Document downloaded from:

http://hdl.handle.net/10251/140861

This paper must be cited as:

Rêgo, T.; Elena-Jiménez, G.; Correia, KC.; Tovar-Pedraza J.M.; Câmara, MPS.; Armengol Fortí, J.; Michereff, SJ.... (06-2). Genetic diversity and population structure of Lasiodiplodia theobromae from different hosts in northeastern Brazil and Mexico. Plant Pathology. 68(5):930-938. https://doi.org/10.1111/ppa.12997

The final publication is available at

https://doi.org/10.1111/ppa.12997

Copyright Blackwell Publishing

Additional Information

This is the peer reviewed version of the following article: Rêgo, T.J.S., Elena, G., Correia, K.C., Tovar‐Pedraza, J.M., Câmara, M.P.S., Armengol, J., Michereff, S.J. and Berbegal, M. (2019), Genetic diversity and population structure of Lasiodiplodia theobromae from different hosts in northeastern Brazil and Mexico. Plant Pathol, 68: 930-938. doi:10.1111/ppa.12997 , which has been published in final form athttps://doi.org/10.1111/ppa.12997. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.

1 # Genetic diversity and population structure of *Lasiodiplodia theobromae* from
2 different hosts in northeastern Brazil and Mexico

3

4 **T. J. S. Rêgo[a], G. Elena[b], K. C. Correia[c], J. M. Tovar-Pedraza[d], M. P. S. Câmara[a], J.**
5 **Armengol[b], S. J. Michereff[c] and M. Berbegal[b*]**

6

7 [a]*Departamento de Agronomia, Universidade Federal Rural de Pernambuco, 52171-900 Recife,*
8 *Brazil;* [b]*Instituto Agroforestal Mediterráneo, Universitat Politècnica de València, 46022*
9 *Valencia, Spain;* [c]*Centro de Ciências Agrárias e da Biodiversidade, Universidade Federal do*
10 *Cariri, 63133-610 Crato, Brazil; and* [d]*Coordinación Culiacán, Centro de Investigación en*
11 *Alimentación y Desarrollo, 80110, Culiacán, Sinaloa, Mexico*

12

13 [*]Corresponding author: Mónica Berbegal - e-mail: mobermar@etsia.upv.es
14 Telephone number: +34 963879254
15 Fax number: +34 963879269
16

17 ## Abstract

18

19 *Lasiodiplodia theobromae* is one the most frequent fungal pathogens associated with dieback,
20 gummosis, leaf spot, stem-end rot, and fruit rot symptoms, in cashew, mango, papaya and
21 grapevine. In this study, the variation in the genetic diversity of 117 *L. theobromae* isolates
22 from northeastern Brazil (n=100) and Mexico (n=17), which were collected from these four
23 crops, was analyzed using microsatellite markers. The results revealed low genetic diversity
24 among *L. theobromae* populations and the existence of two genetic groups. All Mexican isolates
25 were grouped with Brazilian isolates suggesting a low level of differentiation between these
26 populations. Furthermore, no evident host or climate-based population differentiation was
27 observed for *L. theobromae* in Brazil. The populations studied were mostly clonal, but
28 additional studies are needed to better understand the mode of reproduction of the pathogen.
29 The low genetic diversity of *L. theobromae* populations in northeastern Brazil suggests that
30 resistant cultivars could be used as a durable management strategy to reduce the impact of the
31 diseases caused by this pathogen.

32

33 *Keywords*: Microsatellites, population biology, SSR, variability

34

35 ## Introduction

36

37 Brazil ranks third in the world in fruit producing countries with more than 37 million t (FAO
38 2018), representing US$ 9.5 billion in value (IBGE 2018). Fruit crops are widely grown in the
39 country, but the northeastern states are among the largest producers. In 2016, the northeastern
40 region of Brazil produced 8.7 million t of fruit, corresponding to 23.5% of the national
41 production (IBGE 2018). Of this total, 2.1 million t were obtained from cashew, mango, papaya,

42 and grapevine, which are the most economically relevant fruit crops grown in the region (IBGE
43 2018).

44 In recent years, fungal diseases affecting cashew, mango, papaya, and grapevine crops have
45 been reported in northeastern Brazil, with prevalence of infections caused by members of the
46 family Botryosphaeriaceae (Netto *et al*. 2014; Correia *et al*. 2016; Coutinho *et al*. 2017).
47 Botryosphaeriaceae species are important pathogens of native and non-native plants worldwide
48 (Slippers & Wingfield 2007; Philips *et al*. 2013). For instance, research on the identification
49 and prevalence of Botryosphaeriaceae spp. in fruit producing orchards and at post-harvest
50 conditions in northeastern Brazil showed that *Lasiodiplodia theobromae* (Pat.) Griff. & Maubl.
51 (syn. *Botryodiplodia theobromae* Pat.) is present in all areas studied, being the most frequent
52 species associated with dieback, gummosis, leaf spot, stem-end rot, and fruit rot symptoms
53 (Netto *et al*. 2014; Correia *et al*. 2016; Coutinho *et al*. 2017).

54 *Lasiodiplodia theobromae* is widely distributed in the tropics and subtropics, causing
55 important diseases in several species of host plants (Philips *et al*. 2013). This fungus, known as
56 a latent pathogen, has an endophytic infection ability without exhibiting symptoms in colonized
57 plant tissues until conditions are favorable, making its control difficult (Mohali *et al*. 2005;
58 Slippers & Wingfield, 2007). In Brazil, there are few products registered for the management
59 of *L. theobromae*, and are restricted to avocado (*Persea americana*), cotton (*Gossypium* spp.),
60 *Annona* spp., banana (*Musa* spp.), cacao (*Theobroma cacao*), citrus (*Citrus* spp.), coconut
61 (*Cocos nucifera*), and papaya (*Carica papaya*) crops.

62 Studies on population genetics often provide information on the biology and evolution of
63 plant pathogens, sometimes contributing directly to disease management (Milgroom 2015).
64 Knowledge obtained from genetic analysis of plant pathogen populations may be relevant for
65 breeding and fungicide resistance management programs (McDonald & Linde 2002).
66 Population genetics can be used to identify genetic patterns and processes of pathogen
67 emergence and reemergence (Grünwald *et al*., 2017). A better understanding of these processes

68 is necessary to predict and mitigate their impact. Furthermore, population genetics can also be
69 useful to infer the relative contribution of asexual and sexual reproduction in pathogen
70 populations (Milgroom 1996).

