A peer-reviewed version of this preprint was published in PeerJ on 9 February 2017.

View the peer-reviewed version (peerj.com/articles/2931), which is the preferred citable publication unless you specifically need to cite this preprint.

Vásquez-Mayorga M, Fuchs EJ, Hernández EJ, Herrera F, Hernández J, Moreira I, Arnáez E, Barboza NM. 2017. Molecular characterization and genetic diversity of *Jatropha curcas* L. in Costa Rica. PeerJ 5:e2931
https://doi.org/10.7717/peerj.2931

Molecular characterization and genetic diversity of *Jatropha curcas* L. in Costa Rica

Marcela Vásquez-Mayorga [1] , **Eric J Fuchs** [2] , **Eduardo J Hernández** [1] , **Franklin Herrera** [3] , **Jesús Hernández** [4] , **Ilenana Moreira** [5] , **Elizabeth Arnáez** [6] , **Natalia M Barboza** [Corresp. 1, 7, 8]

[1] Centro de Investigación en Biología Celular y Molecular, Universidad de Costa Rica, San Pedro, San José, Costa Rica
[2] Escuela de Biología, Universidad de Costa Rica, San Pedro, San José, Costa Rica
[3] Estación Experimental Fabio Baudrit Moreno, Universidad de Costa Rica, Alajuela, Costa Rica
[4] Ministerio de Agricultura y Ganadería, San José, Costa Rica
[5] Escuela de Biología, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica
[6] Escuela de Biología, Instituto Tecnológico de Costa Rica, Cartago, Cartago, Costa Rica
[7] Escuela de Tecnología de Alimentos, Universidad de Costa Rica, San Pedro, San José, Costa Rica
[8] Centro Nacional en Ciencia y Tecnología de Alimentos, Universidad de Costa Rica, San Pedro, San José, Costa Rica

Corresponding Author: Natalia M Barboza
Email address: natalia.barboza@ucr.ac.cr

We estimated the genetic diversity of 50 *Jatropha curcas* samples from the Costa Rican germplasm bank using 18 EST-SSR, one G-SSR and nrDNA-ITS markers. We also evaluated the phylogenetic relationships among samples using nuclear ribosomal ITS markers. Non-toxicity was evaluated using G-SSRs and SCARs markers. A Neighbor-Joining (NJ) tree and a Maximum Likelihood (ML) tree were constructed using SSR markers and ITS sequences, respectively. Heterozygosity was moderate (He = 0.346), but considerable compared to worldwide values for *J. curcas*. The PIC (PIC = 0.274) and inbreeding coefficient (*f* = -0.102) were both low. Clustering was not related to the geographical origin of accessions. International accessions clustered independently of collection sites, suggesting a lack of genetic structure, probably due to the wide distribution of this crop and ample gene flow. Molecular markers identified only one non-toxic accession (JCCR-24) from Mexico. This work is part of a countrywide effort to characterize the genetic diversity of the *Jatropha curcas* germplasm bank in Costa Rica.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2469v2 | CC BY 4.0 Open Access | rec: 26 Jan 2017, publ: 27 Jan 2017

Molecular characterization and genetic diversity of *Jatropha curcas* L. in Costa Rica

Marcela Vásquez-Mayorga[1], Eric J. Fuchs[2], Eduardo J. Hernández[1], Franklin Herrera[3], Jesús Hernández[4], Ileana Moreira[5], Elizabeth Arnáez[5], Natalia Barboza[1,6,7]

1 Centro de Investigación en Biología Celular y Molecular, Universidad de Costa Rica, San José, Costa Rica

2 Escuela de Biología, Universidad de Costa Rica, San José, Costa Rica

3 Estación Experimental Fabio Baudrit Moreno, Universidad de Costa Rica, Alajuela, Costa Rica

4 Ministerio de Agricultura y Ganadería, Costa Rica, San José, Costa Rica

5 Escuela de Biología, Instituto Tecnológico de Costa Rica, Cartago, Costa Rica

6 Escuela de Tecnología de Alimentos, Universidad de Costa Rica, San José, Costa Rica

7 Centro Nacional en Ciencia y Tecnología de Alimentos, Universidad de Costa Rica, San José, Costa Rica

Corresponding Author:

Natalia Barboza

Ciudad de la Investigación, San Pedro, San José, 11501-2060 Costa Rica

Email address: natalia.barboza@ucr.ac.cr

23 **ABSTRACT**

24 We estimated the genetic diversity of 50 *Jatropha curcas* samples from the Costa Rican
25 germplasm bank using 18 EST-SSR, one G-SSR and nrDNA-ITS markers. We also evaluated the
26 phylogenetic relationships among samples using nuclear ribosomal ITS markers. Non-toxicity was
27 evaluated using G-SSRs and SCARs markers. A Neighbor-Joining (NJ) tree and a Maximum
28 Likelihood (ML) tree were constructed using SSR markers and ITS sequences, respectively.
29 Heterozygosity was moderate (He = 0.346), but considerable compared to worldwide values for *J.*
30 *curcas*. The PIC (PIC = 0.274) and inbreeding coefficient (f = -0.102) were both low. Clustering
31 was not related to the geographical origin of accessions. International accessions clustered
32 independently of collection sites, suggesting a lack of genetic structure, probably due to the wide
33 distribution of this crop and ample gene flow. Molecular markers identified only one non-toxic
34 accession (JCCR-24) from Mexico. This work is part of a countrywide effort to characterize the
35 genetic diversity of the *Jatropha curcas* germplasm bank in Costa Rica.

36

37

38

39

40

41

42 **INTRODUCTION**

43 The use of fossil fuels for energy production is being discouraged because of global warming and
44 fluctuating market prices. This situation has motivated research on alternative fuel sources such as
45 biodiesel from corn or palm oil (*To & Grafton, 2015*). *Jatropha curcas* is being explored as a new
46 biofuel crop (*Islam et al., 2013*). It is planted on approximately 1.8 million ha in Indonesia, China,
47 Brazil and Africa, and has the potential to become a biofuel crop in India and other tropical
48 countries (*Carels, 2013*). *J. curcas,* a member of the Euphorbiaceae family, is native to America
49 and has a pantropical distribution. It grows well under unfavorable climatic and soil conditions,
50 making it an attractive biofuel crop. Average oil content per seed is 40-45 % (*Jongschaap et al.,*
51 *2007*). Biofuel from this species is similar in quality to biofuels derived from conventional crops
52 like canola, linseed and sunflower, and surpasses the quality of biofuels produced from soybean
53 (*Basili & Fontini, 2012*).

54 The potential of *J. curcas* has not been fully exploited, mainly because of its variable and
55 unpredictable oil yield that limits large-scale cultivation. Genetic improvement may alleviate this
56 problem; however, characterization of the available germplasm is needed for breeding programs
57 to be efficient (*King et al., 2015; Mastan et al., 2012*). Over the past decade, *J. curcas* germplasm
58 has been genetically evaluated in India, China, Brazil, Mexico, Costa Rica and Central America
59 (*Avendaño et al., 2015; Basha & Sujatha, 2007; China Plant BOL Group et al., 2011; Montes*
60 *Osorio et al., 2014; Pecina-Quintero et al., 2014; Rosado et al., 2010; Wen et al., 2010*). Molecular
61 markers such as RAPDs, ISSRs, AFLPs, genomic simple sequence repeats (G-SSR) and expressed
62 sequence tags-SSR (EST-SSR) have all been used to assess the genetic diversity of *J. curcas*
63 collections. These studies have revealed low levels of genetic variability in India and Brazil (*China*
64 *Plant BOL Group et al., 2011; Rosado et al., 2010; Sun et al., 2008; Yadav et al., 2011*). Numerous

65 authors consider Mexico and Central America to be the center of origin and diversification
66 (*Abdulla et al., 2009; Basha et al., 2009; Heller, 1996; Openshaw, 2000; Pamidimarri,*
67 *Chattopadhyay & Reddy, 2008; Pamidimarri & Reddy, 2014; Pecina-Quintero et al., 2011;*
68 *Tatikonda et al., 2009*), and high levels of genetic diversity in Guatemala (*Raposo et al., 2014*)
69 and Mexico support this hypothesis (*Ambrosi et al., 2010; Ovando-Medina*, *Adriano-Anaya &*
70 *Vásquez-Ovando, 2013; Ovando-Medina et al., 2011; Pamidimarri & Reddy, 2014; Pecina-*
71 *Quintero et al., 2011*). However, these studies have consistently shown a lack of relationship
72 between the geographic proximity of collection sites and the genetic similarity among accessions,
73 because collection sites rarely represent the place of origin of accessions.
74 Another limitation to large-scale production of *J. curcas* is the possible toxicity of the seed. Both
75 toxic and non-toxic genotypes of *J. curcas* are known (Insanu *et al.,* 2013). Varieties with a non-
76 toxic seed cake would be more readily accepted by local farmers because sub-products of the oil
77 extraction process could be utilized for animal feeding (*Makker & Becker, 2015*). Therefore, it is
78 important to evaluate materials stored in germoplasm banks for non-toxicity. Non-toxic genotypes
79 have been described in Mexican accessions (*Vera-Castillo et al., 2014*). As with variation in
80 toxicity, genotypes with variable seed oil content and number of seeds may surface in other Latin
81 American regions and contribute to *J. curcas* breeding programs all over the world.
82 In the past decade, SSR markers have proven useful for the analysis of genetic diversity. Simple
83 sequence repeats that identify variability in transcribed genomic regions can be found in EST
84 libraries. They may facilitate the identification of functional candidate genes, increase the
85 efficiency of marker-assisted selection and serve as markers for comparative mapping (*Varshney,*
86 *Graner & Sorrells, 2005*). Nuclear ribosomal-DNA internal transcribed spacers (nrDNA ITS) are
87 sequence-based markers that have been used in phylogenetic studies and to assess genetic diversity

