American Journal of Botany

American Journal of Botany: e111–e113. 2010.

AJB Primer Notes & Protocols in the Plant Sciences

Development and characterization of microsatellite loci for *Khaya senegalensis* (Meliaceae)[1]

G. J. Sexton[2,3,4], C. H. Frere[2,4], M. J. Dieters[2], I. D. Godwin[2], and P. J. Prentis[2]

[2]School of Land Crop and Food Sciences, University of Queensland, St Lucia Qld 4072 Australia

- *Premise of the study*: Microsatellite loci were developed to characterize genetic variation and population subdivision in *Khaya senegalensis* (Desr.) A. Juss. (Meliaceae).
- *Methods and Results*: Microsatellite loci were identified from genomic DNA sequences generated using the 454 GS-FLX titanium platform. Primers were designed for 67 tri- and tetranucleotide repeats, of which 20 were selected for 2 multiplexes based on amplification success and band size. Eleven of these loci showed polymorphism in two populations of *Khaya senegalensis* and amplified in individuals from across the species range.
- *Conclusions*: These new microsatellite loci will be useful for investigation of the evolutionary and conservation genetics of *Khaya senegalensis*.

Key words: conservation genetics; *Khaya senegalensis*; microsatellite; next-generation sequencing.

Khaya senegalensis (Desr.) A. Juss., also known as African mahogany or Cailcedrat, is a member of the Meliaceae family, within the subtribe Swietenioideae (Styles, 1972), which contains many of world's most highly valued timber species. This tree species is widely distributed across tropical North Africa, occurring along a climatic gradient (400–1600 mm/yr) between the Sahara Desert in the north to the Equatorial rainforests to the south, and in a continuous band that stretches 5500 km from the Atlantic Coast in the west to the White Nile Valley in the east. Throughout its natural range, *K. senegalensis* has been exploited for timber, which is used for fine furniture, construction, and fuel. Medicinal compounds, extracted from bark, leaves, and seeds, are used to treat a multitude of tropical diseases, and its foliage is used to sustain livestock during the driest periods of the year. The dynamics of such demands have already led to local extinctions (especially in the north of its range), and the species is now classified as vulnerable under the Red List of Threatened Species (IUCN, 2010). Consequently, it is necessary to quantify patterns of genetic diversity and gene flow to develop adequate management strategies for the long-term conservation of this species. Unfortunately, the majority of microsatellite loci developed for related Swietenioideae species that have amplified in *K. senegalensis* were not polymorphic. As a first step toward indentifying the level of genetic variation within and between natural populations of *K. senegalensis*, we isolated and characterized 11 novel and two previously published microsatellite loci (Lemes et al., 2002; White and Powell, 1997).

[1] Manuscript received 9 August 2010; revision accepted 1 September 2010.
[3] Author for correspondence: sexton.gary@gmail.com
[4] Co-first authors.

doi:10.3732/ajb.1000300

METHODS AND RESULTS

Twenty-two individuals of *K. senegalensis* were selected from across the full geographic range of this species and used as the source of DNA in this study. DNA was extracted from dried leaves using a modified CTAB method of (Doyle and Doyle, 1987). Individual DNA extractions were pooled, washed with 70% ethanol, and resuspended in TE buffer. A total of 5 µg of DNA was sent to the University of Florida, checked for quality using their standard procedures, and used to construct 454 GS-FLX titanium libraries following the manufacturer's protocol (454 Life Sciences, Branford, CT). This library was sequenced on a 454 GS-FLX titanium platform at the University of Florida. This sequencing run generating 97 351 reads that met the minimum quality requirements, with an average read length of 346.6 bp. In total, we were able to sequence 33.7Mb of the *K. senegalensis* genome representing 3.95% of the 1C DNA genome content (Ohri et al., 2004).

We searched for pure microsatellite repeats in our sequences, using MSAT-COMMANDER v 0.8.1 (Faircloth, 2008), which is a Python script that identifies all microsatellite markers (SSRs) within a set of sequences. We set the script to identify all possible di-, tri, and tetranucleotide repeats with a minimum of four subunits. While some researchers have employed higher cutoffs (Kantety et al., 2002), relaxing this threshold maximizes SSR discovery while at the same time still producing a high percentage of polymorphic loci. Using our search criteria we identified a total of 12 014 microsatellite repeats. These microsatellite sequences were found in 10 181 different reads. Only 15% of the 10 181 reads containing SSRs were found to have multiple microsatellite repeats with 1317, 205, 30, and 4 reads having two, three, four, and five loci respectively. Di-nucleotide repeats were the most abundant repeat type (5817 reads; 57.13%), followed by trinucleotide repeats (3446 reads; 33.85%), and tetra-nucleotides were the least abundant repeat type (918 reads; 9.02%).