71 On a global scale, genetics of *L. theobromae* populations have been studied primarily using
72 microsatellite or simple-sequence repeat (SSR) markers (Mohali *et al*. 2005; Begoude *et al*.
73 2012; Shah *et al*. 2011). SSR markers have become very valuable for population genetic
74 analysis because they are highly polymorphic, codominant, and locus-specific (Milgroom
75 2015). Rapidly mutating markers like SSRs offer unique advantages since genetic divergence
76 in contemporary populations and within closely related individuals can be observed when it is
77 not visible using other marker systems (Varshney *et al*. 2005). In a preliminary study, Burgess
78 *et al*. (2003) developed eight polymorphic SSR markers for *L. theobromae* and determined that
79 host species seemed to influence the population structure of this pathogen.

80 Other genetic diversity studies of *L. theobromae* in South Africa, Mexico, Venezuela and
81 Punjab suggested that this species predominantly reproduces asexually (Mohali *et al*. 2005;
82 Shah *et al*. 2011). However, there are still few studies involving the population genetics of this
83 pantropical pathogen (Burgess *et al*. 2003; Shah *et al*. 2011; Begoude *et al*. 2012, Mehl *et al*.,
84 2017). Although recent research has studied the identity, distribution, and pathogenicity of *L.*
85 *theobromae* in Brazil (Netto *et al*. 2014; Correia *et al*. 2016; Coutinho *et al*. 2017), currently
86 there is a single study about the genetic diversity and structure of this pathogen isolated from
87 coconut trees (*Cocos nucifera*) populations at a regional level, and compared with other
88 countries.

89 In the present study, the variation in the genetic diversity of *L. theobromae* in northeastern
90 Brazil and Mexico was analyzed using microsatellite markers with the following objectives: (i)
91 to determine the structure and potential subdivision among and within populations based on the
92 country of origin, and based on host and/or climate type within northeastern Brazil and, (ii) to
93 infer about the relative contribution of the reproduction mode of the pathogen on the genetic

94 diversity observed. This information will provide critical insights into the pathogen biology and
95 evolution with subsequent implications for disease management.
96

97 Materials and methods

98

99 Fungal isolates and DNA extraction

100

101 A total of 117 *L. theobromae* isolates from Brazil and Mexico were genotyped in this study
102 (Supplementary Table 1). Isolates from Brazil (n = 100) and Mexico (n = 17) were obtained
103 from the Culture Collection of Phytopathogenic Fungi "Prof. Maria Menezes" (CMM) of the
104 Universidade Federal Rural de Pernambuco, Brazil. The identity of all isolates was confirmed
105 by sequencing a partial region of the translation elongation factor 1-alpha (EF1-α) and the
106 rDNA ITS region as described in previous studies (Netto *et al*. 2014; Correia *et al*. 2016;
107 Coutinho *et al*. 2017). The isolates from Brazil were collected between 2006 and 2013 from
108 cashew (n = 27), mango (n = 21), papaya (n = 29) and grapevine (n = 23), and the isolates from
109 Mexico were collected in 2014 from mango (n = 17) (Supplementary Table 1).
110 The isolates from Brazil were grouped into two populations based on climate types according
111 to Köppen's climate classification map for Brazil (Alvares *et al*. 2013): tropical with dry
112 summer (As) and semiarid (BSh). The As climate is located in coastal regions with a strong
113 rainfall gradient (east to west), from 1,500 to 700 mm concentrated in a rainy season from May
114 to July and a dry season from September to December. The BSh climate is dry, with high indices
115 of insolation during most of the year with a drought period of nine months and a rainy period
116 from February to April (Alvares *et al*. 2013). Both climate types-based populations included
117 isolates from five states of northeastern Brazil: As climate (Ceará, Paraíba, Pernambuco, and
118 Rio Grande do Norte) and BSh climate (Bahia, Pernambuco, and Rio Grande do Norte)

119 (Supplementary Table 1). The Mexican isolates were obtained from mango orchards located in
120 four states (Chiapas, Guerrero, Sinaloa, and Veracruz) (Supplementary Table 1).
121 Pure cultures of *L. theobromae* were grown in potato-dextrose-agar (PDA, Biokar
122 diagnostics, Beauvais, France) medium and incubated at 25°C in the dark for a period of five
123 days. Mycelium was scraped from the colony surface using a sterile pipette tip and genomic
124 DNA was extracted using the E.Z.N.A. Plant Miniprep Kit (Omega Biotek®, Doraville, GA)
125 following the manufacturer's instructions.
126

127 **Haplotype identification**

128

129 Primers for eight previously characterized SSR loci (LAS3&4, Las13&14, LAS15&16,
130 LAS17&18, LAS21&22, LAS27&28, LAS29&30, and LAS 35&36) (Burgess *et al*. 2003) were
131 tested for specific amplification, polymorphism and reproducibility of the allele calls with a
132 selection of 11 *L. theobromae* isolates representing different origins and hosts. Primers were
133 synthesized by Metabion International (Berlin, Germany). All amplifications were performed
134 in 20 µl reaction volume containing 1x PCR buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.2 µM
135 each primer, 0.05 U Horse-Power™ Taq DNA polymerase (Canvax Biotech, S.L. Córdoba,
136 Spain), and 1 µl of DNA template. PCR amplifications were performed on a Veriti 96 Well
137 Thermal Cycler (Applied Biosystems, Foster City, CA) using the following conditions: an
138 initial step of 5 min at 94°C, followed by 35 cycles of denaturation at 94°C for 30 s, annealing
139 at 56°C for 30 s, elongation at 72°C for 30 s and a final extension at 72°C for 6 min.
140 Amplification products were purified (Kit Quickclean, dominion MBL, Córdoba, Spain) and
141 sequenced to validate the variation in the number of SSR repeats at the sequencing service of
142 the Institute of Molecular and Cellular Biology of Plants (IBMCP, UPV-CSIC).
143 Fluorescent dye labeling of PCR fragments was performed for each for each SSR locus using
144 three primers in each reaction: SSR forward primer with M13 tail

145 (TGTAAAACGACGGCCAGT) at its 5′ end, SSR reverse primer, and the fluorochrome 6-
146 carboxyfluorescein (FAM)-labeled M13 primer (TGTAAAACGACGGCCAGT) (Invitrogen,
147 Carlsbad, CA), as described by Schuelke (2000) with minor modifications (Berbegal *et al*.
148 2013). PCR reactions (20 µl) were performed with the following final concentrations: 1x PCR
149 buffer, 0.2 mM dNTPs, 0.04 µM M13-SSRf primer, 0.16 µM SSRr primer, 0.16 µM FAM-M13
150 primer, 0.05 U Horse-Power™ Taq DNA polymerase, and 1 µl of DNA template. The PCR
151 program consisted of an initial step of 2 min for 94°C followed by 35 cycles of denaturation
152 94°C for 30 s, 56°C for 30 s, 72°C 30 s, and final extension 72°C for 6 min.

