Development of microsatellite markers for the endangered Neotropical tree species *Tibouchina papyrus* (Melastomataceae)

M.P.C. Telles[1], F.P. Peixoto[1], J.S. Lima[1], L.V. Resende[1], R.P. Vianello[2], M.E.M.T. Walter[3] and R.G. Collevatti[1]

[1]Laboratório de Genética e Biodiversidade, Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO, Brasil
[2]Laboratório de Biotecnologia,
Empresa Brasileira de Pesquisa Agropecuária Arroz e Feijão,
Santo Antônio de Goiás, GO, Brasil
[3]Departamento de Ciência da Computação, Universidade de Brasília, Brasília, DF, Brasil

Corresponding author: M.P.C. Telles
E-mail: tellesmpc@gmail.com

ABSTRACT. We isolated and characterized 12 microsatellite loci for *Tibouchina papyrus* (Melastomataceae), an endangered species with narrow and disjunct range, endemics to a few localities in "cerrado rupestre" from Central Brazil. These microsatellites were obtained by sequencing of a genomic shotgun library for primer design. Leaves from 96 individuals collected in the three known local populations were genotyped using the 12 primers designed to analyze the polymorphisms at each locus. The number of alleles per locus ranged from one to six; two loci were monomorphic. Among the polymorphic loci, expected heterozygosities ranged from 0.161 to 0.714. Combined paternity exclusion probability was 0.957 and combined genetic identity (0.051) was high for studies

on parentage. *Tibouchina papyrus* is a rare and endemic tree species of outcrop quartzite and sandstone soils, with highly isolated populations, which may have lead to the low degree of polymorphism that we detected. Also, motifs of most loci are larger than dinucleotide, which typically display lower levels of polymorphism.

Key words: Cerrado; "Pau-papel"; Shotgun library; SSR; STR

INTRODUCTION

Microsatellites are one of the most powerful molecular markers to estimate population genetic parameters and understand detailed patterns of gene flow and parentage composition, because of the high genetic information content (Morgante and Olivieri, 1993). Despite the usefulness of microsatellite makers, there are few reports on the development, characterization and use of microsatellite loci in Neotropical tree species (Collevatti et al., 1999; Braga et al., 2007; Reis et al., 2009).

Tibouchina papyrus (Pohl) Toledo (Melastomataceae), known as "pau-papel", is endemic to outcrop quartzite and sandstone soils of the vegetation subtype known as Cerrado rupestre of the Cerrado biome, in central Brazil. The species occurs only in three localities: Serra de Natividade, Southeast Tocantins, Serra Dourada and Serra dos Pirineus, West Goiás. Besides the endemism and rarity, the species is threaten because of the highly unstable habitat with high levels of disturbance caused by fire during the dry season and sandstone and quartzite disruption, mainly during the rainy season (Collevatti et al., 2010).

We are interested in understanding the population genetic structure, patterns of gene flow and mating system of *T. papyrus*, to generate useful information for conservation strategies. As part of this project we report here the development of 12 microsatellite loci for the species.

MATERIAL AND METHODS

Microsatellite isolation for primer design was based on a genomic shotgun library. Total genomic DNA (2.0 μg) from one individual of *T. papyrus* was shared using a sonicator at 120 W for 1 h and 45 min, to obtain fragments of 200 bp to 1.0 kb. Fragments were recovered and cloned into a pMOSBlue dephosphorylated blunt vector using the Blunt-ended PCR Cloning Kit® (GE HealthCare, Sweden). Cloned fragments were transformed into chemically pMOSBlue® competent cells (GE HealthCare) and plated onto Luria-Bertani (LB) plates containing ampicillin and X-Gal. Recombinant clones were grown overnight in liquid ampicillin LB media and plasmid DNA was extracted using standard protocol (Sambrook and Russell, 2001). DNA inserts were sequenced using U19 primer on 3100 automated DNA sequencer (Applied Biosystems, USA) using the DYEnamicET terminator kit (GE Healthcare), according to the manufacturer instruction.

Sequences were analyzed and filtered by their quality and length (phred value ≥20; length ≥150) and analyzed for their nucleotide content. The reads obtained were screened for microsatellites using the WEBSAT software (Martins et al., 2009). Oligonucleotides complementary to the repeats were designed with the Primer 3 software (Rozen and Skaletsky, 2000).

