GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Development of microsatellite markers for *Hancornia speciosa* Gomes (Apocynaceae)

A.J.L. Rodrigues[1], A.T. Yamaguishi[2], L.J. Chaves[3], A.S.G. Coelho[3], J.S. Lima[4,5] and M.P.C. Telles[4]

[1]Universidade Estadual de Goiás, Anápolis, GO, Brasil
[2]Embrapa - Recursos Genéticos e Biotecnologia, Brasília, DF, Brasil
[3]Escola de Agronomia, Universidade Federal de Goiás, Goiânia, GO, Brasil
[4]Instituto de Ciências Biológicas, Universidade Federal de Goiás, Goiânia, GO, Brasil
[5]Programa de Pós-Graduação em Ecologia e Evolução, Universidade Federal de Goiás, Goiânia, GO, Brasil

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

Genet. Mol. Res. 14 (3): 7274-7278 (2015)
Received June 6, 2014
Accepted February 20, 2015
Published July 3, 2015
DOI http://dx.doi.org/10.4238/2015.July.3.2

ABSTRACT. Herein, we describe 34 microsatellite loci developed using an enrichment genomic library for the tree species *Hancornia speciosa* Gomes (Apocynaceae). Thirty-five individuals were genotyped using 34 primers to analyze the polymorphisms at each locus. The number of alleles per locus ranged from 4 to 20. The average number of alleles was 8.11, and the expected heterozygosity ranged from 0.62 to 0.94. These microsatellite primers will be useful in population genetics studies for this species.

Key words: Cerrado; Enrichment genomic library; Mangaba; SSR; STR

INTRODUCTION

Microsatellites are noncoding, repetitive DNA regions that consist of tandemly repeated small motifs. These markers have been an important tool in population genetic studies of natural species due to their reproducibility, multiallelic nature, codominant inheritance, relative abundance, and wide genome coverage (Morgante and Olivieri, 1993; Goldstein and Schlotterer, 1999). Nevertheless, the use of microsatellite loci in neotropical tree species is still scarce, in part due to the difficulties in developing specific markers and transferring them from closely related species (Telles et al., 2011).

Hancornia speciosa Gomes (Apocynaceae) is a widely distributed Neotropical plant species native to Brazil; it occurs in the north, northeast, and central regions of the country, mainly in the Cerrado biome. Its fruits possess agroindustrial potential and stand out for the taste of their derivative products, which are highly appreciated and, thus, easily marketable (Nogueira and Albuquerque, 2006; Silva-Júnior et al., 2006).

Knowledge of the genetic structure of this species may provide important information for developing conservation and management programs, but microsatellite markers are not available for this species. Because the species is monotypic within the genus, the possibility of transferability is potentially low, and the best current technological strategy to deal with lack of markers is to specifically develop them using an enriched genomic library. Thus, the aim of this study was the development of microsatellite loci for this species.

MATERIAL AND METHODS

A microsatellite-enriched library was constructed using total genomic DNA of one individual of *H. speciosa* based on the protocol developed by Rafalski et al. (1996). The genomic DNA was extracted from expanded leaves following the standard 2% CTAB procedure (Doyle and Doyle, 1990); it was then digested by the restriction enzymes *Sal*3AI (GATC), *Mse*I (TAA), and *Tsp*I 509 (AATT). Digested fragments ranging in size between 200 and 800 bp were gel-extracted using the Qiaquick Gel Extraction Kit (Qiagen). The DNA fragments were linked for 12 h at 12°C to the adaptors *Tsp*I, prepared from long (5'-AATTGGAATTCGACTCGCAGCAGCC-3') and short (5'-GGCTGCTGCGAGTCGAATTCC-3') oligos in a reaction including 3 U T4 DNA ligase (Amersham Pharmacia Biotech). The amplified DNA fragments were hybridized with a probe $(TC)_{13}$ linked to biotin. The hybridization mixture was captured by coated magnetic beads (Dynabeads, Streptavidina Bohering Mannhein). The enriched DNA obtained was amplified through 20 cycles using primers complementary to the adaptor sequence. The purified DNA was cloned into a pGEM-T Easy Vector (Promega Corp.), and plasmids were inserted into transformed *Escherichia coli* competent cells. The positive clones were amplified through polymerase chain reaction (PCR) using the M13 primers (forward and reverse) and sequenced on an ABI 3000 semi-automated sequencer.

