Theobroma cacao DNA: Protocols for RFLP Analysis[1]

M.L. Mirazon, G. Gora-Maslak*, L. McHenry*, P.J. Fritz**

ABSTRACT

This paper demonstrates the applicability of restriction fragment length polymorphism (RFLP) technology to *T. cacao* L and gives detailed protocols for its use. Genomic DNA (gDNA) from two cocoa trees in the Penn State Greenhouse collection was hybridized to a cocoa gDNA clone (pTC101) in Southern blot experiments for detection of RFLPs. Two restriction endonucleases were used to digest the plant DNA and in one case a polymorphism was identified. The availability of methods for using RFLPs as new genetic markers offers a new time-saving method for germplasm identification in *T. cacao* as perfection of RFLP methodology will allow knowledge of genetic variability to be determined at the seedling stage. When RFLPs are correlated with desirable agronomic traits, improvement in breeding programs can be expected to follow.

COMPENDIO

Este trabajo demuestra la aplicabilidad de la tecnología del polimorfismo restringido de longitud de fragmentos (PRLF) aplicada al cacao (*Theobroma cacao* L.) y da protocolos detallados para su utilización. El genoma DNA (gDNA), de dos cacaoteros de la colección existente en los invernaderos de Penn State, fue hibridizado a un clone g DNA (pTC101) procedente de los experimentos de Southern para detectar PRLF. Se utilizaron dos endonucleasas de restricción para digerir el DNA; en un caso, se identificó polimorfismo. La disponibilidad de métodos para utilizar los PRLF como herramienta genética, ofrece un método nuevo y rápido para identificar germoplasma de *T. cacao* pues la perfección de la metodología PRLF permite conocer la variabilidad genética a ser determinada en la etapa de plántula. Cuando se correlaciona el PRLF con características agronómicas deseables, se pueden esperar resultados ventajosos en los programas de fitomejoramiento.

INTRODUCTION

DNA markers, first described as tools for genetic analysis in 1974 (14), and later used in linkage studies and in monitoring genetic traits in humans (6), are now being used successfully in higher plants (4, 5, 7, 15, 19). The method depends upon the ability of restriction endonucleases to catalyze cleavage of DNA at specific recognition sites yielding polydeoxynucleotides of defined lengths. The method has been called restriction fragment length polymorphism (RFLP) analysis.

Differences in RFLPs are an expression of the genetic heterogeneity encountered among individuals of a species; they are visualized when genomic DNA is digested with a restriction endonuclease, electophoresed, transferred to a solid support, and hybridized to a labelled probe.

The use of RFLPs as genetic markers has broad potential application in plant genetics, for individual, varietal and parental identification, mapping and monitoring of quantitative traits, analysis of genome organization, measurement of genetic diversity, and development of detailed linkage maps (4, 5, 6, 7, 15, 19). Extensive mapping using RFLPs has been done with maize, tomato, lettuce and peppers (16, 17).

Many of the proposed uses of RFLPs are identical to previously developed approaches using isozymes as molecular markers. Though isozyme markers have been quite useful in attempts to characterize cocoa germplasm (1, 2, 21, 22, 23, 38), their use is limited, first because not all genes code for enzymes and second, because much of a plant genome is composed of non-coding regions. In contrast, a potentially unlimited number of RFLPs exist, which should allow much wider use of the molecular marker approach to identify cocoa strains.

1 Received for publication 13 March 1990.
Published as Paper No. 7877, Journal Series, Pennsylvania Agricultural Experiment Station. This work was supported in part by an endowment from the American Cocoa Research Institute (ACRI) and a grant from the United States Agency for International Development. We thank Dr. Paul Chomet of the Cold Spring Harbor Laboratory, for providing modifications of the DNA isolation method.

* ACRI Cocoa Molecular Biology Laboratory, Department of Food Science, The Pennsylvania State University, University Park, PA 16802, USA.

The use of RFLPs for germplasm identification in *T. cacao* L. would be particularly useful to cocoa breeders, who historically have used morphological markers to identify cocoa varieties (3, 9, 10, 11, 12, 28, 29, 30, 32, 33, 35, 36) Such markers are often subjective, influenced by environment, and are generally applicable only to mature trees In contrast, RFLPs identify genetic differences at the DNA level, are not influenced by the environment, and can be detected at very early stages of plant development, helping to minimize the time and cost of planting and maintaining "unsuitable" material.

