Frei, O.M., Stuber, C.W., and Goodman, M.M. (1986) Yield manipulation from selection on allozyme genotypes in a composite of elite corn lines. Crop Sci., **26**, 917-921.

Lanaud, C. (1986) Utilisation des marqueurs enzymatiques pour l'etude genetique du cacaoyer: *Theobroma cacao* L. II. Étude du polymorphisme de six systemes enzymatiques. Cafe, Cacao, The (France), **30**, 271-280.

Santamour, F.S., and Denuth P. (1980) Identification of callery pear cultivars by peroxidase patterns. J. Hered. **71**, 447- 449.

Shechter, Y., and De Wet, J.M.J. (1975) Comparative electrophoresis and isozyme analysis of seed proteins from cultivated races of sorghum. Amer. J. Bot., **62**(3), 254-261.

Stavely, J.R., and Hanson, E.W. (1967) Electrophoretic comparisons of resistant and susceptible Trifolium pratense noninoculated and inoculated with *Erysiphe polygoni*. Phytopath. **57**, 482-485.

Tanskely, S.D., Medina-Filho H., and Rick, C.M. (1982) Use of a naturally occurring enzyme variation to detect and map genes controlling quarantine traits in an interspecific backcross of tomato. Heredity. **49**, 11-25.

Walters, T.W., Posluszny, U., and Kevan, P.G. (1989) Isozyme analysis of grape (Vitis). I. A practical solution. Can. J. Bot., **67**, 2894-2899.

Yidana, J.A., Kennedy, A.J., and Withers, L. (1987) Variation in peroxidase Isozymes of cocoa (*Theobroma cacao* L.). Proc. 10th Internat. Cocoa Res. Conf., Santo Domingo, 719-723.

CHARACTERISATION OF COCOA CLONES USING DNA-BASED MARKERS

J.Wilde * E.Johnson ** R.Waugh * and W.Powell *

INTRODUCTION

Concern over the genetic erosion of crop plants has led to an increase in plant exploration and in the establishment of numerous genebanks. Widespread deforestation particularly in South and Central America has resulted in the rapid loss of genetic variability for a range of tree species of the tropics. This factor has resulted in cocoa (*Theobroma cacao* L.) being designated as a priority crop for conservation (Anon, 1987). In order to provide a secure conservation base for the future, a sufficiently wide range of diversity is needed and two international "base collections" sited in Trinidad and Costa Rica have been designated. Plant exploration and collection is an important feature of the history of the crop (Allen and Lass, 1983) and many of the accessions collected are maintained by the Cocoa Research Unit, University of West Indies (Kennedy, 1984). The objective of such a living collection is to include with minimum redundancy, the genetic diversity of a crop species and its wild relatives (Brown and Clegg, 1983). This definition emphasises the need for a systematic evaluation and characterisation of the genetic resources present in a living collection.

The difficulty and expense of maintaining perennial tree crops in field genebanks, places great pressure on the evaluation procedure for such crops as cocoa. Traditionally, cocoa genetic resources have been characterised using a combination of morphological and agronomic traits. The effectiveness of using these to estimate genetic diversity has been questioned by several workers (Gottlieb, 1977; Brown, 1979). The long generation time of most perennial crops also means that many of the morphological descriptors can only be assessed at maturity. Furthermore, some collections have become too large and diffuse (Holden, 1984) with the result that utilisation is discouraged rather than assisted. Priority must therefore be given to the development of methods for the precise evaluation of variability in collections. The methods must be cost effective and the technology appropriate to the Institute responsible for managing the collection. In this article we describe our research on the development of molecular tools for the characterisation of cocoa genetic resources.

METHODS OF ANALYSING VARIABILITY

Morphological traits have the disadvantage of being influenced by environmental factors and this may not accurately estimate or represent genetic relationships. These factors have prompted the development of protein and DNA markers.

