Biochemical Systematics and Ecology 59 (2015) 168–173

Contents lists available at ScienceDirect

Biochemical Systematics and Ecology

journal homepage: www.elsevier.com/locate/biochemsyseco

Development and use of novel microsatellite markers from double-enriched genomic libraries in Guatemalan *Jatropha curcas*

Ramon da Silva Raposo [a, 1], Fábio Barros Britto [b, 1], Isis Gomes de Brito Souza [c], Adriana Márcia Ferreira de Carvalho [c], Adilson Kenji Kobayashi [d], Bruno Galvêas Laviola [d], Paul Bentzen [e], Fábio Mendonça Diniz [c, *]

[a] *Universidade de Fortaleza, Núcleo de Biologia Experimental (NUBEX), Fortaleza, CE, 60.811-905, Brazil*
[b] *Universidade Federal do Piauí, Campus Professora Cinobelina Elvas, Caixa Postal 58, Bom Jesus, PI, 64.900-000, Brazil*
[c] *Embrapa Mid-North, CP: 01, Teresina, PI, 64.006-220, Brazil*
[d] *Embrapa Agroenergy, PqEB, Brasilia, DF, 70.770-901, Brazil*
[e] *Department of Biology, Dalhousie University, Halifax, NS, B3H 4J1, Canada*

ARTICLE INFO

Article history:
Received 14 November 2014
Accepted 24 January 2015
Available online 14 February 2015

Keywords:
Short tandem repeats
Energy crop
Physical nut
Genetic studies
Polymorphism

1. Introduction

Jatropha curcas is a perennial bush-tree from the Euphorbiaceae family and possibly native to Central America, Mexico or neighboring parts of the continent. The genus *Jatropha*, which contains approximately 170 known species, is found in almost all tropical and subtropical countries of the world (Fairless, 2007). At present, the species *J. curcas,* rich in oil content, has been seriously considered as a biofuel alternative. One of the major limitations for successful cultivation of *J. curcas*, however, is the selection of planting material due to its narrow genetic base (Ranade et al., 2008).

The development of molecular tools for monitoring and selecting natural populations for breeding and conservation programs has become of paramount importance with the increasing interest in *J. curcas* as a crop for oil seed production. The use of molecular markers to gather genetic information for the estimation of intra-specific diversity, the selection of

* Corresponding author. Tel.: +55 86 3198 0593; fax: +55 86 3198 0500.
E-mail address: fabio.diniz@embrapa.br (F.M. Diniz).
[1] These authors contributed equally to this work.

http://dx.doi.org/10.1016/j.bse.2015.01.019

prospective plant materials and the identification of useful traits in marked-assisted selection can be helpful in efficient breeding strategies (Kumar et al., 2011).

Among the different classes of markers microsatellite DNA, or Simple Sequence Repeats (SSRs), has proved to be of particular interest due to its high levels of inter- and intra-specific polymorphism. Microsatellite loci are present in nuclear and organellar DNA and consist of short stretches of tandemly-repeated DNA with 2–5 base pairs in length. The genetic variation is present in the number of repeats of certain microsatellite loci, a result of slippage of the DNA polymerase during replication (Tautz, 1986). Microsatellites have proven to be extremely useful as nuclear DNA tags and their efficiency is reflected in the fact that polymorphism can be even detected between closely-related individuals (Chen and Cheng, 2013). These multi-allelic markers are easy to score by polymerase chain reaction (PCR)/gel electrophoresis, and allelic variants of a microsatellite locus are codominant and show Mendelian inheritance (Jarne and Lagoda, 1996). So far, only dinucleotide microsatellites are available in the literature for *J. curcas*. However, tri and tetranucleotide repeat microsatellites have the advantage of giving less pronounced "stutter" bands in PCR, more reproducibility, and consequently easier to genotype. Therefore, these types of microsatellite markers may have higher applicability than dinucleotides in population genetic analysis (Ellegren, 2004).

Herein, we describe the isolation and amplification of 23 novel microsatellite loci from *J. curcas*, using a new double-enrichment hybridization procedure, and capture a snapshot of the genetic diversity of Guatemalan accessions referred to in the literature as a putative center of origin in Central America.

