American Journal of Botany

American Journal of Botany: e154–e156. 2012.

AJB Primer Notes & Protocols in the Plant Sciences

Development of microsatellite markers for *Anadenanthera colubrina* (Leguminosae), a neotropical tree species[1]

Juliana Massimino Feres[2,3], Mariza Monteiro[4], Maria I. Zucchi[5], José B. Pinheiro[4], Moacyr A. Mestriner[3], and Ana Lilia Alzate-Marin[2,3,6]

[2]Programa de Pós-graduação em Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes 3900 14049-900, Ribeirão Preto, São Paulo, Brazil; [3]Departamento de Genética, Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Av. Bandeirantes 3900 14049-900, Ribeirão Preto, São Paulo, Brazil; [4]Departamento de Genética, Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Av. Pádua Dias 11, Vila Independência 13400-970, Piracicaba, São Paulo, Brazil; and [5]Agência Paulista de Tecnologia dos Agronegócios, Rodovia SP 127, Km 30 Vila Fátima 13400-970, Piracicaba, São Paulo, Brazil

- *Premise of the study:* We developed and characterized nuclear microsatellite markers for *Anadenanthera colubrina*, a tropical tree species widely distributed in South America.
- *Methods and Results:* Leaf samples of mature *A. colubrina* trees, popularly called "angico," were collected from an area that is greatly impacted by agricultural practices in the region of Ribeirão Preto in São Paulo State in southeastern Brazil. Twenty simple sequence repeat (SSR) markers were developed, 14 of which had polymorphic loci. A total of 96 alleles were detected with an average of 6.86 alleles per polymorphic locus. The expected heterozygosity, calculated at polymorphic loci, ranged from 0.18 to 0.83. Finally, we demonstrated that 18 loci were cross-amplified in *A. peregrina*.
- *Conclusions:* A total of 14 polymorphic markers suggest a high potential for genetic diversity, gene flow, and mating system analyses in *A. colubrina*.

Key words: *Anadenanthera colubrina*; angico; forest conservation; genetic diversity; Leguminosae; simple sequence repeat markers.

Anadenanthera colubrina (Vell.) Brenan (Leguminosae, subfamily Mimosoidae), commonly known as "angico," is a native tree species that can reach up to 35 meters in height. This species is frequently found in montane and submontane semideciduous forests in Brazil, but is also found in Bolivia and Peru (Carvalho, 2002). *Anadenanthera colubrina* is an important pollen source for bees, and the bark of the tree also possesses medicinal properties (Carvalho, 2002; Monteiro et al., 2006). This species is widely used to facilitate the recovery of ecosystems and restoration of riparian forests (Carvalho, 2002).

Microsatellites, or simple sequence repeat (SSR) markers, are among the most suitable molecular markers to study the genetic structure, gene flow, and breeding of tree populations (Chase et al., 1996). Therefore, in this study, we developed and characterized for the first time a set of nuclear microsatellite markers for *A. colubrina* to investigate the genetic structure, gene flow, and mating system in this species in detail. In addition, we evaluated the transferability of these loci to *A. peregrina* (L.) Speg., which is a close relative of *A. colubrina*.

METHODS AND RESULTS

We sampled 51 mature *A. colubrina* trees in two populations (30 individuals from 21°15′38.81″S, 47°49′8.17″W; 21 from 21°9′41.30″S, 47°51′51.33″W) that are located in a large area that is affected by agricultural practices of the sugar cane industry in the region of Ribeirão Preto, São Paulo State, southeastern Brazil. We observed a predominance of this species in areas with deciduous forests that are not suitable for agriculture. All trees were mapped using GPS (GARMIN eTrex Vista Cx, Garmin International, Olathe, Kansas, USA). The leaves were collected and stored at −20°C until DNA extraction. All DNA samples were extracted using a modified cetyltrimethylammonium bromide (CTAB) extraction procedure as described in Alzate-Marin et al. (2009). The modifications included the use of PVP to eliminate the polyphenols, only one chloroform-isoamyl alcohol step, and the addition of RNase immediately after extraction with chloroform. The leaf tissue from one individual tree (UTM 23 K 0202646; 7657299) was used to prepare the microsatellite-enriched library (Billotte et al., 1999). A voucher herbarium specimen was deposited at the Universidade de São Paulo Herbarium (SPF) and recorded as *J. Feres 1* (SPF). The DNA was digested using *Rsa*I, ligated to the Rsa21 and Rsa25 adapters, and amplified using PCR. The PCR products were purified using the QIAquick PCR purification kit (QIAGEN GmbH, Hilden, Germany). Fragments containing microsatellite sequences were selected using hybridization with biotinylated oligonucleotide probes [$(CT)_8$, $(GT)_8$, and $(CTT)_8$] and recovered using streptavidin-coated paramagnetic beads. A second PCR was performed to increase the number of enriched fragments. An aliquot of the PCR product was ligated to the pGEM-T vector

[1]Manuscript received 9 September 2011; revision accepted 29 January 2012.