---

* Cell and Molecular Genetics Department, Scottish Crop Research Institute, Invergowrie, Dundee, DD2 5DA, Scotland

** Cocoa Research Unit, University of West Indies, St Augustine, Trinidad, West Indies

l ..ers

The most commonly used protein markers are isozymes. Isozymes are multiple molecular forms of enzymes that share a common substrate but differ in electrophoretic mobility (Markert and Moller, 1959). They are revealed when tissue extracts are subjected to electrophoresis in various types of gels and subsequently immersed in a solution containing enzyme-specific stains. Twenty-four enzyme systems have been examined in *Theobroma* (Atkinson et al., 1986; Yidana et al., 1987; Lanaud and Berthaud, 1984) and eight have been shown to be polymorphic. Although technically relatively easy to use, the low level of protein polymorphism detected is a limiting factor for fingerprinting cocoa. In addition, a number of different stains are required to identify polymorphism.

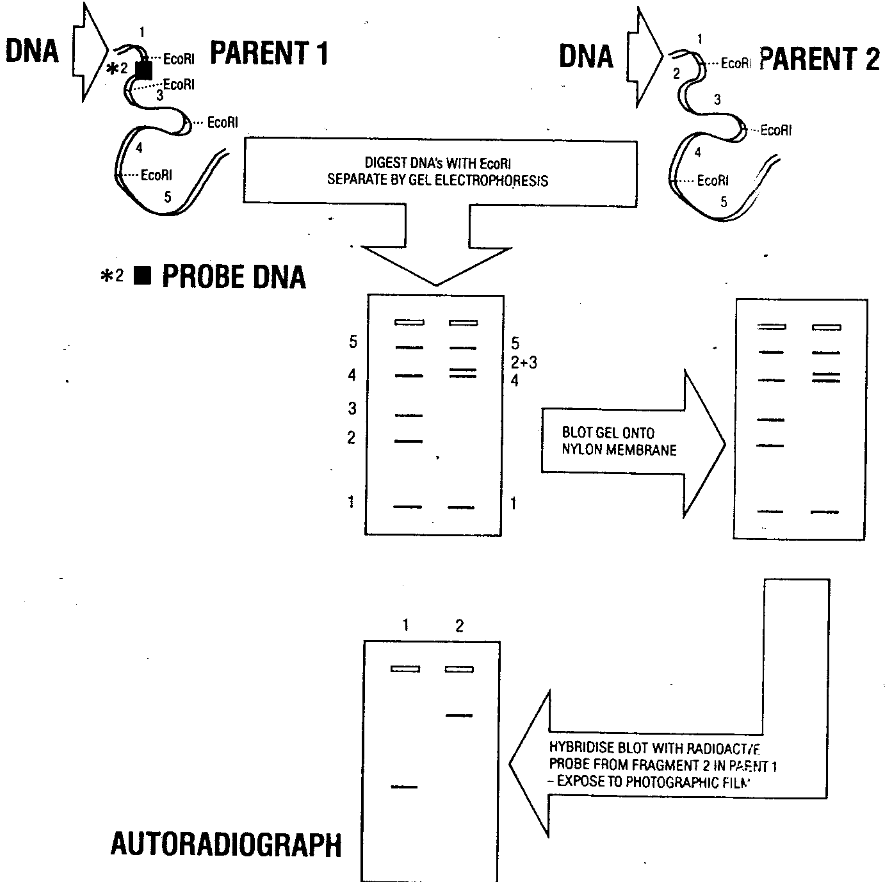
DNA Markers

Advances in molecular biology have resulted in new methods for the detection of polymorphism. Polymorphism in this case is revealed as differences in the lengths of homologous DNA fragments from different individuals. To detect these differences, the most frequently used technology relies on the use of a special class of enzymes called restriction enzymes. These enzymes have the ability to recognize and cleave DNA at target 'restriction sites'. These 'restriction sites' are locations at which exist specific sequences of the four bases, which when strung together, make the long DNA molecules which form the genome. A given restriction enzyme will therefore cleave a large fragment of DNA into smaller fragments. The size and number of fragments generated will reflect the distribution of restriction sites in the DNA and, after separating the fragments according to their length by electrophoresis, can produce a genetic fingerprint of the DNA of the organism under study. However, because the plant genome is so large, the number of fragments generated by restriction enzyme digestion is too great to resolve individual differences by electrophoretic techniques alone. This problem is overcome by exploiting