GMR

Phylogeography and population structure analysis reveals diversity by mutations in *Lasiodiplodia theobromae* with distinct sources of selection

P.H.D. Santos[1], B.M. Carvalho[2], K.P. Aguiar[3], F.A.S. Aredes[2], T.P.S. Poltronieri[1], J.M.S. Vivas[1], V. Mussi Dias[1], G.A. Bezerra[1], D.B. Pinho[4], M.G. Pereira[2] and S.F. Silveira[1]

[1]Laboratório de Entomologia e Fitopatologia,
Centro de Ciências e Tecnologias Agropecuárias,
Universidade Estadual do Norte Fluminense Darcy Ribeiro,
Campos dos Goytacazes, RJ, Brasil
[2]Laboratório de Melhoramento Genético Vegetal,
Centro de Ciências e Tecnologias Agropecuárias,
Universidade Estadual do Norte Fluminense Darcy Ribeiro,
Campos dos Goytacazes, RJ, Brasil
[3]Laboratório de Solos, Centro de Ciências e Tecnologias Agropecuárias,
Universidade Estadual do Norte Fluminense Darcy Ribeiro,
Campos dos Goytacazes, RJ, Brasil
[4]Departamento de Fitopatologia, Universidade de Brasília, DF, Brasil

Corresponding author: P.H.D. Santos
E-mail: pedroh_dias@hotmail.com

Genet. Mol. Res. 16 (2): gmr16029681
Received March 22, 2017
Accepted April 17, 2017
Published April 20, 2017
DOI http://dx.doi.org/10.4238/gmr16029681

ABSTRACT. *Lasiodiplodia theobromae* is a plant pathogen with a worldwide distribution, with low host specificity, causing stem cankers, dieback diseases, and fruit rot in several species of plants. In coconut, this pathogen is reported as the etiological agent of "coconut leaf blight" (CLB) disease, causing several losses in fruit production. The CLB is

an important disease for this crop in Brazil. In our study, we used a phylogeographic approach through the molecular characterization of the translation elongation factor 1-α (TEF1-α) to elucidate the pathogen distribution in Brazil and other countries, besides, search information about diversity sources of this pathogen in coconut palm tree at Brazilian northern, northeast, and southeast. We found that *L. theobromae* diversity is within populations (locations), and populations that are located closest to the center of the tropical zone have more variability as Central Africa, Brazilian Southeast, and Northeast. The widespread distribution could be in part related with long-distance dispersal via global trade of plants and plant products. The entrance route of *L. theobromae* in Brazil probably occurred from Africa route and not occurred once. In Brazil, the diversity of this pathogen in coconut tree could be linked to two agents of selection: high host diversity (in Northeast) and distinct management measures adopted in Southeast. These different sources of selection, mainly the mutations, could be one of the reasons that we found distinct reactions to "coconut leaf blight" chemical control in these regions.

Key words: Phytopathology; Evolution model; Mycology; TEF1-α

INTRODUCTION

The fungus *Lasiodiplodia theobromae* (Pat.) Griffon & Maubl. is a pathogen with a worldwide distribution in tropical and subtropical regions, with low host specificity, causing many diseases in several species of plants (Lima et al., 2013; Muniz et al., 2014). According to the Systematic Mycology and Microbiology Laboratory Fungus-Host Distributions Database, there are around 638 literature records distributed in 71 locations around the world and 360 hosts (Farr and Rossman, 2015) associated with this pathogen.

In coconut crop, this pathogen is reported as the etiological agent of "coconut leaf blight" disease (Souza Filho et al., 1979), affecting the fruit production, causing premature lower leaves abscission, decreasing by 50% photosynthetic area, leaving bunches without support and, consequently, the fruits fall before harvest time (Correia and Costa, 2005; Monteiro et al., 2013).

There are disagreements regarding the fungicides indicated for "coconut leaf blight" chemical control in the Northeast (Benzimidazoles) and Southeast (Triazoles) of Brazil (Ram, 1995; Monteiro et al., 2013). We believed that in Southeast the disease is more aggressive due to isolating adaptation to the climate, host or chemical management, which could select more resistant isolates. Alves et al. (2008) reported the existence of "cryptic" species in *L. theobromae*, and there are indications that coconut isolates from northeastern differ from others described in the same species (Subileau et al., 1994), both by morphology and pathogenically.

Phylogeographic studies have been intensively used to investigate the genetic diversity of plant and animal populations, but fungus investigations are still scarce. This type of studies deals with the principles and processes that govern the geographic distributions of genealogical lineages, especially those within and among closely related species (Avise, 2000). A phylogeographic approach aims to understand genetic diversity in connection with

dispersal history (Avise, 2000). The potential for the use of molecular phylogeography in the investigation of native or introduced species is enormous, mostly because it provides information about genetic diversity not only in the spatial dimension but also in the temporal dimension (Avise, 2000). Molecular phylogeography techniques have been more frequently used in the areas of ecology, conservation genetics, and biogeography.

The utilization of molecular phylogeographic approaches to uncover patterns of genetic diversity in fungus species of agricultural importance is not widespread, which reinforces the need for such studies. In the last decade, some studies have used this approach to reveal patterns of genetic diversity in the plant pathogens *Phakopsora pachyrhizi* (Freire and Oliveira, 2008), *Erysiphe necator* (Brewer and Milgroom, 2010), and *Fusarium* sp (Summerell et al., 2010), but for *L. theobromae*, still now, there is a lack around phylogeographic studies.

In the present study, we used a phylogeographic approach through the molecular characterization of the translation elongation factor 1-α (TEF1-α) to address the following questions: i) How was the spread of *L. theobromae* in continents with a higher incidence of this pathogen? ii) What are the genealogical relationships among *L. theobromae* found in Brazil (Southeast, Northeast, and North) with isolates present in other countries? iii) What are the agents of variation found in coconut isolates from Brazil and why are distinct responses to chemical managements at Southeastern and Northeastern regions?

