FUNCTIONAL PLANT BREEDING JOURNAL
Scientific Journal. ISSN 2595-9433
Volume 3, Number 1, Article n. 4, January/June
D.O.I. http://dx.doi.org/10.35418/2526-4117/v3n1a4
Received: 11/07/2020 - Accepted: 12/23/2020

WIDE GENETIC VARIABILITY WITHIN AND AMONG FAMILIES IN A GERMPLASM COLLECTION OF *Jatropha curcas* L. AS REVEALED BY MICROSATELLITE MARKERS

Erika da Costa Fernandes[1], Luiz Antônio dos Santos Dias[1]*, Eveline Teixeira Caixeta[2], Thais Roseli Corrêa[3], Dandara Rego Muniz[1], Odimar Ferreira de Almeida[1]

1 Departamento de Agronomia, Universidade Federal de Viçosa, 36570-900, Viçosa, MG, Brasil.
2 Embrapa Café. Current address: Bioagro, Universidade Federal de Viçosa, 36570-900, Viçosa, MG, Brasil.
3 Departamento de Fitotecnia e Fitossanidade, Universidade Estadual do Maranhão, 65055-970, São Luís, MA, Brasil.
** Corresponding author: Luiz Antônio dos Santos Dias (lasdias@ufv.br)*

Abstract: *J. curcas* is a species with wide potential for biofuel production. However, there are few breeding programs and little information on its genetic structure. Studies indicate that the species has narrow genetic variability. We quantify genetic variability, and decomposing it within and among 28 families of a genebank by means of microsatellite markers. Thirty-nine pairs of primers were tested, of which six were polymorphic for a total of 18 alleles, with a mean of three alleles/locus. These six markers allowed genetic variability to be estimated within and among the families through estimates of PIC (0.36), expected (He=0.44) and observed (Ho=0.48) heterozygosity, inbreeding coefficient (*f*=-0.03), Shannon-Wiener index (H'=0.71), and the formation of 11 clusters. Bayesian analysis classified the families in four groups. The present study was the first to portray the formation of four groups and detect high genetic variability using only accessions from outside the main center of diversity of the species. Analysis of molecular variance showed that most of the variability (92.4%) is contained within families. There was low differentiation among the families (F_{ST}=0.07). Collection of genotypes within families should be prioritized because they are where greater variability is concentrated. This strategy was used in setting up the present genebank, which prioritized the collection of more plants per family, efficiently bringing together greater variability. For the first time, the diversity statistics revealed high genetic variability to be exploited in this collection, on the contrary to most studies with *J. curcas* that have claimed low genetic diversity in the species.

Keywords: Genetic diversity, molecular marker, germplasm bank, breeding, physic nut.

Introduction

Jatropha curcas L. is an oilseed species with considerable potential for biofuel production due to the high oil content in its seeds, which ranges from 27% to 42% (Freitas et al., 2011; Montes et al., 2014; Edrisi et al., 2015; Freitas et al., 2016; Laviola et al., 2017; Cavalcante et al. 2020; Muniz et al. 2020). In spite of the potential of the crop, information regarding its genetic structure and variability is still lacking, and breeding programs aiming at obtaining new cultivars are rare.

The use of molecular markers is highly recommended for characterization of genetic variability since they are tools that are easily applied and highly reproducible and allow precise and rapid prospection of the whole genome (Caixeta et al., 2013). Among the different molecular markers, microsatellites (SSRs) are the most recommended since they are highly polymorphic, multiallelic, and of codominant inheritance.

Several studies on genetic diversity in *J. curcas* using SSR markers have been conducted (Sudheer Pamidimarri et al., 2009; Rosado et al., 2010; Wen et al., 2010; Na-ek et al., 2011; Bressan et al., 2012; Pecina-Quintero et al., 2014; Salvador-Figueroa et al., 2014; Alves et al., 2015; Sinha et al., 2015; Santos et al., 2016; Ribeiro et al., 2017; Vásquez-Mayorga et al., 2017; Gangapur et al., 2018; Souza et al., 2019).

Results in respect to the genetic variability reported are contradictories (Li et al., 2017). *J. curcas* has been classified as having narrow genetic variability, mainly in studies conducted with accessions from regions that are not considered it centers of diversity. This reduced variability may be related to different aspects, such as the type of marker used (dominant or codominant), the number of markers found, and the limited number of individuals evaluated (Maurya et al., 2015). It is also necessary to consider the origin of these accessions, the center of diversity, the collection method, and, if it be the case, even how the germplasm collection was set up.

When the study of genetic variability is conducted with accessions not coming from the center of diversity, it is highly important to study how the species was introduced in the region or in the country. *J. curcas* was introduced in Brazil mainly for use in establishing hedgerows (Dias et al., 2012). It is thought that colonizers brought plant cuttings that had the desired phenotype, and they were distributed in different regions of Brazil. Studies performed in germplasm banks evaluating the genetic variability among Brazilian accessions of *J. curcas* have reported low variability (Alkimim et al., 2013; Santos et al., 2016). These results may be related to how the plant was introduced.

The germplasm collection of *J. curcas* of UFV was confirmed to have high variability through morpho-agronomic and yield characteristics (Freitas et al., 2011; Freitas et al., 2016) and through microsatellite markers (Souza et al., 2019). However, this last study with microsatellite markers considered only one plant per accession, evaluating only the diversity among 93 accessions.

Decomposition of genetic variability into its components within and among families allows more efficient breeding strategies to be designed that make it possible to obtain greater genetic gain in a shorter time. For that reason, the aim of this study was to quantify the genetic variability and decomposing it into its components within and among 28 families of the germplasm collection of *J. curcas* through microsatellite molecular markers.

Material and methods

Collecting families

The Germplasm Collection of *Jatropha curcas* L. of UFV is at the Araponga Experimental Farm (lat 20º 39' S, long 42º 32' W, at alt 823 m asl) in the municipality of Araponga, Minas Gerais, Brazil. It consists of accessions coming from different geographic regions of Brazil and of the world (Table 1). Each accession is represented by 16 plants constituting families of open pollination, presumed half-sib families. The Collection is composed of 93 families and two of which are considered controls, for a total of 1504 plants. It is set up in five experiments (modules) in a randomized block design with four replications. Plots consist of four plants at a spacing of 2 x 2 m, with two common controls per module (Freitas et al.,

2016). The control families are representatives of the genotypes most planted in Brazil.

Table 1. Identification and provenance of the 28 families of *Jatropha curcas* L. evaluated.

Families	Origin
UFVJC 4	Santa Vitória - MG
UFVJC 5	João Pinheiro - MG
UFVJC 6	João Pinheiro - MG
UFVJC 7	João Pinheiro - MG
UFVJC 10	João Pinheiro - MG
UFVJC 18	Cana Brava II, Montalvânia
UFVJC 19	Cana Brava II, Montalvânia
UFVJC 22	Ipatinga - MG
UFVJC 31	Poté - MG
UFVJC 33	Itaipé - MG
UFVJC 36	Serra da Ibiapaba - CE
UFVJC 41	Jales - SP
UFVJC 43	Matozinhos - MG
UFVJC 47	Pirajaí - SP
UFVJC 53	Barbacena - MG
UFVJC 57	Janaúba - MG
UFVJC 68	São Luís - MA
UFVJC 70	Ariquemes - RO
UFVJC 71	João Pinheiro - MG
UFVJC 80	Unknown
UFVJC 83	Araras - SP
UFVJC 84	Petrolina - PE
UFVJC 86	Jequié - BA
UFVJC 90	Ipiaú - BA
UFVJC 93	Itaitê - BA
UFVJC 98	Iraquara - BA
UFVJC 102	Unknown
UFVJC C1	Sector 2 Bento, Janaúba - MG

For genetic-molecular evaluation within and among families of the Collection, 28 families considered as most divergent were selected, as reported by Souza et al. (2019). It is noteworthy that Souza et al. (2019) evaluated the variability among 93 families of this same Collection. In the present study, each family was represented by seven plants chosen at random among the 16 available, together with three plants from the control families, for a total of 187 plants evaluated. Samples were collected of young, completely developed, healthy leaves of bright green color; they were wrapped in aluminum foil, duly identified, and placed in a styrofoam box with ice for transport to the Biotechnology and Plant Breeding Laboratory of the UFV where they were stored in an ultra-freezer at -80 ºC.

