Universidade de Aveiro Departamento de Biologia
2018

ANABELA RODRIGUES LOPES

Diversidade, espécies crípticas e estratégias sexuais em fungos da ordem *Botryosphaeriales*

Diversity, cryptic species and sexual strategies in fungi of the order *Botryosphaeriales*

Universidade de Aveiro Departamento de Biologia
2018

ANABELA RODRIGUES LOPES

Diversidade, espécies crípticas e estratégias sexuais em fungos da ordem *Botryosphaeriales*

Diversity, cryptic species and sexual strategies in fungi of the order *Botryosphaeriales*

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Biologia, realizada sob a orientação científica do Doutor Artur Jorge da Costa Peixoto Alves, Investigador Principal do Departamento de Biologia da Universidade de Aveiro e do Doutor Alan John Lander Phillips, Investigador do Instituto de Biossistemas e Ciências Integrativas da Faculdade de Ciências da Universidade de Lisboa.

Apoio financeiro da FCT e do FEDER através do programa COMPETE no âmbito do projecto de investigação ALIEN.

Apoio financeiro da FCT e do FSE no âmbito do III Quadro Comunitário de Apoio.

Bolsas com referência:
PTDC/AGR-PRO/2183/2014
POCI-01-0145-FEDER-016788

Bolsa de Doutoramento:
SFRH/BD/85615/2012

Para ti querida mãe…

Por muitas conquistas que faça sem ti a felicidade nunca será completa.

e

Para o mentor do MicroLab, o muito estimado Professor António Correia.

o júri

presidente

Doutor Manuel João Senos Matias
Professor Catedrático da Universidade de Aveiro

Doutor Nelson Manuel Viana da Silva Lima
Professor Catedrático da Universidade do Minho

Doutor Octávio Fernando de Sousa Salgueiro Godinho Paulo
Professor Auxiliar da Universidade de Lisboa

Doutor António Manuel Santos Carriço Portugal
Professor Auxiliar da Universidade de Coimbra

Doutora Isabel da Silva Henriques
Investigadora Auxiliar da Universidade de Aveiro

Doutor Artur Jorge da Costa Peixoto Alves
Investigador Principal da Universidade de Aveiro

agradecimentos

Como um bom doutoramento exige um bom orientador o primeiro agradecimento não poderia deixar de ser para o meu orientador Doutor Artur Alves. Numa conversa recente mencionaste, e bem, que fui eu que escolhi o orientador e não o contrário. A verdade é que se no início trabalhar contigo pareceu ser uma escolha acertada, agora tenho a certeza que foi. Há muitos motivos de agradecimento, desde teres aceitado orientar o meu trabalho ao muito que me ensinaste mas acima de tudo quero agradecer-te por teres compreendido, por me teres dado espaço e liberdade na gestão do meu trabalho que foi fundamental no melhor e no pior momento da minha vida. Obrigada!

Ao meu co-orientador Doutor Alan Phillips cujas experiência, espírito crítico e rigor foram fundamentais na solidez deste trabalho.

Aos co-autores dos trabalhos publicados Doutora Carla Barradas e Doutor Benedetto Linaldeddu pelo valioso contributo. A ti Carla agradeço ainda o companheirismo, amizade e disponibilidade de ajuda.

Às minhas colegas de laboratório Carina, Cátia, Liliana, Marta Alves e Susana pelo companheirismo, espírito de equipa e entreajuda, fundamentais para que esta jornada não tenha sido tão árdua.

Às Doutoras sénior do MicroLab Isabel Henriques, Ana Cristina Esteves, Cláudia Oliveira, Ana Sofia Duarte, Marta Tacão e Anabela Pereira pelo caloroso (re)acolhimento e pela prontidão em ajudar e dar bons conselhos.

Por fim, agradeço ao meu marido pelo amor e apoio incondicional e ao meu pequeno Simão que surgiu durante este percurso para dar mais luz à minha vida.

palavras-chave

Botryosphaeriaceae, *Neofusicoccum*, *Diplodia*, fitopatogénico, filogenia, genes *MAT*, heterotalismo, homotalismo

resumo

Os membros da ordem *Botryosphaeriales* ocorrem associados a inúmeras plantas lenhosas em todo o mundo como endofíticos ou patogénicos. De todas as famílias que compõem a ordem, a família *Botryosphaeriaceae* é a mais antiga e a maior, contendo vários géneros fitopatogénicos relevantes. Embora seja um passo fundamental para estudos em epidemiologia e para o controlo de doenças na planta, a correta identificação de espécies na família *Botryosphaeriaceae* é comumente problemática devido à existência de complexos de espécies crípticas definidos pela sobreposição de características morfológicas. Assim, o desenvolvimento de ferramentas para delimitar de forma clara as espécies é de extrema importância.

Este estudo explorou a diversidade de *Neofusicoccum*, um dos maiores e mais amplamente distribuídos géneros da família *Botryosphaeriaceae*, reconhecido também por possuir um grande número de espécies crípticas. Os isolados foram obtidos não apenas de espécies florestais e agrícolas relevantes, mas também de espécies menos exploradas como as ornamentais. A maioria dos isolados obtidos pertencia às espécies *N. eucalyptorum* e *N. australe*, mas também foram identificados isolados de *N. luteum*, *N. parvum* e *N. kwambonambiense*. Novas associações de hospedeiros foram encontradas, sendo *N. eucalyptorum* (um patogeno de *Eucalyptus* spp.) associado pela primeira vez a um hospedeiro não pertencente à família Myrtacea.

Os genes que determinam a identidade sexual e controlam o ciclo sexual fúngico (genes *MAT*) foram avaliados como potenciais marcadores moleculares para delimitar espécies em *Neofusicoccum* e *Diplodia*, dois géneros que contêm vários complexos de espécies crípticas. Para cada género foi desenvolvido um protocolo baseado na técnica de PCR para permitir facilmente amplificar e sequenciar os genes *MAT*. Foram feitas comparações entre as análises filogenéticas obtidas com os genes *MAT* e os marcadores convencionais factor de alongamento 1-alfa, *internal transcribed spacer* e beta-tubulina, onde se verificou que ambos os genes *MAT* permitem delimitar de forma clara e confiável todas as espécies, representando assim uma ferramenta poderosa para estabelecer limites de espécies, mesmo entre aquelas estreitamente relacionadas.

Apesar da sua relevância económica e ambiental, o conhecimento sobre as estratégias de reprodução sexual utilizadas por estes fungos é muito escasso. Neste estudo mostra-se que o homotalismo é a estratégia predominante entre as espécies de *Neofusicoccum*, não tendo sido encontrada em *Diplodia*, onde todas as espécies estudadas mostraram ser heterotálicas.

Aproveitando a disponibilidade de genomas sequenciados de várias espécies pertencentes à ordem *Botryosphaeriales*, realizaram-se análises da organização genómica dos *loci MAT* e respectivas regiões flanqueadoras. Os resultados mostraram que o heterotalismo predomina entre os membros da ordem. Com poucas exceções, a organização genómica dos *loci MAT* e das regiões flanqueadoras revelou algum grau de semelhança entre as diferentes famílias.

keywords

Botryosphaeriaceae, *Neofusicoccum*, *Diplodia*, phytopatogen, phylogeny, *MAT* genes, heterothallism, homothallism

abstract

Members of *Botryosphaeriales* are associated to numerous woody plants all over the world as endophytes or pathogens. From all families composing the order, the *Botryosphaeriaceae* is the oldest and the largest, containing several genera of relevant phytopathogens. Although it is a fundamental step towards studies on epidemiology and control of a plant disease, correct identification of species in the *Botryosphaeriaceae* is commonly problematic due to the existence of cryptic species complexes defined by the overlapping of morphological features. Thus, the development of tools to clearly delimit species boundaries is of the utmost importance.

This study explored the diversity of *Neofusicoccum*, one of the largest and most widely distributed genus in the *Botryosphaeriaceae*, well known by having a large number of cryptic species. Isolates were obtained, not only from relevant forest and crop tree species, but also from less exploited ornamental species. The majority of the isolates obtained belonged to *N. eucalyptorum* and *N. australe* but isolates of *N. luteum*, *N. parvum* and *N. kwambonambiense* were also identified. Novel host associations were found, with *N. eucalyptorum* (a pathogen of *Eucalyptus* spp.) being reported for the first time on a host outside the family Myrtaceae.

Mating type (*MAT*) genes, which determine sexual identity and control the fungal sexual cycle, were evaluated as molecular markers to delimit species in *Neofusicoccum* and *Diplodia*, two genera that contain several cryptic species complexes. For each genus a PCR-based assay was developed in order to easily amplify and sequence the mating type genes. Comparisons between the phylogenetic analyses obtained with *MAT* genes and the commonly used molecular markers translation elongation factor 1-alpha, internal transcribed spacer of the rDNA and beta-tubulin showed that both *MAT* genes are able to clearly and reliably delimit all species, thus representing a powerful tool to establish species boundaries, even between closely related cryptic species.

Despite their economic and environmental relevance, knowledge about the sexual reproductive (mating) strategies employed by these fungi is very scarce. Here it is shown that homothallism is the predominant strategy among *Neofusicoccum* species, while being absent in *Diplodia* where all studied species were shown to be heterothallic.

