PERGAMON

Phytochemistry 50 (1999) 919–924

PHYTOCHEMISTRY

Customisation of AFLP analysis for cassava varietal identification

Hui-Lee Wong, Hock-Hin Yeoh, Saw-Hoon Lim*

Department of Biological Sciences, National University of Singapore, Kent Ridge, Singapore 119620, Singapore

Received 9 April 1998; revised 12 November 1998

Abstract

Amplified fragment length polymorphism (AFLP) markers were used in the characterization of eight cassava varieties. This non-radioactive AFLP system was customized in terms of the choice of restriction enzymes used and the selection of nucleotides added to the 3′ end of primers. *Eco*RI/*Mse*I and *Hin*dIII/*Mse*I fragments generally gave monomorphic profiles while *Apa*I/*Taq*I fragments produced polymorphic profiles suggesting a genome with high G+C content. It was possible to identify the eight cassava varieties used in this study using CTG as selective bases at the *Taq*I primer. For cassava, the AFLP system provided a higher number of loci detected per run when compared to RAPD. The reliability accompanying AFLP analysis would thus make it suitable for the characterization of cassava varieties.

Keywords: *Manihot esculenta*; Euphorbiaceae; AFLP markers; Selective nucleotides; Varietal identification; Cyanogenic potential

1. Introduction

Cassava (*Manihot esculenta* Crantz) provides dietary carbohydrate to more than five hundred million people in developing countries (Cock, 1985). Cassava food safety, however, is compromised by the varying levels of cyanogenic potential (mainly attributed to linamarin content) in the storage roots (Yeoh & Truong, 1993). Thus, the assessment of genetic variation in this crop not only facilitates efficient conservation and utilization of germ plasm resources, it also assists in breeding strategies for safe consumption.

Determination of genetic variation can be conducted through a range of molecular markers (Paul & Yeoh, 1989; Yeoh & Paul, 1991; Fregene et al., 1994; Marmey, Beeching, Hamon, & Charrier, 1994; Yeoh & Joseph, 1995). Previous work which employed the use of RAPD in detecting DNA polymorphism between varieties showed that primers with higher G and C contents produced better DNA amplification patterns (Wong, Yeoh, Lim, & Looi, 1997). However, RAPD analyses do not always produce a very high number of reproducible fragments per run and this may limit its usefulness in varietal identification. This study thus aims to customize another PCR based technique, amplified fragment length polymorphisms (AFLP). This method is based on the selective amplification of restriction fragments from total genomic DNA with different primer pairs (Vos et al., 1995). We first focused on the restriction enzymes used for AFLP analysis and tested three combinations namely *Apa*I/*Taq*I, *Eco*RI/*Mse*I and *Hin*dIII/*Mse*I. We then added different selective nucleotides to the 3′ ends of the primers used and observed the banding pattern obtained. Two criteria were considered; the ability of the restriction enzyme and primer set to distinguish among genotypes and the number of loci/bands detected per experiment.

2. Results and discussion

Three sets of restriction enzyme combinations were tested in the AFLP analysis of cassava varieties. *Eco*RI/*Mse*I is generally useful for G+C-poor genomes, *Hin*dIII/*Mse*I for genomes with G+C contents of 40–50 mol% and *Apa*1/*Taq*I for G+C-rich genomes (Janssen et al., 1996). In our work, we found that the *Eco*RI/*Mse*I and *Hin*dIII/*Mse*I combinations generally produced monomorphic AFLP profiles (Figure 1a) when amplified with a selection of primers (Table 1). However, the *Apa*I/*Taq*I template, provided strong polymorphic banding patterns (Figure 1b) when amplified with all the primers tested. This suggests that the cassava genome is high in G+C content as enzymes that recognise high G+C stretches, like *Apa*I, are able to restrict the cassava DNA. This agrees well with the earlier observation that primers with high G+C content showed better RAPD patterns

* Corresponding author.

