BioOne COMPLETE

Development and Characterization of Microsatellite Loci for the Moroccan Endemic Endangered Species Argania spinosa (Sapotaceae)

Authors: Bahloul, Yasmina El, Dauchot, Nicolas, Machtoun, Ikrame, Gaboun, Fatima, and Cutsem, Pierre Van

PRIMER NOTE

DEVELOPMENT AND CHARACTERIZATION OF MICROSATELLITE LOCI FOR THE MOROCCAN ENDEMIC ENDANGERED SPECIES *ARGANIA SPINOSA* (SAPOTACEAE)[1]

YASMINA EL BAHLOUL[2,4,5], NICOLAS DAUCHOT[3,5], IKRAME MACHTOUN[2], FATIMA GABOUN[2], AND PIERRE VAN CUTSEM[3]

[2]Plant Breeding Unit Rabat Regional Research Center, National Institute of Agronomy Research, Institut National de la Recherche Agronomique (INRA) Maroc, B.P. 6570 Rabat-Instituts, 10101 Rabat, Morocco; and [3]Unit of Research in Plant Biology, University of Namur, rue de Bruxelles 61, B-5000 Namur, Belgium

- *Premise of the study:* Microsatellite loci were developed for the Moroccan endemic endangered species *Argania spinosa* with a combination of a typical library enrichment procedure and a 454 GS FLX Titanium–based high-throughput sequencing approach.
- *Methods and Results:* A genomic DNA library was enriched and further screened using $(GA)_{15}$, $(GTA)_8$, and $(TTC)_8$ biotin-labeled probes coupled with chemi-luminescence detection. To increase simple sequence repeat (SSR) loci number, an ultra-high-throughput sequencing-based approach was used. Evaluation of all primer pairs was performed with labeled dUTP on an ABI 3130xl sequencer. Eleven polymorphic SSR loci were selected out of 79 SSR regions and extensively characterized on 150 individuals from eight populations. Total alleles ranged from six to 19 alleles per locus while expected heterozygosity ranged from 0.618 to 0.869.
- *Conclusions:* The SSRs developed here will be used to further characterize the genetic diversity of *A. spinosa* across its distribution range, mainly in the southern part of Morocco and southwestern Algeria. They may also be transferable to other Sapotaceae species.

Key words: argan tree; *Argania spinosa*; dUTP labeling; microsatellite; Sapotaceae; simple sequence repeats (SSRs).

Argania spinosa (L.) Skeels is endemic to Morocco and is known worldwide for its oil, which is extremely rich in unsaturated fatty acids and used in the food and cosmetic industries. Argan is the only Sapotaceae species in the region. Argan forest is found primarily in the southwestern part of Morocco and in the Souss-Massa-Drâa region (Msanda et al., 2005); a few populations have also been identified in southwestern Algeria (Kaabèche et al., 2010). Currently, argan is facing critical regeneration problems resulting from overgrazing of young sprouts and trees by goats, coupled with inadequate protection; consequently, argan tree genetic diversity is severely threatened. Molecular analyses of *A. spinosa* are very recent and still quite limited,

mainly due to the lack of species-specific molecular tools. A few studies have tried to evaluate its genetic diversity, first with the analysis of isozymes (El Mousadik and Petit, 1996a) and chloroplast DNA (El Mousadik and Petit, 1996b) and later with interspecific random-amplified polymorphic DNA (RAPDs) and simple sequence repeats (SSRs) initially developed on two different Sapotaceae species (Majourhat et al., 2008). However, only six SSRs were polymorphic, and these exhibited low diversity according to what would be expected in a perennial spontaneous (i.e., natural, nondomesticated) tree. Species-specific codominant molecular markers are thus critically needed.

Here, we developed de novo, highly polymorphic and *A. spinosa*–specific SSR markers and used them to evaluate the genetic diversity of a collection of geographically diversified argan trees.

METHODS AND RESULTS

Argania spinosa genomic DNA (gDNA) was extracted from young dried leaves and ground using a MM400 mixer mill (Retsch, Düsseldorf, Germany) for 2 min at 30 Hz in 1.5-mL centrifugation tubes with two stainless steel beads (Retsch, no. Fr0120). Between 10 and 20 mg of powder was used for gDNA extraction using the NucleoSpin Plant II DNA extraction kit according to the manufacturer's instructions (Macherey-Nagel, Düren, Germany). DNA concentration was adjusted to 50 ng/μL.