Taking advantage of the available genome sequences of several species belonging to the order *Botryosphaeriales*, analyses of the genomic organization of *MAT* loci and their flanking regions were performed. Results showed that heterothallism predominates among members of the order. With few exceptions, genomic organization of the *MAT* loci and flanking regions revealed some degree of gene synteny between the different families.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

act2 – Actin 2

APC5 - Anaphase promoting complex subunit 5

APN – Apurinic/apyrimidic endonuclease

BL – Acronym for Benedetto Linaldeddu culture collection

BLAST - Basic Local Alignment Search Tool

BSR – Biological species recognition

BSC – Biological species concept

Bp – Base pair

CAA – Acronym for Artur Alves culture collection

CAD – Acronym for Antonio Deidda culture collection

CAP – Acronym for Alan Phillips culture collection

CBS – Centraalbureau voor Schimmelcultures (Utrecht, Netherlands) accession number of strain culture

CERC – China Eucalypt Research Center, Beijing, China

CIA30 - Mitochondrial complex I intermediate-associated protein 30

CMM – Acronym for Maria Menezes culture collection

CMW – Tree Pathology Co-operative Program, Forestry and Agricultural Biotechnology Institute, University of Pretoria, South Africa

Cox-VIa - Cytochrome c oxidase subunit VIa

DNA – Deoxyribonucleic acid

dNTP – Deoxynucleotide triphosphates

EDTA - Ethylenediaminetetraacetic acid

GAP - GTPase-activating protein

GCPSR - Genealogical concordance phylogenetic species recognition

GPD - Glyceraldehyde-3-phosphate dehydrogenase

***his3* –** Histone H3

HMG - High mobility group

ICNF - Instituto da Conservação da Natureza e das Florestas

IMI - International Mycological Institute, CBI-Bioscience, Egham, Bakeham Lane, UK

iTOL - Interactive tree of life

ITS rDNA – Nuclear ribosomal internal transcribed spacer

JGI - Joint Genome Institute

LSU – Ribosomal large subunit

MAT – Mating type

MEGA - Molecular evolutionary genetics analysis

MgCl2 – Magnesium chloride

Mito_carr – Mitochondrial carrier protein

ML - Maximum likelihood

MSR – Morphological species recognition

MUCC – Murdoch University Culture Collection, Perth, Australia

NCBI - National Center for Biotechnology Information

NJ - Neighbour-Joining

ORF – Open reading frame

PCR – Polymerase chain reaction

PD - University of California, Davis, USA

PDA – Potato-dextrose agar

PH - Pleckstrin homology

PSR – Phylogenetic species recognition

RPB2 - RNA-polymerase II

RNA - Ribonucleic acid

SAICAR - Phosphoribosylaminoimidazole-succinocarboxamide synthase

SLA - Cytoskeleton assembly control

SNP - Single nucleotide polymorphisms

SSU – Ribosomal small subunit

TAE - Tris-acetate-EDTA buffer

***tef1-α* -** Translation elongation factor 1-alpha

***tub2* –** β-tubulin protein

UCR – College of Natural and Agricultural Sciences, Riverside, California, USA

WAC - Department of Agriculture, Western Australia Plant Pathogen Collection, South Perth, Western Australia

CHAPTER 1

Introduction

The order *Botryosphaeriales*

Several ascomycetes live within plant tissues without causing symptoms of disease. These fungi, called endophytes, are important components of plant microbiomes and could affect plant growth and plant responses to pathogens, herbivores and environmental changes (Porras-Alfaro & Bayman 2011). Some endophytes could be, however, latent or quiescent pathogens meaning that a change in the host or environment could trigger pathogenicity in the fungus resulting in disease symptoms (Porras-Alfaro & Bayman 2011). Fungi that belong to the order *Botryosphaeriales* are a good example of endophytic fungi that are often also latent pathogens (Slippers et al. 2017). The order includes 33 genera known from culture distributed within the families *Aplosporellaceae*, *Botryosphaeriaceae*, *Endomelanconiopsisaceae*, *Melanopsaceae*, *Phyllostictaceae*, *Planistromellaceae*, *Pseudofusicoccumaceae*, *Saccharataceae* and *Septorioideaceae* (Wyka & Broders 2016; Slippers et al. 2017; Yang et al. 2017). From these, the family *Botryosphaeriaceae* is the oldest and the most studied due to its worldwide distribution, ease of dissemination and richness of relevant plant pathogenic genera. Although the majority of *Botryosphaeriaceae* are woody plant colonizers some have the ability to infect non-woody plants and also humans (Dissanayake et al. 2016).

The family *Botryosphaeriaceae*

The *Botryosphaeriaceae* was first introduced by Theissen and Sydow in 1918 as a sub-family in the *Pseudosphaeriaceae* to include the genus *Botryosphaeria*, *Phaeobotryon* and *Dibotryon* (Phillips et al. 2013; Dissanayake et al. 2016). Since that time the taxon was rearranged many times based first on the morphological features alone and later based also on the internal transcribed spacer (ITS) sequence data (Crous et al. 2006; Phillips et al. 2013). Nowadays, based on morphological and molecular data, the *Botryosphaeriaceae* is accepted as a family of the order *Botryosphaeriales* and includes 23 genera known from culture (Slippers et al 2017).

Species of *Botryosphaeriaceae* colonize a huge variety of woody plants worldwide as saprobes, pathogens, endophytes or latent pathogens (Slippers & Wingfield 2007; Phillips et al. 2013). The distribution and diversity of species can be influenced by geographic location,

climatic conditions, and available hosts (Baskarathevan et al. 2012). As endophytes they can be easily moved and introduced into new regions threatening native and cultivated plants (Burgess & Wingfield 2002; Slippers & Wingfield 2007; Slippers et al. 2009). These fungi can enter plant tissues through natural openings (e.g. lenticels, stomata), reproductive structures (e.g. seeds) or wounds and colonize all plant parts (Slippers & Wingfield 2007). Between hosts, the *Botryosphaeriaceae* are predominantly transmitted horizontally via ascospores or conidia but also vertically *via* systematic infection, or *via* asexual sporulation from a seed infection, followed by infection of the growing plant (Slippers & Wingfield 2007). The switch from a latent to a pathogenic lifestyle is usually triggered by stressful environmental conditions such as drought, frost, hail, extreme temperature fluctuations, nutrient deficiencies and damage caused by other pathogens and pests (Slippers & Wingfield 2007). Disease symptoms caused by the *Botryosphaeriaceae* include fruit rots, leaf spots, gummosis, blue-stain of the sapwood, seedling damping-off and collar rot, blight of shoots and seedlings, dieback, cankers and tree death (Figure 1.1) (Slippers & Wingfield 2007). Although they are considered to be opportunistic pathogens, some diseases like panicle and shoot blight of pistachio and *Diplodia* dieback of pines are often aggressive and difficult to control (Burgess & Wingfield 2002; Michailides et al. 2002).

Botryosphaeriaceae could be host specific colonizing only a single host or a single lineage of hosts or could have broad host ranges (Slippers & Wingfield 2007). Species with very wide host and/or geographic ranges are usually the most aggressive pathogens because the interaction with a wider range of plants allows the species to be in contact with hosts that have not coevolved resistance to them (Parker & Gilbert 2004; Slippers et al. 2005). Also, species with broader host ranges are not dependent on the presence of specific hosts facilitating their establishment in new areas (Slippers & Wingfield 2007).

Figure 1.1 - Typical symptoms of disease caused by *Botryosphaeriaceae*. A – Dieback of twigs and branches associated with *N. australe* and *N. luteum* in *Tilia platyphyllos*. B – Internal necrosis in trunk of grapevine associated with *Lasiodiplodia theobromae*.

Within the *Botryosphaeriaceae* the genus *Neofusicoccum*, introduced by Crous et al. (2006) accommodates important species of plant pathogenic fungi, commonly associated with numerous woody hosts worldwide. Of these, *N. parvum* has the widest distribution, host range and proven ability to cause disease (Phillips et al. 2013; Sakalidis et al. 2013). Pathogenicity trials proved that *N. parvum* is an aggressive pathogen on a range of economically important horticultural and forestry plants such as grapevines (Úrbez-Torres et al. 2009; Baskarathevan et al. 2012), blueberries (Espinoza et al. 2009) and eucalyptus (Iturritxa et al. 2011) but also on ornamentals plants (Heath et al. 2011; Zlatković et al. 2016; Pelleteret et al. 2017). In Portugal *N. parvum* has been reported so far as being a pathogen of grapevines (Phillips 2002), *Proteaceae* (Crous et al. 2013), conifers (Alves et al. 2013) and eucalypts (Barradas et al. 2016). In addition to the genus *Neofusicoccum*, the genera *Diplodia* and *Botryosphaeria* include also some important pathogenic species. *Diplodia corticola*, *D. sapinea* and *B. dothidea* are just a few examples (Phillips et al. 2013). Whatever the pathogen is the most important thing to have in mind is that their correct identification is fundamental to implement

the appropriate quarantine decisions and suitable control strategies since different species might have different behaviours (e.g. in their virulence and host specialization) and respond to control measures differently (Wingfield et al. 2012; Slippers et al. 2017). However, this task is sometimes quite difficult due to the common occurrence of cryptic species. Cryptic species are a group of two or more species morphologically indistinguishable that are often erroneously classified under a single species name (Bickford et al. 2007). Morphological similarity is frequently associated to a recent speciation where there was not enough time for diagnosable features to evolve but could be also a result of a strong selective pressure that promotes morphological stasis (Bickford et al. 2007). The existence of these species complexes requires the use of accurate criteria for species recognition and delimitation.