MATERIAL AND METHODS

Sample collection and isolation

We collected samples of coconut trees with "coconut leaf blight" symptoms and categorized them according to the place of origin for isolation in pure culture. The samples were conducted in the Southeast, Northeast, and Northern states of Brazil. Samples of healthy plants were also subjected to isolation, to obtain possible endophyte isolates. Some isolates from other hosts were collected for the collection.

Fungal isolation was done as described by Ismail et al. (2012). We obtained the pure cultures by hyphal tip excision from the colony margins on PDA and subsequent incubation at 25°C in the dark. All pure cultures obtained were deposited in the plant disease clinic collection of Universidade Estadual do Norte Fluminense Darcy Ribeiro.

DNA extraction, amplification of target sequences, and sequencing

Mycelium pure cultures were macerated inside tubes containing pellets of 1.4 mm in diameter and 600 µL Nuclei Lysis Solution; then, the tubes were placed into a cell disruptor (Loccus L-Beader 3) three times at 3700 rpm for 40 s; the other DNA extraction steps were done as reported by Pinho et al. (2013).

Target sequences of the TEF1-α were amplified using primers EF1-728F (5'-CATCGAGAAGTTCGAGAA-3') and EF2 (5'-GGARGTACCAGTSATCATGTT-3') (Jacobs et al., 2004). Polymerase chain reaction (PCR) was performed with 50 ng DNA, 1X PCR buffer, 1.5 U Taq polymerase, 0.06 µM primers (3 pmol/reaction), 0.2 mM of each dNTP, 1.5 mM $MgCl_2$ in a final volume of 50 µL. Amplification was performed in a thermocycler (Veriti® model), with an initial denaturation at 94°C for 2 min; 35 cycles of 30 sat 94°C, 1 min at 55°C, 1 min at 72°C; followed by a final extension step of 3 min at 72°C. PCR amplification

products were visualized and quantified on 1% agarose gel with a mass marker (Kasvi DNA Ladder RTU model K9-100 L).

The amplified products were purified using a commercial purification system Agencourt AMPure XP (Magnetic Stand-96 Ambion), following the manufacturer's recommendations. The sequencing was performed by ACTGene Análises Moleculares Ltda. (Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brazil).

Data analyses

The nucleotide sequences were edited with the DNA Dragon software. All sequences were checked manually, and nucleotides with ambiguous positions were clarified using both primer direction sequences.

The isolate identification was performed by multigene analysis (ITS and TEF1-α) to select only the isolates that belonged to the *L. theobromae* species (data not shown). The analysis was performed as reported by Rosado et al. (2016), and the nucleotide substitution models were selected according to the Akaike information criterion (AIC). The model SYM+I of evolution was used for ITS and HKY+I+G was used for TEF1-α.

We selected 18 isolates that were allocated at the *L. theobromae* clade, and other sequences of *L. theobromae* were selected from GenBank (http://www.ncbi.nlm.nih.gov) that was clustered in groups based on location, totalizing 104 isolates (Table 1).

All the sequences obtained were added and aligned in the Muscle® program implemented in the MEGA software v.6.

To identify individual haplotypes and their frequencies, we used the DnaSP v.5.10.01 software. DnaSP also estimated the haplotype and nucleotide diversity (Nei, 1987).

The genealogical relationships among haplotypes were generated by a median-joining network (MJ) method in the Network 4.6.0.0 software (Bandelt et al., 1999). This program applies statistical parsimony by the algorithm developed by Templeton et al. (1992). The haplotype network estimation considered each of the bases (A, T, G, and C) as a character to identify individual haplotypes and their frequencies. The maximum number of parsimonious connections between the sequences was estimated at 95% of probability for all haplotype connections within the network. Haplotypes were presented as circles, which are color-coded. Circle sizes were proportional to the frequency of isolates in each haplotype.

The genetic structure was estimated for each sampled locality (genetic diversity within and between locations), and analysis of molecular variance (AMOVA) was performed in the ARLEQUIN v.3.1 software (Excoffier et al., 2007). AMOVA estimates the genetic structure using information from the haplotype allelic contents as well as their frequencies (Excoffier et al., 2007).

To relate the phylogeographic haplotype relationships, we plotted in geographic maps the frequency of each haplotype obtained, using the Inkscape software v.0.91.

RESULTS AND DISCUSSION

The haplotype "H3" was the most frequent in locations under study, totalizing 47 isolates in 104 obtained sequences (Table 2). This haplotype was present in eight of ten locations studied. Haplotypes "H1" and "H5" appeared in nine and four locations, respectively. We observed eight singleton haplotypes, i.e., they were identified only once and, therefore,

Table 1. GenBank accession numbers of DNA sequences of *Lasiodiplodia theobromae* used in the analyses.

