ISSN 1479-2621
Plant Genetic Resources: Characterization and Utilization; 1–10
doi:10.1017/S1479262115000179

Cross-species amplification and characterization of new microsatellite markers for the macaw palm, *Acrocomia aculeata* (*Arecaceae*)

Fekadu Gebretensay Mengistu[1]*, Sérgio Yoshimitsu Motoike[1], Eveline Teixeira Caixeta[2], Cosme Damião Cruz[3] and Kacilda Naomi Kuki[1]

[1]*Universidade Federal de Viçosa, Departamento de Fitotecnia, Av. P.H. Rolfs, Campus, Viçosa - MG 36570-000, Brazil,* [2]*Embrapa Café, BIOAGRO, BIOCAFÉ, Universidade Federal de Viçosa, Viçosa - MG 36570-000, Brazil and* [3]*Universidade Federal de Viçosa, Departamento de Biologia Geral, Av. P.H. Rolfs, Campus, Viçosa - MG 36570-000, Brazil*

Received 26 January 2015; Revised 1 April 2015; Accepted 10 April 2015

Abstract

Microsatellites or simple sequence repeats (SSRs) are useful molecular markers allowing for efficient conservation and sustainable use of genetic resources of plant species. Development of SSR marker system for new species is a very expensive task and time consuming. Cross-species amplification of microsatellite loci is considered as a cost-effective approach for developing microsatellite markers for new species. The aim of this work was to examine the transferability of some SSR markers of two *Arecaceae* species (*Astrocaryum aculeatum* and *Elaeis oleifera*), in *Acrocomia aculeata*. Out of the total markers analysed 44% of the markers successfully amplified the genomic DNA in *A. aculeata*, of which 26% were polymorphic detecting a range of three to eight alleles with an average of 4.5 per locus. High average percentage of polymorphic loci ($P = 71.2\%$) per provenance was obtained within a range of 57–100% detecting genetic variation in *A. aculeata* germplasm collections. The polymorphic markers detected a positive inbreeding coefficient ($F > 0$) per locus revealing heterozygote deficiency in the accessions that were analysed. As the cross-amplification was at family level, in which the taxonomic distance is relatively wider between the sources (*A. aculeatum* and *E. oleifera*) and the target (*A. aculeata*) species, the amplification success was relatively low. However, the results are promising and implicated that high cross-amplification success could be achieved at species or genus level in *A. aculeata*. The markers will contribute towards the domestication of the potential macaw palm through realizing various studies such as population genetics, germplasm characterization, genetic improvement and conservation.

Keywords: biodiesel; domestication; macauba; null alleles; SSRs; transferability

Introduction

Microsatellites (simple sequence repeats, SSRs) are the choice of many studies because of their high levels of polymorphism, reproducibility, co-dominant nature and genome and locus-specificity (Powell *et al.*, 1996; Pinheiro *et al.*, 2009).

The high transfer rate of SSRs between related species is an advantage to save a large portion of development time and costs (Barbará *et al.*, 2007). Hence, sourcing of SSR primers from related species could be a cost-effective alternative to develop primers especially for new species

* Corresponding author. E-mail: fgebretensay@yahoo.com

where abundant sequence data are not available (Kantety *et al.*, 2002) and resources for developing new SSR primers are limited (Roa *et al.*, 2000; Pinheiro *et al.*, 2009).

The macaw palm, *A. aculeata* (Jacq.) (Lodd. ex Mart.) – *Arecaceae* ($2n = 2x = 30$) (Abreu *et al.*, 2011), is an emerging perennial palm native to South America. It is monoecious and self-compatible with androgynous inflore scence, which bears a mixed reproductive system, with a predominance of out-crossing (Scariot *et al.*, 1991; Abreu *et al.*, 2012). The combination of two pollination strategies (entomophily and anemophily) with flexible reproductive systems (cross- and self-pollination) suggests that *A. aculeata* can be highly successful in the colonization of new areas, as evidenced by its ample distribution in Brazilian biomes and in the rest of the Neotropics (Scariot *et al.*, 1991).

The species is known for its high oil-producing potential, thereby becoming sources of biofuel for both aviation and automotive sectors (Lanes *et al.*, 2014). It produces up to 25 tonnes/ha of fruits, which can be processed to 4000 kg of vegetable oil and the resulting solid waste of which can be transformed into charcoal and nutritious cakes to generate energy and feed livestock (Tickel, 2000; Moura *et al.*, 2009). The biochemical properties of the oil are suitable for cosmetic industries and for biodiesel production (Fortes and Baugh, 1999; Bora and Rocha, 2004; Hiane *et al.*, 2005). Moreover, this palm has environmental benefits as it can grow in impoverished soils and drought-prevailing areas (Motoike and Kuki, 2009). Hence, *A. aculeata* is a suitable option for production of biodiesel among the common food-based oleaginous plants such as soybean, sunflower and oil palms (Teixeira, 2005).

However, *A. aculeata* has risks of predatory extractivism in natural populations, unsustainable land use and climate change, which potentially threaten its natural genetic diversity (Faleiro *et al.*, 2008; Ribeiro *et al.*, 2011). Our study focused on the development of new microsatellite markers for *A. aculeata* sourcing SSRs from related palm species to contribute for the conservation of genetic resources of this noble palm.

Therefore, this study was carried out to evaluate the transferability of some SSR markers from *Astrocaryum aculeatum* and *Elaeis oleifera* to *A. aculeata*. It was also performed to identify and characterize some polymorphic SSR markers from sets of markers previously developed for *A. aculeata*.

