Genetic variability revealed by microsatellite markers in a germplasm collection of *Jatropha curcas* L. in Brazil: an important plant for biofuels

Romero de Lima Souza¹, Luiz Antônio dos Santos Dias¹*, Thais Roseli Corrêa², Eveline Teixeira Caixeta¹,³, Erika da Costa Fernandes¹, Dandara Rêgo Muniz¹, Lucas Barbosa de Castro Rosmaninho¹ and Poliane Marcele Ribeiro Cardoso¹

**Abstract:** We evaluated the genetic variability of a collection of *Jatropha curcas* germplasm, represented by 93 accessions, using microsatellite markers. Among the 60 markers tested, five of them detected polymorphisms, with a total of 11 alleles and mean of 2.2 alleles per loci. These five markers enabled the quantification of genetic variability through estimates of expected (He=0.42) and observed (Ho=0.64) heterozygosity, Shannon-Weaver index (H’=0.62), coefficient of inbreeding (f=-0.44) and the formation of 11 clusters. Simultaneously, 14 accessions randomly sampled among the 93 and represented by seven plants each, were analyzed with these same five markers to quantify the within and between variability. Most of the genetic variation (92.58%) was contained within the accessions. These analyses revealed, for the first time, expressive genetic variability to be explored in this collection. The accessions UFVJC 05, 07, 12, 18, and 53 presented expressive variability among them with potential for the constitution of a base population for the breeding program.

**Keywords:** Variability, molecular markers, germplasm bank, breeding.

**INTRODUCTION**

Oil price volatility, combined with the need to reduce greenhouse gas emissions, has boosted global demand for biofuels. Among the potential species for the production of biofuels, *Jatropha curcas* L. deserves to be highlighted, as its seeds contain high oil content (36.2%) with the best quality (Freitas et al. 2016). These aspects have led to rapid expansion of cultivated areas and demand for improved cultivars (Sorrel et al. 2010).

Despite its great potential, *J. curcas* is still a species undergoing domestication, but with considerable variability to be explored. However, information regarding genetic variability and population structure is still limited (Bressan et al. 2012) and breeding programs are rare compared to other oilseed species (Dias et al. 2012, Pecina-Quintero et al. 2014).

The initial phase of a breeding program involving any kind of plant species requires a germplasm collection, functioning as a repository of genes for the future development of varieties. Therefore, the success of an improvement program depends on an efficient choice of accessions.
program depends on knowledge of the available genetic variability, which will allow the efficient selection of different genotypes to produce hybrids and similar genotypes to produce lines (Pecina-Quintero et al. 2014).

Molecular markers are suitable tools for the characterization of genetic variability in germplasm collections. Microsatellites, also known as Simple Repeated Sequence (SSR) markers, have high reproducibility, codominant inheritance and high polymorphism. Studies involving these markers have been successfully used to characterize the genetic variability of *J. curcas* (Pamidimarri et al. 2008, Sun et al. 2008, Sudheer et al. 2009, Rosado et al. 2010, Wen et al. 2010, Bressan et al. 2012, Na-ek et al. 2011, Pecina-Quintero et al. 2014, Sinha et al. 2015, Santos et al. 2016, Vásquez-Mayorga et al. 2017, Gangapur et al. 2018). In all of these studies, the genetic variability revealed for this species has been considered low.

The present study aimed to quantify the genetic variability present in a germplasm collection of *J. curcas*, composed of 93 accessions, using microsatellite markers. This study is part of the strategy to identify superior genotypes for cultivar development.

**MATERIAL AND METHODS**

**Plant material**

The germplasm collection of *J. curcas* of the Federal University of Viçosa (UFV), located in the Araponga Experimental Farm (lat 20º 39’ S, long 42º 32’ W and alt 823 m asl), in the municipality of Araponga, MG, Brazil is composed of accessions originating from different Brazilian geographic regions and from abroad (Table 1), all propagated by seeds. Currently, this collection is composed of 93 accessions (1504 plants) and is installed in modules of five trials in randomized block design, with four replications and 4-plant plots in 2 x 2 m spacing, with two common controls (Freitas et al. 2011, Freitas et al. 2016).

Samples of young and fully developed leaves from the 93 accessions were collected, each with seven sub-samples (plants), totaling 658 plants sampled. The leaves were wrapped in identified aluminum foil and placed in styrofoam boxes with ice for transportation to the Federal University of Viçosa, where they were stored at -80 °C.

