



Molecular phylogeny of *Colletotrichum* (Sordariomycetes: Glomerellaceae) inferred from multiple gene sequences

H. Chen^{1,2}, Y. Ma², W.F. Zhang¹, T. Ma^{1,2} and H.X. Wu^{1,2}

¹International Fungal Research & Development Centre,
The Research Institute of Resource Insects, Chinese Academy of Forestry,
Bailongsi, Kunming, China

²The Key Laboratory of Cultivating and Utilization of Resources,
Insects of State Forestry Administration, Kunming, China

Corresponding author: H.X. Wu
E-mail: haixiawu1983@126.com

Genet. Mol. Res. 14 (4): 13649-13662 (2015)

Received February 4, 2015

Accepted May 29, 2015

Published October 28, 2015

DOI <http://dx.doi.org/10.4238/2015.October.28.27>

ABSTRACT. The genus *Colletotrichum* contains many important plant pathogens mainly distributed in tropical areas. Previous studies that utilized morphology or single genes have failed to resolve the phylogenetic relationships among the species. In this study, sequences of β -tubulin, 28S ribosomal DNA, and ITS region from nine species were analyzed separately and combined to establish a fast method to infer the phylogeny of *Colletotrichum* using maximum parsimony, maximum likelihood, and Bayesian inference methods. The tree topologies inferred from the combined data set received higher bootstrap and posterior probability support than those inferred from the individual data sets. Obtained phylogenies highly supported *C. capsici* as the earliest diverging lineage followed by *C. nymphaeae*. The remaining seven species clustered into two distinct clades. Clade 1 consists of two monophyletic subclades: *C. circinans*, *C. trichellum*, and *C. caudatum* form one subclade and three accessions of *C. dematium* form another

subclade. In Clade 2, *C. incarnatum* is in the basal-most clade. Three accessions of *C. musae* and *C. caricae* form a strongly supported clade indicating their close relationship. Spore shape analysis reveals an interesting evolutionary trend in the spore shape from acute- to obtuse-ended conidia and from curved to straight conidia in the sampled group of species. A quick and reliable way to infer the phylogeny of *Colletotrichum* based on combined DNA sequence data is presented in this paper.

Key words: *Colletotrichum*; β -tubulin; 28S rDNA; ITS; Molecular phylogeny

INTRODUCTION

The genus *Colletotrichum* belongs to the family Glomerellaceae (Sordariomycetes, Pezizomycotina, Ascomycota), which includes some of the most important plant pathogens worldwide, causes economically significant diseases in a wide range of hosts, including cereals, legumes, vegetables, perennial crops, and fruits (Moriwaki et al., 2002; Than et al., 2008; Manamgoda et al., 2013). Although mostly located in the tropics, it has a worldwide distribution (Wharton et al., 2001; Wharton and Diéguez-Uribeondo, 2004; Hyde et al., 2009; Prihastuti et al., 2010). *Colletotrichum* was first described by Tode in 1790 under the name *Vermicularia* but was later established as *Colletotrichum* by Corda in 1837 (Wharton et al., 2001; Moriwaki et al., 2002). Despite the original description and excellent illustrations provided by Corda, the genus has been described under several different names such as *Blennorella* and *Gloeosporium*. The names *Colletotrichum* and *Blennorella* were used interchangeably for several centuries for a range of species now accepted in the genus *Colletotrichum* (Sutton, 1992).

Traditionally, identification of *Colletotrichum* species has been based on morphological characters, such as size and shape of conidia and appressoria, existence of setae, the host origin, the teleomorph state, and cultural characteristics (von Arx, 1957; Smith and Black, 1990; Than et al., 2008; Peng et al., 2012, 2013). These characters were used to assign more than 750 *Colletotrichum* species to 11 species, with many taxa treated as synonyms of *C. gloeosporioides* and *C. dematium* (von Arx, 1957; Cai et al., 2009; Moura et al., 2014).

Many problems associated with the taxonomy of *Colletotrichum* species and their evolutionary relationships remain unresolved (Sreenivasaprasad and Talhinhos, 2005; Than et al., 2008). However, molecular biology provides new methods that can be utilized for species delimitation and defining of inter- and intraspecific relationships (Bruns et al., 1991; Lee et al., 2007). The most significant advances in phylogeny have been obtained from approaches based on the analysis of DNA and random amplified polymorphic DNA sequences as well as restriction fragment length polymorphism since the molecular data are not directly influenced by environmental factors (Welsh and McClelland, 1990; Williams et al., 1990; Faisal et al., 2013).

Most fungal phylogenetic studies utilized sequences from the ribosomal gene cluster, and the ribosomal DNA (rDNA) data of *Colletotrichum* species have been used to infer their phylogenetic relationships and for preliminary identification (Sherriff et al., 1994; Cai et al., 2009). In particular, internal transcribed spacer (ITS) sequences have been proven to be useful in studying phylogenetic relationships of *Colletotrichum* species because of their comparative

variability (Sherriff et al., 1994; Sreenivasaprasad et al., 1994, 1996; Moriwaki et al., 2002; Photita et al., 2005), and sequence analysis of β -tubulin gene was used to resolve phylogenetic relationships among *Colletotrichum* species complexes (Sreenivasaprasad and Talhinhas, 2005; Lee et al., 2007; Peng et al., 2013). Other gene fragments such as introns have also been utilized to evaluate the diversity of this genus (Guerber et al., 2003; Peres et al., 2008).

In the present paper, we assessed the interspecific relationships among nine species of *Colletotrichum* collected from different countries using sequences of β -tubulin and 28S rDNA genes and the ITS region. The objectives of this research were: 1) to identify a quick method to infer the phylogenetic relationships in the genus; 2) to provide a reliable phylogenetic framework for nine species of *Colletotrichum*; and 3) to explore the implications of the molecular phylogeny and explore the evolutionary trend of the conidia in *Colletotrichum*, which provide a foundation for accurate diagnosis of pathogens and development of appropriate disease management.

