasized and that together with the statistical qualifications outlined above, these disadvantages of this strategy seriously outweigh the apparent advantages which might otherwise make this the procedure of choice.

Most recently , Keygene have developed a method which is equally applicable universally, which reveals very high levels of polymorphism and which is highly reproducible. This procedure, termed Amplified Fragment Length Polymorphism(AFLP) (Zabeau and vos 1992) is essentially intermediate between RFLPs and RAPDs, in that the first step is

restriction of the genomic DNA but this is then followed by selective rounds of PCR amplifications of the restricted fragments. The fragments are amplified by P33 labeled primers designed to the sequence of the restricted site, plus one of the three additional selected nucleotides. Only fragments containing the restriction site sequence plus the additional nucleotides added on to the primer sequence (upto a maximum of three can be added at either site) the fewer the number of fragments amplified by PCR. This selection is necessary to achieve a total number of fragments within the range that can be resolved on a gel (approximately 150 to 200 fragments). The amplified products are normally seperated on a sequencing gel and visualization after exposure to X-ray film. Recently, the technique has been automated , using fluorescent labeled primers and, therefore high throughput can be achieved Two different types of polymorphism are detected: (1) Point mutation in the restriction sites, or in the selective nucleotides of the primers which result in a signal in one case and absence of a band in the other and (2) small insertions/deletions within the restriction fragment which results in different size bands.

- AFLPs have proven to be different proficient in revealing diversity at below the species level and provide an efficient means of covering large areas of the genome in a single assay. Although we have classified them under arbitrary priming approaches they can be targeted to specific sequences (e.g. VNTRs) if these are used in the primer design. All the evidence use far indicates that they are as reproducible as RFLPs thereby overcoming one of the major problems with RAPDs. They require more DNA (1ug per reaction) and are more technically demanding than RAPDs, requiring experience of sequencing gels, and manually

necessitating the use of radioactivity, but their recent automation and the availability of kits in some species means that the technology can be brought in at a higher level. AFLPs, however, do run into the same problem as RAPDs regarding the type of data generated and the concomitant problems of data analysis for population genetic parameters. Although Keygene are developing means of identifying heterozygotes, AFLPs are essentially a dominant marker system, the identity of the DNA fragment s amplified on the gels is not known, and fragments which migrate to the same molecular weight in the AFLP profile of two different individuals cannot be conclusively interpreted as being the same. Unlike RAPDs, individual bands on an AFLP gel are single DNA fragments (although they may be repeated sequence elements), but the assignment of ales to high and the resultant AFLP patterns very complex. In short AFLPs provide multilocus bi-allelic fingerprints to be subject to considerable analysis by statisticians before the applicability of their data to population analyses can be determined.

3. REVIEW OF LITERATURE

Plant breeding is a process of designing and pursuing a desirable end product (e.g., cultivar, hybrid, synthetic) that represents a compilation of desirable agronomic /economic traits, which may be ranging from simple (qualitative) to complex (quantitative) in their genetic control. Having as much genetic information as possible about major and minor traits and their interactions improves the efficiency and probability of success in achieving an end product with the desired attributes.

Construction of a detailed genetic linkage map for the crop of interest will make available a precise but vast amount of information that plant breeders can use to identify, manipulate, and complement traits to their maximum advantage. In many respects, plant genome maps might be considered analogous to a road map. Specific chromosomes could be thought of as numbered highways and genes located on specific chromosomes comparable to cities and streets within cities. Larger boundaries (country, state, etc.) containing a network of roads may be compared to multiple loci over several chromosomes that govern quantitative traits (QTLs). An effective and efficient way to reach a desired destination is to use a well-developed road map.

It will require several years and a vast amount of resources to map the genomes of major crops. The three major areas having impact on plant breeding are gene action, foreign or exotic genes, and molecular markers.

3.1 Increased understanding and exploitation of gene action

3.1.1 Gene action models

The selection methods that are used by plant breeders for genetic improvement in plant species are developed around the theories and concepts of gene action models. These models have been developed in order to explain the phenotypic expression of traits and the genetic variation observed in populations. The discrete classes of gene action are additive, dominance (including recessiveness, partial, and overdominance), epistasis, and pleiotropy.

In the additive model, the phenotype of the hybrid is intermediate between the two parents. The epistatic gene action model is referred to as interallelic interaction such that the value of alleles present at one locus depends on which allele(s) is (are) present at another locus. In cases where a single gene controls more than one trait, the gene action is referred to as pleiotropy. High-resolution molecular genetic maps should lead to a better understanding and utilization of precise types of gene action and phenomena that result from them.

Transgressive segregation. : Segregants in a F_2 population, whose phenotypic expression for the trait of interest goes beyond one or both of the parents, are referred to as transgressive segregants (Briggs and Knowles, 1967). This is an extremely important phenomenon on which we base many population improvement procedures in cross-pollinated plant species as well as procedures to develop pure-line varieties in self-pollinated species. It is based on additive gene action at individual loci (qualitative traits) and across multiple loci for quantitative traits. The strategy to take advantage of transgresssive segregation is to match parents which posses different "favorable" alleles for the trait(s) of interest so that with crossovers and recombination it is possible to produce progeny that possess the strengths of both the parents with the least

weaknesses of both. With a quantitative trait such as seed yield, several hundred to several thousand progeny have to be evaluated in order to have a reasonable probability of detecting those rare individuals that posses the maximum number of favorable with the minimum number of unfavorable alleles. This is one area where markers linked to QTLs that highly influence the trait would greatly enhance the accuracy and reduce the number of evaluations required to detect progeny with superior gene combinations.

Hybrid Vigor or heterosis: In the early part of this century, maize (*Zea mays* L.) breeders discovered that inbreeding reduced vigor and production of the inbred stocks, but when some combinations of inbreeds were crossed the F_1 hybrid had vigor and production substantially higher than the average of the two parents. Out of these studies the term "hybrid vigor" and subsequently "heterosis" were coined.

It became apparent that there was a strong association between heterozygosity and heterosis. That phenomenon has led to one of the best plant breeding success stories for genetic improvement of crop yields (Duvick,1984). Hybrid varieties have revolutionized corn production substantially in the U.S.A. and the approach has spread to other crops, including. The genetic mechanisms for heterosis are still not clear, but the two most widely accepted theories are dominance and overdominance (Crow, 1964). When inbred lines are crossed together, the F_1 hybrid is heterozygous at all loci for which the genotype of the inbred parents differ. The dominance theory basically states that the different dominant alleles contributed by the inbred parents mask the detrimental effects of the recessive alleles, thus the hybrid has the best strengths of the parents expressed with their weaknesses masked.

The theory of overdominance is that there is an inherent superiority of the heterozygote interaction between the dominant and recessive alleles at each locus) compared to the dominant

homozygote (interaction between the dominant alleles at each locus; Crow, 1964). With the development of genetic maps and genetic markers it should become easier to study the effects of individual as well as sets of genes on the expression of traits (Paterson *et al.* 1991). In return, that information should contribute to a better understanding of the genetic basis of heterotic responses observed and how plant breeders to "design" inbreeds and inbred combinations to further improve performance of hybrids can use that.

Epistasis: As described earlier, epistasis is the interaction between/among alleles at different loci (interallelic interaction). Because of the immediate complexity of the number of combinations of alleles and their effects that are possible with a small number of loci, it has been very difficult to assess epistatic combinations of alleles. For quantitative traits there are many different loci involved in the expression of the trait and there are many interactions taking place to give final expression. Even with genetic maps, it will continue to be very difficult to evaluate large numbers of combinations and the differences elicited with each change. On the other hand, the task will be easier to undertake when the location and functions of genes are better defined. This is where the application of computer and statistical techniques (e.g., Informatics) will greatly facilitate predicted outcomes through stimulating changes of interacting loci and alleles based on gene products and function (Casey 1992). It seems logical that epistatic gene action plays a larger role than we now understand in the final expression of traits, but the degree of complexity will mean that increased understanding will still be slow at best.

Pleiotropy: *It is very difficult to separate Pleiotropy from linkage. Because of the large number of genes contained in crop species and the fact that some of them occur adjacent to each other on a chromosome results in some very tight linkages. These linkages give the appearance that two or more traits controlled by the same gene(s). Very tight linkages necessitate evaluations of a large number of progeny before a crossover type can be detected. With well developed genetic maps, it should be possible to separate some strong associations between traits that are due to linkages of a small number of genes or linkages of QTLs vs. genes that are pleotropic (Paterson et al, 1991). With the information of gene location, function and activation it should be possible to inactivate some genes that are known to control one trait and determine if there is a corresponding lack of expression of the other trait(s). It would be particularly helpful in developing breeding strategies to know if strong associations between desirable and undesirable traits can be broken because they are linked, or cannot because they are due to pleiotropy. In cases of strong associations between two desirable traits pleiotropic control may be better than tight linkage; however, the best strategies to exploit the association would differ with the two scenarios.*

3.2 Foreign or exotic genes

The rapid development of molecular techniques has opened up sources of genes/germplasm to plant breeding that have been unavailable previously through conventional techniques. This is a very exciting and potentially valuable mechanism for crop improvement of the future. Some examples of active research for transferring genes from "foreign or exotic species" are: Bt (*Bacillus thuringiensis*) genes for insect resistance, viral coat protein genes for virus resistance, genes for tolerance to various herbicides, and genes for improved quality of protein. It is easy to

visualize other important agronomic/economic trait possibilities such as: genes for drought tolerance; for tolerance to extreme soil acidity or salinity; transfer potential habit to important annual species, etc. Ideas of transferring genes among species, genera, kingdoms that seemed impossible or too difficult a few years ago are now within the realm of possibility. Once genes are transferred they become a part of the recipient's genome and can be subjected to further modifications and enhancements.

3.3 Molecular markers

In many cases, such as for drought or mold resistance, planned indirect methods of selection (markers) for the traits of interest may be more desirable or effective than direct selection. Some of the reasons for using indirect selection via associated markers may be:

- to identify individuals in early stages of growth for discarding, to conserve resources or to identify individuals for crossing prior to flowering (e.g., backcrossing or population improvement program)
- inaccurate direct measures of the trait expression due to many loci involved (such as QTLs) or due to uneven inoculations/infections/infestations
- difficulties in selecting for several traits simultaneously.

Indirect methods may take the form of morphological markers, biochemical markers (e.g., isozymes), or DNA markers (e.g., RFLP, RAPD, DAF, SSR, AFLP).

In 1865 Mendel determined that genetic factors behave as discrete particles when passed from parent to offspring. His studies on Pea plants marked the beginning of the discipline concerned with the segregation of genes. In the early part of the 20th century, scientists discovered that Mendelian 'factors' controlling inheritance, which we now call as genes, were organized in linear order on cytologically defined structures called chromosomes. Shortly

thereafter the first chromosome map as produced by Strutevant with segregation data derived from studies on Drosophila (Crow and Dove 1988). The markers of this first genetic map were phenotypic traits scored by visual observation of morphological characteristics of the flies.

A major breakthrough occurred when it was realized that genetic maps could be constructed by using pieces of chromosomal DNA as direct markers for segregation pattern of chromosomal segments. In eukaryotes, DNA is condensed with histone and non-histone proteins into thread-like structures called chromosomes. The number of chromosomes varies between species and occasionally within species. At the sub-chromosomal level, several types of organizations are observed. These can be summarized as follows:

- Gene-rich sectors: In large genomes, genes are found clustered in gene-rich sectors especially in regions close to the telomeres. In a number of cases, it is significant that the order of genes, in a sector is conserved between species ('gene synteny'). Genes in a gene-rich sector are interspersed with short repeat sequences, often-transposable elements.
- Tandem repeats: Multiple repeats of essentially the same sequence are found at many locations, especially around the centromeres, telomeres and interstitial locations. These arrays can consist of upto millions of repeat units. Tandem repeats vary according to size and sequence of the repeat unit, the number of repeats found and their distribution throughout the genome. They have therefore received considerable attention as molecular markers.

