Vidyodaya J. of Sci. (2006) Vol. 13, pp 65-73

Preliminary attempt for the development of a molecular marker based map of coconut genome

Dasanayaka PN[1], Everard JMDT[2], Bandaranayaka Ck[2], Karunanayaka EH[3], Nanadasa HG[1]

1. Department of Botany, University of Sri Jayewardenepura, Nugegoda.
2.Coconut Reserach Institute, Lunuwila.
3.Institute of Biochemistry Molecular Biology and Biotechnology, University of Colombo, 94, Kumarathunga Munidasa Mawatha, Colombo 3.

Received on : 10-27-05
Accepted on : 07-06-06

Abstract

A self-pollination programme has been carried out at the Coconut Research Institute (CRI) of Sri Lanka to produce F_2 families of an F_1 cross between tall and dwarf coconut palms. Among the F_2 families generated three were selected for studying the segregation of molecular markers.

Simple sequence repeat (SSR) markers were used for the study due to its high level of heterozygosity, codominant inheritance and PCR-based detection. Thirty five SSR primers were used and screening of polymorphic SSR markers was performed by testing the markers on the parents and F_1 individuals of the progeny.

Among the 35 pairs of primers tested 24 exhibited heterozygocity and gave a consistent pattern of SSR alleles in the segregating populations. 173 individuals of F_2 population were genotyped using these selected 24 primer pairs. Unexpected banding patterns were observed in these 3 families and the data were inadequate to submit to the mapping software JOINMAP for construction of linkage relationships.

This preliminary study helped to realize the potential of mapping the coconut genome and the need to prepare for genome mapping by construction of segregating populations with utmost care.

Key words: Genome mapping, coconut, SSR markers, F_2 population

1. Introduction

Coconut, *Cocos nucifera* L. is a heritage for people in many tropical countries as a principal food commodity as well as a source of foreign exchange. Sri Lanka is the fourth best coconut producing country though its production accounts for about 5% of world's production. Sri Lanka has to increase the productivity to compete with other main producers. Sri Lanka being a small country can no longer expand the

land under coconut and has to depend mainly on increasing productivity and quality of the coconut products. Development of genetically important planting material is an important step towards increasing coconut productivity.

But there are many obstacles for rapid progress of coconut breeding. These obstacles are its long juvenile phase, time and cost limitations, lack of vegetative propagation methods, small number of seed produced per year and the inherent heterozygosity.

With the rapid development of DNA technology, molecular markers have come to the forefront of many crop-breeding programs. DNA polymorphisms can aid conventional breeding programs in two ways, i.e. assessing the genetic diversity and its organization at various levels of the crop of interest and by developing genetic linkage maps. Genetic linkage maps provide an important tool for the location of genes linked to trait of economic importance, which could be used in marker-assisted selection.

Various types of molecular markers have been used in the assessment of coconut genetic diversity (Dasanayaka *et al.*, 2002, Dasanayaka *et al.*, 2003, Everad 1996, Perera *et al.*,1998, Perera *et al.*,2001, Rohde *et al.*, 1992, Teulat *et al.*, 2000). Mapping of the coconut genome too has been attempted by Heran *et al* 2000 and Bandaranayake *et al.*, 2002. The progress of this program has constrained by lack of proper mapping populations and limited number of informative molecular markers available for screening. Therefore, it is important that studies on coconut genome mapping be commenced in the country for gradual development towards achieving a proper map for use in marker-assisted selection of coconut in the near future. The objective of this study was to initiate work towards developing a molecular marker based map of coconut.

2. Materials and methods

Plant material

Material obtained for this study was from self-pollinated progenies of F_1 coconut hybrids established at the CRI of Sri Lanka. Sri Lankan tall and Sri Lankan dwarf have been used as parents of this mapping population to maximize the variations between parents, which helps to produce a reliable segregating mapping population for many loci. Nine different F_2 families derived from nine different F_1 trees which come from four different mother trees and 16 different father trees, have been developed and 3 families out of nine, comprising 76, 65 and 32 individuals respectively were used in this study.