After approximately three days in the ultra-freezer at -80 ºC, the samples were placed in the containers of the freeze dryer (ModulyoD) with their respective identifications. The time necessary for complete freeze drying of the material in this device was 30 hours. After that, the samples were macerated in crucibles and stored in 2-mL microtubes.

DNA extraction

DNA was extracted in the Coffee Biotechnology Laboratory (BioCafé) in the BIOAGRO of the UFV, based on the coffee protocol proposed by Diniz et al. (2005), with modifications. Approximately 50 mg of each macerated leaf tissue sample was placed in 2-mL microtubes that were previously identified. To this was added 1.5 mL of the extraction solution (0.35 M Sorbitol, 0.10 M Tris-HCl pH 8.0, 5 mM EDTA, 2 M NaCl, 2% CTAB (w/v), 0.2 M Tris-HCl, 0.05 M EDTA, 5% sarkosyl (w/v), 1% sodium bisulfate (w/v), 0.1% activated carbon (w/v) and 2% PVP-40 (w/v)) pre-heated to 65 ºC. The tubes were incubated in a water bath at 65 ºC for 40 min and were inverted every 10 minutes.

After incubation, the samples were centrifuged for 10 minutes at 14000 rpm and the supernatant was transferred to new tubes, with the addition of 600 µL of chloroform-isoamyl alcohol (24:1). The tubes were manually inverted until complete homogenization and centrifuged for 10 minutes at 14000 rpm, and the upper (aqueous) phase was removed and transferred to a new tube. This step was repeated. Then 600 µL of the supernatant was transferred to another microtube, and an equal volume (600 µL) of cold isopropanol was added, gently inverting the microtube to homogenize the mixture.

The microtubes were kept for 2 hours at -20 ºC, and then the material was centrifuged for 20 minutes at 14000 rpm. The supernatant was discarded and the pellet washed with 70% and 95% ethanol. The DNA was re-suspended in 200 µL of TE (10 mM Tris-HCl pH 8.0, 0.1 mM EDTA pH 8.0), containing RNAse in the final concentration of 80 µg/µL and incubated in a water bath at 37 °C for 30 minutes. The DNA was precipitated with 40 µL of 5M NaCl, plus 500 µL of 100% isopropanol, incubated for 2 hours at -20 ºC, and then centrifuged for 20 minutes at

14000 rpm. The supernatant was discarded, and after washing with 70% and 95% ethanol, was re-suspended in 200 µL of autoclaved ultrapure water. DNA was quantified using a spectrophotometer *NanoDrop* 2000 from Thermo Scientific. The final concentration was adjusted to 25 ng/µL and the DNA stored at -20 ºC.

Microsatellite molecular markers (SSRs)

Thirty-nine pairs of microsatellite primers developed, validated, and described for *J. curcas* were tested (Table 2). The PCR reactions were performed with 50 ng of DNA for a total volume of 20 µL, containing 1.5 mM $MgCl_2$, 150 µM of each dNTP (dATP, dTTP, dGTP, and dCTP), 0.4 µM of each primer, 1x of Taq buffer with KCl [500 mM KCl, 100 mM Tris-HCl (pH 8.8), 0.8% Nonidet P40], and 1.2 U of Taq DNA Polymerase (Ferramentas Life Science).

Amplifications were made in a thermoscycler (MJ Research, PTC 100), whose programming consisted of initial denaturation at 94 °C for 2 min, with a step of touchdown PCR composed of 7 cycles, beginning with 94 ºC for 30 s, annealing temperature reducing 1 ºC in each cycle (from 57 ºC to 51 ºC) for 30 s, and extension at 72 ºC for 40 s. This was followed by an additional 35 cycles at 94 °C for 30 s, annealing temperature at 51 ºC for 30 s, extension at 72 °C for 30 s, and finishing with an extension cycle of 72 °C for 15 minutes. The amplification products were separated in 1 h of vertical electrophoresis in polyacrylamide gel, 6x denaturing solution, stained with silver nitrate, according to the method described by De Brito et al. (2010).

Statistical analyses

The data were codified as codominant. Thus, when the locus exhibited four alleles, the codes 11, 22, 33, and 44 were attributed to the homozygote forms and 12, 13, 14, 23, 24, and 34 to the heterozygote forms. Only the data generated by polymorphic loci were used for analyses. The Popgene 1.31 software (Yeh and Boyle, 1997) was used to calculate the gene and allele frequencies, the observed number of alleles (na) and effective number of alleles (ne) (Kimura and Crow, 1964), the observed heterozygosity (Ho) and expected heterozygosity (He), also described as Nei's genetic diversity (Nei, 1973), the inbreeding coefficient or Wright's fixation index (f) (Wright, 1978), the Shannon-Wiener diversity index (H') (Lewontin, 1995), and Nei's genetic distance matrix (Nei, 1972). The mean Polymorphic Information Content (PIC) (Botstein et al., 1980) was calculated with the Genes software (Cruz, 2013).

A circular dendrogram was constructed using the unweighted pair group method with arithmetic mean (UPGMA), from Nei's (1972) genetic distance, with the software Mega (Kumar et al., 2016). The Structure software (Pritchard et al., 2000), based on Bayesian statistics, was used to study the population structure. K values ranging from 1 to 20 were tested with 20 replications. Each run was implemented with 500,000 initial interactions and 1,000,000 Markov Chain Monte Carlo (MCMC) simulations. For definition of the best number of genetic groups, the ΔK calculation was made, as described by Evanno et al. (2005), using the Structure Harvester software (Earl and vonHoldt, 2012).

Analysis of molecular variance (AMOVA) was used to reveal the partition of variance in its components within and among families. AMOVA was processed using the data codified according to the method described by Excoffier et al. (1992) using the Genes software (Cruz, 2013).

Results

Of the 39 pairs of primers tested, nine exhibited amplification of the fragments with clear and individualized bands. Of these nine, six were efficient in identifying polymorphism in the 187 plants of the Germplasm Collection of *Jatropha curcas* L. (Table 2) and were used in the analyses of population structure and diversity.

The allele frequencies observed ranged from 0.0080 (SSRY 146 locus) to 0.9572 (SSRY 113 locus). The A_1 allele was the most frequent from each locus and ranged from 0.5027 (SSRY 159) to 0.9572 (SSRY 113). The loci SSRY 159 and SSRY 146 exhibited the best detection of polymorphism because they revealed the greatest number of alleles (4), with wide variation of frequency (Table 3). The six polymorphic loci analyzed allowed identification of 18 alleles (na), with a mean of 3 alleles per locus and rang from 2 to 4. The effective number of alleles (ne) ranged from 1.09 to 2.70, with a mean of 1.93 ± 0.52 (Table 4).

Table 2. Identification and sequence of the microsatellite primers used in evaluation of genetic variability between and within families of *Jatropha curcas* L.