Sixteen DNA samples originating from two populations geographically separated by more than 400 km were used for library construction (Appendix 1).

TABLE 1. Characterization of 11 *Argania spinosa* SSR loci, with 150 argan individuals.

Locus[a]	Primer sequences (5′–3′)[b]	Fluorescent dye	T_m (°C)	Repeat motif	Allele size range (bp)	A	A_r	H_e	H_o	χ^2	GenBank accession no.
ASMS01 (-)	F: GTTTCTTGCAGTTTGAGAATTGAAAGGACAACG #R: CCTCCCATCTAATATTTCTAGTTCCATCC	6-FAM	57	$(CT)_{11}(CA)_{17}$	142–180	18	9.085	0.869	0.887	NS	KC138526
ASMS04 (-)	F: GTTTCTTCAAAAATAAATGTGAATGAGCAC #R: TGATGGGATTATGAGAAAGAG	6-FAM	57	$(GA)_{14}$	309–351	14	5.930	0.787	0.792	NS	KC138527
ASMS19 (- -)	F: GTTTCTTATGTAGGTCGCTGGCATTT #R: AAGCCTGGGATTATGTTATCAG	HEX	57	$(GTA)_8$	194–248	10	5.440	0.740	0.545	*	KC138528
ASMS20 (- -)	F: GTTTCTTAATACTTCAATGCGAAGGTCGTG #R: ATTACTCCCAACCTCAGTCAGC	ATTO-565	53.8	$(CT)_{14}$	198–228	15	6.294	0.726	0.725	NS	KC138529
ASMS31 (- -)	F: GTTTCTTATGTGGGTAAAGCTCTAGAAGG #R: GTGCAAATCAACAGACCTC	6-FAM	57	$(GA)_{15}$	166–184	11	6.991	0.834	0.532	*	KC138530
ASMS2012-04 (+)	#F: CCAATAATAGAAACACCCGAA R: GTTTCTTACTTAGCCCACCTTCCTCTTCT	HEX	55	$(CTT)_7$	303–320	7	3.992	0.618	0.604	NS	KC527586
ASMS2012-10 (+)	#F: AGCTGAGACCATTGCAGTCATA R: GTTTCTTCGATCCAAACAAGAGAGTCATG	ATTO-565	55	$(TTC)_{10}$	236–254	6	4.278	0.624	0.679	NS	KC527587
ASMS2012-22 (+)	#F: TGGGGTCTGCTGTTTAAGTA R: GTTTCTTCCAATCTCCATACGAACCAGAT	6-FAM	55	$(TTC)_8$	117–144	9	5.541	0.721	0.688	NS	KC527588
ASMS2012-34 (++)	#F: CCCATTGTAGACTTCCGCTTAC R: GTTTCTTAACCACAGAGAGCAGCAACTTT	6-FAM	55	$(AG)_{16}$	199–251	16	7.101	0.832	0.864	NS	KC527589
ASMS2012-37 (++)	#F: CGGAAAGGAATTAGGATTTGG R: GTTTCTTCGGTTCGTCTCTTCTCCAGTAT	HEX	55	$(AG)_{15}$	180–230	16	7.210	0.793	0.771	NS	KC527590
ASMS2012-41 (++)	#F: ATTGATTGGAGGCTAGAGTTGG R: GTTTCTTACTGGCTATGGTAGTTGAGGT	ATTO-565	55	$(GA)_{17}$	151–177	14	8.739	0.859	0.814	NS	KC527591

Notes: A = number of alleles; A_r = allelic richness; H_e = expected heterozygosity; H_o = observed heterozygosity; NS = no significant difference; T_m = melting temperature.

[a] SSR loci labeled with (-) or (- -) are two sets of markers designed for multiplex detection but simplex amplification. SSR loci labeled with (+) or (++) are two sets of markers designed for multiplex amplification and detection. ASMS2012 loci were isolated using an NGS-based approach while the other loci were isolated using a typical genomic library enrichment strategy.

[b] Labeled primers are preceded by #.

* Significant difference ($P < 0.05$) between H_e and H_o according to χ^2 test.

