Phylogeography and population biology of *Chrysoporthe austroafricana* and allied species

by

NICOLAAS ALBERTUS VAN DER MERWE

A thesis submitted in partial fulfilment of the degree

Philosophiae Doctor

in the

Faculty of Natural and Agricultural Sciences
Department of Genetics
Forestry and Agricultural Biotechnology Institute
University of Pretoria

South Africa

2012

© University of Pretoria

Declaration

I, the undersigned, hereby declare that the thesis submitted herewith for the degree Philosophiae Doctor to the University of Pretoria, contains my own independent work and has not been submitted for any degree at any other University.

_____

Nicolaas A. van der Merwe

© University of Pretoria

Table of Contents

Acknowledgements **7**

Preface **9**

CHAPTER 1

What is a fungal species? **12**

Introduction 13

A unified species concept 15

Species recognition criteria in the *Ascomycota* 17

Recognizing species using phenotype 17

Biological species recognition 20

Phylogenetic species recognition 22

Evolutionary factors that confound species boundaries and their recognition 23

Sex and hybridization 24

Horizontal gene transfer 26

Incomplete lineage sorting and cryptic speciation complicates recognition 27

What is the best definition of a species? 28

Criterion – A pragmatic approach to species 28

Concept – A philosophical approach to species 29

Prospects 31

Tables 34

Figures 37

Bibliography 45

CHAPTER 2

Primers for the amplification of sequence-characterized loci in *Cryphonectria cubensis* populations **66**

Abstract 67

Introduction 68

Materials and Methods 68

© University of Pretoria

Results 69
Discussion 70
Acknowledgements 71
Tables 72
Bibliography 75

CHAPTER 3

Multigene phylogenetic and population differentiation data confirm the existence of a cryptic species within *Chrysoporthe cubensis* 77

Abstract 78
Introduction 79
Materials & Methods 81
Isolates and DNA extraction 81
Phylogenetic analyses 82
Population genetic analyses 83
Morphology 84
Results 85
Phylogenetic analyses 85
Population differentiation analysis 86
Taxonomy 87
Restriction enzyme-based DNA diagnostic 90
Discussion 90
Acknowledgements 94
Tables 95
Figures 99
Bibliography 126

CHAPTER 4

Continental drift fails to explain allopatric speciation patterns of Diaporthalean tree pathogens 133

Abstract 134
Introduction 135
Materials & Methods 137
Data mining and data sets 137
Calibration of the molecular clock 137
Data analysis 138
Results 140

© University of Pretoria

Data mining and data sets 140
Data analysis 141
Discussion 142
Tables 147
Figures 154
Bibliography 161

CHAPTER 5

Host shifts cause diversity bottlenecks in populations of *Chrysoporthe austroafricana* 170

Abstract 171
Introduction 172
Materials & Methods 174
Fungal cultures and DNA extraction 174
Inference of the relative ages of *C. austroafricana* populations 175
Population biology of *C. austroafricana* 176
Results 179
Inference of the relative ages of *C. austroafricana* populations 179
Population biology of *C. austroafricana* 180
Discussion 182
Acknowledgements 186
Tables 187
Figures 190
Bibliography 196

CHAPTER 6

Host switching between native and non-native trees in a population of the canker pathogen *Chrysoporthe cubensis* from Colombia 201

Abstract 202
Introduction 203
Materials & Methods 204
Chrysoporthe cubensis fruiting structures, isolates and genotyping 204
Population subdivision and sexual reproduction 205
Phenotypic diversity of recombinant progeny 206
Results 208
Chrysoporthe cubensis fruiting structures, isolates and genotyping 208
Population subdivision and sexual reproduction 208

© University of Pretoria

Phenotypic diversity of recombinant progeny 209
Discussion 209
Acknowledgements 211
Tables 212
Figures 217
Bibliography 219
Summary 223
Conclusions and Prospects 224

© University of Pretoria

Acknowledgements

My gratitude for inspiration and assistance reaches far beyond my ability to express it verbally, but I would like to single out the following organizations and individuals without whose assistance this work could never have been completed:

Professors Emma Steenkamp, Brenda Wingfield and Michael Wingfield for their extreme patience, understanding, advice, leadership and above all, friendship. I am inspired by your professionalism and enthusiasm for my work.

Professors Alexander Myburg, Paulette Bloomer and Jaco Greeff, and Drs Martin Coetzee, Marieka Gryzenhout and Magriet van der Nest for psychological support and instilling a sense of urgency while allowing me to take my own path. Your influence on my life is permanent.

My family – mother, father and brother – have walked this road with me from start to finish, and I will be eternally grateful for their emotional and financial support. I hope that I can rely on them for further support as I embark on a life devoted to science.

Institutional and administrative support was provided by the administrative staff of the University of Pretoria and the Forestry and Agricultural Biotechnology Institute: Messrs Cathy Barnard, Mathilde Beresford, Ina Goosen, Eva Müller, Rose Visser, Jenny Hale, Heidi Fysh, support staff at the Hans Merensky Library, and especially Mrs Vivienne Clarence.

Financial support was provided by the National Research Foundation, the Andrew Mellon Foundation, the DST/NRF Center of Excellence in Tree Health Biotechnology (CTHB), the Tree Protection Co-operative Programme (TPCP), the Department of Genetics (University of Pretoria), and the University of Pretoria.

© University of Pretoria

I am also grateful to:

The staff of the culture collection at the Forestry and Agricultural Biotechnology Institute (FABI) for admirable performance while retrieving and depositing isolates;

The students I have had the privilege of mentoring over the past couple of years, for keeping me firmly grounded in knowledge transfer and allowing me to inspire them to greater heights;

Students in Lab Alpha, particularly Lieschen de Vos and Gerda Fourie, are walking the same road as I did and our mutual support and friendship will persist far beyond the completion of this degree;

Prof Anton Ströh, Dean of the Faculty of Natural and Agricultural Sciences (University of Pretoria) and Prof Henk Huismans, previous Head of Department of Genetics (University of Pretoria), for providing me an opportunity to practice science in the Department of Genetics as a member of staff;

The Department of Information Technology Services, where I held a previous post and where I learned how the world outside of science works.

Lastly, to my many friends both in and outside of science… without you, I would still have been clueless about interpersonal relationships and what is deemed important in the real world. I learned my biggest lesson from these friends: how to act like a human being, through the fun and also the hardships. A big thank-you to Craig Firman, Shane Strashoon, Nico van der Merwe and Andre Lamprechts (RIP, it was much too soon) – together we were the fabulous five, now we're just the four musketeers. Also Eduard and Riana Venter, Karen Surridge, Jacques Botha and Kevin Strydom as well as Kevin's mother, father and sisters – so much could be said, but I will just say thank you for your understanding and support.

© University of Pretoria

Preface

The genus *Chrysoporthe* includes several species that are economically important to eucalypt forestry activities in the tropics and sub-tropics of the world. Although these species have defined geographic distributions, they are difficult to separate using morphology or other phenotypic characters. This can be attributed to the overlapping phenotypic values that are generally observed between species. Thus, the best way to identify species is by employing DNA based techniques such as sequencing of polymorphic genes and subsequent phylogenetic analyses.

The main theme of this thesis is the elucidation of phylogeographic patterns displayed by species of *Chrysoporthe*, at global, continental and population scales. Phylogenetics and population genetic methods were employed throughout, in order to provide empirical evidence from which conclusions can be drawn or predictions made. However, due to recent taxonomic changes within both the genus *Chrysoporthe*, and the family to which it belongs (i.e., *Cryphonectriaceae*), it seemed in order to start with a review of the relevant literature pertaining to species concepts and recognition criteria in the fungi.

CHAPTER 1 is a literature review entitled "What is a fungal species?", hinting at the fact that the definition of a fungal species is largely dependent on limitations of taxonomists or the techniques they use. The fungi employ distinguishing means of reproduction, in most cases including a sexual and/or asexual phase, but in some cases they are exclusively sexual or asexual. Thus, the traditional notion of a biological species does not necessarily apply to all fungi, and alternatives must be sought. In the literature review I focus on the currently popular species recognition criteria for *Ascomycota* fungi. In turn, the concept of a species is treated with a philosophical approach, although it is possible to integrate the philosophical and pragmatic approaches to species so that a coherent understanding of what a fungal species is, can be formed.

© University of Pretoria

Throughout the course of scientific progress in the genus *Chrysoporthe*, population diversity has been a central theme. In older publications, diversity was determined using vegetative compatibility groups (VCGs). Although the employment of VCGs equates to a quick, cheap and easy method to determine population diversity, it is limited by the fact that the number of loci in the genome that govern a VCG phenotype is unknown. Thus, the aim of studies in **CHAPTER 2** was to develop polymorphic microsatellite markers that could be used for the determination of population parameters in *Chrysoporthe* spp. In addition to quantifying diversity, these markers could also be used to determine gene flow between populations and population subdivision. Such parameters, supported by phylogenetic information, are necessary for the investigation of phylogeography of incipient species.

Before work for this thesis was initiated, it was thought that *Chrysoporthe cubensis* was a pan-continental species occurring natively in Southeast Asia and South America. However, deployment of microsatellite markers to populations from those areas revealed that the continental populations are genetically isolated. Population genetic data were supported by the application of genealogical concordance phylogenetic species recognition (GCPSR) criteria, which showed that the South American and Southeast Asian lineages are evolving independently. Thus, **CHAPTER 3** is devoted to describing a new cryptic species for the fungus occurring in Southeast Asia.

After it was determined that *C. cubensis* (South America), *C. austroafricana* (Africa) and *C. deuterocubensis* (Southeast Asia) had continental distributions, it became possible to investigate the evolutionary histories of these species. Studies in **CHAPTER 4** sought to describe the ancient history of the genus *Chrysoporthe*, *in lieu* of these three species, by utilizing phylochronology and fossil calibration points. These determinations made it possible to speculate on the causes of the currently observed phylogeographic patterns of these species.

Chrysoporthe austroafricana is a native fungus in Africa, occurring on native *Syzygium* spp. (waterberry) as well as non-native *Eucalyptus* spp. and non-native *Tibouchina* spp. The primary objective in **CHAPTER 5** was, therefore, to determine

© University of Pretoria

whether native *Syzygium* spp. in Africa represent the formative host of *C. austroafricana*, i.e., the host on which the fungus evolved on this continent. Polymorphic DNA markers (developed in Chapter 2) and a divergence date for the *Chrysoporthe* lineage (from Chapter 4) were employed to determine which of the host-based fungal populations had the most ancient common ancestor, and to determine whether sexual outcrossing has played a role in host colonization. These data could additionally provide valuable information about the colonization route of *C. austroafricana* in Africa.

Although it had been shown that *C. austroafricana* jumps between hosts in Africa (Chapter 5), it was unknown whether such host jumps could also be present in the other species of *Chrysoporthe*. Unfortunately, all species of *Chrysoporthe* are not equally well sampled, but a valuable population of *C. cubensis* was sampled from two hosts in Colombia. This population could potentially reveal a host jump in a second *Chrysoporthe* species. The aim in **CHAPTER 6** was to determine whether populations of *C. cubensis* from *Eucalyptus* and *Miconia* hosts in close proximity to each other in Colombia, could reveal a host jump. This information would strengthen previous efforts to detect host jumping in *Chrysoporthe*, and would also be valuable to the forestry industry in South America.

Data and interpretations presented in this thesis provide valuable insight into the evolutionary mechanisms and processes underlying phylogeographic distributions of *Chrysoporthe* species. Additionally, many new questions have arisen during the course of these investigations. Some of these questions are outlined in Discussion sections of each of the chapters, while others have been incorporated into the Conclusions section for the thesis. These questions will hopefully form the basis of future research regarding the evolution, phylogeography and host associations of *Chrysoporthe*.

© University of Pretoria

CHAPTER 1

What is a fungal species?

© University of Pretoria

Introduction

The elucidation of species and the concepts used to delimit them is crucially important to understand the evolution of global biological diversity (De Meeûs *et al*. 2003; Hawksworth 2001; May & Beverton 1990). However, research pertaining to species concepts and most biodiversity studies focus almost exclusively on "higher" eukaryotes such as plants and animals, with an almost complete disregard of microorganisms. For example, unicellular microorganisms only became important in the development of generalized species concepts during the course of the last two decades (Nee 2003; Patterson 1999; Tautz 2003; Wilson 2003). The systematic treatment of parasitic organisms, which constitute approximately 30% of all known eukaryotes (De Meeûs & Renaud 2002), is particularly complex because they often do not comply with established species concepts (De Meeûs 1998; Kunz 2002). In contrast, some progress has been made regarding the systematic treatment and overall classification of eukaryotic microorganisms compared to prokaryotes (i.e., Roselló-Mora & Amann 2001; Schleifer 2009; Sneath 1957) and viruses (Van Regenmortel 1989). But the systematical classification of microorganisms, i.e. the use of relationships to delineate species through time, generally lags far behind that of plants and animals. Subsequently, this inability to accurately recognize and delineate most microbial species obscures the true nature of their biodiversity (reviewed by O'Donnell *et al*. 1994).

The "species problem" can be summarized as the disagreement between biologists about how species should be defined, and what the word "species" means. Hey (2001) pointed out that the species problem is rooted in the failure to recognize that our knowledge is incomplete, and that this limitation can be overcome by enquiry-based research. The problem is also exacerbated by disagreement about how to approach species philosophically. It is, therefore, not the species that is the problem, but the scientists that deal with them, leading Hey (2006) to argue that it is the human element that is overlooked. We do not understand how our own minds work, and Darwin recognized that this is a fundamental problem with defining species when he stated that "No one definition has yet satisfied all naturalists; yet every naturalist knows vaguely what he means when he speaks of a species" (Darwin 1859).

© University of Pretoria

Modern scientific techniques and new theoretical developments have revolutionized systematics. We can now probe deeply into the DNA of organisms and ask questions about relatedness, the influence of the environment, and how species have evolved over time. Additionally, population genetics has advanced to such a degree that it has become possible to dissect adaptation to niches and evolution over short time periods, i.e., biogeographic processes. These techniques are invaluable to understanding how and why speciation takes place, and they contribute to the establishment of the notion that evolutionary processes are predictable (Morris 2010).

The kingdom *Fungi*, and in particular the class *Sordariomycetes* in the phylum *Ascomycota*, includes a wide variety of life history habits, such as saprophytes, plant and human pathogens and endophytes (Taylor *et al*. 1999c; Zhang *et al*. 2006). Due to the variety of niches and lifestyles employed by these fungi, they constitute an ideal group to study adaptation, speciation and biogeography. In the Sordariomycete fungi, a wide range of species problems emerge, including convergence and otherwise cryptic species. As a consequence of being the largest class within the *Ascomycota* (Hawksworth *et al*. 1996), studies on the *Sordariomycetes* provide the best coverage in the fungi of current species concepts, species problems and solutions that might be found to deal with them.

In this review, the purpose is to attempt an understanding of species problems in the fungi by summarizing currently used species recognition criteria. This is followed by a discussion of evolutionary factors that influence how species are delimited. Lastly, pragmatic and philosophical approaches to species are differentiated in order to clarify the perceived "species problem". The focus of the review is restricted to the concepts applicable to the fungi, rather than on how speciation takes place, as this is adequately reviewed elsewhere (Butlin 1987; Coyne 2007; Coyne & Orr 2004; Gavrilets & Waxman 2002; Giraud *et al*. 2008; Johannesson 2001; Mallet *et al*. 2009; McCoy 2003; Summerell *et al*. 2010; Taylor *et al*. 1999b).

© University of Pretoria

A unified species concept

The idea of a unified species concept for all organisms can be attributed to Mayr (1947), but it has gathered notable support in the last 15 years (reviewed by De Queiroz 2007, 2011). Briefly, the concept distinguishes between a higher-level concept of species and operational criteria for species delineation. This higher-level concept can be viewed as an all-inclusive philosophical or mental construct of what species are. This is in contrast to operational criteria, which are those properties that can be used to recognize species. Among others, these properties include reproductive isolation (Avise & Wollenberg 1997; Dobzhansky 1970; Mayr 1942, 1963; Taylor *et al*. 2006; Wright 1943), monophyly (Donoghue 1985; Harrington & Rizzo 1999; Mishler & Brandon 1987) and our ability to distinguish between groups (Hey *et al*. 2003; Nixon & Wheeler 1990). The unified species concept should be separate from the science that needs to be applied in order to realize it. In as far as fungi are concerned, Taylor *et al*. (2000) made it clear that reform with regards to how species are conceived was necessary and that a new, possibly unified, species concept for fungi would have to be based on the acknowledgement of evolutionary, phylogenetic and population genetic information.

The problem with the words concept vs. criterion when dealing with species was originally recognized by Darwin (1859), who noted that the line of descent should be separated from Linnaean rank (cf. Lee 2003). In other words, he suggested that criteria to identify species should be treated independently of evolutionary history. However, during the last century improvements in DNA-based molecular systematics and our understanding of evolutionary mechanisms and processes, has allowed for the development of a so-called unified species concept. This concept essentially defines a species as separately evolving metapopulation lineages (De Queiroz 1998, 2007, 2011; Mayr 1942, 1953; Mayr *et al*. 1953). The advantages of this concept include the fact that the issues of species conceptualization and species delimitation are clearly separated, allowing all of the previously used species concepts to be used as evidence for delimitation. Additionally, evidence for any one of the operational criteria can be regarded as sufficient to delimit species, although several lines of evidence are usually employed, especially when morphological characters fail to yield evidence of delineation such as is true with cryptic species (e.g., Geiser *et al*. 1998; Giraud *et al*.

© University of Pretoria

2008; Harper *et al*. 2009; Koufopanou *et al*. 2001; Pringle *et al*. 2005; Steenkamp *et al*. 2002; Stielow *et al*. 2011). This allows for and encourages development of new operational criteria for certain taxa, and these criteria do not necessarily have to be concordant with those for other taxa at the same or a different rank. In other words, recognition criteria can be applied on a per-species or per-rank basis.

The notion of a species rank in and of itself is steadily losing support in favor of a rank-free phylogenetic taxonomy (Mishler 1999). In such a system, not all monophyletic groups need to be named but those that are named should be uninomial as defined by Mishler (1999) and unranked, although they should be hierarchically nested. It might not be immediately obvious how this system would be advantageous over the currently ranked Linnaean system, were it not for the integration of knowledge about the evolution of taxa. Thus, the Linnaean world-view of fungal taxonomy is being streamlined to align with current evolutionary synthesis (cf. Gilbert *et al*. 1996; Mayr & Provine 1981; Pigliucci 2007; Redecker 2002).

It might be argued that mycologists are at the forefront of taxonomic reform in order to align evolutionary knowledge with a ranked fungal taxonomy. Here they are integrating modern DNA based phylogenies with Linnaean taxonomy (e.g., Hibbett *et al*. 2007; Hibbett & Donoghue 1998; Lumbsch & Huhndorf 2007; Lutzoni *et al*. 2004; Schoch *et al*. 2009a; Schoch *et al*. 2009b; Zhang *et al*. 2006), which leads to the situation where rank and line of descent are unified and indistinguishable. Thus, in mycology, conceptual unification exists not only for species but also in their systematics, nomenclature and delineation, where the most commonly employed operational criteria for species recognition are phenotypic, biological and phylogenetic.

The “one fungus, one name” movement of recent years (Hawksworth 2011) indirectly supports the notion of a unified species concept for the fungi. Many fungi have separate species names for the teleomorph (sexual state) and the anamorph (asexual states). The suggestion is that there must be one name for the holomorph species name, which encompasses the teleomorph and all anamorphs. All states of a fungus have the same genome sequence and thus would group together based on

© University of Pretoria

phylogenetic analysis. Therefore, a single-name system does not present a problem for the species conceptualization, since modern mycological classification is biased towards nucleic acid sequence comparisons and DNA evolution as anchoring evidence for species and rank delimitation (cf. Berbee & Taylor 2010; Hibbett *et al*. 2007; Lutzoni *et al*. 2004; Redecker 2002; Schoch *et al*. 2009a; Schoch *et al*. 2009b; Spatafora *et al*. 2006; Taylor *et al*. 1999b; Taylor *et al*. 2000; Taylor *et al*. 2006; Zhang *et al*. 2006). These notions form an integral part of the Amsterdam Declaration, which aims to guide fungal taxonomists towards a new way of thinking about the nomenclature of pleomorphic fungal species (Hawksworth *et al*. 2011a; Hawksworth *et al*. 2011b; Norvell 2011).

Species recognition criteria in the *Ascomycota*

Recognizing species using phenotype

The basic premise of phenotypic species recognition in fungi is that phenotypic differences between species are fixed (reviewed by Harrington & Rizzo 1999). These include, but are not limited to, macroscopic and microscopic morphology (Hawksworth *et al*. 1996), biochemical differences such as the production of secondary metabolites and pigments (Boysen *et al*. 1995), differences in growth conditions (Harrington 1981; Taylor *et al*. 2006), formative host (i.e., the host on which a pathogen evolved) (cf. Baker *et al*. 2003; Banks & Paterson 2005; Hernandez-Bello *et al*. 2006; Van der Merwe *et al*. 2010) and geographic origin (cf. Sites & Marshall 2003). DNA and RNA molecules could also be viewed as phenotypic characters, but they are specifically excluded from this definition when their nucleic acid sequences are used for phylogenetic inference. Most phenotypic traits are quantitative in the sense that they are controlled by many loci in the genome that interact to produce a continuous distribution of values or states, and their expression is subject to environmental modification (Avise & Wollenberg 1997; Falconer & Mackay 1996; Felsenstein 1988; Hartl & Clark 1989). In other words, most phenotypic traits are plastic in the sense that they are variable according to environmental conditions (reviewed by Slepecky & Starmer 2009; Snell-Rood *et al*. 2010; Soll 2002). Thus, overlapping ranges of trait values or states for phenotypic characters present a challenge when closely related species are to be recognized.

© University of Pretoria

Some morphological traits are diagnostic for species or higher ranks. For instance, Some species in the genus *Ceratocystis* (Order: *Microascales*) produce hat-shaped or pillow-shaped ascospores that are otherwise rarely found in the fungi (Hanlin 1990). In recent times, however, it has become customary to first define closely related and morphologically similar species using other criteria such as phylogenetic monophyly, followed by searching for morphological characteristics that might, at least in part, distinguish them. For instance, various new species have been identified in the *Gibberella fujikuroi* species complex by phylogenetic methods, and these were later corroborated by the discovery of morphological differences between them (reviewed by Kvas *et al*. 2009).

Phenotypic species recognition on its own can fail when host range or the expression of disease symptoms on the host are used as primary operational recognition criteria. Some plant pathogenic fungi, for example, are inescapably bound to a specific host, while others are polyphagous. Among the latter fungi, host jumps are thus fairly common, even in obligate biotrophs, but occur mostly between closely related hosts (cf. Antonovics *et al*. 2002; Pérez *et al*. 2008). Furthermore, multiple fungal species can be associated with the same disease symptom (Baayen *et al*. 2000; Gryzenhout *et al*. 2009; Johanson *et al*. 1998; Van der Merwe *et al*. 2010) and, in extreme cases such as the Mycosphaerella Leaf Blotch disease of *Eucalyptus*, within the same lesion (Crous *et al*. 2001; Crous *et al*. 2009). The occurrence of disease complexes and the potential for host jumps in plant pathogens means that using only pathogenicity towards a specific plant or group of plants for taxonomic differentiation is commonly ineffectual. This is particularly problematic in the face of increasing globalization and anthropogenic movement of pathogens (cf. De Vienne *et al*. 2009; Slippers *et al*. 2005). The potential host ranges of pathogens can be predicted by using phylogenetic signal correlations between pathogen and host phylogenies (De Vienne *et al*. 2009; Garbelotto *et al*. 2010; Gilbert & Webb 2007) and artificial inoculations (cf. Alfenas *et al*. 2005; Pegg *et al*. 2010; Seixas *et al*. 2004). However, while these correlations are useful for studying disease complexes, the wide occurrence of these host-associated properties significantly confounds their value for species recognition

© University of Pretoria

purposes (Baker *et al*. 2003; Britz *et al*. 1999; Engelbrecht & Harrington 2005; Groenewald *et al*. 2006; Koufopanou *et al*. 2001).

Phenotypic species recognition can also fail when used in isolation and where recently diverged or even unrelated but isomorphic (phenotypically indistinguishable) lineages are considered, i.e. in the case of cryptic species (cf. Geiser *et al*. 1998; Johnson *et al*. 2005; Koufopanou *et al*. 2001; Pavlic *et al*. 2008; Pavlic *et al*. 2009; Pringle *et al*. 2005; Steenkamp *et al*. 2002). Such isomorphic lineages can arise due to a speciation event that lacks evolutionary pressures to drive phenotypic divergence (Giraud *et al*. 2008; Kohn 2005). Alternatively, unrelated species could harbor independently evolved convergent traits due to similar ecologies and niche functions. Selective pressure can thus result in convergence of morphology, physiology, behavior, or other phenotypic traits (Bliska & Casadevall 2009; Fontaine *et al*. 2010; Fraser *et al*. 2004). For instance, plant pathogenic oomycetes and filamentous fungi have evolved similar morphologies and lifestyles despite the fact that they are unrelated (Andersson 2006). Well known examples from the *Ascomycetes* are the near-isomorphic nature of *Ophiostoma* and *Gondwanamyces* species inhabiting protea inflorescences in South Africa (Roets *et al*. 2009). In the same fashion, unrelated species can independently evolve physiological traits governed by biochemical networks (cf. Slot & Rokas 2010), which arose due to specific selection pressures or through mechanisms such as horizontal (lateral) gene transfer (see Horizontal Gene Transfer).

In all cases where phenotypic species recognition "fails", species can still be delineated using other operational recognition criteria. In reality, failure of phenotypic species recognition is quite common and this is being increasingly recognized in fungal systematics. Contemporary studies seldom describe species solely on phenotypic characters. This is because fungal systematists usually employ a polyphasic approach to species recognition, encompassing several recognition criteria (e.g., Samson *et al*. 2009; Samson & Varga 2009). Also, to recognize the persistent occurrence of cryptic species in fungal nomenclature, numerous species are and have been characterized and described based primarily on DNA-based information and the

© University of Pretoria

biological species and/or phylogenetic species recognition criteria (Pavlic *et al*. 2008, 2009; Stielow *et al*. 2011).

Biological species recognition

Biological species recognition (BSR) (e.g., Mayr 1942, 1953) assumes that individuals that are able to reproduce sexually and produce viable, fertile offspring should be members of a single species (Brasier 1997; Coyne & Orr 2004; Kohn 2005; Taylor *et al*. 1999a; Taylor *et al*. 2000) (Table 1). As such, the ability of organisms to mate within the same species, but not with other species, forms the foundation of all other species recognition criteria. Sexual reproduction in the form of random mating leads to homogenization of both genetic and phenotypic characters within a species (Ayala 1982; Falconer & Mackay 1996; Hartl & Clark 1989; Hedrick 2000), allowing comparisons to be made between empirically estimated averages of equilibrium states of separate species. Thus, the absence of gene flow between separate species is an indication that the species have possibly diverged to sexually incompatible states, although other factors such as geographic separation may play a role (Barker 2007; Fitzpatrick *et al*. 2009; Liu *et al*. 2009; Mallet *et al*. 2009; Petit & Excoffier 2009).

Due to the homogenizing effect of interbreeding between individuals of the same species, groups of interbreeding individuals can be thought of as cohesive units resulting from a common ancestry. Such cohesion species (compare Table 1) are therefore represented by lineages that are genetically exchangeable and ecologically interchangeable with each other but not with lineages representing other species (cf. De Queiros & Donoghue 1988; Stockman & Bond 2007) (Figure 1). Therefore, cohesion species are essentially genetically isolated segments of lineages in metapopulations that are adapted to a niche, and individuals from one sub-population can function as members of another sub-population (Petit & Excoffier 2009). Gene flow between populations can thus be viewed as evidence that such populations belong to the same species, although this view has been strongly contested (Barker 2007).

BSR may fail or is not practically usable in some fungi due to the prevalence of apparently asexual species (Hawksworth 2011; Kück & Pöggeler 2009; Taylor *et al*.

© University of Pretoria

1999b). Members of these species cannot be mated in the laboratory, either because artificial induction of the sexual phase is poorly understood or impossible, or because they are truly asexual in the sense that the permanent anamorphic (mitosporic) state is an evolutionarily derived character. Well known examples of apparently asexual fungi include *Fusarium oxysporum* (reviewed by Fourie *et al*. 2011), *Aspergillus fumigatus* (Paoletti *et al*. 2005; Pöggeler 2002), *Cochliobolus heterostrophus* (Sharon *et al*. 1996) and *Coccidioides immitis* (reviewed by Taylor *et al*. 1999b). Nevertheless, these fungi possess apparently functional mating-type loci required for sex, or are suspected of having cryptic sex based on population genetic evidence that points to recombination (Burt *et al*. 1996; Geiser *et al*. 1998; Kück & Pöggeler 2009).

Another confounding aspect of applying BSR lies in the differences in mating strategies of sexual fungi. Two major mating strategies exist in these fungi, namely homothallism and heterothallism (Glass *et al*. 1990; Yun *et al*. 2000; Zeyl 2009). Heterothallic individuals contain one of two mating-type idiomorphs in their genomes, and thus they must undergo sexual reproduction with an individual of the opposite mating-type (Clark & Haskins 1998; McGuire *et al*. 2001; Picard *et al*. 1991; Yun *et al*. 2000). In contrast, homothallic individuals contain both mating-type idiomorphs in their genomes and can undergo self-fertilization during sexual reproduction (Beatty *et al*. 1994; Glass *et al*. 1990; Yun *et al*. 2000). These homothallic species present practical problems when biological species recognition is applied, because additional strategies are needed to ascertain whether sexual reproduction was the result of a selfing or an outcrossing event.

Closely related fungi do not necessarily employ the same mating strategy, which also complicates BSR. In other words, both homothallism and heterothallism could be employed by different species of the same genus, e.g., *Calonectria* (Crous 2002; Crous *et al*. 2006; Crous *et al*. 2004). Additionally, even members of the same species could undergo unidirectional mating-type switching, or switching from homothallic to heterothallic mating strategies (Perkins 1987). In *Ceratocystis coerulescens*, for instance, the progeny of a homothallic selfing can be divided into roughly equal numbers of homothallic and self-incompatible or heterothallic individuals (Harrington & McNew 1997; Witthuhn *et al*. 2000). Genera or species that employ such

© University of Pretoria

heterogeneous mating strategies are generally taxonomically "difficult" (cf. Schoch *et al*. 2009b; Strandberg *et al*. 2010), and this can be ascribed primarily to difficulties in applying biological species recognition criteria in these assemblages.

Even though there are problems with BSR, it remains implicit and fundamental to most other species concepts. As a result, most of the practical problems pertaining to this operational character can be overcome by using phylogenetics for species recognition (see below), or to combine the latter with population genetic approaches to allow for the delineation of populations of interbreeding and reproductively isolated individuals. The latter approach has proven highly effective for the fungi. For instance, *Coccidioides posadasii* was delineated from *Coccidioides immitis* based on lack of gene flow, among other criteria (Fisher *et al*. 2002). Several cryptic species of *Aspergillus* have also been identified, although they have not yet been named (Geiser *et al*. 1998; Hawksworth 2011; Pringle *et al*. 2005).

Phylogenetic species recognition

Phylogenetic species recognition (PSR) relies on hierarchical descent (i.e., relationships among species) rather than tokogenetic processes (i.e., relationships among populations within species) (Coyne & Orr 2004; Harrington & Rizzo 1999; Hennig 1950, 1976; Olmstead 1995; Rieppel 2009; Taylor *et al*. 2000). As such, the concept encapsulates a number of sub-concepts or operational criteria that have been reviewed extensively (Dettman *et al*. 2003, 2006; Harrington & Rizzo 1999; Hibbett & Donoghue 1998; Koufopanou *et al*. 2001; Miller & Huhndorf 2004; Taylor *et al*. 2000) (Table 1). However, the basic method of species delimitation revolves around the identification of clades or groups of organisms or their genes that represent evolutionary distinct lineages.

Genealogical concordance phylogenetic species recognition (GCPSR) (Figure 2) (Avise & Wollenberg 1997; Barraclough & Nee 2001; Taylor *et al*. 2000) is currently the most favored delimitation method employed by mycologists (Dettman *et al*. 2006; Fisher *et al*. 2000; Geiser *et al*. 1998; O'Donnell *et al*. 2004). The method relies on concordance, i.e. congruence or agreement, between phylogenetic clades recovered from multiple genomic regions. In cases where genealogies among different regions

© University of Pretoria

are incongruent, the superclade (i.e., the least inclusive clade that eliminates incongruence between genealogies) is used for species delimitation. Thus, using this recognition method, a species is the smallest group of populations that have a common lineage and that share unique and diagnosable phenotypic or DNA sequence characters (compare Table 1, Genealogical species).

The major problem associated with the use of phylogenetic species recognition, or any other species recognition approach, relates to a so-called "Iceberg Effect" illustrated for the fungi by Leslie *et al*. (2001). The "iceberg" in this case describes inadequate sampling of diversity within and between populations of the same species, thus much more is below the surface than is realized. When selective sampling takes place, the chances of sampling all the diversity within a population or species are diminished. This can result in sampling locally adapted populations, which may appear to be genetically isolated populations or even different species.

The choice of marker system is also important when attempting to avoid the artificial recovery of so-called "iceberg" samples. The larger the sample of loci from the genome, the more likely it will be that valleys can be recovered between locally adapted "icebergs". Utilization of greater numbers of loci also has the added benefit of reducing the possibility of inadvertently selecting loci that may be under evolutionary selection. The advantage of GCPSR in fungal species delineation is, thus, that several loci and multiple isolates are used to recognize species, and this method is fairly tolerant of "iceberg" samples.

Evolutionary factors that confound species boundaries and their recognition

Hierarchical evolutionary processes result in well-defined lineages that are relatively easily diagnosable, by using any number of species recognition tools. However, several life history processes can result in the breakdown of boundaries between species. Such processes thus complicate or obfuscate the operational criteria applied to recognize true species.

© University of Pretoria

Sex and hybridization

Biological species recognition requires complete fertility between compatible individuals of the same species. However, in some genera incomplete interfertility between species in the form of incomplete prezygotic and/or complete or incomplete postzygotic isolation (Kohn 2005; Orr & Turelli 2001) confounds this approach. Dobzhansky (1935) defined prezygotic isolation mechanisms as those that prevent the appearance of hybrids, while postzygotic isolation is associated with hybrid sterility. Both mechanisms are thus involved in the continued separation of distinct species. However, artificial hybrids, i.e. those produced by forced mating (Hernandez-Bello *et al*. 2006; Lin *et al*. 2007), whether they are sterile or not, cannot be used as evidence for the lack of species boundaries. For example, the *Gibberella fujikuroi* complex includes numerous examples where distinct biological species can produce viable and fertile hybrid progeny under laboratory conditions (Desjardins *et al*. 2000; Leslie *et al*. 2004; Steenkamp *et al*. 2002). This is because species can emerge for many reasons, and may require more time to develop sexual incompatibility. Thus, artificial hybridization cannot be seen as a true reflection of the natural situation, and this is an oversight that commonly creeps into the mycological literature.

Prezygotic isolation mechanisms in fungi include temporal, geographical and host separation of fertile cycles, and the loss of mate recognition between closely related species that fall into one or more of these categories (Casselton 2002; Engelbrecht & Harrington 2005; Le Gac *et al*. 2007; Paterson 1980). Mate recognition is presumably governed by pheromone molecules that are secreted by one individual and detected by another individual, in order to initiate mating (Bistis 1981, 1983; Glass & Kuldau 1992; Horn *et al*. 2009; Kües & Casselton 1992). As such, failure of recognition is perhaps the last step towards complete genetic isolation of two divergent species (Barton 2001; Brasier 2000; Brasier *et al*. 1999). However, in order for mate recognition to take place, the two individuals must be in close temporal and geographic proximity. This implies that an individual of one species may be able to recognize a mate from another species when the isolating mechanism is not dependent on sexual reproduction. Thus, recently diverged but spatially separated species may still be able to mate when they are brought together under natural or artificial conditions (cf. Brasier 2001; Schardl & Craven 2003).

© University of Pretoria

Postzygotic isolating mechanisms can be divided into two classes applicable to the fungi, namely chromosomal and genic incompatibility (Dobzhansky 1935, 1937). When species with different numbers of chromosomes are mated artificially, the offspring are likely to be infertile due to chromosome number aberrations, resulting in the continued separation of the parental species (reviewed by Dujon 2010). On the other hand, genic causes of incompatibility can be ascribed to the loss of genes or sections of chromosomes that fail to pair during meiosis (Dobzhansky 1935; Maclean & Greig 2010; Xu & He 2011). An example of genic incompatibility is the artificial cross of *Gibberella subglutinans* and *Gibberella circinata*, the progeny of which display gross segregation distortion (De Vos *et al*. 2007), potentially affecting the fitness of the progeny (cf. Wu 2001).

The emergence of sexual incompatibilities can be ascribed to either allopatric speciation or sympatric speciation (cf. Wiens 2004). Allopatric (geographic) species arise due to the spatial separation of populations, which subsequently undergo selection and drift. However, sympatric species diverge while occupying the same geographic area, potentially due to initially small differences in host preferences of sub-populations, or random effects such as differences in allele frequencies over larger distances (parapatric speciation). These forms of speciation have been adequately reviewed elsewhere (Coyne 2007; Fitzpatrick *et al*. 2009; Gavrilets & Waxman 2002; Giraud *et al*. 2008; Johannesson 2001; Kliman *et al*. 2000; Kohn 2005; Mallet *et al*. 2009; McCoy 2003; Ribeiro & Caticha 2008; Wiens 2004). It is, however, important to note that neither allopatric nor sympatric speciation implies the emergence of sexual incompatibilities between the resultant species. Thus, natural interspecies hybrids can occur among some taxa (Brasier 2000; Brasier 2001; Inderbitzen *et al*. 2011; Kohn 2005; Olsen & Stenlid 2002; Schardl & Craven 2003).

Hybridization between taxa that are perceived to be distinct represents one of the most complex problems for any kind of phylogenetic species recognition system. In fungi, natural hybrids are well known (reviewed by Schardl & Craven 2003), and they present a considerable obstacle to the application of traditional species recognition to define species boundaries. Inter-specific hybrids contain alleles from each of the

© University of Pretoria

parental species, and thus they could exhibit both morphological and genetic relatedness to either or both of the parents (cf. Falconer & Mackay 1996; Hartl & Clark 1989; Hedrick 2000). Using isolated, non-integrated methods, such hybrids could be seen as the "missing link" between two well defined species – Bailey's "intergradient forms" (Bailey 2010) – which could result in a breakdown of the species boundary, leading to misclassification of both of the parental species and their hybrid. Therefore, it is only through integrating morphological, biological and genetic methods, such as population genetic analyses, phylogenetics and phylogeography, that these hybrids can be unequivocally identified as such (Brasier 2001; Brasier *et al*. 1999; Hewitt 2001; Olsen & Stenlid 2002; Schardl & Craven 2003).

Horizontal gene transfer

The stable, nonsexual transfer of genes across species boundaries is known as horizontal gene transfer (HGT) (cf. Kidwell & Lisch 1997; Lawrence 1997; Lawrence & Roth 1996). Hybrids are specifically excluded, due to the fact that they originate from sexual reproduction among closely related species. Nonetheless, HGT has played a major role in the evolution of most of the life forms on Earth and is thought to have altered the course of evolution of all groups of eukaryotes very early on (Richards *et al*. 2009). However, more recent horizontal acquisition of genes may be problematic for systematics, because these genes can give rise to apparently convergent phenotypic characters (cf. Slot & Rokas 2010), and their use for phylogenetic inference could result in discordance between gene trees and species trees (Berbee & Taylor 2001; Knowles & Carstens 2007; Rosenberg 2002; Strandberg *et al*. 2010).

Philosophically, HGT alters the view of modification by descent, or the Darwinistic view, as it allows modification of a species without descent. Although no direct evidence exists for HGT in nature, in fungi it can certainly be imitated experimentally and inferred phylogenetically. In nature, it is thought that plasmids, mycoviruses and transposable elements facilitated HGT among fungi (reviewed by Eisen 2000; Rosewich & Kistler 2000). This can have far reaching implications for the fungal phenotype. As suggested by Rosewich and Kistler (2000) HGT has aided fungal

© University of Pretoria

evolution by altering, for example, biosynthetic capabilities (Schmitt & Lumbsch 2009; Walton 2000) and host range of plant pathogens (Mehrabi *et al*. 2011).

During the last few years with the rapidly rising numbers of comparative genomics studies, sparked by improved next-generation sequencing technologies, it has also become abundantly clear that HGT has shaped the structure and development of fungal chromosomes and genomes (Ma *et al*. 2010; Mallet *et al*. 2010; Mehrabi *et al*. 2011; Richards *et al*. 2011). HGT is thus a major driving force of fungal evolution that could severely impact on systematic analyses of species if properties or genes subject to HGT are used for species recognition (cf. Jordan *et al*. 1999). Perhaps the only way to solve such problems is the application of a polyphasic approach based on phylogenetics, phylogenomics, morphology, ecology, etc. (Eisen 1998; Eisen & Fraser 2003; Gräser *et al*. 2008; Richards *et al*. 2009; Samson *et al*. 2009; Samson & Varga 2009).

Incomplete lineage sorting and cryptic speciation complicates recognition

Evolutionary forces act gradually on the genomes of organisms. When an isolating barrier appears that divides a population, the two lineages are at first isomorphic, i.e. they are indistinguishable based on allele frequencies, phylogenetic affiliation or morphology. Bar the collapse of the barrier, the two resultant populations will steadily diverge over time (Coyne 2007; Coyne & Orr 2004). The populations may undergo differential random allele fixation (i.e., genetic drift) or they may be subjected to different selective pressures that drive their divergence (cf. Gillespie 2000; Hedrick 2000; Wakeley 2004). This divergence signal will not be apparent at all loci in the divergent genomes for many generations (Geiser *et al*. 2000; Geiser *et al*. 1998; Taylor *et al*. 1999a). In other words, the incipient species will carry shared fixed loci as well as shared polymorphic loci, until selection and drift had caused differential fixation between them (Figure 3). Thus, in genetically isolated sibling species one can expect to find shared polymorphic loci, loci that are fixed in one species but polymorphic in the other, and loci that are differentially fixed between the species. This process is dependent on and inversely proportional to time since divergence (Taylor *et al*. 2000) if population sizes remain constant. Depending on which stage is

© University of Pretoria

sampled and which gene regions are analyzed, it may appear that polymorphisms are shared between the two lineages, yet these lineages evolve independently.

The problem outlined above illustrates the difficulties associated with incomplete lineage sorting during any systematic approach. If two lineages share the same polymorphic alleles at a locus, the question arises as to how they could be separate species. However, if only one locus is found that is differentially fixed between them, this can be viewed as evidence of genetic isolation, since interbreeding would have quickly randomized the polymorphisms. Our inability to capture sections of genomes that support the fact that two groups of fungal isolates represent independently evolving lineages might be cited as the primary reason for the existence of cryptic species.

What is the best definition of a species?

De Queiroz and his co-workers (De Queiros & Donoghue 1988; De Queiroz 1998, 2007, 2011) proposed that there really is only one species concept, and that all other species concepts should be demoted to operational criteria. However, not all researchers support this view (cf. Hey 2001; Mishler & Donoghue 1982; Sites & Marshall 2003; Wheeler & Nixon 1990). For instance, Hey (2006) argued that the major confusion seems to stem from the historically almost interchangeable use of the words "concept" and "criterion". If the meaning of these words can be clearly defined, progress can perhaps be made towards unification of the idea of a species.

Criterion - A pragmatic approach to species

Science is rooted in empiricism (cf. Harrison 2008; Hooker 1975), which is to say that it is based on sense-experience or observation. Therefore, it is desirable to obtain a set of rules, criteria or recognition mechanisms for species in order to provide an unbiased perspective on biased observations. As the word implies, a species "criterion" is a critical evaluation of the validity of a species based on empirical evidence, i.e. a characteristic. Thus, a species criterion is pragmatic; in other words it is based on practical rather than theoretical considerations (cf. Godfray *et al*. 2008; Quine 1951). The question thus revolves around how to define the rules for species recognition, rather than how to implement them. In a modern interpretation of

© University of Pretoria

Darwin's Origin of Species, Jones (1999) describes a species as a "gene republic", implying that species generally do not exchange genetic material with each other. The widely employed operational species recognition criteria for fungi thus serve to detect any exchanges across "state borders" (i.e., gene flow across species barriers) by targeting specific phenotypic, reproductive or phylogenetic properties or characteristics.

Concept - A philosophical approach to species

When the true meaning of the word "concept" is analyzed, it becomes clear that a concept is a cognitive process, i.e., an act of knowing or understanding (cf. Ghiselin 1987; Goodwin & Johnson-Laird 2011). A species concept can then be decoupled from empiricism, i.e. how to apply it, thus removing the temptation to argue over how to define a species. In order to do this, we must recognize that we are not defining and naming species based on deductive reasoning, i.e., reasoning based on arguments designed to prove that a conclusion follows logically on another set of arguments (Politzer & Macchi 2000). Instead, we are identifying species based on inductive reasoning, i.e., reasoning that is generalized while based on individual instances, such as fungal isolates representing a lineage (i.e., Wilson 1995). Therefore, our definition of a species may or may not coincide perfectly with the real, biological species.

When philosophical arguments over species are sought, it is useful to delimit realms that differentially include concepts, theories, hypotheses, and operational criteria. In such an organization, a concept is an abstract idea or mental symbol, i.e., a cognitive unit of meaning which is usually difficult to verbalize. A theory is then a system of ideas that are useful to explain species but that are independent of the species to be explained, while a hypothesis is a proposed explanation for a species, which can be tested based on experimental observations. The latter forms the basis of the operational criteria for recognizing species, which are also the smallest philosophical aspects of the philosophical argument over species. Therefore, phenotypic, biological and phylogenetic species can be contrasted to concrete organic species, i.e. the real-world, natural Species ("S" in Figure 4) that exist in nature, which are philosophically immutable, whether they are described or not.

© University of Pretoria

The formal definition of a species, i.e. as it stands in the published species description, does not necessarily correlate with the natural Species (Figure 4). Mostly, however, it is expected that there would be at least some overlap between species (i.e., the formal description, operational species or philosophical species) and the natural Species. This overlap can be denoted as the realized species ("RS" in Figure 4), because it is that part of the Species that is diagnosable and recognizable based on the applied species recognition criteria. However, via empiricism, i.e. observation and experimentation, the definition may at one point be both broad enough to encompass the entire Species and specific enough to exclude any other Species. Although it will be impossible to know whether perfect Species-realized species overlap had been achieved, the degree of overlap may change. This is because new evidence or ideas (such as evolutionary factors that confound species boundaries) may be integrated into operational criteria that can cause the overlap to shift. Additionally, current scientific techniques are inadequate to quantify all the natural variants of a species prior to describing it. Such variants may be described as new species if the only distinguishing characteristic is omitted in species quantification, leading to gross discordance between operational criteria and Species.

The pursuit of perfect operational criteria for a species definition, so that the definition is in agreement with the natural Species as much as possible, explains the existence of "lumpers" and "splitters" among taxonomists (Endersby 2009). Such taxonomic actions are guided by available evidence and the technologies used to gather the evidence. Thus, the problems with species and their definition and recognition are rooted in the human element, which is technologically constrained in quantifying its own ideas. Perhaps this is the reason why Bailey (1896) described species as a "human contrivance", alluding to the fact that our understanding of species is potentially far removed from the reality of organic, natural Species.

To anchor and guide our understanding of species, it might be best to view a formal species definition simply as a hypothesis, which could be either supported or rejected based on empirical data (Hey *et al*. 2003). By viewing a species as a hypothesis or at least some level of philosophical entity, it becomes testable using the scientific method, and less prone to interpretation (Ghiselin 1987). For example, any number of

© University of Pretoria

the currently used operational species recognition criteria may be applied to either accept or reject the hypothesis. Species recognition can then be compartmentalized within certain Linnaean ranks such as genera, families, etc. For instance, for one rank a certain set of recognition criteria may be required, while the recognition criteria for another rank could be vastly different. This allows true unification of species recognition across the tree of life, by recognizing that theoretical unity does not necessarily have to imply empirical uniformity (cf. Holsinger 1987; Mishler & Donoghue 1982).

Prospects

The resolution of the species debate coincides with exciting technical achievements in biological science. Most notably, the development of high throughput genome sequencing over the last decade (reviewed by Creer *et al*. 2010; Zhang *et al*. 2011) will have an enormous impact on the methods employed to define species boundaries. Coincident with these developments are major advances in bioinformatics practice and techniques to analyze and interpret such huge volumes of information (cf. Nielsen *et al*. 2010). However, the field of fungal systematics faces similar challenges to those in bioinformatics in the sense that new methods are now required to complement standard analysis techniques with genomics resources. Fungal taxonomists may thus find it prudent to borrow techniques from those working on other organisms.

Regardless of which sequence-based techniques are employed, incomplete lineage sorting will remain problematic during species delimitation. For recently diverged species, it may be extremely laborious to search for differentially fixed loci. In such cases it may be more efficient to perform genome-wide sequence comparisons in the light of phylogenetic species recognition. To this end, *Oryza sativa* (cultivated Asian rice) has recently been distinguished from its wild progenitor species, *O. rufipogon*, using genome-wide multilocus gene genealogies and GCPSR (Yang *et al*. 2011) in order to discount incomplete lineage sorting. This was done because *O. sativa* has hybridized with *O. rufipogon*, yielding the domesticated varieties *japonica* and *indica*. These results show clearly that incomplete lineage sorting can be overcome using GCPSR. Additionally, it may be useful to consider genome-wide single nucleotide polymorphism (SNP) analyses (Shaffer & Thomson 2007; Tian *et al*. 2008; Zhang

© University of Pretoria

2010; Zhang *et al*. 2004), although this approach has not yet been widely applied for fungal species delimitation. Therefore, the suggestion to sequence 1000 fungal genomes (Grigoriev 2011) will certainly contribute towards fungal systematics in the near future.

In addition to employing the nucleotide sequence information associated with whole genomes, it may be possible to use chromosomal synteny for the purposes of species recognition. Although microsynteny (i.e., the collinear chromosomal locations of small sets of genes) (Xu *et al*. 2006) can obscure much of the evolutionary signal hidden in genome organization, it may provide a mechanism to analyze similarities between species instead of differences (cf. Slot & Rokas 2010). In this way it may be possible to use metabolic gene clusters to dissect the evolutionary histories of deep divergences within the fungal tree of life (ex. Koestler & Ebersberger 2011), or to study co-evolution between pathogen and host (ex. Oberhaensli *et al*. 2011). Such information can aid greatly in the resolution of species boundaries in the fungi.

The vast amounts of data generated by new sequencing technologies beg for novel analysis methods. Although standard analysis methods such as multi-gene phylogenetics remain useful and economic (cf. Aguileta *et al*. 2008), new methods can complement current phylogenetic knowledge. In this regard, a composition vector method has recently been devised to analyze many whole fungal genomes without the need to perform sequence alignment (Wang *et al*. 2009). Additionally, Dress *et al*. (2010) exploited the fact that a phylogenetic tree is a mathematical construct known as a directed graph, in order to devise a method to reconstruct phylogenies based purely on hierarchical relationships that are mathematically defined. With adaptation and refinement, this method could potentially be extended to the identification of previously unknown phylogenetic lineages. Undoubtedly, new theoretical and technical advances will certainly build upon existing knowledge in order to speed the elucidation of the fungal tree of life.

Genomics and the emergent technologies and methodologies associated with it will change the way we classify, identify and analyze fungal species. Species complexes are highly prevalent in the fungi (cf. Dettman *et al*. 2006; Grube & Kroken 2000;

© University of Pretoria

Johnson *et al*. 2005; Koufopanou *et al*. 2001; Le Gac *et al*. 2007; Pavlic *et al*. 2008, 2009; Samson & Varga 2009; Stielow *et al*. 2011; Van Wyk *et al*. 2006), but many of these may potentially be resolved using novel techniques. Indeed the current standard for description of a new species is to demonstrate that it conforms to the definition of a species based on multiple species concepts (cf. Couch & Kohn 2002; Engelbrecht & Harrington 2005; Hibbett *et al*. 2011; Linnakoski *et al*. 2010; Venter *et al*. 2002; Wingfield & Van Wyk 1993). The causes of the difficulties associated with fungal species delimitation are multitude, but mostly rectifiable by using sufficient sample sizes to represent the variability within a delineated species appropriately, and using multiple neutral markers to maximize the chances of detecting gene flow or recombination within and among populations.

© University of Pretoria

Table 1 Modern species concepts applicable to the fungi, which are referred to as operational criteria by De Queiroz [modified from De Queiroz (1998, 2007), Harrington & Rizzo (1999) and Sites & Marshall (2003)].

Species recognition criterion ①	Biological basis and definition	Features	References
Morphological / Phenetic / Phenotypic	• Species form a morphological, phenetic or phenotypic cluster based on quantitative differences	• Only morphologically stable characters, i.e. those that do not change under different environmental conditions, are useful • Only numerically definable characters can be used for phenetic species, i.e. numerical taxonomy	Duncan & Baum (1981) Hawksworth *et al.* (1996)* Sokal & Crovello (1970)
Biological	• Species are interbreeding groups of populations that are reproductively isolated from other such populations	• Species are defined in terms of mechanisms that isolate them • Shared ecological and morphological features are assumed to implicitly confirm interbreeding • Sexual reproduction results in viable and fertile offspring • Inapplicable to asexual species	Avise & Wollenberg (1997) Dobzhansky (1970) Mayr (1942) Mayr (1963) Taylor *et al.* (2006)* Wright (1943)
Isolation	• Reproductive isolation based on intrinsic properties (i.e. gene exchange between populations is prevented by one or more reproductively isolating mechanisms), rather than extrinsic properties (i.e. geographic isolation)	• Species are inherently sympatric • Inapplicable to asexual species	Dobzhansky (1970) Engelbrecht & Harrington (2005)* Le Gac *et al.* (2007)* Mayr (1942) Mayr (1963) Mayr (1970)
Recognition	• A species is the most inclusive population of individuals that share a common mate recognition and fertilization system	• Pre- and postzygotic incompatibilities isolate species • Inapplicable to asexual species	Casselton (2002)* Paterson (1980)
Ecological	• A species is a lineage that occupies an adaptive zone that is not significantly different from another lineage in the same range, but which evolves separately from other lineages outside of this range	• Populations of a species share a niche or adaptive zone • Evolution occurs within the constraints of ecological boundaries • Implicitly sympatric • Applicable to sexual and asexual species	Andersson (1990) Mayr (1947) Fournier & Giraud (2007)* Harrington & Rizzo (1999)* Miller & Wenzel (1995) Schluter (2009)* Van Valen (1976)

© University of Pretoria

Evolutionary	• A species is a series of populations that have a common ancestor and have not significantly diverged from each other or the ancestor	• The metapopulation has a unique evolutionary role and historical fate • Applicable to sexual and asexual species	Simpson (1951) Giraud *et al.* (2008)* Taylor *et al.* (1999a; 1999b)*
Cohesion	• The most inclusive population with the potential for phenotypic cohesion, i.e. similarities between populations that enable niche and demographic exchangeability	• Genetic exchangeability is limited by the spread of genetic variants via gene flow • Demographhic exchangeability is limited by the niche that prevents gene flow • Implicitly allopatric, peripatric or parapatric but not sympatric • Applicable to sexual and asexual species	Bond & Stockman (2009) Stockman & Bond (2007) Templeton (1998)
Phylogenetic	• A species is the smallest collection of sexual populations or asexual lineages that share a unique combination of character states	• No phylogenetic structure exists within a species • Monophyly and paraphyly do not apply to a species • Constant character states distinguish between reticulation and hierarchical relationships • Applicable to sexual and asexual species	Avise & Wollenberg (1997) Hibbett & Donoghue (1998)* Miller & Huhndorf (2004)* Nixon & Wheeler (1990) Taylor *et al.* (2000)* Wilson (1995)
Monophyletic	• A species is the smallest monophyletic taxon, and is ranked as a species based on processes that maintain lineages within it	• A clade consists of an ancestor and its descendants, usually displaying shared polymorphism of derived characters (i.e., autapomorphies) • Ranking of a taxon is pluralistic, i.e. it is based on many criteria • Applicable to sexual and asexual species	Donoghue (1985) Harrington & Rizzo (1999)* Mishler & Brandon (1987)
Genealogical	• A species is the smallest exclusive monophyletic group whose genes coalesce more recently with each other than with any organisms outside the group	• Exclusive coalescence of alleles, i.e. all alleles of a gene are descendant from a single ancestral allele which is not shared with other species • Emphasis is on the monophyly of genes rather than organisms • Species exist at the boundary of reticulating and hierarchical relationships • The smallest apomorphic group is a species • Applicable to sexual and asexual species	Avise & Wollenberg (1997) Cummings *et al.* (2008) Dettman *et al.* (2003, 2006)* Koufopanou *et al.* (1997)* O'Donnell *et al.* (2004)* Shaffer & Thomson (2007)
Diagnosable	• A species is an irreducible cluster of organisms that is diagnosably distinct from other such clusters	• Requires a qualitative and fixed difference with the closest relative • Applicable to sexual and asexual species	Nixon & Wheeler (1990) Hey *et al.* (2003)

© University of Pretoria

Genotypic cluster	• A species is a morphologically or genetically isolated group of populations that has few or no intermediates when in contact with other such groups	• Inapplicable to asexual species	Mallet (1995) Mallet *et al.* (2009)
"Polyphasic"	• A species conforms to at least two recognized species concepts	• Usually includes the morphological and/or biological species concept and one of the phylogenetic species concepts • Applicable to sexual and asexual species	Samson & Varga (2009) Gräser *at al.* (2008)*

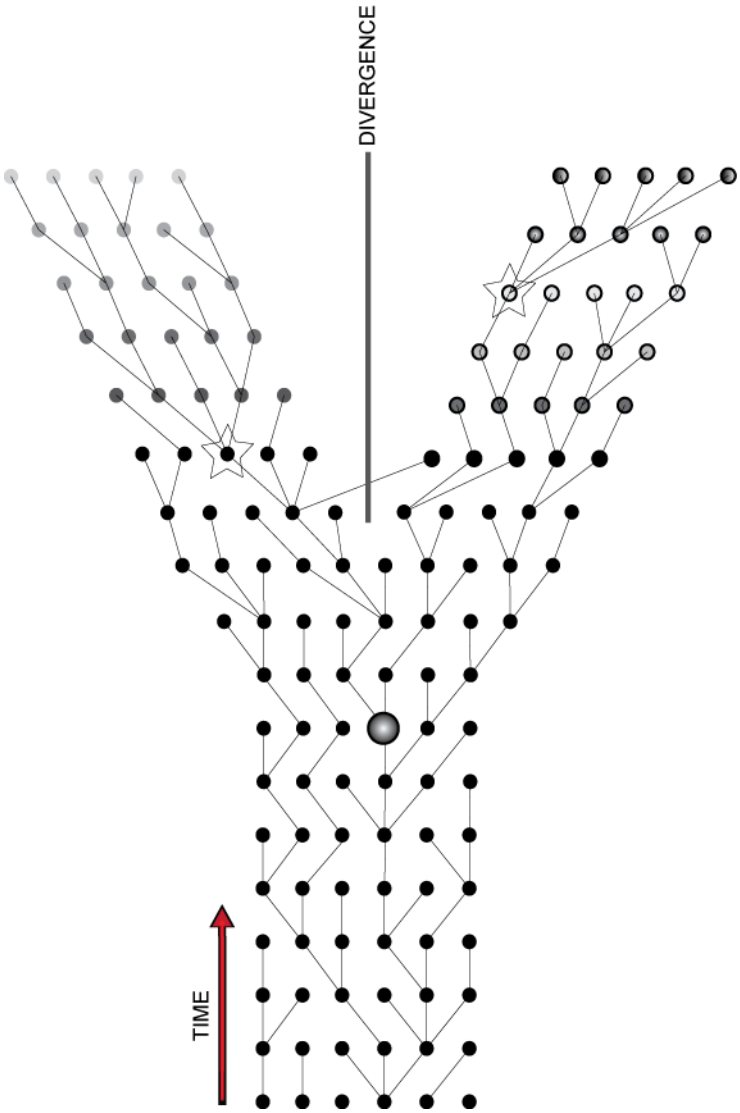
① The concepts and operational criteria listed here apply narrowly or broadly to species of fungi. For a more extensive list, please refer to Mayden (1997), while an exhaustive discussion of each can be found in Coyne & Orr (2004).