Location	Isolate[a]	Host	TEF1-α
Central Africa	CMW28311	*Terminalia ivorensis*	GQ469898
Central Africa	CMW36127	*Adansonia digitata*	KU886993
Central Africa	CMW28571	*Terminalia ivorensis*	GQ469897
Central Africa	LASPTGR42B	*Grevillea robusta*	FJ904889
Central Africa	LASPTAI41	*Grevillea robusta*	FJ904888
Central Africa	LASPTAI38	*Grevillea robusta*	FJ904887
Central Africa	LASPTAI17	*Grevillea robusta*	FJ904886
Central Africa	GrS1-1	*Grevillea robusta*	FJ904863
Central Africa	CBS 190.73	*Persea americana*	EF622048
Central Africa	CMW 10130	*Vitex doniana*	AY236900
Central Africa	G3/20/Grevillea	*Grevillea robusta*	GU130539
Central Africa	1/3/1/Grevillea	*Grevillea robusta*	GQ999856
Central Africa	CMW18420	*Casuarina equisetifolia*	DQ103564
Central Africa	CMW33290	*Adansonia digitata*	KU886956
South Africa	CBS 112874	*Vitis vinifera*	EF622055
South Africa	STE-U 4419	*Vitis vinifera*	AY343368
South Africa	STE-U 5051	*Vitis vinifera*	AY343369
South Africa	CMW18422	*Pinus patula*	DQ103562
South Africa	CMW18425	*Pinus patula*	DQ103561
Africa/Asia	BOT-5	*Mangifera indica*	JN814403
Africa/Asia	BOT-9	*Mangifera indica*	JN814419
Africa/Asia	BOT4	*Mangifera indica*	JN814422
Africa/Asia	BOT-7	*Mangifera indica*	JN814423
Africa/Asia	BOT-6	*Mangifera indica*	JN814426
Africa/Asia	BOT-23	*Mangifera indica*	JN814427
East Asia	JMB-122	*Prunus persica*	HQ660490
East Asia	ML	*Zea mays*	KT985635
East Asia	CRI-LP3	*Ipomoea batatas*	KU870369
East Asia	CRI-LP2	*Ipomoea batatas*	KU870368.1
East Asia	HD1332	*Albizia falcataria*	KU712503.1
East Asia	BL1331	*Albizia falcataria*	KU712500
East Asia	CERC3825	*Rosa rugosa*	KR816842
East Asia	CMW24702	*Eucalyptus* sp	HQ332210
East Asia	CMW24701	*Eucalyptus* sp	HQ332209
East Asia	CERC1985	*Pinus balfouriana*	KP822998
East Asia	CERC1989	*Eucalyptus* sp	KP823000
Southwestern Asia	IRNHM-KB642	*Citrus* sp	KU737511
Southwestern Asia	IRNHM-KB64	*Citrus* sp	KU737510
Southwestern Asia	IRAN 1233C	*Mangifera indica*	GU973860
Southwestern Asia	IRAN 1496C	*Mangifera indica*	GU973861
Southwestern Asia	IRAN 1499C	*Mangifera indica*	GU973862
Southwestern Asia	CJA198	*Mangifera indica*	GU973863
Southwestern Asia	CJA199	*Mangifera indica*	GU973864
Southwestern Asia	CJA279	*Mangifera indica*	GU973865
Southwestern Asia	B0151	*Pinus kesiya*	KM006467
Southwestern Asia	B0281	*Manilkara zapota*	KM006473
Southwestern Asia	B0451	*Syzygium samarangense*	KM006485
Southwestern Asia	APAO-01	*Anacardium occidentale*	LC146471
Brazil Northeast	**CF/UENF417**	*Cocos nucifera*	**KY223714**
Brazil Northeast	CDA472	*Cocos nucifera*	KP308467
Brazil Northeast	COAD1790	*Cocos nucifera*	KP308468
Brazil Northeast	CDA465	*Cocos nucifera*	KP308465
Brazil Northeast	CDA467	*Cocos nucifera*	KP308473
Brazil Northeast	CDA469	*Cocos nucifera*	KP308466
Brazil Northeast	**CF/UENF431**	*Cocos nucifera*	**KY223714**
Brazil Northeast	**CF/UENF432**	*Cocos nucifera*	**KY223715**
Brazil Northeast	**CF/UENF435**	*Cocos nucifera*	**KY223716**
Brazil Northeast	**CF/UENF437**	*Cocos nucifera*	**KY223718**
Brazil Northeast	**CF/UENF430**	*Cocos nucifera*	**KY223713**
Brazil Northeast	**CF/UENF436**	*Cocos nucifera*	**KY223705**
Brazil Northeast	**CF/UENF438**	*Cocos nucifera*	**KY223717**
Brazil Northeast	CMM 0384	*Vitis vinifera*	KJ417876
Brazil Northeast	CMM 0820	*Vitis vinifera*	KJ417877
Brazil Northeast	CMM 0455	*Vitis vinifera*	KJ417878
Brazil Northeast	CMM 0307	*Vitis vinifera*	KJ417879

Continued on next page

Table 1. Continued.

Location	Isolate[a]	Host	TEF1-α
Brazil Northeast	CMM 0310	*Vitis vinifera*	KJ417880
Brazil Northeast	CMM1517	*Manga*	JX464054
Brazil Northeast	CMM4050	*Manga*	JX464024
Brazil Northeast	IBL340	*Spondias purpurea*	KT247472
Brazil Northeast	IBL375	*Talisia esculenta*	KT247473
Brazil Northeast	IBL404	*Anacardium occidentale*	KT247470
Brazil Northeast	IBL405	*Anacardium occidentale*	KT247471
Brazil North	**CF/UENF425**	*Elaeis guineensis*	**KY223705**
Brazil North	**CF/UENF426**	*Cocos nucifera*	**KY223706**
Brazil Southeast	**CF/UENF429**	*Capparis flexuosa*	**KY223712**
Brazil Southeast	**CF/UENF419**	*Cocos nucifera*	**KY223719**
Brazil Southeast	**CF/UENF420**	*Cocos nucifera*	**KY223720**
Brazil Southeast	**CF/UENF421**	*Cocos nucifera*	**KY223708**
Brazil Southeast	**CF/UENF423**	*Cocos nucifera*	**KY223710**
Brazil Southeast	COAD1788	*Cocos nucifera*	KP308476
Brazil Southeast	COAD1789	*Cocos nucifera*	KP308474
Brazil Southeast	CDA444	*Cocos nucifera*	KP308477
Brazil Southeast	CDA425	*Cocos nucifera*	KP308475
Brazil Southeast	CDA450	*Cocos nucifera*	KP308478
Brazil Southeast	CDA455	*Cocos nucifera*	KP308463
Brazil Southeast	**CF/UENF427**	*Persea americana*	**KY223707**
Brazil Southeast	**CF/UENF428**	*Cocos nucifera*	**KY223711**
Brazil Southeast	**CF/UENF422**	*Cocos nucifera*	**KY223709**
Brazil Southeast	CMM3831	*Jatropha curcas*	KF226717
Brazil Southeast	CMM3654	*Jatropha curcas*	KF226716
Brazil Southeast	CMM3612	*Jatropha curcas*	KF226692
Brazil Southeast	CMM3647	*Jatropha curcas*	KF226704
Oceania	MUCC709	*Lysiphyllum cunninghamii*	GU199393
Oceania	CBS 289.56	*Sail-cloth*	EF622050
Oceania	CBS 164.96	*Fruit along coral reef coast*	AY640258
South America	CMW22924	*Schizolobium parahyba*	KF886732
South America	CMW9271	*Schizolobium parahyba*	KF886731
South America	LA-SJ1	*Vitis vinifera*	KM401973
South America	LA-SV1	*Vitis vinifera*	KM401972
South America	LA-SOL1	*Vitis vinifera*	KM401971
South America	LAREP3	*Mangifera indica*	KU507453
South America	LAREP2	*Mangifera indica*	KU507452
South America	LAHUAL3	*Mangifera indica*	KU507450
South America	LASOM3	*Mangifera indica*	KU507444