Thus a molecular marker is a sequence of DNA or a protein which can be readily detected and whose inheritance can be monitored. It is the variation in, or *polymorphism* of, molecular markers, which can be used in genetic diversity studies.

Desirable properties of molecular markers

1. Highly polymorphic behavior.
2. Codominant inheritance (which allows us to discriminate homozygotic and heterozygotic states of diploid organisms.)
3. Occurs throughout the genome.
4. Selectively neutral behavior (no pleotropic effects).
5. Easy, fast and cheap to detect.
6. Reproducible within and between laboratories.

No molecular markers are available yet that fulfills all of these criteria.

3.3.1 Protein based markers

The number of polymorphic morphological markers is limited, especially in intraspecific crosses, and the environment influences their expression. Therefore, more reliable markers such as proteins or, more specifically, allelic variants of several enzymes, so-called isozymes (Tanksley & Orton 1983), other biochemical characteristics, such as lipids or sugars, had to be considered. The multiple forms of an enzyme are of two classes:

- Allozymes: The enzyme is coded for by different alleles *at one* gene locus.
- Isozymes: The enzyme is coded for by alleles *at more than one gene* locus.

For the generation of molecular markers based on protein polymorphism's the most frequently used technique is the electrophoretic separation of proteins on gels and staining.

Interpretation of banding patterns: The principle considerations here are:

- Whether the organism is homozygous or heterozygous at the gene loci.
- The quaternary structure of the enzymes (monomeric, dimeric etc.)

- the number of gene loci
- The number of alleles per locus.

Allozymes are controlled by codominant alleles, which means that it is possible to distinguish between homozygotes and heterozygotes. For monomeric enzymes (i.e. consisting of a single polypeptide), plants homozygous for that locus will produce one band whereas heterozygous individuals will produce two. For dimeric enzymes (i.e. consisting of two polypeptides), plants homozygous for that locus will produce one band whereas heterozygous individuals will produce three owing to random association of the polypeptides. With tetrameric enzymes, heterozygous individuals will produce five bands. For multimeric enzymes, where the polypetides are specified by different loci, the formation of isozymic heteromers can complicate the banding patterns considerably.

Applications

1. Isozyme polymorphism has been used for characterizing/identifying genotypes, for studying population genetics, and for examining geographical patterns of variation.
2. Enzyme electrophoresis has also been very useful in genetic diversity studies, biochemistry, physiology, genetic breeding, etc. as it can directly reveal genetic polymorphism through demonstrating multiple forms of a specific enzyme. Over 30 enzyme systems have been used in plants, and for some crop plants the genes involved have been mapped.

3.3.2 DNA based markers

A major break through occurred when it was realized that genetic maps could be constructed using pieces of chromosomal DNA as direct markers for segregation pattern of chromosomal segments (Bostein et al. 1980). Because each individual's DNA sequence is unique, this

information can be exploited for any study of genetic diversity and relatedness between organisms. A wide variety of techniques to visualize DNA sequence polymorphism have been derived from these techniques.

The term DNA fingerprinting is used to describe a method for the simultaneous detection of many highly DNA loci by hybridisation of specific multilocus "probes" to electrophoretically separated restriction fragments. In recent years, several modifications of the basic technique have appeared and related strategies have been developed. Most importantly DNA polymorphisms became detectable by the Polymerase chain reaction (PCR). Some of the new marker methods are still called DNA fingerprinting, but "DNA profiling", "DNA typing" or more specific terms have also been introduced. According to this definition, DNA fingerprints are mainly obtained by either of two strategies.

- *"classical" hybridization - based fingerprinting* involves cutting of genomic DNA with a restriction enzyme, electrophoretic separation of resulting DNA fragments according to size; and detection of polymorphic multilocus banding patterns by hybridization with a labeled complementary DNA sequence, also called "probe"
- *PCR - based fingerprinting* involves the *in vitro* amplification of particular DNA sequences with the help of specifically or arbitrarily chosen oligonucleotides ("primers") and a thermostable DNA polymerase; the electrophoretic separation of amplified fragments, and the detection of polymorphic banding patterns by such methods as staining.

Restriction Fragment Length Polymorphism (RFLP)

RFLP analysis was one of the first techniques to be used widely to detect variation at the sequence level. It examines the variation in size of specific DNA fragments following digestion

with restriction enzymes. A large number of different restriction enzymes are commercially available. Digestion of a particular DNA molecule with such an enzyme results in a reproducible set of fragments of well-defined lengths. Point mutations within the recognition sequences as well as insertions and deletions will result in an altered pattern of restriction fragments and may thus bring about a screenable polymorphism between genotypes.

This methodology is quite similar to hybridization-based fingerprinting which actually represents a special case of RFLP analysis. Genomic DNA is extracted, digested with restriction enzymes and separated by electrophoresis on a gel. This gel is southern blotted onto a membrane and species specific fragments are made visible by hybridization with a labeled probe.

Analysis of results:

The result is ideally a series of bands on a gel, which can be scored for the presence, or absence of particular bands. Differences between genotypes are usually visualized as an altered pattern of DNA restriction fragments This may result from the point mutations creating or developing restriction sites, or because of reorganization of blocks of DNA, such as deletions or insertions, between restriction sites.

It is clear that the choice of the DNA probe/restriction enzyme combination is crucial in discriminating power of RFLP technology. In general terms, RFLP probes are locus specific - giving rises to easily identifiable codominant markers and is species specific. Traditionally RFLP analysis makes use of the probes obtained from the following sources:

1. Nuclear DNA: these probes are obtained from Genomic libraries & cDNA (complementary DNA) libraries
2. Cytoplasmic DNA: these probes are obtained from mitochondrial DNA and Chloroplast DNA libraries

Advantages of RFLP technique:

- Results are highly reproducible between laboratories.
- RFLP markers usually show co-dominant inheritance.
- Discriminating power - can be at species/population level (single locus probes) or individual level (multi-locus probes)
- Simplicity of the method - given the availability of suitable probes, the technique can be readily be applied to any system.

Disadvantages of RFLP technique:

- Time consuming and expensive to perform - technical expertise required.

- Where no suitable single-locus probes exist, it is time consuming and expensive to identify suitable marker/restriction enzyme combinations from genomic and cDNA libraries.
- Most RFLP work is carried using radioactive labeled probes, and therefore requires expertise in autoradiography. This can be a serious drawback in some situations where special facilities and permits are required to carry out the work.

DNA fingerprinting based on hybridization

The technique of classical DNA fingerprinting is methodologically derived from RFLP analysis and is mainly distinguished from the latter technique by the kind of probe applied to reveal polymorphism's. Two main differences exist between RFLP and hybridization based fingerprinting.

1. DNA fingerprinting makes use of multilocus probes, creating complex banding patterns, whereas RFLP probes are usually locus specific, resulting in an easy to screen co-dominant marker behavior.
2. DNA fingerprinting is mostly performed with non-species specific probes that recognize ubiquitously occurring sequences such as minisatellites, whereas RFLP probes are generally species specific

Two categories of such multilocus probes are mainly used. The first category comprises cloned DNA fragments or oligonucleotides which are complementary to so called "minisatellites" i.e., tandem repeats of a basic motif of about 10 to 60 BP. The second category is exemplified by oligonucleotide probes which are complementary to so called "simple sequences" or "microsatellites" i.e., tandem repeats of very short motifs mostly 1-5 bp.

With both kinds of probes, a high degree of polymorphism between related genotypes is usually observed, which has been exploited for numerous studies in diverse areas of genome analysis.

Molecular markers based on DNA amplification

The analysis of nucleotide sequence variability has been revolutionized by the development of Polymerase Chain Reaction (PCR). This technique allows us to amplify any DNA sequence of interest to high copy numbers, thereby circumventing the need of molecular cloning. Further advancements in this technique has evolved PCR - based markers such as Random Amplified Polymorphic DNA sequences (RAPD) to Amplified Fragment Length Polymorphism (AFLP) and Simple Sequence Repeat markers (SSR) or microsatellites. Currently their potential for use in germplasm characterization and fingerprinting, and ultimately in conservation is widely studied.

Random Amplified Polymorphic DNA (RAPD)

In a large variety of plants and animals it has been shown single arbitrary primers, 8 to 10 nucleotides in length, will produce one to few amplification products (Williams, 1990). The primers are generated with >50% G+C content to ensure efficient annealing, and with sequences that are not capable of internal pairing so as to avoid PCR artifacts. The PCR procedure allows specific amplification of DNA fragments ranging from 200 bp to 3000 bp in length that can be visualized after electrophoresis by staining with eithidium bromide. The key point about this technique is that nothing is known about the identity of the amplification products. The amplification products are however extremely useful as markers in genetic diversity studies. Other important features of the technique are:

- The number of fragments. Many different fragments are normally amplified using each single primer, and the technique has therefore proved a fast method for detecting polymorphism. The majority of commercially produced primers result in 6 to 12 fragments.
- Simplicity of the technique. RAPD analysis does not involve hybridization or autoradiography or high technical expertise. Only minute quantities of target DNA are required. Arbitrary primers can be purchased. Unit costs per assay are low. This has made RAPD analysis very popular.
- RAPD markers are dominant. Amplification either occurs at a locus or it does not, leading to scores of band presence/absence. This means homozygotes and heterozygotes cannot be distinguished.

Problems of reproducibility. RAPD does suffer from sensitivity to changes in PCR conditions resulting in changes to some of the amplified fragments. Reproducible results can be obtained if care is taken to standardize the conditions used.

The various factors, which affect the reproducibility, are:

1. Primer: Primers can be purchased from several manufacturers [e.g., Operon Technologies Inc., U.S.A, UBC, Canada or Pharmacia LKB). Primer concentrations are generally optimal between 0.1 to 2.0 μM. In most species, the majority of RAPD primers result in fragment patterns with 6 to 12 fragments, while a few primers fail to amplify DNA. The G+C content has the highest prediction value; a high G+C content is positively co-related with primer strength.
2. Polymerase: A large number of brands and types of polymerases are available for PCR. Different polymerases often give to different RAPD products. Therefore, the initial choice of

polymerase is important; switching to another type of enzyme is likely to render comparisons with previous experiments impossible.

3. Template concentration: The concentration of the genomic DNA should be determined accurately and the amount of DNA used in the assay should be uniform and well within the experimentally determined reproducibilty ranges (usually 5 to 500 ng).
4. $MgCl_2$ concentration: Strong and reproducible bands are obtained over a wide range of $MgCl_2$ concentrations. A change in concentration often results in a qualitative change of fragment patterns.

Advantages of RAPD technique:

- Fast method for detecting polymorphism's.
- Simple, not technically demanding.
- Relatively cheap to perform (low unit costs).
- Avoids the need for hybridization with radioactive probes.

Disadvantages of RAPD technique:

- Dominant markers.
- Problems with reproducibility - RAPD are sensitive to alterations in PCR conditions.
- Problems with interpreting band patterns e.g., problems of co-migration.

Applications of RAPD markers:

1. Cultivar identification.
2. Genetic mapping.
3. Phylogenetic pedigree and linkage analysis.
4. Population differentiation

5. Estimation of out crossing rates.
6. Identification of duplicates and the establishment of core collections within the germplasm.
7. To determine the extent and role of introgression in the evolution of the species.
8. To detect genetic variations at the intraspecific level between closely related cultivars.
9. Recently it was reported that RAPD primers detected polymorphism among plants generated from tissue culture.

Other techniques using arbitrary primers are

DAF(DNA amplification fingerprinting)

Differences between DAF (Caetano, *et al.* 1991) and RAPD:

- higher primer concentrations in DAF than in RAPD
- Shorter primers are used in DAF (5-8 nucleotides)
- Two - temperature cycle in DAF compared to 3 - temperature cycle in RAPD.
- DAF usually produces more complex banding patterns than RAPD.