* References for the application of each concept as it relates to the fungi.

© University of Pretoria

Figure 1 The speciation process starts with a barrier to reproduction, leading to the divergence of lineages. Each dot represents a gene transmitted from one generation to the next with modification, and the generational tokogenetic relationships (Rieppel 2009) are indicated with solid lines. The most recent common ancestors of each of the resultant species are indicated with stars, while the most recent common ancestor of both lineages is indicated with a large dot.

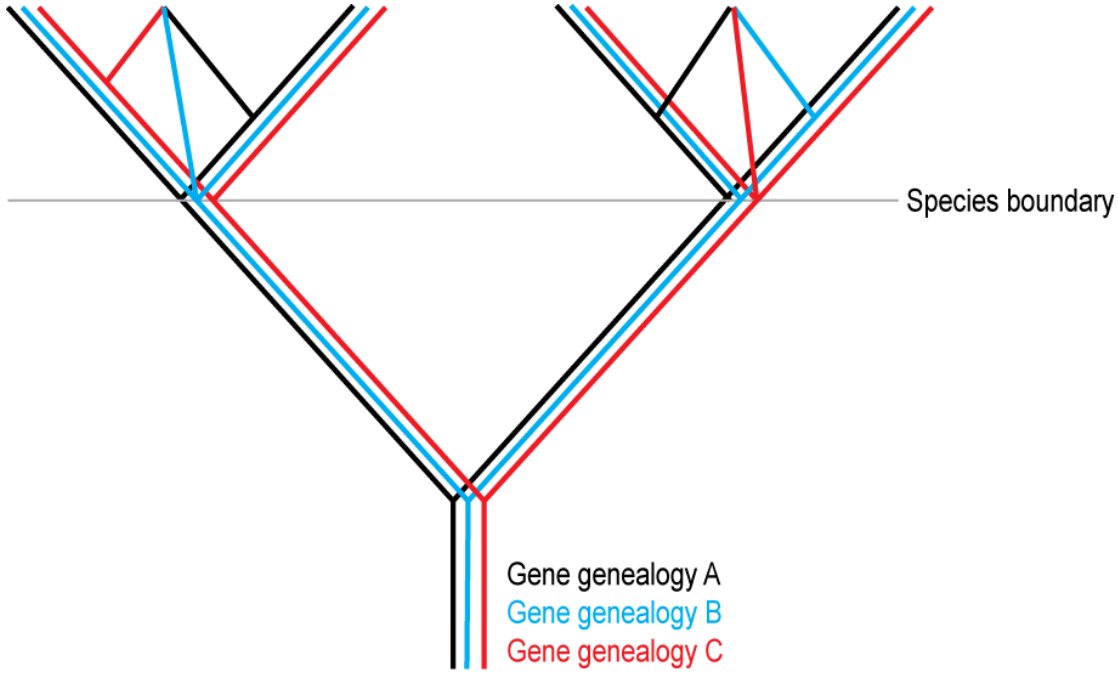
© University of Pretoria



© University of Pretoria

Figure 2 Genealogical concordance phylogenetic species recognition (GCPSR) (Avise & Wollenberg 1997; Sites & Marshall 2003; Taylor *et al.* 2000) relies on the simultaneous analysis of different gene regions, three in this case (gene genealogies A, B and C). The genealogies show a transition from concordance to incongruence at the species boundary. Thus, incongruence (i.e., disagreement) between gene genealogies can be used to delimit species boundaries. Adapted from Taylor *et al.* (2000).

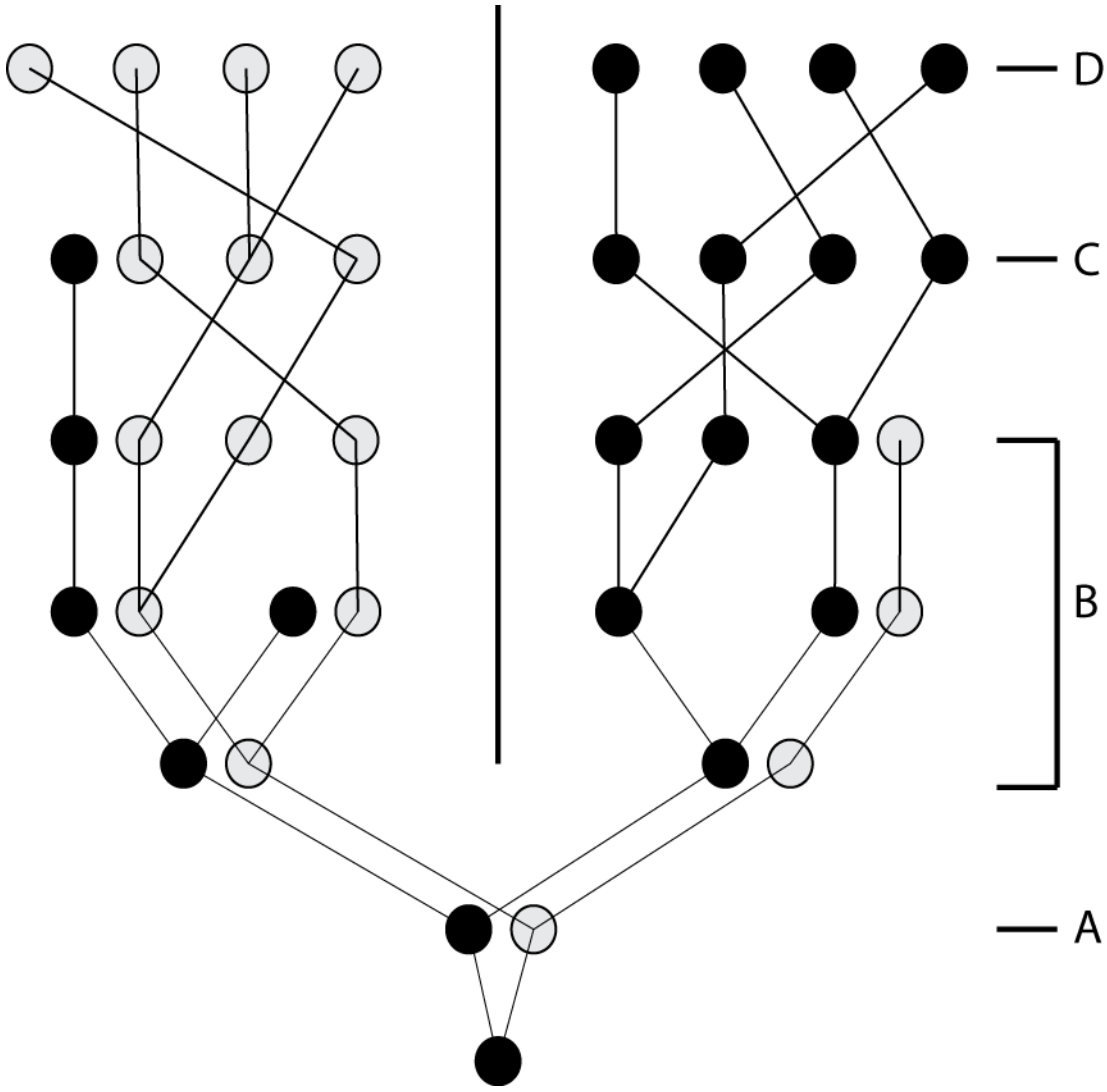
© University of Pretoria



© University of Pretoria

Figure 3 Lineage sorting results from differential fixation of alleles due to random drift or selection. A single locus becomes polymorphic and thus carries two alleles (A), after which the population is divided by an isolating mechanism. (B) Both lineages share the polymorphism, but (C) eventually one of the alleles become fixed first in the one population (one lineage displays incomplete sorting), then in the other. (D) Differential fixation of polymorphisms culminates in complete lineage sorting for alleles at the locus. Modified from Taylor (2000).

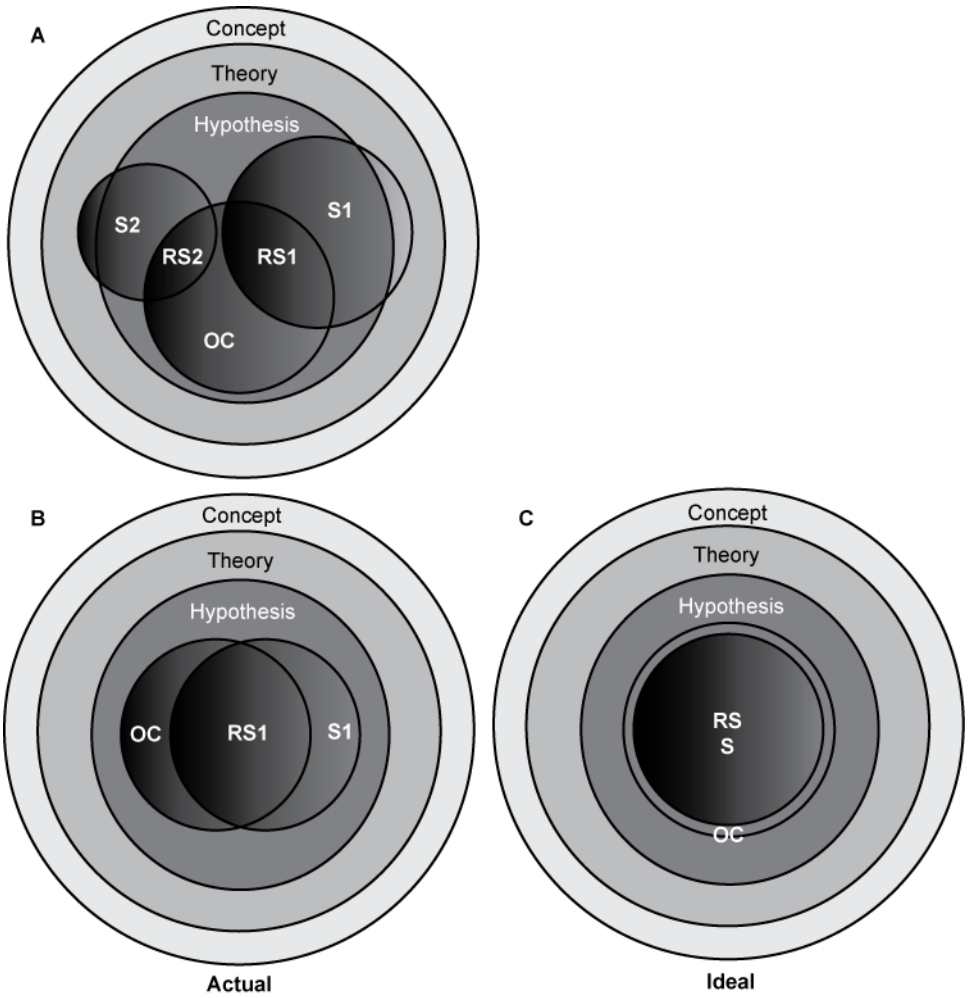
© University of Pretoria



© University of Pretoria

Figure 4 The organization of the levels of understanding regarding "species". The concept is a cognitive understanding of what a species is (i.e., an abstract idea or mental symbol), with mutable influence from other factors. A theory can be developed regarding species status; it is based on the concept, and can be implemented via a hypothesis. The species hypothesis is a more concrete understanding of which criteria would be useful to reject a species; it is based on previous empirical evidence and relies heavily on the theory developed from the concept. (A) Especially in the case of species complexes, the operational criteria (OC) to delimit natural Species 1 (S1) can potentially incorporate a sub-section or the whole of natural Species 2 (S2). The overlap between the operational criteria for these natural Species, and the Species themselves, can be denoted by the Realized Species (RS1 and RS2), i.e. that section of the natural Species that is empirical. However, in this situation it is impossible to separate RS1 and RS2, and thus the two species are treated under one species name. (B) With refinement of the operational criteria and/or species hypothesis, natural Species 2 can be excluded from the operational criteria applied to delimit natural Species 1 (S1). However, since the operational criteria may still exclude certain aspects of S1, the Realized Species does not overlap perfectly with the natural Species. (C) Ideally, the operational criteria should be broad enough to encompass the entire natural Species, but specific enough to exclude any other natural Species.

© University of Pretoria



© University of Pretoria

Bibliography

Aguileta G, Marthey S, Chiapello H, Lebrun M-H, Rudolphe F, Fournier E, Gendrault-Jacquemard A, Giraud T (2008) Assessing the performance of single-copy genes for recovering robust phylogenies. *Systematic Biology* **57**, 613-627.

Alfenas AC, Zauza EAV, Wingfield MJ, Roux J, Glen M (2005) *Heteropyxis natalensis*, a new host of *Puccinia psidii* rust. *Australasian Plant Pathology* **34**, 285-286.

Andersson JO (2006) Convergent evolution: gene sharing by eukaryotic plant pathogens. *Current Biology* **16**, R804-R806.

Andersson L (1990) The driving force: species concepts and ecology. *Taxon* **39**, 375-382.

Antonovics J, Hood M, Partain J (2002) The ecology and genetics of a host shift: *Microbotryum* as a model system. *The American Naturalist* **160**, S40-S53.

Avise JC, Wollenberg K (1997) Phylogenetics and the origin of species. *Proceedings of the National Academy of Science, USA* **94**, 7748-7755.

Ayala FJ (1982) *Population and evolutionary genetics: a primer.* The Benjamin/Cummings Publishing Company, Inc., Menlo Park, CA.

Baayen RP, O'Donnell K, Bonants PJM, Cigelnik E, Kroon LPNM, Roebroeck EJA, Waalwijk C (2000) Gene genealogies and AFLP analyses in the *Fusarium oxysporum* complex identify monophyletic and nonmonophyletic formae speciales causing wilt and rot disease. *Phytopathology* **90**, 891-900.

Bailey LH (1896) The philosophy of species-making. *Botanical Gazette* **22**, 454-462.

Bailey LH (2010) *The outlook to nature (classic reprint).* Forgotten Books.

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.

Banks JC, Paterson AM (2005) Multi-host parasite species in cophylogenetic studies. *International Journal of Parasitology* **35**, 741-746.

Barker MJ (2007) The empirical inadequacy of species cohesion by gene flow. *Philosophy of Science* **74**, 654-665.

© University of Pretoria

Barraclough TG, Nee S (2001) Phylogenetics and speciation. *TRENDS in Ecology and Evolution* **16**, 391-399.

Barton NH (2001) The role of hybridization in evolution. *Molecular Ecology* **10**, 551-568.

Beatty NP, Smith ML, Glass NL (1994) Molecular characterization of mating-type loci in selected homothallic species of *Neurospora*, *Gelasinospora* and *Anixiella*. *Mycological Research* **98**, 1309-1316.

Berbee ML, Taylor JW (2001) Fungal molecular evolution: gene trees and geologic time. In: *The Mycota VII Part B: Systematics and Evolution* (eds. McLaughlin DJ, McLaughlin EG, Lemke PA). Springer-Verlag, Berlin.

Berbee ML, Taylor JW (2010) Dating the molecular clock in fungi – how close are we? *Fungal Biology Reviews* **24**, 1-16.

Bistis GN (1981) Chemotropic interactions between trichogynes and conidia of opposite mating-type in *Neurospora crassa*. *Mycologia* **73**, 959-975.

Bistis GN (1983) Evidence for diffusable, mating-type specific trichogyne attractants in *Neurospora crassa*. *Experimental Mycology* **7**, 292-295.

Bliska JB, Casadevall A (2009) Intracellular pathogenic bacteria and fungi - a case of convergent evolution? *Nature Reviews Microbiology* **7**, 165-171.

Bond JE, Stockman AK (2009) An integrative method for delimiting cohesion species: finding the population-species interface in a group of Californian trapdoor spiders with extreme genetic divergence and geographic structuring. *Systematic Biology* **57**, 628-646.

Boysen M, Skouboe P, Frisvad J, Rossen L (1995) Reclassification of the *Penicillium roqueforti* group into three species on the basis of molecular genetic and biochemical profiles. *Microbiology* **142**, 541-549.

Brasier C (1997) Fungal species in practice: identifying species units in fungi. In: *Species: The units of Biodiversity*. (eds. Claridge MF, Dawah HA, Wilson MR), pp. 135-170. Chapman & Hall, UK.

Brasier C (2000) The rise of the hybrid fungi. *Nature* **405**, 134-135.

Brasier CM (2001) Rapid evolution of introduced pathogens via interspecific hybridization. *BioScience* **51**, 123-133.

© University of Pretoria

Brasier CM, Cooke DE, Duncan JM (1999) Origin of a new *Phytophthora* pathogen through interspecific hybridization. *Proceedings of the National Academy of Science, USA* **96**, 5878-5883.

Britz H, Coutinho TA, Wingfield MJ, Marasas WF, Gordon TR, Leslie JF (1999) *Fusarium subglutinans* f.sp. *pini* represents a distinct mating population in the *Gibberella fujikuroi* species complex. *Applied and Environmental Microbiology* **65**, 1198-1201.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proceedings of the National Academy of Science, USA* **93**, 770-773.

Butlin R (1987) A new approach to sympatric speciation. *TRENDS in Ecology and Evolution* **2**, 310-311.

Casselton LA (2002) Mate recognition in fungi. *Heredity* **88**, 142-147.

Clark JD, Haskins EF (1998) Heterothallic mating systems in the *Echinostelium minutum* complex. *Mycologia* **90**, 382-388.

Couch BC, Kohn LM (2002) A multilocus genealogy concordant with host preference indicates segregation of a new species, *Magnaporthe oryzae*, from *M. grisea*. *Mycologia* **94**, 683-693.

Coyne JA (2007) Sympatric speciation. *Current Biology* **17**, R787-R788.

Coyne JA, Orr HA (2004) *Speciation* Sinauer Associates, Inc., Sunderland, MA.

Creer S, Fonseca VG, Porazinska DL, Giblin-Davis RM, Sung W, Power DM, Packer M, Carvalho GR, Blaxter ML, Lambshead PJD, Thomas WK (2010) Ultrasequencing of the meiofaunal biosphere: practice, pitfalls and promises. *Molecular Ecology* **19**, 4-20.

Crous PW (2002) *Taxonomy and pathology of Cylindrocladium (Calonectria) and allied genera.* APS Press, St. Paul, Minnesota, USA.

Crous PW, Groenewald JZ, Risède J-M, Simoneau P, Hyde KD (2006) *Calonectria* species and their *Cylindrocladium* anamorphs: species with clavate vesicles. *Studies In Mycology* **55**, 213-226.

Crous PW, Groenewald JZ, Risède J-M, Simoneau P, Hywel-Jones NL (2004) *Calonectria* species and their *Cylindrocladium* anamorphs: species with sphaeropedunculate vesicles. *Studies In Mycology* **50**, 415-530.

© University of Pretoria

Crous PW, Hong L, Wingfield BD, Wingfield MJ (2001) ITS rDNA phylogeny of selected *Mycosphaerella* species and their anamorphs occurring on Myrtaceae. *Mycological Research* **105**, 425-431.

Crous PW, Summerell BA, Carnegie AJ, Wingfield MJ, Groenewald JZ (2009) Novel species of *Mycosphaerellaceae* and *Teratosphaeriaceae*. *Persoonia* **23**, 119-146.

Cummings MP, Neel MC, Shaw KL (2008) A genealogical approach to quantifying lineage divergence. *Evolution* **62**, 2411-2422.

Darwin C (1859) *On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life.* London: John Murray.

De Meeûs T, Durand P, Renaud F (2003) Species concepts: what for? *Trends in Parasitology* **19**, 425-427.

De Meeûs T, et al. (1998) Santa Rosalia revisited or why are there so many kinds of parasites in the "garden of earthly delights"? *Parasitology Today* **14**, 10-13.

De Meeûs T, Renaud F (2002) Parasites within the new phylogeny of eukaryotes. *Trends in Parasitology* **18**, 247-251.

De Queiros K, Donoghue MJ (1988) Phylogenetic systematics and the species problem. *Cladistics* **4**, 317-338.

De Queiroz K (1998) The general lineage concept of species, species criteria and the process of speciation. In: *Endless Forms: Species and Speciation*, pp. 57-75. Oxford University Press.

De Queiroz K (2007) Species concepts and species delimitation. *Systematic Biology* **56**, 879-886.

De Queiroz K (2011) Branches in the lines of descent: Charles Darwin and the evolution of the species concept. *Biological Journal of the Linnean Society* **103**, 19-35.

De Vienne DM, Hood ME, Giraud T (2009) Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology* **22**, 2532-2541.

De Vos L, Myburg AA, Wingfield MJ, Desjardins AE, Gordon TR, Wingfield BD (2007) Complete linkage maps from an interspecific cross between *Fusarium circinatum* and *Fusarium subglutinans*. *Fungal Genetics and Biology* **44**, 701-714.

Desjardins AE, Plattner RD, Gordon TR (2000) *Gibberella fujikuroi* mating population A and *Fusarium subglutinans* from teosinte species and maize from Mexico and Central America. *Mycological Research* **104**, 865-872.

© University of Pretoria

Dettman JR, Jacobson DJ, Taylor JW (2003) A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. *Evolution* **57**, 2703-2720.

Dettman JR, Jacobson DJ, Taylor JW (2006) Multilocus sequence data reveal extensive phylogenetic species diversity within the *Neurospora discreta* complex. *Mycologia* **98**, 436-446.

Dobzhansky T (1935) A critique of the species concept in biology. *Philosophy of Science* **2**, 344-355.

Dobzhansky T (1937) Genetic nature of species differences. *The American Naturalist* **71**, 404-420.

Dobzhansky T (1970) *Genetics of the evolutionary process.* Columbia University Press, New York.

Donoghue MJ (1985) A critique of the biological species concept and recommendations for a phylogenetic alternative. *The Bryologist* **88**, 172-181.

Dress A, Moulton V, Steel M, Wu T (2010) Species, clusters and the 'Tree of life': a graph-theoretic perspective. *Journal of Theoretical Biology* **265**, 535-542.

Dujon B (2010) Yeast evolutionary genomics. *Nature Reviews Genetics* **11**, 512-524.

Duncan T, Baum BR (1981) Numerical phenetics: its uses in botanical systematics. *Annual Review of Ecology and Systematics* **12**, 387-404.

Eisen JA (1998) Phylogenomics: improving functional predictions for uncharacterized genes by evolutionary analysis. *Genome Research* **8**, 163-167.

Eisen JA (2000) Horizontal gene transfer among microbial genomes: new insights from complete genome analysis. *Current Opinion in Genetics and Development* **10**, 606-611.

Eisen JA, Fraser CM (2003) Phylogenomics: intersection of evolution and genomics. *Science* **300**, 1706-1707.

Endersby J (2009) Lumpers and splitters: Darwin, Hooker, and the search for order. *Science* **326**, 1496-1499.

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

Falconer DS, Mackay TFC (1996) *Introduction to quantitative genetics*, 4th edn. Longman Group Ltd, Essex, UK.

Felsenstein J (1988) Phylogenies and quantitative characters. *Annual Review of Ecological Systematics* **19**, 445-471.

© University of Pretoria

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

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

Fitzpatrick BM, Fordyce JA, Gavrilets S (2009) Pattern, process and geographic modes of speciation. *Journal of Evolutionary Biology* **22**, 2342-2347.

Fontaine KM, Ahti T, Piercey-Normore MD (2010) Convergent evolution of *Cladonia gracilis* and allies. *The Lichenologist* **42**, 323-338.

Fourie G, Steenkamp ET, Ploetz RC, Gordon TR, Viljoen A (2011) Current status of the taxonomic position of *Fusarium oxysporum* formae specialis *cubense* within the *Fusarium oxysporum* complex. *Infection, Genetics and Evolution* **11**, 533-542.

Fournier E, Giraud T (2007) Sympatric genetic differentiation of a generalist pathogenic fungus, *Botrytis cinerea*, on two different host plants, grapevine and bramble. *Journal of Evolutionary Biology* **21**, 122-132.

Fraser JA, Diezmann S, Subaran RL, Allen A, Lengeler KB, Dietrich FS, Heitman J (2004) Convergent evolution of chromosomal sex-determining regions in the animal and fungal kingdoms. *PLoS Biology* **2**, e384.

Garbelotto M, Linzer R, Nicolotti G, Gonthier P (2010) Comparing the influences of ecological and evolutionary factors on the successful invasion of a fungal forest pathogen. *Biological Invasions* **12**, 943-957.

Gavrilets S, Waxman D (2002) Sympatric speciation by sexual conflict. *Proceedings of the National Academy of Sciences, USA* **99**, 10533-10538.

Geiser DM, Dorner JW, Horn BW, Taylor JW (2000) The phylogenetics of mycotoxin and sclerotium production in *Aspergillus flavus* and *Aspergillus oryzae*. *Fungal Genetics and Biology* **31**, 169-179.

Geiser DM, Pitt JI, Taylor JW (1998) Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. *Proceedings of the National Academy of Sciences, USA* **95**, 388-393.

Ghiselin MT (1987) Species concepts, individuality and objectivity. *Biology and Philosophy* **2**, 127-143.

© University of Pretoria

Gilbert GS, Webb CO (2007) Phylogenetic signal in plant pathogen-host range. *Proceedings of the National Academy of Science, USA* **104**, 4979-4983.

Gilbert SF, Opitz JM, Raff RA (1996) Resynthesizing evolutionary and developmental biology. *Developmental Biology* **173**, 357-372.

Gillespie JH (2000) Genetic drift in an infinite population: the pseudohitchhiking model. *Genetics* **155**, 909-919.

Giraud T, Refrégier G, Le Gac M, De Vienne DM, Hood ME (2008) Speciation in fungi. *Fungal Genetics and Biology* **45**, 791-802.

Glass NL, Kuldau GA (1992) Mating type and vegetative incompatibility in filamentous ascomycetes. *Annual Review of Phytopathology* **30**, 201-224.

Glass NL, Metzenberg RL, Raju NB (1990) Homothallic Sordariaceae from nature: the absence of strains containing only the *a* mating type sequence. *Experimental Mycology* **14**, 274-289.

Godfray HCJ, Mayo SJ, Scoble MJ (2008) Pragmatism and rigour can coexist in taxonomy. *Evolutionary Biology* **35**, 309-311.

Goodwin GP, Johnson-Laird PN (2011) Mental models of boolean concepts. *Cognitive Psychology* **63**, 34-59.

Gräser Y, Scott J, Summerbell R (2008) The new species concept in dermatophytes - a polyphasic approach. *Mycopathologia* **166**, 239-256.

Grigoriev IV (2011) JGI Fungal Genomics Program. U.S. Department of Energy Joint Genome Institute.

Groenewald M, Groenewald JZ, Braun U, Crous PW (2006) Host range of *Cercospora apii* and *C. beticola* and description of *C. apiicola*, a novel species from celery. *Mycologia* **98**, 275-285.

Grube M, Kroken S (2000) Molecular approaches and the concept of species and species complexes in lichenized fungi. *Mycological Research* **104**, 1284-1294.

Gryzenhout M, Wingfield BD, Wingfield MJ (2009) *Taxonomy, phylogeny, and ecology of bark-infecting and tree killing fungi in the Cryphonectriaceae.* APS Press.

Hanlin RT (1990) *Illustrated genera of Ascomycetes.* APS Press, St. Paul.

Harper JT, Gile GH, James ER, Carpenter KJ, Keeling PJ (2009) The inadequacy of morphology for species and genus delineation in microbial eukaryotes: an example from the parabasalian termite symbiont *Coronympha*. *PLoS ONE* **4**, e6577.

© University of Pretoria

Harrington TC (1981) Cycloheximide sensitivity as a taxonomic character in *Ceratocystis*. *Mycologia* **73**, 1123-1129.

Harrington TC, McNew DL (1997) Self-fertility and uni-directional mating-type switching in *Ceratocystis coerulescens*, a filamentous ascomycete. *Current Genetics* **32**, 52-59.

Harrington TC, Rizzo DM (1999) Defining species in the fungi. In: *Structure and Dynamics of Fungal Populations.* (ed. Worrall JJ), pp. 43-71. Kluwer Academic Press.

Harrison P (2008) *The fall of man and the foundations of science.* Cambridge University Press.

Hartl DL, Clark AG (1989) *Principles of Population Genetics, Second Edition.* Sinauer Associates, Inc., Sunderland, Massachusetts.

Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycological Research* **105**, 1422-1432.

Hawksworth DL (2011) Naming *Aspergillus* species: progress towards one name for each species. *Medical Mycology* **49 (Suppl. 1)**, S70-S76.

Hawksworth DL, Crous PW, Redhead SA, Reynolds DR, Samson RA, Seifert KA, Taylor JA, Wingfield MJ, Signatories (2011a) The Amsterdam Declaration on fungal nomenclature. *Mycotaxon* **116**, 491-500.

Hawksworth DL, Crous PW, Redhead SA, Reynolds DR, Samson RA, Seifert KA, Taylor JW, Wingfield MJ, Abaci Ö, Aime C, Asan A, Bai F-Y, De Beer ZW, Begerow D, Berikten D, Boekhout T, Buchanan PK, Burgess T, Buzina W, Cai L, Cannon PF, Crane JL, Damm U, Daniel H-M, Van Diepeningen AD, Druzhinina I, Dyer PS, Eberhardt U, Fell JW, Frisvad JC, Geiser DM, Geml J, Glienke C, Gräfenhan T, Groenewald JZ, Groenewald M, De Gruyter J, Guého-Kellermann E, Guo L-D, Hibbett DS, Hong S-B, De Hoog GS, Houbraken J, Huhndorf SM, Hyde KD, Ismail A, Johnston PR, Kadaifciler DG, Kirk PM, Kõljalg U, Kurtzman CP, Lagneau P-E, Lévesque CA, Liu X, Lombard L, Meyer W, Miller A, Minter DW, Najafzadeh MJ, Norvell L, Ozerskaya SM, Öziç R, Pennycook SR, Peterson SW, Pettersson OV, Quaedvlieg W, Robart VA, Ruibal C, Schnürer J, Schroers H-J, Shivas R, Slippers B, Spierenburg H, Takashima M, Taşkın E, Thines M, Thrane U, Uztan AH, Van Raak M, Varga J, Vasco A, Verkley G, Videira SIR, De Vries RP, Weir BS, Yilmaz N, Yurkov A, Zhang N (2011b) The Amsterdam Declaration on fungal nomenclature. *IMA Fungus* **2**, 105-112.

Hawksworth DL, Kirk P, Sutton B, Pegler D (1996) *Ainsworth and Bisby's Dicttionary of the Fungi.*, 8th edn. CABI, Wallingford, UK.

© University of Pretoria

Hedrick PW (2000) *Genetics of populations.*, 2nd edn. Jones and Bartlett Publishers, Inc., Sudbury, MA.

Hennig W (1950) *Grundzuge einer Theorie der phylogenetischen Systematik.* Deutscher Zentralverlag, Deutsches Entomologisches Institut, Berlin.

Hennig W (1976) *Phylogenetic systematics (translated by D. D. Davis and R. Zangerl).* University of Illinois Press, Urbana.

Hernandez-Bello MA, Chilvers MI, Akamatsu H, Peever TL (2006) Host specificity of *Ascochyta* spp. infecting legumes of the Viciae and Cicerae tribes and pathogenicity of an interspecific hybrid. *Phytopathology* **96**, 1148-1156.

Hewitt GW (2001) Speciation, hybrid zones and phylogeography - or seeing genes in space and time. *Molecular Ecology* **10**, 537-549.

Hey J (2001) The mind of the species problem. *TRENDS in Ecology and Evolution* **16**, 326-329.

Hey J (2006) On the failure of modern species concepts. *TRENDS in Ecology and Evolution* **21**, 447-450.

Hey J, Waples RS, Arnold ML, Butlin RK, Harrison RG (2003) Understanding and confronting species uncertainty in biology and conservation. *TRENDS in Ecology and Evolution* **18**, 597-603.

Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lücking R, Lumbsch HT, Lutzoni F, Matheny PB, McLaughlin DJ, Powell MJ, Redhead S, Schoch CL, Spatafora JW, Stalpers JA, Vilgalys R, Aime MC, Aptroot A, Bauer P, Begerow D, Benny GL, Castlebury LA, Crous PW, Dai Y-C, Gams W, Geiser DM, Griffith GW, Gueidan C, Hawksworth DL, Hestmark G, Hosaka K, Humber RA, Hyde KD, Ironside JE, Kõljalg U, Kurtzman KP, Larsson K-H, Lichtwardt R, Longcore J, Miądlikowska J, Miller A, Moncalvo J-M, Mozley-Standridge S, Oberwinkler F, Parmasto E, Reeb V, Rogers JD, Roux C, Ryvarden L, Sampaio JP, Schußler A, Sugiyama J, Thorn RG, Tibell L, Untereiner WA, Walker C, Wang Z, Weir A, Weiss M, White MM, Winka K, Yao Y-J, Zhang N (2007) A higher-level phylogenetic classification of the Fungi. *Mycological Research* **111**, 509-547.

Hibbett DS, Donoghue MJ (1998) Integrating phylogenetic analysis and classification in fungi. *Mycologia* **90**, 347-356.

Hibbett DS, Ohman A, Glotzer D, Nuhn M, Kirk P, Nilsson RH (2011) Progress in molecular and morphological taxon discovery in Fungi and options for formal classification of environmental sequences. *Fungal Biology Reviews* **25**, 38-47.

© University of Pretoria

Holsinger KE (1987) Pluralism and species concepts, or when must we agree with one another? *Philosophy of Science* **54**, 480-485.

Hooker CA (1975) Philosophy and meta-philosophy of science: empiricism, popperianism and realism. *Synthese* **32**, 177-231.

Horn BW, Ramirez-Prado JH, Carbone I (2009) Sexual reproduction and recombination in the aflatoxin-producing fungus *Aspergillus fumigatus*. *Fungal Genetics and Biology* **46**, 169-175.

Inderbitzen P, Davis RM, Bostock RM, Subbarao KV (2011) The ascomycete *Verticillium longisporum* is a hybrid and a plant pathogen with an expanded host range. *PLoS ONE* **6**, e18261.

Johannesson K (2001) Parallel speciation: a key to sympatric divergence. *TRENDS in Ecology and Evolution* **16**, 148-153.

Johanson A, Turner HC, McKay GJ, Brown AE (1998) A PCR-based method to distinguish fungi of the rice sheath-blight complex, *Rhizoctonia solani*, *R. oryzae* and *R. oryzae-sativae*. *FEMS Microbiology Letters* **162**, 289-194.

Johnson JA, Harrington TC, Engelbrecht CJB (2005) Phylogeny and taxonomy of the North American clade of the *Ceratocystis fimbriata* complex. *Mycologia* **97**, 1067-1092.

Jones S (1999) *Almost like a whale: The Origin of Species updated.* Doubleday, UK.

Jordan LK, Matyunina LV, McDonald JF (1999) Evidence for the recent horizontal transfer of a long terminal repeat retrotransposon. *Proceedings of the National Academy of Science, USA* **96**, 12621-12625.

Kidwell GK, Lisch D (1997) Transposable elements as sources of variation in animals and plants. *Proceedings of the National Academy of Science, USA* **94**, 7704-7711.

Kliman RM, Andolfatto P, Coyne JA, Depaulis F, Kreitman M, Berry AJ, McCarter J, Wakeley J, Hey J (2000) The population genetics of the origin and divergence of the *Drosophila simulans* complex species. *Genetics* **156**, 1913-1931.

Knowles LL, Carstens BC (2007) Delimiting species without monophyletic gene trees. *Systematic Biology* **56**, 887-895.

Koestler T, Ebersberger I (2011) Zygomycetes, Microsporidia, and the evolutionary ancestry of sex determination. *Genome Biology and Evolution* **3**, 186-194.

Kohn LM (2005) Mechanisms of fungal speciation. *Annual Review of Phytopathology* **43**, 279-308.

© University of Pretoria

Koufopanou V, Burt A, Szaro T, Taylor JW (2001) Gene genealogies, cryptic species, and molecular evolution in the human pathogen *Coccidioides immitis* and relatives (Ascomycota, Onygenales). *Molecular Biology and Evolution* **18**, 1246-1258.

Koufopanou V, Burt A, Taylor JW (1997) Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*. *Proceedings of the National Academy of Science, USA* **94**, 5478-5482.

Kück U, Pöggeler S (2009) Cryptic sex in fungi. *Fungal Biology Reviews* **23**, 86-90.

Kües U, Casselton LA (1992) Fungal mating type genes - regulators of sexual development. *Mycological Research* **96**, 993-1006.

Kunz W (2002) When is a parasite species a species? *Trends in Parasitology* **18**, 121-124.

Kvas M, Marasas WFO, Wingfield BD, Wingfield MJ, Steenkamp ET (2009) Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. *Fungal Diversity* **34**, 1-21.

Lawrence JG (1997) Selfish operons and speciation by gene transfer. *Trends in Microbiology* **5**, 355-359.

Lawrence JG, Roth JR (1996) Selfish operons: horizontal transfer may drive the evolution of gene clusters. *Genetics* **143**, 1843-1860.

Le Gac M, Hood ME, Giraud T (2007) Evolution of reproductive isolation within a parasitic fungal species complex. *Evolution* **61**, 1781-1787.

Lee MSY (2003) Species concepts and species reality: salvaging a Linnaean rank. *Journal of Evolutionary Biology* **16**, 179-188.

Leslie J, Zeller K, Wohler M, Summerell B (2004) Interfertility of two mating populations in the *Gibberella fujikuroi* species complex. *European Journal of Plant Pathology* **110**, 611-618.

Leslie JF, Zeller KA, Summerell BA (2001) Icebergs and species in populations of *Fusarium*. *Physiological and Molecular Plant Pathology* **59**, 107-117.

Lin X, Litvintseva P, Nielsen K, Patel S, Floyd A, Mitchell TG, Heitman J (2007) αADα hybrids of *Cryptococcus neoformans*: evidence of same-sex mating in nature and hybrid fitness. *PLoS Genetics* **3**, e186.

Linnakoski R, De Beer ZW, Ahtiainen J, Sidorov E, Niemelä P, Pappinen A, Wingfield MJ (2010) *Ophiostoma* spp. associated with pine- and spruce-infesting bark beetles in Finland and Russia. *Persoonia* **25**, 72-93.

© University of Pretoria

Liu M, Milgroom MG, Chaverri P, Hodge KT (2009) Speciation of a tropical fungal species pair following transoceanic dispersal. *Molecular Phylogenetics and Evolution* **51**, 413-426.

Lumbsch HT, Huhndorf SM (2007) Whatever happened to the pyrenomycetes and the loculoascomycetes? *Mycological Research* **111**, 1064-1074.

Lutzoni F, Kauff F, Cox CJ, CmcLaughlin D, Celio G, Dentinger B, Padamsee M, Hibbett D, James TY, Baloch E, Grube M, Reeb V, Hofstetter V, Schoch C, Arnold AE, Miadlikowska J, Spatafora J, Johnson D, Hambleton S, Crockett M, Shoemaker R, Sung GH, L\"u cR, Lumbsch T, O'Donnell K, Binder M, Diederich P, Ertz D, Gueidan C, Hansen K, Harris RC, Hosaka K, Lim YW, Matheny B, Nishida H, Pfister D, Rogers J, Rossman A, Schmitt I, Sipman H, Stone J, Sugiyama J, Yahr R, Vilgalys R (2004) Assembling the fungal tree of life: progress, classification, and evolution of subcellular traits. *American Journal of Botany* **91**, 1446-1480.

Ma L-J, Van der Does HC, Borkovich KA, Coleman JJ, Daboussi M-J, Di Pietro A, Dufresne M, Freitag M, Grabherr M, Henrissat B, Houterman PM, Kang S, Shim W-B, Woloshuk C, Xie X, Xu J-R, Antoniw J, Baker SE, Bluhm BH, Breakspear A (2010) Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* **464**, 367-373.

Maclean CJ, Greig D (2010) Reciprocal gene loss following experimental whole-genome duplication causes reproductive isolation in yeast. *Evolution* **65**, 932-945.

Mallet J (1995) A species definition for the modern synthesis. *TRENDS in Ecology and Evolution* **10**, 294-299.

Mallet J, Meyer A, Nosil P, Feder JL (2009) Space, sympatry and speciation. *Journal of Evolutionary Biology* **22**, 2332-2341.

Mallet LV, Becq J, Deschavanne P (2010) Whole genome evaluation of horizontal transfers in the pathogenic fungus *Aspergillus fumigatus*. *BMC Genomics* **11**, 171.

May RM, Beverton RJH (1990) How many species? *Philosophical Transactions of the Royal Society of London B* **330**, 293-304.

Mayden RL (1997) A hierarchy of species concepts: the denouement in the saga of the species problem. In: *Species: The Units of Biodiversity* (eds. Claridge MF, Dawah HA, Wilson MR), pp. 381-424. Chapman and Hall, London.

Mayr E (1942) *Systematics and the origin of species.* Colombia University Press, New York.

Mayr E (1947) Ecological factors in speciation. *Evolution* **1**, 263-288.

© University of Pretoria

Mayr E (1953) Concepts of classification and nomenclature in higher organisms and microorganisms. *Annals of the New York Academy of Sciences* **56**, 391-397.

Mayr E (1963) *Animal species and evolution.* Belknap Press of Harvard University Press, Cambridge, Massachusetts.

Mayr E (1970) *Populations, species and evolution.* Belknap Press of Harvard University Press, New York.

Mayr E, Linsley EG, Usinger RL (1953) *Methods and principles of systematic zoology*. McGraw-Hill, New York.

Mayr E, Provine WB (1981) The evolutionary synthesis. *Bulletin of the American Academy of Arts and Sciences* **34**, 17-32.

McCoy KD (2003) Sympatric speciation in parasites - what is sympatry? *Trends in Parasitology* **19**, 400-404.

McGuire IC, Marra RE, Turgeon BG, Milgroom MG (2001) Analysis of mating-type genes in the chestnut blight fungus, *Cryphonectria parasitica*. *Fungal Genetics and Biology* **34**, 131-144.

Mehrabi R, Bahkali AH, Abd-Elsalam KA, Moslem M, Ben M'Barek S, Gohari AM, Jashni MK, Stergiopoulus I, Kema GHJ, De Wit PJGM (2011) Horizontal gene and chromosome transfer in plant pathogenic fungi affecting host range. *FEMS Microbiology Reviews* **35**, 542-554.

Miller AN, Huhndorf SM (2004) Using phylogenetic species recognition to delimit species boundaries within *Lasiosphaeria*. *Mycologia* **96**, 1106-1127.

Miller JS, Wenzel JW (1995) Ecological characters and phylogeny. *Annual Review of Entomology* **40**, 389-415.

Mishler BD (1999) Getting rid of species? In: *Species: New Interdisciplinary Essays* (ed. Wilson R), pp. 305-315. MIT Press, USA.

Mishler BD, Brandon RN (1987) Individuality, pluralism and the phylogenetic species concept. *Biology and Philosophy* **2**, 397-414.

Mishler BD, Donoghue MJ (1982) Species concepts: a case for pluralism. *Systematic Zoology* **31**, 491-503.

Morris SC (2010) Evolution: like any other science it is predictable. *Philosophical Transactions of the Royal Society B* **365**, 133-145.

Nee S (2003) Unveiling prokaryotic diversity. *TRENDS in Ecology and Evolution* **18**, 62-63.

© University of Pretoria

Nielsen CB, Cantor M, Dubchak I, Gordon D, Wang T (2010) Visualizing genomes: techniques and challenges. *Nature Methods Supplement* **7**, S1-S11.

Nixon KC, Wheeler QD (1990) An amplification of the phylogenetic species concept. *Cladistics* **6**, 211-223.

Norvell LL (2011) Melbourne approves a new Code. *Mycotaxon* **116**, 481-490.

O'Donnell AG, Goodfellow M, Hawksworth DL (1994) Theoretical and practical aspects of the quantification of biodiversity among microorganisms. *Philosophical Transactions: Biological Sciences* **345**, 65-73.

O'Donnell K, Ward TJ, Geiser DM, Kistler HC, Aoki T (2004) Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genetics and Biology* **41**, 600-623.

Oberhaensli S, Parlange F, Buchmann JP, Jenny FH, Abbott JC, Burgis TA, Spanu PD, Keller B, Wicker T (2011) Comparative sequence analysis of wheat and barley powdery mildew fungi reveals gene colinearity, dates divergence and indicates host-pathogen co-evolution. *Fungal Genetics and Biology* **48**, 327-334.

Olmstead RG (1995) Species concepts and plesiomorphic species. *Systematic Botany* **20**, 623-630.

Olsen A, Stenlid J (2002) Pathogenic fungal species hybrids infecting plants. *Microbes and Infection* **4**, 1353-1359.

Orr HA, Turelli M (2001) The evolution of postzygotic isolation: accumulating Dobzhansky-Muller incompatibilities. *Evolution* **55**, 1085-1094.

Paoletti M, Rydholm C, Schwier EU, Anderson MJ, Szakacs G, Lutzoni F, Debeaupuis J-P, Latgé JP, Denning DW, Dyer PS (2005) Evidence for sexuality in the opportunistic pathogen *Aspergillus fumigatus*. *Current Biology* **15**, 1242-1248.

Paterson HE (1980) A comment on "Mate Recognition Systems". *Evolution* **34**, 330-331.

Patterson DJ (1999) The diversity of eukaryotes. *American Naturalist* **65**, S96-S124.

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ (2008) Multiple gene genealogies and phenotypic data reveal cryptic species of the Botryosphaeriaceae: A case study on the *Neofusicoccum parvum* / *N. ribis* complex. *Molecular Phylogenetics and Evolution* **51**, 259-268.

© University of Pretoria

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ (2009) Molecular and phenotypic characterization of three phylogenetic species discovered within the *Neofusicoccum parvum*/*N. ribis* complex. *Mycologia* **101**, 636-647.

Pegg GS, Gryzenhout M, O'Dwyer C, Drenth A, Wingfield MJ (2010) The *Eucalyptus* canker pathogen *Chrysoporthe cubensis* discovered in eastern Australia. *Australasian Plant Pathology* **39**, 343-349.

Pérez CA, De Beer ZW, Altier NA, Wingfield MJ, Blanchette RA (2008) Discovery of the eucalypt pathogen *Quambalaria eucalypti* infecting a non-*Eucalyptus* host in Uruguay. *Australasian Plant Pathology* **37**, 600-604.

Perkins DD (1987) Mating-type switching in filamentous ascomycetes. *Genetics* **115**, 215-216.

Petit RJ, Excoffier L (2009) Gene flow and species delimitation. *TRENDS in Ecology and Evolution* **24**, 386-393.

Picard M, Debuchy R, Coppin E (1991) Cloning the mating types of the heterothallic fungus *Podospora anserina*: developmental features of haploid transformants carrying both mating types. *Genetics* **128**, 539-547.

Pigliucci M (2007) Do we need an extended evolutionary synthesis? *Evolution* **61**, 2743-2749.

Pöggeler S (2002) Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. *Current Genetics* **42**, 153-160.

Politzer G, Macchi L (2000) Reasoning and pragmatics. *Mind & Society* **1**, 73-93.

Pringle A, Baker DM, Platt JL, Wares JP, Latgé JP, Taylor JW (2005) Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. *Evolution* **59**, 1886-1899.

Quine WV (1951) Main trends in recent philosophy: two dogmas of empiricism. *The Philosophical Review* **60**, 20-43.

Redecker D (2002) New views on fungal evolution based on DNA markers and the fossil record. *Research in Microbiology* **153**, 125-130.

Ribeiro F, Caticha N (2008) Emergence and loss of assortative mating in sympatric speciation. *Journal of Theoretical Biology* **258**, 465-477.

Richards TA, Soanes DM, Foster PG, Leonard G, Thornton CR, Talbot NJ (2009) Phylogenomic analysis demonstrates a pattern of rare and ancient horizontal gene transfer between plants and fungi. *The Plant Cell* **21**, 1897-1911.

© University of Pretoria

Richards TA, Soanes DM, Jones MDM, Vasieva O, Leonard G, Paszkiewicz K, Foster PG, Hall N, Talbot NJ (2011) Horizontal gene transfer facilitated the evolution of plant parasitic mechanisms in the oomycetes. *Proceedings of the National Academy of Science, USA* **108**, 15258-15263.

Rieppel O (2009) Hennig's enkaptic system. *Cladistics* **25**, 311-317.

Roets F, Wingfield MJ, Crous PW, Dreyer LL (2009) Fungal radiation in the Cape Floristic Region: an analysis based on *Gondwanamyces* and *Ophiostoma*. *Molecular Phylogenetics and Evolution* **51**, 111-119.

Roselló-Mora R, Amann R (2001) The species concept for prokaryotes. *FEMS Microbiology Reviews* **25**, 39-67.

Rosenberg NA (2002) The probability of topological concordance of gene trees and species trees. *Theoretical Population Biology* **61**, 225-247.

Rosewich UL, Kistler HC (2000) Role of horizontal gene transfer in the evolution of fungi. *Annual Review of Phytopathology* **38**, 325-363.

Samson RA, Houbraken J, Varga J, Frisvad JC (2009) Polyphasic taxonomy of the heat resistant ascomycete genus *Byssochlamus* and its *Paecilomyces* anamorphs. *Persoonia* **22**, 14-27.

Samson RA, Varga J (2009) What is a species in Aspergillus? *Medical Mycology* **47**, S13-S20.

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

Schleifer KH (2009) Classification of *Bacteria* and *Archaea*: past, present and future. *Systematic and Applied Microbiology* **32**, 533-542.

Schluter D (2009) Evidence for ecological speciation and its alternative. *Science* **323**, 737-740.

Schmitt I, Lumbsch HT (2009) Ancient horizontal gene transfer from bacteria enhances biosynthetic capabilities of fungi. *PLoS ONE* **4**, e4437.

Schoch C, Crous PW, Groenewald JZ, Boehm EWA, Burgess TI, De Gruyter J, De Hoog GS, Dixon LJ, Grube M, Gueidan C, Harada Y, Hatakeyama S, Hirayama K, Hosoya T, Huhndorf SM, Hyde KD, Jones EBG, Kohlmeyer J, Kruys Å, Li YM, Lücking R, Lumbsch HT, Marvanová L, Mbatchou JS, McVay AH, Miller AN, Mugambi GK, Muggia L, Nelsen MP, Nelson P, Owensby CA, Phillips AJL, Phongpaichit S, Pointing SB, Pujade-Renaud V, Raja HA, Rivas Plada E, Robbertse B, Ruibal C, Sakayaroj J, Sano T, Selbmann L, Shearer CA, Shirouzu T, Slippers B, Suetrong S, Tanaka K, Volkmann-Kohlmeyer B, Wingfield MJ, Wood AR, Woudenberg JHC, Yonezawa H,

© University of Pretoria

Zhang Y, Spatafora JW (2009a) A class-wide phylogenetic assessment of *Dothideomycetes*. *Studies In Mycology* **64**, 1-15.

Schoch CL, Sung GH, López-Giráldez F, Townsend JP, Miadlikowska J, Hofstetter V, Robbertse B, Matheny PB, Kauff F, Wang Z, Gueidan C, Andrie RM, Trippe K, Ciufetti LM, Wynns A, Fraker E, Hodkinson BP, Bonito G, Groenewald JZ, Arzanlou M, De Hoog GS, Crous PW, Hewitt D, Pfister DH, Peterson K, Gryzenhout M, Wingfield MJ, Aptroot A, Suh S-O, Blackwell M, Hillis DM, Griffith GW, Castlebury LA, Rossman AY, Lumbsch HT, Lücking R, Büdel B, Rauhut A, Diederich P, Ertz D, Geiser DM, Hosaka K, Inderbitzen P, Kohlmeyer J, Volkmann-Kohlmeyer B, Mostert L, O'Donnell K, Sipman H, Rogers JD, Shoemaker RA, Sugiyama J, Summerbell RC, Untereiner W, Johnston PR, Stenroos S, Zuccaro A, Dyer PS, Crittenden PD, Cole MS, Hansen K, Trappe JM, Yahr R, Lutzoni F, Spatafora JW (2009b) The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. *Systematic Biology* **58**, 224-239.

Seixas CDS, Barreto RW, Alfenas AC, Ferreira FA (2004) *Cryphonectria cubensis* on an indigenous host in Brazil: a possible origin for eucalyptus canker disease? *Mycologist* **18**, 39-45.

Shaffer HB, Thomson RC (2007) Delimiting species in recent radiations. *Systematic Biology* **56**, 896-906.

Sharon A, Yamaguchi K, Christiansen S, Horwitz BA, Yoder OC, Turgeon BG (1996) An asexual fungus has the potential for sexual development. *Molecular and General Genetics* **251**, 60-68.

Simpson GG (1951) The species concept. *Evolution* **5**, 285-298.

Sites JW, Marshall JC (2003) Delimiting species: a renaissance issue in systematic biology. *TRENDS in Ecology and Evolution* **18**, 462-470.

Slepecky RA, Starmer WT (2009) Phenotypic plasticity in fungi: a review with observations on *Aureobasidium pullulans*. *Mycologia* **101**, 823-832.

Slippers B, Stenlid J, Wingfield MJ (2005) Emerging pathogens: fungal host jumps following anthropogenic introduction. *TRENDS in Ecology and Evolution* **20**, 420-421.

Slot JC, Rokas A (2010) Multiple *GAL* pathway gene clusters evolved independently and by different mechanisms in fungi. *Proceedings of the National Academy of Science, USA* **107**, 10136-10141.

