Molecular Ecology Notes (2006) **6**, 904–906 doi: 10.1111/j.1471-8286.2006.01393.x

PRIMER NOTE

Ten informative markers developed from WRKY sequences in coconut (*Cocos nucifera*)

MARGARITA MAURO-HERRERA,* ALAN W. MEEROW,* JAMES W. BORRONE,* DAVID N. KUHN† and RAYMOND J. SCHNELL*

**USDA-ARS-SHRS, National Germplasm Repository, 13601 Old Cutler Rd., Miami, Florida 33158, USA, †Department of Biological Sciences, Florida International University, Miami, Florida 33199, USA*

Abstract

Coconut (*Cocos nucifera* L.) WRKY sequences containing single nucleotide polymorphisms (SNPs) and one microsatellite repeat were used to develop 10 informative markers. These markers were evaluated in 15 genotypes representing six coconut cultivars. SNP-containing alleles were detected by single-strand conformation polymorphism (SSCP) analysis. The number of detected alleles ranged from two to four. Five pairs of loci were in linkage disequilibrium in the test population. These markers are currently being evaluated in more individuals/cultivars to determine their value in estimating the genetic diversity of this species.

Keywords: coconut, *Cocos nucifera*, SNP, SSCP, WRKY genes

Received 2 February 2006; revision accepted 17 March 2006

The coconut palm (*Cocos nucifera* L.) grows in tropical and subtropical regions and is used for the production of copra, as food, fodder, as an ornamental and has many traditional uses (Chan & Elevitch 2005). In coconut, as in most other crops, diseases have a considerable impact on the quality or yield of the crop (Chan & Elevitch 2005). WRKY transcription factors derive their name from their well-conserved amino acid sequence WRKYGQK and have been reported as directly involved in the response of the plant to biotic and abiotic stresses, as well as in senescence, plant growth and embryo formation (Eulgem *et al*. 2000; Zhang & Wang 2005). WRKY genes are widely represented in plant genomes, with numbers as high as 70 for *Arabidopsis* (Eulgem *et al*. 2000) and up to 109 for rice (Zhang & Wang 2005). The development of molecular markers using WRKY genes as a target sequence was a successful approach in *Theobroma cacao* L. (Borrone *et al*. 2004). Following a similar approach, we have developed 10 WRKY markers in coconut.

By using a degenerate primer pair (Borrone *et al*. 2004), a total of 254 putative WRKY sequences were isolated from two cultivars: 'Green Malayan Dwarf' and 'Atlantic Tall'. These putative WRKY sequences were clustered in 21 WRKY groups/loci based on sequence comparison against GenBank using BLAST. Single nucleotide polymorphisms (SNPs) were identified in four of these WRKY groups/loci and one of them had a microsatellite sequence (Table 1). In addition, we designed primers specific for each one of the other 16 WRKY groups/loci that were monomorphic between these two genotypes and did direct sequencing on four more cultivars: 'Fiji Dwarf', 'Panama Tall', 'Red Spicata' and 'Red Malayan Dwarf'. BigDye Terminator Cycle Sequencing Ready Reaction Kit version 2.0 (Applied Biosystems) was used to sequence the polymerase chain reaction (PCR) products. Sequencing products were analysed on an ABI 3100 (Applied Biosystems). SNPs were identified in five more WRKY groups/loci for a total of 10 polymorphic WRKY loci, out of the 21 identified. Primer pairs were designed for the microsatellite and the SNP-containing regions (Table 1) using the Prime function in GCG (Accelrys). The forward primer targeting the microsatellite was fluorescently labelled with 6-FAM, while for SNPs, forward and reverse primers were labelled with 6-FAM and HEX, correspondingly, for single-strand conformation polymorphism (SSCP) analysis. The 10 primer pairs were evaluated in 15 individuals representing six cultivars: 'Atlantic Tall', 'Fiji Dwarf', 'Green Malayan Dwarf', 'Panama Tall', 'Red Malayan Dwarf' and 'Red Spicata'. Amplification reactions were carried out in a volume of 10 μL containing 2.0 μL of genomic DNA (10–25 ng/μL), 1.0 μL 10× buffer with 15 mM $MgCl_2$, 0.2 μL 10 mM dNTPs, 0.5 U of *Taq* polymerase, 0.15 μL of each forward and reverse

Correspondence: Alan W. Meerow, Fax: (305)969-6410; E-mail: ameerow@saa.ars.usda.gov