88 at the species level in a wide range of taxonomic groups (*Nieto-Feliner & Roselló, 2007*). ITS
89 markers have been developed for *J. curcas* (*Pecina-Quintero et al., 2011*) and are believed to have
90 greater discriminatory capacity than plastid *rbcL* and *matK* markers (*China Plant BOL Group et*
91 *al., 2011*). SSR and ITS markers used together allow for a better understanding of the evolutionary
92 history of undomesticated species such as *J. curcas*.
93 Prior to the current study, the genetic variability of the Costa Rican *J. curcas* germplasm bank had
94 not been studied. The present study explores the molecular diversity of accessions using EST-SSR
95 and G-SSR markers and evaluates the phylogenetic relationships between them using nuclear
96 ribosomal ITS markers. Alleles associated with *J. curcas* toxicity were also evaluated with G-
97 SSRs and sequence characterized amplified region (SCAR) markers.
98

99 **MATERIALS & METHODS**

100 **DNA collection and extraction**

101 The germplasm bank of *Jatropha curcas* in Costa Rica includes accessions from Honduras, Brazil,
102 India, Mexico, El Salvador, Ecuador, Uganda, Colombia and South Africa. Costa Rican accessions
103 include only spontaneously occurring individuals collected in the field and vegetatively
104 propagated. The collection does not include material from commercial plantations (Fig. 1).
105 International accessions originated from seeds collected from spontaneously occurring plants in
106 each country. All accessions are maintained by vegetative propagation in the germplasm bank
107 located at the Fabio Baudrit Experimental Station (10°00'10.3"N, 84°16'17.6"W) at Universidad
108 de Costa Rica.

109 For genetic analysis, two young leaves were collected from each of 50 plants representing different
110 accessions in the germplasm bank (Table 1). Vegetative material was frozen and later lyophilized

111 for DNA extraction. Nucleic acids were extracted using the *Möller et al.* (*1992*) protocol and
112 quantified using a Nanodrop (Thermo Scientific).
113
114 **SSR analysis**
115 Eighteen EST-SSRs and one G-SSR marker (*Wen et al., 2010*) were used to analyze the 50 samples
116 (Table 2). PCR was performed in a final volume of 25 µL with 1 X PCR buffer, 400 µM dNTPs,
117 0.4 µM of each primer, and 1 U Taq DNA polymerase (Thermo Scientific). Amplification
118 conditions included an initial denaturation at 94 °C for 1 min, followed by 35 cycles of 94 °C for
119 1 min, 42-50 °C for 1 min (depending on the annealing temperature of each primer pair) and 72
120 °C for 1 min, with a final extension of 72 °C for 10 min. Results were visualized in 4-6 %
121 polyacrylamide gels dyed with silver nitrate. To confirm the results, duplicates of 10 % of the
122 samples were made. Acrylamide gels were scored manually in the GNU Image Manipulation
123 Program (www.gimp.org) and a data matrix with band size data was created. The SSRs were
124 scored according to amplicon size. The software Microchecker was used to test for null alleles and
125 allelic dropout (*van Oosterhout et al., 2004*).
126
127 **nrDNA-ITS region amplification and sequencing**
128 ITS primers were used to elucidate phylogenetic relationships between accessions. The primer pair
129 JCITS-1-F (5´-ACCTGCGGAAGGATCATTGTCGAAA-3´) and JCITS- 2-R
130 (5´CCTGGGGTCGCGATGTGAGCGT 3´) was used (*Pamidimarri, Chattopadhyay & Reddy,*
131 *2008*) in a PCR reaction with a final volume of 25 µL and a final concentration of 1 X reaction
132 buffer, 1.5 mM $MgCl_2$, 0.2 µM of each primer, 0.2 mM dNTPs, 1 U Taq DNA polymerase
133 (Thermo Scientific) and 50 ng of DNA. The PCR program consisted of an initial denaturation step

of 94 °C for 1 min, followed by 30 cycles of 40 s at 94 °C, 65 °C for 1 min and 72 °C for 1 min. The final extension step was 5 min at 72 °C. The PCR products were sent to Macrogen® for sequencing. The obtained sequences were aligned and submitted to Genbank (www.ncbi.com).

Data analysis

Genetic diversity was quantified as the expected heterozygosity, the Polymorphism Information content (PIC) and the inbreeding coefficient (*f*). The PIC index describes the probability that two random accessions would have different alleles at a random locus (*Smith et al., 2004*). All estimates were calculated using the Powermarker 3.25 software (*Liu & Muse, 2005*). The validity of all estimates was assessed by means of 50 000 bootstraps.

To determine the genetic relation between accessions, we calculated the $\delta\mu^2$ genetic distance between accessions using the POPULATIONS software (http://bioinformatics.org/~tryphon/populations/). The $\delta\mu^2$ metric developed by *Goldstein et al.* (*1995*) is based on the stepwise mutation model and estimated distances are a linear function of divergence time; this distance is preferred for taxa that have diverged widely (*Goldstein et al., 1995*). We used the distance matrix to construct a Neighbor-Joining (NJ) tree using the default parameters in POPULATIONS. The standardized genetic distance matrix was also used to perform a Principal Coordinates Analysis (PCA) in GenAlEx 6.5 (*Peakall & Smouse, 2012*).

For phylogenetic analysis, ITS sequences for *J. curcas* samples from Mexico, India, Cape Verde, Spain, Africa and Madagascar were downloaded from GenBank (http://www.ncbi.nlm.nih.gov). Sequences from Costa Rican samples were edited with BioEdit software version 7.2.5 and aligned using the MAFFT algorithm in the GUIDANCE server. To compare the phylogenetic relationship of GenBank sequences with those from the Costa Rican *J. curcas* germplasm bank, we initially

157 used jModelTest 2.1.7 (*Santorum et al., 2014*) to define the optimum substitution model for all
158 sequences and the GTR model (*Tavaré, 1986*) with uniform rates was chosen. A maximum
159 likelihood (ML) tree was constructed with 3 000 bootstrap replications and expressed as the
160 number of base substitutions per site using MEGA 6.0 (*Tamura et al., 2013*).
161

162 **Toxicity evaluation**

163 Plant toxicity was evaluated by the presence of alleles from three SSR markers associated with
164 lack of toxicity (Table 3) (*Vischi, Raranciuc & Baldini, 2013*). PCR amplification was achieved
165 in a final volume of 25 µL with a final concentration of 1 X reaction buffer, 1.5 mM of $MgCl_2$, 0.2
166 µM of each primer, 0.2 mM dNTPs, 1 U Taq DNA polymerase (Thermo Scientific) and 50 ng/µl
167 DNA sample. The PCR program had an initial denaturation of 94 °C for 1 min, 30 cycles of 40 s
168 at 94 °C, 65 °C for 1 min, 72 °C for 1 min and a final extension of 5 min at 72 °C. Results were
169 visualized by genotyping with fluorescent dyes (FAM, VIC and PET) (*Vischi, Raranciuc &*
170 *Baldini, 2013*) in a 3130 sequencer (Applied Biosystems). Two SCAR markers (ISPJ1 and ISPJ2)
171 were also used to evaluate alleles for toxicity in all accessions following the protocol of *Basha &*
172 *Sujatha* (*2007*). ISPJ1 amplifies a 543 bp fragment and is specific for toxic genotypes, while ISPJ2
173 is specific for non-toxic genotypes and amplifies a 1096 bp fragment. Results were visualized in a
174 2 % agarose gel run for one hour and dyed with GelRed (Biotium). Plants were scored as toxic or
175 non-toxic based on the presence and size of amplicons.
176