Sneath PHA (1957) Some thoughts on bacterial classification. *Journal of General Microbiology* **17**, 184-200.

© University of Pretoria

Snell-Rood EC, Van Dyken JD, Cruickshank T, Wade MJ, Moczek AP (2010) Toward a population genetic framework of developmental evolution: the costs, limits, and consequences of phenotypic plasticity. *BioEssays* **32**, 71-81.

Sokal RR, Crovello TJ (1970) The biological species concept: a critical evaluation. *The American Naturalist* **104**, 127-153.

Soll DR (2002) *Candida* commensalism and virulence: the evolution of phenotypic plasticity. *Acta Tropica* **81**, 101-110.

Spatafora JW, Sung GH, Johnson D, Hesse C, O'Rourke B, Serdani M, Spotts R, Lutzoni F, Hofstetter V, Miadlikowska J, Reeb V, Gueidan C, Fraker E, Lumbsch T, Lücking R, Schmitt I, Hosaka K, Aptroot A, Roux C, Miller AN, Geiser DM, Hafellner J, Hestmark G, Arnold AE, Büdel B, Rauhut A, Hewitt D, Untereiner WA, Cole MS, Scheidegger C, Schultz M, Sipman H, Schoch CL (2006) A five-gene phylogeny of Pezizomycotina. *Mycologia* **98**, 1018-1028.

Steenkamp ET, Wingfield BD, Desjardins AE, Marasas WFO, Wingfield MJ (2002) Cryptic speciation in *Fusarium subglutinans*. *Mycologia* **94**, 1032-1043.

Stielow B, Bratek Z, Orczán AKI, Rudnoy S, Hensel G, Hoffman PG, Klenk H-P, Göker M (2011) Species delimitation in taxonomically difficult fungi: the case of *Hymenogaster*. *PLoS ONE* **6**, e15614.

Stockman AK, Bond JE (2007) Delimiting cohesion species: extreme population structuring and the role of ecological exchangeability. *Molecular Ecology* **16**, 3374-3392.

Strandberg R, Nygren K, Menkis A, James TY, Wik L, Stajich JE, Johannesson H (2010) Conflict between reproductive gene trees and species phylogeny among heterothallic and pseudohomothallic members of the filamentous ascomycete genus *Neurospora*. *Fungal Genetics and Biology* **47**, 869-878.

Summerell BA, Laurence MH, Liew ECY, Leslie JF (2010) Biogeography and phylogeography of *Fusarium*: a review. *Fungal Diversity* **44**, 3-13.

Tautz D, et al. (2003) A plea for DNA taxonomy. *TRENDS in Ecology and Evolution* **18**, 70-74.

Taylor JW, Geiser DM, Burt A, Koufopanou V (1999a) The evolutionary biology and population genetics underlying fungal strain typing. *Clinical Microbiology Reviews* **12**, 126-146.

Taylor JW, Jacobson DJ, Fisher MC (1999b) The evolution of asexual fungi: reproduction, speciation and classification. *Annual Review of Phytopathology* **37**, 197-246.

© University of Pretoria

Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS, Fisher MC (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology* **31**, 21-32.

Taylor JW, Turner E, Townsend JP, Dettman JR, Jacobson D (2006) Eukaryotic microbes, species recognition and the geographic limits of species: examples from the kingdom Fungi. *Philosophical Transactions of the Royal Society B* **361**, 1947-1963.

Taylor TN, Hass H, Kerp H (1999c) The oldest fossil ascomycetes. *Nature* **399**, 648.

Templeton AR (1998) Nested clade analysis of phylogeographic data: testing hypotheses about gene flow and population history. *Molecular Ecology* **7**, 381-397.

Tian C, Gregersen PK, Seldin MF (2008) Accounting for ancestry: population substructure and genome-wide association studies. *Human Molecular Genetics* **17**, R143-R150.

Van der Merwe NA, Gryzenhout M, Steenkamp ET, Wingfield BD, Wingfield MJ (2010) Multigene phylogenetic and population differentiation data confirm the existence of a cryptic species within *Chrysoporthe cubensis*. *Fungal Biology* **114**, 966-979.

Van Regenmortel MHV (1989) Applying the species concept to plant viruses. *Archives of Virology* **104**, 1-17.

Van Valen L (1976) Ecological species, multispecies, and oaks. *Taxon* **25**, 233-239.

Van Wyk M, Roux J, Barnes I, Wingfield BD, Wingfield MJ (2006) Molecular phylogeny of the *Ceratocystis moniliformis* complex and description of *C. tribiliformis sp. nov*. *Fungal Diversity* **21**, 181-201.

Venter M, Myburg H, Wingfield BD, Coutinho TA, Wingfield MJ (2002) A new species of *Cryphonectria* from South Africa and Australia, pathogenic to *Eucalyptus*. *Sydowia* **54**, 98-117.

Wakeley J (2004) Recent trends in population genetics: More data! More math! Simple models! *Journal of Heredity* **95**, 397-405.

Walton JD (2000) Horizontal gene transfer and the evolution of secondary metabolite gene clusters in fungi: a hypothesis. *Fungal Genetics and Biology* **30**, 167-171.

Wang H, Xu Z, Gao L, Hao B (2009) A fungal phylogeny based on 82 complete genomes using the composition vector method. *BMC Evolutionary Biology* **9**, 195.

Wheeler QD, Nixon KC (1990) Another way of looking at the species problem: a reply to De Queiroz and Donoghue. *Cladistics* **6**, 77-81.

© University of Pretoria

Wiens JJ (2004) What is speciation and how should we study it? *The American Naturalist* **163**, 914-923.

Wilson BE (1995) A (not-so-radical) solution to the species problem. *Biology and Philosophy* **10**, 339-356.

Wilson EO (2003) The encyclopaedia of life. *TRENDS in Ecology and Evolution* **18**, 77-80.

Wingfield MJ, Van Wyk PS (1993) A new species of *Ophiostoma* from *Protea* infructescences in South Africa. *Mycological Research* **97**, 709-716.

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

Wright S (1943) Isolation by distance. *Genetics* **28**, 114-138.

Wu C-I (2001) The genic view of the process of speciation. *Journal of Evolutionary Biology* **14**, 851-865.

Xu J-R, Peng Y-L, Dickman MB, Sharon A (2006) The dawn of fungal pathogen genomics. *Annual Review of Phytopathology* **44**, 337-366.

Xu M, He X (2011) Genetic incompatibility dampens hybrid fertility more than hybrid viability: yeast as a case study. *PLoS ONE* **6**, e18341.

Yang C, Sakai H, Numa H, Itoh T (2011) Gene tree discordance of wild and cultivated Asian rice deciphered by genome-wide sequence comparison. *Gene* **447**, 53-60.

Yun SH, Arie T, Kaneko I, Yoder OC, Turgeon BG (2000) Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella*/*Fusarium* species. *Fungal Genetics and Biology* **31**, 7-20.

Zeyl C (2009) The role of sex in fungal evolution. *Current Opinion in Microbiology* **12**, 592-598.

Zhang J (2010) Ancestral informative marker selection and population structure visualization using sparse laplacian Eigenfunctions. *PLoS ONE* **5**, e13734.

Zhang J, Chiodini R, Badr A, Zhang G (2011) The impact of next-generation sequencing on genomics. *Journal of Genetical Genomics* **38**, 95-109.

Zhang N, Castlebury LA, Miller AN, Huhndorf SM, Schoch CL, Seifert KA, Rossman AY, Rogers JD, Kohlmeyer J, Volkmann-Kohlmeyer B, Sung G-H (2006) An overview of the systematics of the Sordariomycetes based on a four-gene phylogeny. *Mycologia* **98**, 1076-1087.

© University of Pretoria

Zhang P, Sheng H, Uehara R (2004) A double classificaton tree search algorithm for index SNP selection. *BMC Bioinformatics* **5**, 89.

© University of Pretoria

CHAPTER 2

Primers for the amplification of sequence-characterized loci in *Cryphonectria cubensis* populations[1,2]

[1] **Van der Merwe, N. A., Wingfield, B. D., and Wingfield, M. J.** (2003) Primers for the amplification of sequence-characterized loci in *Cryphonectria cubensis* populations. *Molecular Ecology Notes*, **3**, 494-497.

[2] The text from this publication is reproduced in full, although it has been reformatted to conform with the layout of the rest of the thesis. Since the time of publication, *Cryphonectria cubensis* had been subdivided into several *Chrysoporthe* species.

© University of Pretoria

Abstract

We describe the development of DNA markers for the fungal pathogen of *Eucalyptus*, *Cryphonectria cubensis*. These markers originated from cloned intershort sequence repeat polymerase chain reactions, which enrich for medium to highly repetitive DNA sequences. In total, 10 markers were isolated, eight of which were polymorphic, and these can subsequently be applied to study populations of *C. cubensis*.

© University of Pretoria

Introduction

Cryphonectria cubensis causes a debilitating stem canker disease on *Euclayptus* trees in many countries of South America (Hodges *et al*. 1979; Van Zyl *et al*. 1998), Africa (Wingfield *et al*. 1989) and Southeast Asia (Davison & Coates 1991). Vegetative compatibility groups have been used to study some populations of this pathogen to consider questions of origin and local genetic diversity (Van Heerden *et al*. 1997; Van Zyl *et al*. 1998; Van Zyl *et al*. 1994; Wingfield *et al*. 1997). This technique makes possible the identification of individual phenotypes that are capable of hyphal anastomosis at any point of contact (Leslie 1993; Newhouse & MacDonald 1991). Screening is thus for genotypes based on an unknown number of loci. Although the technique has many advantages, lack of resolution precludes the possibility of deducing population parameters such as genetic diversity, outcrossing rates and migration patterns (Cortesi *et al*. 1996; Milgroom *et al*. 1993). The aim of the present study was therefore to develop co-dominant DNA markers that would make it possible to analyse populations of *C. cubensis* at a considerably higher level of resolution than was previously possible.

Materials & Methods

Genomic DNA was extracted from three *C. cubensis* isolates (CMW6112, CMW8856 and CMW8890) using a CTAB extraction protocol (Murray & Thompson 1980). An intershort sequence repeat (ISSR) polymerase chain reaction (PCR) was performed on extracted DNA, using four different ISSR primers (Table 1). All ISSR PCR reactions were performed using the protocol of Hantula, Dusabenyagasani & Hamelin (1996). Amplification products were cloned into the pGEMT-Easy plasmid (Promega Corporation), transformed into and propagated in *Escherichia coli* JM109 cells. Plasmids were extracted and purified (Sambrook *et al*. 1989) and insert sizes were determined using a restriction digest with *Eco*RI followed by agarose gel electrophoresis. Inserts of different sizes and smaller than 600 base pairs (bp), were cycle sequenced using the BigDye Dye Terminator Kit (Perkin Elmer) with M13 primers SP6 and T7, followed by electrophoresis using an ABI Prism 377 automated sequencing instrument. Specific primers were designed to flank medium repetitive

© University of Pretoria

sequences present in several inserts, and were used to amplify the loci from genomic DNA of the original *C. cubensis* isolates from South Africa and Mexico, as well as isolates from Colombia, Republic of Congo, Indonesia and Vietnam. Each 25-μL PCR reaction contained 1 ng/ μL genomic DNA, 10 mM $MgCL_2$, 2·5 mM of each dNTP, 1 × PCR buffer (Southern Cross Biotechnology), 0·8 M 2-pyrrolidinone (Chakrabarti & Schutt 2001), 40 mM of each primer, and 1 U SuperTherm *Taq* Polymerase (Southern Cross Biotechnology). PCR was performed on a GeneAmp® PCR System 9700 (Applied Biosystems), and was initiated with denaturation at 94°C for 2 min, followed by 30 cycles of 54°C for 30 s, 72°C for 1 min and 94°C for 1 min. PCR was completed with a final extension at 72°C for 7 min.

Sequences of inserts from which primers were designed were used in BLAST searched of the GenBank database (Altschul *et al*. 1990), to determine their similarity to known sequences. Two separate searches were performed for each sequence: (i) a BLAST search restricted to the fungi and (ii) a BLAST search restricted to *Saccharomyces cerevisiae*.

Results

After agarose gel electrophoresis (1% w/v) and visualisation using ethidium bromide and ultraviolet light, PCR products from isolates CMW6112 and CMW8856 with primers $(CGA)_5$ and $(CAC)_5$ revealed easily distinguishable fingerprints. Primer $(GT)_8$ yielded unique fingerprints for CMW6112 and CMW8890, but primer $(CCA)_5$ did not provide unique fingerprints for the different isolates. After cloning the PCR amplification products into the plasmid, 307 recombinant colonies were recovered. Restriction digests of purified plasmid DNA revealed that 58 clones contained inserts smaller than 600 bp. Of these, 22 inserts were of different size. Primers were designed from these 22 sequences. Only 10 of the 22 primer pairs consistently resulted in the amplification of the desired targets across all isolates during PCR (Table 1).

GeneScan® analysis of 56 test isolates revealed that four of the candidate loci were monomorphic based on size. These loci, namely *COL4*, *COL6*, *SA2* and *SA10*, may still contain sequence polymorphisms not detectable by the GeneScan® technique. The remaining six primer pairs amplified varying numbers of alleles from the 56 test

© University of Pretoria

isolates (Table 2) and therefore, varied in their levels of polymorphism. Primer pairs for loci *SA6* and *SA9* each yielded PCR products from two loci and in both cases these loci were polymorphic.

BLAST searches using the fungal database revealed that sequences for two loci, namely *SA1* and *SA3*, were homologous to a *Neurospora crassa* sequence (GenBank accession AL356815) (Table 1). A search on the *S. cerevisiae* database showed that several of the *C. cubensis* sequences were homologous to sequences of *S. cerevisiae* chromosome XV.

Discussion

In this study, we have succeeded in developing 10 primer sets that amplify loci from *C. cubensis* isolates originating in Colombia, Mexico, Indonesia, Vietnam, South Africa and the Republic of Congo. Six of these primer sets could be used to amplify polymorphic loci from genomic DNA of *C. cubensis*, while four were unsuccessful in targeting size polymorphism. A relatively high success rate was achieved, with ≈7% of all recombinant colonies containing inserts harbouring microsatellite-like sequences. Ultimately, only 3% of the colonies yielded markers that can be consistently amplified by PCR across *C. cubensis* populations. This demonstrates that the method followed in this study sufficiently enriches for repetitive sequences to allow quick and easy cloning of markers. Our results are therefore in accordance with results obtained for *Eucalyptus*, using a similar approach (Van der Nest *et al*. 2000).

Two of the loci (*SA7* and *SA9*), in *C. cubensis* showed significant similarity to a single sequence of *N. crassa*, when a BLAST search was performed on the fungal database. This, together with the fact that both of these sequences also show similarity to *S. cerevisiae* chromosome I, indicates that *SA7* and *SA9* are probably in close proximity to each other in the *C. cubensis* genome. The lack of a genetic map for the *C. cubensis* genome, however, makes it impossible to test this hypothesis.

The markers developed in this study will provide a robust tool for future population studies of *C. cubensis*. These will have a positive impact on the *Eucalyptus* breeding programmes currently being developed, by reducing the impact of Cryphonectria

© University of Pretoria

canker. These markers will also be useful in studies that are currently underway to determine the origin and international movement of *C. cubensis*.

Acknowledgements

This research was supported by financial assistance from the National Research Foundation, the Department of Trade and Industry and the Andrew Mellon Foundation.

© University of Pretoria

Table 1 Primer sequences, their origins and amplicon homologies to sequences deposited in GenBank.

Source isolate ①	**ISSR primer**	**Locus name**	**Fungal BLAST, most homologous result** ②	***S. cerevisiae* BLAST most homologous chromosome (*E*-value)**		**Primer sequence (5'→3')** ③	**GenBank accession no.**	**Amplicon size (bp)** ④
CMW8856	$(CGA)_5$	COL6	AF246264 (1·4)	XIV (5·6)	F	GGCCAGGGCAGAGGTAAGGCAG	AY280941	242
					R	GCTAGAGAGTCAACATGATGTG		
CMW8890	$(GT)_8$	COL7	AJ009934 (1·3)	XVI (1·3)	F	GAACCCCGACTACGTGATTATC	AY280942	175
					R	TGGCACTATATCACCATCACTG		
CMW8890	$(GT)_8$	COL11	AL670003 (0·3)	XVI (4·6)	F	CTCATGGGTCCCTGCATGCGAC	AY280943	262
					R	GTGGCACTACCAGAACATACAG		
CMW6112	$(CAC)_5$	SA1	AF004553 (0·054)	VIII (0·83)	F	GGAATCACCACCACTAGCGTCC	AY280944	320
					R	GTGTCTCCGTTAACGCAGTGGT		
CMW6112	$(CAC)_5$	SA3	AL356324 (0·054)	VIII (0·82)	F	TCACCACCACTGGCGTCCAGAC	AY280945	207
					R	TCGTTATCTTGGTGACTGTAGA		
CMW6112	$(CGA)_5$	SA4	AF281307 (0·012)	X (2·9)	F	CAGAGCATGAGATGAATAGATG	AY280946	163
					R	AGTCAGGCTCTTCACGCTCTGT		
CMW6112	$(GT)_8$	SA6	AF107791 (0·37)	V (0·36)	F	ATCGACGATCAGGTTCTGGATC	AY280947	208
					R	TATTGCGGTAACCCAATTTTCG		
CMW6112	$(CAC)_5$	SA7	AL669986 (0·007)	I (0·11)	F	CTGAGGATGACCTTAAGGATTG	AY280948	232
					R	CCATGCACGGACTGATGCTCAC		
CMW6112	$(CAC)_5$	SA9	AL669986 (0·005)	I (0·073)	F	GCTCGGGCTGCCAATCCTTAAG	AY280949	194
					R	CGCCGAGTTTCTCGCCACCATC		

© University of Pretoria

CMW6112	$(CAC)_5$	SA10	AJ295347 (0·03)	V (0·03)	F	GCCGAGCCATCGCTTTACGAAG	AY280950	184
					R	CCGCCGATGTGCTTCTTGGACG		

① All fungal isolates are maintained in the culture collection of the Tree Pathology Co-operative Programme (TPCP), Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa.

② All results are given as GenBank accession numbers followed by the *E*-value of the GenBank match.

③ 'F' and 'R' denote forward and reverse primers, respectively.

④ Amplicon sizes were experimentally verified using PCR with genomic DNA from the source isolates, and electrophoresed on a polyacrylamide gel.

© University of Pretoria

Table 2 Allele frequencies for polymorphic markers developed during this study, and tested on 56 isolates of *Cryphonectria cubensis*. Allele names are based on the size of each allele, as determined from GeneScan® analysis.

Markers	Alleles	Occurrence	Frequencies
SA1	312	1	0·018
	319	42	0·750
	320	9	0·161
	null	4	0·071
SA4	160	1	0·018
	164	17	0·304
	166	38	0·679
SA6, Locus 1	203	8	0·143
	204	4	0·071
	205	10	0·179
	206	10	0·179
	207	3	0·054
	null	21	0·375
SA6, Locus 2	210	33	0·560
	214	1	0·018
	null	22	0·391
SA9, Locus 1	190	27	0·482
	191	23	0·411
	192	4	0·071
	202	1	0·018
	203	1	0·018
SA9, Locus 2	196	22	0·393
	197	33	0·560
	null	1	0·018
SA10	172	1	0·018
	179	9	0·161
	180	14	0·250
	181	1	0·018
	183	16	0·286
	196	6	0·107
	205	1	0·018
	210	8	0·143
COL3	169	3	0·054
	170	1	0·018
	172	5	0·089
	173	41	0·732
	null	6	0·107

© University of Pretoria

Bibliography

Altschul S, Gish W, Miller W, Myers E, Lipman D (1990) Basic local alignment tool. *Journal of Molecular Biology* **5**, 403-410.

Chakrabarti R, Schutt CE (2001) The enhancement of PCR amplification by low molecular weight amides. *Nucleic Acids Research* **29**, 2377-2381.

Cortesi P, Milgroom MG, Bisiach M (1996) Distribution and diversity of vegetative compatibility types in subpopulations of *Cryphonectria parasitica* in Italy. *Mycological Research* **100**, 1087-1093.

Davison EM, Coates DJ (1991) Identification of *Cryphonectria cubensis* and *Endothia gyrosa* from eucalypts in Western Australia using isozyme analysis. *Australasian Plant Pathology* **20**, 157-160.

Hantula M, Dusabenyagasani M, Hamelin RC (1996) Random amplified microsatellites (RAMS) - a novel method for characterizing genetic variation within fungi. *European Journal of Forest Pathology* **26**, 159-166.

Hodges CS, Geary TF, Cordell CE (1979) The occurrence of *Diaporthe cubensis* on *Eucalyptus* in Florida, Hawaii, and Puerto Rico. *Plant Disease Reporter* **63**, 216-220.

Leslie JF (1993) Fungal vegetative compatibility. *Annual Review of Phytopathology* **31**, 127-150.

Milgroom MG, Lipari SE, Ennos RA, Liu Y (1993) Estimation of the outcrossing rate in the chestnut blight fungus, *Cryphonectria parasitica*. *Heredity* **70**, 385-392.

Murray MG, Thompson WF (1980) Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Research* **8**, 4321-4325.

Newhouse JR, MacDonald WL (1991) The ultrastructure of hyphal anastomosis between vegetatively compatible and incompatible virulent and hypovirulent strains of *Cryphonectria parasitica*. *Canadian Journal of Botany* **69**, 602-614.

Sambrook J, Fritch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY, USA.

© University of Pretoria

Van der Nest MA, Steenkamp ET, Wingfield BD, Wingfield MJ (2000) Development of simple sequence repeats (SSR) markers in *Eucalyptus* from amplified inter-simple sequence repeats (ISSR). *Plant Breeding* **119**, 433-436.

Van Heerden SW, Wingfield MJ, Coutinho T, Van Zyl LM (1997) Population diversity among Venezuelan and Indonesian isolates of *Cryphonectria cubensis*. *South African Journal of Science* **93**, xiv.

Van Zyl LM, Wingfield MJ, Alfenas AC, Crous PW (1998) Population diversity among Brazilian isolates of *Cryphonectria cubensis*. *Forest Ecology and Management* **112**, 41-47.

Van Zyl LM, Wingfield MJ, Kemp GHJ, Alfenas AC, Crous PW (1994) Population diversity in Brazilian isolates of *Cryphonectria cubensis*. *South African Journal of Science* **91**, 7 (Abstract).

Wingfield MJ, Swart WJ, Abear B (1989) First record of Cryphonectria canker of *Eucalyptus* in South Africa. *Phytophylactica* **21**, 311-313.

Wingfield MJ, Van Zyl LM, Van Heerden S, Myburg H, Wingfield BD (1997) Virulence and the genetic composition of the *Cryphonectria cubensis* Bruner population in South Africa. In: *Physiology and Genetics of Tree-Phytophage Interactions* eds. Lieutier F, Mattson WJ, Wagner MR), pp. 163-172. INRA Editions.

© University of Pretoria

CHAPTER 3

Multigene phylogenetic and population differentiation data confirm the existence of a cryptic species within *Chrysoporthe cubensis*[a]

[a] **Van der Merwe NA, Gryzenhout M, Steenkamp ET, Wingfield BD, Wingfield MJ.** 2010. Multigene phylogenetic and population differentiation data confirm the existence of a cryptic species within *Chrysoporthe cubensis*. *Fungal Biology* **114**: 966-979.

© University of Pretoria

Abstract

Chrysoporthe cubensis is one of the most important pathogens of *Eucalyptus*. Based on phylogenetic evidence and geographic origin, isolates of this fungus are known to reside in distinct "South America" and "Southeast Asia" clades. In this study, reproductive isolation amongst these isolates of *C. cubensis* was tested using gene flow statistics for 12 polymorphic loci, and to support these data, phylogenetic affiliations based on gene trees and a multigene phylogeny were used. Gene flow statistics between populations, and relative to the closely related *Chrysoporthe austroafricana*, were low and not significantly different ($P < 0{\cdot}05$). Additionally, phylogenetic analyses of DNA sequence data for four gene regions convincingly distinguished the two sub-clades of *C. cubensis*. Isolates in the Southeast Asian sub-clade are described in the new species, *Chrysoporthe deuterocubensis*. *Chrysoporthe cubensis* and *C. deuterocubensis* represent closely related fungi that are thought to be native to South America and Southeast Asia, respectively. A technique is presented that allows for rapid differentiation between these species and that will aid in quarantine procedures to limit their spread to new environments.

© University of Pretoria

Introduction

Chrysoporthe cubensis causes a serious stem canker disease of *Eucalyptus* (Myrtaceae, Myrtales), commonly known as Chrysoporthe canker (Gryzenhout et al. 2009; Hodges 1980). Until 2004, *C. cubensis* and the closely related *C. austroafricana* were treated as *Cryphonectria cubensis* (Gryzenhout et al. 2004). Their recognition as distinct species in the new genus *Chrysoporthe*, was facilitated by DNA sequence-based phylogenetic analyses. Despite the fact that both of these species are associated with Myrtalean hosts, their geographic distributions do not overlap. *Chrysoporthe cubensis* is considered native to South and Central America and Southeast Asia, due to its association with native woody Melastomataceae (Myburg et al. 1999a; Nakabonge et al. 2006; Roux et al. 2005) such as *Miconia* and *Melastoma* species (Gryzenhout et al. 2009). In contrast, disease surveys on the African continent revealed that *Syzygium* species in the Myrtales (Heath et al. 2006; Nakabonge et al. 2006) are commonly infected by *C. austroafricana*, which suggests an African origin for this fungus (Gryzenhout et al. 2009).

Previous phylogenetic studies based on the rRNA internal transcribed spacer (ITS) regions, β-tubulin and histone *H3* genes have consistently separated *C. cubensis* into two well-supported clades (Gryzenhout et al. 2004; Gryzenhout et al. 2006a; Gryzenhout et al. 2006c; Myburg et al. 2004; Myburg et al. 2002, 2003; Myburg et al. 1999b). One of these, referred to as the South American clade, accommodates isolates from countries in South and Central America, as well as likely introductions into western African countries such as Cameroon, Congo and the Democratic Republic of the Congo (Gryzenhout et al. 2006b; Myburg et al. 2002, 2003; Roux et al. 2003). The second clade accommodates isolates from Southeast Asian countries such as Indonesia and Thailand, as well as likely introductions into Australia, China, Hawaii (Gryzenhout et al. 2006b; Myburg et al. 2002, 2003), and several countries in Eastern Africa (Gryzenhout et al. 2006b; Myburg et al. 2003; Nakabonge et al. 2006). Although isolates in these clades have distinct and non-overlapping geographical distributions (Gryzenhout et al. 2004), they all include native hosts in the Melastomataceae. Where they have been found on trees in the Myrtaceae

© University of Pretoria

such as eucalypts and clove (*Syzygium aromaticum*), these are considered to be host shifts (Slippers et al. 2005) either arising from planting these trees in areas where the fungus occurs on related native Myrtales, or through accidental introductions associated with agriculture and forestry (Gryzenhout et al. 2009; Wingfield 2003). There are also no obvious morphological characters that have been shown to distinguish isolates representing the two phylogenetic clades of *C. cubensis* (Gryzenhout et al. 2004).

So-called cryptic species, or species that are distinct but indistinguishable based on morphology, began to emerge when phylogenetic inference arose as an effective means to characterise fungal taxa (Taylor et al. 1999). Well-known examples of taxa harbouring cryptic species include *Coccidioides immitis* (Burt et al. 1996; Koufopanou et al. 2001), *Aspergillus flavus* (Geiser et al. 1998), *Aspergillus fumigatus* (Pringle et al. 2005), *Fusarium subglutinans* (Steenkamp et al. 2002), *Amanita muscaria* (Geml et al. 2006), *Neofusicoccum parvum* and *N. ribis* (Pavlic et al. 2008, 2009). These species are mainly separated based on DNA sequence comparisons, and in some cases diagnostic morphological characters have later been found to support their separation (Geiser et al. 2000; O'Donnell et al. 2004; Pavlic et al. 2008; Taylor et al. 2000).

In addition to making use of phylogenetic species recognition (De Queiros 2007; Hudson & Coyne 2002), specifically the genealogical concordance version of this approach (Taylor et al. 2000), cryptic species can be separated based on low levels of inter-specific gene flow (Sites & Marshall 2003; Taylor *et al*. 2000). This is because continuous admixture between disparately distributed populations can be detected from discordant genealogies for multiple genetic loci and/or low levels of population differentiation and high numbers of migrants (Fisher et al. 2002; Geiser et al. 1998; Milgroom et al. 2008; Zhou et al. 2007). Conversely, concordance among genealogies for multiple loci and diminished gene flow due to ecological, geographical or historical processes are generally regarded as useful indicators of species divergence (Avise & Wollenberg 1997; Barraclough & Nee 2001).

© University of Pretoria

Chrysoporthe cubensis is an economically important fungal pathogen of substantial quarantine importance. Chrysoporthe canker has had a substantial impact on one of the most important sources of paper pulp in the world, and has distinctly shaped *Eucalyptus* forestry globally (Wingfield 2003). Regulations to control its movement are frustrated by a vague taxonomic definition and the fact that very obvious phylogenetic differences amongst isolates are overlooked due to isolates residing under a single name. The aim of this study was, therefore, to gain a refined understanding of isolates residing in the two phylogenetic clades of *C. cubensis*. This was achieved using gene genealogy and population genetic approaches that are based on DNA sequence information for five nuclear genes and polymorphic marker data.

Materials & Methods

Isolates and DNA extraction

Eight isolates of *C. cubensis* representing the two phylogenetic clades, as well as representatives for the other known species of *Chrysoporthe* (Table 1), were used to construct gene genealogies and a multigene phylogeny. Isolates used for population genetic comparisons included 112 *C. cubensis* isolates obtained from *Eucalyptus* trees and were chosen to represent a wide geographic distribution encompassing the largest possible level of diversity. Of these, a total of 79 isolates potentially represented the South American clade and were obtained from Cuba (10), Colombia (34), Mexico (32), and the Democratic Republic of Congo (three). Populations from Southeast Asia (33 isolates) were represented by 16 isolates from Indonesia and 17 from Vietnam. For comparative purposes, the isolates used for the population genetics analyses included a population of 97 *C. austroafricana* isolates from South and Eastern Africa and included those from *Eucalyptus* spp. in Zambia (5), Mozambique (10) and South Africa (29), *Syzygium* spp. in Mozambique (12) and South Africa (26), and *Tibouchina* spp. in South Africa (15). All isolates are maintained in the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa.

© University of Pretoria

Isolates were grown on 20% w/v malt extract agar or inoculated into 800 µl malt extract broth in 1·5 ml micro-centrifuge tubes. After 1 wk of growth in the dark at 25 °C, fungal mycelium was harvested. Total genomic DNA was extracted using a previously published method based on hexadecyltrimethylammonium bromide (CTAB) and standard phenol-chloroform extractions (Steenkamp et al. 1999).

Phylogenetic analyses

Polymerase chain reactions (PCR) were used to amplify rRNA internal transcribed spacer (ITS) and the intron or non-coding regions of the Actin (ACT), β-tubulin (Bt1 and Bt2 primer sets), and eukaryotic translation elongation factor 1-α (EF-1α) genes (Carbone & Kohn 1999; Glass & Donaldson 1995; White *et al*. 1990) for phylogenetic comparisons using a subset of isolates. Each PCR reaction contained 0·1 U SuperTherm *Taq* DNA polymerase enzyme (Southern Cross Biotechnology, South Africa), 25 mM $MgCl_2$, 2 µM of each primer, 200 µM of each dNTP, 25 ng genomic DNA and 1·5 µl 10 × PCR buffer. Reaction volumes were adjusted to 15 µl using sterile deionised water. Reactions were performed using a GeneAmp® PCR System 9700 thermocycler (Applied Biosystems, USA) with the cycling protocol described by Glass and Donaldson (1995). Reaction annealing temperatures were 55 °C for the ITS and EF-1α loci, and 62 °C for the ACT locus and two regions of the β-tubulin gene (BT1 and BT2). Amplicon sizes were visually confirmed using agarose gel (2% w/v) electrophoresis, after which the PCR products were purified using polyethylene glycol precipitation (Steenkamp et al. 2006). PCR products were sequenced using BigDye® dye terminator chemistry (Applied Biosystems, USA) and an ABI™ Prism® 3500 automated sequencing machine (Applied Biosystems).

DNA sequences for each locus were aligned using Muscle 3·6 (Edgar 2004) and manually adjusted using SeaView 2·2 (Galtier et al. 1996). The alignments were amended with ACT, β-tubulin and ITS sequences from the NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov) nucleotide database for a closely related taxon, *Amphilogia gyrosa*, to serve as outgroup (Ju et al. 2007). The incongruence length difference (ILD) test (Cunningham 1997; Farris et al. 1995)

© University of Pretoria

implemented in PAUP* v. 4·0b10 (Swofford 2002), was used to test whether the resulting alignments represent homogenous partitions. In order to test for phylogenetic signal, the g_1 statistic (Hillis & Huelsenbeck 1992) for each data set was calculated using parsimony methods in PAUP* v. 4·0b10. Individual and combined gene alignments were subjected to maximum likelihood (ML) analyses using PhyML 2·4·5 (Guindon & Gascuel 2003), incorporating the GTR+G+I model of evolution as determined by jModelTest 0·1·1 (Posada 2008). The confidence in branches was tested using 1000 bootstrap replicates for each analysis. Phylogenic trees were displayed and annotated using MEGA 4 (Tamura et al. 2007).

Population genetic analyses

In order to determine whether populations of isolates representing *C. austroafricana* and the two clades of *C. cubensis* were significantly different from each other, population differentiation and gene flow between these species were estimated. For this purpose 12 microsatellite markers (Table) were used following previously published methods (Van der Merwe et al. 2003).

The computer programme MultiLocus 1·3b (Agapow & Burt 2001) was used for all allelic analyses. The population differentiation (θ) (Weir & Cockerham 1984) and theoretical number of migrants per generation [$\hat{M} = \frac{1}{2}\left(\frac{1}{\theta} - 1\right)$] (Slatkin 1995) was estimated between a population of *C. cubensis* isolates from South America, from which the type of *C. cubensis* originates (Bruner 1917), and a population of the known species *C. austroafricana*. This was used as the expected statistic in further analyses, since the species represented by these populations are well defined. Pair wise θ-values were then calculated in all combinations between isolates representing the two clades of *C. cubensis* and *C. austroafricana*.

The significance of equality or positive deviation from the expected differentiation value was determined using a one-tailed *G*-test (Sokal & Rohlf 1994), which is the maximum likelihood statistical significance of deviation,

© University of Pretoria

$$G = 2\sum_i O_i \cdot \ln\left(\frac{O_i}{E_i}\right)$$

where G is the maximum likelihood estimator, O is the observed value, and E is the expected value. The distribution of G is approximately that of χ^2, with the same degrees of freedom. Thus, if isolates representing the Southeast Asian clade of *C. cubensis* represent a distinct species, they should display population differentiation values that are equal to or higher than those obtained between the populations representing *C. cubensis* from South America and *C. austroafricana*. However, if *G*-test values were significantly lower than expected, the null hypothesis of equal or higher levels of differentiation could be rejected.

Population differentiation was further explored using the programme Structure 2·3·1 (Pritchard et al. 2000). Allelic data were subjected to an assignment test with the origin of an isolate dictating the prior to the Markov Chain analysis. Therefore, three populations, *i.e. C. austroafricana* and the two clades of *C. cubensis*, were used as priors. The run length was 1,100,000 with a burn-in of 100,000 iterations. After analysis, the assignments were visualized using a bar plot constrained by the population priors, and a triangle plot to visualize possible admixture between populations. These plots depicted the estimated membership coefficients for each individual to each population prior (*q*) in two different ways (Pritchard et al. 2000). In the bar plot, each individual was represented by a vertical bar partitioned into *K* population priors that indicated the estimated membership of that individual to each prior. The triangle plot depicted *Q*, the probability of an individual's ancestry from population prior *q*, where each individual was indicated with a dot and the distance of the dot from each of the triangle's edges was proportionate to the ancestry vectors for the individual. Therefore, each of the individuals in this analysis would have had *K=3* ancestry vectors adding up to 1.

Morphology

In order to characterize and compare the morphology of representative specimens of the two clades of *C. cubensis*, dried herbarium specimens of *C. cubensis* s. l. bearing fruiting

© University of Pretoria

structures (Gryzenhout et al. 2004) were re-examined microscopically. Some of these specimens are linked to isolates in the two subclades (Gryzenhout et al. 2004). Fruiting structures were mounted in Leica mountant (Setpoint Premier, South Africa) and sectioned at 12–16 µm using a Leica CM1100 cryostat (Setpoint Premier). Sections were mounted in lactophenol and examined using light microscopy and the measurement software Axiovision 4·8 (Carl Zeiss, GmbH). Fifty asci, ascospores, conidiophores and conidia were measured for each specimen, and a range was obtained for ascostromata and conidiomata. Measurements were represented as (minimum–) (mean – SD) (mean + SD) (–maximum) where SD is the standard deviation.

Results

Phylogenetic analyses

After sequencing and alignment of four gene regions from each isolate (Table 1), alignment lengths ranged from 273 bp for Actin to 830 bp for the two β-tubulin regions (Table 2). The total alignment length when gene regions were combined was 1914 characters. These alignments are available in TreeBase (SN4622).

Sequence alignments revealed 13 fixed nucleotide polymorphisms across all gene regions that differentiate isolates representing the Southeast Asian clade of *C. cubensis* from the South American clade and *C. austroafricana* (Table 4). Three nucleotide polymorphisms in the β-tubulin gene differentiated Southeast Asian *C. cubensis* from the others, while three polymorphisms across the four genes were diagnostic for South American *C. cubensis*. Similarly, six polymorphisms across the four genes were characteristic of *C. austroafricana*. Therefore, these fixed nucleotide differences are diagnostic for the different species, either in combination or singly in the case of private polymorphisms.

A partition homogeneity test revealed that all the DNA regions used in this study could be combined (P=0·001) (Cummings et al. 1995). Inspection of the g_1 statistic for each of the four gene regions, as well as the combined data set, revealed that all data sets contained useful phylogenetic signal (P=0·01) (Table 2). ML analysis of individual gene

© University of Pretoria

regions mostly recovered the two clades of *C. cubensis* as separate (Figure 1). Southeast Asian *C. cubensis* isolates formed a separate clade in the β-tubulin and ITS gene genealogies (Figure 1B and E). However, in the EF-1α genealogy, the Southeast Asian and South American *C. cubensis* were difficult to distinguish because they were present in the same clade with no bootstrap-supported partitions (Figure 1C). The Actin genealogy (Figure 1A) did not distinguish between Southeast Asian *C. cubensis* and *C. inopina*. ML analysis of the combined information for the four regions sequenced (Figure 1F) recovered two well-supported and separate clades for the Southeast Asian and South American *C. cubensis* isolates. Using these analyses, isolates representing the South American clade of *C. cubensis* was more closely related to *C. austroafricana* than to isolates in the Southeast Asian clade.

Population differentiation analysis

Differentiation between *C. austroafricana* and South American *C. cubensis* populations, and thus, the expected level of differentiation between two distinct species (θ), was 0·30 ($\hat{M} = 1{\cdot}18$). Differentiation values (θ) were 0·27 ($\hat{M} = 1{\cdot}35$) between South American and Southeast Asian populations of *C. cubensis,* and 0·29 ($\hat{M} = 1{\cdot}25$) between Southeast Asian *C. cubensis* and *C. austroafricana* populations. Analyses using a *G*-test showed that there were no significant differences ($P < 0{\cdot}05$) in the levels of differentiation among the three populations. It was thus possible to reject the null hypothesis that these populations are not significantly different, because the theoretical number of migrants per generation ($\hat{M}$; calculated from the θ value) between the different populations were comparable and similar levels of differentiation were observed among them. Similarly, the results of population assignment tests suggested that populations of *C. austroafricana* and South American and Southeast Asian *C. cubensis* can be readily separated (Figure 3). These data highlighted the fact that the three populations were each characterized by markedly different allelic compositions, although a low level of admixture was detected (Figure 3A). However, the genetic distance between the two populations of *C. cubensis* was comparable to those between the *C. austroafricana* population and the respective *C. cubensis* populations (Figure 3B). This was evident by the reciprocal presence of alleles and nearly identical ancestry vectors for all three populations.

© University of Pretoria

Taxonomy

Phylogenetic and population genetic analyses in this study have provided robust justification to treat the Southeast Asian and South American isolates of *C. cubensis* as distinct taxa. *Chrysoporthe cubensis* was first described from Cuba (Bruner 1917) and this name should be reserved for South American isolates related to those from Cuba. Gryzenhout et al. (2006a) designated an epitype for *C. cubensis* based on an isolate from Cuba and residing in the South American clade of the fungus. Isolates representing the Southeast Asian clade represent a distinct taxon described as follows:

Chrysoporthe deuterocubensis Gryzenh. & M.J. Wingf. sp. nov., MB 516634 Figure 4.

Etymology: The name reflects the fact that the fungus is different yet closely related to *Chrysoporthe cubensis*.

Ascosporae (5.5–)6.5–7.5(–8) × 2–2.5(–3) μm; conidiomata subaurantiaca, brunnea Siennae vel atrofusca, pyriformia, clavata vel pulvinata; conidia (3–)3.5–4.5(–5) × (1.5)2(–2.5) μm; position actinis 475 (G, A); positiones β-tubulinis 546 (C, T), 699 (T, C) 729 (T, C), 1477 (C, T), 1488 (G, A), 1572 (C, T) (TreeBase SN4622); sitibus exceptionis pro *Ava*I, fragmenta 87 bp, 113 bp, et 337 bp ferentibus, et uno pro *Hin*dIII fragmenta 206 bp et 331 bp ferenti.

Ascospores (5.5–)6.5–7.5(–8) × 2–2.5(–3) μm; conidiomata sienna to almost orange to fuscous-black, pyriform to clavate to pulvinate; conidia (3–)3.5–4.5(–5) × (1.5–)2(–2.5) μm; Actin position 475 (G, A); β-tubulin positions 546 (C, T), 699 (T, C) 729 (T, C), 1477 (C, T), 1488 (G, A), 1572 (C, T) (TreeBase SN4622). Two restriction sites for *Ava*I, yielding fragments of 87 bp, 113 bp, and 337 bp, and one restriction site for *Hin*dIII, yielding fragments of 206 bp and 331 bp.

Ascostromata semi-immersed erumpent, top of perithecial bases covered with cinnamon to orange, predominantly prosenchymatous, limited stromatic tissue, ascostroma 120–230 μm high above level of bark, 280–490 μm diam., perithecia valsoid, bases immersed in

© University of Pretoria

bark, fuscous black, extending necks up to 240 µm long emerging through bark covered in umber stromatic tissue of *textura porrecta*, appearing fuscous-black. **Asci** (19–)22–26.5(–28) × (4.5–)5–6.5(–7) µm, fusoid to ellipsoidal, 8-spored. **Ascospores** (5.5–)6.5–7.5(–8) × 2–2.5(–3) µm, hyaline, 1-septate, fusoid to oval, ends tapered, with septum variously placed in the spore but usually central.

Conidiomata occurring on the surface of the ascostroma or as separate structures, superficial to slightly immersed, sienna to almost orange to fuscous-black, with an umber interior when young, pyriform to clavate, sometimes pulvinate, with one to four attenuated necks per structure, conidiomatal base above the bark surface 130–740 µm high, 100–950 µm diam, necks up to 230 µm long, 90–240 µm wide. Conidiomatal locules with even to convoluted inner surface, occasionally multilocular, single locule connected to one or several necks. Stromatic tissue at base of *textura globulosa* with walls of outer cells thickened, neck cells of *textura porrecta*. **Conidiophores** hyaline, with globose to rectangular basal cells that are (2.5–)4–7(–8.5) × (2–)3–4.5(–5.5) µm, branched irregularly at the base or above into cylindrical cells, cells delimited by septa or not, total length of conidiophore (12–)13.5–19(–24.5) µm, conidiogenous cells cylindrical to flask-shaped with attenuated apices, (1.5–)2–2.5(–3) µm wide. **Conidia** (3–)3.5–4.5(–5) × (1.5–)2(–2.5) µm, hyaline, oblong, aseptate, exuded as bright luteous tendrils or droplets.

Cultures white with cinnamon to hazel patches on malt extract agar, fluffy, margin smooth, fast-growing, covering a 90 mm diam plate after a minimum of 5 days at the optimum temperature of 30 ºC (Gryzenhout et al. 2004).

The following **nucleotide characters** are differentially fixed for *C. deuterocubensis* (given as the gene name, the nucleotide position relative to the start codon of the corresponding aligned gene for *Neurospora crassa*, and in parentheses, the nucleotides fixed for *C. deuterocubensis* and *C. cubensis*, respectively): Actin position 475 (G, A); β-tubulin positions 546 (C, T), 699 (T, C) 729 (T, C), 1477 (C, T), 1488 (G, A), 1572 (C, T) (TreeBase SN4622). Additionally, the Bt1 region of the β -tubulin gene of *C.*

© University of Pretoria

deuterocubensis, amplified using primers Bt1a and Bt1b (Glass & Donaldson 1995), contains two restriction sites for *Ava*I, yielding fragments of 87 bp, 113 bp, and 337 bp, and one restriction site for *Hin*dIII, yielding fragments of 206 bp and 331 bp.

Specimens examined: INDONESIA**,** Sumatra, Lake Toba, *Melastoma malabathricum*, May 2005, M.J. Wingfield, **holotype** PREM 58799, living ex-type culture CMW 18515 = CBS 118651 shown to group in the Southeast Asian clade (Gryzenhout et al. 2006a), Lake Toba, Aek Nauli, *Melastoma malabathricum*, Feb. 2004, M.J. Wingfield, PREM 58798, living culture CMW 16192 = CBS 119474; Sulawesi, *Syzygium aromaticum*, 2001, M.J. Wingfield, PREM 57470, cultures CMW 8650 = CBS 115719, CMW 8651 = CBS 115718; Sulawesi, *Syzygium aromaticum*, 2003, M.J. Wingfield, PREM 58018, PREM 58019; Sulawesi, Utard, *Syzygium aromaticum*, 2003, M.J. Wingfield, PREM 58020; Bankals, Selindung, *Eugenia* sp., C.P.A. Bennett, IMI 231648; Sumatra, Kurai, Taji, *Eugenia* sp., C.P.A. Bennett, IMI 231649; Sumatra, *Eucalyptus* sp., 2001, M.J. Wingfield, PREM 57297, cultures from the same area CMW 11288 = CBS 115736, CMW 11289 = CBS 115737, CMW 11290 = CBS 115738; Sumatra, Sei Kabaro, *Eucalyptus* sp., 2001, M.J. Wingfield, PREM 58021, cultures from same area CMW 11289, CMW 11290. MALAYSIA**,** Johar, Kluang, *Eucalyptus aromatica*, 1986, Loh Chow Fong, IMI 304273; Serdang, Fe. Exp. Stn., *Syzygium aromaticum* (as *Eugenia caryophyllata*), 1954, A. Johnston, IMI 58569; *Eugenia* sp., 1954, A. Johnston, IMI 58567, IMI 58568; Jelok Bahang, *Syzygium aromaticum* (as *Eugenia caryophyllata*), 1954, A. Johnston, IMI 58388. SINGAPORE**,** Istana grounds, *Syzygium aromaticum*, 1991, C.P. Yik, dried culture IMI 350626; *Tibouchina urvilleana*, Apr. 2003, M.J. Wingfield, PREM 58797, living culture CMW 12745 = CBS 117837.

Distribution: Countries where the identity has been confirmed based on DNA sequence comparisons: U.S.A. (Hawaii), Tanzania (Zanzibar), Kenya, Malawi, Mozambique, Indonesia, Singapore, Thailand, China (Hong Kong - ITS only), Australia (Gryzenhout et al. 2004; Gryzenhout et al. 2006b; Myburg et al. 1999a; Myburg et al. 2003; Nakabonge et al. 2006; Roux et al. 2003). Isolates from Vietnam, although previously reported (Old et al. 2003), are also for the first time confirmed to belong to *C. deuterocubensis* based on

© University of Pretoria

DNA sequence data. *Chrysoporthe cubensis sensu lato* reported from India, Malaysia and Western Samoa (Hodges et al. 1986; Hodges et al. 1979; Old et al. 2003; Sharma et al. 1985) most likely also reside in *C. deuterocubensis* although sequence data for these isolates are not available.

Restriction enzyme-based DNA diagnostic

To facilitate routine differentiation among *C. cubensis*, *C. deuterocubensis* and *C. austroafricana*, the β -tubulin Bt1 region was subjected to PCR-RFLP (restriction fragment length polymorphism) analysis. For this purpose we used two restriction enzymes *Ava*I and *Hin*dIII. Separate digests with these enzymes revealed that *C. austroafricana*, *C. cubensis* and *C. deuterocubensis* were easily distinguishable from each other (Figure 2). When *Ava*I was used, three bands (87 bp, 113 bp and 337 bp) were observed for *C. austroafricana* and *C. deuterocubensis*, while two bands (87 bp and 440 bp) were observed for *C. cubensis*. Therefore, this enzyme could distinguish *C. cubensis* from *C. austroafricana* and *C. deuterocubensis*. In contrast, *Hin*dIII did not cut for *C. austroafricana* but produced two fragments (206 bp and 331 bp) for each of the other two species. Therefore, *Hin*dIII could distinguish *C. austroafricana* from *C. cubensis* and *C. deuterocubensis*.

Discussion

Results of this study have shown that isolates representing the South American and Southeast Asian clades of *C. cubensis* represent distinct species. Those residing in the Southeast Asian clade have consequently been provided with the name *C. deuterocubensis*. Recognition of these two taxa as distinct species is supported by phylogenetic analyses of four variable gene regions that separated representative isolates of the two species. Both species are also associated with a number of differentially fixed polymorphisms in the five regions examined. Populations linked to these two species from different geographic regions also showed significant differentiation from each other as their distributions do not overlap.

© University of Pretoria

Based on morphology, *C. cubensis* and *C. deuterocubensis* are virtually indistinguishable (Gryzenhout et al. 2004) and perceived differences are usually variable or due to environmental conditions (Gryzenhout et al. 2009). However, the sienna to sometimes orange colour of especially young conidiomata observed in some *C. deuterocubensis* specimens is not common in *C. cubensis*, although mature conidiomata are usually similar in appearance. An alternative and robust approach to distinguish the species is to use variation in the gene encoding β-tubulin, which can either be evaluated directly through sequencing or using the PCR-RFLP procedure described in this study. The latter approach is rapid and will be useful for quarantine purposes where a simple diagnostic is typically required.

The multigene phylogeny presented in this study showed clear separation of *C. deuterocubensis* from *C. cubensis* (Figure 1). However, analyses of the individual regions suggested that the EF-1α region is not sufficiently variable to allow separation of *C. deuterocubensis* from *C. cubensis* (Figure 1C). This was also true for the Actin sequences that did not allow separation of *C. deuterocubensis* and *C. inopina* (Figure 1A). This is probably due to the relatively recent divergence of species in *Chrysoporthe*, and different rates of mutation in the gene regions analyzed. Furthermore, the ITS region failed to recover a statistically supported monophyletic *C. cubensis*. This may be due to incomplete lineage sorting (Dettman *et al*. 2003; Hare & Avise 1998) that is expected to be present when closely related species are considered (Hudson & Coyne 2002; Rosenberg 2003).

Population genetic analysis of alleles for 12 loci in *C. cubensis*, *C. deuterocubensis* and *C. austroafricana* isolates showed that the levels of differentiation between the populations were not significantly different ($P < 0{\cdot}05$). Although the inferred number of migrants between species is relatively high, they are typical for fungi and may be an artefact of close relatedness or incomplete lineage sorting (Liu et al. 2009; Stukenbrock et al. 2006). These data, therefore, confirmed the observation based on multigene phylogenetic inference that *C. deuterocubensis* represents a distinct species. Additionally, the level of population differentiation reported in this study can be used in future studies

© University of Pretoria

considering species delineations in *Chrysoporthe*, given that it is possible to obtain adequate populations. Separate species in *Chrysoporthe* display differentiation values (θ) of *c*. 0·27 – 0·3, while the corresponding number of migrants is *c*. 1·1 – 1·3. When new species are considered and the population differentiation increases above 0·3, the likelihood of complete lineage sorting increases and subsequently, new species can be described based on population genetic data, particularly where phylogenetic data are inconclusive or confusing.

Population assignment tests showed that *C. deuterocubensis* isolates represent a well defined assemblage. However, *C. deuterocubensis* isolates harboured higher frequencies of some alleles that were assigned to *C. cubensis* or *C. austroafricana* (Figure 3). It is, therefore, possible that *C. deuterocubensis* represents an ancestral species, and that the other species are derived from it. This notion is supported by the fact that *C. deuterocubensis* appears basal to *C. cubensis* and *C. austroafricana* when the joint phylogeny of five gene regions is considered (Figure 1F).

Africa is the only continent besides South America that harbours different species of *Chrysoporthe*. *Chrysoporthe austroafricana* has a wide geographical range and is thought to be native because it occurs on native *Syzygium* spp. and has not been found outside of Africa (Heath et al. 2006; Nakabonge et al. 2006). *Chrysoporthe cubensis* has been found in western African countries such as Cameroon, Congo and the Democratic Republic of Congo, while *C. deuterocubensis* is found in the eastern African countries of Zanzibar (Tanzania), Kenya, Malawi and Mozambique (Nakabonge et al. 2006). *Chrysoporthe cubensis* and *C. deuterocubensis* have not been found on any native African hosts in recent surveys (Nakabonge et al. 2006; Roux et al. 2005; Roux et al. 2003) and *C. deuterocubensis* was also shown to have a low genetic diversity (Nakabonge et al. 2007). These facts suggest that *C. deuterocubensis* was introduced into Africa. The same could be true for *C. cubensis* in Africa, although population level studies would be necessary to show this conclusively. If this should be true, *C. cubensis* and *C. deuterocubensis* do not occur naturally in Africa.

© University of Pretoria

Description of *C. deuterocubensis* now extends the number of known species in the genus to eight, including *C. zambiensis*, *C. syzygiicola* (Chungu et al. 2009), *C. austroafricana*, *C. cubensis*, *C. doradensis*, C. *inopina* and *C. hodgesiana*. Four of these have an apparent Central and South American distribution. Together with *C. cubensis* these include *C. doradensis*, *C. inopina* and *C. hodgesiana*, which are currently known only from Colombia and adjacent Ecuador (Gryzenhout et al. 2004; Gryzenhout et al. 2005; Gryzenhout et al. 2006b). All of these species, except *C. doradensis*, have been found on native trees (Gryzenhout et al. 2004; Gryzenhout et al. 2005; Gryzenhout et al. 2006b). This suggests that these species occur naturally in South America, with this continent currently harbouring the most species of *Chrysoporthe*.

The segregation of *C. cubensis* and *C. deuterocubensis* has important quarantine implications. Where these fungi were previously linked to Chrysoporthe canker on *Eucalyptus* with a single species as the causal agent (Gryzenhout et al. 2004; Gryzenhout et al. 2009), the name now encompasses three species including *C. austroafricana*, *C. cubensis* and *C. deuterocubensis*, with geographical ranges that do not overlap. The pathogenicity of *C. cubensis* has been well established in inoculation trials on *Eucalyptus* (Boerboom & Maas 1970; Hodges *et al*. 1976; Wingfield 2003). *Chrysoporthe deuterocubensis* is associated with symptoms in Southeast Asia, which are very similar to those caused by *C. cubensis* in South America. For example, pathogenicity tests on *Eucalyptus* have been conducted with *C. deuterocubensis* in Indonesia with results very similar to those for *C. cubensis* (Wingfield M. J., unpublished). Pegg *et al*. (2010) have also conducted inoculation trials on *Eucalyptus* with *C. deuterocubensis* although the fungus in that study was treated as *C. cubensis*. Both species also have the ability to infect native woody plants in the Melastomataceae and Myrtaceae and could thus cause serious damage if they were accidentally introduced into new environments with native Myrtaceae or Melastomataceae (Gryzenhout et al. 2009).

© University of Pretoria

Acknowledgements

Financial support from Members of the Tree Protection Co-operative Programme (TPCP), the National Research Foundation and the Department of Science and Technology (DST) and the NRF Centre of Excellence in Tree Health Biotechnology (CTHB) made it possible to conduct this research. Ms. Kerry-Anne Pillay is acknowledged for technical support and particularly for generating some DNA sequences used in this study.

© University of Pretoria

Table 1 Isolates of *Chrysoporthe* spp. used in the multigene phylogenetic studies, and the GenBank accession numbers for the sequences included in phylogenetic analyses. *Amphilogia gyrosa* was used as an outgroup taxon (Ju et al. 2007).