DNA Isolation

Leaf materials were obtained from 4 female parents, 3 male parents, 9 F_1 plants and 173 F_2 plants and genomic DNA was isolated using the protocol described by Dasanayake *et al.*, 2003.

Selection of polymorphic SSR primers

Reasonable numbers of SSR primers have been developed by Perera *et al* (1999) and Rivera *et al* (1999) for the identification of microsatellite polymorphisms in coconut. Among them the following SSR primers were used for the analysis: CAC02, CAC03, CAC04, CAC06, CAC08, CAC10, CAC13, CAC20, CAC21, CAC23, CAC38, CAC39, CAC50, CAC52, CAC56, CAC65, CAC68, CAC71, CAC72, CAC74, CAC77, CAC84 (Perera *et al*) CNZ04, CNZ06, CNZ10, CNZ12, CNZ18, CNZ21, CNZ26, CNZ29, CNZ37, CNZ40, CNZ43, CNZ44 and CNZ46 (Rivera *et al*) (Table 1).

The primers were first assayed with parental DNA and their F_1 for detection of polymorphisms. The PCR reactions were performed in 10-µl volumes with 1 µπ forward primer, 1 µπ reverse primer, 1 unit of *Taq* DNA polymerase (Promega) 0.2 mM each of dNTP (Pharmacia), 1x PCR buffer [50 mM potassium choloride, 10 mM Tris-HCl pH 9.0 (at 25°C), 0.1% Triton x-100] supplied with enzyme (Promega), 1.5 mM $MgCl_2$ supplied with enzyme (Promega), 30 ng of template DNA in a PTC-100 thermocycler (MJ Research, Inc) with 30 cycles of three steps of 1 min denaturation at 94°C, 1 min annealing at 51°C -58°C (depending on primer) and 2 minutes elongation at 72°C. The first cycle was preceded by a 3 minutes denaturation at 95°C, and the last cycle followed by a 2 minutes elongation at 72°C. Reaction products were electrophoresed on 6% denaturing polyacrylamide gels and visualized by staining with silver nitrate.

A total of 24 primers that recognized heterozygosity in F_1 were selected and used for screening of three families.

Data analysis

SSR alleles were scored as A for female homozygote B for male homozygote and H for heterozygote.

3. Results

Thirty-five SSR primer pairs were screened to detect polymorphisms in parents and 9 F_1 individuals of the mapping population. Among the primer pairs 26 (74%) detected polymorphisms. These primers were CAC04, CAC06, CAC08, CAC10, CAC20, CAC23, CAC38, CAC50, CAC52, CAC56, CAC65, CAC68, CAC72, CAC77, CNZ04, CNZ06, CNZ10, CNZ12, CNZ18, CNZ21, CNZ29, CNZ37, CNZ40, CNZ43, CNZ44 and CNZ46. Table 2 describes the approximate size of each band, resolve length and the distance of gel progression by amplified DNA of each primer. Figure 1 shows the segregation of SSR markers after amplification by primer pair CNZ06 within the parents and F_1 individuals of the mapping population.

Six monomorphic loci were observed with primer pairs CAC03, CAC13, CAC21, CAC71, CAC84 and CNZ26. PCR did not perform well for primer pairs, CAC02,

CAC39 and CAC74 even after several repeated attempts. Although CAC20 and CNZ18 exhibited polymorphic loci in both parents and F_1 they were excluded from the analysis because of indistinguishable short distance between two alleles of the locus and the faintness alleles on silver stained gels. Finally, 24 primers were selected to genotype all 173 F_2 individuals based on the ability to resolve polymorphic bands, intensity and ability to amplify with all the individuals. A description of band scoring with all the 24 primers against 173 individuals of the three families are given in Table 3. Figure 2 shows segregation of SSR alleles CNZ04 and CNZ12 in some F_2 individuals.

Table 1. Details of the primers used in detection of SSRPs in segregating populations of coconut.