Primers	Sequences 5'-3'	
	Forward	Reverse
[3]CESR 0163	AACCACAGGAGTTGGTAATG	GAAAGAAGCAACAGAAATGG
[3]CESR 0333	ACAATATACAATGGCGATTTC	TAATGAATCTGTAGGACCCG
[3]CESR 0399	CAATGCATGGATCATAAGTG	CTCAAGTCAAATCTGGGAAC
[3]CESR 0836	CTTCCGCCTTCTCTTCCC	CATGGTTATCAGCCCATAGT
[3]CESR 0756	CAGGTTCGTCTTCTTCAACT	ATATATGATCCCGACAACAA
[1]JCENA 87	**ATCTGGAGTGAAACCAAAGA**	**CACATGGTAAGCATTACAAGC**
[1]JCENA 27	CATTTTTCATCAAGGCCTAC	GTATTTCTCCACACGCAACT
[1]JCENA 41	CTTTCTTACCCCTCATCCTT	AAAGCCAGGACATACTTGAA
[1]JCENA 47	GCCCGAGTTCTCTATAAGGT	CCATGGACTATCTCAACTTC
[1]JCENA 63	GCGTGGACTATCTCAACTTC	CTGATTACGCAATGGAACTA
[2]JCDS 10	CATCAAATGCTAATGAAAGTACA	CACACCTAGCAAACTACTTGCA
[2]JCMS 21	TAACCTCTTCCTGACA	ATAGGAAATAAGAGTTCAAA
[2]JCPS 1	GAGGATATTACAGCATGAATGTG	AATCAATCAATCTTTGGCAAA
[3]SSRY 63	TCAGAATCATCTACCTTGGCA	AAGACAATCATTTTGTGCTCCA
[3]SSRY 132	CTTTTTGCCAGTCTTCCTGC	TGTCCAATGTCTTCCTTTCCTT
[3]SSRY 184	TCATCCCAAAAATACCTCTAACA	CTCCGACAAGCATGTGAATG
[3]SSRY 107	CCATTTTCTCTTGCTTCTGTCA	TGGTTTGAACTATAAAATCCTT
[3]SSRY 150	CAATGCAGGTGAAGTGAATACC	AGGGTGCTCTTCAGAGAAAGG
[3]SSRY 153	TTCCAGAAAGACTTCCGTTCA	CTCAACTACTGCACTGCACTC
[3]SSRY 133	AGCATGTCATTGCACCAAAC	CGACTGCATCAGAACAATGC
[3]SSRY 50	CCGCTTAACTCCTTGCTGTC	CAAGTGGATGAGCTACGCAA
[3]SSRY 7	TGCCTAAGGAAAATTCATTCAT	TGCTAAGCTGGTCATGCACT
[3]SSRY 146	TTCCCTCGCTAGAACTTGTC	CTATTTGACCGTCTTCGCCG
[3]SSRY 113	**TTTGCTGACCTGCCACAATA**	**TCAACAATTGGACTAAGCAGC**
[3]SSRY 46	TCAGGAACAATACTCCATCGAA	CGCTAAAGAAGCTGTCGAGC
[3]SSRY 53	CCATGCAGTAGTGCCATCTTT	ATTTTCACCAACCGCAACTC
[3]SSRY 136	CGACTGCATCAGAACAATGC	AGCATGTCATTGCACCAAAC
[3]SSRY 100	**ATCCTTGCCTGACATTTTGC**	**TTCGCAGAGTCCAATTGTTG**
[3]SSRY 127	CTTCGCCCTCTACAAAAGGA	GCTGAACTGCTTTGCCAACT
[3]SSRY 61	GGCTGCTTTACCTTCTACTCAGA	CAAGAACGCCAATATGCTGA
[3]SSRY 4	ATAGAGCAGAAGTGCAGGCG	CTAACGCACACGACTACGGA
[3]SSRY 185	GAAGAAGACGGTTAAAGCAAGTT	ATGCCAGTTTGCTATCCAGG
[3]SSRY 177	ACCACAAACATAGGCACGAG	CACCCAATTCACCAATTACCA
[3]SSRY 151	**AGTGGAAATAAGCCATGTGATG**	**CCCATAATTGATGCCAGGTT**
[3]SSRY 146	**TTCCCTCGCTAGAACTTGTC**	**CTATTTGACCGTCTTCGCCG**
[3]SSRY 159	**CTTATCCTGTCCCCTCCACC**	**GACAATTGCATAGGAAGCACA**
[3]SSRY 119	AACATAGGCATTAAAGTTTGGCA	GCAAATGTGTTTTCAATATAAGGC
[3]SSRY 58	GAAGGACAAGCAAAGAAGCAA	TGGAATCCAATATTGATGACTAAGA
[2]JCDS 24	GGATATGAAGTTTCATGGGACAAG	TTCATTGAATGGATGGTTGTAAGG

Primers sources: [1] Bressan et al. (2012); [2] Sudheer Pamidimarri et al. (2009); [3] Wen et al. (2010).
In bold are the primers used in the present paper.

Table 3. Frequencies of the alleles A_1, A_2, A_3, and A_4 of the six polymorphic loci obtained from microsatellite markers and used in evaluation of genetic variability between and within 28 families of *Jatropha curcas* L.

Loci	A_1	A_2	A_3	A_4
JCENA 87	0.5535	0.4465	-	-
SSRY 113	0.9572	0.0428	-	-
SSRY 159	0.5027	0.0187	0.2059	0.2727
SSRY 146	0.6070	0.0080	0.0187	0.3663
SSRY 100	0.5856	0.0294	0.3850	-
SSRY 151	0.6818	0.2968	0.0214	-

Table 4. Diversity statistics for the 28 families of *Jatropha curcas* L. generated by six microsatellite molecular markers.

Loci	na	ne	Ho	He	f	H'	PIC
JCENA 87	2	1.98	0.83	0.49	-0.68	0.69	0.37
SSRY 113	2	1.09	0.04	0.08	0.48	0.18	0.08
SSRY 159	4	2.70	0.40	0.63	0.36	1.10	0.57
SSRY 146	4	1.99	0.60	0.50	-0.22	0.78	0.40
SSRY 100	3	2.03	0.45	0.51	0.11	0.78	0.41
SSRY 151	3	1.81	0.55	0.45	-0.23	0.70	0.36
Mean	3	1.93	0.48	0.44	-0.03	0.71	0.36
Standard deviation	0.89	0.52	0.26	0.19	0.43	0.30	0.16
Total	**18**						

na: Total number of alleles per locus; ne: Effective number of alleles; Ho: Observed heterozygosity; He: Expected heterozygosity; f: Inbreeding coefficient; H': Shannon-Wiener diversity index; and PIC: Polymorphic information content.

The estimates of observed heterozygosity (Ho) and expected heterozygosity (He) are important for quantifying genetic diversity and the occurrence of random mating within populations. In this study, Ho ranged from 0.04 to 0.83, with a mean of 0.48 ± 0.26, while He ranged from 0.08 to 0.63, with a mean of 0.44 ± 0.19 (Table 4). These estimates confirm the wide heterozygous nature of the Collection. The markers SSRY 113, SSRY 159, and SSRY 100 identified lower heterozygosis in these loci (0.04, 0.40, and 0.45, respectively), in line with positive values of the inbreeding coefficient (0.48, 0.36, and 0.11, respectively). The other markers exhibited negative estimates for the inbreeding coefficients, ranging from -0.68 to -0.22, and mean of -0.03 ± 0.43. Negative values for this coefficient are interpreted as zero inbreeding, i.e., there are no preferential crosses. These results indicate that the families of the Collection come from populations under random mating.

Another important estimator in characterization of diversity is the mean polymorphic information content (PIC), which measures the discriminatory power of the marker, considering not only the number of alleles per locus but also the frequency of the allele. The estimates of PIC ranged from 0.08 (SSRY 113) to 0.57 (SSRY 159). The marker SSRY 113 had the lowest PIC. The values estimated for the Shannon-Wiener index (H') ranged from 0.18 (SSRY 113) to 1.10 (SSRY 159), with a mean of 0.71 ± 0.30 (Table 4), indicating high variability in the families evaluated. Some high values of standard deviation result from the expressive variability contained in the small data set (n = 6). It should be remembered that the standard deviation is a measure of dispersion expressed in the same unit of data and corresponds to the square root of the variance.

