Consolidate microsatellite data on coconut diversity

Appendices

LOA APO/08/007

Luc Baudouin

CR.UPR29/2009-2-2
Mars 2009

Appendix A: Updated information on the marker sets

Table A1: The initial marker set

a) Technical data

Marker	Motif		Forward_Primer	Accession GenBank	reference[1]
CnCirA3	(GT)14	F	AATCTAAATCTACGAAAGCA		2
		R	AATAATGTGAAAAAGCAAAG	AJ458309	
CnCirA9	(GT)9 (GA)8	F	AATGTTTGTGTCTTTGTGCGTGTGT		1
		R	TCCTTATTTTTCTTCCCCTTCCTCA	AJ458310	
CnCirB6	(GT)4 (GT)10 (GA)11	F	GAGTGTGTGAGCCAGCAT		2
		R	ATTGTTCACAGTCCTTCCA	AJ458311	
CnCirB12	(CA)20 (GA)15	F	GCTCTTCAGTCTTTCTCAA		3
		R	CTGTATGCCAATTTTTCTA	AJ458312	
CnCirC7	(GT)7 (GA)16	F	ACCAACAAAGCCAGAGC		2
		R	GCAGCCACTACCTAAAAAG	AJ458314	
CnCirC12	(CA)15 (TA)6	F	ATACCACAGGCTAACAT		2
		R	AACCAGAGACATTTGAA	AJ458315	
CnCirE2	(CT)17 (GT)9	F	TCGCTGATGAATGCTTGCT		2
		R	GGGGCTGAGGGATAAACC	AJ458316	
CnCirE10	(CA)8 (GA)11	F	TTGGGTTCCATTTCTTCTCTCATC		2
		R	GCTCTTTAGGGTTCGCTTTCTTAG	AJ458317	
CnCirE12	(CT)18 (GT)9 (CT)8	F	CAGAAACAGCCAATCAAGCAATA		2
		R	ATCCATAATAGCCACTCAACAAAAA	AJ458318	
CnCirF2	(GT)10 (GA)12	F	GGTCTCCTCTCCCTCCTTATCTA		2
		R	CGACGACCCAAAACTGAACAC	AJ458319	
CnCirG11	(GT)9 (GA)9	F	AATATCTCCAAAAATCATCGAAAG		2
		R	CATCCCACACCCTCCTCT	AJ458320	
CnCirH4'	(CT)8 CGCT (CA)5 (CGCA)6	F	TTAGATCTCCTCCCAAAG		1
		R	ATCGAAAGAACAGTCACG	AJ458321	
CnCirH7	(CT)16 (CA)13	F	GAGATGGCATAACACCTA		2
		R	TGCTGAAGCAAAAGAGTA	AJ458322	
CnCIRC3'	(CA)12(GC)6(AC)10 (AG)12	F	AGAAAGCTGAGAGGGAGATT		
		r	GTGGGGCATGAAAAGTAAC		

b) Diversity and location

	Alleles			Controls		Location	
Marker	number	Size range		T1	T2	Linkage group	Position (cM)
CnCirA3	14	218	254	228 228	240 240		
CnCirA9	13	87	115	089 089	097 097		
CnCirB6	14	196	228	198 204	202 202		
CnCirB12	26	129	189	163 163	169 169	3	139.2
CnCirC7	19	147	189	165 167	161 161		
CnCirC12	14	153	193	167 167	161 183		
CnCirE2	31	115	177	163 163	135 135		
CnCirE10	11	224	246	244 244	238 238		
CnCirE12	5	162	182	174 174	164 164		
CnCirF2	16	187	237	193 193	205 205		
CnCirG11	19	186	216	204 208	194 194		
CnCirH4'	6	218	236	230 230	230 230	3	139.2
CnCirH7	15	87	149	133 133	139 139		
CnCIRC3'	41	164	242	173 206	174 174		

[1] See table A3

Table A2: The extended marker set

a) Technical data

Marker	Motif	Forward_Primer	Accession GenBank	reference[2]
CnCir C5	(TG)9 (GA)28	F ACCAACAAAGCCAGAGC R GCAGCCACTACCTAAAAAG	AJ865094	3
CnCir F3'	(TC)17(CA)16	F CCCTACTACTCCCTCAT R TGCCTAGTCAATCATAC	AJ865103	5
CNZ42	(GA)21	F TGATACTCCTCTGTGATGCTT R GTAGATTGTGGGAGAGGAATG		4
CnCir 215	(CT)17	F TACCACCTAAAGAGGAATG R AGTATCTGGGTTTGGCT	AJ865299	1
CnCir 147	(CT)9(GT)2(CT)6	F TTTCTCACCAACAAATAAAC R CTTGTGTGTTAGGGTCATC	AJ865242	1
CnCir 206	(CT)19	F AAAGAGAACGCAACCA R CAAGTTCCAAAGAACCA	AJ865292	1
mEgCIR3400	(GA)16	F CAATTCCAGCGTCACTATAG R AGTGGCAGTGGAAAAACAGT	AJ578662	1
CnCir E1	(CA)11	F CTTGTTATGTCGTTTGTTG R CTGAGACCCTGTTGATGT	AJ865099	5
mEgCIR3750	(GA)16	F GATGTTGCCGCTGTTTG R CATCCCATTTCCCTCTT	AJ578720	1
Cncir H11	(TC)18	F TCATTCAGAGGACAAAAGTT R TAAAAATTCATAAAGGTAAAA	AJ865105	1
CNZ03	(GA)7	F CATCTTTCATCATTTAGCTCT R AAACCAAAAGCAAGGAGAAGT		4
CnCir 2	(TC)24 (CA)7	F AGTCCTAAAAGTGTTGGC R GTAATCCTATGGCTGCTT	AJ865124	1
CnCir I4	(CT)9(CA)8	F TCCTAGTGCTTATGCTTGAC R TTGATGGTTTGATGTTGAA	AJ865106	1
CnCir 126	(AAAG)5 (GA)10	F TAATGACCTCTGCCG R CCTGATTGGGTGTCTAT	AJ865224	1
mEgCIR2739	(GA)18	F ACCCAATTTGCAACTGA R GAGTGGTTGGGATCAGTAT		1
CnCir 119	(GA)9 AA (AG)19	F AGAACCCTTGCTCCAC R TCCAGCCATTCCATC	AJ865217	1
CNZ40	(CT)20	F CTTGATTGCTATCTCAAATGG R CTGAGACCAAATACCATGTGT		4

b) diversity and location

	Alleles			Controls		Location	
Marker	Number	Size range		T1	T2	Linkage group	Position (cM)
CnCir C5	28	86	140	108 108	132 132	4	64.9
CnCir F3'	34	136	199	176 176	172 172	4	64.9
CNZ42	18	140	192	158 158	166 166	10	75.3
CnCir 215	26	114	176	126 126	141 141	2	40.7
CnCir 147	14	205	233	211 211	223 223	2	38.9
CnCir 206	27	81	135	099 131	107 107	2	31.9
mEgCIR3400	17	118	158	138 142	150 150	4	58.7
CnCir E1	8	215	235	225 233	227 227	4	58.7
mEgCIR3750	9	117	139	129 129	127 127	5	41.8
Cncir H11	30	135	215	151 151	173 173	5	47.9
CNZ03	10	87	127	091 095	089 089	5	47.9
CnCir 2	27	206	270	222 222	248 248	10	43.4
CnCir I4	4	285	292	285 285	285 285	15	68.3
CnCir 126	6	179	190	179 184	184 184	15	73.2
mEgCIR2739	25	187	243	213 213	193 193	1	59
CnCir 119	23	174	228	202 202	214 214	1	65.2
CNZ40	19	122	174	152 152	152 152	10	43.4

[2] See table

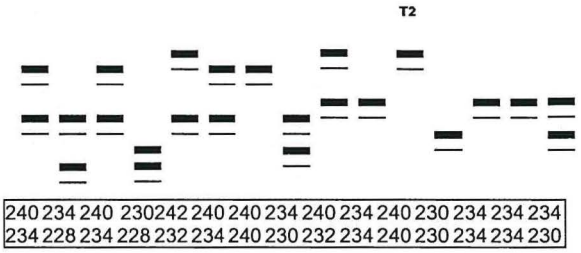
Table A3 : Bibliographic references for the markers

	References
1	Baudouin, L., P. Lebrun, et al. (2005). QTL analysis of fruit components in the progeny of a Rennell Island Tall coconut (Cocos nucifera L.) individual. Theor. Appl. Genet. 112: 258-268.
2	Lebrun, P., A. Berger, et al. (2006). Germplasm characterization, evaluationand use: Biochemical and molecular methods for characterizing coconut diversity in Coconut Genetic Resources (Batugal P., Ramanatha Rao V. and Oliver J. editors), IPGRI
3	Lebrun, P., L. Baudouin, et al. (2001). "Construction of a linkage map of the Rennell Island Tall coconut type (cocos nucifera L.) and QTL analysis for yield characters." Genome 44: 962-970.
4	Teulat, B., C. Aldam, et al. (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.
5	P. Lebrun & L. Baudouin &W. Myrie & A. Berger & M. Dollet (2008). Recent lethal yellowing outbreak: why is the Malayan Yellow Dwarf Coconut no longer resistant in Jamaica?