**DNA extraction**

The DNA extraction process was conducted at the Laboratory of Forest Pathology of the UFV, based on the protocol for eucalyptus, modified from Doyle and Doyle (1990). This modification was due to the maceration of the samples in the following the steps: after removing of the central and secondary veins, the leaves were placed in 2 mL microcentrifuge tubes with metal beads and 700 μL extraction buffer (100 mM Tris-HCl pH 8.0, 20 mM EDTA, 1.4 M NaCl, 2% (w/v) CTAB, 2% (w/v) PVP and 0.4% (v/v) β-mercaptoethanol, the latter separated from the other components). The samples were macerated with Tissuelyser II (Qiagen) and incubated in a water bath at 65 °C for 30 minutes. After incubation, 500 μL of chloroform-isooamyl alcohol (24:1) was added to the tubes, which were manually inverted several times. The tubes were then centrifuged at 12,000 rpm for 5 min. The supernatants were transferred to new tubes and the extraction with chloroform was repeated. The supernatants (~ 500 μL) were transferred to fresh tubes and 0.9 volumes (450 μL) of cold isopropanol were added. The precipitated DNA was washed twice with 500 μL of cold 70% and 95% ethanol. The DNA was dried at room temperature for 1 hour and dissolved in 50 μL of TE (10 mM Tris and 1 mM EDTA, pH 8.0) plus RNAse (10 μg mL⁻¹) for 2 hours at 37 °C and then stored at -20 °C. DNA quantification was performed using a Thermo Scientific’s NanoDrop spectrophotometer and standardized with final concentrations adjusted to 10 ng mL⁻¹.

**Microsatellite molecular markers**

Sixty pairs of microsatellite primers were tested (see Table 2). All primers used were previously reported in the literature surrounding *J. curcas*, some of them drawn based on microsatellite loci derived from genomic sequences and specific ESTs of the species by Bressan et al. (2012), partial genomic sequences developed initially for cassava by Wen et al. (2010), and a number developed by Sudheer Pamidimarri et al. (2009) for differentiation of *J. curcas* toxic and non-toxic genotypes. PCRs (Polymerase Chain Reaction) were performed in a volume of 20 μL containing 50 ng DNA sample, 1x Taq DNA polymerase buffer, 100 μM of each dNTP, 1.5 mM MgCl₂, 0.2 μM of each primer and 1.0 U Taq DNA polymerase (Life Science). Amplifications were performed in an MJ Research PTC 100 thermocycler with denaturation at 94 °C for 3 minutes, followed by 40 cycles of denaturation at 94 °C for 1 min, annealing for 1 min at the specific
temperature of each primer and extension at 72 °C for 1 min. Final extension was performed at 72 °C for 8 min. PCR products were separated on a 6% denaturing polyacrylamide gel and visualized by silver staining solution, according to Creste et al. (2001).

Table 1. Provenance of 93 accessions of *Jatropha curcas* L. evaluated

| Accessions | Location of collection |
|------------|------------------------|
| UFVJC 1    | Santa Vitória, MG      |
| UFVJC 3    | Santa Vitória, MG      |
| UFVJC 4    | Santa Vitória, MG      |
| UFVJC 5    | João Pinheiro, MG      |
| UFVJC 6    | João Pinheiro, MG      |
| UFVJC 7    | João Pinheiro, MG      |
| UFVJC 8    | João Pinheiro, MG      |
| UFVJC 9    | João Pinheiro, MG      |
| UFVJC 10   | João Pinheiro, MG      |
| UFVJC 11   | João Pinheiro, MG      |
| UFVJC 12   | João Pinheiro, MG      |
