

CHLOROPLAST DNA VARIATION IN COCONUT IS OPPOSITE TO ITS NUCLEAR DNA VARIATION

By

Lalith Perera¹

ABSTRACT

The pattern of world distribution of two major fruit morphotypes of coconuts has led to development of theories on origin, domestication and dissemination of coconut. Results of recent nuclear DNA analyses are in agreement with these theories with several other new insights. Compared to the plant nuclear genome however, the plant organelle genomes, the chloroplast genome and the mitochondrial genome are highly conserved and are maternally inherited in most angiosperms. Therefore, most useful information have come from regions of DNA located in organelle genome for studying phylogeny in angiosperms and for deducing historical information and evolutionary history of populations such as past migration routes and colonization dynamics. This study was aimed to determine the feasibility of developing polymorphic cytoplasmic markers, particularly the chloroplast markers. Chloroplast DNA variation of coconut from all coconut growing regions in the world assessed by both restriction digestions and physical separation of PCR products obtained with universal primers, by chloroplast microsatellites and by sequencing showed no variation. This tends to suggest that coconut may have gone through a severe cytoplasmic bottleneck and

¹ Genetics and Plant Breeding Division, Coconut Research Institute, Lunuwila, Sri Lanka

only one chloroplast type may have participated in the colonization process.

INTRODUCTION

Morphological studies of coconuts have identified two main groups of coconuts, which Harries (1978) described, one as coconuts naturally evolved from wild coconuts ('Niu kafa') and the other, the domesticated coconuts ('Niu vai') selected for increased nut water content as a source of uncontaminated and fresh drinking water. The Niu vai type is predominant in Southeast Asia and the Pacific and the Niu kafa are predominant in Indian Ocean islands and in the East Africa. It is further suggested that coconut in West Coast of America is of the type of Niu vai coconuts disseminated by early man and the coconuts in Caribbean and West Coast of Africa are of mainly Niu kafa type. These theories are based on morphological studies and the past and present literature available and hence evidence based on DNA relationship would be interesting to re-examine these various hypotheses. Recent nuclear DNA analysis of coconut (Rohde *et al.*, 1995; Lebrun *et al.*, 1998, Perera, 2000) from around the globe indicated a very high level of nuclear diversity in coconut. These studies also identified two main groups of coconut parallel to what Harries proposed in 1978 and the results were largely in agreement with the Harries' (1978) proposed theory of natural and human assisted dissemination.

Compared to the plant nuclear genome however, the plant organelle genomes, the chloroplast genome (cpDNA) and the mitochondrial genome (mtDNA) are highly conserved, non-

recombinant and are uni-parentally inherited (Palmer, 1985) and in most angiosperms, they exhibit maternal inheritance (Whatley, 1982; Sederoff, 1987). Thus, in higher plants, they provide ideal tools for high resolution of maternal lineage. They are regarded as the most suitable units in studying phylogeny in angiosperms and for deducing historical information and evolutionary history of populations such as past migration routes and colonization dynamics from their present day geographical distributions (King and Ferris, 1998; Demesure *et al.*, 1996; Soranzo, 1999). These polymorphisms also provide complementary information to the information obtained from nuclear DNA polymorphisms.

There are many different approaches to analyze chloroplast and mitochondrial genomes. The restriction fragment analysis (cpRFLP) of chloroplast DNA or mitochondrial DNA (cpDNA or mtDNA) was the most commonly used method in exploiting cytoplasmic DNA variation. However, the relatively conserved gene order of the organelle molecule between species has facilitated the design of universal primers for PCR and these universal primers are the sequences that are homologous to the most conserved coding regions of chloroplast DNA (cpDNA) or mitochondrial DNA (mtDNA), but amplify the most diverse non-coding regions (Demesure *et al.*, 1995; Dumolin *et al.*, 1997a; Taberlet *et al.*, 1991). Most of the phylogenetically informative variation in organelle genome occurs in non-coding regions and therefore, the use of universal PCR primers that amplify specific target sequences of the organelle genome is very useful and has become the most effective approach in detecting cytoplasmic diversity (Ferris *et al.*, 1993). Polymorphism occurs within the amplified regions can be detected in