153 Samples were prepared in 96-well plates using 0.1 µl of GeneScanTM500 LIZ® internal
154 standard (Applied Biosystems, Foster City, CA), 10.9 µl of Hi-Di Formamide (Applied
155 Biosystems), and 1 µl of the SSR-PCR product. A negative control was included in each plate.
156 Capillary electrophoresis to determine the sizes of the fluorescent labeled fragments was
157 performed on a 3130xl/Genetic Analyzer® sequencer (Applied Biosystems) at the sequencing
158 service of the Institute of Molecular and Cellular Biology of Plants (IBMCP, UPV-CSIC).
159 Allele sizes were determined using the GeneMarker V2.6.3 software package (Applied
160 Biosystems). The analysis was replicated for a subset of 11 isolates to confirm reproducibility
161 of results.

162

163 **Data analysis**

164

165 *Genetic diversity and mode of reproduction*. For each population defined either by country of
166 origin or climate within northeastern Brazil, the total number of alleles and allele frequencies
167 at each SSR locus were estimated. A multilocus genotype (MLG) was constructed for each
168 isolate by combining data for single SSR alleles and expected multilocus genotype (eMLG)
169 based on rarefaction were calculated using the R package poppr V.2.3.0 (R Core Team 2017;
170 Kamvar *et al*. 2015). Given the clonality observed analyses were conducted for the clone-

171 corrected data set, only one isolate of each MLG for each geographical location and host by
172 population were considered. The Stoddart and Taylor's diversity index (G) (1988) and evenness
173 index E_5 (Grünwald *et al*. 2003) were calculated using the same R package.

174 The standardized index of association (rbarD) as an estimate of linkage disequilibrium, was
175 calculated to investigate the mode of reproduction (Agapow & Burt 2001; Kamvar *et al*. 2015).
176 The expectation of rbarD for a randomly mating population is zero and significant deviation
177 from this value would suggest clonal reproduction. The test was not performed in Mexico
178 population with a very low sample size after clone correction due to lack of statistical power
179 (Fincham & Day, 1963). Significance was tested based on 999 permutations and conducted in
180 the R package poppr using the clone corrected data (Kamvar *et al*. 2015).

181 *Population structure*. The standardized measure of genetic differentiation G´st described by
182 Hedrick (2005) is suitable to estimate subdivision among populations. It ranges from 0 to 1,
183 independent of the extent of population genetic variation and loci mutation rates (Hedrick,
184 2005). Pairwise G´st values within the clone corrected data were calculated using the R
185 packages strataG V.1.0.5 (Archer *et al*. 2017) and mmod V.1.3.3 (Winter 2012). Statistical
186 significance was calculated based on 1000 permutations. Analysis of Molecular Variance
187 (AMOVA) was performed using the R package ade4 V.1.7-5 (Dray & Dufour 2007; Excoffier
188 *et al*. 1992). AMOVA was calculated on clone corrected data to estimate the variance explained
189 by populations defined by country of origin or individuals within populations. Statistical
190 significance was tested using 999 permutations.

191 Discriminant analysis of principal components (DAPC) was performed to infer clusters of
192 populations without considering previous geographic/climate types-based assignment criteria.
193 This approach is based on sequential K-means clustering of principal component analysis
194 (PCA) and discriminant analysis (DA) (Jombart *et al*. 2010). DAPC was conducted with the R
195 package adegenet V. 2.0.1 (Jombart 2008) using the Bayesian information criterion (BIC) to
196 infer the optimal number of groups. Important advantages of the DAPC are that it maximizes

197 variation between the groups, minimizing the within-group genetic variability and it does not
198 require assumptions regarding evolutionary models (Jombart *et al*. 2010).
199 To assess the relationships among MLGs, minimum spanning network (MSN) was
200 constructed from the clone-corrected dataset including isolates from Brazil and Mexico.
201 Bruvos´s genetic distance matrix and MSN were generated using the R package poppr V.2.3.0
202 (Kamvar *et al*. 2015; R Core Team 2017). Genetic distance described by Bruvo et al. (2004),
203 takes SSR repeat number into account being a distance of 0.1 equivalent to one mutational step
204 (one repeat). The neighbor-joining algorithm based on Bruvo's distance was also run using the
205 same R package with 1,000 bootstrap replicates. In the resulting tree, DAPC clustering was
206 represented to visualize how both clustering patterns hold up.
207
208 **Results**
209
210 Six of the eight primer pairs successfully amplified SSR loci for *L. theobromae* from Brazil and
211 Mexico. Allele sizes determined by DNA sequencing of 11 representative isolates were
212 consistent with allele sizes determined by fragment analysis. In addition, fragment analysis
213 replication for these 11 isolates showed consistent results, confirming reliability and
214 reproducibility of the method. Three loci (Las3&4, Las21&22, and Las35&36) were
215 polymorphic for the 117 isolates genotyped. The number of observed alleles per locus was 4
216 resulting in a total of 11 MLGs (Table 1). The Brazilian population exhibited 5 MLGs, the
217 Mexican population 4 MLGs, and both populations shared 2 MLGs (Figure 1A).
218 A clone correction of data was performed to remove the bias of resampled MLG in the
219 analysis, resulting in 42 representative isolates of the two climate types (17 in BSh and 25 in
220 As) of northeastern Brazil and 10 representative isolates of Mexico. Of the 6 MLGs observed
221 in the Mexican population, the most frequent one was MLG11 (40%), followed by MLG7
222 (20%); the remaining 4 MLGs were observed once (10%) (Figure 1A). In Brazil, 6 (MLG1,

223 MLG3, MLG4, MLG5, MLG6, and MLG7) of the 7 MLGs were present in the two climate
224 types studied (Figure 2B). The most frequent MLG was MLG7 (47.6%), followed by MLG6
225 (21.4%), and the less frequent MLGs were MLG1 and MLG2 (4.8%), with genotype MLG2
226 exclusive of the As climate population (Figure 2B).

227 In general, populations showed low genetic diversity. The genotypic richness (R) for the
228 Mexican population was higher (0.56) than was observed for the Brazilian population (0.15)
229 (Table 1). Evenness (E_5) values were 0.66 and 0.79 for Brazilian and Mexican populations,
230 respectively (Table 1). A high level of genotypic diversity (G) was observed for the Mexican
231 population (4.17), whereas the Brazilian population exhibited lower diversity (3.42) (Table 1).
232 When considering the Brazilian populations defined by climate types, R and E_5 values were
233 low and similar for both populations, ranging from 0.31 to 0.25 and from 0.71 to 0.66, for BSH
234 and As populations respectively (Table 1). Genotypic diversity values observed were also low
235 for populations of different climate types in Brazil (3.32 and 3.42 for BSh and As populations,
236 respectively) (Table 1).