| UFVJC 13   | Taurá, MG              |
| UFVJC 14   | Ólhos D’água, MG       |
| UFVJC 15   | Veredas, MG            |
| UFVJC 16   | Jaíba, MG              |
| UFVJC 17   | Montalvânia, MG        |
| UFVJC 18   | Cana Brava II-Montalvânia, MG |
| UFVJC 19   | Cana Brava II-Montalvânia, MG |
| UFVJC 20   | Poções, MG             |
| UFVJC 22   | Ipatinga, MG           |
| UFVJC 23   | Don Lara-Caratinga, MG |
| UFVJC 24   | Santa Luzia-Caratinga, MG |
| UFVJC 25   | Imbé-Caratinga, MG     |
| UFVJC 28   | Poté, MG               |
| UFVJC 29   | Poté, MG               |
| UFVJC 30   | Poté, MG               |
| UFVJC 31   | Poté, MG               |
| UFVJC 32   | Poté, MG               |
| UFVJC 33   | Itaipé, MG             |
| UFVJC 34   | Itaipé, MG             |
| UFVJC 35   | Ervália, MG            |
| UFVJC 36   | Serra da Ipiapaba, MG  |
| UFVJC 37   | Janaúba, MG            |
| UFVJC 38   | Petrolina, PE          |
| UFVJC 39   | Natal, RN              |
| UFVJC 40   | Formoso, TO            |
| UFVJC 41   | Jales, SP              |
| UFVJC 42   | Dourados, MS           |
| UFVJC 43   | Matozinhos, MG         |
| UFVJC 44   | São Carlos, SP         |
| UFVJC 45   | Barra do Bugre, MT     |
| UFVJC 46   | Barra do Bugre, MT     |
| UFVJC 47   | Pirajai, SP            |
| UFVJC 48   | Getulina, SP           |
| UFVJC 49   | Bocaíuva, MG           |
| UFVJC 50   | Bocaíuva, MG           |
| UFVJC 51   | Rio Pompa, MG          |
| UFVJC 52   | Barbacena, MG          |
| UFVJC 53   | Barbacena, MG          |
| UFVJC 54   | Barbacena, MG          |
| UFVJC 55   | Janaúba, MG            |
| UFVJC 56   | Janaúba, MG            |
| UFVJC 57   | Janaúba, MG            |
| UFVJC 58   | Janaúba, MG            |
| UFVJC 59   | Janaúba, MG            |
| UFVJC 60   | Pompéu, MG             |
| UFVJC 61   | Rio Grande do Sul      |
| UFVJC 62   | Juiz de Fora, MG       |
| UFVJC 63   | UFVJC 64               |
| UFVJC 65   | Desconhecido           |
| UFVJC 66   | Desconhecido           |
| UFVJC 67   | Desconhecido           |
| UFVJC 68   | São Luiz, MA           |
| UFVJC 70   | Ariquemes, RO          |
| UFVJC 71   | João Pinheiro, MG      |
| UFVJC 72   | Camboja                |
| UFVJC 73   | Camboja                |
| UFVJC 74   | Camboja                |
| UFVJC 75   | Bomfim, MG             |
| UFVJC 79   | Jordânia, MG           |
| UFVJC 80   | Desconhecido           |
| UFVJC 81   | Jordânia, MG           |
| UFVJC 82   | Jordânia, MG           |
| UFVJC 83   | Araras, SP             |
| UFVJC 84   | Petrolina, PE          |
| UFVJC 85   | Jataí, GO              |
| UFVJC 86   | Jequié, BA             |
| UFVJC 87   | Jequié, BA             |
| UFVJC 88   | Jequié, BA             |
| UFVJC 89   | Jitaúna, BA            |
| UFVJC 90   | Ipiáu, BA              |
| UFVJC 91   | Apuarema, BA           |
| UFVJC 92   | Itaitê, BA             |
| UFVJC 93   | Itaitê, BA             |
| UFVJC 94   | Andarai, BA            |
| UFVJC 95   | Andarai, BA            |
| UFVJC 96   | Mucugê, BA             |
| UFVJC 97   | Mucugê, BA             |
| UFVJC 98   | Iraquara, BA           |
| UFVJC 99   | Iraquara, BA           |
| UFVJC 100  | Souto Soares, BA       |
| UFVJC 101  | Wagner, BA             |
| UFVJC 102  | Desconhecido           |