Criteria for species recognition

Species are considered the currency of biology (Agapow et al. 2004). Precise species identification is a fundamental issue for studies in biodiversity, conservation, ecology, genetics, evolution and epidemiology (Agapow et al. 2004; Bickford et al 2007). However, decisions on how to define species boundaries, especially in fungi, are often problematic. The criteria used to delimit and identify species are many and have changed over time (De Queiroz 2007; Mallet 2007). In fungal taxonomy three criteria for species recognition are most commonly applied, namely the biological (BSR), morphological (MSR), and phylogenetic (PSR) species recognition criteria (Taylor et al. 2000). From these, the most popular is perhaps the BSR. The biological species recognition criterion is based on the biological species concept (BSC) which recognizes a species as a group of individuals with the capacity of interbreed but are reproductively isolated from other such groups (Taylor et al. 2000). Although apparently simple this approach has, however, several limitations that unable the use of the BSR as universal yardstick to delimit species. For example, the existence of hybrids proves that some different species of fungi are able to cross (Joly et al. 2006; De Vos et al. 2011; Cruywagen et al. 2017; Rodrigues-Galvez et al. 2017); the mechanisms of reproductive isolation are different among taxa (Giraud et al. 2008); many fungi are known only in their asexual states and it is not possible to determine if they are able to reproduce sexually (Taylor et al. 2000); homothallic species could undergo sexual reproduction without a partner (Taylor

et al. 2000); several heterothallic fungi cannot be grown and/or cross-mated in culture (Taylor et al. 2000; Bihon et al. 2014; Amorim et al. 2017).

For many decades the most commonly used species criterion was based on the morphology (MSR) (Taylor et al. 2000). Several phenotypic characters but also the production of secondary metabolites and/or the presence of pigments have been widely used to diagnose the species which nowadays allow making comparisons among existing taxa and between new and existing taxa (Taylor et al. 2000). In the family *Botryosphaeriaceae*, for instance, the morphological characters of asexual reproductive structures are the most often used due to their higher variability, frequency in nature, and ease to induce in culture (Slippers & Wingfield 2007). Conidial features such as pigmentation, wall thickness, ornamentation, maturation, septation, shape and size (length, width, l/w and l×w) are the most informative characters but also the presence or absence of paraphyses in the conidiomata can be useful to identify the genus (Slippers & Wingfield 2007; Phillips et al. 2013). However, in the last two decades, with the use of other species delimitation criteria some weaknesses of the MSR criterion were unveiled. The first one concerns the existence of cryptic species. In the presence of a complex of morphological identical species the MSR criterion showed not to be able to distinguish them, hiding the true biodiversity (Taylor et al. 2000). A good example is the *Neofusicoccum parvum-N. ribis* complex. This complex is now recognized as a group of several different species, most of them indistinguishable by the traditional morphological characters (Pavlic et al. 2009a, b; Sakalidis et al. 2011). More examples of complexes of cryptic species can be found within the genus *Lasiodiplodia* (Damm et al. 2007; Alves et al. 2008) and *Diplodia* (Figure 1.2) (Phillips et al. 2012; Alves et al. 2014). Less frequent but also found are the morphological similarities shared between species of different genera such as *Botryosphaeria* and *Neofusicoccum* (Fig 1.2) (Abdollahzadeh et al. 2013; Slippers et al. 2017). In these cases the species delimitation based on morphology alone is also almost impossible.

The second weakness of the MSR is related to the reproductive structures. Almost all fungi have two reproductive cycles (asexual and sexual) exhibiting different morphologies for the reproductive structures (Figure 1.2). Also, the same species could have more than one asexual morphology (Shenoy et al. 2007; Wingfield et al. 2012; Martin et al. 2013). The discovery of the different morphologies at different times, in different places and by different

investigators resulted, along the years, in the attribution of more than one name to the same species, which overestimated the true biodiversity (Shenoy et al. 2007; Wingfield et al. 2012; Crous et al. 2015). Thus, although MSR could be useful in some cases, it should be used always with caution and never alone. In the case of taxa belonging to the family *Botryosphaeriaceae* the typical culture morphology (grey to black aerial mycelium and grey to indigo-grey or black pigment visible on the reverse side of Petri dishes) could be useful for a primary grouping of related isolates from a broad sampling (Slippers & Wingfield 2007).

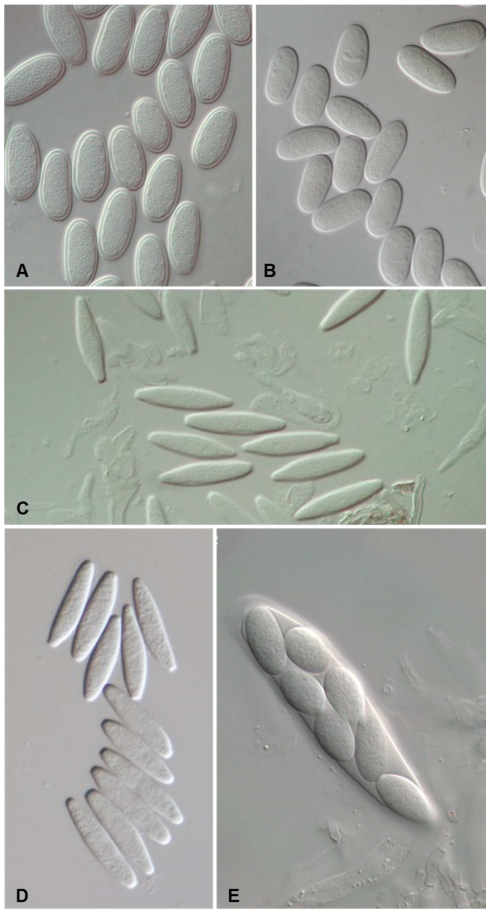

Figure 1.2 - Similar conidia morphology within the same genus (A - *Diplodia malorum*, B – *D. mutila*) and between genera (C – *Botryosphaeria dothidea*, D – *Neofusicoccum australe*). The same species may have two morphologically different reproductive structures (asexual (D) and sexual (E) morphologies of *Neofusicoccum australe*).

The limitations and inefficiencies related to BSR and MSR triggered nowadays the massive use of the phylogenetic species recognition criterion (PSR) (Taylor et al. 2000), resulting in the recognition of a huge number of cryptic species (Damm et al. 2007; Alves et al. 2008; Pavlic et al. 2009a, b; Sakalidis et al. 2011; Phillips et al. 2012; Alves et al. 2014). The PSR is based on the phylogenetic analysis of variable characters, usually DNA sequences of specific genes or genomes (Taylor et al. 2000; Cai et al. 2011). The molecular revolution in fungal taxonomy started in 1990 when White et al. designed primers to amplify nuclear rDNA genes. Nowadays the number of sequences deposited in the public databases places the internal transcribed spacer (ITS) as the most popular genetic marker in fungal species identification and molecular taxonomy. The ITS region consists of two regions, ITS1 and ITS2, separated by the 5.8S gene and is located between the 18S (SSU) and 28S (LSU) genes in the nrDNA repeat unit (White et al. 1990). The availability of universal primers able to bind to DNA of most fungal taxa, the multi-copy structure in the genome that allows an efficient amplification even from samples with initial low amount of DNA, the relatively small length of the region that allows an ease amplification and Sanger sequencing and the good resolution power to species discrimination in most fungal taxa due to its high inter- and low intra-species divergence were enough valid reasons to considered the ITS region as a barcode for fungi (Schoch et al. 2012; Crous et al. 2015; Irinyi et al. 2016). However, in spite of having all these advantages the ITS region misses the biggest issue that a barcode marker should have that is to allow a straightforward identification, i.e., be unique to a single species and constant within each species. Several cryptic species, e.g., were overlooked in studies where only the ITS was used (Denman et al. 2000; Smith et al. 2001; Zhou and Stanosz 2001; Pavlic et al. 2007). A single gene or region analysis would be straightforward only if its evolutionary history reflects all the entire organism evolution, which not happen in the case of ITS nor in any other region (Aguileta et al. 2008). In order to minimize this problem Taylor et al. (2000) proposed the use of a Genealogical Concordance Phylogenetic Species Recognition (GCPSR) where genealogies of multiple independent loci are compared. According to this criterion, conflict among gene genealogies is likely to be due to recombination among individuals within a species, imposing the limits of species in the nodes where multiple trees display congruence (Taylor et al. 2000).

Due to the advantages described above to the use of ITS, this region is almost all the times used in the GCPSR criterion given a good discrimination at the generic and/or family level (Crous et al. 2015). The problem lies on the choice of additional loci to compare in order to get an accurate discrimination at the species level. The use of genes whose evolutionary histories include duplications, horizontal transfer, lineage sorting, or selection-based biases may result in discrepancies between gene and species trees. Also, the performance of the different protein-coding genes in deriving a reliable phylogeny is highly variable, resulting that the same group of genes could be used in different taxa with different phylogenetic resolution power (Aguileta et al. 2008). Several genes were tested for their potential as a supplement to ribosomal genes, being the translation elongation factor 1-alpha (*tef1-α*) considered one good option (Stielow et al. 2015). However, other protein-coding genes such as β-tubulin (*tub2*) gene, the second largest subunit of RNA-polymerase II (*RPB2*), actin (*act2*), and histone H3 (*his3*) are also useful (Crous et al. 2015; Stielow et al. 2015).

In the family *Botryosphaeriaceae* the most common molecular markers used are the ITS and *tef1-α* (Barradas et al. 2016; Correia et al. 2016; González-Domínguez et al. 2016; Slippers et al. 2017). Although these two regions could give satisfactory results in some cases, it is more frequent and advisable the usage of at least one more, being *tub2, RPB2* and/or LSU the most applied genes (Jami et al. 2014; Júnior et al. 2016; Cruywagen et al. 2017; Lawrence et al. 2017; Moral et al. 2017; Netto et al. 2017; Slippers et al. 2017). Combining slow evolving loci with loci with high rates of evolution will provide not only the reconstruction of deep phylogenetic relationships but also reflect recent evolutionary and speciation events (Schmitt et al. 2009; Stielow et al. 2015). The phylogenetic analyses with multiple genes are usually performed in two different ways. In the concatenation approach the phylogenetic reconstruction is done after the gene sequences are concatenated head-to-tail to form a super-gene alignment. On the other hand, in the consensus phylogeny approach the phylogenies are inferred and analysed separately for each gene (Gadagkar et al. 2005). Although both methods are commonly accepted, the usage of the concatenation approach should be done only if the individual gene phylogenies have similar topologies. Otherwise, the true evolutionary relationships among closely related taxa could be masked and the species poorly discriminated (Liu et al. 2016; Slippers et al. 2017).

