Mycologia, 106(2), 2014, pp. 224–242. DOI: 10.3852/106.2.224

Intraspecific and intragenomic variability of ITS rDNA sequences reveals taxonomic problems in *Ceratocystis fimbriata* sensu stricto

T.C. Harrington[1]
Department of Plant Pathology and Microbiology, Iowa State University, Ames, Iowa 50011

M.R. Kazmi
Australian Centre for International Agricultural Research, Australian High Commission, Diplomatic Enclave No. 1, Islamabad, Pakistan

A.M. Al-Sadi
Department of Crop Sciences, Sultan Qaboos University, P.O. Box 34, AlKhoud 123, Oman

S.I. Ismail
Department of Plant Pathology and Microbiology, Iowa State University, Ames, Iowa 50011

Abstract: Fourteen new species in the Latin American Clade (LAC) of the *Ceratocystis fimbriata* complex recently were distinguished from *C. fimbriata* sensu stricto largely based on variation in ITS rDNA sequences. Among the 116 isolates representing the LAC, there were 41 ITS haplotypes. Maximum parsimony (MP) analysis of ITS sequences produced poorly resolved trees. In contrast, analyses of mating-type genes (*MAT1-1-2* and *MAT1-2-1*) resolved a single MP tree with branches of high bootstrap and posterior probability support. Four isolates showed intragenomic variation in ITS sequences. Cloning and sequencing of PCR products from a single haploid strain identified two or more ITS sequences differing at up to 16 base positions and representing two described species. Isolates from introduced populations that appeared to be clonal based on microsatellite markers varied at up to 14 bp in ITS sequence. Strains of seven Brazilian ITS haplotypes and an isolate from *Ipomoea batatas* (on which the species name *C. fimbriata* was based) were fully interfertile in sexual crosses. These analyses support three phylogenetic species that differ in pathogenicity: *C. platani*, *C. cacaofunesta* and *C. colombiana*. Five ITS species (*C. manginecans*, *C. mangicola*, *C. mangivora*, *C. acaciivora*, *C. eucalypticola*) appear to be ITS haplotypes that have been moved from or within Brazil on nursery stock. The taxonomic status of other species delineated primarily by ITS sequences (*C. diversiconidia, C. papillata*, *C. neglecta*, *C. ecuadoriana*, *C. fimbriatomima*, *C. curvata*) needs further study, but they are considered doubtful species.

Key words: *Eucalyptus* spp., *Mangifera indica*, *Punica granatum*

Introduction

Tandemly repeated copies of similar or identical rDNA and universal primers flanking the highly variable internal transcribed spacers (ITS) have made sequences of this region of great use in delimiting and identifying fungal species (Schoch et al. 2012). However, intraspecific and intragenomic variation in the rDNA array can occur (Cigelnik 1997, Lachance et al. 2003, O'Donnell and Smith et al. 2007, Nilsson et al. 2008, James et al. 2009, Linder and Banik 2011). The ITS region has been problematic in delimiting species in the Latin American Clade (LAC) of the *Ceratocystis fimbriata* Ell. & Halst. complex (Harrington et al. 2011). Fourteen new species have been described recently in the LAC, in which there is limited morphological variation, physiological (host) specialization, isozyme variation or DNA sequence variation, aside from ITS sequences (Engelbrecht and Harrington 2005; Johnson et al. 2005; Thorpe et al. 2005; van Wyk et al. 2007, 2009, 2010, 2011a, b, 2012; Harrington et al. 2011).

Members of the LAC are primarily soilborne and cause root and corm rots or lethal wilt diseases on a number of exotic hosts in southeastern USA, the Caribbean, Mexico, Central America and South America (CABI 2005, Harrington 2000, Johnson et al. 2005). Clear host specialization to natives has been demonstrated for only two species, *C. cacaofunesta* Engelbr. & TC Harr and *C. platani* Engelbr. & TC Harr, which appear to have adapted to *Theobroma* spp. (cacao) and *Platanus occidentalis* (American sycamore) in the Upper Amazon and eastern USA respectively (Baker et al. 2003; Engelbrecht et al. 2004, 2007; Engelbrecht and Harrington 2005).

Although the LAC appears to be native to the Americas, many haplotypes of the LAC have been moved around the world in propagation material, such as storage roots of *Ipomoea batatas* (sweet potato), corms of *Colocasia esculenta* (taro or inhame) and other Araceae, and cuttings for rooting or grafted nursery stock of *Eucalyptus* spp., *Mangifera indica* (mango), *Ficus carica* (fig) and cacao (Harrington 2000, 2013; Baker et al. 2003; Engelbrecht et al. 2004,

Submitted 7 Jun 2013; accepted for publication 20 Aug 2013.
[1] Corresponding author. E-mail: tcharrin@iastate.edu

2007; Engelbrecht and Harrington 2005; Thorpe et al. 2005; Ferreira et al. 2010, 2011, 2013; Harrington et al. 2011). The species name *C. fimbriata* was based on the widely distributed *Ipomoea* haplotype (Steimel et al. 2004, Engelbrecht and Harrington 2005), and some consider the name *C. fimbriata ss* to apply solely to the *Ipomoea* strain (van Wyk et al. 2007, 2009, 2010, 2011a, b, 2012). The MAT2 strains of *C. fimbriata* are self-fertile due to unidirectional mating-type switching, so even introduced populations that are reproducing sexually may be genetically uniform, except for their ITS sequences (Harrington and McNew 1997; Witthuhn et al. 2000; Engelbrecht et al. 2004, 2007; van Wyk et al. 2006; Ocasio-Morales et al. 2007; Ferreira et al. 2010). Ceratocystis wilt of mango was known only in Brazil, and two ITS haplotypes commonly spread in mango nursery stock in Brazil were recently named *C. mangicola* M. van Wyk & M. J. Wingfield and *C. mangivora* M. van Wyk & M. J. Wingfield respectively (van Wyk et al. 2011a). A new ITS haplotype on mango in Pakistan and Oman was named *C. manginecans* M. van Wyk, Al Adawi & M. J. Wingfield (van Wyk et al. 2007). Other recently described species appear to be ITS haplotypes common on *Eucalyptus* spp. and *Acacia* spp. in Asia (*C. acaciivora* Tarigan & M. van Wyk) and Africa (*C. eucalypticola* M. van Wyk & M. J. Wingfield) (Tarigan et al. 2011, van Wyk et al. 2012).

The greatest genetic variability in the LAC appears to be in northern South America, where a large number of new species have been distinguished based on ITS sequences. Soilborne populations of *C. fimbriata* in Colombia cause an important disease on *Coffea arabica* (coffee) and appear to be uniquely aggressive to Asian *Citrus* spp., and it attacks some native hosts (Marin et al. 2003). The *Coffea*/*Citrus* pathogen in Colombia was named as two species, *C. colombiana* M. van Wyk & M. J. Wingfield and *C. papillata* M. van Wyk & M. J. Wingfield (van Wyk et al. 2010). Other new species from wounds or stumps of plantation trees in Venezuela, Colombia and Ecuador include *C. neglecta* M. van Wyk, Jol. Roux & Rodas (Rodas et al. 2008), *C. fimbriatomima* M. van Wyk & M. J. Wingfield (van Wyk et al. 2009), and *C. curvata* M. van Wyk & M. J. Wingfield, *C. ecuadoriana* M. van Wyk & M. J. Wingfield, and *C. diversiconidia* M. van Wyk & M. J. Wingfield (van Wyk et al. 2011b). With the exceptions of *C. platani* and *C. cacaofunesta*, none of the above species has been shown to have distinguishing phenotype and would not be considered phylogenetic species in the traditional sense (Harrington and Rizzo 1999): '' … the smallest aggregation of populations with a common lineage that share unique, diagnosable phenotypic characters.''

This study focuses on ITS sequence variation in the LAC, including variation in ITS sequences from single-ascospore strains. An earlier phylogenetic comparison of ITS and MAT2 haplotypes in Brazil (Harrington et al. 2011) is expanded to include isolates from outside Brazil and sequences of a MAT1 gene. The taxonomic implications of ITS hypervariability is discussed.

MATERIALS AND METHODS

Isolates.—In addition to the isolates and sequences studied earlier (Harrington et al. 2011), Brazilian isolates from *Eucalyptus* spp., mango, *Acacia* spp., *Crotolaria* sp. (sunhemp) and *Cajanus cajan* (pigeonpea) were subjected to DNA sequencing. Isolates from *Citrus* spp., coffee, and *Inga* sp. in Colombia (Baker et al. 2003) also were included, as well as Asian isolates: an Indian isolate from *Punica granatum* (pomegranate) (Somasekhara 1999), a Pakistani isolate from *Dalbergia sissoo* (Indian rosewood) (Poussio et al. 2010) and mango isolates from Pakistan and Oman (TABLE I).

ITS sequences.—New sequences were generated with PCR and direct DNA sequencing of the PCR products with primers ITS1F and ITS4 and these cycling conditions: 85 C for 2 min, 95 C for 95 s, 36 cycles of 58 C for 1 min, 72 C for 80 s, and 95 C for 70 s, followed by 52 C for 1 min and 72 C for 15 min (Harrington et al. 2011). Some isolates had ITS PCR products that did not produce clean sequencing results with direct sequencing, so the PCR products were cloned into pGEM-T easy vector (Promega Corp., Madison, Wisconsin). Plasmids were extracted with Illustra™ plasmidPrep Mini Spin Kit (GE Healthcare Life Sciences, Buckinghamshire, UK), and the cloned fragments were sequenced at the Iowa State University DNA Synthesis and Sequencing Facility with primers T7-2 and SP6. For three of these isolates, single ascospore strains were derived from the original field isolates by dispersing an ascospore mass in a light oil (Isopar M, ExonMobile Chemical, Houston, USA) and spreading the spore suspension over the plate; individual germlings were subcultured to fresh plates for growth and DNA extraction (Harrington and McNew 1997). One isolate (C2059) that could not be directly sequenced for ITS no longer produced perithecia and ascospores, so a single conidium strain was studied. The generation of PCR products, cloning and sequencing of these fragments were as described above, although Takara *Ex Taq* polymerase (Clontech, Mountain View, California) was used to generate some of the PCR products to compare their sequences with those using Promega GoTaq polymerase.