177 **RESULTS**

178 **Genetic diversity analysis**

179 Genetic diversity was estimated using 18 SSR-ESTs and one G-SSR marker. Average
180 heterozygosity (He) was 0.346 ± 0.062 (± SD). Polymorphism information contents (PIC) ranged
181 from 0.042 to 0.677, with a mean PIC of 0.274 ± 0.165. We did not find evidence of inbreeding f
182 = -0.102 ± 0.346 (Table 4).
183 The NJ tree did not show a clear clustering pattern (Fig. 2) and clusters did not reflect geographic
184 proximity. Accessions from countries located in close proximity such as Colombia (JCCR-38) and
185 Ecuador (JCCR-25) did not seem to be genetically close to each other. In contrast, samples from
186 distant locations clustered together, for example, India (JCCR-INDIA) and Costa Rica (JCCR-14);
187 South Africa (JCCR-47) and Honduras (JCCR-2); Brazil (JCCR-16) and Costa Rica (JCCR-MIR);
188 and Mexico (JCCR-24) and Ecuador (JCCR-25). Within Costa Rica, there was no evidence of
189 geographic structure. Samples collected from sites separated by more than 300 km grouped
190 together (JCCR-20, JCCR-7).
191 Our PCA analysis produced comparable results. Genetic information accounted for 41.26 % of the
192 observed variance; the first two components explained 19.17 % and 11.86 % of the total variance,
193 respectively. We did not observe distinct groups of accessions in a biplot of the first two
194 components (Fig. 3). Accessions from the same country did not group together, such as those from
195 Mexico (JCCR-24 and JCCR-31) and India (JCCR-INDIA and JCCR-27 INDIA). Costa Rican
196 accessions were scattered throughout the plot without any discernible pattern. These results were
197 congruent with our NJ tree.
198
199 **nrDNA-ITS sequence analysis**
200 As with the NJ and PCA analysis, the observed patterns from the Maximum Likelihood (ML) tree
201 did not reflect the geographic origin of the accessions (Fig. 4) and in most cases, the clades grouped
202 in polytomies. International and Costa Rican accessions from this research were scattered

203 throughout the tree (JCC-38, JCCR-2, JCCRLEP, JCCR33-1, JCCR31, JCCR21, JCCR19,
204 JCCR9, JCCR8, JCCR7, JCCR3), they grouped with sequences from Mexico and with individuals
205 from Spain, Cape Verde, Africa and Madagascar (GenBank accession numbers EU700449,
206 EU70055, EU70046, EU70045, respectively). It is important to mention that the observed
207 clustering of germplasm independent of geographical origin could be an artifact of unequal
208 sampling, as the majority of the samples used in this research were from Costa Rica and only one
209 or two representatives from each of the other countries were included.

210 **Toxicity evaluation**

211 A single Mexican accession (JCCR-24) was identified as non-toxic by the SSR primer set (JCT-
212 31, JCT-27 and JcSSR-26) and both SCAR primers ISPJ1 and ISPJ2. ISPJ1 also identified non-
213 toxic genotypes in two other accessions: JCCR-32 and JCCR-43. Using ISPJ2, accessions JCCR-
214 22 and JCCR25 were identified as non-toxic. These results need to be corroborated using *in vivo*
215 assays.

216

217 **DISCUSSION**

218 **Genetic diversity of *Jatropha curcas***

219 Our study is the first to provide genetic diversity estimates for Costa Rican *J. curcas* samples.
220 Based on morphological and molecular evidence, different authors support the idea that Mexico
221 and Central America may be the center of origin and diversification for *J. curcas* (*Pamidimarri &
222 Reddy, 2014; Pecina-Quintero et al., 2014)*. *Pecina-Quintero et al.* (*2014*) found high genetic
223 diversity of *J. curcas* in Mexico. *Grativol et al.* (*2010*) analyzed 332 accessions from 12 locations
224 in Brazil using ISSR primers and reported lower genetic diversity than that reported in Mexico.
225 *Pamidimarri & Reddy* (*2014*) used RAPD and AFLPs to analyze the molecular diversity of 42

226 Indian accessions of *J. curcas* and found a mean percentage of polymorphism (PP) of 26.47. In
227 the same study, the average PP of Mexican accessions was 33.18. The mean PP of the germplasm
228 samples excluding the Indian accessions was 35.86, supporting the hypothesis that Indian
229 germplasm is less diverse than germplasm in other regions of the world (*Colombo, Second &*
230 *Charrier, 2000; Ram, Kumar & Bhatt, 2004*). *Montes et al*. (*2014*) also found higher genetic
231 diversity in Mexican and Central American accessions compared to those from other parts of the
232 world. Given the observed levels of genetic diversity, Costa Rica may be a secondary center of
233 origin or diversification for this species (*Pamidimarri & Reddy 2014; Pecina-Quintero et al.,*
234 *2014*). However, comparisons are limited since few studies have evaluated genetic diversity of *J.*
235 *curcas* germplasm using EST-SSR markers. *Wen et al.* (*2010*) evaluated 45 accessions from
236 Indonesia, Grenada, South America and two Chinese provinces, and found a mean genetic
237 diversity of 0.3819. The most diverse locations were South America and Yunnan with H= 0.33
238 and H= 0.3473, respectively. In another study, 50 EST-SSR markers were used to evaluate 25
239 Indian accessions and an average He of 0.30 was found (*Yadav et al., 2011*). As in our study,
240 accessions clustered independently of geographic origin. Our PIC estimates (PIC= 0.274 $\pm$ 0.165)
241 were comparable to those obtained by *Yadav et al.* (*2011*) (PIC= 0.25 ± 0.16) and are considered
242 moderately informative (*Botstein et al., 1980*).
243 In Costa Rica, *Jatropha curcas* is typically not cultivated commercially. Plants usually grow as
244 hedgerows and are occasionally reproduced by farmers through cuttings. Our samples represent
245 plants growing spontaneously in the field; no commercially grown material was included in the
246 study. Therefore, our genetic diversity estimates represent the standing natural variation of this
247 species. However, since EST-SSR markers were developed from expressed sequence tag libraries,
248 they reside within genes and are subject to selection, which reduces unfavorable polymorphisms

249 (*Cova et al., 2012; Ellis & Burke 2007*). EST-SSR markers are less polymorphic than genomic
250 SSRs (*Song et al., 2012*) and consequently, genetic diversity may have been underestimated in this
251 study. *J. curcas* is a predominantly outcrossing species (tm= 0.683) (*Bressan et al., 2013*). As
252 expected, we found no significant evidence of inbreeding (f = -0.102). High rates of gene flow
253 should produce low levels of inbreeding, which would result in the low structure suggested by our
254 NJ and ML clustering. Low inbreeding coefficients were also estimated in Mexico and South
255 America (*Ambrosi et al., 2010*). Although inbreeding was negligible in the present study, in other
256 parts of the world such as India and Brazil, lower genetic diversity has been attributed to increased
257 selfing or a high paternity correlation due to the spread of introductions across the country through
258 vegetative propagation, recent common ancestry, drift, and intensive selection of the currently
259 cultivated materials since the time of introduction (*Basha & Sujatha, 2007; Bressan et al., 2013;*
260 *Rosado et al., 2010*).

261

262 **Phylogenetic analysis of *Jatropha curcas***

263 Several studies of *J. curcas* have shown that collection sites do not necessarily reflect the genetic
264 origin of accessions (*Ambrosi et al., 2010; Maghuly et al., 2015*). Our NJ tree and PCA analysis
265 (Fig.2 and Fig.3) showed no correlation between genetic similarity and geographic proximity. For
266 example, the two Mexican accessions, JCCR-24 and JCCR-31, clustered in different putative
267 groups. Also, the two Indian accessions (JCCR-27 and JCCR-INDIA) did not seem to be related.
268 *J. curcas* is widely cultivated and plants are exchanged commonly. Accessions from the same
269 country may come from diverse origins and thus may be placed in different clades. Our analysis
270 suggests that collection sites may not necessarily represent local germplasm, but genetically
271 distinct lineages from different geographic regions. Material exchanges between American,

272 African and Asian collections have occurred commonly over the last 200 years (*Heller, 1996*) and
273 may have resulted in founder effects in Africa and Asia (*Henning, 2007; Lengkeek, 2007;*
274 *Pamidimarri & Reddy 2014*). For example, according to *Pamidimarri & Reddy* (*2014*), Portuguese
275 seafarers introduced accessions from Mexico and Central America to India through two dispersal
276 routes: one brought *J. curcas* through Africa, Madagascar and finally to India, while the other
277 passed through Spain on its way to India. These migration routes support our findings of a widely
278 dispersed plant with little geographic structure.
279 The observed ML tree topology may be an artifact of the low level of genetic structure seen in our
280 other analysis. Concurrently, low levels of inbreeding suggest considerable gene flow may be
281 occurring in *J. curcas* in Costa Rica. Although our samples did not cluster similarly across analysis,
282 a general lack of group structure was maintained throughout. The lack of consensus between
283 clustering algorithms may be attributed to the nature of the different markers used. We analyzed
284 multiple EST-SSR loci distributed throughout the genome (*Davies & Bermingham, 2002;*
285 *Pamidimarri & Reddy, 2014*), and we are confident that we have accounted for a significant
286 portion of the genetic variability in this species. Costa Rican diversity estimates may be improved
287 by enriching the germplasm bank with more accessions from the Caribbean and southern parts of
288 the country.
289
290 **Toxicity evaluation of *J. curcas***
291 Our results show that only one of the accessions, Mexican JCCR-24, had all of the alleles that
292 indicate non-toxicity. This accession was previously confirmed as non-toxic by *in vivo* evaluation
293 (*Basha & Sujatha, 2007*). JCCR-24 could be used as a parental plant in a breeding program to
294 obtain dual-purpose non-toxic plants, thereby increasing the attractiveness of *J. curcas* as a biofuel

295 plant (*King et al., 2013*). In other samples, only one or two toxic alleles were detected, depending
296 on the primers used. This may have been due to variations present in Costa Rican genotypes. SSR
297 markers are very polymorphic *(Powell et al., 1996)*. The primers used to detect non-toxicity were
298 developed from Mexican, Asian and African accessions (*Basha & Sujatha, 2007; Phumichai et*
299 *al., 2011*) and their ability to detect toxicity may differ for genotypes from other parts of the world.
300 It is possible that other Costa Rican accessions have alleles that have not yet been identified as
301 indicative of non-toxicity, and in this case, the non-toxic nature of the accessions would have been
302 overlooked. *In vivo* evaluations are needed to confirm this hypothesis.