Appendix B: Annotated gel pictures

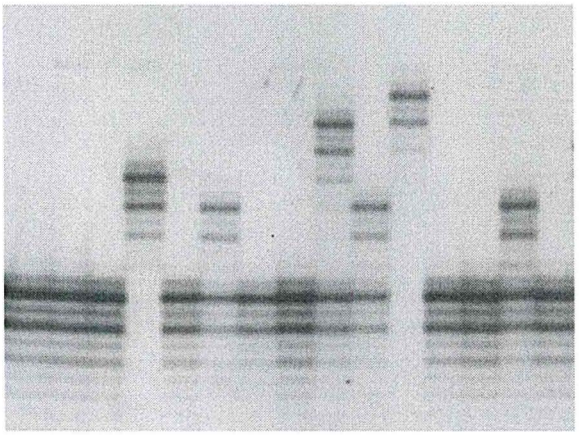
B1: Initial kit (silver staining)

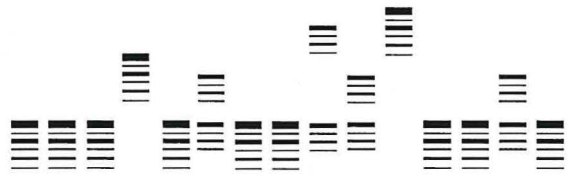
CnCir A3



240	234	240	230	242	240	240	234	240	234	240	230	234	234	234
234	228	234	228	232	234	240	230	232	234	240	230	234	234	230

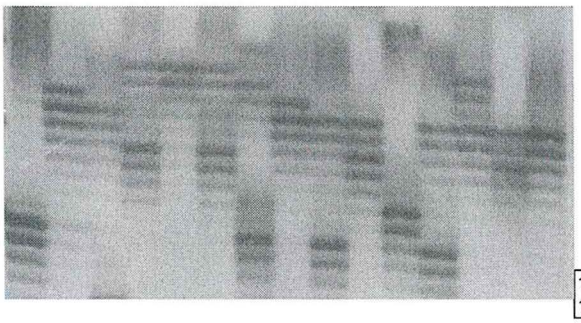
CnCir A9

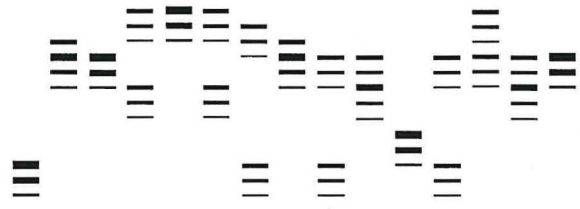




89	89	89	97	89	95	89	89	101	95	103	89	89	95	89
89	89	89	97	89	89	89	89	89	89	103	89	89	89	89

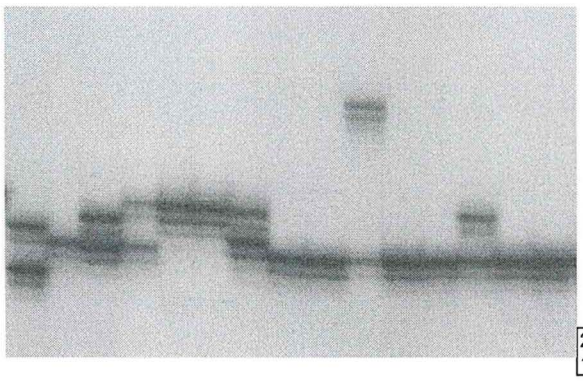
CnCir B12





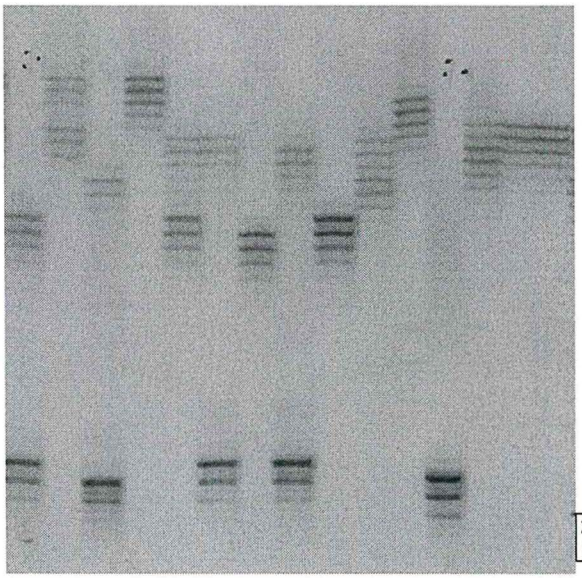
159	173	173	169	179	169	159	173	159	169	163	159	173	169	173
159	175	173	179	179	179	177	175	173	173	163	173	179	169	173

CnCir B6



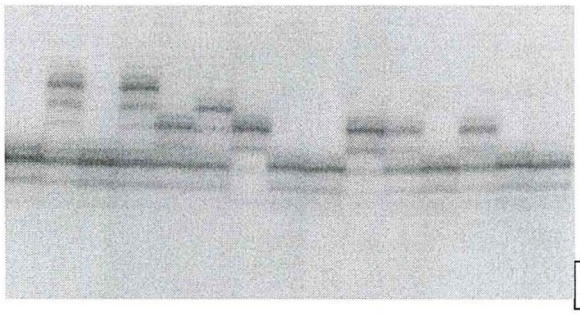
204 202 206 208 208 208 208 202 202 226 202 202 208 202 202
198 202 202 202 208 208 204 202 202 202 202 202 202 202 202

CnCir C3'



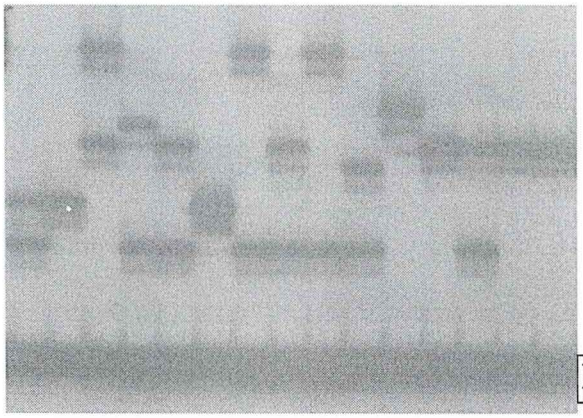
206 228 212 228 218 218 204 216 206 218 224 174 220 220 220
176 220 174 226 206 176 204 176 206 212 224 174 216 218 218

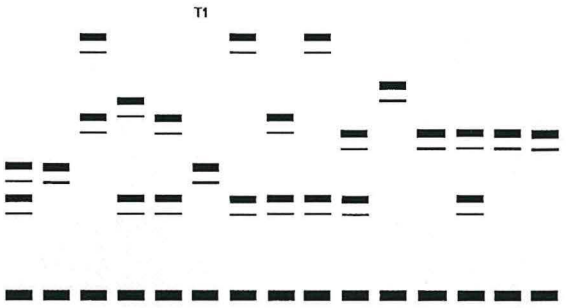
CnCir C7



157 165 157 165 161 163 161 157 157 161 161 157 161 157 157
157 157 157 157 157 157 161 157 157 161 157 157 157 157 157

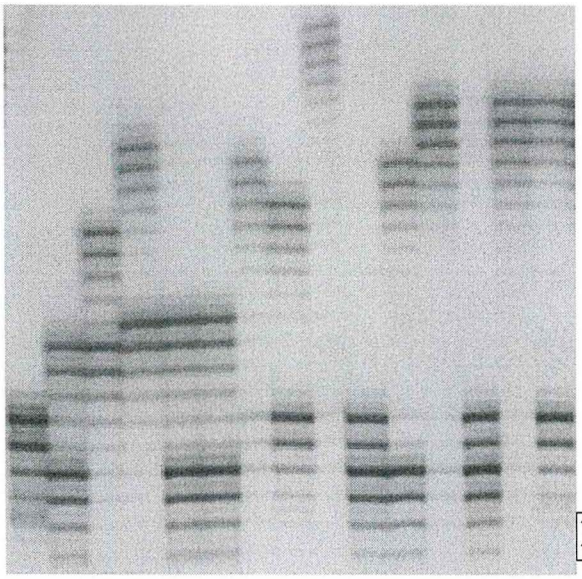
CnCir C12

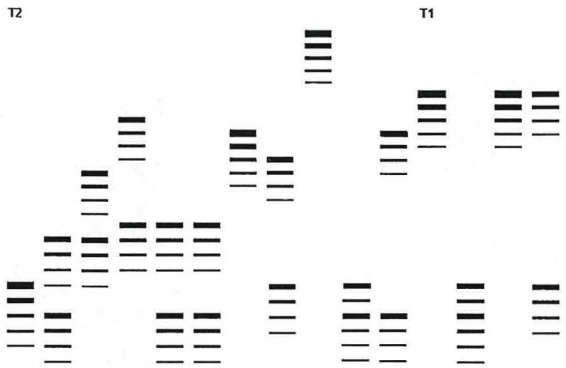




158 158 174 166 164 158 174 164 174 162 168 164 164 164 164
154 158 164 154 154 158 154 154 154 154 168 164 154 164 164

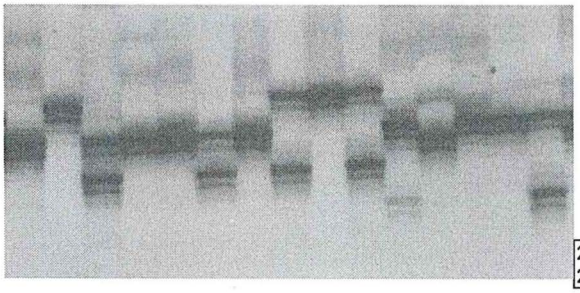
CnCir E2

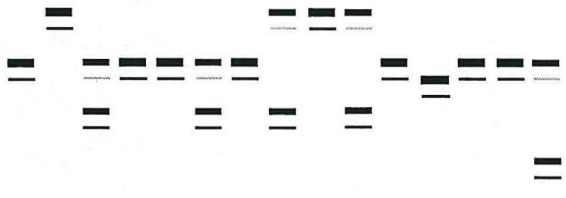




135 141 151 159 143 143 157 153 173 135 157 163 135 163 163
135 131 141 143 131 131 153 135 173 131 131 163 131 161 135

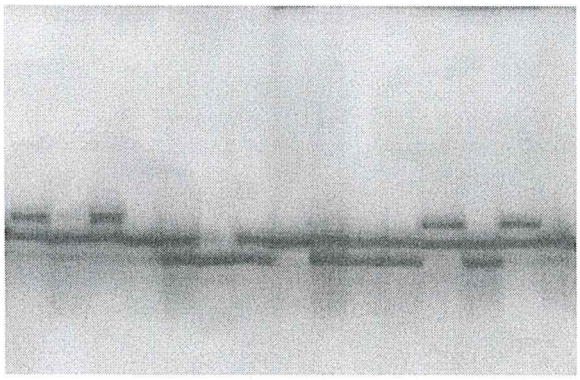
CnCir E10

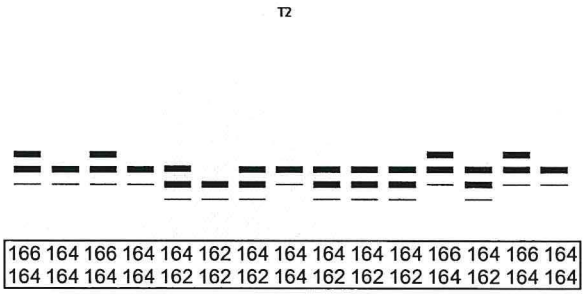




238 244 238 238 238 238 238 244 244 244 238 236 238 238 238
238 244 232 238 238 232 238 232 244 232 238 236 238 238 226