Materials and methods

Plant material and DNA isolation

Leaf samples of 192 *A. aculeata* germplasm accessions were obtained from the *ex situ* plant collection, Macaúba Active Germplasm Bank (BAG-Macaúba respository #: 084/2013/CGEN/MMA) located in the experimental farm of the Universidade Federal de Viçosa in the municipality of Araponga (20°40′1″S, 42°31′15″W), State of Minas Gerais, Brazil (Table 1). The accessions, obtained from six States in Brazil (Fig. 1), were developed from seeds germinated using a pre-germination protocol as described in patent INPI 014070005335 (Motoike *et al.*, 2007). The accessions represent six provenances containing 41 different families, each having a maximum of five individual samples (Table 1).

Genomic DNA (gDNA) was isolated from each individual sample according to the CTAB (Cetyl Tri-methyl Ammonium Bromide) method (Doyle and Doyle, 1990), with some modifications as described by Lanes *et al.* (2013). Isolated DNA samples were quantified using Multiscan™ GO Microplate Spectrophotometer (Thermo Fisher Scientific OY, Ratasite, Finland) at absorbance of 260 and 280 nm. The integrity of the DNA samples was confirmed by electrophoresis on 2% agarose gel and the working concentration was adjusted to 30 ng/μl.

Condition of polymerase chain reaction (PCR) and electrophoresis

The PCR was performed according to Nucci *et al.* (2008), Ramos *et al.* (2012) and Zaki *et al.* (2012) with minor modifications (Table 2). Totally, 20 SSR markers, originally developed for *E. oleifera* (Zaki *et al.*, 2012), 14 developed for *A. aculeatum* (Ramos *et al.*, 2012) and three identified from *A. aculeata* (Nucci, 2007) were used in the reactions. Of the total number of primers that amplified the target microsatellite loci, the sequences of only 18 were listed (Table 2).

The amplification cycles were carried out on a PCR thermal cycler (Applied Biosystem® Verti® cycler, Thermo Fisher Scientific Brand, USA) programmed as follows: initial denaturation step at 94°C for 5 min followed by 30 cycles of denaturation at 94°C for 1 min; annealing at primer-specific annealing temperature for 1 min (Table 2); extension at 72°C for 1 min and the final extension at 72°C for 8 min (Nucci *et al.*, 2008). PCR products were denatured in a bromophenol blue dye solution at 95°C for 5 min on the thermal cycler immediately before electrophoresis on 6% polyacrylamide gel in 1 × TBE (Tris–Borate–EDTA) buffer solution at 60 W for 1 h and 40 min.

Polyacrylamide gel staining

After electrophoresis, the PCR products were visualized in polyacrylamide gels stained with silver nitrate

Table 1. List of 41 *Acrocomia aculeata* families assessed in the Macaúba Active Germplasm Bank (BAG-Macaúba)

No.	Family code	State/provenance[a]	Coordinates[b]		No.	Family code	State/provenance[a]	Coordinates[b]	
			Latitude	Longitude				Latitude	Longitude
1	BGP99	PA	S 06 03 58.0	W 49 33 39.0	22	BGP11	MG	S 19 14 01.2	W 43 03 28.4
2	BGP82	PE	S 07 14 23.0	W 36 46 55.0	23	BGP9	MG	S 19 33 12.0	W 46 51 10.1
3	BGP124	PB	S 08 48 49.0	W 36 57 14.0	24	BGP78	MG	S 18 51 25.6	W 46 52 55.2
4	BGP51	SP	S 21 32 04.6	W 48 44 24.7	25	BGP37	MG	S 18 40 51.3	W 46 33 41.4
5	BGP34	SP	S 22 25 10.8	W 50 34 43.1	26	BGP33	MG	S 19 19 40.3	W 46 38 11.5
6	BGP47	SP	S 22 29 14.2	W 50 46 16.2	27	BGP21	MG	S 19 31 15.9	W 46 31 42.2
7	BGP20	MG	S 16 39 52.7	W 43 53 58.9	28	BGP2	MG	S 20 39 20.4	W 43 18 45.2
8	BGP27	MG	S 16 21 20.7	W 44 25 30.5	29	BGP76	MG	S 19 41 51.4	W 43 11 27.7
9	BGP22	MG	S 17 25 54.0	W 45 08 59.5	30	BGP25	MG	S 17 06 54.6	W 43 49 16.4
10	BGP16	MG	S 16 26 07.6	W 44 00 50.5	31	BGP64	MG	S 16 44 12.7	W 43 51 54.9
11	BGP49	MG	S 20 38 58.0	W 44 01 15.5	32	BGP105	MS	S 20 29 52.5	W 55 18 39.3
12	BGP10	MG	S 21 03 12.9	W 44 16 28.2	33	BGP102	MS	S 20 30 38.6	W 55 37 59.7
13	BGP68	MG	S 21 11 27.6	W 44 19 29.7	34	BGP104	MS	S 20 27 55.9	W 55 46 41.7
14	BGP3	MG	S 21 09 52.2	W 44 08 49.5	35	BGP117	MS	S 20 27 56.5	W 55 46 38.2
15	BGP51	MG	S 21 17 20.5	W 44 49 12.6	36	BGP118	MS	S 20 50 22.3	W 55 54 53.3
16	BGP5	MG	S 19 05 02.0	W 44 39 13.9	37	BGP112	MS	S 20 50 16.5	W 55 54 51.8
17	BGP14	MG	S 19 56 29.0	W 44 36 12.0	38	BGP106	MS	S 21 28 42.3	W 56 10 03.6
18	BGP18	MG	S 19 52 34.0	W 43 52 20.5	39	BGP92	MS	S 21 28 45.7	W 56 10 06.6
19	BGP24	MG	S 19 53 20.2	W 43 41 11.5	40	BGP103	MS	S 21 42 04.8	W 57 50 39.0
20	BGP1	MG	S 20 17 42.6	W 43 42 30.9	41	BGP119	MS	S 21 42 06.0	W 57 50 32.4
21	BGP52	MG	S 20 50 13.1	W 42 54 27.3					

[a] States include: PA = Pará, PE = Pernambuco, PB = Paraiba, SP = São Paulo, MG = Minas Gerais and MS = Mato Grosso do Sul. [b] Coordinates are in degrees, minutes and seconds for both the latitude (S = South) and longitude (W = West).

