Genetic variation in populations of the cacao wilt pathogen, *Ceratocystis cacaofunesta*

C. J. B. Engelbrecht[a], T. C. Harrington[a]*, A. C. Alfenas[b] and C. Suarez[c]

[a]*Department of Plant Pathology, Iowa State University, 351 Bessey Hall, Ames, IA 50011, USA;* [b]*Department de Fitopatologia, Universidade Federal de Viçosa, Viçosa, MG, Brazil; and* [c]*Estacion Experimental Tropical, Instituto Nacional Autonomuode Investigaciones Agropecuarias, Pichilingue, Ecuador*

Ceratocystis cacaofunesta (= *Ceratocystis fimbriata*) causes a lethal wilt disease of cacao (*Theobroma cacao*) in Latin America. Polymorphic microsatellite markers, (CAT)$_5$ nuclear DNA fingerprints and *Hae* III mitochondrial DNA fingerprints were used to compare genetic diversity among isolates of *C. cacaofunesta* collected from populations in western Ecuador, Costa Rica, Colombia, and Rondônia and Bahia in Brazil. Microsatellite markers and nuclear DNA fingerprints separated Ecuadorian isolates from isolates of the other four populations, and these two major groups correspond to genetic lineages already identified from ITS-rDNA sequences and intersterility groupings. Mitochondrial DNA fingerprints also demonstrated substantial diversity and split the Ecuadorian isolates into two groups. All marker types showed limited variation in the Colombian, Costa Rican and Bahian populations, as might be expected for introduced populations that have gone through recent genetic bottlenecks. In contrast, the Rondonian and western Ecuadorian populations showed gene diversity values similar to natural populations of other *Ceratocystis* species. The Rondonian population was the only sampled population in the native range of *T. cacao* (the Upper Amazon), and the putatively introduced populations were more closely related to the Rondonian population than to the western Ecuadorian population. The Ecuadorian population is in an area with other native *Theobroma* species, which may serve as natural hosts.

Keywords: *Ceratocystis fimbriata* complex, fungal population genetics, genetic diversity, *Herrania* spp., *Theobroma cacao*

Introduction

Ceratocystis cacaofunesta is a serious pathogen of cacao (*Theobroma cacao*) and related *Herrania* spp., causing wilt and death of infected trees. The cacao pathogen is a member of the Latin American clade of the *C. fimbriata* species complex, which has substantial genetic variation and a wide range of hosts (Harrington, 2000; Barnes *et al.*, 2001; Baker *et al.*, 2003; Marin *et al.*, 2003; Steimel *et al.*, 2004; Engelbrecht & Harrington, 2005). The North American clade also contains several host-specialized species, but members of this clade are primarily wound colonizers that do not kill their hosts (Johnson *et al.*, 2005). *Ceratocystis cacaofunesta* was recently differentiated from *C. fimbriata* based on minor morphological differences, pathogenicity to cacao, ITS-rDNA sequences and intersterility with other members of the Latin American clade, including isolates from *T. cacao* that proved to be non-pathogenic in inoculation studies (Baker *et al.*, 2003; Engelbrecht & Harrington, 2005). The cacao pathogen

has only been identified in Central and South America and may be indigenous to this region (Harrington, 2000; Baker *et al.*, 2003; Engelbrecht & Harrington, 2005).

Ceratocystis cacaofunesta typically enters cacao plants through fresh wounds, such as pruning or pod harvesting wounds (Malaguti, 1952), and moves through the host in the secondary xylem. Ambrosia beetles of the genus *Xyleborus* often attack the wood of infected trees (Saunders, 1965), liberating frass that includes aleurioconidia and conidia (asexual spores) as the beetles excavate their galleries (Iton & Conway, 1961). This frass may be carried by wind or rainsplash to wounds on other trees (Iton, 1960). Machete blades are another efficient means of spreading the fungus (Malaguti, 1952). Like other members of the *C. fimbriata* complex, *C. cacaofunesta* is homothallic through unidirectional mating type switching (Witthuhn *et al.*, 2000), and introduced populations could persist in an essentially clonal manner through asexual reproduction and production of sexual spores through selfing.

The genus *Theobroma* is indigenous from southern Mexico to the Amazon Basin (Purseglove, 1968), with many species on each side of the Andes (Baker, 1953). It is thought that the rising of the Andes may have facilitated speciation within the genus by separating previously

*E-mail: tcharrin@iastate.edu

widespread *Theobroma* populations (Cuatrecasas, 1964). *Theobroma cacao* is probably indigenous to the Upper Amazon region (Cheesman, 1944) but was domesticated in pre-Columbian times in Central America (Bergmann, 1969). An alternative hypothesis is that cacao is indigenous to the region from southern Mexico to Guyana-Amazon (Cuatrecasas, 1964), but most genetic studies point to the Upper Amazon as the centre of diversity of this species (Laurent *et al.*, 1993, 1994a,b; Figueira *et al.*, 1994; N'goran *et al.*, 1994; Ronning & Schnell, 1994; Warren, 1994; Whitkus *et al.*, 1998; Marita *et al.*, 2001). It is reasonable to hypothesize that a host-specialized cacao pathogen like *C. cacaofunesta* may have originated in the native range of *T. cacao* in the Upper Amazon, but an origin on other *Theobroma* species west of the Andes is also possible.

Ceratocystis wilt of cacao was first reported on cacao in western Ecuador in 1918 (Rorer, 1918). It caused extensive damage in Colombia after 1940, and in Venezuela in 1958 (Thorold, 1975). It also caused severe disease in Costa Rica in 1958, though it may have been present as early as 1926 (Thorold, 1975). The disease caused serious losses in Trinidad in the late 1950s (Spence & Moll, 1958). Reports of the disease stretch from Guatemala (Schieber & Sosa, 1960) and Central America to northern South America, including the Peruvian Amazon (Soberanis *et al.*, 1999), Ecuador, Colombia and Venezuela (Thorold, 1975). In Brazil, the disease was reported in the south-western Amazon (Rondônia) in 1978 (Bastos & Evans, 1978) and more recently in Bahia (Bezerra, 1997), which is out of the native range of *T. cacao*.

Comparing genetic diversity of populations of *C. cacaofunesta* may help to distinguish introduced from indigenous populations, as was recently done with *C. platani* (Engelbrecht *et al.*, 2004). The objective of this study was to evaluate genetic diversity in, and possible origins for, *C. cacaofunesta* in Ecuador, Costa Rica, Colombia and Brazil using polymorphic microsatellite markers and nuclear and mitochondrial DNA fingerprints.