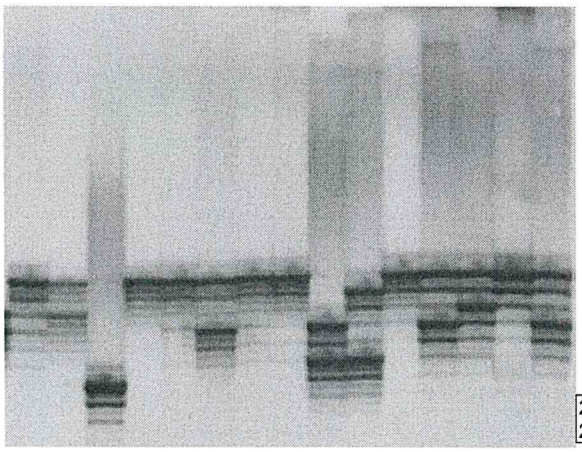
CnCir E12

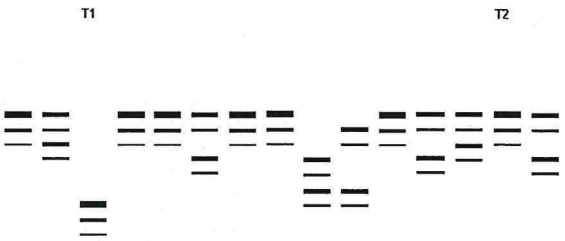




166 164 166 164 164 162 164 164 164 164 164 166 164 166 164
164 164 164 164 162 162 162 164 162 162 162 164 162 164 164

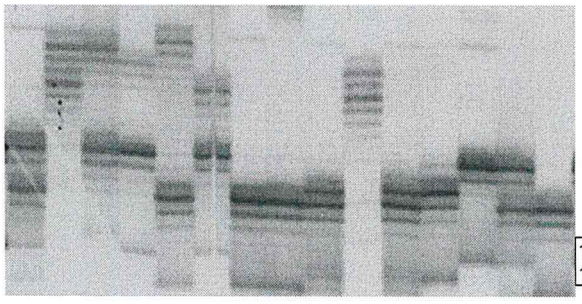
CnCir F2

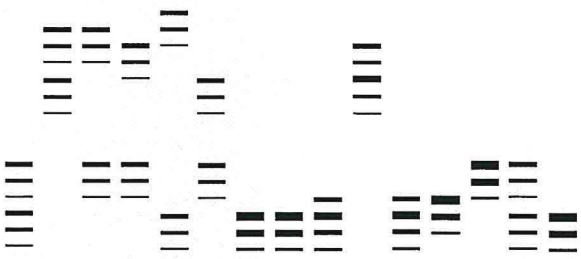




205 205 193 205 205 205 205 205 199 203 205 205 205 205 205
205 201 193 205 205 199 205 205 195 195 205 199 201 205 199

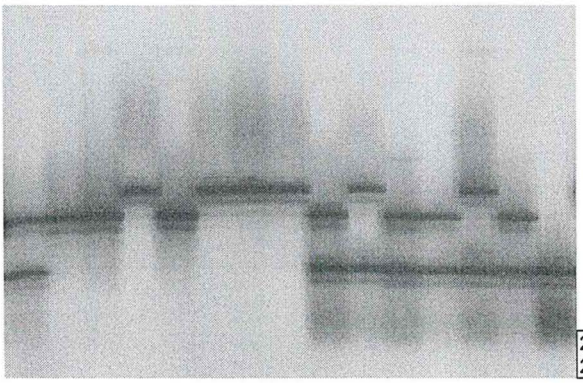
CnCir G11

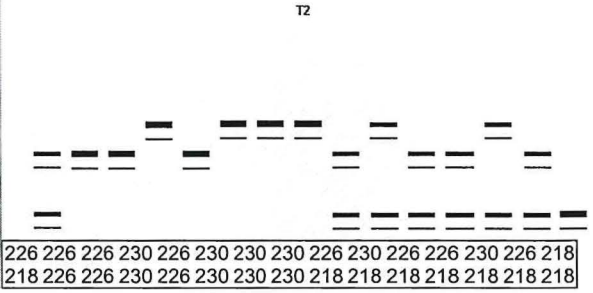




194 210 210 208 212 204 188 188 190 208 190 190 194 194 188
188 204 194 194 188 194 188 188 188 204 188 190 194 188 188

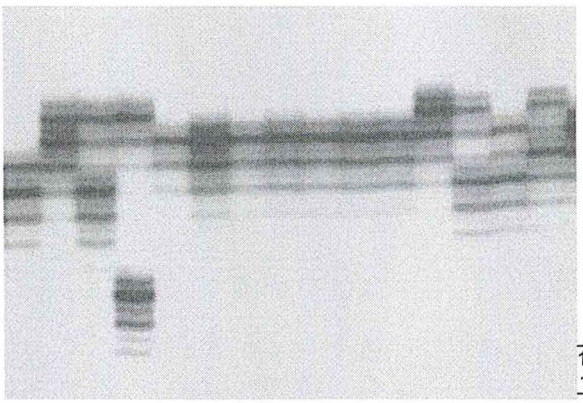
CnCir H4'

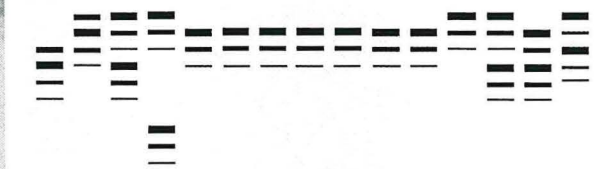




226 226 226 230 226 230 230 230 226 230 226 226 230 226 218
218 226 226 230 226 230 230 230 218 218 218 218 218 218 218

CnCir H7

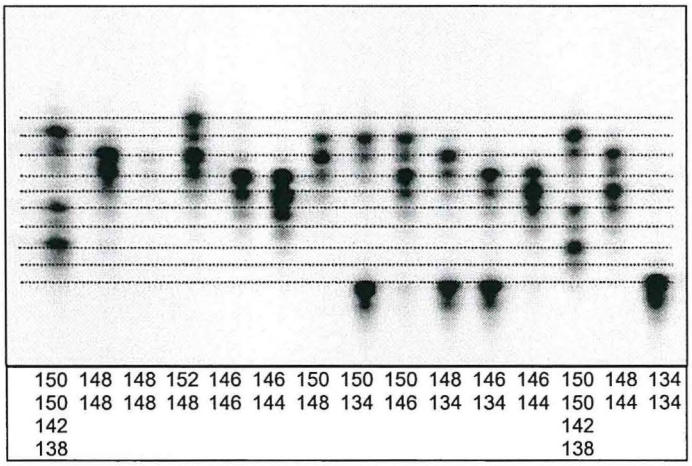




137 141 141 141 139 139 139 139 139 139 139 141 141 139 141
135 137 135 127 139 139 139 139 139 139 139 141 135 135 137

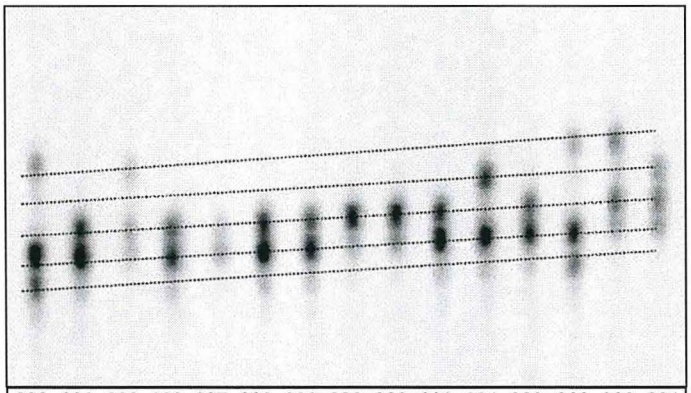
B2: Additional markers of the extended kit (Fluorescence)

mEg3400



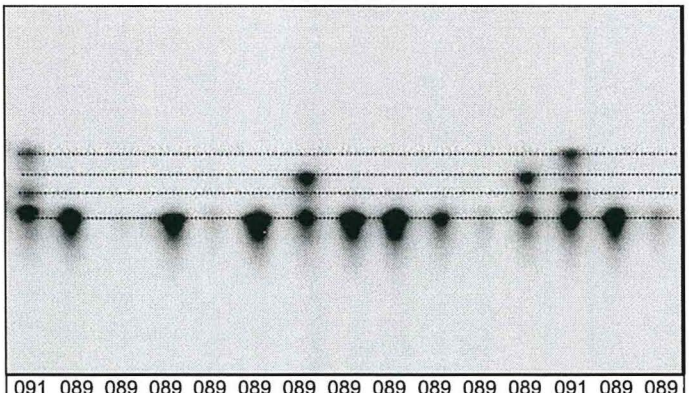
150	148	148	152	146	146	150	150	150	148	146	146	150	148	134
150	148	148	148	146	144	148	134	146	134	134	144	150	144	134
142												142		
138												138		

CnCirE1



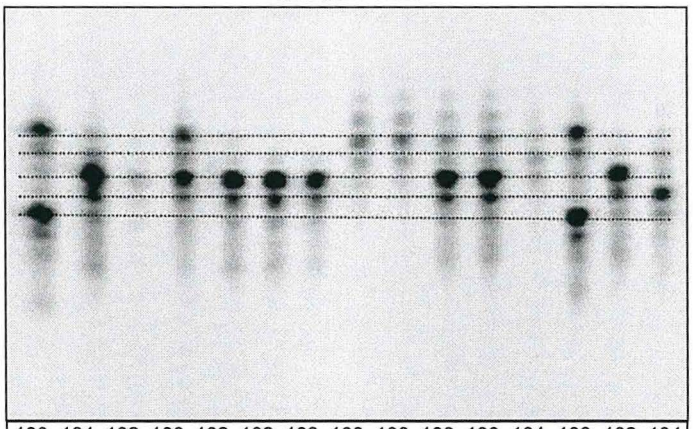
233	229	233	229	227	229	229	229	229	229	231	229	233	233	231
227	227	229	227	227	227	227	229	229	227	227	227	227	229	229
227												227		
225												225		

CNZ03



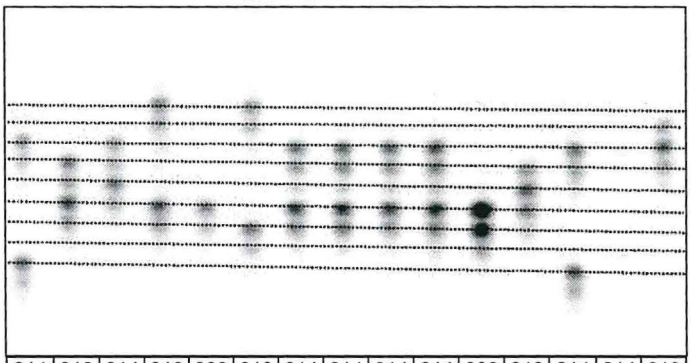
091	089	089	089	089	089	089	089	089	089	089	089	091	089	089
089	089	089	089	089	089	093	089	089	089	089	093	089	089	089
089												089		
095												095		