To characterize the polymorphism at the developed loci, 96 individuals of *T. papyrus* were genotyped. Total genomic DNA was extracted from expanded leaves following the standard CTAB 2% procedure (Doyle and Doyle, 1990). Microsatellite amplifications were performed in a 15-µL reaction volume containing 15 ng template DNA, 0.9 µM of each primer, 1 U Taq DNA polymerase (Phoneutria, BR), 250 µM of each dNTP, 0.25 µg BSA and 1X reaction buffer (10 mm Tris-HCl, pH 8.3, 50 mm KCl, 1.5 mm $MgCl_2$). Polymerase chain reactions (PCRs) were carried out in PE 9700 thermal cycler (Applied Biosystems) under the following conditions: 96°C for 2 min (one cycle), 94°C for 1 min, 54° to 66°C (according to each primer annealing temperature; see Table 1) for 1 min, 72°C for 1 min (30 cycles), and 72°C for 10 min (one cycle). Polymorphisms were detected on 6% denaturing polyacrylamide gels stained with silver nitrate (Creste et al., 2001) and sized by comparison to a 10-bp DNA ladder standard (Invitrogen, USA). The number of alleles per locus*,* observed and expected heterozygosities under Hardy-Weinberg (Nei, 1978), and inbreeding coefficient were estimated (Weir and Cockerham, 1984). Analyses were performed with FSTAT 2.9.3.2 (Goudet, 2002) and randomization-based tests with Bonferroni's correction were performed to test for deviation from Hardy-Weinberg expectations and for linkage disequilibrium (Goudet et al., 1996). We also estimated the probability of genetic identity (Paetkau et al., 1995) and paternity exclusion probability (Weir, 1996), for each polymorphic locus and overall loci, using the Identity 1.0 software (Wagner and Sefc, 1999).

RESULTS AND DISCUSSION

We sequenced 1344 clones from the genomic library of *T. papyrus* that generated 927 valid reads. From these sequences, 232 (25% of valid reads; 17.3% of the clones) presented microsatellites: 139 had mononucleotide microsatellites; 22 had dinucleotides; 27 had trinucleotides; 29 had tetranucleotides; 10 had pentanucleotides, and 5 had hexanucleotides. Primers were designed for 19 (2.0% of the valid reads) fragments containing microsatellites (8.2%), and 12 amplified clearly interpretable products using a single PCR protocol (Table 1). From these 12 microsatellite loci, 2 were di-, 6 tri-, 2 tetra-, and 2 hexanucleotides (Table 1).

Only two loci presented no polymorphism and expected heterozygosities were low, ranging from 0.161 to 0.714 (Table 1). All polymorphic loci are in Hardy-Weinberg equilibrium ($P < 0.05$) and all pairs of loci are in linkage equilibrium ($P < 0.05$). Probability of identity was high (0.051) and probability of paternity exclusion was equal to 0.957 (Table 1) showing that further studies on parentage analysis and fine-scale genetic structure will require a search for more polymorphic loci.

Tibouchina papyrus is an endemic and rare species, which may lead to the low polymorphism found in this study, compared to other Cerrado tree species (Collevatti et al., 1999; Braga et al., 2007). Besides that, most loci are tri-, tetra- and hexanucleotides that typically present lower levels of polymorphisms due to lower mutation rates when compared to dinucleotides (Chakraborty et al., 1997). Also, the number of alleles per locus is positively correlated with the number of repeat motifs (Weber, 1990; Taramino and Tingey, 1996; Brandström and Ellegren, 2008). Although the number of sequences surveyed in this study was limited, our results showed that the *T. papyrus* genome has a low abundance of microsatellites (only 2.0% of the sequences presented microsatellites) and also that repeat length is short, which may limit polymorphism.

Table 1. Characterization of the 12 microsatellite loci developed for *Tibouchina papyrus*, based on 96 individuals.

Locus	Repeat motif	Primer sequence (5'- 3')	T_a	Size range (bp)	A	H_O	H_E	f	I	Q	GenBank
Tpap02	AC(5)	F-CTCAGTTGGGAAACGGTCAT R-GTAGAAAATTCCGGATGCGA	60	335	1	0.000	0.000	-	-	-	HQ606062
Tpap04	ATA(4)	F-AATGTCAGTCGAAGGCTGGT R-ATTTGCATGCTCATTTGGTT	54	136/151	4	0.136	0.714	0.810	0.242	0.457	HQ606063
Tpap05	GTT(5)	F-CGCTGCCTTCTTTACGATTC R-CTAACCCCACGCCAGATTT	58	348/351	2	0.135	0.161	0.166^{NS}	0.742	0.074	HQ606064
Tpap06	TAA(4)	F-CAAGCGCATAAACATAAACTCG R-ACAAAACCAAGAAATTGCGG	56	137/143	2	0.156	0.298	0.480	0.625	0.126	HQ606065
Tpap08	CGG(6)	F-TAGCGGGAGATCATTCCAAG R-GGTAGGTGCAGAGCATGTGA	68	220/232	5	0.411	0.673	0.394	0.256	0.427	HQ606066
Tpap09	TAATAG(3)	F-CCCCGAGAGACATAGAATGC R-CACGTCACAAGCGACATTTA	54	143	1	0.000	0.000	-	-	-	HQ606067
Tpap12	GA(6)	F-CCGACGGGGTCATTTTATTA R-AAAGTCGGGTCAATCGTGTC	60	151/173	6	0.178	0.705	0.749	0.234	0.462	HQ606068
Tpap15	TGC(5)	F-GACTAGCGGATGGAAATGGA R-GGGCAAACAAGGTGAGGATA	66	225/228	2	0.041	0.457	0.910	0.608	0.176	HQ606069
Tpap16	AAAT(3)	F-CACCAAAATGAGGGAGGAAA R-CCATCAATCCTTGTTGTCCC	58	257/261	2	0.135	0.358	0.626	0.604	0.147	HQ606070
Tpap17	AGAAGG(3)	F-GGAGGATGGTGATAGGGGTT R-ATGCTAACGAACAACCCGAC	65	242/254	3	0.145	0.551	0.738	0.442	0.274	HQ606071
Tpap18	AGTC(3)	F-CCACGTCACTCGATTACCCT R-CAAGGGACTTTTGGTTTCCA	54	188/192	2	0.052	0.425	0.879	0.601	0.167	HQ606072
Tpap19	ACC(5)	F-TGGAAACCAAAAGTCCCTTG R-AGAGAGAGAGAGACTGGTGGA	64	116/128	4	0.125	0.497	0.751	0.425	0.258	HQ606072
All loci						0.151	0.483	0.689	0.051	0.957	