Materials and methods

Fungal isolates

Isolates of *C. cacaofunesta* were collected from dead or dying *Theobroma cacao* trees. Two isolates were collected from *Herrania* spp., which are closely related to *Theobroma*. All isolates are listed in Table 1 and are held in the culture collection of TC Harrington, Iowa State University.

Mycelium was grown in 20 mL liquid broth (2% malt extract and 0·2% yeast extract) in 125 mL flasks at room temperature for two weeks, and the DNA was extracted and analyzed following the methods of DeScenzo & Harrington (1994).

Microsatellite markers

Sixteen PCR-based microsatellite markers (Steimel *et al.*, 2004) were used to compare genetic variability. A fluores-cently-labelled primer and an unlabelled primer flanking each of 16 three- or four-base pair repeat microsatellite regions were used in the PCR amplification reactions, as described in Steimel *et al.* (2004). PCR products were analyzed using an ABI Prism 377 DNA sequencer (Applied Biosystems Inc), with band size determination using standards, ABI GeneScan Analysis Software v3·1·2 and Genotyper 2·0 software (Applied Biosystems Inc.). Each unique product length (differing by more than 1 bp) was considered a different allele.

Nuclear DNA fingerprinting

To examine variability in nuclear DNA, total genomic DNA was restricted with *Pst*I and electrophoresed on 1% agarose gels (19·5 × 25·5 cm) (Bio-Rad Certified Molecular Biology agarose, Bio-Rad) for 1600 volt hours, with *Hind*III-restricted Lambda DNA as size standards (Promega). Gels were dried, hybridized with the oligonucleotide (CAT)₅ labelled with ³²P, and washed following the methods of DeScenzo & Harrington (1994). Labelled bands were visualized by exposing the gels to a phosphor screen for 4–6 days and then scanning the screens with a Storm PhosphorImager (Molecular Dynamics). Several representative isolates were extracted and fingerprinted two or more times to ensure consistency in results.

Quantity One quantitation software (Bio-Rad) was used to estimate the length (in bp) of each fingerprint band based on comparison with the Lambda *Hind*III size standards. The DNA restrictions and fingerprinting were repeated at least once for representative isolates. Only bands that were consistently visible and unambiguously scored in all replicate gels were analyzed.

Mitochondrial DNA fingerprinting

*Hae*III cuts GC-rich DNA and allows visualization of the AT-rich mitochondrial DNA (Wingfield *et al.*, 1996). Twenty-five micrograms of extracted total genomic DNA was restricted overnight at 37°C with *Hae*III in a total volume of 500 μL (50 μL 10X buffer, 1 mᴍ spermidine, 10 units μL⁻¹ *Hae*III and 5 μL RNAse). Restricted DNA was precipitated, dried, and washed as in Wingfield *et al.* (1996). Washed DNA was resuspended in 37·5 μL of 0·10 mᴍ Tris HCl pH 8·0, and 10 μg of this resuspended DNA was elecrophoresed on a 1% agarose gel for 1600 volt hours, with *Hind*III-restricted Lambda DNA (Promega) as size standard. Separated bands of DNA were visualized using ethidium bromide and UV light. Several representative isolates were extracted and fingerprinted two or more times for comparison. Analyses were conducted as described above.

Analysis

Different fragment sizes of microsatellite markers were considered different alleles. Each fingerprint band was considered a locus with two possible alleles, present or absent. The dataset was corrected for clonal reproduction

Table 1 Origin and genotypes of isolates of *Ceratocystis cacaofunesta* used for genetic analyses

Population	Location collected	Isolate numbers	Genotype
Ecuador	Pichilingue, Ecuador	C1691	I
		C1690	II
		C1833 (= CBS 115169), C1834	III
		C1835	IV
		C1695 (= CBS 115163)	V
		C1751, C1756	VI
	Ecuador (unknown location)	C1004 (= CBS 153·62)	VII
Rondônia, Brazil	Rondônia, Brazil (site 1)	C1983 (= CBS 115172)	VIIIa
		C1984	VIIIb
	Rondônia, Brazil (site 2)	C2031	IX
Costa Rica	La Lola, Matine, Costa Rica	C1548 (= CBS 114722)	X
	La Lola, Matine, Costa Rica	C1547	XI
	Bri Bri, Costa Rica	C1639	XII
	Cahuita, Costa Rica	C1640	XIII
	Turrialba, Costa Rica	C1778[a], C1779	XIV
		C1776	XV
		C1549, C1775	XVI
		C1550, C1634, C1635, C1636, C1638, C1642[a], C1777	XVII
		C1637	XVIII
	Costa Rica (unknown location)	C940 (= CBS 152·62)	XIX
Colombia	Palestina, Caldas, Colombia	C1935	XX
		C1947	XXI
		C1937	XXII
		C1936	XXIII
Bahia, Brazil	Novo Oriente, Bahia, Brazil	C1595	XXIV
		C1593	XXV
		C1594, C1596	XXVI
	Uruçuca, Bahia, Brazil, site 1	C1719, C1720, C1723	XXVI
		C1718	XXVII
		C1722	XXVIII
		C1721	XXIX
	Uruçuca, Bahia, Brazil, site 2	C1728, C1729	XXVI
	Uruçuca, Bahia, Brazil, site 3	C1733–C1735, C1737	XXVI
		C1736	XXX
	Camacã, Bahia, Brazil, site 1	C1585, C1586	XXVI
		C1587	XXXI
	Camacã, Bahia, Brazil, site 2	C1725–C1727	XXVI
		C1724	XXXII
	Camacã, Bahia, Brazil, site 3	C1730–C1732	XXVI
	Itabuna, Bahia, Brazil, site 1	C1599, C1601	XXVI
	Itabuna, Bahia, Brazil, site 2	C1597, C1600	XXXIII
	Itabuna, Bahia, Brazil, site 3	C1598	XXXIV
	Ubaitaba, Bahia, Brazil	C1738–C1749[b]	XXVI

[a]C1642 and C1778 were collected from *Herrania* spp.
[b]C1749 was collected from soil under a diseased cacao tree.

by removing isolates from each collection site that were genetically identical except where noted. In analyses concerning the Amazonian population (all populations except C940 and the Ecuadorian isolates), the dataset was clone-corrected by removing duplicate isolates that were genetically identical within a population rather than within a site.

Genotypic diversity within the five populations was calculated as the number of genotypes found using each marker type individually and using all three marker types together. Nei's gene diversity for each marker type and for all marker types together for each population was calculated with clone-corrected data using PopGen 1·32 software (Yeh & Boyle, 1997).