For analysis of diversity among the 187 plants, disregarding family structure, the UPGMA dendrogram was constructed using Nei's genetic distance matrix (Nei, 1972) (Figure 1). Eleven clusters were formed, separated in the following manner: Cluster I (red, 45 plants), Cluster II (green, 17 plants), Cluster III (dark blue, 2 plants), Cluster IV (purple, 3 plants), Cluster V (brown, 12 plants), Cluster VI (gray, 14 plants), Cluster VII (light green, 5 plants), Cluster VIII (yellow, 45 plants), Cluster IX (blue, 27 plants), Cluster X (light blue, 10 plants), and Cluster XI (orange, 7 plants). This therefore showed wide variability among plants and, in addition, that many plants belonging to the same family were distributed in different clusters. The controls, which represent the commercially planted genotypes, were separated into different clusters: UFVJCT1_3 and UFVJCT1_5 in the yellow cluster and UFVJCT1_4 in the green cluster. The UFVJC 43 and UFVJC 41 families had the largest number of plants (5 plants each) in the same cluster (cluster I). This indicates that these families have lower variability among their individuals compared to the others. In contrast, the individuals of the UFVJC 19 family had the highest variability, with distribution in six different clusters (I, IV, V, VIII, IX, and X). The plants UFVJC 10_7, UFVJC 19_1, and UFVJC 31_2 formed a distinct cluster (cluster IV), which indicates greater similarity among these plants, although they come from different families.

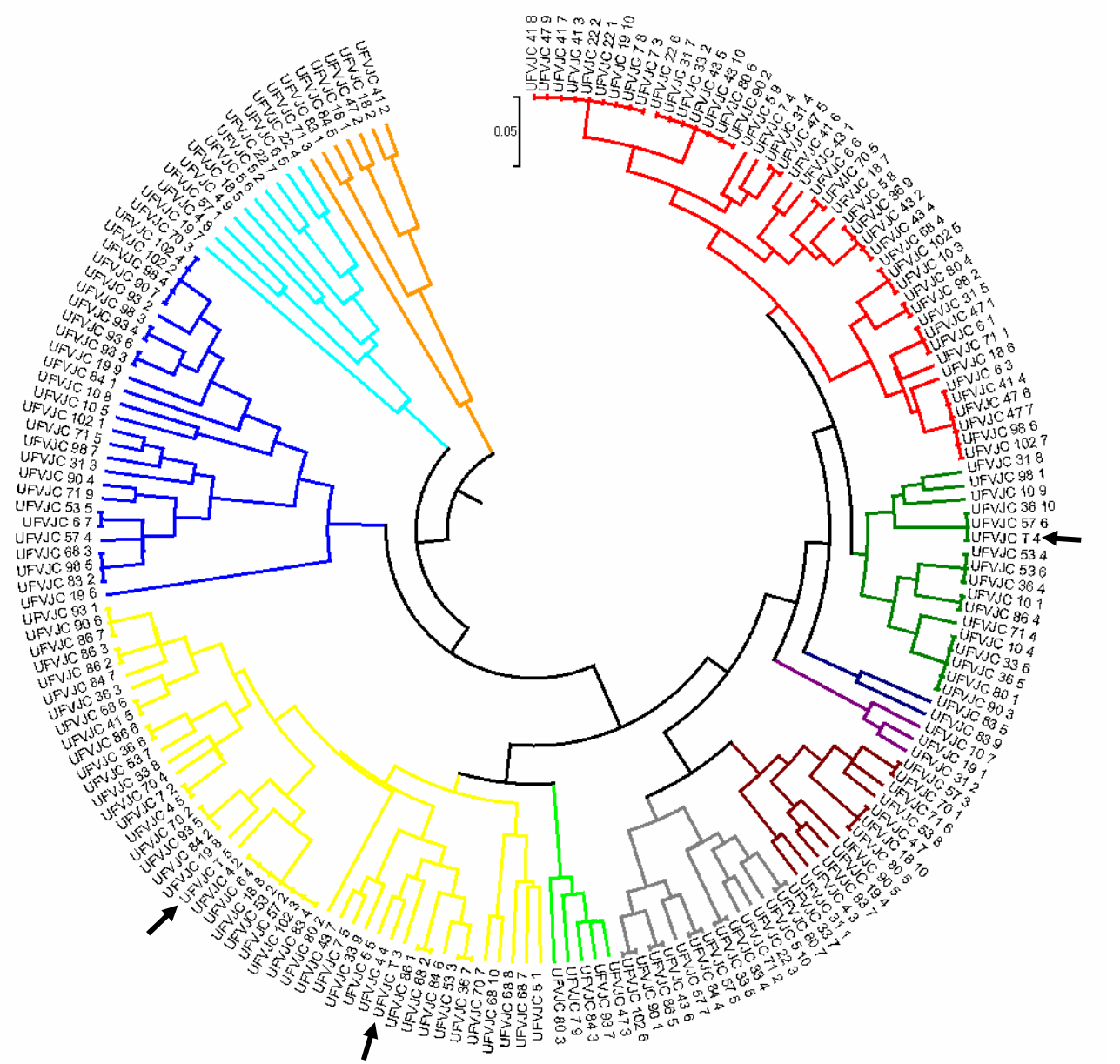

Figure 1. UPGMA dendrogram, based on Nei's genetic distance matrix, for evaluation of diversity among 187 *Jatropha curcas* L. plants.

Analysis with the Structure software allowed the structure of the population to be ascertained and the ancestry of the individuals constituting the families to be visualized. Based on calculation of ΔK, k = 4 (four distinct groups) was the optimal number to classify all the families and study the population structure (Figure 2), according to the method proposed by Evanno et al. (2005).

The proportion of alleles coming from each genetic group (four colors) and for each individual (vertical bar) is shown in Figure 3. Analysis of the structure of the population in the four groups reveals some individuals that have greater variation in division of the alleles, indicating admixture in the ancestry of the populations.

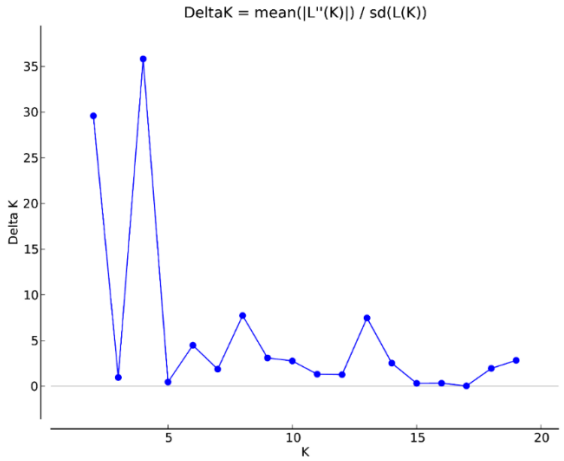

Figure 2. Estimated values of Δk for characterization of the best number of groups according to the method proposed by Evanno et al. (2005).