Sequencing of mating-type genes.—Portions of two mating-type genes and their flanking regions from the mating-type locus were amplified and sequenced for phylogenetic analyses. The genes were identified and PCR primers were developed from sequencing the mating-type locus of a self-fertile strain (C1099) of *C. fimbriata* from *Ipomoea* (Harrington unpubl) and comparing the DNA sequences and putative translated amino acid sequences to that of *C.*

TABLE I. Geographic and host records of ITS-rDNA haplotypes of *Ceratocystis fimbriata* isolates from Brazil and other countries

	Brazilian records			Records from other countries		
ITS rDNA Haplotype[a]	States	Hosts	Studied cultures and/or ITS sequence accessions[b]	Countries	Hosts	Studied cultures and/or ITS sequence accessions[b]
1	São Paulo	*Ficus carica*	C1852, C1853, C1856, C1857 (HQ157542), C1858, C1895, C1899			
1a				Asia, Oceania, St Vincent, USA	*Ipomoea batatas*	C1418 (AY157956), C1476 (= ICMP 8579, AY157957)
1b-e (*C. eucalypticola*)	Bahia	*Eucalyptus* spp.	C2123, C2124	Congo, South Africa	*Eucalyptus* spp.	AF395685, FJ236721-4, FJ236733-4
2	São Paulo	*Mangifera indica*	C1655 (HQ157546)			
3	Bahia, Rio de Janeiro	*Eucalyptus* spp., *M. indica*	C1440 (HQ157544), C1441, C1444, C1451, C1665, C1986, C2015, C2021, C2059ssc19 (= C2976)	Pakistan, Thailand, Uganda	*Eucalyptus* spp., *M. indica*	C2759ssc12 (= C2977, CBS 135868), AF395686-7, FJ236731-2, FJ236737-9
3a	Distrito Federal	*M. indica*	C2176			
4	Bahia, Distrito Federal	*Cajanus cajan*, *Eucalyptus* spp.	C1442 (= CBS 115174, HQ157545), C1450, C2174			
4a	Pernambuco	*M. indica*	C1968, C1969			
5	Bahia, Paraná, Rio de Janeiro	*Acacia* sp., *Eucalyptus spp.*, *M. indica*	C1345 (AY157966), C1352, C1985, C1988, C2007, C2018, C2020, C2023, C2026, C2028-9, 2042, C2059ssc11 (= C2976), C2114, C2116, FJ236715	China, Indonesia, South Africa, Thailand, Uruguay	*Acacia mearnsii*, *Eucalyptus* spp.	C1181, AF453438-40, EU588655-6, FJ236716-20, FJ236728-30, FJ236735-6, FJ236740-1, FJ236743-4, JQ862733-6
5a				Uruguay	*Eucalyptus* spp.	FJ236726-7
5b (*C. curvata*)				Ecuador	*Eucalyptus* spp.	FH151438, J151437
6 (*C. acaciivora*)	Bahia, Rio de Janeiro, São Paulo	*Eucalyptus* spp., *M. indica*	C1656, C1784-5, C2054, C2055 (HQ157548), C2056-7	India, Indonesia, Oman, Pakistan	*Acacia* spp., *Dalbergia sissoo*, *Eucalyptus* spp., *M. indica*, *Punica granatum*	C1213ssc2 (= C2974), C2530, C2730, C2751, C2753, C2755-6, C2758, C2759ssc7 (= 2977, CBS 135868), C2760, C2762, C2785ssc8 (= C2975, CBS 135988), EU588657, EU588660-1, FJ236742
6a	Rio de Janeiro	*M. indica*	C2093			
7	São Paulo	*M. indica*	C1889 (HQ157547)			
7a				India	*P. granatum*	C1213ssc3 (= C2974)

TABLE I. Continued

ITS rDNA Haplotype[a]	Brazilian records			Records from other countries		
	States	Hosts	Studied cultures and/or ITS sequence accessions[b]	Countries	Hosts	Studied cultures and/or ITS sequence accessions[b]
7b (*C. manginecans*)				Indonesia, Oman, Pakistan	*Acacia* sp., *D. sissoo, M. indica*	C2529, C2752, C2754, C2757 (= CBS 135866), C2759ssc1 (= 2977, CBS 135868), C2761, C2785ssc1 (= C2975, CBS 135988), C2786, C2787 (= CBS 135867), AY953383-5, EF433300-5, EU588658-9, EU588662-5
8	São Paulo	*F. carica*	C1782 (= CBS 115166, AY526292), C1783, C1848-1851, C1854-5			
8a, b (*C. papillata*)				Colombia	*Coffea arabica, Citrus* spp.	AY233867-8
8c, d (*C. neglecta*)				Colombia	*Eucalyptus* spp.	EF127990-1
8e, f (*C. ecuadoriana*)				Ecuador	*Eucalyptus* spp.	FJ151432, FJ151434
8g (*C. fimbriatomima*)				Venezuela	*Eucalyptus* spp.	EF190963-4
9	Rio de Janeiro, São Paulo	*Colocasia esculenta, Annona sp., M. indica*	C1554, C1577, C1558 (= CBS 115175, AY157965), C1589-90, C1592, C1671, C1914 (HQ157540), C2112			
10, 10a, 10b (*C. mangivora*)	São Paulo	*M. indica*	C994 (= CBS 600.70, AY157964, EF042605), C1847, FJ200261-8			
10c	Pernambuco	*M. indica*	C1970			
11	Rio de Janeiro	*C. esculenta*	C1865 (= CBS 114713, AY526286)			
12	São Paulo	*C. esculenta*	C1907, C1916, C1926 (HQ157541)			
13	São Paulo	*C. esculenta*	C1905 (= CBS 115171, AY526288), C1915, C1917-1918, C1920-1922			
14, 14a (*C. mangicola*)	Rio de Janeiro, São Paulo	*M. indica*	C1657, C1688 (= CBS 114721, AY526291), C2092, C2094, EU200256-60			
14b	Paraná	*Eucalyptus* sp.	C1987			

TABLE I. Continued

	Brazilian records			Records from other countries		
ITS rDNA Haplotype[a]	States	Hosts	Studied cultures and/or ITS sequence accessions[b]	Countries	Hosts	Studied cultures and/or ITS sequence accessions[b]
14c	Distrito Federal	*Crotolaria niger*, *Cajanus cajan*	C2173, C2175			
15	Pará	*Gmelina arborea*	C918, C920, C925 (= CBS 115173, AY157967)			
16	Pará	*G. arborea*	C924 (HQ157539)			

[a] ITS haplotype numbers follow the designations of Harrington et al. (2011), and new haplotypes are followed by lowercase letters. Species names in parentheses are considered synonyms of *C. fimbriata* sensu stricto.

[b] Isolates numbers are those of the collection of T.C. Harrington, Iowa State University or those also deposited in CBS = Centraalbureau voor Schimmelcultures or in ICMP = International Collection of Microorganisms. Isolate numbers followed by ssc represent ITS rDNA sequences recovered from cloned fragments of PCR products from DNA extracted from single-spore strains. Other numbers are GenBank accessions.

eucalypti (Witthuhn et al. 2000) and other ascomycota (e.g. accession numbers EKJ71585, AF182425). The *MAT1-1-2* and *MAT1-1-1* genes are associated with the MAT1 phenotype, and all tested strains in the *C. fimbriata* complex, whether self-fertile or self-sterile, contain these genes (Witthuhn et al. 2000). The *MAT1-2* gene confers the MAT2 phenotype in pairings, but strains with this gene also are capable of selfing due to unidirectional mating-type switching (Harrington and McNew 1997, Witthuhn et al. 2000). Homothallic strains have the MAT2 gene (*MAT1-2*) in the mating-type locus, flanked by the two MAT1 genes, *MAT1-1-1* and *MAT1-1-2*.

The primers CFMAT1-F (5′-CAGCCTCGATTGAKGGTATGA-3′) and CFMAT1-R (5′-GGCATTTTTACGCTGGTTAG-3′) were used to amplify and sequence a 1022 bp region of *MAT1-1-2*. A BLASTx query (Altschul et al. 1997) with the 1022 bp sequence from isolate C1418 found partial homology with the amino acid sequences of the *MAT1-1-2* gene of *Claviceps purpurea* (BAD72603) and *Gibberella zea* (XP389067), starting near amino acid position 208 of *C. purpurea* and continuing 642 bp beyond amino acid position 350 of *G. zeae* toward the *MAT1-2* gene. The 1022 region used for phylogenetic analysis begins approximately 476 bp into the first exon of the putative *MAT1-1-2* coding region and extends beyond a fourth and final exon for another 43 bp toward the *MAT1-2* gene, which points in the opposite direction.

Primers X9978R1R (5′-GCTAACCTTCACGCCAATTT-3′) and CFM2-1F (5′-AGTTACAAGTGTTCCCAAAAG-3′) amplify and sequence a 1102 bp region that begins 652 bp into the first exon and runs 84 bp beyond the third and final exon toward the *MAT1-1-2* gene (Harrington et al. 2011). For some isolates the forward primer X9978a (5′-GCTAACCTTCACGCCAATTTTGCC-3′) produced a better PCR product for sequencing.

The thermo-cyler settings for amplifying the MAT1 and MAT2 regions included: initial denaturation at 94 C for 2 min, with 36 cycles of 94 C for 1 min, 58 C for 1 min, 72 C for 2 min and a final extension of at 72 C for 10 min. The amplified products were sequenced with the PCR primers.

Phylogenetic analyses.—Isolate C1963 of *C. variospora* (R.W. Davidson) C. Moreau was used as outgroup taxon; it is in the North American Clade of the *C. fimbriata* complex and well outside the LAC based on rDNA and isozyme analyses (Johnson et al. 2005). The sequences were manually aligned, with some ambiguity in the alignment of the ITS sequences with the outgroup taxon due to indels (short regions of insertions or deletions). Some of the ingroup alignment of the ITS dataset also was ambiguous, but in these cases, shifting the alignments did not affect significantly the topology of the generated trees. For the MAT1 and MAT2 sequences, no ambiguity in alignment was seen among sequences of ingroup isolates.

The aligned sequences were analyzed for maximum parsimony (MP) using PAUP 4.0b10 (Swofford 2002). Gaps were treated as a fifth base, all characters had equal weight, and the heuristic searches used simple stepwise addition and tree-bisection-reconnection. Bootstrap analyses also were conducted in PAUP with 1000 replications. Because

of the large number of most parsimonious trees found with the ITS dataset, the maximum number of trees in the ITS bootstrap analysis was set to 100 and only 100 bootstrap replications were run. For the combination of sequences of the two MAT genes with the ITS sequences, the maximum number of trees in the bootstrap analysis was set to 1000.

Posterior probability (Bayesian) estimates for branches were determined with Mr. Bayes 3.2.1 (Ronquist and Huelsenbeck 2003), in which gaps are treated as missing data. The settings included a general time-reversible (GTR) model and gamma distribution. For the ITS dataset, there were 1 000 000 generations, a diagnosis frequency of 5000 and a sample frequency of 1000. The first 25% of samples were discarded (burninfrac = 0.25), and Bayesian posterior probability estimates were calculated by majority rule consensus of the trees after burn-in. The sequences of the two mating-type genes were analyzed separately and combined, with these settings: number of generations 1 000 000, diagnostic frequency 5000 and sample frequency 500. The combined MAT1/MAT2/ITS dataset was analyzed as in the MAT1/MAT2 analyses.