The authors thank FAEPA and CAPES-PROEX. J.M.F. was supported by FAPESP (grant no. 2009/14200-6). A.L.A.M. was supported by grants from CNPq (PDS 150277/2009-1, PV 300140/2011-8). M.I.Z. and J.B.P. are fellow CNPq researchers. The authors thank Marco Aurélio C. dos Reis and Leonardo Borges for collaboration regarding sample collections. J.M.F., M.A.M., and A.L.A.M. are members of Centro de Estudos e Extensão Florestal da Universidade de São Paulo (CEEFLORUSP).

[6]Author for correspondence: anaalzate@rge.fmrp.usp.br

doi:10.3732/ajb.1100446

(Promega, Madison, Wisconsin, USA) and transformed into *Escherichia coli* XL-1 Blue. A total of 174 clones were sequenced using the ABI 377 Genetic Analyzer and BigDye Terminator Kit (Applied Biosystems, Foster City, California, USA).

All 51 samples were used for primer screening. The microsatellite loci were individually amplified in 10 μL reactions, as previously described by Feres et al. (2009), in a final volume of 10 μL containing 0.3 μM of each primer, 1 U *Taq* DNA polymerase (Biotools, Madrid, Spain), 0.25 mM of each dNTP, 1× $MgCl_2$-free reaction buffer (75 mM Tris-HCl [pH 9.0], 50 mM KCl, and 20 mM $(NH_4)_2SO_4$), 1.5 mM $MgCl_2$, and 2.5 ng of template DNA. Amplifications were performed using a Mastercycler (Eppendorf AG, Hamburg, Germany) with the following conditions: 96°C for 4 min; 30 cycles at 94°C for 40 s, 50–58°C (Table 1) for 1 min, and 72°C for 1 min; and 72°C for 7 min. PCR products were separated using a 10% denaturing polyacrylamide gel and stained with silver nitrate. Allele sizes were estimated using the original DNA from the library and a 10-bp ladder standard (Invitrogen, Carlsbad, California, USA). A total of 38 clones possessing microsatellite repeats (21.8%) were identified, and only 20 of these clones were suitable for primer design. The other 18 clones that were found to be inappropriate for primer design presented microsatellites located too close to the end of an insert to accommodate primer design, presented an unsuitable flanking sequence, or presented chimeric sequences (in which one of the flanking regions matches that of another clone). All primer pairs amplified. However, six loci were monomorphic (Acol 01, Acol 03, Acol 04, Acol 06, Acol 07, and Acol 08), whereas 14 loci were polymorphic (Table 1). Among the amplified loci, 17 were dinucleotides, one was trinucleotide (Acol 02), one was tetranucleotide (Acol 04), and one had compound loci (Acol 19). The most commonly repeated motif was CA/TG, which was expected because the motif CT/GT was used for the enrichment of the library.

Statistical analyses were performed for the 14 identified polymorphic loci. Polymorphic loci were characterized based on the number of alleles per locus (A) and the expected (H_e) and observed (H_o) heterozygosities for each locus, averaged over all loci, and Wright's fixation index calculated using GDA 1.1 software (Lewis and Zaykin, 2002). The FSTAT software package version 2.9.3.2 (Goudet, 2002) was used to evaluate linkage disequilibrium for all loci, and sequential Bonferroni correction was applied for multiple comparisons (Rice, 1989). The cumulative exclusion probabilities for the first and second parents and the null allele frequency were calculated using CERVUS 3.0 (Kalinowski et al., 2007). Pairwise comparisons among loci did not indicate linkage disequilibrium after applying Bonferroni correction (95%, $\alpha = 0.05$) for multiple comparisons. These results suggest that these loci are suitable for parentage analysis. A total of 96 alleles were identified, ranging from three to 14 alleles per locus (average: 6.86) (Table 2). The genetic diversity ranged from 0.18 (Acol 05) to 0.83 (Acol 15 and Acol 16) with a mean of 0.59. The observed heterozygosities ranged from 0.14 (Acol 02 and Acol 14) to 0.83 (Acol 12) with a mean of 0.41. These results consolidate the positive and significant fixation index (Table 2), which could be due to mating system patterns and/or the Wahlund effect (i.e., population substructure). However, the loci presented evidence of a potential effect of null alleles. The cumulative exclusion probabilities for all 14 loci were high for the first ($P_{exclu(1)} = 0.988$) and second ($P_{exclu(2)} = 0.999$) parents, indicating that these sets were suitable for paternity analysis (Table 2).

TABLE 1. Analysis parameters for 20 loci of *Anadenanthera colubrina* that were analyzed in 51 individual samples.