2. Materials and methods

2.1. Sample collection, tissue preparation and DNA extraction

Fresh young plant leaves of *J. curcas* were collected from the Biocombustibles de Guatemala S.A. (Guatemala City, Guatemala), stored into a plastic bag with silica gel to dehydrate and kept on ice. Approximately 20 mg of silica gel-dried young leaves were grounded in 2.0 mL microtubes containing ceramic grinding beads (CK28, BIOAMERICA) with two 16 s pulses at 5200 rpm with 10 s intervals on a Precellys®24 Tissue Homogenizer (BERTIN TECHNOLOGIES). High molecular weight DNA was extracted from these using the *DNeasy Plant Mini Kit* (QIAGEN) according to manufacturer's instructions. Extracts were checked for quality and quantity on a 1% agarose gel electrophoresis with ethidum bromide. DNA concentration was measured using NanoDrop 2000 Spectrophotometer (Thermo Scientific) as well. DNA was maintained at −20 °C until further analyses.

2.2. Genomic library construction and SSR isolation

Two genomic libraries enriched for tetranucleotide microsatellites were prepared as outlined in Glenn and Schable (2005) and Diniz et al. (2007) with modifications. In both libraries DNA (≈200 ng/μl) was digested separately with *Hin*cII, *Rsa*I, *Bst*UI and *Hae*III (NEW ENGLAND BIOLABS) in the presence of bovine serum albumin (BSA) overnight. Digests were ligated to double stranded SNX linkers (SNX-For: 5′-CTAAGGCCTTGCTAGCAGAAGC-3′ and SNX-Rev: 5′-pGCTTCTGCTAGCAAGGCCTTAGAAAA-3′) with T4 ligase (NEW ENGLAND BIOLABS) in the presence of *Xmn*I.

Linker-ligated inserts were amplified by a symmetric PCR with the SNX-For linker as primer. Cycle number was optimized to minimize over-amplification of products. PCR conditions were as follows: [1 × 95 °C for 5 min, 20 × (95 °C for 45 s, 62 °C for 1 min, and 72 °C for 2 min), 1 × 72 °C for 30 min]. Amplifications were carried out in 50-μL reaction volume containing 20–100 ng DNA, 1 × Thermopol buffer (+1.5 mM MgCl), 50 μM each dNTP, 0.5 U *Taq* DNA polymerase (Thermopol, NEW ENGLAND BIOLABS), 0.3–0.5 μM of SNX-For on an MJ Research DNA Engine Tetrad PTC-225 thermalcycler.

PCR products (linker-ligated DNA) were recovered using Qiaquick PCR Purification columns. They were enriched for repeats by subtractive hybridization with biotinylated tetranucleotide probes [i.e., Mix A: 2.5 μM $(ACTC)_6$; 2.5 μM $(ACTG)_6$; 2.5 μM $(ACCT)_6$ 2.5 μM $(ACAG)_6$ and Mix B: 2.5 μM $(AATG)_6$; 2.5 μM $(AAAC)_6$; 2.5 μM $(AATC)_6$; 2.5 μM $(AAAG)_6$] (Operon Technologies, Alameda, Calif.) bound to magnetic beads (DYNAL BIOTECH Inc., Lake Success, NY).

Enriched-DNA recovered from the beads was amplified with SNX-For linker to generate double stranded DNA. PCR conditions were similar to those described above, except the number of cycles, which was 10in this case. Amplified-enriched DNA from this hybridization/enrichment step was cleaned with Qiaquick purification columns with a final elution of 50 μL. Then, a second round of hybridization was performed using a small fraction of the post-enrichment amplified inserts, using the same probe mixtures and hybridization conditions described in this section. This double-enrichment procedure was performed to increase the chance of recovering SSR-containing sequences within DNA fragments (Diniz et al., 2007).

Amplified-enriched DNA from the double-enrichment procedure was cleaned with Purelink™ PCR purification kit (INVITROGEN) and ligated into Qiagen pDrive™ Vector (QIAGEN PCR CLONING KIT). The cloning vector was transformed into NEB 5-alpha competent *Escherichia coli* (DH5″ derivative; NEB) and plated on imMedia Amp Blue agar (Invitrogen), and then grown overnight at 37 °C. After the cloning procedure, positive colonies were identified in a white (positive) and blue (negative) screen. The positive colonies were transferred to Luria–Bertani (LB) broth with ampicillin (100 μg/mL) and allowed to grow for 15 hours with constant shaking (225 rpm). Then, 1.0 μL of LB broth containing positive bacteria was used in PCR amplifications of target sequences, using M13 forward (−20) and reverse (−40) primers. Plasmid DNA of appropriate length (500–1000 bp inserts) was prepared using a Qiaprep spin miniprep kit (QIAGEN) and cycle-sequenced [96 °C for 3 min, 40 × (96 °C for 20s, 52 °C for 20s and 60 °C for 4 min] in one direction with a universal primer (T7/M13F) using the BigDye