237 Global population from Brazil showed significant deviation in rbarD value from the null
238 hypothesis of recombination supporting clonal reproduction (Table 1). However, when
239 considering populations defined by climate types in Brazil, As population showed consistent
240 results with the global population but for BSh population the null hypothesis could not be
241 rejected (P = 0.016) indicating linkage equilibrium or presence of sexual reproduction (Table
242 1).

243 The minimum spanning network (MSN) revealed two main groups when considering the
244 global population of Brazil and Mexico. The first group included MLGs of both countries and
245 the second group only two MLGs exclusive of Brazil (MLG3 and MLG4) (Figure 1A). The
246 MSN performed between the isolates belonging to the two different climate types of
247 northeastern Brazil showed one main group which was present in both populations including

248 the most frequent MLGs (MLG6 and MLG7), and a second small group with two MLGs being
249 one MLG exclusive of the As population (MLG2) (Figure 1B).

250 The Bruvo´s distance-based dendrogram using Neighbor-Joining grouped the 11 genotypes
251 of the global population into two groups. The cluster I consisted of two MLGs and cluster II
252 included nine MLGs. Multivariate analysis showed the presence of five subgroups represented
253 by colors in the dendrogram, where cluster I grouped one subgroup (green) and cluster II
254 grouped four (purple, blue, black and red) (Figure 2A). For northeastern Brazil populations,
255 two main groups were also identified. Cluster I was represented by two MLGs (MLG1 and
256 MLG2), whereas the rest of the MLGs were grouped in the Cluster II. The multivariate analysis
257 showed four subgroups in the dendrogram, where cluster I grouped one subgroup (blue) and
258 cluster II grouped three (red, green, and purple) (Figure 2B).

259 Pairwise G´st values calculated on the clone corrected data showed very low genetic
260 differentiation among Mexico and Brazil populations and within Brazilian populations (G´st =
261 0.104 and G´st = 0.014, respectively, $P > 0.01$). Results of population subdivision analysis
262 based on G´st were consistent with those obtained by the AMOVA. Analysis of molecular
263 variance on the clone corrected data revealed only 3.4% of variation between Mexican and
264 Brazilian populations ($P = 0.16$). None of the calculated values were statistically significant
265 showing that further sampling is likely needed.

266

267 Discussion

268

269 To the best of our knowledge, this is the first study that aims to investigate the genetic structure
270 and the mode of reproduction of a collection of *L. theobromae* isolates obtained from different
271 hosts in Brazil and Mexico. These populations were defined based on the country of origin and

272 climate types within northeastern Brazil, and the analysis was performed using the
273 microsatellite markers developed by Burgess *et al*. (2003).

274 Minimum spanning network analysis and discriminant analysis of principal components
275 identified two distinct genetic clusters in the Brazilian and Mexican populations of *L.*
276 *theobroma*e, the first one with nine MLGs including the two most common MLGs identified in
277 this study (MLG11 and MLG10), and the second one containing two MLGs only found in Brazil
278 (MLG3 and MLG4). A limitation of this work that need to be stated is the unbalanced number
279 of isolates from Mexico, which ideally should have represented more hosts than only mango.
280 Studies based on culture collections are sometimes associated with this kind of issues. However,
281 the ideal strategy of sampling and sample size are often difficult to put in practice and depend
282 on many factors including the questions being addresses, the type of genetic markers used and
283 the resources available to support the project (Milgroom, 2015)

284 All Mexican isolates were grouped with Brazilian isolates suggesting the existence of a low
285 level of differentiation between these populations. These results agree with those obtained by
286 Burgess *et al*. (2003) indicating that geographical location did not determine the genetic
287 relationship between isolates of *L. theobromae*. No evident population differentiation based on
288 host or climate was observed for *L. theobromae* in Brazil. These results are consistent with
289 those found in previous studies where the relationships between *L. theobromae* isolates from
290 different hosts and geographical origins did not contribute in the differentiation of pathogen
291 haplotypes (Begoude *et al*. 2012; Mohali *et al*. 2005).

292 *Lasiodiplodia theobromae* spores develop in a sticky matrix, are relatively large (21.5– 31.5
293 × 13–17 µm) (Phillips *et al*. 2013), and can be dispersed by wind or rain splash (Úrbez-Torres
294 2011). In this context, the spores are not likely to be spread over large distances (Mehl *et al*.
295 2017). A likely explanation for the genetic similarity among the studied populations would be
296 the movement of infected plant material (Begoude *et al*. 2012). *Lasiodiplodia theobromae* is
297 recognized as a latent pathogen that could be dispersed when seeds or asymptomatic plant

298 material are moved between different regions (Burgess *et al*. 2016; Mohali *et al*. 2005). In a
299 recent study, Mehl *et al*. (2017) concluded that *L. theobromae* isolates collected from different
300 hosts and countries of the world represent a single species distributed worldwide, without
301 apparent phylogeographic structure, and suggested that the most likely explanation for this
302 result is the long-distance human dispersal of this pathogen.

303 Seven MLGs were obtained using the SSR markers in the populations of *L. theobromae* from
304 Brazil. The lack of diversity found in this study emphasizes the low genetic diversity of this
305 fungus in northeastern Brazil. This is contrast with results about population genetics of *L.*
306 *theobromae* from coconut trees in Brazil and other countries based on EF1-α sequences analysis
307 (Santos *et al*., 2017). In this study, relatively high levels of genetic diversity were observed in
308 population from northeastern and southeast Brazil. However, for this specific and according to
309 authors, the variability of *L. theobromae* haplotypes would be associated to a selection process
310 experienced by the pathogen due high host diversity and control management strategies adopted
311 by different regions in Brazil (Santos *et al*., 2017).

312 Some other studies have analyzed the genetic structure of *L. theobromae* populations from
313 different countries and hosts with variable results. Begoude *et al*. (2012) reported high to
314 moderate levels of genetic diversity of *L. theobromae* collected from *Theobroma cacao* and
315 *Terminalia* spp. in Cameroon. Mohali *et al*. (2005) reported low levels of genotypic diversity
316 of *L. theobromae* collected from *Pinus caribaea* var. *hondurensis*, *Eucalyptus urophylla* and
317 *Acacia mangium* in Venezuela, and Shah *et al*. (2011) reported high levels of genetic diversity
318 of *L. theobromae* collected from *Pyrus* sp. in India. These variable results might be related to
319 the intrinsic characteristics of the different countries and hosts used in each study.