*Accessions in bold were utilized to quantify the within and between variability.*
Variability between and within accessions
To quantify the genetic variability among the 93 accessions, one individual was selected from the seven collected, based on the highest concentration and the best DNA quality. After that, the most polymorphic primers were selected to analyze the variability between and within the accessions. To quantify the genetic variability between and within accessions, DNA from seven plants was extracted from 14 accessions randomly selected from the 93 accessions (Table 2).

Statistical analyses
The markers were coded as codominant, assigning numbers to the alleles. Thus, when a locus presented three alleles, the codes 11, 22, 33 were attributed to the homozygotes and 12, 13, and 23 for the heterozygotes. Popgene software version 1.31 (Yeh et al. 1999) was used to estimate the genetic variability statistics such as allele frequencies, number

---

**Table 2. SSR primers tested in the evaluation of the genetic variability of *Jatropha curcas* L. accessions**

| Name      | Forward primer                      | Reverse primer                      |
|-----------|-------------------------------------|-------------------------------------|
| JCENA 27  | CATTATCTCCTGACAGGCCTAC              | GTATTCTCCACACGCACT                  |
| JCENA 41  | CATTTCCTCCTCCTCTCTCT                | AAAAGGCAAGCAATCACTGGA               |
| JCENA 47  | GGCAGAGGCTTCCTCTAAGGGT              | CCAAGAGAATAGAGGAATGGC               |
| JCENA 87  | ATCGAGAGGAACCAAGACAGG               | CACAGGTAGAGCAGGAGACT                |
| JCD 10    | CATGCAAGATGTCATGAGTGG               | CACAACAGGAAACTGTTGGTGGCA            |
| JCD 24    | GATGTAGAATCGCTTGGGCAAG              | TCATTGATGTTGATGGTGTATGG             |
| JCD 41    | AACACCTAGGGCCACAGGT                 | TCATTGATGTTGATGGTGTATGG             |
| JCD 58    | TCGTAGTTGTTGTTGCAAC                 | AGGCTAGCTGTGAACGTTG                 |
| JCD 66    | CTCAGAGTGATGAGATTG                | TCTCCACTAAGAGCTGTTGGGCA            |
| JCP 6     | CAAAGAGGCTTGCCTTGG                 | AGGCTAGCTGACATTAGTGC               |
| JCP 9     | GTACTCTGAGATCTCTCTGAAACTAAGCAG     | TATCTCTTGTCAAGAAAAATGGAT            |
| JCP 20    | ACAGCAAGAGTGCAACAAACTCCTA          | TACCTGAGATGATGGCATG                 |
| JCMS 30   | GGGAAGAGGCTTGGGC                   | TATCTCTTCACATAAAATCAGTCA           |
| CESR 0163 | ACACGAGATGTTGTAAGT                 | GAAAGAAGCAACAGAAGATG                |
| CESR 0231 | GCTTTAACAAACCAAGACTC               | AATCTGATGTTGACTGTGGA                |
| CESR 0235 | TGCAGATGACGACACTT                   | ACCTGATGTTGACTGTGGA                |
| CESR 0290 | TCAATCTAAACACAAACAC                | CTAAATCTACACAGAGTGGG               |
| CESR 0293 | CAAGAGACAGACTGTTCAGTC              | TATGACCTACAAAGGGTTGCA              |
| CESR 0303 | AACGGCTTATCTGGCAATGGA               | TACTGCTCTGTCCTGATC                  |
| CESR 0312 | AAGAAACAAACAAATCTG                 | TAAATCTGATAGCGGTG                   |
| CESR 0333 | ACATCTCAGATGAGCTG                   | TAATGACCTACAGGTCGTG                 |
| CESR 0361 | AGGCAGATATGTGTAAGTTA               | AGGCAAGATATGATGTAAGT                |
| CESR 0370 | ATCAATGAACTGTGGAGG                  | TCCAGACTTCCAAACTTACCA              |
| CESR 0382 | ATGACCTACAAAGGAGCT                 | CCAATGACACCTGCGAAGTT               |
| CESR 0386 | ATTACTGTGAGAAGAG GCCA              | AGACAGCCACCTACAGATAA                |
| CESR 0398 | CAATCTCTGTGGCTGTCGA                 | CAAGACACTGAAACCTACGTG               |
| CESR 0399 | CAATCTATCTGGCAGTTCG                 | CTAATGACAAATCTGAGGCCA              |
| CESR 0494 | GCCAAAGGCTTCTGAAACA                 | TAAATCTGATAGCGGTG                   |
| CESR 0498 | GCTGGTACATGGCTCACTC                 | TAAATCTGATAGCGGTG                   |
| CESR 0718 | ATGAGTGATGAGAGGCTG                  | ATGCAGCTTCTGTCCT                  |
| CESR 0719 | ATGAGTGATGAGAGGCTG                  | GCAAGAATCTGAGAGAGCTG                 |
| CESR 0733 | CAAACAGATCAAGACGAGGTGGA             | AAATTGATGGTGGTGATG                  |
| CESR 0756 | CAGGTGGTCTCCTTCTTACT                | ATATATGACACCTGCGAAGTT              |
| CESR 0804 | GCTGTATGATGATGATGATGATGCTT         | CTAAATCTGACATTACTGGA               |
| CESR 0820 | GCTGATGATGATGATGATGATGCTT         | CTAAATCTGACATTACTGGA               |
| CESR 0844 | GAAATAGAAGAGAAGGAGGCA              | AAATTGATGGTGGTGATG                  |
| CESR 0889 | GCCTGTGATGATGATGATGATGCTT         | TCGAGTATGATGATGATGATG               |
| CESR 0934 | GTATTATCTCGCCACTTGT                | TTTGTATTTGGTTCACACGGA              |
Genetic variability revealed by microsatellite markers in a germplasm collection of *Jatropha curcas* L. in Brazil: an important...

of alleles (na), number of effective alleles (ne), observed (H₀) and expected (Hₑ) heterozygosity, Shannon-Weaver index (H'), Nei's (1972) genetic distance of and the coefficient of inbreeding (f). Analysis of molecular variance (AMOVA, Excoffier et al. 1992) and polymorphism information content (PIC, Botstein et al. 1980) were performed using Genes software (Cruz 2013). For the construction of a circular dendrogram, the Mega7 (Kumar et al. 2016) software was used to perform the UPGMA clustering algorithm from Nei's (1972) genetic distance. Interpretation of the dendrogram was conducted taking into consideration high-change points of cluster fusion.

**RESULTS**

Among the 60 microsatellite markers tested in the 93 accessions, five were able to detect polymorphisms (Table 2, underlined); therefore, these were used in the analysis of genetic variability.

**Genetic variability among the 93 accessions**

The locus CESR 0756 was the only one among the five analyzed which allowed the detection of three alleles in the 93 accessions. The locus JCENA 87 enabled the detection of less polymorphism according to allele frequency, although two alleles were detected, similar to the JCD 10, SSRY 107 and SSRY 127 loci. In JCENA 87, the allele A₁ frequency was almost 100%. The five microsatellite loci used to evaluate the accessions generated a total of 11 alleles (na), with an average of 2.2 alleles per locus. The number of effective alleles (ne) ranged from 1.01 to 2.71, with a mean of 1.91 (Table 3). Because the A₂ allele of locus JCENA87 presented a frequency lower than 0.05 (P < 0.05) it can be considered to be rare, with occurrence in only one individual of the accession UFVJC18.