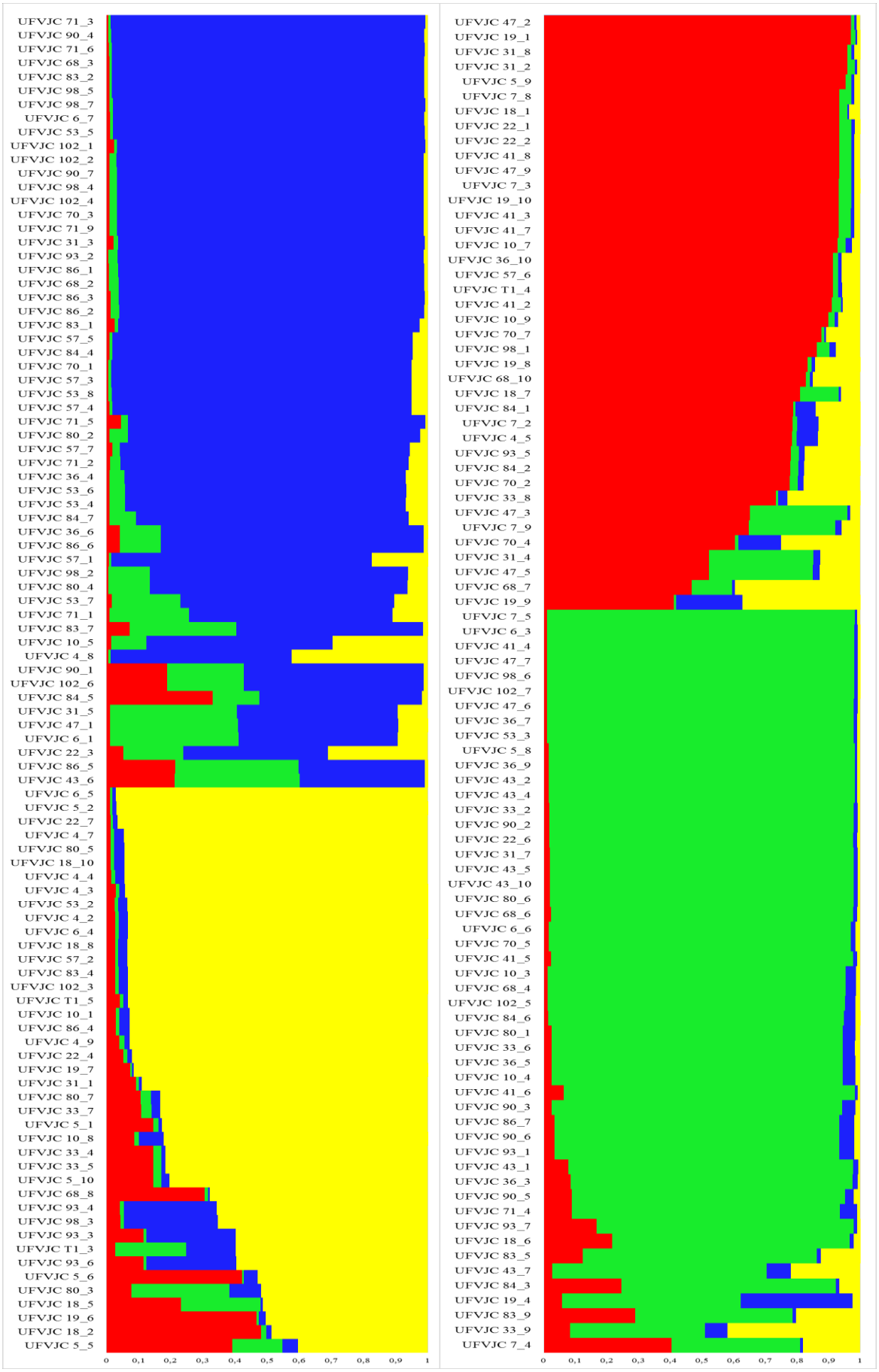

Figure 3. Display of population structure with separation in four groups of the 187 plants of the Collection of *Jatropha curcas L.*

Individuals from different families were distributed in the four groups, and some plants had similar proportions of alleles, such as UFVJC 47_5 and UFVJC 31_4.

The groups were formed by families coming from different states of Brazil, such that there was no parallelism between the regions of collection and the formation of groups. Few families had all their individuals joined in only one group; among them are UFVJC 86 (blue group) coming from Jequié, Bahia state; UFVJC 4 (yellow group) from Santa Vitória, Minas Gerais state; and UFVJC 41 and 47 (red group) from Jales and Pirajaí, São Paulo state, respectively. The three plants of the control family were separated in the yellow group (UFVJC T1_3 and UFVJC T1_5) and red group (UFVJC T1_4). Evaluation of the proportion of the alleles corresponds with analysis of the dendrogram, since each one of the three plants had different proportions.

Analysis of molecular variance (AMOVA) with the six markers showed that 92.40% of the genetic variation is contained within the families (Table 5). Only 7.59% of the total variation occurs among the families. The estimator F_{ST} represents differentiation among families, and the value found (F_{ST}=0.07) indicates low differentiation among them.

Table 5. Analysis of molecular variance (AMOVA) of 28 families of *Jatropha curcas* L.

Sources of variation	df	SS	MS	Variance components	% of variation	F_{ST}
Among families	27	15.79	0.58	0.03	7.59	0.07
Within families	159	60.05	0.37	0.37	92.40	
Total	**187**	**75.84**	**0.40**	**0.40**		

Discussion

The existence of genetic variability in germplasm collections is important in breeding programs for selection of parents with the aim of obtaining the best combination of favorable alleles and exploiting the manifestation of maximum heterosis. In addition, knowledge of variability is important to maintain genetic resources available. Decomposition of variability into its components within and among families allows efficient strategies to be designed for collecting germplasm and assisting selection. Few studies of diversity within and among families of *Jatropha curcas* L. have been developed with microsatellite molecular markers (SSRs) (Montes Osorio et al., 2014; Salvador-Figueroa et al., 2014; Maurya et al., 2015). This is a pioneering study with the aim of evaluating the genetic variability families and within a large number of families (28), with a total of 187 plants, coming from different Brazilian states.

The diversity statistics estimated in the present study corroborate those reported in recent studies on *J. curcas*. The mean number of alleles per locus (na) was 3, similar to that reported by Santos et al. (2016). In that study, 43 genotypes were evaluated with 11 SSR markers, and 2 to 5 alleles were identified. Del Angel et al. (2016) used six SSR markers and identified from 2 to 5 alleles, with a mean of 3 alleles/locus. Bressan et al. (2013) used five SSR markers to test the reproductive system and reported variation from 2 to 6 alleles, with a mean of 4.7 alleles/locus.

The effective number of alleles (ne) quantifies the alleles that had significant frequency in the population. The values of ne found (1.09 to 2.70) were greater than those reported by Wen et al. (2010) and Salvador-Figueroa et al. (2014); however, they were similar to those reported by One et al. (2014). Wen et al. (2010) obtained ne values that ranged from 1.45 to 1.67 in their study with 45 accessions evaluated with SSR markers. Salvador-Figueroa et al. (2014) found ne values from 1.50 to 1.82. One et al. (2014) evaluated a F_2 population with SSR markers and reported ne from 1.35 to 2.71.

The polymorphic information content (PIC) indicates the discriminatory power of the marker. PIC values ranged from 0.08 to 0.57, with a mean of 0.36 ± 0.16, and, for that reason, are considered moderately informative (Botstein et al., 1980). Vásquez-Mayorga et al. (2017) found similar values (PIC from 0.04 to 0.67) upon evaluating 50 accessions from Costa Rica; however, they had a lower mean value (0.27), exhibiting moderate informativity. Montes et al. (2014) studied

diversity in 70 accessions coming from Africa, Asia, and South America and found lower values of PIC (0.01 to 0.50, and mean of 0.15). Montes et al. (2014) associated these values with the high homozygosity in the accessions evaluated and they suggested the occurrence of high rates of self-pollination in them. Nevertheless, the values found in the estimates of observed heterozygosity (0.04 to 0.83) in that study show high variability in the accessions.

The estimates of observed heterozygosity (Ho) and expected heterozygosity (He) allow not only genetic diversity to be evaluated, but also the structure of the populations to be inferred. The mean values obtained in this study (Ho=0.48 and He=0.44) were of the same magnitude as those found by One et al. (2014) (Ho=0.47 and He=0.43) and by Santos et al. (2016) (Ho=0.64 and He=0.40). Santos et al. (2016) evaluated 48 accessions from the germplasm collection of EPAMIG through SSR and ISSR markers.

The estimates of the inbreeding coefficient (*f*), also called the fixation index, are used to identify the occurrence of preferential crosses among related individuals in the population, which, consequently, alter the frequency of heterozygotes. Negative values for this estimate are interpreted as zero inbreeding, such that in the population there is no occurrence of mating between related individuals (Cruz et al., 2011).

