British Mycological Society promoting fungal science

journal homepage: www.elsevier.com/locate/funbio

High gene flow and outcrossing within populations of two cryptic fungal pathogens on a native and non-native host in Cameroon

Aime Didier BEGOUDE BOYOGUENO[a,b], *Bernard SLIPPERS*[c], *Guillermo PEREZ*[a], *Michael J. WINGFIELD*[a], *Jolanda ROUX*[a,*]

[a]*Department of Microbiology and Plant Pathology, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa*

[b]*Laboratoire Régional de Lutte Biologique et de Microbiologie Appliquée, Institut de Recherche Agricole pour le Développement (IRAD), Nkolbisson, BP 2067, Yaoundé, Cameroun*

[c]*Department of Genetics, Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa*

ARTICLE INFO

Article history:
Received 3 May 2011
Received in revised form
5 December 2011
Accepted 5 December 2011
Available online 27 December 2011
Corresponding Editor:
Andrew N. Mille

Keywords:
Lasiodiplodia
Microsattelites
Population diversity
Taxonomy
Taungya
Terminalia
Cacao

ABSTRACT

In this study, we determined the genetic diversity of 126 isolates representing both *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae*, collected from *Theobroma cacao* and *Terminalia* spp. in Cameroon, using simple sequence repeat (SSR) markers. SSR alleles showed clear genetic distinction between *L. theobromae* and *L. pseudotheobromae*, supporting their earlier separation as sister species. Both *L. theobromae* and *L. pseudotheobromae* populations from Cameroon had high levels of gene diversity, moderate degrees of genotypic diversity, and high levels of gene flow between isolates from *T. cacao* and *Terminalia* spp. There was no evidence for geographic substructure in these populations across the region studied, and the SSR alleles were randomly associated in both species, suggesting outcrossing. The significant levels of aggressiveness, evolutionary potential represented by high levels of diversity, outcrossing and gene flow between geographically and host defined populations, identify these fungi as high-risk pathogens for their native and non-native hosts in Cameroon.

Introduction

Theobroma cacao is native to South America (Purseglove 1968) and was introduced into West Africa towards the end of the 19th century (Havinden 1970). It has become one of the most important cash crops in Cameroon and other West African countries (Havinden 1970). Traditionally, *T. cacao* is planted in the shade of forest trees. Various timber and fruit trees are also intercropped with *T. cacao*. In Cameroon, some of the most popular timber trees planted as a shade crop for *T. cacao* include *Terminalia ivorensis* and *Terminalia superba*. These native tree species are used to establish a 'taungya' agri-sylvicultural system where the production of timber is combined with that of *T. cacao* (Lawson 1995; Norgrove & Hauser 2002).

* *Corresponding author*. Tel.: +27 124203938/9; fax: +27 124203960.
E-mail address: jolanda.roux@fabi.up.ac.za

1878-6146/$ – see front matter © 2011 The British Mycological Society. Published by Elsevier Ltd. All rights reserved.
doi:10.1016/j.funbio.2011.12.001

Insect pests and pathogens are recognized as important constraints to the productivity of agroforestry systems (Epila 1986; Rao et al. 2000). Schroch et al. (2000) provided an extensive review of pests and diseases in agroforestry systems in the humid tropics and highlighted the fact that latent pathogens of one crop could move to other crops grown in association with it. The Botryosphaeriaceae provide an excellent example of latent pathogens of woody plants that move between hosts (Slippers & Wingfield 2007). Because native *Terminalia* spp. and non-native *T. cacao* trees occur in close association in plantations, it is possible that pathogens such as the Botryosphaeriaceae can move between these trees.

Knowledge of the genetic structure of pathogen populations is essential to predict disease epidemics and to develop effective strategies for disease management (McDonald & McDermott 1993; McDonald & Linde 2002a; Ma & Michailides 2005; Burlakoti et al. 2008). Population genetic tools also make it possible to determine the type of reproduction in micro-organisms and they provide insight into the adaptive potential as well as the evolution of a pathogen. Few studies have been conducted on the population genetics of fungi in the Botryosphaeriaceae (Burgess et al. 2004; Mohali et al. 2005; Burgess et al. 2006a), and none of these studies have considered the structure of two closely related species from two different hosts that occur sympatrically. The closely related species, *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae*, that occur on closely associated hosts, provide a useful model to better understand the ecology of the interaction between pathogens and hosts in the taungya system involving *Terminalia* spp. and *T. cacao*.

In this study the genetic diversity and structure of populations of *L. theobromae* and *L. pseudotheobromae* from non-native *T. cacao* and native *Terminalia* spp. in Cameroon were analyzed using polymorphic microsatellite DNA markers. The specific objectives were to: (i) test the integrity of species boundaries between *L. theobromae* and *L. pseudotheobromae*, (ii) determine whether there is population structure in *L. theobromae* and *L. pseudotheobromae* populations from non-native *T. cacao* and native *Terminalia* spp. in Cameroon, (iii) determine the level of gene flow between isolates of these species from different hosts, and (iv) consider the possible mode of reproduction of *L. theobromae* and *L. pseudotheobromae*.

Materials and methods

Fungal isolates

A total of 126 *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae* isolates, collected from two different regions in Cameroon (Table 1), were used for population analyses in this study. Of these, 42 isolates were previously obtained from asymptomatic bark and branches of *Terminalia* spp. in Dec. 2007 and Jan. 2008 (Begoude et al. 2011). The remaining 84 isolates were collected in Nov. 2008 from *Theobroma. cacao* trees showing symptoms of dieback. The trees at Mbalmayo and Nkoemvone were growing as understory to the *Terminalia* trees sampled previously and the same number of trees (20) was sampled at each site. All the collection sites occurred within an area of 250 km^2. One isolate per tree was selected to be used in the population genetic studies. For isolation of fungi from *T. cacao*, the technique described in Begoude et al. (2010) was used. Single conidial cultures were prepared for all isolates and duplicates of these cultures are maintained in the Culture Collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa.

DNA extraction, PCR reactions, and DNA sequencing

To identify isolates collected from *Theobroma cacao*, conidial morphology and DNA sequence data for the Internal Transcribed Spacer regions (*ITS*) of the nrDNA, including the 5.8S gene region was used. Procedures and protocols for genomic DNA extraction and sequencing of representative isolates of the Botryosphaeriaceae from *T. cacao* were as described in Begoude et al. (2010). The identity of isolates representing *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae* was confirmed by comparing the *ITS* sequences of isolates obtained in this study with corresponding sequences in GenBank for isolates CBS 164.96 and CMW 9074, representing *L. theobromae*, and isolates CBS 116459 and CBS 447.62, representing *L. pseudotheobromae*. The phylogenetic analyses for all the datasets were performed using the maximum parsimony (MP) option, with trees generated by heuristic searches with random stepwise addition of 1000 replicates, tree bisection and reconnection (TBR) as branch swapping algorithms, and random taxon addition for the construction of most parsimonious trees. The support for branches of the most parsimonious trees was assessed with a 1000 bootstrap replication (Felsenstein 1985).

Simple sequence repeat (SSR)-PCR and GENESCAN analyses

Thirteen PCR-based SSR microsatellite markers (Burgess et al. 2003) were employed to study the population structures of *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae*

Table 1 – Source of *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae* isolates in Cameroon.