the double-stranded nature of DNA (maintained by complementary base pairing) using a technique called DNA:DNA hybridisation. In this technique a small piece of cloned DNA from the plant under study (usually about $^1/_{10}$ - $^1/_5$ the size of the average restriction fragment length) is labelled with radioisotopes, made single-stranded, and used to 'probe' membrane filters on to which the total digested genomic DNA has been transferred. The probe molecule only binds tightly to the few sequences in the digest which have an exactly complementary DNA sequence. Because the probe is radioactive, the fragments to which it has bound can be visualised simply by exposing the hybridisation filter to photographic X-ray film. The differences detected between individuals are called Restriction Fragment Length Polymorphisms or RFLPs. The approach is illustrated diagrammatically in Figure 1.

DNA from several Amelonado trees has been isolated and digested with four restriction enzymes (BamHI, BglII, EcoRV and HindIII). The DNA was separated by electrophoresis and transferred to a membrane for subsequent hybridisation with radioactively labelled chloroplast specific probes. The polymorphism detected is illustrated in Figure 2 and indicates that the three cocoa genotypes differ in terms of their chloroplast DNA types. This is the first report of chloroplast DNA polymorphism in *Theobroma*. Although RFLPs can be used to characterise cocoa genetic resources, the technology is labour intensive requiring relatively large amounts of uncontaminated DNA which is difficult to obtain from cocoa. In addition, the isotopic labelling of a probe for

use in RFLP studies can be a limiting factor particularly in developing countries. However, a DNA-based fingerprinting method requires just two steps: isolation of genomic DNA and determination of sequence variability.

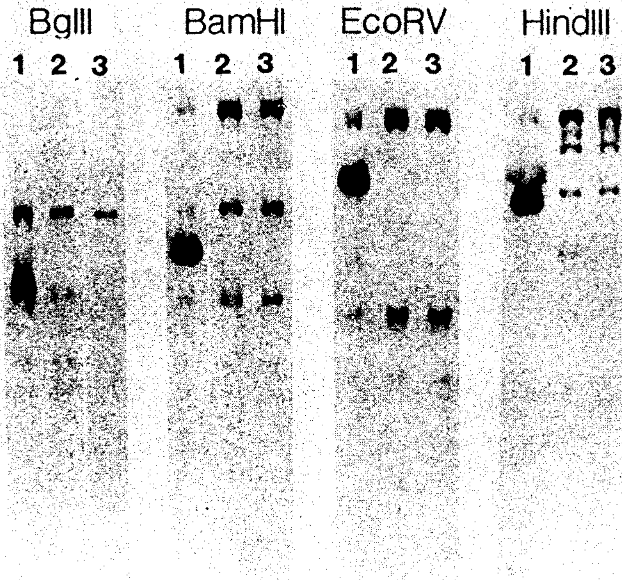
FIGURE 1: PROCEDURES FOR RFLP GENETIC FINGERPRINTING: DNA IS EXTRACTED AND CUT INTO FRAGMENTS BY RESTRICTION ENDONUCLEASES. THE DNA FRAGMENTS ARE SEPARATED ELECTROPHORETICALLY BY SIZE, TRANSFERRED TO A MEMBRANE AND HYBRIDISED TO RADIOACTIVE PROBES. HYBRIDISING FRAGMENTS ARE REVEALED BY AUTORADIOGRAPHY



POLYMERASE CHAIN REACTION (PCR)