RESULTS

Cloned PCR fragments of ITS.—Amplified ITS rDNA products could be directly sequenced for all but four of the tested isolates: C2059 (from mango in Rio de Janeiro), C2759 (from Indian rosewood in Pakistan), C2785 (from mango in Oman) and C1213 (from pomegranate in India). In these cases repeated sequencing attempts with both primers ITS1-F and ITS4 gave electropherograms that initially showed readable peaks, but then superimposed peaks followed from a fixed base position. When the PCR products were cloned and individual cloned fragments sequenced with the vector primers T7-2 and SP6, clear reads were found, but more than one ITS sequence was identified from among the cloned fragments from a single PCR product (TABLE II). Single-ascospore or single-conidium strains were generated from each of these four isolates, and sequencing the cloned PCR fragments from these single spore strains also identified more than one ITS sequence from a given strain. Cloned fragments amplified from a single conidium strain (from C2059) and a single ascospore strain (from C2759) gave similar variation in sequences whether the PCR product was generated with GoTaq DNA polymerase or *Takara Ex Taq* DNA polymerase.

Some of sequenced fragments appeared to have random, single-base substitutions that were found in the more conserved regions (SSU, LSU, 5.8s) as well as in the ITS1 and ITS2 regions. These base substitutions were unique to a single cloned fragment and were not found in other LAC isolates that were directly sequenced. They occurred in less than 0.3% of the base reads. To determine whether this was a normal PCR error rate or represented intragenomic polymorphisms for the ITS region, the PCR products of two isolates that had been successfully sequenced by direct sequencing (C2756 and C2757, both from mango in Pakistan but with different ITS sequences) were similarly cloned and sequenced. For isolate C2756, nine cloned fragments of 665 bp had 12 unique, single-base substitutions, for a random substitution rate of 0.20%. For the single ascospore strain from C2757, the 661 bp fragment of four clones showed nine unique base substitutions, for an estimated random substitution rate of 0.34%. Although it is possible that they were intragenomic polymorphisms (Ganley and Kobayashi 2007, Simon and Weiss 2008), such unique base substitutions were considered errors during PCR and were not considered to be inherent variation in sequence among the rDNA copies and thus are not illustrated.

The base substitutions (TABLE II) were found not only in more than one of the 68 sequenced cloned fragments but also in other isolates of *C. fimbriata* that were directly sequenced. The 537 bp alignment of the 68 sequenced fragments from the four isolates included 17 variable base positions and 18 unique ITS sequences (TABLE II). For each of the four isolates, about half of the cloned fragments had one sequence, a second sequence occurred in about half that number of fragments, and the other unique sequences occurred in only one or two of the cloned fragments. Ten of the 17 variable sites consisted of either a gap or a base (a single-base indel). The base substitutions of only four sites appeared to be independent from the base substitutions of the other sites. Three of these four variable sites were associated with stretches of mononucleotide repeats: two or three Cs (position 40), three or four As (position 130) and six or seven As (position 188). The insertion of a C at position 473 was unique to some isolates from Oman and Pakistan (all other ITS sequences of the LAC had a gap at this position) but was not associated with multiple Cs.

Among the 68 cloned fragments, the character states within one region in ITS1 (base positions 114–141 in the alignment, except base position 130) and within two regions of ITS2 (base positions 363–370 and 511–529 respectively) appeared to be tightly linked. Most cloned fragments had either T, T, G, G, T or –, –, A, A, – at the respective base positions 114, 120, 122, 128, and 141 (TABLE II). The only exception was in the sequences from isolate C1213, in which all cloned fragments had a T at position 141. Similarly, all 68 cloned fragments had either TTATTCT– or CTCTTTTG at positions 363–370. Last, all cloned fragments had either G, C, T at positions 511, 512 and 529 or had gaps at these

TABLE II. Variable site positions of ITS-rDNA sequences recovered from four isolates of *Ceratocystis fimbriata* by cloning and sequencing of 68 PCR fragments

Isolate No.	Country of Origin	ITS Haplo-type[a]	No. of Cloned Fragments	Base Position in Aligned ITS-rDNA Sequence[b] 40 114—141 188 363—370 473 511—529
C2059	Brazil	3	12	- TAGTCTTCGCCACTGTAAAACTCTTTTT - TTATTCT- - --TCAACTTTTGTTGAAC-
		5	6	C TAGTCTTCGCCACTGT-AAACTCTTTTT A CTCTTTTG - GCTCAACTTTTGTTGAACT
		3x	2	C TAGTCTTCGCCACTGT-AAACTCTTTTT A TTATTCT- - --TCAACTTTTGTTGAAC-
		5x	2	C TAGTCTTCGCCACTGT-AAACTCTTTTT A CTCTTTTG - --TCAACTTTTGTTGAAC-
		3w	1	C TAGTCTTCGCCACTGT-AAACTCTTTTT A TTATTCT- - GCTCAACTTTTGTTGAACT
		3y	1	C TAGTCTTCGCCACTGT-AAACTCTTTTT - TTATTCT- - GCTCAACTTTTGTTGAACT
		3z	1	- TAGTCTTCGCCACTGTAAAACTCTTTTT - TTATTCT- - GCTCAACTTTTGTTGAACT
		5w	1	C TAGTCTTCGCCACTGT-AAACTCTTTTT A CTCTTTTG - --TCAACTTTTGTTGAAC-
C2759	Pakistan	7b	10	- -AGTCT-CACCACTATAAAACTCTTTT- - CTCTTTTG C GCTCAACTTTTGTTGAACT
		3	5	- TAGTCTTCGCCACTGTAAAACTCTTTTT - TTATTCT- - --TCAACTTTTGTTGAAC-
		5z	2	C TAGTCTTCGCCACTGT-AAACTCTTTTT - CTCTTTTG - GCTCAACTTTTGTTGAACT
		6z	2	- TAGTCTTCGCCACTGT-AAACTCTTTTT A CTCTTTTG C GCTCAACTTTTGTTGAACT
		3z	1	- TAGTCTTCGCCACTGTAAAACTCTTTTT - TTATTCT- - GCTCAACTTTTGTTGAACT
		5v	1	- TAGTCTTCGCCACTGTAAAACTCTTTTT A CTCTTTTG - --TCAACTTTTGTTGAAC-
		5y	1	C TAGTCTTCGCCACTGT-AAACTCTTTTT - CTCTTTTG - --TCAACTTTTGTTGAACT
		6	1	- TAGTCTTCGCCACTGT-AAACTCTTTTT A CTCTTTTG - GCTCAACTTTTGTTGAACT
		6y	1	- TAGTCTTCGCCACTGT-AAACTCTTTTT - CTCTTTTG C GCTCAACTTTTGTTGAACT
		7y	1	- -AGTCT-CACCACTATAAAACTCTTTT- - CTCTTTTG - GCTCAACTTTTGTTGAACT
		7z	1	- -AGTCT-CACCACTATAAAACTCTTTT- A CTCTTTTG C GCTCAACTTTTGTTGAACT
C2785	Oman	7b	3	- -AGTCT-CACCACTATAAAACTCTTTT- - CTCTTTTG C GCTCAACTTTTGTTGAACT
		6	1	- TAGTCTTCGCCACTGT-AAACTCTTTTT A CTCTTTTG - GCTCAACTTTTGTTGAACT
C1213	India	7a	7	- -AGTCT-CACCACTAT-AAACTCTTTTT - CTCCTTTG - --TCAACTTTTGTTGAAC-
		6	4	- TAGTCTTCGCCACTGT-AAACTCTTTTT A CTCTTTTG - GCTCAACTTTTGTTGAACT
		7x	1	- -AGTCT-CACCACTAT-AAACTCTTTTT - CTCTTTTG - GCTCAACTTTTGTTGAACT

[a] ITS haplotypes numbered as in TABLE I, or unique haplotypes designated by the number of the most-similar haplotype followed by v, w, x, y or z.

[b] Variable regions of the aligned 531 bp region (including gaps) of portions of ITS 1 (base positions 40, 114–141, 188) and ITS 2 (base positions 363–370, 473, 511–529), with invariable positions within linked regions in gray.

respective positions, except for one cloned fragment from C2759 that had –, –, T.

Comparisons of ITS sequences among isolates.—Representative ITS sequences of the recently described species were aligned with those of Brazilian isolates and newly studied isolates from Pakistan, Oman and India (TABLE I). The most common ITS sequences from the cloned fragments of the four isolates (TABLE II) also were included in the ITS analysis, and these sequences are marked (FIG. 1) with an asterisk. Much of the variation among the ITS sequences was as described above, that is three main regions of insertions of tightly associated character states (positions 114–141, 363–370, and 511–529 in TABLE II) and the regions of mononucleotide repeats. Another polymorphic indel region among the *C. fimbriata* sequences began at position 89 in the alignment (TABLE II), immediately preceding GGG, where a 4 bp insertion of GAGA or GGGA or a deletion was found.

Many of the ITS sequences of the recently described species that were deposited in GenBank were truncated by 25 bases at the beginning of ITS1 and 97 bases at the end of ITS2, leaving 203 aligned bp from ITS1 and 108 aligned bp from ITS2 for analysis. This eliminated one of the common indels (positions 511–529 in TABLE II) from phylogenetic analyses. With the truncated alignment, 41 haplotypes were considered representative of *C. fimbriata*, including 16 Brazilian haplotypes identified in Harrington et al. (2011), and the designation numbers (FIG. 1, TABLE I) are consistent with those in the earlier study. The 25 new haplotypes are designated with the same numbers followed by small letters to indicate their similarity to one of the original 16 haplotypes (TABLE I, FIG. 1).

Of the 311 aligned bases of the ITS1 and ITS2 regions (excluding the 5.8S gene), 175 were constant, 39 were parsimony uninformative and 97 characters were parsimony informative. The MP analysis resulted in a total of 1324 trees of 291 steps, and one of those trees is illustrated (FIG. 1). The homoplasy index (HI) was high (0.4192 with all characters and 0.4880 excluding uninformative characters; consistency index was 0.5808 and 0.5120 respectively), the retention index (RI) was 0.8806 and the rescaled consistency index (RC) was 0.5114. The ITS sequences of two *C. diversiconidia* isolates differed substantially from the rest of the LAC (FIG. 1). Long branch lengths and high bootstrap support values were seen for *C. curvata*, *C. colombiana* and *C. cacaofunesta*, but these three taxa were found among the ITS haplotypes of *C. fimbriata* (FIG. 1). Clades of other species (*C. papillata*, *C. fimbriatomima*, *C. neglecta*, *C. ecuadoriana*) were found near the base of the tree and generally had strong bootstrap support. Moderate bootstrap support was found for groupings of three taro isolates (haplotype ITS9), two mango isolates (ITS10) and three fig isolates (ITS8) of *C. fimbriata* from Brazil.

