Isolation and characterization of polymorphic microsatellites in *Cocos nucifera* L.

R. Rivera, K.J. Edwards, J.H.A. Barker, G.M. Arnold, G. Ayad, T. Hodgkin, and A. Karp

Abstract: Microsatellites or simple sequence repeats (SSRs) were isolated from coconut (*Cocos nucifera)* and tested for polymorphism on restricted germplasm. Sequencing of 197 clones from a cv. Tagnanan Tall-enriched genomic library showed that 75% contained a microsatellite, of which 64% were dinucleotide (GA/CT, CA/GT and GC/CG), 6% were trinucleotide, and 30% were compound repeats. Of 41 primer pairs tested on Tagnanan Tall genomic DNA, 38 gave the expected size product, two amplified two loci, and another gave a multilocus pattern. On 20 coconut samples, the 38 SSRs detected 198 alleles (average: 5.2 alleles per microsatellite). Genetic diversity ($D = 1 - \Sigma p_i^2$) values ranged from 0.141 to 0.809. Heterozygotes were present at high frequencies among some dwarf samples. Analysis of similarity matrices based either on shared alleles at each locus (simple matching coefficient) or on allele bands across all loci (Jaccard coefficient) showed similar results. Dwarfs grouped separately from talls and showed less genetic diversity. In a wider test on 40 samples, 8 SSRs detected 64 alleles (average: eight alleles per microsatellite). These results indicate the high potential of microsatellites to detect genetic diversity in coconut germplasm.

Key words: molecular markers, microsatellite, SSR, *Cocos nucifera*, coconut.

Résumé : Des microsatellites ont été isolés chez le cocotier (*Cocos nucifera*) et examinés pour leur polymorphisme parmi un ensemble limité de germoplasme. Le séquençage de 197 clones génomiques d'une banque enrichie du cv. Tagnanan Tall a montré que 75% de ceux-ci contenaient un microsatellite dont 64% étaient de type dinucléotidique (GA/CT, CA/GT et GC/CG), 6% de type trinucléotidique et 30% de type composé. Un ensemble de 41 paires d'amorces ont été testées sur l'ADN génomique du cv. Tagnanan Tall et 38 de celles-ci ont généré des produits de la taille attendue, deux ont amplifié deux loci et une autre paire a généré de multiples produits. Sur une collection de 20 échantillons de cocotiers, les 38 microsatellites retenus ont détecté 198 allèles (5,2 allèles par microsatellite en moyenne). Les valeurs de diversité génique ($D = 1 - \Sigma p_i^2$) variaient de 0,141 à 0,809. Des hétérozygotes étaient présents à des fréquences élevées parmi certains échantillons de matériel nain. Des analyses de matrices de similarité fondées sur les allèles partagés pour chaque locus (coefficient simple d'identité) ou sur les allèles présents chez tous les loci (coefficient de Jaccard) ont généré des résultats semblables. Les cocotiers nains ont été groupés séparément des cocotiers de grande taille et montraient moins de diversité génétique. Sur une collection plus vaste composée de 40 échantillons, huit microsatellites ont détecté 64 allèles (une moyenne de huit allèles par locus). Ces résultats mettent en lumière le fort potentiel des microsatellites pour la détection de diversité génétique chez le cocotier.

Mots clés : marqueurs moléculaires, microsatellites, *Cocos nucifera*, cocotier.

[Traduit par la Rédaction]

Introduction

Coconut (*Cocos nucifera* L.) is one of the most important plantation crops in the wet tropics. Its food and industry-based products play a very important role in the economies of many developing countries. Examples include national consumption in Indonesia and the export by the Philippines of coconut oil, copra, and dessicated coconut. Coconut is also important for the rural economy. It is adapted to conditions under which other crops would not grow well and is suitable for intercropping. It provides growers, the majority of whom farm small holdings, with cash resources without the need for high investment.

Cocos nucifera ($2n = 2x = 32$) is a member of the monocotyledonous family Aracaceae (Palmaceae). It is the only species of the genus *Cocos* belonging to the subfamily Cocoideae which includes 27 genera and 600 species. The crop is widely believed to have evolved in the Asian-Pacific area (Harries 1990), although there are conflicting theories regarding the origin and domestication of coconut (Whitehead 1976; Harries 1978). There are two major groups based on stature: (1) tall palms, also referred to as var. *typica* (Nar.), and (2) dwarf palms, also referred to as var. *nana* (Griff) (Santos et al. 1996). Tall palms are widely planted

Corresponding Editor: J.P. Gustafson.

Received June 8, 1998. Accepted December 15, 1998.

R. Rivera. Philippine Coconut Authority, Zamboanga Research Centre, P.O. Box 356, Zamboanga City, Philippines.
K.J. Edwards, J.H.A. Barker, G.M. Arnold, and A. Karp.[1] IACR-Long Ashton Research Station, Department of Agricultural Sciences, University of Bristol, Long Ashton, Bristol, BS41 9AF, U.K.
G. Ayad, and T. Hodgkin. International Plant Genetic Resources Institute, Via delle Sette Chiese 142, 00145 Rome, Italy.

[1]Author to whom all correspondance should be addressed.

both for household and commercial use and grow to a height of 20–30 m. They are slow to mature, flowering 6–10 years after planting, and are long lived with an economic life of about 60–70 years. They are normally cross-pollinating and therefore considered to be heterozygous. Dwarf palms, on the other hand, are believed to be mutants from tall types and have short statures of 8–10 m when 20 years old. They start bearing fruit in the 3rd year from planting, when they are scarcely 1 m tall. They have a short productive life of 30–40 years, are normally self-pollinating, and, therefore, are considered to be homozygous.

Coconut faces strong competition from other oil crops, such as oil palm. The cultivation of coconut is currently in relative decline in many countries due to low productivity, increased demand for timber, and devastating diseases. For successful economic production, old palms have to be replaced by locally adapted high-yielding coconut cultivars that are tolerant to local pests and diseases. Germplasm collections that contain significant amounts of genetic diversity within and among coconut populations are a prerequisite for an effective crop improvement programme.