Region	Location	Host	Number of isolates	
			L. theobromae	*L. pseudotheobromae*
Centre	Mbalmayo	*Terminalia ivorensis*	16	22
		Theobroma cacao	16	8
	Nkong	*T. cacao*	9	4
South	Nkoemvone	*T. cacao*	28	18
		Terminalia superba	–	5

isolates. The PCR reactions and conditions were the same as those described by Burgess *et al.* (2003). The DNA concentration of the PCR products was measured visually against the intensity of a 100 bp marker (Roche Molecular Biochemicals) on 2 % agarose gels, exposed to Ultra-violet (UV) illumination.

PCR products were multiplexed for GENESCAN analysis based on the approximate sizes of the PCR products and type of fluorescent label attached to the primer (Burgess *et al.* 2003). Each sample mix contained 1 μl of combined DNA, 0.14 μl 1 × loading buffer and 1 μl internal standard GENE-SCAN-500 LIZ (Applied Biosystems, Warrington, UK). Fluorescent-labelled SSR-PCR products were separated on an ABI Prism 3100 sequencer (Applied Biosystems, Warrington, UK). Allele sizes were determined by comparing the mobility of the SSR products with those of the LIZ internal size standard using a combination of the GENESCAN 2.1 analysis software (Applied Biosystems) and GENOMAPPER V3.5 (Applied Biosystems).

Statistical analyses

Isolates that contained the same alleles at each locus potentially represented clones. The inclusion of multiple clonal representatives can strongly distort estimates of population genetic parameters (Frantz *et al.* 2006). Therefore, duplicates of each multilocus genotype were discarded from the analyses to provide a clone-corrected dataset.

Bayesian clustering analyses

The software programme Structure version 2.2 (Pritchard *et al.* 2000; Falush *et al.* 2003) was used to infer the population structure of all isolates, without any *a priori* knowledge of population subdivision, using a Bayesian model-clustering algorithm. This algorithm assumes a model where there are *K* populations or clusters, in which a set of allele frequencies at each locus characterizes each population. Individuals in the same sample are probabilistically distributed to *K* clusters, or jointly to two or more clusters if their genotypes indicate that they are admixed, regardless of their region or host origin. Each cluster is characterized by a set of allele frequencies at each locus. Loci are assumed to be at Hardy–Weinberg equilibrium, unlinked and at linkage equilibrium. The model with admixture was applied in all simulations as this model is recommended for situations where little is known about the existence of admixture (Falush *et al.* 2003). Priors were assumed uniform for the vectors of proportions q_i of the individual *i*'s genome deriving from each cluster. Iteration parameters were set to 950 000 Monte Carlo Markov Chain (MCMC) iterations preceded by a burn-in period of 50 000 iterations and 20 independent simulations were performed to test for the consistency of the results. The number of clusters, *K*, was varied from 1 to 10. Individuals were assigned to a single cluster when their proportion of ancestry in that cluster was greater than 80 %. This threshold was determined after analyzing the distribution of mean ancestry coefficients for each *K*.

Gene and genotypic diversity

Single alleles were assigned a different letter for each of the loci. For each isolate, a data matrix of multistate characters, each state corresponding to a different locus, was compiled for the polymorphic loci (e.g. ABCDE), thus providing each isolate with a haplotype. A number was assigned to each haplotype and the equation $\hat{G} = 1/\Sigma p_i^2$, developed by Stoddart & Taylor (1988), was applied to estimate the genotypic diversity ($\hat{G}$). In this equation, p_i stands for the observed frequency of the ith phenotype. The maximum percentage of genotypic diversity (G_{max}), obtained from the equation $G_{max} = \hat{G}/N \times 100$ (where *N* is the population size), was used to compare the genotypic diversities between populations (Chen *et al.* 1994). Allelic frequency, as well as the number of alleles at each locus, was calculated and gene diversity determined, using the program POPGENE version 1.31 (Yeh *et al.* 1999) based on the equation $H = 1-\Sigma x_k^2$, where x_k is the frequency of the kth allele (Nei 1973). Chi-square tests for differences in allele frequencies were calculated for each locus across clone-corrected collections. The software Programme Multilocus version 1.2 (Agapow & Burt 2000) available at http://www.agapow.net/software/multilocus/ was used to plot the genotypic diversity against the number of loci with 1000 resampling repetitions, in order to determine whether the isolates and microsatellite markers used were sufficient to recover the maximum genotypic diversity.

Genetic differentiation and gene flow

The genetic differentiation among populations was assessed in POPGENE, using Nei's (1973) G_{ST} statistic, which varies between zero and one. POPGENE was also used to estimate the number of migrants (N_m) exchanged among the populations for each generation from the estimate of G_{ST} where $N_m = 0.5(1-G_{ST})/G_{ST}$ (McDonald & McDermott 1993). Populations that are completely genetically isolated would have N_m values of zero and G_{ST} values tending towards one (Hartl & Clark 1989). The software programme GENALEX 6.2 (Peakall & Smouse 2006) was used to analyze the molecular variance (AMOVA) among populations of *Lasiodiplodia* spp. from *Theobroma cacao* and *Terminalia* spp. from different locations and hosts.

Linkage disequilibrium

The multilocus linkage disequilibrium for each clone-corrected population was tested with the Index of Association (I_A) (Maynard Smith *et al.* 1993). The I_A provides information related to whether two different individuals which possess the same allele at one locus, will more likely possess the same allele at another locus (Fournier & Giraud 2008). The tests were performed on a data matrix of multistate characters using the program Multilocus (Agapow & Burt 2000). For any pair of individuals, the number of loci at which they differ was calculated and its variance was compared with the expected value ($I_A = 0$). There is no linkage disequilibrium when the observed data fall within the distribution range of the recombined data, but the population is most likely

influenced by clonal reproduction if the observed data fall outside the distribution range with a significant value of $P < 0.05$.

Results

Fungal isolates

A total of 16 *Lasiodiplodia theobromae* isolates and 26 *Lasiodiplodia pseudotheobromae* isolates were obtained from *Terminalia ivorensis* and *Terminalia superba* in a previous study (Begoude *et al.* 2011). An additional 84 isolates of *Lasiodiplodia* were collected from *Theobroma cacao* in the present study.

Lasiodiplodia isolates from *T. cacao* were identified to species level using DNA sequence data for the *ITS* and 5.8S gene regions. The *ITS* dataset comprised 114 sequences of which 97 originated from *Terminalia* spp. and *T. cacao* and 17 sequences were retrieved from GenBank. Of the 486 characters present in the *ITS* sequence dataset, 34 were parsimony informative. The MP analyses generated two trees with identical topology [Tree length (TL) = 129, Consistency index (CI) = 0.698, Retention index (RI) = 0.839, Rescaled CI (RC) = 0.585]. These analyses revealed that all 97 isolates from *Terminalia* spp. and *T. cacao* belonged to the clades accommodating either *L. theobromae* [Bootstrap support (BS) = 55 %] or *L. pseudotheobromae* (BS = 77 %) (Fig 1). Of the isolates from *T. cacao*, 54 represented *L. theobromae* and 33 *L. pseudotheobromae* (Fig 3).