Table 1 Description of the 10 primer pairs (markers) developed from coconut (*Cocos nucifera*) WRKY sequences

Locus	GenBank Accession no.	Type of polymorphism	Repeat unit and SNP(s) position	Blast analysis: homology to known WRKY sequences
CnWRKY-01	DQ307149	SSR	$(AG)_{11}$	XP_475778.1 contains WRKY DNA-binding domain (*Oryza sativa*)
		SNP	C/T : 103	NP_197989.2 WRKY50 transcription factor (*Arabidopsis thaliana*)
CnWRKY-02	DQ307150	SNP	A/C : 899	BAB61056.1 WRKY DNA-binding protein (*Nicotiana tabacum*) XP_483175.1 WRKY DNA-binding protein (*Oryza sativa*)
CnWRKY-03	DQ307151	SNP	A/G : 100	AAT46067.1 DNA-binding potein WRKY2 (*Vitis vinifera*) ABA98690.1 Transcription factor WRKY1 (*Oryza sativa*)
CnWRKY-05	DQ307152	SNP	C/T : 391	AAQ57653.1 WRKY13 (*Theobroma cacao*)
CnWRKY-06	DQ307153	SNP	C/T : 795	AAQ57650.1 WRKY12 (*Theobroma cacao*) AAP85545.1 WRKY DNA-binding protein (*Glycine max*)
CnWRKY-10	DQ307154	Probably SNP(s)	Polymorphism detected via SSCP	AAT84156.1 Transcription factor WRKY07 (*Oryza sativa*) CAA88326.1 DNA-binding protein (*Avena fatua*)
CnWRKY-13	DQ307155	Probably SNP(s)	Polymorphism detected via SSCP	AAR99334.1 WRKY DNA-binding protein (*Brassica rapa*) NP_001031766.1 WRKY18 transcription factor (*Arabidopsis thaliana*)
CnWRKY-16	DQ307156	SNP(s)	A/G : 428 C/G : 525	DAA05094.1 WRKY transcription factor 29 (*Oryza sativa*) AAL61859.1 WRKY57; transcription factor (*Arabidopsis thaliana*)
CnWRKY-19	DQ307157	SNP	A/G : 74	NP_568995.2 WRKY51; transcription factor (*Arabidopsis thaliana*) AAU44313.1 WRKY transcription factor 67 (*Oryza sativa*)
CnWRKY-21	DQ307158	SNP(s)	C/G : 447 C/T : 486 A/G : 552	NP_176982.1 WRKY9; transcription factor (*Arabidopsis thaliana*) AAQ57638.1 WRKY5 (*Theobroma cacao*)

Locus	Primer sequence (5′–3′)	T_a (°C)	*n*	Size (bp)	No. of Alleles*	H_E	H_O	f	LD†
CnWRKY-01	F: TACGATGGAACCGAGCCCCCAA R: TGCAAACGAAATTTGAGCCTGCGA	54	14	205	3	0.489	0.214	0.571	CnWRKY-10 CnWRKY-13
CnWRKY-02	F: GGAACCCAAATCCGAGGTAAGTC R: TGAGATCATGAGATGCCCTCTCAA	50	14	227	2	0.071	0.071	0.000	
CnWRKY-03	F: CTTGTATGCTGTACCTGCT R: CTCCTGCATATTTCCATAGA	55	15	233	3	0.645	0.266	0.595	CnWRKY-06 CnWRKY-21
CnWRKY-05	F: TCTCTGTAAGCAGTAGTGCGACC R: ACGCACGTTACACTTGGGAGT	50	15	179	3	0.535	0.333	0.385	
CnWRKY-06	F: TTGATGTTGGTCTTGTTGGT R: TGTGCATTTGTAGTAACTCCT	52	15	173	2	0.331	0.133	0.605	CnWRKY-16
CnWRKY-10	F: GGAGCATGACTACCTTGGATCTT R: CCCGCAGTGGTGCATTTGTA	57	15	201	4	0.448	0.133	0.709	
CnWRKY-13‡	F: GCCCCTTCATGTCCGGTTAAGA R: CGGCACTCCTTTGAACCTAATGGTA	50	15	188	4	0.443	0.266	0.407	
CnWRKY-16	F: ATACTGCTACTCGTTCTGTT R: TCTGTACGTTTCCTCTCAC	50	15	229	2	0.239	0.133	0.450	
CnWRKY-19	F: CAGCCCAAATCCAAGGTATG R: GATGATGATGATGGGAGCAAG	58	14	171	2	0.137	0.000	1.000	
CnWRKY-21	F: GCACAAGTCTCATTGCCTATT R: TTTGGTTAGGTCCTTTGAGC	58	15	181	3	0.351	0.133	0.629	

n, number of individuals tested; H_E, expected heterozygosity; H_O, observed heterozygosity; f, fixation index estimate.
*Alleles detected via SSCP analysis, except for CnWRKY-01, which was assayed via fragment analysis (microsatellite conditions).
†In LD with the stated locus ($P < 0.05$).
‡This primer pair generated two amplification products in some genotypes, only the SSCP alleles corresponding to the fragment with the expected fragment size were scored.