AP-PCR - arbitrarily primed polymerase chain reaction

Differences between AP-PCR (Welsh and Mc Clelland, 1990) and RAPD:

- In AP-PCR the amplification is in three parts each with its own stringency and concentrations of constituents.
- High primer concentrations are used in the first PCR cycles.
- primers of variable length, and often designed for other purposes are arbitrarily chosen for use (e.g. M13 universal sequencing primer).

MAAP (multiple arbitrary amplicon profiling) encompasses all these closely related techniques, but which is not commonly used.

DNA markers based on sequence tagged sites

As more information is becoming available from different sources can be located in widely available databases; it can be used for developing new strategies for the analysis of genetic variation. A sequence-tagged site (STS) is a general term given to a marker, which is defined by its primer sequences.

Microsatellite DNA as a Genetic Marker

Microsatellite DNA, also called short tandem repeats (STR) or simple sequence repeats (SSRs), are tandem repetitive DNA sequences with core sequences of two to five base pairs. One example of such a sequence is dinucleotide, (AT) n; where n equals the number of times the sequence of AT is repeated. Microsatellite DNA is flanked by unique and often conserved DNA sequences. The repetitive and conserved DNA sequences as a whole is generally referred to as a variable number tandem repeat (VNTR) locus.

In studying these repetitive regions, Polymerase Chain Reaction (PCR) primers can be developed targeting specifically to the conserved sequences flanking the repetitive region. Southern analysis can also be used by hybridizing clones to the unique region of the locus. Polymorphism in VNTR's may be due to the differences in the number of repeating sequences. More than two alleles are present in a locus. An example is the SAT1 locus found in soybean in which 25 alleles were found at this single locus. In humans as many as 80 alleles have been documented at one locus. Polymorphism in Restriction Fragment Length Polymorphism (RFLP) from either low copy sequence or cDNA clones is often the result of the presence or absence of a restriction site. Thus, in most instances only two alleles exist at a locus. Either a cut occurs which results in a short fragment, or no cut occurs in which a larger fragment is found. Since microsatellites can find more alleles at a locus than RFLP's, former is more informative.

SSRs offer a potentially attractive combination of features that are useful as molecular markers:

- SSRs have been reported to be highly polymorphic in plants, and thus highly informative, providing many different alleles for each marker screened, even among closely related individuals.
- SSRs can be analyzed by a rapid, technically simple, and inexpensive PCR-based assay that requires only small quantities of DNA.
- SSRs are co-dominant and simple Mendelian segregation has been observed.
- SSRs are both abundant and uniformly dispersed in both human and plant genomes.

Microsatellite DNA markers are useful in many types of studies. They can be used in pedigree analysis to determine kinship among individuals, fingerprinting, forensics, genetic-mapping, and phylogenetic analysis. Genetic mapping is used particularly in crop species with low polymorphism such as wheat and soybean. Since microsatellite DNA changes rapidly during the course of evolution, and is not influenced by selection, phylogenetic analysis can be conducted and also can be used as an evolutionary timeclock by measuring the gain or loss of repeats in a genera over evolutionary time and can possibly detect when speciation occurs.

Anchored microsatellite oligonucleotides

Variants of STMS technique have been developed using anchored microsatellite oligonuleotides as primers which direct the amplification of genomic DNA segments other than the repeat region itself. These approaches use oligonucleotides based on a simple sequence repeat (SSR) anchored to their 5' or 3' ends by 2 to 4 arbitrarily chosen nucleotides which trigger site-specific annealing. These initiates PCR amplification of genomic segments, which are flanked by inversely, oriented, closely spaced, repeat sequences. Specifically, intersimple sequence repeat (ISSR) primers are

anchored to their 3' ends and amplify segments between ISSRs. Such anchored microsatellite markers are usually dominant.

Microsatellite sequences are more useful than minisatellites in these and STMS protocols; many minisatellites are too long to allow amplification using current technology and they are not spread as evenly over the genome as microsatellites. However, the core sequences of both types of microsatellites sequence may be used.

An example of STS, based on the RAPD technique, is sequence characterized amplified regions (SCARs). These markers are generated by cloning and sequencing RAPD fragments, which are of particular interest. When the sequence is known, it is then possible to design primers which are longer than usual RAPD primers (24-mer oligonucleotides) and which are exactly complementary to the ends of the original RAPD fragment. When these primers are used in a PCR, single loci are identified which correspond to the original fragment. These loci are called SCARs. SCARs offer several advantages over RAPD and other arbitrarily primed methods, principally that the results are highly reproducible (longer primers used) and the markers are co-dominant.

In another technique called cleaved amplified polymorphic sequence (CAPS) or PCR-RFLP, PCR primers are constructed for a particular locus. The PCR amplified product is digested with a restriction enzyme and visualized on an agarose gel using eithidium bromide staining. As with RFLP, polymorphism's are detected by differences in restriction fragment sizes.

Amplified Fragment Length Polymorphism (AFLP)

The AFLPs were initially named to rhyme with RFLP as "Amplified Fragment Length Polymorphism" but subsequently it was realized that AFLP involves the detection of "presence or

absence" of restriction fragments rather than differences in their lengths. The AFLP approach was developed by a private company *Keyegene* in Netherlands led by Dr. Marc Zabeau, which holds the patent for this technology (Vos *et al.* 1995). The primary reason for the rapid acceptance of AFLP technology is due to its ability to detect a large number of polymorphic DNA markers rapidly and in a reproducible manner. These fingerprints may be used as a tool for determining the identity of a specific DNA sample or to assess the relatedness between samples. Fingerprints are also used as source for genetic markers to generate linkage maps or to identify molecular markers linked to phenotypic traits and/or genetic loci. Polymorphisms detected in DNA fingerprints obtained by restriction cleavage can result from alterations in the DNA sequence including mutations abolishing or creating a restriction site, and insertions, deletions, or inversions between two restriction sites. The DNA polymorphisms identified using AFLP are typically inherited in Mendelian fashion and may therefore be used for typing, identification of molecular markers, and mapping of genetic loci.

The AFLP approach is conceptually simple and combines both RFLP and PCR techniques. The various steps involved are:

Restriction Endonuclease Digestion

To prepare an AFLP template, genomic DNA is isolated digested with two restriction endonucleases simultaneously. This step generates the required substrate for ligation and subsequent amplification.

The restriction fragments for the amplification is generated by two restriction endonucleases: *Eco*RI and *Mse*I. *Eco*RI has a 6-bp recognition site; *Mse*I has a 4 bp recognition

site. When used together, these enzymes generate small DNA fragments that will amplify well and are in the optimal size range (< 1 kb) for separation on denaturing polyacryamide gels. Due to primer design and amplification strategy, these *Eco*RI-*Mse*I fragments are preferentially amplified (rather than *Eco*RI - *Eco*RI and *Mse*I-*Mse*I).

The success of the AFLP technique is dependent upon the complete restriction digestion. Therefore, much care should be taken to isolate high quality genomic DNA, intact without contaminating nucleases or inhibitors.

Ligation of adapters

Following heat inactivation of the restriction endonucleases, the genomic DNA fragments are ligated to *Eco*RI and *Mse*I adapters to generate template DNA for amplification. This common adapter sequences flanking variable genomic DNA sequences serve as primer binding sites on these restriction fragments. Using this strategy, it is possible to amplify many DNA fragments without having prior sequence knowledge.

Amplification reactions

PCR is performed in two consecutive reactions. In the first reaction called preamplification, genomic DNAs are amplified with AFLP primers each has one selective nucleotide. The PCR products of the preamplification reaction are diluted and used as a template for the selective amplification using 2 AFLP primers, each containing 3 selective nucleotides. (The *Eco*RI selective primer is ^{32}P- or ^{33}P-labeled before amplification.) This two step amplification strategy results in

consistently cleaner and more reproducible fingerprints with the added benefit of generating enough templates DNA for thousands of AFLP reactions.

The most important factor determining the number of restriction fragments amplified in a single AFLP reaction is the number of selective nucleotides in the selective primers. Plants having genomes ranging in size from 5 X 10^8 to 6 X 10^9 bp, the number of fragments amplified per sample/ primer pair averages 50, but may range from as low as 10 to ~100 depending on the sequence context of the selective nucleotides, and the complexity of the genome.

A second factor in determining the number of restriction fragments is the C and G composition of the selective nucleotides. In general, the more Cs and Gs used as selective nucleotides in the amplification primers, the fewer the DNA fragments amplified. Also, the smaller the genome being analyzed, the fewer fragments and the simpler the fingerprint.

Separation of amplified fragments on denaturing polyacrylamide gels

Products from the selective amplification are separated on a 6% denaturing polyacrylamide (sequencing) gel. The resultant banding pattern obtained after autoradiography can be analyzed for polymorphism's either manually or using analytical software.

Interpretation of results

Individual band intensity, size distribution of amplified products, and overall pattern should be the same for AFLP analysis with the same primer pairs and the same DNA template, and will vary between different genomic DNA samples and different primer pairs. Fingerprints of related plants

should display common bands, as well ad some differences in banding pattern due to DNA polymorphism's.

The total number of bands, as well as the number of polymorphisms will depend on the crop variety, complexity of the genome and the primer pair used. Some primer pair combinations may result in either too few or too many bands for a particular sample. In case of too few bands, using a primer pair containing fewer Gs and Cs in selective nucleotides is recommended. In the case of too many bands, select a primer pair containing more Gs and Cs in the selective nucleotides.

The primary reason for the superiority of AFLP approach is that detects very large number of DNA bands enabling identification of many polymorphic markers. Routinely about 50-100 bands are observed in each lane of a gel and thus enables rapid creation of very high-density genetic maps rapidly. For instance, in genomes such as barley with large genome with low polymorphism rate, the use of AFLP approach enabled scientists to develop a more informative and enriched genetic map (Becker et al. 1995). The AFLP does not necessarily offer higher rates of polymorphism but is more efficient than RFLP, RAPD or microsatellite approaches of detecting polymorphic DNA. AFLPs detect more point mutations than RFLPs, enable detection of very large number of polymorphic DNA markers than RFLP or RAPDs, and are simpler than microsatellites as no prior sequence information is needed.

The AFLP markers are dominant markers similar to RAPDs but *Keygene* scientists are developing densitometric software that may discriminate between heterozygotes and homozygotes based on allelic density. Imagining software is also being developed by Keygene to analyze the AFLP bands which can be difficult to be done manually. Although AFLP approach is highly

informative, a few criticisms of this technique include the use of multiple procedures, expensive, cumbersome and laborious protocol. Although the use of radioactivity to detect DNA in AFLPs is one major drawback that may limit its use, Guohao He at the Center for Plant Biotechnology Research at Tuskegee University and Dr. Susan McCouch at Cornell University have developed non-radioactive silver staining protocols to detect AFLP markers with no major loss in sensitivity.

3.4 Applications of DNA marker technologies

It is evident that the development of DNA markers has revolutionized the construction of genetic maps in plants and the utilization of genetic maps in studies of plant evolution, systematics, and practical applications such as plant breeding. DNA markers allow direct access to any part of a plant genome, and they liberate researchers from having to deal with plant genes through the fog of phenotype, many steps away from the gene itself. Technology for the utilization of DNA markers is evolving rapidly at the present time, and further advances are sure to occur soon. Some of these will involve making the process of developing and utilizing DNA markers technically simple, less expensive, and more capable of automation. To be practical on a large scale for plant breeding applications, and particularly in developing countries, the detection procedures for DNA markers need to be developed which do not require the use of radio-isotopes, southern blots, DNA sequencing gels and the like. PCR based methods such as RAPD analysis seem to provide part of the answer, but these procedures are still very expensive because of high reagent costs. Simplified DNA analysis seem to be possible with PCR, and even tissue squashes may suffice for DNA isolation (Langridge *et al.* 1991.)