The decision about which species recognition criterion should be applied depends on the type of organism, its history of speciation and the degree of achieved divergence (Cai et al. 2011). In fungi, is clear that the GCPSR is the most widely used and accepted criterion to infer about phylogenetic relationships and species boundaries. However, searches for new molecular markers have been made in order to reach the most reliable results especially in groups were cryptic species and/or hybrids are frequent. Among them the mating type genes are being one of the most studied.

Mating type genes

Sexual reproduction is considered a major driving force for diversification in natural populations of fungi (Butler 2007). Beyond the genetic variation the evolutionary benefits that promote sexual reproduction are DNA maintenance and repair during meiosis and elimination of deleterious mutations (Lee et al. 2010; Heitman et al. 2013). Sexual reproduction is also widely recognized as one of the crucial processes in fungal pathogens for habitat adaptation and host specificity (McDonald & Linde 2002; Hsueh & Heitman 2008).

In fungi, the sexual cycle is regulated by mating type genes (Debuchy et al. 2010; Lee et al. 2010; Ni et al. 2011; Bennett & Turgeon 2016). Besides involvement in fungal mating, *MAT* genes have been shown to be involved with hyphal morphology, conidia formation, and production of secondary metabolites (Böhm et al. 2013; Becker et al. 2015). Filamentous ascomycete fungi present a bipolar system, which is defined by the presence of a single locus (*MAT1*) that harbours two mating-types known as *MAT1-1* and *MAT1-2* (Turgeon & Yoder 2000; Debuchy et al. 2010; Ni et al. 2011). Although the study of Martin et al. (2010) suggest that core motifs in some *MAT* genes may have a common evolutionary origin, the term 'idiomorph' has been adopted to emphasize that they are not obviously related by structure or common descent despite occupying the same locus (Metzenberg & Glass 1990). Due to their low similarity idiomorphs do not usually recombine (Turgeon 1998).

The molecular structure of the *MAT* locus was first characterized in *Saccharomyces cerevisiae* (Astell et al. 1981), and mating-type genes were first cloned and sequenced in the filamentous ascomycete *Neurospora crassa* (Glass et al. 1988). Since then, *MAT* genes have been identified and characterized in an increasing number of filamentous ascomycetes

(Inderbitzin et al. 2005; Kanematsu et al. 2007; Pöggeler et al. 2011; Palmer et al. 2014; De Miccolis Angelini et al. 2016), including putatively asexual species (Pöggeler 2002; Groenewald et al. 2006; Santos et al. 2010; Bolton et al. 2012; Bohm et al. 2013). Although the gene composition of the *MAT* locus may vary greatly among species, two master *MAT* genes are consistently found in filamentous ascomycetes (Li et al. 2010). In the *MAT1-1* idiomorph the mandatory gene *MAT1-1-1* encodes an alpha box protein whereas the mandatory gene *MAT1-2-1* characterize the *MAT1-2* idiomorph and encodes for a protein with a high mobility group (HMG) domain (Coppin et al. 1997; Kronstad & Staben 1997; Debuchy & Turgeon 2006; Lee et al. 2010; Li et al. 2010). Beyond the main genes also additional genes has been found in the idiomorphs. In the *MAT1-1* idiomorph of Dothideomycetes, for example, may be present the gene *MAT1-1-8*, while in the *MAT1-2* idiomorph could be located the additional gene *MAT1-2-5* (Wilken et al. 2017).

The presence and expression of genes from both *MAT1-1* and *MAT1-2* idiomorphs play an essential role in the initiation and full process of sexual reproduction (Ni et al. 2011). Thus, fungal mating systems can be classified based on the genic content of the *MAT1* locus. Homothallic species contain sequences encoding both *MAT1-1* and *MAT1-2* specific genes that can be either located in close proximity or apart (Figure 1.3) (Lin & Heitman 2007; Debuchy et al. 2010; Ni et al. 2011; Bennett & Turgeon 2016). On the contrary, heterothallic species possess genes from only one of the two idiomorphs (Figure 1.3). In these species, the combined expression of genes from both idiomorphs requires the presence of two opposite mating type partners (Lin & Heitman 2007; Debuchy et al. 2010; Ni et al. 2011; Bennett and Turgeon 2016). While heterothallism represents a relatively simple and well-understood opposite mate interaction, homothallism constitutes a variety of distinct strategies that collectively allow for single individuals to sexually reproduce independently of an opposite mating partner (Wilson et al. 2015). Apart from homothallism and heterothallism also pseudohomothallism could be found in some filamentous ascomycetes such as species of *Neurospora* and *Podospora* (Figure 1.3) (Debuchy et al. 2010; Billiard et al. 2012; Bennett & Turgeon 2016). In this case self-fertility is the result of the packaging of two independent and opposite mating type nuclei within a single spore. Functionally, it thus represents a heterothallic system that is able to occur within a single originating cell (Debuchy et al. 2010;

Billiard et al. 2012; Bennett & Turgeon 2016). Finally, unisexual reproduction was also reported in filamentous ascomycetes (e.g. *Huntiella moniliformis* and *Neurospora africana*) (Figure 1.3) (Wilson et al. 2015; Bennett & Turgeon 2016). This extremely rare event describes an atypical system where species are able to complete an entire sexual cycle when only genes from a single *MAT* idiomorph are expressed (Wilson et al. 2015; Bennett & Turgeon 2016).

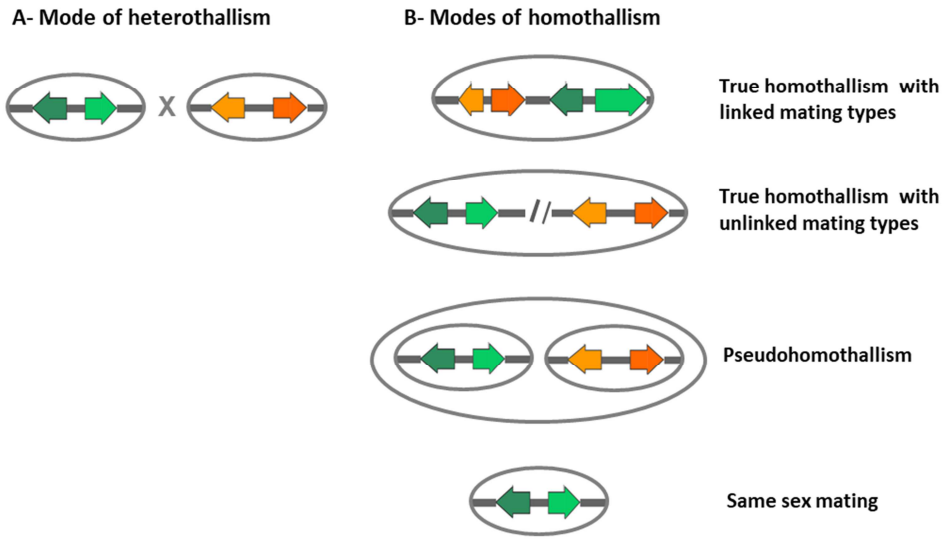

Figure 1.3 - Sexual reproductive strategies of ascomycetes.

The structure of the mating type locus in most heterothallic ascomycetes is relatively conserved in opposition to more variable organization found in the mating type loci of homothallic fungi (Debuchy & Turgeon 2006). However, the flanking regions are usually conserved in content and gene order among most homothallic and heterothallic ascomycetes (Butler 2007; Debuchy et al. 2010). Open reading frames (ORFs) such as DNA lyase (*APN1/2*), anaphase promoting complex subunit 5 (*APC5*), cytochrome c oxidase subunit VIa (*Cox-VIa*) and cytoskeleton assembly control (*SLA2*) are commonly found close to fungal *MAT* genes (Butler 2007; Debuchy et al. 2010; Lee et al. 2010). In contrast, ORFs like GTPase-activating protein (*GAP*), mitochondrial complex I intermediate-associated protein 30 (*CIA30*) and mitochondrial carrier protein (*Mito_carr*) have been associated to *MAT* genes in a

few cases (Cozijnsen & Howlett 2003; Bihon et al. 2014; Vaghefi et al. 2015; Amorim et al. 2017).

The rapid increasing number of studies on fungal mating systems emphasizes the important role that *MAT* genes play in fungal biology. Information regarding structure and organization of *MAT* loci can be used to establish whether a fungal species reproduces in a homothallic or heterothallic way. Furthermore, such knowledge can be used to develop molecular markers in order to determine the mating types of isolates, replacing the laborious and time consuming process of developing and crossing mating tester strains (Scherrer et al. 2005; Ramirez-Prado et al. 2008; Santos et al. 2010; Brewer et al. 2011). The ability to identify mating types using molecular markers is useful when it is necessary to select isolates of opposite mating type for genetic experiments, but also to infer about the existence of sexual reproduction in natural populations of species known only by their asexual morphs (Groenewald et al. 2006; Stergiopoulos et al. 2007; Kück & Pöggeler 2009; Duong et al. 2015). The discovery of sexual cycles is important in many fungi of clinical or industrial relevance (Seidl et al. 2009; Kück & Böhm 2013) and in the phytopathogenic fungi due to its serious implications for disease management (McDonald & Linde 2002). Recombining phytopathogenic populations often exhibit a high degree of genotype diversity which can increase the mean fitness of the next generation, accelerating adaptation to changes in the environment and improving their chances for long-term survival (McDonald & Linde 2002; Hsueh & Heitman 2008). Sexual fruiting structures can also function as survival structures in times of adverse environmental conditions, prior to the release of ascospores as primary inoculum and dispersal of the pathogen over long distances (Vaghefi et al. 2015). Knowledge of the reproductive capacities of a pathogen is thus crucial to understand the disease epidemiology and the subsequent development of effective and sustainable management strategies (McDonald & Linde 2002).