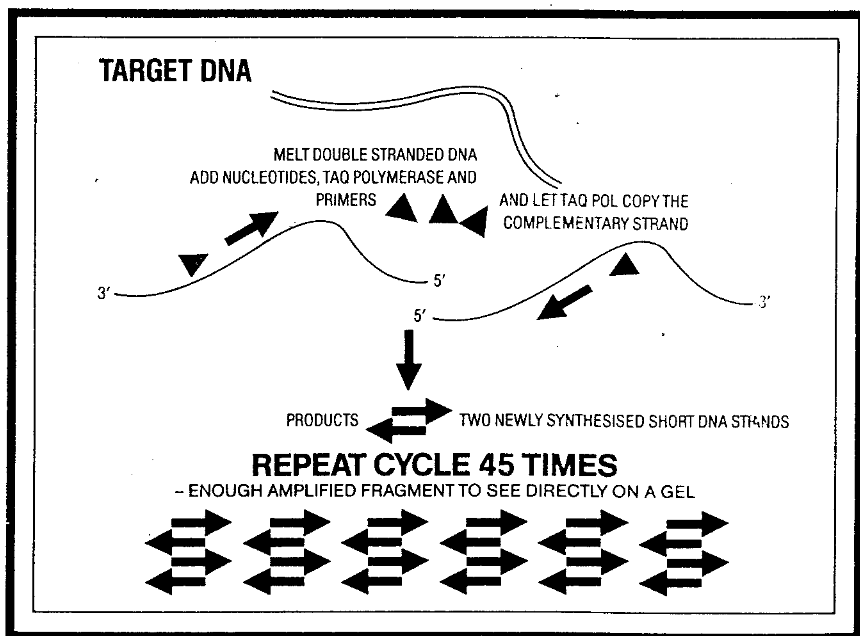
The PCR is based on a thermostable DNA polymerase enzyme (*taq*) which allows the selective amplification of a specific DNA fragment from a small amount of starting DNA. The technique depends on the ability of short, single-stranded DNA molecules called 'primers' to locate and bind to complementary sequences in the extremely long DNA molecule which comprises the plants genome. The *taq* polymerase catalyses the synthesis of a complementary DNA strand in one direction starting from the site where the primer has bound. When two primer binding sites are located on opposite strands of a DNA molecule within a given distance and in the right orientation (i.e. pointing together), repeating the PCR process through several cycles (which is possible because the *taq* is thermostable) results in an exponential amplification of the DNA between the two primer binding sites. The process is outlined in Figure 3.

FIGURE 2: DNA HYRBRIDISATION ANALYSIS OF THREE COCOA CLONES DIGESTED WITH FOUR RESTRICTION ENZYMES AND PROBED WITH A CHLOROPLAST SPECIFIC SEQUENCE



In the original form of PCR, it was necessary to know the sequences on either side of the target DNA before synthesizing complementary primers or oligonucleotides. However, recently it has been demonstrated (Williams et al., 1990; Wilde et al., 1991) that single short oligonucleotides of arbitrary sequence may be used to specifically amplify DNA fragments which can be visualised simply after electrophoresis by staining with ethidium bromide. The primers used have normally been only ten nucleotides long (10-ners). This procedure, known as Randomly Amplified Polymorphic DNA (RAPiD) analysis, has been successfully used in our laboratory to genetically fingerprint a range of *Theobroma* clones. The range of different phenotypes identified with a single primer for 13 *Theobroma* genotypes is shown in Figure 4. To date, our results demonstrate that the level of variability detected in *Theobroma* using RAPiDs is far greater than that observed with other polymorphism assays. Furthermore the RAPiD analysis is reproducible and allows discrimination both between, and within, *Theobroma* groups.

