

## Development and characterization of microsatellite markers in *Tectona grandis* (Linn. f)

Daniel Verhaegen, Daniel Ofri, Inza Fofana, Mireille Poitel, Alexandre Vaillant

► **To cite this version:**

Daniel Verhaegen, Daniel Ofri, Inza Fofana, Mireille Poitel, Alexandre Vaillant. Development and characterization of microsatellite markers in *Tectona grandis* (Linn. f). *Molecular Ecology Notes*, Wiley-Blackwell, 2005, 5, pp.945-947. <cirad-00845617>

**HAL Id: cirad-00845617**

**<http://hal.cirad.fr/cirad-00845617>**

Submitted on 17 Jul 2013

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

## PRIMER NOTE

# Development and characterization of microsatellite markers in *Tectona grandis* (Linn. f)

DANIEL VERHAEGEN,\* DANIEL OFORI,† INZA FOFANA,‡ MIREILLE POITEL\* and ALEXANDRE VAILLANT\*

\*CIRAD, Forestry Department, Research Unit 'Genetic Diversity and Breeding of Forest Trees', Baillarguet TA10/C, 34398

Montpellier cedex 5, France, †FORIG Forestry Research Institute of Ghana UST, PO Box 63 Kumasi, Ghana, ‡SODEFOR Forest

Development Society BP 3770 Abidjan 01, Côte d'Ivoire

## Abstract

**Teak (*Tectona grandis* Linn. f) is a hardwood species of worldwide reputation. It occurs naturally in India, Thailand, Myanmar and Laos. Fifteen highly polymorphic microsatellite loci were developed from a genomic library enriched for AG/TC repeats. Primers to amplify these loci were tested on 265 individual trees from different teak populations. An average of 14.7 alleles per locus were detected and expected heterozygosity ranged from 0.223 to 0.879. These loci constitute a powerful tool in investigating the mating system, gene flow, parentage analysis and population dynamics for use in sustainable management of teak forests and for *in situ* conservation purposes.**

*Keywords:* conservation, genetic diversity, microsatellite, *Tectona grandis*

*Received 27 May 2005; revision accepted 30 June 2005*

The natural distribution of *Tectona grandis* covers parts of India, Myanmar, Thailand and Laos. Appreciated as an extraordinarily durable building timber, teak is now coveted worldwide. Currently there is a serious deficit in timber resources and, consequently, harvesting of teak from natural forests is banned or severely restricted in all the countries of its natural range. In spite of its great interest, no nuclear microsatellite markers have yet been identified for this species. This study describes the development of 15 microsatellite markers for *Tectona grandis*.

Our genomic library was constructed using a DNA sample from an individual belonging to one Tanzanian population. Total genomic DNA was extracted from 300 mg of freeze-dried leaf sample using standard methods (Saghai-Marouf *et al.* 1984). Ten micrograms of teak purified DNA were used to construct a (GA)<sub>n</sub> repeat enriched genomic library, according to the protocol of Billote *et al.* (1999). DNA fragments containing (GA) microsatellite sequences were selected using biotin-labelled microsatellite oligoprobe and streptavidin-coated magnetic beads, following a hybridization-based capture methodology adapted from Kijas *et al.* (1994). We selected and cultured

864 bacterial clones containing plasmids (size range 500–900 bp) with inserts that gave strong hybridization signals. These clones were sequenced by Genoscope Centre National de Séquençage (<http://www.genoscope.cns.fr/>). Because of the difficulty of sequencing long fragments, only 480 clones could be exploited. A first check of the presence of simple sequence repeat (SSR) loci showed that the library created from this protocol gave a very good enrichment rate, with 417 sequenced clones (86.9%) containing a microsatellite (GA)<sub>n</sub>. A total of 143 primers pairs were defined with OLIGO 4.06 software (National Biosciences). Fifteen SSR loci were selected among these primers pairs on the basis of good polymerase chain reaction (PCR) product quality and clear polymorphism detection (Table 1). Loci with only one or two closest and strongly amplified fragments were chosen and microsatellite polymorphisms were screened on six genotypes of teak belonging to different origins.

A total of 265 trees were sampled: 71, 64, 36 and 94 individuals from India, Thailand–Laos, Indonesia and Africa, respectively. Automated infrared fluorescence DNA sequencing was used to find the allele variability according to Steffens *et al.* (1993). Microsatellite loci were amplified using PCR in 15 µL reaction volume containing: 25 ng of genomic DNA in a 1× reaction buffer (10 mM Tris-HCl,

Correspondence: Daniel Verhaegen, Fax: +33 (0)467593733; E-mail: daniel.verhaegen@cirad.fr

**Table 1** Development, screening and polymorphism of 15 nuclear microsatellites for *Tectona grandis* Accession no. in the EMBL Nucleotide Sequence Database