The characterization and evaluation of coconut populations have relied mostly on morphological and agronomic traits (e.g., Akpan 1994; Sugimura et al. 1997). These procedures are time- and labour-consuming and do not provide an accurate measure of genetic diversity because many characters exhibit complex inheritance and are influenced by both environmental and genetic factors. Characterizations of coconut germplasm using isozymes (Carpio 1982), fruit components (Ashburner et al. 1997*a*), and cartenoid differences (Fernando et al. 1997) have also been reported. More recently, with the advent of molecular marker techniques, genetic diversity in coconut germplasm has been assessed at the DNA level using randomly amplified polymorphic DNA (RAPD) (Ashburner et al. 1997*b*), restriction fragment length polymorphism (RFLP) (Lebrun et al. 1998), and amplified fragment length polymorphism (AFLP) (Perera et al. 1998).

Microsatellites, or simple sequence repeats (SSRs), are becoming increasingly attractive markers in molecular breeding and diversity assessment (Morgante and Olivieri 1993; Powell et al. 1996) but have not yet been developed or applied using coconut. SSRs are short tandemly repeated sequence motifs of approximately 1–8 bp in length, which are scattered throughout the genome and can vary between individuals in repeat length. High frequencies of polymorphism have been described for SSRs in several plant species (Saghai-Maroof et al. 1994; Gupta et al. 1996; Chase et al. 1996). Primer pairs designed for the flanking sequences can be used in PCR reactions for site-specific amplification of the microsatellite, thereby producing sequence-tagged microsatellite markers (Morgante and Olivieri 1993; Powell et al. 1996). In the present study, we describe the development of coconut microsatellite primers, show their high degree of polymorphism in a limited set of germplasm, and discuss their potential for application in the genetic conservation of coconut.

Materials and methods

Plant materials

The plant materials used were obtained from the coconut gene bank of the Philippine Coconut Authority–Zamboanga Research Center (PCA–ZRC) in San Ramon, Zamboanga City, Philippines, and comprised a limited set of 20 cultivars (10 tall and 10 dwarf cultivars) used for testing all the microsatellites (Table 1, Group 1) and a more extensive set of 40 cultivars used for a wider screen with eight microsatellites (Table 1, Group 2). Cultivars present in each set were represented by different individual samples. All the materials are included in the Coconut Genetic Resources Database (CGRD) of the International Coconut Genetic Resources Network (COGENT), where details of their passport and characterization data can be found. A sample of at least 1 g/palm was taken from the youngest leaves (spear leaf) of the coconut palm. Each leaf sample (midrib removed) was placed in a 1.5-mL Eppendorf tube and sent to Long Ashton Research Station, where they were stored at –21°C.

DNA extraction

DNA was extracted from 125 mg frozen leaf material and purified using the Nucleon Phytopure kit (Scotlab, U.K.) with the addition of 2-mercaptoethanol in Reagent One.

Construction of small-insert genomic library

A small-insert library (250–700 bp) enriched for several microsatellite types was constructed by the method of Edwards et al. (1996) using genomic DNA prepared from coconut cv. Tagnanan Tall. For long-term storage at –21°C, 1536 colonies were picked from the primary transformation coconut plates into 16 microtitre plates.

Sequencing of clones and primer design

Individual library clones were cultured overnight in 5 mL LB medium (Sambrook et al. 1989) with ampicillin (100 µg/mL). Plasmid DNA from individual clones was isolated using the Wizard™ Minipreps DNA Purification system (Promega). Manual sequencing was performed using the AmpliCycle™ Sequencing kit (Perkin–Elmer; containing M13 (*lacZ*) forward and reverse primers, and AmpliTaq®DNA). The resulting DNA sequences were examined for the presence of microsatellite and the flanking regions were used for primer design. Automated sequencing was performed using the ABI Prism™ 377 DNA sequencer (Perkin–Elmer). All primers were synthesised by Genosys Biotechnologies (Europe) Ltd.

PCR assay for single-locus profiles and separation of amplified DNA fragments

PCR amplification was performed in 0.5-mL thin-walled tubes (NBL Gene Sciences Ltd.) using a Perkin–Elmer 480 thermal cycler. For the preliminary PCR, the reaction mixture (25 µL) contained 10 ng template DNA, 50 ng forward primer, 50 ng reverse primer, 200 µM of each dNTP (Promega), 1 unit *Taq* DNA polymerase (Gibco-BRL), 20 mM Tris–HCl (pH 8.4), 50 mM KCl, and 1.5 mM $MgCl_2$. Reactions were overlayed with one drop of mineral oil. The PCR regime consisted of an initial denaturation (94°C for 2 min), 35 cycles each consisting of 40 s denaturation (94°C), 1 min annealing (52°C), and 2 min elongation (72°C). At the end of the final run, an extension period of 10 min at 72°C was included. The PCR products were separated on a 2% agarose gel with ethidium bromide (100 µg/L). The gels were run at a constant voltage of 100 v for 2–3 h in 1× TBE buffer and observed under UV light.

Using the gradient program available with the Robocycler™ Gradient 40 thermal cycler (Stratagene), the optimal annealing temperature of each primer pair was determined. Once known, the forward primer was radiolabelled with [γ-^{33}P]dATP by T4 kinase and used in similar PCR reactions. For screening PCR products, the 25-µL reactions were mixed with 10 µL of stop solution (98% formamide, containing 10 mM EDTA, 0.01% (w/v) xylene cyanol, and 0.01% (w/v) Bromophenol blue) and denatured at 94°C for

Table 1. Coconut materials used in the study for testing the microsatellite primer sets.