different ways for example length polymorphism by restriction site analysis [PCR-RFLP or CAPS (Cleaved Amplified Polymorphic Sequence; Konieczny and Ausubel, 1993)] and analysis of Single Nucleotide Polymorphisms (SNPs) by physical separation of PCR-amplified products based on the variation in mobility of small polymorphic single-stranded DNA fragments in non-denaturing acrylamide gels [(SSCP: Single Strand Conformation Polymorphism (Orita *et al.*, 1989)] and by direct sequencing (Soranzo, 1999). In addition, recently mononucleotide microsatellites have been identified in the chloroplast genome (cpSSRs) of tobacco, liverwort, black pine and rice and their length polymorphism have been well demonstrated (Powell *et al.*, 1995, 1996). Both inter- and intra-specific variation have been detected and this has allowed a greater resolution in the cytoplasmic analysis than can be achieved with conventional RFLP studies (Powell *et al.*, 1995; Provan *et al.*, 1997; Provan *et al.*, 1998; Soranzo, 1999).

The aim of this study was to determine the feasibility of developing polymorphic chloroplast markers in coconut to allow characterization, population differentiation and reconstruction of the dissemination route of coconuts.

MATERIALS AND METHODS

Plant materials and DNA extraction

The materials for PCR-RFLP (or CAPS) used were tall ecotypes 'Samboanga', 'Mapanget', 'Malayan tall' and 'Thai tall' and dwarf ecotypes 'Categan', 'Bali yellow' and 'Eo brown', representing

Southeast Asia; tall ecotypes 'Andaman Ordinary', and 'Sri Lanka Ordinary' and dwarf ecotype 'Sri Lanka green' representing South Asia; and tall ecotypes 'Rennell' and 'Vanuatu' and dwarf ecotype 'Madang brown' representing the Pacific Islands. Tall ecotypes 'Comoro' and 'Mozambique' represented East Africa while tall ecotype 'West African' and dwarf ecotype 'Cameroon red' represented West Africa. Tall ecotypes 'Panama Monagre' and 'Panama Aquadulce' and dwarf ecotype 'Brazilian green' represented America. Hundred and thirty (130), genotypes representing 51 tall ecotypes and 43 dwarf ecotypes from across all coconut growing regions in the world were used for the SSCP and cpSSR studies. Tall ecotypes 'Samboanga', 'Mapanget', 'Rennell', 'Andaman Ordinary', 'Sri Lanka Ordinary', 'West African', 'Mozambique' and 'Panama Monagre' and dwarf ecotypes 'Bali yellow', 'Categan', 'Madang brown', 'Cameroon red' and 'Brazilian green' representing all geographical regions of coconut growing areas were used for direct PCR sequencing. DNA extractions were carried out as described in Perera (1999).

Amplification of chloroplast region

Non-coding regions of cpDNA were amplified using three pairs of universal primers described by Taberlet *et al.* (1991) and two pairs of universal primers described by Demesure *et al.* (1995).

PCR reactions were carried out in a total volume of 20 μ l consisting of 2 μ l of 10x PCR buffer (Boehringer), 0.2 μ l of *Taq* polymerase (Boehringer) (5U/ μ l), 2 μ l of 200 mM dNTPs, 2 μ l of forward primer (10⁻⁵ M), 2 μ l of reverse primer (10⁻⁵ M), 9.8 μ l of H₂O and 2 μ l of total genomic DNA (10ng). PCR conditions were as follows; 94^o C for 4 min, 30 cycles of 94^o C for 1 min, at annealing

temperatures of 50-58° C for 1 min and 72° C for 2 min. An aliquot of each sample was mixed with 1 μ l of loading dye and was run on a 1.6% agarose gel in 0.5x TBE buffer with marker VIII (Boehringer Mannheim) to confirm the amplification.

Restriction digestion of PCR amplified products

Restriction enzymes *AluI*, *BamHI*, *RsaI*, *Sau3AI*, *ApaI*, *BglII*, *ClaI*, *KpnI*, *PvuII*, *SmaI*, *TaqI*, *EcoRI*, *MseI*, *HindIII*, *EcoRV* and *PstI* were used to cleave PCR amplified products. A 10 μ l aliquot was digested overnight in a total volume of 20 μ l consisting 2 μ l of 1x restriction buffer (Gibco BRL), and 5-6 units of restriction enzyme (Gibco BRL). Restriction fragments were separated on 2% agarose gel in 0.5 x TBE and stained with ethidium bromide and visualized under UV light.