The purpose of this report is to give detailed protocols for RFLP analysis in cocoa and to demonstrate the general applicability of RFLP analysis to the cocoa plant.

MATERIALS AND METHODS

Source of plant material

The plants used in this study were growth in the cocoa greenhouse maintained at The Pennsylvania State University (PSU) Presently, the collection includes 27 five-year-old trees that originated from seed propagation of clones EQX-100 and UF-667 For the present study PSU 3 (EQX-100 origin) and PSU 20 (UF-667 origin) were used.

DNA preparation

Total DNA was obtained from PSU 20 and PSU 3 leaf tissue by a modification of the method of Malmberg *et al* (25): 1 g of fresh or frozen tissue was ground, using a mortar and pestle, to a fine powder in the presence of either liquid nitrogen or dry ice in sufficient amount to ensure that the grinding would occur without thawing of the leaf tissue. The powdered material was added to 6 ml of ice-cold extraction buffer (2.52 g urea, 0.35 M NaCl, 50 mM Tris-Cl pH 8 0, 20 mM EDTA, 1% Sarkosyl) in a small beaker on ice The importance of maintaining the tissue in a frozen state until the first centrifugation cannot be over-emphasized The mixture was transferred to a 15 or 30 ml Corex tube, the tube was placed in crushed ice, and the mixture dispersed with a Tissumizer (Tekmar Co., Cincinnati, OH) according to the following schedule: 30 s burst, 30 s wait, 30 s burst. The mixture was than centrifuged 5 min at 10 000 xg at 4°C. The supernatant was decanted and strained through Miracloth (CalBiochem, San Diego, CA) Subsequently an equal volume of phenol/chloroform: isoamyl alcohol, 24:1 (32) was added and mixed gently for 10 min at room temperature, and centrifuged at 5 000 xg for 5 min. The aqueous (upper) phase was removed (making sure that none of the interphase was picked up) and mixed 10 min at room temperature with an equal volume of chloroform: isoamyl alcohol, 24:1, followed by centrifugation The aqueous phase was removed and 1/10 vol of 4.4 M ammonium acetate, pH 5 0, plus 1 vol ice-cold isopropanol were added and mixed well At this point a white, stringy cloud of DNA formed After standing in ice for 30 min, the DNA mass could be removed with a glass stirring rod The DNA was then washed in 75% ethanol and resuspended in 200 µl of 10 mM Tris-Cl pH 8 0/1 mM EDTA pH 8 0 (TE). This volume usually yields a DNA concentration of about 0 5 g/l when starting with 1 g of leaf tissue. DNA does not always dissolve readily; usually several hours or perhaps overnight may be necessary for complete solubility

Preparation of partial cocoa leaf genomic library for probe selection

Hpa II-digested genomic DNA was ligated to Acc I-digested pUC 13 (37) using T4 ligase, and then transformed (8, 27) into *E. coli* strain DH5 (BRL, Bethesda, MD) Transformants were plated on LB/Amp/X-Gal agar plates (8, 26, 27) to identify clones Positive clones were first identified by blue-white selection, then screened for the presence of inserts by extracting DNA with a rapid plasmid isolation procedure (8, 27) and subsequent analysis by gel electrophoresis. Plasmid DNA containing useful probes was then prepared by a large-scale plasmid procedure (8, 26, 27, 31). Plasmids containing cocoa DNA are designated pTC (plasmid *T. cacao*) followed by a number Numbers 1-99 have been reserved for chloroplast or mitochondrial DNA. A useful method for preparing probes suitable for plant RFLP analysis has recently been published (18).

Probe labelling

Total plasmid DNA was nick-translated (BRL, Bethesda, MD, or IBI, New Haven, CT, kit) in a 50 x 1 reaction mixture containing 1 3 g probe DNA, 25 mM Tris-Cl pH 7 8, 50 mM, 10 mM MgCl, 100 g bovine serum albumin (BSA), 1 mM DTT, 500 M dGTP, dCTP, dTTP, 400 M biotinylated dATP (BRL), 4 ng DNase I and 4 units DNA Polymerase I Incubation was for 2 h at 15°C, after which the labeled probe was separated from unicorporated nucleotides by chromatography on G-50 Sephadex. Radioactive probes can be prepared by substituting ^{32}P-dATP for the biotinylated nucleotide The biotinylated probes can be stored for approximately two years.

