American Journal of Botany

American Journal of Botany: e106–e108. 2011.

AJB Primer Notes & Protocols in the Plant Sciences

Isolation of SSR markers for two African tropical tree species, *Erythrophleum suaveolens* and *E. ivorense* (Caesalpinioideae)[1]

Jérôme Duminil[2,3,7], Guillaume K. Koffi[2], Gabriel Debout[2], Federico Sebastiani[4], Giovanni G. Vendramin[4], Myriam Heuertz[2,5,6], Santiago C. González-Martínez[6], and Olivier J. Hardy[2]

[2]Service Evolution Biologique et Ecologie, CP160/12, Faculté des Sciences, Université Libre de Bruxelles, 50 Av. F. Roosevelt, 1050 Brussels, Belgium; [3]Laboratoire de Foresterie des Régions tropicales et subtropicales, Unité de Gestion des Ressources forestières et des Milieux naturels, Gembloux Agro-Bio Tech, Université de Liège, Passage des Déportés, 2 - 5030 Gembloux, Belgium; [4]Istituto di Genetica Vegetale, Consiglio Nazionale delle Ricerche, Via Madonna del Piano 10 50019 Sesto Fiorentino (Firenze), Italy; [5]Department of Biodiversity and Conservation, Real Jardín Botánico, CSIC, Plaza de Murillo, 2, E-28014 Madrid, Spain; and [6]Department of Forest Ecology and Genetics, Centro de Investigación Forestal, INIA, Carretera de La Coruña km 7.5, E-28040 Madrid, Spain

- *Premise of the study:* To characterize the level of genetic diversity and gene flow, as well as to identify unambiguously two African tropical tree species, *Erythrophleum ivorense* and *E. suaveolens*, we have developed a set of nuclear SSR (Simple Sequence Repeats) markers.
- *Methods and Results:* Nine SSRs that display polymorphism in both species were identified. The nine newly developed SSR markers can be amplified in only two multiplexed reactions. Levels of polymorphism were assessed in two populations per species, yielding two to fifteen alleles per locus in *E. ivorense* and three to sixteen alleles per locus in *E. suaveolens*.
- *Conclusions:* The SSR markers developed here are promising to study the spatial distribution of genetic diversity and the genetic delimitation of two *Erythrophleum* species from central Africa.

Key words: gene flow; genetic diversity; species identification; spatial genetic structure; tropics.

Despite the urgency of conserving highly diverse African rainforests, information about the evolution of these ecosystems is still scarce. In particular, few data are currently available to test hypotheses about the processes that shaped diversity. In this context, approaches relying on the characterization of spatial genetic diversity are very useful (Escudero et al., 2003). Although molecular markers have already proven their usefulness in the Neotropics, their application in African rainforest taxa is still rare.

The two sister tree species *Erythrophleum ivorense* A. Chev. (syn. *E. micranthum* Harms) and *Erythrophleum suaveolens* (Guill. & Perr.) Brenan (syn. *E. guineense* G. Don.) represent a model of choice to understand the influence of historical factors on the rainforest as they are geographically widespread in tropical Africa and are ecologically distinct from each other. Although they are difficult to distinguish in the field, the use of chloroplast DNA (cpDNA) markers allows species distinction (Duminil et al., 2010). cpDNA, however, has limited polymorphism and represents only one of the genealogical processes characterizing the evolution of *E. ivorense* and *E. suaveolens.* This illustrates the need to develop highly polymorphic biparental markers, such as nuclear microsatellites, to understand the evolutionary processes shaping current levels of genetic diversity in tropical forest trees.

METHODS AND RESULTS

Genomic DNA from one individual of *E. ivorense* (coordinates of the individual N5.06948 E8.86007) and one individual of *E. suaveolens* (N3.69430 E13.67498) was purified and then pooled for the following analyses. The procedure to isolate microsatellite markers is reported in Albaladejo et al. (2010). Briefly, genomic DNA was sheared with a Nebulizer Kit (Invitrogen, Carlsbad, California, USA) applying high-pressure nitrogen to obtain DNA fragments ranging between 500 and 1000 bp. Sheared DNA was repaired with the End-It DNA End-Repair Kit (Epicentre Biotechnologies, Madison, Wisconsin, USA), and then the fragments were routinely cloned using the Zero Background Cloning Kit (Invitrogen). Ninety-six clones were sequenced with the M13 universal primer and using the Big Dye Terminator v.3.1 Cycle Sequencing Kit (Applied Biosystems, Foster City, California, USA) and run on an ABI 3730 DNA Analyzer (Applied Biosystems) automatic sequencer. Sequences were checked for the occurrence of di-, tri-, or tetranucleotide repeats with the SPUTNIK software available at http://www.cbib.ubordeaux2.fr/pise/sputnik.html. Thirty-three di- or trinucleotide repeat microsatellite regions were detected, and for each region a primer pair was designed using the software OLIGO 3.3 (Rychlik and Rhoads, 1989). Primer pairs were then tested on both species (identified

[1] Manuscript received 24 November 2010; revision accepted 12 January 2011.