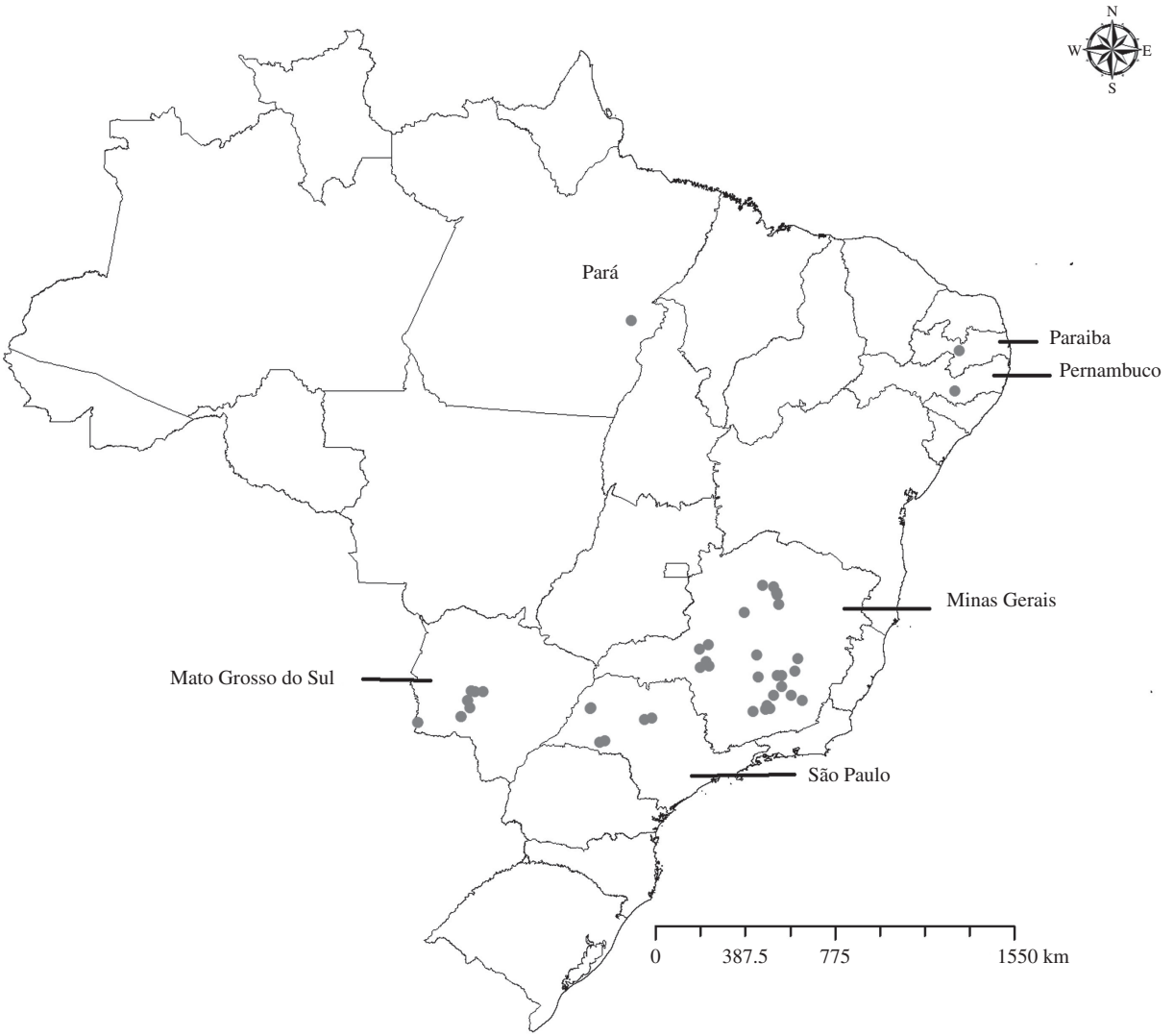

Fig. 1. Map of Brazil showing the six geographical states, where the original plant materials were collected. The states include: Pará; Pernambuco; Paraiba; São Paulo; Minas Gerais and Mato Grosso do Sul. Araponga is a city in Minas Gerais State, where the germplasm bank is located from which the experimental plant materials were obtained.

($AgNO_3$) according to Brito *et al.* (2010). The gels were immersed and agitated in several colouring steps in different solutions at different concentrations and durations until all allelic bands were totally visible for evaluation. Finally, the stained gels were allowed to dry in the air and scanned for documentation and DNA fragments were scored as co-dominant alleles for data analysis.

Data analyses

Co-dominant data were analysed using the GENES statistical software program (Cruz, 2013) to estimate allelic diversity, heterozygosity and polymorphism level of the SSR markers. The number of alleles per locus (A) was determined by quantifying the number of different alleles amplified by each marker analysed with 192 individuals. The total number of alleles per provenance (N_t) was determined by summing the number of alleles amplified by each locus (A). Hence, the average number of alleles per provenance (N_a) was calculated from the total number of alleles detected in the provenances by the loci that were analysed (Cruz *et al.*, 2011). Effective numbers of alleles (N_e), which are used to make sampling in successive generations of a given population, were determined by quantifying the number of alleles amplified by the polymorphic loci out of the total number of loci analysed with a criterion that alleles have a frequency of <0.95 (Cole, 2003).