The regions containing microsatellites were analyzed according to the quality of the flanking regions and repetitive sequences. Primer pairs for these regions were designed with the Primer3 software (Rozen and Skaletsky, 2000). The Tandem Repeats Finder software (Benson, 1999) was used to identify microsatellite regions with length ≥18 bp. The quality of the sequences was analyzed with the Phred software (Ewing et al., 1998) and aligned with the BioEdit 5.0.0 (Hall, 1999) and Clustal X (Thompson et al., 1997) software programs.

Thirty-five individuals of *H. speciosa* from different populations were genotyped to characterize the polymorphism at the developed loci. Microsatellite amplifications were performed in 25 µL volume reactions containing 10 ng DNA, 1X reaction buffer, 1.5 mM $MgCl_2$, 0.4 mM dNTP, 0.1 µM primer, and 1 U Taq DNA polymerase. PCRs were carried out in a PTC-100 thermal cycler (MJ Research Inc.) with the following conditions: 94°C for 5 min (one cycle), 94°C for 1 min, 48° to 64°C (according to each primer annealing temperature, see Table 1) for 1 min, 72°C for 1 min (30 cycles), and 72°C for 7 min (one cycle). Amplified products were separated in 6% denaturing polyacrylamide gels stained with silver nitrate (Creste et al., 2001) and sized by comparison with 10- and 100-bp DNA ladder standards (Invitrogen, Frederick, MD, USA).

The number of alleles per locus (N_A) and the observed (H_O) and expected heterozygosities (H_E) under Hardy-Weinberg equilibrium (Nei, 1978) were estimated with the GDA software (Lewis and Zaykin, 2001).

Table 1. Characterization of the 34 microsatellite loci developed for *Hancornia speciosa*, based on 35 individuals.