Microsatellite PCR amplification

Eleven of the 13 pairs of microsatellite primer pairs developed by Burgess *et al.* (2003) successfully amplified DNA markers for *Lasiodiplodia theobromae* from Cameroon. Among these primer pairs, five were polymorphic for *L. theobromae* in Cameroon (Table 2). Nine of the 13 microsatellite primer pairs previously developed for *L. theobromae* (Burgess *et al.* 2003) successfully amplified the expected loci in Cameroonian isolates of *Lasiodiplodia pseudotheobromae*, among which five were polymorphic (Table 3). PCR products from primer pairs *las15* & *las16*, *las27* & *las28*, and *las29* & *las30*, which were polymorphic among isolates of *L. pseudotheobromae*, were monomorphic among isolates of *L. theobromae*. Overall, seven primer pairs were polymorphic among isolates of both species (Fig 2) and five primer pairs were polymorphic among isolates of only one of the species.

Statistical analyses

Bayesian clustering analyses

The Bayesian inference of the population structure was performed with 21 unique haplotypes representing all the multilocus genotypes inferred with seven polymorphic loci among isolates of *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae*. These samples included isolates from *Terminalia* spp. and *Theobroma cacao* from all the locations sampled. The distribution of the maximum likelihood was the highest for $K = 2$ with an assignment rate value of 98.8 %. The first cluster included all the genotypes of isolates representing *L. pseudotheobromae* while the second cluster consisted of genotypes of *L. theobromae*. There was no subdivision in the population according to either host or location. Separate investigation of the population structure within each species, showed that the distribution of the maximum likelihood was the highest for $K = 1$, indicating a high degree of admixture, which suggests that neither the host nor the geographic location influenced the population structure within the species.

Gene diversity

The allelic diversity of 69 isolates of *Lasiodiplodia theobromae* from *Terminalia* spp. and *Theobroma cacao* were analyzed at five polymorphic loci (Table 2). The number of alleles ranged from two to four per locus. A total of 13 alleles were produced across populations from *Terminalia* spp. and *T. cacao*, of which nine alleles were observed across isolates from *Terminalia* spp. and all 13 alleles were observed among isolates from *T. cacao*. Four unique alleles, with low frequency (8–24 %), were observed in isolates from *T. cacao*. The mean total gene diversity (H), calculated using the allele frequencies across all isolates of *L. theobromae* was 0.46, which was similar to the gene diversity observed in isolates from *T. cacao* and higher than the gene diversity observed in isolates from *Terminalia* spp. (Table 2).

Fifty-three isolates of *Lasiodiplodia pseudotheobromae* from *Terminalia* spp. and *T. cacao* were analyzed at five polymorphic loci after clone correction of populations (Table 3). The number of alleles ranged from two to six per locus. A total of 14 alleles were identified across isolates from both hosts in which 13 alleles were observed among isolates from *Terminalia* spp. and 12 alleles were observed in isolates from *T. cacao*. Two unique alleles, with low frequency (16 %), were observed in isolates from *Terminalia* spp. and only one unique allele, with low frequency (12 %), was observed in isolates originating from *T. cacao*. The mean total gene diversity across all isolates of *L. pseudotheobromae* was 0.445, which was similar to the gene diversity observed in isolates from each host (Table 3).

Genotypic diversity

Among the 69 isolates of *Lasiodiplodia theobromae*, 26 different multilocus genotypes were discriminated. Of these genotypes, 19 were unique to the sampled localities (three in Nkong, six in Mbalmayo and ten in Nkoemvone) whereas seven genotypes, representing 60.9 % of the isolates collected, were shared among the three localities (Fig 3). Where *Terminalia* spp. and *Theobroma cacao* occurred in the same area, such as in Mbalmayo, of 14 genotypes found in the area, three genotypes were shared between both hosts, representing 65.6 % of the isolates collected.

When considering isolates of *L. theobromae* from the two hosts separately, six genotypes were found amongst isolates from *Terminalia* spp. and 25 genotypes were found amongst isolates from *T. cacao* (Table 4). Of these genotypes, only one genotype (33.3 % of the isolates) was unique to *Terminalia* spp. whereas 16 genotypes (64 % of the isolates) were unique to the *T. cacao* population. The percentage of shared genotypes between *Terminalia* spp. and *T. cacao* represented 19 % of all the genotypes observed. The most common genotypes (AABAA)

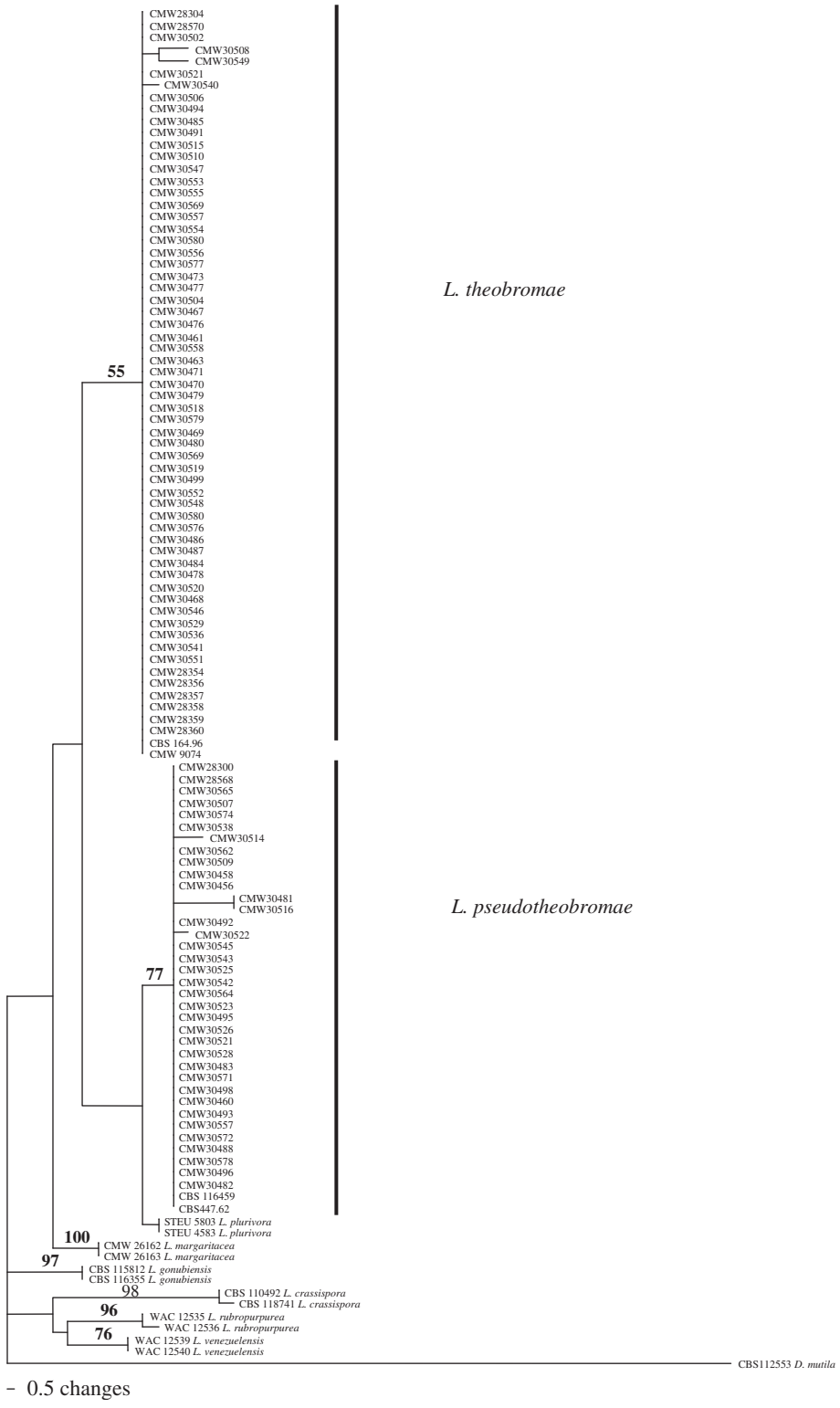

Fig 1 – MP phylogram of *Lasiodiplodia theobromae* and *L. pseudotheobromae* from this study obtained with sequences of *ITS*. Bootstrap support (%) from 1000 replications is given on the branches.