Molecular markers for estimating genetic diversity

Genetic diversity - caused by selection and various mutational and sexual events - rests on genome changes ranging from a single base-pair exchange to rearrangments of entire chromosomes. In closely related genomes, differences may occur every 100 bp (Soller & Beckman 1983). These DNA polymorphism's are exploited by an ever increasing number of molecular marker techniques for the differentiation between individuals, accessions and species of plants, pathogens and pests. Their higher resolution compared with all other markers makes them a valuable tool for varietal and parental identification for the protection of breeder's rights.

DNA markers further add to the repertoire of tools for the determination of the evolutionary relationship between plant species and families. For example, using repetitive DNA (Jung *et al.* 1993) was able to elucidate the evolutionary relationship between several species in the *Beta*.

DNA fingerprinting with minisatellites (Jeffreys *et al.* 1985) or simple synthetic oligonucleotides (Tautz & Renz 1984) has also found widespread application in the differentiation of species. Even a minisatellite-like sequence present in the genome of the M13 phage has been found useful (Rogstadt *et al.* 1988; Weising & kahl 1990). This probe was used to examine the gene flow and genetic diversity in coastal seagrass populations in California, revealing more sexual than clonal propagation in ecologically important and genetically heterogeneous species (Alberte *et al.* 1994) Using human minisatellite probes, molecular taxonomy has possible with crop species such as rice (Dallas 1988), tomato (Brown & Tanksley 1993) and grape (Thomas *et al.* 1993). Microsatellites have also been used in various genera (weising et al. 1989, 1991a), including Brassica (Poulsen et al. 1994), *Beta* (Schmidt *et al.* 1993), *Cicer* (Weising *et al.* 1992; Sharma *et al.* 1995), *Musa* (Kaemmer *et al.* 1992) and *tomato* (Kaemmer *et al.* 1995).

Canadian scientists have used DAF for the management and maintainenece of their genetic diversity. These studies revealed a much higher level of diversity of Douglas fir in coastal and interior regions of Canada than observed in earlier allozyme studies (Carlson *et al.* 1994)

Molecular markers allow the relationships between chromosomes of related species to be determined. By examining the segregation of heterologous DNA markers, chromosomes of different species can be ordered into synthetic groups so that the probes derived from one organism can be used in related organisms. For example, comparative genetic mapping with RFLP markers has shown that tomato and potato are nearly identical in the order of marker loci (Bonierbale *et al.* 1988; Tanksley *et al.* 1992). Conservation of loci has also been found between maize and sorghum (Whitkus *et al.* 1992) and between rice, wheat and barley chromosomes (Ahn *et al.* 1993).

3.4.2 Marker - assisted Breeding (MAS)

The use of molecular markers enables the breeder to connect the gene action underlying a specific phenotype with the distinct regions of the genome in which the gene resides. Once markers for an interesting trait are established, these should allow the prediction of the yield or resistance of individual offspring derived from a cross, solely by the markers distribution pattern in the offspring's genome. Molecular markers then would have considerable impact on breeding economically important crops, because they provide, together with genetic engineering techniques, access to hitherto unavailable genetic resources for crop improvement program's. Besides the exploitation of genomic polymorphisms for germplasm utilization and protection of varieties, the breeder's interest in molecular markers currently focuses on three major issues:

1. The acceleration of the introgression of single resistance genes for plant pathogens such as viruses, bacteria, fungi, nematodes or insects, from wild species or cultivated donor lines into otherwise superior cultivars.
2. The accumulation (pyramiding) of major and/or minor resistance genes into cultivars to generate multiple and more durable (horizontal) resistance's against several pathotypes of the same pathogen.
3. The improvement if the agronomic value of crops by breeding for quantitatively inherited traits, such as yield, fruit solids and protein content, or drought and cold tolerance.

3.4.3 Resistance breeding

The main advantage of using molecular markers for the introgression of resistance genes to cultivars is a gain in time (Tanksley *et al.* 1989; Melchinger 1990). Gene introgression is normally conducted by crossing a resistant donor line with an agronomically superior cultivar, only retaining the desired resistance gene. The use of DNA markers could speed up this process by three plant generations, allowing selection of the resistant offspring that contain the lowest amounts of the donor genome in every generation (Tanksley *et al.* 1989).

Quantitative Trait Loci (QTLs) - a challenge for genomic analysis: Many ergonomically interesting traits, such as yield or tolerance to biotic stresses, are controlled by polygenes, with every gene contributing only a few percent to the expression of the trait. Tagging of polygenes with molecular markers requires a saturated linkage map with a marker spacing of no more than 20 cM and at least 250 F_2 individuals from a cross between parental lines that differ markedly with respect to the trait in question (Paterson *et al.* 1988; Tanksley 1993). First the offspring are tested for the trait and their genotype determined for every marker locus. Then the likelihood that the

observed data rely of the presence of a QTL is calculated, against that no QTL is present, using specially designed computer software such as *MAPMAKER* (Lander *et al.* 1987; Paterson *et al.* 1988).

3.4.4 Map-based cloning of ergonomically interesting genes

The detection and cloning of distinct genes of unknown sequence and function, when only their involvement in specific traits and their chromosomal location is known, has been termed "reverse genetics". In, contrast to conventional approaches, where a gene is cloned on the basis of its known product or sequence and then localized to a chromosomal region, this strategy starts with the localization of a gene on a specific chromosomal region by determining the linkage of the phenotype it specifies to a set of flanking molecular markers. These linked markers are then used as starting points for physically mapping the gene-flanking region with pulsed field gel electrophoresis and rare cutting restriction enzymes. Large-scale restriction site mapping is necessary because physical and genetic distances between markers may vary over several orders of magnitude (Sehgal et al. 1992). This could cause severe problems if the cloning of the region is intended. Physical maps are especially useful in polyploid crops such as soybean, where duplicated sequences could prevent the assignment of markers to a single distinct location (Funke et al. 1993).

The utility of maps and molecular markers will continue to increase. The ability to rapidly construct genetic maps has made possible applications that were unthinkable using conventional mapping techniques. Comparative mapping of different crop plants (Bonierbale *et al*, 1988) will provide useful information about the location of important genes, because it is likely that there will be enough conservation of synthetic blocks so that genes located in one plant will have the same flanking markers in another plant. Comparative mapping of crop plants and their wild

relatives will be a valuable tool for phylogenetic analysis, as well as being useful in introgression studies.

4. MATERIALS AND METHODS

4.1 Plant material

Table 1. List of genotypes used for the study of DNA polymorphism.

Set	Genotype	Comment
1	ICGV 86031	Medium partition, Medium Water use efficiency,Medium duration (100-110days),pest resistant)
2	ICGV 86707	High WUE, Low partition, Med duration (110-120 days)
3	CHICO	High Partition, low WUE, short duration(90-100 days) ,dormant.
4	TAG24	High WUE, high partition.
5	TMV 2	High partition, Med WUE, short duration(100 days), Susceptible to ELS& LLS

DNA isolation

Total plant DNA was isolated from young leaves of field grown groundnut plants. The CTAB method of DNA extraction was followed (Saghai-Maroof *et al.*1984). The frozen leaf tissue was ground well and immediately mixed with homogenizing buffer (1.0 M Tris, pH 8.0, 5.0 M NaCl, 0.5 M EDTA, 2% β- mercaptoethanol, 2% CTAB and incubated at 65°C for 3 hours). This was followed by phenol-chloroform extraction. An equal volume of phenol: chloroform (1:1) was added to the slurry, mixed gently, and centrifuged at 12,000 g for 10 min in a Sorvall RC 2 centrifuge. The upper aqueous layer was transferred to a separate tube and equal volume choloroform: isoamylalcohol (24:1) added, mixed well, and centrifuged at 12,000 g for 10 min. The aqueous layer was again removed and the DNA was precipitated with 0.6 volumes of

isopropanol. DNA was spooled using a glass rod; washed twice with 70% ethanol and suspended in $T_{50}E_{10}$ buffer (50 mM Tris-HCl, 10 mM EDTA, pH 8.0). This dissolved DNA was treated with RNase (100 µl/ml) at 37°C for 1 hour. After the RNase treatment, an equal volume of phenol: chloroform (24:1) was added to the solution, mixed well and centrifuged. The upper aqueous phase was transferred to a separate tube and the DNA was precipitated with 2.5 volumes of absolute ethanol. The DNA pellet was washed with 70% ethanol lyophilized and suspended in $T_{10}E_1$ buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0).

In all the cases the purity and quantity of the DNA samples was estimated using an UV spectrophotometer (Beckman Inst., USA) and with ethidium bromide stained agarose gel.

4.2 RAPD analysis

Step 1: PCR amplification

PCR was performed according to the protocols of (Williams *et al.* 1990). Random decamer primers used in this study and the were purchased from Operon Technologies, Inc., USA. The various primers used for this study are listed in Table 1.

PCR reaction was carried out in 15 µl reaction mixture containing:

* 5.0 µl of DNA (5ng/µl)
* 2.0 µl of 10x PCR buffer
* 2.0 µl of 25mM $Mgcl_2$
* 2.0 µl of Taq DNA polymerase (1 unit)
* 1.0 µl of dNTPs(0.1mM)
* 1.0 µl of RAPD primer
* 5.0µl of DDwater

The amplification was performed in a *Perkin Elmer Gene Amp* PCR system 9600 programmed for 45 cycles (92°C for 1 min for template denaturation, 52°C for 1 min for primer

annealing and 72°C for 2 min for primer extension) using the fastest available temperature transitions.

Step 2: Electrophoresis

The amplified DNA fragments were mixed with 2 µl of 6x loading dye {Xyelene cyanole (25 mg) Ficoll (type 400) (1.5 GM) for 10 ml}. The sample was electrophoresed on 1.2% agarose gels at a constant voltage of 42 v and 70 v for 3-4 hours. The gels were stained with ethidium bromide (5 mg/ml) and photographed under UV illumination.

Step 3: Scoring of gels

The presence of a DNA band was scored as 1 and absence as 0. The polymorphism in an accession was detected as, presence of a band, which is shared with a different accession analyzed.

Step 4: Cluster analysis

Similarity index matrices were generated based on the proportion of common restriction digestion fragments between two genotypes (Nei 1987) using

$$F = \frac{2\,M_{xy}}{M_x + M_y}$$

Where 'F' is de similarity index, M_x is the number of bands in genotype x, M_y is the number of bands in accession y, and M_{xy} is the number of bands common to both x and y. Cluster analysis of data was analysis of the data for 5 groundnut genotypes was carried out using the statistical software package GENSTAT.

4.4 AFLP analysis

AFLP analysis was performed by using AFLP analysis system I kit (Life Technologies Inc, U.S.A) following manufacturer's instructions.

Step 1: Restriction digestion and adapter ligation

Genomic DNA (100 ng in 16 µl) was double-digested with 2 µl *EcoRI* and *MseI*and 5 µl of reaction buffer made to a final volume of 25 µl with AFLP grade water. This was incubated at 37 ^{0}C for 2hours and then denatured the enzymes at 70^{0}C for 15 minutes. The DNA fragments are ligated using 1 unit of T4 DNA ligase and 24 µl of adapter ligation solution (provided along with the kit) at 20 ^{0}C for 2 h.

Step 2: Preamplification

The ligation mixture was diluted 10-fold with sterile distilled water and the fragments were preamplified for 20 cycles using *Perkin Elmer 9600 Gene Amp* system. A total of 5 µl of double-digested and adapter ligated DNA was amplified in a final volume of 50 µl containing 40 µl of pre-amp primer mix (*EcoRI+A and MseI*+C), 5 µl of 10x PCR buffer for AFLP (both provided in kit) and 1 unit of *Taq* polymerase.