TEF1-α = translation elongation factor 1-α. [a]Isolates obtained in this study are highlighted in bold. CF/UENF = Coleção Clínica Fitossanitária at Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF).

were exclusive from the location. It is important to highlight that the host origin of isolates did not contribute to the haplotype grouping as reported by Mohali et al. (2005).

It is common that ancestral haplotypes were the most frequent and found covering the majority of studied areas (Templeton et al., 1992; Freire and Oliveira, 2008). The high frequencies observed to haplotype "H3" is an indicative that this haplotype is ancestral in relation with the other haplotypes identified in this study. This fact was reinforced by the type species (CBS 164.96) present in this haplotype.

In general, we observed diversity for most locations in this study; a fact that could be connected to the widespread of the pathogen around the world (Mohali et al., 2005; Marsberg et al., 2017).

In Brazil, all locations showed more than one haplotype and the majority were singletons (Table 2). In North of Brazil, we observed the presence of two singleton haplotypes, resulting in low diversity. On the other hand, Brazilian Southeast and Northeast showed six haplotypes in each, indicating diversity for these regions.

Table 2. Distribution of the 14 *Lasiodiplodia theobromae* haplotypes between 104 analyzed TEF-1α sequences.

Haplotypes	Total	Central Africa	South Africa	Africa/ Asia	North Brazil	Southeast Brazil	Northeast Brazil	South America	Southwestern Asia	East Asia	Oceania
H1	18	3	1	1	1	6	3	1	1	1	
H2	1				1						
H3	47	4		5		7	9	7	8	6	1
H4	4					1	1	1	1		
H5	16		2			2	9			3	
H6	1					1					
H7	1					1					
H8	6	1	2						2		1
H9	5	4						1			
H10	1						1				
H11	1						1				
H12	1	1									
H13	1	1									
H14	1									1	
Total	104	14	5	6	2	18	24	10	12	11	2

AMOVA, based on F_{ST}, showed that most of the observed variance was within populations (91.66%), indicating a greater variability within populations than among populations (Table 3) as reported by Peixouto et al. (2015).

The results of AMOVA, together with those of Al-Sadi et al. (2013), indicated as a possible cause for this variability within populations the frequent movement of pathogen inoculum across geographical locations. Moreover, this variability was consistent with the high number of singletons uncovered during sequence analysis (Perneel et al., 2006).

Most of the populations shared the haplotype "H3" (Table 2), which support low differences found among populations (Table 3). The known low host specificity of *L. theobromae* could in part induce to low genetic structure between populations, and this could be explained by the fact that the fungus can colonize many hosts in a determinant area (Al-Sadi et al., 2013; Marsberg et al., 2017).

Table 3. Analysis of molecular variance (AMOVA), among *Lasiodiplodia theobromae* isolate sequences on different geographic locations.

Source	d.f.	Sum of squares	Variance components	Variation (%)	F_{ST}
Among populations	9	5.89	0.03126	8.34	0.083*
Within populations	94	32.295	0.34356	91.66	
Total	103	38.183	0.3749		

*Significant at 0.05 probability level.

The high difference found within populations (Table 3) can be attributed to population gene flow, increasing the diversity within populations (Templeton, 1998; Mohali et al., 2005). A high gene flow prevented local adaptation, reducing the fixation of alleles that were favored under local conditions, hindering the process of speciation. On the other hand, gene flow generated new polymorphisms in the population and increased the effective size of the local population and their ability to withstand random changes in gene frequencies, opposing the genetic drift and generating new combinations of genes, in which natural selection can act (Balloux and Lugon-Moulin, 2002). These facts can be considered for the structure of these populations, explaining its higher diversity within them.

The results could also be influenced by the sample size, but some authors reported that the sample size does not directly affect the genetic diversity observed (Wang et al., 2006). Freire and Oliveira (2008), working with 99 isolates of *Phakopsora pachyrhizi*, and Jorge et

al. (2015), working with 436 isolates, reported the similar genetic diversity and population structure for the Asian soybean rust in their study. This fact reinforced the idea that the sample size does not directly affect the genetic diversity as commented above.

One of our objectives refers to the study of the *L. theobromae* evolutionary history and dispersion in world and Brazil. Increasing the total number of sequences by location would be uninformative since the probability that the network obtained by this study have significant alterations would be very small. In fact, we probably could observe new haplotypes, and more haplotype by location, but the conclusions about the history and dispersion would be the same or similar.

Nucleotide diversity estimates the probability that two randomly chosen homologous nucleotides will be distinct, bearing in mind the number of mutations between haplotypes, and is equivalent to the level of polymorphism within a population (Nei, 1987). Haplotype diversity is defined as the probability that two individuals are randomly chosen to have distinct haplotypes. This variable is equivalent to genetic diversity, having as the difference the replacement of the genotypic frequency by the haplotype frequency (Nei and Li, 1979).