Weir and Cockerham's theta (θ), which is comparable to Wright's F_{ST} (Weir & Cockerham, 1984), was calculated among pairs of populations using the program Multilocus (Mac version 1·21, Department of Biology, Imperial College at Silwood Park, UK). Gene flow between pairs of populations (N_m) was estimated based on the average coefficient of gene differentiation (G_{ST}) across all loci using PopGen 1·32. For both of these measures, values

Table 2 Microsatellite alleles and nuclear and mitochondrial DNA fingerprint bands for the 35 genotypes identified in populations of *Ceratocystis cacaofunesta*

Genotype	Number of isolates	Microsatellite alleles[a]	Nuclear DNA bands[b]	Mitochondrial DNA bands[b]
I	1	ACDBDABAAAAADDAA	0000010100110101101001100010110	1101010101110001001001001010010
II	1	ABABDABAEAAADDAA	0000010100101000101001100010110	0010010100100001010010010
III	2	ABGBDABAAAAADDAA	0000010100110000101001100110110	0010010100000011010010011
IV	1	ABABDABAAAAADDAA	0000010100110000101001100110110	0010010101000010010010010
V	1	ACDBDABAAAAADDAA	0100010100100000101001100010110	0010010001000100000010010
VI	2	ABABDABAEAAADDAA	0100010100100000101001100010110	0010010100100001010010010
VII	1	ABABDABAAAAADCAA	0000010100110000000001100111110	1101010111000100010010010
VIIIa	1	BABACECBBBBBAABA	0000101000001000001010010001100	0010110110101000000101000
VIIIb	1	BABACFCBBBBBADBA	0000101000001000001010010001100	0010010110101000000101000
IX	1	CAJBCECBBBBBAACB	1001001000001010001000001001101	1011010111010011001010010
X	1	BAEACCCBCBCBCAAA	1010011001111010000100100101110	0111110100011010010011100
XI	1	BAEACCCBCBCBCAAA	1010011000011010000100100101010	0011110100011011000001000
XII	1	BAFACCCBCBCBCAAA	1010011000011010000100100101110	0110010100011100100000000
XIII	1	BAIACCCBCBCBCAAA	1010011000011010000100100101110	0011110100011011000001000
XIV	2	BAEACCCBCBCBCAAA	1010011000011010000100100101110	0110010100011010100001100
XV	1	BAEACCCBCBCBCAAA	1010011000011010000100100101110	0110110100011010010011100
XVI	2	BAFACCCBCBCBCAAA	1010011000011010000100100101110	0110110100011010010011100
XVII	7	BAEACCCBCBCBCAAA	1010011000011010000100100101110	0111110100011010010011100
XVIII	1	BAGACCCBCBCBCAAA	1010011000011010000100100101110	0111110100011010010011100
XIX	1	BBCBDABAFAAADCAA	0000010100101000000001000110110	1101010101000010010010011
XX	1	BAFACCCBCBCBCAAA	1010011000011010000100100101110	0110010100011011000001000
XXI	1	BAGACCCBCBDBCAAA	1010011000011010000100100101110	0110010100011011000001000
XXII	1	BAFACCABCBCBCAAA	1010011000011010000100100101110	0110010100011011001001000
XXIII	1	BAHACCCBCBDBCAAA	1010011000011010000100100101110	0110010100011011001001000
XXIV	1	BACACCCBCBBCBABA	1001001001001000011000010101101	0110110100110010010000001
XXV	1	BACACCCBCBBCBABA	1001001001001000011000010101101	0110101100110011010010000
XXVI	33	BACACCCBCBBCBABA	1001001001001000011000010101101	0110110100110011010010000
XXVII	1	BACACCCBCBBCBABA	1001001001001000011000010101101	0110110100110111010010000
XXVIII	1	BACACCCBCBBCBABA	1001001001001000011000010101101	0110010100111011011010100
XXIX	1	BACABCCBCBDBBCBABA	1001001001001000011000010101101	0110010100111011001010010
XXX	1	BACACCCBCBBCBABA	1001001000001000011000010101101	0110110100110011010010000
XXXI	1	BACACCCBCBBCBABA	1001001001001000011010010101101	0110110100011010011000010
XXXII	1	BACAACDBCBBCBABA	1000001001001000011000010101101	0110010100111011001000010
XXXIII	2	BACACCCBCBBCBABA	1001001001001000011000010101101	0110010100111011010010000
XXXIV	1	BACACBCBCBBCBABA	1001001001001000011000010101101	0110010100111011010010000

[a]Microsatellite alleles (different letters) for 16 microsatellite loci AAG8, AAG9, CAA9, CAA10, CAA15, CAA38, CAA80, CAG5, CAG15, CAG900, CAT1, CAT3k, CAT9X, CAT1200, GACA650, GACA60, in increasing size of amplified product.
[b]Polymorphic bands are listed smallest to largest, present = 1, absent = 0.

obtained from microsatellite markers and nuclear DNA fingerprints were nearly identical and were averaged to obtain an overall nuclear DNA value.

To test for random mating within the Ecuadorian, Bahian and Amazonian (all populations except Ecuador combined) populations, linkage disequilibrium was analyzed using the index of association (I_A) statistic in Multilocus. Randomization procedures were used (1000 replications, without replacement) to test the significance of the observed I_A value.

Relationships among isolates were also examined using genetic distance matrices and UPGMA (unweighted pair group method with arithmetic mean) trees generated using PAUP* (Swofford, 1998). Nei's genetic distance between populations and UPGMA dendrograms comparing populations were calculated using PopGen 1·32. Bootstrap values for the population trees were calculated from 100

bootstrap replicates of 100 UPGMA trees using SEQ-BOOT in PHYLIP version 3·6 (Felsenstein 1989, 1993).

Results

All 16 microsatellite markers were polymorphic for the isolates tested. Each locus had between two and 10 alleles (Table 2), and together they resolved 20 microsatellite genotypes among the 77 isolates tested from the five populations (Table 3). No microsatellite genotype was found in more than one population.