MATERIAL AND METHODS

Different strains of nine species of *Colletotrichum* were collected from eight countries from different parts of the world (Table 1). Cultures were transferred aseptically to a potato dextrose agar (PDA) (Oxoid Ltd., Hampshire, UK) and incubated at 25°C (Than et al., 2008). Morphological characters of the cultures were examined under a compound microscope (Nikon 80i, Tokyo, Japan) and three subcultures from each isolate were prepared and used in subsequent analysis. Detailed information of the strains is shown in Table 1. All the isolates were subcultured and stored at room temperature at the International Fungal Research and Development Centre (Kunming, China).

Table 1. Sample information of nine species of *Colletotrichum* used in this study.

Name	Site	Identifier/Supplier	Host	Original code
<i>C. musae</i>	Hong Kong, China	Hellen	<i>Musa</i> sp	6,636
<i>C. musae</i>	Mae Taeng, Thailand	PoPo Than	<i>Musa</i> sp	10,854
<i>C. musae</i>	Mae Taeng, Thailand	PoPo Than	<i>Musa</i> sp	10,853
<i>C. dematium</i>	Wellington, New Zealand	J.M. Dingley	<i>Crinum</i> sp	10,413
<i>C. dematium</i>	Auckland, New Zealand	J.M. Dingley	<i>Aucuba</i> sp	10,400
<i>C. dematium</i>	Wellington, New Zealand	G. Laundon	<i>Ranunculus repens</i> L.	10,403
<i>C. capsici</i>	Sansai, Thailand	PoPo Than	<i>Capsicum annum</i>	10,858
<i>C. circinans</i>	Japan	NIAS	<i>Allium cepa</i> L.	10,375
<i>C. caudatum</i>	Queensland, Australia	BRIP	<i>Aucuba</i> sp	10,448
<i>C. nymphaeae</i>	Netherlands, Herwijnen	H.A. van Kesteren	<i>Nupharluteum</i>	10,655
<i>C. incarnatum</i>	East Africa, Boukoko	J. Nicot	<i>Theobroma cacao</i>	10,732
<i>C. caricae</i>	Laboulaye, Argentina	CBS	<i>Ficus carica</i>	10,742
<i>C. trichellum</i>	Japan	NIAS	<i>Hedera helix</i>	10,378

DNA extraction

Each isolate was grown on PDA for 7-10 days at 25°C. Mycelium was scraped from the surface of the culture plate and used for DNA extraction. Total DNA was extracted using a Biospin Kit (Bioer Technology Co., Ltd., Hangzhou, China) according to the manufacturer protocol. DNA extracts were preserved at -20°C until their use.

PCR and sequencing

Amplification reactions were performed in a final volume of 50 μ L containing 25 ng DNA, 5 μ L 10X PCR buffer (Invitrogen, Carlsbad, CA, USA), 2.5 mM MgCl₂, 0.5 mM dNTPs, 1 mM of each primer, 1 U Taq polymerase (Promega, Madison, WI, USA), and 29 μ L ultrapure sterile water. The primers used for amplification and their sequences are listed in Table 2.

Table 2. Primers used for amplification of three DNA regions used in this study with references.

Gene	Name	Primer sequence (5'-3')	Reference
β -tubulin	T1	AACATGCGTGAGATTGTAAGT	O'Donnell and Cigelnik (1998)
	β t2b	ACCCTCAGTGTAGTGACCCTTGCC	Lee et al. (2007)
28S rDNA	23032/1	TGAAAAGAACTTTGAAAAGAGA	Johnston and Jones (1997)
	28487	CTTGGTCCGTGTTCAAGACG	Johnston and Jones (1997)
ITS	ITS4	TCCTCCGCTTATTGATATGC	White et al. (1990)
	ITS5	GGA AGT AAA AGT CGT AAC AAG G	White et al. (1990)

DNA amplifications were carried out in a Bio-RAD MyCycler thermal cycler (Hercules, CA, USA) using different thermal regime for different regions. Amplification of β -tubulin sequences consisted of initial denaturation at 95°C for 4 min, 35 cycles of denaturation at 64°C for 1 min, annealing at 72°C for 2 min, and extension at 72°C for 7 min, and ended by cooling at 4°C. Amplification cycles of 28S rDNA consisted of one cycle of initial denaturation at 94°C for 1 min, 35 cycles of denaturation at 94°C for 1 min, annealing at 55°C for 1 min, and extension and 72°C for 1 min, followed by a final step at 72°C for 3 min. The ITS regions were amplified using the following cycle conditions: initial denaturation at 94°C for 5 min, 30 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 30 s, and extension at 72°C for 50 s, followed by one cycle at 72°C of 7 min. PCR products were purified with a QIAquick PCR purification Kit (Qiagen, Valencia, CA, USA). Sequencing of both strands was carried out on a CEQ2000XL automated sequencer (Beckman Coulter, Inc., Fullerton, CA, USA) with the same primers used for amplification. The resulting sequences were assembled using Bioedit version 7.0.5.3 (Hall, 1999) and deposited in GenBank (accession No. GQ369588-GQ379701) (Table 3).

Phylogenetic analyses

Alignments of the individual data matrices were generated using similarity index calculated at the nucleotide level with ClustalW version 1.81 implemented in the Lasergene DNASTAR software package (Thompson et al., 1997). The data sets were subjected to Bayesian phylogenetic inference, maximum likelihood (ML) analysis, and equal weights parsimony (MP) analyses to estimate the tree topology. For Bayesian inference, metropolis-coupled Markov Chain Monte Carlo simulations were carried out with MrBayes 3.1.2 (Huelsenbeck and Ronquist, 2001) to estimate the posterior probability distributions of the tree topologies. The best evolutionary models for ML analyses were generated using hierarchical likelihood ratio test estimated by Modeltest 3.06 (Posada and Crandall, 1998). Statistical support for each node was evaluated by bootstrap analysis (Felsenstein, 1985). Parsimony analyses were conducted in the PAUP 4.0b software (Swofford, 2001) using heuristic search algorithm (Hillis

et al., 1996) with 1000 replicates of random-addition-sequence (Wu et al., 2011, 2014). The resulting trees were rooted using outgroup taxon. Tree visualization and drawing were carried out in TreeView version 1.5.2 (Page, 1996).