Locus	Repeat motif	Primer sequence (5'-3')	Expected size (bp)	Allele range (bp)	Ta (°C)	N_A	H_E
HS 01	$(GCA)_6(TC)_{20}(GCA)_8$	F: GTGTCTTCCATCCGAGCTTAAC R: TTTCCCAGAAAGGAGAGGTACA	275	250-310	50	13	0.89
HS 02	$(CT)_{14}$	F: AATTCAACCCTTCTGCGAATC R: CACCAGGAACGATCAGGAAG	169	80-150	50	4	0.68
HS 03	$(CT)_5(CT)_6$	F: CACTCTCTCTCAGCATTTCCCT R: TGCATAGAAGGAGAAGAAGAAGC	158	120-180	58	7	0.81
HS 04	$(AG)_{17}$	F: GCGAGTCGAATTCCAATTACTC R: AAAGTCCAGTAATAGCGCCAAA	247	80-200	54	13	0.84
HS 05	$(GA)_{15}(TGC)_6$	F: GGGTGTACTGCCACAAGGTACT R: GTGTCTTCCATCCGAGCTTAAC	177	200-300	58	7	0.73
HS 06	$(GA)_{14}$	F: CGGCTGTATTAAGTCTATTGCCA R: CCCTGCCACTCTCTTTTCC	107	100-150	50	7	0.75
HS 07	$(CT)_6(TC)_5(CA)_6$	F: ACTCACACAAAGCTCACACACC R: AGGTAATCAGTTCTGGGGAGGT	261	220-260	56	5	0.69
HS 08	$(CA)_6(CT)_{17}$	F: AATGTAGAGGTGAACGAGTGGG R: TACACCCTGCTCATCGTTTATG	248	200-250	48	12	0.89
HS 09	$(CT)_7(CT)_{10}(CT)_6$	F: TGCAAACCCTCGTTTATTTCTT R: ATTGTGTGTGTTGTGTGTGGG	262	200-300	64	8	0.81
HS 10	$(CT)_{14}(CT)_8$	F: ACAAATCAATGAGGAGGTGCTT R: TAACTATGTGCAACCGCAAGAC	135	100-200	52	8	0.85
HS 11	$(GA)_{17}$	F: GTGATATTTCGTGCTCTCCAAG R: CTCTGCCACTGTGCAACC	106	100-200	50	8	0.77
HS 12	$(CT)_{22}$	F: CAAACCCTCGTTCTTCTCTTCT R: GATGTCGCAACTCGAGCA	214	200-250	56	5	0.66
HS 13	$(CT)_{12}$	F: CTGGGGTACTTCAGCAAATCAC R: CATCAAAGACCGTTGTCTCCTT	101	100-150	56	8	0.84
HS 14	$(TC)_{14}$	F: GAGCAGGAGTCAGGAAAATCAC R: ACAGTGAAGGGGCAATGAAG	126	100-200	56	6	0.71
HS 15	$(GA)_{16}$	F: GTGAGTGTTGTGCGTGTGTGT R: CTTCTTCTTCCTCTCGCGGT	116	150-300	60	5	0.62
HS 16	$(GA)_{12}$	F: CGTTGGTAGCGGCTGTATTAAG R: CCCCTCCTGCCACTCTCT	114	100-150	48	7	0.72
HS 17	$(GA)_{16}$	F: ACTCGAGCAGAAGAAGCAAATC R: ACACACCCTCATCAGCCC	116	100-200	54	11	0.82
HS 18	$(AG)_{14}$	F: ATTCATGCTCCACTGGCTTC R: GACCACAGCTAGTGACGTGTTC	207	300-350	50	11	0.87
HS 19	$(TTC)_5(GCC)_6$	F: ATTCTGCAAACCCTCATTTC R: AAAGAGAGAGTGTGTGTGCG	109	350-400	56	4	0.63
HS 20	$(CT)_{11}$	F: CTAACCCACTACAACTCTGGGC R: ATCCGATTTAGGACATTGGGT	140	200-250	50	9	0.84

Continued on next page

Table 1. Continued.

Locus	Repeat motif	Primer sequence (5'-3')	Expected size (bp)	Allele range (bp)	Ta (°C)	N_A	H_E
HS 21	$(AG)_{13}$	F: GATTGTGTGTGGTGTGTGTGG R: CCCTCGTTCTTCTCTTCTTCTTC	207	200-300	64	7	0.76
HS 22	$(GA)_{12}$	F: GGACGAAACGAAATGGAGAGTA R: AGTAAAGACACGTCATCCCCAC	160	180-220	56	7	0.74
HS 23	$(CGC)_6$	F: TGCAAACCCTCATTTCTTTTCTTC R: GGAGCAAATCGGGAAGCC	134	180-250	56	9	0.84
HS 24	$(AG)_{12}$	F: GCTAAATCAAGCAAACCTCGAC R: AAAGCAGTCCATGATCCATTTC	152	200-300	58	7	0.81
HS 25	$(TC)_{11}$	F: CACGGCTGTCTGTCGTTG R: ACTCGAGCAGAAGAAGCAAATC	118	80-150	56	6	0.65
HS 26	$(CT)_{10}$	F: CAAACAAGCTTTATGTGGGTCA R: AGCTCAAGGAAGTGGGATCTAA	172	200-300	58	9	0.80
HS 27	$(GA)_{14}$	F: TATAGTGGTCCTGCACCCTTGT R: TTTTCCCTTGTGCTTCGC	128	100-150	54	20	0.94
HS 28	$(CT)_{15}$	F: TCTATTGGAGTGTTTCCCTCAC R: AGATTCAAGAGAAAACGGAAAA	129	250-300	56	5	0.76
HS 29	$(CT)_{15}$	F: GTGAGCTTGTGCTACTCTCCCT R: TTGTCACCGTGTTTCTAGCAGT	178	300-400	50	6	0.68
HS 30	$(AG)_{10}$	F: GAGGAATCTCAGCCAAGTCCTA R: CCCAGCCTCTACAAACTCTCTG	169	180-200	56	10	0.88
HS 31	$(AG)_{19}$	F: GTCGATCCGATTTAGGACATTG R: GCACACACCCACACTCCA	176	50-100	56	5	0.70
HS 32	$(CT)_9$	F: CTTTTCTGCAACCCTCATTCTT R: ACTCGAGCTGAGAAAACGAATC	185	200-300	52	14	0.90
HS 33	$(AG)_{24}$	F: CGTTGGTAGCGGCTGTATTAAG R: CACTCTCTTTTCCCGATTTTCC	132	80-120	56	9	0.83
HS 34	$(TGG)_5$	F: GCATCTTGAGAAGAAAGAGGGA R: TGTAGATGGAATCTTGCCACTG	144	180-200	60	4	0.65
Average	-	-	-	-	-	8.11	0.77