T_a = annealing temperature; A = number of alleles; H_O = observed heterozygosity; H_E = expected heterozygosity; f = fixation index; I = probability of genetic identity; Q = paternity exclusion probability. Value followed by NS is not significant ($P > 0.0042$), Bonferroni's correction for a nominal value of 5%.

ACKNOWLEDGMENTS

REFERENCES

Braga AC, Reis AMM, Leoi LT, Pereira RW, et al. (2007). Development and characterization of microsatellite markers for the tropical tree species *Tabebuia aurea* (Bignoniaceae). *Mol. Ecol. Notes* 7: 53-56.

Brandström M and Ellegren H (2008). Genome-wide analysis of microsatellite polymorphism in chicken circumventing the ascertainment bias. *Genome Res.* 18: 881-887.

Chakraborty R, Kimmel M, Stivers DN, Davison LJ, et al. (1997). Relative mutation rates at di-, tri-, and tetranucleotide microsatellite loci. *Proc. Natl. Acad. Sci. U. S. A.* 94: 1041-1046.

Collevatti RG, Brondani RV and Grattapaglia D (1999). Development and characterization of microsatellite markers for genetic analysis of a Brazilian endangered tree species *Caryocar brasiliense*. *Heredity* 83 (Pt 6): 748-756.

Collevatti RG, Lima JS, Soares TN and Telles MPC (2010). Spatial genetic structure and life history traits in Cerrado tree species: inferences for conservation. *Nat. Conserv.* 8: 54-59.

Creste S, Tulmann-Neto A and Figueira A (2001). Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Mol. Biol. Rep.* 19: 299-306.

Doyle JJ and Doyle JL (1990). Isolation of plant DNA from fresh tissue. *Focus* 12: 13-15.

Goudet J (2002). FSTAT 2.9.3.2.: A Program to Estimate and Test Gene Diversities and Fixation Indices. Available at [http://www.unil.ch/izea/softwares/fstat.html]. Accessed April 25, 2010.

Goudet J, Raymond M, Meeus T and Rousset F (1996). Testing differentiation in diploid populations. *Genet. Res.* 144: 1933-1940.

Martins WS, Lucas DC, Neves KF and Bertioli DJ (2009). WebSat - a web software for microsatellite marker development. *Bioinformation* 3: 282-283.

Morgante M and Olivieri AM (1993). PCR-amplified microsatellites as markers in plant genetics. *Plant J.* 3: 175-182.

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89: 583-590.

Paetkau D, Calvert W, Stirling I and Strobeck C (1995). Microsatellite analysis of population structure in Canadian polar bears. *Mol. Ecol.* 4: 347-354.

Reis AMM, Braga AC, Lemes MR, Gribel ROG, et al. (2009). Development and characterization of microsatellite markers for the Brazil nut tree *Bertholletia excelsa* Humb. & Bonpl. (Lecythidaceae). *Mol. Ecol. Resour.* 9: 920-923.

Rozen S and Skaletsky HJ (2000). Primer3: Bioinformatics Methods and Protocols. In: Methods in Molecular Biology (Krawetz S and Misener S, eds.). Humana Press, Totowa, 365-386. Available at [http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi.]. Accessed February 8, 2010.

Sambrook J and Russell DW (2001). Molecular Cloning: a Laboratory Manual. 3rd edn. Cold Spring Harbor, New York.

Taramino G and Tingey S (1996). Simple sequence repeats for germplasm analysis and mapping in maize. *Genome* 39: 277-287.

Wagner HW and Sefc KM (1999). Identity 1.0. Centre for Applied Genetics. University of Agricultural Sciences, Vienna. Available at [http://www.boku.ac.at/zag/forsch/identity.htm]. Accessed April 25, 2010.

Weber JL (1990). Informativeness of human (dC-dA)n.(dG-dT)n polymorphisms. *Genomics* 7: 524-530.

Weir BS (1996). Genetic Data Analysis II - Methods for Discrete Population Genetic Data. Sinauer Associates Inc., Sunderland.

Weir BS and Cockerham C (1984). Estimating F-statistics for the analysis of population structure. *Evolution* 38: 1358-1370.