Taxon	Origin	Isolate Number ①	ACT	BT1	BT2	EF1α	ITS
Chrysoporthe austroafricana	South Africa	CMW9327	GQ290173	GQ290185	GQ290194	GQ290151	GQ290158
..	..	CMW10192	GQ290163	GQ290176	GQ290187	GQ290138	AY214299
Chrysoporthe cubensis	Colombia	CMW10028	GQ290161	GQ290175	GQ290186	GQ290137	GQ290153
..	Republic of Congo	CMW10669	GQ290171	GQ290177	GQ290188	GQ290140	GQ290154
..	Brazil	CMW10778	GQ290165	GQ290178	GQ290189	GQ290141	GQ290155
..	Mexico	CMW12734	GQ290159	DQ368791	GQ290191	GQ290146	DQ368769
Chrysoporthe deuterocubensis	Australia	CMW2631	GQ290174	GQ290184	AF543825	GQ290149	GQ290157
..	Indonesia	CMW8650	GQ290172	AY084024	GQ290193	GQ290150	AY084001
..	Singapore	CMW12745	GQ290160	GQ290183	DQ368781	GQ290147	DQ368764
..	Thailand	CMW17178	GQ290164	DQ368785	GQ290192	GQ290148	DQ368766
Chrysoporthe doradensis	Ecuador	CMW11287	GQ290167	GQ290179	GQ290190	GQ290142	GQ290156
Chrysoporthella hodgesiana	Colombia	CMW9995	GQ290162	AY956978	AY956977	GQ290152	AY956969
..	..	CMW10625	GQ290170	AY262391	AY956979	GQ290139	AY262399
Chrysoporthe inopina	..	CMW12727	GQ290169	GQ290180	DQ368806	GQ290143	DQ368777
..	..	CMW12731	GQ290168	GQ290182	DQ368811	GQ290145	DQ368779
Amphilogia gyrosa	Taiwan	BCRC34145	EF025600	EF025615	EF025615	-	EF026147

① CMW - culture collection of the Forestry and Agricultural Biotechnology Institute, University of Pretoria, Pretoria, South Africa; BCRC – Bioresource Collection and Research Center, Taiwan.

© University of Pretoria

Table 2 Statistics resulting from parsimony and maximum likelihood phylogenetic analyses.

Statistic	Actin	β-tubulin	EF1-α	ITS	Combined
Taxa	16	16	15	16	16
Aligned characters	275	830	327	490	1914
Parsimony-informative	6	23	11	6	53
g_1-statistic ①	-1·029	-0·728	-0·948	-0·584	-0·921

① Homoplasious nucleotides are not included in calculation of the g_1-statistic.Values were compared with critical values published by Hillis and Huelsenbeck (1992). All values were significant at P = 0·01, indicating that phylogenetic signal was present in all data sets.

© University of Pretoria

Table 3 Primers for polymorphic DNA markers used in this study.

Locus Name ①	Dye Label	Bin Size (bp)	Primer Name	Primer Sequence (5'-3')
CcPMC	VIC	190-212	PMCF	ttgcgtatggaaatgacg
			PMCR	atggcgcttgtatagagca
CcPMG	6-FAM	197-297	PMGF	tgattcacgtctattgccac
			PMGR	gttaagttctcggtgaatcg
COL6	6-FAM	260-270	COL6F	ggccagggcagaggtaaggcag
			COL6R	gctagagagtcaacatgatgtg
COL7	VIC	173-174	COL7F	gaaccccgactacgtgattatc
			COL7R	tggcactatatcaccatcactg
COL11	VIC	258-267	COL11F	ctcatgggtccctgcatgcgac
			COL11R	gtggcactaccagaacatacag
SA1	NED	300-320	SA1F	ggaatcaccaccactagcgtcc
			SA1R	gtgtctccgttaacgcagtggt
SA3	6-FAM	200-215	SA3F	tcaccaccactggcgtccagac
			SA3R	tcgttatcttggtgactgtaga
SA4	PET	150-200	SA4F	cagagcatgagatgaatagatg
			SA4R	agtcaggctcttcacgctctgt
SA6 ②	PET	209-221, 316-365	SA6F	atcgacgatcaggttctggatc
			SA6R	tattgcggtaacccaattttcg
SA9 ②	NED	190-200, 203-215	SA9F	gctcgggctgccaatccttaag
			SA9R	cgccgagtttctcgccaccatc

① Loci in bold were published in Van der Merwe et al. (2003). *CcPMC* and *CcPMG* were developed according to previously described methods (Van der Merwe et al. 2003).
② Primer pairs for markers *SA6* and *SA9* each amplify two polymorphic loci.

© University of Pretoria

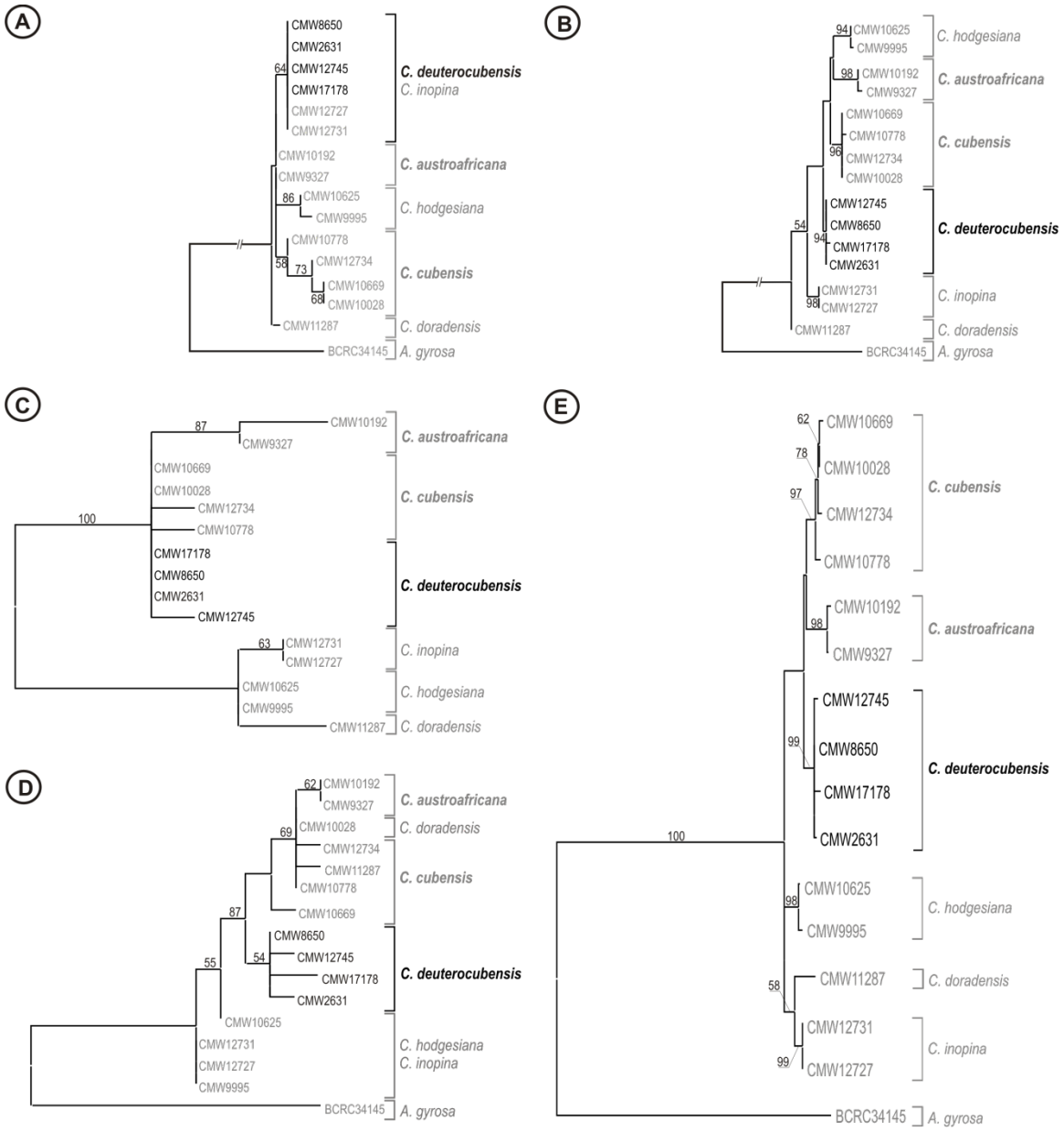
Table 4 Fixed nucleotide polymorphisms between *C. austroafricana*, *C. cubensis* (South American clade) and *C. deuterocubensis* (Southeast Asian clade). Nucleotide positions are relative to the start codon of the corresponding gene from *Neurospora crassa* or (*) alignment positions due to a corresponding gap in the *N. crassa* sequence. Shaded nucleotides differentiate *C. deuterocubensis* (Southeast Asian clade) from *C. austroafricana* and *C. cubensis* (South American clade), while nucleotides in bold were exploited in a PCR-RFLP diagnostic technique.

Species	Isolate	ACT				β-tubulin							EF-1α		ITS		
			*			1	**1**	1	1	**1**	1	1		2	2	* 2	2
		4	5	6	7	4	**4**	4	4	**4**	5	5	4	1	1	2	2
		7	5	9	2	7	**8**	9	9	**9**	5	7	1	0	7	0	1
		5	0	9	9	7	**8**	0	3	**9**	4	2	8	7	8	5	0
C. austroafricana	CMW9327	G	T	C	C	C	**G**	C	C	**C**	C	C	A	T	A	-	A
	CMW10192	G	T	C	C	C	**G**	C	C	**C**	C	C	A	T	A	-	A
C. cubensis	CMW10028	A	T	C	C	T	**A**	A	A	**T**	T	T	C	C	A	-	A
	CMW10669	A	T	C	C	T	**A**	A	A	**T**	T	T	C	C	A	-	A
	CMW10778	A	T	C	C	T	**A**	A	A	**T**	T	T	C	C	A	-	A
	CMW12734	A	T	C	C	T	**A**	A	A	**T**	T	T	C	C	A	-	A
C. deuterocubensis	CMW2631	G	C	T	T	C	**G**	A	A	**T**	T	C	C	C	G	T	G
	CMW8650	G	C	T	T	C	**G**	A	A	**T**	T	C	C	C	G	T	G
	CMW12745	G	C	T	T	C	**G**	A	A	**T**	T	C	C	C	G	T	G
	CMW17178	G	C	T	T	C	**G**	A	A	**T**	T	C	C	C	G	T	G

© University of Pretoria

Figure 1 Maximum likelihood phylogenies obtained from (A) Actin, (B) β-tubulin (BT1 and BT2 regions), (C) Elongation Factor 1-α, and (D) ITS gene sequences. (E) is the multilocus phylogeny when all genes were combined and analyzed with maximum likelihood. (C) was midpoint rooted, while the other phylogenies were rooted with *Amphilogia gyrosa*. Bootstrap values above 50% (1000 replicates) are indicated above each branch.

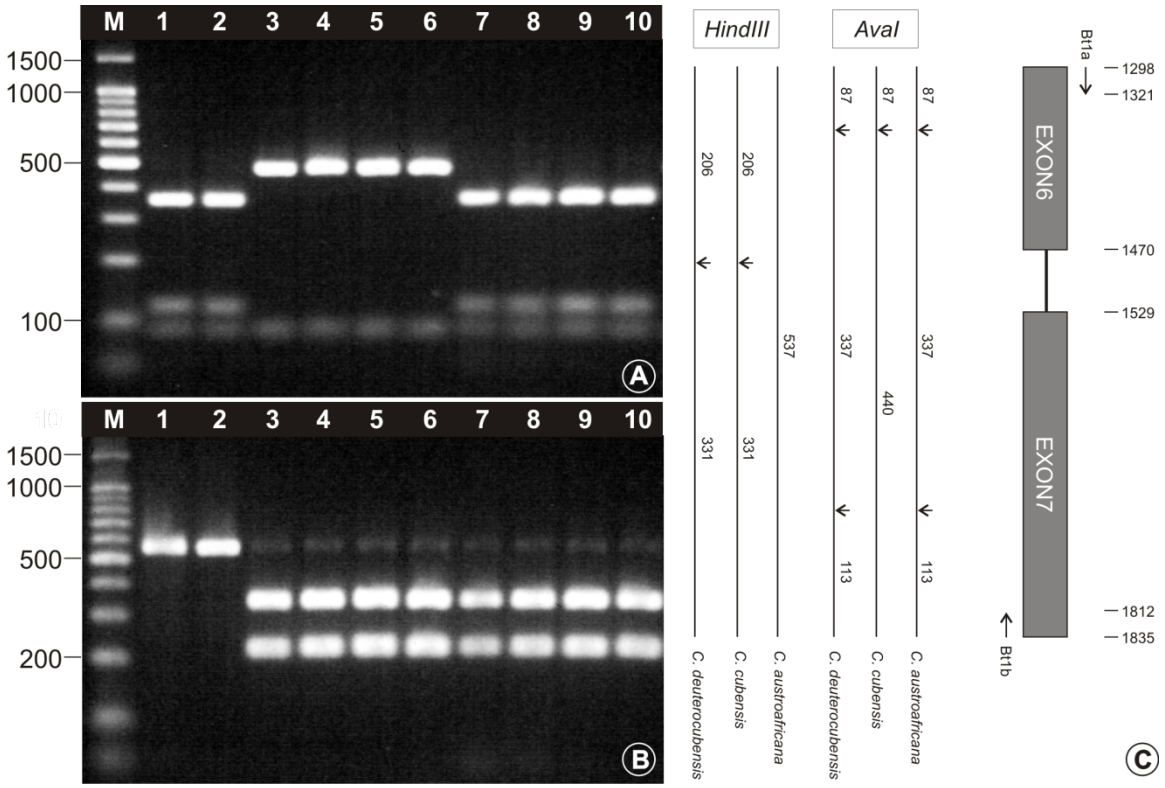
© University of Pretoria



© University of Pretoria

Figure 2 Digests of β-tubulin gene PCR products (primers Bt1a to Bt1b) for *C. austroafricana* (lanes 1 and 2), *C. cubensis* (lanes 3–6) and *C. deuterocubensis* (lanes 7–10). Lane 1, CMW9327; 2, CMW10192; 3, CMW10028; 4, CMW10669; 5, CMW10778; 6, CMW12734; 7, CMW2631; 8, CMW8650; 9, CMW12745; 10, CMW17178. Lane M is a 100bp DNA ladder. (A) Restriction profile using *Alu*I; (B) restriction profile using *Hin*dIII; (C) restriction maps showing restriction site differences between the three fungal species. The gene map displays coding sequence as grey boxes and intron sequence as a solid line. Primer positions are indicated with arrows on opposite strands of the expected amplicon of 537 bp. Position numbers start at 1 for the first coding base of the *Bml* gene for *Neurospora crassa* (Glass & Donaldson 1995).

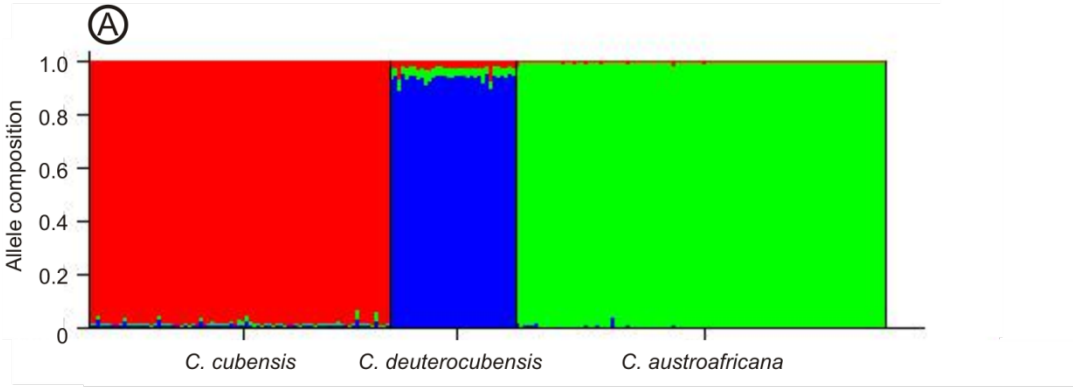
© University of Pretoria

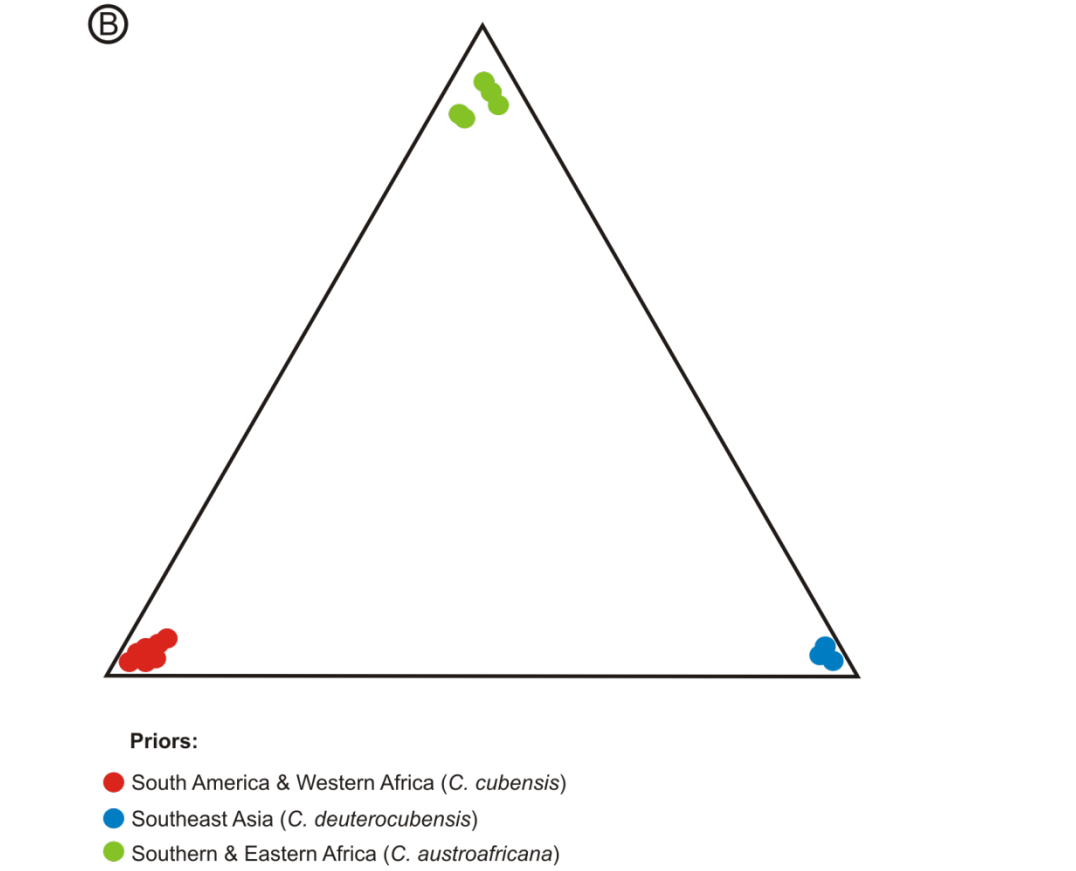


© University of Pretoria

Figure 3 Visualization of population assignment tests for allelic data of *Chrysoporthe austroafricana* (green), *C. cubensis* (red) and *C. deuterocubensis* (blue) genotypes. (A) Bar plot displaying individuals (X-axis) and the percentage of alleles from each of the priors that each individual possesses (Y-axis). (B) Triangle plot of genotypes displaying the absence of genotype admixture between *C. austroafricana* (southern and western Africa), *C. cubensis* (South America) and *C. deuterocubensis* (Southeast Asia).

© University of Pretoria

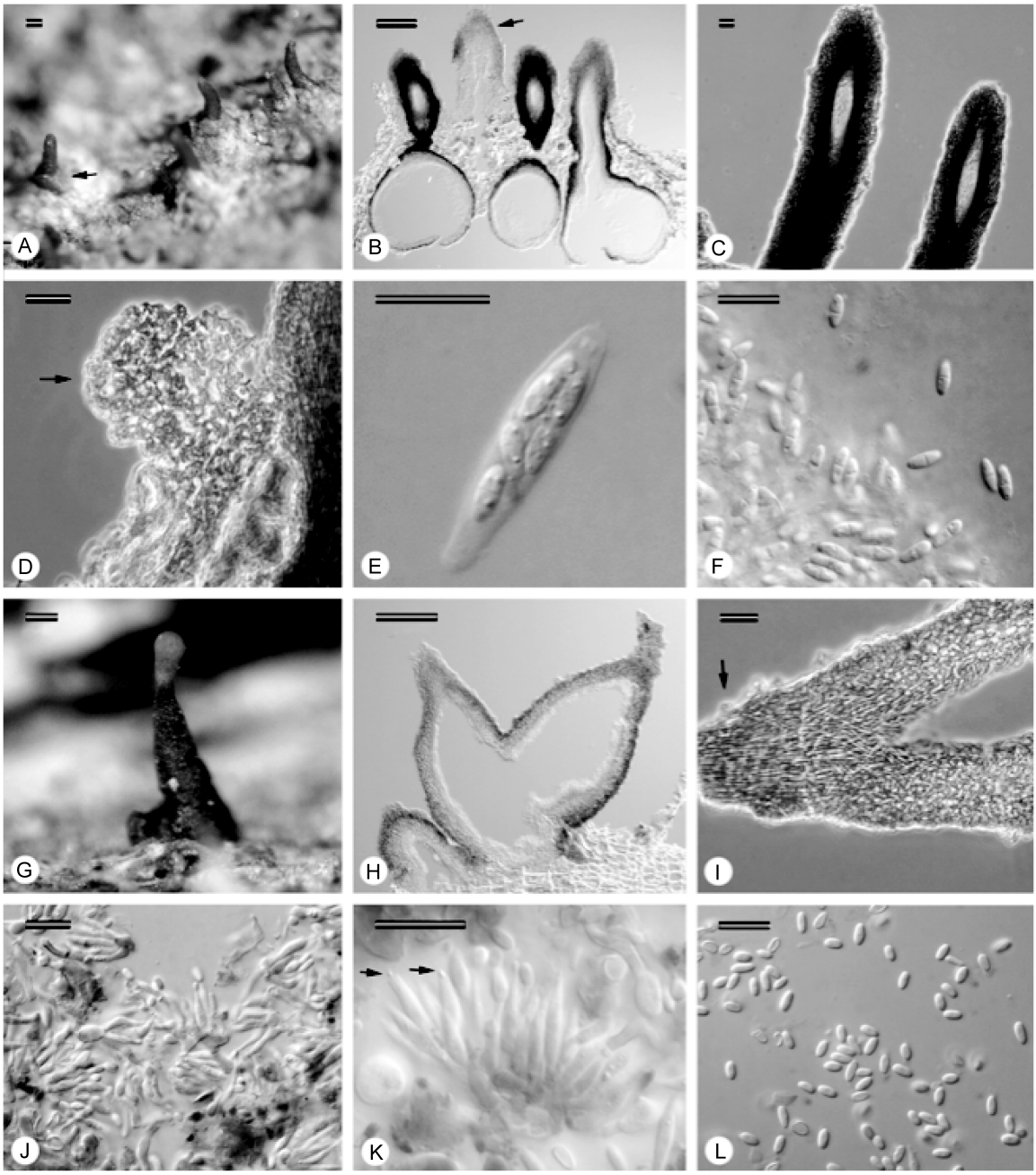




© University of Pretoria

Figure 4 Fruiting structures of *Chrysoporthe deuterocubensis*. (A) Ascostromata on bark (arrow indicates stromatic tissue). (B) Longitudinal section through ascostroma. (C) Black perithecial necks covered with stromatic tissue. (D) Prosenchymatous stromatic tissue (arrow) of ascostroma. (E) Ascus. (F) Ascospores. (G) Conidioma on bark. (H) Vertical section through conidiomata. (I) Tissue of *textura globulosa* for the conidiomal base and of *textura porrecta* for the neck (arrow). (J–K) Conidiophores. (L) Conidia. Scale bars A–B, G–H = 100 μm; C–D, I = 20 μm; E–F, J–L = 10 μm.

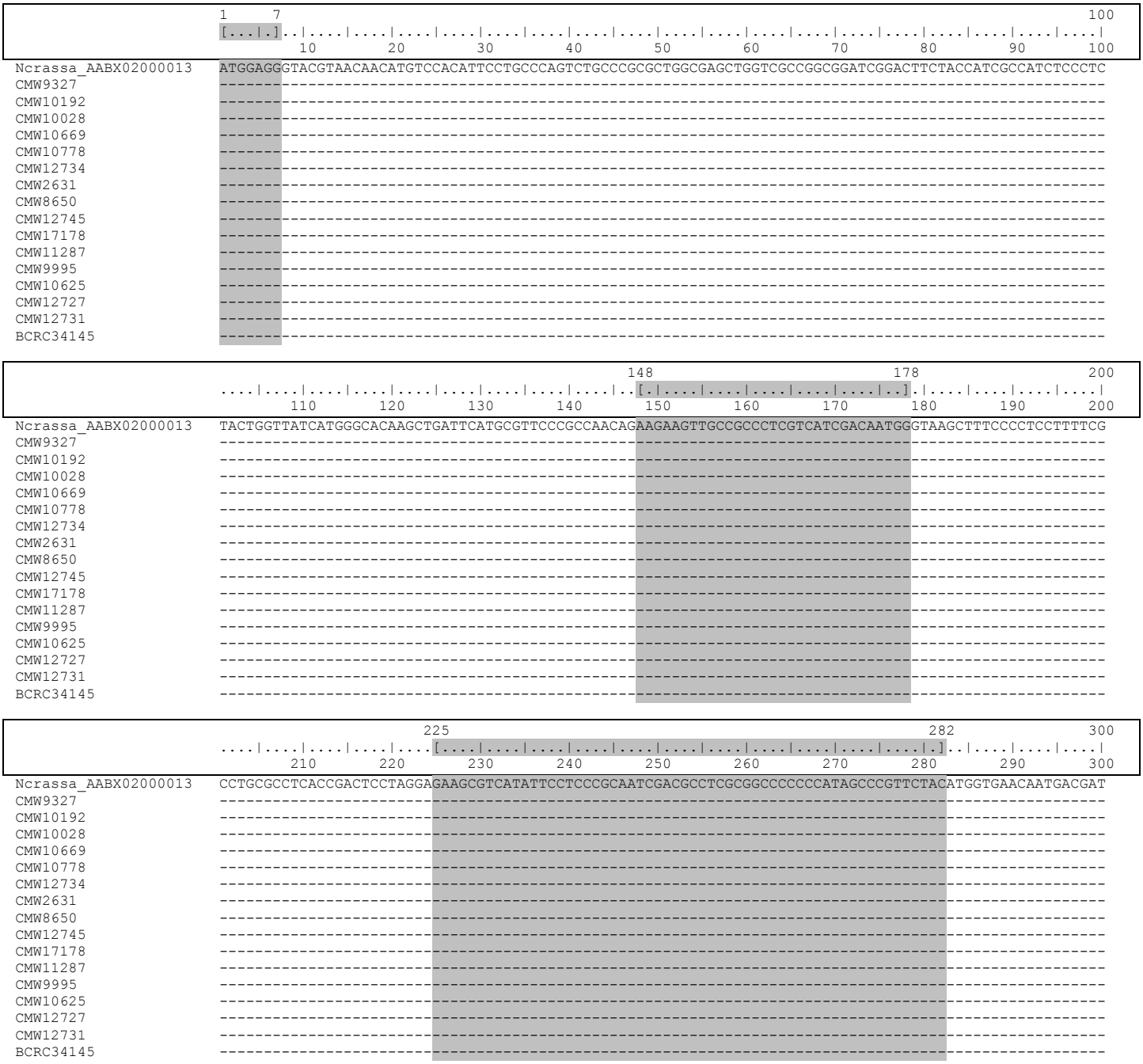
© University of Pretoria



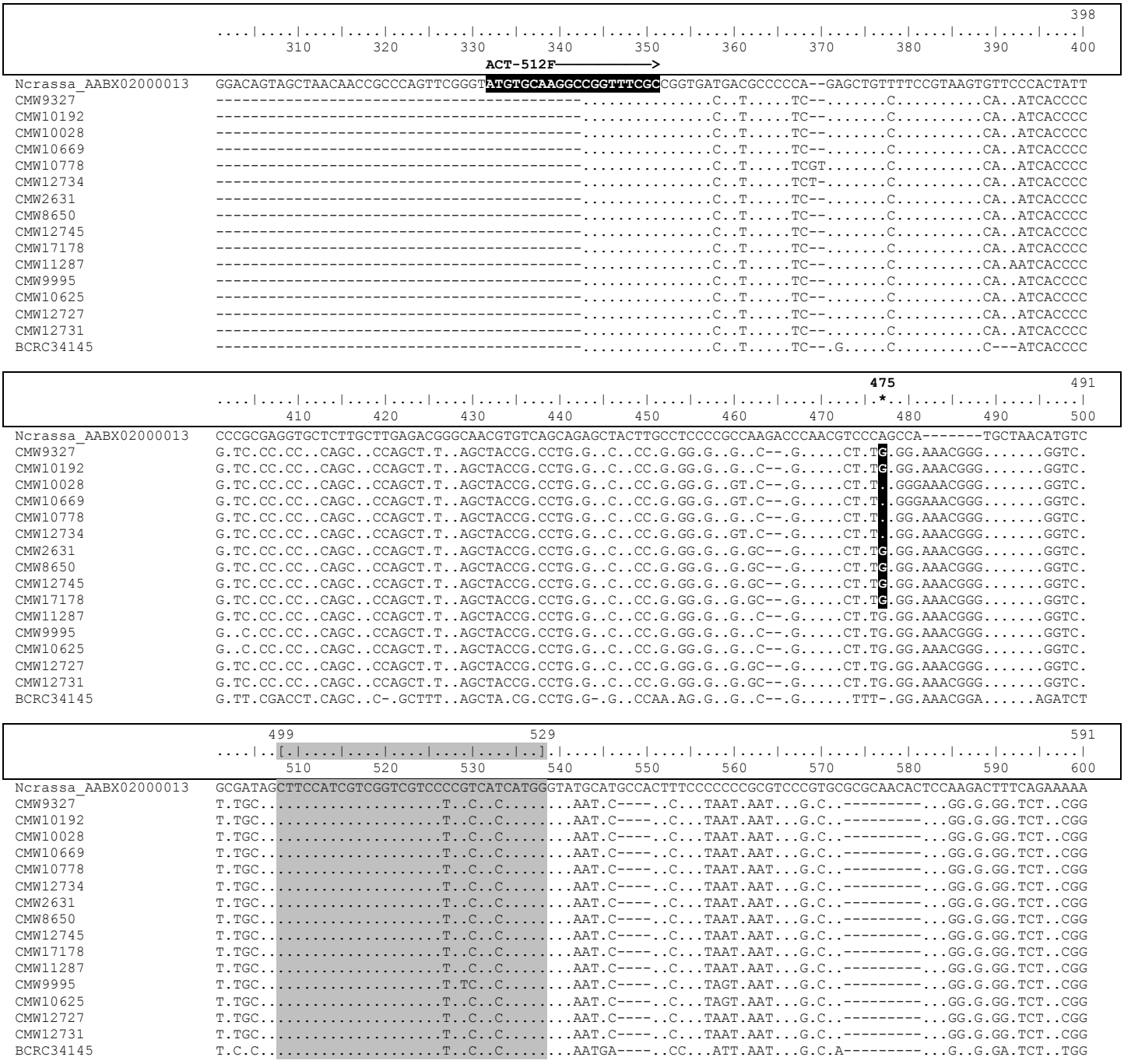
© University of Pretoria

Suppl. Figure 1 Alignment of *Chrysoporthe* spp. and *Neurospora crassa* (GenBank Accession AABX02000013) actin genes. Numbers above the ruler are counted from the start of Exon 1 of the *N. crassa* gene and exclude gaps, while those below are alignment positions. PCR primers are indicated in inverted colors on the *N. crassa* sequence. Exons (1–5 of 7) are shaded, and the nucleotide at position 475 relative to the *N. crassa* sequence can differentiate *C. austroafricana* (G) from *C. cubensis* and *C. deuterocubensis* (A).

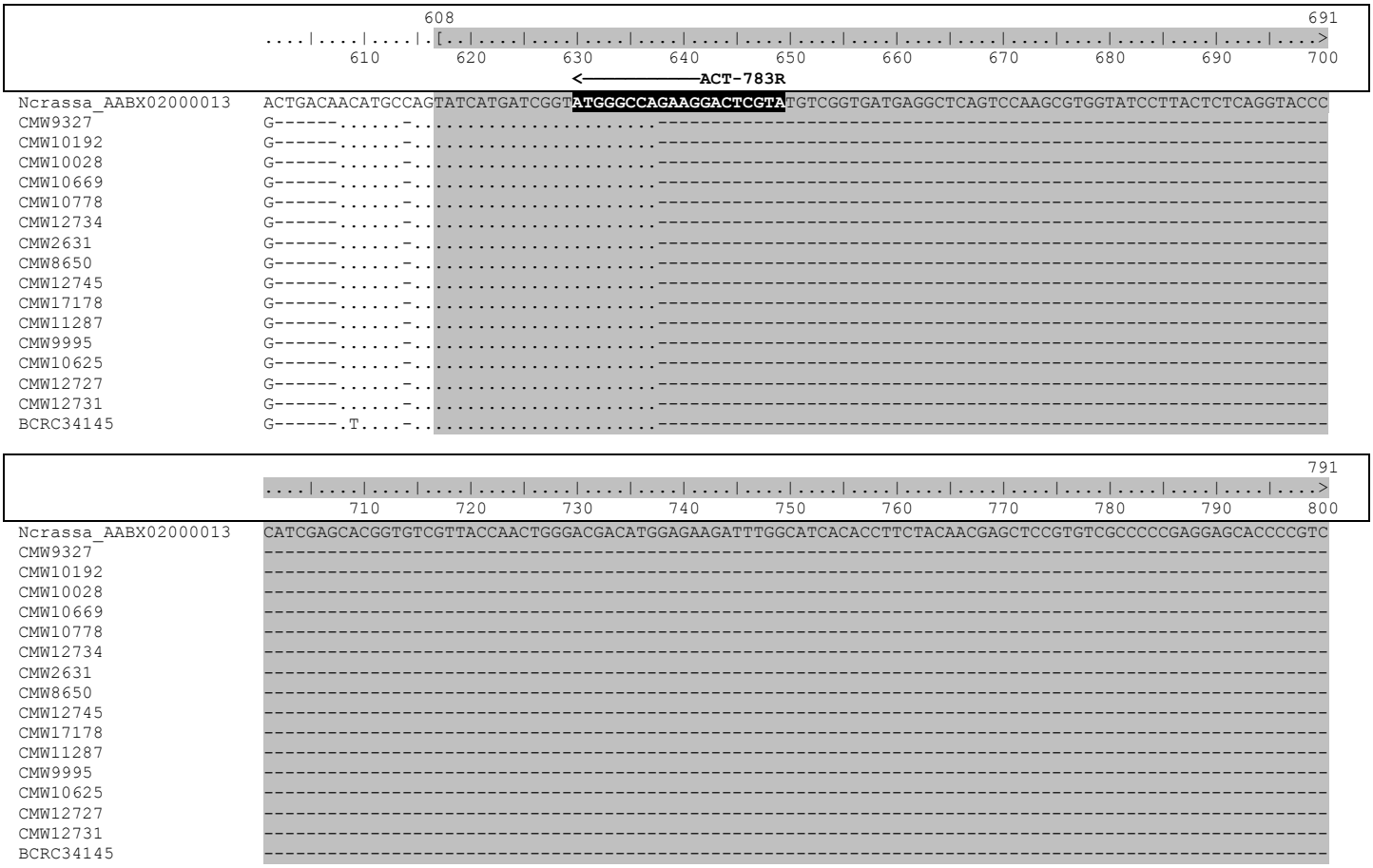
© University of Pretoria



© University of Pretoria



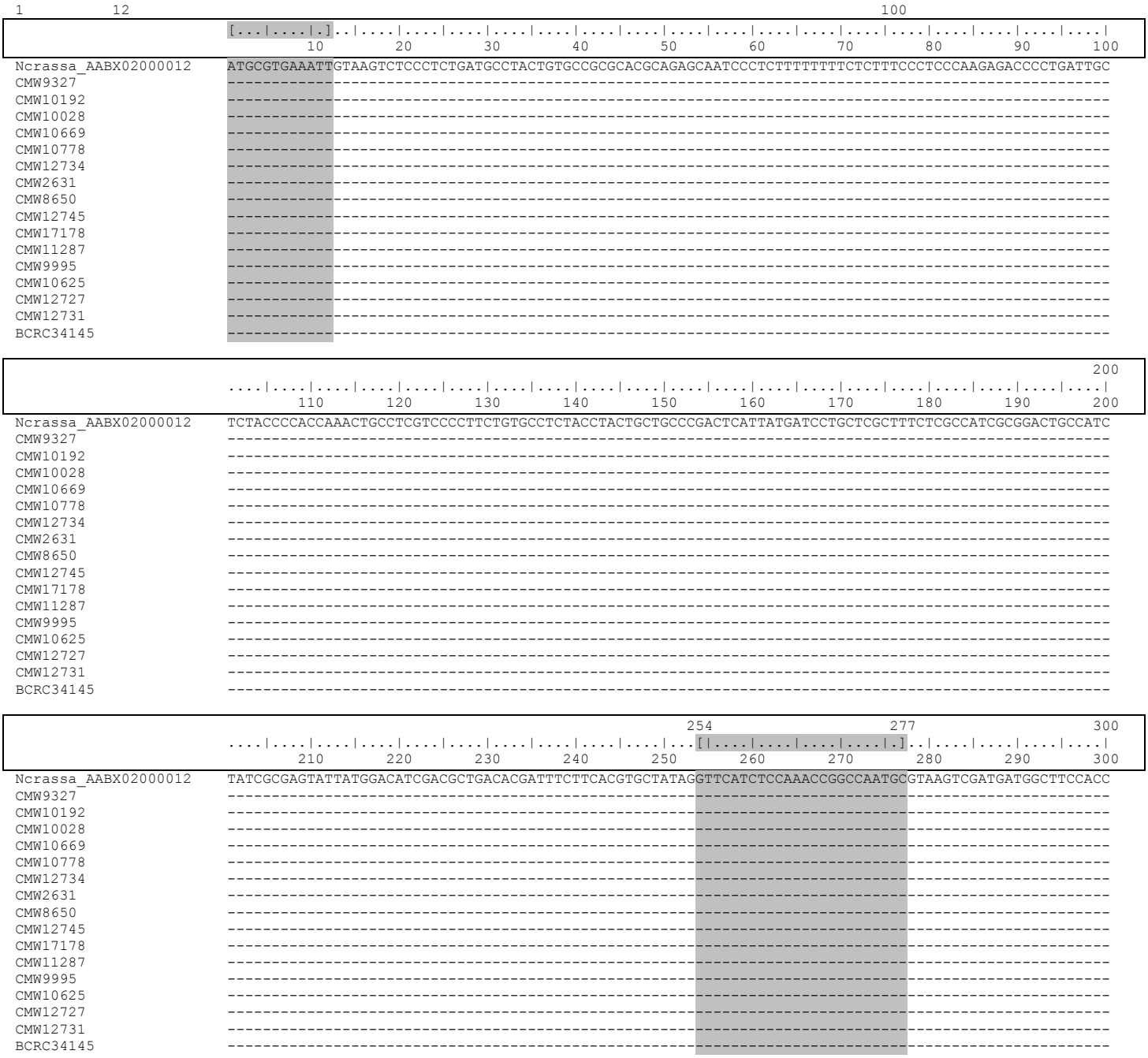
© University of Pretoria



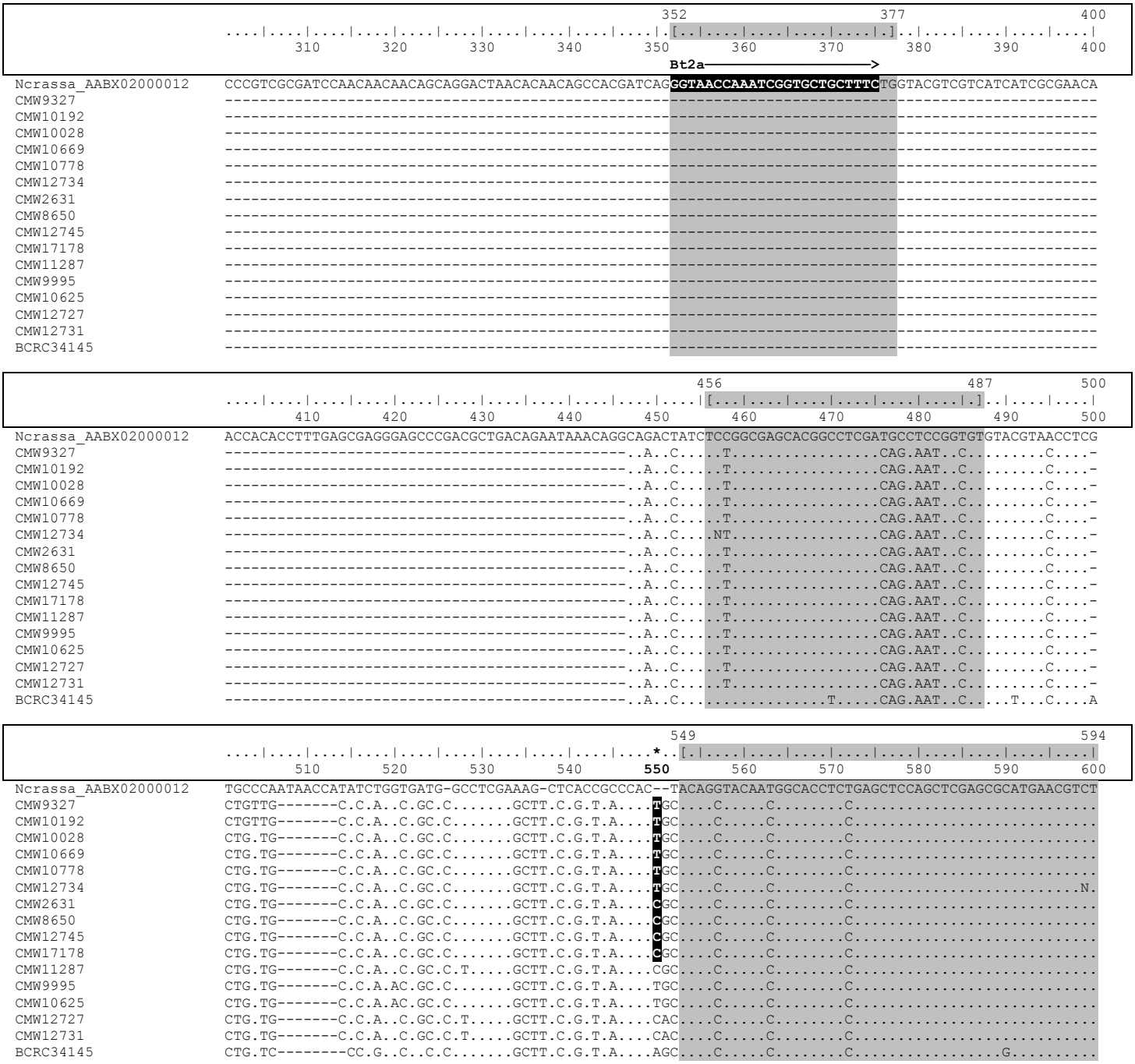
© University of Pretoria

Suppl. Figure 2 Sequences of the BT1 (primers Bt1a to Bt1b) and BT2 (primers Bt2a to Bt2b) regions of the β-tubulin gene aligned to GenBank Accession AABX02000012, the β-tubulin gene from *Neurospora crassa*. Sequence positions above the ruler correspond to the sequence positions of the *N. crassa* gene relative to the start of Exon 1 and excluding gaps, while those under the ruler correspond to alignment positions for this study. Exons 1–7 are indicated with shaded stretches of the alignment, while PCR primers are indicated in inverted colors on the *N. crassa* sequence. Diagnostic nucleotide polymorphisms in the BT1 region (primers Bt1a to Bt1b, positions 1326–1878 in the alignment) that distinguish *C. austroafricana*, *C. cubensis* and *C. deuterocubensis* from each other are indicated in inverted colors. *Ava*I (positions 1402–1407 and 1535–1540 in the alignment) and *Hin*dIII (positions 1516–1521 in the alignment) restriction sites are indicated with inverted grey shading.