320 *Botryosphaeria rhodina* (Berk. & M. A. Curtis) Arx. was reported as the teleomorph of *L.*
321 *theobromae*. However, the connection between *L. theobromae* and its sexual morph has not
322 been proven conclusively (Phillips *et al*. 2013). Furthermore, in recent years, cryptic species
323 have been described from isolates previously identified as *L. theobromae* (Burgess *et al*. 2006;

324 Alves *et al*. 2008; Cruywagen *et al*. 2017), leading to uncertainty on the accuracy of information
325 regarding the sexual status of *L. theobromae* (Begoude *et al*. 2012). In the present study, linkage
326 disequilibrium analysis showed evidence of recombination in *L. theobromae* population of the
327 BSh climate type. However, in nature, it is difficult to observe sexual structures of *L.*
328 *theobromae* and the fungus appears to exist predominantly in an asexual form (Mohali *et al*.
329 2005).

330 In Brazil, the BSh climate type is characterized by annual mean precipitation of less than
331 800 mm (Alvares *et al*. 2013). The conditions of temperature, luminosity and relative humidity
332 in this climatic region, as well as the use of technologies, such as irrigation management,
333 facilitate an intensive fruit production throughout the year. However, it is likely that stress
334 conditions experienced by plants are extended to plant pathogens by forcing them to recombine.
335 Thus, our results suggest that *L. theobromae* may develop sexual state in its life cycle under
336 such conditions. In a previous study performed by Begoude *et al*. (2012), using SSR markers
337 with populations of *L. theobromae* and *L. pseudotheobromae* from *Terminalia* spp. and *T.*
338 *cacao*, these authors observed that these two species develop regular sexual reproduction in
339 humid forests areas with bimodal rainfall in Cameroon. Considering the results obtained, it is
340 difficult to reach conclusions about the species mode of reproduction and further research is
341 needed to analyze such question

342 Overall, in the present study, the population structure and mode of reproduction of *L.*
343 *theobromae* were explored using the microsatellite markers developed by Burgess *et al*. (2003).
344 The results showed low genetic diversity, and two genetic clusters were identified, without
345 structure based on the geographic region, host of origin or climate type within Brazil.
346 Populations studied were mostly clonal but further studies are needed to better understand the
347 mode of reproduction of the pathogen. Our results suggest that the management of *L.*
348 *theobromae* infections will not likely be so challenging as for other Botryosphaeriaceae species.
349 For example, development of resistant cultivars for commercial cultivation could be used to

350 reduce the impact of diseases caused by *L. theobromae*. Since populations of this latent
351 pathogen present low diversity in northeastern Brazil, it is possible that the durability of
352 resistant cultivars would be relatively long-lived (McDonald & Linde 2002).
353

354 ## Acknowledgments

355

356 This study was financial supported by Coordenação de Aperfeiçoamento de Pessoal de Nível
357 Superior – CAPES, "Ciência sem Fronteiras – CAPES" (process number 88881.132070/2016-
358 01) and the Universitat Politècnica de València. G. Elena was supported by the Spanish post-
359 doctoral grant Juan de la Cierva-Formación. We thank Maela León and Valentin Garrigues
360 (Universitat Politècnica de València, Valencia, Spain) for laboratory support.
361
362

363 **Conflicts of Interest:** The authors declare no conflicts of interest.
364

365 ## References

366

367 Agapow PM, Burt A, 2001. Indices of multilocus linkage disequilibrium. *Molecular Ecology*
368 *Notes* **1**, 101–2.
369 Alvares CA, Stape JL, Sentelhas PC, Gonçalves JLM, Sparovek G, 2013. Köppen's climate
370 classification map for Brazil. *Meteorologische Zeitschrift* **22**, 711–28.
371 Alves A, Crous PW, Correia A, Phillips AJL, 2008. Morphological and molecular data reveal
372 cryptic speciation in *Lasiodiplodia theobromae*. *Fungal Diversity* **28**, 1–13.
373 Archer FI, Adams PE, Schneiders BB, 2017. STRATAG: An R package for manipulating,
374 summarizing and analysing population genetic data. *Molecular Ecology* **17**, 5–11.

375 Arnaud-Haond S, Duarte CM, Alberto F, Serrao EA, 2007. Standardizing methods to address
376 clonality in population studies. *Molecular Ecology* **16**, 5115-139.
377 Begoude ADB, Slippers B, Perez G, Wingfield MJ Roux J, 2012. High gene flow and
378 outcrossing within populations of two cryptic fungal pathogens on a native and non-native
379 host in Cameroon. *Fungal Biology* **116**, 343–53.
380 Berbegal M, Pérez-Sierra A, Armengol J, Grünwald NJ, 2013. Evidence for multiple
381 introductions and clonality in Spanish populations of *Fusarium circinatum*.
382 *Phytopathology* **103**, 851–61.
383 Bruvo R, Michiels NK, D'Souza TG, Schulenburg H, 2004. A simple method for the calculation
384 of microsatellite genotype distances irrespective of ploidy level. *Molecular Ecology* **13**,
385 2101-106.
386 Burgess TI, Crous CJ, Slippers B, Hantula J, Wingfield, MJ, 2016. Tree invasions and
387 biosecurity: Eco-evolutionary dynamics of hitchhiking fungi. *AoB Plants 8*:plw076.
388 Burgess TI, Barber PA, Mohali S, Pegg G, de Beer W, Wingfield MJ, 2006. Three new
389 *Lasiodiplodia* spp. from the tropics, recognized based on DNA sequence comparisons and
390 morphology. *Mycologia* **98** 423–35.
391 Burgess TI, Wingfield MJ, Wingfield BD, 2003. Development and characterization of
392 microsatellite loci for the tropical tree pathogen *Botryosphaeria rhodina*. *Molecular
393 Ecology Notes* **3**, 91–4.
394 Correia KC, Silva MA, Morais Júnior MA, Armengol J, Phillips AJL, Câmara MPS, Michereff
395 SJ, 2016. Phylogeny, distribution and pathogenicity of *Lasiodiplodia* species associated
396 with dieback of table grape in the main Brazilian exporting region. *Plant Pathology* **65**,
397 92–03.
398 Coutinho IBL, Freire FCO, Lima CS, Lima JS, Gonçalves FJT, Machado AR, Silva AMS,
399 Cardoso JE, 2017. Diversity of genus *Lasiodiplodia* associated with perennial tropical fruit
400 plants in northeastern Brazil. *Plant Pathology* **66**, 90–04.