Locus name	Accession no. EMBL Database	Primer sequence	Repeat motif	Allele size range (bp)	$T_a$ (°C)	$N$	$A$	$H_O$	$H_E$	$P$ value
CIRAD1TeakA06	AJ968929	F: 5'-CAAAACAAAACCAATAGCCAGAC-3' R: 5'-TTTCATCATCATCATCAACATCC-3'	(GA) <sub>15</sub>	191–229	51	265	20	0.774	0.857	0.3166 (NS)
CIRAD1TeakB03	AJ968930	F: 5'-AACACCCCTCCTCTCTTCACTA-3' R: 5'-CACTACCCTCATCATCAACACA-3'	(TC) <sub>5</sub> TG(TC) <sub>8</sub> (AC) <sub>5</sub> (N) <sub>65</sub> (AC) <sub>14</sub>	236–266	51	265	15	0.709	0.814	0.0029 (*)
CIRAD1TeakF05	AJ968931	F: 5'-CTTCTGCAACCCTTTTTCAC-3' R: 5'-AGCCATATCTTCCTTTTCTCT-3'	(GA) <sub>20</sub> GT(GA) <sub>3</sub>	249–279	51	265	15	0.649	0.753	0.0000 (**)
CIRAD1TeakG02	AJ968932	F: 5'-TTAACGCCAAATCCCAAAG-3' R: 5'-CACAAAGAGAACCAGCAG-3'	(TC) <sub>10</sub>	166–170	51	265	3	0.185	0.223	0.3611 (NS)
CIRAD1TeakH10	AJ968933	F: 5'-CGATACCTGCGATGCGAAGC-3' R: 5'-CGTTGAATACCCGATGGAGA-3'	(TC) <sub>16</sub>	225–273	51	261	20	0.843	0.879	0.0743 (NS)
CIRAD2TeakB07	AJ968934	F: 5'-GGTGCTGATGATTTTGAGTT-3' R: 5'-CTAAGGAGTGAGTGGAGTTT-3'	(TC) <sub>14</sub>	129–157	51	265	13	0.694	0.778	0.0036 (*)
CIRAD2TeakC03	AJ968935	F: 5'-AGTGGGATGTGGTTAGAAGC-3' R: 5'-AAATGGTCATCAGTGTGAGAA-3'	(GA) <sub>17</sub>	269–313	51	265	17	0.679	0.852	0.0000 (**)
CIRAD3TeakA11	AJ968936	F: 5'-AAACCATGACAGAAACGAATC-3' R: 5'-TTGGGAATGGGAGGAGAAGT-3'	(GA) <sub>16</sub>	263–291	51	265	15	0.600	0.767	0.1744 (NS)
CIRAD3TeakB02	AJ968937	F: 5'-ATGAAGACAAGCTGGTAGCC-3' R: 5'-GGAAGACTGGGAATAACACG-3'	(TC) <sub>11</sub> GC(TC) <sub>4</sub> (N) <sub>62</sub> (AC) <sub>7</sub>	216–252	51	264	18	0.697	0.807	0.0210 (NS)
CIRAD3TeakDa09	AJ968938	F: 5'-CTCGTTCTTTCCACATT-3' R: 5'-ATCATCGCGCATCGTCAA-3'	(AC) <sub>10</sub>	198–222	51	265	9	0.325	0.408	0.0001 (**)
CIRAD3TeakE06	AJ968939	F: 5'-GCGTCAACCACCTCAACCACCAG-3' R: 5'-CCTATTTTCTTCCCTCCCTTCT-3'	(GA) <sub>10</sub> CA(GA) <sub>2</sub>	204–228	51	264	11	0.356	0.547	0.0305 (NS)
CIRAD3TeakF01	AJ968940	F: 5'-GCTCTCCACCAACCTAAACAA-3' R: 5'-AAAACGTCTCACCTTCTCACT-3'	(TC) <sub>16</sub>	198–234	51	265	18	0.774	0.825	0.1416 (NS)
CIRAD4TeakDa12	AJ968941	F: 5'-CGCACACCAGTAGCAGTAGCC-3' R: 5'-GCCGAAAAAAGAAAAACAAA-3'	(GA) <sub>4</sub> (N) <sub>5</sub> (GA) <sub>11</sub> A(GA) <sub>4</sub>	129–171	51	265	20	0.528	0.737	0.0000 (**)
CIRAD4TeakF02	AJ968942	F: 5'-CCGGTAAAAAGGTGTGTCA-3' R: 5'-GAGTGGAGGTGCTAATGGA-3'	(TC) <sub>4</sub> (AC) <sub>3</sub> (N) <sub>16</sub> (TC) <sub>11</sub>	217–243	51	265	14	0.547	0.626	0.9997 (NS)
CIRAD4TeakH09	AJ968943	F: 5'-GCAAACCAACCTTACT-3' R: 5'-CCGTTAGCACTCCATT-3'	(GA) <sub>14</sub>	149–175	51	264	13	0.379	0.550	0.0112 (NS)

Annealing temperature ( $T_a$ ), number of individuals ( $N$ ), total number of alleles ( $A$ ), observed heterozygosity ( $H_O$ ), expected heterozygosity ( $H_E$ ) under HWE,  $P$  values for the HWE test, significance threshold adjusted using sequential Bonferroni correction:  $P$  (5%)  $\leq$  0.0040.