Funded by the Gembloux Agricultural University (FUSAGx, Belgium) via project PPR 10.000, the FNRS (grants FRFC 2.4.576.07.F and MIS F.4.519.10.F) and by the IFORA (Iles Forestières Africaines) project financed by the French ANR (Agence Nationale de la Recherche) under the ANR-BIODIV program.

[7] Author for correspondence: jduminil@ulb.ac.be

doi:10.3732/ajb.1000482

TABLE 1. Characteristics of the nine polymorphic SSR markers developed for *E. suaveolens* and *E. ivorense*. [a]number of repeats found in the clone that corresponds to the accession number.

Primer name	Primer sequence (5′–3′)	Repeat[a]	Fragment size range (bp)	Annealing temperature	GenBank Accession No.
Ery-01	F: TGAATGTGTTGACCGTTGAT	$(TA)_{16}$	212–254	53.4	FR725943
	R: ATATGCGTCAAAATGTCTACG				
Ery-03	F: CTCTGGCAAATCCACTCTTC	$(TG)_7$	132–176	54.5	FR725945
	R: CAACAACCGATCAAACACAC				
Ery-04	F: GGTGTTAGTTTCTGTTTTCC	$(TA)_9$	136–174	49.5	FR725946
	R: GTGAAACAAAGATGTGGAGT				
Ery-06	F: AAACAGTACCAAACCCGC	$(TA)_{12}$	114–144	53.4	FR725948
	R: AAATGAGACCCATAGAAAAATAG				
Ery-07	F: ATGGTCAACAGTTCATATTTG	$(ATA)_5(ATT)_8(A)_{10}$	140–200	49.9	FR725949
	R: AGACATACCAGTCAATAAGTAAAC				
Ery-14	F: TAACTCTGGCACAAATAGGA	$(TA)_9$	266–318	49.8	FR725951
	R: CTAATATGTTCTATATGTTGTCCG				
Ery-17	F: AACTTAGTGCGAGATGAAAC	$(TC)_9$	128–140	49.8	FR725953
	R: CTTCTTCACACTTTTAATTTTC				
Ery-18	F: CTTTTGTCACTTTAAGATTG	$(GT)_8$	190–228	44.8	FR725954
	R: CTAAGCGGTATGATGTATAG				
Ery-23	F: CTTTGCCCTCAGTCCTCAGA	$(GA)_7$	108–128	60	FR725956
	R: AAGCTCAATACGACATCGGG				

following Duminil et al., 2010) using 16 individuals from four populations, two of each species: 'Korup' (N5.06675-E8.85860) and 'Forêt classée de la Mondah' (N0.57659-E9.33319) for *E. ivorense*, and 'Bertoua' (N4.33877-E14.43997) and 'Lastourville' (S0.29629-E12.75119) for *E. suaveolens*. Of the 33 primer pairs, 23 generated amplification products of the expected sizes in at least one of the two species. Finally nine primer pairs displayed polymorphism in both species and were used in the analyses (namely, Ery-01, Ery-03, Ery-04, Ery-06, Ery-07, Ery-14, Ery-17, Ery-18, and Ery-23; Table 1). Eight additional primer pairs amplified in both species but were not polymorphic (Appendix S1).

The characterization of microsatellite loci requires "a multistep screening process to evaluate candidate loci for inclusion in a genetic study" (Selkoe and Toonen, 2006). This includes checking for Hardy–Weinberg Equilibrium (HWE), linkage disequilibrium between pairs of loci, and the presence of null alleles. To perform these tests, additional individuals sampled in the four selected populations were analyzed (the number of individuals per population is reported in Table 2). SSR multiplexes were set according to the amplification sizes and melting temperatures of the primers using Multiplex Manager v1.1 (Holleley and Geerts, 2009). Two different multiplex sets were developed allowing, respectively, the amplification of five (Ery-03, Ery-04, Ery-14, Ery-18,

TABLE 2. Average genetic diversity for two populations of *E. ivorense* and *E. suaveolens* using the nine newly developed polymorphic SSR markers. A = number of alleles, A_E = effective number of alleles, H_o = observed heterozygosity, H_e = expected heterozygosity, F = fixation index (* indicates F significantly higher than zero after Bonferroni corrections with an adjusted nominal level at 0.00556), *null* = expected null allele frequency, na = not available.