To quantify the genetic variability between 93 accessions, estimates of expected (He) and observed (Ho) heterozygosity are important. Ho values were found to be ranging from 0.02 to 0.92, with an average of 0.64. He values spanned 0.02 to 0.64, with a mean of 0.42 (Table 3), which indicated a possible heterozygous origin of the collection accessions. For all loci (except JCENA 87), Ho was higher than He, revealing an excess of heterozygotes relative to that expected when in

| Name     | Forward primer         | Reverse primer         |
|----------|------------------------|------------------------|
| CESR 1041| TTGCTGAAGCCCTTTCTAT    | CAGTGTGGAGATGATACGCA   |
| CESR 1042| TGGATCTCCCTAGAAACAC    | TTGGTCGTCAATCTTCTGC   |
| CESR 1044| TGGTGAAGCTAAGGATTTCC   | CCATCTCTTTCTCTTTTG    |
| CESR 1050| TTTGACACTACAGTGGC      | ATAAAACCTCAAGGAGAAG   |
| CESR 1055| TTTGAAGACTGCAATACC     | GTCACAAGACCGACAGCT    |
| NS. 260  | TCAAGCTGTATGTTGAGTCA   | AGGGAGGAACACCTTCTCTA  |
| NS. 308  | GGAATGGTGTATGCTTCTCC   | GCAGTTGAGCTTCTCACAA   |
| NS. 689  | AGGATGAGTGAGACAAAGA    | CAGACTGGAATGCAATTTTC  |
| NS. 720  | CCATTACTACATGGAGCTTTCC | GGAAATTGTTATGTCCTTCT  |
| SSRY 4   | ATAGAGGAGAAGTGGAGCCG   | CTAAGCAGAAGCTACTGGA   |
| SSRY 7   | TGCTTAAGGAAATTCATTCT   | TGCTAAGCTGATCGACT    |
| SSRY 107 | CCATTCTTCTCTCTTCCTCA   | TGTTGAGGCTCTATAAATCT  |
| SSRY 112 | CGCAAGCTAAATCGGAGCTA   | ACAATCAAGAGCTCTGTAATC |
| SSRY 113 | TTGGCTGAAGCTGCCACAATA  | TCAACATTGACTGAGACAG   |
| SSRY 127 | CTCGGGTCTTACAAGAAGGA   | GCTGAACTGCTTGGCAACT   |
| SSRY 133 | AGCATGCTATTGCCAACC     | GCAGTGTACAGAAATGCTC   |
| SSRY 146 | TTCCCTGCTGAACTTCTGTC   | CTATTGGACCTCTTCGCGG   |
| SSRY 150 | CAATGAGGTGAAGTGAATACC  | AGGGTGCTCTTGAGAAAGG   |
| SSRY 151 | AGTGAAGAAATGCGATTGATG  | CCCATATATCGTTGAGTT    |
| SSRY 153 | TTCCGAGAAGACTTCCGTTCA  | CTAACTACTGACTGACTC    |
| SSRY 159 | CTATGCTTGGGCTCTTCACC   | GACATGACTAGGGAGACACA   |
| SSRY 177 | ACCCAAGGAAATCGGACAGAG  | CACCCAATCTACCTACTACA   |
Hardy-Weinberg equilibrium. For genetic improvement purposes, this result demonstrated the presence of heterozygotes able to be explored in this collection.

A parameter of great importance in assessing genetic variability in populations, by measuring the level of homozygosity, is the coefficient of inbreeding ($f$). It is essential to verify the existence of crossing among related individuals. Here, $f$ values ranged from -0.005 to -0.85 (Table 3). Notably, negative $f$ values are interpreted as null inbreeding, suggesting that there were no crosses between related individuals in the collection.

The Shannon-Weaver index ($H'$) presented values ranging from 0.03 to 1.05, and the mean value found was 0.62 (Table 3). This mean value of 0.62 reveals the existence of high genetic variability in the collected accessions of *J. curcas*, more than sufficient for the continuity of the breeding program.

All 93 accessions were clustered by the UPGMA method (Figure 1) and the formation of the following 11 clusters was observed: cluster I (orange, 23 accessions), cluster II (purple, 15 accessions), cluster III (pink, 10 accessions), cluster VI (blue, 14 accessions), cluster V (yellow, 4 accessions), cluster VI (dark green, 5 accessions), cluster VII (red, 4 accessions), cluster VIII (light green, 6 accessions), cluster IX (dark blue, 10 accessions) and two single clusters (X and XI) formed by accessions UFVJC 53 and UFVJC 7, respectively, which stood out from the others and were collected in Barbacena (MG) and João Pinheiro (MG), respectively. The accessions UFVJC 61 and 70 collected in Rio Grande do Sul (RS) and Ariquemes (RO) were grouped in the same cluster II.

