Conservation Genet Resour
DOI 10.1007/s12686-009-9058-z

TECHNICAL NOTE

Development and characterization of microsatellite markers for castor (*Ricinus communis* L.), an important oleaginous species for biodiesel production

Miklos Maximiliano Bajay · José Baldin Pinheiro · Carlos Eduardo Araújo Batista · Márcia Barreto Medeiros Nobrega · Maria Imaculada Zucchi

Received: 24 June 2009 / Accepted: 30 June 2009

Abstract Castor (*Ricinus communis L.*) is an important oleaginous plant from both economic and social points of view. The seeds contain an oil with excellent properties for industrial uses. This paper presents the main results of a study aiming to develop microsatellite markers for castor. Twelve new polymorphic microsatellite markers were isolated and characterized in 38 genotypes accessions from the castor germplasm of the Brazilian Agricultural Research Company (EMBRAPA). Knowledge on the genetic diversity of castor can be used to gain a better understanding on genetic diversity conservation, and germplasm management, guiding breeding programs and conservation strategies.

Keywords Germplasm · Genetic diversity · SSRs and conservation

M. M. Bajay · J. B. Pinheiro · C. E. A. Batista
Departamento de Genética, Universidade de São Paulo-USP, Av. Pádua Dias, 11, Vila Independência, C. P. 83, Piracicaba, SP 13400-970, Brazil

M. B. M. Nobrega
Empresa Brasileira de Pesquisa Agropecuária, Centro Nacional de Pesquisa de Algodão, OSWALDO CRUZ, 1143. CENTENÁRIO, C. P. 174, Campina Grande CEP 58428-095, Brazil

M. I. Zucchi (✉)
Centro de Pesquisa e Desenvolvimento em Recursos Genéticos Vegetais, Instituto Agronômico IAC, Av. Barão de Itapura, 1481, Centro 13020902, C. P. 28, Campinas, SP 13075-630, Brazil
e-mail: mizucchi@iac.sp.gov.br

Castor (*Ricinus communis* L., Euphorbiaceae, $2n = 20$), whose origin center is probably Ethiopia, is cultivated in many tropical and subtropical regions of the world (Govaerts et al. 2000). The seeds, leaves and stem of castor are poisonous when consumed by humans and livestock, because of the presence of ricin, the major toxic alkaloid present in the plant (Zimmerman et al. 1958). In the other hand, the seeds have a very high concentration (more 45%) of excellent oil, ideal for the production of biodiesel (Jeong and Park 2009). Is the only vegetable oil soluble in alcohol, presenting high viscosity and requiring less heating than others oils during the production of biodiesel (Beltrão 2003). Castor oil has many other uses ranging from cosmetics and, medicines to lubricants utilized in aircraft and space rockets.

Loss of genetic diversity is a problem in the case of some important species for agriculture, since ancient cultivars (or landraces) and wild relatives of domesticated species are being lost as modern varieties become adopted by farmers. This has led to calls for genetic conservation of crop germplasm (Frankel and Bennet 1970). Castor diversity is well represented in Brazilian germplasms collections, but there is little information on the molecular characterization of the accessions, and diversity loss is a serious risk. Managers of collections strive to accumulate and maintain the biological diversity of crops and other economically important plant species (Allan et al. 2007).

The development of microsatellite markers for this species would greatly facilitate the work of several researches involved in its germplasm diversity conservation. Aiming to provide researchers with these tools, in this work we provide 12 microsatellite markers for castor.

Genomic DNAs were extracted from fresh leaf samples using the CTAB method (Doyle and Doyle 1987). A microsatellite-enriched library was obtained using

Published online: 25 July 2009

Table 1 Characteristics of 12 microsatellite loci from *Ricinus communis L.* forward (F) and reverse (R) primer sequence, repeat motif, temperature of annealing number of alleles (N), product size range in base pairs, observed (H_o) and expected (H_e) heterozigosities, FIS, PIC and *P* value HWE