Nuclear (CAT)$_5$ fingerprinting resolved 34 consistently scorable bands (Fig. 1a, Table 2). Thirty-one of these bands were polymorphic, and bands of sizes of 6·5, 7·3 and 11·2 kb were present in all isolates. Other bands that could not be consistently scored between gel runs of the same isolates were excluded from the analysis. Fourteen different nuclear

Table 3 Number of genotypes identified in populations of *Ceratocystis cacaofunesta*

Population	No. of isolates	No. of collection sites	No. of genotypes				Isolates after clone correction[a]
			Microsatellites	(CAT)₅	*Hae*III	All markers	
Ecuador	9	1	5	5	6	7	7
Rondônia	3	2	3	2	3	3	3
Colombia	4	1	4	1	2	4	4
Costa Rica	17	4	4	3	5	9	9
Bahia	44	9	4	4	8	11	19
all non- Ecuadorian populations	68	16	15	9	18	27	35
all populations	77	17	20	14	24	34	42

[a]Data were clone-corrected by removing duplicate isolates within a site that had the identical genotype.
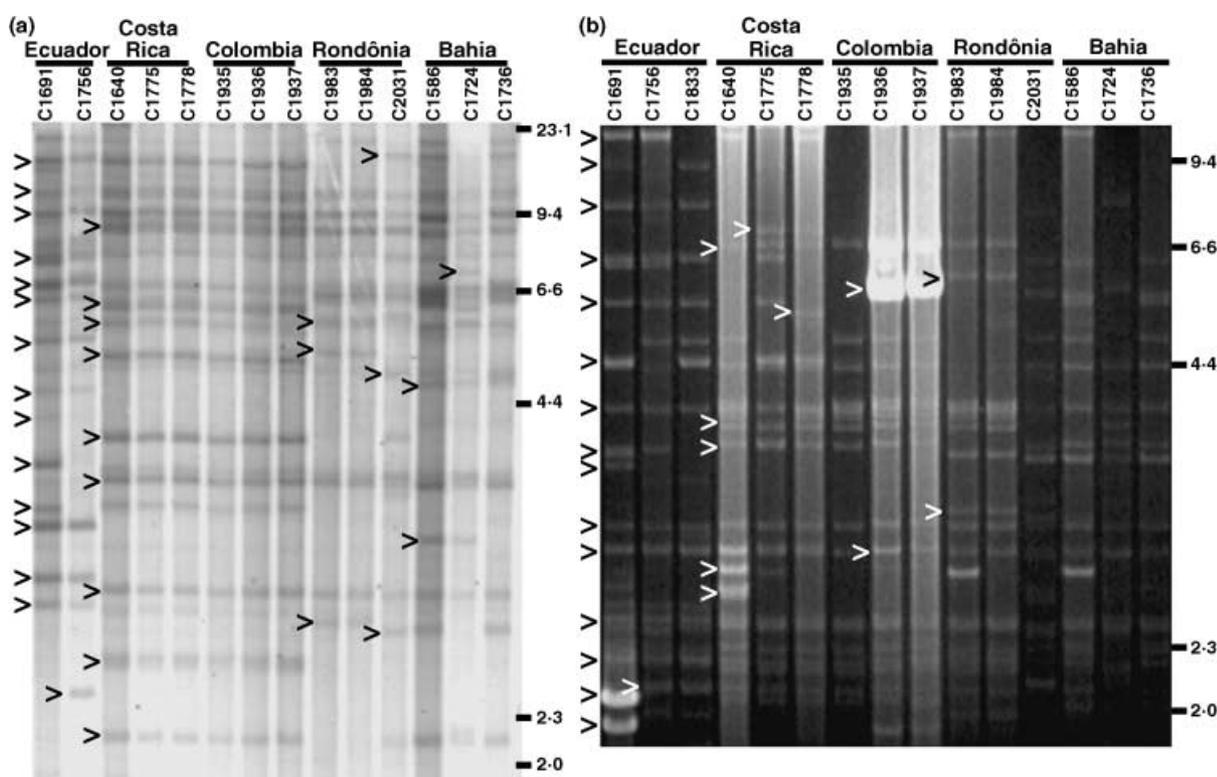
Figure 1 (a) Representative (CAT)₅ nuclear DNA fingerprints and (b) *Hae*III mitochondrial DNA fingerprints of *Ceratocystis cacaofunesta* isolates from populations in Ecuador, Costa Rica, Colombia and Brazil (Rondônia and Bahia). Arrowheads indicate polymorphic bands. Sizes (kb) of standards are given on the right.

fingerprint patterns were found. Isolates from Colombia had an identical (CAT)₅ fingerprint (Table 3), which was the same as that of 15 of the 17 isolates from Costa Rica.

*Hae*III mitochondrial DNA fingerprints yielded 30 consistently scorable bands (Fig. 1b). Bands of sizes of 2·3, 2·4, 2·5 and 4·0 kb were monomorphic, and the remaining bands were polymorphic among the isolates tested (Table 2). A total of 24 mitochondrial fingerprint patterns were found, and none were shared between populations. Bands were not as clearly separated as with the (CAT)₅ fingerprinting, so it is possible that some co-migrating bands that were scored as identical in different

populations were actually of different sizes and were not homologous.

All three combined marker types resolved 34 total genotypes among the 77 isolates tested (Table 2). None of the multilocus genotypes were shared among populations. However, only seven unique genotypes were found among the nine isolates from Ecuador, nine unique genotypes were found among the 17 isolates tested from Costa Rica, and 19 unique genotypes were found among the 44 isolates from Bahia (Table 3). Further analyses of populations used clone-corrected data, i.e. only unique genotypes from a particular population were used.

Table 4 Nei's gene diversity of clone-corrected populations of *Ceratocystis cacaofunesta* and other *Ceratocystis* species

Species	Population	Number of isolates after clone correction	Nei's gene diversity (H)			
			Microsatellites	$(CAT)_5$	*Hae*III	All markers
C. cacaofunesta	Ecuador	7	0·1020	0·0840	0·1548	0·1137
	Rondônia	3	0·1979	0·1498	0·2146	0·1639
	Colombia	4	0·0938	0·0000	0·0172	0·0253
	Costa Rica	9	0·0386	0·0174	0·1345	0·0647
	Bahia	19	0·0308	0·0088	0·0917	0·0437
	all non-Ecuadorian populations	35	0·2440	0·2148	0·2146	0·2207
	all	42	0·4129	0·2787	0·2239	0·2928
C. platani[a]	Eastern USA	20	0·2178	0·1683	0·2006	0.1911
	California, USA	4	0·0625	0·0000	0·0000	0.0137
	Europe	24	0·0191	0·0000	0·0000	0.0042
C. albofundus[b]	South Africa	31		0·2137	0·2490	
C. virescens[c]	North America	16		0·0935	0·0928	

[a]Data from Engelbrecht *et al.*, 2004.
[b]Data from Roux *et al.*, 2001.
[c]Data from Harrington *et al.*, 1998.