accounted for 14.5 % of all the isolates included in the *L. theobromae* population, while genotypes occurring only once were most abundant amongst isolates from *T. cacao* (Table 4). Overall, low values were generated for the genotypic diversities in each population, ranging from 28.57 % on the *Terminalia* spp. to 32.12 % for isolates from *T. cacao*. These values of genotypic diversity are also reflected by the low number of single isolate genotypes (12 and 23.2 % for each host, respectively).

Table 2 – Allele frequencies at five SSR loci for *Lasiodiplodia theobromae* populations collected from *Terminalia* spp. and *Theobroma cacao* in Cameroon.

Locus	Allele length	Allele configuration	*Terminalia* spp.	*T. cacao*
las21–22	383	A	–	0.24
	388	B	1	0.76
las23–24	454	A	–	0.08
	458	B	0.333	0.48
	461	C	0.667	0.36
	463	D	–	0.08
las25–26	417	A	0.333	0.2
	420	B	0.5	0.48
	421	C	0.167	0.32
las35–36	376	A	0.333	0.4
	379	B	0.667	0.6
las37–38	117	A	1	0.8
	135	B	–	0.2
No isolates			**16**	**53**
No alleles			**9**	**13**
No. unique alleles			**0**	**4**
Polymorphic loci			**3**	**5**
H			**0.3**	**0.484**

Ten different multilocus genotypes were detected amongst the 57 isolates of *Lasiodiplodia pseudotheobromae* from *Terminalia* spp. and *T. cacao*. Among these genotypes, three were unique to the localities of Mbalmayo (one) and Nkoemvone (two), whereas seven genotypes, representing 87.7 % of the isolates collected, were shared amongst the three sampled locations (Fig 3). Of seven genotypes obtained in Mbalmayo, where isolates were collected from both *Terminalia* spp. and *T. cacao*, four were shared between the hosts and represented 90 % of all the isolates collected in this location. Similarly, two of nine genotypes obtained in Nkoemvone, representing 65.2 % of the isolates collected in the area, were shared between *Terminalia* spp. and *T. cacao*.

Table 3 – Allele frequencies at five SSR loci for *Lasiodiplodia pseudotheobromae* populations collected from *Terminalia* spp. and *Theobroma cacao* in Cameroon.

Locus	Allele length	Allele configuration	*Terminalia* spp.	*T. cacao*
las15–16	351	A	0.833	0.625
	353	B	0.167	0.375
las21–22	383	A	0.833	0.625
	388	B	0.167	0.375
las25–26	415	A	0.333	0.25
	417	B	0.667	0.75
las27–28	458	A	0.167	-
	463	B	0.167	0.375
	466	C	0.167	-
	471	D	0.167	0.125
	474	E	0.333	0.375
	477	F	-	0.125
las29–30	180	A	0.833	0.875
	188	B	0.167	0.125
No isolates			**26**	**31**
No alleles			**13**	**12**
No. unique alleles			**2**	**1**
Polymorphic loci			**5**	**5**
H			**0.41**	**0.44**

When considering isolates from different hosts, regardless of their locality of origin, six different genotypes of *L. pseudotheobromae* were found among isolates from *Terminalia* spp. and eight genotypes were found among isolates from *T. cacao* (Table 5). Among these genotypes, two (33.3 % of the isolates) were unique to *Terminalia* spp. and four genotypes (50 % of the isolates) were unique to the *T. cacao* population. The percentage of shared genotypes between *Terminalia* spp. and *T. cacao* represented 40 % of all the genotypes observed for *L. pseudotheobromae*. The most common genotypes (BAAEA) accounted for 30.2 % of the isolates included in the population of *L. pseudotheobromae* and the genotypes occurring only once were rare in populations from both *Terminalia* spp. and *T. cacao* (Table 5). The overall genotypic diversities calculated for each population were low, ranging from 17.57 % on *Terminalia* spp. to 15.97 % for the *T. cacao* population. This was also reflected in the low number of single isolate genotypes (7.6 and 6.5 % for each host, respectively).

Genetic differentiation and gene flow

The measure of genetic differentiation between populations of *Lasiodiplodia theobromae* from *Terminalia* spp. and *Theobroma cacao* reflected a lack of substructuring in the *L. theobromae* population. The values obtained for χ^2 tests revealed no significant differences ($P > 0.05$) in allele frequencies at any loci for populations from either the *Terminalia* spp. or *T. cacao* (Table 6). These results were further supported by very low G_{ST} values (0.046), indicating that most of the gene diversity is found within the subpopulations (*Terminalia* spp. and *T. cacao*). This was also true when comparing populations of *L. theo-bromae* from different hosts at different locations. Consequently, a low level of differentiation exists in populations of *L. theobromae* from *Terminalia* spp. and *T. cacao*. The number of migrants (N_m) exchanged between populations per generation was estimated at 10.47.

Similar to *L. theobromae*, the measure of genetic differentiation between populations of *Lasiodiplodia pseudotheobromae* from *Terminalia* spp. and *T. cacao* showed a lack of substructuring. The values obtained for χ^2 tests revealed no significant differences ($P > 0.05$) in allele frequencies at any loci for either the *Terminalia* spp. or *T. cacao* populations of *L. pseudotheobromae* (Table 6). There was only 3.5 % genetic diversity distributed between populations from *Terminalia* spp. and *T. cacao* and no difference was observed after comparing populations of *L. pseudotheobromae* from different hosts at different locations. This indicated that most of the genetic variation is distributed within each subpopulation. Therefore, a low level of differentiation also exists in populations of *L. pseudotheobromae* from *Terminalia* spp. and *T. cacao*. The number of migrants (N_m) exchanged between populations per generation was estimated at 13.83.