303

304 **ACKNOWLEDGEMENTS**

305 We thank the Estación Experimental Fabio Baudrit Moreno (EEFBM) of the Universidad de Costa
306 Rica for the collection and maintenance of the germplasm bank.

307

308 **ADDITIONAL INFORMATION**

309 **DNA Submission to GenBank**

310 The following information was supplied regarding the submission of DNA sequences: GenBank
311 accession no.**:** KU561375, KU561376, KU561377, KU561378, KU561379, KU561380,
312 KU561381, KU561382, KU561383, KU561384, KU561385, KU561386, KU561387, KU561388,
313 KU561389, KU561390, KU561391, KU561393, KU561394, KU561395, KU561396, KU561397,
314 KU561398, KU561399, KU561400, KU561401, KU561402, KU561403, KU561404, KU561405,
315 KU561406, KU561407, KU561408, KU561409, KU561410, KU561411, KU561412, KU561413,
316 KU561414, KU561415, KU561416, KU561417, KU561418, KU561419, KU561420, KU561421,
317 KU561422, KU561423, KU561424.

318

319 **REFERENCES**

320 **Abdulla JM, Janagoudar BS, Biradar DP, Ravikumar RL, Koti RV, Patil SJ. 2009.** Genetic
321 diversity analysis of elite *Jatropha curcas* L. genotypes using randomly amplified polymorphic
322 DNA markers. *Journal of Agricultural Science* 22:293-295.

323 **Ambrosi DG, Galla G, Purelli M, Barbi T, Fabbri A, Lucretti S, Sharbel TF, Barcaccia G.**
324 **2010.** DNA Markers and FCSS Analysis Shed Light on the Genetic Diversity and Reproductive
325 Strategy of *Jatropha curcas* L. *Diversity* 2:810-836 DOI10.3390/d2050810.

326 **Avendaño R, García Díaz E, Valdez-Melara M, Chaves Solano N, Mora Villalobos A,**
327 **Aguilar Cascante F, Williamson Benavides B, Solís-Ramos L. 2015.** Genetic Diversity
328 Analysis of *Jatropha* Species from Costa Rica Using AFLP Markers. *American Journal of Plant*
329 *Sciences* 6:2426-2438 DOI10.4236/ajps.2015.614245.

330 **Basili M, Fontini F. 2012.** Biofuel from *Jatropha curcas*: Environmental sustainability and option
331 value. *Ecological Economics* 78:1-8 DOI10.1016/j.ecolecon.2012.03.010.

332 **Basha SD, Sujatha M. 2007.** Inter and intrapopulation variability of *Jatropha curcas* (L.)
333 characterized by RAPD and ISSR markers and development of population-specific SCAR
334 markers. *Euphytica* 156:375-386 DOI 10.1007/s10681-007-9387-5.

335 **Basha SD, Francis G, Becker K, Makkar HPS, Sujatha M. 2009.** A comparative study of
336 biochemical traits and molecular markers for assessment of relationships between *Jatropha curcas*
337 germplasm in China. *Biomass & Bioenergy* 7:1-12 DOI10.1016/j.plantsci.2009.03.008.

338 **Botstein D, White RL, Skolnick M, Davis RW. 1980.** Construction of a genetic linkage map in
339 man using restriction fragment length polymorphisms. *The American Journal of Human Genetics*
340 32:314-331.

341 **Bressan EA, Sebbenn AM, Rodrigues Ferreira R, Lee TSG, Figueira A. 2013.** *Jatropha curcas*
342 L. (Euphorbiaceae) exhibits mixed mating system, high correlate mating and apomixis. *Tree*
343 *Genetics & Genomes* 9:1089-197 DOI 10.1007/s11295-013-0623-y.
344 **Carels N. 2013.** Towards the domestication of *Jatropha*: The integration of science. In: Bahadur
345 B, Sujatha M, Carels N (ed) *Jatropha* Challenges for a New Energy Crop: Volume II. Springer,
346 New York, 263-299.
347 **China Plant BOL Group, De-Zhu L, Lian-Ming G, Hong-Tao L, Hong W, Xue-Jun G, Jian-**
348 **Quan L, Zhi-Duan C, Shi-Liang Z, Shi-Lin C, Jun-Bo Y, Cheng-Xin F, Chun-Xia Z, Hai-Fei**
349 **Y, Ying-Jie Z, Yong-Shuai S, Si-Yun C, Lei Z, Kun W, Tuo Y, Guang-Wen D. 2011.**
350 Comparative analysis of a large dataset indicates that internal transcribed spacer (ITS) should be
351 incorporated into the core barcode for seed plants. *PNAS* 108:19641-19646 DOI
352 10.1073/pnas.1104551108.
353 **Colombo C, Second G, Charrier A. 2000.** Diversity within American cassava germplasm based
354 on RAPD markers. *Genetics and Molecular Biology* 23:189–199.
355 **Cova V, Perini D, Soglio V, Komjanc M, van de Weg E, Glessler C, Gianfranceschi L. 2012.**
356 Exploiting expressed sequence tag database for mapping markers associated with fruit
357 development and fruit quality in apple. *Molecular Breeding* 29:699-715.
358 **Davies N, Bermingham E. 2002.** The historical biogeography of two Caribbean butterflies
359 (Lepidoptera: heliconiidae) as inferred from genetic variation at multiple loci. *Evolution* 56:573–
360 589.
361 **Ellis JR, Burke JM. 2007.** EST-SSRs as a resource for population genetic analysis. *Heredity*
362 99:125–132.

363 **Goldstein DB, Ruiz Linares A, Cavalli Sforza LL, Feldman MW. 1995.** An evaluation of
364 genetic distances for use with microsatellite loci. *Genetics* 139:463-471.
365 **Grativol C, da Fonseca Lira-Medeiros C, Hemerly AS, Ferreira PCG. 2010.** High efficiency
366 and reliability of inter-simple sequence repeats (ISSR) markers for evaluation of genetic diversity
367 in Brazilian cultivated *Jatropha curcas* L. accessions. *Molecular Biology Reports* 38:4245-4256
368 DOI 10.1007/s11033-010-0547-7.
369 **Heller J. 1996.** Physic Nut *Jatropha curcas* L. *Promoting the Conservation and use of*
370 *Underutilized and Neglected Crops.* Rome: IPGRI, 10–11.
371 **Henning RK. 2007.** *Jatropha curcas* L. In: van der Vossen HAM, Mkamilo GS, ed. Plant
372 Resources of the Tropical Africa: Volume I. Vegetable oils PROTA Fondation, Wageningen, The
373 Netherlands, 14:116–122.
374 **Insanu M, Dimaki C, Wilkins R, Brooker J, van der Linde P, Kayser O. 2013.** Rational use
375 of *Jatropha curcas* L. in food and medicine: from toxicity problems to safe applications.
376 Phytochemistry review 12:107–119.
377 **Islam A, Taufiq-Yap YH, Chu CH, Ravindra P, Chan ES. 2013.** Transesterification of palm
378 oil using KF and NaNO3 catalysts supported on spherical millimetric γ-Al2O3. *Renewable Energy*
379 59:23–29.
380 **Jongschaap REE, Corré WJ, Bindraban PS, Brandenburg, WA. 2007.** Claims and Facts. In:
381 *Jatropha curcas* L.: Global *Jatropha curcas* Evaluation, Breeding and propagation programme.
382 Report 158 Plant Research International BV, Wageningen, The Netherlands, 42.
383 **King AJ, Montes LR, Clarke JG, Affleck J, Li Y, Witsenboer H, van der Vossen E, van der**
384 **Linde P, Tripathi Y, Tavares E, Shukla P, Rajasekaran T, van Loo EN, Graham IA. 2013.**
385 Linkage mapping in the oilseed crop *Jatropha curcas* L. reveals a locus controlling the

386 biosynthesis of phorbol esters which cause seed toxicity. *Plant Biotechnology J* 11:986–996 DOI
387 10.1111/pbi.12092.

388 **King AJ, Montes LR, Clarke JG, Itzep J, Perez CA, Jongschaap RE, Visser RG, van Loo**
389 **EN, Graham I. 2015.** Identification of QTL markers contributing to plant growth, oil yield and
390 fatty acid composition in the oilseed crop *Jatropha curcas* L. *Biotechnology for Biofuels* 8:1-17
391 DOI 10.1186/s13068-015-0326-8.

392 **Lengkeek AG. 2007.** The *Jatropha curcas* agroforestry strategy of Mali Biocarburant SA. In:
393 Proceedings of the FACT Seminar on *Jatropha curcas* L. Agronomy and Genetics. Wageningen,
394 The Netherlands, 597.