RESULTS

Characteristics of colonies cultured on PDA

The differences in growth speed in various *Colletotrichum* isolates were evident after 7 days of culture at 25°C (Figure 1). At this temperature, all isolates were clearly differentiated into two main types. The first type comprised the slow growing *C. nymphaeae*, *C. incarnatum*, and *C. trichellum* (the colonies were less than 5 cm in diameter). The other type included the fast growing *C. circinans* as well as the isolates of *C. musae*, *C. dematium*, *C. capsici*, *C. caudatum*, and *C. caricae* (the colonies were more than 5 cm in diameter).

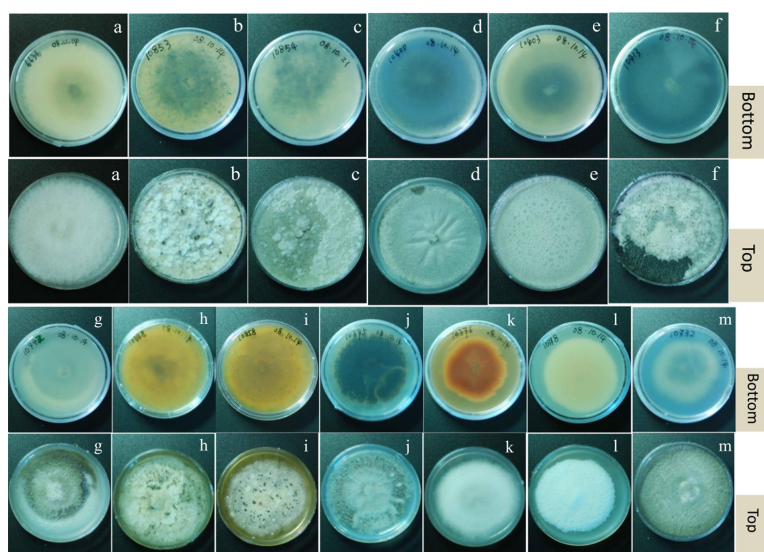


Figure 1. Colony characteristics of nine species of *Colletotrichum* cultured on PDA 7 days after inoculation. **a.-c.** *C. musae*; **d.-f.** *C. dematium*; **g.** *C. caricae*; **h.** *C. caudatum*; **i.** *C. capsici*; **j.** *C. circinans*; **k.** *C. nymphaeae*; **l.** *C. trichellum*; **m.** *C. incarnatum*.

Colony characteristics revealed a wide range of variation between different isolates even in the same species (Figure 1). All isolates could be separated into two groups. One group was characterized by upper side white, grey, and orange, cottony, and zonate and reverse side cream, yellow, or orange; some isolates were darker in the center (6636; 10,655) or contained black spots (10,858; 10,378). The other group was characterized by upper side olive-green, white, and gray, felty; reverse side cream, yellow-black, or gray-olive, mostly zonate, and darker in the center (10,400; 10,403; 10,375; 10,372; 10,448) or having black spots (10,853; 10,854; 10,732; 10,742). Morphological characters of the species are summarized in Table 3.

Table 3. Morphological characters and GenBank accession Nos. of *Colletotrichum* isolates used in the study.

No.	Name	Morphologic characters	GenBank accession No.		
			β -tubulin	28S rDNA	ITS
1	<i>C. musae</i>	White, hairy, flat, superficial, reverse cream, darker at center, zonate	GQ369602	GQ379688	GQ369588
2	<i>C. musae</i>	White, felty, powdery, fruiting bodies black with orange exudates, reverse cream, darker at center	GQ369603	GQ379699	GQ369589
3	<i>C. musae</i>	Grey, felty, powdery, fruiting bodies black with brown and orange exudates, reverse cream	GQ369604	GQ379700	GQ369590
4	<i>C. dematium</i>	White, felty, hairy, superficial, radiate, reverse gray-olive	GQ369610	GQ379693	GQ369591
5	<i>C. dematium</i>	Grey, felty, reverse gray-olive	GQ369611	GQ379695	GQ369592
6	<i>C. dematium</i>	White, felty, reverse cream, darker at center	GQ369608	GQ379692	GQ369593
7	<i>C. capsici</i>	White with black spot, cottony, zonate, reverse yellow with black spots	GQ369609	GQ379701	GQ369594
8	<i>C. circinans</i>	White with black spot, felty, edge not smooth, reverse black	GQ369614	GQ379690	GQ369595
9	<i>C. trichellum</i>	White, cottony, thick, reverse cream	GQ369613	GQ379691	GQ369597
10	<i>C. caudatum</i>	Grey with black spot, felty, edge smooth, reverse yellow, darker at center	GQ369612	GQ379694	GQ369598
11	<i>C. nymphaeae</i>	White, cottony, reverse cream, zonate, darker at center	GQ369607	GQ379697	GQ369599
12	<i>C. incarnatum</i>	Grey, felty, superficial, reverse cream, zonate, darker at center	GQ369606	GQ379689	GQ369600
13	<i>C. caricae</i>	Grey with white, felty, superficial, reverse cream, zonate	GQ369605	GQ379698	GQ369601

Molecular data sets

The data set of aligned β -tubulin sequences was 739 bp long, of which 363 (49.1%) sites were variable and 198 (26.8%) sites were parsimony informative. The aligned ITS sequences were 699 bp long, of which 294 (42.1%) sites were variable and 63 (9.0%) sites were parsimony informative. The average genetic (Tamura-Nei) distance within the in-group taxa was 44.1% for β -tubulin and 16.8% for ITS data sets. There was no excess of AT residues in β -tubulin (46.7%) data set, whereas the genomic base composition was nearly balanced in the ITS (47.8%) data set. The mean transition-to-transversion ratio (ti/tv ratio) was 1.09 and 1.29 for β -tubulin and ITS sequences, respectively.