Mating type genes have been sequenced in an increasing number of fungi, not only to investigate their reproduction mode, but also to resolve their phylogenetic history and species boundaries (Barve et al. 2003; O'Donnell et al. 2004; Du et al. 2005; Voight et al. 2005; Yokoyama et al. 2006; Rau et al. 2007; Wik et al. 2008; Arzanlou et al. 2010; Santos et al. 2010; Strandberg et al. 2010; Pöggeler et al. 2011; Geng et al. 2014; Kashyap et al. 2014). Due

to their fast evolution rates, high interspecific variation and low dissimilarity within species (Turgeon 1998) *MAT* genes are suitable candidates for phylogenetic analysis of closely related species, giving better resolution than other single phylogenies such as ITS (Barve et al. 2003; Du et al. 2005; Yokoyama et al. 2006; Rau et al. 2007).

Although essential to understand the biology, evolution, genetics and epidemiology of fungi, the knowledge of mating type loci in the family *Botryosphaeriaceae* is restricted so far to the species *Diplodia sapinea* (Bihon et al. 2014), leaving much to explore.

Objectives

The wide range of hosts and geographic areas colonized by species of the family *Botryosphaeriaceae* together with the ability to cause serious disease or even plant death are enough reasons to study this taxon in all possible fields. However, despite the attention given to this group in the last decades some questions remain to be addressed. Thus, the main objectives of this work were:

1. Characterize the diversity of species associated with several woody and non woody hosts in Portugal;
2. Unveil cryptic species;
3. Uncover sexual strategies and characterize mating type loci.

Thesis outline

Regarding the first main objective we focused on the diversity of *Neofusicoccum* species due to its recognizable capacity to colonize a wide range of geographical areas and hosts, the presence of relevant plant pathogenic species, as well as the common occurrence of cryptic species. Thus, in Chapter 2 the *Neofusicoccum* species isolated from several diseased and healthy forest, crop and ornamental plant species are identified.

The second major objective concerned the existence of numerous cryptic species in genera that belong to the family and the need for their correct identification. Since several studies proved that mating type genes are good molecular markers to delimit species boundaries the purpose was to test their usefulness in the important phytopathogenic genera

Neofusicoccum (Chapter 3) and *Diplodia* (Chapter 4). The specific objectives were similar for both cases. First, to analyse the available genomes of each genus to obtain the mating type genes sequences; second, to develop a PCR-based assay to amplify and sequence them; third, to evaluate the usefulness of *MAT* genes as phylogenetic markers to establish species boundaries and to compare the results with those obtained with the conventional molecular markers; and finally, to retrieve information about sexual strategies.

Beyond the uncovering of sexual strategies in *Diplodia* and *Neofusicoccum* species, and through the analyses of all available genomes in the databases, the sexual reproductive strategies of species belonging to other families in the order *Botryosphaeriales* were unveiled. To achieve this, the *MAT* locus of each genome was identified and characterized. Also, a characterization of flanking regions and comparison between loci was carried out (Chapter 5).

References

Abdollahzadeh J, Zare R, Phillips AJL, 2013. Phylogeny and taxonomy of *Botryosphaeria* and *Neofusicoccum* species in Iran, with description of *Botryosphaeria scharifii* sp. nov. *Mycologia* **105**:210-220.

Agapow PM, Bininda-Emonds OR, Crandall KA, Gittleman JL, Mace GM, Marshall JC, Purvis A, 2004. The impact of species concept on biodiversity studies. *The Quarterly Review of Biology* **79**:161–179.

Aguileta G, Marthey S, Chiapello H, Lebrun M-H, Rodolphe 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.

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

Alves A, Barradas C, Phillips AJL, Correia A, 2013. Diversity of *Botryosphaeriaceae* species associated with conifers in Portugal. *European Journal of Plant Pathology* **135**:791-804.

Alves A, Linaldeddu BT, Deidda A, Scanu B, Phillips AJL, 2014. The complex of *Diplodia* species associated with *Fraxinus* and some other woody hosts in Italy and Portugal. *Fungal Diversity* **67**:143-156.

Amorim R, Savi DC, Ferreira-Maba L, Aluizio R, Goulin EH, Takita MA, Machado MA, Glienke C, 2017. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in the citrus pathogen *Phyllosticta citricarpa*. *European Journal of Plant Pathology* **147**:325-337.

Arzanlou M, Crous PW, Zwiers LH, 2010. Evolutionary dynamics of mating-type loci of *Mycosphaerella* spp. occurring on banana. *Eukaryotic Cell* **9**: 164-172.

Astell CR, Ahlstrom-Jonasson L, Smith M, Tatchell K, Nasmyth KA, Hall BD, 1981. The sequence of the DNAs coding for the mating-type loci of *Saccharomyces cerevisiae*. *Cell* **27**:15-23.

Barradas C, Phillips AJL, Correia A, Diogo E, Bragança H, Alves A, 2016. Diversity and potential impact of *Botryosphaeriaceae* species associated with *Eucalyptus globulus* plantations in Portugal. *European Journal of Plant Pathology* **146**:245-257.

Barve MP, Arie T, Salimath SS, Muehlbauer FJ, Peever TL, 2003. Cloning and characterization of the mating type (*MAT*) locus from *Ascochyta rabiei* (teleomorph: *Didymella rabiei*) and a *MAT* phylogeny of legume-associated *Ascochyta* spp. *Fungal Genetics and Biology* **39**:151-167.

Baskarathevan J, Jaspers MV, Jones EE, Ridgway HJ (2012) Incidence and distribution of botryosphaeriaceous species in New Zealand vineyards. *European Journal of Plant Pathology* **132**:549–560.

Becker K, Beer C, Freitag M, Kück U, 2015. Genome-wide identification of target genes of mating-type α-domain transcription factor reveals functions beyond sexual development. *Molecular Microbiology* **96**:1002-1022.

Bennett RJ, Turgeon BG, 2016. Fungal sex: the *Ascomycota*. *Microbiology Spectrum* **4**:FUNK-0005-2016.

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I, 2007. Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* **22**:148-155.

Bihon W, Wingfield MJ, Slippers B, Duong TA, Wingfield BD, 2014. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. *Fungal Genetics and Biology* **62**:55-61.

Billiard S, López-Villavicencio M, Hood M, Giraud T, 2012. Sex, outcrossing and mating types: unsolved questions in fungi and beyond. *Journal of Evolutionary Biology* **25**:1020-1038.

Böhm J, Hoff B, O'Donnell CM, Wolfers S, Klix V, Binger D, Zadra I, Kürnsteiner H, Pöggeler S, Kück U. 2013. Sexual reproduction and mating type-mediated strain development in the penicillin-producing fungus *Penicillium chrysogenum*. *Proceedings of the National Academy of Sciences of the United States of America* **110**:1476–1481.

Bolton MD, Secor GA, Rivera V, Weiland JJ, Rudolph K, Birla K, Rengifo J, Campbell LG, 2012. Evaluation of the potential for sexual reproduction in field populations of *Cercospora beticola* from USA. *Fungal Biology* **116**:511–521.

Brewer MT, Cadle-Davidson L, Cortesi P, Spanu PD, Milgroom MG, 2011. Identification and structure of the mating-type locus and development of PCR-based markers for mating type in powdery mildew fungi. *Fungal Genetics and Biology* **48**: 704-713.

Burgess TI, Wingfield MJ, 2002. Quarantine is important in restricting the spread of exotic seed-borne tree pathogens in the southern hemisphere. *International Forestry Review* **4**:56–65.

Butler G, 2007. The evolution of *MAT*: the ascomycetes. In: Heitman J, Kronstad JW, Taylor J, Casselton L (eds), *Sex in fungi: molecular determination and evolutionary implications*. ASM Press, Washington, DC, pp. 3-18.

Cai L, Giraud T, Zhang N, Begerow D, Cai G, Shivas RG, 2011. The evolution of species concepts and species recognition criteria in plant pathogenic fungi. *Fungal Diversity* **50**:121-133.

Coppin E, Debuchy R, Arnaise S, Picard M, 1997. Mating types and sexual development in filamentous ascomycetes. *Microbiology and Molecular Biology Reviews* **61**: 411-428.

Correia KC, Silva MA, de Morais Jr MA, Armengol J, Phillips AJL, Câmara MPS, Michereff SJ, 2016. Phylogeny, distribution and pathogenicity of *Lasiodiplodia* species associated with dieback of table grape in the main Brazilian exporting region. *Plant Pathology* **65**:92-103.

Cozijnsen AJ, Howlett BJ, 2003. Characterisation of the mating-type locus of the plant pathogenic ascomycete *Leptosphaeria maculans*. *Current Genetics* **43**:351-357.

Crous PW, Slippers B, Wingfield MJ, Rheeder J, Marasas WFO, Phillips AJL, Alves A, Burgess T, Barber P, Groenewald JZ, 2006. Phylogenetic lineages in the *Botryosphaeriaceae*. *Studies in Mycology* **55**:235–253.