CNZ42



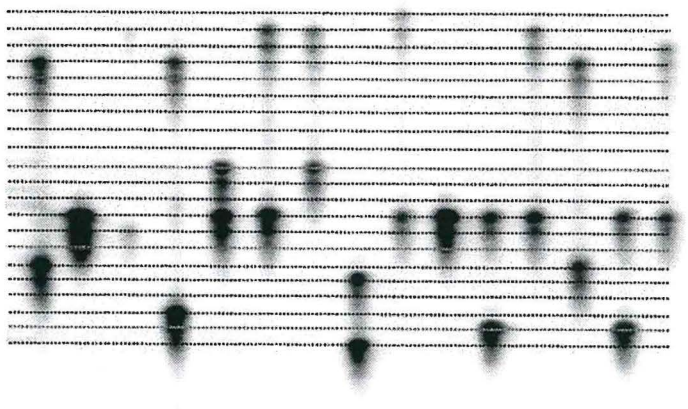
166	164	162	166	162	162	162	166	166	166	166	164	166	162	164
166	162	162	162	162	162	162	166	166	162	162	162	166	162	160
158												158		
156												156		

CnCir 119



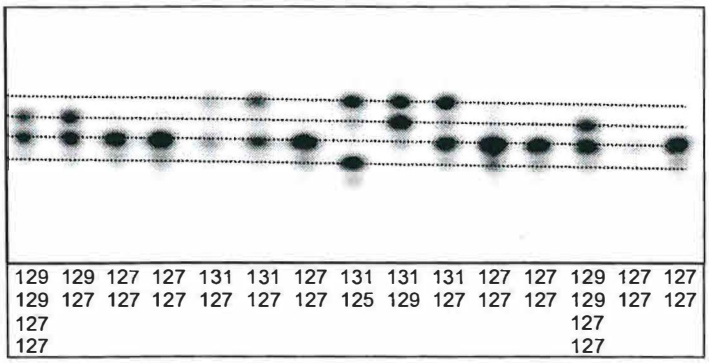
214	212	214	218	208	218	214	214	214	214	208	212	214	214	216
214	208	210	208	208	206	208	208	208	208	206	210	214	202	214
202												202		
202												202		

CnCir C5



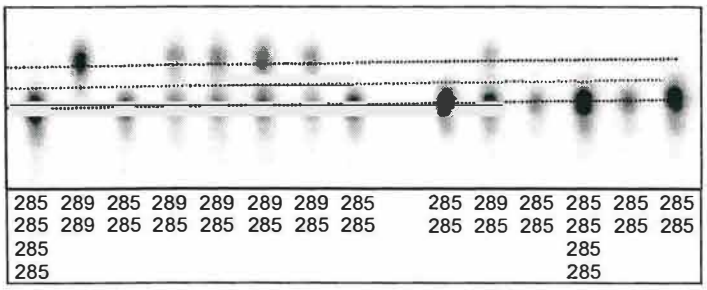
132	114	136	132	120	136	136	106	138	114	114	136	132	114	134
132	114	112	102	114	114	120	098	114	114	100	114	132	100	114
108												108		
108												108		

mEg3750



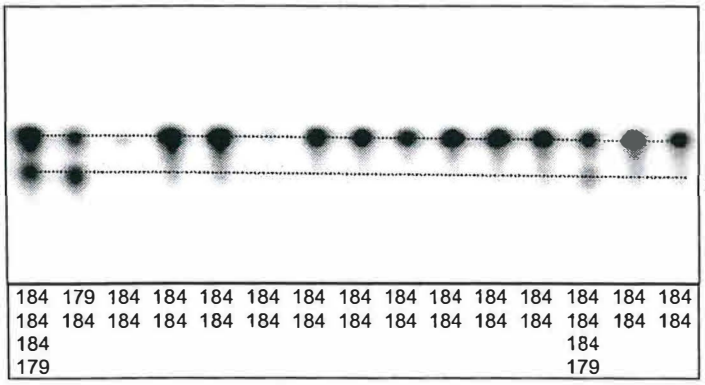
129	129	127	127	131	131	127	131	131	131	127	127	129	127	127
129	127	127	127	127	127	127	125	129	127	127	127	129	127	127
127												127		
127												127		

CnCirI4



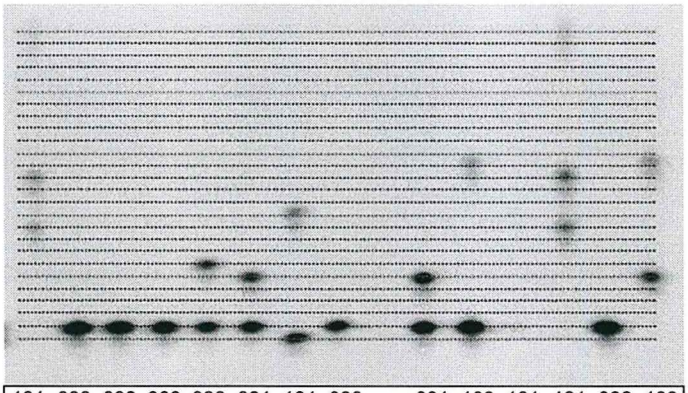
285	289	285	289	289	289	289	285	285	289	285	285	285	285	
285	289	285	285	285	285	285	285	285	285	285	285	285	285	
285											285			
285											285			

CnCir126



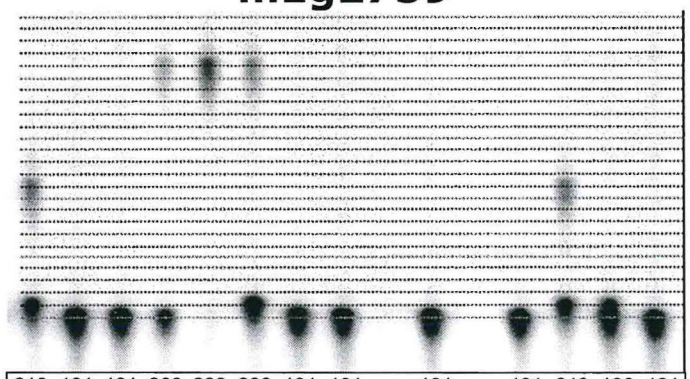
184	179	184	184	184	184	184	184	184	184	184	184	184	184	184
184	184	184	184	184	184	184	184	184	184	184	184	184	184	184
184												184		
179												179		

CnCir206



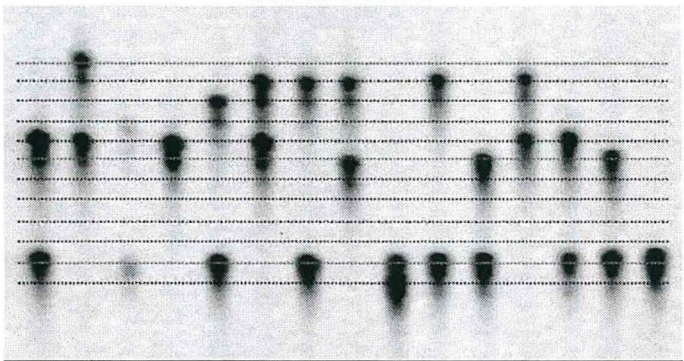
131	083	083	083	093	091	101	083		091	109	101	131	083	109
107	083	083	083	083	083	081	083		083	083	083	107	083	091
107												107		
099												099		

mEg2739



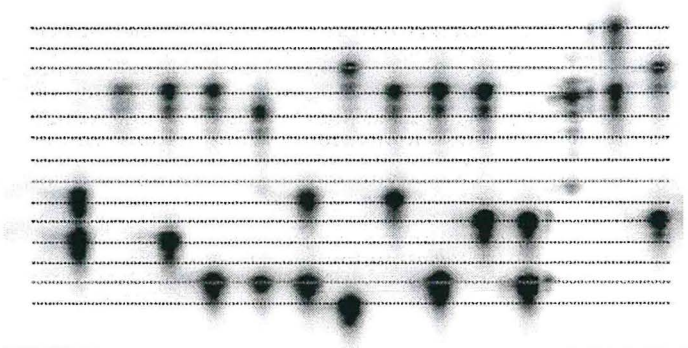
213	191	191	233	233	233	191	191		191		191	213	193	191
213	191	191	191	233	193	191	191		191		191	213	191	191
193												193		
193												193		

CnCir147



223	231	225	223	227	229	229	229	211	229	221	229	223	221	211
223	223	211	223	211	223	211	221	209	211	211	223	223	211	211
211												211		
211												211		

CNZ40



138	148	148	148	146	148	150	148	148	148	136	152	154	150
134	148	134	130	130	138	128	138	130	136	130	152	148	136
											152		
											152		

Appendix C: Technical manual

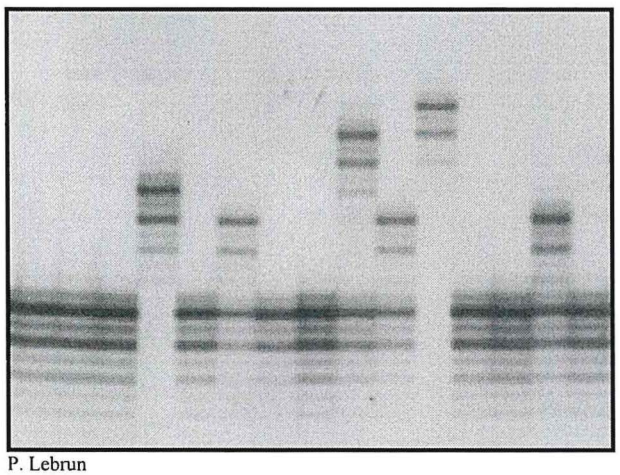
R. Bourdeix

Coconut microsatellite kit

Training Session
April 15- 24, 2002
Cirad
Montpellier, France

Updated version October 2006

A Laboratory Manual



P. Lebrun

Ipri, Cogent, Burotrop, Cirad, CEE

I. Standardized techniques

1a: Collecting and storing plant material for molecular biology

1b: DNA extraction

1c: Quantification and dilution

1d: PCR Amplification

1e: Denaturing polyacrylamide gel electrophoresis

1f: Silver staining

1g: Reading and encoding the data

II. Microsatellite markers and standards

Allele size for the standard genotypes T1 et T2

I. Standardized techniques

1a: Collecting and storing plant material for molecular biology

COLLECTION OF PLANT MATERIAL

- If possible, collect the samples early in the morning, when the weather is not too warm.
- For each sample, collect several leaflets from the basis part of leaf 1,2 or 3 in order to obtain approximately 10-20 g fresh weight. Leaflets should be **green (not yellow)**, healthy, undamaged and without fungi or algae.
- Avoid crushing or damaging the leaflets whilst collecting.
- Label the samples correctly (: cultivar name, plantation date, tree identification and country of origin, collection date, name of the collector;) This is very important!
- Roll the leaflets in a bag plastic without crushing them.
- Leaves can be kept fresh in the refrigerator for a maximum of 4 weeks provided that they are not allowed to freeze or dry. For this reason, they must be placed into a sealed plastic bag.
- As soon as possible after collecting continue with the sections below.