401 Cruywagen EM, Slippers B, Roux J, Wingfield MJ, 2017. Phylogenetic species recognition and
402 hybridisation in *Lasiodiplodia*: a case study on species from baobabs. *Fungal Biology* **121**,
403 420–36.

404 Dray S, Dufour AB, 2007. The ade4 package: implementing the duality diagram for ecologists.
405 *Journal of Statistical Software* **22**, 1–20.

406 Excoffier L, Smouse PE, Quattro J, 1992. Analysis of molecular variance inferred from metric
407 distances among DNA haplotypes: Application to human mitochondrial DNA restriction
408 data. *Genetics* **131**, 479–91.

409 FAO, 2018. FAOSTAT. [http://www.fao.org/faostat/en/#data/QC]. Accessed 08 august 2018.

410 Fincham JRS, Day PR, 1963. *Fungal Genetics*. Ed. Blackwell Scientific Publication, Oxford,
411 UK.

412 Grünwald NJ, Everhart SE, Knaus BJ, Kamvar ZN, 2017. Best practices for population genetics
413 analyses. *Phytopathology* **107**, 1000-1010.

414 Grünwald NJ, Goodwin SB, Milgroom MG, Fry WE, 2003. Analysis of genotypic diversity
415 data for populations of microorganisms. *Phytopathology* **93**, 738–46.

416 Hedrick PW, 2005. A standardized genetic differentiation measure. *Evolution* **59**, 1633- 638.

417 IBGE. 2018. Instituto Brasileiro de Geografia e Estatística.
418 [https://agenciadenoticias.ibge.gov.br/agencia-noticias/2013-agencia-de-noticias/releases
419 /16814-pam-2016-valor-da-producao-agricola-nacional-foi-20-maior-do-que-em-2015.ht
420 ml]. Accessed 08 august 2018.

421 Jombart T, 2008. Adegenet: R package for the multivariate analysis of genetic markers.
422 *Bioinformatics* **24**, 1403–405.

423 Jombart T, Devillard S, Balloux F, 2010. Discriminant analysis of principal components: a new
424 method for the analysis of genetically structured populations. *BCM Genetics* **11**, 94.

425 Kamvar ZN, Brooks JC, Grunwald NJ, 2015. Novel R tools for analysis of genome-wide
426 population genetic data with emphasis on clonality. *Frontiers in Genetics* **6**, 208.

427 McDonald BA, Linde C, 2002. Pathogen population genetics, evolutionary potential, and
428 durable resistance. *Annual Review of Phytopathology* **40**, 349-79.
429 Mehl J, Wingfield MJ, Roux J, Slippers B, 2017. Invasive everywhere? Phylogeographic
430 analysis of the globally distributed tree pathogen *Lasiodiplodia theobromae. Forests* **8**,
431 145.
432 Milgroom MG, 1996. Recombination and the multilocus structure of fungal populations.
433 *Annual Review of Phytopathology* **34**, 457-477.
434 Milgroom MG, 2015. *Population biology of plant pathogens: genetics, ecology, and evolution*.
435 American Phytopathological Society Press, St. Paul, MN.
436 Mohali S, Burgess TI, Wingfield MJ, 2005. Diversity and host association of the tropical tree
437 endophyte *Lasiodiplodia theobromae* revealed using simple sequence repeat markers.
438 *Forest Pathology* **35**, 385-96.
439 Netto MSB, Assunção IP, Lima GSA, Marques MW, Lima WG, Monteiro JHA, Balbino VQ,
440 Michereff SJ, Phillips AJL, Câmara MPS, 2014. Species of *Lasiodiplodia* associated with
441 papaya stem-end rot in Brazil. *Fungal Diversity* **67**, 127–41.
442 Phillips A, Alves A, Abdollahzadeh J, Slippers B, Wingfield M, Groenewald J, Crous PW,
443 2013. The Botryosphaeriaceae: genera and species known from culture. *Studies in*
444 *Mycology* **76**, 51–167.
445 R Core Team. 2017. R: A language and environment for statistical computing. R Foundation
446 for Statistical Computing, Vienna, Austria [http://www.R-project.org].
447 Santos PHD, Carvalho BM, Aguiar KP, Aredes FAS, Poltronieri TPS, Vivas JMS, Mussi Dias
448 V, Bezerra1 GA, Pinho DB, Pereira MG, Silveira SF, 2017. Phylogeography and
449 population structure analysis reveals diversity by mutations in *Lasiodiplodia theobromae*
450 with distinct sources of selection. *Genetics and Molecular Research* **16**, 1-14.
451 Schuelke M, 2000. An economic method for the fluorescent labeling of PCR fragments. *Nature*
452 *Biotechnology* **18**, 233–34.

453 Shah MD, Verma KS, Singh K, Kaur R, 2011. Genetic diversity and gene flow estimates among
454 three populations of *Botryodiplodia theobromae* causing dieback and bark canker of pear
455 in Punjab. *Archives of Phytopathology and Plant Protection* **44**, 951–60.
456 Slippers B, Crous PW, Jami F, Groenewald JZ, Wingfield MJ, 2017. Diversity in the
457 Botryosphaeriales: looking back, looking forward. *Fungal Biology* **121**, 307–21.
458 Slippers B, Wingfield MJ, 2007. Botryosphaeriaceae as endophytes and latent pathogens of
459 woody plants: diversity, ecology and impact. *Fungal Biology Reviews* **21**, 90–06.
460 Stoddart JA, Taylor JF, 1988. Genotype diversity: estimation and prediction in samples.
461 *Genetics* **118**, 705–11.
462 Úrbez-Torres JR, 2011. The status of Botryosphaeriaceae species infecting grapevines.
463 *Phytopathologia Mediterranea* **50**, S5–45.
464 Varshney RK, Graner A, Sorrells ME, 2005. Genomics assisted breeding for crop improvement.
465 *Trends in Plant Science* **10**, 621-30.
466 Winter DJ, 2012. Mmod: an R library for the calculation of population differentiation statistics.
467 *Molecular Ecology Resources* **12**, 1158–160.