Loci	Primer sequences (5′–3′)	Repeat motif	Size (bp)	T_a (°C)	GenBank accession no.
Acol 01	F: GGTGAGATAGCTGAAGTTCG	$(AT)_6$	181	50	JN098377
	R: GAATTTCAACTCCCCATG				
Acol 02	F: TGGCTATCACTCTCGGACTTC	$(CTT)_5$	181	57	JN098378
	R: CATAATGCAAACCGATGACC				
Acol 03	F: TTAAGACCCCGCACAAAC	$(AATT)_4$	150	50	JN098379
	R: CTATTCAAGTTCCTGCAAATTA				
Acol 04	F: GATGGTGGGCAAAGTGAG	$(AT)_3G(AT)_2T(AT)_2$	246	50	JN098380
	R: GAAGCCGCATGTATAAGAGT				
Acol 05	F: TGGACTTTTGCACTGGACAC	$(GA)_2A(GA)_7$	234	58	JN098381
	R: CCCACCCACCTAGAACTGTC				
Acol 06	F: TGGGAATGTGCATGAGAAAG	$(AT)_4AG(AT)_2$	166	58	JN098382
	R: CTCGTGAATTTTCGGGTGT				
Acol 07	F: GCTTGCCTTATTATCACACCA	$(CA)_4$	166	57	JN098383
	R: GGGGAAACTTGGTAACATGG				
Acol 08	F: ATGTTAGTGTGATGGCC	$(CA)_4$	222	53	JN098384
	R: CTGACTGAAAGGCTTAA				
Acol 09	F: CCAGGGTCCTCTCAGATTG	$(CT)_{10}(AT)_7$	160	58	JN098385
	R: ATGCTCCTCTCAACACACC				
Acol 10	F: CGTATGTAGGTGGACTTAATATGC	$(TG)_3C(GT)_{15}(GA)_{20}$	190	54	JN098386
	R: GCATGAAAGAAAGCTAATCACTC				
Acol 11	F: GGAAATGCAAGATCAGAGGCTC	$(CA)_9$	200	60	JN098387
	R: GTTTGGTTGATTGAGGATGTCCT				
Acol 12	F: CCCCACATTTAGAACAGGAG	$(TA)_9$	118	58	JN098388
	R: CCACCTTGGATGATGTTTC				
Acol 13	F: CTTATTGCCTACTCCCTAACC	$(AG)_{15}$	139	58	JN098389
	R: GCATGGGGTCTTGATCC				
Acol 14	F: GCGATACCCCCTTGATG	$(GA)_{11}$	183	54	JN098390
	R: GGCGAATGACCATCTTATC				
Acol 15	F: CATATGTCCAATCAGGTTAGAC	$(GA)_{10}$	124	58	JN098391
	R: CTCATTTGGTGACTGTAAGC				
Acol 16	F: AAGGTCCAAGGGTTATGC	$(TC)_{11}(CA)_5AA(CA)_5$	220	58	JN098392
	R: GGACGTTCTCTTGTCATGC				
Acol 17	F: ACTCTCGACAACAAGTAATTCAAC	$(AG)_9$	182	58	JN098393
	R: CAACAGCCAATGACAATGATG				
Acol 18	F: GTCAGGTATTCAATCTTCTACGT	$(TG)_{10}$	160	52	JN098394
	R: CACCATACCTCCACCTCATC				
Acol 19	F: GACTGAAACCATTCTTTTATC	$(CT)_5(CCCTCT)_4(CT)_{16}(AC)_7$	163	52	JN098395
	R: TGTTGCGTATGTGGGTG				
Acol 20	F: CAACCGCCACCATCAAAC	$(GT)_4GGG(TG)_6$	136	56	JN098396
	R: AGTATACCAGTAGATCACAAGGATG				

Note: T_a = annealing temperature.

TABLE 2. Statistical parameters from 14 polymorphic loci of *Anadenanthera colubrina*.