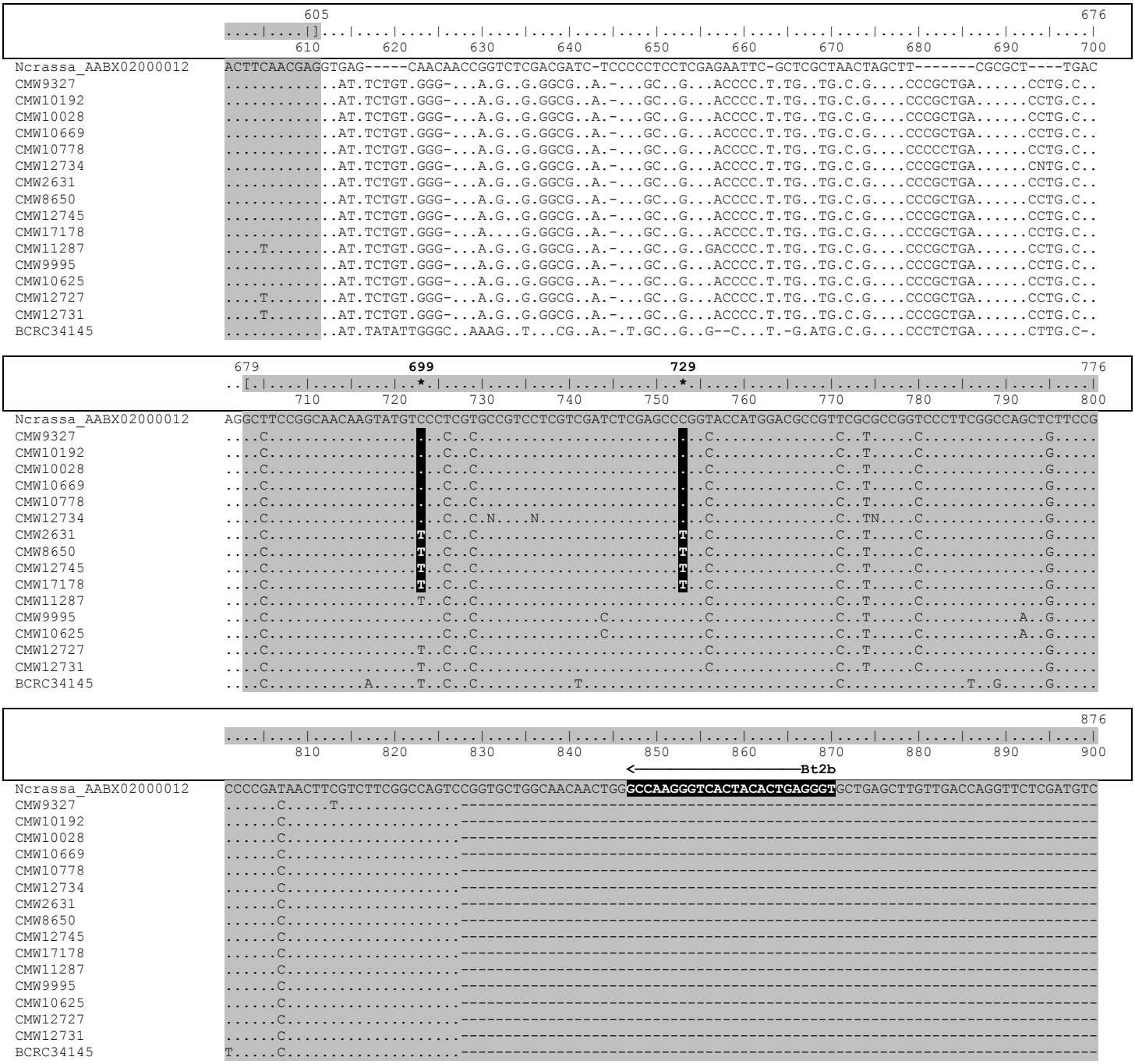
© University of Pretoria



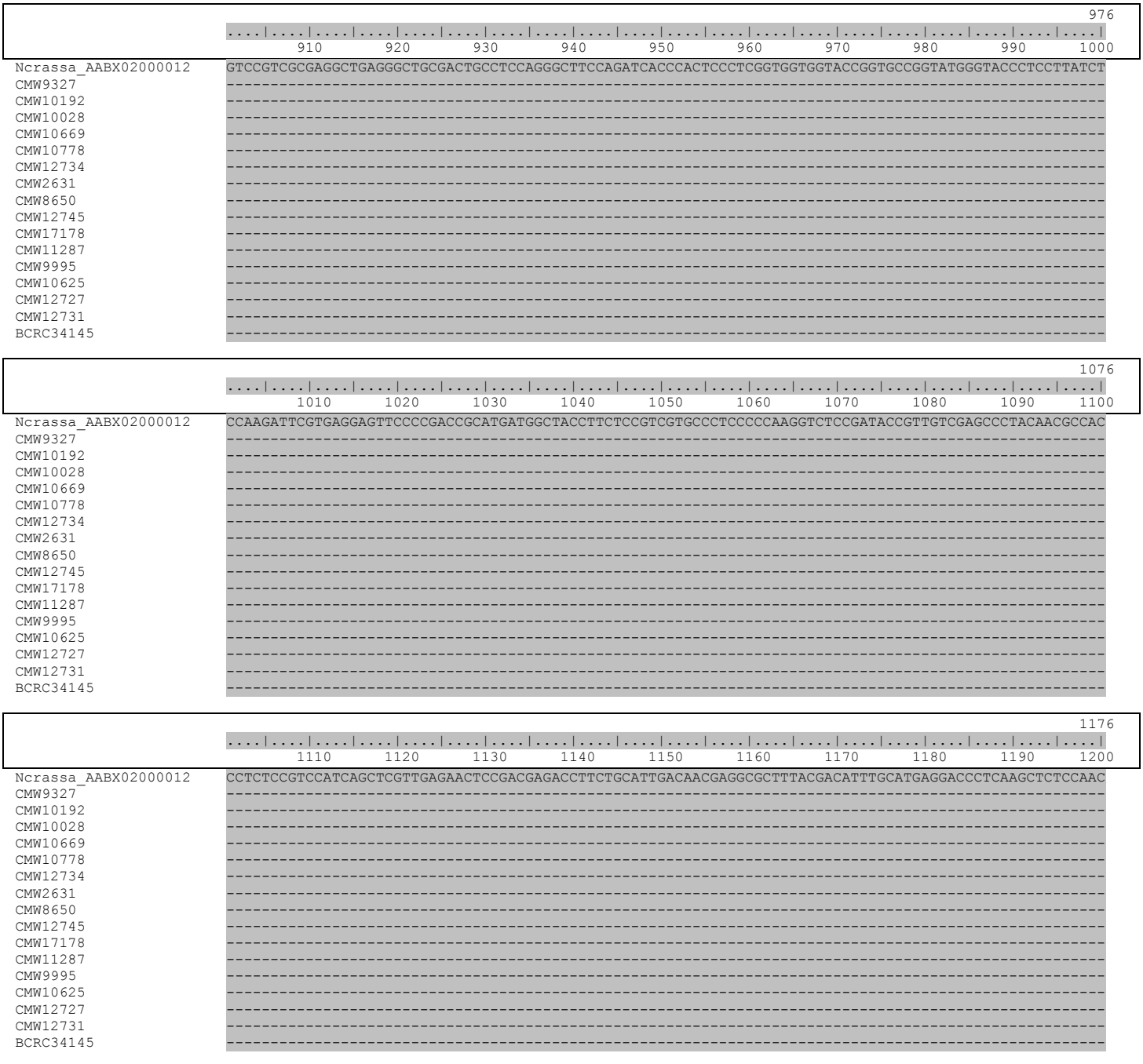
© University of Pretoria



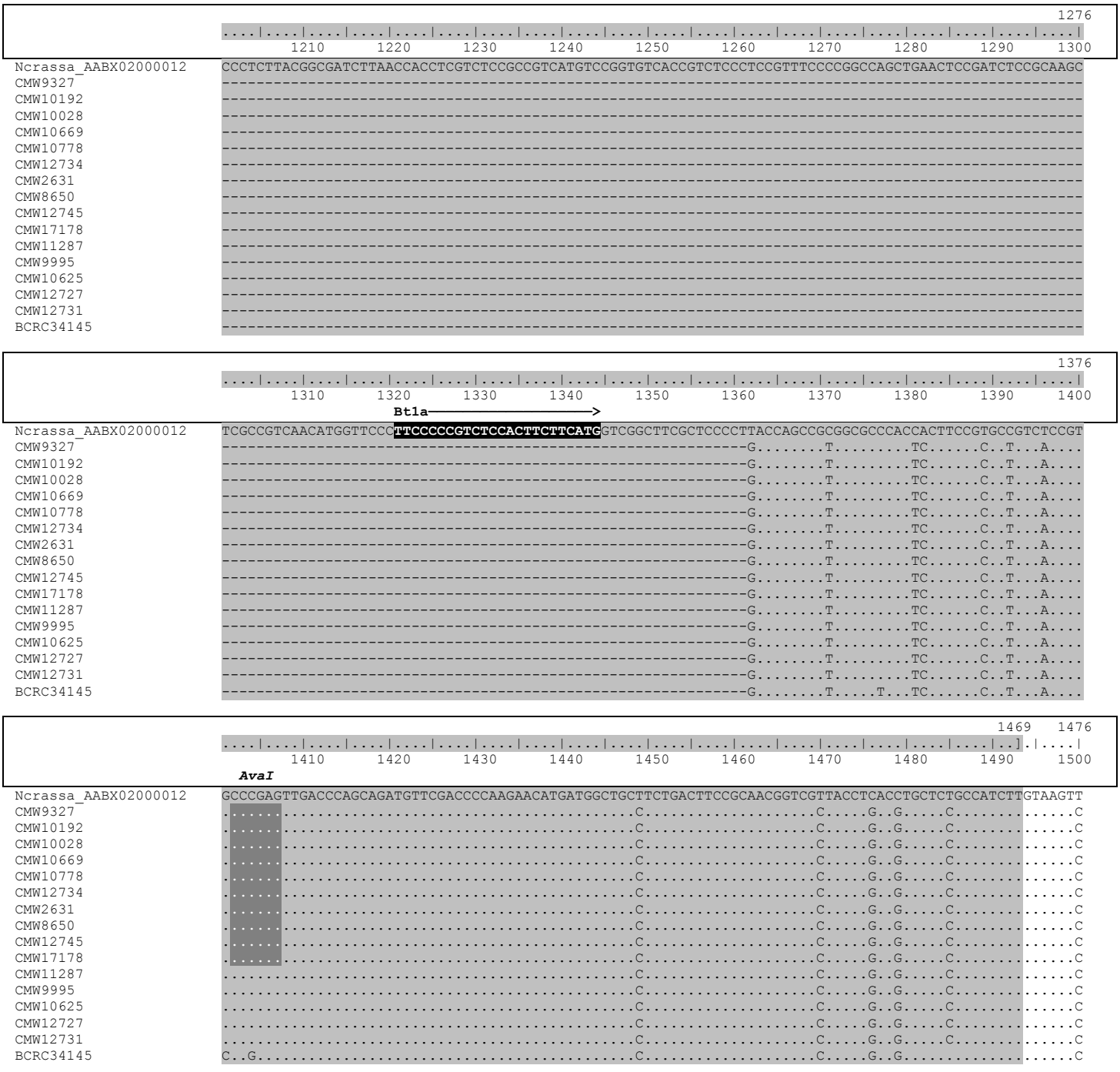
© University of Pretoria



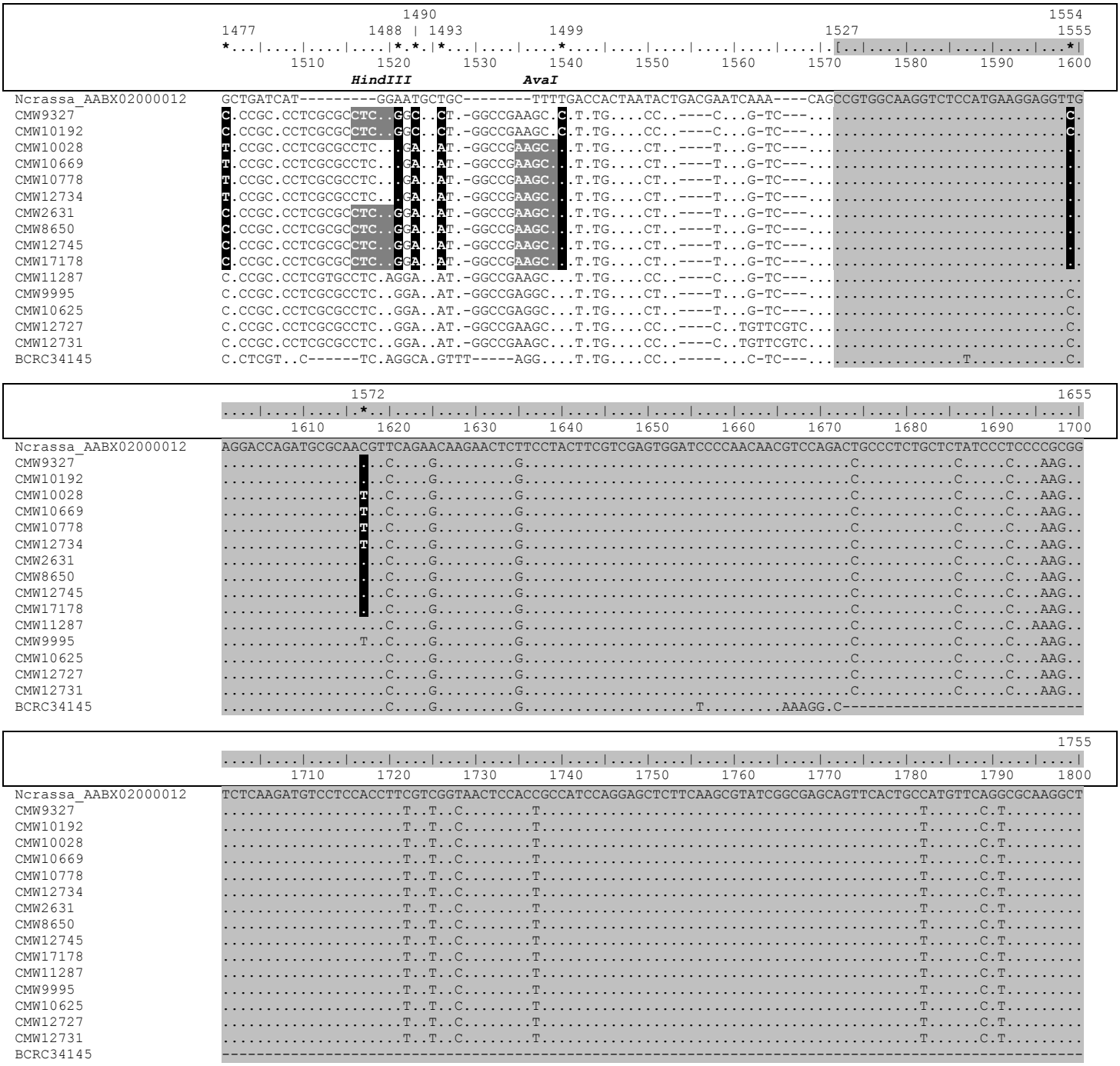
© University of Pretoria



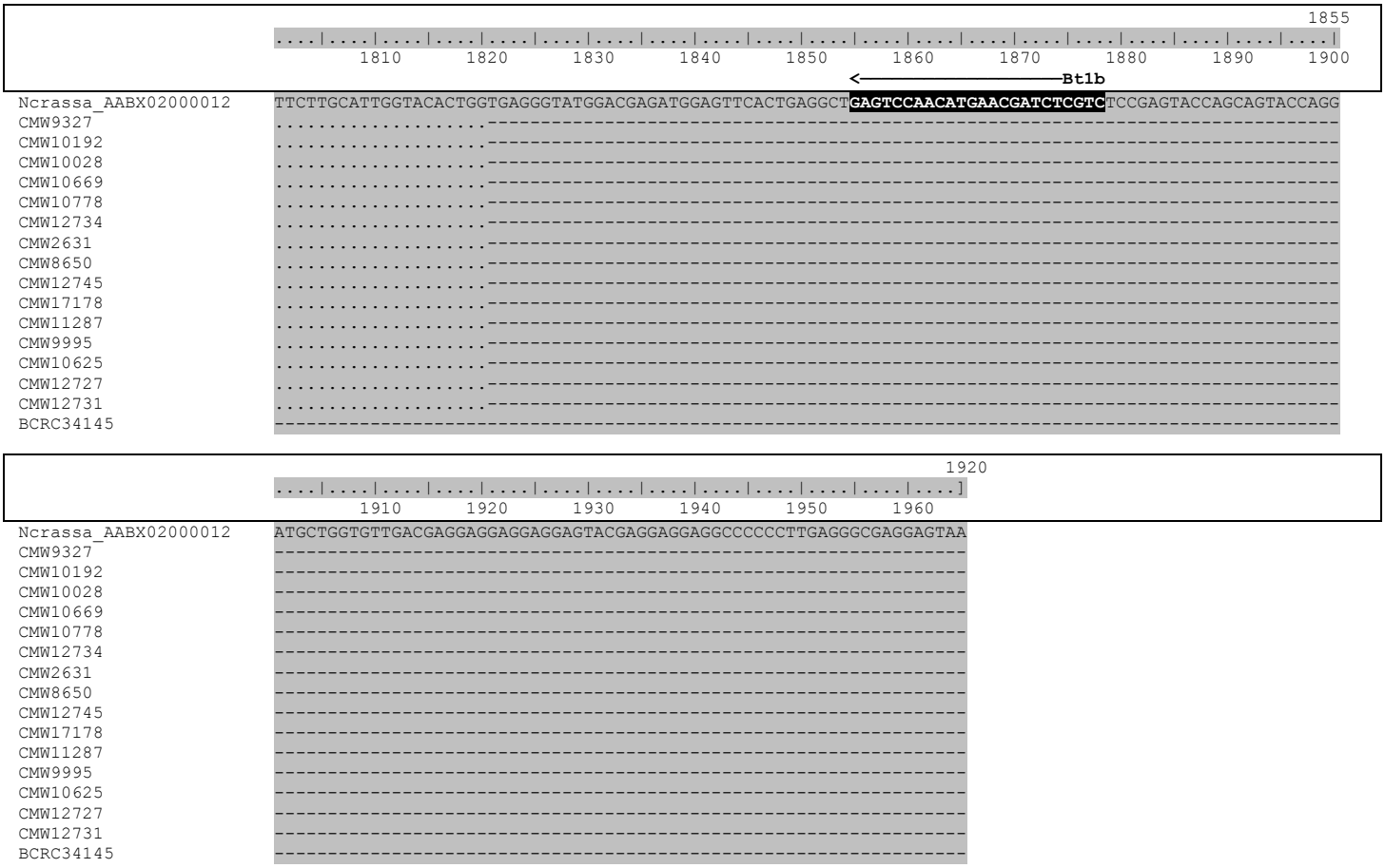
© University of Pretoria



© University of Pretoria



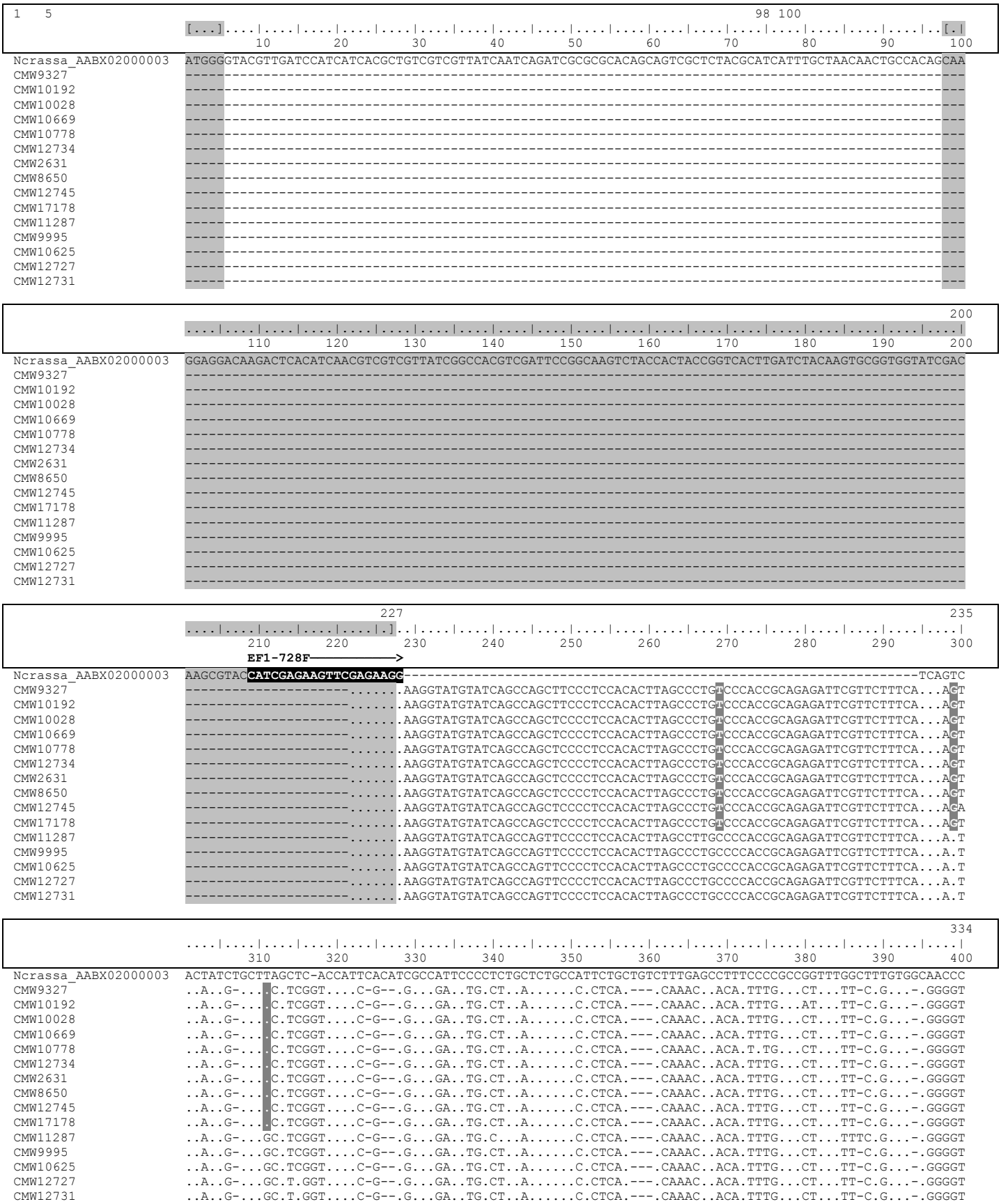
© University of Pretoria



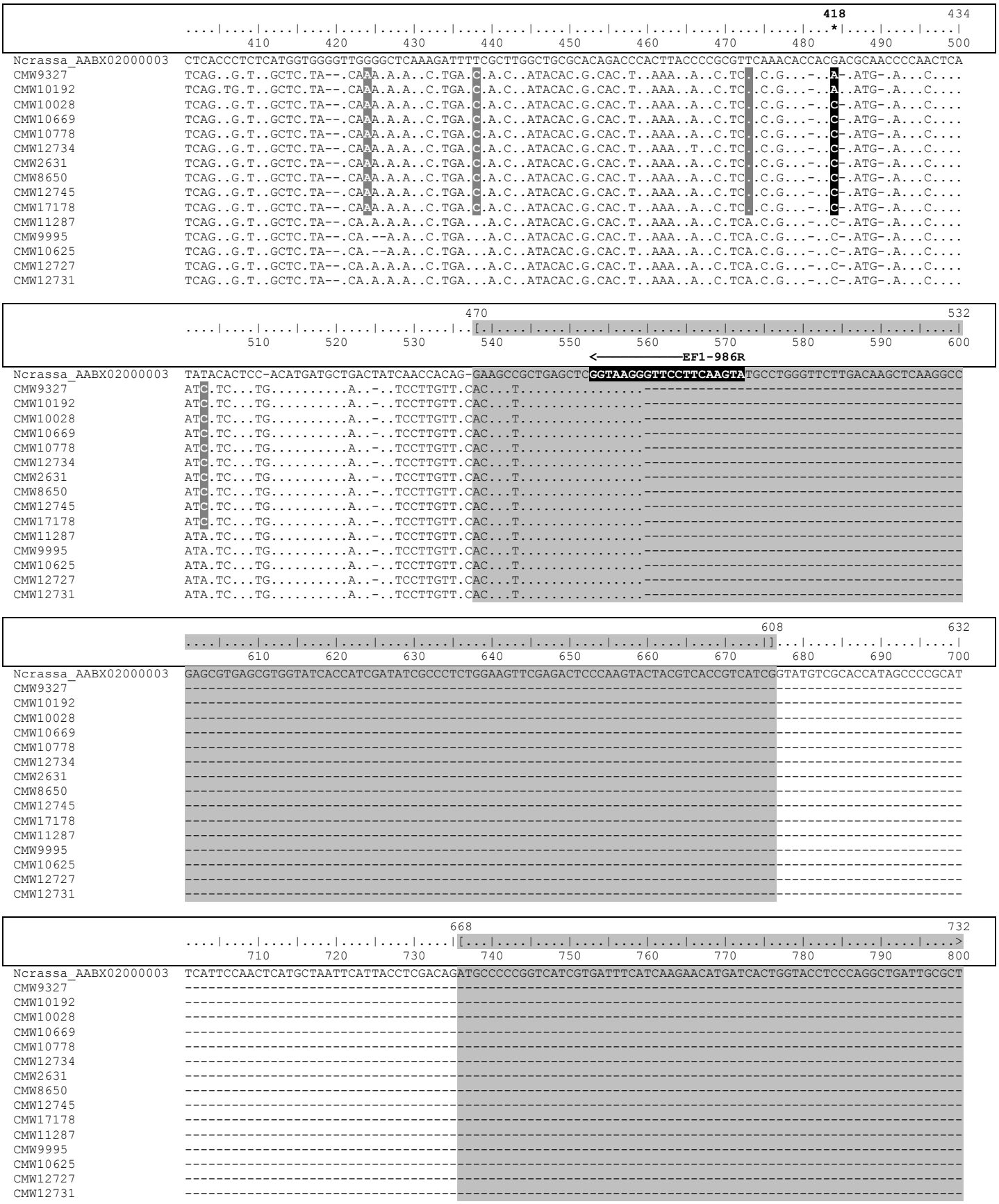
© University of Pretoria

Suppl. Figure 3 Sequence alignment of the EF-1α gene from *Neurospora crassa* (GenBank Accession AABX02000003) with *Chrysoporthe* sequences used in this study. Sequence positions above the ruler correspond to the sequence of the *N. crassa* gene, counted from the first codon and excluding gaps, while those below it are alignment positions. Exons 1–4 are indicated with shaded sequence, while the positions of the PCR primers EF1-728F and EF1-986R are shown in inverted colours on the *N. crassa* sequence. Although the *N. crassa* intron between exons 2 and 3 cannot be reliably aligned to *Chrysoporthe* sequence, it contains nucleotides that are diagnostic for the *C. austroafricana*, *C. cubensis* and *C. deuterocubensis* clade (indicated with inverted light grey nucleotides). The nucleotide at position 484 distinguishes *C. austroafricana* (A) from *C. cubensis* and *C. deuterocubensis* (C) and is indicated in inverted dark grey.

© University of Pretoria



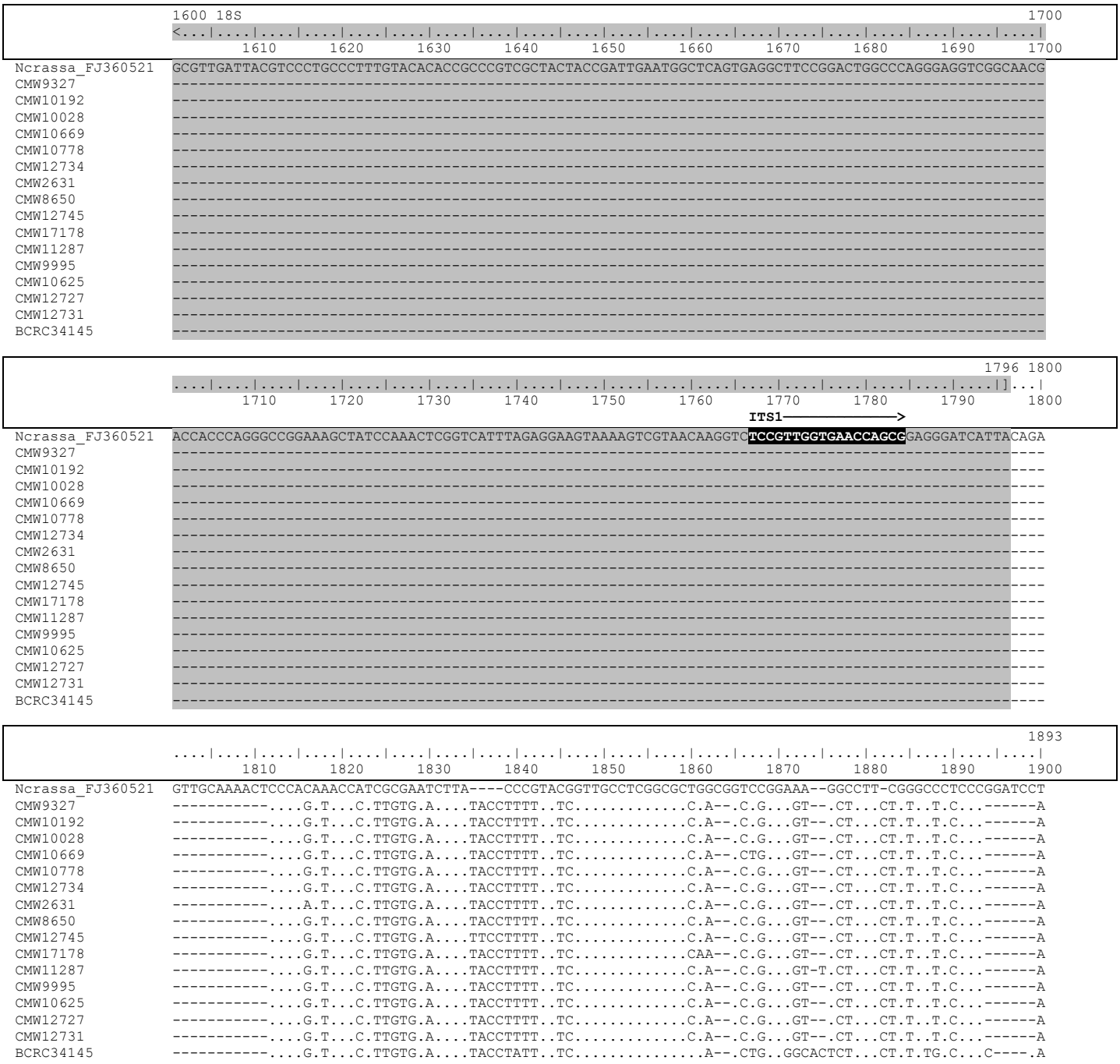
© University of Pretoria



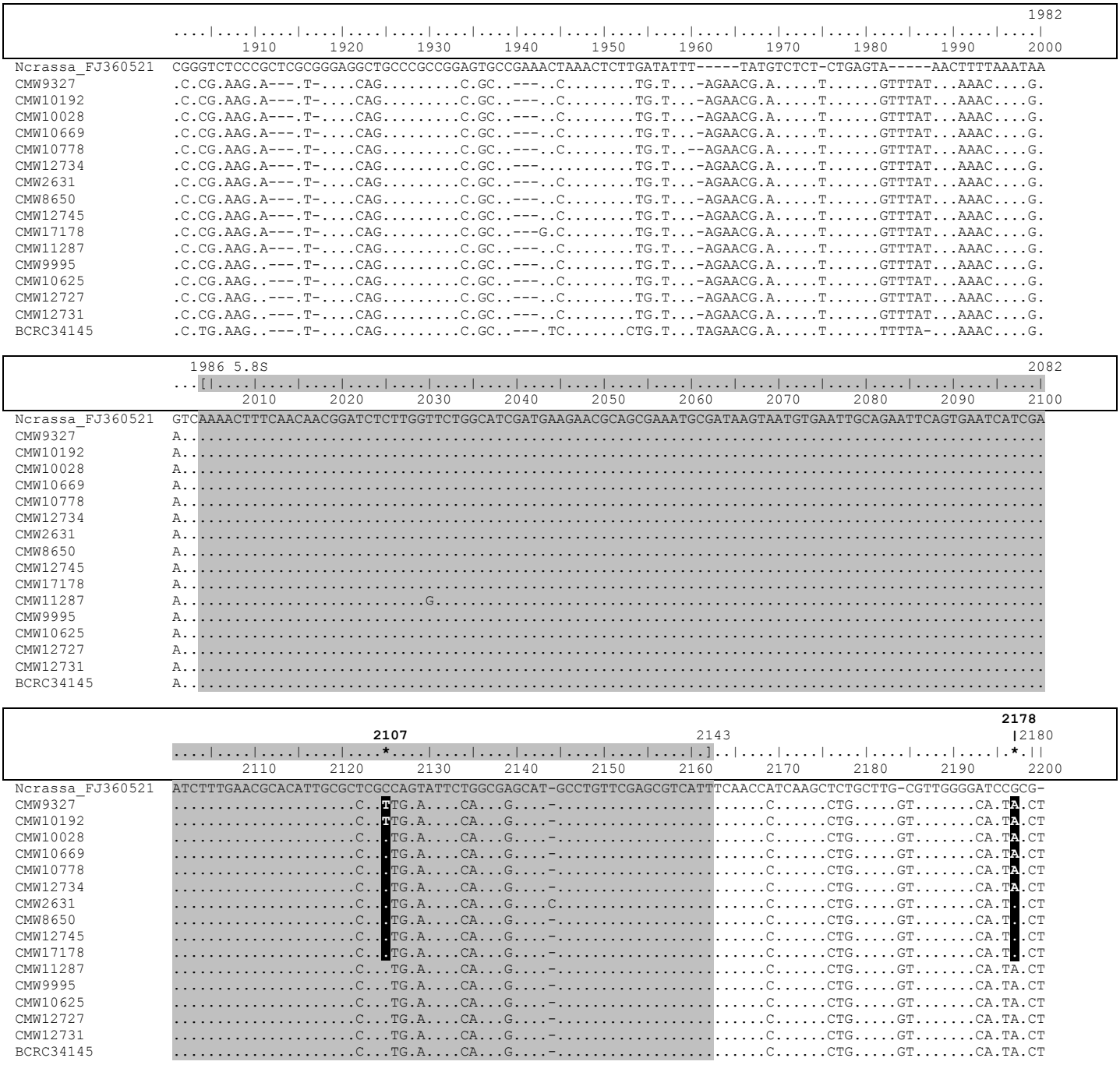
© University of Pretoria

Suppl. Figure 4 Alignment of the ITS1, 5.8S and ITS4 regions of the ribosomal DNA operon sequenced in this study with their homolog in *Neurospora crassa* (GenBank Accession FJ360521). Numbers above the ruler indicate bases from the start of the *N. crassa* 18S gene, while those underneath are alignment positions. The 18S, 5·8S and 28S genes are shaded. Nucleotides that distinguish between *C. austroafricana*, *C. cubensis* and *C deuterocubensis* are highlighted in inverted colors. Locations of the PCR primers ITS1 and ITS4 are indicated on the *N. crassa* sequence.

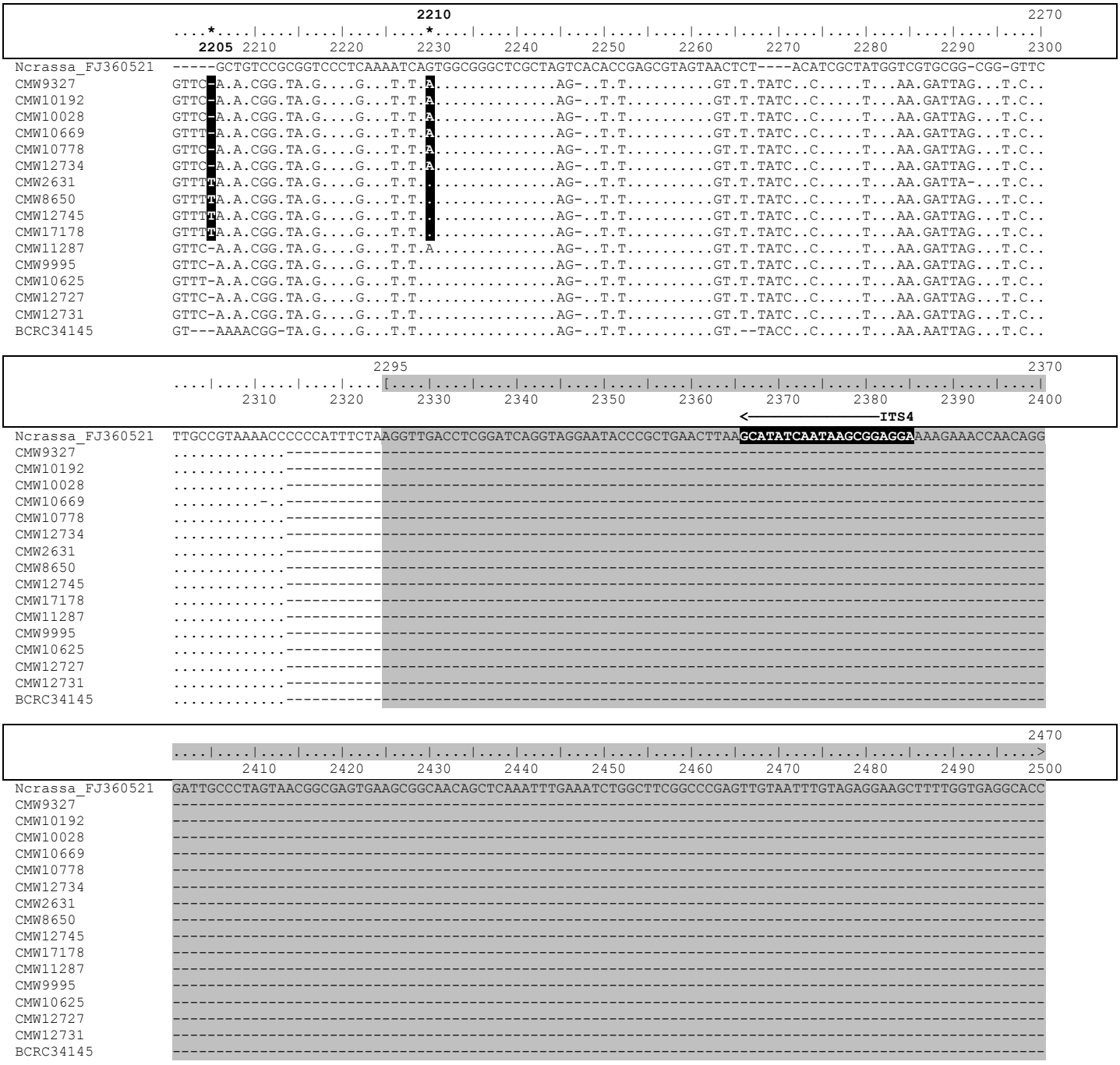
© University of Pretoria



© University of Pretoria



© University of Pretoria



© University of Pretoria

Bibliography

Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* **1**, 101-102.

Avise JC, Wollenberg K (1997) Phylogenetics and the origin of species. *Proceedings of the National Academy of Science, USA* **94**, 7748-7755.

Barraclough TG, Nee S (2001) Phylogenetics and speciation. *TRENDS in Ecology and Evolution* **16**, 391-399.

Boerboom JHA, Maas PW (1970) Canker of *Eucalyptus grandis* and *E. saligna* in Surinam caused by *Endothia havanensis*. *Turrialba* **20**, 94-99.

Bruner SC (1917) Una enfermedad gangrenosa de los eucaliptos. *Estacion Experimental Agronomica Bulletin* **37**, 1-33.

Burt A, Carter DA, Koenig GL, White TJ, Taylor JW (1996) Molecular markers reveal cryptic sex in the human pathogen *Coccidioides immitis*. *Proceedings of the National Academy of Science, USA* **93**, 770-773.

Carbone I, Kohn LM (1999) A method for designing primer sets for speciation studies in filamentous ascomycetes. *Mycologia* **91**, 553-556.

Chungu D, Gryzenhout M, Muimba-Kankolongo A, Wingfield MJ, Roux J (2009) Taxonomy and pathogenicity of two novel *Chrysoporthe* species from *Eucalyptus grandis* and *Syzygium guineense* in Zambia. *Mycological Progress* **9**, 379-393.

Cummings MP, Otto SP, Wakeley J (1995) Sampling properties of DNA sequence data in phylogenetic analysis. *Molecular Biology and Evolution* **12**, 814-822.

Cunningham CW (1997) Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution* **14**, 733-740.

De Queiros K (2007) Species concepts and species delimitation. *Systematic Biology* **56**, 879-886.

Dettman JR, Jacobson DJ, Taylor JW (2003) A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. *Evolution* **57**, 2703-2720.

© University of Pretoria

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797.

Farris JS, Källersjo M, Kluge AG, Bult C (1995) Testing significance of incongruence. *Cladistics* **10**, 315-319.

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

Galtier N, Gouy M, Gautier C (1996) SeaView and Phylo_Win: two graphic tools for sequence alignment and molecular phylogeny. *Computational and Applied BioScience* **12**, 543-548.

Geiser DM, Dorner JW, Horn BW, Taylor JW (2000) The phylogenetics of mycotoxin and sclerotium production in *Aspergillus flavus* and *Aspergillus oryzae*. *Fungal Genetics and Biology* **31**, 169-179.

Geiser DM, Pitt JI, Taylor JW (1998) Cryptic speciation and recombination in the aflatoxin-producing fungus *Aspergillus flavus*. *Proceedings of the National Academy of Sciences, USA* **95**, 388-393.

Geml J, Laursen GA, O'Neill K, Nusbaum HC, Taylor DL (2006) Beringian origins and cryptic speciation events in the fly agaric *Amanita muscaria*. *Molecular Ecology* **15**, 225-239.

Glass NL, Donaldson GC (1995) Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Applied and Environmental Microbiology* **61**, 1323-1330.

Gryzenhout M, Myburg H, Van der Merwe NA, Wingfield BD, Wingfield MJ (2004) *Chrysoporthe*, a new genus to accommodate *Cryphonectria cubensis*. *Studies In Mycology* **50**, 119-142.

Gryzenhout M, Myburg H, Wingfield BD, Montenegro F, Wingfield MJ (2005) *Chrysoporthe doradensis* sp. nov. pathogenic to *Eucalyptus* in Ecuador. *Fungal Diversity* **20**, 39-57.

Gryzenhout M, Myburg H, Wingfield BD, Wingfield MJ (2006a) Cryphonectriaceae (Diaporthales), a new family including *Cryphonectria*, *Chrysoporthe*, *Endothia* and allied genera. *Mycologia* **98**, 239-249.

Gryzenhout M, Rodas CA, Portales JM, Clegg P, Wingfield BD, Wingfield MJ (2006b) Novel hosts of the *Eucalyptus* canker pathogen *Chrysoporthe cubensis* and a new *Chrysoporthe* species from Colombia. *Mycological Research* **110**, 833-845.

© University of Pretoria

Gryzenhout M, Wingfield BD, Wingfield MJ (2006c) New taxonomic concepts for the important forest pathogen *Cryphonectria parasitica* and related fungi. *FEMS Microbiology Letters* **258**, 161-172.

Gryzenhout M, Wingfield BD, Wingfield MJ (2009) *Taxonomy, phylogeny, and ecology of bark-infecting and tree killing fungi in the Cryphonectriaceae*. APS Press.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate phylogenies by maximum likelihood. *Systematic Biology* **52**, 696-704.

Hare MP, Avise JC (1998) Population structure in the American oyster as inferred by nuclear gene genealogies. *Molecular Biology and Evolution* **15**, 119-128.

Heath RN, Gryzenhout M, Roux J, Wingfield MJ (2006) Discovery of the *Cryphonectria* canker pathogen on native *Syzygium* species in South Africa. *Plant Disease* **90**, 433-438.

Hillis DM, Huelsenbeck JP (1992) Signal, noise, and reliability in molecular phylogenetic analyses. *Journal of Heredity* **83**, 189-195.

Hodges CS (1980) The taxonomy of *Diaporthe cubensis*. *Mycologia* **72**, 542-548.

Hodges CS, Alfenas AC, Ferreira FA (1986) The conspecificity of *Cryphonectria cubensis* and *Endothia eugeniae*. *Mycologia* **78**, 343-350.

Hodges CS, Geary TF, Cordell CE (1979) The occurrence of *Diaporthe cubensis* on *Eucalyptus* in Florida, Hawaii, and Puerto Rico. *Plant Disease Reporter* **63**, 216-220.

Hodges CS, Reis MS, Ferreira FA, Henfling JDM (1976) O cancro do eucalipto causado por *Diaporthe cubensis*. *Fitopatalogia Basileira* **1**, 129-167.

Hudson RR, Coyne JA (2002) Mathematical consequences of the genealogical species concept. *Evolution* **56**, 1557-1567.

Ju Y-M, Hsieh H-M, Ho M-C, Szu D-H, Fang M-J (2007) *Theissenia rogersii* sp. nov. and phylogenetic position of *Theissenia*. *Mycologia* **99**, 612-621.

Koufopanou V, Burt A, Szaro T, Taylor JW (2001) Gene genealogies, cryptic species, and molecular evolution in the human pathogen *Coccidioides immitis* and relatives (Ascomycota, Onygenales). *Molecular Biology and Evolution* **18**, 1246-1258.

Liu M, Milgroom MG, Chaverri P, Hodge KT (2009) Speciation of a tropical fungal species pair following transoceanic dispersal. *Molecular Phylogenetics and Evolution* **51**, 413-426.

© University of Pretoria

Milgroom MG, Sotirovski K, Spica D, Davis JE, Brewer MT, Milev M, Cortesi P (2008) Clonal population structure of the chestnut blight fungus in expanding ranges in southeastern Europe. *Molecular Ecology* **17**, 4446-4458.

Myburg H, Gryzenhout M, Heath R, Roux J, Wingfield BD, Wingfield MJ (1999a) Cryphonectria canker on *Tibouchina* in South Africa. *Mycological Research* **106**, 1299-1306.

Myburg H, Gryzenhout M, Wingfield BD, Stipes RJ, Wingfield MJ (2004) Phylogenetic relationships of *Cryphonectria* and *Endothia* species, based on DNA sequence data and morphology. *Mycologia* **96**, 990-1001.

Myburg H, Gryzenhout M, Wingfield BD, Wingfield MJ (2002) ß-tubulin and histone *H3* gene sequences distinguish between *Cryphonectria cubensis* from South Africa, Asia, and South America. *Canadian Journal of Botany* **80**, 590-596.

Myburg H, Gryzenhout M, Wingfield BD, Wingfield MJ (2003) Conspecificity of *Endothia eugeniae* and *Cryphonectria cubensis*: a re-evaluation based on morphology and DNA sequence data. *Mycoscience* **44**, 187-196.

Myburg H, Wingfield BD, Wingfield MJ (1999b) Phylogeny of *Cryphonectria cubensis* and allied species inferred from DNA analysis. *Mycologia* **91**, 243-250.

Nakabonge G, Gryzenhout M, Wingfield BD, Wingfield MJ, Roux J (2007) Genetic diversity of *Chrysoporthe cubensis* in eastern and southern Africa. *South African Journal of Science* **103**, 261-264.

Nakabonge G, Roux J, Gryzenhout M, Wingfield MJ (2006) Distribution of Chrysoporthe canker pathogens on *Eucalyptus* and *Syzygium* spp. in eastern and southern Africa. *Plant Disease* **90**, 734-740.

O'Donnell K, Ward TJ, Geiser DM, Kistler HC, Aoki T (2004) Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genetics and Biology* **41**, 600-623.

Old KM, Wingfield MJ, Yuan ZQ (2003) A manual of diseases of eucalypts in South-East Asia. Centre for International Forestry Research, Jakarta, Indonesia.

© University of Pretoria

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ (2008) Multiple gene genealogies and phenotypic data reveal cryptic species of the Botryosphaeriaceae: A case study on the *Neofusicoccum parvum* / *N. ribis* complex. *Molecular Phylogenetics and Evolution*.

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ (2009) Molecular and phenotypic characterization of three phylogenetic species discovered within the *Neofusicoccum parvum*/*N. ribis* complex. *Mycologia* **101**, 636-647.

Pegg GS, Gryzenhout M, O'Dwyer C, Drenth A, Wingfield MJ (2010) The *Eucalyptus* canker pathogen *Chrysoporthe cubensis* discovered in eastern Australia. *Australasian Plant Pathology* **39**, 343-349.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.

Pringle A, Baker DM, Platt JL, Wares JP, Latgé JP, Taylor JW (2005) Cryptic speciation in the cosmopolitan and clonal human pathogenic fungus *Aspergillus fumigatus*. *Evolution* **59**, 1886-1899.

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

Rosenberg NA (2003) The shapes of neutral gene genealogies in two species: probabilities of monophyly, paraphyly, and polyphyly in a coalescent model. *Evolution* **57**, 1465-1477.

Roux J, Meke G, Kanyi B, Mwangi L, Mbaga A, Hunter GC, Nakabonge G, Heath RN, Wingfield MJ (2005) Diseases of plantation forestry trees in eastern and southern Africa. *South African Journal of Science* **101**, 1-5.

Roux J, Myburg H, Wingfield BD, Wingfield MJ (2003) Biological and phylogenetic analyses suggest that two *Cryphonectria* spp. cause cankers of *Eucalyptus* in Africa. *Plant Disease* **87**, 1329-1332.

Sharma JK, Mohanan C, Florence EJM (1985) Occurrence of Cryphonectria canker disease of *Eucalyptus* in Kerala, India. *Annals of Applied Biology* **106**, 265-276.

Sites JW, Marshall JC (2003) Delimiting species: a renaissance issue in systematic biology. *TRENDS in Ecology and Evolution* **18**, 462-470.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

© University of Pretoria

Slippers B, Stenlid J, Wingfield MJ (2005) Emerging pathogens: fungal host jumps following anthropogenic introduction. *TRENDS in Ecology and Evolution* **20**, 420-421.

Sokal RR, Rohlf FJ (1994) *Biometry: the principles and practice of statistics in biological research*., 3rd edn. Freeman & Co., San Francisco.

Steenkamp ET, Wingfield BD, Coutinho TA, Wingfield MJ, Marasas WFO (1999) Differentiation of *Fusarium subglutinans* f. sp. *pini* by histone gene sequence data. *Applied and Environmental Microbiology* **65**, 3401-3406.

Steenkamp ET, Wingfield BD, Desjardins AE, Marasas WFO, Wingfield MJ (2002) Cryptic speciation in *Fusarium subglutinans*. *Mycologia* **94**, 1032-1043.

Steenkamp ET, Wright J, Baldauf SL (2006) The protistan origins of animals and fungi. *Molecular Biology and Evolution* **23**, 93-106.

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

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.0 beta.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.

Taylor JW, Jacobson DJ, Fisher MC (1999) The evolution of asexual fungi: reproduction, speciation and classification. *Annual Review of Phytopathology* **37**, 197-246.

Taylor JW, Jacobson DJ, Kroken S, Kasuga T, Geiser DM, Hibbett DS, Fisher MC (2000) Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology* **31**, 21-32.

Van der Merwe NA, Wingfield BD, Wingfield MJ (2003) Primers for the amplification of sequence-characterized loci in *Cryphonectria cubensis* populations. *Molecular Ecology Notes* **3**, 494-497.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenies. In: *PCR protocols: a guide to methods and applications*. (eds. Innes

© University of Pretoria

MA, Gelfand DH, Sninsky SS, White TJ), pp. 315-322. Academic Press Inc., San Diego, California.

Wingfield MJ (2003) Daniel McAlpine Memorial Lecture. Increasing threat of diseases to exotic plantation forests in the Southern Hemisphere: lessons from Cryphonectria canker. *Australasian Plant Pathology* **32**, 133-139.

Zhou X, Burgess TI, De Beer ZW, Lieutier F, Yart A, Klepzig K, Carnegie A, Mena Portales J, Wingfield BD, Wingfield MJ (2007) High intercontinental migration rates and population admixture in the sapstain fungus *Ophiostoma ips*. *Molecular Ecology* **16**, 89-99.

© University of Pretoria

Chapter 4

Continental drift fails to explain allopatric speciation patterns of Diaporthalean tree pathogens

© University of Pretoria

Abstract

Fungal plant pathogens are generally believed to have evolved by allopatric speciation, which was driven by the emergence of population differentiation due to geographic barriers to gene flow. Chronometric phylogenies of the fungi support this view, because most of the fungal stem lineages, *i.e.* ordinal and familial lineages were present at the end of the Gondwanan break-up. The Diaporthales is one such order, which includes numerous saprobes and plant pathogens such as the Cryphonectriaceae including important tree pathogens such as the chestnut blight fungus, *Cryphonectria parasitica*, and *Chrysoporthe* spp., that kill *Eucalyptus* spp. and other Myrtales. Species of *Chrysoporthe* with an apparent continental origin have been identified using DNA sequence phylogenies, but these species are exceptionally closely related with very short terminal branches. In this study we test whether the continental species of *Chrysoporthe* evolved via allopatric speciation. Calibrated phylogenies based on *rrl* (18S rDNA) data, as well as ITS rDNA and β-tubulin sequences, were used to estimate the time to the most recent common ancestor of the genus *Chrysoporthe*. Results showed that *Chrysoporthe* emerged *c*. 7 Ma, placing its evolution well after the break-up of Gondwana. Therefore, we considered the evolution of the Myrtales and their spread around the world during the break-up of Gondwana. The data were sufficient to show that *Chrysoporthe* species have evolved after Myrtalean stem lineages were already present. Thus, we believe that the ancestor of *Chrysoporthe* probably moved around the world by utilising dispersal of ancestral Myrtales after the break-up of Gondwana. This would have been followed by allopatric speciation in the various regions where the fungus became established. Furthermore, trans-oceanic dispersal can explain the presence of all the *Chrysoporthe* species in the regions where they are currently found.

© University of Pretoria

Introduction

The ascomycete genus *Chrysoporthe* includes numerous species that are important canker pathogens of trees in the Myrtales. *Chrysoporthe austroafricana* (Gryzenhout *et al*. 2004) is a pathogen of Myrtaceae and Melastomataceae on the African continent (Gryzenhout *et al*. 2009; Heath *et al*. 2006; Myburg *et al*. 1999; Nakabonge *et al*. 2007; Wingfield *et al*. 1989) where it causes a serious stem canker disease on commercially-grown *Eucalyptus* spp. (Nakabonge *et al*. 2006; Roux *et al*. 2005; Roux *et al*. 2003; Van Staden *et al*. 2004; Wingfield 1999; Wingfield *et al*. 1997). *Chrysoporthe cubensis* occurs on Myrtaceae and Melastomataceae in Africa, but is native to South America and Central America where it is associated with basal stem cankers of *Eucalyptus* spp. (Gryzenhout *et al*. 2004; Gryzenhout *et al*. 2009; Hodges *et al*. 1979; Hodges & Reis 1974; Hodges *et al*. 1976; Van der Merwe *et al*. 2001). Recently, *Chrysoporthe deuterocubensis* was described for a closely related fungus that is common and probably native on Myrtales in Southeast Asia (Van der Merwe *et al*. 2010). The *Chrysoporthe* anamorph, *Chrysoporthella hodgesiana*, infects native *Tibouchina* and *Miconia* species in Colombia (Gryzenhout *et al*. 2004). Two of the remaining species in the genus, *C. doradensis* and *C. inopina*, cause stem cankers on *Eucalyptus* spp. in Ecuador and *Tibouchina* spp. in Colombia, respectively (Gryzenhout *et al*. 2005a; Gryzenhout *et al*. 2006b).

The genus *Chrysoporthe* resides in the Cryphonectriaceae, order Diaporthales (Gryzenhout *et al*. 2006a; Gryzenhout *et al*. 2009), which also includes *Cryphonectria parasitica*, the causal agent of chestnut blight (Anagnostakis 1987; Merkel 1905). Most other members of *Cryphonectria* are non-pathogenic saprobes of woody plants, and they include *C. macrospora*, *C. nitschkei* and *C. radicalis* (Gryzenhout *et al*. 2009; Myburg *et al*. 2004). In addition to *Chrysoporthe* and *Cryphonectria*, the Cryphonectriaceae includes other genera such as *Amphilogia*, *Cryptodiaporthe*, *Holocryphia*, *Microthia, Rostraureum*, *Ursicollum*, and *Wuestneia* (Gryzenhout *et al*. 2005b; Gryzenhout *et al*. 2009). Most of the species in these genera are regarded as saprobic, but may be opportunistic or weak pathogens such as is true for *Microthia havanensis*, *Rostraureum tropicale* and *Ursicollum fallax* (Gryzenhout *et al*. 2005b). Unfortunately, many of the genera in the Cryphonectriaceae are monotypic, suggesting that the family as a whole is vastly undersampled (Gryzenhout *et al*. 2009).

© University of Pretoria

A number of studies have been undertaken to understand the phylogeography of *Chrysoporthe cubensis* and its closest relatives (Myburg *et al*. 2002; Myburg *et al*. 1999). Based on the geographic distributions of these fungi, it has been suggested that *Chrysoporthe* spp. have specific continental origins (Gryzenhout *et al*. 2006c; Nakabonge *et al*. 2007; Van der Merwe *et al*. 2010). While *C. austroafricana* occurs exclusively on the African continent, *C. cubensis* and *C. deuterocubensis* occur in South America and Southeast Asia, respectively, but also disjunctly with introductions into Western and Eastern Africa (Van der Merwe *et al*. 2010). The origins of these species have not been closely examined, but various authors have suggested endemism for each of the species (Gryzenhout *et al*. 2009; Hodges 1980; Seixas *et al*. 2004). To explain the continental associations of these species it can be argued that the ancestor of *Chrysoporthe* evolved before the fragmentation of Gondwana (Upchurch 2008), leading to initial vicariance, *i.e*. population differentiation due to the emergence of geographic or ecological barriers (Rosen 1978). Allopatric speciation (Johannesson 2001; Voelker 1999) would follow naturally after geographic vicariance by continental drift.

Closely related tree pathogenic fungi that display clear phylogeographic separation are thought to have speciated mainly through allopatry, as many examples exist where closely related fungi are geographically separated (Dettman *et al*. 2003; Koufopanou *et al*. 1997; O'Donnell *et al*. 2000; O'Donnell *et al*. 2004). It is generally believed that native species have diverged by continental drift or other more recent geographic and climatic phenomena, including changes in the distributions of host plants (Anderson *et al*. 2004; Dimichele *et al*. 2009). Conversely, sympatric speciation in fungi presents difficult theoretical questions (Giraud *et al*. 2008), but can usually be explained by anthropogenic dispersal linked to human activity (Coetzee *et al*. 2001; Slippers *et al*. 2005), or host and habitat specificity (Giraud 2006; Giraud *et al*. 2006). Examples of sympatric speciation in fungal pathogens include *Mycosphaerella graminicola*, which causes leaf spot of wheat (Stukenbrock *et al*. 2007), and *Ascochyta* spp. causing blight of legumes (Abbo *et al*. 2007).

In this study we tested the hypothesis that the three most important species of *Chrysoporthe* have continental origins and arose through allopatric speciation after the break-up of Gondwana. For this purpose, we used a coalescent approach incorporating speciation (Doolittle *et al*. 1996; Douzery *et al*. 2004; Roger & Hug 2006) and calibration based on fossil evidence. Overall, the aim was to gain an understanding of the time of divergence of

© University of Pretoria

these pathogenic fungal assemblages, and attempt to infer the conditions that led to their appearance. This information could be helpful to pre-emptively plan for imminent climate change and rapid species divergence, which could lead to serious new tree disease problems.

Materials & Methods

Data mining and data sets

In order to apply fossil calibration points to the deeper branches of the fungi, a data set including most of the major lineages of the Ascomycetes, Basidiomycetes, as well as some animals, was compiled. To accomplish this, we utilized previously published small subunit (18S) ribosomal RNA (rRNA) gene (*rrl*) data sets (Berbee 1996; Berbee & Taylor 1995; Mori *et al*. 2000; Padovan *et al*. 2005) and focused on the Sordariomycetes while incorporating a diverse sample of fungal lineages. These data sets were amended with *rrl* sequences for representatives of the Diaporthales (Table 1). In order to resolve genera and species within the Diaporthales, a second data set was compiled using sequence information for the internal transcribed spacer (ITS) rRNA region (White *et al*. 1990) and the β-tubulin gene (Glass & Donaldson 1995) (Table 2). The ITS region spanned the ITS1 and ITS2 internal transcribed spacers as well as the 5·8S gene, while the β-tubulin sequences spanned from the 3' end of exon 4 to the 5' end of exon 6 and included two introns as well as the entire exon 5. All *rrl*, ITS and β-tubulin sequences were obtained from the NCBI online database (http://www.ncbi.nlm.nih.gov), and were aligned using Muscle 3·52 (Edgar 2004) and manually adjusted using SeaView 2·2 (Galtier *et al*. 1996).

Calibation of the molecular clock

Doolittle et al. (1996) dated the divergence of animals and fungi at *c*. 965 Mya (million years ago) and based their analysis on protein sequences of 57 enzymes, which was supported independently by Lynch (1999). The earliest date of emergence for the Glomalean lineages is 460 Mya (Remy *et al*. 1994; Taylor *et al*. 2005), based on fossil Glomalean fungi preserved in dolomite from the Ordovician period (490 – 443 Mya) (Redecker *et al*. 2000). These lineages, together with two *Mucor* species, represented the Glomeromycetes and Zygomycetes in our *rrl* phylogeny. The earliest evidence of clamp connections, associated with the Basidiomycota, is 290 Mya (Dennis 1969, 1970, 1976; Osborn *et al*. 1989) at the

© University of Pretoria

boundary between the Permian and Pennsylvanian Carboniferous epochs. Polypore fossils, represented by *Boletus satanas*, *Cantharellus tubaeformis, Chroogomphus vinicolor*, and *Lentinellus ursinus* in our *rrl* phylogeny, have been found in amber from the Early Cretaceous period (144 – 99 Mya), and these were used as a 118 Mya calibration point to their most recent common ancestor (Smith *et al*. 2004). The discovery of Sordariomycete fossils showing flask-shaped and ostiolate perithecia, and description of *Paleopyrenomycites devonicus* (Taylor *et al*. 1999; Taylor *et al*. 2005) from the Early Devonian Rhynie chert (408 – 360 Mya), provided a calibration point for the divergence of the Sordariomycetes (Pyrenomycetes) (Taylor & Berbee 2006; Taylor *et al*. 1999). No fossils representing any of the other ascomycete lineages are currently available.

Divergence times from the *rrl* chronogram were subsequently used to calibrate the ITS+β-tubulin chronogram of the Diaporthales. First, the divergence between the Diaporthales and the Sordariales (*Neurosopora crassa*) was used as an anchor to the tree. Second, the time to the most recent common ancestor (t_{MRCA}) of the Diaporthales was used to calibrate the root of this order.

Data analysis

Congruence of data sets and the evolutionary model – Phylogenetic congruence of ITS and β-tubulin data sets, representing the Diaporthales and including the Cryphonectriaceae, was evaluated using a partition homogeneity test (Cunningham 1997; Farris *et al*. 1995; Huelsenbeck *et al*. 1996) as implemented in PAUP* v. 4·0b10 (Swofford 2002). This was done in order to determine whether the ITS and β-tubulin data sets represent homogenous partitions. Additionally, the best model of nucleotide substitution was determined for all data sets using jModelTest v. 0·1 (Guindon & Gascuel 2003; Posada 2008).

The presence of a constant molecular clock (null hypothesis) in all data sets was investigated by analysing the χ^2 significance of tree likelihood statistics ($2\Delta L$) (Huelsenbeck & Crandall 1997; Huelsenbeck & Rannala 1997). All data sets were subjected to maximum likelihood analyses using PAUP* with and without a constant molecular clock enforced. The log-likelihood of the tree resulting from each analysis was recovered, and the difference in log-

© University of Pretoria

likelihood statistics for the two trees for each data set was used to determine whether a constant molecular clock is present in the data.

Maximum likelihood analyses – Phylogenetic analyses were performed in order to define groups of taxa that would be expected during coalescent analysis, and also to identify preliminary calibration points for the ITS+β-tubulin phylogeny. These initial analyses of the data sets involved exclusion of nucleotide positions that could not be aligned from the *rrl* data set, followed by maximum likelihood estimation using PhyML (Guindon & Gascuel 2003) with a GTR+I+G model of evolution as was determined with jModelTest. Confidence in branches was determined using 1000 bootstrap replicates under the same tree building parameters previously used.

Calibration points as priors – In order to maximise the accuracy of the relaxed molecular clock, calibration points in the *rrl* phylogeny were defined using the age of the representative fossil as the median. The standard deviation (σ) was adjusted so that the 95% confidence of the normal distribution associated with each calibration median included the end points of the period to which it belongs. The maxima of these distributions were represented by fossil ages. We followed this protocol due to the inaccuracy of fossil dating, although previously estimated ages of fossils and phylogenetic divergences usually fall within the same period. Therefore, priors were provided for the currently estimated age of each fossil, but the recovered age after analysis was allowed to slide within the normal distribution of likelihood. In this methodology, the published age of a fossil will have the highest likelihood, and the likelihood will decrease as the recovered age nears the 95% borders of the normal distribution. For the ITS+β-tubulin phylogeny, the medians and standard deviations of secondary calibration points were used as they were recovered from the *rrl* phylogeny.

Coalescent analyses – The *rrl* and ITS+β-tubulin data sets were successively subjected to Markov Chain Monte Carlo simulations incorporating the coalescent process, using BEAST 1·5·1 (Drummond *et al*. 2006; Drummond & Rambaut 2007). An uncorrelated lognormal relaxed molecular clock (Drummond *et al*. 2006) was used to translate substitutions along branches into time periods (Yang 2007). Tree priors consisted of the previously mentioned calibration points, as well as nucleotide frequencies, gamma distribution shapes and proportions of invariant sites as obtained from jModelTest, under a GTR+I+G evolutionary

© University of Pretoria

model. The tree priors were set with the Yule speciation process (Aldous 2001; Drummond *et al*. 2006). A chain length of 1×10^{10} steps, with a burn-in of 1×10^{9} steps, was executed on a high performance computing cluster using the Sun Grid Engine software (Sun Microsystems, Inc.). Analyses were performed in array format, and divided between five computing nodes, each with a unique random seed number. All analyses were repeated in order to mitigate the effect of random seed number on the Markov Chain. The analysis output files were combined and parsed to consensus tree format using LogCombiner and TreeAnnotator supplied with BEAST 1·5·1. Consensus trees were analysed with FigTree 1·1·2 (Drummond & Rambaut 2007) and further annotated using the tree editor in MEGA 4·0 (Tamura *et al*. 2007).

Results

Data mining and data sets

In our *rrl* data set, animals were represented by a chordate (*Mus musculus*), a flatworm (*Bivitellobilharzia nairi*) and a nematode (*Strongyloides stercoralis*), while the fungi were represented by 101 species of Zygomycota, Basidiomycota and Ascomycota (Table 1). Another deep divergence exists between the Ascomycota and Basidiomycota (Padovan *et al*. 2005; Taylor & Berbee 2006). Thus, in the *rrl* data set, the Basidiomycota were represented by 14 species, including four species of the polypore families of Auriscalpiaceae (*Lentinellus ursinus*), Boletaceae (*Boletus satanas*), Cantharellaceae (*Cantharellus tubaeformis*) and Gomphidiaceae (*Chroogomphus vinicolor*). The Ascomycota included 82 species representing three subphyla, namely Taphrinomycotina, Saccharomycotina and Pezizomycotina. The Taphrinomycotina are also known as Archiascomycetes, and they represent the oldest known group of Ascomycota.

In total, 104 taxa were included in the *rrl* data set, resulting in a sequence alignment of 3316 characters after exclusion of ambiguously aligned regions. The ITS data set included 38 taxa (663 characters) while the β-tubulin data set represented 25 taxa (420 characters). The combined ITS+β-tubulin data set represented 38 taxa with an alignment length of 1083 characters (Table 4).

© University of Pretoria

Data analysis

Congruence of data sets and the evolutionary model – A partition homogeneity test of the ITS and β-tubulin data sets revealed that these gene regions are evolutionarily congruent (P = 0·01) (Table 4). Therefore, they were combined in subsequent analyses (Cummings *et al*. 1995). Tests for the best evolutionary model for each of the data sets showed that the GTR+I+G model was most appropriate for both the *rrl* and combined ITS+β-tubulin data sets. The null hypothesis that a constant molecular clock is present, based on 2ΔL values of maximum likelihood trees, could be rejected in all data sets.