Crous PW, Denman S, Taylor JE, Swart L, Bezuidenhout CM, Hoffman L, Palm ME, Groenewald JZ, 2013. Cultivation and Disease of *Proteaceae*: *Leucadendron*, *Leucospermum*, and *Protea*: Second Edition. CBS Biodiversity Series, 360 pages.

Crous PW, Hawksworth DL, Wingfield MJ, 2015. Identifying and Naming Plant-Pathogenic Fungi: Past, Present, and Future. *Annual Review of Phytopathology* **53**:1-21.

Cruywagen EM, Slippers B, Roux J, Wingfield MJ, 2017. Phylogenetic species recognition and hybridisation in *Lasiodiplodia*: a case study on species from baobabs. *Fungal Biology* **121**:420-436.

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

De Miccolis Angelini RM, Rotolo C, Pollastro S, Faretra F, 2016. Molecular analysis of the mating type (*MAT1*) locus in strains of the heterothallic ascomycete *Botrytis cinerea*. *Plant Pathology* **65**:1321-1332.

De Queiroz K, 2007. Species Concepts and Species Delimitation. *Systematic Biology* **56**:879–886.

De Vos L, Van der Nest MA, Van der Merwe NA, Myburg AA, Wingfield MJ, Wingfield BD, 2011. Genetic analysis of growth, morphology and pathogenicity in the F1 progeny of an interspecific cross between *Fusarium circinatum* and *Fusarium subglutinans*. *Fungal Biology* **115**:902-908.

Debuchy R, Turgeon BG, 2006. Mating-Type structure, evolution, and function in euascomycetes. In: Kües, Fisher (eds), *The Mycota I - Growth, Differentiation and Sexuality*. Springer-Verlag, Berlin, Heidelberg, pp. 293-323.

Debuchy R, Berteaux-Lecellier V, Silar P, 2010. Mating systems and sexual morphogenesis in ascomycetes. In: Borkovich KA, Ebbole E (eds), *Cellular and Molecular Biology of Filamentous Fungi*. ASM Press, Washington, DC, pp. 501-535.

Denman S, Crous PW, Taylor JE, Kang JC, Pascoe I, Wingfield MJ, 2000. An overview of the taxonomic history of *Botryosphaeria* and a re-evaluation of its anamorphs based on morphology and ITS rDNA phylogeny. *Studies in Mycology* **45**:129–140.

Dissanayake AJ, Phillips AJL, Li XH, Hyde KD, 2016. *Botryosphaeriaceae*: Current status of genera and species. *Mycosphere* **7**:1001-1073.

Du M, Schardl CL, Nuckles EM, Vaillancourt LJ, 2005. Using mating- type gene sequences for improved phylogenetic resolution of *Colletotrichum* species complexes. *Mycologia* **97**:641-658.

Duong TA, de Beer ZW, Wingfield BD, Eckhardt LG, Wingfield MJ, 2015. Microsatellite and mating type markers reveal unexpected patterns of genetic diversity in the pine root-infecting fungus *Grosmannia alacris*. *Plant Pathology* **64**:235-242.

Espinoza JG, Briceño EX, Chávez ER, Úrbez-Torres JR, Latorre BA, 2009. *Neofusicoccum* spp. associated with stem canker and dieback of blueberry in Chile. *Plant Disease* **93**:1187-1194.

Gadagkar SR, Rosenberg MS, Kumar S, 2005. Inferring Species Phylogenies From Multiple Genes: Concatenated Sequence Tree Versus Consensus Gene Tree. *Journal of Experimental Zoology (Molecular and Developmental Evolution)* **304B**:64-74.

Geng Y, Li Z, Xia L-Y, Wang Q, Hu X-M, Zhang X-G, 2014. Characterization and phylogenetic analysis of the mating-type loci in the asexual ascomycete genus *Ulocladium*. *Mycologia* **106**:649-665.

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, Vollmer SJ, Staben C, Grotelueschen J, Metzenberg RL, Yanofsky C, 1988. DNAs of the two mating-type alleles of *Neurospora crassa* are highly dissimilar. *Science* **241**:570–573.

Groenewald M, Groenewald JZ, Harrington TC, Abeln ECA, Crous PW, 2006. Mating type gene analysis in apparently asexual *Cercospora* species is suggestive of cryptic sex. *Fungal Genetics and Biology* **43**: 813-825.

Heath RN, Roux J, Slippers B, Drenth A, Pennycook SR, Wingfield BD, Wingfield MJ, 2011. Occurrence and pathogenicity of *Neofusicoccum parvum* and *N. mangiferae* on ornamental *Tibouchina* species. *Forest Pathology* **41**:48-51.

Heitman J, Sun S, James TY, 2013. Evolution of fungal sexual reproduction. *Mycologia* **105**:1-27.

Hsueh YP, Heitman J, 2008. Orchestration of sexual reproduction and virulence by the fungal mating-type locus. *Current Opinion in Microbiology* **11**: 517–524.

Inderbitzin P, Harkness J, Turgeon BG, Berbee ML, 2005. Lateral transfer of mating system in Stemphylium. *Proceedings of the National Academy of Sciences of the United States of America* **102**:11390-11395.

Irinyi L, Lackner M, De Hoog GS, Meyer W, 2016. DNA barcoding of fungi causing infections in humans and animals. *Fungal Biology* **120**:125-136.

Iturrixa E, Slippers B, Mesanza N, Wingfield MJ, 2011. First report of *Neofusicoccum parvum* causing canker and die-back of Eucalyptus in Spain. *Australasian Plant Disease Notes* **6**:57–59.

Jami F, Slippers B, Wingfield MJ, Gryzenhout M, 2014. Botryosphaeriaceae species overlap on four unrelated, native South African hosts. *Fungal Biology* **118**:168-179.

Joly DL, Langor DW, Hamelin RC, 2006. Molecular and morphological evidence for interspecific hybridization between *Cronartium ribicola* and *C. comandrae* on *Pinus flexilis* in southwestern Alberta. *Plant Disease* **90**:1552-1552.

Júnior AFN, Fisher IH, Bragança CAD, Júnior NSM, Amorim L, 2016. *European Journal of Plant Pathology* **144**:271-287.

Kanematsu S, Adachi Y, Ito T, 2007. Mating-type loci of heterothallic *Diaporthe* spp.: homologous genes are present in opposite mating-types. *Current Genetics* **52**:11-22.

Kashyap PL, Rai S, Kumar S, Srivastava AK, Anandaraj M, Sharma AK, 2014. Mating type genes and genetic markers to decipher intraspecific variability among *Fusarium udum* isolates from pigeon pea. *Journal of Basic Microbiology* **54**:1-11.

Kronstad JW, Staben C, 1997. Mating type in filamentous fungi. *Annual Review of Genetics* **31**: 245-276.

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

Kück U, Böhm J, 2013. Mating type genes and cryptic sexuality as tools for genetically manipulating industrial molds. *Applied Microbiology and Biotechnology* **97**:9609-9620.

Lawrence DP, Hand FP, Gubler WD, Trouillas FP, 2017. *Botryosphaeriaceae* species associated with dieback and canker disease of bay laurel in northern California with the description of *Dothiorella californica* sp. nov. *Fungal Biology* **121**:347-360.

Lee SC, Ni M, Li W, Shertz C, Heitman J, 2010. The evolution of sex: a perspective from the fungal kingdom. *Microbiology and Molecular Biology Reviews* **74**: 298-340.

Li W, Metin B, White TC, Heitman J, 2010. Organization and evolutionary trajectory of the mating type (*MAT*) locus in dermatophyte and dimorphic fungal pathogens. *Eukaryotic Cell* **9**:46-58.

Lin X, Heitman J, 2007. Mechanisms of homothallism in fungi and transitions between heterothallism and homothallism. In: Heitman J, Kronstad JW, Taylor J, Casselton L (eds), *Sex in fungi: molecular determination and evolutionary implications*. ASM Press, Washington, DC, pp. 35–57.

Liu F, Wang M, Damm U, Crous PW, Cai L, 2016. Species boundaries in plant pathogenic fungi: a *Colletotrichum* case study. *BMC Evolutionary Biology* **16**:81.

Porras-Alfaro A, Bayman P, 2011. Hidden Fungi, Emergent Properties: Endophytes and Microbiomes. *Annual Review of Phytopathology* **49**:291–315.

Mallet J, 2007. Species, concepts of. In Levin S et al. (eds), *Encyclopedia of Biodiversity*. Volume 5 (2nd edition). Academic Press. pp. 427-440. Online update 1, pp. 1-15, Elsevier, Oxford.

Martin T, Lu SW, van Tilbeurgh H, Ripoll DR, Dixelius C, Turgeon BG, Debuchy R, 2010. Tracing the origin of the fungal alpha 1 domain places its ancestor in the HMG-box super- family: implication for fungal mating-type evolution. *PLoS One* **5**: e15199.

Martin SH, Steenkamp ET, Wingfield MJ, Wingfield BD, 2013. Mate-recognition and species boundaries in the ascomycetes. *Fungal Diversity* **58**:1-12.

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

Metzenberg RL, Glass NL, 1990. Mating type and mating strategies in *Neurospora*. *BioEssays* **12**:53–59.

Michailides TJ, 2002. Panicle and shoot blight. In: Teviotdale BL, Michailides TJ, Pscheidt JW (eds), *Compendium of nut crop diseases in temperate zones*. American Phytopathological Society, St. Paul, pp 68–69.

Moral J, Agustí-Brisach C, Pérez-Rodríguez M, Xaviér C, Raya MC, Rhouma A, Trapero A, 2017. Identification of Fungal Species Associated with Branch Dieback of Olive and Resistance of Table Cultivars to *Neofusicoccum mediterraneum* and *Botryosphaeria dothidea*. *Plant Disease* **101**: 306-316.