The heterozygosity level of each locus was determined. Expected heterozygosity (H_E) per locus was estimated from the frequency of alleles detected per locus, while observed heterozygosity (H_o) was determined by

Table 2. List of SSR primer sequences used to amplify the target microsatellite loci in the cross-amplification

Locus	Forward primer (5′–3′)	Reverse primer (5′–3′)	Repeat motif	Accession name	T_a (°C)	Source[a]
Aac04	F: GCATTGTCATCTGCAACCAC	R: GCAGGGGCCATAAGTCATAA	(GT)7(GA)16	GF111930	60	1
Aac08	F: CGCACGTACACACACACACAT	R: GCAGGGGCCATAAGTCATAA	(CA)11	GF111934	57	1
Aac10	F: AGCCGTGAGTGAACTGCTTT	R: AAGCCCAAACTTCTTCCTCG	(CT)7	GF111936	60	1
Aac11	F: AAAGGAACAACCCAAGAGGG	R: TGGGGAGTGGACGTAAGTGT	(AC)5	GF111937	60	1
Aac12	F: GCTCTGTAATCTCGGCTTCCT	R: TCCAGTTCAAGCTCTCTCAGC	(GC)5(AC)3	GF111938	60	1
Aac13	F: CTAGACAACCCAAGAGAGGGG	R: TTGGAGAGTGGATGTAGGTGC	(CA)7	GF111939	60	1
sMo00020	F: CCTTTCTCTCCCTCTCCTTTTG	R: CCTCCCTCCCTCTCACCATA	(AG)15	Pr009947964	58	2
sMo00027	F: TTACAGTTGAGGCAGTATGTCAAT	R: CTGTATGTCAAACCTTCTGCAC	(TC)14	Pr009947965	50	2
sMo00055	F: GGCATTTCAGATAACGACAAA	R: GCACCCAAGTCTCTCTACCTC	(GA)11	Pr010315684	54	2
sMo00130	F: TAAGCAAAAGATCAGGGCACTC	R: GCTGGTGAAAATAGGTTTACAAAG	(AAG)11	Pr009947968	56	2
sMo00137	F: AGGAAGGAGAAGGCAGATGAACAG	R: CTTTGGATTTGAGCAGAGGAAG	(AAAT)6	Pr010315687	54	2
sMo00018	F: TTAAATGAGAGAGACGAGGAC	R: TGGAGCCATGAGAAAGAGTA	(CT)14	Pr009947963	54	2
sMo00141	F: ACTTGACATACAGGTTCCACTGA	R: CCTGCTACCTCCTAATTCTATCAAA	(TTCTT)5	Pr010317029	56	2
sMo00147	F: TACCCAATCCCACCGAGTTA	R: CGTCTCCACTGAACCACAAAA	(AAAAG)5	Pr010317030	54	2
sMo00161	F: ACTGTTTCGTCAAGCATTTG	R: ATCAAGAGAAGGTCGTGTCAG	(TG)8(AG)8	Pr010317032	54	2
Aacu38	F: TTCTCAGTTTCGTGCGTGAG	R: GGGAGGCATGAGGAATACAA	(TC)15	–	56	3
Aacu45	F: CAGACTACCAGGCTTCCAGC	R: TCATCATCGCAGCTTGACTC	(CGAC)5	–	56	3
Aacu74	F: TACTGTTGTGCCAAGTCCCA	R: GAGCACAAGGGGGATATCAA	(CA)15	–	56	3

T_a, annealing temperature.
[a]1, Ramos *et al.* (2012); 2, Zaki *et al.* (2012); 3, Nucci (2007).

calculating the number of heterozygotes observed out of the total number of individuals analysed (Nei, 1978). H_E and H_o were used to estimate the inbreeding coefficient (F) per locus to determine whether there are excess (or any deficiency) of heterozygotes per locus (Hartl and Clark, 1997).

The informativeness of the SSR markers was estimated in the analyses by calculating the polymorphic information content (PIC) of each locus, which was computed from the frequency of alleles detected per locus according to Bostein *et al.* (1980). To estimate the proportions of polymorphic loci detected in each provenance, the percentage of polymorphic loci (P) per provenance was determined by quantifying the number of polymorphic loci obtained out of the total number of loci analysed based on a criterion when a locus with most common alleles has a frequency of less than 0.95 (Cole, 2003).

The null allele test was performed to check for evidence of null alleles detected by the loci using the FreeNA (Chapuis and Estoup, 2007) computer program based on the Expectation Maximization (EM) algorithm as described in Dempster *et al.* (1977), which estimated the null allele frequency (p) for each locus across the provenances analysed. A test for deviations from Hardy–Weinberg Equilibrium (HWE) was performed for each locus across the provenances at a significance level of $\alpha = 0.05$ in GENEPOP (Raymond and Rousset, 1995).

Results

Cross-amplification and polymorphism

In the cross-amplification, 15 of the total SSR markers (44%) from *A. aculeatum* and *E. oleifera* successfully produced amplicons and were able to amplify the gDNA in *A. aculeata* (Table 3). However, only four (26%) of the markers were polymorphic and they detected a range of three to eight alleles with an average of 4.5 per locus (Table 3). Besides, three new polymorphic SSR markers were identified from sets of SSR markers previously designed for *A. aculeata*. Hence, a total of 38 alleles were amplified by the seven polymorphic SSR markers with an average of 5.4 per locus (Table 3).