The cycle profile was

* 94^{0}C for 30 sec
* 56^{0}C for 60 sec
* 72^{0}C for 60 sec

50 dilution is performed by transferring 3 µl of preamplification mix into a tube containing 147 µl of TE buffer. This is sufficient for 30 selective amplifications. Both unused diluted and undiluted reactions can be stored at -20°C

Step3: Primer labeling

Primer labeling is performed by phosphorylating the 5' end of the *Eco*RI primers with [γ^{32}P or γ^{33}P]ATP and T4 polynucleotide kinase. ^{33}P-labelled primers are preferred because they give better resolution of the PCR products on the gels. Also, the reaction products are less prone to degradation due to autoradiolysis.

Labeling reaction with ^{33}P

* *Eco*RI primer	18 µl
* 5X kinase buffer	10 µl
* [γ^{32}P]ATP(3,000 Ci/mmol)	20 µl
* T4 polynucleotide kinase	2 µl
* Total volume	50 µl

Labeling reaction with ^{33}P

* *Eco*RI primer	18 µl
* AFLP-grade water	10 µl
* 5X kinase buffer	10 µl
* [γ^{33}P]ATP (2,000 Ci/mmol)	10 µl
* T4 polynucleotide kinase	2 µl
* Total volume	50 µl

The reagents were Mix gently and centrifuge briefly to collect the contents to tube. Incubate the reaction at 37°C for 1hr.

Heat inactivates the enzyme at 70°C for 10 min. Centrifuge briefly to collect the contents.

Step 4: Selective amplification

Table 4 Primer combinations used for selective amplification

*Eco*RI	*Mse*I
E-ACA [γ^{33}P]labeled	M-CAA
E-ACA [γ^{33}P]labeled	M-CAC
E-ACA [γ^{33}P]labeled	M-CAG
E-ACA [γ^{32}P] labeled	M-CTG

1 For each primer pair, add the following components to a 1 5-ml microcentrifuge Tube and label it "Mix1"

- Labeled *Eco*RI primer 5 µl
- *Mse*I primer (contains dNTPs) 45 µl
- total volume (sufficient for 20 reactions 50 µl

2. Add the following components to another 1 5-ml microcentrifuge tube and label it as "Mix2"

- AFLP grade water 79 µl
- 10X PCR buffer for AFLP 20 µl
- Taq DNA polymerase (1unit/µl) 1 µl
- Total volume 100 µl

3 Each AFLP amplification is assembled by combining the following in a 0 2 -0 5ml thin-walled microcentrifuge tube

- diluted template DNA 2 5 µl
- Mix1(primers/dNTPs) 2 5 µl
- Mix2(Taq DNA polymerase/buffer) 5 0 µl
- Total volume 10 0 µl

4. Mix gently and centrifuge briefly to collect reaction.

PCR temperature profile was, one cycle at 94° C for 30 sec, 65° C for 30sec and 72°C for 60 sec, followed by lowering the annealing temperature during each cycle by 0.7°C for 12 cycles. This gives a touch down phase of 13 cycles. Then the reaction was amplified for 23 cycles at 94°C for 30 sec, 56°Cfor 30 sec and 72° C for 30 sec

Step 5: Electrophoresis

Following amplification, reaction products were mixed with an equal volume of formamide dye (98% deionised formamide, 10 mM EDTA (pH 8 0), 0 025% bromo phenol blue and 0.025% xyelene cyanol as tracking dyes) The resulting mixtures were heated for 5 min at 90°C and then quickly cooled on ice Each sample (6µl) was loaded on a 6% denaturing (sequencing) polyacrylamide gel (20 1 acrylamide bis, 7 5 M urea, 1xTBE buffer) Electrophoresis was performed at constant power of 1500 volts for 2h After electrophoresis, gels were dried and autoradiographed

Step 6: Scoring of gels:

The presence of DNA band was scored as 1 and absence as 0 The polymorphism in an accession was detected as the difference between the distance traveled by DNA band on agarose gel

Table: Comparision of AFLP with DAF & RAPD markers

Characteristic	AFLP	RAPD	DAF
Principle involved	PCR amplification & restricted digestion	PCR amplification with random primers	PCR amplification with random primers
Type of polymorphism	Single base changes; insertions; deletions	Single base substitutions insertions; deletions	Amplification of base pairs with insertions
Genomic abundance	Low	High	High
Level of polymorphism	Very high	Medium	High
Inheritance	Dominant	Dominant	Dominant
Amount of DNArequired	5 ng	10-25 ng	50-100ng
Sequence information required	Yes	No	Yes
Radioactive detection required	Yes	No	Yes
Development costs	High	Medium	High
Start up costs	High	Low	High
Detection	Autoradiography; Silver staining	Ethidium bromide; silver staining; fluorescence	silver staining ; Autoradiography, fluorescence

5. RESULTS AND DISCUSSION

5.1 Detection of variability among Groundnut genotypes using RAPDs

Five Groundnut lines ICGV 86031, ICGV 86707, Chico, TMV 2, and TAG 24, were screened with 48 oligonucleotide primers. A total of 96 amplified DNA fragments were identified when PCR amplified products were separated on 1.2% agarose gels, out of these 10 were polymorphic. Five Operon primers GN-39, B-11, B-13, V-4, and Gn-20 showed good polymorphism. Overall the use of 5 selected primers produced an average of 19.2 bands/primer/genotyrpe out of which 2bands were polymorphic(10.4%). Among the five primers, the primer GN-39 was better than other five in distinguishing may of the groundnut genotypes. Cluster analysis was carried out our using the data obtained by screening five groundnut genotypes with five primer pairs. A dendrogram based on RAPD placed the genotypes into two groups based on banding pattern. The genotypes Chico and TAG 24 formed a distinct group, where as the genotypes TMV 2, ICGV 86707 and ICGV 86031 formed a separate group which can be divided into two sub groups; Sub-Group I ICGV 86707 and ICGV 86031, & TMV 2 in other Sub-Group II. The genotype TMV 2 is more diverse when compared other genotypes as this genotype amalgamated far from others.

RAPD analysis is simple and fast; it involves PCR amplification followed by gel electrophoresis of genomic DNA It requires very little amount of DNA (25 ng per reaction) and analysis is free from radioactive materials. As the primers used are of 10 bp length, the conditions for PCR amplification such as annealing temperature, concentration of $MgCl_2$ and dNTPs and G+C content of the primers are crucial to get reproducible results.

Table 5. RAPD data of 5 Groundnut genotypes:

		Groundnut genotypes					
Primer	Distance migrated from the well	ICGV 86031	ICGV 86707	Chico	TAG 24	TMV 2	Remarks
OPB 11	1.2	1	1	1	1	1	N
	1.3	1	1	1	1	1	N
	1.4	1	1	1	1	1	N
	1.5	0	0	1	0	0	P
	1.7	1	1	1	1	1	N
OPB 13	1.3	1	1	1	1	1	N
	1.4	1	1	1	1	1	N
	1.6	1	1	1	1	1	N
	1.7	1	1	1	1	1	N
	1.9	0	0	0	1	0	P
OPGN 39	1.0	1	1	1	1	1	N
	1.1	1	0	1	1	1	P
	1.4	1	0	1	1	1	P
	1.6	1	1	1	1	1	N
	1.7	0	0	0	0	1	P
OPGN 20	1.5	1	1	1	1	0	P
	1.8	1	1	1	1	0	N
	1.9	1	1	1	1	1	N
	2.1	1	1	1	1	1	N
	2.4	0	0	0	0	1	P
OP V4	2.5	0	1	1	0	0	P
	3.1	0	1	1	1	0	P
	3.6	1	1	1	1	1	N
	3.7	1	1	1	1	1	N
	3.9	0	0	0	1	0	P

N = Not polymorphic; P = Polymorphic

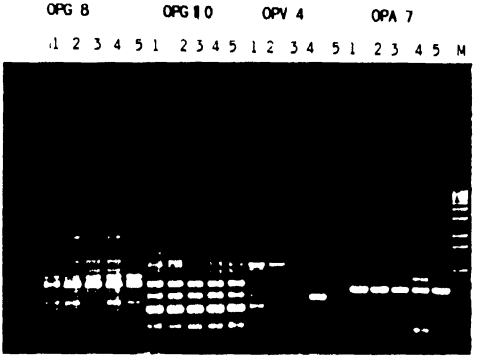


Figure 1.
RAPD profile of 5 Groundnut genotypes
Lane 1: ICGV86031; 2: ICGV 86707 ; 3: CHICO; 4: TAG 24; 5: TMV-2
M: Lambda Hind III marker

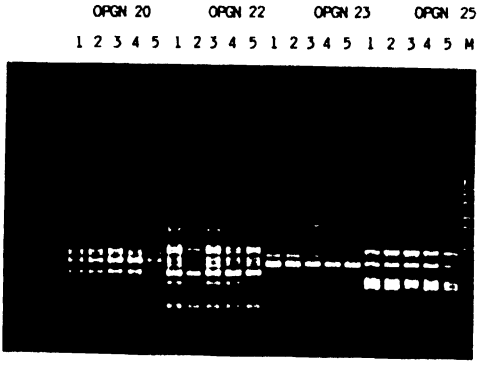


Figure 2.
RAPD profile of 5 Groundnut genotypes
Lane 1: ICGV86031; 2: ICGV 86707 ; 3: CHICO; 4: TAG 24; 5: TMV-2
M: Lambda Hind III marker

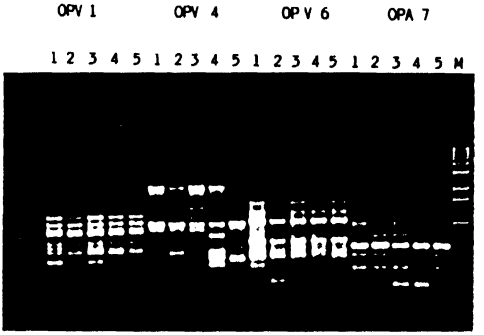


Figure 3.
RAPD profile of 5 Groundnut genotypes
Lane 1: ICGV86031; 2: ICGV 86707 ; 3: CHICO; 4: TAG 24; 5: TMV-2
M: Lambda Hind III marker

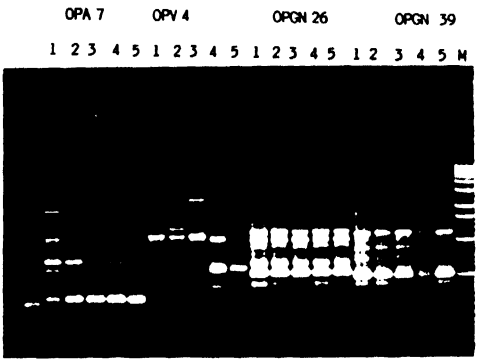


Figure 4.
RAPD profile of 5 Groundnut genotypes
Lane 1: ICGV86031; 2: ICGV 86707 ; 3: CHICO; 4: TAG 24; 5: TMV-2
M: Lambda Hind III marker

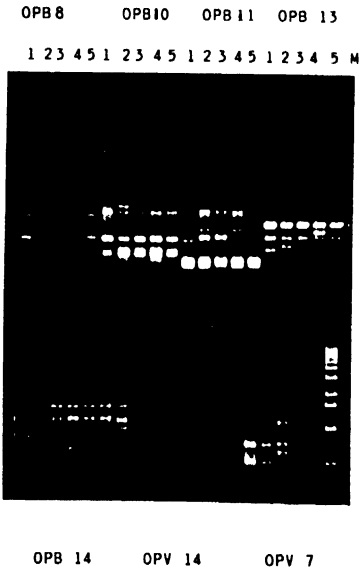


Figure 5.
RAPD profile of 5 Groundnut genotypes
Lane 1: ICGV86031; 2: ICGV 86707 ; 3: CHICO; 4: TAG 24; 5: TMV-2
M: Lambda Hind III marker