### Genetic variability within and between 14 accessions

The same five polymorphic SSR loci were used on 14 accessions, randomly selected on 93, for evaluating the variability within and between them. A mean of 1.86 alleles per locus was found. The analysis of polymorphism information content (PIC) allowed quantification of the genetic polymorphism of each locus in the accessions evaluated. The highest PIC value observed was 0.35 for the locus SSRY 107, while the lowest value occurred for the locus JCENA 87 which did not distinguish within and between accessions (Table 3). The mean value of PIC (0.20) indicated a moderate level of polymorphism of the analyzed loci, with the exception of the locus JCENA 87.

Values of $H'_e$ ranged from 0 for the locus JCENA 87 to 0.46 in the locus SSRY 107, with mean of 0.25. $H'_o$ ranged from 0 for the locus JCENA 87 to 0.79 at the locus SSRY 107, with an average of 0.42 (Table 3).

The proportion of variability within and between the 14 accessions was evaluated by AMOVA. Most of the genetic variation (92.58%) was within the populations (Table 4). Thus, it can be concluded that there is considerable genetic variability within populations.

### DISCUSSION

The characterization of germplasm collections using molecular markers is an important strategy for the success of
Genetic variability revealed by microsatellite markers in a germplasm collection of *Jatropha curcas* L. in Brazil: an important...

breeding programs. Molecular markers have been used in several studies to characterize accessions of *J. curcas*. Sun et al. (2008) used SSR and AFLP markers to characterize 58 accessions from China, and found a low level of polymorphism. Tatikonda et al. (2009) evaluated the genetic variability of 48 accessions from six different regions of China using AFLP markers. Sudheer et al. (2009) used SSR markers to characterize germplasm from India. Santos et al. (2016) evaluated genetic variability using SSR and ISSR markers in 48 accessions from northern Minas Gerais state, Brazil.

The evaluation of genetic variability, by means of microsatellite markers, in 93 accessions from the UFV’s *J. curcas* collection was performed for the first time. This rich collection presents accessions from several Brazilian regions and abroad and its genetic evaluation is important for both genetic conservation and development of improved cultivars.

Regarding the genetic variability statistics obtained here, the number of alleles per loci (na), ranged from 2 to 3 (mean=2.2), was similar to that found by several other authors. Santos et al. (2016), using 11 SSR markers, found 2 to 5 alleles per locus, whereas Bressan et al. (2012) found 2 to 8 alleles per locus. However, Na-ek et al. (2011) identified only 1.4 alleles per locus, after evaluating 32 accessions. Rosado et al. (2010) found 1 to 2 alleles per locus in accessions in other Brazilian collection, and of the six microsatellite markers selected by them, four were monomorphic. According to Cruz et al. (2011) it is important to have polymorphic loci that have sufficient numbers of alleles to infer the genetic variability of a population in relation to another or its own over time, especially when subjected to evolutionary forces that promote differentiation.

The number of effective alleles (ne) is a measure that quantifies the alleles with significant frequency in a population. In the present study, ne ranged from 1.01 to 2.71. This result was superior to those found by Pecina-Quintero et al.

**Figure 1.** Dendrogram based on the UPGMA cluster analysis among 93 accessions of *Jatropha curcas* L. using five microsatellite loci.
In addition to estimating the number of polymorphic loci, other quantitative measures can be adopted, such as Ho and He, which allow to infer about the genetic structure of the population. The means of Ho and He in the present study (Ho=0.64 and He=0.40) were similar to those found by Santos et al. (2016). Bressan et al. (2012) also reported similar values (Ho=0.53 and He=0.66) for these measures, although with He higher than Ho, in a direction contrary to that verified in the present study and by Santos et al. (2016).

The coefficient of inbreeding or fixation index \( f \) is a parameter of importance in breeding programs that aim at the development of superior cultivars, as it allows measurement of the level of homozygosity in the population. In our study, the values of \( f \) varied from -0.01 to -0.85, being close to expected measures, given the genetic nature of the accessions. Vásquez-Mayorga et al. (2017), in evaluating accessions from Costa Rica, found negative values (-0.10) for \( f \), evidencing that there was no crossing between relatives among accessions. Cruz et al. (2011) stated that negative values of the inbreeding coefficient are common when the Ho values are greater than the expected heterozygosity, suggesting an excess of heterozygous loci. The negative \( f \) values should be interpreted as estimates of null inbreeding, that is, there was no crossing among related individuals. *J. curcas* is a monoic species, so it is expected that the loci are in the heterozygous state due to the mating system by allogamy, as evidenced by Sun et al. (2008), Rosado et al. (2010) and Wen et al. (2010).

The Shannon-Weaver index \( H' \) has been used in genetic studies as a measure of genetic variability within populations and resembles a genotype richness index. Here, \( H' \) values varied from 0.03 to 1.05 (mean of 0.62), which revealed the existence of high genetic variability among our 93 accessions of *J. curcas* collection. Wen et al. (2010) also used this index to verify genotypic richness in 45 accessions. These authors found an average value of 0.55 using SSR markers and suggested a high level of genetic variability to be explored in five *J. curcas* populations.

