Industrial Crops and Products 49 (2013) 136–142

Contents lists available at SciVerse ScienceDirect

Industrial Crops and Products

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

Microsatellite polymorphism in *Jatropha curcas* L.—A biodiesel plant

Ramanuj Maurya, Astha Gupta, Sunil Kumar Singh, Krishan Mohan Rai, Chandrawati, Samir V. Sawant, Hemant Kumar Yadav*

CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow 226001, India

ARTICLE INFO

Article history:
Received 5 February 2013
Received in revised form 19 April 2013
Accepted 19 April 2013

Keywords:
Jatropha curcas
Enriched library
Genomic SSR
Biodiesel
Genetic diversity

ABSTRACT

We developed and characterized 1207 SSRs to enrich the validated markers repertoire of *Jatropha curcas*. A total of 248 polymorphic SSRs were identified with a panel of 7 accessions of *J. curcas* including some exotic accessions. Furthermore, 179 and 331 SSRs were found polymorphic among parental lines of NBRI-J05 × EC643912 and Chhatrapati × *Jatropha integerrima* used in developing two mapping population respectively. The number of alleles varied from 2 to 5 with an average of 2.24 ± 0.55 and 2.42 ± 0.62 alleles/SSR for CA and GA enriched library respectively. Most of the SSRs had lower PIC value (less than 0.30) and the maximum PIC value was observed for JGM_A281, JGM_A326 (0.63) followed by JGM_B300 (0.62), JGM_B361 (0.58), JGM_A244 (0.55), JGM_B595 (0.55) and JGM_B176 (0.55). The genetic similarity coefficient among the accessions of *J. curcas* and one accession of *J. integerrima* ranged from 0.11 to 0.92 with an average of 0.57 ± 0.31. The phenogram classified all the 7 accessions of *J. curcas* in one cluster and the *J. integerrima* remained as an out group. The BLASTX analysis of SSR containing sequences showed maximum similarity of 50% with *Ricinus communis* (Euphorbiaceae) followed by *Populus trichocarpa* (22%), *Vitis vinifera* (16%) and *Arabidopsis* spp. (4.5%). This study may enrich the validated repertoire of SSR markers in *J. curcas* and could be used in various genetic studies including construction of linkage map, diversity analysis, and QTL/association mapping.

1. Introduction

Jatropha curcas L., commonly known as physic nut, belongs to the family Euphorbiaceae and has small genome size of 416 Mb (Carvalho et al., 2008). It is reported to be native to tropical America, but also widely distributed in other tropical and sub-tropical areas of the world, especially in Africa, India and Southeast Asia (Heller, 1996; Rao et al., 2008). Traditionally, this plant is used as a fence plant, to prevent and/or control erosion and reclaim land. Recently, *J. curcas* has been projected as a promising and potential source of biodiesel as crude oil from its seeds meets the fuel quality of rapeseed (www.fact-foundation.com) which can be easily converted to biodiesel with US and European standards (Azam et al., 2005). However, *J. curcas* is still considered as undomesticated/semi domestic plant with various negative features with major knowledge gaps regarding basic genetics, ecological and agronomic properties (Achten et al., 2008; Fairless, 2007). Considering the importance of the crop, there is a need for a better understanding of various basic and applied aspects of this plant so as to improve or develop stable high yielding varieties. Traditional approaches for genetic improvement of polygenic traits mainly rely on phenotypic and pedigree information (Falconer et al., 1996) that are labor and time intensive. Therefore, the molecular breeding approach is being considered for genetic improvement for most of the economically important crop plants as faster alternatives. The basic requirements of any marker assisted breeding program to be successful include availability of reliable molecular marker system, and markers tightly linked to QTLs of desired traits. Among various DNA markers, SSRs and SNPs have been widely applied for estimation of genetic diversity, construction of linkage map and association/QTL mapping to tag the target traits. As far as *J. curcas* is concerned, a limited number of SSRs and SNPs marker are available as the information on quantitative genetics. Although, Sato et al. (2011) reported genome sequence of *J. curcas* and identified large number of SSRs, but the details of these SSRs are not available with the public domain. Further, they characterized only 100 SSRs (<0.3% reported) over 12 accessions. So, there is still need of larger numbers of validated markers. In recent past, various efforts have been made to develop and utilize different molecular markers in *J. curcas* which majorly include RAPD (Ganesh et al., 2008; Rafii et al., 2012), SPAR (Ranade et al., 2008), ISSR (Grativol et al., 2010; Kumar et al., 2009) and AFLP (Tatikonda et al., 2009; Shen et al., 2010) markers. However, some reports on SSRs and SNPs markers are also available (Sudheer et al., 2010; Yadav et al., 2011; Wang et al., 2011; Gupta et al., 2012a). Wang et al. (2011) have constructed linkage map of *J. curcas* with 216 SSRs and 290 SNPs using