Locus	GenBank accession	Primer nucleotide sequence (5′–3′)	Repeat motif	T (°C)	Size range (bp)	No. of aleles	H_o	H_e	FIS	PIC	*P* value HWE
Rco2	GF100470	F: CTAGCTTTGGGGCACAGTC	$(AC)_{12}$	60	130–152	4	0.0526	0.1953	0.737	0.1881	0.0011*
		R: GGAAAATAGGTGCGTATGAAAC									
Rco3	GF100460	F: GATGTGAGCCCATTATGCTG	$(GA)_{22}$	60	230–240	2	0.0000	0.1884	1.000	0.1703	0.0000*
		R: TCAGAAATACCTCTAGGCGACA									
Rco8	GF100461	F: CGTGTGTCTGTGTGCATGTC	$(TG)_{10}$	60	280–288	4	0.1081	0.3579	0.705	0.3235	0.0002*
		R: CCTCAACCCTTTGCTGTTTC									
Rco11	GF100471	F: GCGTGGACTAACTTCAAGCA	$(TC)_{10}(GT)_6$	60	240–250	2	0.0789	0.3168	0.757	0.2663	0.0000*
		R: CCCCATTAGCATCGAGAAAG									
Rco12	GF100462	F: AAGACTGCACCTCCTCCTA	$(TG)_8 8(GA)_9$	62	220–224	2	0.1053	0.4114	0.750	0.3265	0.0000*
		R: TGCTGGAACAAACCCTGATA									
Rco13	GF100463	F: GGTGCTTCCAGAAATTCAGTT	$(GA)_{23}$	62	226–254	5	0.0833	0.7126	0.886	0.6597	0.0000*
		R: GGAGGGGAAAGACAGGATTC									
Rco15	GF100464	F: CACGCACGTTAAAGCAAACT	$(AG)_{18}$	60	220–230	4	0.0000	0.4945	1.000	0.431	0.0000*
		R: GCGAAGAAACCAAAATGGAG									
Rco22	GF100465	F: ATCCGCCGACAATAGCAG	$(AAAC)_3(AC)_9(TC)_5$	62	220–230	2	0.0789	0.2836	0.728	0.2433	0.0000*
		R: GCAACACTCTCTTCCCTGAA									
Rco23	GF100466	F: CATGGATGTAGAGGGTCGAT	$(GA)_{15}(AG)_8$	62	240–260	3	0.0526	0.6163	0.917	0.5427	0.0001*
		R: CAGCCAAGCCAAAGATTTTC									
Rco26	GF100467	F: TTGCTTGTCAAAGGGGAGTT	$(CT)_{19}$	62	260–274	5	0.0526	0.4865	0.895	0.4583	0.0000*
		R: TCATTTTGAGGGAGAAACCA									
Rco29	GF100468	F: GGAGAAAAGAAAGGGAGAAGG	$(GA)_7$	60	210–220	2	0.0000	0.2659	1.000	0.2307	0.0000*
		R: GCCAAAAGCACACTTAATTTGA									
Rco30	GF100469	F: TGAAACTTTGGAGCTTGGAGA	$(AG)_{19}$	60	220–240	5	0.1081	0.6709	0.843	0.6275	0.0000*
		R: GGTCCCACACATTCATACACA									

Forward (F) and reverse (R) primer sequence, temperature of annealing (T), number of alleles (N), observed (H_o) and expected (H_e) heterozigoties, FIS, PIC and *P* value HWE

* Departs significantly from HWE at $P < 0.05$ after Bonferroni correction

adapted protocols from Billotte et al. (1999). Genomic DNA of one genotype of *R. communis* was digested with Rsa I (Invitrogen), enriched in microsatellite fragments using $(CT)_8$ and $(GT)_8$ motifs. The enriched fragments were cloned into pGEM-T (Promega) and ligation products were used to transform Epicurian Coli XL1-Blue *Escherichia coli* competent cells. The positive clones were selected using the b—galactosidase gene and then grown overnight with ampicillin. A total of 96 clones were sequenced in an ABI 377 automated sequencer (PE Applied Biosystems) using BigDye terminator cycle sequencing kit (Applied Biosystems). About 12 pairs of primers were designed for SSR flanking regions using Primer 3 and tested in 38 morphologically divergent genotypes of *R. communis* from the germplasm collection of EMBRAPA. PCR reactions were performed in a 25 µl reaction volume of buffer [20 mM] Tris-HCl (pH 8.4), 50 mM KCl and 1.5 mM MgCl] containing 50 ng of genomic DNA, 0.8 lM of each primer, 150 lM dNTPs and 1 U Taq DNA polymerase (Invitrogen) using a BIO-RAD My Cicler termocycler. The PCR program consisted of an initial denaturing step at 94°C for 1 min followed by 35 cycles of amplification [94°C (1 min), 1 min at the specific annealing temperature of each primer pair (Table 1), 72°C (1 min)] and a final elongation step at 72°C for 10 min.