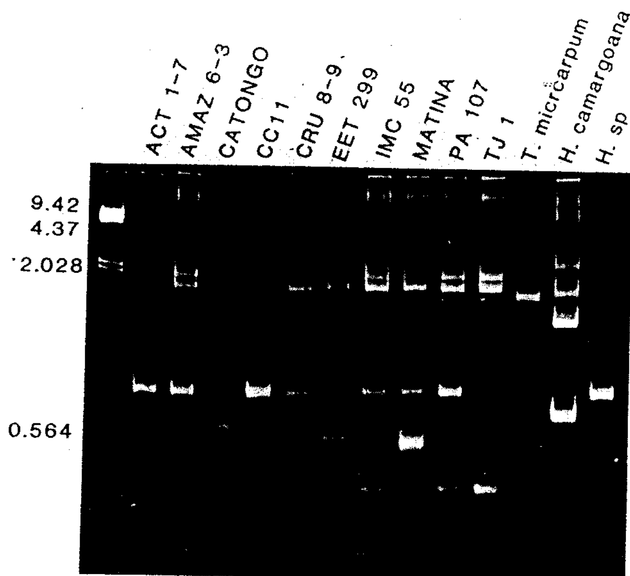
FIGURE 3: PCR-BASED TECHNIQUE FOR THE PRECISE AMPLIFICATION OF GENOMIC DNA



CONCLUSIONS AND FUTURE PLANS

Conventional RFLP analysis is labour intensive and requires relatively large quantities (5 ug) of uncontaminated DNA, which is difficult to obtain from cocoa (Couch and Fritz, 1990). PCR-based polymorphic assays are technically simple, easily automated and require only small amounts (100 ng) of DNA which can be extracted using mini-prep procedures from juvenile cocoa tissue. The methodology is not dependent on radio-isotopes and is therefore easily transferable to laboratories without facilities for radio-isotopes. In addition, the primers identified in our studies could represent a series of descriptors which would facilitate the characterisation and evaluation of cocoa genetic resources in field genebanks. The exploitation of RAPiDs would assist in the confirmation of accession identity, elimination of duplicates, allow the rationalisation of existing genebanks and permit future cocoa collecting expeditions to be targeted towards geographical areas possessing maximum levels of genetic diversity.

FIGURE 4: RAPiD ANALYSIS OF GENOMIC DNA USING A SINGLE SHORT OLIGONUCLEOTIDE PRIMER PRODUCING A GENETIC FINGERPRINT FOR 13 COCOA GENOTYPES



Presently, facilities are being established at the Cocoa Research Unit, University of West Indies, Trinidad for the evaluation of cocoa germplasm via the use of RAPiD markers. This important initiative has a crucial role to play in the future cost-effective characterisation and management of cocoa genetic resources.

REFERENCES

Anon. (1987) Genetic resources of cocoa. Report of the first meeting of the Cacao working group, IBPGR, Rome AGP/IBPGR/80/56.

Allen, J. and Lass, R.A. (1983) London Cocoa Trade Amazon Project, Final Report, Phase 1. Cocoa Growers' Bull., 34.

Atkinson, M.D., Withers, L.A. and Simpson, M.J.A. (1986) Characterisation of cacao germplasm using isoenzyme markers. Euphytica, **35**, 741-750.

Brown, A.D.H. (1979) Enzyme polymorphism in plant populations. Theor. Pop. Biol., **15**, 1-42.

Brown, A.D.H. and Clegg, M.T. (1983) Isozyme assessment of plant genetic resources. In: Isozymes current topics in biological and medical research. Vol. II. M.C. Rattazzi, J.G. Scandalios and G.S. Whitt (eds), 285-295, New York Press.

Couch, J.A. and Fritz, P.J. (1990) Isolation of DNA from plants high in polyphenolics. Plant Mol. Biol. Rep., **8**, 20-24.

Gottlieb, L.D. (1977) Electophoretic evidence and plant systematics. Ann. Missouri. Bot. Gard., **64**, 161-180.

Holden, J.H.W. (1984) The second ten years. In: Crop Genetic Resources: conservation and evaluation. J.W.Holden and J.T.Williams (eds) Allen & Unwin, London, 277-285.

Kennedy, A.J. (1984) The International Cocoa Genebank, Trinidad: Its role in future cocoa research and development. Proc. 9th Internat. Cocoa Res. Conf., Lome, Togo, 63-69.