Single Strand Conformation Polymorphism (SSCP)

SSCP analysis was carried out with the PCR amplified products from chloroplast intergenic spacer between the *trnL* (UAA) 3' exon and *trnF* (GAA). Primer E (Table 1) was radioactively labeled using 32 P- [ATP]. PCR products were denatured at 95° C for 5 minutes and held at 4° C. Products were electrophoresed on a 40cm vertical mutation detection acrylamide gel (MDE) (Flowgen) (12.5ml of MDE gel, 34.5ml of SDW, 3ml of 10x TBE, 200 μ l of APS and 20 μ l of TEMED) at 5 W for 17 h and exposed to Kodak X-Omat 100 film for 2 days.

Sequencing

PCR products obtained from intergenic spacer between trnT-trnL exon, trnL intron, intergenic spacer between trnS- trnT, intergenic spacer between trnC – trnD and intergenic spacer between trnL exon and trnF from 12 ecotypes were sequenced directly and sequences obtained were aligned to see if there is any variation in the DNA sequence level.

Chloroplast microsatellite (cpSSR) study

cpSSR primers from other monocotyledonous species

Chloroplast SSR primers, which have been developed for monocotyledons such as rice, barley and wheat were used to amplify coconut. The locus, repeat type, primer sequences and the expected product size of those cpSSR primers are found in Perera (1999). The PCR conditions were as follows: a denaturing step at 94°C for 5 min followed by 30 cycles of denaturing step at 94°C for 1 min, primer annealing step at primer specific temperature for 1 min and primer extension step at 72°C for 2 min. PCR was carried in a total volume of 10 μ l containing 1 x PCR buffer (10mM Tris-HCl, 1.5mM MgCl₂, 50 mM KCl, pH 8.3), 200 μ M dNTPs, 10 pmol [³²P] end-labelled forward primer, 10 pmol reverse primer, 0.1 U Taq polymerase (Boehringer Mannheim) and 20ng genomic DNA. Reaction products were separated on 6% polyacrylamide gels in 1x TBE buffer and visualised by autoradiography.

Coconut specific chloroplast microsatellite (coconut cpSSR)

Coconut Chloroplast sequences were amplified using the universal primers described previously and double stranded PCR products were purified using QIAquick spin columns (QIAGEN) and were sequenced. Three microsatellite motifs were observed in the sequenced PCR amplified chloroplast regions of intergenic spacer between *trnT* and *trnL* ((T)₁₃), *trnS* and *trnT* ((G)₉) and *trnC* and *trnD* ((A)₁₂) and cpSSR primers specific to coconut were designed (Table 2). These primers were tested on 130 coconut genotypes as well as four other Palmae species; palmyra palm (*Borassus flabellifer*), fishtail palm or toddy Palm (*Caryota urens*), arica palm or betel palm (*Areca catechu*) and oil palm (*Elaeis guineensis*).

RESULTS

Amplification of chloroplast regions and restriction analysis

All five universal chloroplast primer combinations amplified coconut DNA at their expected product size. However all restriction digests tested were monomorphic across the sub-set of genotypes tested (Figure 1).

Single Strand Conformation Polymorphism (SSCP) and sequencing

This methodology again yielded no variation.

Sequencing

Sequenced data were checked against the sequence information held in the EMBL database and confirmed that the

sequences were from the correct chloroplast regions. When aligned the sequences they all were found to be identical.

cpSSR study

The cross amplification of cpSSR primers in coconut generated PCR products in the expected size. No length polymorphisms were observed. The three coconut specific cpSSR primers (Table 2) tested on coconut as well as related Palmae species again showed no length polymorphisms (Figure 2).

DISCUSSION AND CONCLUSIONS

The aim of this study was to identify a cytoplasmic marker that could distinguish between domesticated ('Niu vai') and naturally evolved coconut ('Niu kafa'), which would improve the understanding of the pattern of dissemination of coconut from its center of origin. Using all the available methodology to study the chloroplast genome (including available universal primers and SSRs and sequencing) no polymorphism was detected between coconuts from different geographical regions. Although the entire chloroplast genome was not analyzed, this study involved the analysis of non-coding regions of the chloroplast genome which have been shown to be highly variable at both inter and intraspecific level (King and Ferris, 1998; Ferris *et al.*, 1993; Demesure *et al.*, 1996; Dumolin *et al.*, 1997b). Moreover, cpSSRs have shown to be the most polymorphic chloroplast marker system distinguishing between individuals within the same population in several species (Powell *et al.* 1995; Provan *et al.*, 1997, Provan *et al.* 1998, Provan *et al.*, 1999) and were even