0031-9422/99 - see front matter
PII: S0031-9422(98)00628-1

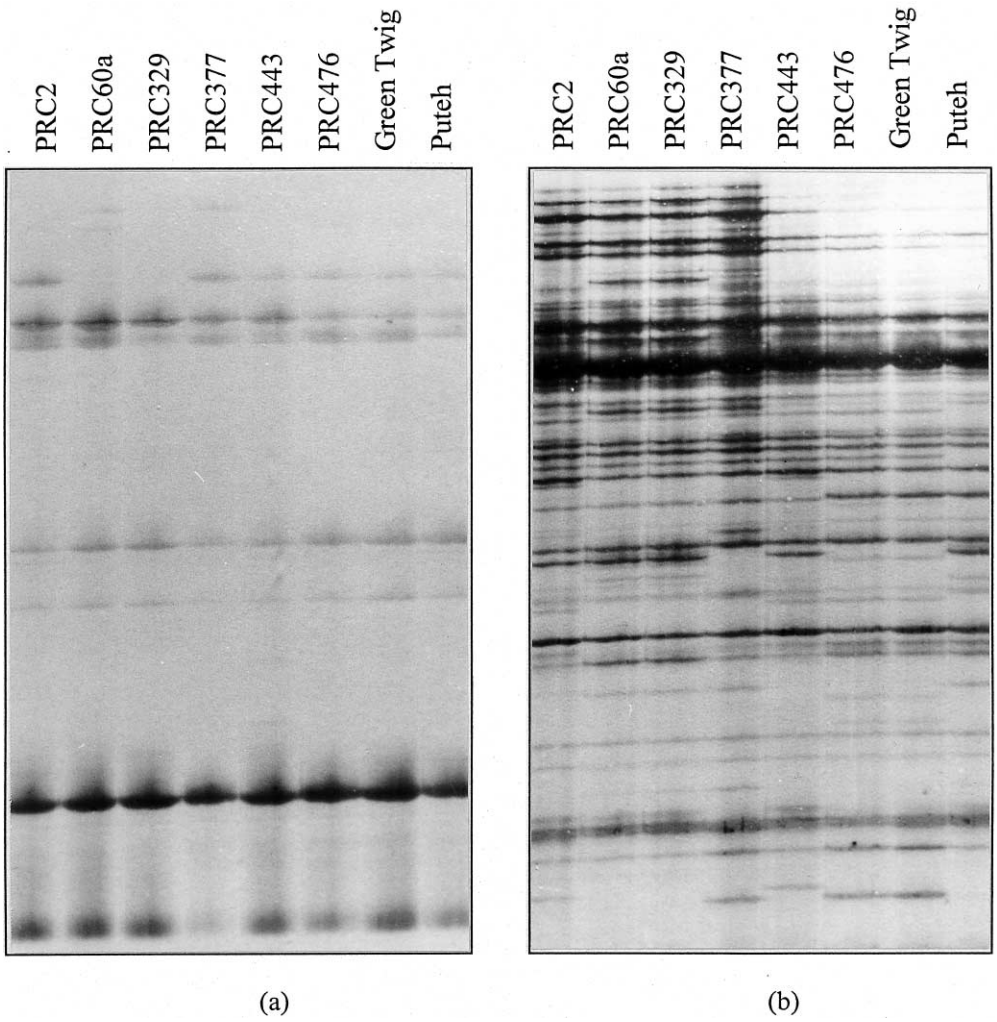

Fig. 1. AFLP analyses of cassava varieties. (a) Profiles generated using *Eco*RI-TC/*Mse*I-CCC and (b) profiles generated using *Apa*I-ATT/*Taq*I-CTG.

as they would find complementary sites in the G+C rich genomic DNA (Wong, Yeoh, Lim, & Looi, 1997). In addition, the quality of the AFLP patterns amplified is dependent upon the selective nucleotides added to the 3′ ends of the primers. In general, profiles with the strongest heterogeneity among the varieties examined and the highest number of bands were amplified with *Taq*I primers possessing 'CTG' at the 3′ end (Table 2). Unlike RAPD (Wong, Yeoh, Lim, & Looi, 1997, two individual AFLP experiments carried out using the *Taq*I-CTG primer generated profiles that distinguished all the different cassava varieties analysed Table 2. This would suggest that AFLP is suitable for fingerprinting cassava germplasm collections. Although AFLP at first glance appears to be a more tedious method of generating molecular markers, the reliability accompanying AFLP analysis would thus make it far more useful than RAPD in this respect. In addition, AFLP produced on average a higher number of loci detected per run (18.1 ± 12.7 bands) compared to RAPD analysis with custom-made primers (5.9 ± 2.2 bands) (Wong, Yeoh, Lim, & Looi, 1997.