Linkage disequilibrium

The I_A calculated for populations of *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae* were −0.153 ($P = 0.99$) and 0.069 ($P = 0.4$), respectively. These values did not significantly

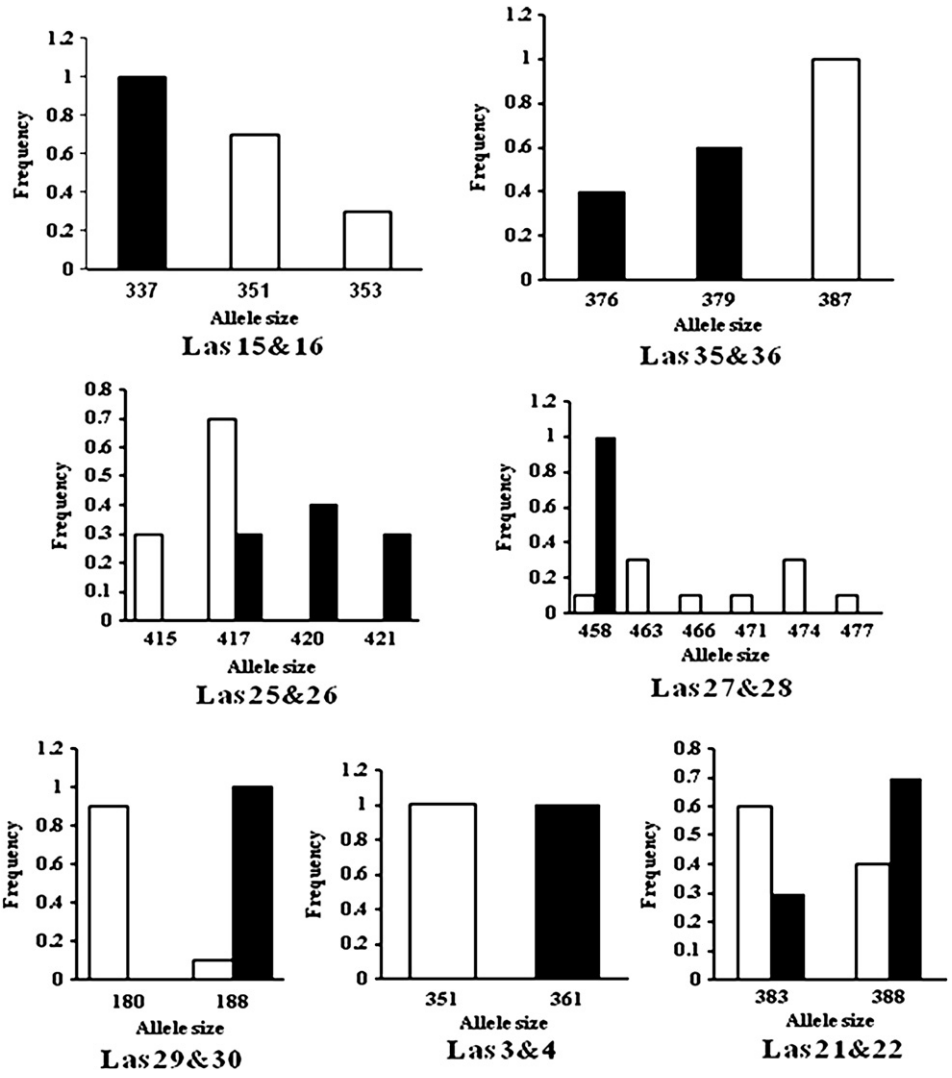

Fig 2 – Distribution of alleles showing the size of PCR product at seven microsatellite loci for *Lasiodiplodia pseudotheobromae* (white) and *L. theobromae* (black).

deviate from the expected value when there is no linkage disequilibrium. This suggests that alleles are randomly associated, as would be expected for outcrossing populations.

Discussion

Genetic diversity analyses for *Lasiodiplodia theobromae* and *Lasiodiplodia pseudotheobromae* in Cameroon showed that there is a complete lack of gene flow between these recently described taxa in this area. This supports the previous segregation of *L. pseudotheobromae* from *L. theobromae* as distinct cryptic species based on divergence in sequences of their nuclear genes (Alves *et al.* 2008). Both *L. theobromae* and *L. pseudotheobromae* have a wide distribution globally and they share similar biological and ecological characteristics. Since their description as cryptic species, these fungi have not been studied in areas where they occur on the same host or environment, which is where possible hybridization of two closely related fungal species might be expected (Schardl & Craven 2003).

A number of lines of evidence supported the distinction of *L. theobromae* and *L. pseudotheobromae*. Of the 22 alleles that were detected, 11 and six were unique to populations of *L. pseudotheobromae* and *L. theobromae*, respectively. A number of these were fixed in either species, and any recombination between them would have shared these alleles. The five alleles that were shared between these species occurred in significantly different frequencies. It was thus no surprise that the Bayesian clustering algorithm used in this study clustered all the individuals in one or the other of the clusters representing the two species. The alleles shared between populations of *L. theobromae* and *L. pseudotheobromae* in this study reflects its recent separation (Carbone & Kohn 2004).

Delimitating boundaries between sister species with low levels of genetic divergence is challenging. Two methods are used in species delimitation, one of which encompasses tree-based approaches that delimit species as historical lineages (Goldstein & Desalle 2000). The other method includes non tree-based analyses where information regarding the

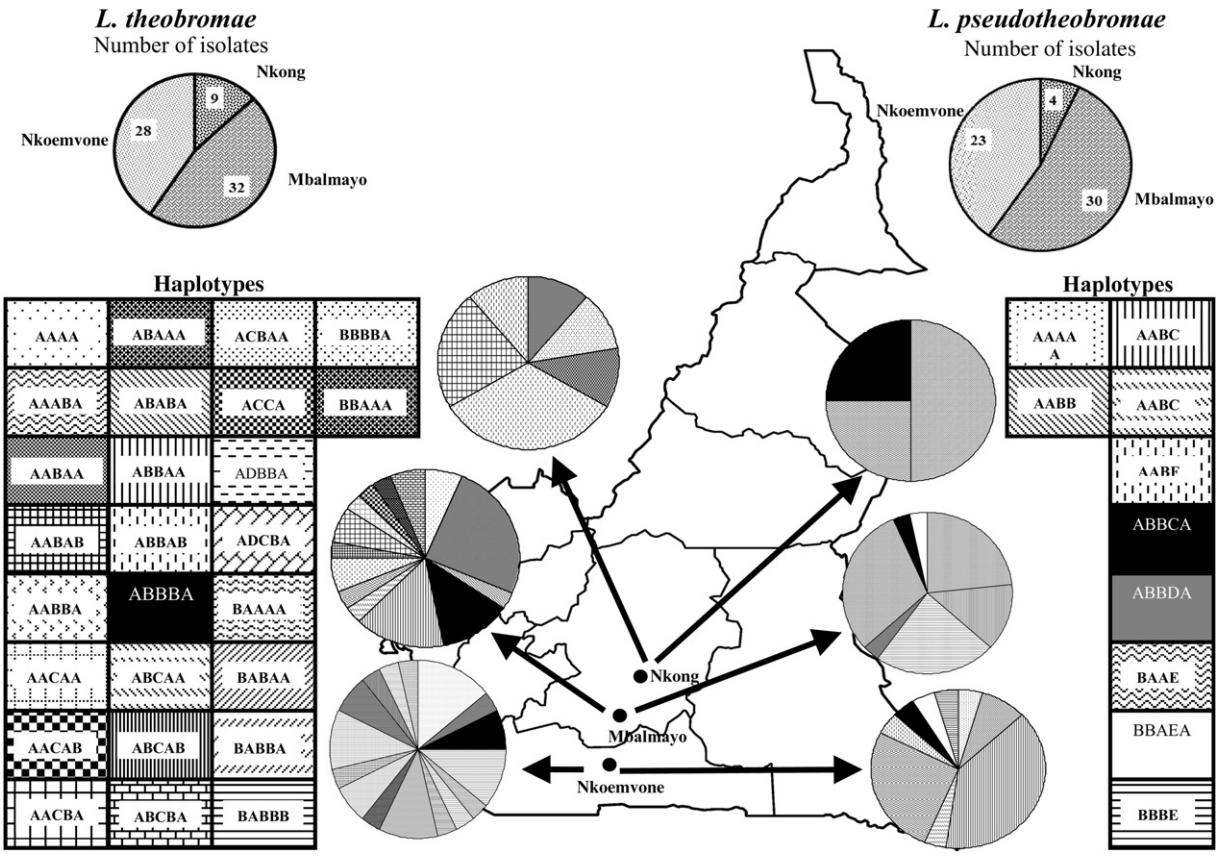

Fig 3 – Map of collecting sites and distribution of the 26 and ten haplotypes of *Lasiodiplodia theobromae* and *L. pseudotheobromae*, respectively among the three locations. Each pie chart linked with arrows represents a collecting site and its haplotypes. The pie charts topping the left and right sides represent the number of isolates of *L. theobromae* and *L. pseudotheobromae*, respectively collected per locality.