395 **Liu K, Muse SV. 2005.** PowerMarker: an integrated analysis environment for genetic marker
396 analysis. *Bioinformatics* 21:228-2129 DOI 10.1093/bioinformatics/bti282.

397 **Maghuly F, Jankowicz-Cieslak J, Parbinger S, Till BJ, Laimer M. 2015.** Geographic origin is
398 not supported by the genetic variability found in a large living collection of *Jatropha curcas* with
399 accessions from three continents. *Biotechnology Journal* 10:536-551 DOI
400 10.1002/biot.201400196.

401 **Makker HPD, Becker K. 2015.** *Jatropha curcas*, a promising crop for the generation of biodisel
402 and value-added coproducts. *European Journal of Lipid Science and Technology* 11:773-787 DOI
403 10.1002/ejlt.200800244.

404 **Mastan S, Sudheer P, Rahman H, Ghosh A, Rathore M, Ravi Prakash C, Chikara J. 2012.**
405 Molecular characterization of intra-population variability of *Jatropha curcas* L. using DNA based
406 molecular markers. *Molecular Biology Reports* 39:4383-4390 DOI 10.1007/s11033-011-1226-z.

407 **Möller EM, Bahnweg G, Sandermann H, Geiger HH. 1992.** A simple and efficient protocol for
408 isolation of high molecular weight DNA from filamentous fungi, fruit bodies, and infected plant
409 tissues. *Nucleic Acids Research* 22: 6115-6116**.**

410 **Montes Osorio LR, Torres Salvador AF, Jongschaap REE, Azurdia C, Berduo J, Trindade**
411 **LM, Visser RGF, Loo EN van. 2014.** High level of molecular and phenotypic biodiversity in
412 *Jatropha curcas* from Central America compared to Africa, Asia and South America. *BMC Plant*
413 *Biology* 14:19 DOI10.1186/1471-2229-14-77.

414 **Nieto-Feliner G, Roselló JA. 2007.** Better the devil you know? Guideline for insightful utilization
415 of nrDNA ITS in species-level evolutionary studies in plants. *Molecular Phylogenetics and*
416 *Evolution* 44:911-919.

417 **Openshaw K. 2000.** A review of *Jatropha curcas*: an oil plant of unfulfilled promise. *Biomass &*
418 *Bioenergy* 19:1-15 DOI 10.1016/S0961-9534(00)00019-2.

419 **Ovando-Medina I, Espinosa GF, Nuñez FJ, Salvador FM. 2011.** Genetic variation in Mexican
420 *Jatropha curcas* L. estimated with seed oil fatty acids. *Journal of Oleo Science* 60:301-311.

421 **Ovando-Medina I, Adriano-Anaya L, Vásquez-Ovando A. 2013.** Genetic diversity of *Jatropha*
422 *curcas* in southern Mexico. In: Bahadur B, Sujatha M, Carels N (ed) *Jatropha* Challenges for a
423 New Energy Crop. Springer, New York, pp 263-299,

424 **Pamidimarri SDV, Chattopadhyay B, Reddy MP. 2008.** Genetic divergence and phylogenetic
425 analysis of genus *Jatropha* based on nuclear ribosomal DNA ITS sequence. *Molecular Biology*
426 *Reports* 36:1929-1935 DOI 10.1007/s11033-008-9401-6.

427 **Pamidimarri SDV, Reddy MO. 2014.** Phylogeography and molecular diversity analysis of
428 *Jatropha curcas* L. and the dispersal route revealed by RAPD, AFLP and nrDNA-ITS analysis.
429 *Molecular Biology Reports* 41:3225-34 DOI 10.1007/s11033-014-3185-7.

430 **Peakall R, Smouse PE. 2012.** GenAlEx 6.5: genetic analysis in Excel. Population genetic
431 software for teaching and research-an update. *Bioinformatics* 28:2537-2539.
432 **Pecina-Quintero V, Amaya JL, Zamarripa A, Montes N, Núñez C, Solís J. 2011.** Molecular
433 characterization of *Jatropha curcas* L. genetic resources from Chiapas, México through AFLP
434 markers. *Biomass & Bioenergy* 35:1897-1905 DOI 10.1016/j.biombioe.2011.01.027.
435 **Pecina-Quintero V, Anaya-Lopez JL, Zamarripa-Colmenero A, Nunez-Colin CA. 2014.**
436 Genetic structure of *Jatropha curcas* L. in Mexico and probable center of origin. *Biomass &*
437 *Bioenergy* 60:147–155 DOI 10.1016/j.biombioe.2013.11.005.
438 **Phumichai C, Phumichai T, Kongsiri N, Wongkaew A, Sripichit P, Kaveeta R. 2011.** Isolation
439 of 55 microsatellite markers for *Jatropha curcas* and its closely related species. *Biologia*
440 *Plantarum*
441 55:387-390.
442 **Powell W, Morgante M, Andre C, Hanafey M, Vogel J, Tingey S, Rafalski A. 1996.** The
443 comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis.
444 *Molecular Breeding* 2:225-238.
445 **Ram J, Kumar A, Bhatt J. 2004.** Plant diversity in six forest types of Uttaranchal, Central
446 Himalaya, India. *Current Science* 86:975-978.
447 **Raposo RS, Souza IGB, Veloso MEC, Kobayashi AK, Laviola BG, Diniz FM. 2014.**
448 Development of novel simple sequence repeat markers from a genomic sequence survey database
449 and their application for diversity assessment in *Jatropha curcas* germplasm from Guatemala.
450 *Genetics and molecular research* 13:6099-6106 DOI 10.4238/2014.August.7.25.
451 **Rosado TB, Laviola BG, Faria DA, Pappas MR, Bhering LL, Quirino B, Grattapaglia D.**
452 **2010.** Molecular markers reveal limited genetic diversity in a large germplasm collection of the

453 biofuel crop *Jatropha curcas* L. in Brazil. *Crop Science* 50:2372-2382 DOI
454 10.2135/cropsci2010.02.0112.
455 **Santorum JM, Darriba D, Taboada GL, Posada D. 2014.** *jmodeltest*.org: selection of nucleotide
456 substitution models on the cloud. *Bioinformatics* 30:1310–1311 DOI
457 10.1093/bioinformatics/btu032.
458 **Smith JSC, Chin ECL, Shu H, Smith OS, Wall SJ, Senior ML, Mitchell SE, Kresovich S,**
459 **Ziegle J. 1997.** An evaluation of the utility of SSR loci as molecular markers in maize (*Zea mays*
460 L.): comparisons with data from RFLPs and pedigree. *Theoretical and Applied Genetics* 95:163–
461 173.
462 **Song YP, Jiang XB, Zhang M, Wang ZL, Bo WH, An XM, Zhang DQ, Zhang ZY. 2012.**
463 Differences of EST-SSR and genomic-SSR markers in assessing genetic diversity in poplar.
464 *Forestry Studies in China* 14:1-7 DOI 10.1007/s11632-012-0106-5.
465 **Sun QB, Li LF, Li Y, Wu GJ, Ge XJ. 2008.** SSR and AFLP markers reveal low genetic diversity
466 in the biofuel plant *Jatropha curcas* in China. *Crop Science* 48:1865-1871 DOI
467 10.2135/cropsci2008.02.0074.
468 **Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013.** MEGA6: Molecular
469 Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30:2725-2729 DOI
470 10.1093/molbev/mst197.
471 **Tatikonda L, Wani PS, Kannan S, Beerelli N, Sreedevi KT, Hoisington AD. 2009.** AFLP based
472 molecular characterization of an elite germplasm collection of *Jatropha curcas* L., biofuel plant.
473 *Plant Science* 176:505-513.
474 **Tavaré S. 1986.** Some probabilistic and statistical problems in the analysis of DNA sequences. In:
475 American Mathematical Society. *Lectures on Mathematics in the Life Sciences* 17:57-86.

476 **To H, Grafton Q. 2015.** Oil prices, biofuels production and food security: past trends and future
477 challenges. *Food Security* 7:323-336 DOI 10.1007/s12571-015-0438-9
478 **van Oosterhout C, Hutchinson WF, Wills DP, Shipley P. 2004.** MICRO-CHECKER: software
479 for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes*
480 4:535-538 DOI 10.1111/j.1471-8286.2004.00684.x.
481 **Varshney RK, Graner A, Sorrells ME. 2005.** Genic microsatellite markers in plants: features
482 and applications. *Trends in Biotechnology* 23:48-55 DOI 10.1016/j.tibtech.2004.11.005.
483 **Vera-Castillo YB, Cuevas JA, Valenzuela-Zapata AG, Urbano B, Gonzalez-Andres F. 2014.**
484 Biodiversity and indigenous management of the endangered non-toxic germplasm of Jatropha
485 curcas L. in the Totonacapan (Mexico), and the implications for its conservation. *Genetic*
486 *Resources and Crop Evolution* 61:1263-1278 DOI: 10.1007/s10722-014-0109-2.
487 **Vischi M, Raranciuc S, Baldini M. 2013.** Evaluation of genetic diversity between toxic and non
488 toxic *Jatropha curcas* L. accessions using a set of simple sequence repeat (SSR) markers. *African*
489 *Journal of Biotechnology* 12: 265-274 DOI 10.5897/AJB12.2656.
490 **Wen M, Wang H, Xia Z, Zou M, Lu C, Wang W. 2010.** Development of EST-SSR and genomic
491 SSR markers to assess genetic diversity in *Jatropha curcas* L. *BMC Research Notes* 3:4210 DOI
492 10.1186/1756-0500-3-42**.**
493 **Yadav HK, Ranjan A, Asif MH, Mantri S, Sawant SV, Tuli R. 2011.** EST-derived SSR markers
494 in *Jatropha curcas* L. development, characterization, polymorphism, and transferability across the
495 species/genera. *Tree Genetics & Genomes* 7:207-219 DOI: 10.1007/s11295-010-0326-6.
496
497
498

499

500

List of tables

Table 1. Germplasm bank identification number, geographical collection sites and Genbank Accession number of nrDNA-ITS region for each of the *J. curcas* accessions in the Fabio Baudrit Experimental Station used in the study.