The cluster analysis following the NTSYS-UPGMA algorithm (Rohlf, 1987) was generated with the Jaccard coefficient based on all the markers in Table 2. The dendrogram produced grouped PRC2, PRC443 as one cluster, PRC 476, Green Twig and Putih as another and PRC60, PRC329 and PRC377 as the third group (Figure 2). The close affinity between the following pairs, PRC2 and PRC443, PRC476 and Green Twig and PRC60a and PRC329 was also observed in the RAPD analysis of these varieties (Wong, Yeoh, Lim, & Looi, 1997.

When the AFLP products of each primer pair were analysed individually, it was found that PRC60 and PRC329 consistently grouped closest together. This is an interesting observation as both these varieties possess significantly higher cyanogenic potential (more than 200 mg linamarin kg^{-1} fr. wt.) compared to the other varieties tested (less than 150 mg linamarin kg^{-1} fr. wt.) (Wong, Yeoh, Lim, & Looi, 1997. One of the primer pairs, *Apa*I-ATT/*Taq*I-CTG generated only 11 markers but was able to distinguish these two varieties from the other cassava varieties (Figure 3).

Table 1
Adaptors and PCR primers used in the study. Primers for AFLP consist of a core region and an extension (E) of 1 to 3 nucleotides

Adapters and primers used	Adapter and primer sequences
*Apa*I adapter	5′-TCGTAGACTGCGTACAGGCC-3′ 3′-CATCTGACGCATGT-5′
*Apa*I primer core sequences	5′-GACTGCGTACAGGCCCE-3′
*Apa*I selective nucleotides added (E)	ATT, CCC, CT, CTA, CTG, GAC, TTG
*Eco*RI adapter	5′-CTCGTAGACTGCGTACC-3′ 3′-CTGACGCATGGTTAA-5′
*Eco*RI primer core sequences	5′-GACTGCGTACCAATTCE-3′
*Eco*RI selective nucleotides added (E)	C, CCC, TC
*Hin*dIII adapter	5′-CTCGTAGACTGCGTACC-3′ 3′-CTGACGCATGGTCGA-5′
*Hin*dIII primer core sequences	5′-GACTGCGTACCAGCTTE-3′
*Hin*dIII selective nucleotides added (E)	ACC, CCC,CCT,GAC
*Taq*I adapter	5′-GACGATGAGTCCTGAC-3′ 3′-TACTCAGGACTGGC-5′
*Taq*I primer core sequences	5′-CGATGAGTCCTGACCGAE-3′
*Taq*I selective nucleotides added (E)	CCC, CTG, GAC, GTA

Table 2
Number of fragments detected after amplification with primers with different selective nucleotides added

Primers with selective nucleotides added to the 3′ end	Number of polymorphic fragments obtained	Ability to differentiate all the varieties analysed
*Apa*I-ATT/*Taq*I-CTG	11	no
*Apa*I-CT/*Taq*I-CTG	35	no
*Apa*I-CTA/*Taq*I-CTG	27	yes
*Apa*I-GAC/*Taq*I-CTG	40	yes
*Apa*I-ATT/*Taq*I-GAC	6	no
*Apa*I-CCC/*Taq*I-GAC	7	no
*Apa*I-CTG/*Taq*I-GAC	2	no
*Apa*I-TTG/*Taq*I-GAC	17	no
*Apa*I-ATT/*Taq*I-CCC	27	no
*Apa*I-CTG/*Taq*I-CCC	8	no
*Apa*I-CTA/*Taq*I-GTA	19	no

Total bands = 199.