In the Bayesian analysis, the level of convergence from two parallel runs after 1 000 000 generations had a mean standard deviation of split frequencies of 0.0354. Moderate to strong posterior probability values (FIG. 1) were seen for some putative species: *C. curvata* (100), *C. neglecta* (100), *C. ecuadoriana* (99), *C. platani* (100) and *C. diversiconidia* (100). Only weak posterior probability support was seen for *C. colombiana*, *C. cacaofunesta, C. papillata* and *C. fimbriatomima*. The posterior probability values were higher for the branch supporting three taro isolates (ITS12) and the branch grouping isolates from *Gmelina arborea* in Pará (ITS15 and ITS16) with some mango isolates from São Paulo (ITS14 = *C. mangivora*). Five recently described species (*C. eucalypticola*, *C. acaciivora*, *C. manginecans*, *C. mangivora*, *C. mangicola*) had either little bootstrap support and/or low posterior probability support (FIG. 1).

Sequences of isolates from different hosts were found scattered throughout the ITS tree (FIG. 1). Isolates from mango had 13 different ITS haplotypes, isolates from *Eucalyptus* spp. had 10 ITS haplotypes, isolates from taro had four haplotypes, isolates from *Acacia* spp. had three haplotypes and isolates from fig had two haplotypes (TABLE I). Twenty-four of the 41 ITS haplotypes were found in Brazil and all but eight in South America. The ITS haplotypes (ITS1b-e) representing *C. eucalypticola* in South Africa (van Wyk et al. 2012) also were found in a *Eucalyptus* isolate from Congo and in two isolates from a *Eucalyptus* plantation in Bahia, but the similar ITS3 and ITS5 haplotypes were more common among *Eucalyptus* isolates from Brazil (TABLE I). The *C. acaciivora* ITS haplotype (ITS6) from Indonesia was found in Bahia, Rio de Janeiro and São Paulo in Brazil, as well as in India, Indonesia, Oman and Pakistan (FIG. 1, TABLE I).

The ITS7b haplotype of *C. manginecans* from Oman and Pakistan and Indonesia was not found in Brazil but was similar to the common Brazilian haplotypes ITS5, ITS6, and ITS7 (FIG. 1). Further, the Brazilian haplotype ITS6 is strongly associated with the ITS7b haplotype in the Oman and Pakistan mango populations (Al Adawi et al. 2013). Two of the three isolates from diseased mango in Oman (C2786 from the Shinas District, and C2785 from the Barka District) had the ITS7b sequence based on direct sequencing, and a single ascospore strain from

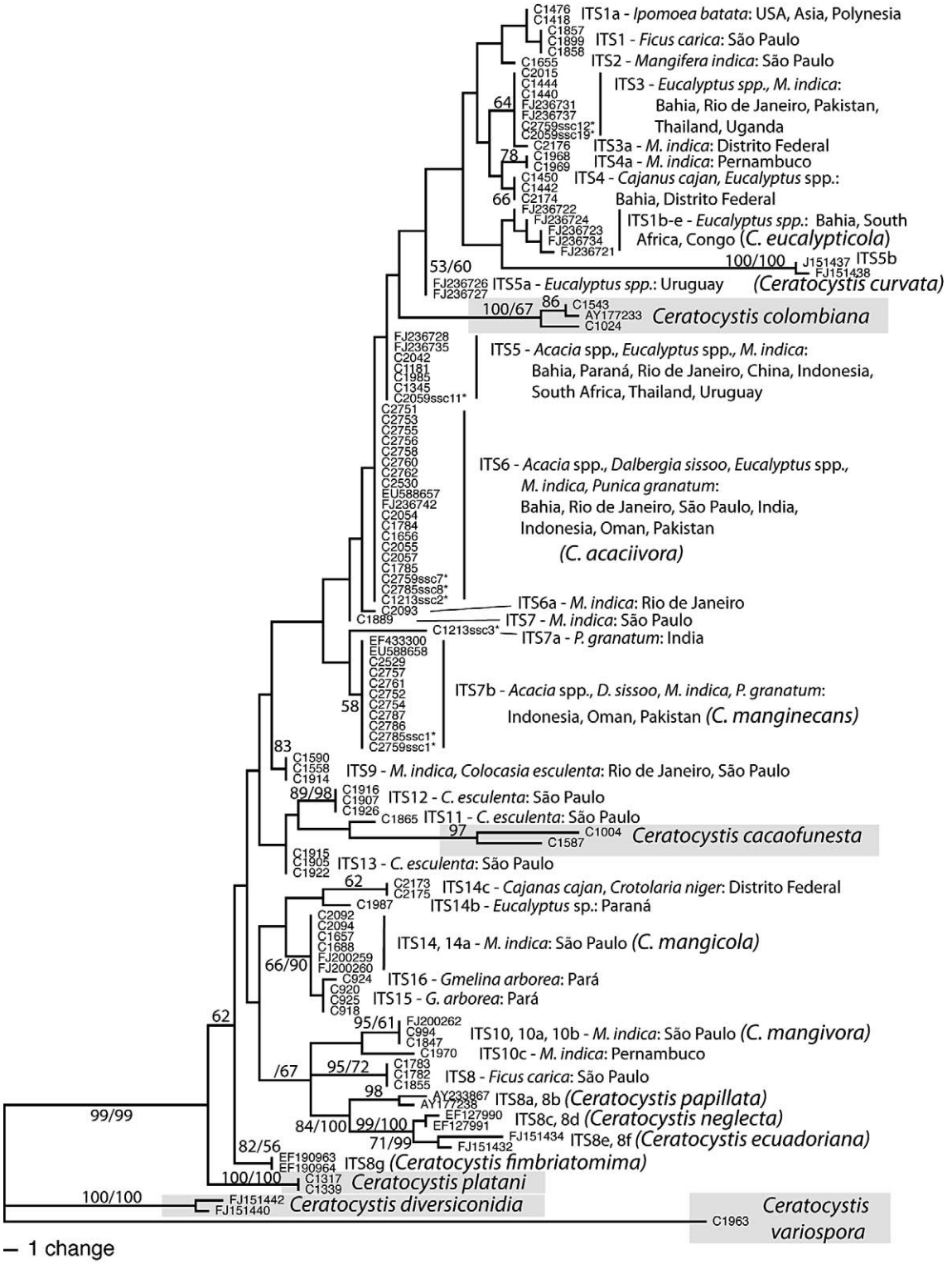

FIG. 1. One of 1324 most parsimonious trees based on the ITS rDNA sequences of representative isolates of the Latin American Clade (LAC) of the *Ceratocystis fimbriata* complex. The tree is rooted to *C. variospora* of the North American Clade, and all other isolates are considered to be in the LAC. Species names thought to be synonyms of *C. fimbriata* are in parentheses, and the ITS haplotypes of other accepted species are in gray. The ITS haplotypes of *C. fimbriata* are numbered following the numerical designations of Harrington et al. (2011), with new haplotypes designated by a lowercase letter after the number. Host genera and countries or states (Brazil) of origin are given for the *C. fimbriata* haplotypes. Isolate numbers (beginning with ''C'') or GenBank accession numbers are given for the representatives of each hapotype.

C2785 (Mahda District) had both the ITS7b and ITS6 sequences among the rDNA repeats (TABLE II). Of the 13 studied mango isolates from Pakistan, five had the ITS7b sequence and eight had the ITS6 sequence based on direct sequencing. Moving from south to north in Pakistan, one isolate from Sindh province had the ITS7b sequence and one had the ITS6 sequence; five isolates from the Bahwalpur and Multan region had the ITS6 sequence and three had the ITS7b sequence; two isolates from Faisalabad had the ITS6 sequence; and one isolate from Islamabad had the ITS7b sequence. Isolate C2759 from Indian rosewood in Islamabad had a mixture of ITS sequences, including ITS7b, ITS6, and the common ITS3 sequence (TABLE II). The ITS6 sequence and the similar ITS7a sequence were found in the pomegranate isolate from India (FIG. 1, TABLE II).

ITS sequences compared with other species criteria.—The illustration (FIG. 1) was duplicated and simplified (FIG. 2) to show only the haplotypes thought to represent *C. fimbriata* sensu stricto and illustrate other relationships among the ITS haplotypes. Based on the data (TABLE II), those pairs of ITS haplotypes found in a single ascospore isolate (isolate numbers marked with an asterisk) are connected with solid lines with double arrowheads (FIG. 2). For instance, the ITS6 haplotype (*C. acaciivora*) was found in single ascospore strains that also yielded the ITS3, ITS7a and/or ITS7b (*C. manginecans*) haplotypes. A mango isolate from Rio de Janeiro (C2059) yielded both the ITS3 and ITS5 haplotypes.

Some populations have been shown to be nearly clonal based on microsatellite markers (Ferreira et al. 2010, Al Adawi 2011) but have different ITS haplotypes, and these haplotypes are connected with dotted lines and double arrowheads (FIG. 2). The population on *Ficus* appears to be derived from introductions of a single genotype of *C. fimbriata* in nursery stock sold to a small number of fig growers in São Paulo because 19 of 20 isolates from these plantings have identical microsatellite alleles at 14 loci (Ferreira et al. 2010). However, these same isolates have ITS sequences (ITS1 and ITS8) that differ at 14 base positions (seven in ITS1 and seven in ITS2). The ITS1 haplotype is similar to that of the *Ipomoea* strain, while the ITS8 haplotype is near that of *C. mangivora*, *C. papillata*, *C. neglecta* and *C. ecuadoriana* (FIG. 1). The Oman and Pakistan isolates also have nearly identical microsatellite markers (Al Adawi 2011), yet both the ITS6 (*C. acaciivora*) and ITS7b (*C. manginecans*) sequences can be found in Pakistan and Oman, and both sequences can be recovered from single-spore strains from each country (FIG. 2, TABLE II). The pomegranate isolate from India (C1213) also has the microsatellite alleles of the typical Pakistani and Omani isolates (Harrington unpubl), but cloned fragments from the PCR product of this isolate yielded the ITS6, ITS7a and ITS7x sequences (FIG. 2, TABLE II).