TO DRY THE LEAFLETS

- Cut the leaflets into pieces, about 20cm long (6 pieces per sample are enough), do not remove the midrib or edge of leaflets.
- If you have an oven, put the leaflet pieces horizontally into the oven and leave to dry for 3 days at 33°C.
- If you do not have an oven, place the leaflet pieces horizontally on absorbent paper, in a cold or refrigerated room if possible, and leave to dry for at least one week.
- DO NOT PLACE THE LEAFLETS ON TOP OF EACH OTHER WHILST DRYING.
- After drying, put leaflets into paper bags, making sure that the samples are correctly labelled.
- Labels MUST include: cultivar name, plantation date, tree identification and country of origin., collection date, name of the collector;
- Before sending, keep dried leaflets in a dry place.

TO SEND THE LEAFLETS IMMEDIATELY WITHOUT DRYING

- Bend the entire leaflet so that it will fit into a plastic bag (cold storage bags are recommended). Put the leaflet into the plastic bag and close tightly, making sure that the bag is hermetically sealed (airtight).
- Before sending, store the bags in a refrigerator. When they are fresh, green leaflets can be conserved for 15-20 days at 4°C.

DO NOT STORE DRIED LEAFLETS IN PLASTIC BAGS.

DO NOT MIX SAMPLES FROM DIFFERENT TREES.

1b DNA extraction

1. The Extraction buffer should be pre-warmed at 74°C in a water bath before use.
2. Remove the midrib.
3. Weigh 500 mg (450-550 mg) of fresh leaf sample (250 mg dried leaf sample or 200 mg lyophilized leaf).
4. Put purified sand (500-600 mg) and Polyvinylpolypyrrolidone (PVP) (130-160 mg) in a mortar.

If it's possible to use liquid Nitrogen

5. Put liquid nitrogen in the mortar and cut rapidly the weighed sample into very small pieces (approximately 1 mm).
6. Carefully add liquid nitrogen and crush the sample to a fine powder with the pestle. Crush gently at first to avoid splashing.
7. Transfer the powder (green) to a plastic test tube and place it in liquid nitrogen.
8. Step no. 5, 6 and 7 should be done quickly to prevent oxidation/browning of the sample.
9. Add 5 ml pre-warmed *Extraction buffer to the powder and mix thoroughly (preferably using a vortex shaker).
10. Incubate at 74°C for 30 minutes. Mix by inverting or shaking every 5 minutes.

If it's not possible to use liquid Nitrogen

5. Crush the sample with few buffer (1 ml) to obtain a paste. (The paste should not become brown).
6. Then add 4 ml of buffer, mix until obtaining a liquid and uniform green paste.
7. Pour the green solution into a 15 ml test tube.
8. Incubate at 74°C for 30 minutes. Mix by inverting or shaking every 5 minutes.

In both cases

11. Allow the samples to cool to room temperature.
12. Add 5 ml *Chloroform/Isoamyl alcohol (*CIAA; 24:1 v/v) to each sample and mix by inverting 50 times. CAUTION: Care should be taken when handling this chemical as it is poisonous.
13. Centrifuge at 4000 rpm for 15 minutes.
14. Using a pipette, transfer the supernatant to a new test tube.
15. Add 0.7 ml isopropanol per 1.0 ml supernatant and mix gently by inverting to precipitate the DNA.
16. Remove the precipitated DNA using a glass hook, dry it onto an absorbant paper and transfer to a microtube containing 1.0 ml TE/2* (Tris-EDTA Buffer diluted 2 times in water).

Reagents required for DNA extraction

Extraction buffer

To make 100 ml Extraction buffer, weigh the following chemicals, dissolve each separately in a small amount of distilled water, then mix them together and make the final volume to 100 ml by adding distilled water.

	Final concentration	For 100 ml
Tris	100 mM (pH 8 with HCl)	1.21 g
NaCl	1.4 M	8.1 g
EDTA	20 mM	0.8 g
MATAB	2%	2.0 g
PEG 6000	1%	1.0 g
Sodium sulphite	0.5%	0.5 g
Distilled H_2O	to	100 ml

If using concentrate solutions

	For 500ml solution	For 1 litre solution
Tris 1M	50 ml	100 ml
NaCl 5M	140 ml	280 ml
EDTA 500mM	20 ml	40 ml
MATAB	10 g	20 g
PEG	5 g	10 g
Na Sulfite	2.5 g	5 g
Distilled H_2O	to 500ml	to 1000 ml

Chloroform-isoamyl alcohol (CIAA; 24:1 v/v)

To 24 ml chloroform, add 1 ml isoamylalcohol and mix thoroughly.

TE/2 buffer

	Final concentration	For 10 ml
Tris	5 mM	6 mg
EDTA	0.5 mM	2 mg
	pH 8	

1c: Quantification and dilution

The quality and concentration of the DNA is determined by agarose gel electrophoresis. For this, 1 µl of each DNA sample is mixed with 8 µl distilled water and 1 µl 10x *DNA loading buffer and then loaded into the wells of a 1% agarose gel (1 g agarose per 100 ml 1x *TBE buffer containing 0.5 µg/ml ethidium bromide). **CAUTION:** ethidium bromide is highly toxic!

Known standard concentrations of DNA are also loaded onto the agarose gel for comparisons to be made. There are many different standards that can be used, for example, the 1 kb ladder (ref Invitrogene n° 15615-024) or λ uncut DNA (Sigma; D-3779), loaded at several different concentrations (1, 0.5, 0.25 and 0.1 µg). The agarose gel is then run in 1x TBE buffer at 90 V for 2 h or 30 V overnight and the DNA is observed on a UV transilluminator and images recorded.

The DNA samples are compared with the known standards and then a proportion is diluted to 5ng/µl with TE/2 buffer. This is the working solution for the PCR amplification. DNA stock solutions and working solutions are stored at -20°C.

Reagents required for agarose gel electrophoresis

10x DNA loading buffer

	Final concentration		For 250 ml
Tris (HCl to pH 8)	100 mM		3 g
EDTA	100 mM		10 g
Xylene cyanol	0.25%		0.625 g
Bromophenol blue	0.25%		0.625 g
Saccharose	50 %		125 g
Distilled water		to	250 ml

10x TBE buffer

	Final concentration	For 1 litre
Tris	0.89 M	108 g
Boric ac	0.89 M	55 g
EDTA	0.025 M	93 g
Distilled water	to	1000 ml

For 1x TBE buffer, the 10x TBE buffer is diluted 10-fold with distilled water.

1d: PCR Amplification

Primer stock solutions

- Primers can be obtained, purified by desalting, from many commercial companies for use in the microsatellite assay.
- Each primer is diluted to a stock concentration of 100 µM (using distilled water). Stock solutions are stored at –18°C.
- Stock solutions are further diluted to a working concentration of 1 µM (using distilled water). Working solutions are stored at 4°C.
- For PCR amplification, the working solutions for the primers and the DNA are 1 µM and 5 ng/µl, respectively.

PCR assay

Aliquot DNA

- Use 96 well PCR plates. Pipette 5 µl of DNA working solution (5 ng/µl = 25 ng) for each sample into the bottom of each well. As a negative control, set up one well more where NO DNA is added and instead put 5 µl distilled water into the well. ALWAYS include a negative control, no products should be amplified in this control. Seal plates tightly and place on ice whilst making the following master mix.

Standard PCR Master Mix

- The amounts shown below are for one DNA sample only but these amounts can be multiplied by the no. of samples to be tested + 2 (to be sure the volume is enough). This is done so that all the reagents added in the PCR amplification are consistent across all the samples tested.

	Per 25 µl PCR reaction	Final concentration
10x PCR Buffer	2.50 µl	1x
dXTP (2.5mM)	2.00 µl	200µM
Forward primer (1 µM)	5.00 µl	0.2 µM
Reverse primer (1 µM)	5.00 µl	0.2 µM
$MgCl_2$ (50mM)	0.25 µl	0.5+1.5=2 mM
***Taq* polymerase (2 U/µl)**	0.50 µl	1 U
H_2O	4.75 µl	
Total volume of PCR mix:	20.0 µl	

- Add the reagents together, as listed above, in a 5 ml test tube, mix thoroughly and centrifuge briefly (full speed for 10 secs).
- Add 20 µl of PCR master mix to each well containing DNA and to the negative control and seal plates tightly.

NOTE: Once the *Taq* polymerase has been added to the master mix, it is important to work quickly so that the PCR amplification can be started as soon as possible.

1d: PCR Amplification (cont'd)

- PCR plates are immediately loaded into the PCR thermal cycler (for example model PTC 100 by MJ Research, Inc)
- All primers used in this kit will anneal at 51°C using the following PCR programme:

1 cycle:	94°C for 5 minutes
35 cycles:	94°C for 30 seconds 51°C for 1 minute 72°C for 1 minute
1 cycle:	72°C for 30 minutes

1e: Denaturing polyacrylamide gel electrophoresis

- PCR products (microsatellites) are separated on 5% denaturing polyacrylamide gels using 1x TBE buffer.