Alleles amplified by the polymorphic loci were detected in the six provenances at different proportions from 14 (35.9%) to 33 (84.6%) (Table 3). Out of the total number of alleles, a range of 2.4–6.8 effective numbers of alleles (N_e) with an average of four per provenance were obtained (Table 3). The number of polymorphic loci across the provenances varied from 3.99 ($P = 57\%$) to 7 ($P = 100\%$) with an average of 4.98 ($P = 71.2\%$) per provenance (Table 3). Observed heterozygosity (H_o) ranged from 0.01 (sMo00137) to 0.61

Table 3. Characteristics of 15 SSR markers identified from *Astrocaryum aculeatum* (Ramos *et al.*, 2012) and *Elaeis oleifera* (Zaki *et al.*, 2012), which showed amplification in *Acrocomia aculeata* accessions and three SSR markers screened from sets of markers developed for *A. aculeata* (Nucci, 2007)

			Number of alleles amplified per provenance									
Locus	Amplicon (bp)	*A*	PA	PE	PB	SP	MG	MS	H_o	H_E	*F*	PIC
Aac04[a]	258–306	8	4	5	4	7	7	6	0.61	0.72	0.15	0.68
Aac08	428	1	1	1	1	1	1	1	–	–	–	–
Aac10	134	1	1	1	1	1	1	1	–	–	–	–
Aac11	302	1	1	1	1	1	1	1	–	–	–	–
Aac12[a]	229–247	4	2	1	2	3	2	3	0.06	0.31	0.81	0.27
Aac13	355	1	1	1	1	1	1	1	–	–	–	–
sMo00020[a]	242–250	3	1	1	1	3	2	3	0.03	0.09	0.67	0.08
sMo00027	293	1	1	1	1	1	1	1	–	–	–	–
sMo00055	273	1	1	1	1	1	1	1	–	–	–	–
sMo00130	276	1	1	1	1	1	1	1	–	–	–	–
sMo00137[a]	179–187	3	1	1	1	2	2	2	0.01	0.11	0.91	0.09
sMo00138	305	1	1	1	1	1	1	1	–	–	–	–
sMo00141	441	1	1	1	1	1	1	1	–	–	–	–
sMo00147	310	1	1	1	1	1	1	1	–	–	–	–
sMo00161	221	1	1	1	1	1	1	1	–	–	–	–
Aacu38[a]	316–346	6	4	2	2	4	7	6	0.13	0.64	0.80	0.58
Aacu45[a]	260–284	5	4	2	2	2	5	4	0.30	0.38	0.21	0.34
Aacu74[a]	278–313	9	1	2	2	6	8	6	0.26	0.45	0.42	0.42
N_t		38	17 (43.6)	14 (35.9)	14 (35.9)	27 (69.2)	33 (84.6)	30 (76.9)	–	–	–	–
N_a		5.4	2.4	2.0	2.0	3.9	4.7	4.3	–	–	–	–
N_e		–	3.5	2.8	2.4	4.2	6.8	4.3	–	–	–	–
P		–	3.99 (57)	3.99 (57)	4.97 (71)	5.95 (85)	3.99 (57)	7 (100)	–	–	–	–
Mean H_o, H_E, *F* and PIC		–	–	–	–	–	–	–	0.20	0.39	0.57	0.35

A, number of alleles per locus; H_o, observed heterozygosity; H_E, expected heterozygosity; *F*, inbreeding coefficient; PIC, polymorphic information content; N_t, total number of alleles over polymorphic loci and provenance (numbers in brackets are percentages); N_a, average number of alleles per polymorphic loci and per provenance; N_e, effective number of alleles per provenance; *P*, percentage of polymorphic loci in brackets (numbers outside brackets are numbers of polymorphic loci); Provenances: PA = Pará, PE = Pernambuco, PB = Paraiba, SP = São Paulo, MG = Minas Gerais, MS = Mato Grosso do Sul.

[a] Rows refer to the polymorphic loci.

(Aac04) with an average of 0.20 per locus, while expected heterozygosity (H_E) ranged from 0.09 (sMo00020) to 0.72 (Aac04) with an average of 0.39 per locus (Table 3). A range of positive inbreeding coefficients (F) were obtained from 0.15 (Aac04) to 0.91 (sMo00137) averaging 0.57 per locus, while PIC of the markers varied from 0.08 (sMo00020) to 0.68 (Aac04) with an average of 0.35 per locus (Table 3).

The polymorphism and heterozygosity levels of the seven polymorphic SSR markers that amplified the target microsatellite loci in *A. aculeata* were depicted in a figure (Fig. S1, available online). This figure demonstrated the allelic profiles of the loci amplified on 12 selected families of *A. aculeata* containing 48 accessions from the six provenances and elucidated the capability of the markers to distinguish between heterozygote and homozygote individuals in *A. aculeata*.

The test for evidence of null alleles identified the presence of null alleles across the loci presented in the form of allelic frequency (*p)* in which the intensity of the null alleles varied among the loci and the provenances analysed (Table 4). A complementary test for HWE showed significant deviations of the loci in certain provenances suspected of having null alleles (Table 4). The combination of the latter two test results showed that loci with a strong null allele frequency are significantly deviated from HWE (Table 4).

Discussion

The cross-amplification showed the ability of the polymorphic SSR markers to amplify the target microsatellite sequences in *A. aculeata*. The low percentage of the cross-amplification (26%) however could be attributed to the relatively wider taxonomic distance between the sources (*A. aculeatum* and *E. oleifera*) and the target (*A. aculeata*) species analysed in the study, since all the species are from different genera (Baker *et al.*, 2010). Nevertheless, when we look at the taxonomic relatedness between the species, both *A. aculeatum* and *A. aculeata* belong to the same sub-tribe (Bactridinae) and hence, in our investigation, provided a relatively higher percentage of cross-amplification (33%) when compared with *E. oleifera* (22%) (Table 3), which belongs to a different sub-tribe (Elaeidinae) (Baker *et al.*, 2010).