Some of the studied isolates had been used in interfertility tests and could be considered members of a biological species (Ferreira et al. 2010). Fully interfertile strains produced an abundance of perithecia with copious ascospore masses at the apex of the perithecia, the ascospores readily germinated, and progeny showed evidence of segregation of characters. The dashed, straight lines to the right (FIG. 2) connect representatives of ITS haplotypes in Brazil that were interfertile with each other and a tester from *Ipomoea* (Ferreira et al. 2010). Summarizing the fully interfertile crosses: the female *Ipomoea* strain from C1418 (ITS1a) × C1440 (ITS3), C1590 (ITS9), C1907 (ITS12), C1657 (ITS14) and C924 (ITS16); the female tester from C1440 (ITS3) × C1858 (ITS1), C1907 and C1926 (ITS12) and C1657; the female tester from C1907 (ITS12) × C1926 (ITS12), C1657 and C925 (ITS15); the female tester from C1657 (ITS14) × C1440, C1907 and C925; and the female testers from C920 and C925 (ITS15) × C1858, C1440, C1783 (ITS8), C918, C1590, C1657, C1907, C1926 (ITS15), C924 and C925. None of the above strains was interfertile with female tester strains from C1317 (*C. platani*) or C1587 (*C. cacaofunesta*).

Phylogenetic analyses of mating-type genes.—An alignment of 1039 characters of one of the MAT1 genes (*MAT1-1-2*) showed only limited variation among the sequences of 62 isolates: 877 characters were constant, 138 were parsimony uninformative, and 24 were parsimony informative. Ten MAT1 haplotypes were identified among the 49 isolates of *C. fimbriata* (TABLE III). Two other MAT1 haplotypes were found in *C. colombiana*, two in *C. cacaofunesta* and one in *C. platani*. Maximum parsimony analysis found 12 trees

←

Isolate numbers followed by ssc, a number and an asterisk indicate a sequence of a single spore strain that yielded more than one ITS sequence. Bootstrap values/posterior probability values greater than 50% are indicated on appropriate branches.

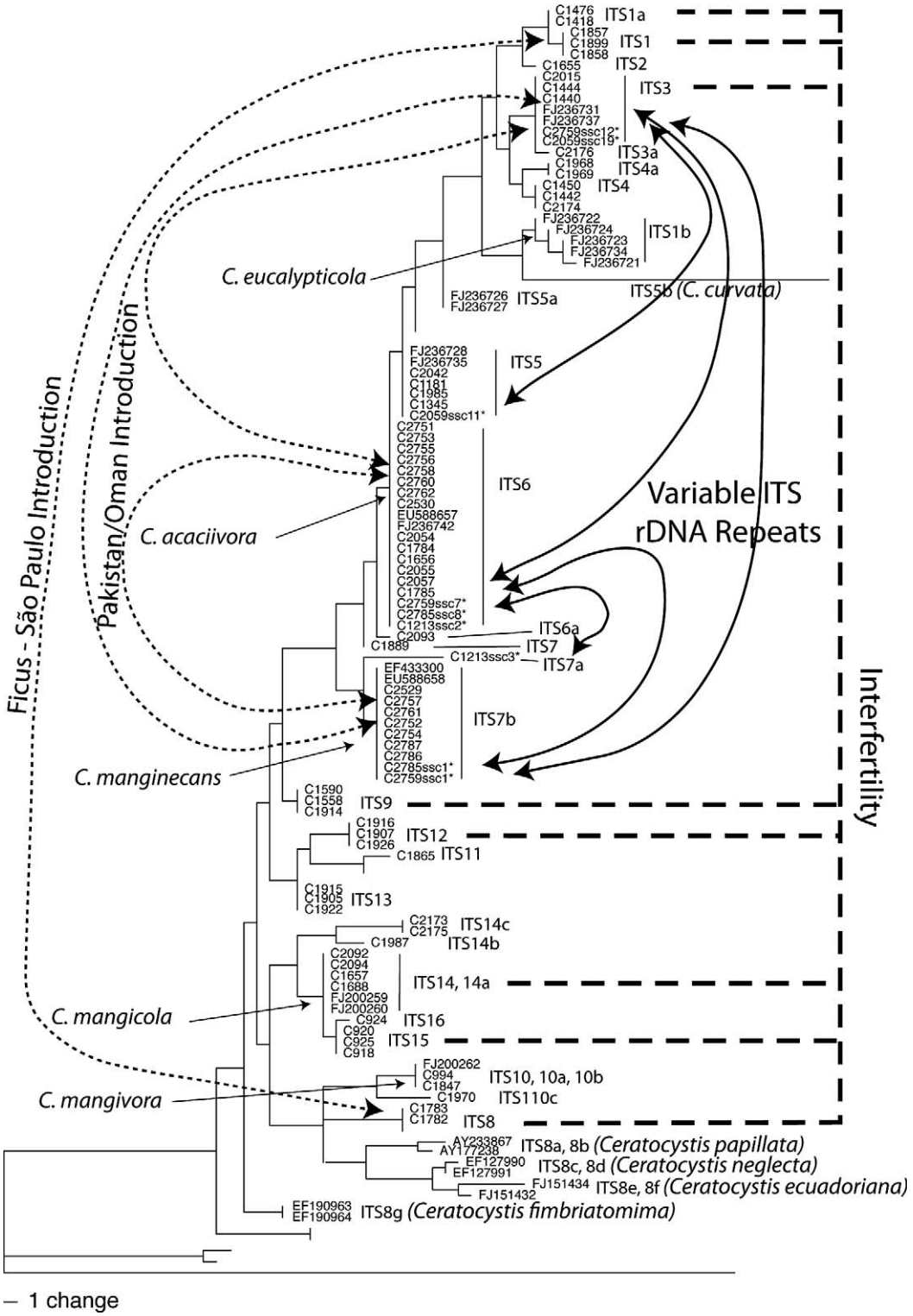

FIG. 2. Intraspecific relationships among *Ceratocystis fimbriata* ITS haplotypes based on interfertility (connected by dashed lines on the right, from Ferreira et al. 2010), identical or nearly identical microsatellite markers (connected by dashed, curved lines with arrowheads, from Ferreira et al. 2010 and Al Adawi 2011), and single-spore isolates with more than one ITS sequence due to intragenomic variation (connected by curved lines with arrowheads, TABLE II). The ITS haplotypes of *C. fimbriata* are in the skeletal ITS phylogenetic tree (FIG. 1). Putative species names thought to represent Brazilian ITS haplotypes are to the left of the tree. Other putative species names needing further taxonomic study are in parentheses to the right of the tree.

TABLE III. Representative *MAT1-1-2* (MAT1) and *MAT1-2* (MAT2) sequences for 12 mating-type haplotypes of isolates of *Ceratocystis fimbriata* and for *C. cacaofunesta*, *C. colombiana*, *C. platani* and *C. variospora*

Mating type haplotypes	ITS rDNA haplotypes	Representative isolate	Hosts	Location	*MAT1-1-2* GenBank accessions	*MAT1-2* GenBank accessions
1	15, 16	C925 (CBS 115173)	*Gmelina arborea*	Pará, Brazil	KF482983	HQ157549
2	4a	C1968	*Mangifera indica*	Pernambuco, Brazil	KF482984	HQ157553
3a	1e, 3, 3a, 4, 5	C1442 (= CBS 115174)	*Eucalyptus* spp.	Bahia, Brazil	KF482985	HQ157550
3b	2, 5, 6, 7, 10c, 14, 14c	C1688 (= CBS 114721)	*M. indica*	São Paulo, Brazil	KF482986	‘‘
4a	1, 8, 10	C994 (= CBS 600.70)	*M. indica*	São Paulo	KF482987	HQ157551
4b	14	C2092	*M. indica*	Rio de Janeiro, Brazil	‘‘	KF482997
4c	14	C2094	*M. indica*	Rio de Janeiro	‘‘	KF482998
5a	9	C1558 (= CBS 115175)	*M. indica*	Rio de Janeiro	KF482988	HQ157552
5b	11, 12, 13	C1905 (= CBS 115171)	*Colocasia esculenta*	São Paulo	KF482989	‘‘
5c	4b	C1987	*Eucalyptus* sp.	Paraná, Brazil	KF482990	‘‘
7	3, 6, 7a, 7b	C2759 (= C2977, CBS 135868)	*Dalbergia sissoo*	Pakistan	KF482991	KF482999
8	1a	C1476 (= ICMP 8579)	*Ipomoea batatas*	Papua New Guinea	KF482992	KF483000
C. cacaofunesta		C1004 (= CBS 153.62)	*Theobroma cacao*	Ecuador	KF482993	KF483001
C. colombiana		C1543 (= CBS 135861)	*Coffea arabica*	Colombia	KF482994	KF483002
C. platani		C1317 (= CBS 115162)	*Platanus occidentalis*	North Carolina, USA	KF482995	KF483003
C. variospora		C1963 (= CBS 135862)	*Prunus* sp.	Iowa, USA	KF482996	KF483004

[a] Isolates numbers are those of the collection of T.C. Harrington, Iowa State University and equivalent numbers in CBS = Centraalbureau voor Schimmelcultures or ICMP = International Collection of Microorganisms from Plants, Landcare Research, New Zealand.

of 166 steps, with a HI = 0.012 using all characters and HI = 0.0714 excluding uninformative characters, RI = 0.9789, and RC = 0.9672.

The aligned sequences of 1141 characters of the MAT2 gene (*MAT1-2*) also showed limited variation among the 62 sequenced isolates: 835 characters were constant and 251 were parsimony uninformative, leaving 55 parsimony informative characters. Nine MAT2 haplotypes were identified among the 49 isolates of *C. fimbriata* (TABLE III), two in *C. cacaofunesta* and one each in *C. colombiana* and *C. platani*. Maximum parsimony analysis found two trees of 330 steps, with a HI of 0.0455 using all characters and HI = 0.1899 excluding uninformative characters, RI = 0.9598, and RC = 0.9162.

When the data for the MAT1 and MAT2 genes were combined, 12 mating-type haplotypes were identified among the isolates considered to be *C. fimbriata* (TABLE III). The designation of the haplotypes follows designations of the five MAT2 haplotypes in Harrington et al. (2011). Maximum parsimony analysis of the combined dataset yielded a single tree of 477 steps (FIG. 3), with HI = 0.0356 (HI = 0.1604 excluding uninformative characters), RI = 0.9636 and RC = 0.9293. The likelihood tree from Bayesian analysis had the identical topology of the MP tree. Bootstrap values and posterior probability values both indicated that the combined tree was robust, with moderate to strong support for most branches (FIG. 3). There was weak support for the grouping of *C. colombiana* and *C. platani*, but each species was well supported at the base of the tree (FIG. 3). There was good support for *C. cacaofunesta*, which was basal to the branch connecting the rest of the isolates (*C. fimbriata*), which had bootstrap (88) and posterior probability (99) support.