Table 2. EST-SSR and G-SSR primers used for evaluation of *Jatropha curcas* germplasm.

Table 3. Primers used to evaluate toxicity of accessions in the germplasm bank.

*NA: No amplification expected

Table 4. Parameters of genetic diversity and information contents of eighteen EST-SSR markers and one SSR used to analyze 50 *J. curcas* accessions.

*He: Heterozygosity, PIC: Polymorphism information content, f: inbreeding coefficient

List of figures

Fig. 1. Map of Costa Rica with sites where *Jatropha curcas* accessions were collected. A) North Pacific Region. B) Central Pacific Region.

Fig. 2. Neighbor-Joining (NJ) tree based on a genetic distance matrix from 19 microsatellite data from 50 germplasm accessions of *J. curcas*.

Fig. 3. Principal coordinate analysis created with GenAlEx based on the $\delta\mu^2$ genetic distance estimated in the POPULATIONS software with 19 microsatellites from 50 germplasm accessions of *J. curcas* from the Costa Rican germplasm bank evaluated in this study.

Fig. 4. Maximum likelihood phylogenetic tree generated with 60 *Jatropha curcas* samples. Analyzed sequences represent the nrDNA-ITS region. The tree was constructed with the GTR model using a jModelTest analysis with 3000 bootstraps and uniform substitution rates in the MEGA 6.0 software. Only bootstrap values higher than 50 % are shown. The bar indicates the

523 substitutions per site. Blue boxes show the sequences obtained from the current work in Costa
524 Rica. Other accessions were obtained from Genbank for comparison.

Figure 1

Map of Costa Rica

Map of Costa Rica with sites where *Jatropha curcas* accessions were collected. A) North Pacific Region. B) Central Pacific Region

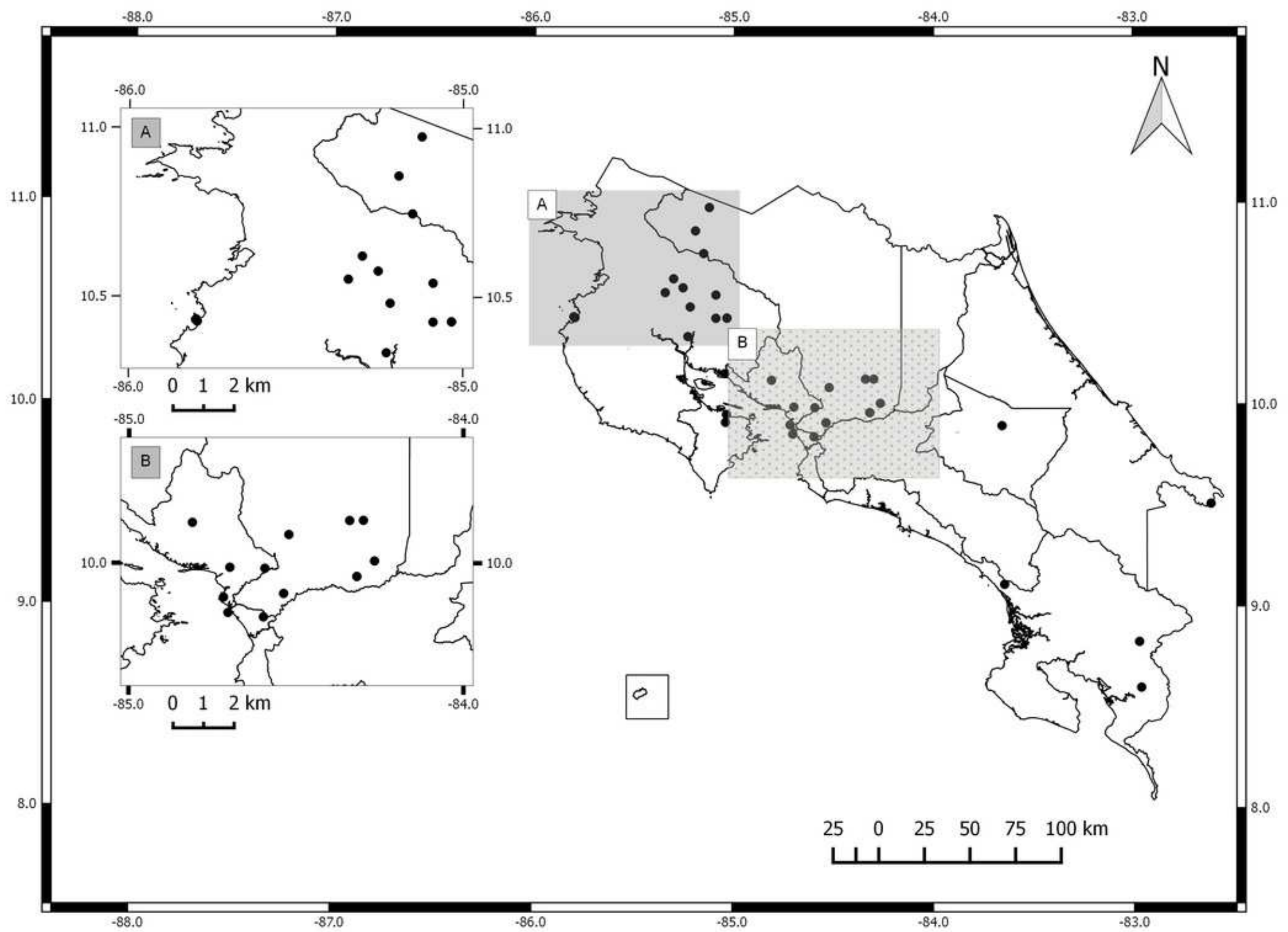

Figure 2

Neighbor-Joining (NJ) tree

Neighbor-Joining (NJ) tree based on a genetic distance matrix from 19 microsatellite data from 50 germplasm accessions of *J. curcas*.