level of gene flow is the main basis to determine boundaries between species (Sites & Marshall 2003, 2004). Application of microsatellite markers, which fall within the latter category, represents a powerful approach to demarcate barriers to gene flow between individuals of closely related fungal species (also see Fisher *et al.* 2000, 2002). This approach is also illustrated in this study where the distinction between *L. theobromae* and *L. pseudotheobromae* using a non tree-based method with microsatellite markers, provided strong additional support for their separation based on phylogenetic analyses of *ITS* and *TEF1-α* loci (Alves *et al.* 2008).

No evidence of host or geography linked population structure was observed for either *L. theobromae* or *L. pseudotheobromae* in Cameroon. The population of *L. theobromae* and *L. pseudotheobromae* on *Theobroma cacao* and native *Terminalia* spp. appeared to be totally integrated, suggesting that the movement of these pathogens between the hosts may be symmetrical (Hayden *et al.* 2007; Fournier & Giraud 2008). The maximum distance between collection sites was ~200 km. Agro-ecology in Cameroon is subdivided into five zones based on vegetation and climatic conditions (http://www.irad-cameroon.org/carte_us.php). The collection sites in this study occurred within zone five, characterized by humid forests with bimodal rainfall. The hosts are fairly continuous over the area studied, providing a possible explanation for the connectedness of the populations. Endophytic Botryosphaeriaceae infections of woody hosts are thought to develop over time by horizontal transmission through wind- or water-dispersed spores (Arnold & Herre 2003a, b; Slippers & Wingfield 2007). Movement of infected material would provide another explanation for genetic similarity between populations. When establishing cacao in the shade of thinned forest trees, farmers obtain seedlings from a centralized seedling distributor (Sonwa 2002). This could have contributed to the spread of the pathogens as endophytes over large areas.

We detected no restriction in the movement of *L. theobromae* and *L. pseudotheobromae* between *Terminalia* spp. and *T. cacao*. Most alleles were found in isolates on both hosts, and at similar frequencies. This result is not unexpected, because both *L. theobromae* and *L. pseudotheobromae* commonly occur in the tropics and subtropics on a wide diversity of hosts with no observed specialization (Mohali *et al.* 2005; Alves *et al.* 2008). Many cases of shared genotypes were also observed on both *Terminalia* spp. and *T. cacao*, possibly representing clonal lineages. These occurred either in the same field or among locations (over scales of a few metres to 200 km). Both populations displayed a high allelic diversity and almost all of the alleles were shared between isolates from *T. cacao* and *Terminalia* spp. As explained above, this could reflect natural spread through asexual conidia in a step-wise manner, or

Table 4 – Genotype estimation from multilocus profiles generated from five SSR loci for *Lasiodiplodia theobromae*.

Genotypes	*Terminalia* spp.	*T. cacao*
AAAAA	2	4
AABAA	5	5
ABABA	1	
ABBBA	3	3
ABBAA	4	1
ABCAA	1	3
BAAAA		1
BABBA		1
BABBB		1
BBBBA		1
BABAA		3
BBAAA		1
ACBAA		1
ACCAA		1
AAABA		1
AABBA		5
AABAB		1
AACBA		5
AACAA		4
AACAB		1
ABAAA		1
ABBAB		1
ABCBA		4
ABCAB		1
ADBBA		1
ADCBA		1
N	**16**	**53**
N(g)	**6**	**25**
$\hat{G}$	**4.751**	**17.024**
$\hat{G}$(%)	**28.57**	**32.12**

N, number of isolates.
N(g), number of genotypes.
$\hat{G}$, Genotypic diversity (Stoddart & Taylor 1988).
$\hat{G}$(%), % max diversity.

Table 5 – Genotype estimation from multilocus profiles generated from five SSR loci for *Lasiodiplodia pseudotheobromae*.

Genotypes	*Terminalia* spp.	*T. cacao*
AAAAA	1	
AABBA	6	5
AABCA	5	8
AABCB	6	2
ABBDA	1	
BAAEA	7	9
AABFA		1
ABBCA		3
BBAEA		3
BBBEA		1
N	**26**	**31**
N(g)	**6**	**8**
$\hat{G}$	**4.56**	**4.95**
$\hat{G}$(%)	**17.57**	**15.97**

N, number of isolates.
N(g), number of genotypes.
$\hat{G}$, Genotypic diversity (Stoddart & Taylor 1988).
$\hat{G}$(%), % max diversity.

spread by direct transport of infected material from a central location. Any new genotypes introduced on one host, or emerging on one of them, are likely to freely move between hosts. Given the distant relationship between *T. cacao* and *Terminalia* spp., this is expected to also reflect the situation of populations of these fungi on other hosts.

The geographic and host origin of *L. theobromae* and *L. pseudotheobromae* remain unresolved. Native populations or populations closest to their centre of origin generally have high levels of genetic diversity while introduced populations often exhibit lower levels of diversity (McDonald & Linde 2002a; Stukenbrock *et al.* 2006). This study revealed high genetic diversity in populations of both *L. theobromae* and *L. pseudotheobromae* from non-native *T. cacao* and native *Terminalia* spp. in Cameroon, suggesting that this might be the region of origin of these fungi. It is, however, difficult to apply this rule in the case of these fungi, due to the possibility of numerous introductions through anthropogenic action and high levels of gene flow across large areas and between hosts, which could influence the levels of diversity. As example, very high levels of genetic diversity exist in populations of *Diplodia pinea* in many regions of the world, making it difficult to predict its origin (Burgess *et al.* 2004; Bihon *et al.* 2012). The only other population study of *L. theobromae*, compared isolates on three non-native hosts in Venezuela and reported lower diversities than those found in Cameroon (Mohali *et al.* 2005).

At the time of its description, *L. pseudotheobromae* was known only from a few hosts and it had a limited known geographic distribution (Alves *et al.* 2008). Its known distribution in Africa has, however, expanded rapidly since that time and studies are beginning to reflect the common occurrence of *L. pseudotheobromae* in tropical environments (van der Walt 2008; Begoude *et al.* 2010, 2011; Mehl *et al.* 2011). It is highly likely that *L. pseudotheobromae* is much more common, over a wide geographic area, than has been reflected in recent reports. This, together with its proven pathogenicity (Begoude *et al.* 2010) makes it an important pathogen. The ability to distinguish it from *L. theobromae* is thus of critical importance for programs focussed on selection and breeding for resistance.