- The gel mix is made as follows:

		For 4000 ml gel solution
Urea (Merck 1.08.488.1000)		1812 g
10 x TBE		200 ml
40% (19:1 ratio) acrylamide (Eurobio 018808)		500 ml
Distilled water	to	4000 ml final volume

- Only adjust the final volume when the urea has completely dissolved
- When ready to pour gel, add 60 ml gel solution, 250 µl (0.1 g ml^{-1}) ammonium persulfate + 80 µl TEMED and mix briefly. Be quick to pour gel.
- Allow the gel to set for at least 1 hour.
- Pre-warm gel for 20-30 minutes at 55V
- Add an equal volume of 2X *gel loading buffer to each PCR well from the amplification, denature the PCR products at 94°C for 3 minutes before loading gel.
- Using a 62-well comb makes it possible to use a multichannel pipette for loading the samples.
- For each gel (for each microsatellite marker used with the same set of samples), the loading pattern must be identical. For a 62-well gel, the recommended scheme is:

  Marker-Size Ladder, sample1, sample2...sample10, T1, sample11, ...sample 20, T2, sample 21,...sample 30, T1, sample 31,...sample 40, T2, sample 41,...sample 50, T1, sample 51...sample 54.

  Where T1 and T2 are the two standard DNA (T1= WAT4 tree M63 2407 and T2 = MYD tree 092 0715, both from M. Delorme research station in Côte d'Ivoire)

- The running conditions are: 120 minutes at 55V per gel (110V for two gels).

1f: Silver staining

The separated PCR products are visualised using the Promega SILVER SEQUENCE™ DNA Sequencing System (Cat no.: Q4130) as follows:

1. Agitate the gel in all the solutions using a flat-bed stirrer.
2. Fix gel in 2 L 10% acetic acid for 30 min. Retain the acid and use to stop the staining reaction later on in step 7.
3. Wash in 2 L distilled water for 2 min., three times.
4. Place in 2 L Silver Nitrate (2 g) with 3 ml 37% formaldehyde for 30 min. (add the formaldehyde only just before it is required).
5. Rinse briefly in distilled water - 10 secs between entering the water and entering the next solution.
6. Place in 2 L cold Sodium Carbonate (60 g) with 3 ml 37% formaldehyde and 400 µl Sodium Thiosulphate (add sodium thiosulphate only just before it is required) and leave until no further bands appear at the top of the gel.
7. Stop the reaction with 1 L 10% acetic acid (saved from step 2).
8. Once the fizzing has stop, rinse in 2 L distilled water for a minimum of 10 min.
9. Allow gel to dry on bench overnight before scanning.

Quality of water is very important for the silver staining step
Do not make up solutions containing the formaldehyde or sodium thiosulphate until just before they are required

1g: Reading gels and scoring the data

Reading gels

- Check the orientation of the gel. Marker/size ladder MUST be on the left.
- Locate the 5 DNA standard controls (3 times T1 and 2 times T2).
- Determine the allele (bands/products) sizes in these controls. These are known genotypes
- Compare the allele sizes in each sample with those in the standard controls.
- Have at least two different people reading the gel independently, to ensure accuracy.

Scoring and analysing the data with Geneclass2.0

The geneclass2.0 software is available on internet: www.montpellier.inra.fr/URLB. The first page is in French but the English version is down. You can find help online.

Geneclass 2.0 is able to understand different data formats (Genetix, Fstat, Genepop) and has also its own formats. To enter new data, the following one is convenient (the data have been truncated to the first 5 loci:

List of loci

```
                E12         A9          B12         C3b         A3
```

Beginning of the « data » section

```
DATA
```

A bracket precedes the population name,

```
{MXPT1
```

the identifier of an individual is followed by the list of its alleles (2 per locus). alleles are coded with 3 characters

```
        362     166   166   093   103   135   135   218   218   240   232
        363     164   166   089   103   135   135   216   218   234   234
        364     164   164   097   093   135   135   220   220   234   234
        365     166   166   093   103   179   179   220   218   234   240
        366     166   166   089   089   135   173   220   220   234   234
```

End of first population (reverse bracket)

```
}
```

next population

```
{MXPT2  367     164   164   089   103   159   159   220   218   234   240
        368     164   166   089   103   135   177   220   225   240   240
        369     164   164   097   103   135   153   220   225   240   240
        370     164   164   103   103   135   153   212   176   240   240
        371     164   164   097   089   135   135   225   225   240   240
}
```

A comment line begins with an asterisk « * »

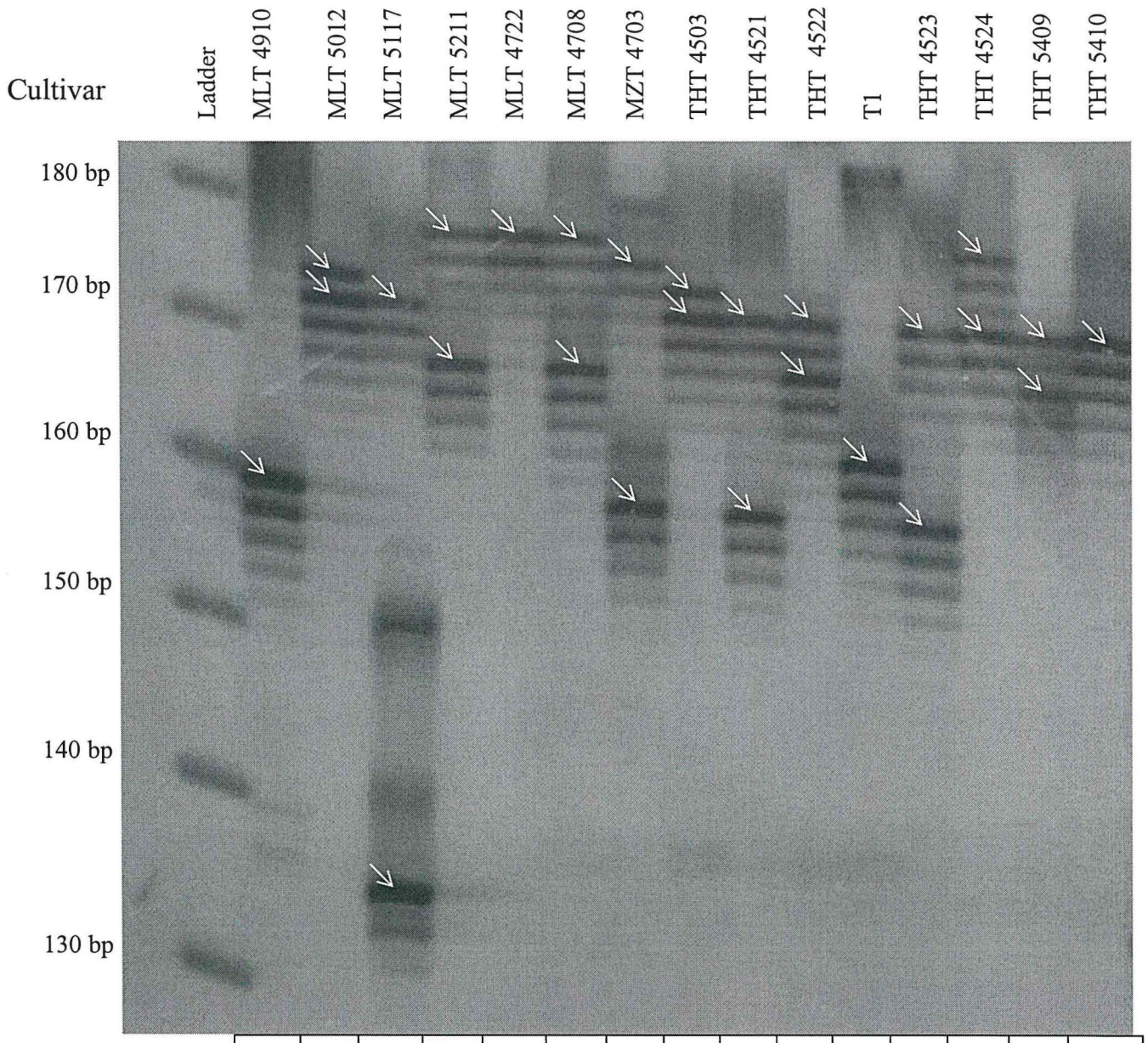
```
*Peru Tall
{PET    357     166   166   103   103   165   165   220   220   234   234
        358     166   166   103   103   173   165   220   220   234   234
        359     164   166   103   103   173   173   220   220   234   234
```

An individual with missing data (NIL)

```
        360     NIL   NIL   103   103   173   173   220   220   NIL   NIL
}
```

End of the last population

Example of autoradiogramme



Genotype	159 159	173 175	135 173	169 179	179 179	169 179	159 177	173 175	159 173	169 173	163 163	159 173	173 179	169 173	173 173

II. Microsatellite markers and standards

Allele size for the standard genotypes T1 et T2

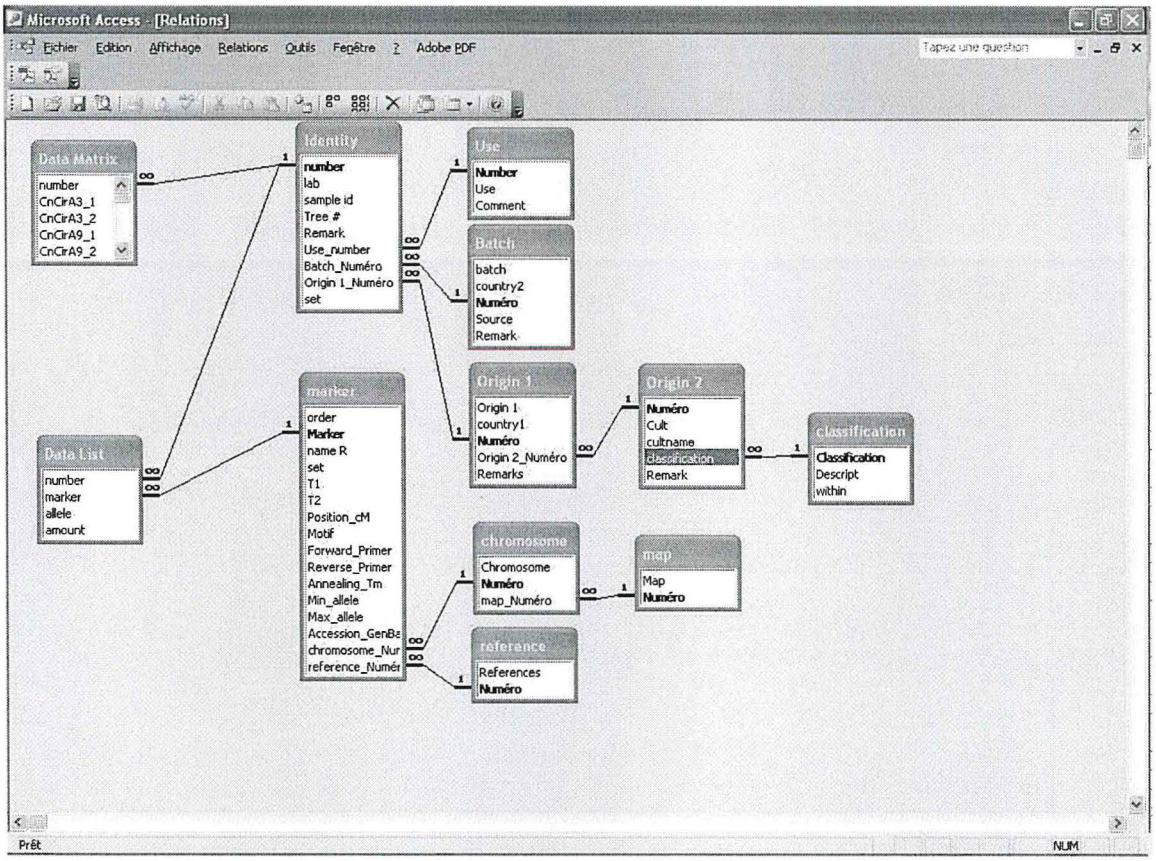
Name of the SSR	Genotype T1 (WAT4)	Genotype T2 (MYD)
CnCir A3	228 228	240 240
CnCir A9	097 097	089 089
CnCir B6	196 204	202 202
CnCir B12	163 163	169 169
CnCir C3'	176 206	174 174
CnCir C7	165 167	161 161
CnCir C12	167 167	183 183
CnCir E2	163 163	135 135
CnCir E10	244 244	238 238
CnCir E12	174 174	164 164
CnCir F2	193 193	205 205
CnCir G11	204 208	194 194
CnCir H4'	230 230	230 230
CnCir H7	133 133	139 139