Table 5 Estimate of differentiation (theta, θ), above diagonal, and estimate of gene flow (N_m), below diagonal, between populations of *Ceratocystis cacaofunesta* using nuclear DNA fingerprints and microsatellites markers combined

	Ecuador	Colombia	Costa Rica	Rondônia	Bahia
Ecuador		0·8532	0·8916	0·7830	0·9405
Colombia	0·1200		0·0702	0·7098	0·9182
Costa Rica	0·1134	4·9581		0·8236	0·9289
Rondônia	0·2210	0·2996	0·2666		0·8530
Bahia	0·0852	0·1165	0·0766	0·3585	

Based on Nei's gene diversity from clone-corrected populations (Table 4), the populations from Ecuador and Rondônia had the highest gene diversity values, while those from Costa Rica, Colombia and Bahia had lower diversity values. The non-Ecuadorian population (all isolates except those from Ecuador and isolate C940) had a high gene diversity value. Table 4 also lists the gene diversity values of other homothallic *Ceratocystis* species based on comparable markers.

Weir and Cockerham's theta (θ) was used to estimate the degree of differentiation between populations, using combined nuclear DNA markers (Table 5) from the clone-corrected dataset. Theta varies from zero (no differentiation among populations) to one (populations completely differentiated). Based on nuclear DNA markers, the Costa Rican and Colombian populations were the most similar (θ = 0·07), while the Bahian and Ecuadorian populations were the most differentiated (θ = 0·94). Other than the comparison between Costa Rican and Colombian populations, all θ values were above 0·7, indicating differentiation among the populations. The estimate of gene flow (N_m) was also used to make pairwise comparisons of populations with the two nuclear markers together using the clone-corrected dataset (Table 5). The value of N_m ranges from zero (no gene flow) to infinity (complete gene flow between populations). Analysis of nuclear DNA markers revealed that the Costa Rican and Colombian populations

had the highest amount of probable gene flow ($N_m = 4·96$), while the Costa Rican and Bahian population comparison had the least ($N_m = 0·07$). Most other values of N_m were also relatively low, not exceeding 0·34.

The UPGMA trees constructed using microsatellite and $(CAT)_5$ fingerprint data separated the isolates into two distinct groups, one containing all Ecuadorian isolates and the other containing isolates from Costa Rica, Brazil and Colombia (Fig. 2a,b). In both cases, the branches separating the Ecuadorian and non-Ecuadorian isolates had strong bootstrap support. With both nuclear marker types, there was moderate support of branches grouping the Bahian isolates, and the Colombian and Costa Rican isolates also grouped together. The three Rondonian isolates were divergent, but the nuclear DNA fingerprints of the Rondonian isolates grouped them with the Bahian isolates. A UPGMA tree constructed from *Hae*III mitochondrial DNA fingerprints showed similar trends, with the Bahian isolates grouping together and the Costa Rican and Colombian isolates grouping together, but the Ecuadorian isolates splitting into two groups (Fig. 2c). However, bootstrap support for the branches was generally low.

The UPGMA trees constructed using populations instead of isolates showed trends similar to those found when analyzing by isolate, but there was little bootstrap support for branches connecting populations (Fig. 3). Nuclear and mitochondrial markers showed slightly different
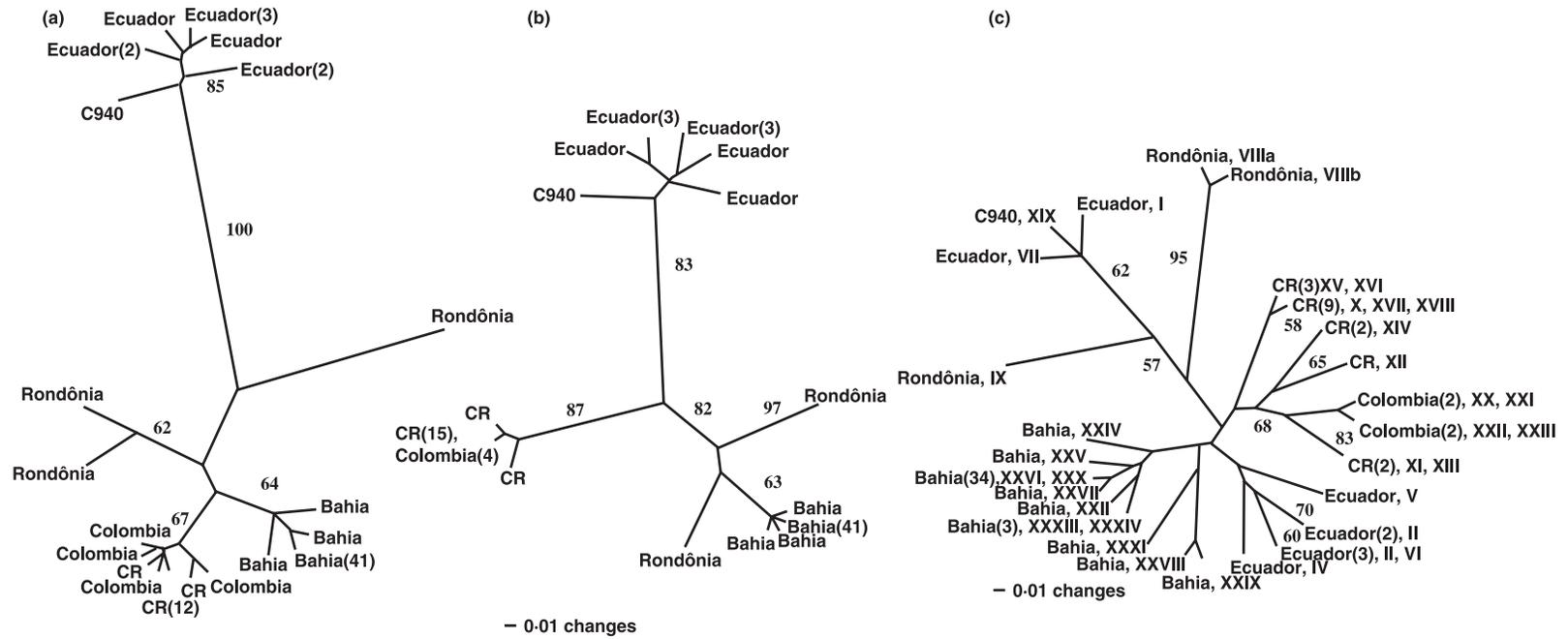
Figure 2 Dendrograms generated by UPGMA (unweighted pair group method, arithmetic mean) based on Nei's genetic distance among *Ceratocystis cacaofunesta* genotypes from populations in Ecuador, Costa Rica, Colombia and Brazil (Rondônia and Bahia), generated from (a) microsatellite data, (b) (CAT)$_5$ nuclear DNA fingerprint data and (c) *Hae*III mitochondrial DNA fingerprint data. Bootstrap values greater than 50 are shown. Roman numerals in (c) correspond to genotypes in Table 2.
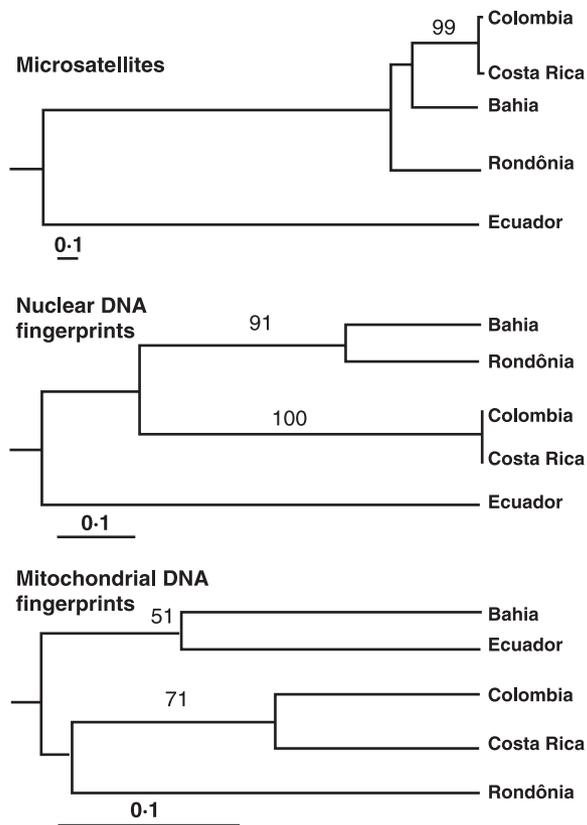
Figure 3 Dendrograms generated by UPGMA (unweighted pair group method, arithmetic mean) based on Nei's genetic distance among five populations of *Ceratocystis cacaofunesta* using microsatellites, nuclear fingerprint and mitochondrial fingerprint markers. Scale bars show genetic distance. Bootstrap values above 50 are shown above the branches.