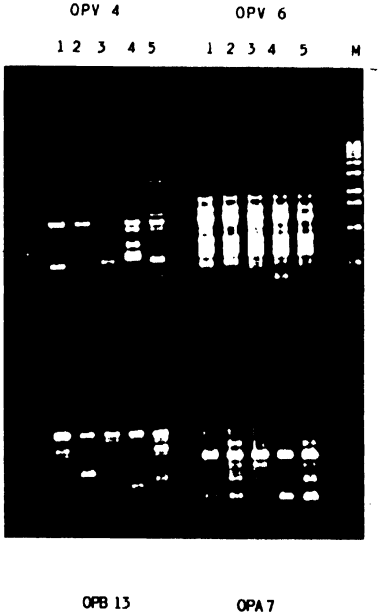


Figure 6.
RAPD profile of 5 Groundnut genotypes
Lane 1: ICGV86031; 2: ICGV 86707 ; 3: CHICO; 4: TAG 24; 5: TMV-2
M: Lambda Hind III marker

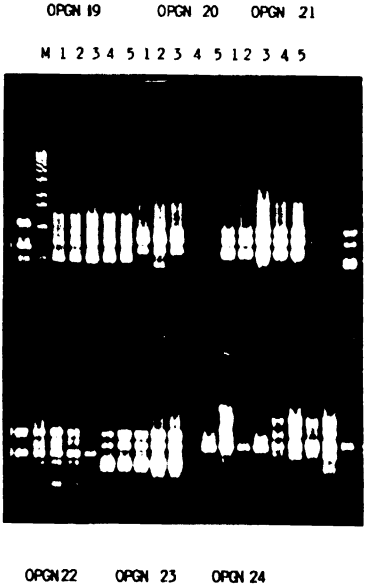
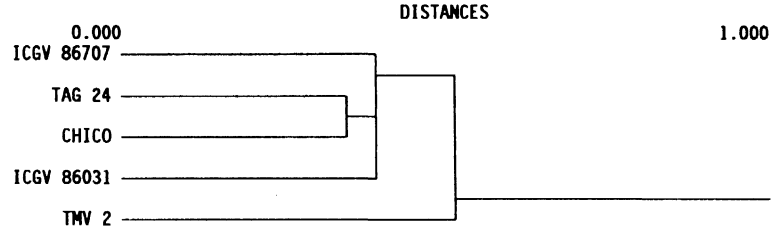


Figure 7.
RAPD profile of 5 Groundnut genotypes
Lane 1: ICGV86031; 2: ICGV 86707 ; 3: CHICO; 4: TAG 24; 5: TMV-2
M: Lambda Hind III marker

DISTANCE METRIC IS EUCLIDEAN DISTANCE
SINGLE LINKAGE METHOD (NEAREST NEIGHBOR)

TREE DIAGRAM



SIMILARITY MATRIX COMPUTED, LENGTH = 15

***** SIMILARITY MATRIX *****

1	----				
2	76.5	----			
3	85.3	85.3	----		
4	85.3	79.4	88.2	----	
5	73.5	61.8	70.6	70.6	----
	1	2	3	4	5

Similarity matrix of 5 groundnut genotypes produced by cluster analysis

1 ----
2 76.5 ----
3 85.3 85.3 ----
4 85.3 79.4 88.2 ----
5 73.5 61.8 70.6 70.6 ----

1 2 3 4 5

1: ICGV 86031; 2: ICGV 86707; 3: Chico; 4: TAG 24 and 5: TMV 2

Table: Primer code, sequence and polymorphic bands with percentages

Primer Code	Primer Sequence	Anneal Temp (°C)	No. of Bands	No. of Bands Polymorphic	% Bands Polymorphic
OPB - 11	GTAGACCCGT	72	22	1	4.7
OPB - 13	TTCCCCCGCT	72	21	1	4.7
OPGN - 20	TCCCGACCTC	72	19	3	15.7
OPGN - 39	GACAGGAGGT	72	19	2	10.5
OPV - 4	CCCCTCACGA	72	16	3	18.7

Table: Sequence of primers used for RAPD analysis

Code	Sequence	Code	Sequence
OPA-07	5'GAAACGGGTG		
OPA-09	5'GGGTAACGCC	OPV-12	5' ACCCCCCACT
OPA-10	5'GTGATCGCAG	OPV-07	5' GAAGCCAGCC
OPA-13	5'CAGCACCCAC	OPGN-09	5' GGTCACCTCA
OPA-18	5'AGGTGACCGT	OPGN-12	5' TATGCCGCCA
OPA-19	5'CAAACGTCGG	OPGN-19	5' AGAGATGCCC
OPA-20	5'GTTGCGATCC	OPGN-20	5' GTTTCCGGTG
OPB-04	5'GGACTGGAGT	OPGN-22	5' AGAGATGCCC
OPB-02	5'TGATCCTGGG	OPGN-23	5' GTTAGCGGCG
OPB-05	5'TGCGCCCTTC	OPGN-25	5' GGCTCGTACC
OPB-06	5'TGCTCTGCCC	OPGN-26	5' CCACAATGGG

OPB-08	5'GTCCACACGG	OPGN-27	5' GGTGGGTGCT
OPB-10	5'CTGCTGGGAC	OPGN-28	5' GTGCGGACAG
OPB-11	5'GTAGACCCGT	OPGN-29	5' GGACGGGTGC
OPB-13	5'TTCCCCCGCT	OPGN-31	5' CCACGAGCAT
OPB-14	5'TCCGCTCTGG	OPGN-33	5' CGGGAGACCC
OPB-16	5'TTTGCCCGGA	OPGN-34	5' GCATGGAGCT
OPV-01	5'TGACGCATGG	OPGN-39	5' GACAGGAGGT
OPV-03	5'CTCCCTGCAA	OPI – 20	5' AAAGTGCGGG
OPV-04	5'CCCCTCACGA	OPI – 18	5' TGCCCAGCGT
OPV-06	5'ACGCCCAGGT	OPG –08	5' TCACGTCCAC
OPV-08	5'GGACGGCGTT	OPG –10	5' AGGGCCGTCT
OPV-06	5'ACGCCCAGGT		

5.2 Detection of variability among groundnut genotypes using AFLP markers

Amplified Fragment Length Polymorphism (AFLP) kit supplied by Gibco-BRL worked well for detecting the polymorphism in 10 Groundnut genotypes. Analysis of 10 genotypes of Groundnut with 4 AFLP primers pairs (E-ACA, M-CAA; E-ACA, M-CAC; E-ACA, M-CAG; E-ACA M-CTG) identified a total of 76 fragments, of which 11 were polymorphic. The first primer pair is used for the 5 wild species which showed high polymorphism and the markers are difficult to scorable. Of this the primer combination E-ACA-M-CTG showed more polymorphic than the other primer combinations. Cluster analysis was carried out using the data obtained by screening 10 genotypes with 3 primer combinations. A dendogram based on AFLP markers placed the genotypes into 3 major groups based on the pattern. The genotypes JL 24(group I), ICGV86448, KADIRI 3, ICGV 92209, ICGV 86325 form (group II), ICGV93044, ICG 4906, ICGV 44 ICG1171, TMV10 form (group III) . Of these second major group is divided into 2 subgroups. Subgroup I includes ICGV86448 and KADIRI 3, Subgroup II includes ICGV92209 and ICGV 86325. The third major group was divided into three subgroups. Subgroup I includes ICGV93044, ICG4906, Subgroup II includes ICGV44 &ICG1171 and last Sub group III TMV10. The cluster analysis indicate the genotype JL24 is more diverse as it was amalgamated far from the other genotypes.

The cluster analysis is done for other different genotypes 10 genotypes when screened with one primer combination E-ACA-M-CTT showed of total of 37 bands and of these 7 polymorphic markers are detected and Dendogram is generated based on the banding pattern. The genotypes are placed in three major groups group I includes ICG 6280, CHICO, ICG 86031, TMV10, group

II includes ICG5094, ICG8263, group III includes ICG1712, ICG 15222, ICG4906 & TMV2NLM. Of this the Second major group is divided into two subgroups. Subgroup I (ICG5094&ICG 8263) and the other subgroupII ICG1712 . From the above all the genotypes are diversified and are differing from each others . one primer corresponding to the EcoRI adapter and four primers corresponding to MseI primers was tested in 4 combinations of peanut genotypes. The DNA polymorphism was detected in the peanut with all the primer pairs tested. Theses 4 AFLP primer pairs cummulatively detected 18 polymorphic loci (Table 1a) and this rate is comparable to the 10 markers detected by RAPD approach from 48 primers. An average of 27.5 bands per primer pair were detectable and 4.5 bands per primer pair were polymorphic (16.3%) . Among the AFLP primers corresponding to the EcoRI adapter E-ACA with MseI primer CTG & CTT showed superior in identifying polymorphism in peanut. Two AFLP gels were shown in the figure 1& 2 , and in the first AFLP gel 5 wild species are repeated with E-ACA-MCTT and the bands are difficult to score because of high markers. The AFLP approach was more efficient in detecting DNA polymorphism in peanut as 22.2% of the AFLP primer pairs identified polymorphism compared to 10.4% of the RAPD primers. However , both AFLP and RAPD showed fairly simlar levels of polymorphism an average of 4.5 polymorphic markers per primer pair in AFLP versus 2polymorphic markers per primer in RAPD. But with the AFLP primer every one of it showed polymorphic markers, and · have compared three genotypes CHICO, ICGV86031, TMV2 NLM with the both AFLP and RAPD and found more polymphic markers are detected with AFLP . However, AFLP approach is generally more tolerant to PCR reaction conditions and thus higher reproducible rate compared to RAPD and DAF procedures. Lack of detectable DNA markers in peanut has hindered marker assisted genetic studies of this crop. Halward et al. (1991 and 1992) employed 16 primers and found no polymorphism among

cultivated peanut although the wild species displayed considerable diversity. After that Prakash et al.(1996-97) study demonstrates for the first time demonstrated that DNA variation among cultivated species using different primers with AFLP and DAF proceedures. My study demonstrates that DNA variation exists among peanut genotypes and can be detected using AFLP and RAPD techniques including the use of informative primers, and although this study has identified polymorphism but level of genetic variation is very low when compard to other crops. Because of earlier studies it is desairable to construct a genetic map of the cultivated peanut , especially alien (wild) species have not been used much in breeing . Development of genetic map also may enrich the existing map of Arachis and thus facilitate an accelerated improvement of this crop. Availability of molecular markers enable to detail invesigation of the peanut genome with immediate practical applications in cultivar identification through DNA finger printing , genetic diversity and in understanding the domestic history of this crop.

Table-1. Groundnut genotypes used for AFLP analysis:

S.No	Accession No	Traits
1	ICGV 92209	Succeptible to Rosette virus
2	Kadiri	Low yield
3	ICGV 86325	High yield
4	ICGV 88448	High O/L ratio
5	JL 24	Low O/L ratio
6	TMV 10	High oil
7	ICG 1171	Low oil
8	ICGV 93044	High seed mass
9	ICG 4906	Low seed mass
10	ICGV 44	Succeptible to rust
11	ICGS 44x A.batizicoi	Rust
12	A.batizicoi	Rust
13	A.correntina	Rust
14	A.duranensis	Rust
15	ICGV 44x A.batizicoi	Rust

Table. 8. AFLP analysis of 10 groundnut genotypes

Primer combination I E-ACA-M-CTG				Primer combination II E-ACA-M-CAC			Primer combination III E-ACA-M-CAG			
DISTANCE MOVED FROM THE WELL										
	cm									
Sample No	4.0	5.0	7.3	12.0	18.5	4.2	8.8	14.8	4.5	8.5	11.0
1	0	0	1	1	1	1	0	1	1	1	1
2	0	0	1	1	1	1	0	1	0	1	1
3	1	0	1	1	1	1	0	1	1	1	1
4	0	0	1	1	1	1	1	1	0	1	1
5	1	0	1	1	0	1	1	1	1	0	0
6	0	1	0	1	0	0	1	0	1	1	1
7	1	1	1	0	0	0	1	0	0	1	1
8	0	0	1	1	0	0	1	0	1	1	1
9	0	0	1	1	0	0	1	0	1	1	1
10	1	1	1	1	1	1	1	0	1	1	1

Table 1a: Number of detectable and polymorphic DNA fragments among AFLP primer Combinations detecting DNA polymorphism in cultivated peanut.