* Corresponding author. Tel.: +91 522 2297938; fax: +91 522 2205836.
E-mail address: h.yadav@nbri.res.in (H.K. Yadav).

0926-6690/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.indcrop.2013.04.034

backcross mapping population derived from *J. curcas* and *Jatropha integerrima*. Later, Sun et al. (2012) used linkage map developed by Wang et al. (2011) and reported QTLs associated with growth and seed traits with deployment of 105 SSR markers. Till date, most of the markers studies in *J. curcas* have been performed with limited numbers of markers. For better understanding of polygenic traits, construction of dense map and fine QTL mapping a large number of workable and validated markers are required. Thus, in view of the previous report of low level of genetic diversity and limited number of validated markers, there is a need to enrich the genetic pool, develop and validate large number of polymorphic markers. Therefore, the present investigation was undertaken to develop a set of SSRs from genomic libraries and to thereafter characterize, validate and identify polymorphic SSRs in *J. curcas*.

2. Materials and methods

2.1. Plant materials and DNA isolation

The plant materials include 7 accessions of *J. curcas* i.e. Hansraj, Chhatrapati, CRIDA-C-16, RRL-Mon-C-1, NBRI_J05, EC685205 and EC643912 and 1 accession of *J. integerimma*. All these accessions (except EC685205 and EC643912) of indigenous origin and collected from different eco-geographical and agro-climatic zones of India. These are reported to be elite and diverse accessions of *J. curcas* from India (Yadav et al., 2011). The accessions EC685205 and EC643912 are exotic collections and were procured from South Africa and Mexico respectively. The Mexican accession is reported to be non toxic as it contains low amount of phorbol esters (Makkar et al., 1998). The NBRI-J05 × EC643912 and Chhatrapati × *J. integerrima* were used as parental lines for developing intraspecific and interspecific mapping population respectively at National Botanical Research Institute (NBRI), Lucknow, India. The DNA from young leaves was extracted using Qiagen DNeasy Plant Mini kit (Qiagen, Valencia, California) as per manufacturer's instructions. The quality of DNA was checked on 0.8% agarose gel and the concentration was determined using a Nanodrop spectrophotometer ND1000 (Nanodrop Technologies, DE, USA). Finally, the DNA was normalized to 10 ng/μl for PCR amplification.

2.2. SSR enriched genomic library and sequencing

Two genomic libraries enriched for CA and GA repeat motifs were custom produced by Genetic Identification Services (GIS, Chatsworth, CA, USA). The SSR enriched DNA fragments were cloned in *Hin*dIII restriction site of the plasmid vector pUC19. The recombinant plasmid was transformed into ElectroMax™ DH5α-E™ electrocompetent *Escherichia coli* cells (Invitrogen) and plasmid DNA was isolated from transformed cells following standard alkaline lysis mini prep protocol. The sequencing was carried out with M13 primer (5′ACGACGTTGTAAAACGACGG-3′) and Big Dye Terminator Cycle Sequencing Kit 3.1 on ABI 3730xl DNA Analyzer (Applied Biosystems, Foster City, CA, USA).