Loci	*A*	H_e	H_o	*F*	$P_{exclu(1)}$	$P_{exclu(2)}$	*NAF*	T_a (°C) *A. peregrina*[a]
Acol 01[b]	—	—	—	—	—	—	—	50
Acol 02	4	0.31	0.13	0.56*	0.951	0.835	0.396	57
Acol 03[b]	—	—	—	—	—	—	—	NA
Acol 04[b]	—	—	—	—	—	—	—	50
Acol 05	3	0.18	0.15	0.13*	0.983	0.911	0.061	58
Acol 06[b]	—	—	—	—	—	—	—	58
Acol 07[b]	—	—	—	—	—	—	—	57
Acol 08[b]	—	—	—	—	—	—	—	52
Acol 09	6	0.48	0.36	0.26*	0.869	0.706	0.571	58
Acol 10	9	0.78	0.21	0.72*	0.582	0.401	0.043	52
Acol 11	3	0.34	0.31	0.10	0.940	0.843	0.060	60
Acol 12	6	0.75	0.83	–0.10	0.654	0.478	0.120	58
Acol 13	6	0.78	0.59	0.24*	0.623	0.444	0.173	58
Acol 14	3	0.18	0.13	0.25*	0.983	0.907	0.052	54
Acol 15	14	0.82	0.72	0.12*	0.513	0.341	0.155	58
Acol 16	8	0.82	0.60	0.26*	0.798	0.523	0.139	57
Acol 17	8	0.67	0.50	0.24*	0.714	0.524	0.304	NA
Acol 18	7	0.67	0.36	0.46*	0.736	0.561	0.125	51
Acol 19	10	0.78	0.60	0.23*	0.588	0.409	0.587	53
Acol 20	9	0.64	0.17	0.73*	0.764	0.592	0.571	56
Average	6.86	0.59	0.41	0.31*	—	—	—	—
Total	96				0.988	0.999	—	

Note: *A* = allele number; *F* = fixation index; H_e = expected heterozygosity; H_o = observed heterozygosity; NA = not amplified; *NAF* = null alleles frequency; $P_{exclu(1)}$ = first parental exclusion probability; $P_{exclu(2)}$ = second parental exclusion probability.

[a] Annealing temperatures of 18 loci cross-amplified in *A. peregrina*.

[b] Monomorphic loci in *A. colubrina.*

* Significant deviation ($P < 0.05$).

Additionally, transferability of the 20 primer pairs was tested in 10 individuals of *A. peregrina* using the same PCR and polymorphism detection protocol described above. A voucher herbarium specimen was deposited for this species at the Universidade de São Paulo Herbarium (SPF) and recorded as *J. Feres 2* (SPF). A total of 18 markers (90%) were cross-amplified in *A. peregrina* (Table 2), and 14 had more than one allele.

CONCLUSIONS

Our results provide new variable microsatellite markers for *A. colubrina* that may be used for the precise estimation of parameters such as genetic diversity, population structure, gene flow, and mating system. Species of the genus *Anadenanthera* are difficult to distinguish among each other. Therefore, since the markers we developed cross-amplified in *A. peregrina*, they might prove useful to identify distinguishing features among species.

LITERATURE CITED

ALZATE-MARIN, A. L., M. C. GUIDUGLI, H. H. SORIANI, C. A. MARTINEZ, AND M. A. MESTRINER. 2009. An efficient and rapid DNA minipreparation procedure suitable for PCR/SSR and RAPD analyses in tropical forest tree species. *Brazilian Archive of Biology and Technology* 52: 1217–1224.

BILLOTTE, N., P. J. L. LAGODA, A. M. RISTERUCCI, AND F. C. BAURENS. 1999. Microsatellite-enriched libraries: Applied methodology for the development of SSR markers in tropical crops. *Fruits* 54: 277–288.

CARVALHO, P. E. R. 2002. Angico-branco. Circular técnica 56. EMBRAPA-FLORESTAS, Colombo, Paraná, Brazil. Website http://ainfo.cnptia.embrapa.br/digital/bitstream/item/42015/1/CT0056.pdf [accessed 22 February 2012].

CHASE, M., R. KESSELI, AND K. BAWA. 1996. Microsatellite markers for population and conservation genetics of tropical trees. *American Journal of Botany* 83: 51–57.

FERES, J. M., M. L. L. MARTINEZ, C. A. H. MARTINEZ, M. A. MESTRINER, AND A. L. ALZATE-MARIN. 2009. Transferability and characterization of nine microsatellite markers for the tropical tree species *Tabebuia roseo-alba. Molecular Ecology Resources* 9: 434–437.

GOUDET, J. 2002. FSTAT: A program to estimate and test gene diversities and fixation indices (version 2.9.3.2). Website http://www2.unil.ch/popgen/softwares/fstat.htm [accessed 22 March 2011].

KALINOWSKI, S. T., M. L. TAPER, AND T. C. MARSHALL. 2007. Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Molecular Ecology* 16: 1099–1106.

LEWIS, P. O., AND D. ZAYKIN. 2002. Genetic Data Analysis: Computer program for the analysis of allelic data. Website http://hydrodictyon.eeb.uconn.edu/people/plewis/software.php [accessed 22 March 2011].

MONTEIRO, J. M., F. C. B. R. ALMEIDA, U. P. ALBUQUERQUE, F. P. LUCENA, T. N. FLORENTINO, AND L. C. OLIVEIRA. 2006. Use and traditional management of *Anadenanthera colubrina* (Vell.) Brenan in the semi-arid region of northeastern Brazil. *Journal of Ethnobiology and Ethnomedicine* 2: 6.

RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution; International Journal of Organic Evolution* 43: 223–225.