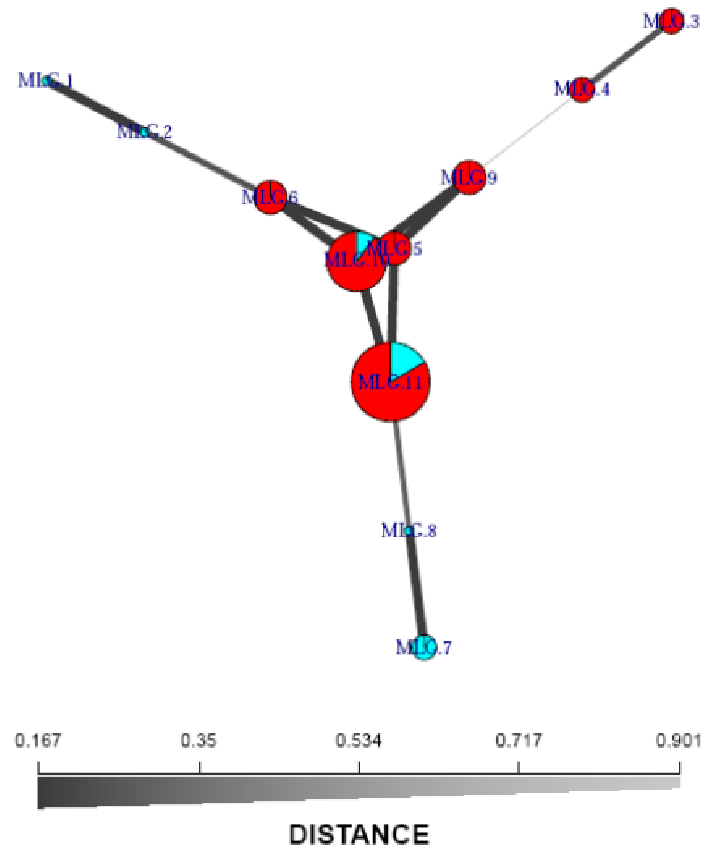
468 **Figure 1** Minimum spanning network from the clone corrected data showing the relationships
469 among the individual multilocus genotypes (MLGs). **A**. The populations from Brazil (BR) and
470 Mexico (MX). **B**. The northeastern Brazil population of *Lasiodiplodia theobromae* defined by
471 climate types (Alvares *et al*. 2013). Each node represents a different MLG. Distances and
472 thickness of the lines between nodes are proportional to Bruvo's distance (Bruvo *et al*. 2004).
473 Node colors and sizes correspond the population studied and number of individuals,
474 respectively
475
476 **Figure 2** Neighbor-Joining dendrogram of Bruvo´s genetic distance (Bruvo *et al*. 2004) among:
477 **A**, seven observed multilocus genotypes in Brazil (MLG11, MLG10, MLG5, MLG6, MLG9,
478 MLG3, MLG4) and six MLGs in Mexico (MLG11, MLG10, MLG7, MLG1, MLG2, MLG8),
479 and **B**, seven observed MLGs in northeastern Brazil. The isolates are colored according to PCA
480 analyses results: **A**, in green, MLGs assigned to cluster I, in purple, blue, black and red, MLGs
481 assigned to cluster II; **B**, in blue, MLGs assigned to cluster I, in red, green and purple, isolates
482 assigned to cluster II. Support values greater than 70% using 1,000 bootstrap samples are shown
483
484