reported to detect polymorphism to a greater extent in crops in which no chloroplast polymorphism was detected with any other technique (Provan *et al.*, 1997). However, cpSSRs did not show any variation in coconut. Successful cross amplification of cpSSRs has been reported in other studies, for example rice cpSSR primers were successfully used to amplify *Hordeum vulgare* (Provan *et al.*, 1999) and *Eucalyptus* (J. Russell, personal communication) and potato cpSSR primers has been successfully used to amplify Jute (R. Meyer, personal communication). The results of this study suggested that there is no chloroplast variation in coconut or if there is any variation it is very low and we do not yet have the tools to detect such variation. Similar observations of no chloroplast variation have been reported among a group of 12 pearl Millet (*Pennisetum glaucum*) cultivars (Clegg *et al.*, 1984). Those observations have forced Clegg and others to suggest a dramatic reduction of variability upon domestication (Clegg *et al.*, 1984). In order to clarify the issue further investigation was carried by Gepts and Clegg (1989) using additional genotypes including 13 wild lines representing most of the geographic distribution of wild pearl Millet and they still found no variation in their extended sample. The results of those two studies have led to the conclusion of possible bottleneck induced at some stage of the evolutionary history of pearl millet perhaps owing to climatic fluctuation. A study conducted by Waters and Schall (1991) on two populations of *Pinus torreyana* that have been isolated for more than 8000 years found no variation over 150 restriction sites and concluded that these populations are not relicts of a once highly polymorphic populations but rather relicts of a once highly monomorphic population. In contrast, studies in oaks (Ferris *et al.*, 1993, 1995, 1998), common beech (Demesure *et al.*, 1995) and black

alder (King and Ferris, 1998), using chloroplast markers from the same chloroplast DNA regions used in this study observed higher level of intra-specific variation.

Factors such as genetic bottlenecks induced during the evolution of a crop either as a result of fluctuating weather changes or domestication (Clegg *et al.*, 1984b), founder effects and natural selection may play a role in reducing chloroplast DNA variation. Since there is no variation found between naturally evolved coconuts and domesticated coconuts and between tall and dwarfs coconut, it is reasonable to assume a reduction of chloroplast variability has occurred before coconuts were domesticated, cultivated and disseminated. It can be assumed that only one type of chloroplast was successful in colonizing coconuts. It is possible that the bottleneck occurred due to selection for a particular chloroplast type of the wild coconut adapted to coastal conditions. However, three cpSSR loci, which were monomorphic across four other Palmae species, tend to suggest an extreme bottleneck encountered long before coconut originated, leaving a narrow cytoplasmic base for the family Palmae. In contrast, the higher level of diversity observed at the nuclear genome in coconut is the result of founding effect, genetic drift, mutation and recombination.

In summary, levels of cytoplasmic diversity in coconut assessed by using both restriction digestions of PCR products obtained with universal primers for the chloroplast genomes and cpSSRs showed no variation. This suggests that coconut may have gone through a severe cytoplasmic bottleneck.

REFERENCES

- Clegg MT, Brown AHD, Whitfeld PR (1984b) Chloroplast DNA diversity in wild and cultivated barley - Implication and conservation. *Genetic Research* **43**: 339-343
- Clegg MT, Rawson JRY, Thomas K (1984) Chloroplast DNA variation in pearl millet and related species. *Genetics* **106**: 449-461
- Demesure B, Comps B, Petit RJ (1996) Chloroplast DNA phylogeography of the common beech (*Fagus sylvatica* L.) in Europe. *Evolution* **50**: 2515-2520
- Demesure B, Sodzi N, Petit RJ (1995) A set of universal primers for amplification of polymorphic noncoding regions of mitochondrial and chloroplast DNA in plants. *Molecular Ecology* **4**: 129-131
- Dumolin S, Demesure B, Fineschi S, LeCorre V (1997b) Phylogenetic structure of white oaks throughout the European continent. *Genetics* **146**: 1475-1487
- Dumolin S, Pemonge MH, Petit RJ (1997a) An enlarged set of consensus primers for the study of organella DNA in plants. *Molecular Ecology* **6**: 393-397
- Ferris C, Oliver RP, Davy AJ, Hewitt GM (1993) Native oak chloroplasts reveal an ancient divide across Europe. *Molecular Ecology* **2**: 337-344
- Ferris C, Oliver RP, Davy AJ, Hewitt GM (1995) Using chloroplast DNA to trace postglacial migration routes of oaks into Britain. *Molecular Ecology* **4**: 731-738