On this aspect, we noted a usual pattern between populations; in general, populations that were located closest to the center of the tropical zone presented more variability (Central Africa, Brazilian Southeast and Northeast) (Tables 2 and 4). This evidence, still now, has not yet been reported in any study with this pathogen, giving originality to this study, contributing to the understanding of the special ecological conditions of this pathogen. The haplotype diversity ranged from 0.33 to 1. The lower value (0.33), attributed to Egypt region, which belongs as much to Africa as to Asia, could be in part associated with the environmental conditions and host scarcity to the pathogen establishment. On the other hand, the value 1 was observed in populations wherein all haplotypes were distinct, i.e., high genetic diversity within the locality.

Table 4. Haplotype diversity (*H*), nucleotide diversity (π) and its standard deviation (SD) of *Lasiodiplodia theobromae* in 10 locations around the world.

Location	N	S	K	*H* (SD)	π (SD)
East Asia	11	10	4	0.67 (0.123)	0.00871 (0.0047)
Southwestern Asia	12	7	5	0.67 (0.141)	0.006 (0.0019)
South Africa	5	4	3	0.8 (0.164)	0.008 (0.0016)
Central Africa	14	25	6	0.835 (0.062)	0.02 (0.0092)
Northeast Brazil	24	15	6	0.728 (0.058)	0.0083 (0.0027)
Southeast Brazil	18	24	6	0.76 (0.07)	0.011 (0.0055)
North Brazil	2	2	2	1 (0.5)	0.0074 (0.0024)
South America	9	6	4	0.694 (0.147)	0.0048 (0.0021)
Oceania	3	5	3	1 (0.272)	0.012 (0.004)
Africa/Asia	6	2	2	0.33 (0.22)	0.0023 (0.0015)
Total	104	30	14	0.74 (0.035)	0.0077 (0.0012)

N = number of individuals; S = number of polymorphic sites; K = number of haplotypes.

The North region in Brazil showed high values of diversity presenting only two haplotypes. In other words, these haplotypes are highly distinct. We found high diversity values to Southeast and Northeast, but in this case, these regions presented more haplotypes than the North region (Table 4).

For nucleotide diversity, locations that exhibited substantially the same haplotype frequency not necessarily have the same value, since the estimate also takes into account the amount of mutation between these haplotypes. For example, East Asia (11 haplotypes)

and Southwestern Asia (12 haplotypes) had the same proportion of haplotype and thus has the same value of haplotype diversity (0.67) (Table 4). However, Southwestern Asia had nucleotide diversity around 0.006, while East Asia had 0.00871 (Table 4).

As seen in Figure 1, between the haplotypes identified in East Asia (Table 2) three had mutation events for each haplotype, since the haplotypes identified in Southwestern Asia differed in four mutation events, explaining this difference in nucleotide diversity.

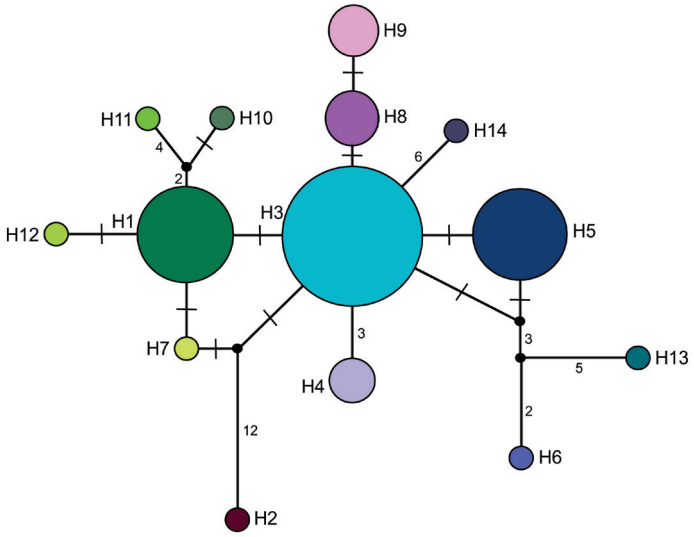

Figure 1. Median-joining haplotype network based on the TEF-1α gene for *Lasiodiplodia theobromae.* Haplotypes are distinguished by colors, and the circle sizes indicate haplotype frequencies. Branch numbers indicate the total of mutation events and connection bars between two neighbor haplotypes represent a single mutation event.

We obtained a single network containing all 14 haplotypes for the TEF-1α gene (Figure 1). The estimation procedure detected four nodes (represented by black smaller circles) in which none of the 104 sequences could be allocated. According to Templeton (1998), these nodes represent intermediate haplotypes, which are inferred between two nearest neighbor haplotypes in a network, which differ from two or more mutation events. These nodes refer to extinct or not sampled haplotypes.

It is expected that ancestral haplotypes occupy the center of the network, and descendants or derived haplotypes are found in network tips. The network obtained, pointed as the more central, the haplotype "H3" (Figure 1). This haplotype in the network also was the most frequent, as noted in Table 2, reinforcing the evidence of its ancestry.

The haplotype "H3" showed ramifications for haplotypes that occupied the network tips (Figure 1). Templeton et al. (1992) and Freire and Oliveira (2008) reported the presence of ramifications as something expected in networks, because these haplotypes undergo for several mutations at different times, generating several descendant haplotypes. Haplotypes that occupied the network tips are recent, and this evidence is reinforced by most of them present as singletons (Templeton et al., 1992).

The ancestral haplotype "H3" was present in most of the sampled locations (Figure 2). The two other more frequent haplotypes ("H1" and "H5") were founded in many sampled

locations. We believe that the haplotype "H3" could be similar to the first specimen described as *Botryodiplodia theobromae* (Patouillard and Lagerheim, 1892) in Ecuador, and the one found years later for Griffon and Maublanc (1909) in central Africa renaming the fungus to *Lasiosiplodia theobromae*.

According to the coalescence theory (Templeton et al., 1992; Avise, 2000), it is expected the ancestral haplotype to be present in most of the geographical locations while new haplotypes have a more restricted distribution. Thus, the distribution observed in Figure 2 strengthens the evidence mentioned above, the haplotype "H3" would have probably the ancestral haplotype, and haplotypes with more restricted distribution would be descendants.