In the 253-bp-long 28S rDNA data set, 17 (6.7%) sites were variable and 11 (4.35%) sites were parsimony-informative. The average genetic distance between in-group taxa was 3.2%. The genomic base composition was slightly GC-rich (40.7%) and the mean ti/tv ratio was 1.40.

Table 4 shows sequence characteristics of the three datasets independently and combined. Comparison of the four data sets revealed that β -tubulin sequences had the largest proportions of variable and informative sites and higher Tamura-Nei distance (2.82-fold) than the other data sets. Sequences of the 28S rDNA region showed the lowest percentage of variable and informative sites, although the mean ti/tv ratio was higher than in other two sequenced regions. In addition, the sequence divergence estimate was larger in ITS than in 28S rDNA.

Table 4. Summary of statistics for each individual and combined phylogenetic datasets.

Gene name	Variable site	Pi*	Total	AT %	ti/tv*	Tamura-Nei
β -tubulin	363 (49.1%)	198 (26.8%)	739	46.7	1.09	0.441
ITS	294 (42.1%)	63 (9.0%)	699	47.8	1.29	0.168
28S rRNA	17 (6.7%)	11 (4.35%)	253	40.7	1.40	0.032
Combined	674 (39.9%)	309 (13.90%)	1691	47.3	1.45	0.156

*Pi = parsimony informative site; ti/tv = mean transition-to-transversion ratio.

Phylogenetic analysis

The difference in the genetic base composition between the sequences of the three regions

is reflected in the evolutionary models found to best fit each data set. Thus, the simple HKY and SYM models were chosen for the nearly balanced genetic base composition of β -tubulin and ITS data sets, whereas somewhat more complex GRT model was chosen for the 28S rDNA dataset.

The majority rule consensus tree from the combined Bayesian analyses is presented in Figure 2. From all trees generated from molecular data used three different analysis strategies. The MP, ML, and Bayesian analyses of the three dataset produced nearly identical tree topologies; they differed in bootstrap support (BS) values for some of the branches. The trees inferred from the combined dataset resolved *C. capsici* as the earliest diverging lineage followed by *C. nymphaeae* with high support (Bayesian posterior probability [PP] = 1; MP BS = 100%; ML BS = 64%). The remaining seven species formed two distinct clades. Clade 1 comprises two monophyletic clades; one includes *C. circinans*, *C. trichellum*, and *C. caudatum*, and the other three strains of *C. dematium* (PP = 1; MP BS = 93%; ML BS = 85%). Clade 2 includes the basal *C. incarnatum* and a monophyletic clade comprised of three strains of *C. musae* and *C. caricae* (Figure 2).

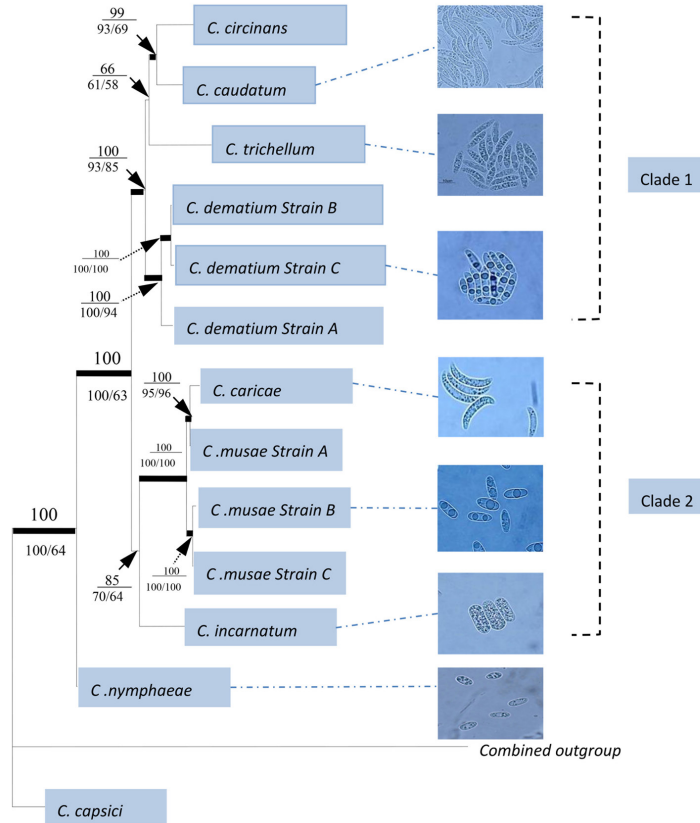


Figure 2. Majority-rule consensus tree resulting from the mixed-model Bayesian analysis of combined dataset (1691 bp; model = SYM + G) illustrating the relationship between nine species of *Colletotrichum*; *Physcia ainophila* was used as outgroup. Branches represented with dashed lines were not recovered in the Bayesian consensus tree topology and are based on the maximum likelihood tree topology. Numbers above internodes indicate Bayesian posterior probabilities; numbers below internodes indicate non-parametric bootstrap proportions for parsimony analysis (left) and likelihood analysis (right). Thickened branches indicate Bayesian posterior probabilities ≥ 0.95 . The photos linked with species name by blue dotted lines represent the morphology of spores.

DISCUSSION

Which DNA sequence data is prior to phylogenetic analysis in the genus *Colletotrichum*?