Primer	Expected product size(bp)	Used Annealing Temp (^{0}C)	Primer	Expected product size(bp)	Used Annealing Temp (^{0}C)
CAC02	220	58	CAC72	130	58
CAC03	190	58	CAC74	176	53.5
CAC04	200	58	CAC77	131	53.5
CAC06	158	58	CAC84	160	58
CAC08	190	58	CNZ04	162	53.5
CAC10	198	58	CNZ06	85	58
CAC13	160	58	CNZ10	148	53
CAC20	128	52	CNZ12	214	58
CAC21	154	58	CNZ18	102	55
CAC23	192	53.9	CNZ21	255	58
CAC38	155	53.5	CNZ26	232	53
CAC39	153	53.8	CNZ29	135	51
CAC50	153	53.5	CNZ37	232	53
CAC52	156	53.5	CNZ40	151	53
CAC56	154	53.5	CNZ43	197	53.5
CAC65	151	53.5	CNZ44	165	58
CAC68	142	58	CNZ46	116	51
CAC71	183	54			

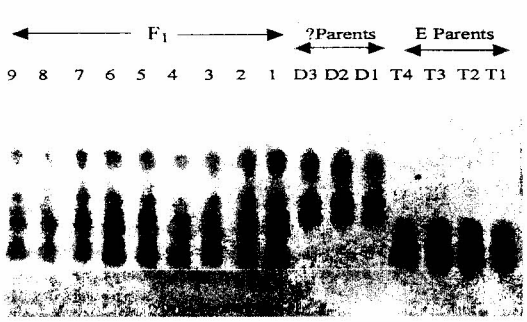

Figure 1: Simple sequence repeat (SSR) alleles detected by primer CNZ06, in parents and F_1 individuals of the cross dwarf green x tall coconut, visualized in

silver stained 6% denaturing polyacrylamide gels. Lanes T1, T2, T3 and T4 refer to the 4 female parents used for the production of mapping population. D1, D2 and D3 refer to the male parents used for production of mapping populations. Lanes 1-9 refer to 9 F_1 individuals.

Some markers identified contaminants in populations investigated. CNZ21, CNZ10, CAC56 markers identified contaminants in all the families studied. Marker CNZ12 identified contaminants in families 1 and 2. CAC65 identified contaminants in family 2 and CNZ04 identified in family 3. CAC43 identified contaminants only in family1. CNZ 204, CAC08, CAC38, CNZ37, CNZ29, CNZ68 identified contaminants only in family 2.

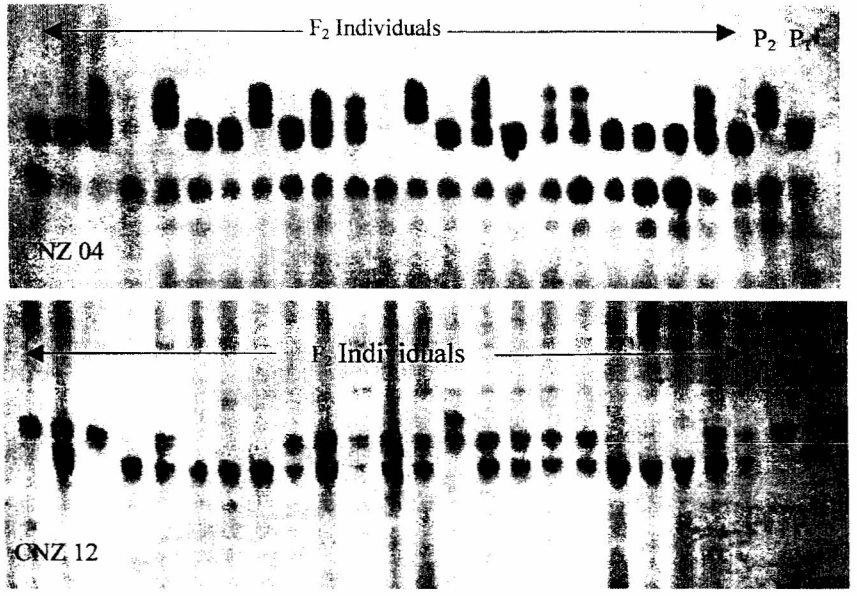

Figure 2: Segregation of SSR alleles CNZ04, and CNZ12 in F_2 individuals derived by selfing Dwarf Green x Tall F_1 coconuts. P_1 and P_2 are female and male parents respectively.