There was little similarity between the topology of the ITS tree (FIG. 1) and the mating-type tree (FIG. 3). Isolates of some ITS haplotypes had more than one mating haplotype (FIG. 3, TABLE III): ITS5 haplotype with mating haplotypes 3a and 3b; ITS6 (*C. acaciivora*) haplotype with mating haplotypes 3b and 7; and ITS14 (*C. mangicola*) haplotype with mating haplotypes 3b, 4b and 4c. The taro ITS haplotypes (ITS11, ITS12, ITS13) grouped with two mango isolates from northeastern Rio de Janeiro (ITS9) and an isolate from *Eucalyptus* sp. in Paraná (ITS14b) in the mating-type tree, and this branch was basal to the *Ipomoea* isolates and the rest of the *C. fimbriata* haplotypes (FIG. 3). Mating haplotypes 4a and 4c were found in isolates that had haplotype ITS10 (*C. mangivora*), ITS14 (*C. mangicola*) and the two *Ficus* ITS haplotypes (ITS1 and ITS8). Most of the isolates from *Eucalyptus* spp. had mating haplotype 3a, which included isolates of haplotypes ITS3, ITS3a, ITS4 and ITS5 (FIG. 3) and the Bahian *Eucalyptus* isolate C2123 that had the ITS1e haplotype (*C. eucalytpicola*). The similar mating haplotype 3b had isolates of haplotypes ITS2, ITS5, ITS6 (*C. acaciivora*), ITS7, ITS10c, ITS14 (*C. mangicola*) and ITS14c. Mating haplotype 7 was represented by isolates from Pakistan and Oman populations (haplotypes ITS6 and ITS7b, *C. acaciivora* and *C. manginecans* respectively) and the pomegranate isolate from India (ITS6 and ITS7a), and this mating haplotype was similar to the mating haplotype of isolate C2092 (ITS14 = *C. mangicola*) from mango in Rio de Janeiro (TABLE III).

Combined mating type and ITS datasets.—In spite of the obvious incongruence of the mating type and ITS trees, the ITS sequences were added to the mating-type dataset of 62 isolates to see whether the topology or robustness of the tree would be affected. While the analysis of the MAT1 and MAT2 sequences yielded a single MP tree (FIG. 3), there was a substantial loss of resolution with the addition of the ITS dataset. The combined dataset of 2641 bp had 106 characters that were parsimony informative and yielded 308 MP trees of 586 steps. The homoplasy index was much higher than with the sequences of the mating-type genes alone (HI = 0.1177, HI = 0.3670 excluding uninformative characters), and the RI (= 0.8991) and RC (= 0.7933) were lower.

The MP tree with the MAT1/MAT2/ITS sequences looked similar to the MP tree generated with only the mating-type genes with respect to the *C. colombiana*, *C. platani* and *C. cacaofunesta* branches but with lower bootstrap/posterior probability values (81/97, 82/100, 79/100 for the three respective species). Most of the *C. fimbriata* groupings found in the mating-type analysis were lost with the addition of the ITS data (trees not shown). Bootstrap and posterior probability (pp) values with the mating type plus ITS dataset were similar to those in the ITS trees (FIG. 1), with moderate to strong bootstrap/posterior probability support (0.9 or greater for both) for only the branch grouping the two *Ipomoea* isolates (ITS1a, 100/99), the branch grouping the *Gmelina* isolates (ITS15 and ITS16, 95/100) and the branch connecting two *Ficus* isolates (C1782 and C1783 of the ITS8 haplotype) with two isolates of the ITS14 haplotype (C994 and C1847, *C. mangivora*) (91/99).

DISCUSSION

There is too much variability and apparent randomness in ITS rDNA sequences of the LAC of the *C. fimbriata* complex to be used confidently in phylogenetic analyses or in delimiting species (Harrington et al. 2011). The inferred relationships among popula-

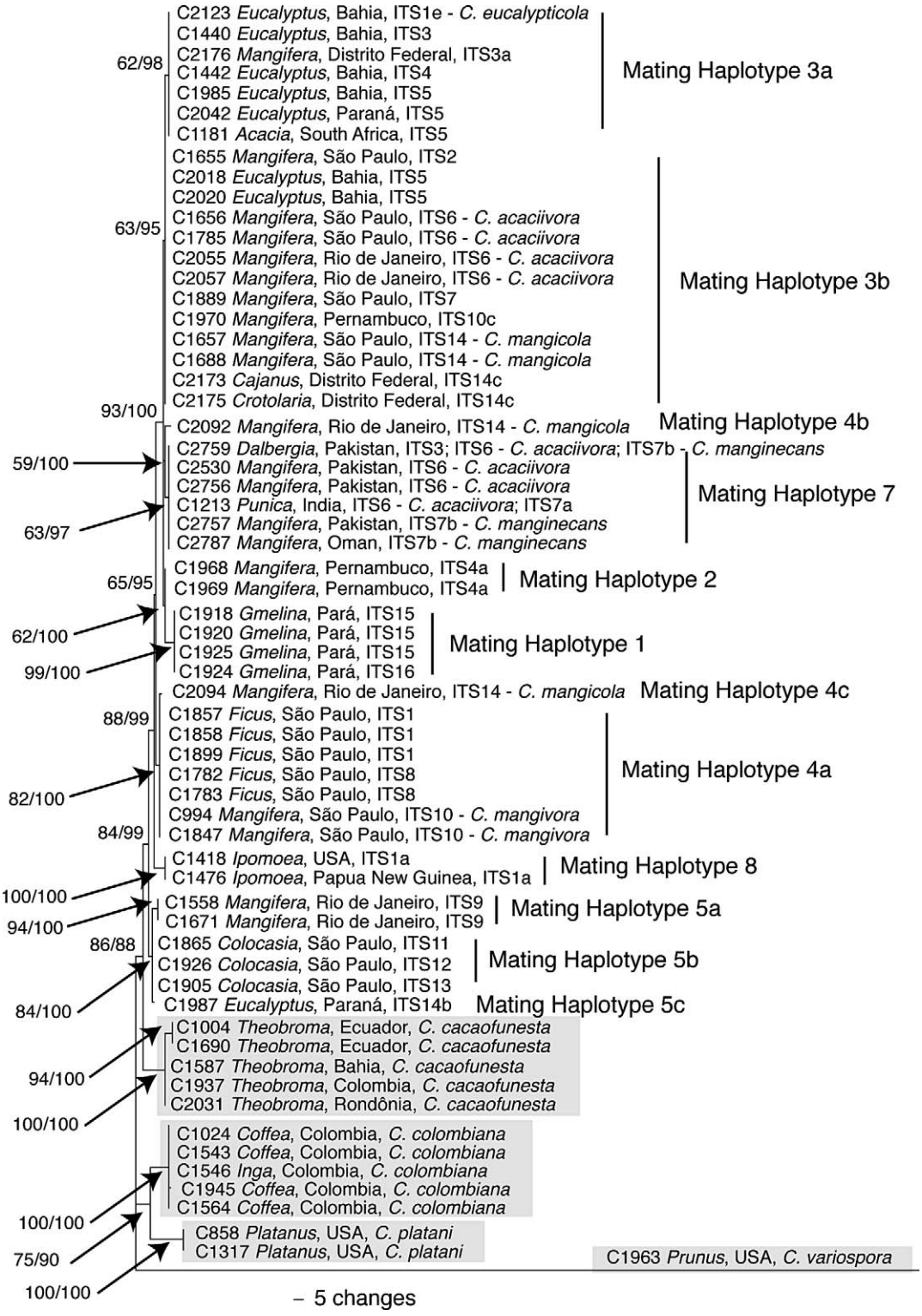

FIG. 3. The single most parsimonious tree based on the combined dataset of the *MAT1-1-2 gene* (MAT1) and the *MAT1-2* gene (MAT2) of representative isolates of the Latin American Clade (LAC) of the *Ceratocystis fimbriata* complex. Isolate number, host genus, country or state (within Brazil) and ITS haplotype (and equivalent species name) are given for each sequenced isolate. Accepted species other than *C. fimbriata* (*C. platani*, *C. colombiana*, *C. cacaofunesta*) and the outgroup taxon (*C. variospora*) are in gray, while other mating haplotypes (unique MAT1/MAT2 sequences) of *C. fimbriata* are numbered following the MAT2 haplotype nomenclature of Harrington et al. (2011). Bootstrap values/posterior probability values greater than 50% are indicated on appropriate branches.

tions and putative species using ITS sequences differed greatly from those derived from sequences of mating-type genes. Strains with different ITS sequences representing different putative species were fully interfertile in mating studies. Isolates from the same introduced population that had gone through severe genetic bottlenecks based on microsatellite markers had ITS sequences that differed by up to 14 bp and would be considered two species if the ITS species concept were applied. Two or more ITS sequences representing different ITS species were obtained from a single ascospore strain.

Most of the intragenomic and intraspecific variation in ITS sequences was of two types. There were regions of mononucleotide repeats, such as the A repeat of 6–13 bp in ITS1. Slippage during replication or unequal crossovers during meiosis could produce such variation among the ITS repeats. More important, stretches in ITS1 and ITS2 appeared to be allelic, that is the characters within the stretches appeared to be linked when comparing individual isolates from an introduced population or cloned fragments from single ascospore strains. The allele-like stretches of linked characters in the ITS region were found in both southern (Brazil) and northern (Ecuador and Colombia) populations, so it is possible that distantly related strains of *C. fimbriata* or closely related species crossed early in the evolutionary history of the LAC. Lineage sorting of such divergent ITS sequences may be confounding phylogenetic inference and species delimitation.

The tandem rDNA repeats in a single haploid genome should be nearly identical in sequence due to concerted evolution (Ganley and Kobayashi 2007), although minor variation among the repeats may not always be evident by direct sequencing (Simon and Weiss 2008). If there is variation in one of the repeats due to a mutation or a crossover during sexual reproduction, then concerted evolution should homogenize the repeats to either correct the mutations or change all the repeats to the new sequence (Ganley and Kobayashi 2007). The ITS repeats appear to be mostly homogeneous for the majority of isolates of *C. fimbriata* because direct sequencing of the ITS region was possible for all but four of the isolates tested. The exceptional cases suggest recent recombinations within the rDNA array without sufficient time for concerted evolution to homogenize the repeats.

All members of the *C. fimbriata* complex are homothallic through uni-directional mating-type switching (Harrington and McNew 1997, Witthuhn et al. 2000), but some outcrossing occurs in natural populations (Engelbrecht et al. 2004, 2007; Ferreira et al. 2010). In the absence of dispersal by humans, populations of *C. fimbriata* are isolated due to limited dispersal in soil or by insects, so accumulations of mutations may result in highly differentiated populations (Ferreira et al. 2010, Harrington et al. 2011). Recently introduced populations of *C. fimbriata* may have two or more ITS haplotypes, and sexual crosses between haplotypes could result in within-repeat recombination and long stretches within ITS1 and ITS2 that may vary substantially among the repeats of a single genome.