Among the accessions that showed the greatest genetic diversity, as revealed by the UPGMA algorithm (Figure 1), we highlight: UFVJC 05, 07, 12, 18, and 53, collected in João Pinheiro, Montalvânia and Barbacena (MG). The greater variability present in the accessions collected in Minas Gerais ratifies the study presented by Dias et al. (2012) that considered the State of Minas Gerais as a secondary center of genetic variability in *J. curcas*.

The formation of 11 clusters with the UPGMA algorithm evidences the expressive genetic variability and structuring of the collection. Sun et al. (2008), evaluating 58 accessions using microsatellite markers, found low genetic variability. Rosado et al. (2010) analyzed the genetic variability of 192 accessions by means of RAPD and SSR markers, finding limited variability among the accessions. Na-ek et al. (2011) evaluated 32 plants from different regions of the world with five SSR markers and also recorded low genetic variability. Naresh et al. (2015) assessed genetic variability among 14 accessions from India using RAPD markers and observed considerable variability among accessions.

The collection of seeds led to the formation of our genebank, prioritizing higher number of seeds per accessions, may be the differential of it in terms of expressive variability. Dias and Kageyama (1991) reported that knowledge of the level of genetic variation and its distribution within and between populations is critical. It is possible to better target the breeding strategies to be adopted, in order to maximize genetic gains through the selection cycles. In *J. curcas*, some studies have also detected a greater amount of genetic variability within populations (Bhering et al. 2015, Piotto et al. 2015, Sinha et al. 2015). This high concentration of genetic variation within populations implies sampling more plants per population, as recommended by Dias and Kageyama (1991) and Bhering et al. (2015). This strategy was effectively practiced in our collection, when priority was given to collecting more seeds per accession. The present study confirmed the accuracy of our accession collection process, in which each accession was represented by 16 plants. Previous studies

---

**Table 4. Analysis of molecular variance within and between 14 accessions of *Jatropha curcas* L.**

| Source of variation | df | Sums of Square | Variation (%) | \( \Phi_{ST} \) |
|---------------------|----|---------------|---------------|----------------|
| Between populations | 13 | 3.86          | 7.42          | 0.0742         |
| Within populations  | 82 | 15.7          | 92.58         |                |
| Total               | 95 | 19.56         | 100.00        |                |

(2014) (ne from 1.06 to 1.25) evaluating genetic variability in nine *J. curcas* populations from Mexico. Wen et al. (2010), in a study of genetic variability of five *J. curcas* populations, found values for ne ranging from 1.45 to 1.68.
Genetic variability revealed by microsatellite markers in a germplasm collection of *Jatropha curcas* L. in Brazil: an important...

employing molecular markers evaluated collections with a reduced number of plants per accession. This is possibly the main reason why they have systematically revealed low genetic variability in the species.

**CONCLUSION**

The five microsatellite markers selected (JCENA 87, JCDS 10, SSRY 107, SSRY 127 and CESR 0756) were able to detect genetic variability among *J. curcas* accessions, and to enable genetic evaluation through diversity statistics.

The accessions UFVJC 05, 07, 12, 18 and 53 exhibited expressive diversity, being able to comprise a base population for breeding.

Our results demonstrated the existence of molecular genetic variability within and between and accessions, indicating that our germplasm collection can be used as a base for breeding program.

**ACKNOLEGMENTS**

Thanks are due to CNPq, FAPEMIG and CAPES for financial support.

**REFERENCES**

Bhering LL, Peixoto LA, Leite NLSF and Laviola BG (2015) Molecular analysis reveals new strategy for data collection in order to explore variability in *Jatropha*. *Industrial Crops and Products* 74: 898-902.

Botstein D, White RL, Skolnick M and Davis RW (1980) Construction of genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics* 32: 314-331.

Bressan EA, Scotton DC, Ferreira RR, Jorge EC, Sebbenn AM, Gerald LTS and Figueira A (2012) Development of microsatellite primers for *Jatropha curcas* (Euphorbiaceae) and transferability to congeners. *American Journal of Botany* 99: 237-239.

Creste S, Tulmann Neto A and Figueira A (2001) Detection of single sequence repeat polymorphisms in denaturing polyacrylamide sequencing gels by silver staining. *Plant Molecular Biology Reporter* 19: 299-306.

Cruz CD, Ferreira FM and Pessoni LA (2011) *Biometria aplicada ao estudo da diversidade genética*. Universidade Federal de Viçosa, Viçosa, 620p.

Cruz CD (2013) Genes - a software package for analysis in experimental statistics and quantitative genetics. *Acta Scientiarum* 35: 271-276.

Dias LAS and Kageyama PY (1991) Variação genética em espécies arbóreas e consequências para o melhoramento florestal. *Agrotrópica* 3: 119-127.

Dias LAS, Missio RF and Dias DCFS (2012) Antiquity, botany, origin and domestication of *Jatropha curcas* (Euphorbiaceae), a plant species with potential for biodiesel production. *Genetics and Molecular Research* 11: 2719-2728.

Excoffier L, Smouse PE and Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131: 479-491.