Calibrated trees – Coalescent analysis of the *rrl* data set yielded a chronometric tree that was concordant with previous analyses of similar data sets (Figure 1) (Berbee 1996; Berbee & Taylor 1995, 2001; Padovan *et al*. 2005). These calibrations provided rate-smoothing capability across the phylogenetic tree (Corruccini *et al*. 1980; Kasuga *et al*. 2002; Roger & Hug 2006; Takezaki *et al*. 1995). Also, the validity of fossil dates was strengthened by defining calibrations as priors to a normal distribution that includes the boundaries of the time period in which each fossil occurs (Thorne & Kishino 2002; Welch & Bromham 2005). To overcome the problem of rate heterogeneity, we used a nonparametric relaxed clock algorithm coupled with a coalescence algorithm (Aldous 2001; Douzery *et al*. 2004; Drummond *et al*. 2006; Drummond & Rambaut 2007), which applies a unique substitution rate to each lineage. Douzery et al. (2004) showed that an auto-correlated substitution rate model avoids overestimations of deep divergence times when data are not clock-like.

The estimated time to the most recent common ancestor of the Ascomycetes and Basidiomycetes was 712 Mya (Table 5). Estimated divergence times were within the same ranges as previously recovered, with the Sordariomycetes (Pyrenomycetes) at *c*. 410 (419·36–399·81) Mya, and the Diaporthales at *c*. 136 (216·62–67·22) Mya (Figure 1, Table 5). The most recent common ancestor of the Diaporthales and *Neurospora crassa* was at 366 (414·18–294·86) Mya, and this was used as a secondary calibration point in the ITS+β-tubulin analysis.

The calibration points for the *rrl* and ITS+β-tubulin phylogenies are summarised in Table 3. When the Diaporthales were analysed using combined ITS and β-tubulin sequence data, six of the nine families were recovered (Figure 2). The Cryphonectriaceae appeared to have

© University of Pretoria

diverged from their common ancestor *c*. 82·5 (110·25–56·6) Mya, while *Chrysoporthe* spp. and *Cryphonectria* spp. are *c*. 7·1 (13·07–2·34) and 39·1 (57·77–21·72) million years old respectively. Two *Endothia* spp. could not be properly resolved into a single clade, and grouped with the Diaporthaceae and Cryphonectriaceae, respectively. These results are consistent with previous phylogenetic studies of the Diaporthales (Castlebury *et al*. 2002; Rossman *et al*. 2007).

Discussion

The *rrl* chronogram generated in this study is congruent with those produced using other methods (Berbee & Taylor 2001; Padovan *et al*. 2005). The coalescent approach presented here may in fact yield better results than constant clock maximum likelihood methods (cf. Berbee & Taylor 2010; Kuhner *et al*. 1998; Meligkotsidou & Fearnhead 2005; Zhaxybayeva & Gogarten 2004). We could, for example, place the discomycete *Leifidium tenerum* within the expected Lecanoromycetes clade (Wedin 1993) and not together with other Eurotiomycetes, as previously reported (Padovan *et al*. 2005). The chronogram generated from more variable sequences, i.e. ITS and β-tubulin, also revealed the expected fungal groups and allowed estimation of emergence dates that correlated with known geographic and biological information. Therefore, a calibrated tree generated using coupled coalescence and maximum likelihood, i.e. Yule speciation, appears to yield the best estimation of evolution in the Ascomycetes.

The recovered emergence date for the Diaporthales was 143 (162·04–124·76) Mya based on ITS+β-tubulin sequences, at the border between the Late Jurassic and Early Cretaceous Periods of the Mesozoic Era (Figure 2). During this time, the super-continent Pangaea continued to break up into Gondwana and Laurasia, and this led to increased regional differences in climate and the fauna and flora inhabiting these regions (Dimichele *et al*. 2009). The Diaporthales include both general and specific associations with other organisms such as plants, and opportunistic pathogens of humans, *e.g*. *Phaeoacremonium* which causes phaeohyphomycosis (Crous *et al*. 1996). Consequently, this fungal order could not have emerged prior to vascular plants (*c*. 420–410 Mya) (Kedrick & Crane 1997; Steemans *et al*. 2009) because the required climatic and organismal diversity would have been lacking. This dependence of Sordariomycetes on other organisms in ecosystems is elegantly illustrated by the specificity and diversity of *Gondwanamyces* and *Ophiostoma* (species of the sister order

© University of Pretoria

Ophiostomatales), to *Protea* spp. in the Cape Floristic Region (Roets *et al*. 2009). Surprisingly, there is no evidence for allopatric speciation processes in the Diaporthales, although many of the families are restricted to specific climatic regions such as the tropics, or specific groups of hosts such as hardwoods (Gryzenhout *et al*. 2009; Rossman *et al*. 2007).

The Valsaceae (65·6 Mya, 103·01–32·04) and Diaporthaceae (21·9 Mya, 41·21–6·36) had a common ancestor *c*. 114 Mya, during the Early Cretaceous epoch. During this time, Gondwana started to break up into Samafrica and East Gondwana, and the South Atlantic ocean started to appear (Upchurch 2008). Therefore, this event represents the first opportunity for lineages within the Diaporthales to diverge allopatrically. Valsaceae occur on woody angiosperms and occasionally on conifers in temperate regions of the world (Spielman 1985). In contrast, species in the Diaporthaceae are known from an extremely wide range of vascular plants (Rossman *et al*. 2007). Therefore, our results imply that the Valsaceae and Diaporthaceae diverged from their common ancestor due to host preferences, perhaps initially in sympatry, but not allopatrically.

The Schizoparmaceae and Cryphonectriaceae had a common ancestor approximately 94 Mya (125–64), during the Late Cretaceous. Similarly, the Gnomoniaceae and Melanconidaceae diverged from each other *c*. 76 Mya (110–43). Overlapping divergence estimates suggest that lineages of these four families started to take shape during the same time period, hinting at a geological or climatic change that might have prompted it.

Previous studies on the phylogeny of *Chrysoporthe* spp. have shown that four main groups exist in this assemblage (Gryzenhout *et al*. 2004; Myburg *et al*. 2002; Myburg *et al*. 1999; Van der Merwe *et al*. 2010). These can be partitioned into a South American clade representing *C. cubensis*, a Southeast Asian clade (*C. deuterocubensis*), and separate clades representing *C. austroafricana* from Africa and *C. hodgesiana* from South America (Gryzenhout *et al*. 2004; Gryzenhout *et al*. 2006c; Van der Merwe *et al*. 2010). In the present study, these clades were dated and found to all have emerged well after the break-up of Gondwana (Figure 2). Therefore, the continental distributions observed in phylogenies of *Chrysoporthe* are not as a result of geographic vicariance caused by continental drift, and other mechanisms of allopatric or sympatric speciation should therefore be considered for them.

© University of Pretoria

Following previous hypotheses, *C. cubensis sensu stricto* had its origins in South America (Seixas *et al*. 2004), *C. deuterocubensis* originated in Southeast Asia (Gryzenhout *et al*. 2009; Hodges 1980; Van der Merwe *et al*. 2010), and *C. austroafricana* originated in Africa (Gryzenhout *et al*. 2004; Gryzenhout *et al*. 2009). These hypotheses have emerged from the fact that these fungi occur on native hosts in the regions in which they are found (Davison & Coates 1991; Gryzenhout *et al*. 2006b; Hodges 1980; Seixas *et al*. 2004; Wingfield *et al*. 2001), and the high level of genetic diversity observed in the fungal populations (Conradie *et al*. 1992; Van Heerden *et al*. 1997; Van Zyl *et al*. 1994). Our results indicate that Asian *C. deuterocubensis* (*c*. 5·3 Mya) is older than the South American *C. cubensis* (*c*. 1·39 Mya). Therefore, we believe that the stem lineage of *Chrysoporthe* must have migrated around the southern hemisphere, followed by allopatric speciation that was driven by factors other than continental drift. Colonisation of South America, Africa and Southeast Asia by this stem lineage could have been favoured by the presence of Myrtalean hosts on those continents. However, the difference in species composition of these plants must have driven the evolution of *Chrysoporthe* species, resulting in reproductive isolation from the founder population and subsequent species replacement (Ribeiro & Caticha 2008). This hypothesis agrees with the phylogeography of the Myrtales, which is marred with geographic vicariance and post-vicariant long-distance dispersal around the Southern Hemisphere (Sytsma *et al*. 2004).

Several important geological phenomena could have had an influence on the evolution of the Diaporthales, and specifically the Cryphonectriaceae. For instance, during the past 110 million years the earth's axis tilted through 20°, with a single period of shifting between 150–80 Mya (Prévot *et al*. 2000). Such changes could gradually shift floristic distributions on land masses, and would have a large impact on the evolution of fungi associated with flora that are bound by climatic regions. Interestingly, this is exactly the period during which the Diaporthales emerged (Figure 2, Table 5). During this time, which falls within the Jurassic (175–140 Mya) and Early Cretaceous (145·5–112 Mya), the continent of Gondwana started to divide into Samafrica (South America and Africa) and East Gondwana (Dalziel *et al*. 2000; Upchurch 2008). The first Myrtalean lineages appeared in Africa during the Upper Cretaceous, *i.e*. 100–60 Mya. Towards the end of the Gondwanan break-up, between 80 and 30 Mya, the first Myrtaceae appeared in Australia (Sytsma *et al*. 2004). Stem lineages of the

© University of Pretoria

Diaporthales were already present when the Atlantic Ocean started to form *c*. 135 Mya (Dalziel *et al*. 2000). This would have presented an opportunity for allopatric speciation in the Diaporthales, but such processes are not obvious in the evolution of this group of fungi.

When the evolution of the genus *Chrysoporthe* is considered (7·07 Mya, 13·07–2·34), it is noteworthy that by the time its most recent ancestors appeared, the continents were close to where they are presently (Dalziel *et al*. 2000). Also during this time, around 4–3 Mya, the Indonesian seaway started to close and that event brought about a period of aridification in eastern Africa (Cane & Molnar 2001). This was also a period of deglaciation of Antarctica (3 Mya) (Barrett *et al*. 1992), which brought about a rise in sea levels. These changes drastically influenced the circulation of oceanic currents (Cane & Molnar 2001; Peltier 1988; Philander & Fedorov 2003), which fuelled the gradual increase in temperatures over the Southern Hemisphere. Conceivably, the stem lineage of *Chrysoporthe* could have migrated on fleshy Myrtalean fruits (Sytsma *et al*. 2004) via these seaways around the Southern Hemisphere. Importantly, even though the phylogeography of the Myrtales and *Chrysoporthe* spp. appear to be parallel, dispersal in their evolutionary histories are temporally separated. Evolution in *Chrysoporthe* is, therefore, not a direct consequence of Myrtalean phylogeography.

Prior to 7 Mya, the stem lineage of *Chrysoporthe* must have emerged. We believe that this fungus evolved on Myrtaceae in Australia, because that is where the first Myrtaceae appeared 80–30 Mya (Sytsma *et al*. 2004). Around 7 Mya, and probably due to changes in sea currents and climate change over the Southern Hemisphere (Barrett *et al*. 1992; Cane & Molnar 2001; Cerling *et al*. 1997; Peltier 1988; Philander & Fedorov 2003), Myrtalean plants spread around the world, probably carrying the stem lineage of *Chrysoporthe* along with them. The stem lineage would, therefore, have been carried to Africa and South America, where it would have encountered native Myrtalean species. This would subsequently have been followed by allopatric speciation, resulting in *C. cubensis* in South America, *C. austroafricana* in Africa and *C. deuterocubensis* in Southeast Asia (Figure 3). Lastly, between 1·39 Mya and present, *C. cubensis* and *C. deuterocubensis* were introduced from South America and Southeast Asia, respectively, into Africa. This event must have occurred over water, and probably anthropogenically during trade, since no land bridges were present between the continents.

© University of Pretoria

In this study we have shown that neither the Diaporthales nor any of the families therein, including the Cryphonectriaceae, display evidence of allopatric speciation due to continental drift. Rather, allopatric speciation in this group of fungi was shaped by a confluence of factors including host preference, the ability of hosts to migrate around the world, geological change, and climate change. Results showed that the most recent common ancestor of *Chrysoporthe* existed *c*. 13·07–2·34 Mya. This estimate is well after the break-up of Gondwana between 175 Mya and 30 Mya (McLoughlin 2001; Sanmartin & Ronquist 2004), implying that *Chrysoporthe* spp. did not have Gondwanan origins. However, allopatric speciation in the Myrtales, *i.e*. the preferred hosts of *Chrysoporthe* spp., and their spread around the world provide convincing evidence that the continental species of *Chrysoporthe* have diverged allopatrically. Future studies should focus on investigating selective pressures other than geography that could aid in shaping the evolutionary trajectory of fungi in the Diaporthales.

© University of Pretoria

Table 1 Animal and fungal taxa used for construction of the 18S phylogeny, and the GenBank accession numbers for their *irl* (18S) ribosomal DNA sequences.

Taxon	18S Accession No.
Fungi	
Ascomycota	
Pezizomycotina: Dothideomycetes	
Botryosphaeria ribis **(Loculomycete)**	U42477
Dothidea insculpta **(Loculomycete)**	U42474
Jahnula siamensiae **(Loculomycete)**	AF438180
Mycosphaerella mycopappi **(Loculomycete)**	U43449
Myriangium duriaei **(Loculomycete)**	AY016347
Pyrenophora trichostoma **(Loculomycete)**	U43459
Pezizomycotina: Eurotiomycetes (Plectomycetes)	
Aspergillus fumigatus	M60300
Catapyrenium lachneum **(Loculomycete)**	AF412410
Coccidioides immitis	M55627
Eremascus albus	M83258
Exophiala (*Phaeococcomyces*) *exophialae* **(Loculomycete)**	X80709
Penicillium chrysogenum	M55628
Pezizomycotina: Lecanoromycetes (Discomycetes)	
Anamylopsora pulcherrima	AF119501
Cyphelium inquinans	U86695
Diploschistes rampoddensis	AF274111
Graphis scripta	AF038878
Lasallia rossica	AF088238
Lecanora dispersa	L37535
Leifidium tenerum	U70959
Pertusaria saximonticola (*P. saximontana*)	AF113720
Pezizomycotina: Leotiomycetes (Discomycetes)	
Blumeria graminis f. sp. *hordei*	AB033480
Cyttaria darwinii	U53369
Graphium rubrum	AB278753
Sclerotinia sclerotiorum	X69850
Thelebolus stercoreus	U49936
Pezizomycotina: Pezizomycetes (Discomycetes)	
Microstoma floccosum	AF006313
Morchella elata	L37537
Otidea onotica	AF006308
Pithya cupressina	AF006316
Tuber gibbosum	U42663
Pezizomycotina: Sordariomycetes (Pyrenomycetes)	
Amphisphaeria umbrina	AF225207
Ascovaginospora stellipala	U85087
Ceratocystis fimbriata	U43777
Chaetomium elatum	M83257
Colletotrichum gloeosporioides	M55640
Cornuvesica falcata	AY271797
Halosarpheia spartinae	AF352076
Hypocrea lutea	D14407
Hypomyces chrysospermus	M89993
Kionochaeta spissa	AB003790
Microascus cirrosus	M89994
Nais inornata	AF050482
Neurospora crassa	X04971
Ophiostoma piliferum	AJ243295
Ophiostoma ulmi	M83261
Podospora anserina	X54864

© University of Pretoria

Pseudallescheria boydii	U43914
Verticillium dahliae	U33637
Xylaria carpophila	Z49785
Pezizomycotina: Sordariomycetes: Diaporthales	
Amphilogia gyrosa	FJ176835
Amphiporthe castanea	AF277117
Apiosporopsis carpinea	AF277110
Apognomonia supraseptata	AF277118
Chrysoporthe cubensis	DQ862047
Cryphonectria havanensis	L42440
Cryphonectria parasitica	L42441
Cryphonectria radicalis	L42442
Cryptodiaporthe corni	AF277119
Diaporthe eres	DQ471015
Diaporthe phaseolorum	L36985
Discula campestris	AF277107
Discula destructiva	AF429719
Discula fraxinea (*Gnomoniella fraxini*)	AF277106
Discula quercina	AF277108
Endothia gyrosa	DQ471023
Gnomonia setacea	AF277121
Leucostoma persoonii	M83259
Melanconis alni	DQ862052
Melanconis marginalis	AF277122
Melanconis stilbostoma	DQ862054
Pilidiella eucalyptorum	EU754051
Plagiostoma euphorbiae	DQ862055
Prosopidicola mexicana	AY720717
Valsa ambiens	DQ862065
Valsa ambiens subsp. *leucostomoides*	AF277120
Valsella salicis	DQ862057
Saccharomycotina	
Citeromyces matritensis	AB018176
Dekkera anomala	X83828
Dipodascopsis uninucleata	U00969
Galactomyces geotrichum	U00974
Saccharomyces cerevisiae	J01353
Zygosaccharomyces cidri	X91085
Taphrinomycotina (Archiascomycetes)	
Pneumocystis carinii	X12708
Protomyces macrosporus	D85143
Saitoella complicata	D12530
Taphrina deformans	U00971
Basidiomycota	
Boletus satanas	M94337
Cantharellus tubaeformis	AF026636
Chroogomphus vinicolor	M90822
Coprinus cinereus	M92991
Cronartium ribicola	M94338
Cryptococcus podzolicus	AB032645
Exobasidium rostrupii	AJ271380
Filobasidiella neoformans	X60183
Lentinellus ursinus	U59081
Leucosporidium scottii	X53499
Russula compacta	U59093
Suillus cavipes	M90828
Tilletia caries	U00972
Ustilago hordei	U00973
Glomeromycota	
Gigaspora rosea	X58726

© University of Pretoria

Glomus intraradices	X58725
Glomus mosseae	Z14007
Zygomycota	
Mucor mucedo	X89434
Mucor racemosus	X54836
Animalia	
Metazoa: Coelomata (Chordata)	
Mus musculus	BK000964
Metazoa: Acoelomata (Trematoda)	
Bivitellobilharzia nairi	AY829261
Metazoa: Pseudocoelomata (Nematoda)	
Strongyloides stercoralis	M84229

© University of Pretoria

Table 2 Fungal species and isolates used in estimation of divergence times within the Diaporthales, and the Genbank accession numbers for their internal transcribed spacer regions and β-tubulin (exons 4−6) gene regions.

Fungal Species	Culture Number	ITS	β-tubulin 2
Diaporthales: Cryphonectriaceae			
Amphilogia gyrosa	CMW 10471	AF452116	AF525716
Aurapex penicillata	CMW 10030	AY214311	AY214275
Celoporthe dispersa	CMW 9976	DQ267130	DQ267142
Chrysoporthe austroafricana	CMW 2113	AF046892	AF273462
Chrysoporthe cubensis	CMW 14394	DQ368773	DQ368799
Chrysoporthe deuterocubensis	CMW 8651	AY084002	AY084014
Chrysoporthe doradensis	CMW 11286	AY214290	AY214254
Chrysoporthe inopina	CMW 12727	DQ368777	AH015657
Chrysoporthella hodgesiana	CMW 10641	AY692322	AY692325
Cryphonectria macrospora	CMW 10463	AF368331	AH011608
Cryphonectria nitschkei	CMW 10518	AF452118	AF525713
Cryphonectria parasitica	CMW 1652	AF046902	AF273468
Cryphonectria radicalis	CMW 13754	AY697932	AH014584
Cryptodiaporthe corni	AR 2814	DQ120762	AH015163
Endothia gyrosa	CMW 2091	AF046905	AF543822
Endothia singularis	CMW 10465	AF368323	AH011599
Holocryphia eucalypti	CMW 7033	DQ368727	AH015661
Microthia havanensis	CMW 11301	AY214323	AY214287
Rostraureum tropicale	CMW 9971	AY167435	AY167430
Ursicollum fallax	CMW 18124	DQ368757	AH015660
Diaporthales: Diaporthaceae			
Diaporthe ambigua	CMW 5587	AF543818	AF543822
Diaporthe eres	CBS 345.94	AY853215	—
Diaporthales: Gnomoniaceae			
Cryptodiaporthe aesculi	AR 3640	DQ313557	—
Ditopella ditopa	CBS 109748	DQ323526	—
Gnomonia gnomon	CBS 199.53	AY818956	—
Phragmoporthe conformis	CBS 109783	DQ323527	—
Plagiostoma euphorbiae	CBS 340.78	DQ323532	—
Diaporthales: *Harknessia* Complex			
Harknessia eucalypti	CBS 342.97	AY720745	AY720777
Harknessia eucalyptorum	CBS 113620	AY720746	AY720778
Wuestneia molokaiensis	CBS 114877	AY720749	—
Diaporthales: Melanconidaceae			
Melanconis stilbostoma	CBS 109778	DQ323524	—
Diaporthales: Valsaceae			
Valsa ambiens	CBS 191.42	AY347330	—
Valsa eugeniae	CMW 8648	AY347344	—
Valsa viridistroma ①	CMW 10454	AF452120	AF525711
Diaporthales: *Schizoparmaceae*			
Schizoparme straminea	STE-U 3932	AY339348	—
Pilidiella eucalyptorum	STE-U 3327	AY339314	—
Hypocreales: Hypocreaceae			
Hypomyces chrysospermus	SWA3	AY344799	—
Sordariales: Sordariaceae			
Neurospora crassa	ICMP 6360	AY681193	AY681226

① Previously known as *Endothia viridistroma*, but recently transferred to *Valsa* (Adams *et al.* 2006).

© University of Pretoria

Table 3 Summary of calibration points for the *rrl* and ITS+β-tubulin phylogeny.

Taxon / Designation	Epoch	Period	Calibration (Ma)	Median (Ma)	σ ①
		rrl			
Tree Height ②		Cryogenian / Tonian	965	965	10
Glomales		Ordovician	460	460	20
Sordariomycetes ③	Early	Devonian	416–407	411·5	9
Basidiomycota ④	Cisuralian	Permian	290	290	10
Polypores ⑤	Early	Cretaceous	145·5–112	128·5	17
		ITS+β-tubulin[f]			
Diaporthales/Sordariales divergence	Early	Carboniferous	295 – 414	366	25
Diaporthales (tMRCA)	Early	Cretaceous	67 – 217	136	30

① Standard Deviation

② Animal-fungal divergence.

③ Sordariomycete fossils showing flask-shaped and ostiolate perithecia (Taylor *et al*. 2005).

④ Fossilized clamp connections (Dennis 1969, 1970, 1976; Osborn *et al*. 1989).

⑤ Polypore fossils from the Early Cretaceous (Smith *et al*. 2004).

⑥ Calibration points for the ITS+β-tubulin tree were recovered from the *rrl* chronogram.

© University of Pretoria

Table 4 Tree statistics gathered from maximum likelihood analyses of both the *rrl* and ITS+β-tubulin data sets. The ITS and β-tubulin data sets were analyzed separately, but also as a combined data matrix.

Statistic	***rrl***	**ITS**	**β-tubulin**	**Combined ITS+β-tubulin**
Best evolutionary model	GTR+G+I	GTR+G+I	TVM+G+I	GTR+G+I
H_0: Constant molecular clock ①	Rejected	Rejected	Rejected	Rejected
Confidence in congruence (*P*-value) ②	–	–	–	0·01
Total included characters	3316	663	420	1083
Maximum Likelihood:				
Number of trees (no clock)	1	1	1	5
Number of trees (constant clock)	2	19	3	3

① The null hypothesis that a constant molecular clock is present will be rejected when the Δ log likelihood between an analysis without a clock and one where the constant clock is present does not follow a χ^2 distribution (Gaunt & Miles 2002; Yang 2007).

② A partition homogeneity test was used to determine the probability of congruence between data sets when they are combined.

© University of Pretoria

Table 5 Recovered divergence times and confidence intervals of important fungal groups, based on *rrl* and ITS+β-tubulin chronometric analyses.

Taxon	Median (t_{MRCA}) ①	95% Confidence (Ma)
	***rrl* data set**	
Animals / Fungi	952·53	885·99 – 1022·48
Basidiomycota / Ascomycota	712·20	536·77 – 918·93
Basidiomycota	295·74	276·51 – 314·78
Ascomycota	632·06	464·29 – 832·89
Pezizomycotina	490·79	420·36 – 591·24
Sordariomycetes	409·72	399·81 – 419·36
Pezizomycetes	283·62	132·82 – 480·17
Sordariales / Diaporthales ②	365·58	294·86 – 414·18
Hypocreales	248·66	148·35 – 349·59
Xylariales	221·88	77·11 – 386·63
Sordariales	206·22	107·81 – 317·44
Microascales	159·57	83·42 – 243·24
Phyllachorales	144·75	39·28 – 241·60
Diaporthales	135·78	67·22 – 216·62
Ophiostomatales	18·99	0·82 – 51·41
Leotiomycetes	229·11	85·11 – 425·61
Lecanoromycetes	225·01	122·70 – 342·26
Dothideomycetes	215·28	96·50 – 349·40
Eurotiomycetes	211·64	109·88 – 329·40
Saccharomycotina	407·69	241·79 – 582·78
Taphrinomycotina	355·15	137·56 – 591·92
	ITS+β-tubulin data set	
Sordariales / Diaporthales ②	361·84	342·36 – 381·56
Diaporthales	143·31	124·76 – 162·04
Cryphonectriaceae	82·50	56·60 – 110·25
Melanconidaceae	75·91	43·3 – 110·72
Valsaceae	65·60	32·04 – 103·01
Gnomoniaceae	44·13	21·29 – 69·89
Cryphonectria spp.	39·07	21·72 – 57·77
Schizoparmaceae	36·44	9·00 – 70·18
Diaporthaceae	21·87	6·36 – 41·21
Harknessia complex	21·2	6·01 – 39·73
Chrysoporthe spp.	7·07	2·34 – 13·07
C. austroafricana / *C. cubensis*	1·39	0·07 – 3·38

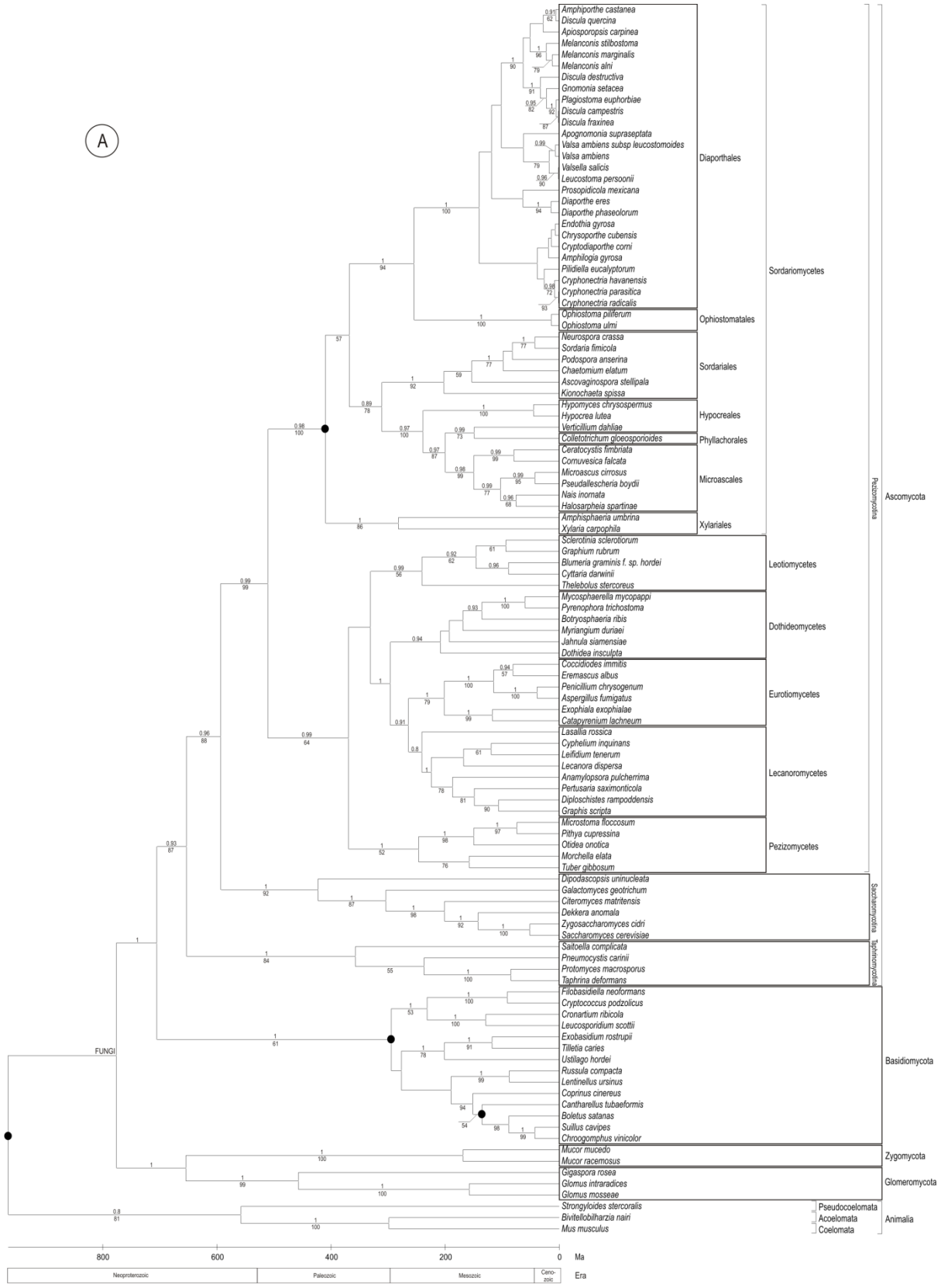
① The t_{MRCA} (time to most recent common ancestor) is given in millions of years ago. The median is considered a more accurate measure than the mean, due to the influence of skewed normal distributions on mean values.

② The divergence time of the Sordariales (*Neurospora crassa*) and Diaporthales was used as a secondary calibration point in the ITS+β-tubulin chronogram.

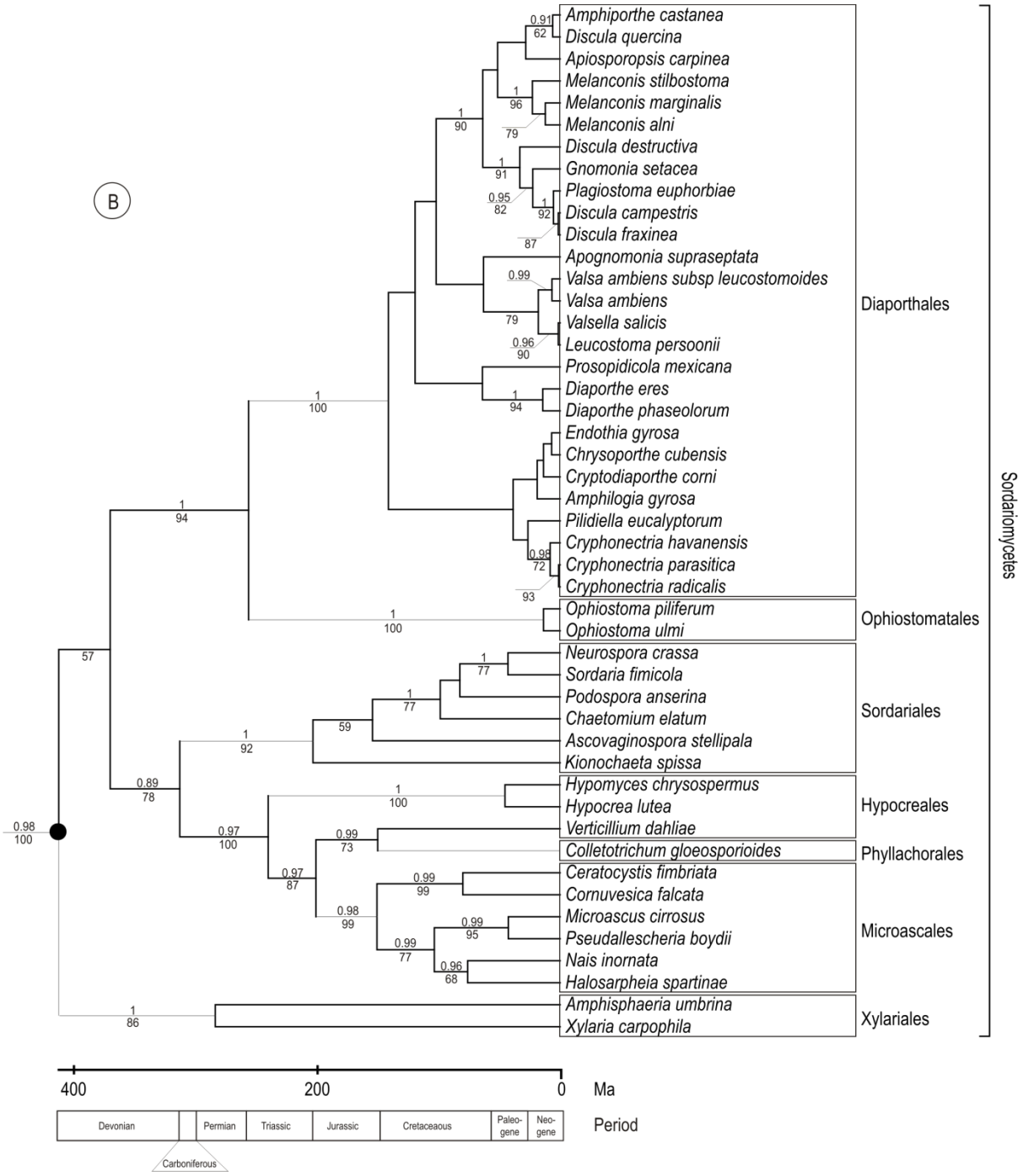
© University of Pretoria

Figure 1 (A) Ultrametric chronogram produced using *rrl* sequence data from selected taxa. Calibration points are indicated with filled circles. (B) Section of the *rrl* chronogram depicting the Sordariomycetes. The calibration point for *Paleopyrenomycites devonicus* is indicated by a filled circle. Divergences used for secondary calibration of the ITS+β-tubulin chronogram are indicated with open circles. In both (A) and (B), posterior probabilities above 0·8 are displayed above branches, while bootstrap confidence above 50% (maximum likelihood, 1000 repetitions) is displayed below.

© University of Pretoria



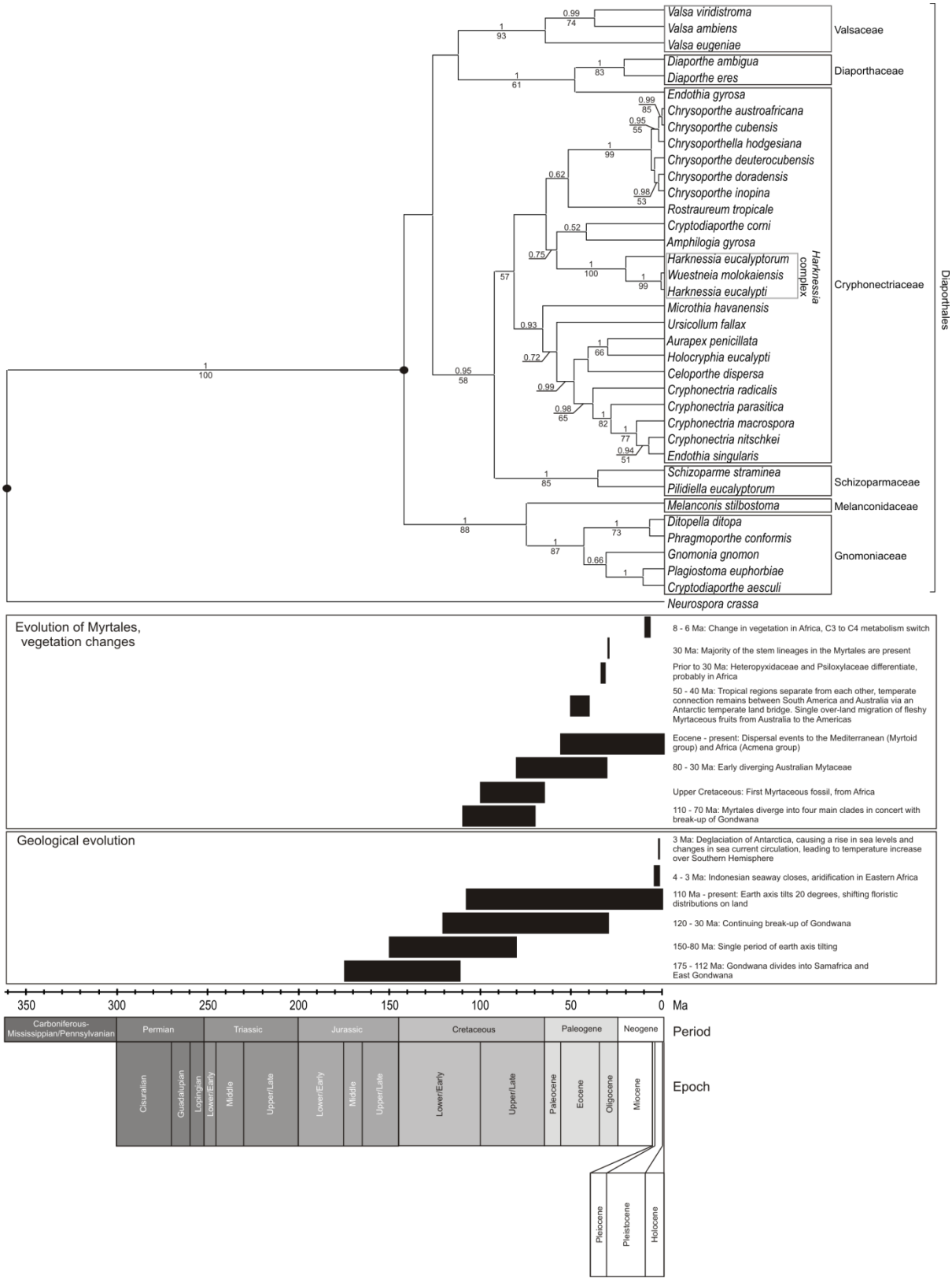
© University of Pretoria



© University of Pretoria

Figure 2 Chronogram of the ITS+β-tubulin data set representing the Diaporthales. Secondary calibration points are marked with filled circles. Posterior probabilities above 0·5 are indicated above branches while maximum likelihood bootstrap confidence (1000 repetitions) above 50% are indicated below branches.

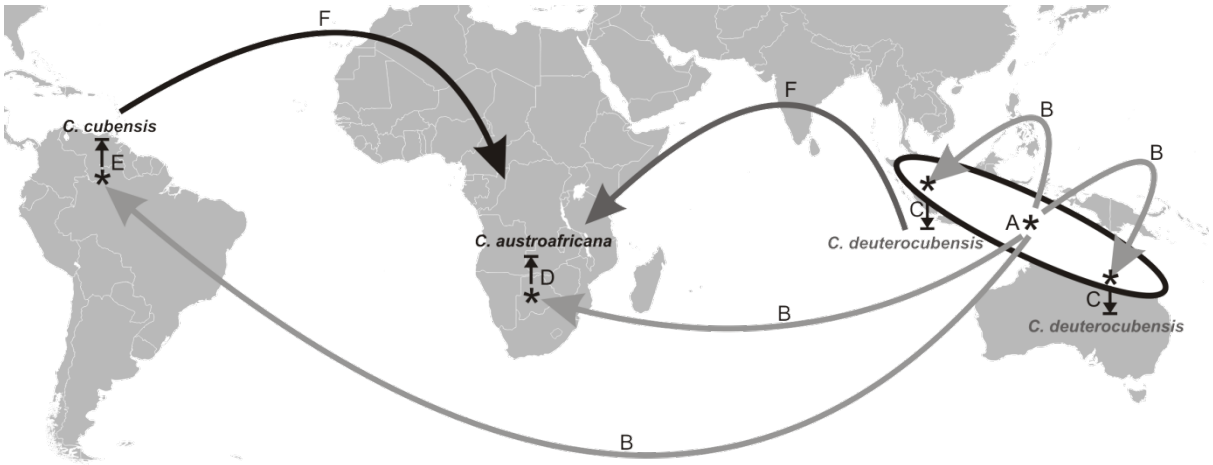
© University of Pretoria



© University of Pretoria

Figure 3 Evolution and trans-oceanic dispersal of *Chrysoporthe* species around the world. (A) Prior to 7 Ma, the stem lineage (asterisk) of *Chrysoporthe* may have evolved in Southeast Asia, probably on a Myrtaceous host in Australia (Sytsma *et al.* 2004). (B) The stem lineage dispersed throughout Southeast Asia and to Africa and South America, *c.* 7 Ma. (C, D, E) Allopatric speciation took place in Southeast Asia, Africa and South America, resulting in *Chrysoporthe deuterocubensis*, *C. austroafricana* and *C. cubensis*, respectively. (F) Trans-oceanic dispersal and introduction of *C. cubensis* and *C. deuterocubensis* to the African continent (between 1·39 Ma and present), probably on Myrtalean hosts.

© University of Pretoria



© University of Pretoria

Bibliography

Abbo S, Frenkel O, Sherman A, Shtienberg D (2007) The sympatric Ascochyta pathosystems of Near Eastern legumes, a key for better understanding of pathogen biology. *European Journal of Plant Pathology* **119**, 111-118.

Adams GC, Roux J, Wingfield MJ (2006) *Cytospora* species (*Ascomycota*, *Diaporthales*, *Valsaceae*): introduced and native pathogens of tree in South Africa. *Australasian Plant Pathology* **35**, 521-548.

Aldous DJ (2001) Stochastical models and descriptive statistics for phylogenetic trees, from Yule to today. *Statistical Science* **16**, 23-34.

Anagnostakis SL (1987) Chestnut blight: the classical problem of an introduced pathogen. *Mycologia* **79**, 23-37.

Anderson PK, Cunningham AA, Patel NG, Morales FJ, Epstein PR, Daszak P (2004) Emerging infectious diseases of plants: pathogen pollution, climate change and agrotechnology drivers. *TRENDS in Ecology and Evolution* **19**, 535-544.

Barrett PJ, Adams CJ, McIntosh WC, Swisher CC, Wilson GS (1992) Geochronological evidence supporting Antarctic deglaciation three million years ago. *Nature* **359**, 816-818.

Berbee ML (1996) Loculoascomycete origins and evolution of filamentous ascomycete morphology based on 18S rRNA gene sequence data. *Molecular Biology and Evolution* **13**, 462-470.

Berbee ML, Taylor JW (1995) From 18S ribosomal sequence data to evolution of morphology among the fungi. *Canadian Journal of Botany* **73**, S677-S683.

Berbee ML, Taylor JW (2001) Fungal molecular evolution: gene trees and geologic time. In: *The Mycota VII Part B: Systematics and Evolution* (eds. McLaughlin DJ, McLaughlin EG, Lemke PA). Springer-Verlag, Berlin.

Berbee ML, Taylor JW (2010) Dating the molecular clock in fungi – how close are we? *Fungal Biology Reviews* **24**, 1-16.

Cane MA, Molnar P (2001) Closing of the Indonesian seaway as a precursor to east African aridification around 3-4 million years ago. *Nature* **411**, 157-162.

Castlebury LA, Rossman AY, Jaklitsch WJ, Vasilyeva LN (2002) A preliminary overview of the Diaporthales based on large subunit nuclear ribosomal DNA sequences. *Mycologia* **94**, 1017-1031.

© University of Pretoria

Cerling TE, Harris JM, MacFadden BJ, Leakey MG, Quade J, Eisenmann V, Ehleringer JR (1997) Global vegetation change through the Miocene/Pliocene boundary. *Nature* **389**, 153-158.

Coetzee MPA, Wingfield BD, Harrington TC, Steimel J, Coutinho TA, Wingfield MJ (2001) The root rot fungus *Armillaria mellea* introduced into South Africa by early Dutch settlers. *Molecular Ecology* **10**, 387-396.

Conradie E, Swart WJ, Wingfield MJ (1992) Susceptibility of *Eucalyptus grandis* to *Cryphonectria cubensis*. *European Journal of Forest Pathology* **22**, 312-315.

Corruccini RS, Baba M, Goodman M, Ciochon RL, Cronin JE (1980) Non-linear macromolecular evolution and the molecular clock. *Evolution* **34**, 1216-1219.

Crous PW, Gams W, Wingfield MJ, Van Wyk PS (1996) *Phaeoacremonium* gen. nov. associated with wilt and decline diseases of woody hosts and human infections. *Mycologia* **88**, 786-796.

Cummings MP, Otto SP, Wakeley J (1995) Sampling properties of DNA sequence data in phylogenetic analysis. *Molecular Biology and Evolution* **12**, 814-822.

Cunningham CW (1997) Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution* **14**, 733-740.

Dalziel IWD, Lawver LA, Murphy JB (2000) Plumes, orogenesis, and supercontinental fragmentation. *Earth and Planetary Science Letters* **178**, 1-11.

Davison EM, Coates DJ (1991) Identification of *Cryphonectria cubensis* and *Endothia gyrosa* from eucalypts in Western Australia using isozyme analysis. *Australasian Plant Pathology* **20**, 157-160.

Dennis RL (1969) Fossil mycelium with clamp connections from the Middle Pennsylvanian. *Science* **163**, 670-671.

Dennis RL (1970) A Middle Pennsylvanian basidiomycete mycelium with clamp connections. *Mycologia* **62**, 578-584.

Dennis RL (1976) *Palaeosclerotium*, a Pennsylvanian Age fungus combining features of modern ascomycetes and basidiomycetes. *Science* **192**, 66-68.

Dettman JR, Jacobson DJ, Taylor JW (2003) A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. *Evolution* **57**, 2703-2720.

Dimichele WA, Montanez IP, Poulsen CJ, Tabor NJ (2009) Climate and vegetational regime shifts in the late Paleozoic ice age earth. *Geobiology* **7**, 200-226.

© University of Pretoria

Doolittle RF, Feng D-F, Tsang S, Cho G, Little E (1996) Determining divergence times of the major kingdoms of living organisms with a protein clock. *Science* **271**, 470-477.

Douzery EJP, Snell EA, Bapteste E, Delsuc F, Philippe H (2004) The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proceedings of the National Academy of Science, USA* **101**, 15386-15391.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**, e88.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797.

Farris JS, Källersjo M, Kluge AG, Bult C (1995) Testing significance of incongruence. *Cladistics* **10**, 315-319.

Galtier N, Gouy M, Gautier C (1996) SeaView and Phylo_Win: two graphic tools for sequence alignment and molecular phylogeny. *Computational and Applied BioScience* **12**, 543-548.

Gaunt MW, Miles MA (2002) An insect molecular clock dates the origin of the insects and accords with palaeontological and biogeographic landmarks. *Molecular Biology and Evolution* **19**, 748-761.

Giraud T (2006) Selection against migrant pathogens: the immigrant inviability barrier in pathogens. *Heredity* **97**, 316-318.

Giraud T, Refrégier G, Le Gac M, De Vienne DM, Hood ME (2008) Speciation in fungi. *Fungal Genetics and Biology* **45**, 791-802.

Giraud T, Villaréal L, Austerlitz F, Le Gac M, Lavigne C (2006) Importance of the life cycle in host race formation and sympatric speciation in parasites. *Phytopathology* **96**, 280-287.

Glass NL, Donaldson GC (1995) Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Applied and Environmental Microbiology* **61**, 1323-1330.

Gryzenhout M, Myburg H, Van der Merwe NA, Wingfield BD, Wingfield MJ (2004) *Chrysoporthe*, a new genus to accommodate *Cryphonectria cubensis*. *Studies In Mycology* **50**, 119-142.

Gryzenhout M, Myburg H, Wingfield BD, Montenegro F, Wingfield MJ (2005a) *Chrysoporthe doradensis* sp. nov. pathogenic to *Eucalyptus* in Ecuador. *Fungal Diversity* **20**, 39-57.

© University of Pretoria

Gryzenhout M, Myburg H, Wingfield BD, Montenegro F, Wingfield MJ (2005b) *Rostraureum tropicale* gen. sp. nov. (*Diaporthales*) associated with dying *Terminalia ivorensis* in Ecuador. *Mycological Research* **109**, 1029-1044.

Gryzenhout M, Myburg H, Wingfield BD, Wingfield MJ (2006a) Cryphonectriaceae (Diaporthales), a new family including *Cryphonectria*, *Chrysoporthe*, *Endothia* and allied genera. *Mycologia* **98**, 239-249.

Gryzenhout M, Rodas CA, Portales JM, Clegg P, Wingfield BD, Wingfield MJ (2006b) Novel hosts of the *Eucalyptus* canker pathogen *Chrysoporthe cubensis* and a new *Chrysoporthe* species from Colombia. *Mycological Research* **110**, 833-845.

Gryzenhout M, Wingfield BD, Wingfield MJ (2006c) New taxonomic concepts for the important forest pathogen *Cryphonectria parasitica* and related fungi. *FEMS Microbiology Letters* **258**, 161-172.

Gryzenhout M, Wingfield BD, Wingfield MJ (2009) *Taxonomy, phylogeny, and ecology of bark-infecting and tree killing fungi in the Cryphonectriaceae*. APS Press.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate phylogenies by maximum likelihood. *Systematic Biology* **52**, 696-704.

Heath RN, Gryzenhout M, Roux J, Wingfield MJ (2006) Discovery of the *Cryphonectria* canker pathogen on native *Syzygium* species in South Africa. *Plant Disease* **90**, 433-438.

Hodges CS (1980) The taxonomy of *Diaporthe cubensis*. *Mycologia* **72**, 542-548.

Hodges CS, Geary TF, Cordell CE (1979) The occurrence of *Diaporthe cubensis* on *Eucalyptus* in Florida, Hawaii, and Puerto Rico. *Plant Disease Reporter* **63**, 216-220.

Hodges CS, Reis MS (1974) Identification do fungo causador de cancro de *Eucalyptus* spp. no Brazil. *Brazil Florestal* **5**, 19.

Hodges CS, Reis MS, Ferreira FA, Henfling JDM (1976) O cancro do eucalipto causado por *Diaporthe cubensis*. *Fitopatalogia Basileira* **1**, 129-167.

Huelsenbeck JP, Bull JJ, Cunningham CW (1996) Combining data in phylogenetic analysis. *TRENDS in Ecology and Evolution* **11**, 152-158.

Huelsenbeck JP, Crandall KA (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecological Systematics* **28**, 437-466.

Huelsenbeck JP, Rannala B (1997) Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* **276**, 227-232.

© University of Pretoria

Johannesson K (2001) Parallel speciation: a key to sympatric divergence. *TRENDS in Ecology and Evolution* **16**, 148-153.

Kasuga T, White TJ, Taylor JW (2002) Estimation of nucleotide substitution rates in eurotiomycete fungi. *Molecular Biology and Evolution* **19**, 2318-2324.

Kedrick P, Crane PR (1997) The origin and early evolution of plants on land. *Nature* **389**, 33-39.

Koufopanou V, Burt A, Taylor JW (1997) Concordance of gene genealogies reveals reproductive isolation in the pathogenic fungus *Coccidioides immitis*. *Proceedings of the National Academy of Science, USA* **94**, 5478-5482.

Kuhner MK, Yamato J, Felsenstein J (1998) Maximum likelihood estimation of population growth rates based on the coalescent. *Genetics* **149**, 429-434.

Lynch M (1999) The age and relationships of the major animal phyla. *Evolution* **53**, 319-325.

McLoughlin S (2001) The breakup history of Gondwana and its impact on pre-Cenozoic floristic provincialism. *Australian Journal of Botany* **49**, 271-300.

Meligkotsidou L, Fearnhead P (2005) Maximum likelihood estimation of coalescence times in genealogical trees. *Genetics* **171**, 2073-2084.

Merkel HW (1905) A deadly fungus on the American chestnut. In: *New York Zoological Society*, pp. 97-103.

Mori Y, Sato Y, Takamatsu S (2000) Molecular phylogeny and radiation time of Erysiphales inferred from the nuclear ribosomal DNA sequences. *Mycoscience* **41**, 437-447.

Myburg H, Gryzenhout M, Wingfield BD, Stipes RJ, Wingfield MJ (2004) Phylogenetic relationships of *Cryphonectria* and *Endothia* species, based on DNA sequence data and morphology. *Mycologia* **96**, 990-1001.

Myburg H, Gryzenhout M, Wingfield BD, Wingfield MJ (2002) ß-tubulin and histone *H3* gene sequences distinguish between *Cryphonectria cubensis* from South Africa, Asia, and South America. *Canadian Journal of Botany* **80**, 590-596.

Myburg H, Wingfield BD, Wingfield MJ (1999) Phylogeny of *Cryphonectria cubensis* and allied species inferred from DNA analysis. *Mycologia* **91**, 243-250.

Nakabonge G, Gryzenhout M, Wingfield BD, Wingfield MJ, Roux J (2007) Genetic diversity of *Chrysoporthe cubensis* in eastern and southern Africa. *South African Journal of Science* **103**, 261-264.

Nakabonge G, Roux J, Gryzenhout M, Wingfield MJ (2006) Distribution of Chrysoporthe canker pathogens on *Eucalyptus* and *Syzygium* spp. in eastern and southern Africa. *Plant Disease* **90**, 734-740.

© University of Pretoria

O'Donnell K, Kistler HC, Tacke BK, Casper HH (2000) Gene genealogies reveal global phylogeographic structure and reproductive isolation among lineages of *Fusarium graminearum*, the fungus causing wheat scab. *Proceedings of the National Academy of Science, USA* **97**, 7905-7910.

O'Donnell K, Ward TJ, Geiser DM, Kistler HC, Aoki T (2004) Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genetics and Biology* **41**, 600-623.

Osborn JM, Taylor TN, White JF (1989) *Paleofibulus* gen. nov., a clamp-bearing fungus from the Triassic of Antarctica. *Mycologia* **81**, 622-626.

Padovan ACB, Sanson GFO, Brunstein A, Briones MRS (2005) Fungi evolution revisited: application of the penalized likelihood method to a Bayesian fungal phylogeny provides a new perspective on phylogenetic relationships and divergence dates of Ascomycota groups. *Journal of Molecular Evolution* **60**, 726-735.

Peltier WR (1988) Lithospheric thickness, antarctic deglaciation history, and ocean basin discretization effects in a global model of postglacial sea level change: a summary of some sources of nonuniqueness. *Quarternary Research* **29**, 93-112.

Philander SG, Fedorov AV (2003) Role of tropics in changing the response to Milankovich forcing some three million years ago. *Paleoceanography* **18**, 1045-1056.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.

Prévot M, Mattern E, Camps P, Daignierès M (2000) Evidence for a 20° tilting of the Earth's rotation axis 110 million years ago. *Earth and Planetary Science Letters* **179**, 517-528.

Redecker D, Kodner R, Graham LE (2000) Glomalean fungi from the Ordovician. *Science* **289**, 1920-1921.

Remy W, Taylor TN, Hass H, Kerp H (1994) Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proceedings of the National Academy of Science, USA* **91**, 11841-11843.

Ribeiro F, Caticha N (2008) Emergence and loss of assortative mating in sympatric speciation. *Journal of Theoretical Biology* **258**, 465-477.

Roets F, Wingfield MJ, Crous PW, Dreyer LL (2009) Fungal radiation in the Cape Floristic Region: an analysis based on *Gondwanamyces* and *Ophiostoma*. *Molecular Phylogenetics and Evolution* **51**, 111-119.

Roger AJ, Hug LA (2006) The origin and diversification of eukaryotes: problems with molecular phylogenetics and molecular clock estimation. *Philosophical Transactions of the Royal Society B* **361**, 1039-1054.

© University of Pretoria

Rosen DE (1978) Vicariant patterns and historical explanation in biogeography. *Systematic Zoology* **27**, 159-188.

Rossman AY, Farr DF, Castlebury LA (2007) A review of the phylogeny and biology of the Diaporthales. *Mycoscience* **48**, 135-144.

Roux J, Meke G, Kanyi B, Mwangi L, Mbaga A, Hunter GC, Nakabonge G, Heath RN, Wingfield MJ (2005) Diseases of plantation forestry trees in eastern and southern Africa. *South African Journal of Science* **101**, 1-5.

Roux J, Myburg H, Wingfield BD, Wingfield MJ (2003) Biological and phylogenetic analyses suggest that two *Cryphonectria* spp. cause cankers of *Eucalyptus* in Africa. *Plant Disease* **87**, 1329-1332.

Sanmartin I, Ronquist F (2004) Southern hemisphere biogeography inferred by event-based models: plant versus animal patterns. *Systematic Biology* **53**, 216-243.

Seixas CDS, Barreto RW, Alfenas AC, Ferreira FA (2004) *Cryphonectria cubensis* on an indigenous host in Brazil: a possible origin for eucalyptus canker disease? *Mycologist* **18**, 39-45.

Slippers B, Stenlid J, Wingfield MJ (2005) Emerging pathogens: fungal host jumps following anthropogenic introduction. *TRENDS in Ecology and Evolution* **20**, 420-421.

Smith SY, Currah RS, Stockey RA (2004) Cretaceous and Eocene poroid hymenophores from Vancouver Island, British Colombia. *Mycologia* **96**, 180-186.

Spielman LJ (1985) A monograph of *Valsa* on hardwoods in North America. *Canadian Journal of Botany* **63**, 1355-1378.

Steemans P, Le Hérissé A, Melvin J, Miller MA, Paris F, Verniers J, Wellman CH (2009) Origin and radiation of the earliest vascular land plants. *Science* **324**, 353.