The positive *f* in three of the six markers evaluated may be related to the occurrence of self-pollination among some plants of the Collection. Alves et al. (2015) identified variations in the crossing rate (tm=0.88) and occurrence of self-pollination (s=0.12) in *J. curcas* populations. Mixed crossing rates (tm=0.68 and s=0.32) were also found by Bressan et al. (2013). This dynamic process in reproduction of the species can generate plants with different levels of inbreeding, such as identified in these studies. However, based on the mean value of *f* (-0.03) of the present study, it can be affirmed that the inbreeding level was not significant. It should be noted that these values are for the markers and might not reflect the values in the population. A similar value for *f* (*f*=-0.10) was reported by Vásquez-Mayorga et al. (2017).

The Shannon-Wiener index (H') is highly used as a measure of diversity within populations. In its calculation, it considers the frequency of the genotype and the number of genotypic classes for a given locus as a probability. For that reason, it is considered an index of genotypic richness. The values found for H' ranged from 0.18 to 1.10 (mean 0.71), which revealed the existence of high genetic variability within the 28 families of the Collection. Wen et al. (2010) reported H' values from 0.41 to 0.50, and although the accessions they studied were lower than those described in the present study, these authors considered them as having high variability. A H' value of 0.55 was found by Arolu et al. (2012) when they evaluated 48 accessions. However, One et al. (2014) obtained H' values ranging from 0.13 to 1.04, in line with the present study.

The formation of 11 clusters by UPGMA based on Nei's genetic distance matrix shows wide variability among the families of the collection, as previously reported by Freitas et al. (2011) and Freitas et al. (2016), using morphoagronomic and grain yield data, and by Souza et al. (2019) using SSR markers. Most of the accessions evaluated come from the North region of the state of Minas Gerais, corroborating the idea of this region as a secondary center of diversity of the species, as proposed by Dias et al. (2012).

Analysis of the population structure carried out on the Structure software showed distribution of the families in four groups (k=4), as well as admixture in the ancestry of some individuals. However, this fact was not related to the region of collection or to the order of the families. Similar results (k=4) were reported by Maurya et al. (2015) upon evaluating 96 accessions of *J. curcas* from different regions of the world, including toxic and non-toxic genotypes. In that case, the authors attributed this admixture to phylogenetic affinity of the accessions. Ovando-Medina et al. (2011), studying five populations from Mexico, found the formation of five groups (k=5), which were related to the geographic origin of the populations, although they also identified the occurrence of admixture in the ancestry of the accessions. To explain this mixture, the authors raised hypotheses such as the biogeography of the species, which led to the occurrence of crosses between individuals from different origins, a fact

that would sustain the idea of the heterozygotic origin of the population. Both the aforementioned studies contain accessions from the center of diversity, which may have contributed to the formation of the high number of groups. The present study is the first to portray the formation of four groups and detect high genetic variability using only accessions from outside the main center of diversity of the species.

Santos et al. (2016) used SSR and ISSR markers and detected the formation of three gene groups (k=3) upon evaluating the genetic diversity of 48 accessions from the EPAMIG Collection, all from Minas Gerais. However, Santos et al. (2016) reported low genetic variability among the accessions. Alkimim et al. (2013) evaluated the genetic diversity among 46 accessions of this same Collection with RAPD and ISSR markers and confirmed low genetic variability among the accessions. The results of the two previous studies may be attributed to the discriminatory power of the markers used, namely ISSR and RAPD. Since they are dominant, they do not allow identification of heterozygotes, and thus they become less informative than the codominant microsatellite markers. This contrast shows the efficacy of the setup strategy of the UFV Collection in representing each family by 16 plants. This strategy made it possible to gather greater genetic variability in the collection, making it more representative of the diversity that exists in the species.

Analysis of Molecular Variance (AMOVA) revealed that most of the genetic variation found is contained in the families (92.4%), with a differentiation index (F_{ST}) of 0.07 among families. Ovando-Medina et al. (2011) reported 87.8% of the variation occurring within the populations, with F_{ST} of 0.04 upon evaluating five populations in Mexico, indicating high variability and significant differentiation among them. Maurya et al. (2015) detected 94.0% of the total variation within the populations upon studying 96 accessions through SSR markers, with a differentiation index of 0.06, considered high, in spite of reporting low variability in the accessions studied. Mavuso et al. (2016) used ISSR markers to evaluate 13 populations of *J. curcas* and quantified the variation within the populations as 68.7%. Arolu et al. (2012) reported 94% of the variation occurring within the populations which was composed by 48 accessions. In contrast, Grativol et al. (2011) quantified the genetic diversity of 332 accessions of *J. curcas* coming from eight states of Brazil using ISSR markers and reported low variability, and identified that most of it occurred among accessions (37.4%) and among regions (33.4%). The authors attributed this low genetic divergence to the consequences of selections made by producers for the purpose of obtaining higher yields. Those studies report similar results, contrasting with those of the present study, where low differentiation was shown among the families of the UFV Collection and high genetic variability, with variation within families as its most important component.

Freitas et al. (2011) and Freitas et al. (2016) conducted the first studies on genetic diversity of the UFV Collection of *J. curcas* using morpho-agronomic and grain yield characteristics. Souza et al. (2019) confirmed this diversity among the accessions by means of microsatellite markers and revealed the formation of new clusters. In the present study, the high variability among the previously described families was confirmed with the formation of 11 clusters. In addition, it was possible to show that this wide genetic diversity occurs within families (92.4%). This result confirms the success of the strategy used in formation of the present collection, in which the collection of a greater number of seeds per families was endorsed instead of a greater number of accessions. The lack of adoption of this collection strategy in the formation of other collections may be considered the reason why other studies of diversity report low genetic variability in the species.

Based on the results presented in this study, we suggest that the step of selection of genotypes in breeding programs be conducted in experiments with a large number of replications and in varied conditions of environments, corroborating Santos et al. (2016). Due to the high genetic variability within families, selection should be practiced with greater emphasis within families, selecting and collecting more plants per family.

Final considerations

The six microsatellite molecular markers used (JCENA 87, SSRY 113, SSRY 159, SSRY

146, SSRY 100, and SSRY 151) were efficient in quantifying the genetic diversity within and among the 28 families of the *Jatropha curcas* UFV genebank. It was found that the variance within families is responsible for 92.4% of the total variability. Therefore, greater attention is recommended in the step of selection in breeding programs, giving emphasis to selecting more plants per family.

These results are the first to show the components of variance within and among families of *Jatropha curcas* L. coming from different regions of Brazil.

Acknowlegments

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

References

ALKIMIM, E.R.; SOUSA, T.V.; SOARES, B.O.; SOUZA, D.A.; JUHÁSZ, A.C.P.; NIETSCHE, S.; COSTA, M.R. 2013. Genetic diversity and molecular characterization of physic nut genotypes from the active germplasm bank of the Agricultural Research Company of Minas Gerais, Brazil. **African Journal of Biotechnology**, 12:907-913. http://dx.doi.org/10.5897/AJB12.2388

ALVES, P.F.; SEBBENN, A.M.; MANOEL, R.O.; CAMBUIM, J.; MORAES, M.A.; JUNIOR, E.F.; KUBOTA, T.Y.K.; PUPIN, S.; MORAES, M.L.T. 2015. Sistema de reprodução em uma população base de *Jatropha curcas* L.. **Scientia Forestalis**, 43:427-434.

AROLU, I.W.; RAFII, M.Y.; HANAFI, M.M.; MAHMUD, T.M.M.; LATIF, M.A. 2012. Molecular characterization of Jatropha curcas germplasm using inter simple sequence repeat (ISSR) markers in Peninsular Malaysia. **Australian Journal of Crop Science**, 6:1666-1673.

BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.W. 1980. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. **American Journal of Human Genetics**, 32:314-331.