Analyses of the linkage disequilibrium amongst alleles at the SSR loci in populations of *L. theobromae* and *L. pseudotheobromae* from *Terminalia* spp. and *T. cacao* suggest that both species undergo regular sexual reproduction. However, despite this evidence of sexual reproduction, sexual states for this group of fungi are rarely seen. *Botryosphaeria rhodina* (Berk. & M.A. Curtis) Arx [which is no longer a valid name following Crous *et al.* (2006)] is frequently reported as the sexual state of *L. theobromae*. However, this connection has not been confirmed since Stevens (1925, 1926) reported *B. rhodina* as the teleomorph of *L. theobromae*. The recent description of a number of species that were previously confused with *L. theobromae* (Pavlic *et al.* 2004; Burgess *et al.* 2006b; Damm *et al.* 2007; Alves *et al.* 2008; Pavlic *et al.* 2008; Begoude *et al.* 2011) sheds doubt on the accuracy of those discoveries. Although the purported sexual state might be known for *L. theobromae*, there is no morphological evidence for such a state in *L. pseudotheobromae*.

Populations of both *L. theobromae* and *L. pseudotheobromae* exhibited a low degree of genotypic diversity. The level of genotypic diversity obtained for *L. theobromae* isolates from both *Terminalia* spp. and *T. cacao*, suggests that a low level of

Table 6 – Gene diversity (*H*) for the five SSR loci across clone-corrected populations of *Lasiodiplodia theobromae* and *L. pseudotheobromae* in Cameroon.

	Gene diversity (*H*)							
	L. theobromae		*L. pseudotheobromae*		*L. theobromae*		*L. pseudotheobromae*	
Loci	*Terminalia* spp.	*T. cacao*	*Terminalia* spp.	*T. cacao*	χ^2	*df*	χ^2	*df*
las15–15			0.28	0.47			0.7	1
las21–22	0.00	0.36	0.28	0.47	1.8	1	0.7	1
las23–24	0.44	0.63			2.3	3		
las25–26	0.61	0.63	0.44	0.37	0.8	2	0.1	1
las27–28			0.78	0.69			3.9	5
las29–30			0.28	0.22			0.04	1
las35–36	0.44	0.48			0.1	1		
las37–38	0.00	0.32			1.4	1		
N	6	25	6	8				
Mean	0.30	0.48	0.41	0.44				

recombination takes place in this fungus. These results are different from those of Mohali *et al.* (2005) who found very low levels of recombination for *L. theobromae* from *Pinus* sp., *Acacia* sp. and *Eucalyptus* sp. in Venezuela. In the *L. theobromae* population, the number of alleles and genotypes observed in isolates from *T. cacao* was higher than that found in isolates from native *Terminalia* spp., resulting in a higher genetic diversity for isolates from *T. cacao*. This could, however, be explained by the larger number of isolates collected from *T. cacao*.

Populations of *L. pseudotheobromae* were characterized by a low number of single isolate genotypes. Indeed, 60 % of the total genotypes occurred more than once and the proportion of the most common genotype (30.2 %) was high, resulting in a clonal fraction of 82.4 %. These results indicate the presence of a high proportion of widely distributed clonal genotypes across both *Terminalia* spp. and *T. cacao*, despite some evidence of recombination. Although similar observations were made for *L. theobromae*, the frequency of recombination was higher than that in *L. pseudotheobromae*.

Lasiodiplodia theobromae and *L. pseudotheobromae* correspond to the highest category of evolutionary risk for plant pathogens as defined by McDonald & Linde (2002a, b). These authors define this category by fungi capable of both sexual and asexual reproduction. This is because the combination of alleles generated through regular recombination with the highest level of fitness could be increased rapidly through asexual reproduction (Ciampi *et al.* 2008). This risk is increased in cases where there is high gene flow over large areas (as we observed here), because genotypes with high levels of fitness can spread rapidly. Together with data from previous studies showing the pathogenicity of these fungi (Begoude *et al.* 2010), they clearly pose a significant threat to both native *Terminalia* spp., and introduced *T. cacao*. Given their wide host and geographic ranges, this is most likely also true for other native and non-native hosts growing in close proximity to each other.

Acknowledgements

This study was financially supported by the DST/NRF Centre of Excellence in Tree Health Biotechnology (CTHB) and the University of Pretoria, South Africa. We also thank the Institute of Agricultural Research for Development (IRAD), the International Institute of Tropical Agriculture (IITA) in Cameroon for logistic support. Mr Michael Mbenoun and other colleagues at IRAD are gratefully acknowledged for providing us with fungal isolates.

REFERENCES

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

Agapow PM, Burt A, 2000. *Multilocus* 1.2. Department of Biology, Imperial College, Silwood Park, Ascot.

Arnold EA, Herre EA, 2003a. Canopy cover and leaf age affect colonization by tropical fungal endophytes: ecological pattern and process in *Theobroma cacao* (Malvaceae). *Mycologia* **95**: 388–398.

Arnold EA, Mejia LC, Kyllo D, Rojas EI, Maynard Z, Robbins N, Herre EA, 2003b. Fungal endophytes limit pathogen damage in a tropical tree. *Proceedings of the National Academy of Sciences* **100**: 15649–15654.

Begoude BAD, Slippers B, Wingfield MJ, Roux J, 2010. Botryosphaeria associated with *Terminalia catappa* in Cameroon, South Africa and Madagascar. *Mycological Progress* **9**: 101–123.

Begoude BAD, Slippers B, Wingfield MJ, Roux J, 2011. Identity and pathogenicity of botryosphaeriaceous fungi on *Terminalia* species in Cameroon. *Forest Pathology* **41**: 281–292.

Bihon W, Slippers B, Burgess T, Wingfield MJ, Wingfield BD, 2012. Diverse sources of infection and cryptic recombination revealed in South African Diplodia pinea populations. *Fungal Biology* **116:** 112–120.

Burgess TI, Wingfield MJ, Wingfield BD, 2003. Development and characterization of microsatellite loci for the tropical tree pathogen *Botryosphaeria rhodina*. *Molecular Ecology Notes* **3**: 91–94.

Burgess TI, Gordon TR, Wingfield MJ, Wingfield BD, 2004. Geographic isolation of *Diplodia scrobiculata* and its association with native *Pinus radiata*. *Mycological Research* **108**: 1399–1406.

Burgess TI, Sakaladis M, Hardy GEStJ, 2006a. Gene flow of the canker pathogen *Botryosphaeria australis* between *Eucalyptus globules* plantations and native eucalypt forests in Western Australia. *Australasian Ecology* **31**: 559–566.

Burgess TI, Barber PA, Mohali S, Pegg G, de Beer W, Wingfield MJ, 2006b. Three new *Lasiodiplodia* spp. from the tropics, recognized based on DNA sequence comparisons and morphology. *Mycologia* **98**: 423–435.

Burlakoti RR, Ali S, Secor GA, Neate SM, McMullen MP, Adhikari TB, 2008. Genetic relationship among populations of *Gibberella zeae* from barley, wheat, potato, and sugar beet in the upper Midwest of the United States. *Phytopathology* **98**: 969–976.

Carbone I, Kohn L, 2004. Inferring process from pattern in fungal population genetics. *Applied Mycology and Biotechnology* **4**: 1–29.

Chen RS, Boeger JM, McDonald BA, 1994. Genetic stability in a population of a plant pathogenic fungus over time. *Molecular Ecology* **3**: 209–218.