trends. In each analysis, isolate C940 from Costa Rica was excluded because this isolate is genetically similar to the Ecuadorian isolates in phylogenetic analyses (Baker *et al*., 2003) and intersterility studies (Engelbrecht & Harrington, 2005). The other Costa Rican isolates and the Colombian population grouped together strongly based on the microsatellite and nuclear DNA fingerprint data, and the Bahian and Rondonian populations were strongly linked based on nuclear DNA fingerprints. There was moderate support for the branch connecting the Costa Rican and Colombian populations in the mitochondrial DNA analysis.

The index of association (I_A) was used to test for random mating in the Ecuadorian, Bahian and Amazonian (non-Ecuadorian) clone-corrected populations. The value of I_A is expected to be zero in a randomly mating population, and 1000 randomized replicates were performed to assess whether the value obtained was significantly different from zero. Nuclear markers in the Ecuadorian population showed index of association values consistent with an outcrossing population ($I_A = -0.08333$, $P = 0.595$ for microsatellites, and $I_A = 0.3741$, $P = 0.168$ for nuclear DNA fingerprints). Analysis of the group of populations other than the Ecuadorian population showed evidence of

clonal reproduction (microsatellite $I_A = 2.1337$, $P < 0.001$; nuclear DNA fingerprint $I_A = 7.6492$, $P < 0.001$). The Bahian population showed evidence of clonal reproduction using microsatellites ($I_A = 0.6372$, $P = 0.04$). Analysis of nuclear DNA fingerprints of the Bahian population showed no evidence of clonal reproduction ($I_A = -0.1111$, $P = 1.000$), but this lack of resolution may be due to the small number of polymorphic loci (three) in the Bahian population.

Discussion

Despite the limitations of small sample sizes due to the rarity of the disease in some locations, the gene diversity values for the Rondonian and Ecuadorian populations of *C. cacaofunesta* are similar to those of indigenous populations of other homothallic *Ceratocystis* species analyzed using the same markers. Although the data may be considered preliminary, the diversity values found in Rondônia and Ecuador are similar to those found in a native population of *C. platani* from the eastern USA (Engelbrecht *et al*., 2004). They are also comparable to values obtained from a native population of *C. albofundus* in South Africa (Roux *et al*., 2001) and those of a native population of *C. virescens* in eastern North America (Harrington *et al*., 1998). All of these *Cerato-cystis* species can produce perithecia and ascospores through selfing via unidirectional mating type switching (Harrington & McNew, 1997). Thus, even natural populations would be expected to have relatively low genetic diversity (Harrington *et al*., 1998), and introduced populations are expected to be essentially clonal and have very limited genetic diversity (Engelbrecht & Harrington, 2004).

Gene diversity values of two introduced populations of *C. platani* in Modesto, California and southern Europe (Engelbrecht & Harrington, 2004) were similar to those found in the *C. cacaofunesta* populations in Costa Rica, Colombia and Bahia, Brazil, suggesting that these three populations on cacao are the result of introductions. The Costa Rican and Colombian isolates share nuclear DNA fingerprints and group closely together in UPGMA analyses using all the markers, and they may have been derived from one interbreeding population based on analyses of nuclear markers. Except for the Costa Rica and Colombia populations, the samples of populations of *C. cacaofunesta* are highly differentiated from each other as shown by the theta values based on nuclear markers. This further supports the conclusion that the Costa Rican and Colombian isolates are from the same original population, perhaps on cuttings of cacao from the Upper Amazon, and it is possible that the Colombian population of *C. cacaofunesta* was introduced from Costa Rica.

The Bahian, Costa Rican and Colombian populations are likely to be introduced as all three locations are beyond the native range of cacao. Infested cuttings of *T. cacao* appear to be the means of introduction of *C. cacaofunesta* (Harrington, 2000). The disease has only been known in Bahia since 1997 (Bezerra, 1997), when the fungus was recognized in grafted cuttings in an experiment station.

The source of those cuttings is not known, but it is possible that they originated from the cacao experiment station in Rondônia where two isolates of *C. cacaofunesta* were collected for this study. The low level of gene flow between Bahia and the Costa Rican/Colombian populations suggests that the Bahian population was the result of a separate, independent introduction of a limited number of genotypes.