As cited by Rossetto (2001), the success in the cross-species amplification of any DNA sequence is inversely related to the evolutionary distance between two species. According to Rossetto (2001), in plants, there is a higher average rate of success in cross-species amplification of gSSR (genomic SSR) markers at subgenus (89.8%) and genus (76.4%) levels than at family

Table 4. Locus by provenance table of estimated null allele frequencies (p)[a] and Hardy–Weinberg deviation test

	PA		PE		PB		SP		MG		MS	
Locus	p	HW	p	HW	p	HW	P	HW	p	HW	p	HW
Aacu45	0.000	0.619ns	0.000	1.000ns	0.290	0.030*	0.000	0.142ns	0.129	0.000*	0.148	0.000*
Aacu38	0.419	0.000*	0.326	0.000*	0.290	0.000*	0.290	0.000*	0.307	0.000*	0.258	0.000*
Aacu74[b]	–	–	0.000	0.123ns	0.290	0.142ns	0.210	0.440ns	0.611	0.000*	0.409	0.000*
Aac04	0.000	0.840ns	0.000	0.213ns	0.000	0.370ns	0.194	0.860ns	0.172	0.000*	0.101	0.000*
Aac12[b]	0.326	0.010*	–	–	0.290	0.040*	0.638	0.000*	0.056	0.025*	0.206	0.000*
sMo00137[b]	–	–	–	–	–	–	0.816	0.000*	0.991	0.000*	0.935	0.000*
sMo00020[b]	–	–	–	–	–	–	0.896	0.000*	0.995	0.000*	0.887	0.000*

PA, Pará; PE, Pernambuco; PB, Paraíba; SP, São Paulo; MG, Minas Gerais; MS, Mato Grosso do Sul; ns, not significant.
*Significant deviation from HWE at $\alpha = 0.05$. [a]Frequency estimates (p) were based on the EM algorithm as described in the study by Dempster *et al.* (1977).
[b]No information (–) generated for loci with less than two alleles at the corresponding provenances.

(35.2%) level. Using 20 of the SSR markers tested in this study (sourced from *E. oleifera* sets of SSRs), a high cross-amplification percentage (83.3%) was reported when analysed with *Elaeis guineensis* accessions (Zaki *et al.*, 2012). Similarly, using the other 14 markers (sourced from *A. aculeatum* sets of SSRs), between 50 and 93% cross-amplification was reported in accessions of four *Astrocaryum* species (Ramos *et al.*, 2012). These high cross-amplification percentages in the latter two studies were due to the fact that the cross-amplifications were at genus and species levels, respectively.

Despite the low rate of cross-amplification obtained in the present study, the number of alleles amplified by the polymorphic markers was comparable with that in the study by Nucci *et al.* (2008), who first characterized polymorphic SSR markers for *A. aculeata*. Nucci *et al.* (2008) reported a range of two to eight alleles with an average of five per locus; while as already reported in this study, a range of three to nine alleles were detected by the markers with an average of 5.4 per locus (Table 3). Out of the total number of alleles, a range of 2.4–6.8 effective numbers of alleles (N_e) were obtained in the provenances with an average of four per provenance (Table 3). According to Staub (1994), N_e represents measures of the number of equally frequent alleles, which are maintained in a population and used to make proper sampling in several generations. Similar results were obtained for heterozygosity and PIC when compared with the study by Nucci *et al.* (2008) ($H_o = 0.27$; $H_E = 0.51$ and PIC = 0.48) (Table 3).

However, the number of alleles reported for the four SSRs (Aac04, Aac12, sMo00020 and sMo00137) was higher in other studies when the markers were analysed with accessions of *A. aculeatum* (average of six alleles per locus; Ramos *et al.* (2012)), and *E. oleifera* and *E. guineensis* (average of 7.5 alleles per locus; Zaki *et al.* (2012)). Similarly, average H_o, H_E and PIC per locus reported in Zaki *et al.* (2012) ($H_o = 0.63$; $H_E = 0.82$ and PIC = 0.48) and Ramos *et al.* (2012) ($H_o = 0.87$ and $H_E = 0.76$) were comparatively higher. As these markers were originally designed for the latter three species, a higher number of alleles, higher heterozygosity and PIC were expected than when analysing with *A. aculeata*. These explain the species, genome and locus specificity of the SSRs as already discussed by Powell *et al.* (1996) and Pinheiro *et al.* (2009); which increased the success of cross-amplification.

The average polymorphic loci ($P = 71.2\%$) per provenance obtained in our study imply that SSRs are powerful molecular markers for assessing genetic diversity in *A. aculeata* populations. Besides, the markers used in this study are informative based on Bostein *et al.*'s (1980) classification of SSRs using PIC values (Table 3). Nevertheless, the polymorphism level reported here is still relatively lower than a RAPD (random amplified polymorphic DNA) polymorphism (79%) obtained in the same species (Oliveira *et al.*, 2012). This could be explained by the lack of amplification products in some samples at certain SSR loci in the cross-amplification. Lack of amplification of an allele in certain accessions can be the result of null alleles potentially caused by PCR failure, the quality and quantity of DNA used, heterozygote deficits, scoring errors due to stuttering bands or other possible causes (Roa *et al.*, 2000; Dakin and Avise, 2004). Hence, null alleles can no longer be detected by the primer, leading to an influence on the polymorphism level of SSR markers.

Our tests for the null allele showed evidence for the presence of null alleles across the loci (Table 4). The null allele frequency varied highly among the loci across the provenances demonstrating that the effects of the null alleles are locus specific (Dakin and Avise, 2004). As we observed from our results, loci with high heterozygote deficits showed stronger null allele frequency than loci with less heterozygote deficits (Tables 3 and 4). According to Chakraborty *et al.* (1992), inbreeding causes significant heterozygote deficits relative to HWE that might be misconstrued as evidence for null alleles. Our test for HWE showed significant deviation by most of the loci across the provenances analysed (Table 4). As discussed in the following paragraph, heterozygote deficits were detected in most of the loci analysed, which were potentially caused by inbreeding. Hence, when we compare the two results of null allele and HW tests, loci with a strong null allele frequency deviated significantly from HWE owing to their heterozygote deficits, which are one of the potential causes of null alleles leading to low polymorphism (Dempster *et al.*, 1977; Chakraborty *et al.*, 1992).