Netto MSB, Lima WG, Correia KC, da Silva CFB, Thon M, Martins RB, Miller RNG, Michereff SJ, Câmara MPS, 2017. Analysis of phylogeny, distribution, and pathogenicity of *Botryosphaeriaceae* species associated with gummosis of *Anacardium* in Brazil, with a new species of *Lasiodiplodia*. *Fungal Biology* **121**:437-451.

Ni M, Feretzaki M, Sun S, Wang X, Heitman J, 2011. Sex in fungi. *Annual Review of Genetics* **45**:405-430.

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

Palmer JM, Kubatova A, Novakova A, Minnis AM, Kolarik M, Lindner DL, 2014. Molecular characterization of a heterothallic mating system in *Pseudogymnoascus destructans*, the fungus causing white-nose syndrome of bats. *G3 (Bethesda)* **4**:1755-1763.

Parker IM, Gilbert GS, 2004. The evolutionary ecology of novel plant-pathogen interactions. *Annual Review of Ecology, Evolution, and Systematics* **35**: 675–700.

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

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ, 2009b. 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.

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ, 2007. *Botryosphaeriaceae* occurring on native *Syzygium cordatum* in South Africa and their potential threat to Eucalyptus. *Plant Pathology* **56**:624–636.

Pelleteret P, Crovadore J, Cochard B, Pasche S, Bovigny P-Y, Chablais R, Lefort F, 2017. Urban London plane tree dieback linked to fungi in the *Botryosphaeriaceae*. *Urban Forestry & Urban Greening* **22**:74-83.

Phillips AJL, 2002. *Botryosphaeria* species associated with diseases of grapevines in Portugal. *Phytopathologia Mediterranea* **41**:3-18.

Phillips AJL, Lopes J, Abdollahzadeh J, Bobev S, Alves A, 2012. Resolving the *Diplodia* complex on apple and other *Rosaceae* hosts. *Persoonia* **29**:29-38.

Phillips AJL, Alves A, Abdollahzadeh J, Slippers B, Wingfield MJ, Groenewald JZ, Crous PW, 2013. The *Botryosphaeriaceae*: genera and species known from culture. *Studies in Mycology* **76**:51-167.

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

Pöggeler S, O'Gorman CM, Hoff B, Kück U, 2011. Molecular organization of the mating-type loci in the homothallic ascomycete *Eupenicillium crustaceum*. *Fungal Biology* **115**: 615-624.

Ramirez-Prado JH, Moore GG, Horn BW, Carbone I, 2008. Characterization and population analysis of the mating-type genes in *Aspergillus flavus* and *Aspergillus parasiticus*. *Fungal Genetics and Biology* **45**:1292-1299.

Rau D, Attene G, Brown AHD, Nanni L, Maier FJ, Balmas V, Saba E, Schäfer W, Papa R. 2007. Phylogeny and evolution of mating-type genes from *Pyrenophora teres*, the causal agent of barley ''net blotch''disease. *Current Genetics* **51**:377-392.

Rodríguez-Gálvez E, Guerrero P, Barradas C, Crous PW, Alves A, 2017. Phylogeny and pathogenicity of *Lasiodiplodia* species associated with dieback of mango in Peru. *Fungal Biology* **121**:452-465.

Sakalidis ML, Hardy GESJ, Burgess T, 2011. Use of the Genealogical Sorting Index (GSI) to delineate species boundaries in the *Neofusicoccum parvum–Neofusicoccum ribis* species complex. *Molecular Phylogenetics and Evolution* **60**:333-344.

Sakalidis ML, Slippers B, Wingfield BD, Hardy GE St J, Burguess TI, 2013. The challenge of understanding the origin, pathways and extent of fungal invasions: global populations

of the *Neofusicoccum parvum – N. ribis* species complex. *Diversity and Distributions* **19**: 873–883.

Santos JM, Correia VG, Phillips AJL, 2010. Primers for mating-type diagnosis in *Diaporthe* and *Phomopsis*: their use in teleomorph induction in vitro and biological species definition. *Fungal Biology* **114**:255-270.

Scherrer S, Zippler U, Honegger R, 2005. Characterisation of the mating-type locus in the genus *Xanthoria* (lichen-forming ascomycetes, Lecanoromycetes). *Fungal Genetics and Biology* **42**:976-988.

Schmitt I, Crespo A, Divakar PK, Fankhauser JD, Herman-Sackett E, Kalb K, Nelsen MP, Nelson NA, Rivas-Plata E, Shimp AD, Widhelm T, Lumbsch HT, 2009. New primers for promising single-copy genes in fungal phylogenetics and systematics. *Persoonia* **23**:35-40.

Schoch CL, Seifert KA, Huhndorf S, Robert V, Spouge JL, Levesque CA, Chen W, Fungal Barcoding Consortium, 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America* **109**:6241-6246.

Seidl V, Seibel C, Kubicek CP, Schmoll M, 2009. Sexual development in the industrial workhorse *Trichoderma reesei*. *Proceedings of the National Academy of Sciences of the United States of America* **106**:13909-13914.

Shenoy BD, Jeewon R, Hyde KD, 2007. Impact of DNA sequence-data on the taxonomy of anamorphic fungi. *Fungal Diversity* **26**:1-54.

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

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

Slippers B, Burgess TI, Pavlic D, Ahumada R, Maleme H, Mohali SR, Rodas CA, Wingfield MJ. 2009. A diverse assemblage of *Botryosphaeriaceae* infects Eucalyptus in native and non-native environments. *Southern Forests* **71**:101-110.

Slippers B, Crous PW, Jami F, Groenewald JZ, Wingfield MJ, 2017. Diversity in the *Botryosphaeriales*: Looking back, looking forward. *Fungal Biology* **121**:307-321.

Smith H, Crous PW, Wingfield MJ, Coutinho TA, Wingfield BD, 2001. *Botryosphaeria eucalyptorum* sp. nov., a new species in the *B. dothidea*- complex on Eucalyptus in South Africa. *Mycologia* **93**:277-285.

Stergiopoulos I, Groenewald M, Staats M, Lindhout P, Crous PW, De Wit PJGM, 2007. Mating-type genes and the genetic structure of a world-wide collection of the tomato pathogen *Cladosporium fulvum*. *Fungal Genetics and Biology* **44**:415-429.

Stielow JB, Lévesque CA, Seifert KA, Meyer W, Irinyi L, Smits D, Renfurm R, Verkley GJM, Groenewald M, Chaduli D, Lomascolo A, Welti S, Lesage-Meessen L, Favel A, Al-Hatmi AMS, Damm U, Yilmaz N, Houbraken J, Lombard L, Quaedvlieg W, Binder M, Vaas LAI, Vu D, Yurkov A, Begerow D, Roehl O, Guerreiro M, Fonseca A, Samerpitak K, van Diepeningen AD, Dolatabadi S, Moreno LF, Casaregola S, Mallet S, Jacques N, Roscini L, Egidi E, Bizet C, Garcia-Hermoso D, Martín MP, Deng S, Groenewald JZ, Boekhout T, de Beer ZW, Barnes I, Duong TA, Wingfield MJ, de Hoog GS, Crous PW, Lewis CT, Hambleton S, Moussa TAA, Al-Zahrani HS, Almaghrabi OA, Louis-Seize G, Assabgui R, McCormick W, Omer G, Dukik K, Cardinali G, Eberhardt U, de Vries M, Robert V, 2015. One fungus, which genes? Development and assessment of universal primers for potential secondary fungal DNA barcodes. *Persoonia* **35**:242-263.

Strandberg R, Nygren K, Menkis A, James TY, Wik L, Stajich JE, Johannesson, 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.

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.

Turgeon BG, 1998. Application of mating type gene technology to problems in fungal biology. *Annual Review of Phytopathology* **36**:115-137.

Turgeon BG, Yoder OC, 2000. Proposed nomenclature for mating type genes of filamentous ascomycetes. *Fungal Genetics and Biology* **31**:1-5.

Úrbez-Torres JR, Gubler WD, 2009. Pathogenicity of *Botryosphaeriaceae* species isolated from grapevine cankers in California. *Plant Disease* **93**:584-592.

Vaghefi N, Ades PK, Hay FS, Pethybridge SJ, Ford R, Taylor PW, 2015. Identification of the *MAT1* locus in *Stagonosporopsis tanaceti*, and exploring its potential for sexual reproduction in Australian pyrethrum fields. *Fungal Biology* **119**:408-419.

Voigt K, Cozijnsen AJ, Kroymann J, Pöggeler S, Howlett BJ, 2005. Phylogenetic relationships between members of the crucifer pathogenic *Leptosphaeria maculans* species complex as shown by mating type (*MAT1-2*), actin and β-tubulin sequences. *Molecular Phylogenetics and Evolution* **37**:541-557.

White TJ, Bruns T, Lee S, Taylor J, 1990. Amplification and direct sequencing of fungal genes for phylogenies. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, (eds), *PCR protocols: a guide to methods and applications*. Academic Press, California, pp. 315−322.

Wik L, Karlsson M, Johannesson H, 2008. The evolutionary trajectory of the mating-type (*mat*) genes in *Neurospora* relates to reproductive behavior of taxa. *BMC Evolutionary Biology* **8**:109-120.

Wilson AM, Wilken PM, van der Nest MA, Steenkamp ET, Wingfield MJ, Wingfield BD, 2015. Homothallism: an umbrella term for describing diverse sexual behaviours. *IMA Fungus* **6**:207-214.