4. Discussion

Rohde *et al.*, (1999) established an F_1 population of 52 genotypes from a cross between a Malayan Yellow Dwarf (MYD) and a Laguna Tall (LAGT) genotype for linkage mapping and QTL analysis. Herran *et al.*, (2000) applied the above family for mapping the coconut genome. These genotypes were scored for the character germination time to commence QTL analysis. For molecular marker linkage mapping they have used 382 markers resulting from PCR of 28, 24, 5 and 44 primer combinations of RAPD, ISTR, ISSR and AFLP respectively. Of these 211 segregating fragments came from LAGT parent and 149 came from MYD parent.

However, 36 fragments showed significant deviations from the expected pattern of segregation. The map constructed covered the entire genome with 16 linkage groups representing the 16 chromosomes of coconut (2n = 32). The respective lengths of the genomes corresponded to LAGT and MYD were 2226 cM and 1266 cM respectively with 4-32 markers per linkage group. They have also identified six loci linked to early germination in coconut which is genetically correlated to early flowering and yield providing an opportunity for MAS in coconut.

Table 2. The approximate size of each band, resolved length and the distance of gel progression by amplified DNA of each primer used for the development of map.

Primer	Band size (bp)	Distance between alleles and length of gel progress	Remarks
CAC04	200	0.5 cm at 12 cm	Alleles were very clear and scored easily
CAC06	158	0.5 cm at 15 cm	Alleles were very clear and scored easily
CAC08	190	0.3 cm at 12 cm	Alleles were very clear and scored easily
CAC10	198	0.1 cm at 10.5 cm	Alleles were very clear and scored easily
CAC23	192	0.3 cm at 14 cm	Alleles were very clear and scored easily
CAC38	155	0.4 cm at 9.5 cm	Alleles were very clear and scored easily
CAC50	153	0.3 cm at 9 cm	Alleles were clear and scored easily
CAC52	156	0.1 cm at 11 cm	Alleles scoring were difficult due to very short distance between them.
CAC56	154	0.3 cm at 11.5 cm	Alleles were not so clear and scoring was not so easy
CAC65	151	0.6 cm at 10.5 cm	Alleles were very clear and scored very easily
CAC68	142	0.5 cm at 13.5 cm	Alleles were clear. But alleles haven't amplified in some individuals.
CAC72	130	0.2 cm at 12 cm	Alleles were not so clear due short distance between them and scoring was not easy.
CAC77	131	0.4 cm at 13 cm	Alleles were clear and could score.
CNZ04	162	1.5 cm at 11.5 cm	Alleles were not so strong but could score due to higher distance between them.
CNZ06	85	1.8 cm at 16 cm	Alleles were clear and could score very easily.
CNZ10	148	2.0 cm at 12 cm	Different alleles have amplified with this primer and scoring was not so easy.
CNZ12	214	0.4 cm at 11 cm	Alleles were very clear and could score easily.
CNZ21	255	0.5 cm at 13.5 cm	Alleles were clear and could score easily
CNZ29	135	2.5 cm at 17.5 cm	Alleles were very clear and could score easily
CNZ37	232	0.6 cm at 10 cm	Alleles were very clear and could score easily
CNZ40	151	1.0 cm at 15.5 cm	Alleles were clear and could score easily
CNZ43	197	1.0 cm at 11 cm	Alleles were not so clear in some individuals
CNZ44	165	0.4 cm at 11.5 cm	Alleles were clear and could score easily
CNZ46	116	1.0 cm at 17.5 cm	Alleles were not so clear and scoring was difficult.

In the present study an F_2 population segregating from F_1 individuals of the cross, dwarf green x Sri Lanka tall were attempted as a preliminary exercise toward generating a genome map of coconut. Sri Lanka tall and Sri Lanka dwarf have been selected as parents of this mapping population to maximize the variation between

parents and thereby to produce a reliable segregating mapping population for many loci. As all the P_2 palms has died, DNA extracted from another 3 dwarf palms was used for genotyping. This procedure would not add any different P_2 alleles because almost all loci in dwarf coconuts are identically homozygous having undergone self-pollination over several generations.