2.3. SSR identification, primer designing and PCR amplification

The sequence data obtained were checked manually and vector sequences were removed. The redundant sequences were identified by comparison using stand-alone BLAST (2.2.1.2) and removed. The unique sequences were subjected to SSR search by web based program SSRIT (http://www.gramene.org/db/markers/ssrtool). The basic search criteria for SSRs were a minimum of five repeat and maximum motif length was six. The primer pairs flanking the SSRs were designed using PRIMER3 software (http://frodo.wi.mit.edu/primer3). The primers were synthesized with an additional 18 base (5′-TGTAAAACGACGGCCAGT-3′) tag at 5′ end to all the forward primers as M13 tail (Eurofins, Germany). The PCR amplification was carried out in 10 μl reaction volume containing 10 ng of genomic DNA, 1× PCR master mix (Fermentas Inc, USA), 0.1 μl (5 pmol/μl) of forward primer (tailed with M13 tag), 0.3 μl (5 pmol/μl) each of both normal reverse and M13 tag (labeled with either 6-FAM NED, VIC and PET) using Applied Biosystems Veriti PCR machine. The PCR conditions was as follows: initial denaturation for 5 min at 95 °C, followed by 35 cycles of denaturation for 30 s at 94 °C, annealing for 45 s at 48–52 °C (primer specific) and extension for 30 s at 72 °C. Subsequently, 10 cycles of denaturation for 30 s at 94 °C, annealing for 45 s at 53 °C, extension for 45 s at 72 °C followed by final extension for 15 min at 72 °C was performed. After PCR amplification confirmation on 1.5% agarose gel, post PCR multiplex sets was prepared based on fluorescence labeled primers. For post PCR multiplexing, 1 μl of 6-FAM and 2 μl of each VIC, NED and PET labeled PCR product representing different SSRs were combined with 13 μl of water. 1 μl of this mixed product was added to 10 μl Hi-Di formamide containing 0.25 μl GeneScan™ 600 LIZ® as internal size standard, denatured for 5 min at 95 °C, quick chilled on ice for 5 min and loaded on ABI 3730xl DNA Analyzer. The fragment analysis was performed by GeneMapper v4.0 software (Applied Biosystems, Foster City, CA, USA).

2.4. Data acquisition and statistical analyses

The allelic data of polymorphic SSRs were subjected to statistical analysis using PowerMarker (Liu and Muse, 2005) to calculate observed heterozygosity (Ho), gene diversity or expected heterozygosity (He), major allele frequency and polymorphic information content (PIC) value. The PIC value was calculated following Botstein et al. (1980) as follow:

$$\mathrm{PIC} = 1 - \left[\sum_{i=1}^{n} Pi^2\right] - \left[\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} 2Pi^2Pj^2\right]$$

where *Pi* and *Pj* are the frequencies of *i*th and *j*th allele.

Further, pair-wise genetic similarities among all the accessions using Jaccard's coefficient was also calculated and a dendrogram was prepared based on unweighted pair-group method of arithmetic average (UPGMA) using NTSYS-pc v.2.02e software (Rohlf, 2000).

2.5. Functional annotation and GO analysis

The SSR containing sequences were annotated against NCBI nr protein database (NCBI nr, release: 20th Dec, 2011) using BLASTX with a criteria of minimum e-value of 1e−5 and minimum alignment length 50% of the query sequence. The annotations were further classified on the basis of their plant specific associations. The Gene Ontology (GO) analysis (http://arabidopsis.org/tools/bulk/go/index.jsp/), for functional annotations, was performed on the basis of TAIR GO annotations (ftp://ftp.arabidopsis.org/home/tair/Sequences/blast_datasets/TAIR10_blastsets/TAIR10_pep_20101214). The GO terms associated with *Arabidopsis* loci (best BLASTX hit) were assigned for annotations of corresponding sequences and categorized under molecular function, biological process and cellular component categories.

3. Results

3.1. Characterization and polymorphism evaluation of dinucleotide SSRs

The SSRs were developed from genomic libraries enriched with CA (designated as Lib A) and GA (designated as Lib B) repeat units.