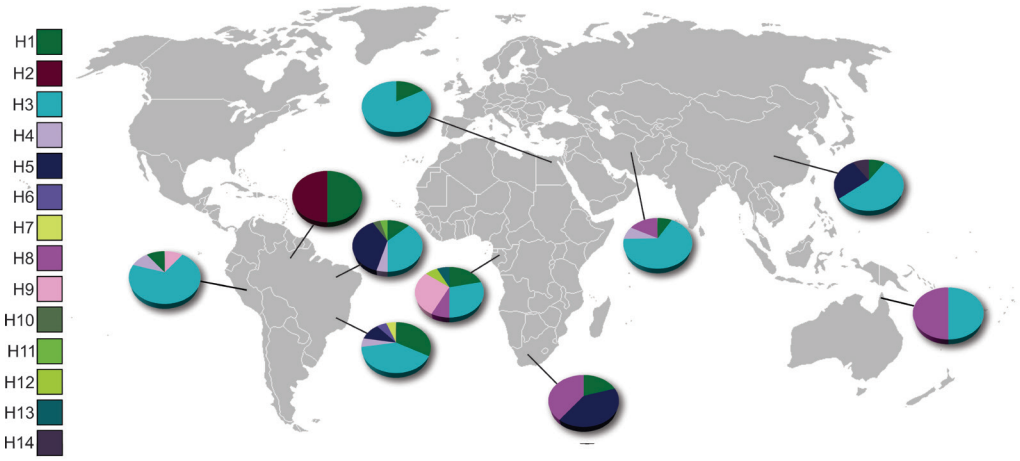

Figure 2. Geographic distribution of the 14 haplotypes for the TEF-1α gene of *Lasiodiplodia theobromae*. Areas of pie charts represent composite haplotype frequency. Haplotype colors and codes are as in the network shown in Figure 1.

We believed that this widespread distribution could be in part relations with anthropogenic long-distance dispersal via global trades of plants and plant products, started in the period of trade Navigations, where many products and plants were brought to these regions. Furthermore, the low host specificity contributed positively to the pathogen establishment in distinct locations.

In Brazilian Southeast and Northeast we notice the same pattern, but as previously mentioned these locations, as well as Central Africa, showed high haplotype variability. This current diversity has most likely resulted from these regions have been for many years the main route for the navigation traffic, which included plant material. The local weather conditions probably contributed to the higher haplotype variability since its occurrence is higher in tropical and subtropical regions (Burgess and Wingfield, 2002).

The pathogen entrance in Brazil could be by the bordering countries (South America) as well as Africa, due to the great products exchanged between these continents. We did not observe the ancestral haplotype "H3" in northern Brazil (Figure 2) (nearest to South America countries), and this fact led us to believe that the physical barriers of neighboring countries could have prevented or delayed the entry of this pathogen. Haplotypes "H3" and "H1" were shared by the northeastern and southeastern regions of Brazil, as well as the central region of Africa, leading us to believe that this might be the most likely entrance route of *L. theobromae* in the country and this event not occurred once.

Figure 3 shows a map of Brazil with *L. theobromae* haplotypes, formed only by the isolates found in coconut (Table 1). The same higher haplotype variability was observed in Southeast and Northeast locations, indicating that the host origin does not have an influence on haplotype grouping as reported above.

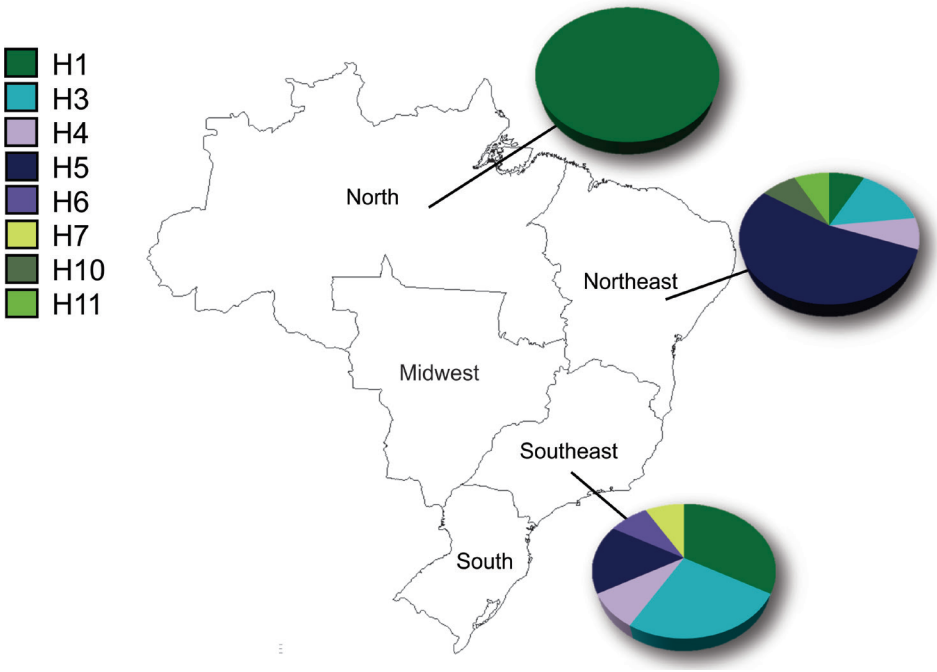

Figure 3. Geographic distribution of the Brazilian coconut *Lasiodiplodia theobromae* haplotypes for the TEF-1α gene. Areas of pie charts represent composite haplotype frequency.

The North region showed only one coconut haplotype ("H1") indicating that this haplotype is more recent in this region than others studied (Figure 3); this could be in part associated with the endophytic behavior of the pathogen (Mohali et al., 2005; Alves et al., 2008). The other fact could be that the disease often occurs when some stress is detected by the plant, mainly hydric stress (Phillips et al., 2005; Marsberg et al., 2017). We believe that as Brazilian North is a region with favorable conditions for water supplementation, "coconut leaf blight" occurs occasionally and the fungus remains endophytically inside the host. Thus, the need for pathogen specialization decreases, contributing to its low variability.