	A	A_E	H_o	H_E	F	*null*		A	A_E	H_o	H_E	F	*null*
Erythrophleum ivorense							*Erythrophleum suaveolens*						
Korup (N = 51)							Bertoua (N = 53)						
Ery-01	12	8.06	0.773	0.876	0.121	0.018	Ery-01	9	4.17	0.708	0.760	0.085	0.019
Ery-03	10	5.13	0.745	0.805	0.075	0.013	Ery-03	5	2.20	0.500	0.546	0.103	0.013
Ery-04	15	3.92	0.449	0.745	0.400*	0.170	Ery-04	4	1.17	0.100	0.146	0.322	0.000
Ery-06	8	5.88	0.667	0.830	0.202	0.051	Ery-06	9	6.49	0.727	0.846	0.189	0.064
Ery-07	11	5.00	0.400	0.800	0.503*	0.220	Ery-07	11	6.17	0.800	0.838	0.084	0.023
Ery-14	4	3.79	0.143	0.736	0.818*	0.319	Ery-14	9	5.49	0.565	0.818	0.299*	0.118
Ery-17	3	2.23	0.429	0.551	0.224*	0.109	Ery-17	4	1.62	0.438	0.383	–0.148	0.000
Ery-18	10	5.62	0.351	0.822	0.576*	0.257	Ery-18	4	2.29	0.500	0.563	0.133	0.000
Ery-23	6	4.22	0.745	0.763	0.024	0.000	Ery-23	5	2.56	0.500	0.609	0.194	0.065
Mean	8.78	4.87	0.52	0.77	0.33	0.13	*Mean*	6.67	3.57	0.54	0.61	0.14	0.03
(±SD)	(±3.90)	(±1.63)	(±0.22)	(±0.09)	(±0.27)	(±0.12)	*(±SD)*	(±2.78)	(±2.05)	(±0.21)	(±0.24)	(±0.14)	(±0.04)
Forêt classée de la Mondah (N = 23)							Lastourville (N = 26)						
Ery-01	13	8.20	0.643	0.878	0.276*	0.092	Ery-01	16	9.43	0.651	0.894	0.274*	0.126
Ery-03	4	2.04	0.591	0.511	–0.036	0.000	Ery-03	3	2.42	0.646	0.586	–0.103	0.000
Ery-04	6	3.25	0.500	0.692	0.270*	0.096	Ery-04	3	1.13	0.060	0.115	0.481*	0.077
Ery-06	3	3.50	0.143	0.714	0.813*	0.315	Ery-06	12	7.52	0.682	0.867	0.216*	0.088
Ery-07	5	4.52	0.389	0.779	0.537*	0.228	Ery-07	14	10.64	0.830	0.906	0.085	0.034
Ery-14	na	na	na	na	na	na	Ery-14	12	9.43	0.313	0.894	0.653*	0.304
Ery-17	2	1.85	0.500	0.460	–0.128	0.000	Ery-17	4	1.32	0.190	0.240	0.207	0.044
Ery-18	5	2.28	0.364	0.562	0.367*	0.149	Ery-18	9	3.83	0.580	0.739	0.217	0.063
Ery-23	4	2.03	0.455	0.508	0.078	0.046	Ery-23	7	2.31	0.400	0.567	0.296	0.071
Mean	5.25	3.46	0.45	0.64	0.27	0.12	*Mean*	8.89	5.34	0.48	0.65	0.26	0.09
(±SD)	(±3.37)	(±2.13)	(±0.15)	(±0.15)	(±0.31)	(±0.11)	*(±SD)*	(±4.91)	(±3.88)	(±0.26)	(±0.30)	(±0.22)	(±0.09)

and Ery-23) and four loci (Ery-01, Ery-06, Ery-07, and Ery-17) using three different dyes for the forward-labeled primers (NED, FAM, and HEX). Polymerase chain reactions (PCRs) were carried out in a Biometra TProfessional Thermocycler. PCRs were performed in a total volume of 10 μL containing 1.5 μL of template DNA (10–100 ng), 4 μL of 2× QIAGEN Multiplex PCR Master Mix, either 0.5 μL of primer mixI or 0.7 μL of primer mixII, and 4 μL (for the first primer mix) or 3.8 μL (for the second) of ddH_2O. Primer mixI was prepared from primer stock solutions at 100 μM, taking 1.1 μL of both primers for the locus Ery3 (NED), 0.9 μL for Ery4 (HEX), 1.8 μL for Ery14 (HEX), 3.3 μL for Ery18 (HEX), 0.6 μL for Ery23 (FAM), and 84.6 μL of H_2O. Primer mixII was composed of 0.7 μL of both primers for the locus Ery1 (FAM) 0.7 μL for Ery6 (FAM), 2.1 μL for Ery7 (NED), 2.7 μL for Ery17 (HEX), and 87.6 μL of H_2O. The PCR cycling profile included an initial step of 15 min at 95°C followed by 30 cycles of 30 s at 94°C, 60 s at 56°C, and 45 s at 72°C, followed by a final incubation at 60°C for 10 min. Amplified fragments were run on a 3100 Genetic Analyzer (Applied Biosystems). The lengths of the fragments were determined by comparison with the GeneScan 500 ROX Size Standard (Applied Biosystems) using the Genemapper v3.0 software (Applied Biosystems).