	EcoRI primer	
MseI primer	ACA	% polymorphic Loci
CTG	30(5)	16.6
CAC	18(3)	16.6
CAG	25(3)	12
CTT	38(7)	18.4

Table-9: AFLP profile for 10 genotypes

Primer Combination E-ACA-M-CTT

	Distance moved from the well						
	Band I	Band II	Band III	Band IV	Band V	Band IV	Band VII
Sample no	3.0cm	3.5cm	4.5cm	7.0cm	11.0cm	14.0cm	19.5cm
1	0	0	1	1	1	1	1
2	0	0	0	0	0	0	1
3	0	0	1	1	0	1	1
4	0	0	1	1	1	1	1
5	1	1	1	1	1	1	1
6	0	0	1	1	1	0	0
7	1	1	1	1	1	0	0
8	1	0	0	1	1	0	0
9	0	1	1	0	1	0	0
10	0	1	1	0	0	0	0

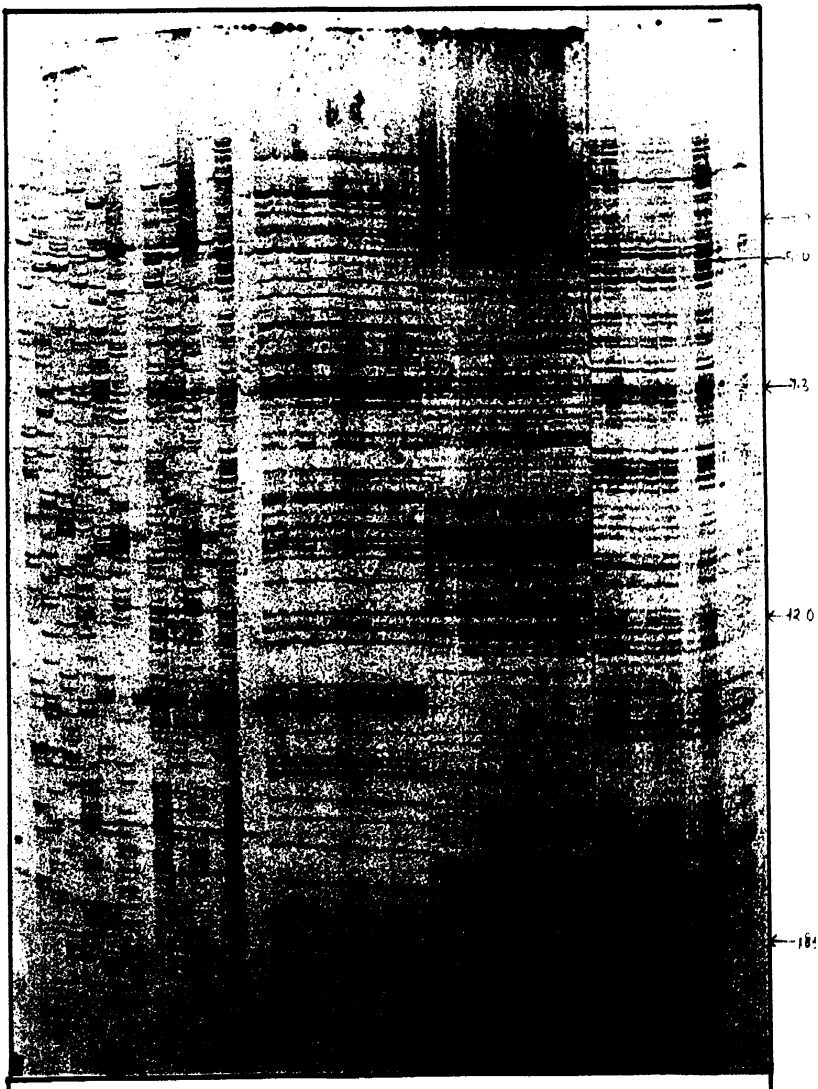


Figure 1. AFLP profile of the 10 groundnut genotypes of table 1.
Primer combination used E-ACA-M-CAA, E-ACA-M-CAC, E-ACA-M-CAG, E-ACA-M-CTG

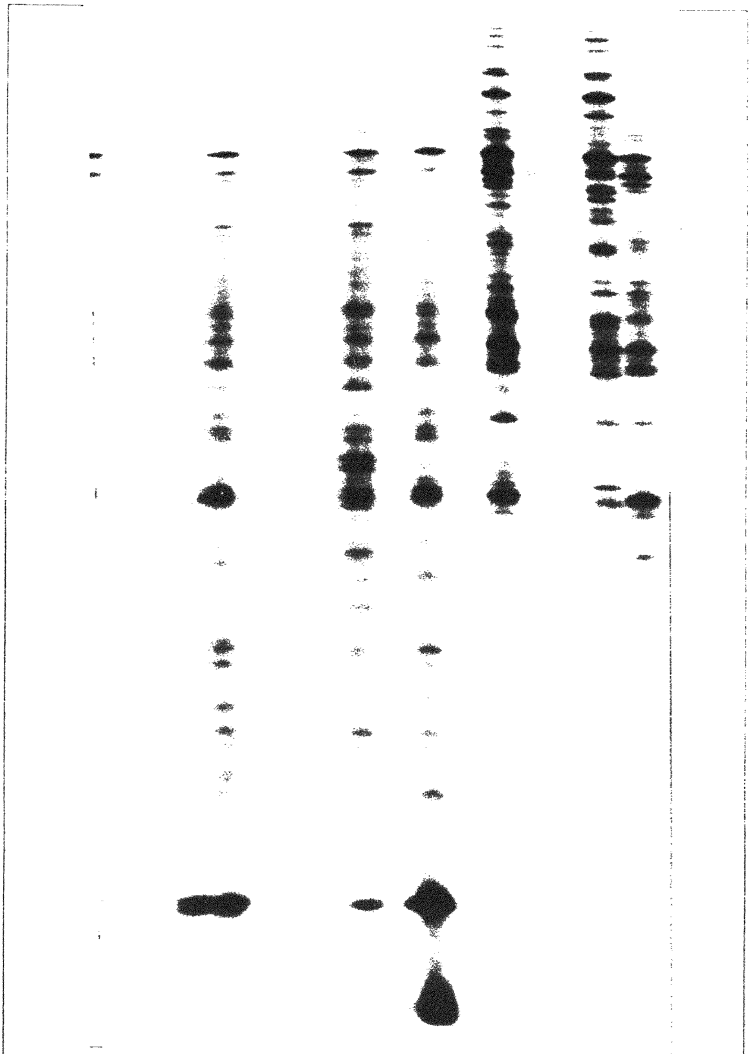


Figure 2:
AFLP Profile of 10 genotypes -1: ICG1712; 2: ICG6280; 3: ICG8263; 4:ICG5094 5: ICG15222; 6: TMV10; 7: ICG4906, 8: ICG 86031; 9: TMV2 NLM; 10:CHICO
Primer Combination Used : E-ACA-M-CTT

DISTANCE METRIC IS EUCLIDEAN DISTANCE
SINGLE LINKAGE METHOD (NEAREST NEIGHBOR)

TREE DIAGRAM

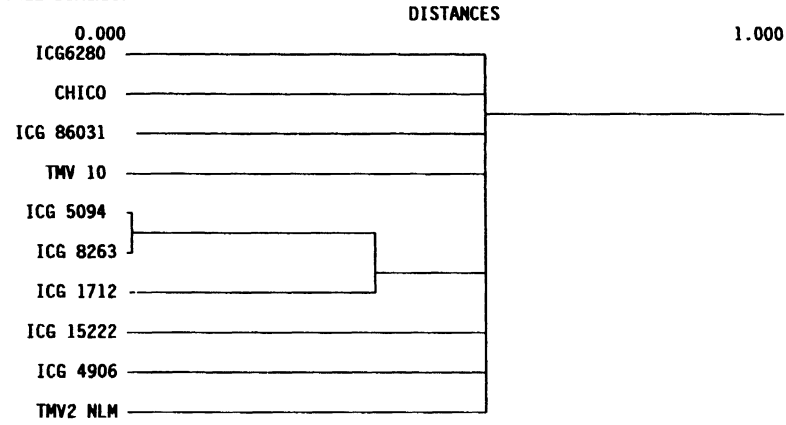


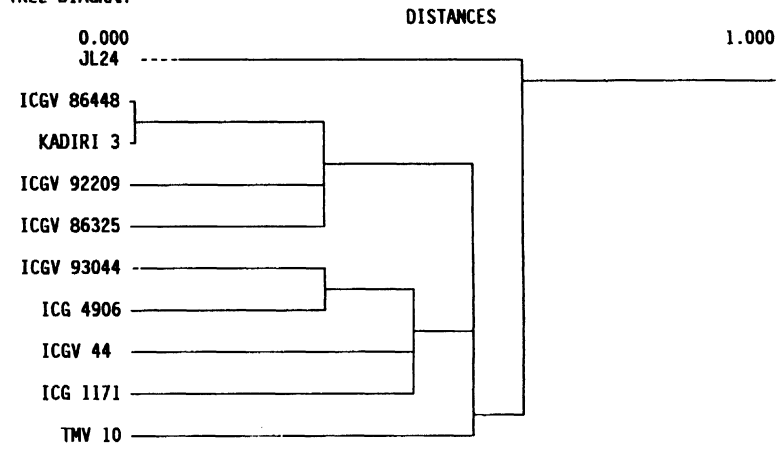
Figure 5.2.1 Dendrogram based on AFLP data of 10 groundnut genotypes using the primer combination E-ACA-MCTT

***** SIMILARITY MATRIX *****

1	----									
2	42.9	----								
3	85.7	57.1	----							
4	100.0	42.9	85.7	----						
5	71.4	14.3	57.1	71.4	----					
6	71.4	42.9	57.1	71.4	42.9	----				
7	42.9	14.3	28.6	42.9	71.4	71.4	----			
8	42.9	42.9	28.6	42.9	42.9	71.4	71.4	----		
9	42.9	42.9	28.6	42.9	42.9	71.4	71.4	42.9	----	
10	42.9	71.4	57.1	42.9	14.3	71.4	42.9	71.4	42.9	----
	1	2	3	4	5	6	7	8	9	10

DISTANCE METRIC IS EUCLIDEAN DISTANCE
SINGLE LINKAGE METHOD (NEAREST NEIGHBOR)

TREE DIAGRAM



SIMILARITY MATRIX COMPUTED, LENGTH = 55

Figure 5.2.2 Dendrogram based on AFLP data of 10 groundnut genotypes using primer combinations E-ACA-M-CAC; E-ACA-M-CAG; and E-ACA-M-CTG

***** SIMILARITY MATRIX *****

1	----									
2	90.9	----								
3	90.9	81.8	----							
4	90.9	100.0	81.8	----						
5	54.5	45.5	63.6	45.5	----					
6	36.4	45.5	27.3	45.5	27.3	----				
7	45.5	36.4	54.5	36.4	54.5	72.7	----			
8	72.7	63.6	63.6	63.6	45.5	63.6	72.7	----		
9	72.7	63.6	63.6	63.6	45.5	63.6	72.7	100.0	----	
10	63.6	54.5	72.7	54.5	54.5	54.5	81.8	72.7	72.7	----
	1	2	3	4	5	6	7	8	9	10

5.3 Silver Staining

The DNA Silver Staining System provides the reagents required to visualize short tandem repeat (STR) products that have been amplified with the GenePrint(TM) STR System and separated by denaturing polyacrylamide gel electrophoresis. The DNA Silver Staining System detects bands in a polyacrylamide gel using sensitive silver stains protocol (1). This system eliminates the expense and hazards associated with the use of radioisotopes. The simple staining process can be completed in 90 minutes. Multiple film prints can be made immediately following staining or after the gel has air-dried overnight.