Regardless of the source of the ITS hypervariability, the taxonomy of the LAC requires a more rigorous species concept. Morphological differences are not sufficient to delimit taxa in the group (Engelbrecht and Harrington 2005, van Wyk et al. 2011b). The combined MAT1/MAT2 dataset showed minor sequence variation, but few of the haplotypes have phenotypic autapomorphies that could define phylogenetic species (Harrington and Rizzo 1999), and evidence for biological species is limited. Both *C. platani* and *C. cacaofunesta* are monophyletic and differ in physiology (unique pathogenicity), and they are not interfertile with tester strains of *C. fimbriata* from Brazil and elsewhere (Baker et al. 2003, Engelbrecht and Harrington 2005, Ferreira et al. 2010). Although genotypes of both *C. cacaofunesta* and *C. platani* have been moved elsewhere around the world by humans, these pathogens show their greatest genetic variation within the native range of their hosts, that is on cacao in the Upper Amazon and on sycamore in eastern USA respectively (Engelbrecht et al. 2004, Engelbrecht et al. 2007). Host-specialization may have been a driving force in relatively recent speciation events.

Some populations of the LAC in Colombia are soilborne and appear to be uniquely aggressive on coffee and citrus, as well as on hosts native to Colombia (Marin et al. 2003). Inoculation of coffee plants revealed a wide range of aggressiveness (Marin et al. 2003), as might be expected for natural populations (Harrington et al. 2011). The ITS sequences of Colombian isolates (e.g. isolate C1543, sequence AY157961) from Baker et al. (2003) were similar to those of the recently described *C. colombiana* (van Wyk et al. 2010), and these isolates are monophyletic and distinct from *C. fimbriata* in MAT1 and MAT2 analyses. Further, mating studies suggest that *C. colombiana* is a distinct biological species (Harrington unpubl). *C. papillata* also was described as an aggressive pathogen on coffee and citrus in Colombia (van Wyk et al. 2010), but it appears to be *C. fimbriata* or perhaps *C. colombiana* based on ITS sequences. *C. neglecta*, *C. ecuadoriana*, *C. curvata* and *C. fimbriatomima* (van Wyk et al. 2010, 2011b) also require further study,

but at present these are considered synonyms of *C. fimbriata*.

Aside from *C. cacaofunesta*, Brazilian populations on various exotic hosts have been shown to be interfertile in sexual crosses with *C. fimbriata* strains from *Ipomoea* (Ferreira et al. 2010) and vary greatly in aggressiveness to exotic, cultivated crops, although they do not appear to be host specialized (Baker et al. 2003, Zouza et al. 2004, Thorpe et al. 2005, Silveira et al. 2006, Harrington et al. 2011). Thus, the Brazilian isolates and the *Ipomoea* haplotype are of a single biological and phylogenetic species. Five recently described species in the LAC appear to be Brazilian haplotypes of *C. fimbriata* that have been moved in cuttings or grafted nursery stock (Ferreira et al. 2010, 2011).

Two common genotypes of *C. fimbriata* on mango in São Paulo recently were named *C. mangicola* and *C. mangivora* based on ITS sequences (van Wyk et al. 2011a), but they appear to be merely genotypes that have been moved around the country in mango nursery stock (Rossetto and Ribeiro 1990, Ferreira et al. 2010). The ITS sequences of a collection of isolates from *Eucalyptus* spp. in Africa, Asia and South America were found to be similar to an isolate from Brazil, and it was suggested that they may represent a single species. The new name *C. eucalypticola* was applied only to the South African ITS haplotypes (van Wyk et al. 2012), but ITS sequences of isolates from Congo (ITS1e), Uganda (ITS3), Thailand (ITS3 and ITS5), Indonesia (ITS5 and ITS6), China (ITS5) and Uruguay (ITS5 and ITS5a) were deposited in GenBank as representative of *C. eucalypticola*. Isolates from *Eucalyptus* in Bahia (C2123, C2124) have the ITS1e sequence of *C. eucalypticola* and mating haplotype 3b, one of the most common mating haplotypes of isolates from Brazil. Our South African isolate from *Acacia mearnsii* (C1181, CMW4101) has the ITS5 haplotype, which, along with ITS haplotypes 3 and 6, are predominant ITS haplotypes on *Eucalyptus* in Brazil.

Recent epidemics of Ceratocystis wilt on mango and Indian rosewood in Pakistan (Fateh et al. 2006, Poussio et al. 2010) and on mango in Oman (Al Adawi et al. 2006, 2013) appear to be due to an introduction because microsatellite analyses suggest little genetic variation among the isolates (Al Adawi 2011, Harrington unpubl), and the MAT1 and MAT2 sequences of these isolates and the Indian isolate from pomegranate (Somasekhara 1999) are identical. Initially only one ITS sequence (ITS7b) was reported from the Oman and Pakistan populations, and this ITS haplotype was named *C. manginecans* (van Wyk et al. 2007). The *C. manginecans* haplotype and a second ITS haplotype (ITS6) were found on *Acacia magnum* in Indonesia, and the second haplotype was described as *C. acaciivora* (Tarigan et al. 2011), whose ITS6 sequence was found earlier on *Eucalyptus* in Indonesia (FJ236742, deposited in 2008 as *C. eucalypticola*). Consistent with our results with single-spore isolates yielding two or more ITS sequences (mostly sequences ITS3, ITS5, ITS6, ITS7b), Al Adawi et al. (2013) reported that both the ITS6 and ITS7b sequences were found within individual isolates from Pakistan and Oman by repeatedly conducting direct sequencing of PCR products. *C. manginecans* and *C. acaciivora* clearly are the same species, and analyses of mating-type genes places them among Brazilian isolates of *C. fimbriata*. The ITS3, ITS5, ITS6 and ITS7b haplotypes in Asia, Africa and Uruguay likely have their origin in Brazil, perhaps in cuttings of *Eucalyptus* (Ferreira et al. 2011, 2013; Harrington et al. 2011; Harrington 2013), but *Acacia* spp., mango, pomegranate and other hosts also appear to be susceptible.

The convenience of amplification and sequencing of rDNA regions has greatly facilitated modern fungal systematics and identification of species, but inferring evolutionary histories or delimiting species based on this region alone may be misleading (O'Donnell and Cigelnik 1997, Lachance et al. 2003, Smith et al. 2007, James et al. 2009, Linder and Banik 2011), and barcode identifications of species with ITS sequences (Nilsson et al. 2008, Schoch et al. 2012) may not be feasible for all fungal groups. Ambiguous alignments and a high degree of homoplasy are particular problems with ITS analyses in the *C. fimbriata* complex (Harrington et al. 2011). Problems also may arise if ITS sequences are combined with sequences of less variable genetic regions because the variability of the ITS sequences largely will determine the topology of the multigene trees. This was seen in the combined MAT1/MAT2/ITS trees and in combined ITS/translation elongation factor 1 alpha/beta-tubulin trees of the *C. fimbriata* complex (van Wyk et al. 2011b). No separate comparison of sequences of the translation elongation factor 1-alpha or beta tubulin with a range of representatives of the LAC has been presented, but the limited published data show that sequences from these regions vary little across the LAC (Harrington 2009, van Wyk et al. 2011b).

Limited variation in DNA sequences from coding regions supports the concept that the LAC is a recently derived group and that most of the genetic variation is among populations, which are geographically isolated due limited dispersal by insects and because they are soilborne (Harrington et al. 2011). Most branches of the combined MAT1/MAT2 tree may represent geographically isolated populations or

select genotypes moved by humans in vegetatively propagated material. Generally, populations of *C. fimbriata* show a wide range of aggressiveness to different cultivated crops and are not distinguished by host specialization or morphology (Baker et al. 2003, Marin et al. 2003, Thorpe et al. 2005, Harrington et al. 2011). Without distinguishing phenotype, localized lineages would not be considered phylogenetic species in the traditional sense (Harrington and Rizzo 1999). Paraphyletic species could result if an isolated population of unique lineage acquired a fixed phenotypic character of ecological importance. The MAT1/MAT2 dataset gave a robust tree that, along with host-range studies and interfertility tests, supports just three phylogenetic species (*C. cacaofunesta*, *C. platani*, *C. colombiana*) at the base of the LAC and outside *C. fimbriata*. There are likely other cryptic species in the LAC and perhaps within *C. fimbriata* as currently defined, but these species should be delimited by more than ITS sequences alone.

ACKNOWLEDGMENTS

The technical assistance of Maria Ferreira, Qiong Huang, Doug McNew and Leonardo Oliveira is greatly appreciated. Elizabeth Alvarez, Y.M. Somasekhara, Nilton Jungeira, Acelino Alfenas, Edson Furtado and numerous others kindly provided isolates or aided in the collection of isolates.

LITERATURE CITED

Al Adawi AO. 2011. Studies on *Ceratocystis* spp. associated with mango dieback in Oman [doctoral dissertation]. Univ. of Pretoria Press. 193 p.

———, Barnes I, Khan IA, Al Subhi AM, Al Jahwari AA, Deadman ML, Wingfield BD, Wingfield MJ. 2013. *Ceratocystis manginecans* associated with a serious wilt disease of two native legume trees in Oman and Pakistan. Australas Plant Pathol 42:179–193, doi:10.1007/s13313-012-0196-5

———, Deadman ML, Al Rawahi AK, Al Maqbali YM, Al Jahwari AA, Al Saadi BA, Al Amri IS, Wingfield MJ. 2006. Aetiology and causal agents of mango sudden decline disease in Sultanate of Oman. Eur J Plant Pathol 116:247–254, doi:10.1007/s10658-006-9056-x

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402, doi:10.1093/nar/25.17.3389

Baker CJ, Harrington TC, Krauss U, Alfenas AC. 2003. Genetic variability and host specialization in the Latin American clade of *Ceratocystis fimbriata*. Phytopathology 93:1274–1284, doi:10.1094/PHYTO.2003.93.10.1274

CAB International. 2005. *Ceratocystis fimbriata* [original text prepared by TC Harrington & C Baker, revised by T.C. Harrington]. In: Crop Protection Compendium. Wallingford, UK: CAB International (Revised CD version).