Although only three isolates were obtained from Rondônia, they were genetically unique, and the Rondonian population had gene diversity values similar to native populations of selfing *Ceratocystis* species. These *C. cacaofunesta* isolates were collected from within the native range of cacao and near its centre of genetic diversity (Laurent *et al.*, 1993, 1994a,b; Figueira *et al.*, 1994; N'goran *et al.*, 1994; Ronning & Schnell, 1994; Warren, 1994; Whitkus *et al.*, 1998; Marita *et al.*, 2001). It is possible that these isolates represent an indigenous population of *C. cacaofunesta* on *T. cacao*. The populations in Costa Rica, Colombia and Bahia were more similar to the population in Rondônia than to the population in Ecuador, so it is likely that Rondônia or elsewhere in the Upper Amazon is the source of the introduced populations. Unfortunately, only three isolates of *C. cacaofunesta* from the Upper Amazon were collected, two closely related isolates from the experimental plantation in Ouro Preto do Oeste, and another from a small farm less that 50 km west, near the Bolivian border. Apparently, the most susceptible cultivars of *T. cacao* planted in this region have died and the disease is rare in the surviving cultivars, making it very difficult to obtain isolates there. The initial plantings in the 1970s of 'Theobahia' genotypes in Ouro Preto do Oeste died quickly (Luiz Carlos Almeida, personal communication). Theobahia genotypes proved to be highly susceptible to *Ceratocystis* wilt in Bahia after introduction of *C. cacaofunesta* in 1997.

Like the Rondonian population, the Ecuadorian isolates had gene diversity values similar to indigenous populations of homothallic *Ceratocystis* species, even though they were collected outside *T. cacao*'s native range. It is possible that they represent an indigenous population that is pathogenic to another member of the genus *Theobroma*, or to a related genus, native to Ecuador west of the Andes (Baker, 1953; Purseglove, 1968). All the Ecuadorian isolates were collected from plantings of cacao in an experimental station in western Ecuador, where seven genotypes were identified among the nine isolates, suggesting a relatively high level of genetic diversity there. Attempts to locate the disease in commercial plantings in Ecuador failed, perhaps because commercial cultivars in the region are resistant to the disease. Cacao wilt was first described from this cacao-growing region (Rorer, 1918) and was initially devastating, so growers are likely to use resistant cacao cultivars. Isolate C940, obtained from a culture collection, was reportedly collected in Costa Rica, but it groups genetically with the Ecuadorian population (Baker *et al.*, 2003; Engelbrecht & Harrington, 2005), and it may represent an introduction from Ecuador.

In addition to the nuclear and mitochondrial markers used here, the western Ecuadorian and Upper Amazon genotypes are distinguished by ITS-rDNA sequences (Baker *et al.*, 2003), sequence analyses of mating type genes (unpublished), and intersterility (Engelbrecht & Harrington, 2005). These two sublineages might be considered different species, but they overlap in morphological characters (Engelbrecht & Harrington, 2005) and do not differ in pathogenicity to cacao seedlings (Baker *et al.*, 2003). Using the phylogenetic species concept of Harrington & Rizzo (1999), these sublineages are not recognized as distinct species because they cannot be distinguished by diagnostic phenotypic characters. The sublineages may have arisen from diversification of an original population that was divided by the Andean geographical barrier.

Genetic markers such as DNA fingerprints and microsatellite markers can help to distinguish indigenous populations from introduced ones, as has been done with *C. platani* populations in the eastern USA and Europe (Engelbrecht *et al.*, 2004). The data here support the hypotheses that the Costa Rican, Colombian and Bahian populations of *C. cacaofunesta* resulted from introductions, probably on cuttings, while the Rondonian and Ecuadorian populations may be indigenous. The Rondonian population is within the natural range of *T. cacao*, and the Upper Amazon was the source of *T. cacao* germplasm used elsewhere in Latin America and the Caribbean (Baker, 1953). The experiment station in Rondônia was not in existence when ceratocystis wilt first appeared in Costa Rica and Colombia, so the source of *C. cacaofunesta* in those countries was likely to be elsewhere in the Upper Amazon. However, the Bahian population of *C. cacaofunesta* is genetically closely related to the Rondonian population, and the disease in Bahia appeared in an experiment station (Bezerra, 1997) operated by the same federal agency that operates the Rondonian experiment station where the isolates were collected (Bastos & Evans, 1978). Great care should be taken in moving cuttings from the Upper Amazon or other locations in Latin America where *C. cacaofunesta* is present. The disease has caused dramatic losses in Bahia, and similar losses could develop if it was moved to cacao-growing regions of western Africa and Southeast Asia.

Acknowledgements

References

Baker CJ, Harrington TC, Krauss U, Alfenas AC, 2003. Genetic variability and host specialization in the Latin American clade of *Ceratocystis fimbriata*. *Phytopathology* **93**, 1274–84.

Baker RED, 1953. Anglo-American cacao collecting expedition. *Archives of Cocoa Research* **1**, 127–54.

Barnes I, Gaur A, Burgess T, Roux J, Wingfield BD, Wingfield MJ, 2001. Microsatellite markers reflect intra-specific relationships between isolates of the vascular wilt pathogen *Ceratocystis fimbriata*. *Molecular Plant Pathology* **2**, 319–25.

Bastos CN, Evans HC, 1978. Ocorrência de *Ceratocystis fimbriata* Ell & Halst. na Amazônia Brasileira. *Acta Amazonica* **8**, 543–4.

Bergmann JF, 1969. The distribution of cacao cultivation in pre-Columbian America. *Annals of the Association of American Geographers* **59**, 85–96.

Bezerra JL, 1997. *Ceratocystis fimbriata* causing death of budded cocoa seedlings in Bahia, Brazil. *Incoped Newsletter* **1**, 6.

Cheesman EE, 1944. Notes on nomenclature, classification and possible relationships of cacao populations. *Tropical Agriculture* **21**, 144–59.

Cuatrecasas J, 1964. Cacao and its allies. A taxonomic revision on the genus *Theobroma*. *Contributions from the National Herbarium (USA)* **35**, 379–614.

DeScenzo RA, Harrington TC, 1994. Use of (CAT)$_5$ as a DNA fingerprinting probe for fungi. *Phytopathology* **84**, 534–40.

Engelbrecht CJB, Harrington TC, 2005. Intersterility, morphology, and taxonomy of *Ceratocystis fimbriata* on sweet potato, cacao, and sycamore. *Mycologia* **97**, 57–69.

Engelbrecht CJB, Harrington TC, Steimel J, Capretti P, 2004. Genetic variation in eastern North American and putatively introduced populations of *Ceratocystis fimbriata* f. *platani*. *Molecular Ecology* **13**, 2995–3005.