Ta = annealing temperature; N_A = number of alleles; H_E = expected heterozygosity.

RESULTS AND DISCUSSION

From the 1056 positive clones, we sequenced a total of 576. It was possible to design primers in the conservative region for 74 sequences with different compositions, sizes, and repeat motifs. Dinucleotide regions, which typically exhibit higher polymorphism than that of other regions (Chakraborty et al., 1997), were preferentially chosen for amplification. From these 74 microsatellite loci, 34 were selected for synthesis and amplification (Table 1). The selection of some trinucleotide regions occurred in an attempt to obtain greater coverage of the genome.

All loci presented polymorphisms, with the N_A per locus ranging from 4 to 20. Average allelic richness was 8.11 and the H_E was high, ranging from 0.62 to 0.94 (Table 1). Thus, the microsatellites developed herein revealed high genetic diversity in the species studied and may allow individual discrimination, with many potential applications for population studies (i.e., pollen dispersal, population structure, and gene flow). Based on the above parameters, we conclude that the set of molecular markers developed herein is adequate for further use in population genetics studies for this species, allowing both the understanding of microevolutionary processes and the establishment of more effective conservation strategies for this species.

REFERENCES

Benson G (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 19: 573-580.

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

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.

Ewing B, Hillier L, Wendl MC and Green P (1998). Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.* 8: 175-185.

Goldstein DB and Schlotterer C (1999). Microsatellites: evolution and applications. Oxford University Press, New York.

Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95-98.

Lewis PO and Zaykin D (2001). Genetic Data Analysis: computer program for the analysis of allelic data. Version 1.0 (d16c). Available at [http://www.eeb.uconn.edu/people/plewis/software.php].

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.

Nogueira RJMC and Alburquerque MB (2006). Ecofisiologia. In: A cultura da Mangaba (Silva-Júnior JF and Lédo AS, eds.). 1st edn. Embrapa Tabuleiros Costeiros, Brasília, 36-42.

Rafalski JA, Vogel JM, Morgante M, Powell W, et al. (1996). Generating and using DNA markers in plants. In: Non mammalian genomic analysis: a practical guide. Academic Press, New York, 75-134.

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.

Silva Júnior JF, Araújo IA, Barreira Neto M, Espíndola ACM, et al. (2006). Recursos genéticos nos tabuleiros costeiros e baixada litorânea do Nordeste. In: A cultura da Mangaba (Silva-Júnior JF and Lédo AS, eds.). 1st edn. Embrapa Tabuleiros Costeiros, Brasília, 58-72.

Telles MPC, Peixoto FP, Lima JS, Resende LV, et al. (2011). Development of microsatellite markers for the endangered Neotropical tree species *Tibouchina papyrus* (Melastomataceae). *Genet. Mol. Res.* 10: 321-325.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.