Stukenbrock EH, Banke S, Javan-Nikkhah M, McDonald BA (2007) Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Molecular Biology and Evolution* **24**, 398-411.

Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (* and other methods). Version 4.0 beta.

Sytsma KJ, Litt A, Zjhra ML, Pires C, Nepokroeff M, Conti E, Walker J, Wilson PG (2004) Clades, clocks and continents: historical and biogeographical analysis of Myrtaceae, Vochysiaceae, and relatives in the Southern Hemisphere. *International Journal of Plant Science* **165**, S85-S105.

Takezaki N, Rzhetsky A, Nei M (1995) Phylogenetic test of the molecular clock and linearized trees. *Molecular Biology and Evolution* **12**, 823-833.

© University of Pretoria

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596-1599.

Taylor JW, Berbee ML (2006) Dating divergences in the fungal tree of life: review and new analyses. *Mycologia* **98**, 838-849.

Taylor TN, Hass H, Kerp H (1999) The oldest fossil ascomycetes. *Nature* **399**, 648.

Taylor TN, Hass H, Kerp H, Krings M, Hanlin RT (2005) Perithecial ascomycetes from the 400 million year old Rhynie chert: an example of ancestral polymorphism. *Mycologia* **97**, 269-285.

Thorne JL, Kishino H (2002) Divergence time and evolutionary rate estimation with multilocus data. *Systematic Biology* **51**, 689-702.

Upchurch P (2008) Gondwanan break-up: legacies of a lost world? *TRENDS in Ecology and Evolution* **23**, 229-236.

Van der Merwe NA, Gryzenhout M, Steenkamp ET, Wingfield BD, Wingfield MJ (2010) Multigene phylogenetic and population differentiation data confirm the existence of a cryptic species within *Chrysoporthe cubensis*. *Fungal Biology* **114**, 966-979.

Van der Merwe NA, Myburg H, Wingfield BD, Wingfield MJ (2001) Identification of *Cryphonectria cubensis* from Colombia based on rDNA sequence data. *South African Journal of Science* **97**, 295-296.

Van Heerden SW, Wingfield MJ, Coutinho T, Van Zyl LM (1997) Population diversity among Venezuelan and Indonesian isolates of *Cryphonectria cubensis*. *South African Journal of Science* **93**, xiv.

Van Staden V, Erasmus BFN, Roux J, Wingfield MJ, Van Jaarsveld AS (2004) Modelling the spatial distribution of two important South African plantation forestry pathogens. *Forest Ecology and Management* **187**, 61-73.

Van Zyl LM, Wingfield MJ, Kemp GHJ, Alfenas AC, Crous PW (1994) Population diversity in Brazilian isolates of *Cryphonectria cubensis*. *South African Journal of Science* **91**, 7 (Abstract).

Voelker G (1999) Dispersal, vicariance, and clocks: historical biogeography and speciation in a cosmopolitan passerine genus (*Anthus*: Motacillidae). *Evolution* **53**, 1536-1552.

Wedin M (1993) A phylogenetic analysis of the lichen family Sphaerophoraceae (Caliciales); a new generic classification and notes on character evolution. *Plant Systematics and Evolution* **187**, 213-241.

Welch JJ, Bromham L (2005) Molecular dating when rates vary. *TRENDS in Ecology and Evolution* **20**, 320-327.

© University of Pretoria

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenies. In: *PCR protocols: a guide to methods and applications*. (eds. Innes MA, Gelfand DH, Sninsky SS, White TJ), pp. 315-322. Academic Press Inc., San Diego, California.

Wingfield MJ (1999) Pathogens in exotic plantation forestry. *International Forestry Review* **1**, 163-168.

Wingfield MJ, Rodas C, Myburg H, Venter M, Wright J, Wingfield BD (2001) Cryphonectria canker on *Tibouchina* in Colombia. *Forest Pathology* **31**, 1-10.

Wingfield MJ, Swart WJ, Abear B (1989) First record of Cryphonectria canker of *Eucalyptus* in South Africa. *Phytophylactica* **21**, 311-313.

Wingfield MJ, Van Zyl LM, Van Heerden S, Myburg H, Wingfield BD (1997) Virulence and the genetic composition of the *Cryphonectria cubensis* Bruner population in South Africa. In: *Physiology and Genetics of Tree-Phytophage Interactions* eds. Lieutier F, Mattson WJ, Wagner MR), pp. 163-172. INRA Editions.

Yang Z (2007) *Computational molecular evolution* Oxford University Press Inc., New York.

Zhaxybayeva O, Gogarten JP (2004) Cladogenesis, coalescence and the evolution of the three domains of life. *Trends in Genetics* **20**, 182-187.

© University of Pretoria

CHAPTER 5

Host shifts cause diversity bottlenecks in populations of *Chrysoporthe austroafricana*

© University of Pretoria

Abstract

Chrysoporthe austroafricana is fungal pathogen of native *Syzygium* and non-native *Eucalyptus* (Myrtaceae) as well as non-native *Tibouchina* (Melastomataceae) in Africa. The fungus occurs in the sub-Saharan parts of the continent, and has been suggested to have evolved on native *Syzygium* species. The aim of this study was to test this hypothesis by inferring the migration routes of the pathogen in Africa. Population genetic data based on 12 microsatellite loci for populations of the fungus originating in southern Africa were used to achieve this goal. The results of calibrated phylogenetic analysis on a collection of 167 isolates indicated that those from *Syzygium* in Mozambique represented the oldest lineage, thus supporting the notion that *C. austroafricana* evolved on *Syzygium*. Data for populations from three different hosts, namely *Syzygium*, *Eucalyptus* and *Tibouchina*, also suggested that host jumps may have resulted in important diversity bottlenecks. While populations from *Syzygium* had the highest gene diversity, those from non-native *Tibouchina* were much less diverse. A similar trend could not be observed for populations defined in terms of geography. Our findings thus support the view that *C. austroafricana* is native to Africa and that the genetic diversity in this fungus is determined by host species. These data also indicate that non-native hosts impose a selective pressure on the native population of the pathogen, which results in reduced numbers of genotypes on these non-native hosts.

© University of Pretoria

Introduction

Chrysoporthe austroafricana forms part of an assemblage of tree killing fungal pathogens that are distributed around the world in temperate to tropical climates (Gryzenhout *et al*. 2009). The genus *Chrysoporthe* includes eight species of tree and shrub pathogens, and new species and hosts of these fungi are frequently being discovered (e.g., Chungu *et al*. 2009; Gryzenhout *et al*. 2006; Heath *et al*. 2006; Rodas *et al*. 2005; Van der Merwe *et al*. 2010). *Chrysoporthe austroafricana* specifically, is the causal agent of Chrysoporthe canker on non-native plantation *Eucalyptus* trees in Africa (Wingfield *et al*. 1989; Wingfield *et al*. 1997). The disease is responsible for significant losses in yield due to basal cankers on the tree stems, which girdle trees and cause stem breakage as well as tree mortality (Wingfield *et al*. 1989). Two close relatives of *C. austroafricana*, namely *C. cubensis* and *C. deuterocubensis*, occur in similar climatic regions in South America and Southeast Asia, respectively (Van der Merwe *et al*. 2010). These species also cause canker diseases of *Eucalyptus* (c.f. Old *et al*. 2003; Pegg *et al*. 2010), although cankers are not only basal but also found higher up on the stems of trees (Gryzenhout *et al*. 2009).

The distribution of *Chrysoporthe austroafricana* is limited to Africa (Gryzenhout *et al*. 2009) where it appears to have a relatively narrow host range that is restricted to members of the closely related (Conti *et al*. 1997) plant families Myrtaceae and Melastomataceae. Apart from its non-native Myrtaceae host *Eucalyptus*, *C. austroafricana* is also known to infect native *Syzygium* (Lucas *et al*. 2005; Sytsma *et al*. 2004), and non-native *Tibouchina* (Melastomataceae) (Myburg *et al*. 1999). When the pathogen (as *Cryphonectria cubensis*) was first discovered in South Africa (Wingfield *et al*. 1989), it was suggested that the fungus had been introduced into the country as it was known only from non-native *Eucalyptus* trees deployed in forestry plantations. However, the fungus was later found on native *Syzygium* species (Heath *et al*. 2006), which suggested that *C. austroafricana* probably had an African origin (Gryzenhout *et al*. 2009). The results of a recent phylogeography study of *C. austroafricana* and its two close relatives, *C. cubensis* and *C. deuterocubensis* (Chapter 4, this thesis), have suggested that the stem lineage of *Chrysoporthe* moved with their Myrtalean hosts from Southeast Asia to South America and Africa, followed by allopatric speciation and the emergence of *C. deuterocubensis*, *C. cubensis* and *C. austroafricana*, respectively.

© University of Pretoria

Although relatively large collections of *C. austroafricana* are available from *Eucalyptus*, *Syzygium* and *Tibouchina* (Nakabonge *et al*. 2006; Van Heerden *et al*. 2005), the population biology of the fungus is not well understood. Populations of *C. austroafricana* and its close relatives have been assessed for diversity based on vegetative incompatibility (Van Heerden & Wingfield 2001; Van Zyl *et al*. 1998) and polymorphic DNA markers (Nakabonge *et al*. 2007). Although vegetative compatibility provides a simple method to assess genotypes, it is inappropriate for fine scale studies of populations where gene diversity or migration is of interest. This is because the incompatibility reaction is governed by an unknown number of loci and alleles where the interactions that yield incompatibility are not well understood in *Chrysoporthe* spp. Genetic markers such as microsatellites are also insufficient when populations are assessed in a phylogenetic context, due to high levels of size homoplasy (Estoup *et al*. 2002). An additional problem with population genetic studies of *Chrysoporthe* spp. is that populations are defined based on their geographic origins, and few (Van Heerden *et al*. 1997) comparisons between populations exist.

Fungal population diversity is traditionally assessed in terms of geographic separation or host preference (McDonald & Linde 2002; Zhou *et al*. 2007). When geographically defined populations are considered, the presence of unique alleles in populations can be used as evidence of population subdivision based on an isolation-by-distance model (Zhou *et al*. 2007). Similarly, subdivision in host-based populations can be used to invoke selection pressure as a prerequisite to evolution (Baker *et al*. 2003), or gene flow can be used to illustrate that a host jump has occurred (Woolhouse *et al*. 2005). Although fungal pathogen populations have previously been compared based on host association and transcontinental geography (Atallah *et al*. 2010; Zhou *et al*. 2007), the differences in population structure of a pathogen or pest, simultaneously considering spatiotemporal and host parameters, are rarely considered (Franke *et al*. 2009).

The overall aim of this study was to evaluate the population biology of *C. austroafricana* by specifically considering the influence of geography and host on population structure. Two hypotheses regarding the population genetics and evolution of the fungus were tested. The first was that population differentiation based on geography would be larger than that based on the hosts from which the fungi were isolated. This would be in accordance with the notion

© University of Pretoria

of isolation-by-distance, as well as the population structures of many other plant pathogenic fungi. The second hypothesis was that native *Syzygium* species are the natural hosts of *C. austroafricana* in Africa as suggested previously (Gryzenhout *et al*. 2009; Heath *et al*. 2006) (Chapter 4, this thesis), while *Tibouchina* and *Eucalyptus* represent secondary hosts that were colonized via host jumps. A consequence of the second hypothesis would, therefore, be a reduced genetic diversity as well as reduced numbers of unique alleles (i.e., a founder effect) in the populations of the pathogen associated with these hosts. We also exploited the opportunity to compare the relative ages of populations on the different hosts in order to corroborate information gathered from population genetic data, which were used to formulate a theory regarding the colonization of Africa by *C. austroafricana*.

Materials & Methods

Fungal cultures and DNA extraction

Chrysoporthe austroafricana isolates were made from single conidia, which were obtained from pycnidia on cankers of infected *Eucalyptus*, *Syzygium* and *Tibouchina* trees in South Africa, Mozambique, Malawi, Zambia and Madagascar. For the few instances where perithecia occurred on *Syzygium* hosts, single ascospore isolates were obtained. A single perithecium or pycnidium from a single canker on a tree was dissected at the base using a sterile scalpel. Ascospore or conidial masses were removed using a sterile needle and suspended in 1 ml of sterile distilled water. One hundred micro-litres of the suspension was streaked onto 20% w/v malt extract agar (MEA) and plates were incubated in the dark at 25 °C for eight hours. Single germinating ascospores or conidia were cut from the MEA plates using a sterile needle and dissection microscope, and each was transferred to a 20% w/v potato dextrose agar (PDA) plate and allowed to grow for seven days at 25 °C in the dark.

Two overlapping sets of isolates were used. The first set was used to infer the relative ages of *C. austroafricana* lineages and included 69 isolates of the fungus from various African countries (South Africa, Malawi, Mozambique, Zambia and Madagascar), and representatives of *C. cubensis* and *C. deuterocubensis* that were used for outgroup purposes. A second set of isolates was used to determine various population genetic parameters based on microsatellite data (see below), and included 167 isolates from *Eucalyptus*, *Syzygium* and *Tibouchina* trees in Mozambique, South Africa and Zambia (Table). Isolates from Malawi and Madagascar

© University of Pretoria

were excluded from the population genetics set due to the limited number of samples from these regions. All isolates used in this study were deposited into the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, Pretoria, South Africa.

Fungal cultures were inoculated into 0·8 ml malt extract broth (20% w/v) in 1·5 ml micro-centrifuge tubes and incubated in the dark at 25°C for seven days. Tubes were shaken daily to allow aeration and even growth of fungal mycelium. Total genomic DNA was extracted using a CTAB (hexadecyltrimethylammonium bromide) and phenol-chloroform extraction method (Steenkamp *et al*. 1999), and the DNA concentration was standardized to 25 ng/µl after spectrophotometric quantification.

Inference of the relative ages of *C. austroafricana* populations

To determine the relative ages of the populations of *C. austroafricana* examined in this study, the DNA sequence information for five genomic regions that either do not code for proteins or are rich in introns were used. This was accomplished with primer set ACT-512F+ACT-783R (Carbone & Kohn 1999) that targets an intron in the gene encoding actin (Act), primer set EF1-728F+EF1-986R (Carbone & Kohn 1999) that targets two introns in the Elongation Factor 1-α (EF) gene, and primer sets Bt1a+Bt1b and Bt2a+Bt2b (Glass & Donaldson 1995), which respectively target three (BT1) and one (BT2) intron in the gene encoding β-tubulin. For amplification and sequencing of the internal transcribed spacer regions (ITS) of the ribosomal RNA operon we used primer set ITS1+ITS4 (White *et al*. 1990). Amplification reactions were performed following previously published conditions (Glass & Donaldson 1995; Van der Merwe *et al*. 2010). PCR products were purified by polyethylene glycol precipitation (Hartley & Bowen 1996) and sequenced in both directions using the original PCR primers and the standard BigDye® dye terminator protocol (Applied Biosystems, USA) and an ABI™ Prism® 3700 automated DNA sequencer (Applied Biosystems).

The resulting sequences for the five regions were aligned using Muscle 3·6 (Edgar 2004) and corrected where necessary using SeaView 2·2 (Galtier *et al*. 1996). A partition homogeneity test as implemented in PAUP* 4·0b10 (Sinauer Associates, Inc.) was performed for each pair-wise combination of sequenced regions, as well as for a combination of all five regions. Maximum likelihood (ML) phylogenetic analyses were performed for each of the gene

© University of Pretoria

regions, as well as the combined data set, using PAUP*. These analyses employed the best-fit substitution models as determined using jModelTest (Posada 2008). All data sets employed the Generalized Time Reversible (GTR; Tavaré 1986) model, while the EF, ITS and combined data set, as well as the BT2 regions of the β-tubulin gene, also included gamma correction for among-site rate variation (G) and a proportion of invariable sites (I). ML analysis was done once without a constant molecular clock, and once with a constant clock enforced. The χ^2 significance of tree likelihood ($2\Delta L$) statistics was used to determine whether a constant molecular clock was present in the data sets (Huelsenbeck & Crandall 1997; Huelsenbeck & Rannala 1997).

Markov Chain Monte Carlo simulations of the combined data set was performed with BEAST 1·6·0 (Drummond *et al*. 2006; Drummond & Rambaut 2007). For this Bayesian analysis, the root height was calibrated to 7·07 Mya (Chapter 4, this thesis) while the divergence between *C*. *cubensis* and *C*. *austroafricana* was calibrated to 1·39 Mya. An uncorrelated lognormal relaxed molecular clock incorporating the Yule speciation algorithm (Aldous 2001) was employed. The same substitution model as for the combined ML analysis was employed, and the time to the most recent common ancester (t_{MRCA}) was tracked for groups of isolates from different regions and hosts. Analyses were performed using a high performance computing environment, with ten million steps per computational node over ten nodes, yielding a different random seed number for each node.

Population biology of *C. austroafricana*

The aligned sequence data sets for the five gene regions that were used to infer the relative ages for the populations of *C*. *austroafricana* were analyzed for recombination (reviewed by Nordborg 2000). Ancestral recombination graphs (ARGs) were reconstructed using the BEAGLE module in SNAP Workbench (Price & Carbone 2005). For this purpose, sequences were collapsed to haplotypes while recoding indels and excluding sites that violate the infinite-sites model.

To further study the population biology of *C*. *austroafricana*, we utilized ten PCR primer sets that were developed previously (Van der Merwe *et al*. 2010; Van der Merwe *et al*. 2003). Because two of these primer sets (SA6F+SA6R and SA9F+SA9R) each targeted an additional locus, the ten primer sets allowed us to study 12 polymorphic loci. These

© University of Pretoria

polymorphic markers were amplified as described previously (Van der Merwe *et al*. 2003) and analyzed using a LIZ500 internal size standard (Applied Biosystems, USA), an ABI™ Prism® 3700 automated DNA analyzer and GeneScan software (Applied Biosystems).

The markers analyzed for each isolate were scored with the GeneMapper™ 3·5 software (Applied Biosystems) based on the sizes of amplified PCR products. The resulting alleles for each isolate were recorded in haplotype data matrices, which were subjected to standard population genetic analyses. For genotype related analyses, the matrices were translated into single characters for each allele, producing a multi-character genotype or multi-locus haplotype for each isolate. In order to quantify the genetic distances between populations, the numbers of private alleles in each population were manually calculated.

The amount of diversity sampled within each population was evaluated using MultiLocus 1·3b (Agapow & Burt 2001) by modeling the observed diversity against the number of loci, with 1000 independent repetitions. When a graph depicting these values is drawn and it reaches a plateau, this indicates that the population was adequately sampled and that the use of additional isolates or polymorphic loci will not increase the estimated genetic diversity (Agapow & Burt 2001). Multilocus linkage disequilibrium (Nei & Li 1973) within each population was assessed using the Index of Association (Agapow & Burt 2001) (I_A) as implemented in the MultiLocus 1·3b software package. The observed I_A value of each population was compared to 10000 randomizations of the data set. The I_A statistic was normalized to yield $\bar{r}_d$, which enables comparison of indices of disequilibrium across multiple studies and equals zero at gametic equilibrium. Randomizations also provide a confidence level, which is high ($P \leq 0{\cdot}05$) if the observed data deviates significantly from the null hypothesis of random association of alleles, and low ($P > 0{\cdot}05$) when the null hypothesis can be accepted.

Haplotype diversity (gene diversity, H) (Nei 1973) within populations was calculated with the formula

$$H = 1 - \sum p_{ij}^{2}$$

where p_i is the frequency of the i^{th} haplotype and j is the total number of haplotypes. The genetic distance (Nei 1972) expressed as Weir and Cockerham's (Weir & Cockerham 1984), was calculated in all possible pair wise combinations using MultiLocus 1·3b. This

© University of Pretoria

statistic is a variation on Wright's F_{ST} (Agapow & Burt 2001; Weir & Cockerham 1984; Weir & Hill 2002; Wright 1951), and is low (0·0 to 0·05) when there is little differentiation, between 0·05 to 0·15 when there is moderate differentiation, and above 0·15 when there is a high level of differentiation between populations. Confidence in the obtained values was tested with one million randomizations of alleles within populations. Additionally, populations from different hosts and geographic locations were combined and compared in the same fashion.

The theoretical number of migrants ($\hat{M}$, estimated from Weir and Cockerham's θ) (Slatkin 1995) for haploid populations was estimated with the formula

$$\hat{M} = \tfrac{1}{2}\left(\tfrac{1}{\theta} - 1\right)$$

This statistic allows testing of isolation by distance among populations. It is, therefore, expected to be negatively correlated with the logarithm of the geographic distance between populations (Milgroom & Lipari 1995; Slatkin 1995).

The genotypic diversity of populations was estimated as $\hat{G}$ (McDonald *et al*. 1994), which is the normalized G_{ST} (Stoddart & Taylor 1988), and indicates the percentage of maximum diversity sampled:

$$G = \frac{1}{\sum_{x=0}^{N}\left[f_x\left(\frac{x}{N}\right)^2\right]}$$

$$\hat{G} = \frac{G}{N}$$

where N is the sample size, and f_x is the number of genotypes observed x times in the population. This statistic makes it possible to compare diversities between populations with unequal numbers of individuals, and minimizes the dependence of diversity estimates on sample size. The statistical significance of differences in $\hat{G}$ values between populations was evaluated using a *t*-test with $(N_1 + N_2 - 2)$ degrees of freedom:

$$t = \frac{\left|\frac{G_1}{N_1} - \frac{G_2}{N_2}\right|}{\sqrt{\frac{\mathrm{Var}(\hat{G}_1)}{{N_1}^2} + \frac{\mathrm{Var}(\hat{G}_2)}{{N_2}^2}}}$$

where

© University of Pretoria

$$\mathrm{Var}(\hat{G}) = \frac{4}{N} G^2 \left[G^2 \sum_{i=1}^{K} {p_i}^3 - 1 \right]$$

and G is the genotypic diversity obtained for each data set, K is the number of genotypes in the sample, p_i is the frequency of the i^{th} genotype in the sample, and $\hat{G}$ is the maximum likelihood estimator for Stoddart and Taylor's G_{ST} (Stoddart & Taylor 1988). These t-test values were compared to a table of critical values for two-tailed t-tests at a 95% level of confidence ($P = 0{\cdot}05$).

The probability of sampling a private allele from any random individual in a subpopulation (φ) was defined as the proportion of private alleles in the subpopulation to the subpopulation size, normalized with the total number of alleles over all loci in the metapopulation. Thus,

$$\varphi = \frac{\rho_S}{N_S} \bullet \frac{N_M}{\alpha_M}$$

where ρ_S is the number of private alleles in the subpopulation, N_S is the number of individuals in the subpopulation, α_M is the total number of alleles over all loci in the metapopulation, and N_M is the number of individuals in the metapopulation. This statistic gives an indication of the level of uniqueness of the subpopulation, relative to the metapopulation.

Results

Inference of the relative ages of *C. austroafricana* populations

Phylogenetic analyses were based on the DNA sequence data for five genomic regions in 69 isolates of *C. austroafricana*, as well as two isolates of *C. deuterocubensis* and one isolate of *C. cubensis*, to represent an outgroup. The total alignment was 1442 characters in length, and the ACT, BT1, BT2, EF and ITS data sets respectively contained 235, 467, 401, 327 and 479 characters. Based on the results of partition homogeneity tests, all these data sets represented homogeneous partitions ($P > 0{\cdot}01$) (Cunningham 1997) with P-values ranging from 0·014 for the BT1-ITS partition pair, to 1·00 for the ACT combinations with ITS, BT1 and BT2.

Comparison of the likelihood ratios determined from ML analyses of the various data sets revealed that substitution rates in none of the data sets conformed to that predicted under the constant molecular clock hypothesis. All further analyses were, therefore, based on a

© University of Pretoria

lognormal clock in order to translate nucleotide substitutions into branch lengths and divergence times. The resulting chronogram showed that the outgroup taxa (*C. cubensis* and *C. deuterocubensis*) grouped separate from *C. austroafricana* isolates (Figure 1). Isolate CMW13876 from *Syzygium* in Mozambique formed the root of *C. austroafricana* (1 Mya), which was separated into two clades (Clade A and Clade B, Figure 1). Both clades consisted of isolates from several countries and hosts, but Clade B1 consisted of isolates from *Eucalyptus* in Zambia and grouped apart from other isolates in Clade B. Similarly, Clade A1 grouped apart from other isolates in Clade A and consisted of isolates from *Eucalyptus* and *Syzygium* in Malawi, Mozambique and Zambia. The one isolate from *Syzygium* in Madagascar (CMW27300) grouped with five isolates from *Syzygium* in South Africa, but appeared to be ancestral to that clade of South African isolates. Isolates from *Tibouchina* were interspersed throughout the tree, with no clear pattern of evolution.

Population biology of *C. austroafricana*

DNA sequence alignments for ACT, BT1, BT2, EF and ITS were individually submitted to ARG reconstruction using BEAGLE. Ancestral recombination events were detected in the BT1 region of β-tubulin, as well as the EF data set. In both these cases, it appeared that a genotype from *Syzygium* recombined with one from *Eucalyptus* to produce offspring that could colonize all three hosts (BT1, Figure 2C), or only *Eucalyptus* (EF, Figure E). All other data sets, i.e. ACT, ITS and BT2 (Figure 2A,B,D), produced ARGs with no recombination but only mutation and divergence.

Various alleles were detected for all 12 of the examined microsatellite markers among the population genetics data set of 167 isolates. However, null alleles were observed for some loci. Since these loci could not be amplified after several attempts, null alleles were treated as an additional allele in subsequent analyses. All loci were polymorphic across the combined data set. The number of alleles per locus ranged from three for *CcPMG* to 10 for *SA4*. The results of randomization tests also indicated that the sampling was sufficient for all the populations, and that additional isolates would not have increased the accuracy of statistical inference (Agapow & Burt 2001). The only exception was for the population from *Eucalyptus* in Zambia (N = 6), which was not adequately sampled (Table 1). As a result, statistics from the Zambian population were not regarded as significant in subsequent population genetic analyses.

© University of Pretoria

When populations were compared based on geographic origin (South Africa, Mozambique and Zambia), no private alleles were detected in the population from Zambia ($\varphi = 0$). However, the population from Mozambique had 11 private alleles ($\varphi = 0.741$), while the one from South Africa had 20 private alleles ($\varphi = 0.444$). When populations were divided based on host of origin, a more defined pattern became evident. Populations from *Syzygium*, *Eucalyptus* and *Tibouchina* had 16 ($\varphi = 0.598$), five ($\varphi = 0.205$) and two ($\varphi = 0.186$) private alleles, respectively. Therefore, the probabilities of sampling an individual harboring a private allele from these three populations were 59·8%, 20·5% and 18·6% respectively.

Gametic disequilibrium was evaluated by determining the significance of deviation from the null hypothesis of random mating (gametic equilibrium) (Agapow & Burt 2001). The null hypothesis could not be rejected for the populations from *Syzygium* in Mozambique ($N = 19$) and from *Eucalyptus* in Zambia ($N = 6$). All other populations displayed *P*-values ranging from 0·017 to 0·683 (Table 1), which indicated that these populations could be regarded as non-randomly mating.

Nei's gene diversity (H) for all populations was relatively high, and ranged from 0·296 for the population from *Tibouchina* in South Africa, to 0·459 for the population from *Eucalyptus* in Mozambique. However, gene diversity of this population was similar to that for the population from *Syzygium* in the same country ($H = 0{\cdot}442$). Gene diversities for populations from the three hosts in South Africa showed an apparent gradient, with the population from *Syzygium* displaying the highest gene diversity and the one from *Tibouchina* the lowest (Table 1).

A gradient was also observed in the values for the maximum percentage of genotypic diversity ($\hat{G}$) and the φ-values derived from private alleles in the populations. However, a *t*-test revealed that there were no significant differences between genotypic diversities for any of the populations. This discrepancy between genotypic diversity and gene diversity values were informative due to the large variances associated with *G*-values, which could result in acceptance of the null hypothesis that populations are not different. These large variances are an inherent problem with the currently available methods of quantifying diversity and differentiation (Jost 2008).

© University of Pretoria

Population differentiation analyses revealed that the highest number of migrants per generation ($\hat{M}$ = 15·63) was between the *Syzygium* and *Eucalyptus* populations in Mozambique (Table 1, Table 3). When countries of origin were compared, the highest number of migrants was between South Africa and Zambia ($\hat{M}$ = 8·27). However, the small population size for Zambia could have influenced this result. When hosts were compared, the highest number of migrants was detected between *Syzygium* and *Eucalyptus* ($\hat{M}$ = 8·12), while the numbers of migrants between *Syzygium* and *Tibouchina* ($\hat{M}$ = 1·20) and between *Eucalyptus* and *Tibouchina* ($\hat{M}$ = 1·13) were lower. The mean number of migrants between populations defined by country of origin was 3·67, while the same statistic for populations defined by host was 2·25. These statistics thus suggest that more differentiation existed between host-based populations, i.e. there were fewer migrants, than between those defined by geographic criteria (Table 3).

Discussion

The results of this study showed that the hosts from which *C. austroafricana* isolates originated played an important role in the population biology of the pathogen. Based on allele frequencies, populations of *C. austroafricana* showed a lower level of population subdivision between countries than between hosts (Table 3). This is in contrast to previous studies where geography was used as the criterion to define populations of *Chrysoporthe* (Nakabonge *et al*. 2007; Van Heerden *et al*. 1997). However, this result is not surprising, given the differences between the hosts from which *C. austroafricana* isolates were collected. These differences include the fact that both *Syzygium* and *Eucalyptus* are in the family Myrtaceae, while *Tibouchina* is in the Melastomataceae, as well as the fact that *Syzygium cordatum* and *S. guineense* are native to Africa while *Eucalyptus* and *Tibouchina* are not (Gryzenhout *et al*. 2009).

When populations of *C. austroafricana* were defined based on host of origin, population structure was revealed by significant levels of population subdivision. The population from *Syzygium* had the highest number of private alleles ($\varphi = 0.598$), while the one from *Tibouchina* had the lowest ($\varphi = 0.186$). Therefore, diversity of *C. austroafricana* populations is closely linked to the host of origin, i.e. the native Myrtaceae hosts (*S. cordatum* and *S.*

© University of Pretoria

guineense) harbored the most diverse *C. austroafricana* population that also displayed the highest number of private alleles. Conversely, the non-native host in the Melastomataceae, i.e. *Tibouchina*, harbored the least diverse *C. austroafricana* population with the lowest number of private alleles. The population from *Eucalyptus* ($\varphi = 0.205$) was intermediate between those on *Syzygium* and *Tibouchina*, and this observation can be linked to the fact that *Eucalyptus* is a non-native host in the Myrtaceae. Therefore, it has biological and geographical characteristics in common with both of the other hosts (Sytsma *et al*. 2004).

Our findings suggest that *C. austroafricana* employes different reproductive modes on native and introduced hosts. Tests for gametic disequilibrium revealed that the population originating from *Syzygium* in Mozambique was in gametic equilibrium, i.e. alleles were randomly associated, while those from *Eucalyptus* and *Tibouchina* in all areas were in gametic disequilibrium (Table 1). The population originating from *Syzygium* included some isolates that were made from sexual spores. However, only asexual fruiting structures have ever been observed on *Eucalyptus* and *Tibouchina*. This observation raises questions about the ability of *C. austroafricana* to reproduce sexually on the two non-native hosts. Surprisingly, the population from *Syzygium* in South Africa did not display random association, even though a few sexual fruiting structures were observed on this host. The most parsimonious explanation for these observations is that *C. austroafricana* originated in Mozambique, or another area to the north on native hosts, and during its southerly migration the population became depleted of genetic diversity. This would result in a loss of statistical significance of the test for gametic disequilibrium due to sexual reproduction between closely related individuals. Unfortunately we could not detect this trend of reduced genotypic diversity in a southerly direction, which is probably a limitation imposed by the number of microsatellite loci used in this study.

Examination of the private alleles of *C. austroafricana* isolated from the different hosts revealed that the emergence of new alleles is more likely on the host where sexual reproduction occurs. This is because the numbers of private alleles correspond to the levels of sexual reproduction on the different hosts. More alleles that are private were detected among isolates collected from *Syzygium* than from *Eucalyptus* or *Tibouchina*. Thus, the high numbers of private alleles in the *C. austroafricana* population from *Syzygium* is indicative of both sexual reproduction as well as the ancestral status of this population. The same situation

© University of Pretoria

existed in *Rhynchosporium secalis* when populations of this fungus were analyzed for diversity and private alleles (Linde *et al*. 2009), where ancestral and sexually reproducing populations exhibited higher numbers of private alleles.

The population genetic data emerging from this study suggest that *C*. *austroafricana* jumped from its native Myrtaceae hosts to introduced Myrtalean hosts. Although we could not determine the direction of gene flow between the hosts, it seems logical that gene flow should occur from the population with the highest number of private alleles to the one with the lowest number of private alleles. Based on this assumption, gene flow probably occurred from the *Syzygium* populations to the *Eucalyptus* and *Tibouchina* populations. Gene flow was also detected between the *Eucalyptus* and *Tibouchina* populations. Accordingly, the population from *Syzygium* most likely represents the parental population, while the populations from *Eucalyptus* and *Tibouchina* are secondary populations that resulted from host jumps. These jumps were associated with genetic bottlenecks due to the two new hosts and corresponding reductions in the genetic diversity and numbers of private alleles.

Gene flow inferences were corroborated with the detection of ancestral recombination among lineages of *C*. *austroafricana* (Figure 2). Interestingly, both of the detected recombination events occurred between a lineage from *Syzygium* and one from *Eucalyptus*, although it is possible that these lineages could occur on other hosts. The descendents of these recombination events were detected on all three of the examined hosts, illustrating that the generation of new genotypes through sexual recombination is necessary for the colonization of another host.

The detection of ancestral recombination among lineages of *C*. *austroafricana* (Figure 2) supports the notion that this fungus occasionally outcrosses in nature. This is despite the fact that most of the examined populations of *C*. *austroafricana* were in linkage disequilibrium, and apparently do not reproduce sexually (Table 1). Nordborg (2000) suggested that linkage disequilibrium should be more extensive in self-fertilizing populations, while alleles will be more randomly associated when sexual outcrossing occurs frequently. Therefore, the observed linkage disequilibrium in the populations of *C*. *austroafricana* is probably due to a homothallic mode of sexual reproduction on native or preferred hosts, while the few instances

© University of Pretoria

of linkage equilibrium reflect the fact that sexual outcrossing is more common in these populations.

Although *Eucalyptus* has been on the African continent for less than 200 years (Poynton 1959), the time to the most recent common ancestor for isolates of *C. austroafricana* from *Eucalyptus* is much older. It can be hypothesized that repeated colonization of *Eucalyptus* from native *Syzygium* trees would have increased the number of distinct lineages on the new host, thereby increasing the time to coalescence for those lineages. The phylogenetic tree presented in this study (Figure 1) supports this hypothesis. From this tree, it can be inferred that *Eucalyptus* was colonized at least six different times, and this is based on a relatively small number of isolates. Similarly, *Tibouchina* was colonized at least three different times. Therefore, the coalescent ages of populations on non-native hosts are not related to the amount of time available for colonization. In other words, the age of the host population does not set a limit on the coalescent age of the pathogen population.

Based on the results of the chronometric analysis, the direction of host colonization was inferred (Figure 1). The most ancestral lineage of *C. austroafricana* was one from *Syzygium* in Mozambique, which is in accordance with the currently accepted notion that *C. austroafricana* must have evolved on this native host in Africa (Gryzenhout *et al*. 2009). The population from *Syzygium* in Mozambique represented the oldest composite lineage within the sample (Table 2; t_{MRCA} = 1·027 Mya, 0·520 – 1·542 95% HPD). Following the emergence of clades A and B, the roots of each are represented exclusively by isolates from *Syzygium* and *Eucalyptus* (sub-clade A1) and by isolates from *Eucalyptus* (sub-clade B1), respectively. These patterns may be affected by insufficient sampling of *C. austroafricana* from the different hosts, but *Eucalyptus* appears to have been colonized before *Tibouchina*. In addition, our strategy did not allow us to determine whether each of the non-native hosts were colonized independently, or whether cross-colonization would be a more appropriate hypothesis. Nevertheless, the fungus repeatedly jumped hosts to *Eucalyptus* and *Tibouchina* (Table 1; Table 3). The corresponding loss in genetic diversity could have been caused by stricter selection pressures imposed by the newly colonized hosts, selection due to a change in climate (Jackson *et al*. 2009; Jump *et al*. 2006; Vacher *et al*. 2008), or repeated re-establishment of populations from one host to another (Thrall & Burdon 1997).

© University of Pretoria

The overall results of this study showed that a calibrated phylogenetic tree could be combined and integrated with traditional population genetic data in order to clarify the evolution and movement of populations of a fungal species. Following this approach, we demonstrated that *C. austroafricana* jumps hosts frequently and that sexual reproduction primarily takes place on the host on which it evolved. For the first time, evidence was presented that the tree hosts of the *C. austroafricana* fungus may impose an important selective pressure by diminishing the populations' gene diversities. In turn, the genotypes present on *Eucalyptus* and *Tibouchina* are maintained by asexual propagation, suggesting that the fungus compensates for the selective pressures imposed by the non-native hosts by limiting sexual reproduction. Thus, host jumping causes genetic bottlenecks in populations of *C. austroafricana*. Contrary to previous studies that assumed an isolation-by-distance model and defined populations based on geographic origins, our results indicated that populations of *C. austroafricana* displayed marginally higher levels of differentiation when host rather than geographic region defined them.

Acknowledgements

Financial support for this work was provided by the National Research Foundation (NRF, South Africa), the Department of Science and Technology (DST), and members of the Tree Protection Co-operative Programme and the DST/NRF Center of Excellence in Tree Health Biotechnology. We also thank Dr. G. Nakabonge who contributed to the set of isolates used in this study.

© University of Pretoria

Table 1 Population genetic statistics for populations obtained from different hosts in South Africa, Zambia and Mozambique.

	Mozambique		South Africa			Zambia ①
Statistic	***Eucalyptus***	***Syzygium***	***Eucalyptus***	***Syzygium***	***Tibouchina***	***Eucalyptus***
Number of individuals (*N*)	21	19	39	53	29	6
Adequate sampling	Yes	Yes	Yes	Yes	Yes	No
Gametic disequilibrium ②	$P = 0{\cdot}017$	$P = 0{\cdot}299$	$P < 0{\cdot}0001$	$P < 0{\cdot}0001$	$P < 0{\cdot}0001$	$P = 0{\cdot}683$
Gene (haplotype) diversity (*H*)	0·459	0·442	0·357	0·400	0·296	0·313
Genotypic diversity ($\hat{G}$) ③	1·000	1·000	0·7959	0·8689	0·6744	1·000

① The statistics gathered for the Zambian population were regarded as insignificant due to the small sample size (*N*).

② Gametic disequilibrium is expressed as the probability of rejecting gametic equilibrium (random association of alleles) and this is an indication of the reproductive strategy of the population. When gametic equilibrium can be rejected, the population is in significant deviation from random recombination. Randomly mating populations are boxed.

③ We use *G* to indicate the genotypic diversity, and $\hat{G}$ for the normalized genotypic diversity to allow comparisons between populations.

© University of Pretoria

Table 2 Time to the most recent common ancestor (t_{MRCA} in Mya) for individual populations by country-and-host pair, as well as by country and by host separately. Sorted according to the mean t_{MRCA}, from oldest to youngest.

Population	t_{MRCA} Mean	t_{MRCA} Range (95% HPD)	Effective Sample Size (ESS)
Mozambique: *Syzygium*	1·027	0·520 – 1·542	696·79
Mozambique: *Eucalyptus*	0·701	0·289 – 1·171	541·99
South Africa: *Syzygium*	0·697	0·276 – 1·162	551·73
South Africa: *Eucalyptus*	0·663	0·254 – 1·137	592·74
South Africa: *Tibouchina*	0·659	0·251 – 1·137	585·93
Zambia: *Eucalyptus*	0·650	0·241 – 1·123	571·85
Mozambique	1·027	0·519 – 1·542	696·20
South Africa	0·706	0·289 – 1·169	543·56
Zambia	0·650	0·241 – 1·123	571·85
Madagascar ①	0·410	0·009 – 0·284	—
Malawi ②	0·064	0·002 – 0·171	1027·12
Syzygium	1·027	0·519 – 1·543	696·57
Eucalyptus	0·704	0·289 – 1·172	540·96
Tibouchina ③	0·659	0·251 – 1·137	585·93
Root Height ④	7·046	6·681 – 7·410	19212·80
C. austroafricana / *C. cubensis* divergence ④	1·449	1·075 – 1·821	3256·27

① Only one isolate from *Syzygium* in Madagascar was included. The t_{MRCA} and 95% HPD values for the divergence of this isolate from South African *Syzygium* isolates was obtained from the chronogram.

② Only two isolates from *Syzygium* and one isolate from *Eucalyptus* were used to represent the Malawi population. Due to limitations of the analysis technique, the ages of isolates from the different hosts in Malawi could not be reliably determined.

③ The *Tibouchina* collection of fungal isolates originated only in South Africa.

④ The root height of the tree, as well as the divergence between *C. austroafricana* and *C. cubensis*, were used as calibration points for the chronogram (Chapter 4, this thesis).

© University of Pretoria

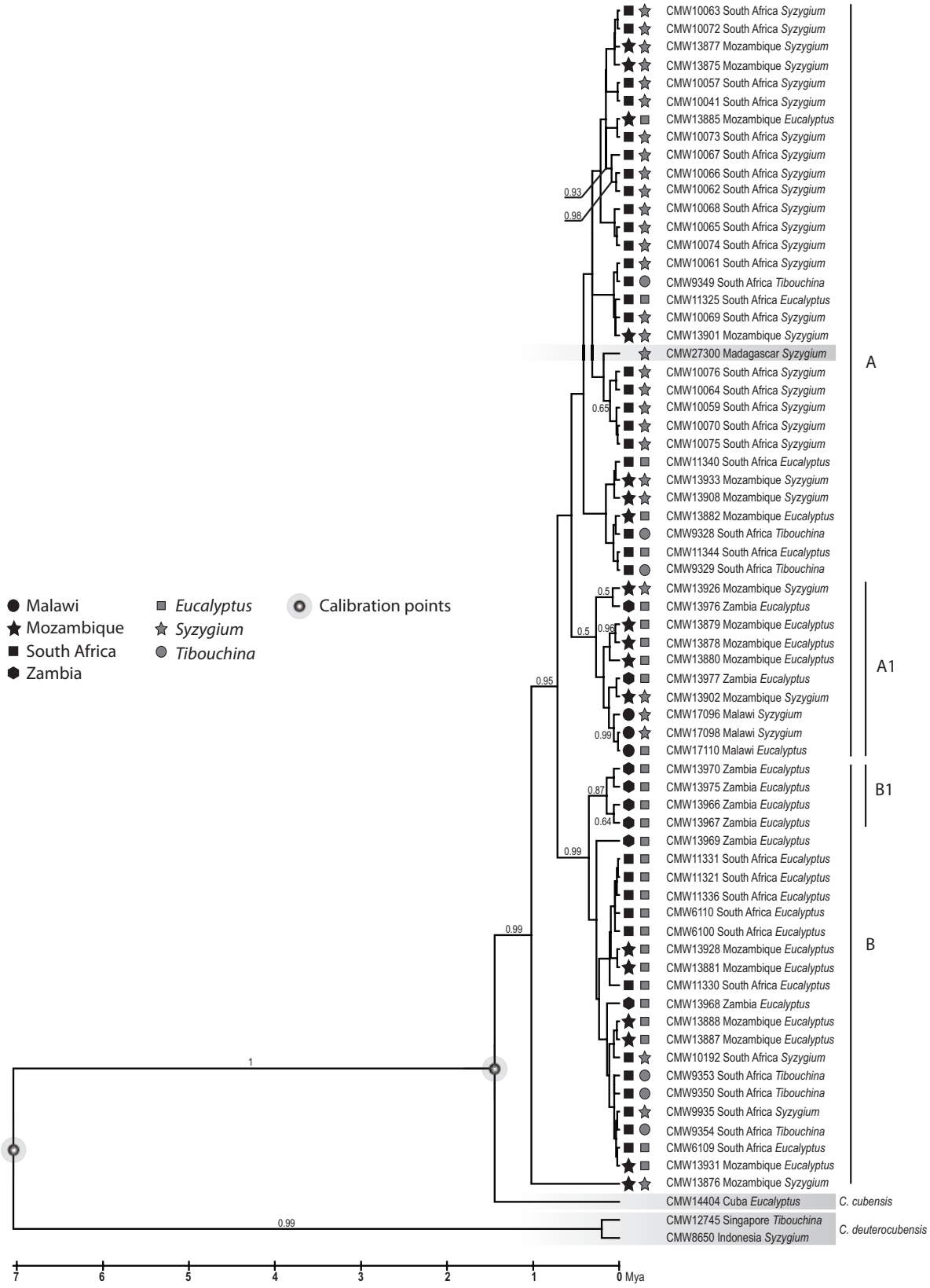
Table 3 Summary of pair wise population differentiation between populations of *C. austroafricana* defined based on country of origin, and host of origin.

Comparison	***N***	***θ***	***P***	t_{MRCA}
Mozambique vs. South Africa	161	0·127	< 0·001	3·44
Mozambique vs. Zambia	46	0·151	0·001	2·81
South Africa vs. Zambia	127	0·057	0·043	8·27
Mean		0·120	< 0·001	3·67
Eucalyptus vs. *Syzygium*	138	0·058	< 0·001	8·12
Eucalyptus vs. *Tibouchina*	95	0·306	< 0·001	1·13
Syzygium vs. *Tibouchina*	101	0·295	< 0·001	1·20
Mean		0·182	< 0·001	2·25

© University of Pretoria

Figure 1 Chronogram resulting from analysis with a lognormal molecular clock. Two calibration points were applied: one for the tree height (7·07 Mya) and one for the divergence between *C. cubensis* and *C. austroafricana* (1·39 Mya). Posterior probabilities are indicated above or below the relevant branches.

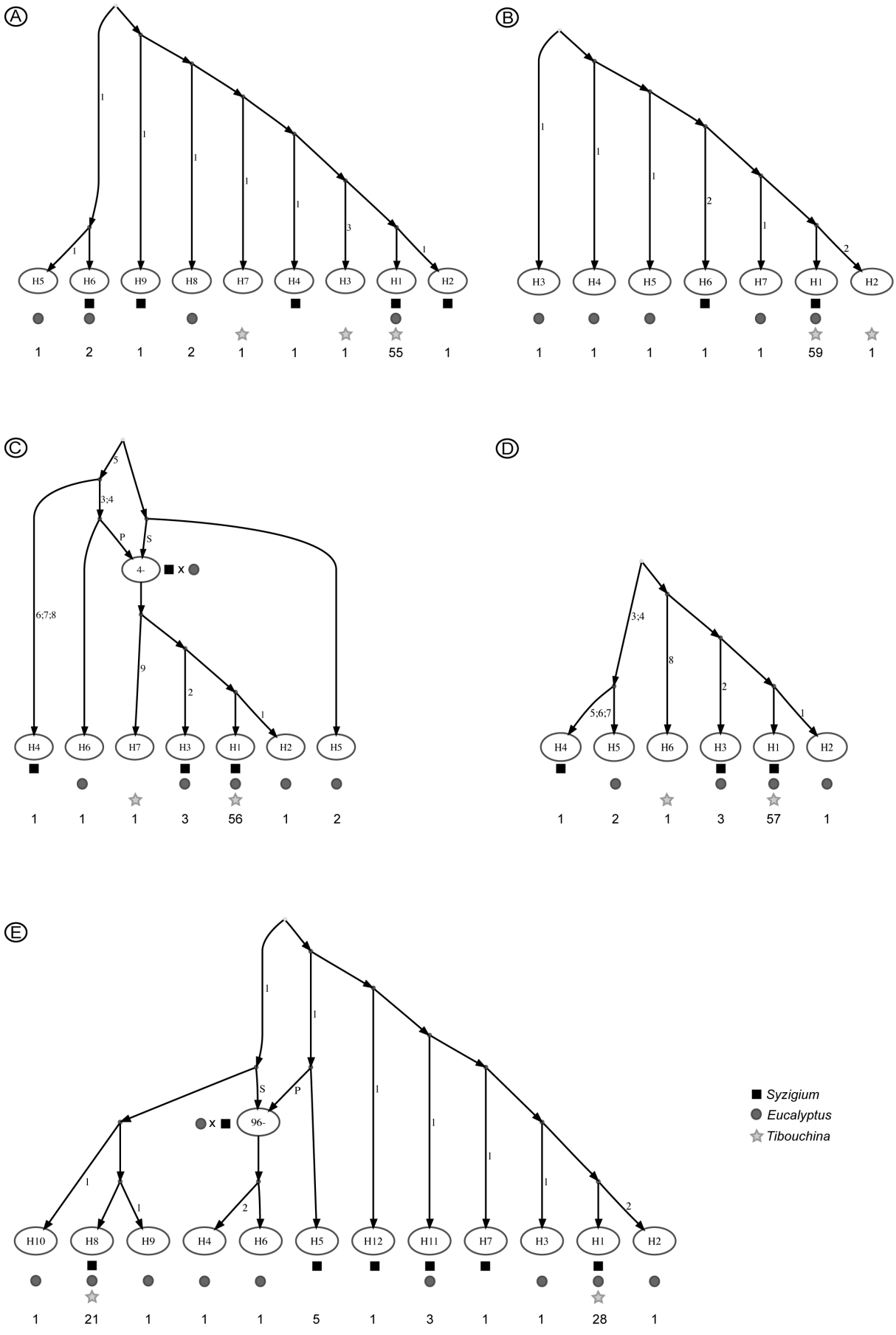
© University of Pretoria



© University of Pretoria

Figure 2 Ancestral recombination graphs (ARGs) for the (A) Actin, (B) ITS, (C) BT1 and (D) BT2 regions of the β-tubulin gene, and (E) Elongation Factor 1-α sequences. Haplotypes are indicated at the bottom of each ARG with an H followed by the number of the haplotype. Each haplotype is composed of isolates that originated on *Syzygium* (black square), *Eucalyptus* (dark grey circle) or *Tibouchina* (light grey star), while the number of isolates in each haplotype is indicated at the bottom of the ARG. Recombination events in the interior of ARGs are indicated with ovals containing the position number in the sequence alignment where a recombination event occurred. One parent contributed the prefix (P) while the other contributed the suffix (S) of each recombinant genotype, and the likely origin of each parent is indicated next to the recombination event. Numbers along the edges of ARGs indicate the number of mutations that are necessary for a lineage to coalesce with its most recent common ancestor.

© University of Pretoria



© University of Pretoria

Suppl Table 1 Tree statistics gathered from maximum likelihood analyses.

Statistic	ACT	BT1	BT2	EF	ITS	Combined
Number of sequences	68	68	69	69	69	74
Alignment length	235	467	401	327	479	1442
Best evolutionary model	GTR	GTR	GTR+I+G	GTR+I+G	GTR+I+G	GTR+I+G
-ln L	401·299	745·82	655·985	653·121	799·699	6518·819
fA	0·1722	0·1863	0·1643	0·2384	0·2265	0·1962
fC	0·3597	0·3471	0·3868	0·3237	0·2628	0·3426
fG	0·2787	0·2516	0·2744	0·1854	0·2403	0·2443
fT	0·1893	0·2151	0·1745	0·2525	0·2704	0·2169
p-inv	—	—	0·003	0·506	0·666	0·0
gamma	—	—	97·817	1·419	97·981	0·499

© University of Pretoria

Suppl Table 2 Pair wise population differentiation (θ, top right; $\hat{M}$, bottom left) between region/host-based populations of *C. austroafricana*.

		Mozambique		South Africa			Zambia
		Eucalyptus	*Syzygium*	*Eucalyptus*	*Syzygium*	*Tibouchina*	*Eucalyptus*
Mozambique	*Eucalyptus*	-	0·031 ($P = 0·046$)	0·188 ($P < 0·001$)	0·157 ($P < 0·001$)	0·365 ($P < 0·001$)	0·176 ($P < 0·001$)
	Syzygium	15·63	-	0·229 ($P < 0·001$)	0·172 ($P < 0·001$)	0·389 ($P < 0·001$)	0·171 ($P < 0·001$)
South Africa	*Eucalyptus*	2·16	1·68	-	0·112 ($P < 0·001$)	0·353 ($P < 0·001$)	0·102 ($P = 0·006$)
	Syzygium	2·69	2·41	3·96	-	0·311 ($P < 0·001$)	0·092 ($P = 0·012$)
	Tibouchina	0·87	0·79	0·92	1·01	-	0·363 ($P < 0·001$)
Zambia	*Eucalyptus*	2·34	2·42	4·40	4·94	0·88	-

© University of Pretoria

Bibliography

Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* **1**, 101-102.

Aldous DJ (2001) Stochastical models and descriptive statistics for phylogenetic trees, from Yule to today. *Statistical Science* **16**, 23-34.

Atallah ZK, Maruthachalam K, Du Toit L, Koike ST, Davis RM, Klosterman SJ, Hayes RJ, Subbarao KV (2010) Population analyses of the vascular plant pathogen *Verticillium dahliae* detect recombination and transcontinental gene flow. *Fungal Genetics and Biology* **47**, 416-422.

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.

Carbone I, Kohn LM (1999) A method for designing primer sets for speciation studies in filamentous ascomycetes. *Mycologia* **91**, 553-556.

Chungu D, Gryzenhout M, Muimba-Kankolongo A, Wingfield MJ, Roux J (2009) Taxonomy and pathogenicity of two novel *Chrysoporthe* species from *Eucalyptus grandis* and *Syzygium guineense* in Zambia. *Mycological Progress* **9**, 379-393.

Conti E, Litt A, Wilson PG, Graham SA, Briggs BG, Johnson LAS, Systma KJ (1997) Interfamilial relationships in Myrtales: molecular phylogeny and patterns of morphological evolution. *Systematic Botany* **22**, 629-647.

Cunningham CW (1997) Can three incongruence tests predict when data should be combined? *Molecular Biology and Evolution* **14**, 733-740.

Drummond AJ, Ho SYW, Phillips MJ, Rambaut A (2006) Relaxed phylogenetics and dating with confidence. *PLoS Biology* **4**, e88.

Drummond AJ, Rambaut A (2007) BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evolutionary Biology* **7**, 214.

Edgar RC (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797.

Estoup A, Jarne P, Cornuet JM (2002) Homoplasy and mutation model at microsatellite loci and their consequences for population genetics analysis. *Molecular Ecology* **11**, 1591-1604.

© University of Pretoria

Franke J, Gebhardt S, Menz G, Helfrich HP (2009) Geostatistical analysis of the spatiotemporal dynamics of powdery mildew and leaf rust in wheat. *Phytopathology* **99**, 974-984.

Galtier N, Gouy M, Gautier C (1996) SeaView and Phylo_Win: two graphic tools for sequence alignment and molecular phylogeny. *Computational and Applied BioScience* **12**, 543-548.

Glass NL, Donaldson GC (1995) Development of primer sets designed for use with the PCR to amplify conserved genes from filamentous ascomycetes. *Applied and Environmental Microbiology* **61**, 1323-1330.

Gryzenhout M, Rodas CA, Portales JM, Clegg P, Wingfield BD, Wingfield MJ (2006) Novel hosts of the *Eucalyptus* canker pathogen *Chrysoporthe cubensis* and a new *Chrysoporthe* species from Colombia. *Mycological Research* **110**, 833-845.

Gryzenhout M, Wingfield BD, Wingfield MJ (2009) *Taxonomy, phylogeny, and ecology of bark-infecting and tree killing fungi in the Cryphonectriaceae*. APS Press.

Hartley JL, Bowen H (1996) PEG precipitation for selective removal of small DNA fragments. *Focus Journal* **18**, 27.

Heath RN, Gryzenhout M, Roux J, Wingfield MJ (2006) Discovery of the *Cryphonectria* canker pathogen on native *Syzygium* species in South Africa. *Plant Disease* **90**, 433-438.

Huelsenbeck JP, Crandall KA (1997) Phylogeny estimation and hypothesis testing using maximum likelihood. *Annual Review of Ecological Systematics* **28**, 437-466.

Huelsenbeck JP, Rannala B (1997) Phylogenetic methods come of age: testing hypotheses in an evolutionary context. *Science* **276**, 227-232.

Jackson ST, Betancourt JL, Booth RK, Gray ST (2009) Ecology and the ratchet of events: climate variability, niche dimensions, and species distributions. *Proceedings of the National Academy of Science, USA* **106**, 19685-19692.

Jost L (2008) G_{ST} and its relatives do not measure differentiation. *Molecular Ecology* **17**, 4015-4026.

Jump AS, Hunt JM, Martínez-Izquierdo JA, Peñuelas J (2006) Natural selection and climate change: temperature-linked spatial and temporal trends in gene frequency in *Fagus sylvatica*. *Molecular Ecology* **15**, 3469-3480.