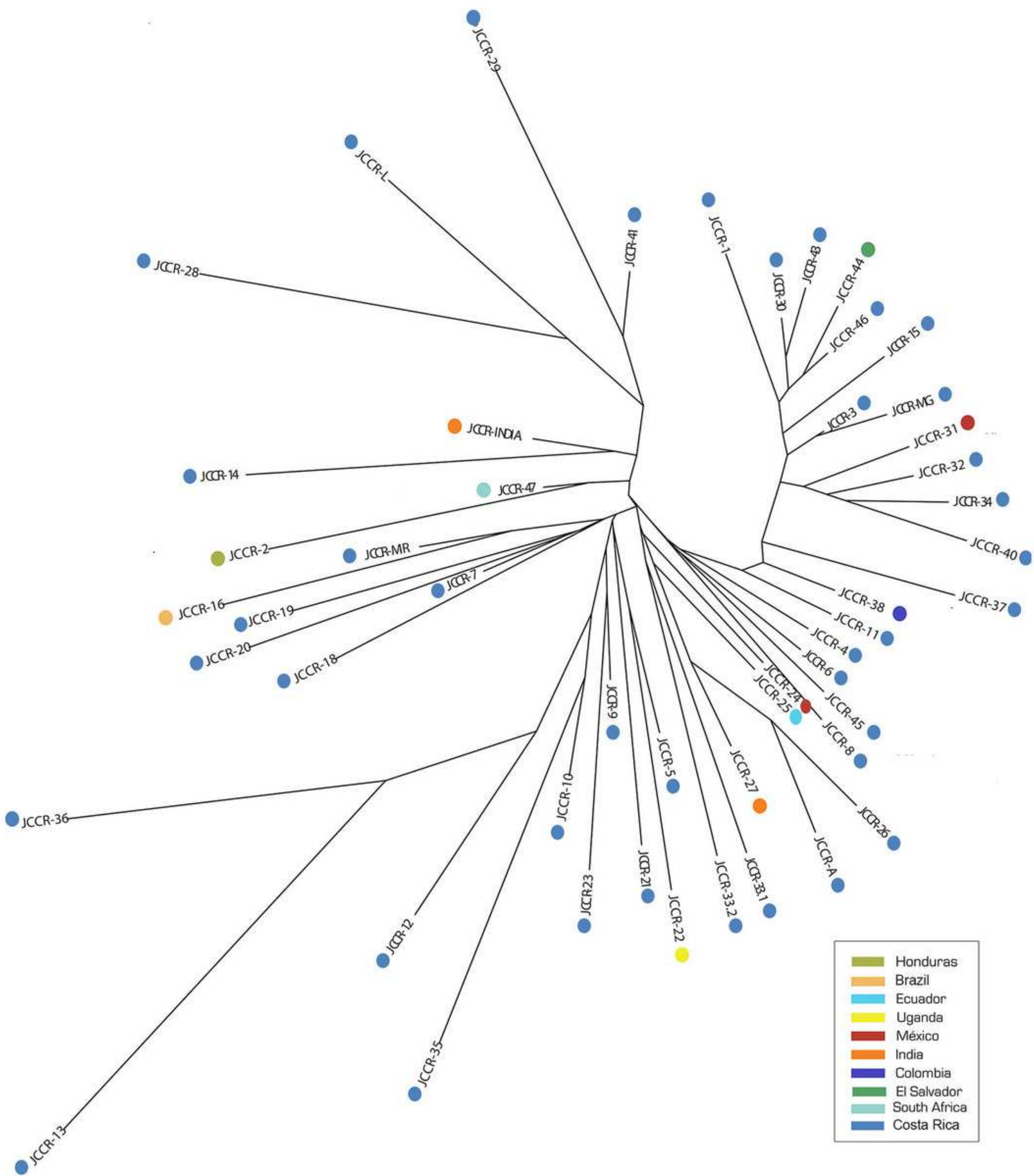
JCCR-29
JCCR-L
JCCR-41
JCCR-1
JCCR-28
JCCR-30
JCCR-43
JCCR-44
JCCR-46
JCCR-15
JCCR-3
JCCR-MG
JCCR-INDIA
JCCR-31
JCCR-14
JCCR-32
JCCR-47
JCCR-34
JCCR-40
JCCR-2
JCCR-MR
JCCR-7
JCCR-37
JCCR-38
JCCR-16
JCCR-19
JCCR-11
JCCR-20
JCCR-4
JCCR-18
JCCR-6
JCCR-24
JCCR-25
JCCR-45
JCCR-9
JCCR-8
JCCR-5
JCCR-27
JCCR-10
JCCR-26
JCCR-36
JCCR-A
JCCR23
JCCR-21
JCCR-33.1
JCCR-33.2
JCCR-22
JCCR-12
JCCR-35
JCCR-13
Honduras
Brazil
Ecuador
Uganda
México
India
Colombia
El Salvador
South Africa
Costa Rica

Figure 3

Principal coordinate analysis

Principal coordinate analysis created with GenAlEx based on the δμ^2 genetic distance estimated in the Populations software with 19 microsatellites from 50 germplasm accessions of *J. curcas* from the Costa Rican germplasm bank evaluated in this study.

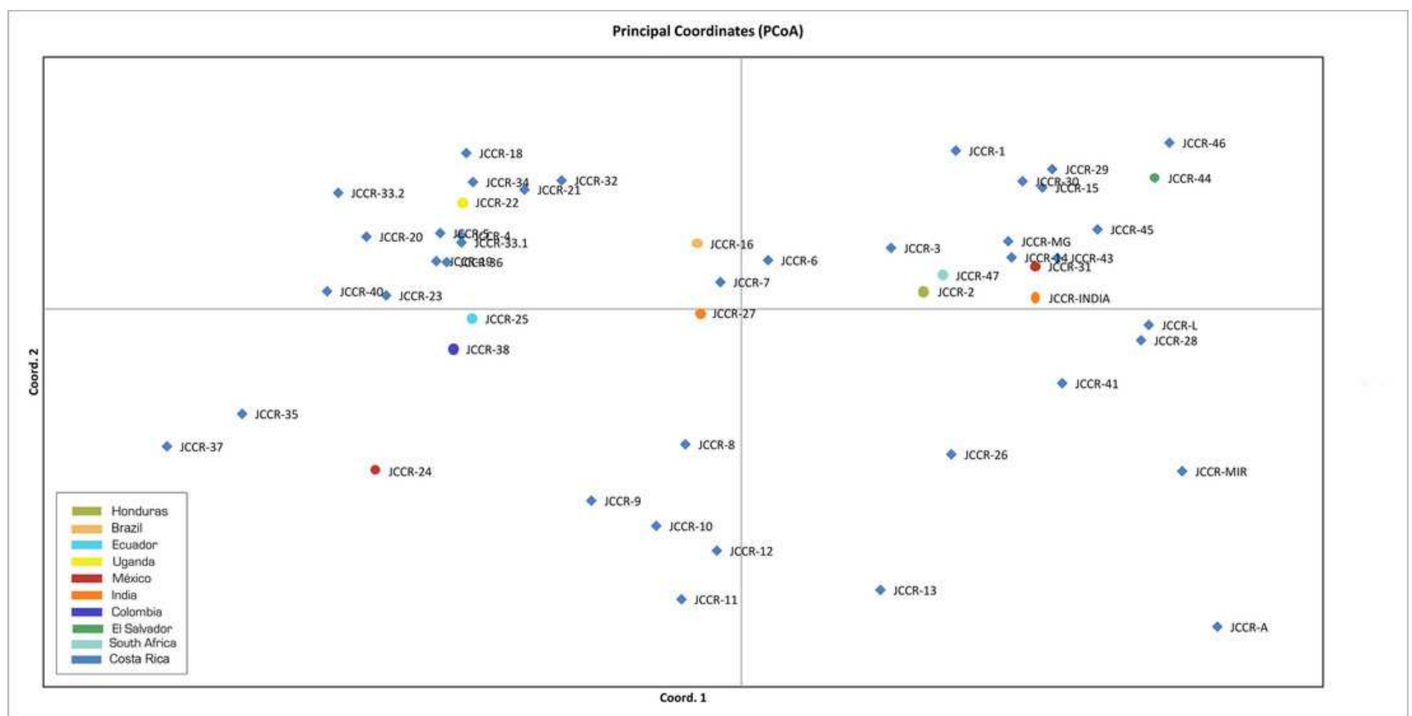

Figure 4

Maximum likelihood phylogenetic tree

Maximum likelihood phylogenetic tree generated with 60 *Jatropha curcas* samples. Analyzed sequences represent the nrDNA-ITS region. The tree was constructed with the GTR model using a jModelTest analysis with 3000 bootstraps and uniform substitution rates in the MEGA 6.0 software. Only bootstrap values higher than 50 % are shown. The bar indicates the substitutions per site. Blue boxes show the sequences obtained from the current work in Costa Rica. Other accessions were obtained from Genbank for comparison.

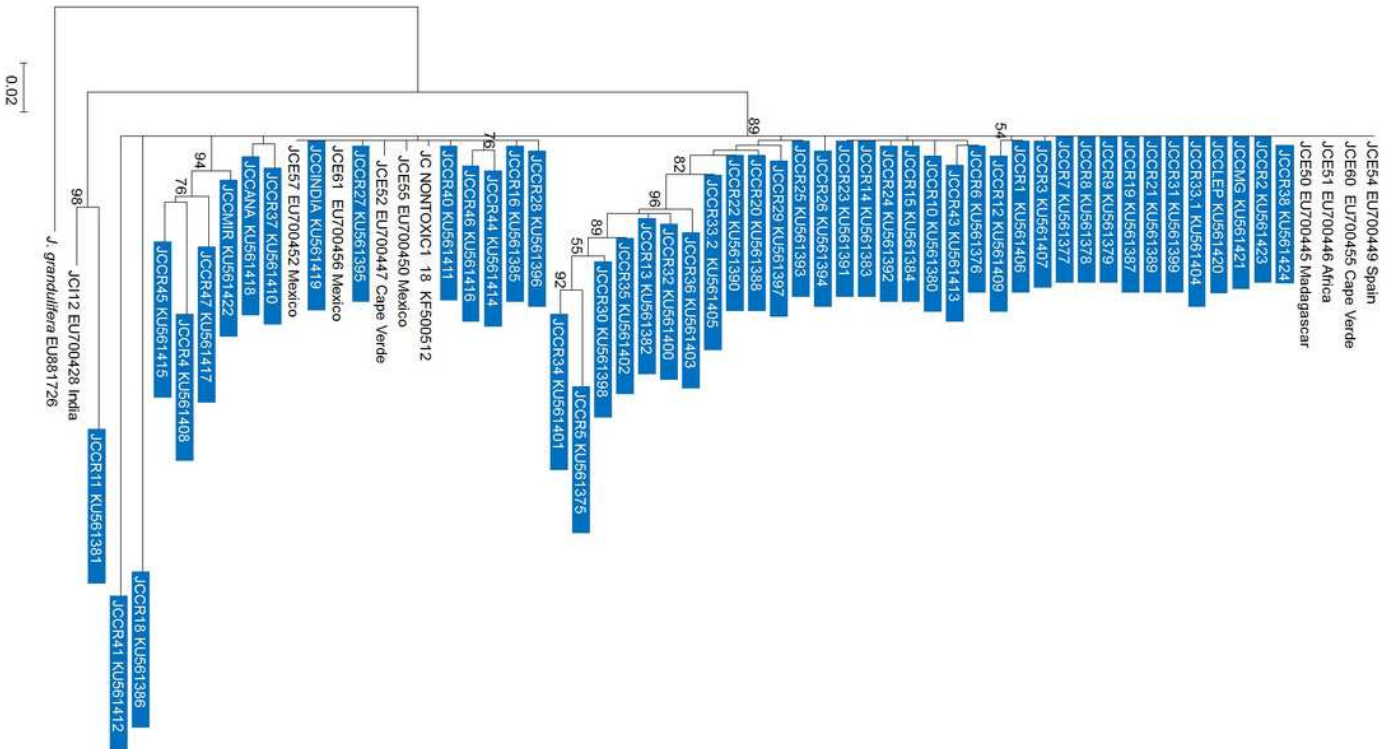

Table 1(on next page)

Germplasm bank information

Germplasm bank identification number, geographical collection sites and Genbank Accession number of nrDNA-ITS region for each of the *J. curcas* accessions in the Fabio Baudrit Experimental Station used in the study.