RFLP analysis

Hybridization to probe. Two g of leaf DNA from PSU and 2 g from PSU 20 were digested with a series of restriction endonucleases. Digestion products were separated by electrophoresis in 1% agarose gels, stained with ethidium bromide to ensure proper digestion, transferred to nitrocellulose sheets by the Southern procedure (34), and hybridized with biotin labeled pTC101 (see below). Leaf DNA fragments hybridizing to the probe were visualized by appearance of blue bands. Details of the hybridization and visualization procedures are described below

Nitrocellulose (NC) sheets containing digested cocoa DNA were prehybridized 3 h at 42°C in 100 l/cm^2 of a solution containing 50% formamide, 5X SSC (0.75 M NaCl, 80 mM sodium citrate), 5X Denhardt's (0.1% Ficoll 400, 0.1% polyvinylpyrrolidone, 0.1% BSA), 25 mM sodium phosphate pH 6.5 and 0.5 mg/ml denatured salmon sperm DNA (Sigma Chem. Co., St. Louis, Missouri). Prehybridization solution was replaced by hybridization solution: 45% formamide, 5X SSC, 1X Denhardt's, 20 mM sodium phosphate pH 6.5, 0.2 mg/ml denatured salmon sperm DNA, 5% dextran sulfate and denatured (10 min 100°C), labelled probe. Overnight hybridization was done at 42°C. The sheets were then washed twice in 2X SSC, 0.1% SDS-3 min at RT, twice in 0.2X SSC, 0.1% SDS 3 min at RT, twice in 0.16X SSC, 0.1% SDS 15 min at 42°C, and rinsed in 2X SSC. Application of the non-radioactive DNA detection system was according to the manufacturer's instructions (BRL). The NC sheets were washed 1 min in 0.1 M Tris-Cl (pH 7.5)/0.15 M NaCl (buffer 1), incubated 1 h at 65°C in 3% BSA in buffer 1 (buffer 2), re-hydrated 10 min at room temperature in buffer 2 and incubated 10 min with 1 g streptavidin-alkaline phosphatase conjugate (BRL) per ml buffer 1. The sheets were then washed twice in buffer 1 and once in 0.1 M Tris-Cl (pH 9.5)/0.1 M NaCl/ 50 mM $MgCl_2$ (buffer 3).

Visualization of polymorphisms (BluGene,™ BRL): To visualize hybridizations with the biotinylated probe, the NC sheets were incubated in heat-sealed plastic bags, 3 h in 330 mg nitro-blue tetrazolium (NBT)/166 mg of 5-bromo-4-chloro-3-indolyl phosphate (BCIP) per ml buffer 3. The reaction was carried out in the dark. To terminate color development, the sheets were washed in 20 mM Tris-Cl (pH 7.5)/0.5 mM EDTA (stop buffer).

Although the manufacturers of the DNA detection system recommend storing the sheets dry by baking them 1-2 min at 80°C under vacuum, we observed that this causes weak bands to fade; we therefore prefer to store the sheets in stop buffer in a sealed plastic bag.

RESULTS

Selection of probes and restriction endonucleases: A partial genomic library was constructed in the 2.7 kilobase pair plasmid pUC 13. Inserts averaging 1 kb or less were cloned in the lac Z gene (13). Initially several clones were selected as potentially useful probes but only one proved to contain a low copy number sequence (i.e. a sequence occurring only once or a few times in the cocoa genome). Clones containing repetitive cocoa DNA gave smears instead of discrete bands and thus could not be used in the search for polymorphisms. Fig. 1 shows that this recombinant plasmid (pTC101) has a cocoa insert of approximately 400 bp.

The choice of restriction endonucleases was initially based on their cost and ability to catalyze digestion of cocoa DNA to completion, since incomplete digestion also results in smears upon hybridization with the probe. The enzymes selected for analysis were EcoRI and Hind III.