Coconut cultivars		Code	Origin (noncountry listings are from the Philippines)
Group 1 (for screening of 41 SSR primer sets)			
1.	San Isidro Green Dwarf	1-SNID	Agusan del Sur
2.	Tagnanan Tall	2-TAGT	Davao del Norte
3.	Pilipog Green Dwarf	3-PILD	Davao City,
4.	Kapatagan Green Dwarf	4-KAPD	Lanao del Norte
5.	Sto. Niño Green Dwarf	5-SÑOD	Davao del Norte
6.	Mangipod Green Dwarf	6-MNDD	Davao City
7.	San Ramon Tall	7-SNRT	Zamboanga City
8.	Baybay Tall	8-BAYT	Baybay, Leyte
9.	Zamboanga Tall	9-ZAMT	Zamboanga City
10.	Kinabalan Green Dwarf	10-KIND	Davao del Sur
11.	Agta Tall	11-AGDT	Davao City
12.	Catigan Green Dwarf	12-CATD	Davao City
13.	Spicata Tall	13-SPIT	Davao City
14.	Baguer Green Dwarf	14-BAGD	North Cotabato
15.	Banigan Green Tall	15-BNGT	Zamboanga del Norte
16.	Laguna Tall	16-LAGT	Davao City
17.	Tacunan Green Dwarf	17-TACD	Davao del Sur
18.	Tahiti Tall[a]	18-PYT	Ivory Coast
19.	Bañga Green Dwarf	19-BÑAD	South Cotabato
20.	Rennel Island Tall[a]	20-RIT	Solomon Islands
Group 2 (for screening of eight polymorphic SSR primer sets)[b]			
1.	Kapatagan Green Dwarf	KAP	Lanao del Norte
2.	Mangipod Green Dwarf	MNDD	Davao City
3.	San Isidro Green Dwarf	SNID	Agusan del Sur
4.	Sri Lanka Green Dwarf[a]	SGD	Sri Lanka
5.	Tupi Green Dwarf	TUP	South Cotabato
6.	Cameron Red Dwarf[a]	CRD	Cameroon
7.	Galas Green Dwarf	GAL	Dipolog City, Zamboanga del Norte
8.	Equatorian Green Dwarf[a]	EGD	Ivory Coast
9.	Aromatic Green Dwarf	ARO	Davao City
10.	Malayan Yellow Dwarf	MYD	Davao City
11.	New Buswang Green Dwarf	BUS	Kalibo, Aklan
12.	Santo Nizo Green Dwarf	SND	Davao del Norte
13.	Malayan Red Dwarf	MRD	Davao City
14.	Tacunan Green Dwarf	TACD	Davao del Sur
15.	La Victoria Dwarf	VIC	Zamboanga del Sur
16.	Pilipog Green Dwarf	PIL	Davao City
17.	Magtuod Green Dwarf	MAG	Davao City
18.	Catigan Green Dwarf	CAT	Davao City
19.	San Ramon Tall	SNR	Zamboanga City
20.	Markham Valley Tall[a]	MVT	Papua New Guinea
21.	Macapuno Tall	MAC	Davao City
22.	Zamboanga Tall	ZAMT	Zamboanga City
23.	Agta Tall	AGAT	Davao City
24.	Gatusan Tall	GAT	Davao City
25.	Bago Oshiro Tall	BAO	Davao City
26.	Baybay Tall	BAYT	Baybay, Leyte
27.	Polynesian Tall[a]	PYT	Ivory Coast
28.	Laguna Tall	LAGT	Davao City
29.	Loong Tall	LON	Concepcion, Iloilo
30.	Tagnanan Tall	TAGT	Mabini, Davao del Norte
31.	Vanuatu Tall[a]	VTT	Vanuatu
32.	Karkar Tall[a]	KKT	Papua New Guinea
33.	Rennel Island Tall	RIT	Solomon Islands

Table 1 (*concluded*).

Coconut cultivars		Code	Origin (noncountry listings are from the Philippines)
34.	Tampakan Tall	TPK	South Cotabato
35.	West African Tall[a]	WAT	Ivory Coast
36.	Aguinaldo Tall	AGD	Zamboanga del sur
37.	Venus Tall	VEN	Zamboanga del Norte
38.	Culaman Tall	CUL	Davao del Sur
39.	Gazelle Tall[a]	GPT	Papua New Guinea
40.	Banigan Green Tall	BNGT	Zamboanga del Norte

[a]Foreign coconut populations included in the PCA-ZRC gene bank.
[b]Although some cultivars are in common with Group 1, different individual samples were taken.

Table 2. PCR annealing temperatures, repeat types, allele numbers, genetic diversity (*D*), and product sizes of 38[a] coconut SSRs[b] tested on 20 coconut cultivars.