Forty four strains of *Colletotrichum* from various areas in Korea were divided into three genotypes, *C. gloeosporioides*, *C. psidii*, and *C. acutatum*, using different molecular markers. Among these markers, sequences of the protein coding gene β -tubulin have been used to differentiate strains and species of *Colletotrichum* (O'Donnell et al., 1998; Lee et al., 2007) and to carry out phylogenetic analyses of the *C. acutatum* species complex (Sreenivasaprasad and Talhinhas, 2005). Thus, apple and pepper strains of *C. acutatum* were resolved in two clades based on β -tubulin sequence data (Lee et al., 2007). The sequences of β -tubulin are 3.5-fold more informative than mitochondrial small subunit ribosomal RNA sequences and therefore have been proposed as a suitable marker for studying close relationships in *Fusarium* (O'Donnell et al. 1998). In the present study, the phylogenetic relationships between nine species complexes of *Colletotrichum* were not clearly resolved by using β -tubulin sequence data although the sequences had the highest percentage (26.8%) of parsimony-informative sites (Table 4). The clade containing *C. incarnatum* and *C. nymphaeae* did not receive the same support in Bayesian and MP analyses and was unresolved in ML analysis. Furthermore, the positions of *C. circinans* and *C. trichellum* remained unresolved in the phylogenetic tree inferred from the β -tubulin dataset (Figure 3).

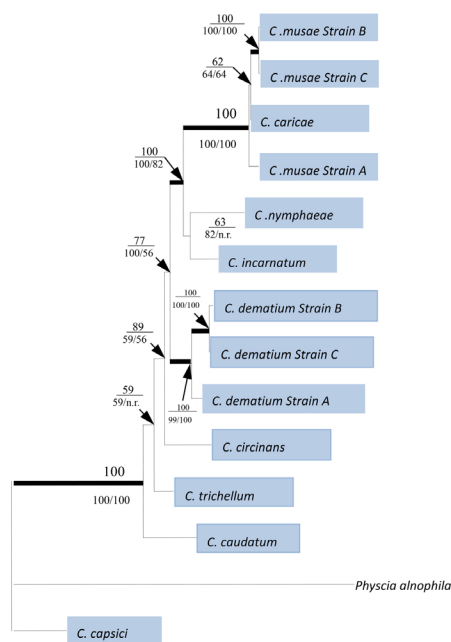


Figure 3. Majority-rule consensus tree resulting from Bayesian analysis of β -tubulin sequence data (739 bp; model = HKY + G) illustrating the relationship between nine species of *Colletotrichum*; *Physcia ainophila* was used as outgroup. Branches represented with dashed lines were not recovered in the Bayesian consensus and are based on the maximum likelihood tree topology. Numbers above internodes indicate Bayesian posterior probabilities; numbers below internodes indicate non-parametric bootstrap proportions for the parsimony analysis (left) and likelihood analysis (right). Nodes not recovered in one of the three analyses are indicated by “n.r.”. Thickened branches indicate Bayesian posterior probabilities ≥ 0.95 .

Several recent publications have reported on the use of rDNA sequence data in species delimitation in *Colletotrichum* (Sherriff et al., 1994; Crouch and Beirn, 2009). The analysis of sequence data from the D2 region of the large subunit rDNA resolved 16 distinct groups among the New Zealand fruit-rotting isolates including the accepted species *C. coccodes*, *C. musae*, *C. orbiculare*, *C. acutatum*, and *C. gloeosporioides*, which were separated into four broad groups, whereas the relationships of *C. coccodes* remained unresolved (Sherriff et al., 1994; Crouch and Beirn, 2009). The rDNA sequence data were unable to resolve the apparently stable morphological and cultural subgroups recognized within the sequenced groups (Johnston and Jones, 1997). However, Sherriff et al. (1994) showed that D2 region of the rDNA provided sufficient variation to discriminate between the groups of *Colletotrichum* defined by the combined D1 and D2 regions. In the present study, the relationship between three accessions of *C. dematium*, *C. capsici*, and *C. caudatum* remained unresolved, and the clade with *C. circinans* and *C. trichellum* was weakly supported in topologies based on the 28S rDNA dataset. These results indicate that rDNA sequences alone are not suitable for phylogenetic studies in the genus *Colletotrichum* (Figure 4).

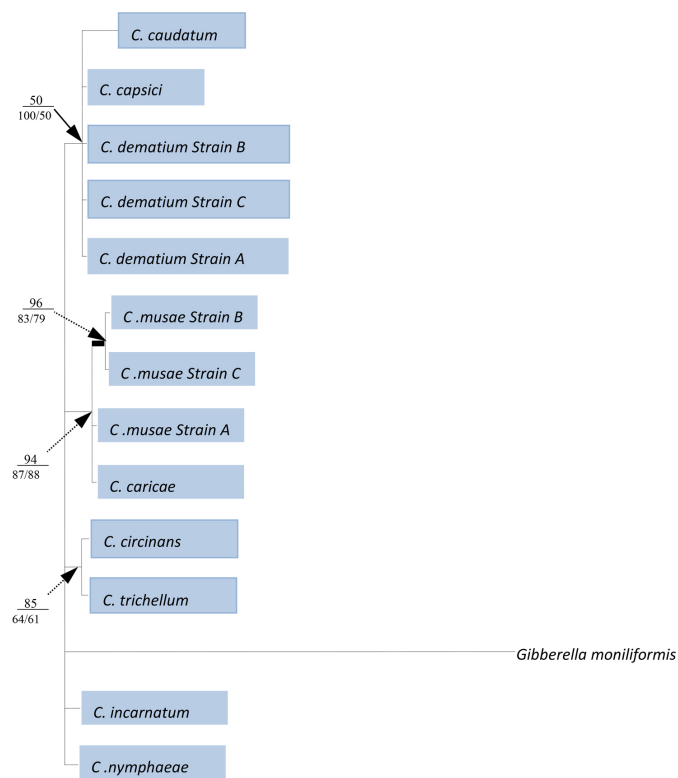


Figure 4. Majority-rule consensus tree resulting from Bayesian analysis of 28S rDNA sequence data (253 bp; model = GTR + G) illustrating the relationships between nine species of *Colletotrichum*; *Gibberella moniliformis* was used as outgroup. Branches represented with dashed lines were not recovered in the Bayesian consensus and are based on the maximum likelihood tree topology. Numbers above internodes indicate Bayesian posterior probabilities; numbers below internodes indicate non-parametric bootstrap proportions for the parsimony analysis (left) and likelihood analysis (right). Thickened branches indicate Bayesian posterior probabilities ≥ 0.95 .