Detection of a polymorphism between two randomly selected trees: Fig. 2 shows the electrophoretic pattern of cocoa leaf DNA from trees PSU 3 and PSU 20 digested with EcoRI and Hind III. A replica of the gel was obtained by transferring the DNA to nitrocellulose paper followed by hybridization to a biotinylated probe (pTC101). Fig. 3 shows the resulting hybridization pattern. When digested with EcoRI, DNA from PSU 20 exhibits three distinct fragments of approximately 3.8, 2.1 and 0.99 kbp. By contrast, DNA from PSU 3 digested with the same enzyme exhibits only two of those fragments (2.1 and 0.99 kbp). Thus one genetic difference between the two trees can be detected by this combination of probe and enzyme. In contrast, the trees are indistinguishable when the DNA is digested by Hind III, since both show the same hybridization pattern, with fragments of approximately 6.5, and 1.0 kb (Fig. 3 lanes C and D). The sizes of the restriction fragments were calculated with the aid of a computer sizer program (Map, BIONET, Intelligenetics, Mountainview, CA).

DISCUSSION

Our results indicate that this method is applicable to *T. cacao* and that a polymorphism can be detected between two randomly selected trees. The detected genetic variation could result from various kinds of

genotypic alterations, for example, one or more bases could differ, resulting in the loss of a cleavage site or formation of a new one; alternatively, insertions or deletions of blocks of DNA within a fragment could alter its size. For instance, digestion with Eco RI revealed a 3.8 kbp fragment that is present in PSU 20 DNA, but not in PSU 3 DNA. This might be due to a DNA insertion in PSU 20 that contained part of the probe sequence. It is also possible that one of the cleavage sites that defines the 3.8 kb fragment is missing in PSU 3 DNA and that the probe sequence contained in such fragment is present in the 2.1 kbp fragment. The fact that some hybridization bands are stronger than others in the same DNA lane may imply that they contain more copies of the cloned DNA sequence than weaker hybridization bands. Or, the weaker bands could contain only part of the target sequence, resulting in a partial hybridization of those fragments, while the stronger bands could contain the entire cloned sequence, resulting in a total hybridization and a stronger signal (Fig. 3 lane B).

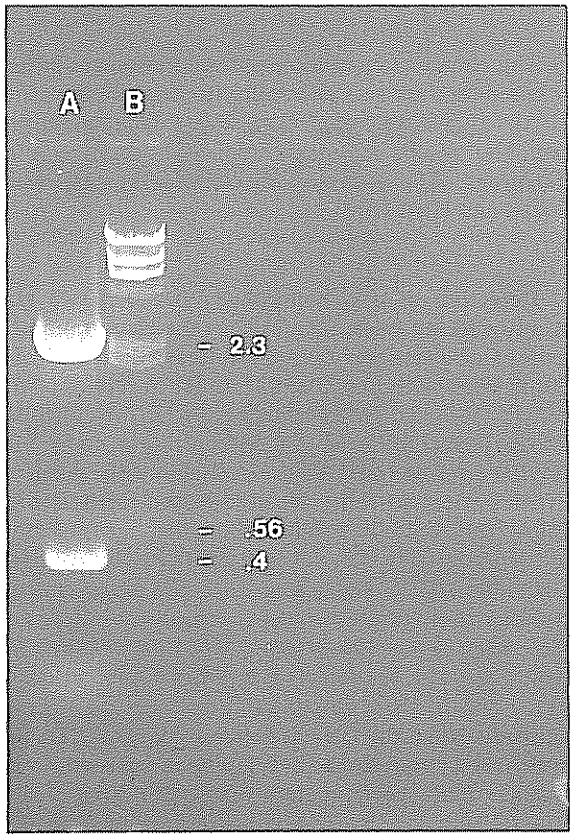

Fig. 1. Preparation of probe. Experimental details of DNA preparation are given in materials and methods. Electrophoretic analysis of the clone pTC101. 14 g of plasmid DNA were cut with 15 U of EcolRI and Hind III (lane A). Lane B contains 1.2 g of the DNA fragment size standard DNA digested with Hind III. The DNA was loaded onto a 1% agarose gel, electrophoresed for 1.5 hr (50 MA)and stained with ethidium bromide (6.0 mg/ml). Photography was done using a Polaroid MP-4 camera, a 325 nm Transilluminator and a red filter. Film Exposure (Polaroid type 55) was for 1.5 min.