Expected (H_e) and observed (H_o) heterozygosities were estimated using Excel Microsatellite Toolkit 3.1.1 (Park, 2001). Genotypic disequilibrium between all pairs of loci, numbers of alleles, effective numbers of alleles, and inbreeding coefficients were calculated for each locus in each population with Fstat 2.9.3.2 (Goudet, 1995). Frequencies of null alleles were estimated in FREENA (Chapuis and Estoup, 2007) following the Expectation Maximization (EM) method by Dempster et al. (1977).

The two multiplexes successfully amplified four and five loci, respectively. As shown in Table 2, both species showed similar numbers of alleles per locus (from 2 to 15 alleles for *E. ivorense*, from 3 to 16 for *E. suaveolens*) and similar levels of expected heterozygosity (mean ± SD = 0.708 ± 0.137 for *E. ivorense*, mean ± SD = 0.629 ± 0.186 for *E. suaveolens*). All populations showed significant departures from HWE for multiple loci. *Erythrophleum ivorense* displayed higher heterozygote deficit than *E. suaveolens*, likely due to higher null allele frequencies. Indeed, *E. ivorense* exhibited higher null allele frequencies over all loci (mean ± SD = 0.123 ± 0.111) than *E. suaveolens* (mean ± SD = 0.062 ± 0.072). No genotypic disequilibrium between loci was observed after applying the Bonferroni correction for multiple test comparisons.

CONCLUSIONS

Despite a relatively high frequency of null alleles at some loci in *E. ivorense*, the SSR markers developed here are efficient to estimate genetic diversity in *Erythrophleum* species. Their use at larger spatial scales will provide detailed information about the distribution of genetic diversity in both species. Fine-scale genetic structure studies will enable us to estimate levels of historical gene flow in these species. This information is useful for building and testing hypotheses on the history of African tropical forests in response to climatic changes.

LITERATURE CITED

ALBALADEJO, R. G., F. SEBASTIANI, S. C. GONZALEZ-MARTINEZ, J. P. GONZALEZ-VARO, G. G. VENDRAMIN, AND A. APARICIO. 2010. Isolation of microsatellite markers for the common Mediterranean shrub *Myrtus communis* (Myrtaceae). *American Journal of Botany* 97: E23–E25.

CHAPUIS, M.-P., AND A. ESTOUP. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24: 621–631.

DEMPSTER, A. P., N. M. LAIRD, AND D. B. RUBIN. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society B. Statistical Methodology* 39: 1–38.

DUMINIL, J., M. HEUERTZ, J. L. DOUCET, N. BOURLAND, C. CRUAUD, F. GAVORY, C. DOUMENGE, ET AL. 2010. CpDNA-based species identification and phylogeography: application to African tropical tree species. *Molecular Ecology* 19: 5469–5483.

ESCUDERO, A., J. M. IRIONDO, AND M. E. TORRES. 2003. Spatial analysis of genetic diversity as a tool for plant conservation. *Biological Conservation* 113: 351–365.

GOUDET, J. 1995. FSTAT (Version 1.2): A computer program to calculate F-statistics. *The Journal of Heredity* 86: 485–486.

HOLLELEY, C. E., AND P. G. GEERTS. 2009. Multiplex Manager 1.0: A crossplatform computer program that plans and optimizes multiplex PCR. *BioTechniques* 46: 511–517.

PARK, S. D. E. 2001. Trypanotolerance in West African cattle and the population genetic effects of selection [Ph.D. thesis] University of Dublin.

RYCHLIK, W., AND R. E. RHOADS. 1989. A computer program for choosing optimal oligonucleotides for filter hybridization, sequencing and in vitro amplification of DNA. *Nucleic Acids Research and Exploration* 17: 8543–8551.

SELKOE, K. A., AND R. J. TOONEN. 2006. Microsatellites for ecologists: A practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9: 615–629.