No.	Name	Predicted length (bp)	Type of repeat	Annealing temp. (°C)	No.of alleles	*D*	Observed length (bp)
1	CNZ01	122	$(CT)_{15}$ $CA)_9$	55	6	0.643	110–128
2	CNZ02	138	$(GA)_{15}$	53	6	0.489	138–158
3	CNZ03	91	$(GA)_7$	51	3	0.141	90–93
4	CNZ04	162	$(CT)_{29}$ $TT(CA)_{10}$	53	7	0.764	130–166
5	CNZ05	163	$(CT)_{17}(GT)_7$	53	5	0.454	153–210
6	CNZ06	85	$(CT)_{15}$	53	6	0.666	69–97
7	CNZ09	137	$(CT)_{13}(CA)_6$ $(CT)_{19}(CA)_8$	53	5	0.445	122–141
8	CNZ10	148	$(CT)_{18}(GT)_{17}$	53	7	0.741	108–152
9	CNZ12	214	$(CT)_{15}$	53	3	0.535	218–229
12	CNZ16	186	$(GTT)_6(GCT)_2$ $(CTT)_8$ TTTGTGC $(GT)_8CT(GT)_7CT$ $(GT)_6$	55	5	0.480	163–207
13	CNZ17	107	$(CA)_{12}$ $(GA)_{11}$	55	4	0.729	93–113
14	CNZ18	102	$(CT)_{15}$ $TT(CT)_3$	55	8	0.767	101–121
15	CNZ19	181	$(CT)_{15}$ $(CA)_5$ CT $(CA)_3(CT)_2(CA)_6$	55	4	0.606	172–191
16	CNZ20	94	$(CA)_{13}$	55	3	0.554	88–94
17	CNZ21	255	$(CT)_{30}$	55	8	0.684	232–268
18	CNZ22	240	$(GT)_{18}$	55	6	0.654	181–275
19	CNZ23	163	$(GA)_{18}$	53	3	0.366	145–164
20	CNZ24	257	$(CT)_{19}$ $(GT)_{19}$	53	4	0.186	249–267
21	CNZ26	232	$(GT)_{17}$ $(GA)_{19}$	53	7	0.744	209–252
22	CNZ29	135	$(GT)_{22}(GA)_2$ $CA(GA)_{11}$	51	6	0.715	105–157
23	CNZ31	189	$(GT)_{14}$	51	3	0.584	182–195
24	CNZ32	148	$(GA)_{18}$	55	4	0.528	138–153
25	CNZ33	164	$(GT)_{11}$	51	2	0.489	152–169
26	CNZ34	140	$(CT)_{15}(GT)_7$	51	5	0.591	131–168
27	CNZ37	232	$(CT)_7CC(CT)_{13}(GT)_5$ $N_{23}(GT)_{29}$	53	9	0.776	193–251
28	CNZ40	151	$(CT)_{20}$	53	4	0.578	136–156
29	CNZ42	161	$(CT)_{16}$	55	3	0.405	164–173
30	CNZ43	197	$(GA)_{21}$	51	8	0.809	175–219
31	CNZ44	165	$(GA)_{15}$	53	3	0.411	151–170
33	CNZ46	116	$(CT)_{24}$	51	6	0.653	101–120
34	CN1H2	203	$(GA)_{18}$	52	4	0.669	191–207
35	CN1C6	183	$(GT)_{11}$ $TT(GT)_5$	52	4	0.689	176–186
36	CN2A5	1889	$(CT)_{30}$	52	5	0.346	163–191
37	CN1G4	129	$(CT)_{15}$	52	5	0.645	114–142
38	CN11E6	95	$(CT)_{21}$	52	5	0.713	93–128
39	CN2A4	88	$(CT)_{15}TT(CT)_3$	52	8	0.775	117–152
40	CN11E10	137	$(GT)_{22}(GA)_{14}$	52	6	0.715	85–105
41	CN11A10	111	$(CT)_{30}$	52	9	0.679	81–119

[a]The following SSRs are not listed: CNZ13 (no. 10) and CNZ14 (no. 11), as each produced a two loci pattern, and CNZ45 (no. 32) which produced a multilocus pattern.
[b]All primer sequences can be obtained from the corresponding author by request.