Amplification products were resolved by electrophoresis in 7% denaturing polyacrylamide gels and visualized by silver-staining (Creste et al. 2001). The allele scoring was done using the 10 pb DNA Ladder (Invitrogen) as size standards.

The number of alleles observed for each loci ranged from two to five, with an average of 3.3 alleles per locus.

Descriptive statistics and the test for Hardy–Weinberg Equilibrium were performed using Tools for Genetic Population Analysis (TFPGA) (Miller 1997). The observed and expected heterozygosities ranged from 0.000 to 0.108 (0.060 on average) and 0.188 to 0.712 (0.416 on average), respectively. The twelve loci depart significantly from Hardy–Weinberg Equilibrium ($P < 5\%$) after Bonferroni correction. However, results were obtained from a germplasm collection, which may not be in HWE. Indeed, there is absolutely no reason to expect HWE in germplasm collections, once genotypes are collected in different populations and, since the moment of sampling, they did not breed anymore. The low observed heterozygosity values suggest the predominance of autogamy in this species, which is know to have a mixed mating system, being both self- and cross-pollinated by wind (Allan et al. 2007). Hardy–Weinberg equilibrium is expected in panmitic populations. However, if the species is selfing (autogamous) or mixed mating system, HWE equilibrium will not be attained, population will reach Wrights' equilibrium (Ritland and Jain 1981; Weir 1996).

The linkage disequilibrium was tested adjusted *P*-value for 5% nominal level using the Fstat (Goudet 1995) and no disequilibrium was detected among all loci.

The development of molecular markers SSR for *Ricinus communis L.* is essential for the ongoing research on this culture in Brazil and worldwide. The generated information is of great importance for the identification, rational exploitation and conservation of the genetic variability of this species.

References

Allan G, Willians A, Rabinowicz PD, Chan AP, Ravel J, Keim P (2007) Worldwide genotyping of castor bean germplasm (*Ricinus communis L.*) using AFLPs and SSRs. Genet Resour Crop Evol. doi:10.1007/s10722-007-9244-3

Beltrão NEM, Melo FB, Cardoso GD, Severino LS (2003) Mamona: Árvore do Conhecimento e Sistemas de Produção para o Semi-Árido Brasileiro. Campina Grande: Embrapa – CNPA, 2003. 19 p (Embrapa CNPA. Circular Técnica, 70)

Billotte N, Lagoda PJR, Risterucci AM, Baurens FC (1999) Microsatellite-enriched libraries: applied methodology for the development of SSR markers in tropical crops. Fruits 54:277–288

Creste S, Tulmann Neto A, 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, Doyle JL (1990) Isolation of plant DNA fresh tissue. Focus 12:13–15

Frankel OH, Bennet E (1970) Genetic resources in plants—their exploration and conservation. F.A. Davis, Philadelphia

Goudet J (1995) FSTAT (version 1.2): a computer program to calculate F-statistics. J Hered 86:485–486

Govaerts R, Frodin DG, Radcliffe-Smith A (2000) World checklist and bibliography of Euphorbiaceae (with Pandaceae). Redwood Books Limited, Trowbridge, Wiltshire

Jeong GT, Park DH (2009) Optimization of biodiesel production from castor oil using response surface methodology. Appl Biochem Biotechnol 156:431–441

Miller MP (1997) Tools for Population Genetic Analysis (TFPGA), Version 1.3. Department of Biological Sciences, Northern Arizona University, AZ, USA

Ritland K, Jain S (1981) A model for the estimation of outcrossing rate and gene frequency using independent loci. Hereditary 47:35–52

Weir BS (1996) Genetic Data Analysis II: Methods for Discrete Population Genetic Data (Sinauer, Sunderland, MA)

Zimmerman LH, Miller MD, Knowles PF (1958). Castorbeans in California. Univ. Calif. Agr. Expt. Sta., Davis, Cir. 468:1–12