The DNA Silver Staining System contains all reagents necessary to analyze 10 sequencing size gels, each 31cm x 38cm x 0.4mm (width x height x thickness). The protocols detailed in this manual describe polyacrylamide gel preparation, gel electrophoresis, silver staining and exposure of film.

Each component of the DNA Silver Staining System is packaged for the convenient preparation of two liters of each required solution. However, the system may be used to stain smaller gels by decreasing each solution volume to one liter. To do this, use 10g of silver nitrate to prepare the staining solution and store the remaining 10g for future use. Open one ampule of the 37% formaldehyde (3ml), and use 1.5ml for the staining solution and 1.5ml for the developer solution. Open one ampule of sodium thiosulfate (10mg/ml), and use 200µl for the developer solution. Store the remaining volume for future use.

1. Water quality is extremely important for good staining. Use ultrapure water (e.g., Milli-Q®) or double distilled water. If the water contains contaminants, the gel may not develop or the top bands on the gel may be the only bands that appear and the bottom of the gel will be blank.

2. The purity (grade) of the sodium carbonate contained in the developer solution is critical. We recommend using anhydrous sodium carbonate, A.C.S. certified, from Fisher Scientific (Cat # S263-500).

3. Perform steps involving formaldehyde solutions in a chemical hood

4. Chill the developer solution to 4-10°C.

5. Save the fix/stop solution from Step A.4 a, below, to use in Step A 4.g.

6. Use 2 liters of each solution per gel for each step (for a 54 1 x 43.5 x 13cm tray).

7. Gently agitate during each step.

8. The duration of the rinse step (Step A.4 e, below) is critical. The total time from immersion in deionized H_2O to immersion in developer solution should be less than 20 seconds. If the deionized H2O rinse step exceeds 20 seconds, repeat the staining procedure, beginning with Step A.4 d.

9. After use of the staining solution, waste silver can be recovered from the solution for recycling. Silver is precipitated as AgCl by the addition of NaCl. This material can be collected by filtration or allowed to settle out by gravity.

Procedure

1. After electrophoresis, empty the buffer chambers and carefully loosens the gel clamps. Remove the glass plates from the apparatus.

2. Place the gel (glass plates) on a flat surface. Remove the comb and the side spacers. Use a plastic wedge to carefully separate the two glass plates. The gel should be strongly affixed to the small glass plate.

Note: Using a metal spatula to pry the plates apart may damage the plates.

3. Place the gel (attached to the small plate) in a shallow plastic tray (e.g. NALGENE® wash tub).

4. To silver stain, follow the steps listed below:

Step	Solution	Time
a.	fix/stop solution	20 minutes
b.	deionized H2O	2 minutes
c	repeat Step 2, twice	2 x 2 minutes
d.	solution	30 minutes
e.	deionized H2O	10 seconds (Critical step, see Note V.8, above)
f.	developer solution (4-10°C)	2-5 minutes (Until alleles and ladders are visible)
g.	fix/stop solution**	5 minutes
h.	deionized H2O	2 minutes

**Add directly to developer solution to stop developing reaction

5. Position the gel (small plate) upright and allow to dry overnight

Note: To create film prints of the gel immediately, cover the gel with plastic wrap and proceed to Section VI.

VI. Exposure of Film

A direct image may be produced using STR Electrophoresis Duplicating Film (EDF). Use of this film allows the generation of multiple permanent images with more control over band and background intensity than does development of the gel alone. A relatively light silver-stained gel and relatively dark film exposure generally yield better results than a dark silver-stained gel and a light film exposure.

Procedure

1. .Place the gel (attached to the small plate) on a white light box.

2. In the dark (or with a safelight), find the notched edge of the film and position the film on the gel with the notched edge on top, with the notch closer to the upper left corner.

Note: EDF is sided, and it is important that the notch be positioned as described. From this orientation, the film may be rotated, if desired, but not turned over.

3. Turn on the white light box and expose the film for 10 to 60 seconds, depending on the gel background level and the intensity of the white light. (This step must be optimized for individual light boxes.)

4. Develop the film manually using procedures for development of autoradiographic film.

Note: The film may jam in some automatic film processors.

The following were used

a. - 3 minutes in Kodak® GBX Developer
b. - 1 minute wash in water
c - 3 minutes in Kodak® GBX Fixer
d - 1 minute wash in water

5 If the film is clear with very little signal, decrease the white light exposure time If the film appears brown or black, increase the white light exposure time.

VII. Reuse of Glass Plates

1. To remove the bound gel, immerse the plate and affixed gel in a 10% sodium hydroxide solution for 2 hours to overnight. Discard the gel and clean the glass plate with deionized H_2O and a detergent such as Liqui-Nox® The 10% sodium hydroxide solution may be reused for additional gels.

2. To remove the glass plate treatments (Gel Slick(TM) or Bind Silane) immerses the plate(s) in 10% sodium hydroxid solution for 1 hour Thoroughly rinse the plate(s) with deionized H2O and clean with a detergent.

IX. Composition of Solutions

0.5% acetic acid in 95% ethanol:

Add 1ml of glacial acetic acid to 199ml of 95° ethanol

40% acrylamide: bis (19:1):

380g acrylamide
20g bisacrylamide

Dissolve in **600ml** of deionized H2O. Bring volume to 1 liter with deionized H2O

10% ammonium persulfate:

Add 0.5g of ammonium persulfate to 5ml of deionized H2O. Use 500µl for one acrylamide gel solution (75ml). Store the remaining volume in 500µl aliquots at -20°C.

Developer solution:

60g sodium carbonate,(anhydrous Na2CO3)
3ml 37% formaldehyde(H2CO)

400µl 10mg/ml sodium thiosulfate($Na_2S_2O_3 \cdot 5H_2O$)
2000ml deionized H_2O

Prepare the solution and chill to 10°C before use. Use high quality deionized H_2O and sodium carbonate. Prepare fresh before each use.

Fix/stop solution (10% acetic acid):

200ml glacial acetic acid
1800ml deionized H_2O

STR 3X loading solution:

10mM NaOH
95% formamide
0.05% bromophenol blue
0.05% xylene cyanol FF

Staining solution:

2g silver nitrate ($AgNO_3$)
3ml 37% formaldehyde (H_2CO)
2000ml deionized H_2O

0.5X TBE:

Add 50ml of 10X TBE to 950ml of deionized H_2O.

10X TBE:

107.8g Tris base
7.44g EDTA (disodium salt, dihydrate)
~55.0g boric acid

Dissolve the Tris base and EDTA in 800ml deionized H_2O. Slowly add the boric acid and monitor the pH until the desired pH of 8.3 is obtained. Bring the volume to 1 liter with deionized H_2O.

X. Reference

1. Bassam, B.J., Caetano-Anolles, G. and Gresshoff, P.M. (1991) "Fast and Sensitive Silver Staining of DNA in Polyacrylamide Gels," Anal. Biochem. 196, 80-83.

6. REFERENCES

Anolles, G. &P.M.Gresshoff, 1994,DNA amplification fingerprinting using arbitrary mini-hairpin oligonucleotide Primers Bio/Technology12:619-623.

A.K.Singh, Genetic Resources Division, ICRISAT, Asia Center, Patancheru 502324 Applications of Molecular Techniques in Plant Genetic Resources characteristion, Conservation and utilization

B.J.Bassam&P.M.Gresshoff, 1991.DNA amplification using very short arbitrary oligonucleotide primers Bio/Technology9:553-557

Becker JP, Vos M, Kuiper F, Salamini & Heun M 1995. Combined mapping of AFLP and RFLP markers in barley Molecular and General Genetics 249: 65-73

Briggs FN & Knowles PF 1967. Introduction to plant breeding Reinhold, New York

Caetano-Anolles G, Bassam BJ & Gresshoff PM 1991 Bio/Technology 9: 553-557

Casey D 1992. Primer on molecular genetics U S. Department of Energy, Washington, D C

Duvick DN, 1984 In. Genetic contributions to yield grains of five major crop plants Fehr WR (ed) special Publ. No. 7, crop science Society of America, Madison, WI 15-47.

Funke RP, Kolchinsky A & Gresshoff PM 1993. Physical mapping of a region in the soyabean (*Gycine max*) genome containing duplicated sequences. Plant Molecular Biology 22: 437-446.

G.Caetano-Anolles&P.M.gresshoff, 1991,Fast and sensitive silver staining of DNA in polyacrilamide gels, Analytical Biochemistry196: 80-83.

G.Ne, N.Knight&C.S.Prakash, 1996. An evaluation of new gel matrices for the separation of PCR-amplified DNA fragments. Molecular Biotechnology 5:63-65.

G.M.Gracia, H.T.Stalker, and G.Kochert, Introgession analysis of an interspecific hybrid population in peanuts (Arachis hypogea L.) use RFLP and RAPD markers

H.T.Stalker, E.A.LaRue & G.Kochert, 1991.Genetic variation detectable with molecular markers among unadaptable germ-plasm resources of peanut and related wild species. Genome 34:1013-1020.

Kaemmer D, Weising K, Beiermann B, Borner T, Epplen JT & Kahl G 1995 Oligonucleotide fingerprinting of tomato DNA Plant Breeding 114: 12-14

Meksem K, Leister D, Peleman J, Zabeau M & Gebhardt C 1995. A high-resolution map of the vicinity of the R1 locus on chromosome V of potato based on RFLP and AFLP markers Mol Gen Genet 249: 74-81

Melchinger AE 1990. Use of Molecular markers in breeding for oligogenic disease resistance Plant Breeding 104: 1-19

P.Skroch, J.Nienhuis, Impact of scoring error and reproducibility of RAPD data on RAPD based estimates of genetic distance.

Paterson AH, Tanksely SD & Sorells ME 1991 DNA markers in plant improvement. Adv Agron 46: 39-90

Schmidt T, Boblenz, K & Weising K 1993 Use of highly repeated DNA polymorphisms for genome diagnosis and evolutionary studies in genus Beta In DNA fingerprinting: State of the science, eds Pena, S.D.J., Chakabrorty, R., Epplen, J.T. & Jeffteys, A.J. pp. 454-459 Basel: Birkhauser Velag

Stalker, H.T., 1992.Utilization Arachis germplasm resources In: S.N.Nigam (Ed.), Groundnut, and A Groundnut Perpective.pp.281-296 ICRISAT, Patancheru, India

T.Staslker&G.Kochert,1993.Development of an RFLP Linkage map in diploid peanut species.Theor Appl Genetics379-384

T.Stalker, E.LaRue &G.Kochert, 1992.Use of single primer DNA amplications in genetic studies of peanut (Arachis hypogea L.). Plant Molecular Biology 18:315-325.

Tanksley SD & Orton TJ 1993 Isozymes in plant genetics and breeding Amesterdam: Elsevier Science

Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M, Frijters A, Pot J, Peleman J, Kuiper M, *et al.* 1995. AFLP: a new technique for DNA fingerprinting Nucleic Acids. Research 21: 4407-4414

Williams, J.G.K., A.R.Kubelik, K.J.Livak, J.A.Rafalski & S.V.Tingley, 1990.DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. Nucleic Acids Research 18:6531-6535.

Weising K, Beyermann, B, Ramser J & Kahl G 1991a. Plant DNA fingerprinting with radioactive and dioxigenated oligonucleotide probes complementary to simple repetitive sequences Electrophoresis 12: 159-169

Weising K & Kahl G 1990 DNA fingerprinting in plants - the potential of a new method. Biotechnology Forum Europe 7: 230-235.

Williams J, Kubelic AR, Livak KJ, Rafalski JA & Tingey SV 1990. DNA polymorphism amplified by arbitrary primers are useful as genetic markers Nucleic Acid Research 18 7213-7218

Winter P & Khal G 1995. Molecular marker technologies for plant improvement World Journal of Microbiology and Biotechnology vol 11: 438-448*