\*Significance threshold at 5%.

\*\*Significance threshold at 1%.

NS non significant.

50 mM KCl, 2 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 0.10 µM of forward primer, 0.06 µM of reverse primer, 0.10 µM of IRdye M13/700 or M13/800 and 0.13 U/µL *Taq* DNA polymerase (Invitrogen). The amplifications were carried out with a thermal cycler Stratagene RoboCycler Gradient 96 under the following conditions: denaturation at 94 °C for 4 min; 30 cycles of denaturation at 94 °C for 30 s, annealing at 51 °C for 45 s and extension at 72 °C for 45 s; and a final extension at 72 °C for 5 min. The reverse PCR primers were probed with a 19 base extension at its 5' tail end with the sequence 5'-CACGACGTTGTAACGAC-3'. This sequence is complementary to an IR-labelled universal M13 forward sequencing primer, which is included in the PCR. During PCR, the tailed primer generates a complementary sequence, which is subsequently utilized for priming in the amplification reaction thereby generating IR-labelled PCR products. The samples were electrophoresed on an IR DNA analyser (LI-COR) that can detect primer-labelled extension products at two different wavelengths (IRDye 700 nm and IRDye 800 nm). This allowed loading of a multiplex of four PCR products in one well. The individuals trees were genotyped using SAGA software (LI-COR).

Data analysis was performed with GDA (Lewis & Zaykin 2001) and MICRO-CHECKER (Van Oosterhout *et al.* 2004) softwares. The number of alleles ranged from three to 20 per locus and was clearly higher in India and Africa, with respective means of 11.5 and 10.8 alleles, than in Indonesia and Thailand–Laos, with 5.5 and 5.1 alleles, respectively. Heterozygosity ranged from 0.185 to 0.843, with most values above 0.60. In all cases, the observed heterozygosity was lower than the expected heterozygosity (Table 1). Six of the 15 loci showed significant deviation from Hardy–Weinberg equilibrium (HWE), which has been interpreted as the result of a Wahlund effect because the total population was structured in several subpopulations. The presence of null alleles was suggested for loci 3TE6, 4TD12 and for loci 1TB3 in natural populations of India and Thailand, respectively. However, the observed segregation in two half-sib families from these populations confirmed the observed homozygosity for these three homozygous loci (data not shown). Microsatellite genotypes were tested for linkage disequilibrium for all pairs of loci within each population. Two pairwise disequilibria between 3TE6/4TF2 and 1TF5/1TG2 were found at the 1% level using the

sequential Bonferroni correction, in Thailand and Indonesia/Africa populations, but for these loci this disequilibrium could largely be due to the Hardy–Weinberg disequilibrium present in the teak populations studied and to the subpopulation structure in these populations.

The 15 microsatellite loci reported here constitute a powerful tool in investigating the mating system, gene flow, parentage and population dynamics for use in sustainable management of natural teak forests and for *in* or *ex situ* conservation of teak genetic resources.

### Acknowledgements

We are grateful to Genoscope (Centre National de Séquençage) for sequencing the teak DNA clones. We would like to thank the European Commission project 'Waft' (Inco-Dev ICA4-CT-2001-10090), which supported this research in Ghana, in Ivory Coast and in the laboratory. Many thanks to the SODEFOR and FORIG which supplied the teak plant material. We are very grateful to Mireille Poitel for her assistance in creating the microsatellite bank and for help with population genotyping.

### References

- Billote N, Lagoda PJJ, Risterucci AM, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. *Fruits*, **54**, 277–288.
- Kijas JMH, Fowler JCS, Garbett CA, Thomas MR (1994) Enrichment of microsatellites from the Citrus genome using biotinylated oligonucleotide sequences bound to streptavidin-coated magnetic particles. *BioTechniques*, **16**, 657–662.
- Lewis PO, Zaykin D (2001) *GENETIC DATA ANALYSIS: Computer program for the analysis of allelic data, version 1.0 (d16c)*. Free program distributed by the authors over the Internet from <http://lewis.eeb.uconn.edu/lewishome/software.html>.
- Saghai-Marooif MA, Soliman KM, Jorgensen RA, Allard RW (1984) Ribosomal DNA spacer length polymorphisms in barley: Mendelian inheritance, chromosomal location, and population dynamics. *Proceedings of the National Academy of Sciences, USA*, **81**, 8014–8018.
- Steffens DL, Sutter SL, Roemer SC (1993) An alternate universal forward primer for improved automated DNA sequencing of M13. *BioTechniques*, **15**, 580–582.
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*, **4**, 535–538.