We screened our newly identified tri and tetranucleotide microsatellite loci for enough flanking region to design primers and flanking region similarities among loci using the Perl script MICROFAMILY V 1.2 (Meglécz, 2007). This decreased the number of loci for which we could design primers by approximately 75%. For the remaining sequences we used an automated design process in Batch Primer with the default settings with following modifications: (1) GC content above 30%; (2) melting temperatures (tm) between 53°C and 64°C; (3) not more than 2°C in annealing temperature between primer pairs; (4) a maximum poly n of five nucleotides; and (5) a minimum amplicon size of 100 bp in length. When all primer design criteria were satisfied for a sequence a single primer pair with the highest Primer 3 assigned score was chosen. We identified 93 potentially amplifiable loci meeting our criteria, of which 67 primer pairs were ordered.

TABLE 1. Characterization of 11 microsatellite markers isolated from *Khaya senegalensis* and two microsatellite markers previously published (Ks008: Lemes et al., 2002; Ks022: White and Powell, 1997). Shown for each primer pair are the forward and reverse sequence, repeat type, size of the original fragment (bp), annealing temperature when run individually (Ta) and the GenBank accession number.

Primer	Sequences (5′–3′)	Repeat	Allele Size (bp)	Ta (°C)	GenBank Accession No.
Ks008	F:GCGCGATTGATTGACTTC* R:GCGCTTAGCATTATTCTCC	$(AG)_{19}$	249–275	56.0	(AF428115)
Ks022	F:TGGAGTAAAGTCGAGGGCTG* R:GGCTGGATATGGCACTTGTT	$(AAT)_5$ $(CAT)_3$	205–225	60.0	(AJ000410)
Ks031	F:CAAAATATTTAACTAGCCATCCA* R:CTCAAGTCCAATTCAATTCAT	$(AAAT)_5$	266–282	55.5	HQ200180
Ks040	F:ACCATGCAATGCTAACTTTAT* R:CTTTTGTGTCCAAAAATTGAC	$(AAC)_6$	242–47	54.0	HQ200181
Ks051	F:CTTAAGACATTTGACCAATCC* R:GTCGGACAATTGTGTTTTAGT	$(AAG)_6$	129–170	54.0	HQ200182
Ks056	F:CACTCAGACGCAATAAATTAAA* R:GCGATTTTTATATCTGGTTTTC	$(AAG)_6$	94–108	55.0	HQ200183
Ks062	F:ATCCAAAACGCTTCTAAAGTC* R:GGATGTTATGGTGAAAAATTG	$(AAT)_{12}$	85–117	55.5	HQ200184
Ks063	F:CAATATAAGGGACAATACTCTCA* R:CAACATAGATCCATCGTGAGT	$(AAT)_{12}$	206–268	53.0	HQ200185
Ks070	F:TGGGGTAGTTTTAAAGTTGTG* R:AAGGTTGGAAGAGTAATTTGC	$(AAT)_{10}$	118–150	54.5	HQ200186
Ks071	F:TATAAGTGAATAAACCGCGTA* R:TATAGCCTATCTTGTCGCAAA	$(AAT)_9$	125–150	53.0	HQ200187
Ks077	F:AAACGTGCTGGCTAGTTATTA* R:TGTAAGTCAACAGCAAGAAAA	$(CTT)_{11}$	90–121	54.5	HQ200188
Ks079	F:TTTCAACTCTTCAATCTTCATCT* R:GGCACTACCAATATTTTGTTTT	$(CTT)_8$	87–115	55.0	HQ200189
Ks086	F:ACAACTCTTTTACGTCCACCT* R:CATCATCTTCTTCTGTTACGG	$(GCT)_7$	111–128	55.5	HQ200190

* Denotes Tagged primers.
Ta, optimal annealing temperature.

Primer pairs were initially screened for amplification success and polymorphism using eight individuals from across the range of *K. senegalensis*. Standard polymerase chain reaction (PCR) conditions were carried out in a 10-μL reaction containing 20 ng of DNA, 1 μL of 10× PCR buffer, 0.3 μM of each primer, 0.25 mM of each DNTP, 2.5 mM of $MgCl_2$, and 0.2 U of IMMOLASE DNA polymerase (Bioline, Alexandria, NSW). Thermocycling conditions used consisted of denaturing at 94°C for 10 min; 35 cycles at 94°C for 1 min, 56°C for 30 s, and 72°C for 1 min; with a final extension of 72°C for 20 min. PCR products and hyperladder IV (Bioline) were separated on an 8% polyacrylamide gel and stained using ethidium bromide. Using these PCR conditions, 48 of the 67 primer pairs were successfully amplified, with amplicon sizes matching expected sizes. Of these loci, 22 polymorphic primer pairs were chosen for further polymorphism testing in two populations of *K. senegalensis* from Ghana (Wa: 10°00′N, 2°30′W) and Senegal (Kedougou: 12°35′N, 12°08′W), and a further eight accessions spanning the entire species' range.