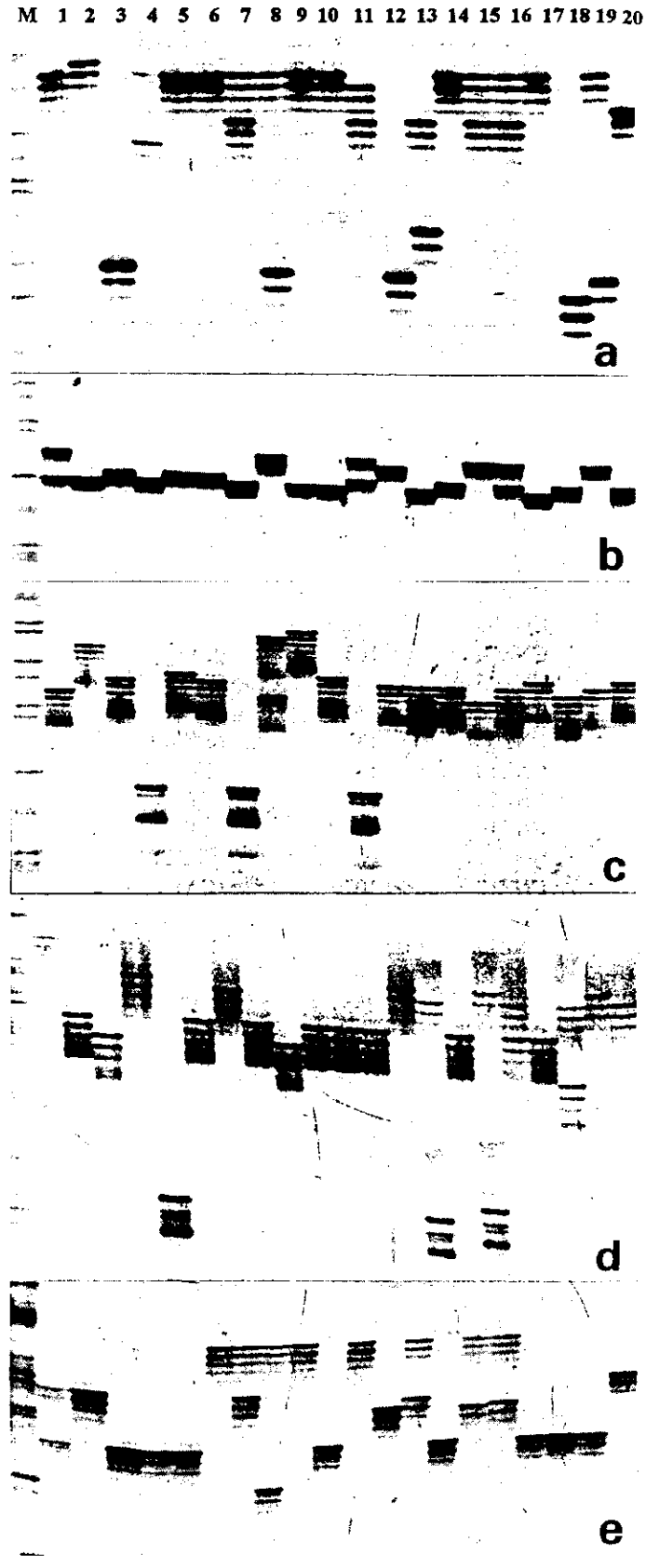

Fig. 1. Examples of microsatellite polymorphisms detected among 20 coconut samples using primers for loci (*a*) CNZ11410; (*b*) CN1C4; (*c*) CNZ43; (*d*) CNZ10; and (*e*) CNZ29. Genomic DNA was amplified from samples of the Group 1 cultivars 1–20 in Table 1: ((1) San Isadora Green Dwarf; (2) Tagnanan Tall; (3) Pilipog Green Dwarf ;(4) Kapatagan Green Dwarf; (5) Sto. Niñi Green Dwarf; (6) Mangipod Green Dwarf; (7) San Ramon Tall; (8) Baybay Tall; (9) Zamboanga Tall; (10) Kinabalan Green Dwarf; (11) Agta Tall; (12) Catigan Green Dwarf; (13) Spicata Tall; (14) Baguer Green Dwarf; (15) Banigan Green Tall; (16) Laguna Tall; (17) Tacunan Green Dwarf ; (18) Tahiti Tall; (19) Bañga Green Dwarf; (20) Rennel Island Tall). Lane M was the sequence marker used for sizing the alleles (two sequence markers are present in (*d*). Where stutter bands were observed (as in *a*, *c*, *d*, and *e*) the darkest band was scored as the allele. Some loci were largely free from stutter (see locus CN1C4 (*b*)). When two dark bands appeared together, (lanes 4, 7, and 11 for locus CNZ43 (*c*)), they were scored as a single allele, using the higher molecular weight band for the sizing since individuals heterozygous for alleles of this kind were observed (lanes 13 and 15, locus CNZ10 (*d*)). Lane 19 for locus CN11410 (*a*) is an example of a locus heterozygous in a dwarf. Nearly all other heterozygotes appear in tall samples.

3 min. A 4-μL aliquot was loaded onto denaturing polyacrylamide gels (6% acrylamide/bisacrylamide 40:2 or 50% Sequagel stock solution diluted 10-fold with 1× TBE – 7 M urea in 1× TBE buffer) and electrophoresed at 55 W constant power for 2 h. The gels were transferred to 3 mm paper and dried in a Bio-Rad 583 gel drier at 80°C for 30 min. Once dried, the gel was exposed to Kodak Biomax MR film for 15 h.

Data analysis

Microsatellite loci were scored individually and the different alleles were recorded for each sample screened. Two different intersample similarity matrices were constructed. The first was based on shared alleles using the simple matching coefficient, calculated at each locus separately and then meaned across loci. The second treated each allele for every locus as a separate band and the Jaccard coefficient was used as the measure of similarity, analagous to the manner in which RAPD and AFLPs are scored. In both cases, a cluster analysis was performed on the similarity matrix using the unweighted pair group method with arithmetic averages (UPGMA) and the resultant dendrograms (phenograms) were constructed. The similarity matrices were also put into a principle coordinates analysis (PCO) and scores for the resultant first three components were plotted pair-wise. Genetic diversity ($D = 1 - \sum P_i^2$) values were calculated according to Nei (1973).

Results

Sequencing of 197 clones of a coconut (cv. Tagnanan Tall) genomic library highly enriched for several microsatellites indicated that 148 colonies (75%) contained at least one microsatellite. Of the 148 clones, 40, 23, and 1% contained dinucleotide GA/CT, CA/GT, and GC/CG repeats, respectively, while 30 and 6% of clones, respectively, contained compound (containing perfect and imperfect repeats) and trinucleotide (GCC/CTT) repeats.

Primers flanking the repeat were designed from clones in which the SSR was of a size greater than 12 bp and positioned sufficiently distant from the cloning site. Forty-one primer pairs (8 designed from manual sequencing and 33 from automated sequencing) were tested for their ability to amplify Tagnanan Tall genomic DNA with the expected product (Table 2). Six microsatellites had CA/GT repeats ranging from 11 to 22, 19 had GA/CT repeats ranging from 7 to 30, 2 had trinucleotide GCC repeats both with a size of 6, and 14 had compound (both perfect and imperfect) repeats with core repeats ranging from 22 to 54. Perfect