We found six haplotypes for Brazilian Northeast, where the haplotype "H5" showed the highest frequency, followed by the ancestral haplotype "H3" (Figure 3). This location has two singleton haplotypes that only occur there, and probably arise recently in this region. As we mentioned above, this location has some haplotype variability and in this case, using only coconut pathogens, this could be linked to the fact that in Northeast the most of the coconut trees cultivated are from Giant variety (Monteiro et al., 2013) that has high diversity in populations (Perera et al., 2000). In this case, the pathogen has to adapt to the host variability, which acts as a source of selection, increasing the fungus variability. In this scenario, we believe that the high host variability could influence the fungus variability.

In Southeast, as in Northeast, we found six haplotypes; two of them only occur in this region ("H6" and "H7"). The ancestral haplotype ("H3") had the highest frequency (Figure 3). In Southeast, most of the coconut trees cultivated are from Green Dwarf variety (Monteiro et al., 2013), and according to Santos (2016), this variety has low diversity.

The low host diversity would result in low diversity for *L. theobromae* haplotypes, which were not found in our study. It is known that in Southeast regions the chemical management adopted are distinct than the measures adopted in Northeast (Monteiro et al., 2013). We believed that these distinct measures in Southeast could be the source of selection of this pathogen at this region.

L. theobromae has a predominantly asexual mode of reproduction, and this contributes to low haplotype diversity (Mohali et al., 2005; Marsberg et al., 2017). In our study, we found some diversity between haplotypes and two hypothesis could be formed to explain these results: i) the sexual phase even uncommon could occur in many hosts or ii) the mutation rate could in part contribute to the fungal diversity, since the large number of individuals generated by mutations could increase the mutation rate, fixing the mutants in populations by distinct agents of selection.

When we found populations with high recently selection pressure, it could generate a linkage disequilibrium state, and the recombination could randomize the alleles in gametes, returning to the equilibrium (Hartl and Grant, 2007). As the sexual phase are uncommon and hardly occurs in Southeast and Northeast regions, we believe that the recombination would be unfeasible. The mutation hypothesis looked to be more acceptable for the high diversity in Brazilian Southeast and Northeast.

We believe that these distinct sources of selection could be one of the reasons that we found distinct reactions to "coconut leaf blight" chemical control at these regions under study.

In the present study, we showed that the *L. theobromae* diversity is within populations (locations) and this is linked to the high influence of mutations in populations. We also concluded that populations located closest to the center of the tropical zone have more variability as Central Africa, Brazilian Southeast, and Northeast.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

Research supported by CAPES, CNPq, FAPERJ, and UENF.

REFERENCES

Al-Sadi AM, Al-Wehaibi AN, Al-Shariqi RM, Al-Hammadi MS, et al. (2013). Population genetic analysis reveals diversity in *Lasiodiplodia* species infecting date palm, citrus, and mango in Oman and the UAE. *Plant Dis.* 97: 1363-1369 https://doi.org/10.1094/PDIS-03-13-0245-RE.

Alves A, Crous PW, Correia A and Phillips AJL (2008). Morphological and molecular data reveal cryptic speciation in *Lasiodiplodia theobromae. Fungal Divers.* 28: 1-13.

Avise JC (2000). Phylogeography: the history and formation of species. Harvard University Press, Cambridge.

Balloux F and Lugon-Moulin N (2002). The estimation of population differentiation with microsatellite markers. *Mol. Ecol.* 11: 155-165 https://doi.org/10.1046/j.0962-1083.2001.01436.x.

Bandelt HJ, Forster P and Röhl A (1999). Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16: 37-48 https://doi.org/10.1093/oxfordjournals.molbev.a026036.

Brewer MT and Milgroom MG (2010). Phylogeography and population structure of the grape powdery mildew fungus, *Erysiphe necator*, from diverse Vitis species. *BMC Evol. Biol.* 10: 268 https://doi.org/10.1186/1471-2148-10-268.

Burgess T and Wingfield MJ (2002). Quarantine is important in restricting the spread of exotic seed-borne tree pathogens in the southern hemisphere. *Int. For. Rev.* 4: 56-65.

Correia MS and Costa JLS (2005). Dispersão anemófila do fungo *Lasiodiplodia theobromae* em plantações de coqueiro. *Fitopatol. Bras.* 30: 150-154 https://doi.org/10.1590/S0100-41582005000200008.

Excoffier L, Laval G and Schneider S (2007). Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evol. Bioinform. Online* 1: 47-50 10.1111/j.1755-0998.2010.02847.x.

Farr DF and Rossman A (2015). Systematic Mycology and Microbiology Laboratory Fungus-Host Distributions Database. [http://nt.ars-grin.gov/fungaldatabases/]. Accessed September 21, 2016.

Freire M and Oliveira L (2008). Evolutionary history of *Phakopsora pachyrhizi* (the Asian soybean rust) in Brazil based on nucleotide sequences of the internal transcribed spacer region of the nuclear ribosomal DNA. *Mol. Biol.* 931: 920-931 10.1590/s1415-47572008005000026.

Griffon, E, Maublanc A (1909). Sur une maladie du cacaoyer. *Bull. la Société Mycol. Fr*. 25: 51-58.

Hartl DL and Grant A (2007). Principles of Population Genetics. 4th edn. Sinauer Associates, Sunderland.

Ismail AM, Cirvilleri G, Polizzi G, Crous PW, et al. (2012). *Lasiodiplodia* species associated with dieback disease of mango (*Mangifera indica*) in Egypt. *Australas. Plant Pathol.* 41: 649-660 https://doi.org/10.1007/s13313-012-0163-1.