Felsenstein J, 1989. PHYLIP – phylogeny inference package (Version 3.2). *Cladistics* **5**, 164–6.

Felsenstein J, 1993. PHYLIP (phylogeny inference package) version 3.5c. Distributed by the author. Department of Genetics, University of Washington, Seattle, WA, USA.

Figueira A, Janick J, Levy M, Goldsbrough P, 1994. Reexamining the classification of *Theobroma cacao* L. using molecular markers. *Journal of the American Society of Horticultural Science* **119**, 1073–82.

Harrington TC, 2000. Host specialization and speciation in the American wilt pathogen *Ceratocystis fimbriata*. *Fitopatologia Brasileira* **25**, 262–3.

Harrington TC, McNew DL, 1997. Self-fertility and uni-directional mating-type switching in *Ceratocystis coerulescens*, a filamentous ascomycete. *Current Genetics* **32**, 52–9.

Harrington TC, Rizzo DM, 1999. Defining species in the fungi. In: Worral JJ, ed. *Structure and Dynamics of Fungal Populations*. Dordrecht, the Netherlands: Kluwer Academic Press, 43–71.

Harrington TC, Steimel J, Kile G, 1998. Genetic variation in three *Ceratocystis* species with outcrossing, selfing and asexual reproductive strategies. *European Journal of Forest Pathology* **28**, 217–26.

Iton EF, 1960. Studies on a wilt disease of cacao at River Estate. II. Some aspects of wind transmission. In: *Annual Report on Cacao Research, 1959–1960*. St Augustine, Trinidad: Imperial College of Tropical Agriculture, University of the West Indies, 47–58.

Iton EF, Conway GR, 1961. Studies on a wilt disease of cacao at River Estate III. Some aspects of the biology and habits of *Xyleborus* spp. and their relation to disease transmission. In: *Annual Report on Cacao Research, 1959–1960*. St. Augustine, Trinidad: Imperial College of Tropical Agriculture, University of the West Indies, 59–65.

Johnson JA, Harrington TC, Engelbrecht CJB, 2005. Phylogeny and taxonomy of the North American clade of the *Ceratocystis fimbriata* complex. *Mycologia* **97**, 1067–92.

Laurent V, Risterucci AM, Lanaud C, 1993. Variability for nuclear ribosomal genes within *Theobroma cacao*. *Heredity* **71**, 96–103.

Laurent V, Risterucci AM, Lanaud C, 1994a. Genetic diversity in cocoa revealed by cDNA probes. *Theoretical and Applied Genetics* **88**, 193–8.

Laurent V, Risterucci AM, Lanaud C, 1994b. RFLP study of genetic diversity of *Theobroma cacao*. *Angew Botany* **68**, 36–9.

Malaguti G, 1952. *Ceratostomella fimbriata* en el cacao de Venezuela. *Acta Científica Venezolana* **3**, 94–7.

Marin M, Castro B, Gaitan A, Preisig O, Wingfield BD, Wingfield MJ, 2003. Relationships of *Ceratocystis fimbriata* isolates from Colombian coffee-growing regions based on molecular data and pathogenicity. *Journal of Phytopathology* **151**, 395–405.

Marita JM, Nienhuis J, Pirese JL, Aitken WM, 2001. Analysis of genetic diversity in *Theobroma cacao* with emphasis on witches' broom disease resistance. *Crop Science* **41**, 1305–16.

N'goran JAK, Laurent V, Risterucci AM, Lanaud C, 1994. Comparative genetic analysis of *Theobroma cacao* L. using RFLP and RAPD markers. *Heredity* **73**, 589–97.

Purseglove JW, 1968. *Tropical Crops. Dicotyledons II*. New York, USA: Wiley.

Ronning CM, Schnell RJ, 1994. Allozyme diversity in a germplasm collection of *Theobroma cacao* L. *Journal of Heredity* **85**, 291–5.

Rorer JB, 1918. *Enfermedades y plagas del cacao en el Ecuador y metodos modernos apropiados al cultivo del cacao*. Guayaquil, Ecuador: Associacion de Agricultores.

Roux J, Harrington TC, Steimel JP, Wingfield MJ, 2001. Genetic variation in the wattle wilt pathogen *Ceratocystis albofundus*. *Mycoscience* **42**, 327–32.

Saunders JL, 1965. The *Xyleborus-Ceratocystis* complex of cacao. *Cacao* **10**, 7–13.

Schieber E, Sosa ON, 1960. Cacao canker in Guatemala incited by *Ceratocystis fimbriata*. *Plant Disease Reporter* **44**, 672.

Soberanis W, Rios R, Arevalo E, Zuniga L, Cabezas O, Krauss U, 1999. Increased frequency of phytosanitary pod removal in cacao (*Theobroma cacao*) increases yield economically in eastern Peru. *Crop Protection* **18**, 677–85.

Spence JA, Moll ER, 1958. Preliminary observations on a wilt condition of cocoa. *Journal of the Agricultural Society of Trinidad* **58**, 349–59.

Steimel J, Engelbrecht CJB, Harrington TC, 2004. Development and characterization of microsatellite markers for the fungus *Ceratocystis fimbriata*. *Molecular Ecology Notes* **4**, 215–8.

Swofford DL, 1998. *PAUP*. Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4*. Sunderland, Massachusetts, USA: Sinauer Associates.

Thorold CA, 1975. *Diseases of Cocoa*. Oxford, UK: Oxford University Press.

Warren JM, 1994. Isozyme variation in a number of populations of *Theobroma cacao* L. obtained through various sampling regimes. *Euphytica* **72**, 121–6.

Weir BS, Cockerham CC, 1984. Estimating F-statistics for the analysis of population structure. *Evolution* **38**, 1358–70.

Whitkus R, de la Cruz M, Mota-Bravo L, Gómez-Pompa A, 1998. Genetic diversity and relationships of cacao (*Theobroma cacao* L.) in southern Mexico. *Theoretical and Applied Genetics* **96**, 621–7.

Wingfield BD, Harrington TC, Steimel J, 1996. A simple method for detection of mitochondrial DNA polymorphisms. *Fungal Genetics Newsletter* **43**, 56–60.

Witthuhn RC, Harrington TC, Wingfield BD, Steimel JP, Wingfield MJ, 2000. Deletion of the *MAT-2* mating type gene during uni-directional mating-type switching in *Ceratocystis*. *Current Genetics* **38**, 48–52.

Yeh FC, Boyle TJ, 1997. Population genetic analysis of co-dominant and dominant markers and quantitative traits. *Belgian Journal of Botany* **129**, 157.