BRESSAN, E.A.; SEBBENN, A.M.; FERREIRA, R.R.; LEE, T.S.G.; FIGUEIRA, A. 2013. *Jatropha curcas* L. (Euphorbiaceae) exhibits a mixed mating system, high correlated mating and apomixis. **Tree Genetics & Genomes**, 9:1089-1097. http://dx.doi.org/10.1007/s11295-013-0623-y

BRESSAN, E.A.; SCOTTON, D.C.; FERREIRA, R.R.; JORGE, E.C.; SEBBENN, A.M.; GERALD, L.T.S.; FIGUEIRA, A. 2012. Development of microsatellite primers for *Jatropha curcas* (Euphorbiaceae) and transferability to congeners. **American Journal of Botany**, 99: e237-e239. http://dx.doi.org/10.3732/ajb.1100532

CAIXETA, E.T.; FERRÃO, L.F.V.; MACIEL-ZAMBOLIM, E. 2013. Marcadores moleculares. In Borém, A.; Fritsche-Neto, R. (Org.). **Biotecnologia aplicada ao melhoramento de plantas.** Suprema, Visconde do Rio Branco, p. 31-68.

CAVALCANTE, N.B.; SANTOS, A.D.C.; ALMEIDA, J.R.G.S. 2020. The genus Jatropha (Euphorbiaceae): A review on secondary chemical metabolites and biological aspects. **Chemico-Biological Interactions**, 318:108976. http://dx.doi.org/10.1016/j.cbi.2020.108976

CRUZ, C.D. 2013. GENES - a software package for analysis in experimental statistics and quantitative genetics. **Acta Scientiarum. Agronomy**, 35:271-276. http://dx.doi.org/10.4025/actasciagron.v35i3.21251

CRUZ, C.D.; FERREIRA, F.M.; PESSONI, L.A. 2011. **Biometria aplicada ao estudo da diversidade genética**. Suprema, Visconde do Rio Branco, 620p.

DE BRITO, G.G.; CAIXETA, E.T.; GALLINA, A.P.; ZAMBOLIM, E.M.; ZAMBOLIM, L.; DIOLA, V.; LOUREIRO, M.E. 2010. Inheritance of coffee leaf rust resistance and identification of AFLP markers linked to the resistance gene. **Euphytica**, 173:255-264. http://dx.doi.org/10.1007/s10681-010-0119-x

DEL ANGEL, I.Z.; GARCÍA-PÉREZ, E.; HERNÁNDEZ, D.G.; PÉREZVÁZQUEZ, A.; RESÉNDIZ, C.Á. 2016. Genetic diversity of *Jatropha curcas* L. in Veracruz State, Mexico, and its relationships with the content of phorbol esters. **Global Advanced Research Journal of Agricultural Science**, 5(5):149-158.

DIAS, L.A.S.; MISSIO, R.F.; DIAS, D.C.F.S. 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. http://dx.doi.org/10.4238/2012.June.25.6

DINIZ, L.E.C.; SAKIYAMA, N.S.; LASHERMES, P.; CAIXETA, E.T.; OLIVEIRA, A.C.B.; ZAMBOLIM, E.M.; LOUREIRO, M.E.; PEREIRA, A.A.; ZAMBOLIM, L. 2005. Analysis of AFLP markers associated to the Mex-1 resistance locus in Icatu progenies. **Crop Breeding and Applied Biotechnology**, 5: 387-393. http://dx.doi.org/10.12702/1984-7033.V05N04A03

EARL, D.A.; VONHOLDT, B.M. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. **Conservation Genetics Resources**, 4:359-361. http://dx.doi.org/10.1007/s12686-011-9548-7

EDRISI, S.A.; DUBEY, R.K.; TRIPATHI, V.; BAKSHI, M.; SRIVASTAVA, P.; JAMIL, S.; SINGH, H.B.; SINGH, N.; ABHILASH, P.C. 2015. *Jatropha curcas* L.: A crucified plant waiting for resurgence. **Renewable and Sustainable Energy Reviews**, 41:855-862. http://dx.doi.org/10.1016/j.rser.2014.08.082

EVANNO, G.; REGNAUT, S.; GOUDET, J. 2005. Detecting the number of clusters of individuals using the software structure: a simulation study. **Molecular Ecology**, 14: 2611-2620. http://dx.doi.org/10.1111/j.1365-294X.2005.02553.x

EXCOFFIER, L.; SMOUSE, P.E.; QUATTRO, J.M. 1992. Analysis of Molecular Variance inferred from metric distances among DNA haplotypes: Application to human mitochondrial DNA restriction data. **Genetics**, 131:479-491.

FREITAS, R.G.; DIAS, L.A.S.; CARDOSO, P.M.R.; EVARISTO, A.B.; SILVA, M.F.; ARAÚJO, N.M. 2016. Diversity and genetic parameter estimates for yield and its components in *Jatropha curcas* L. **Genetics and Molecular Research**, 15: gmr7540. http://dx.doi.org/10.4238/gmr.15017540

FREITAS, R.G.; MISSIO, R.F.; MATOS, F.S.; RESENDE, M.D.V.; DIAS, L.A.S. 2011. Genetic evaluation of *Jatropha curcas*: an important oilseed for biodiesel production. **Genetics and Molecular Research**, 10:1490-1498. http://dx.doi.org/10.4238/vol10-3gmr1146

GANGAPUR, D.R.; AGARWAL, P.; AGARWAL, P.K. 2018. Molecular markers for genetic diversity studies in Jatropha (*Jatropha curcas* L.). **AIMS Environmental Science**, 5:340-352. http://dx.doi.org/10.3934/environsci.2018.5.340

GRATIVOL, C.; DA FONSECA LIRA-MEDEIROS, C.; HEMERLY, A.S.; FERREIRA, P.C.G. 2011. High efficiency and reliability of inter-simple sequence repeats (ISSR) markers for evaluation of genetic diversity in Brazilian cultivated *Jatropha curcas* L. accessions. **Molecular Biology Report**, 38: 4245-4256. http://dx.doi.org/10.1007/s11033-010-0547-7

KIMURA, M.; CROW, J.F. 1964. The number of alleles that can be maintained in a finite population. **Genetics**, 49:725-738.

KUMAR, S.; STECHER, G.; TAMURA, K. 2016. MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets. **Molecular Biology and Evolution**, 33:1870-1874. http://dx.doi.org/10.1093/molbev/msw054

LAVIOLA, B.G.; RODRIGUES, E.V.; TEODORO, P.E.; AZEVEDO PEIXOTO, L.; BHERING, L.L. 2017. Biometric and biotechnology strategies in Jatropha genetic breeding for biodiesel production. **Renewable and Sustainable Energy Reviews**, 76:894-904. http://dx.doi.org/10.1016/j.rser.2017.03.116

LEWONTIN, R.C. 1995. The apportionment of human diversity. In Dobzhansky, T.; Hecht, M.K.; Steere, W.C. (Eds.). **Evolutionary biology**. Springer US, Volume 6, Boston, p. 381-398. http://dx.doi.org/10.1007/978-1-4684-9063-3_14

LI, H.; TSUCHIMOTO, S.; HARADA, K.; YAMASAKI, M.; SAKAI, H.; WADA, N.; ALIPOUR, A.; SASAI, T.; TSUNEKAWA, A.; TSUJIMOTO, H. ANDO, T.; TOMEMORI, H.; SATO, S.; HIRAKAWA, H.; QUINTERO, V.P.; ZAMARRIPA, A.; SANTOS, P.; HEGAZY, A.; ALI, A.M.; FUKUI, K. 2017. Genetic tracing of *Jatropha curcas* L. from its Mesoamerican origin to the world. **Frontiers in Plant Science**, 8:1539. http://dx.doi.org/10.3389/fpls.2017.01539

MAURYA, R.; GUPTA, A.; SINGH, S.K.; RAI, K.M.; CHANDRAWATI; KATIYAR, R.; SAWANT, S.V.; YADAV, H.K. 2015. Genomic-derived microsatellite markers for diversity analysis in *Jatropha curcas*. **Trees**, 29:849-858. http://dx.doi.org/10.1007/s00468-015-1166-7