Figure 1A

POPULATION

BR

MX



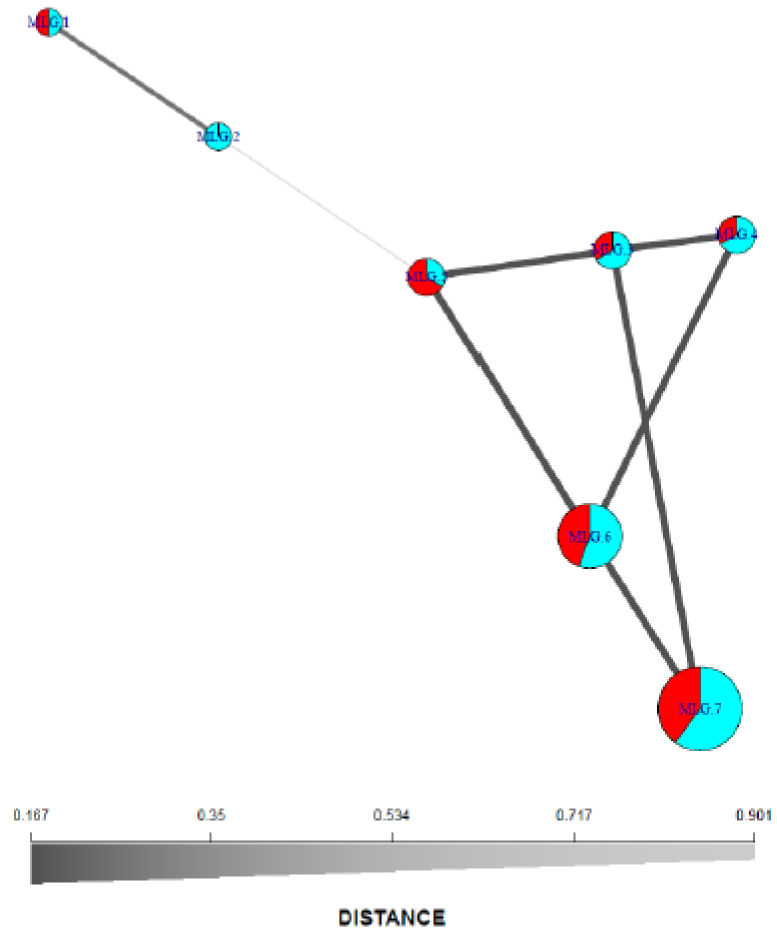
485

Figure 1B

POPULATION

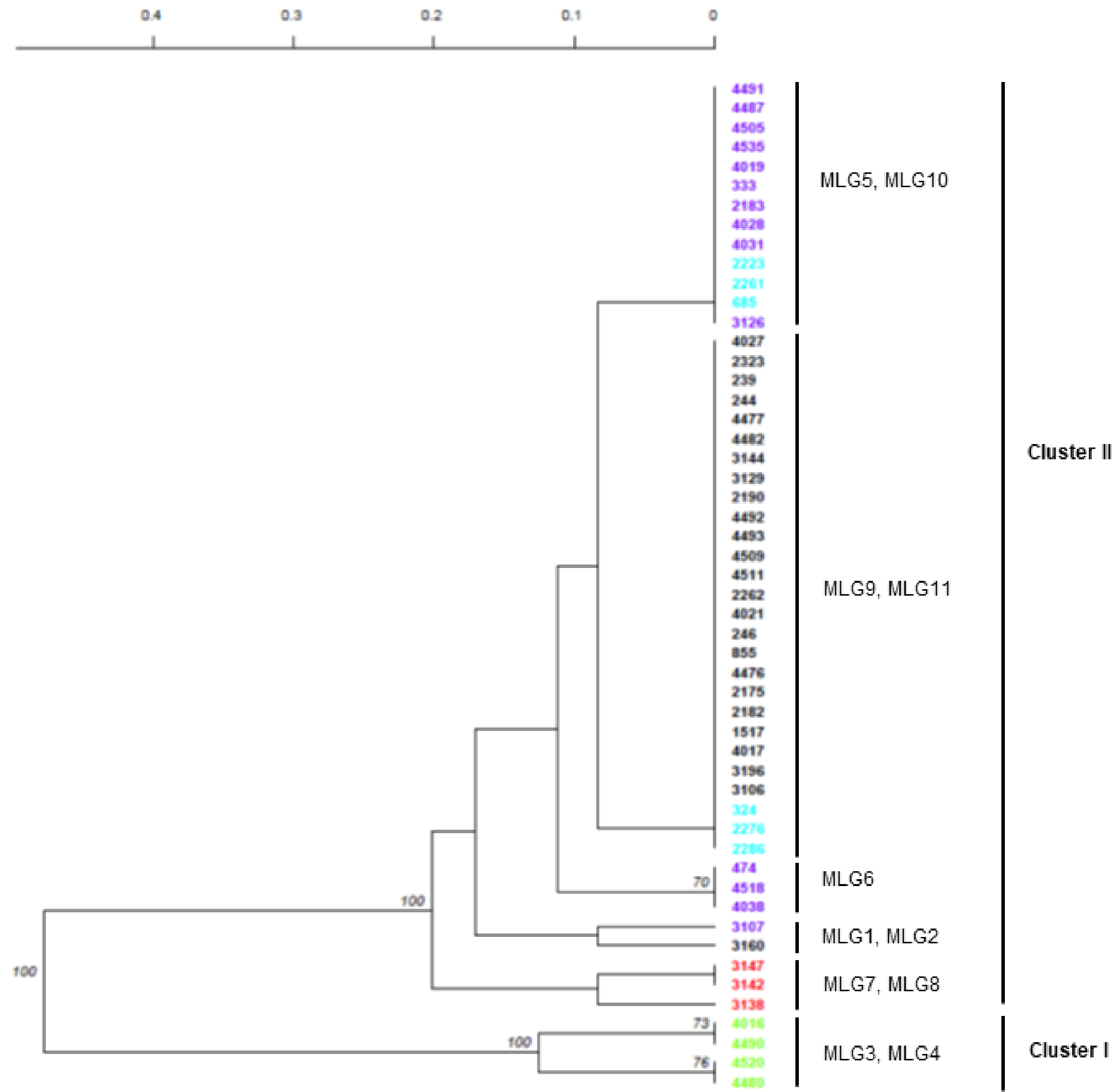
BSh

As

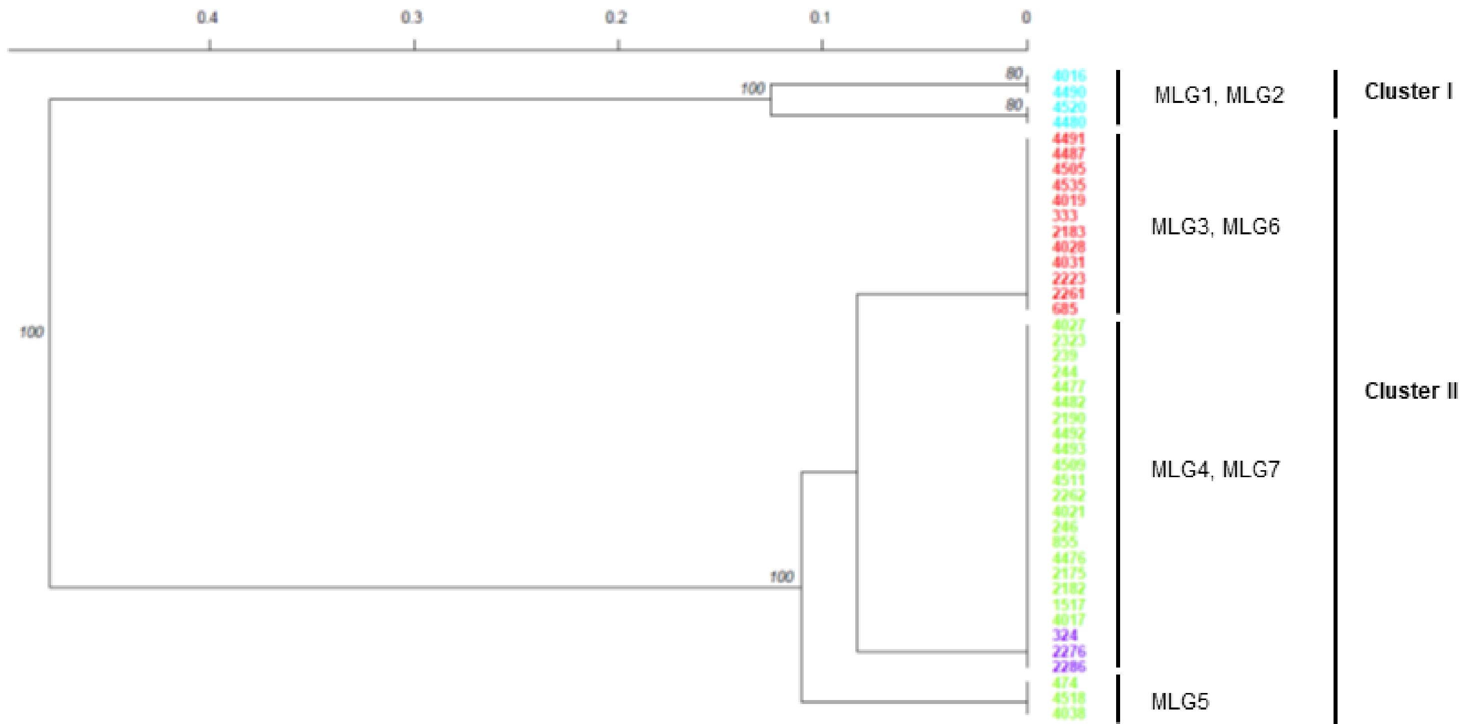


486

487



488


0.4
0.3
0.2
0.1
0
100
80
80
100
100
4016
4490
4520
4480
4491
4487
4505
4535
4019
333
2183
4028
4031
2223
2261
685
4027
2323
239
244
4477
4482
2190
4492
4493
4509
4511
2262
4021
246
855
4476
2175
2182
1517
4017
324
2276
2286
474
4518
4038
MLG1, MLG2
MLG3, MLG6
MLG4, MLG7
MLG5
Cluster I
Cluster II