Ciampi MB, Meyer MC, Costa MJN, Zala M, McDonald BA, Ceresini PC, 2008. Genetic structure of populations of *Rhizoctonia solani* anastomosis group-1 IA from soybean in Brazil. *Phytopathology* **98**: 932–941.

Crous PW, Slippers B, Wingfield MJ, Rheeder J, Marasas WFO, Phillips AJL, Alves A, Burgess T, Barber P, Groenewald JZ, 2006. Phylogenetic lineages in the Botryosphaeriaceae. *Studies in Mycology* **55**: 239–257.

Damm U, Crous PW, Fourie PH, 2007. Botryosphaeriaceae as potential pathogens of *Prunus* species in South Africa, with descriptions of *Diplodia africana* and *Lasiodiplodia plurivora* sp. nov. *Mycologia* **99**: 664–680.

Epila JSO, 1986. The case for insect pest management in agroforestry research. *Agricultural Systems* **19**: 37–54.

Falush D, Stephens M, Pritchard JK, 2003. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Fisher MC, Koenig G, White TJ, Taylor JW, 2000. A test for concordance between the multilocus genealogies of genes and microsatellites in the pathogenic fungus *Coccidioides immitis*. *Molecular Biology and Evolution* **17**: 1164–1174.

Fisher MC, Koenig G, White TJ, Taylor JW, 2002. Molecular and phenotypic description of *Coccidioides posadasii* sp. nov., previously recognized as the non-California population of *Coccidioides immitis*. *Mycologia* **94**: 73–84.

Felsenstein J, 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**: 783–791.

Fournier E, Giraud T, 2008. Sympatric genetic differentiation of a generalist pathogenic fungus, *Botrytis cinerea*, on two different host plants, grapevine and Bramble. *Journal of Evolutionary Biology* **21**: 122–132.

Frantz A, Plantegenest M, Mieuzet L, Simon JC, 2006. Ecological specialization correlates with genotypic differentiation in sympatric host-populations of the pea aphid. *Journal of Evolutionary Biology* **19**: 392–401.

Goldstein PZ, Desalle R, 2000. Phylogenetic species, nested hierarchies, and character fixation. *Cladistics* **16**: 364–384.

Hartl DL, Clark AG, 1989. *Principles of Population Genetics*. Sinauer Associates, Sunderland, MA.

Hayden HL, Cozijnsen AJ, Howlett BJ, 2007. Microsatellite and minisatellite analysis of *Leptosphaeria maculans* in Australia reveals regional genetic differentiation. *Phytopathology* **97**: 879–887.

Havinden MA, 1970. The history of crop cultivation in West Africa: a bibliographical guide. *The Economic History Review* **23**: 532–555.

Lawson GJ, 1995. *Growth of Indigenous Tree Plantations in the Mbalmayo Forest Reserve. Report*. Institute of Terrestrial Ecology, Edinburgh.

Ma Z, Michailides TJ, 2005. Genetic structure of *Botrytis cinerea* from different host plants in California. *Plant Disease* **89**: 1083–1089.

Maynard Smith J, Smith NH, O'Rourke M, Spratt BG, 1993. How clonal are bacteria? *Proceedings of the National Academy of Sciences* **90**: 4384–4388.

McDonald BA, McDermott JM, 1993. Population genetics of plant pathogenic fungi. *Bioscience* **43**: 311–319.

McDonald BA, Linde C, 2002a. The population genetics of plant pathogens and breeding strategies for durable resistance. *Euphytica* **124**: 163–180.

McDonald BA, Linde C, 2002b. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Review of Phytopathology* **40**: 349–379.

Mehl JWM, Slippers B, Roux J, Wingfield MJ, 2011. Botryosphaeriaceae associated with *Pterocarpus angolensis* (kiaat) in South Africa. *Mycologia* **103**: 534–553.

Mohali S, Burgess TI, Wingfield MJ, 2005. Diversity and host association of the tropical tree endophyte *Lasiodiplodia theobromae* revealed using simple sequence repeat markers. *Forest Pathology* **35**: 385–396.

Nei M, 1973. Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Sciences* **70**: 3321–3323.

Norgrove L, Hauser S, 2002. Measured growth and tree biomass estimates of *Terminalia ivorensis* in the 3 years after thinning to different stand densities in an agrisilvicultural system in southern Cameroon. *Forest Ecology and Management* **166**: 261–270.

Pavlic D, Slippers B, Coutinho TA, Gryzenhout M, Wingfield MJ, 2004. *Lasiodiplodia gonubiensis* sp. nov., a new *Botryosphaeria* anamorph from native *Syzygium cordatum* in South Africa. *Studies in Mycology* **50**: 313–322.

Pavlic D, Wingfield MJ, Barber P, Slippers B, Hardy GEStJ, Burgess TI, 2008. Seven new species of the Botrysophaeriaceae from baobab and other native trees in Western Australia. *Mycologia* **100**: 851–866.

Peakall R, Smouse PE, 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* **6**: 288–295.

Pritchard JK, Stephens M, Donnelly P, 2000. Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Purseglove JW, 1968. *Tropical Crops: Dicotyledons*. Longman Group, London, p. 719.

Rao MR, Singh MP, Day R, 2000. Insect pest problems in tropical agroforestry systems: contributory factors and strategies for management. *Agroforestry Systems* **50**: 243–277.

Schroch G, Krauss U, Gasparotto L, Duarte Aguilar JA, Vohland K, 2000. Pests and diseases in agroforestry systems of the humid tropics. *Agroforestry Systems* **50**: 199–241.

Schardl CL, Craven KD, 2003. Interspecific hybridization in plant associated fungi and oomycetes: a review. *Molecular Ecology* **12**: 2861–2873.

Sites Jr JW, Marshall JC, 2003. Delimiting species: a Renaissance issue in systematic biology. *Trends in Ecology and Evolution* **18**: 462–470.

Sites Jr JW, Marshall JC, 2004. Operational criteria for delimiting species. *Annual Review of Ecology and Evolutionary Systematics* **35**: 199–227.

Slippers B, Wingfield MJ, 2007. Botryosphaeriaceae as endophytes and latent pathogens of woody plants: diversity, ecology and impact. *Fungal Biology Reviews* **21**: 90–106.

Sonwa DJ, 2002. *Etude de cas d'aménagement forestier exemplaire en Afrique centrale: les systems agroforestiers cacaoyers, Cameroun. Document de travail FM/12F*. Service de la mise en valeur des ressources forestières, Division des ressources forestières. FAO, Rome.

Stevens NE, 1925. The life history and relationships of *Diplodia gossypina*. *Mycologia* **17**: 191–201.

Stevens NE, 1926. Two species of *Physalospora* on *Citrus* and other hosts. *Mycologia* **18**: 206–217.

Stoddart JA, Taylor JF, 1988. Genotypic diversity: estimation and prediction in samples. *Genetics* **118**: 705–711.

Stukenbrock EH, Banke S, McDonald BA, 2006. Global migration patterns in the fungal wheat pathogen *Phaeosphaeria nodorum*. *Molecular Ecology* **15**: 2895–2904.

van der Walt FJJ. 2008. Botryosphaeriaceae Associated with *Acacia spp*. in South Africa with a Special Reference to *A. mellifera*. M.sc. Thesis. University of Pretoria, South Africa.

Yeh FC, Yang R-C, Boyle T, 1999. *POPGENE. Version 1.31. Microsoft Windows based Freeware for Population Genetic Analysis*. University of Alberta, Alberta.