primer (10 μM each) and 6.4 μL dH_2O. Amplification conditions were 40 s at 94 °C, 35 cycles of 25 s at 94 °C, 40 s at annealing temperature (T_a, Table 1), 1 min at 72 °C; and a final 5 min extension at 72 °C. The SSCP technique was performed as in Kuhn *et al*. (2005). Two temperatures (19 °C and 25 °C) were used for allele detection. SSCP data from the ABI 3100 were analysed using GENESCAN 3.7 and GENOTYPER 3.7 (Applied Biosystems). Amplification

products from the microsatellite locus were analysed on an ABI PRISM 3730 Genetic Analyser, and allele calls were made using GENEMAPPER 3.5 (Applied Biosystems).

One of the primer pairs (CnWRKY-13) amplified more than one locus (three alleles were observed in some SSCP electropherograms). Therefore, the samples were analysed under fragment analysis (microsatellite) conditions. Two amplification products (one of the expected size) were detected in some individuals. Only the alleles that corresponded to the fragment of the expected size were scored in the SSCP analysis.

Genetic data were analysed with GDA version 1.1 (Lewis & Zaykin 2001) to calculate descriptive statistics (Table 1). GENEPOP version 3.4 (Raymond & Rousset 1995) was used to estimate the probability of linkage disequilibrium (LD) between loci. Hardy–Weinberg equilibrium tests were not done as they are not useful in cultivated materials.

Each primer pair flanking SNP-containing regions identified between two and four alleles in the cultivars evaluated via SSCP analysis (Table 1). The microsatellite primer pair (CnWRKY-01) detected three alleles. Overall, the observed heterozygosity was lower than the expected heterozygosity, and fixation indexes higher than 0.5 were observed with several markers (Table 1). This level of inbreeding can be explained by the mating system of some of the cultivars evaluated (primarily selfing), added to the fact that they were chosen from a germplasm collection that very likely has undergone some inbreeding due to the limited number of individuals per cultivar. In addition, five marker pairs were in LD (Table 1), which could be caused by the level of inbreeding of the cultivars. However, it could also point to clustering or association of these markers. Although WRKY genes are generally distributed across the genome/chromosomes, clusters of WRKY genes have been reported in chromosome 1 of rice (Zhang & Wang 2005). We are currently evaluating these markers in more individuals/cultivars to determine their value in estimating the genetic diversity of this species.

References

Borrone JW, Kuhn DN, Schnell RJ (2004) Isolation, characterization, and development of WRKY genes as useful genetic markers in *Theobroma cacao*. *Theoretical and Applied Genetics*, **109**, 495–507.

Chan E, Elevitch CR (2005) *Cocos nucifera* (coconut). In: *Species Profiles for Pacific Island Agroforestry* (ed. Elevitch CR). URL: http://www.traditionaltree.org. PermanentAgricultureResources (PAR), Hôlualoa, Hawai'i.

Eulgem T, Rushton PJ, Robatzek S, Somssich IE (2000) The WRKY superfamily of plant transcription factors. *Trends in Plant Science*, **5**, 199–206.

Kuhn DN, Borrone J, Meerow AW, Motamayor JC, Brown JS, Schnell RJ (2005) Single-strand conformation polymorphism analysis of candidate genes for reliable identification of alleles by capillary array electrophoresis. *Electrophoresis*, **26**, 112–125.

Lewis PO, Zaykin D (2001) GENETIC DATA ANALYSIS: computer program for the analysis of allelic data. Version 1.1. Free program distributed by the authors over the Internet from http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php.

Raymond M, Rousset F (1995) GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *Journal of Heredity*, **86**, 248–249.

Zhang Y, Wang L (2005) The WRKY transcription factor superfamily: its origin in eukaryotes and expansion in plants. *BMC Evolutionary Biology*, **5**, 1.