- Gepts P, Clegg MT (1989) Genetic diversity in pearl millet (*Pennisetum glaucum* [L.] R. Br.) at the DNA sequence level. *Journal of Heredity* **80**: 203-208
- Harries HC (1978) The evolution dissemination and classification of *Cocos nucifera* L. *Botanical Review* **44**: 205-317
- King RA, Ferris C (1998) Chloroplast DNA phylogeography of *Alnus glutinosa* (L.) Gaertn. *Molecular Ecology* **7**: 1151-1161
- Konieczny A, Ausubel F (1993) A procedure for mapping *Arabidopsis* mutations using codominant ecotype specific PCR-based markers. *Plant Journal* **4**: 403-410.
- Lebrun P, N'Cho NP, Seguin M, Grivet L, Baudouin L (1998) Genetic diversity in coconut (*Cocos nucifera* L.) revealed by restriction fragment length polymorphism (RFLP) markers. *Euphytica* **101**: 103-108
- Orita M, Suzuki Y, Sekiya T, Hayashi K (1989) Rapid and sensitive detection of point mutations and DNA polymorphism using the polymerase chain reaction. *Genomics* **5**: 874-879
- Perera L. (1999) Assessing genetic diversity of coconut (*Cocos nucifera* L.) using molecular markers. Ph.D thesis. University of Dundee, UK.
- Perera L, Russell JR, Provan J, Powell W (2000) Use of Microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (*Cocos nucifera* L.). *Genome* **43**: 15-21
- Powell W, Machray GC, Provan J (1996) Polymorphism revealed by Simple Sequence Repeats. *Trends in Plant Science* **1**: 215-222

- Powell W, Morgante M, Andre C, McNicol JW, Machray GC, Doyle JJ, Tingey SV, Rafalski JA (1995) Hypervariable microsatellites provide a general source of polymorphic DNA markers for chloroplast genome. *Current Biology* **5**: 1023-1029
- Provan J, Corbett G, McNicol JW, Powell W (1997) Chloroplast variability in wild and cultivated rice (*Oryza* spp.) revealed by polymorphic chloroplast simple sequence repeats. *Genome* **40**: 104-110
- Provan J, Russell JR, Booth A, Powell W (1999) Polymorphic chloroplast simple sequence repeat primers for systematic and population studies in the genus *Hordeum*. *Molecular Ecology* **8**: 505-511
- Provan, J, Soranzo N, Wilson NJ, McNicol JW, Morgante M, Powell W (1998) The use of uniparentally inherited simple sequence repeats in plant population studies and systematics. In: *Advances in Plant Molecular Systematics*. Hollingsworth P, Gornall R, Bateman R (Eds), Chapman and Hall, London. pp 35-50
- Rohde W, Kullaya A, Rodriguez J, Ritter E (1995) Genome analysis of *Cocos nucifera* L. by PCR amplification of spacer sequences separating a subset of *copia*-like 16RI repetitive elements. *Journal of Genetics and Breeding* **49**: 179-186
- Sederoff RR (1987) Molecular mechanism of mitochondrial-genome evolution in higher plants. *American Naturalist* **130**: 530-545
- Soranzo N (1999) Genetic variation in native European populations of *Pinus sylvestris* (L.). Ph.D Thesis. University of Dundee, UK.

- Taberlet P, Gielly L, Pautou G (1991) Universal primers for amplification of 3 noncoding regions of chloroplast DNA. *Plant Molecular Biology* **17**: 1105-1109
- Waters ER, Schall BA (1991) No variation is detected in the chloroplast genome of *Pinus torreyana*. *Canadian Journal of Forest Research* **21**: 1832-1835
- Whatley, JM (1982) Ultra structure of plastid inheritance: green algae to angiosperms. *Botanical Review* **57**: 527-569

ACKNOWLEDGEMENT

This work was funded by the Commonwealth Scholarship Commission of the UK through the British Council and carried out at the Scottish Crop Research Institute, Scotland under the guidance of Dr. Joanne Russell, Dr. Jim Provan and Professor Wayne Powell. Coconut leaf materials for this study were organized by the COGENT/IPGRI and were supplied by the COGENT member countries. The cooperation of the staff of the Genetics and Plant Breeding division of the Coconut Research Institute towards this work is highly appreciated.

Table 1. Sequences of the six universal chloroplast primers for the amplification of five non-coding regions of cpDNA (Taberlet *et al.*, 1991 and Demesure *et al.*, 1995).