ITS sequence alone should not be used to infer evolutionary relationships in the genus *Colletotrichum*, nor should they be employed to describe novel *Colletotrichum* taxa. It is well established that the ITS sequences often generate poorly supported, unresolved phylogenetic trees in the genus, particularly when distant taxa are sampled (Crouch and Beirn, 2009). In the present study, in the phylogenetic tree inferred from the ITS sequence dataset, three strains of *C. musae* remained unresolved, whereas *C. capsici* and *C. nymphaeae* formed a monophyletic, strongly supported group (MP BS = 100%), indicating that the latter two species exhibited high sequence identity in the ITS region. Furthermore, the support of many branches was below 70% and the topology was poorly resolved (Figure 5).

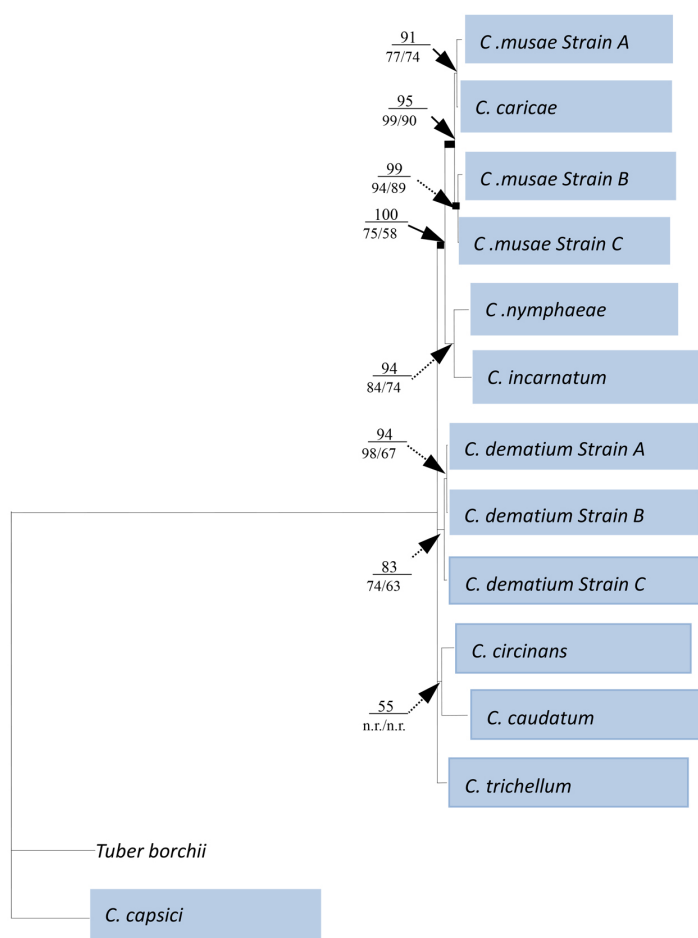


Figure 5. Majority-rule consensus tree resulting from Bayesian analysis of ITS sequence data (699 bp; model = K80 + G) illustrating the relationship between nine species of *Colletotrichum*; *Tuber borchii* was used as outgroup. Branches represented with dashed lines were not recovered in the Bayesian consensus and are based on the maximum likelihood topology. Numbers above internodes indicate Bayesian posterior probabilities; numbers below internodes indicate non-parametric bootstrap proportions for the parsimony analysis (left) and likelihood analysis (right). Nodes not recovered in one of the three analyses are indicated by “n.r.” Thickened branches indicate Bayesian posterior probabilities ≥ 0.95 . The species in blue are all sequenced successfully for the three regions.

Similar to the topology inferred from β -tubulin and ITS sequences, the combined sequence data supported *C. capsici* as the most basal and thus the earliest diverging lineage. The clade containing *C. circinans*, *C. caudatum*, and *C. trichellum* was recovered in the Bayesian, ML, and MP analyses of the combined dataset, whereas this clade was not resolved in any of the topologies inferred from the single datasets. The bootstrap supports of each node in the phylogenetic trees were higher in Bayesian, MP, and ML analysis of the combined dataset (Figure 2). The results obtained herein suggest that the inter-species relationship of *Colletotrichum* are better resolved by the combined β -tubulin, ITS, and 28S rRNA sequence data than by the individual sequence dataset.

Phylogeny

C. capsici and *C. dematium* have often been confused because both species produce similar falcate conidia and are isolated from many types of plants (Moriwaki et al., 2002; Sutton, 1992). However, *C. capsici* and *C. dematium* are clearly divergent lineages as supported by the tree sequence datasets. β -Tubulin, ITS, and combined datasets highly supported the basal position of *C. capsici* in the tree topologies (PP = 1; ML BS = 100%; ML BS = 100%), indicating that *C. capsici* is the most primitive member of the nine species of *Colletotrichum*.

β -tubulin-, ITS-, and 28S rDNA- inferred tree topologies support the sister relationship of *C. nymphaeae* and *C. incarnatum*, while *C. nymphaeae* is sister to *C. capsici* in the phylogeny. In contrast, *C. incarnatum* was sister to *C. musae* albeit with low support in all tree topologies inferred from the combined datasets using the three algorithms (PP = 0.85; MP BS = 70%; ML BS = 64%), suggesting their close relationship (Figure 2). The positions of the latter two species remain unresolved.