Table 3: Segregation of 24 pairs of microsatellite alleles in three F_2 families arose from self-pollination of three individual tall x dwarf green (F_1) coconut palms at the Coconut Research Institute.

Primer	Number of individuals											
	Family 1				Family 2				Family 3			
	P_1 type	F_1 type	P_2 type	Not Scored	P_1 type	F_1 type	P_2 type	Not Scored	P_1 type	F_1 type	P_2 type	Not Scored
CAC04	19	37	15	05	22	31	10	02	08	16	05	03
CAC06	21	32	17	06	14	27	19	05	09	14	06	03
CAC08	23	33	16	04	20	36	07	02	07	20	04	01
CAC10	21	25	26	04	19	29	12	05	08	18	02	04
CAC23	25	40	09	02	20	38	05	02	10	17	03	02
CAC38	20	39	15	02	20	28	14	03	13	11	05	03
CAC50	31	12	30	03	23	27	11	04	04	09	16	03
CAC52	25	29	19	03	00	20	43	02	08	17	04	03
CAC56	18	40	13	05	12	20	25	08	12	06	10	04
CAC65	22	38	14	02	22	31	08	04	10	15	04	03
CAC68	15	43	16	02	14	19	17	15	07	17	05	03
CAC72	17	37	19	03	18	37	08	02	07	16	09	00
CAC77	12	36	26	02	12	21	30	02	05	14	12	01
CNZ04	29	29	17	01	17	34	05	09	10	14	06	02
CNZ06	32	31	13	00	23	26	16	00	08	17	04	03
CNZ10	00	40	33	03	16	37	12	00	02	21	09	00
CNZ12	19	39	11	07	25	31	07	02	11	17	02	02
CNZ21	22	12	39	03	20	19	24	02	08	12	10	02
CNZ29	27	29	17	03	30	28	06	01	12	15	03	02
CNZ37	00	28	46	02	21	29	11	04	08	17	05	02
CNZ40	16	45	13	02	22	32	09	02	11	14	03	04
CNZ43	00	21	53	02	01	31	29	04	05	14	12	01
CNZ44	22	41	13	00	21	36	06	02	10	18	02	02
CNZ46	20	36	14	06	20	34	03	08	07	20	03	02

SSR markers were the choice, due to its high level of heterozygosity, availability of sequence information (Perera *et al* 1999 and Rivera *et al* 1999) and the codominant and PCR based nature. Segregation of twenty-four SSR markers in 173 F_2 individuals of coconut however, was found inadequate to generate a linkage map with JOINMAP and any other software available.

Among 35 pairs of primers tested 24 gave a consistent pattern of SSR alleles in the segregating populations. The study however, found among 35 primer pairs tested only a few as ideal for use in terms of band clarity and ease of recognition of polymorphisms. Among the good primers are CNZO6, CNZ12, CNZ29, CACO4, CACO8, CAC23, CAC38, CAC65 and CNZ37. In addition in the order of

preference the following primer pairs were also acceptable for screening coconut genotypes CAC10, CAC50, CAC77, CNZ04, CNZ40, CNZ44, CAC52, CAC56, CAC68, CAC72, CNZ10, CNZ21, CNZ43 and CNZ46.

The biggest obstacle for mapping coconut genome is the lack of proper segregating populations. In the current assay the data obtained with markers CAC04, CAC08, CAC38, CAC43, CAC56, CAC65, CNZ04, CNZ10, CNZ12, CNZ21, CNZ29, CNZ37, CNZ68 indicated presence of contaminants in the families assessed. The markers CNZ10, CAZ21 and CAC56 mostly identified contaminants and they identified contaminants in all the studied families. Further genotyping was not proceeded due to this unexpected banding patterns of contaminants. These contaminations may have occurred as illegitimate pollinations by negligence or by mites invading through pollination bags. Another likely mishap is inappropriate numbering and mixing of seeds in the nursery. Therefore, although the potential of mapping the coconut genome is realized in the present study the need to prepare for genome mapping by construction of segregating populations with utmost care is seen as a high priority. Coconut being a large perennial crop and the time gap of pollination to a scorable seedling is about 2-3 years much emphasis should be given for constructing families with accurate identification. This should be the prime task of the Coconut Research Institute prior to commencement of coconut genome mapping.