The protocol of Glenn and Schable (2005) was slightly modified to prepare genomic libraries enriched for $(GA)_{15}$, $(GTA)_8$, and $(TTC)_8$ repeats. Biotin-labeled oligonucleotides were used for library enrichment and Southern analysis, which were performed with the Biotin Luminescent Detection Kit (no. 11 811 592 910, Roche Diagnostics GmbH, Mannheim, Germany) and a Bio-Rad Fluor-S MAX MultiImager (Bio-Rad Laboratories, Hercules, California, USA). Positive colonies containing an insert larger than 400 bp were sequenced by Macrogen (Seoul, Korea) on an ABI 3730xl genetic analyzer (Applied Biosystems, Foster City, California, USA) with BigDye Terminator version 3.1 chemistry (Applied Biosystems). SSR motifs were confirmed in 31 nonredundant sequences. Alternatively, to increase the number of SSR loci, gDNA enrichment and direct sequencing using 454 GS FLX Titanium (Roche, Basel, Switzerland) were conducted as described earlier (Micheneau et al., 2011).

A total of 79 primer pairs (31 enriched sequences and 48 sequences developed through next-generation sequencing) were designed with Primer3 (Rozen and Skaletsky, 2000) with the following design parameters: GC content 40–60 (optimum 50), melting temperature (T_m) 55–65°C (optimum 60°C), primer size 18–27 bp (optimum 22 bp), GC clamp 1. A typical pigtail (GTTTCTT-) was added to the 5′ end of the forward or reverse primers. Primers were synthesized at Integrated DNA Technologies (Leuven, Belgium). They were first tested at 55°C on control gDNA. Markers giving no or multiple amplification products were discarded. For the remaining markers, each locus was analyzed on a set of geographically distant argan trees. For this step, amplification was performed using labeled dUTP nucleotides. A 50-μL reaction mixture contained 10 μL of 5× colorless GoTaq reaction buffer (Promega Corporation, Madison, Wisconsin, USA), 10 pmol of each primer, 5 μmol dATP, 5 μmol dCTP, 5 μmol dGTP, 4 μmol dTTP, 0.4 μmol dUTP-DY681 (Dyomics, Jena, Germany), 1.5 unit GoTaq polymerase (Promega Corporation), and 250 ng of gDNA, and was brought to volume with nuclease-free water (Promega Corporation). The amplification program was 94°C for 2 min; followed by 35 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 30 s; followed by a final elongation step at 72°C for 7 min and then hold at 10°C. Amplification products were analyzed on a CEQ 8000 fragment analyzer (Beckman Coulter, Brea, California, USA), using MapMarker MM-D1 as size standard (BioVentures, Murfreesboro, Tennessee, USA) and Frag-4 separation method with a separation time of 100 min. Amplifications were performed on iCycler (Bio-Rad Laboratories). Out of the first 79 primer pairs tested, 11 fulfilled selection criteria regarding polymorphism, lack of background amplification, and ease of scoring. A table summarizing the evaluation results of all 79 primer pairs is available as Appendix S1.

Validated SSR markers were resynthesized by Eurofins MWG Operon (Ebersberg, Germany) and labeled with 6-FAM, Atto 565, or HEX fluorescent dyes. Amplifications were performed in 10-μL volumes using 0.3 unit of GoTaq DNA polymerase (no. M300, Promega Corporation), 2 μL of 5× GoTaq PCR buffer (no. M7921, Promega Corporation), 10–50 ng of gDNA, 0.1 μL of 10 mmol/L dNTP mix (no. R0192, Thermo Fisher Scientific, Waltham, Massachusetts, USA), and 2 pmol of each primer. ASMS markers (Table 1) were amplified individually and further pooled for separation and detection on an ABI 3130xl genetic analyzer (Applied Biosystems), while ASMS2012 markers were multiplexed for amplification in two triplex reactions. All amplifications were performed according to the melting temperatures reported in Table 1.

One hundred fifty argan trees originating from seven populations located in southern Moroccan argan forests and one population from the Oued Grou relict area, in northwestern Morocco, were used for in-depth testing of the newly developed SSRs (Appendix 1). Allele number (A), allelic richness (A_r), and expected (H_e) and observed (H_o) heterozygosities were calculated using FSTAT software version 2.9.3.2 (Goudet, 1995). Results revealed high allele numbers (Table 1). A total of 136 alleles were scored. A ranged from six (ASMS2012-10) to 18 (ASMS01). A_r varied from 3.992 to 9.085. H_e and H_o ranged from 0.618 to 0.869 and 0.532 to 0.887, respectively. No significant difference was observed between H_o and H_e for nine loci, indicating that populations are approaching Hardy–Weinberg equilibrium. A difference has been noticed for loci ASMS19 and ASMS31, where H_e was higher than H_o, indicating a slight lack of heterozygotes. This might be due to natural selection of loci near the two SSRs or to the presence of null alleles.