Freitas RG, Missio RF, Matos FS, Resende MDV and Dias LAS (2011) Genetic evaluation of *Jatropha curcas*: an important oilseed for biodiesel production. *Genetics and Molecular Research* 10: 1490-1498.

Freitas RG, Dias LAS, Cardoso PMR, Evaristo AB, Silva MF and Araújo NM (2016) Diversity and genetic parameter estimates for yield and its components in *Jatropha curcas*. *Genetics and Molecular Research* 15: 1-10.

Gangapur DR, Agarwal P and Agarwal PK (2018) Molecular markers for genetic diversity studies in *Jatropha* (*Jatropha curcas* L.). *AIMS Environmental Science* 5: 340-352.

Kumar S, Stecher G and Tamura K (2016) MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33: 1870-1874.

Na-ek Y, Wongkaew A, Phumichai T, Kongsiri N, Kaveeta R, Reewongchai T and Phumichai C (2011) Genetic diversity of physic nut (*Jatropha curcas* L.) revealed by SSR markers. *Journal of Crop Science and Biotechnology* 14: 10-110.

Naresh B, Medipally SR, Dhanasri G and Devi P (2015) Genetic diversity of *Jatropha curcas* L. accessions and screening for high biodiesel production. *International Journal of Pure & Applied Bioscience* 3: 124-135.

Nei M (1972) Genetic distance between populations. *American Naturalist* 106: 283-292.

Pamidimarri DVNS, Singh S, Mastan SG, Patel J and Reddy MP (2008) Molecular characterization and identification of markers for toxic and non-toxic varieties of *Jatropha curcas* L. using RAPD, AFLP and SSR markers. *Plant Molecular Biology Report* 36: 1357-1364.

Pecina-Quintero V, Anaya-López JL, Zamarripa-Colmenero A, Núñez-Colín CA, Montes-García N, Solís-Bonilla JL and Jiménez-Becerril MF (2014) Genetic structure of *Jatropha curcas* L. in Mexico and probable centre of origin. *Biomass & Bioenergy* 60: 147-155.

Pioto F, Costa RS, França SC, Gavioli EA, Bertoni BW and Zingaretti SM (2015) Genetic diversity by AFLP analysis within *Jatropha curcas* L. populations in the State of São Paulo, Brazil. *Biomass & Bioenergy* 80: 316-320.
Rosado TB, Laviola BG, Faria DA, Pappas MR, Bhering LL, Quirino B and Grattapaglia D (2010) Molecular markers reveal limited genetic diversity in a large germplasm collection of the biofuel crop \textit{Jatropha curcas} L. in Brazil. \textit{Crop Science} 50: 2372-2382.

Santos DN, Ferreira JL, Setoaw TA, Cançado GMA, Pasqual M, Londe LCN, Saturnino HM and Vendrame WA (2016) Genetic structure from the oldest \textit{Jatropha} germplasm bank of Brazil and contribution for the genetic improvement. \textit{Anais da Academia Brasileira de Ciências} 88: 2363-2374.

Sinha P, Islam MA, Negi MS and Tripathi SB (2015) First identification of core accessions of \textit{Jatropha curcas} from India based on molecular genetic diversity. \textit{Journal of Plant Genetic Resources} 14: 77-80.

Sorrell S, Miller R, Bentley R and Speirs J (2010) Oil futures: a comparison of global supply forecasts. \textit{Energy Policy} 38: 4990-5003.

Sudheer PDVN, Singh S, Mastan SG, Patel J and Reddy MP (2009) Molecular characterization and identification of markers for toxic and non-toxic varieties of \textit{Jatropha curcas} L. using RAPD, AFLP and SSR markers. \textit{Molecular Biology Reports} 36: 1357-1364.

Sun QB, Li LF, Li Y, Wu GJ and Ge XJ (2008) SSR and AFLP markers reveal low genetic diversity in the biofuel plant \textit{Jatropha curcas} in China. \textit{Crop Science} 48: 1865-1871.

Tatikonda L, Wani SP, Kannan S, Beereelli N, Sreedevi TK, Hoisington DA, Devi P and Varshney RK (2009) AFLP-based molecular characterization of an elite germplasm collection of \textit{Jatropha curcas} L., a biofuel plant. \textit{Plant Science} 176: 505-513.

Vásquez-Mayorga M, Fuchs EJ, Hernández EJ, Herrera F, Hernández J, Moreira I, Arnáez E and Barboza NM (2017) Molecular characterization and genetic diversity of \textit{Jatropha curcas} L. in Costa Rica. \textit{Peer J} 5: 29-31.

Wen M, Wang H, Xia Z, Zou M, Lu C and Wang W (2010) Development of EST-SSR and genomic-SSR markers to assess genetic diversity in \textit{Jatropha curcas} L. \textit{BMC Research Notes} 3: 42-50.

Yeh FC, Boyle T, Rongca Y, Ye Z and Xiyan JM (1999) \textit{POPGENE}. \textit{Microsoft window-based freeware for population genetic analysis}. Version 1.31, Manual.