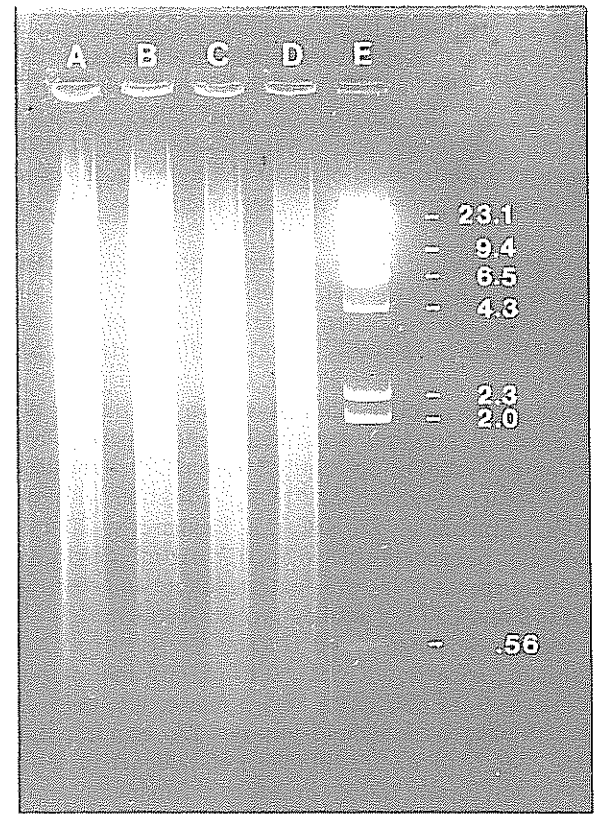

Fig. 2. Electrophoresis of cocoa genomic DNA. 1% agarose gel (stained with ethidium bromide at 6.0 mg/ml) containing 1 g PSU 3 DNA (lanes A & C), 1 g PSU 20 DNA (lanes B and D), cut with 20 U EcoRI (lanes A and B) and 20 U Hind III (lanes C and D). Lane E contains a fragment size standard (DNA cut with Hind III). Electrophoresis was for 2.5 hr at 120 V (85 mA). Photography was as described in legend to fig. 1.

When we compared the sensitivity of the biotin non-radioactive labelling system to the classical radioactive labelling system (data not shown), we saw no differences. The suitability of biotinylated probes should lead to extensive use of this extremely sensitive technique in the cocoa growing countries, where research facilities may not be equipped for the use of radioactive materials. The use of non-radioactive probes as an alternative for detection of specific DNA sequences has been discussed (20, 24).

Further work is in progress to expand the application of RFLP analysis to the cocoa plant by preparing

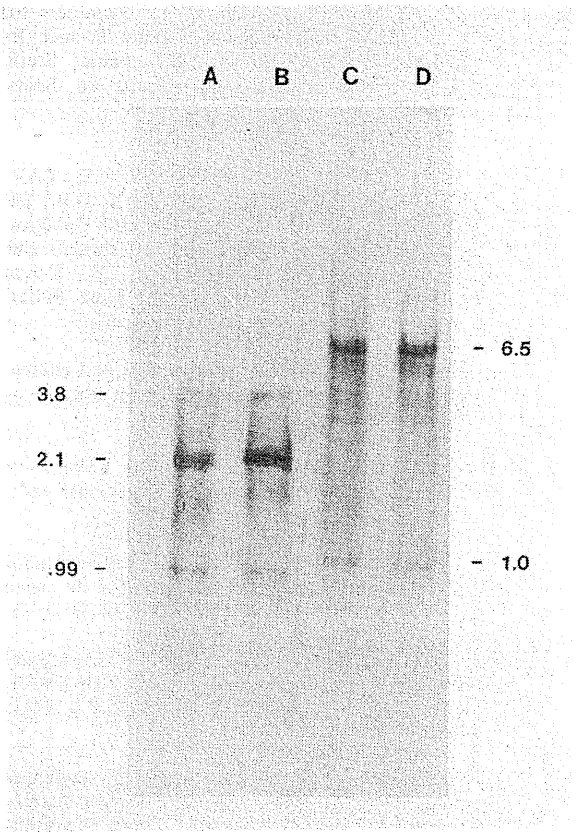

Fig. 3. RFLP analysis of cocoa genomic DNA. The DNA from the agarose gel shown in fig. 2 was transferred to nitrocellulose and hybridized with biotinylated pTC101.

more probes and by using more enzymes. The technique promises to be a powerful tool for germplasm identification in *T. cacao,* to say nothing of its uses in selecting superior trees at the seedling stage.