This preliminary study has surfaced much useful information especially on the most suitable SSR markers for coconut genome mapping. This study generated information on markers that can be applied to any other studies of coconut genome due to their high reproducibility.

5. Acknowledgement:

This work was supported by SAREC grant for Capacity Building in Biotechnology. The authors greatly acknowledge Dr. L. Perera, Dr C. Perera and staff of the Genetics and Plant Breeding Division of the Coconut Research Institute of Sri Lanka and staff of the Institute of Biochemistry, Molecular Biology and Biotechnology Sri Lanka for their cooperation.

6. References

Bandaranayake C (2002) Linkage mapping and QTL analysis of *Arabidopsis thaliana* and *Cocos nucifera* L, PhD thesis, University of Birmingham, UK.

Dasanayake, P.N., Everard, J.M.D.T., Karunanayake, E.H. and Nandadasa, H.G. (2002) Genetic relatedness of coconut (*Cocos nucifera* L.) germplasm in Sri Lanka as revealed by Amplified Fragment Length Polymorphisms ($AFLP_s$) 58th annual session of SLAAS (ISSN1391-023X pg 151 D-414).

Dasanayake, P.N., Everard, J.M.D.T., Karunanayake, E.H. and Nandadasa, H.G. (2003) characterization of coconut Germplasm by microiatellite markers. Tropical agricultural Research vol.15:51-60.

Everard JMDT (1996) Use of molecular markers for breeding of the coconut palm *Cocos nucifera*. MSc thesis, University of New England, Armidale, NSW, Australia.

Herran A, Estioko L, Becker D, Rodriguez MJB, Rohde W, Ritter E, (2000) Linkage mapping and QTL analysis in coconut (*Cocos nucifera* L.) Theor Appl Genet 101:292-300.

Nei M, Li WH (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proceedings of National Academy of Science, USA. 76:5269-5273.

Perera L, Russell JR, Provan J, Powell W (1998) Evaluating genetic relationships between indigenous coconut (*Cocos nucifera* L.) accessions from Sri Lanka by means of AFLP profiling. Theor Appl Genet 96: 545-550.

Perera L, Russell JR, Provan J, Powell W (1999) Identification and characterisation of microsatellite loci in coconut (cocos nucitera L.) and the analysis of coconut populations in Sri Lanka. Molecular Ecology **8**:344-346

Perera L, Russell JR, Provan J, Powell W (2001) levels And distribution of genetic diversity of coconut (*Cocos nucifera* L., var. *Typica* form *typica*) from Sri Lanka assessed by microsatellite markers. Euphytica 122: 381-389.

Rivera R, Edwards KJ, Barker JHA, Arnold GM, Ayad G, Hodgkin T, Karp A (1999) Isolation and characterization of polymorphic microsatellites in *Cocos nucifera* L. Genome 42:668-675.

Rohde W, Becker D, Kullaya A, Rodriguez MJB, Herran A, Ritter E (1999) Analysis of coconut germplasm biodiversity by DNA marker technologies and construction of a first genetic linkage map. In: Oropeza C, Verdeil JL, Ashburner GR, Cardega R, Santamarma JM (eds) Current advances in coconut biotechnology. Kluwer Academic Publ, Dordrechtm the Netherlands, pp 99-120.

Rohde W, Salamini F, Ashburner GR, Randles JW (1992) An *Eco*RI repetitive sequence family of the coconut palm *Cocos nucifera* L. shows sequence homology to *copia*-like elements. j.Genet and Breed. 46:391-394.

Teulat B, Aldam C, Trehin R, Lebrun P, Barker JHA, Arnold GM, Karp A, Baudouin L., Rognon F (2000) An analysis of genetic diversity in coconut (*Cocos nucifera*) populations from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. Theor Appl Genet 100: 764-771.