Although *A. aculeata* has an apparently mixed mating system (Scariot *et al.*, 1991; Colombo *et al.*, 2013), the monoecious nature of its inflorescence could favour more inbreeding, which resulted in reduced average heterozygosity per locus (Table 3). This was explained by positive inbreeding coefficients ($F > 0$) obtained across all loci (Table 3). Substantial positive inbreeding coefficient indicates the presence of inbreeding in a given population (Hartl and Clark, 1997). A case study on jewelweed, *Impatiens capensis*, a common monoecious woodland flower in the Eastern US, demonstrated that the frequency of the homozygous genotypes increased with each generation of selfing, and the frequency of heterozygotes decreased (Stratton, 2008).

As described in the study by Zaki *et al.* (2012), amplification may not always represent functional SSRs for

target species from cross-amplification. The PCR products of the markers need to be cloned and sequenced with selected individuals and the amplicon sequence needs to be aligned with the original sequence from which the primers were designed. This may reveal the level of similarity and differences between the original gene sequence and the amplicon amplified by this specific locus. With this, markers can be identified based on sequence similarity in order to represent functional SSRs in the cross transferability. Therefore, it would be more appreciated if a further study will be conducted to confirm the inter-species transferability of the markers identified in this study by sequencing the amplicons and comparing with the original sequences from which the primers were designed.

We conclude that our investigation demonstrated the potential use of cross-species amplification to transfer SSR markers from two species of *Arecaceae* (*A. aculeatum* and *E. oleifera*) to the newly emerging potential oil palm species, *A. aculeata*. Along with the previously developed SSR markers (Nucci *et al.*, 2008), the SSRs identified in the present study (Aac04, Aac12, sMo00020, sMo00137, Aacu38, Aacu45 and Aacu74) could provide invaluable support to realize various studies towards the domestication of *A. aculeata*. They could be used to study population genetics, germplasm characterization, genetic improvement and conservation. The results also indicated that sourcing sets of SSRs from closely related taxa such as species or genus in *Arecaceae* could increase the success of transferring more number of functional markers. This would further facilitate comparative genetic studies between related species in *Arecaceae,* thus making the results directly comparative and the research more cost effective.

Supplementary material

To view supplementary material for this article, please visit http://dx.doi.org/10.1017/S1479262115000179

Acknowledgements

The authors would like to acknowledge Professor Carlos Nick for his technical help in obtaining leaf samples from Macaúba Active Germplasm Bank (BAG-Macaúba), Araponga, the State of Minas Gerais, Brazil. They are also grateful to Dra Telma, Renata D. Freitas and Éder Lanes for their unconditional assistance in Plant Biotechnology Laboratory, Universidade Federal de Viçosa. This work was financially supported by Petróleo Brasileiro S. A. (Petrobras); The academy of sciences for the developing world (TWAS) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

Abreu SI, Carvalho CR, Carvalho GMA and Motoike SY (2011) First karyotype, DNA C-value and AT/GC base composition of macaw palm (*Acrocomia aculeata*, Arecaceae) - a promising plant for biodiesel production. *Australian Journal of Botany* 59: 149–155.

Abreu AG, Priollig RH, Azevedo-Filho JA, Nucci SM, Zucchi SM, Coelho RM and Colombo CA (2012) The genetic structure and mating system of *Acrocomia aculeata* (Arecaceae). *Genetics and Molecular Biology* 35: 119–121.

Baker WJ, Norup MV, Clarkson JJ, Couvreur TLP, Dowe JL, Lewis CE, Pintaud JC, Savolainen V, Wilmot T and Chase MW (2010) Phylogenetic relationships among arecoid palms (Arecaceae: Arecoideae). *Annals of Botany* 108: 1417–1432.

Barbará T, Palma-Silva C, Paggi GM, Bered F, Fay MF and Lexer C (2007) Cross-species transfer of nuclear microsatellite markers: potential and limitations. *Molecular Ecology* 16: 3759–3767.

Bora PS and Rocha RVM (2004) Macaíba palm: fatty and amino acids composition of fruits. *Ciência e Tecnologia de Alimentos* 4: 158–162.

Bostein D, Whiter L, Skolnick M and Davis RW (1980) Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *The American Journal of Human Genetics* 32: 314–331.

Brito GG, Caixeta ET, Gallina PA, Zambolim EM, Zambolim L, Diola V and Loureiro ME (2010) Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. *Euphytica* 173: 255–264.

Chakraborty R, De Andrade M, Daiger SP and Budowle B (1992) Apparent heterozygote deficiencies observed in DNA typing data and their implications in forensic applications. *Annals of Human Genetics* 56: 45–57.

Chapuis MP and Estoup A (2007) Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24: 621–631.

Cole CT (2003) Genetic variation in rare and common plants. *Annual Review of Ecology, Evolution and Systematics* 34: 213–237.

Colombo CA, Nucci SM, Priolli RHG, Zucchi MI, Carhalho CRL, Chorfl BLH, Siqueira WG and Azvedo-Filho JÁ (2013) *Genetic Diversity of Macaw Palm* (Acrocomia aculeata) *by Microsatellite Markers*. San Diego, CA: International Plant and Animal Genome Conference. Available at http://www.intlpag.org.

Cruz CD (2013) *GENES V9.0. Software Package for Analysis in Quantitative Genetics and Experimental Statistics*. Universidade Federal de Viçosa, Brazil.

Cruz CD, Ferreira FM and Pessoni LA (2011) Biometria aplicada ao estudo da diversidade genética. Viçosa: Editora Universidade Federal de Viçosa, p. 622.

Dakin EE and Avise JC (2004) Microsatellite null alleles in parentage analysis. *Heredity* 93: 504–509.

Dempster AP, Laird NM and Rubin DB (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 39: 1–38.