1

Germplasm code	Country	Location	Genbank accesion number
JCCR-1	Costa Rica	Orotina	KU561406
JCCR-3	Costa Rica	Orotina	KU561407
JCCR-4	Costa Rica	San Mateo	KU561408
JCCR-5	Costa Rica	Cañas	KU561375
JCCR-6	Costa Rica	San Antonio	KU561376
JCCR-7	Costa Rica	Cañas	KU561377
JCCR-8	Costa Rica	Abangares	KU561378
JCCR-9	Costa Rica	Guapinol	KU561379
JCCR-10	Costa Rica	Cañas	KU561380
JCCR-11	Costa Rica	San Antonio	KU561381
JCCR-12	Costa Rica	San Mateo	KU561409
JCCR-13	Costa Rica	Orotina	KU561382
JCCR-14	Costa Rica	Turrubares	KU561383
JCCR-15	Costa Rica	Abangares	KU561384
JCCR-18	Costa Rica	Bagaces	KU561386
JCCR-19	Costa Rica	Cañas	KU561387
JCCR-20	Costa Rica	San Vito	KU561388
JCCR-21	Costa Rica	Capulin	KU561389

JCCR-23	Costa Rica	San Antonio	KU561391
JCCR-26	Costa Rica	Capulin	KU561394
JCCR-28	Costa Rica	Upala	KU561396
JCCR-29	Costa Rica	Bagaces	KU561397
JCCR-30	Costa Rica	Coto 54	KU561398
JCCR-32	Costa Rica	Bagaces	KU561400
JCCR-33.1	Costa Rica	Los Santos	KU561404
JCCR-33.2	Costa Rica	Los Santos	KU561405
JCCR-34	Costa Rica	FabioBaudrit	KU561401
JCCR-35	Costa Rica	Abangares	KU561402
JCCR-36	Costa Rica	Turrubares	KU561403
JCCR-37	Costa Rica	Unknown	KU561410
JCCR-40	Costa Rica	Lagunilla	KU561411
JCCR-41	Costa Rica	CATIE	KU561412
JCCR-43	Costa Rica	Turrubares	KU561413
JCCR-45	Costa Rica	Unknown	KU561415
JCCR-46	Costa Rica	Diquis	KU561416
JCCR-MIR	Costa Rica	Miramar	KU561422
JCCR-ANA	Costa Rica	Anabel	KU561418
JCCR-MG	Costa Rica	Montaña Grande	KU561421
JCCR-LEP	Costa Rica	Lepanto	KU561420

JCCR-2	Honduras	Unknown	KU561423
JCCR-16	Brazil	Unknown	KU561385
JCCR-22	Uganda	Unknown	KU561390
JCCR-24	Mexico	Unknown	KU561392
JCCR-31	Mexico	Unknown	KU561399
JCCR-25	Ecuador	Manabí	KU561393
JCCR-27	India	Unknown	KU561395
JCCR-INDIA	India	Unknown	KU561419
JCCR-38	Colombia	Unknown	KU561424
JCCR-44	El Salvador	Unknown	KU561414
JCCR-47	South Africa	Unknown	KU561417

2

Table 2(on next page)

Primers used for evaluation of J. curcas

EST-SSR and G-SSR primers used for evaluation of *Jatropha curcas* germplasm.

1

ID	Forward Primer	Reverse Primer	T_A (°C)*	Expected size (bp)
JESR-001	AACCACAGGAGTTGGTAATG	GAAAGAAGCAACAGAAATGG	50	307
JESR-028	ACTTCCTTCAGATCATGCAC	CTGGGTAATCTTGTTCCAAA	52	292
JESR-047	GTTGATACTGGAAGTGAGCC	TGTGTTCAAAGGTGATGAGA	52	398
JESR-086	TCCCTCTCCTTCAGATTAAA	ATGATAGCCAAACAGCAACT	54	333
JESR-092	CTCTGAGAATTGAACCATCC	GGGAACAAAGAAATTACTGG	54	378
JESR-093	CACCTCCCATTAGGGTTT	CTAATCGACGCTGATAATCC	54	239
JESR-095	AATGAGTCTGACAATCAGGG	GCATGCTCTGTTCTGCTT	54	336
JESR-096	ACACAAACACAATCAACAGC	CGCGACTCACTTTGTATGTA	54	244
JESR-098	AGATCACAAGGATCACAAGG	GCAGTTGTCAAACACTAGCA	54	290
JESR-099	ATAATGGCAAACAAGTGGTC	TGGTAGTGTTGTTCTTGCAG	54	305
JESR-101	ATCCTAACACAGTTGCCATC	AAACTCAACCAAACCACAAC	54	230
JESR-102	ATCCTTCTGCAGTAGCCATA	TTATATGCTACACATCAACCTG	54	278
JESR-103	CAAGTTCGAGGAGTACAAGG	TGTTACAACGAGATGAGTGC	54	292
JESR-104	CCACAGTTCATCCTCAATTT	GATATTCACTCTGGAACCCA	54	308
JESR-118	CTAAAGGCTGTGAAGAAGGA	TCCGAGCCAATTTCTTATTA	54	276
JESR-161	AAGAAGTGTATGGGTTGCAC	TACGATACCTAGGGCTACGA	56	323
JESR-162	ACTGATGGGTATGTGAGAGG	TTCTTCATCATGGCTACCTT	56	220
JESR-163	CAGAAACGGAGAGGTCTG	AGATTGGAAGAGGAGAGGAG	56	144
JESR-164	AGCCCAGTCTCGCGGAAG	CAGTTCCCTTCAGAAGCTC	56	231
JESR-178	CTTTAGTCCACCTCAAGTGC	TGCAGCAATCAACTCTACTG	56	375
JSSR-203	ATCCTTGCCTGACATTTTGC	TTCGCAGAGTCCAATTGTTG	55	210

2 * T_A = annealing temperature

3

4

5

Table 3 (on next page)

Primers used to evaluate toxicity of accessions

Primers used to evaluate toxicity of accessions in the germplasm bank.

Primer ID	Primer sequences	T_m (°C)	Expected size non-toxic (bp)	Expected size toxic (bp)
JCT27	F: 5´-CATTAGAATGGACGGCTA-3´ R: 5´-GCGTGAAGCTTTGATTTGA-3´	60	259	253
JcSSR-26	F: 5´-CATACAAAGCCTTGTCC-3´ R: 5´-AACAGCATAATACGACTC-3´	55	210	230
JCT31	F: 5´-TGGAAAACGAATGAGGCTCT-3´ R: 5´-GGACACTCTGGAAAGGAACG-3´	59	214	208
ISPJ1	F: 5´-GAGAGAGAGAGAGAGGTG-3´ R-5´-GAGAGAGAGAGAGAAAACAAT-3´	54	NA*	543
ISPJ2	F- 5´GAGAGAGAGAGTTGGGTG-3´ R-5´AGAGAGAGAGAGCTAGAGAG-3´	54	1096	NA

*NA: No amplification expected

16

17

18

Table 4 (on next page)

Parameters of genetic diversity information obtained

1

SSR name	He*	PIC*	f*
JEST-01	0.375	0.551	0.411
JEST-28	0	0.043	0.001
JEST-47	0.050	0.048	-0.0129
JEST-86	0	0.336	0.001
JEST-92	0.814	0.370	-0.654
JEST-93	0.327	0.250	-0.160
JEST-95	0.217	0.175	-0.111
JEST-96	0.380	0.260	-0.225
JEST-98	0.043	0.114	0.650
JEST-99	0.280	0.236	-0.110
JEST-101	0.500	0.357	-0.062
JEST-102	0.325	0.235	-0.182
JEST-118	0.313	0.228	-0.175

JEST-161	0.043	0.042	-0.011
JEST-162	0.500	0.677	0.300
JEST-163	0.245	0.192	-0.129
JEST-164	0.721	0.355	-0.555
JEST-178	0.512	0.364	-0.058
JSSR-203	0.933	0.375	-0.862
Mean	0.346	0.274	-0.102
SD	0.062	0.165	0.346

* He: Heterozygosity, PIC: Polymorphism information content, f: inbreeding coefficient