Forward primers of the 22 loci were labeled with a FAM, PET, NED, or VIC florescent dye and run in two multiplexes using the Qiagen Multiplex kit (QIAGEN, Hilden, Germany) following the manufacturer's instructions with the same PCR conditions. Amplified products were combined with internal size standard LIZ 500 (–250) (Applied Biosystems, Foster City, California, USA) and separated on an ABI 3730 capillary sequencer at the Australian Genome Research Facility (Applied Biosystems). Microsatellite loci were scored using GeneMapper software version 4 (Applied Biosystems).

Thirteen polymorphic markers (Table 1) amplified consistently over all individuals and showed clearly defined banding patterns with a maximum of two alleles/locus in an individual. Preliminary population genetic analyses were performed in Genalex 6.3 (Peakall and Smouse, 2006). Expected heterozygosity

TABLE 2. Characterization of 13 primers from two populations of *Khaya senegalensis* (Ghana and Senegal) including seven samples (included in "Total sample") representing the total distribution of the species. Number of alleles and Expected Heterozygosity, (H_e) are shown for each primer pair. Sample size for each population is in brackets.

	Ghana (N = 12)		Senegal (N = 15)		Total sample (N = 35)	
Locus name	No. of Alleles	H_e	No. of Alleles	H_e	No. of Alleles	*He*
ks008	5	0.566	5	0.538	8	0.551
ks040	2	0.219	3	0.418	3	0.380
ks062	9	0.858	12	0.904	12	0.900
ks071	3	0.542	3	0.624	5	0.624
ks077	8	0.810	8	0.817	9	0.820
ks086	5	0.663	5	0.516	6	0.560
Ks051	7	0.733	10	0.771	11	0.741
ks022	5	0.472	4	0.551	6	0.549
ks031	4	0.229	4	0.553	5	0.445
Ks063	11	0.889	10	0.855	17	0.911
ks070	7	0.788	9	0.838	11	0.837
ks079	10	0.861	8	0.839	11	0.868
ks056	5	0.639	4	0.700	6	0.702

ranged from 0.219 to 0.889 and 0.418 to 0.904, in the Ghana and Senegal populations, respectively (Table 2). The observed number of alleles per locus ranged from 5 to 17 alleles in the entire collection, and from 2 to 11 and 3 to 12 in the Ghana and Senegal populations, respectively (Table 2). A total of 110 alleles were found in the total sample collection of 35 individuals. Markers Ks041 and KS031 showed a significant departure from Hardy-Weinberg equilibrium (HWE) due to heterozygote deficiency in both populations. All other markers were in HWE using sequential bonferoni correction.

CONCLUSIONS

The newly developed microsatellite loci showed high levels of polymorphism in *K. senegalensis*. These markers provide an excellent resource to quantify levels of genetic variation and patterns of population structure in this species. This information will aid in the design of conservation strategies to ensure the long-term persistence of *K. senegalensis* and provide a means for identifying elite individuals within tree breeding programs.

LITERATURE CITED

DOYLE, J. J., AND J. L. DOYLE. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochemical Bulletin* 19: 11–15.

FAIRCLOTH, B. C. 2008. Msatcommander: Detection of microsatellite repeat arrays and automated, locus-specific primer design. *Molecular Ecology Resources* 8: 92–94.

IUCN 2010. IUCN Red List of Threatened Species. Version 2010.3.

KANTETY, R. V., M. LA ROTA, D. E. MATTHEWS, AND M. E. SORRELLS. 2002. Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology* 48: 501–510.

LEMES, M. R., R. P. V. BRONDANI, AND D. GRATTAPAGILA. 2002. Multiplexed systems of microsatellite markers for genetic analysis of mahogany, *Swietenia macrophylla* king (Meliaceae), a threatened neotropical timber species. *The Journal of Heredity* 93: 287–291.

MEGLÉCZ, E. 2007. MicroFamily: A computer program for detecting flanking region similarities among different microsatellite loci. *Molecular Ecology Notes* 7: 18–20.

OHRI, D., A. BHARAGAVA, AND A. CHATTERJEE. 2006. Nuclear DNA amounts in 112 species of tropical Hardwoos—New estimates. *Plant Biology* 3: 555–561.

PEAKALL, R., AND P. E. SMOUSE. 2006. GENALEX 6: Genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6: 288–295.

STYLES, B. T. 1972. The flower biology of the Meliaceae and its bearing on tree breeding. *Silvae Genetica* 21: 175–182.

WHITE, G., AND W. POWELL. 1997. Isolation and characterization of microsatellite loci in *Swietenia humilis* (Meliaceae): An endangered tropical hardwood species. *Molecular Ecology* 6: 851–860.