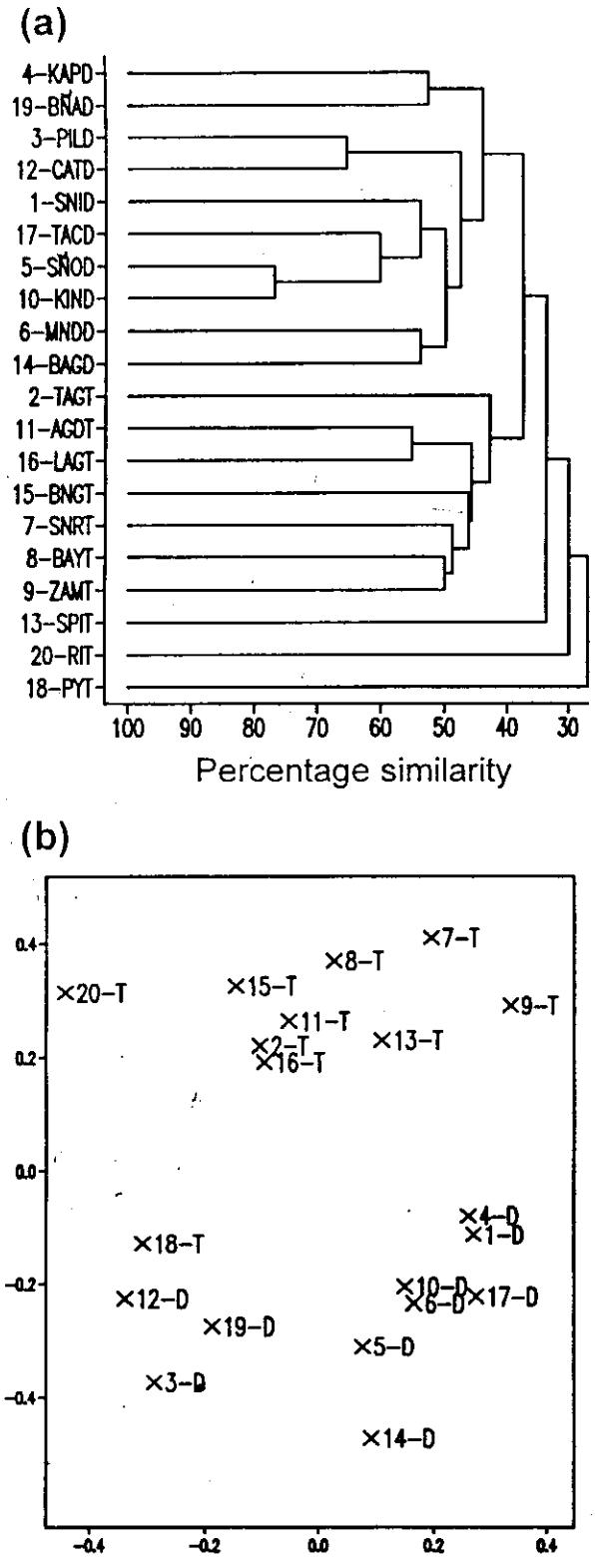

Fig. 2. Dendrogram (*a*) and PCO plot (*b*) from the similarity matrix constructed using a simple matching coefficient (i.e., allele sharing). The dendrogram results from UPGMA cluster analysis of the similarity matrix. The PCO plot shows the scores for components 1 versus 2 resulting from a PCO of the similarity matrix. Numbers given on the *y* axes of the dendrogram and for the points on the PCO plots refer to the clone numbers listed in Table 1. Because of space restrictions, only the clone numbers and the letter T (for tall) and D (for dwarf) are given in the PCO plot.

microsatellites were found in 32 clones, while 9f clones contained interrupted repeats, 5 of which were compound, i.e., consisted of more than one motif.

Using the optimum annealing temperature for each primer pair, PCR was performed using the DNA extracted from 20 coconut populations (Table 1, Group 1). The results obtained from the 41 SSR primer pairs are shown in Table 2. All but three SSR markers were polymorphic with up to nine different alleles observed in the 20 coconut cultivars. The exceptions were CNZ14, which detected two monomorphic bands, CNZ45, which amplified as a complex multiloci SSR, and CNZ13, which appeared to amplify two loci, with all individuals homozygous at each locus. For the remaining 38 SSRs, the number of alleles detected ranged from 2 to 9, while the number of genotypes ranged from 3 to 11. Examples of polymorphisms detected in the 20 samples with different microsatellites are shown in Fig 1. The 38 SSR markers detected a total of 198 alleles, giving an average of 5.2 alleles per microsatellite. Genetic diversity ($D = 1 - \Sigma p_i^2$) values ranged from 0.141 to 0.809 (Table 2).

Interestingly, although heterozygosity values were higher (as expected) in tall than in dwarf cultivars, in three dwarf cultivars higher frequencies of heterozygous loci were detected. Thus, out of the 38 microsatellites tested, heterozygotes were present for 11 loci in 1-SNID and for 4 loci each in 5-SÑOD, 4-KAPD, 17-TACD, and 19-BÑAD.

Similarity matrices were constructed using (*i*) a simple matching coefficient (i.e., allele sharing) and (*ii*) the presence or absence of bands (allele frequencies) using the Jaccard coefficient (Sneath and Sokal 1973). Cluster analyses were performed using UPGMA and phenograms were constructed. A PCO was also carried out on the similarity matrices, with scores on the resultant first three components plotted pair-wise in each case. Results were similar for both analyses using a simple matching coefficient and the presence or absence of bands, and are shown for the simple matching coefficient only (Fig. 2). On both the phenogram (Fig. 2*a*) and PCO (Fig. 2*b*), samples from the dwarf cultivars were grouped separately from those from the tall cultivars, which showed higher genetic diversity than the dwarf.

In addition to the 20 samples, as a wider test screen, 40 coconut samples (Table 1, Group 2) were screened with eight SSRs. The total number of alleles detected was 64, with an average of 8 per microsatellite, and genetic diversity (*D*) ranged from 0.579 to 0.790 (Table 3).

Discussion

A set of 38 informative microsatellite (SSR) markers have been isolated in coconut from an enriched library constructed from Tagnanan Tall genomic DNA using the previously described method of Edwards et al. (1996). A high percentage of clones (>60%) contained either GA/CT, CA/GT, or GC/CG dinucleotide repeats, which is similar to earlier reports of microsatellites in other plant species, such as apple (Guilford et al. 1997), oak (Dow et al. 1995; Steinkellner et al. 1997), wheat (Bryan et al. 1997), rice (Panaud et al. 1995), *Pinus* (Smith and Devey 1994), and maize (Chin et al. 1996).

Table 3. Results of testing eight microsatellite loci[a] on 40 different coconut samples.

No.	SSR Name	Predicted product length	Type of repeat	Annealing temperature	(°C)	No. of alleles	*D*
1	CN1H2	203	$(GA)_{18}$	52	7	0.723	182–211
2	CN1C6	183	$(GT)_{11}$ $TT(GT)_5$	52	4	0.718	175–185
3	CN2A5	1889	$(CT)_{30}$	52	6	0.579	151–191
4	CN1G4	129	$(CT)_{15}$	52	8	0.756	114–139
5	CN11E6	95	$(CT)_{21}$	52	10	0.751	86–128
6	CN2A4	88	$(CT)_{15}TT(CT)_3$	52	7	0.717	88–110
7	CN11E10	137	$(GT)_{22}(GA)_{14}$	52	13	0.790	91–150
8	CN11A10	111	$(CT)_{30}$	52	9	0.739	65–141

[a]All primer sequences can be obtained from the corresponding author by request.