It is possible to multiplexe CnCir A9 and CnCir C7 for running on the gel.
It is possible to multiplexe CnCir E2 and CnCir H4' for running on the gel

Addendum to the technical manual : protocole analyses for fluorescence marking for use with Li-cor

For each SSR marker, one of the primers was designed with a 5'end M13 extension (CACGACGTTGTAAAACGAC) (Steffens et al, 1993). PCR reactions were carried out in 384 wells plates in a 10µl final volume of reaction mixture containing: 25 ng of template DNA, 1X buffer (10mM Tris-HCL pH 8, 50mM KCl and 2 mM $MgCl_2$), 0.2 mM dNTP mix, 0.1U/µl Taq polymerase, 0.08 µM of the M13 labelled primer, 0.1µM of the other primer and 0.1 µM of M13 primer-fluorescent dye IR700 or IR800 (Biolego, The Netherlands). The PCR program started with an initial step of denaturalization at 94°C for 5 min, then 35 cycles of 94°C for 30 s, 51°C for 1min, 72°C 1 min, 72°C for 30 min and stopped after the final elongation at 72°C for 5 min. Each mixture of PCR products contained one IR700 and IR800 labelled M13 reverse complement extensions, diluted to one-fourth with formamid blue; 1.1 µl of the PCR products were separated on 6.5 % on denaturing polyacrylamide gels in a 1X TBE buffer and detected by the IR fluorescence scanning system of sequencer (Li-cor 4300).

Appendix D : Structure of the data base



description of the tables

Table	Comments	Refers to
Identity	Identifies all trees sampled. Each individual is identified by the laboratory where it was analyzed and by a unique number within this laboratory. If available, identification of the tree in the field (or a waypoint) can be stored. Further information is given in the tables represented to the right.	Use Batch Origin1
Use	Describes the possible uses of the molecular data. In principle, only the individuals whose "use" field are set to "diversity" are to be used for diversity studies	
Batch	Identifies set of samples that were sent together to the laboratory. Usually (but not always) the batch descriptors mentions what the sample was supposed to be **before** the analysis.	
Origin 1	Gives a description of the genetic origin of the samples, as confirmed by the analysis. Usually, this corresponds to the "population" level..	Origin 2
Origin 2	Same as origin 1 but corresponds to broader genetic units. Usually (but not always, this corresponds to a cultivar. For diversity studies, the level indicated in the "origin 2" level should be considered as the relevant operational taxonomic unit (OTU).	
Data matrix	A tabular representation of the molecular data. Each marker is allocated two columns. This format is convenient for data inspection and may be used to transfer data to various genetic software.	Identity
Data	A four column representation of the molecular data. This format is convenient to make elementary calculations within the data base.	Identity Marker
Marker	A description of alleles	Chromosome References
Chromosome	List of linkage groups were markers were detected	Map
Map	List of linkage maps (only one for the moment	
References	Bibliographic references	

Appendix E: Summary of the contents of the data base

This appendix lists operational taxonomic units (OTU) represented in the database with the number of individuals studied with either marker set. With a few exceptions, the OTUs correspond to cultivars. There are classified using the three-level classification based on molecular and geographic data.

Group A: Pacific coconuts (Talls and Dwarfs)

Comprises Tall and Dwarf cultivars from South-East Asia, Melanesia, Micronesia and Polynesia.

A1: South East Asian Dwarfs

Table 1: Malayan type Dwarfs (A1a)

classification	Cult	Cultname	GCP	KIT	Total
A1a	CGD	Chowghat Green Dwarf		4	4
	GGBD	Gangabondam Green Dwarf		4	4
	MGD	Malayan Green Dwarf	5		5
	MIGD	Marshall Islands Green Dwarf	3	11	14
	MRD	Malayan Red Dwarf	2	9	11
	MYD	Malayan Yellow Dwarf	2	66	68
	PGD	Sri Lanka Green Dwarf	1	21	22
	RBD	Raja Brown Dwarf	3	9	12
	SKD	Salak Green Dwarf		5	5
Total A1a		*9 OTUs*	*16*	*129*	*145*

Table 2. Filipino type Dwarfs (A1b)

classification	Cult	cultname	GCP	KIT	Total
A1b	AROD	Aromatic Green Dwarf	6	2	8
	BGD	Brazilian Green Dwarf		19	19
	CATD	Catigan Green Dwarf	2	4	6
	MGPD	Mangipod Green Dwarf		5	5
	PILD	Pilipog Green Dwarf	1	3	4
	TACD	Tacunan Green Dwarf	5		5
	THD	Thailand Green Dwarf	10		10
	XGD	Xiem Green Dwarf		2	2
Total A1b		*8 OTUs*	*24*	*35*	*59*
Total A1		**17 OTUs**	**40**	**164**	**204**

A2 : South Pacific Dwarfs

Note that this section includes only self-pollinating Dwarfs. The Niu Leka Dwarf is in group A6.

Table 3: South Pacific Dwarfs (A2)

classification	Cult	Cultname	GCP	KIT	Total
A2	COD	Chowghat Orange Dwarf		5	5
	CRD	Cameroon Red Dwarf	2	3	5
	GPTr	Gazelle Red		4	4
	NICD	Nicaragua Green Dwarf		4	4
	PRD	Pemba Red Dwarf		6	6
	PYD	Papua New Guinea Yellow Dwarf		2	2
	RARD	Rabaul Red Dwarf		2	2
	RTB	King Coconut	2		2
	TRD	Tahiti Red Dwarf	1	4	5
	TRT	Talasea Red Tall		2	2
	VRD	Vanuatu Red Dwarf	1	4	5
Total A2		**11 OTUs**	**6**	**36**	**42**

A3: South-East Asian Talls

Table 4: Continental type Talls (A3a)

classification	Cult	cultname	GCP	KIT	Total
A3a	DAUT	Dau Tall	6		6
	HAIT	Hainan Tall	7		7
	KAT	Cambodia Tall	10		10
	LPT	Lubuk Pakam Tall	5		5
	MLT	Malayan Tall	10		10
	SAT	Sawarna Tall	5		5
	SKT	Sarawak Tall	3		3
	TAAT	Ta Tall	7		7
	THT	Thailand Tall	18		18
Total A3a		*9 OTUs*	*71*		*71*

Table 5: Indonesian type Talls (A3b)

classification	Cult	Cultname	GCP	KIT	Total
A3b	HK	Kopyor coconut	5		5
	MYT	Mamuya Tall		5	5
	PUT	Palu Tall	5		5
	TGT	Tenga Tall	5		5
	TKT	Takome Tall	5		5
Total A3b		*5 OTUs*	*20*	*5*	*25*

Table 6: Filipino type Talls (A3c)

classification	Cult	Cultname	GCP	KIT	Total
A3c	BALT	Ballesteros Tall	7		7
	BAYT	Baybay Tall	8		8
	MACT	Macapuno Tall	5		5
	MXPT	Mexican Pacific Tall	39		39
	PADT	Pandan Tall	6		6
	SNRT	San Ramon Tall	6		6
	TAGT	Tagnanan Tall	14		14
Total A3c		*7 OTUs*	*85*		*85*
Total A3		**21 OTUs**	**176**	**5**	**181**

A4: Melanesian Talls

Table 7: North New Guinean type (A4a)

classification	Cult	Cultname	GCP	KIT	Total
A4a	ELT	East Sepik Tall	8		8
	KKT	Karkar Tall	22	2	24
	MDT_R	Madang Tall Red	4	1	5
	MDT_Y	Madang Tall Yellow	4		4
Total A4a		*4 OTUs*	*38*	*3*	*41*

Table 8: South New Guinean Type (A4b)

classification	Cult	Cultname	GCP	KIT	Total
A4b	BBRT	Baibara Tall	5		5
	HLT	Hisihu Tall	5		5
	KWT	Kiwai Tall	4		4
	MBT	Milne Bay Tall	3	1	4
	OLT	Oro Tall	4		4
	PLT	Poligolo Tall	3		3
	VLT	Vailala Tall	10		10
Total A4b		*7 OTUs*	*34*	*1*	*35*

Table 9: PNG insular Type (A4c)

classification	Cult	Cultname	GCP	KIT	Total
A4c	GPT	Gazelle Peninsula Tall	19	1	20
	GPTy	Gazelle Yellow	5		5
	MAT	Manus Tall	5		5
	NLT	Namatanai Tall	10		10
	WLT	West New Britain Tall	9	1	10
Total A4c		*5 OTUs*	*48*	*2*	*50*

Table 10: Markham Valley Tall (A4d)

classification	Cult	Cultname	GCP	KIT	Total
A4d	MVT	Markham Valley Tall	20		20
Total A4d		*1 OTU*	*20*		*20*

Table 11: *Vanuatu-Solomon type* (A4e)

classification	Cult	Cultname	GCP	KIT	Total
A4e	NCT	New Caledonia Tall	7		7
	RIT	Rennell Island Tall	12		12
	SIT	Solomon Island Tall	6		6
	VTT	Vanuatu Tall	335	100	435
Total A4e		*4 OTUs*	*360*	*100*	*460*
Total A4		**23 OTUs**	**500**	**110**	**610**