Measurement of accurate genetic relationships requires the marker system to provide an unbiased estimate of total genome variance and to be sufficiently abundant to minimize errors due to sampling variance (Spooner et al., 1996). This study thus assessed the ability of AFLP to compare and differentiate cassava varieties. We utilised the flexibility of the AFLP assay to customize its parameters, namely, choice of restriction enzymes, design of 3′ end selective nucleotides and type of amplified fragment visualization (using silver staining) for optimal efficacy. This study recommends *Apa*I/*Taq*I as the choice of restriction enzymes to generate templates for AFLP

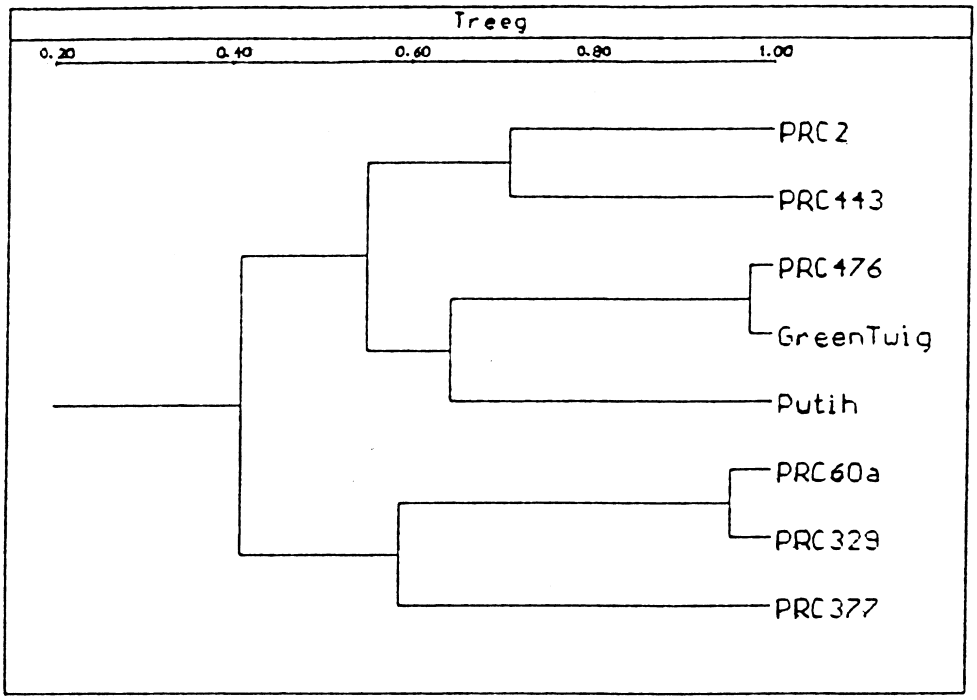

Fig. 2. Dendrogram showing the relationship between cassava varieties constructed using AFLP markers from Table 2.