MAVUSO, C.; WU, Y.-P.; CHEN, F.-C.; HUANG, B.-H.; LIN, S.-J. 2016. Genetic diversity analysis of *Jatropha curcas* L. accessions cultivated in Taiwan using inter simple sequence repeats (ISSR) markers. **Agroforestry Systems**, 90:417-431. http://dx.doi.org/10.1007/s10457-015-9864-0

MONTES, J.M.; TECHNOW, F.; MARTIN, M.; BECKER, K. 2014. Genetic diversity in *Jatropha curcas* L. assessed with SSR and SNP markers. **Diversity**, 6:551-566. http://dx.doi.org/10.3390/d6030551

MONTES OSORIO, L.; TORRES SALVADOR, A.; JONGSCHAAP, R.E.; AZURDIA PEREZ, C.; BERDUO SANDOVAL, J.; TRINDADE, L.; VISSER, R.G.; VAN LOO, E. 2014. High level of molecular and phenotypic biodiversity in Jatropha curcas from Central America compared to Africa, Asia and South America. **BMC Plant Biology,** 14: 77. http://dx.doi.org/10.1186/1471-2229-14-77

MUNIZ, D.R.; ZAIDAN, I.R.; DIAS, L.A.S.; LEITE, J.P.V.; DINIZ, J.A. 2020. Biocide potential of Jatropha curcas L. extracts. **Journal of Biology and Life Science**, 11:138-154. http://dx.doi.org/10.5296/jbls.v11i2.17341

NA-EK, Y.; WONGKAEW, A.; PHUMICHAI, T.; KONGSIRI, N.; KAVEETA, R.; REEWONGCHAI, T.; PHUMICHAI, C. 2011. Genetic diversity of physic nut (*Jatropha curcas* L.) revealed by SSR markers. **Journal of Crop Science and Biotechnology**, 14:105-110. http://dx.doi.org/10.1007/s12892-011-0008-4

NEI, M. 1973. Analysis of gene diversity in subdivided populations. **Proceedings from the National Academy of Science of the USA**, 70: 3321-3323. http://dx.doi.org/10.1073/pnas.70.12.3321

NEI, M. 1972. Genetic distance between populations. **American Naturalist**, 106: 283-292. http://dx.doi.org/10.1086/282771

ONE, K.T.; TANYA, P.; MUAKRONG, N.; LAOSATIT, K.; SRINIVES, P. 2014. Phenotypic and genotypic variability of F_2 plants derived from *Jatropha curcas* × *integerrima* hybrid. **Biomass and Bioenergy**, 67:137-144. http://dx.doi.org/10.1016/j.biombioe.2014.04.036

OVANDO-MEDINA, I.; SÁNCHEZ-GUTIÉRREZ, A.; ADRIANO-ANAYA, L.; ESPINOSA-GARCÍA, F.; NÚÑEZ-FARFÁN, J.; SALVADOR-FIGUEROA, M. 2011. Genetic diversity in *Jatropha curcas* populations in the State of Chiapas, Mexico. **Diversity**, 3:641-659. http://dx.doi.org/10.3390/d3040641

PECINA-QUINTERO, V.; ANAYA-LÓPEZ, J.L.; ZAMARRIPA-COLMENERO, A.; NÚÑEZ-COLÍN, C.A.; MONTES-GARCÍA, N.; SOLÍS-BONILLA, J.L.; JIMÉNEZ-BECERRIL, M.F. 2014. Genetic structure of *Jatropha curcas* L. in Mexico and probable centre of origin. **Biomass and Bioenergy**, 60:147-155. http://dx.doi.org/10.1016/j.biombioe.2013.11.005

PRITCHARD, J.K.; STEPHENS, M.; DONNELLY, P. 2000. Inference of population structure using multilocus genotype data. **Genetics**, 155:945-959.

RIBEIRO, D.O.; SILVA-MANN, R.; ALVARES-CARVALHO, S.V.; SOUZA, E.M.S.; VASCONCELOS, M.C.; BLANK, A.F. 2017. Genetic variability in *Jatropha curcas* L. from diallel crossing. **Genetics and Molecular Research**, 16:gmr16029651. http://dx.doi.org/10.4238/gmr16029651

ROSADO, T.B.; LAVIOLA, B.G.; FARIA, D.A.; PAPPAS, M.R.; BHERING, L.L.; QUIRINO, B.; GRATTAPAGLIA, D. 2010. Molecular markers reveal limited genetic diversity in a large germplasm collection of the biofuel crop *Jatropha curcas* L. in Brazil. **Crop Science**, 50:2372-2382. http://dx.doi.org/10.2135/cropsci2010.02.0112

SALVADOR-FIGUEROA, M.; MAGAÑA-RAMOS, J.; VÁZQUEZ-OVANDO, J.A.; ADRIANO-ANAYA, M.L.; OVANDO-MEDINA, I. 2014. Genetic diversity and structure of *Jatropha curcas* L. in its centre of origin. **Plant Genetic Resources**, 13:9-17. http://dx.doi.org/10.1017/S1479262114000550

SANTOS, D.N.D.O.S.; FERREIRA, J.L.; SETOTAW, T.A.; CANÇADO, G.M.A.; PASQUAL, M.; LONDE, L.C.N.; SATURNINO, H.M.; VENDRAME, W.A. 2016. Genetic structure from the oldest Jatropha germplasm bank of Brazil and contribution for the genetic improvement. **Anais da Academia Brasileira de Ciências**, 88: 2363-2374. http://dx.doi.org/10.1590/0001-3765201620160167

SINHA, P.; ISLAM, M.A.; NEGI, M.S.; TRIPATHI, S.B. 2015. First identification of core accessions of *Jatropha curcas* from India based on molecular genetic diversity. **Plant Genetic Resources**, 14:77-80. http://dx.doi.org/10.1017/S1479262115000088

SOUZA, R.L.; DIAS, L.A.S.; CORRÊA, T.R.; CAIXETA, E.T.; FERNANDES, E.C.; MUNIZ, D.R.; ROSMANINHO, L.B.C.; CARDOSO, P.M.R. 2019. Genetic variability revealed by microsatellite markers in a germplasm collection of *Jatropha curcas* L. in Brazil: an important plant for biofuels. **Crop Breeding and Applied Biotechnology**, 19:337-346. http://dx.doi.org/10.1590/1984-70332019v19n3a46

SUDHEER PAMIDIMARRI, D.V.N.; SINGH, S.; MASTAN, S.G.; PATEL, J.; REDDY, M.P. 2009. Molecular characterization and identification of markers for toxic and non-toxic varieties of *Jatropha curcas* L. using RAPD, AFLP and SSR markers. **Molecular Biology Reports**, 36:1357-1364. http://dx.doi.org/10.1007/s11033-008-9320-6

VÁSQUEZ-MAYORGA, M.; FUCHS, E.J.; HERNÁNDEZ, E.J.; HERRERA, F.; HERNÁNDEZ, J.; MOREIRA, I.; ARNÁEZ, E.; BARBOZA, N.M. 2017. Molecular characterization and genetic diversity of *Jatropha curcas* L. in Costa Rica. **PeerJ**, 5:e2931. http://dx.doi.org/10.7717/peerj.2931

WEN, M.; WANG, H.; XIA, Z.; ZOU, M.; LU, C.; WANG, W. 2010. Development of EST-SSR and genomic-SSR markers to assess genetic diversity in *Jatropha curcas* L. **BMC Research Notes**, 3:42. http://dx.doi.org/10.1186/1756-0500-3-42

WRIGHT, S. 1978. **Evolution and genetics of populations. Variability within and among natural populations**. University of Chicago Press, Volume 4, Chicago, 590p.

YEH, F.C.; BOYLE, J. 1997. POPGENE Version 1.32. Microsoft Windows-based software for population genetics analysis (Computer Program). Available at https://sites.ualberta.ca/~fyeh/popgene_download.html