A5: Micronesian Talls

Table 12: Micronesian Talls (A5)

classification	Cult	Cultname	GCP	KIT	Total
A5	KIT	Kiribati Tall	21	4	25
	MIT	Marshall Island Tall	5		5
	TUVT	Tuvalu Tall	17	1	18
Total A5		**3 OTUs**	**43**	**5**	**48**

A6: Polynesian Talls (includes the cross-pollinating Dwarf Niu Leka)

Table 13: Polynesian Talls (A6)

classification	Cult	Cultname	GCP	KIT	Total
A6	COKT	Cook Island Tall	5	2	7
	NLAD	Niu Leka Dwarf	7		7
	NNMT	Niu Ni Magimagi Tall	4		4
	RTMT	Rotuman Tall	5		5
	TAT	Tahitian Tall	4	1	5
	TONT	Tonga Tall	5		5
Total A6		**6 OTUs**	**30**	**3**	**33**

A7: Panama Tall type

Table 14: Panama Tall type (A7)

classification	Cult	Cultname	GCP	KIT	Total
A7	PET	Peru Tall		3	3
	PNT	Panama Tall	59	42	101
	PNT01	Panama Tall Aguadulce	26	1	27
	PNT03	Panama Tall Costa Rica	18	1	19
Total A7		**4 OTUs**	**103**	**47**	**150**
Total Group A		**85 OTUs**	**898**	**366**	**1264**

Group B: Indo-Atlantic Talls

B1: Typical Indo Atlantic Talls

Table 15: Typical Indo Atlantic Talls (B1)

classification	Cult	Cultname	GCP	KIT	Total
B1	BRT	Brazilian Tall	72	97	169
	CALT	Calangute Tall	3	2	5
	CKT	Cameroon Kribi Tall	5		5
	DRT	Dominican Republic Tall		96	96
	ECT	Indian East Coast Tall	7		7
	JMT	Jamaica Tall	4	1	5
	LCT	Laccadive Ordinary Tall	3	2	5
	LMT	Laccadive Micro Tall	5		5
	MXAT	Mexican Atlantic Tall	9	2	11
	SCT	Seychelles Tall		33	33
	SKGT	Sakhi Gopal Tall	5		5
	SLT	Sri Lanka Tall	17	1	18
	WAT	West African Tall	24		24
	WCT	Indian West Coast Tall	6		6
Total B1		**14 OTUs**	**160**	**234**	**394**

B2: Introgressed Indo-Atlantic Talls

Table 16: Introgressed Indo-Atlantic Talls (B2)

classification	Cult	Cultname	GCP	KIT	Total
B2	ADOT	Andaman Ordinary Tall	5		5
	CMRT	Comoros Tall		10	10
	CMT	Comoro Moheli Tall	7		7
	EAT	East African Tall	18		18
	KPDT	Kappadam Tall	6		6
	MDGT	Madagascar Tall		35	35
	MZT	Mozambique Tall	97	14	111
	SLT12	Sri Lanka Tall Margaret	2	1	3
	SNRT01	San Ramon Tall Clovis	5		5
Total B2		**9 OTUs**	**140**	**60**	**200**
Total Group B		**23 OTUs**	**300**	**294**	**594**
Grand Total		**107 OTUs**	**1198**	**660**	**1858**

Table 17: Cultivars represented by only one individual

classification	Cult	Cultname	GCP	KIT	Total
A1a	TURD	Tuvalu Red Dwarf		1	1
A1b	BAGD	Baguer Green Dwarf		1	1
	KAPD	Kapatagan Green Dwarf	1		1
	KIND	Kinabalan Green Dwarf		1	1
	MNDD	Mindanao Dwarf		1	1
	SNID	San Isidro Green Dwarf		1	1
	SNOD	Santo Nino Dwarf		1	1
A3c	AGAT	Agta Tall		1	1
	SPIT	Spicata Tall		1	1
	ZAMT	Zamboanga Tall		1	1
A4e	NKFT	Tall type niu Kafa		1	1
Total			**1**	**10**	**11**

Appendix F: Review of the post-training study reports

Country Institution	LOA	Phase and date of last report consulted	Number of trees sampled analyzed	Comments
Results were obtained and may be included into the database				
India CPCRI	APO/03/071	Phase I 30/11/2004	152 analyzed 100 OK	5 collection accessions and 12 farmer varieties. Data were checked and found to be of good quality. At least part of these data could be included.
India CPCRI	APO/04/056	Phase II 26/07/2005	57 analyzed 57 OK	6 collection accessions. Good quality results
Mexico	APO/03/068	Phase I 13/12/2004	140 analyzed 100 OK	10 collection accessions with 9 to 11 primers, 4 farmer varieties with 6 primers Good quality data.
Sri Lanka CRI	APO/04/052	Phase II 30/10/2005	60 analyzed 60 OK	5 collection accessions. Data were checked and found correct.
Brazil Embrapa	APO/04/049	Phase II 17/02/2006	96 analyzed, 72 probably OK	14 accessions of exotic and domestic germplasm in collection represented by 72 individual were analyzed. 24 individuals from farmer's field were added. Data not included in the report and not checked, The dendrogram suggests they should be correct.
Results were obtained but data require further checking				
Côte d'Ivoire CNRA	APO/04/063	Phase II 30/10/2005	197 analyzed	Mainly trees from the collection. Some of them come from Ghana. At several loci, there are discrepancies with what is usually observed at Cirad we don't recommend to include these data unless results are checked.
Tanzania NCDP	APO/03/070	Phase II 30/10/2005	172 analyzed	Six markers were analyzed, we failed to find a correspondence between the data provided and those usually observed in similar germplasm. Data not available, Needs to be checked again.
No results available to us				
Indonesia Balitka	APO/03/072	Phase I 27/12/2004	0	no results
Jamaica CIB	APO/03/076	Phase II 01/07/2005	35	from collection. It is not clear whether results have been obtained
Philippines PCA	APO/03/073	Phase I 12/11/2004	0	At this stage, it seems that no analyzes were done.

Appendix G: Questionnaire

Introduction

Several years ago, the Cogent steering committee proposed that Cirad should serve as a repository for molecular marker data. In fact, molecular markers are an essential tool for the management and utilization of germplasm collection. Their applications in this domain include cultivar genetic structure characterization, assessing the quality of the methods used for rejuvenation and assessing genetic erosion from one generation to another.

This is why Cirad developed a microsatellite kit for coconut cultivar identification under a project sponsored by IPGRI and COGENT. This kit was released in 2002 and comprised 14 microsatellite markers, a data base of 599 individuals characterized with these markers, a manual and statistical software (*geneclass 2*). It was presented to the representatives of 9 producing countries during a training session held in 2003. Since that time, the number of individuals was increased to more than 1000 as well as the number of markers (30 widely used). A large number of population have been sampled, but often with small sample sizes.

Other institutions, including yours have also conducted molecular marker studies. At the request of the COGENT co-ordinator Maria Luz George, I am undertaking to collect information on these studies in order to make them broadly available to coconut breeders, molecular geneticists and genebank curators. This will make it possible to plan the practical use of molecular markers to coconut germplasm survey, maintenance and utilization.

With this survey, I wish to collect several types of information:

- on the molecular marker studies you conducted on coconut,
- on the molecular markers you used,
- on the molecular data matrix (genotyping data), especially on those made with the kit.
- on the papers you published based on these results,
-
-

For the moment, we are especially interested by the data obtained with the microsatellite kit, since they allow a direct comparison between the results obtained in different research centres. However, other molecular data can also be included in the repository.

I thank you in advance for devoting a few minutes to answer the following questionnaire.

Luc Baudouin

Questionnaire

Your name	Bee F. Gunn and Kenneth Olsen
Institution	Australian National University and Washington Univ. St. Louis
Email address	bee.gunn@gmail.com and kolsen@wustl.edu
Date	19/11/2008

1) The studies

Please, list below the molecular marker studies you conducted in coconut. A study is defined by its purpose, a genotype set and a marker set.

Study #	Study name or acronym	Purpose of the study[1]	Brief description	Number of populations or cultivars	Total number of individuals	Marker type(s)[2]	Bibliographic reference(s)[3]
1	Human exploration and dispersal history of the coconut	diversity	comparisons of genotypic diversity of cultivated vs putative wild progenitors of the coconut worldwide and tracing migration routes of the coconut	19	1,238	SSR	
2							
3							
4							
5							
6							

[1]E.g. diversity study, linkage mapping based on mapping population, association study, functional genomics etc.
[2]E.g. microsatellites (SSR), RFLP, RAPD, AFLP, ISTR, DArT etc.
[3]All bibliographic references are to be listed in section 5 below. Please, indicate here the reference number.

2) the markers

Please, describe the different marker sets

Marker type	Number of markers[1]	Total number of alleles[2] or of bands[3]	Bibliographic reference(s)[4]
SSR	10	24,760	

[1]Number of primers, primer pairs or restriction enzymes etc. according to the marker type
[2]For codominant markers
[3]For presence/absence markers
[4]All bibliographic references are to be listed in section 5 below. Please, indicate here the reference number.

3) Studies conducted using the microsatellite kit for characterizing coconut diversity

This marker set was developed by Cirad under an IPGRI-Cogent project.. It is made of the following 14 markers.

CnCirA3, CnCirA9, CnCirB12, CnCirB6, CnCirC12, CnCirC3', CnCirC7, CnCirE10, CnCirE12, CnCirE2, CnCirF2, CnCirG11, CnCirH4' and CnCirH7

More than 2000 individuals representing more than 100 cultivars were characterized with this marker set and a data matrix will be made available soon in a *public* database. for about 1500 of them.

- Did you study coconut genetic diversity using these markers?
  Yes/no
- If so, indicate here the study name(s) as in section 1 above
- **Human exploration and dispersal history of the coconut**
- Are you willing to join the genotyping data into the above described database?
  Yes/no
- If not, could you please explain why?

4) Your suggestions for improving synergy among coconut molecular geneticists, breeders and genebank managers
I would like to suggest a list of repositories for coconut specimens with variety names, including fruits, and silica-gel dried leaf material for DNA.

5) Scientific publications

Please list your publications on coconut molecular marker studies

Article #	Reference
1	
2	
3	
4	
5	
6	
7	
8	

Apendix

Example of descriptors for microsatellite markers

Marker name
Chromosome (linkage group)
Map
Position (cM)
Motif
Forward Primer
Reverse Primer
Annealing Temperature
Min_allele
Max_allele
Accession GenBank
References