One advantage that microsatellites offer as molecular markers is the polymorphism that they can reveal and particularly the high numbers of allelic variants that may be detected (Morgante and Olivieri 1993; Powell et al. 1996). A preliminary screen of 20 coconut samples from the PCA gene bank indicated that the SSRs isolated here conform to this expectation. High levels of polymorphism were detected with an average of 5.2 alleles per microsatellite and genetic diversity values (*D*) from 0.141 to 0.809. This was further confirmed by a wider screen of 40 samples using only eight of the SSRs, where an average of eight alleles per SSR were detected.

As a first basic check that the diversity revealed by the SSRs was consistent with known aspects of coconut germplasm, genetic similarity matrices were constructed and put into clustering alogorithms to assess the relationships among the clones analysed. Not surprisingly, cluster analyses using either allele sharing or the presence or absence of microsatellites as input data separated tall and dwarf cultivars clearly, and also revealed greater diversity among the talls. A more surprising feature, however, was the detection of heterozygous loci at quite high frequencies in some of the dwarf cultivars. This may suggest that outcrossing has occurred in the origin of these cultivars.

The value of SSRs for germplasm characterization has now been widely demonstrated in many crops (e.g., Maughan et al. 1995; Provan et al. 1996; Taramino and Tingey 1996; Taramino et al. 1997; Russell et al. 1997; Guilford et al. 1997; Jarret et al. 1997). The work undertaken here was part of a collaborative effort with COGENT, which plans to collect threatened and useful coconut genetic diversity to augment national collections. For this endeavour, sensitive and reliable techniques for assessing genetic diversity will be an important adjunct for the establishment of effective collecting and gene-banking strategies. There are many possible alternative methods for detecting genetic diversity using molecular markers, but only a few are suitable for application in wide-scale efforts spanning several continents. For effective uptake by the COGENT network, it is imperative that techniques provide results that are reproducible by different laboratories and can be analysed using standardized scoring and data analysis methods. Ideally, they should also provide data that can be easily entered into databases and necessary materials should be easily exchangeable and not too technically demanding to apply. Finally, the data derived from the molecular techniques should help to provide the right type of information for the design of effective conservation strategies.

Microsatellites clearly have many features which help meet these criteria: (1) they are mostly single-locu codominant markers; (2) they provide high quality allelic information; (3) their data can be analysed as allele frequencies permitting direct computation in population genetics; (4) they can be detected by a simple PCR-based assay; (5) small amounts of material are required for the assays, which are less affected by DNA quality than, for example, RFLPs or AFLPs ; (6) they are highly reproducible, and can be exchanged between laboratories and, (7), they are suitable for use in databases.

The high reproducibility of microsatellites in network activities has recently been demonstrated (Jones et al. 1997) and there are increasing numbers of examples of their use for identification of cultivars and for the construction of databases (e.g., Thomas et al. 1994; Guilford et al. 1997; Russell et al. 1997; Bowers and Meredith 1997). In light of this, and considering the ease of assay and informativeness of microsatellites as population tools, we believe that the identification of a highly polymorphic set of coconut microsatellites will make an important contribution to the evaluation of both in situ and ex situ genetic resources of this important plantation crop. Work is under way to use these microsatellites in a more extensive germplasm screen.

Acknowledgements

This work was support by the Holdback Fund Facility of the Overseas Development Administration, U.K. The authors acknowledge the valuable input of Pons Batuga of the International Coconut Genetic Resources Network (COGENT) and of the Philippine Coconut Authority. Thanks are also given to V. Ramanatha Rao of IPGRI for helpful discussions. IACR receives grant-aided support from the Biotechnology and Biological Sciences Research Council (BBSRC) of the U.K.

References

Akpan, E.E.J. 1994. Evaluation of tall coconut (*Cocos nucifera* L) genotypes within the Nigerian coconut germplasm bank. Oleagineux, **49**:13–20.

Ashburner, G.R., Thompson, W.K., Halloran, G.M., and Foale, M.A. 1997*a*. Fruit component analysis of south Pacific coconut palm populations. Genet. Resour. Crop Evol. **44**: 327–335.

Ashburner, G.R., Thompson, W.K., and Halloran, G.M. 1997*b*. RAPD analysis of South Pacific coconut palm populations. Crop Sci. **37**: 992–997.

B[illegible]s J.E., and Meredith, C.P. 1997. The parentage of a classic [illegible]pe wine, Cabernet Sauvignon. Nat. Genet. **16**: 84–87.

B[illegible]n, G.J., Collins, A.J., Stephenson, P. Orry, A., Smith, J.B., and Gale, M.D. 1997. Isolation and characterisation of microsatellites from hexaploid bread wheat. Theor. Appl. Genet. **94**: 557–563.

Carpio, C.B. 1982. Biochemical studies of *Cocos nucifera* L. Kalikasan. Philipp. J. Biol. **11**: 319–338.

Chase, M., Kessli, R., and Bawa, K. 1996. Microsatellite markers for population and conservation genetics of tropical trees. Am. J. Bot. **83**: 51–57.

[illegible]in, E.C.L., Senior, M.L., Shu, H., and Smith, J.S.C. 1996. Maize simple sequence repetitive DNA sequences: abundance and allele variation. Genome, **39**: 866–873.

Dow, B.D., Ashley, M.V., and Howe, H.F. 1995. Characterisation of highly variable (GA/CT)n microsatellites in the bur oak, *Quercus macrocarpa*. Theor. Appl. Genet. **91**: 137–141.

Edwards, K.J.E., Barker, J.H.A., Daly, A., Jones, C., and Karp, A. 1996. Microsatellite libraries enriched for several microsatellite sequences in plants. BioTechniques, **20**: 758–760.