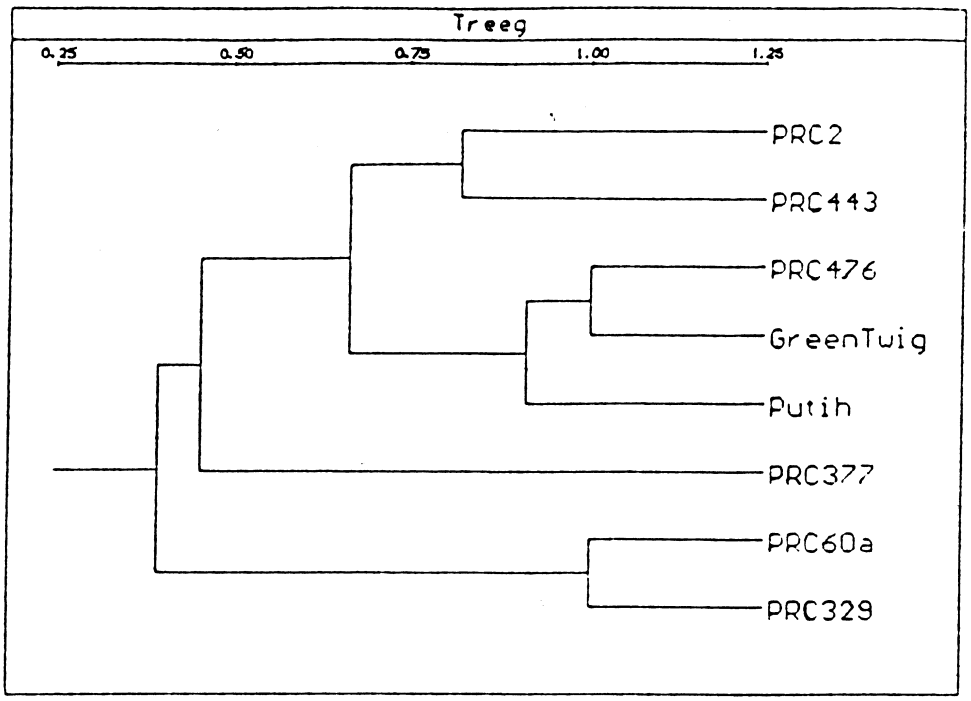

Fig. 3. Dendrogram showing the clustering of PRC60a and PRC329 away from the other varieties analysed. Primer pairs *Apa*I-ATT/*Taq*I-CTG was used in the AFLP analysis.

fragments. The *Taq*I specific primer with CTG added to the 3′ end shows promise as a polymorphic parameter in generating profiles for varietal differentiation in cassava.

3. Experimental

3.1. Restriction–ligation

Procedures were modified from Janssen et al. (1996). Total genomic DNA (0.5 μg) isolated from leaves of each cassava variety was digested with 5 U of either *Eco*RI + *Mse*I, *Hin*dIII + *Mse*I or *Apa*I + *Taq*I restriction enzyme combinations. The double digestions were performed at 37°C for 3 h in the buffer recommended by the manufacturer. For the *Taq*I + *Apa*I combination however, digestion with *Taq*I at 65°C preceded *Apa*I restriction at 37°C; both digestions were carried out for 3 h. Adapter sequences obtained from Janssen et al. (1996) were designed to destroy restriction sites and prevent fragment-to-fragment ligation during the concurrent restriction–ligation reaction Table 1. The adapters were non-phosphorylated to avoid adaptor-to-adaptor ligation. Working solutions of adapters were prepared at 50 ng/μl in 5 mM TrisAc (pH 7.5 at 25°C), 5 mM MgOAc and 10 mM KOAc. Double stranded dephosphorylated adapters (25 ng) were then ligated to the appropriate restriction fragments (10 ng) in the presence of 50 pmol of rare cutters (*Apa*I, *Eco*RI and *Hin*dIII) and 500 pmol of frequent cutters (*Mse*I and *Taq*I), 66 units T4 DNA ligase and 1 μg of BSA.

3.2. Preamplification

Core sequences of primers are as described by Janssen et al. (1996) Table 1. The restriction–ligation mixture was then amplified with upstream adapters, 4 mM dNTP, 0.4 U Taq polymerase and 0.1% Triton X-100 in 27 mM MgCl, 50 mM KCl, 10 mM Tris–HCl (pH 9.0) in a final volume of 20 μl. The preamplification cycle profile was one cycle of 120 s at 94°C followed by 35 cycles of 30 s at 94°C, 30 s at 50°C and 60 s at 72°C. Preamplified PCR products were diluted ten-fold with Tris–HCl and EDTA (pH 7.5) of which 1 μl was used for selective amplification.

3.3. AFLP-PCR

The AFLP-PCR reactions were performed in a final volume of 20 μl consisting of 5 ng of either *Eco*RI, *Hind*III or *Apa*I selective primers, 30 ng of either *Mse*I or *Taq*I primers Table 1, 4 mM dNTP, 0.4 U Taq polymerase in similar salt conditions to the preamplification step using the following touchdown PCR cycle: cycle 1: 60 s at 94°C, 30 s at 65°C and 60 s at 72°C; cycles 2–12: 30 s at 94°C, 30 s at annealing temperatures decreasing at 0.7°C for following cycles, starting at 64.3°C and 60 s at 72°C; cycles 13–24: 30 s at 94°C, 30 s at 56°C and 60 s at 72°C (Janssen et al., 1996). The resultant PCR products were then added to an equivolume of 98% (w/v) formamide, 10 mM EDTA and 0.1% (w/v) bromophenol blue. Reaction mixtures were heated at 95°C for 120 s and immediately put on ice prior to electrophoresis to prevent secondary structure formation.