The analyses of the combined ITS and 28S rDNA D2 sequence data resolved *C. circinans* closely related to *C. trichellum* and both only distantly related to *C. caudatum* (Moriwaki et al., 2002). In the present study, *C. circinans* and *C. trichellum* formed a weakly supported clade based on 28S rDNA sequence data, while *C. circinans* was resolved sister to *C. caudatum* in the ITS-inferred tree topology. The analyses of β -tubulin sequences did not resolve any of the two sister relationships, whereas in the topologies inferred from the combined dataset, *C. circinans*, *C. trichellum*, and *C. caudatum* formed a monophyletic clade, which was sister to the three populations of *C. dematium* with high support (PP = 1; MP BS = 93%; ML BS = 85%).

In the phylogenetic trees based on combined sequence data, *C. caricae* and three strains of *C. musae* formed a clade that received high support (PP = 1; MP BS = 100%; ML BS = 100%), suggesting their close relationship. Among the three strains of *C. musae*, two of which originate from Mae Taeng, Thailand (Table 1) formed a clade, whereas strain A originating from Hongkong was sister to *C. caricae* in all tree topologies. The two species have a more derived position on the tree topologies inferred from β -tubulin and combined sequence data, indicating that they are the youngest members among the nine species.

The spore shape revealed some interesting trends in different species. Thus, the spores of *C. acutatum* isolated from *Lupinus* spp and other hosts had obtuse ends (in 0 to 28% of spores), one obtuse and one acute end (in 40 to 100% of spores) and acute ends (in 0 to 60% of spores), while the spores of *C. gloeosporioides* isolates have obtuse ends (in 68 to 100% of spores) (Talhinhas et al., 2002). The isolates (236) comprising 25 species of *Colletotrichum* in

Japan were separated into three phylogenetic groups (PG): PG1 produces conidia with an acute end, PG2 produces conidia with an obtuse end, and PG3 of *C. theae-sinensis* has small conidia. The evolutionary trend of the conidial shape is toward small conidia, and from conidia with obtuse ends toward conidia with acute ends (Moriwaki et al., 2002).

In the present study, *Colletotrichum* species produced visible spores in multiple shapes, except *C. circinans* and *C. capsici*, these two species produced only one shape. Mass fusiform spores of *C. dematium* usually have both ends acute with one side curved (Figure 2), and conidia of *C. caricae* and *C. caudatum* are mostly falcate with acute ends. Ovoid conidia are always found in *C. nymphaeae* and *C. incarnatum* with obtuse ends, while fusiform spores of *C. trichellum* and *C. musae* have both ends acute but the conidia are not curved (Figure 2). Based on the phylogeny inferred from the combined dataset, the species with acute ends and curved conidia such as *C. caricae*, *C. dematium*, *C. caudatum*, and *C. trichellum* are mostly with the most divergent position in the phylogenetic tree, whereas the species with obtuse ends and straight conidia such as *C. musae*, *C. nymphaeae*, and *C. incarnatum* are resolved basal to the rest of the species and thus represent the primitive species of *Colletotrichum*. These results suggest that the evolutionary direction of the spore shape is from conidia with obtuse ends to conidia with acute ends and from straight to curved conidia. As for the spore size, there does not appear to be a clear evolutionary trend. Further analyses that will include additional species are necessary to confirm the evolutionary trends in the genus *Colletotrichum*.

Conflicts of interest

The authors declare no conflict of interest.

ACKNOWLEDGMENTS

The authors would like to thank Dr. Kevin D. Hyde for reviewing an earlier version of the manuscript. Also we express our gratitude to Dr. PoPo Than for donation of *Colletotrichum* isolates. Research supported by the Fundamental Research Funds for the Central Non-profit Research Institution (#CAFYBB2014ZD005, #riricaf2014001Z), the Introduction Project of International Advanced Technology of Forestry (“948” Program) (#2014-4-66) and the Grant for Science and Technology Project of Yunnan Province (#2013FB087).

REFERENCES

- Bruns TD, White TJ and Taylor JW (1991). Fungal molecular systematics. *Annu. Rev. Ecol. Syst.* 22: 525-564.
- Cai L, Hyde KD, Taylor PWJ, Weir BS, et al. (2009). A polyphasic approach for studying *Colletotrichum*. *Fungal Divers.* 39: 183-204.
- Crouch JA and Beirn LA (2009). Anthracnose of cereals and grasses. *Fungal Divers.* 39: 19-44.
- Faisal PM, Prema R, Nagendran K, Gandhi K, et al. (2013). A specific and sensitive method for the detection of *Colletotrichum musae* in banana fruit. *Rev. Iberoam. Micol.* pii: S1130-1406(13)00016-8. [Epub Ahead of Print].
- Felsenstein J (1985). Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39: 783-791.
- Guerber JC, Liu B, Correll JC and Johnston PR (2003). Characterization of diversity in *Colletotrichum acutatum sensu lato* by sequence analysis of two gene introns, mtDNA and intron RFLPs, and mating compatibility. *Mycologia* 95: 872-895.
- Hall TA (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41: 95-98.
- Hillis DM, Moritz C and Mable BK (1996). *Molecular Systematic*. 2nd edn. Sinauer Associates, Sunderland.
- Huelsenbeck JP and Ronquist F (2001). MRBAYES: Bayesian inference of phylogenetic trees. *Bioinformatics* 17: 754-755.