Fernando, W.M.U., Perera, L., and Peries, R.R.A. 1997. An overview of breeding research in coconut—The Sri Lankan experience. Outlook Agric. **26**: 191–198.

Guilford, P., Prakash, S., Zhu, J.M., Rikkerink, E., Gardiner, S., Bassett, H., and Forster, R. 1997. Microsatellites in *Malus x domestica* (apple): abundance, polymorphism and cultivar identification. Theor. Appl. Genet. **94**: 249–254.

Gupta, P.K., Balyan, H.S., Sharma, P.C., and Ramesh, B. 1996. Microsatellites in plants: a new class of molecular markers. Curr. Sci. **70**: 45–54.

Harries, H.C. 1978. The evolution dissemination and classification of *Cocos nucifera* L. Bot. Review **44**: 205–317.

Harries, H.C. 1990. Malesian origin for a domestic *Cocos nucifera*. *In* The plant diversity of Malesia. *Edited by* P. Baas, K. Kalkman, R. Geesink. Kluwer, Dordrecht, the Netherlands. pp. 351–357.

Jarret, R.L., Merrick, L.C., Holms, T., Evans, J., and Aradhya, M.K. 1997. Simple sequence repeats in watermelon (*Citrellus lanatus*) (Thunb.) Matsum. & Nakai). Genome, **40**: 433–441.

Jones, C.J., Edwards, K.J., Castaglione, S., Winfield, M.O., Sala, F., Van De Wiel, C., Bredemeijer, G., Vosman, B., Matthes, M, Daly, A., Brettschneider, R., Bettini, P., Buiatti, M., Maestri, E., Malcevschi, A. Marmiroli, N., Aert, R., Volckaert, G., Rueda, J., Linacero, R., Vazquez, A., and Karp, A. 1997. Reproducibility testing of RAPD, AFLP & SSR markers in plants by a network of European laboratories. Mol. Breeding, **3**: 381–390.

Lebrun, P., N'Cho, Y.P., Seguin, M., Grivet, L., and Baudouin, L. 1998. Genetic diversity in coconut (*Cocos nucifera* L.) revealed by restriction fragment length polymorphism (RFLP) markers. Euphytica, **101**: 103–108.

Maughan, P.J., Saghai-Maroof, M.A., and Buss, G.R. 1995. Microsatellite and amplified sequence length polymorphisms in cultivated and wild soybean. Genome, **38**: 715–723.

Morgante, M., and Olivieri, A.M. 1993. PCR-amplified microsatellites as markers in plants. Plant J. **3**: 175–182.

Nei, M. 1973. Analysis of gene diversity in subdivided populations. Proc. Natl. Acad. Sci. U.S.A. **70**: 3321–3323.

Panaud, O., Chen, X.L., and McCouch, S.R. 1995. Frequency of microsatellite sequences in rice (*Oryza sativa* L.) Genome, **38**: 1170–1176.

Perera, A.U., Russell, J.R., Provan, J., McNicol, J.W., and 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.

Powell, W., Machray, G.C., and Provan, J. 1996. Polymorphism revealed by simple sequence repeats. Trends Plant Sci. **1**: 215–222.

Provan, J., Powell, W., and Waugh, R. 1996. Microsatellite analysis of relationships within cultivated potato (*Solanum tuberosum*). Theor. Appl. Genet. **92**: 1078–1084.

Russell, J., Fuller, J., Young, G., Thomas, B., Taramino, G., Macaulay, M., Waugh, R., and Powell, W. 1997. Discriminating between barley genotypes using microsatellite markers. Genome, **40**: 442–450.

Saghai-Maroof, M.A., Biyashev, R.M., Yang, G.P., Zhang, Q., and Allard, R.W. 1994. Extraordinarily polymorphic microsatellite DNA in barley: species diversity, chromosomal locations and population dynamics. Proc. Natl. Acad. Sci. U.S.A. **91**: 5466–5470.

Sambrook, J., Fritsch, E.F., and Maniatis, T. 1989. Molecular cloning. A laboratory manual. Cold Spring Harbour Laboratory Press, Cold Spring Habour, N.Y.

Santos, G.A., Batugal, P.A., Othman, A., Baudouin, L., and Labouisse J.P. 1996. Manual on standardised techniques in coconut breeding. IPGRI–COGENT publication. Stamford Press, Singapore.

Smith, D.N., and Devey, M.E. 1994. Occurrence and inheritance of microsatellites in *Pinus radiata*. Genome, **37**: 977–983.

Sneath, P.H.A., and Sokal, R.R. 1973. Numerical taxonomy: the principles and practice of numerical classification. Freeman, San Francisco, Calif.

Steinkellner, H., Fluch, S., Turetschek, E., Lexer, C., Streiff, R., Kremer, A., Burg. K., and Glössl, J. 1997. Identification and characterisation of $(GA/CT)_n$ microsatellite loci from *Quercus petraea*. Plant Mol. Biol. **33**: 1093–1096.

Sugimura. Y., Itano, M., Salud, C.D., Otsuji, K., and Yamaguchi, H. 1997. Biometric analysis on diversity of coconut palm: cultivar classification by botanical and agronomical traits. Euphytica, **98**: 29–35.

Taramino, G., and Tingey, S. 1996. Simple sequence repeats for germplasm analysis and mapping in maize. Genome, **39**: 277–287.

Taramino, G., Tarchini, R., Ferrario, S., Lee, M., and Pe, M.E. 1997. Characterisation and mapping of simple sequence repeats (SSRs) in *Sorghum bicolor*. Theor. Appl. Genet. **95**: 66–72.

Thomas, M.R., Cain, P., and Scott, N.S. 1994. DNA typing of grapevines: a universal methodology and database for describing cultivars and evaluating genetic relatedness. Plant Mol. Biol. **25**: 939–949.

Whitehead, R.A. 1976. Coconut. *In* Evolution of crop plants. *Edited by* N.W. Simonds. Longman, London. pp. 221–225.