LITERATURA CITED

1. AMEFIA, Y.K., CILAS, C.; DJIEKPOR, E.K.; PATRIOT, M. 1984. Etude du polymorphisme enzymatique chez le cacaoyer. I. Mise au point d'une méthode d'extraction et mise en evidence d'un locus spécifiant une esterase. Café, Cacao, Thé 28:89-94.

2. ATKINSON, M.D.; WITHERS, L.; SIMPSON, M.J.A. 1986. Characterization of cacao germplasm using isozyme markers. A preliminary survey of diversity using starch gel electrophoresis and standardization of the procedure. Euphytica 35:741-750.

3. BARTLEY, B.G. 1986. Seed size inheritance. St. Augustine, Tri.; Annual Report on Cacao Research 1965. Imperial College of Tropical Agriculture. p. 23-25.

4. BECKMANN, J.S.; SOLLER, M. 1983. Genetic polymorphism in varietal identification and genetic improvement. Theoretical and Applied Genetics 67:25-33.

5. BECKMANN, J.S.; SOLLER, M. 1983. RFLPs in genetic improvement: methodologies, mapping and cost. Theoretical and Applied Genetics 67:35-43.

6. BOTSTEIN, D.; WHITE, R.; SKOLNICK, M.; DAVIS, R.W. 1980. Construction of a genetic linkage map in man using RFLPs. American Journal of Human Genetics. 32:314-331.

7. BURR, B.; EVOLA, S.V.; BURR, F.A.; BECKMAN, J.S. 1983. The application of RFLP to plant breeding. Genetic Engineering v.5, p. 45-59.

8. DAVIS, R.W.; BOTTSTEIN, D.; ROTH, J. 1980. A Manual for Genetic Engineering. Advanced Bacterial Genetics. N.Y., Cold Spring Harbor.

9. ENRIQUEZ, G.A.; SORIA, V.J. 1966. Estudio de la variabilidad de varias características de las mazorcas de cacao (*Theobroma cacao* L.). Fitotecnia Latinoamericana 3(1-2):99-118.

10. ENRIQUEZ, G.A.; SORIA, V.J. 1967a. Selección y estudio de los caracteres útiles de la flor para la identificación y descripción de cultivares de cacao. Cacao (C.R.) 12(1):8-16.

11. ENRIQUEZ, G.A.; SORIA, V.J. 1967. A study of certain leaf characteristics of cacao (*Theobroma cacao* L.) Tropical Agriculture (Tri.) 44(2):117-123.

12. ENRIQUEZ, G.A.; SORIA V.J. 1968. The variability of certain bean characteristics of cacao (*Theobroma cacao* L.). Euphytica 17(1):114-120.

13. GLOVER, D.M. 1985. DNA Cloning. IRL Press Ltd. Oxford, England, v.1, p. 89-100.

14. GRODZICKER, T.; WILLIAMS, J.; SHARP, J.; SAMBROOK, J. 1974. Physical mapping of temperature-sensitive mutations of adenoviruses. Symp. Quant. Biol. 39:439-446.

15. HELENTJARIS, T.; KING, G.; SLOCUM, M.; SIEDENSTRANG, L., WEGMAN, S. 1985. RFLPs as probes for genetic diversity and their development as tools for applied plant breeding. Plant Mol. Biol. Rep. 5:109-118.

16. HELENTJARIS, T.; SLOCUM, M.; WRIGHT, S.; A.; NIENHUIS, J. 1986. Construction of genetic linkage maps in maize and tomato using RFLPs. Theoretical and Applied Genetic 72:761-769.

17. LANDRY, S.; KESSELI, R.V.; FARRARA, B.; MICHELMORE, R.W. 1987. A genetic map of

lettuce with RFLP, isozyme, disease resistance and morphological markers. Genetics 116:331-337.

18. LANDRY, B.S.; MICHELMORE, R.W. 1985. Selection of probes for restriction fragment length analysis from plant genomic clones. Plant Mol. Biol. Rep. 3:174-179.

19. LANDRY, B.S.; MICHELMORE, R.W. 1987. Methods and applications of RFLP analysis to plants. In Tailoring Genes for Crop Improvement p. 25-44.