- Hyde KD, Cai L, Cannon PF, Crouch JA, et al. (2009). *Colletotrichum* - names in current use. *Fungal Divers.* 39: 147-183.
- Johnston PR and Jones D (1997). Relationships among *Colletotrichum* isolates from fruit-rots assessed using rDNA sequences. *Mycologia* 89: 420-430.
- Lee DH, Kim DH, Jeon YA, Uhm JY, et al. (2007). Molecular and cultural characterization of *Colletotrichum* spp. causing bitter rot of apples in Korea. *Plant Pathol. J.* 23: 37-44.
- Manamgoda DS, Udayanga D, Cai L, Chukeatirote E, et al. (2013). Endophytic *Colletotrichum* from tropical grasses with a new species *C. endophytica*. *Fungal Divers.* 61: 107-115.
- Moriwaki J, Tsukiboshi T and Sato T (2002). Grouping of *Colletotrichum* species in Japan based on rDNA sequences. *J. Gen. Plant Pathol.* 68: 307-320.
- Moura HF, Vasconcelos IM, Souza CE, Silva FD, et al. (2014). Proteomics changes during the incompatible interaction between cowpea and *Colletotrichum gloeosporioides* (Penz.) Penz and Sacc. *Plant Sci.* 217-218: 158-175.
- O'Donnell K, Cigelnik E and Nirenberg HI (1998). Molecular systematics and phylogeography of the *Gibberella fujikuroi* species complex. *Mycologia* 90: 465-493.
- Page RDM (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput. Appl. Biosci.* 12: 357-358.
- Peng LJ, Yang YL, Hyde KD, Bahkali AH, et al. (2012). *Colletotrichum* species on *Citrus* leaves in Guizhou and Yunnan provinces, China. *Cryptogam. Mycol.* 33: 267-283.
- Peng LJ, Sun T, Yang YL, Cai L, et al. (2013). *Colletotrichum* species on grape in Guizhou and Yunnan provinces, China. *Mycoscience* 54: 29-41.
- Peres NA, Mackenzie SJ, Peever TL and Timmer LW (2008). Postbloom fruit drop of citrus and key lime anthracnose are caused by distinct phylogenetic lineages of *Colletotrichum acutatum*. *Phytopathology* 98: 345-352.
- Photita W, Taylor PWJ, Ford R, Hyde KD, et al. (2005) Morphological and molecular characterization of *Colletotrichum* species from herbaceous plants in Thailand. *Fungal Divers.* 18: 117-133.
- Prihastuti H, Cai L, Crouch JA, Phoulivong S, et al. (2010). Neotypification of *Colletotrichum falcatum*, the causative agent of red-rot disease in sugarcane. *Sydowia* 62: 283-293.
- Posada D and Crandall KA (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* 14: 817-818.
- Sherriff C, Whelan MJ, Arnold GM, Lafay JF, et al. (1994). Ribosomal DNA sequence analysis reveals new species groupings in the genus *Colletotrichum*. *Exp. Mycol.* 18: 121-138.
- Smith BJ and Black LL (1990). Morphological, cultural, and pathogenic variation among *Colletotrichum* species isolated from strawberry. *Plant Dis.* 74: 69-76.
- Sreenivasaprasad S and Talhinhas P (2005). Genotypic and phenotypic diversity in *Colletotrichum acutatum*, a cosmopolitan pathogen causing anthracnose on a wide range of hosts. *Mol. Plant Pathol.* 6: 361-378.
- Sreenivasaprasad S, Mills PR and Brown AE (1994). Nucleotide sequence of the rDNA spacer 1 enables identification of isolates of *Colletotrichum* as *C. acutatum*. *Mycol. Res.* 98: 186-188.
- Sreenivasaprasad S, Mills PR, Meehan BM and Brown AE (1996). Phylogeny and systematics of 18 *Colletotrichum* species based on ribosomal DNA spacer sequences. *Genome* 39: 499-512.
- Sutton BC (1992). The genus *Glomerella* and its anamorph *Colletotrichum*. In: *Colletotrichum* biology, pathology and control (Bailey JA and Jeger MJ, eds.). CAB International, Wallingford, 1-26.
- Swofford DL (2001). PAUP*: Phylogenetic analysis using parsimony (and other methods) 4.0b8. Sinauer Associates, Sunderland.
- Talhinhas P, Sreenivasaprasad S, Neves-Martins J and Oliveira H (2002). Genetic and morphological characterization of *Colletotrichum acutatum* causing anthracnose of lupins. *Phytopathology* 92: 986-996.
- Than PP, Jeewon R, Hyde KD, Pongsupasamit S, et al. (2008). Characterization and pathogenicity of *Colletotrichum* species associated with anthracnose on chilli (*Capsicum* spp.) in Thailand. *Plant Pathol.* 57: 562-572.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, et al. (1997). The Clustal_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* 25: 4876-4882.
- von Arx JA (1957). Die Arten der Gattung *Colletotrichum* Cda. *Phytopathol. Z.* 29: 413-468.
- Welsh J and McClelland M (1990). Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res.* 18: 7213-7218.
- Wharton PS and Diéguez-Uribeondo J (2004). The biology of *Colletotrichum acutatum*. *An. Jard. Bot. Madr.* 61: 3-22.
- Wharton PS, Julian AM and O'Connell RJ (2001). Ultrastructure of the infection of *Sorghum bicolor* by *Colletotrichum sublineolum*. *Phytopathology* 91: 149-158.
- White T, Bruns T, Lee S and Taylor J (1990). Amplification and direct sequencing of fungal ribosomal RNA genes for applications (Innis MA, Gelfand DH, Sninsky JJ and White TJ eds.). Academic Press, New York.
- Williams JGK, Kubelik AR, Livak KJ, Rafalski JA, et al. (1990). DNA polymorphisms amplified by arbitrary primers are

- useful as genetic markers. *Nucleic Acids Res.* 18: 6531-6535.
- Wu HX, Schoch CL, Boonmee S, Bahkali AH, et al. (2011). A reappraisal of Microthyriaceae. *Fungal Divers.* 51: 189-248.
- Wu HX, Li YM, Ariyawansa HA, Li WJ, et al. (2014). A new species of *Microthyrium* from Yunnan, China. *Phytotaxa* 176: 213-218.