Chen SF, van Wyk M, Roux J, Wingfield MJ, Xie YJ, Zhou XD. 2013. Taxonomy and pathogenicity of *Ceratcystis* species on *Eucalyptus* trees in South China, including *C. chinaeucensis* sp. nov. Fungal Divers 58:267–279, doi:10.1007/s13225-012-0214-5

Engelbrecht CJB, Harrington TC. 2005. Intersterility, morphology, and taxonomy of *Ceratocystis fimbriata* on sweet potato, cacao, and sycamore. Mycologia 97:57–69, doi:10.3852/mycologia.97.1.57

———, ———, Alfenas AC, Suarez C. 2007. Genetic variation in populations of the cacao wilt pathogen, *Ceratocystis cacaofunesta*. Plant Pathol 56:923–933, doi:10.1111/j.1365-3059.2007.01735.x

———, ———, Steimel J, Capretti P. 2004. Genetic variation in eastern North American and putatively introduced populations of *Ceratocystis fimbriata* f. *platani*. Mol Ecol 13:2995–3005, doi:10.1111/j.1365-294X.2004.02312.x

Fateh FS, Kazmi MR, Ahmad I, Ashraf M. 2006. *Ceratocystis fimbriata* isolated from vascular bundles of declining mango trees in Sindh, Pakistan. Pak J Bot 38:1257–1259.

Ferreira EM, Harrington TC, Thorpe D J, Alfenas AC. 2010. Genetic diversity and interfertility among highly differentiated populations of *Ceratocystis fimbriata* in Brazil. Plant Pathol 59:721–735, doi:10.1111/j.1365-3059.2010.02275.x

Ferreira MA, Harrington TC, Alfenas AC, Mitzubuti ESG. 2011. Movement of genotypes of *Ceratocystis fimbriata* within and among *Eucalyptus* plantations in Brazil. Phytopathology 101:1005–1012, doi:10.1094/PHYTO-01-11-0015

———, ———, Gongora-Canul CC, Mafia RG, Zauza EAV, Alfenas AC. 2013. Spatial-temporal patterns of Ceratocystis wilt in Eucalyptus plantations in Brazil. Forest Pathol 43:153–164, doi:10.1111/efp.12013

Ganley ARD, Kobayashi T. 2007. Highly efficient concerted evolution in the ribosomal DNA repeats: total rDNA repeat variation revealed by whole-genome shotgun sequence data. Genome Res 17:184–191, doi:10.1101/gr.5457707

Harrington TC. 2000. Host specialization and speciation in the American wilt pathogen *Ceratocystis fimbriata*. Fitopatol Brasil 25:262–263.

———. 2009. The genus *Ceratocystis*. Where does the oak wilt fungus fit? In: Appel DN, Billings RF, eds. Proceedings of the 2nd National Oak Wilt Symposium. June 4–7, 2007. Austin, Texas: Texas Forest Service Publication 166. p 21–35.

———. 2013. Ceratocystis diseases. In: Gonthier P, Nicolotti G, eds. Infectious Forest Diseases. Wallingford, UK: CAB International. p 230–255.

———, McNew DL. 1997. Self-fertility and uni-directional mating-type switching in *Ceratocystis coerulescens*, a filamentous ascomycete. Curr Genet 32:52–59, doi:10.1007/s002940050247

———, Rizzo DM. 1999. Defining species in the fungi. In: Worrall JJ, ed. Structure and Dynamics of Fungal Populations. Dordrecht, Netherlands: Kluwer Press. p 43–71.

———, Thorpe D J, Alfenas AC. 2011. Genetic variation and variation in aggressiveness to native and exotic hosts among Brazilian populations of *Ceratocystis fimbriata*. Phytopathology 101:555–566, doi:10.1094/PHYTO-08-10-0228

James SA, O'Kelly MJT, Carter DM, Davey RP, van Oudenaarden A, Roberts IN. 2009. Repetitive sequence variation and dynamics in the ribosomal DNA array of *Saccharomyces cerevisiae* as revealed by whole-genome resequencing. Genome Res 19:626–635, doi:10.1101/gr.084517.108

Johnson JA, Harrington TC, Engelbrecht CJB. 2005. Phylogeny and taxonomy of the North American clade of the *Ceratocystis fimbriata* complex. Mycologia 97:1067–1092, doi:10.3852/mycologia.97.5.1067

Lachance MA, Daniel HM, Meyer W, Prasad GS, Gautam SP, Boundy-Mills K. 2003. The D1/D2 domain of the large-subunit rDNA of the yeast species *Clavispora lusitaniae* is unusually polymorphic. FEMS Yeast Res 4:253–258, doi:10.1016/S1567-1356(03)00113-2

Linder DL, Banik MT. 2011. Intragenomic variation in the ITS rDNA region obscures phylogenetic relationships and inflates estimates of operational taxonomic units in genus *Laetiporus*. Mycologia 103:731–740, doi:10.3852/10-331

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. J Phytopathol 151:395–405, doi:10.1046/j.1439-0434.2003.00738.x

Nilsson RH, Kristiansson E, Ryberg M, Hallenberg N, Larsson KH. 2008. Intraspecific ITS variability in the kingdom Fungi as expressed in the international sequence databases and its implications for molecular species identification. Evol Bioinform 4:199–201.

Ocasio-Morales RG, Tsopelas P, Harrington TC. 2007. Origin of *Ceratocystis platani* on native *Platanus orientalis* in Greece and its impact on natural forests. Plant Dis 91:901–904, doi:10.1094/PDIS-91-7-0901

O'Donnell K, Cigelnik E. 1997. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. Mol Phylogenet Evol 7:103–116, doi:10.1006/mpev.1996.0376

Poussio GB, Kazmi MR, Aken C, Fateh GS. 2010. First record of *Ceratocystis fimbriata* associated with shisham (*Dalbergia sissoo*) decline in Pakistan. Australas Plant Dis Notes 5:63–65, doi:10.1071/DN10023

Rodas CA, Roux J, van Wyk M, Wingfield BD, Wingfield MJ. 2008. *Ceratocystis neglecta* sp. nov., infecting *Eucalyptus* trees in Colombia. Fungal Div 28:73–84.

Ronquist F, Huelsenbeck JP. 2003. MRBAYES 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19:1572–1574, doi:10.1093/bioinformatics/btg180

Rossetto CJ, Ribeiro IJA. 1990. Seca da mangueira XII. Recomendações de controle. Rev Agric 65:173–180.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque A, Chen W, Fungal Barcode Consortium. 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. Proc Natl Acad Sci 109:6241–6246, doi:10.1073/pnas.1117018109

Silveira SF, Harrington TC, Mussi-Dias V, Engelbrecht CJB, Alfenas AC, Silva CR. 2006. *Annona squamosa*, a new host of *Ceratocystis fimbriata*. Fitopatol Brasil 31:394–397, doi:10.1590/S0100-41582006000400010

Simon UK, Weiss M. 2008. Intragenomic variation of fungal ribosomal genes is higher than previously thought. Mol Biol Evol 25:2251–2254, doi:10.1093/molbev/msn188

Smith ME, Douhan GW, Rizzo DM. 2007. Intraspecific and intrasporocarp ITS variation of ectomycorrhizal fungi as assessed by rDNA sequencing of sporocarps and pooled ectomycorrhizal roots from *Quercus* woodland. Mycorrhiza 18:15–22, doi:10.1007/s00572-007-0148-z

Somasekhara YM. 1999. New record of *Ceratocystis fimbriata* causing wilt of pomegranate in India. Plant Dis. 83:400–400, doi:10.1094/PDIS.1999.83.4.400B

Steimel J, Engelbrecht CJB, Harrington TC. 2004. Development and characterization of microsatellite markers for the fungus *Ceratocystis fimbriata*. Mol Ecol Notes 4:215–218, doi:10.1111/j.1471-8286.2004.00621.x

Swofford DL. 2002. PAUP* 4.0b10a: phylogenetic analysis using parsimony (and other methods). Sunderland, Massachusetts: Sinauer Associates.

Tarigan M, Roux J, van Wyk M, Tjahjono B, Wingfield MJ. 2011. A new wilt and dieback disease of *Acacia mangium* associated with *Ceratocystis manginecans* and *C. acaciivora* sp. nov. in Indonesia. S Afr J Bot 77:292–304, doi:10.1016/j.sajb.2010.08.006

Thorpe DJ, Harrington TC, Uchida JY. 2005. Pathogenicity, internal transcribed spacer-rDNA variation, and human dispersal of *Ceratocystis fimbriata* on the family Araceae. Phytopathology 95:316–323, doi:10.1094/PHYTO-95-0316

van Wyk M, van der Merwe NA, Roux J, Wingfield BD, Kamgan GN, Wingfield MJ. 2006. Population genetic analyses suggest that the *Eucalyptus* pathogen *Ceratocystis fimbriata* has been introduced into South Africa. S Afr J Sci 102:259–263.

———, Al Adawi AO, Khan IA, Deadman ML, Al Jahwari AA, Wingfield BD, Ploetz R, Wingfield MJ. 2007. *Ceratocystis manginecans* sp. nov., causal agent of a destructive mango wilt disease in Oman and Pakistan. Fungal Divers 27:213–230.

———, Wingfield BD, Al Adawi AO, Rossetto CJ, Ito MF, Wingfield MJ. 2011a. Two new *Ceratocystis* species associated with mango disease in Brazil. Mycotaxon 117:381–404, doi:10.5248/117.381

———, ———, Wingfield MJ. 2011b. Four new *Ceratocystis* spp. associated with wounds on *Eucalyptus*, *Schizolobium* and *Terminalia* trees in Ecuador. Fungal Divers 40:103–117.

———, ———, Marin M, Wingfield MJ. 2010. New *Ceratocystis* species infecting coffee, cacao, citrus and native trees in Colombia. Fungal Divers 40:103–117, doi:10.1007/s13225-009-0005-9

———, ———, Mohali M, Wingfield MJ. 2009. *Ceratocystis fimbriatomima*, a new species in the *C. fimbriata* sensu lato complex isolated from *Eucalyptus* trees in Venezuela. Fungal Divers 34:173–183.

———, Roux J, Nkuekam GK, Wingfield BD, Wingfield MJ. 2012. *Ceratocystis eucalypticola* sp. nov. from *Eucalyptus* in South Africa and comparison to global isolates from this tree. IMA fungus. 3:45–58, doi:10.5598/imafungus.2012.03.01.06

Witthuhn RC, Harrington TC, Wingfield BD, Steimel J, Wingfield MJ. 2000. Deletion of the *MAT-2* mating-type gene during unidirectional mating-type switching in *Ceratocystis*. Curr Genet 38:48–52, doi:10.1007/s002940000131

Zauza EAV, Alfenas AC, Harrington TC, Mizubuti ES, Silva JF. 2004. Resistance of Eucalyptus clones to *Ceratocystis fimbriata*. Plant Dis 88:758–760, doi:10.1094/PDIS.2004.88.7.758