20. LANGER, P.R.; WALDROP, A.A.; WARD, D.C. 1981. Enzymatic synthesis of biotin-labelled polynucleotides: Novel nucleic acid affinity probes. Proceedings of the National Academy of Sciencies of the United States of America 78:6633-6637.

21. LANAUD, C.; BERTHAUD, J. 1984. Mise en évidence de nouveaux marqueurs génétiques chez *Theobroma cacao* L. par les téchniques d'electrophorese. In International Cocoa Conference (9, Togo, Africa). Proceedings p. 249-253.

22. LANAUD, C. 1986. Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer: *Theobroma cacao* L. I. Contrôle génétique et "linkage" de neuf marquers enzymatiques. Café, Cacao, Thé 30:259-270.

23. LANAUD, C. 1986. Utilisation des marqueurs enzymatiques pour l'étude génétique du cacaoyer: *Theobroma cacao* L. II. Etude du polymorphisme de six systèmes enzymatiques. Café, Cacao, Thé (France) 30:271-280.

24. LEARY, J.J.; BRIGATI, D.J.; WARD, D.C. 1983. Rapid and sensitive colorimetric method for visualizing biotin-labelled DNA probes hybridized to DNA or RNA immobilized on nitrocellulose: Bio-blots. Proceedings of the National Academy of Sciences of the United States of America 80:4045-4049.

25. MALMBERG, R.; MESSING, J.; SUSSEX, I. 1984. Molecular Biology of Plants - A Laboratory Course Manual. N.Y., Cold Spring Harbor Laboratory p. 86.

26. MANIATIS, T.; FRITSCH, E.; SAMBROOK, J. 1982. Molecular Cloning. A Laboratory Manual. N.Y., Cold Spring Harbor Laboratory.

27. MILLER, J. 1972. Experiments in Molecular Genetics. N.Y. Cold Spring Harbor Laboratory.

28. OSTENDORF, F.W. 1957. Identifying characters for cacao clones. In Reunião do Comite Tecnico Interamericano de Cacau 6., 1956, Bahía, Brasil. Salvador, Bra. Instituto de Cacau da Bahia. p. 89-110.

29. POUND, F.J.; COPE, F.W.; CHEESMAN, E.E.; BARTLEY, B.G.; JOLLY, A.C.; MURRAY, D.B.; DE VERTEUIL, L.L. 1986. Selection and Genetics. Fifty Years of Cocoa Research in Trinidad and Tobago: 1930-1980. St. Augustine, Tri., Cocoa Research Unit, The University of West Indies p. 3-31.

30. RUINARD, J. 1964. Variability of various pod characters as a factor in cacao selection. Euphytica 13(1):19-23.

31. SCHRADER, N.I.; HALLEY, B.W. 1980. Laboratory methods for hormone action and molecular endocrinology. Tex., Baylor University.

32. SORIA, V.J.; ESQUIVEL, O. 1968. Algunos resultados del programa de mejoramiento genético de cacao en el IICA, Turrialba. Cacao (C.R.) 13(2):19-23.

33. SORIA, V.J.; OCAMPO, F.; PAEZ, G. 1974. Parental influence of some cacao clones on the yield performance of their progenies. Turrialba (C.R.) 24(1): 58-65.

34. SOUTHERN, E.M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. Journal of Molecular Biology (England) 98:503-517.

35. TOXOPEUS, H. 1969. The second Nigerian breeding program. In Conference Internacional de Pesquisa em Cacau (2., 1967 Salvador e Itabuna., Bra.) Memorias. Bahia, Bra., CEPLAC p. 129-132.

36. VELLO, F. 1965. Estudos sobre a segregação do carater cor de amendoas de cacao (*Theobroma cacao*). In Informe Annual 1965. Itabuna Bra., Centro de Pesquisa do Cacau p. 10.

37. VIEIRA, J.; MESSING, I. 1982. The pUC plasmids, an M13 mp7-derived insertion mutagenesis and sequencing with synthetic universal primers. Genetics 19:259-268.

38. YIDANA, J.A.; KENNEDY, A.J.; WITHERS, L. 1987. Variation in peroxidase isozymes of cocoa (*Theobroma cacao* L.). In International Cocoa Research Conference 10. (Santo Domingo, R.D.). Proceedings.