3.4. Electrophoresis and visualization

Separation of the amplified fragments were carried out in 6% PAGE/7 M urea denaturing gels prepared by adding 15 ml of 40% stock acrylamide mix (38% acrylamide and 2% bisacrylamide) to 85 ml of 7 M urea/1 × TBE stock (42.4 g urea in 100 mM Tris, 100 mM boric acid, 2 mM EDTA, pH 8.0). The reaction was catalyzed by 400 μl of 10% ammonium persulfate and 40 μl TEMED. The soln was poured in 38 × 50 cm sequencing plates (Hoeffer, California, USA), separated by 0.4 mm thick spacers and combs and left for at least 1 h at room temperature for polymerization. The short plate was coated with Bindsilane as per manufacturer's instructions (Promega, Madison, USA). Samples (2 μl) were loaded and the gel was electrophoresed in 1 × TBE buffer at 80 W for 150 min. DNA fragments were visualized by silver nitrate staining (Promega, Madison, USA) with the following modifications to the manufacturer's instruction. Throughout the staining procedure, glassware was used. Rinsing of the gel prior to development/reduction of silver nitrate was extended to 10 s with agitation and development of the gels was maintained at below 10°C. Gels adhering to the glassware were dried in a 37°C oven for an hour. For permanent records, the positive image of the AFLP profile was captured on APC sequencing-sized film (Promega, Madison, USA).

3.5. Data analysis

AFLP banding patterns were scored manually as dominant markers and each locus was treated as a separate character. Thus, two lanes possessing a particular band were considered to carry the same allele at that locus. NTSYS software (Rohlf, 1987) was used to compute the simple matching coefficient of similarity (SIMQUAL) from the manually transcribed input file where presence bands were annotated as '1' and the absence as '0'. The similarity matrices constructed pairwise were used for clustal analysis generated by unweighted pair group method arithmetic average (Rohlf, 1987). Genetic relationships were visualized as dendrograms performed by SAHN option (sequential, agglomerative, hierarchical and nested clustering method as described by Sneath & Sokal, 1973).

Acknowledgements

This project was supported by a research grant (RP940302) from the National University of Singapore.

References

Cock, J. H. (1985). *Cassava. New potential for a neglected crop*. Boulder: Westview Press.

Fregene, M. A., Vargas, J., Ikea, J., Angel, F., Tohme, J., Asiedu, R. A., Akorada, M. O., & Roca, W. M. (1994). *Theor. Appl. Genet.*, *89*, 719.

Janssen, P., Coopman, R., Huys, G., Swings, J., Bleeker, M., Vos, P., Zabeau, M., & Kersters, K. (1996). *Microbiology*, *142*, 1881.

Marmey, P., Beeching, J. R., Hamon, S., & Charrier, A. (1994). *Euphytica*, *74*, 203.

Paul, K., & Yeoh, H. H. (1989). *Phytochemistry*, *28*, 1043.

Rohlf, F. J. (1987). *NTSYS-pc: numerical taxonomy and multivariate analysis system*. New York: Applied Biostatistics, Exeter Publishing.

Spooner, D. M., Tivang, J., Nienhuis, J., Miller, J. T., Douches, D. S., & Contreras, M. A. (1996). *Theor. Appl. Genet.*, *92*, 540.

Sneath, P. H. A. & Sokal, R. R. (1973). *Numerical taxonomy: the principles and practice of numerical classification*. San Franscisco: W.H. Freeman.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., Van de Lee, T., Hornes, M., Freijters, A., Pot, J., Peleman, J., Kuiper, M., & Zabeau, M. (1995). *Nucl. Acids Res.*, *23*, 4407.

Wong, H. L., Yeoh, H. H., Lim, S. H., & Looi, K. C. L. (1997). *Phytochemistry*, *46*, 805.

Yeoh, H. H., & Joseph, S. S. (1995). *Acta Physiol. Plant.*, *17*, 381.

Yeoh, H. H., & Paul, K. (1991). *Plant Var. Seeds*, *4*, 73.

Yeoh, H. H., & Truong, V. D. (1993). *Food Chem.*, *47*, 295.