Argan tree is considered as a spontaneous natural species with no genetic change due to human selection. The observed levels of genetic diversity are thus explained by the combined effects of mutation, genetic drift, gene flow, and natural selection.

CONCLUSIONS

In our study, 150 argan trees were genotyped with 11 SSR markers. Samples originated from eight geographically separated regions with no possibility of natural gene exchange. The 11 microsatellite markers used in this study were very informative and showed a high ability to detect polymorphism within analyzed genotypes. Outcomes of this research provide the first codominant markers specific to *A. spinosa* and will help in elaborating strategies to better protect this threatened species. They will contribute more efficiently to evaluate the genetic diversity and to understand gene flow among and between argan forest areas. Moreover, developed markers may constitute valuable tools for evaluation of genetic diversity in other Sapotaceae species.

LITERATURE CITED

El Mousadik, A., and R. J. Petit. 1996a. High level of genetic differentiation for allelic richness among populations of the argan tree [*Argania spinosa* (L.) Skeels] endemic to Morocco. *Theoretical and Applied Genetics* 92: 832–839.

El Mousadik, A., and R. J. Petit. 1996b. Chloroplast DNA phylogeography of the argan tree of Morocco. *Molecular Ecology* 5: 547–555.

Glenn, T. C., and N. A. Schable. 2005. Isolating microsatellite DNA loci. *Methods in Enzymology* 395: 202–222.

Goudet, J. 1995. FSTAT: A computer program to calculate *F*-statistics, version 1.2. *Journal of Heredity* 86: 485–486.

Kaabèche, M., A. Benkheira, and B. de Foucault. 2010. L'Arganeraie d'Algérie: Structure, écologie, syntaxonomie, dynamique. *Acta Botanica Gallica* 157: 563–572.

Majourhat, K., Y. Jabbar, A. Hafidi, and P. Martínez-Gómez. 2008. Molecular characterization and genetic relationships among most common identified morphotypes of critically endangered rare Moroccan species *Argania spinosa* (Sapotaceae) using RAPD and SSR markers. *Annals of Forest Science* 65: 805.

Micheneau, C., G. Dauby, N. Bourland, J. L. Doucet, and O. J. Hardy. 2011. Development and characterization of microsatellite loci in *Pericopsis elata* (Fabaceae) using a cost-efficient approach. *American Journal of Botany* 98: e268–e270.

Msanda, F., A. El Aboudi, and J. P. Peltier. 2005. Biodiversité et biogéographie de l'arganeraie marocaine. *Cahiers Agricultures* 14: 357–364.

Rozen, S., and H. Skaletsky. 2000. Primer3 on the WWW for general users and for biologist programmers. *In* S. Misener and S. A. Krawetz [eds.], Methods in molecular biology, vol. 132: Bioinformatics methods and protocols, 365–386. Humana Press, Totowa, New Jersey, USA.

APPENDIX 1. Location information for Moroccan populations of *Argania spinosa* used in this study.[a]

Population	Site name	Region	N	Latitude (°N)	Longitude (°W)	Altitude (m)
POP01*	AGD	Agadir	23	30°24′56.3″	009°36′17.3″	32
POP02*	OG1	Rabat oued grou	29	33°55′48.2″	006°45′00.0″	429
POP03	ESSII	Essaouira	20	31°55′35.3″	009°24′28.9″	108
POP04	AG10S1	Agadir	20	30°32′14.1″	009°41′42.3″	34
POP05	AG10S5	Agadir	21	29°02′54.0″	009°58′07.6″	412
POP06	B2-02	Agadir	17	30°11′52.5″	009°18′53.5″	178
POP07	B3-09	Agadir	12	30°04′28.0″	009°09′07.4″	563
POP08	EI	Essaouira	8	31°24′58.8″	009°42′58.5″	97

Note: N = sample size.

[a] No herbarium vouchers were collected due to the sensitive status of the species.

* Population used for DNA library generation.