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

Faleiro FG, Costa AM, Karia CT, Andrade RP, Junqueira NTV, Pereira AV, Pereira EB and Sano SM (2008) Molecular markers and geographic information systems as a tool to study native plant species in the Brazilian Savannas. *The*

2nd International Symposium in Tropical Savannas. 1217 October 2008, ParlaMundi, Brazilia. Available at http://www.cpac.embrapa.br.

Fortes ICP and Baugh PJ (1999) Study of analytical on-line pyrolysis of oils from Macauba fruit (*Acrocomia sclerocarpa* M.) via GC/MS. *Journal of the Brazilian Chemical Society* 10: 469–477.

Hartl DL and Clark AG (1997) *Principles of Population Genetics*. USA: Sinauer Associates, p. 542.

Hiane PA, Ramos-Filho MM, Ramos MIL and Macedo MLR (2005) Bocaiúva, *Acrocomia aculeata* (Jacq.) Lodd., pulp and kernel oils: characterization and fatty acid composition. *Brazilian Journal of Food Technology* 8: 256–259.

Kantety RV, Rota ML, Matthews DE and Sorrells ME (2002) Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant Molecular Biology* 48: 501–510.

Lanes ECM, Nick C, Kuki KN, Freitas RD and Motoike SY (2013) Genomic DNA isolation of *Acrocomia aculeata* (Arecaceae) from leaf and stipe tissue samples for PCR analysis. *Genetics and Molecular Research* 12: 3905–3911.

Lanes ECM, Costa PMA and Motoike SY (2014) Alternative fuels: Brazil promotes aviation biofuels. *Nature* 511: 31.

Motoike SY and Kuki KN (2009) The potential of macaw palm (*Acrocomia aculeata*) as source of biodiesel in Brazil. *International Review of Chemical Engineering* 1: 632–635.

Motoike SY, Lopes FA, Sá Júnior AQ, Carvalho M and Oliveira MAR (2007) Processo de Germinação e Produção de Sementes Pré-Germinadas de Palmeiras do Gênero *Acrocomia*. Submetido à Lei de Patentes. Protocolo INPI: 014070005335.

Moura EF, Motoike SY, Ventrella MC, Sá Junior AQ and Carvalho M (2009) Somatic embryogenesis in macaw palm (*Acrocomia aculeata*) from zygotic embryos. *Scientia Horticulturae* 119: 447–454.

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

Nucci SM (2007) Development, characterization and analyses of the use of microsatellite markers in macaw palm population genetics MSc Thesis, University of Sâo Paulo.

Nucci SM, Azevedo-Filho A, Colomo AC, Priolli GHR, Coelho MR, Mata TL and Zucchi IM (2008) Development and characterization of microsatellites markers from the macaw. *Molecular Ecology Resources* 8: 224–226.

Oliveira DA, Júnior-Melo AF, Brandão MM, Rodrigues LA, Menezes EV and Ferreira PRB (2012) Genetic diversity in populations of *Acrocomia aculeata* (Arecaceae) in the northern region of Minas Gerais. *Genetics and Molecular Research* 11: 531–538.

Pinheiro F, Palma-Silva C, Barros F and Cozzolino S (2009) Cross-amplification and characterization of microsatellite loci for the Neotropical orchid genus *Epidendrum*. *Genetics and Molecular Biology* 32: 337–339.

Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S and Rafalski A (1996) The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Molecular Breeding* 2: 225–238.

Ramos SF, Vasconcelos de Macêdo JL, Lopes MTG, Batista JS, Formiga KM, Pimentel da Silva P, Saulo-Machado AC and Veassey EA (2012) Microsatellite loci for Tucumã of Amazonas (*Astrocaryum aculeatum*) and amplification in other Arecaceae. *American Journal of Botany* 99: e508–e510.

Raymond M and Rousset F (1995) GENEPOP V1.2. Population genetics software for exact tests and ecumenicism. *Heredity* 86: 248–249.

Ribeiro LM, Souza PP, Rodrigues AG, Oliveira TGS and Garcia QS (2011) Overcoming dormancy in macaw palm diaspores, a tropical species with potential for use as bio-fuel. *Seed Science and Technology* 39: 303–317.

Roa AC, Chavarriaga-Aquire P, Dugue MC, Maya MM, Bonierbale MW, Iglesias C and Tohme JM (2000) Cross species-amplification of cassava (*Manihot esculenta*) (Euphorbiaceae) microsatellites: allelic polymorphism and degree of relationships. *American Journal of Botany* 87: 1647–1655.

Rossetto M (2001) Sourcing of SSR markers from related plant species. In: Henry RG (ed.) *Plant Genotyping: The DNA Finger Printing of Plants*. New York: CAB International, pp. 211–224.

Scariot AO, Lleras E and Hay JD (1991) Reproductive biology of the palm *Acrocomia aculeata* in Central Brazil. *Biotropica* 23: 12–22.

Staub JE (1994) Evolution and population. *Crossover: Concepts and Applications in Genetics, Evolution, and Breeding. An Interactive Computer-based Laboratory Manual*. USA: University of Wisconsin Press, p. 359.

Stratton D (2008) Case Studies in Ecology and Evolution (DRAFT). Non-random mating, Inbreeding and Population Structure, available at http://www.uvm.edu

Teixeira LC (2005) Potencialidades de oleaginosas para produção de biodiesel. In: Lacerda V (ed.) *Produção de Oleaginosas para Bioiesel*. Belo Horizonte: Informe Agropecuário, pp. 18–27.

Tickel J (2000) From the Fryer to the Fuel Tank: The Complete Guide to Using Vegetable Oil as an Alternative Fuel. Florida: Tickel Energy Consulting, p. 155.

Zaki NM, Singh R, Rosli R and Ismail I (2012) *Elaeis oleifera* genomic-SSR markers: exploitation in oil palm germplasm diversity and cross-amplification in Arecaceae. *International Journal of Molecular Sciences* 13: 4069–4088.