Jacobs K, Bergdahl DR, Wingfield MJ, Halik S, et al. (2004). *Leptographium wingfieldii* introduced into North America and found associated with exotic *Tomicus piniperda* and native bark beetles. *Mycol. Res.* 108: 411-418 https://doi.org/10.1017/S0953756204009748.

Jorge VR, Silva MR, Guillin EA, et al. (2015). The origin and genetic diversity of the causal agent of Asian soybean rust, *Phakopsora pachyrhizi*, in South America. *Plant Pathol.* 64: 729-737 https://doi.org/10.1111/ppa.12300.

Lima JS, Moreira RC, Cardoso JE, et al. (2013). Caracterização cultural, morfológica e patogênica de *Lasiodiplodia theobromae* associado a frutíferas tropicais. *Summa Phytopathol.* 39: 81-88 https://doi.org/10.1590/S0100-54052013000200001.

Marsberg A, Kemler M, Jami F, Nagel JH, et al. (2017). *Botryosphaeria dothidea*: a latent pathogen of global importance to woody plant health. *Mol. Plant Pathol.* 18: 477-488 https://doi.org/10.1111/mpp.12495.

Mohali S, Burgess TI and Wingfield MJ (2005). Diversity and host association of the tropical tree endophyte *Lasiodiplodia theobromae* revealed using simple sequence repeat markers. *For. Pathol.* 35: 385-396 https://doi.org/10.1111/j.1439-0329.2005.00418.x.

Monteiro CM, Caron ES, da Silveira SF, et al. (2013). Control of foliar diseases by the axillary application of systemic fungicides in Brazilian coconut palms. *Crop Prot.* 52: 78-83 https://doi.org/10.1016/j.cropro.2013.05.013.

Muniz CR, da Silva GF, Souza MT, Jr., Freire FC, et al. (2014). *Agrobacterium tumefaciens*-mediated transformation of *Lasiodiplodia theobromae*, the causal agent of gummosis in cashew nut plants. *Genet. Mol. Res.* 13: 2906-2913 https://doi.org/10.4238/2014.February.21.8.

Nei M (1987). Molecular Evolutionary Genetics. Columbia University Press, New York.

Nei M and Li W-H (1979). Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proc. Natl. Acad. Sci. USA* 76: 5269-5273 https://doi.org/10.1073/pnas.76.10.5269.

Patouillard NT, Lagerheim G (1892). Champignons de l'Equateur (Pugillus II). *Bull. la Société Mycol. Fr.* 8: 113-140.

Peixouto YS, Dórea Bragança CA, Andrade WB, Ferreira CF, et al. (2015). Estimation of genetic structure of a *Mycosphaerella musicola* population using inter-simple sequence repeat markers. *Genet. Mol. Res.* 14: 8046-8057 https://doi.org/10.4238/2015.July.17.13.

Perera L, Russell JR, Provan J and Powell W (2000). Use of microsatellite DNA markers to investigate the level of genetic diversity and population genetic structure of coconut (*Cocos nucifera* L.). *Genome* 43: 15-21. https://doi.org/10.1139/g99-079

Perneel M, Tambong JT, Adiobo A, Floren C, et al. (2006). Intraspecific variability of *Pythium myriotylum* isolated from cocoyam and other host crops. *Mycol. Res.* 110: 583-593 https://doi.org/10.1016/j.mycres.2005.12.002.

Phillips A, Alves A, Correia A and Luque J (2005). Two new species of *Botryosphaeria* with brown, 1-septate ascospores and *Dothiorella* anamorphs. *Mycologia* 97: 513-529 https://doi.org/10.1080/15572536.2006.11832826.

Pinho DB, Dutra DC and Pereira OL (2013). Notes on *Ceratocystis paradoxa* causing internal post-harvest rot disease on immature coconut in Brazil. *Trop. Plant Pathol.* 38: 152-157 https://doi.org/10.1590/S1982-56762013000200010.

Ram C (1995). Eficiência do controle químico da queima das doenças foliares em coqueiro (*Cocos nucifera*) em Sergipe.

Fitopatol. Bras. 20: 248-250.

Rosado AWC, Machado AR, Freire FCO and Pereira OL (2016). Phylogeny, Identification, and Pathogenicity of *Lasiodiplodia* Associated with Postharvest Stem-End Rot of Coconut in Brazil. *Plant Dis.* 100: 561-568 https://doi.org/10.1094/PDIS-03-15-0242-RE.

Santos PHAD (2016). Melhoramento genético do coqueiro (*Cocos nucifera* L.): Capacidade combinatória e diversidade genética via rad-sequencing. Doctoral's Thesis. Universidade Estadual do Norte Fluminense Darcy Riberiro. Available at [http://uenf.br/posgraduacao/gmp/wp-content/uploads/sites/6/2016/02/Tese-DS-Pedro-Henrique-Araujo-Diniz-Santos.pdf].

Souza Filho BF, Santos HP and Robbs CF (1979). Etiologia da queima das folhas do coqueiro. *Fitopatol. Bras.* 4: 5-10.

Subileau C, Renard JL and Lacoste L (1994). *Botryosphaeria cocogena* nov. sp. agent causal du dessechement foliaire du cocotier au Brésil. *Mycotaxon* 41: 5-14.

Summerell BA, Laurence MH, Liew ECY and Leslie JF (2010). Biogeography and phylogeography of *Fusarium*: a review. *Fungal Divers.* 44: 3-13 https://doi.org/10.1007/s13225-010-0060-2.

Templeton AR (1998). Nested clade analyses of phylogeographic data: testing hypotheses about gene flow and population history. *Mol. Ecol.* 7: 381-397 https://doi.org/10.1046/j.1365-294x.1998.00308.x.

Templeton AR, Crandall KA and Sing CF (1992). A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics* 132: 619-633.

Wang Y, Tan ZM, Zhang DC, Murat C, et al. (2006). Phylogenetic and populational study of the Tuber indicum complex. *Mycol. Res.* 110: 1034-1045 https://doi.org/10.1016/j.mycres.2006.06.013.