Primer position	Reference	Genome	Sequence 5'- 3'
TrnT (A)	Taberlet et al., 1991	Chloroplast	CATTACAAATGCGATGCTCT
TrnL 5' Exon (B)	Taberlet et al., 1991	Chloroplast	TCTACCGATTTCGCCATATC
TrnL 5' Exon (C)	Taberlet et al., 1991	Chloroplast	CGAAATCGGTAGACGCTACG
TrnL 3' Exon (D)	Taberlet et al., 1991	Chloroplast	GGGGATAGAGGGACTTGAAC
TrnL 3' Exon (E)	Taberlet et al., 1991	Chloroplast	GGTTCAAGTCCCTCTATCCC
TrnF (F)	Taberlet et al., 1991	Chloroplast	ATTTGAACTGGTGACACGAG
TrnC	Demesure et al., 1995	Chloroplast	CCAGTTCAAATCTGGGTGTC
TrnD	Demesure et al., 1995	Chloroplast	GGGATTGTAGTTCAATTGGT
TrnS	Demesure et al., 1995	Chloroplast	CGAGGGTTCGAATCCCTCTC
TrnT	Demesure et al., 1995	Chloroplast	AGAGCATCGCATTTGTAATG

Table 2. Locus name, repeat motif, sequence, annealing temperature (T_M) and expected product size of coconut specific microsatellite loci.

Locus	Region	Repeat motif	Sequence (5'-3')	T_M	Product size (bp)
CNCPSS R1	trnT-trnL	(T) ₁₃	TGGAATACTGGAACGGTCGA AGATAGAATCCCAATTAGAATG	58	110
CNCPSS R2	trnS- trnT	(G) ₁₂	TCGTTAGGTTCAAGTCTGGC GTGCAAACCCCGAGATATTA	60	230
CNCPSS R3	trnC - trnD	(A) ₁₂	CAGACTCACGAGAGGCTTTG GCAAAGGAATGCTCCTAATG	60	198

Figure 1. Example of digestive profiles of PCR amplified chloroplast region (Intergenic spacer between *trnL* exon and *trnF* with *Bam*HIII).

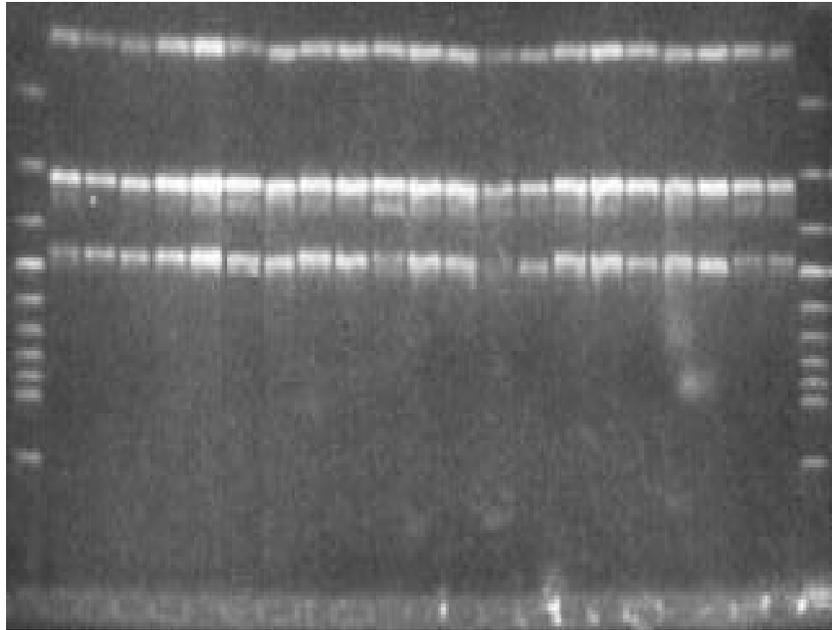
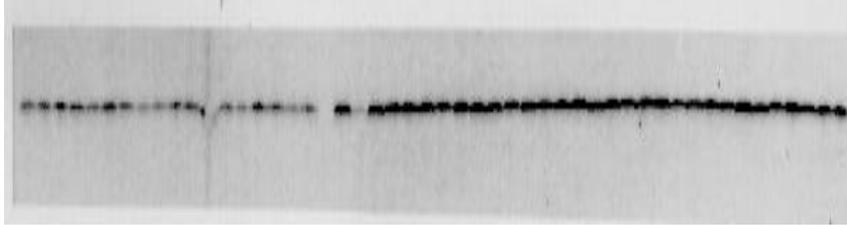


Figure 2. Examples of monomorphic banding pattern of cpSSR loci
CNCPSSR1 (A) and CNCPSSR2 (B).

(A)



(B)