Linde CC, Zala M, McDonald BA (2009) Molecular evidence for recent founder populations and human-mediated migration in the barley scald pathogen *Rhynchosporium secalis*. *Molecular Phylogenetics and Evolution* **51**, 454-464.

© University of Pretoria

Lucas EJ, Belsham SR, Lughada EMN, Orlovich DA, Sakuragui CM, Chase MW, Wilson PG (2005) Phylogenetic patterns in the fleshy-fruited Myrtaceae - preliminary molecular evidence. *Plant Systematics and Evolution* **251**, 35-51.

McDonald BA, Linde C (2002) Pathogen population genetics, evolutionary potential and durable resistance. *Annual Review of Phytopathology* **40**, 349-379.

McDonald BA, Miles J, Nelson LR, Pettway RE (1994) Genetic variability in nuclear DNA in field populations of *Stagonospora nodorum*. *Phytopathology* **84**, 250-255.

Milgroom MG, Lipari SE (1995) Population differentiation in the chestnut blight fungus, *Cryphonectria parasitica*, in Eastern North America. *Phytopathology* **85**, 155-160.

Myburg H, Gryzenhout M, Heath R, Roux J, Wingfield BD, Wingfield MJ (1999) Cryphonectria canker on *Tibouchina* in South Africa. *Mycological Research* **106**, 1299-1306.

Nakabonge G, Gryzenhout M, Wingfield BD, Wingfield MJ, Roux J (2007) Genetic diversity of *Chrysoporthe cubensis* in eastern and southern Africa. *South African Journal of Science* **103**, 261-264.

Nakabonge G, Roux J, Gryzenhout M, Wingfield MJ (2006) Distribution of Chrysoporthe canker pathogens on *Eucalyptus* and *Syzygium* spp. in eastern and southern Africa. *Plant Disease* **90**, 734-740.

Nei M (1972) Genetic distance between populations. *The American Naturalist* **106**, 283-292.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Science, USA* **70**, 3321-3323.

Nei M, Li WH (1973) Linkage disequilibrium in subdivided populations. *Genetics* **75**, 213-219.

Nordborg M (2000) Linkage disequilibrium, gene trees and selfing: an ancestral recombination graph with partial self-fertilization. *Genetics* **154**, 923-929.

Old KM, Wingfield MJ, Yuan ZQ (2003) A manual of diseases of eucalypts in South-East Asia. Centre for International Forestry Research, Jakarta, Indonesia.

Pegg GS, Gryzenhout M, O'Dwyer C, Drenth A, Wingfield MJ (2010) The *Eucalyptus* canker pathogen *Chrysoporthe cubensis* discovered in eastern Australia. *Australasian Plant Pathology* **39**, 343-349.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**, 1253-1256.

Poynton RJ (1959) *Bulletin No. 38. Notes on exotic forest trees in South Africa*, 2nd edn. The Government Printer, Pretoria.

Price EW, Carbone I (2005) SNAP: workbench management tool for evolutionary population genetic analysis. *Bioinformatics* **21**, 402-404.

© University of Pretoria

Rodas CA, Gryzenhout M, Myburg H, Wingfield BD, Wingfield MJ (2005) Discovery of the *Eucalyptus* canker pathogen *Chrysoporthe cubensis* on native *Miconia* (*Melastomataceae*) in Colombia. *Plant Pathology* **54**, 460-470.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Steenkamp ET, Wingfield BD, Coutinho TA, Wingfield MJ, Marasas WFO (1999) Differentiation of *Fusarium subglutinans* f. sp. *pini* by histone gene sequence data. *Applied and Environmental Microbiology* **65**, 3401-3406.

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

Sytsma KJ, Litt A, Zjhra ML, Pires C, Nepokroeff M, Conti E, Walker J, Wilson PG (2004) Clades, clocks and continents: historical and biogeographical analysis of Myrtaceae, Vochysiaceae, and relatives in the Southern Hemisphere. *International Journal of Plant Science* **165**, S85-S105.

Tavaré S (1986) Some probabilistic and statistical problems in the analysis of DNA sequences. *American Mathematical Society: Lectures on Mathematics in the Life Sciences* **17**, 57-86.

Thrall PH, Burdon JJ (1997) Host-pathogen dynamics in a metapopulation context: the ecological and evolutionary consequences of being spatial. *Journal of Ecology* **85**, 743-753.

Vacher C, Vile D, Helion E, Piou D, Desprez-Loustau M-L (2008) Distribution of parasitic fungal species richness: influence of climate versus host species diversity. *Diversity and Distributions* **14**, 786-798.

Van der Merwe NA, Gryzenhout M, Steenkamp ET, Wingfield BD, Wingfield MJ (2010) Multigene phylogenetic and population differentiation data confirm the existence of a cryptic species within *Chrysoporthe cubensis*. *Fungal Biology* **114**, 966-979.

Van der Merwe NA, Wingfield BD, Wingfield MJ (2003) Primers for the amplification of sequence-characterized loci in *Cryphonectria cubensis* populations. *Molecular Ecology Notes* **3**, 494-497.

Van Heerden SW, Amerson HV, Preisig O, Wingfield BD, Wingfield MJ (2005) Relative pathogenicity of *Cryphonectria cubensis* on *Eucalyptus* clones differing in their resistance to *C. cubensis*. *Plant Disease* **89**, 659-662.

Van Heerden SW, Wingfield MJ (2001) Genetic diversity of *Cryphonectria cubensis* isolates in South Africa. *Mycological Research* **105**, 94-99.

Van Heerden SW, Wingfield MJ, Coutinho T, Van Zyl LM (1997) Population diversity among Venezuelan and Indonesian isolates of *Cryphonectria cubensis*. *South African Journal of Science* **93**, xiv.

© University of Pretoria

Van Zyl LM, Wingfield MJ, Alfenas AC, Crous PW (1998) Population diversity among Brazilian isolates of *Cryphonectria cubensis*. *Forest Ecology and Management* **112**, 41-47.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Weir BS, Hill WG (2002) Estimating F-statistics. *Annual Review of Genetics* **36**, 721-750.

White TJ, Bruns T, Lee S, Taylor J (1990) Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenies. In: *PCR protocols: a guide to methods and applications*. (eds. Innes MA, Gelfand DH, Sninsky SS, White TJ), pp. 315-322. Academic Press Inc., San Diego, California.

Wingfield MJ, Swart WJ, Abear B (1989) First record of Cryphonectria canker of *Eucalyptus* in South Africa. *Phytophylactica* **21**, 311-313.

Wingfield MJ, Van Zyl LM, Van Heerden S, Myburg H, Wingfield BD (1997) Virulence and the genetic composition of the *Cryphonectria cubensis* Bruner population in South Africa. In: *Physiology and Genetics of Tree-Phytophage Interactions* eds. Lieutier F, Mattson WJ, Wagner MR), pp. 163-172. INRA Editions.

Woolhouse MEJ, Haydon DT, Antia R (2005) Emerging pathogens: the epidemiology and evolution of species jumps. *TRENDS in Ecology and Evolution* **20**, 238-244.

Wright S (1951) The genetical structure of populations. *Annals of Eugenics* **15**, 323-354.

Zhou X, Burgess TI, De Beer ZW, Lieutier F, Yart A, Klepzig K, Carnegie A, Mena Portales J, Wingfield BD, Wingfield MJ (2007) High intercontinental migration rates and population admixture in the sapstain fungus *Ophiostoma ips*. *Molecular Ecology* **16**, 89-99.

© University of Pretoria

Chapter 6

Host switching between native and non-native trees in a population of the canker pathogen *Chrysoporthe cubensis* from Colombia

© University of Pretoria

Abstract

Chrysoporthe cubensis is a fungal pathogen of Myrtales in South and Central America, and causes severe stem cankers and often mortality in commercially propagated *Eucalyptus* trees. *Chrysoporthe cubensis* also occurs on native trees in South America, and these trees could represent the source of the pathogen. The purpose of this study was to test this hypothesis by investigating populations originating from two adjacent *Eucalyptus* (Myrtaceae) plantations in Colombia, and wild *Miconia rubiginosa* trees (Melastomataceae) growing alongside these stands. Polymorphic microsatellite markers were used to quantify allele sizes in 20 and 39 isolates from the two *Eucalyptus* stands and 32 isolates from adjacent *M. rubiginosa* trees. Gene and genotypic diversities were calculated from these data, and population differentiation and assignment tests were performed to ascertain whether the populations were genetically different. Results showed that there were no differences between any of the populations using these techniques, and that they can be treated as a single population. Therefore, the results support the hypothesis that host switching has occurred in *C. cubensis* in Colombia.

© University of Pretoria

Introduction

Chrysoporthe canker is an economically important stem disease of plantation-grown *Eucalyptus* species in the tropics and sub-tropics (Wingfield 1999, 2003). The disease is caused by three species of *Chrysoporthe* that have non-overlapping geographic ranges. These include *Chrysoporthe austroafricana* that occurs in Africa (Gryzenhout et al. 2004), *C. cubensis* in South and Central America (Hodges *et al*. 1979; Hodges & Reis 1974; Hodges *et al*. 1976; Van der Merwe *et al*. 2010; Van der Merwe *et al*. 2001), and *C. deuterocubensis* in Southeast Asia (Myburg et al. 2002; Pegg et al. 2010; Van der Merwe et al. 2010). In addition to causing similar symptoms on *Eucalyptus*, these species are closely related, morphologically very similar, and they cause similar symptoms on *Eucalyptus* (Myburg et al. 2002; Myburg et al. 1999; Van der Merwe et al. 2010; Van der Merwe et al. 2001). They can be differentiated by their host ranges in their native areas (Van der Merwe et al. 2010; Wingfield 1999, 2003), small differences in culture morphology (Gryzenhout et al. 2009), and fixed housekeeping gene DNA polymorphisms (Van der Merwe et al. 2010).

Chrysoporthe cubensis has been reported from Mexico, Colombia, Venezuela, Argentina and various other countries in South America (Gryzenhout et al. 2006; Gryzenhout et al. 2009; Van der Merwe et al. 2001; Van Heerden et al. 1997; Van Zyl et al. 1998; Wingfield et al. 2001). Some of the native hosts of this pathogen in South America include *Miconia rubiginosa* and *M. theaezans* (Myrtales: Myrtaceae) (Rodas et al. 2005; Van der Merwe et al. 2010), and several *Tibouchina* spp. (Myrtales: Melastomataceae) (Wingfield et al. 2001). Likewise, Seixas et al. (2004) found *C. cubensis* on *Tibouchina granulosa* in Brazil and performed an evaluation of the susceptibility of a variety of species in the Myrtales.

Previous population genetic studies on *C. cubensis* from South America have utilized vegetative compatibility groups (VCGs) (Van Heerden et al. 1997; Van Zyl et al. 1998). The VCG diversities of these populations were very high, and comparable to diversities observed for the closely related African species, *C. austroafricana* (Van Heerden & Wingfield 2001). These data were later interpreted as an indication that *C. cubensis* is native to South America, and this hypothesis was supported by the occurrence of the pathogen on hosts that are native to that region (Gryzenhout et al. 2009). Population diversities of *C. cubensis* from the different native hosts have not been assessed, but would be useful to confirm that the

© University of Pretoria

pathogen is native to South and Central America and they would inform efforts to breed for resistant *Eucalyptus* planting stock.

The goal of this study was to consider the population biology of *C. cubensis* obtained from native and non-native hosts in South America. For this purpose we focused on the isolates obtained from cankers on native *Miconia rubiginosa* (Melastomataceae) in Colombia (Rodas et al. 2005) and from cankers on commercially propagated *Eucalyptus* trees, which were growing alongside and in close proximity to these Melastomataceae. We employed polymorphic microsatellite markers to recover allele frequencies from *C. cubensis* populations isolated from both hosts, and tested for population differentiation and gene flow among the hosts. The assumption was that if population subdivision was low and gene flow between the hosts was high, the probability of host switching was also high.

Materials & Methods

Chrysoporthe cubensis fruiting structures, isolates and genotyping

Isolates of *C. cubensis* were obtained from two adjacent *E. grandis* × *E. urophylla* compartments on Vanessa farm, near Timba in Colombia, and from cankers on *Miconia rubiginosa* (a woody shrub) growing in close proximity to the *Eucalyptus* trees. On cankers from *Eucalyptus*, abundant perithecia were observed, while pycnidia were observed only on cankers where no perithecia were found. Only pycnidia were found on the cankers from *Miconia*. In total, 59 *Eucalyptus* trees were sampled from adjacent stands, i.e. 20 from Stand A and 39 from Stand B, while 32 *Miconia* plants occurring in and around the *Eucalyptus* stands were sampled.

Isolations were performed by sectioning through the sexual and asexual fruiting structures embedded in the bark tissue, using a sterile scalpel. The pale yellow spore masses from single perithecia or pycnidia were lifted from the fruiting structures using a sterile surgical needle, and suspended in 1 ml sterile water. A ten-times dilution was made from the original suspension, and 100 µl was spread over the surface of 20% w/v malt extract agar (Merck) in Petri dishes. The plates were incubated at 25 °C in the dark for 6-8 hours, after which they were examined using a dissection microscope. Single germinating ascospores or conidia were

© University of Pretoria

cut from the medium using a sterile surgical needle, transferred to 20% w/v potato dextrose agar (Difco Laboratories), and allowed to grow at 25 °C in the dark for 1 wk.

For the purpose of population genetic analyses, a fruiting structure was sampled from individual plants, and single ascospore or single conidium cultures were made. This resulted in 59 single ascospore or single conidium isolates from *Eucalyptus* and 32 single conidium isolates from *Miconia*. Each of these isolates was inoculated into 800 µl 20% w/v malt extract broth (Merck) in micro-centrifuge tubes and allowed to grow for 1 wk in the dark at 25 °C, with daily manual shaking to allow aeration. Mycelium was harvested by centrifugation, after which a modification of the CTAB (hexadecyltrimethylammonium bromide) DNA extraction protocol (Steenkamp et al. 1999) was used to extract total genomic DNA from each isolate. Twelve microsatellite loci were PCR-amplified using ten labeled primer pairs previously developed by Van der Merwe et al. (2010). Amplification products were subjected to genotyping using an ABI™ Prism© 3100 automated DNA sequencer (Applied Biosystems, USA), and allele sizes for each locus were determined using the Genotyper computer software package.

Population subdivision and sexual reproduction

Microsatellite loci having a >50% incidence of null alleles, or being monomorphic throughout the sample of fungal isolates, were removed prior to further analyses. Additionally, all remaining null alleles were treated as missing data in population genetic analyses.

A test for population subdivision was performed using Structure 2·3·1 (Pritchard et al. 2000). Five analyses, with priors of K ranging from K=1 to K=5 and ten million MCMC steps, were performed to determine the optimal number of subpopulations based on the posterior likelihood values. These analyses were repeated in order to minimize the effect of the random seed number. The level of uniqueness of each subpopulation was calculated with the formula

$$\varphi = \frac{\rho_S}{N_S} \bullet \frac{N_M}{\alpha_M}$$

where φ is the level of uniqueness of the subpopulation, ρ_S is the number of private alleles in the subpopulation with N_S individuals, and α_M is the total number of alleles over all loci in

© University of Pretoria

the metapopulation with N_M individuals. Null alleles were treated as missing data in this calculation, in order to exclude the possibility of false negative results.

Isolates were divided into three sub-populations: two sub-populations from *Eucalyptus* (Stand A and Stand B sub-populations), and one sub-population from *M. rubiginosa*. Gene diversity ($\overline{H}$) (Nei 1973) and genotypic diversity (G) (Stoddart & Taylor 1988) was calculated for each of the three sub-populations. The maximum likelihood estimator of genotypic diversity ($\hat{G}$) was calculated by normalizing G with the sample size. Genetic distance (Nei 1972) between the two sub-populations from *Eucalyptus*, and the combined sub-populations from *Eucalyptus* vs. the sub-population from *M. rubiginosa* was calculated using the computer software MultiLocus 1.3b (Agapow & Burt 2001) and expressed as Weir and Cockerham's θ (Weir & Cockerham 1984), which is a variation on Wright's F_{ST} (Agapow & Burt 2001; Weir & Cockerham 1984; Weir & Hill 2002). This measure gives an indication of population differentiation and ranges between 0 (low level of differentiation) and 1 (high level of differentiation). The theoretical number of migrants ($\hat{M}$) was estimated from the θ values (Slatkin 1995) in order to yield an indication of the level of gene flow and migration between the sub-populations, and between the two hosts.

In order to test for gametic disequilibrium, i.e., the inferred level of outcrossing, the Index of Association (I_A) (Agapow & Burt 2001) for each of the three subpopulations was estimated using MultiLocus. The measure was estimated using one million randomizations of each of the two data sets, normalization to yield $\bar{r}_d$, and comparison of the randomized distributions to the observed level of gametic disequilibrium. Randomization provides a confidence level to aid in acceptance or rejection of the null hypothesis of random mating and gametic equilibrium, while $\bar{r}_d$ enables comparisons of gametic disequilibrium between different studies.

Phenotypic diversity of recombinant progeny

To determine whether the sexual fruiting structures on *Eucalyptus* arose as a result of self-fertilization or outcrossing, one perithecium was sampled from each of 30 trees originating from *Eucalyptus* Stand B. From each of these perithecia, 12 single ascospore isolates were prepared as described above. These isolates were subjected to vegetative compatibility group (VCG) and inter-short sequence repeat PCR reaction (ISSR) assays. For VCG assays, the

© University of Pretoria

twelve isolates from a perithecium were paired against each other in all possible combinations, resulting in 144 pairings per perithecium. Isolates were placed on PDA-BCG plates (20% w/v potato dextrose agar, 50 mg/L bromocresol green) (Powell 1995) in pair-wise combinations, mycelium side down and approximately 3 cm apart and allowed to grow at 25 ºC for 10 days in the dark. Mycelial incompatibility interactions were evaluated by the identification of a dark line and yellow discoloration of the medium at the zone of contact between two isolates.

For the ISSR assays, genomic DNA was extracted from each of the 360 single ascospore isolates as described above. These DNA samples were used as templates in ISSR reactions with primers BDB$(ACA)_5$, DHB$(CGA)_5$ and HV$(GT)_8$ (Hantula et al. 1996), in order to quantify the number of genotypes within each perithecium. Standard PCR reactions were performed with annealing at 47 °C (BDB$[ACA]_5$), 59 °C (DHB$[CGA]_5$) or 56 °C (HV$[GT]_8$), and 30 cycles (Van der Merwe et al. 2003). Ten µl of each PCR reaction was electrophoresed on a 1% w/v agarose gel and amplicons were visualized using ethidium bromide and ultraviolet light (Sambrook et al. 1989).

To confirm the homothallic nature of *C. cubensis* and that self-fertilization can occur (Hodges et al. 1979), a single ascospore isolate from *Eucalyptus* Stand B was artificially inoculated onto *Eucalyptus* twigs. Artificial wounds (5 mm^2) were made with a sterile scalpel blade on five fresh twigs each of *E. grandis* and hybrid *Eucalyptus grandis* × *Eucalyptus camaldulensis*, respectively. The twigs were approximately 7 mm in diameter, and the bark was 2 mm thick. Twigs were surface sterilized with 70% (v/v) ethanol, and the ends were sealed with molten candle wax. An MEA plug was cut from a fresh three-day-old culture and placed into the wound, with the mycelium side facing the cambium. Wounds were covered with masking tape, and each twig was placed in a clean Petri dish lined with moist filter paper. Petri dishes were incubated at 25 °C for 6 wk in the dark. Once perithecia were observed, two fruiting structures were sampled and 12 single ascospore isolates were made from each. These isolates were included in the ISSR assays.

© University of Pretoria

Results

Chrysoporthe cubensis fruiting structures, isolates and genotyping

In total, 29 single ascospore isolates were collected from *Eucalyptus* trees in Stand A, 30 single ascospore isolates from trees in the nearby Stand B, and 32 single conidium isolates from *M. rubiginosa* plants growing in close proximity to these trees. Nine of the ten microsatellite loci (Van der Merwe et al. 2010) were polymorphic, and the number of alleles per locus, excluding null alleles, varied from two for locus *CcPMC* to 11 for locus *SA4*. Based on these nine polymorphic loci, 56 multilocus genotypes were identified among the isolates of *C. cubensis* collected from *Eucalyptus*. These included 20 from Stand A and 37 from Stand B. Isolates from Stands A and B shared a single genotype. The isolates from 32 pycnidia from cankers on *M. rubiginosa* represented 26 multilocus genotypes, and none of these were coincident with those recovered from *Eucalyptus*.

Population subdivision and sexual reproduction

Structure analyses revealed that *C. cubensis* isolates from *Eucalyptus* and *M. rubiginosa* had the same overall allelic compositions and represented a single population, i.e. K=1 (Table 1). The levels of uniqueness (φ) for the three sub-populations were 0·2486, 0·5304 and 0·17 for those from *M. rubiginosa* and *Eucalyptus* Stand A and Stand B, respectively (Table 2). The population from *Eucalyptus* Stand A had the highest levels of genotypic diversity, uniqueness, gene diversity, and private alleles (Table 2, Suppl Table 1). Conversely, the population from *Eucalyptus* Stand B had the lowest levels of genotypic diversity, uniqueness and private alleles, while its gene diversity was comparable to that of the population from *M. rubiginosa*.

Population differentiation tests showed the two sub-populations from *Eucalyptus* to be the least differentiated ($\theta = 0.026$) with the highest number of migrants ($\hat{M} = 19.05$). The sub-population from *Miconia* was more differentiated from *Eucalyptus* Stand B ($\theta = 0.194; \hat{M} = 2.08$) than from Stand A ($\theta = 0.112; \hat{M} = 3.96$). All three populations were in gametic disequilibrium ($P < 0.001$) (Figure 1).

© University of Pretoria

Phenotypic diversity of recombinant progeny

Among the 30 perithecia examined, 28 appeared to be the result of self-fertilization. These perithecia produced progeny representing single vegetative compatibility groups (VCG), while progeny representing multiple distinct VCGs were detected for only two of the perithecia (Table 3). Similarly, multiple and distinct ISSR profiles were generated for the isolates from each of these two perithecia, while those originating from the 28 other perithecia each represented single ISSR profiles. Furthermore, inoculation of a single ascospore isolate onto twigs of *E. grandis* gave rise to perithecia that produced progeny representing a single VCG and ISSR profile (Table). However, no fruiting structures were observed on hybrid *E. grandis* × *E. camaldulensis* twigs after eight weeks.

Discussion

In this study we showed that *C. cubensis* is capable of switching between non-native *Eucalyptus grandis* × *E. urophylla* and native *M. rubiginosa* in Colombia. Host switching is important in the evolution and epidemiology of plant pathogens, and is thought to occur frequently (Slippers et al. 2005; Wingfield 2003; Woolhouse et al. 2005). Such host switching events could greatly contribute to the ability of a pathogen to become epidemic, and are thought to be primarily associated with a change in genomic sequence at one or more loci (Woolhouse et al. 2005).

Self-fertilization and outcrossing are expected to have markedly different effects on the population biology of a fungus. In this study, we confirmed the homothallic nature of *C. cubensis* by showing that phenotypic characters do not segregate in homothallic progeny. Tests between single ascospore isolates from 30 perithecia collected from *Eucalyptus* revealed that only two perithecia contained more than one vegetative compatibility group. This result would have been expected if most of the sexual events were the result of self-fertilization in homothallic individuals, but not if outcrossing was the predominant mode of reproduction (c.f. Milgroom et al. 2008). However, the fact that two perithecia contained genetically different progeny provides evidence that different parental individuals were involved during fertilization. Such outcrossing events may lead to the generation of a large number of distinct genotypes. Indeed, the number of genotypes sampled from the populations

© University of Pretoria

was very large (Table 2), suggesting that infrequent outcrossing may be sufficient to result in a large genotypic diversity.

Clonality of the *C. cubensis* populations in Colombia supports the notion that sexual reproduction of the fungus in this region is facilitated primarily by self-fertilization. All three *C. cubensis* populations were in gametic disequilibrium ($P < 0.0001$), which is a widespread phenomenon in fungi that can reproduce asexually via conidia, or have mixed mating systems (Milgroom et al. 2008). Such high levels of allelic association among loci can also be linked to migration and the establishment of diversity deficient founder populations in new areas or on new hosts (Linde et al. 2009). The high levels of gametic disequilibrium and limited genotypic diversity, in lieu of non-differentiating allelic compositions in the Colombian *C. cubensis* populations, thus suggest that only a subset of isolates were able to switch between non-native *Eucalyptus* and native *M. rubiginosa*.

A sub-population of isolates from one of the *Eucalyptus* stands displayed greater gene diversity than the other two sub-populations, suggesting that a reduction in gene diversity may be associated with spread and host switching. Even though this may be a sampling artifact, the largest gene and genotypic diversities were nonetheless observed in a *C. cubensis* population from *Eucalyptus* Stand A, while the same statistic for a population from the other stand as well as from *Miconia* were lower. These results could be interpreted in two different ways. First, they could indicate that the largest number of sexual outcrossing events of *C. cubensis* occurs in *Eucalyptus* Stand A, resulting in a relatively more diverse population. This population could then have invaded Stand B, as well as nearby *Miconia* plants, and these notions are supported by gene flow statistics. The second interpretation is that the reduction in genotypic diversity is due to an invasion taking place in a certain direction, namely from an external source which was not sampled, to *Eucalyptus* Stand A, to Stand B and to *M. rubiginosa*. Such invasion-related reductions in diversity are well recognized in other invasive plant pathogens such as *Verticillium dahliae* (Atallah et al. 2010) and *Cryphonectria parasitica* (Dutech et al. 2008; Yan et al. 2007).

The original host of *C. cubensis* in Colombia probably resides in the native Melastomataceae. Rodas et al. (2005) showed that a *C. cubensis* isolate from *M. rubiginosa* in Colombia was more pathogenic on *Tibouchina* spp. (Melastomataceae) than on either *M. rubiginosa* or

© University of Pretoria

Eucalyptus, while another isolate from *M.rubiginosa* was less pathogenic on *Eucalyptus* than on *M. rubiginosa*. In the current study, gene diversity values observed in all three sub-populations were comparable, and there was no statistically significant subdivision between populations from either host. If *C. cubensis* had jumped from *M. rubiginosa* to *Eucalyptus*, the gene diversity should have decreased in that direction (Woolhouse et al. 2005), which is in contrast to our results. Also, a sexually reproducing Ascomycete is likely to complete all stages of its life cycle on the host on which it evolved. However, sexual fruiting structures of *C. cubensis* were not observed on *M. rubiginosa* and all of the populations were in gametic disequilibrium. This suggests that both hosts (*Eucalyptus* and *M. rubiginosa*) are probably being opportunistically infected by the fungus from a yet to be sampled population on a native Melastomataceae host.

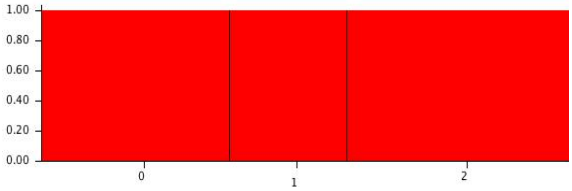
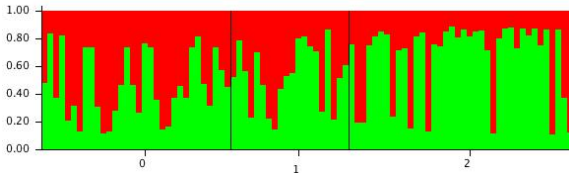
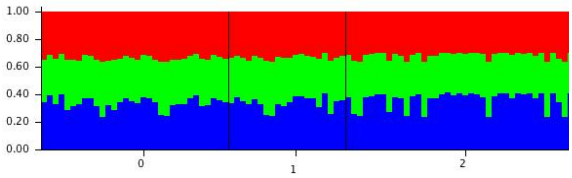
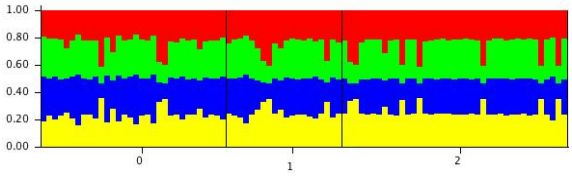
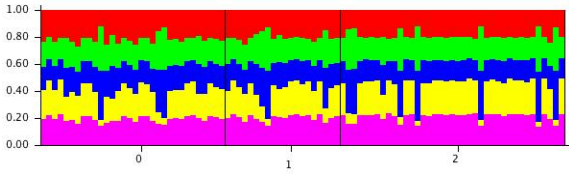
The question remains as to why sexual outcrossing appears to be present in some *C. cubensis* populations, while gametic equilibrium is absent. Based on the Hardy-Weinberg equilibrium principle, one generation of random mating can homogenize the population (Hardy 1908). Out of 30 perithecia, we found only two that were the result of outcrossing, suggesting that mating in these populations is not random. However, if we consider the possibility that numerous rounds of self-fertilization produced the previous generation in these populations, it is likely that those genotypes would have been representative of the founder population, which was in gametic disequilibrium after a host switch. A subsequent environmental change could have initiated favorable conditions for rare outcrossing events between closely related individuals, giving rise to the sampled generations of the pathogen. Of these recombinant genotypes, some would have persisted on the original host, while others could have acquired the ability to also infect related plant species as has been shown for other plant pathogens (De Vienne et al. 2009). Therefore, occasional outcrossing appears to be an economical strategy for *C. cubensis* to expedite the invasion of different host genotypes, or different hosts, without losing infectiousness towards the formative host.

Acknowledgements

Financial support from Members of the Tree Protection Co-operative Programme (TPCP), the National Research Foundation and the Department of Science and Technology (DST) and the NRF Centre of Excellence in Tree Health Biotechnology (CTHB) made it possible to conduct this research.

© University of Pretoria

Table 1 Results from Structure analyses of microsatellite alleles from all included isolates with a range of *K* (number of populations) priors in order to detect population substructure.

Bar plot ①	*K*	lnPr(*X* | *K*)	Mean of lnPr(*X* | *K*)	Variance of lnPr(*X* | *K*)
	1	-525·1	-516·5	17·0
	2	-514·9	-449·7	130·5
	3	-533·3	-412·3	242·0
	4	-517·3	-386·4	261·9
	5	-545·5	-373·1	344·8

① Bar plots display individuals from *Miconia* as population 0, while those from the two *Eucalyptus* plantations are shown as populations 1 (Stand A) and 2 (Stand B).

© University of Pretoria

Table 2 Population genetic statistics for the three subpopulations from *Miconia rubiginosa* and *Eucalyptus grandis* × *Eucalyptus urophylla* stands A and B.

Statistic	***M. rubiginosa***	***Eucalyptus* Stand A**	***Eucalyptus* Stand B**
Number of isolates	32	20	39
Number of genotypes	26	20	37
Genotypic diversity, $\hat{G}$	31.37%	100%	28.47%
Number of alleles ①	29	31	30
Private alleles	6	8	5
Uniqueness, φ	0·2486	0·5304	0·1700
Gene diversity	0·3983	0·4450	0·3990
Gametic equilibrium	No	No	No

① Total number of alleles over all loci for each subpopulation. The total number of alleles for the metapopulation was 46.

© University of Pretoria

Table 3 Multilocus haplotypes and VCG phenotypes of *C. cubensis* isolates from sexual fruiting structures on cankers in *Eucalyptus* Stand B.

Perithecium Number	Number of multilocus haplotypes		Average Clonal Fraction (*Cf*) ①	Number of VCG phenotypes ②
	Primer $(ACA)_5$	Primer $(GT)_8$		
Per7	6	4	75%	3
Per15	6	5	67%	6
Per1-6,8-14,16-30 ③	1	1	100%	1
Self1	1	1	100%	1
Self2	1	1	100%	1

① $Cf = \frac{[N-(C-1)]}{N}$ where N is the sample size and C is the number of distinct multilocus haplotypes within each perithecium. Clonal fractions were averaged between the two primers. A clonal fraction of 100% indicates that all isolates within the sample had the same multilocus haplotype.

② For Per7, two VCGs consisted of three isolates each, while one VCG consisted of seven isolates. For Per15, three VCGs consisted of a single isolate each, one of two isolates, one of three isolates and one VCG consisted of four isolates.

③ From 28 of the 30 perithecia, a single multilocus haplotype and a single VCG phenotype was recovered per perithecium.

© University of Pretoria

Suppl Table 1 Calculation of gene diversity and the levels of uniqueness of each subpopulation. Private alleles are indicated in bold.

Locus	Statistic	Allele	*f* ①	*f* ②	*f* ③
CcPMB		**213**	-	**0.0625**	-
		214	1.0000	0.9375	1.0000
	H ④		0	0.1172	0
CcPMC		**193**	-	**0.1429**	-
		196	0.1538	0.2143	0.2857
		197	0.6923	0.2143	0.1905
		203	-	0.1429	0.1905
		212	0.1539	0.2856	0.3333
	H		0.4734	0.7857	0.7347
COL11		260	0.7931	0.8000	0.6176
		261	0.1724	0.2000	0.3824
		262	**0.0345**	-	-
	H		0.3401	0.3200	0.4723
SA1		316	0.8333	1.0000	0.9723
		317	0.1667	-	0.0277
	H		0.2778	0	0.0539
SA3		**202**	-	-	**0.0323**
		203	-	0.6364	0.8709
		214	1.0000	0.3636	0.0968
	H		0	0.4630	0.2311
SA4		**152**	-	-	**0.0526**
		153	0.1000	0.1667	0.1842
		156	**0.0333**	-	-
		162	**0.0333**	-	-
		165	0.4333	0.2778	0.6316
		168	**0.1333**	-	-
		172	-	**0.0556**	-
		173	0.0333	0.1667	-
		174	0.2000	0.1667	0.0526
		185	-	**0.0556**	-
		193	0.0335	0.1109	0.0790
	H		0.7400	0.8210	0.5554
SA6-1		316	-	0.7000	0.8948
		318	-	-	**0.0526**
		319	**0.5000**	-	-
		348	-	**0.1000**	-
		350	**0.5000**	-	-
		355	-	**0.1000**	-
		362	-	**0.1000**	-
		365	-	-	**0.0526**
	H		0.5000	0.4800	0.1938
SA6-2		209	0.0476	-	0.2858
		210	0.5238	0.7500	0.2381
		211	0.0476	0.1250	0.1429
		213	0.1429	-	0.0476
		214	0.2381	0.1250	0.1905
		217	-	-	**0.0951**
	H		0.6440	0.4063	0.7936
SA9-2		**204**	-	**0.0714**	-
		211	0.1875	0.2143	0.2500

© University of Pretoria

	212	-	0.0714	0.0357
	213	0.0625	-	0.1071
	214	0.1875	0.5715	0.6072
	215	0.5625	0.0714	-
$\overline{H}$		0.6094	0.6122	0.5561
$\overline{H}$		0.3983	0.4450	0.3990
φ		0.2486	0.5304	0.1700

① Frequency in the population from *Miconia rubiginosa*

② Frequency in the population from *Eucalyptus* Stand A

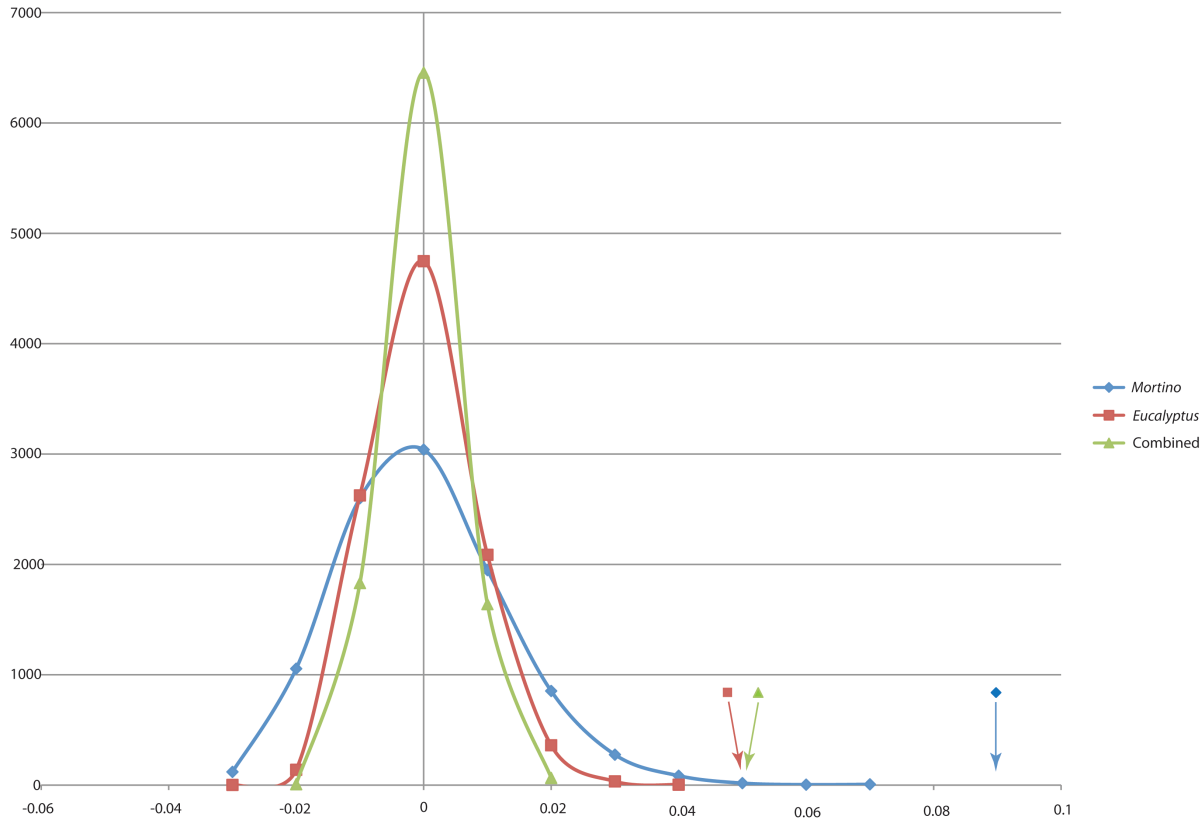
③ Frequency in the population from *Eucalyptus* Stand B

④ $H = 1 - \sum_k x_k^2$ where H is the gene diversity and x_k is the frequency of the k^{th} allele

© University of Pretoria

Figure 1 Graph of $\bar{r}_d$ values resulting from 10000 randomizations of each of the *C. cubensis* populations from *Miconia* (diamonds) and *Eucalyptus* (squares) in Colombia, as well as the two populations combined (triangles). The graph indicates the number of times (y-axis) that each of the $\bar{r}_d$ categories (x-axis) was observed. Arrows indicate the observed $\bar{r}_d$ values for each of the data sets ($P < 0{\cdot}0001$), indicating that the null hypothesis of random association of alleles can be rejected. Therefore, the populations are in gametic disequilibrium.

© University of Pretoria



© University of Pretoria

Bibliography

Agapow PM, Burt A (2001) Indices of multilocus linkage disequilibrium. *Molecular Ecology Notes* **1**, 101-102.

Atallah ZK, Maruthachalam K, Du Toit L, Koike ST, Davis RM, Klosterman SJ, Hayes RJ, Subbarao KV (2010) Population analyses of the vascular plant pathogen *Verticillium dahliae* detect recombination and transcontinental gene flow. *Fungal Genetics and Biology* **47**, 416-422.

De Vienne DM, Hood ME, Giraud T (2009) Phylogenetic determinants of potential host shifts in fungal pathogens. *Journal of Evolutionary Biology* **22**, 2532-2541.

Dutech C, Rossi J-P, Fabreguettes O, Robin C (2008) Geostatistical genetic analysis for inferring the dispersal pattern of a partially clonal species: example of the chestnut blight fungus. *Molecular Ecology* **17**, 4597-4607.

Gryzenhout M, Myburg H, Van der Merwe NA, Wingfield BD, Wingfield MJ (2004) *Chrysoporthe*, a new genus to accommodate *Cryphonectria cubensis*. *Studies In Mycology* **50**, 119-142.

Gryzenhout M, Rodas CA, Portales JM, Clegg P, Wingfield BD, Wingfield MJ (2006) Novel hosts of the *Eucalyptus* canker pathogen *Chrysoporthe cubensis* and a new *Chrysoporthe* species from Colombia. *Mycological Research* **110**, 833-845.

Gryzenhout M, Wingfield BD, Wingfield MJ (2009) *Taxonomy, phylogeny, and ecology of bark-infecting and tree killing fungi in the Cryphonectriaceae.* APS Press.

Hantula M, Dusabenyagasani M, Hamelin RC (1996) Random amplified microsatellites (RAMS) - a novel method for characterizing genetic variation within fungi. *European Journal of Forest Pathology* **26**, 159-166.

Hardy GH (1908) Mendelian proportions in a mixed population. *Science* **28**, 49-50.

Hodges CS, Geary TF, Cordell CE (1979) The occurrence of *Diaporthe cubensis* on *Eucalyptus* in Florida, Hawaii, and Puerto Rico. *Plant Disease Reporter* **63**, 216-220.

Hodges CS, Reis MS (1974) Identification do fungo causador de cancro de *Eucalyptus* spp. no Brazil. *Brazil Florestal* **5**, 19.

Hodges CS, Reis MS, Ferreira FA, Henfling JDM (1976) O cancro do eucalipto causado por *Diaporthe cubensis*. *Fitopatalogia Basileira* **1**, 129-167.

© University of Pretoria

Linde CC, Zala M, McDonald BA (2009) Molecular evidence for recent founder populations and human-mediated migration in the barley scald pathogen *Rhynchosporium secalis*. *Molecular Phylogenetics and Evolution* **51**, 454-464.

Milgroom MG, Sotirovski K, Spica D, Davis JE, Brewer MT, Milev M, Cortesi P (2008) Clonal population structure of the chestnut blight fungus in expanding ranges in southeastern Europe. *Molecular Ecology* **17**, 4446-4458.

Myburg H, Gryzenhout M, Wingfield BD, Wingfield MJ (2002) ß-tubulin and histone *H3* gene sequences distinguish between *Cryphonectria cubensis* from South Africa, Asia, and South America. *Canadian Journal of Botany* **80**, 590-596.

Myburg H, Wingfield BD, Wingfield MJ (1999) Phylogeny of *Cryphonectria cubensis* and allied species inferred from DNA analysis. *Mycologia* **91**, 243-250.

Nei M (1972) Genetic distance between populations. *The American Naturalist* **106**, 283-292.

Nei M (1973) Analysis of gene diversity in subdivided populations. *Proceedings of the National Academy of Science, USA* **70**, 3321-3323.

Pegg GS, Gryzenhout M, O'Dwyer C, Drenth A, Wingfield MJ (2010) The *Eucalyptus* canker pathogen *Chrysoporthe cubensis* discovered in eastern Australia. *Australasian Plant Pathology* **39**, 343-349.

Powell WA (1995) Vegetative incompatibility and mycelial death of *Cryphonectria parasitica* detected with a pH indicator. *Mycologia* **87**, 738-741.

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

Rodas CA, Gryzenhout M, Myburg H, Wingfield BD, Wingfield MJ (2005) Discovery of the *Eucalyptus* canker pathogen *Chrysoporthe cubensis* on native *Miconia* (*Melastomataceae*) in Colombia. *Plant Pathology* **54**, 460-470.

Sambrook J, Fritch EF, Maniatis T (1989) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY, USA.

Seixas CDS, Barreto RW, Alfenas AC, Ferreira FA (2004) *Cryphonectria cubensis* on an indigenous host in Brazil: a possible origin for eucalyptus canker disease? *Mycologist* **18**, 39-45.

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. *Genetics* **139**, 457-462.

Slippers B, Stenlid J, Wingfield MJ (2005) Emerging pathogens: fungal host jumps following anthropogenic introduction. *TRENDS in Ecology and Evolution* **20**, 420-421.

© University of Pretoria

Steenkamp ET, Wingfield BD, Coutinho TA, Wingfield MJ, Marasas WFO (1999) Differentiation of *Fusarium subglutinans* f. sp. *pini* by histone gene sequence data. *Applied and Environmental Microbiology* **65**, 3401-3406.

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

Van der Merwe NA, Gryzenhout M, Steenkamp ET, Wingfield BD, Wingfield MJ (2010) Multigene phylogenetic and population differentiation data confirm the existence of a cryptic species within *Chrysoporthe cubensis*. *Fungal Biology* **114**, 966-979.

Van der Merwe NA, Myburg H, Wingfield BD, Wingfield MJ (2001) Identification of *Cryphonectria cubensis* from Colombia based on rDNA sequence data. *South African Journal of Science* **97**, 295-296.

Van der Merwe NA, Wingfield BD, Wingfield MJ (2003) Primers for the amplification of sequence-characterized loci in *Cryphonectria cubensis* populations. *Molecular Ecology Notes* **3**, 494-497.

Van Heerden SW, Wingfield MJ (2001) Genetic diversity of *Cryphonectria cubensis* isolates in South Africa. *Mycological Research* **105**, 94-99.

Van Heerden SW, Wingfield MJ, Coutinho T, Van Zyl LM (1997) Population diversity among Venezuelan and Indonesian isolates of *Cryphonectria cubensis*. *South African Journal of Science* **93**, xiv.

Van Zyl LM, Wingfield MJ, Alfenas AC, Crous PW (1998) Population diversity among Brazilian isolates of *Cryphonectria cubensis*. *Forest Ecology and Management* **112**, 41-47.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358-1370.

Weir BS, Hill WG (2002) Estimating F-statistics. *Annual Review of Genetics* **36**, 721-750.

Wingfield MJ (1999) Pathogens in exotic plantation forestry. *International Forestry Review* **1**, 163-168.

Wingfield MJ (2003) Increasing threat of diseases to exotic plantation forests in the Southern Hemisphere: lessons from Cryphonectria canker. *Australasian Plant Pathology* **32**, 133-139.

Wingfield MJ, Rodas C, Myburg H, Venter M, Wright J, Wingfield BD (2001) Cryphonectria canker on *Tibouchina* in Colombia. *Forest Pathology* **31**, 1-10.

Woolhouse MEJ, Haydon DT, Antia R (2005) Emerging pathogens: the epidemiology and evolution of species jumps. *TRENDS in Ecology and Evolution* **20**, 238-244.

Yan B, Li Z, Huang H, Qin L (2007) Genetic diversity and population differentiation of chestnut blight fungus, *Cryphonectria parasitica*, in China as revealed by RAPD. *Biochemical Genetics* **45**, 487-506.

© University of Pretoria

© University of Pretoria

Summary

Chrysoporthe canker is an important stem disease of commercially planted eucalypts in the tropics and sub-tropics. The disease is caused by several species of *Chrysoporthe*, which appear to have continental distributions and also occur on native hosts in each of their respective areas. The primary aim of this thesis was to elucidate the phylogeography and population biology of *Chrysoporthe* spp. that are important to commercial forestry. This was achieved by using several modern DNA based techniques, including multigene phylogenetic analyses, population genetic determinations using polymorphic microsatellite markers, and phylochronometric estimations. It was found that *Chrysoporthe cubensis* from South America is genetically distinct from *C. cubensis* occurring in Southeast Asia. The Asian form of the fungus was thus described as *Chrysoporthe deuterocubensis*. Calibrated chronometric phylogenies showed that the genus *Chrysoporthe* was approximately 7 million years old, placing its emergence well after the breakup of the Gondwana supercontinent. Therefore, continental drift did not play a role in the establishment of current geographic distributions of *Chrysoporthe* species. In Africa, *C. austroafricana* resulted from the subsequent allopatric speciation process. This fungus occurs on native *Syzygium* trees and two non-native trees, including *Eucalyptus*. Population genetic data showed that populations of *C. austroafricana* from *Syzygium* were more diverse than those on the other hosts, while a chronometric phylogeny also showed this population to be the oldest. Therefore, *Syzygium* represents the formative host of *C. austroafricana*, *i.e.*, the host on which the fungus evolved, while the other hosts have been colonized via host jumping. This capability was also investigated in a population of *C. cubensis* from Colombia, occurring on *Eucalyptus* and *Miconia* hosts. Population genetic estimates showed that this population is continually jumping between hosts, but it was not possible to determine if either of these hosts represented a formative host. Taken together, this study considerably advanced the understanding of population and phylogeographic processes that shaped the evolution of *C. cubensis*, *C. austroafricana* and *C. deuterocubensis*.

© University of Pretoria

Conclusions and Prospects

Chrysoporthe species are notorious pathogens of commercially planted *Eucalyptus* trees. This property has in the past spurred a substantial body of research regarding fungal pathogenicity, population diversity, host tolerance, and taxonomy of *Chrysoporthe* and related taxa. However, phylogeography of *Chrysoporthe* species remained neglected, presumably because of the large sample sizes required for such studies. In this thesis, a general phylogeographic study of three *Chrysoporthe* species was attempted.

Prior to this study, *Chrysoporthe cubensis* was thought to be native in both South America and Southeast Asia, and introduced into Africa. However, in **CHAPTER 3,** I was able to show that the Southeast Asian form of the fungus is a different species to *C. cubensis*, and named this new species *Chrysoporthe deuterocubensis*. Thus, *Chrysoporthe* species seem to have continental origins. The problem with the original misclassification of *C. deuterocubensis* appeared to be that inordinately large emphasis was placed on morphology and the phylogenetics of limited gene sets. These characteristics, although important, cannot reliably separate species in the *Chrysoporthe* complex. However, by employing a population genetic approach to show that gene flow does not occur, as well as genealogical concordance phylogenetic species recognition (GCPSR) using multiple gene regions, cryptic species can be separated. The application of these methods may in future lead to the identification of many more species in *Chrysoporthe*.

The continental distributions of *C. cubensis*, *C. austroafricana* and *C. deuterocubensis* raised questions as to how they have evolved. It was unclear which of these species were ancestral, and which might be derived species. Additionally, no information was available on possible modes of speciation, and it was though that the allopatric speciation patterns must have been established at the time when the Gondwana supercontinent started to break up. In **CHAPTER 4** it was shown that this assumption was incorrect, and that species of *Chrysoporthe* emerged long after the break-up of Gondwana. Additionally, *C. deuterocubensis* seemed to be the oldest species of the three, although it was not the progenitor of *C. austroafricana* and *C.*

© University of Pretoria

cubensis. Rather, the stem lineage of *Chrysoporthe* probably moved over temporary land bridges on ancient *Myrtales* hosts, into Africa and South America. After the land bridges disappeared, allopatric speciation could take place because genetic contact between the three populations was no longer possible. However, all three species eventually replaced the original stem lineage. Thus, the most recent common ancestor of any of the three species can be traced to a maximum of seven million years ago, when the last land bridges disappeared.

Since it was known how long *C. austroafricana* had been present in Africa, the question arose as to how it proceeded to colonize the continent. This question was investigated in **CHAPTER 5**, which dealt with gene flow between populations of the fungus on three different hosts. It was shown that *C. austroafricana* probably evolved on *Syzygium* hosts in South-Eastern Africa. The fact that the fungus is capable of infecting and causing disease on *Eucalyptus* (*Myrtaceae*) and *Tibouchina* (*Melastomataceae*) seems to be due to the fact that these hosts are related to *Syzygium*. The fungus appears to continually jump between hosts, although not all genotypes appear to have this capability because jumping is associated with a reduction in genetic variability. Additionally, the fungus only reproduces sexually on *Syzygium* and not on *Eucalyptus* or *Tibouchina*. Therefore, neither *Eucalyptus* nor *Tibouchina* contribute to the evolution of *C. austroafricana*, and the presence of the fungus on these hosts is largely coincidental.

When the ecology of *C. austroafricana* in Africa is compared to *C. cubensis* in South America, it is immediately obvious that many characteristics overlap. For instance, both species occur on *Myrtaceae* and *Melastomataceae* in both regions. The focus of **CHAPTER 6** was, therefore, to determine whether the same genetic patterns can be observed between *C. austroafricana* and *C. cubensis*. Population genetic data on a population of *C. cubensis* from *Eucalyptus* (*Myrtaceae*) and *Miconia* (*Melastomataceae*) showed that there are indeed ecological differents between populations on different hosts. For instance, *C. cubensis* can produce perithecia on *Eucalyptus*, but does not seem capable of doing so on *Miconia*. Further investigation into allele patterns and genetic diversity revealed that there are no differences between the two host-based populations. Therefore, in contrast to *C. austroafricana*, it

© University of Pretoria

does not appear that the host exerts a significant selective pressure on the fungus. Rather, we currently think that these two hosts in Colombia are infected by the same fungal population, which possibly originates from another source. The fact that the fungus can produce perithecia on *Eucalyptus* may be related to its ancient evolutionary history, although these speculations must be investigated.

Studies in this thesis presented convincing evidence regarding the phylogeography and population biology of three *Chrysoporthe* species. Although much work has been done, many questions still remain. The advancement of our understanding of these fungi also spurred new questions, which would previously not have been viable.

It is now evident that separate species of *Chrysoporthe* occur natively in Africa (*C. austroafricana*), South America (*C. cubensis*) and Southeast Asia (*C. deuterocubensis*). However, only the economically important species at the present time were considered in this study. *Chrysoporthe austroafricana* can be thought of as a sibling species of other presumably native *Chrysoporthe* species in Africa, such as *Chrysoporthe zambiensis*. It is possible that *C. cubensis* and *C. deuterocubensis*, in the same way as *C. austroafricana*, may have sibling species that are also native in their respective environments. Such species may be recovered by extensive sampling of trees in the *Myrtaceae* and *Melastomataceae* (both *Myrtales*) in all areas where *Chrysoporthe* spp. are known to occur. A better understanding of native *Chrysoporthe* species may lead to a more refined model regarding the ancient evolutionary history of the genus. This, in turn, may lead to better predictions regarding the capabilities of species to hybridize, native and non-native host species that may be at risk, as well as biodiversity information that may aid in the development of better quarantine measures.

Due to the fact that *Chrysoporthe* species are homothallic and can self-fertilize, it is difficult to use biological species recognition to delimit species in this genus. For this reason, it may be useful to develop laboratory tester strains that are differentially deficient for the mating type genes of corresponding *MAT1-1* and *MAT1-2* mating-type idiomorphs in closely related heterothallic fungi. Such engineered strains can

© University of Pretoria

then be used for several purposes, including the determination of sexual compatibility between putative species.

When host jumping occurs in *C. austroafricana*, it is associated with a loss in genetic variation. In other words, only some genotypes are able to jump hosts. Thus, the largest number of genotypes can be found on the formative host, while those on secondary hosts are presumably a sub-set of the genotypes present on the formative host. This leads to questions regarding why only some genotypes are able to jump, which in turn leads to speculation about the selective pressures exerted by the host on the fungus. If host-induced selection is present, then certain regions of the genome, or the regulation of gene sets, should be associated with host jumping capabilities. A possible method to test this hypothesis would be to make a cross between a genotype that has been found only on one host and another genotype that has been found only on another host. If the hypothesis is correct, then the progeny from such a cross should segregate according to host preference. It could then be possible to use "host of origin" or "preferred host" as a phenotypic character in order to create a genetic map so that putative quantitative trait loci (QTLs) may be identified. The resultant F1 progeny could also be used to map other traits, such as growth rate and pathogenicity towards different hosts. The identification of such regions of the genome would greatly advance our understanding of pathogenicity and host preference, not only in *Chrysoporthe* but possibly also in other fungi.

Diversity bottlenecks associated with host jumps, and the absence of sexual reproduction on non-formative hosts of *C. austroafricana*, provide an opportunity for biological control using hypovirulence. Double-stranded RNA hypoviruses, such as the *Cryphonectria parasitica* hypovirus, cause a reduction in pathogenicity of infected fungal individuals and can spread though a population through hyphal anastomosis. However, for hypovirus biocontrol to succeed, every vegetative compatibility group of the target population must be transfected. Since the number of VCGs of *C. austroafricana* on *Eucalyptus* in Africa is presumed to be limited, this strategy becomes feasible. Before the commencement of the current work, an isolate of *C. austroafricana* was transfected with the *C. parasitica* hypovirus. However, it was believed that the strategy would not be viable due to numerous VCGs present on

© University of Pretoria

Eucalyptus. Now that it is known that the VCGs (genotypes) on *Eucalyptus* are possibly selectively adapted to the host, we can assume that this number will not change drastically over time. Thus present study has thus shown that it might in fact be useful to consider biological control of *C. austroafricana* using hypovirulence.

The existence of microsatellite markers for population genetic studies is a marked improvement over vegetative compatibility tests. However, the limited number of markers may yield skewed results when gametic disequilibrium tests are to be performed. The problem may be remedied with the identification of neutral single nucleotide polymorphisms (SNPs) across the genome. This would be greatly aided by the availability of a reference genome sequence, which could be used to assemble RNA-seq or RAD-seq data for the identification of SNPs. Such SNPs may also be associated with defined traits such as host preference, and could lead to an in-depth genome wide association study. I believe that this would be the first time such a study had been performed in a fungus.

© University of Pretoria