Wingfield MJ, De Beer ZW, Slippers B, Wingfield BD, Groenewald JZ, Lombard L, Crous PW, 2012. One fungus, one name promotes progressive plant pathology. *Molecular Plant Pathology* **13**:604-613.

Wyka SA, Broders KD, 2016. The new family Septorioideaceae, within the Botryosphaeriales and *Septorioides strobi* as a new species associated with needle defoliation of *Pinus strobus* in the United States. *Fungal Biology* **120**:1030-1040.

Yokoyama E, Arakawa M, Yamagishi K, Hara A, 2006. Phylogenetic and structural analyses of the mating-type loci in *Clavicipitaceae*. *FEMS Microbiology Letters* **264**:182-191.

Zhou S, Stanosz GR, 2001. Relationships among *Botryosphaeria* species and associated anamorphic fungi inferred from the analyses of ITS and 5.8S rDNA sequences. *Mycologia* **93**:516–527.

Zlatković M, Keča N, Wingfield MJ, Jami F, Slippers B, 2016. *Botryosphaeriaceae* associated with the die-back of ornamental trees in the Western Balkans. *Antonie van Leeuwenhoek* **109**:543-564.

CHAPTER 2

Diversity and phylogeny of *Neofusicoccum* species occurring in forest and urban environments in Portugal

Lopes A, Barradas C, Phillips AJL, Alves A, 2016. Diversity and phylogeny of *Neofusicoccum* species occurring in forest and urban environments in Portugal. *Mycosphere* **7**(7):906–920, Doi 10.5943/mycosphere/si/1b/10

Abstract

A collection of *Neofusicoccum* isolates was obtained from a large number of plant species, showing dieback and canker symptoms, in forest and urban environments in Portugal. A total of 351 isolates was characterised by BOX-PCR fingerprinting to evaluate their overall genetic diversity. Representatives of each group identified in this analysis were selected for multilocus sequence analyses, using sequences of the ribosomal internal transcribed spacer region (ITS rDNA) and partial sequences of the translation elongation factor 1-alpha (*tef1-α*) and β-tubulin (*tub2*). Phylogenetic analysis of multilocus sequence data identified five species within the collection of isolates, namely *N. australe*, *N. eucalyptorum*, *N. kwambonambiense*, *N. luteum*, and *N. parvum*. Of these *N. australe* and *N. eucalyptorum* were the most frequent representing the vast majority of the isolates. Several new fungus-host associations were established for all of the *Neofusicoccum* species found. Here we report for the first time the occurrence of *N. eucalyptorum* on a host outside the family Myrtaceae. The results of this study show that the genus *Neofusicoccum* appears to be common and widespread on a broad range of hosts representing a potential threat to susceptible plants. Additionally, ornamental plants in urban environments are shown to be hosts of a diverse assemblage of *Neofusicoccum* species.

Keywords – *Botryosphaeriaceae*, endophytic, host-association, ornamentals, pathogenic

Introduction

The genus *Neofusicoccum* is a member of the *Botryosphaeriaceae* (Botryosphaeriales, Dothideomycetes) comprising numerous species found on a wide range of plant hosts of agricultural, forestry, ecological and economic importance (Crous et al. 2006; Slippers & Wingfield 2007; Slippers et al. 2013). Host infections are thought to occur predominantly through horizontal transmission, i.e. individual infections *via* spores (ascospores or conidia) but also through the seeds (vertical transmission). Inside the plant, they have the ability to

colonize without producing any external symptoms, remaining inside the host as endophytes (Slippers & Wingfield 2007). Endophytism could be, however, an important feature since these fungi can be moved easily and inconspicuously around the world in seeds, cuttings and even fruits, subsequently infecting both native and non-native trees in their new environments (Burgess et al. 2005; Slippers et al. 2013). The change from endophytic to pathogenic phase is often related to stress such as drought, extreme temperature fluctuations, nutrient deficiencies and mechanical injuries. Infected plants can exhibit a multiplicity of disease symptoms such as fruit rots, leaf spots, seedling damping-off and collar rot, cankers, blight of shoots and seedlings, gummosis, blue-stain of the sapwood, dieback and tree death (Slippers & Wingfield 2007).

Neofusicoccum is known to include a large number of phylogenetically closely related and morphologically similar cryptic species rendering phenotypic characteristics such as morphology, growth and culture appearance inadequate for species identification. Thus, species discrimination is based on a multilocus sequencing approach (Pavlic et al. 2009a, 2009b).

Within the 27 species currently accepted in the genus some are known to have very wide host and geographic ranges while others show some host preference. For example, *N. parvum* was reported from 90 hosts in 29 countries on six continents by Sakalidis et al. (2013). In contrast, *N. eucalypticola* and *N. mangiferae* were associated only with *Eucalyptus* spp. and *Mangifera indica* respectively and were geographically more restricted (Phillips et al. 2013). In general, species of *Neofusicoccum* are a constant presence in almost all kind of woody plants. For instance, they are frequently isolated from eucalypts (Iturritxa et al. 2011), almond (Inderbitzin et al. 2010), avocado (McDonald & Eskalen 2011), walnut (Yu et al. 2015), grapevine (Mondello et al. 2013; Berraf-Tebbal et al. 2014), olive (Triki et al. 2015), blueberry (Pérez et al. 2014), mango (Ismail et al. 2013), rubber tree (Ngobisa et al. 2013) and peach (Thomidis et al. 2011). Although these fungi have been relatively well studied on economically important crops, much less is known about their prevalence in others groups of plants namely the gymnosperms (Slippers et al. 2005; Golzar & Burgess 2011) and ornamental species (Zlatković et al. 2016). Although of low economic value, ornamental plant species

should not be ignored as they provide a wide range of ecosystem services (Zlatković et al. 2016).

In Portugal the forest area occupies 35.4% of the total territory according to the National Forest Inventory 2013 (ICNF, 2013). *Eucalyptus globulus*, a non-native species, is the most abundant followed by *Quercus suber* and *Pinus pinaster*. Apart from the forest species Portugal has important crops such as grapevine and olive among others. In spite of this, knowledge about the diversity of the genus *Neofusicoccum* in Portugal is scarce. The few known studies were done on grapevines (Phillips 2002), conifers (Alves et al. 2013) and more recently on eucalypts (Barradas et al. 2016).

To gain further knowledge about the diversity of the genus *Neofusicoccum* in Portugal the main goal of this study was to identify the species associated with a wide diversity of plants. For this, a survey was performed on several species of forest and crop trees and also on ornamental species.

Materials and methods

Fungal isolation and morphological characterization

Isolates were obtained from samples collected from plants in natural forest ecosystems and ornamentals planted in urban environments (e.g. parks, gardens, streetscapes). The following hosts were sampled: *Acacia longifolia*, *Aesculus hippocastanum*, *Castanea sativa*, *Ferula communis*, *Fraxinus angustifolia*, *Fraxinus excelsior*, *Fraxinus ornus*, *Hydrangea macrophylla*, *Malus domestica*, *Melia azederach*, *Olea europaea*, *Populus alba*, *Populus tremula*, *Pyracantha coccinea*, *Quercus robur*, *Rosa* sp., *Tilia platyphyllos* and *Ulmus minor*. This study also included isolates from diseased and healthy *Eucalyptus globulus* that were previously obtained by Barradas et al. (2016). The remaining isolates were obtained from hosts showing disease symptoms on stems, trunks and branches such as canker and dieback (Figure 2.1). Samples from diseased plants were initially screened for the presence of fruiting bodies (ascomata and/or conidiomata) and when present single spore isolations were made as described previously (Phillips et al. 2013). In the absence of fruiting bodies isolations were made by plating pieces of diseased tissues following surface sterilization as described by

Alves et al. (2013). Cultures were routinely grown and maintained on half-strength potato-dextrose agar (PDA) (HIMEDIA, India).

Figure 2.1 - a. *Aesculus hippocastanum* with trunk canker. **b.** Detailed view of trunk canker. **c,d.** *Tilia platyphyllos* with symptoms of dieback of twigs and branches associated with *N. luteum* and *N. australe*. **e.** Ascus and ascospores of *N. australe* on *Ferula communis*. **f.** Developing conidia of *N. luteum* on *Melia azedarach*. **g.** Conidia of *N. australe* from *Acacia longifolia*.

To assign isolates to the genus *Neofusicoccum* conidial micromorphological characteristics and mode of conidiogenesis were observed with a Nikon 80i microscope and images captured with a Nikon DS-Ri1 camera. Isolates were induced to sporulate by plating

them on ¼ strength PDA (Merck, Germany) containing sterilised pine needles and incubating at room temperature (about 20–25ºC) under diffused daylight until pycnidia developed. For microscopy, pycnidia were mounted in 100% lactic acid.

Molecular characterization

Genomic DNA was extracted from fresh mycelium grown on half-strength PDA plates for 5 d at approximately 23°C, according to Alves et al. (2004). All PCR reactions were carried out in 25 µL reaction mixtures with NZYTaq 2× Green Master Mix (2.5 mM MgCl2; 200 µM dNTPs; 0.2 U/µL DNA polymerase) (Lisbon, Portugal) in a Bio-Rad C-1000 TouchTM Thermal Cycler (Hercules, CA, USA). Negative controls with sterile water instead of the template DNA were used in every PCR reaction.

BOX-PCR fingerprinting was done as described previously (Alves et al. 2007) using primer BOXA1R. Representatives of each group identified in this analysis were selected for multilocus sequence analyses (Table 2.1). Primers ITS1 and ITS4 (White et al. 1990) were used for amplification and sequencing of the ITS region of the ribosomal RNA as described by Alves et al. (2004). Part of the translation elongation factor 1-alpha (*tef1-α*) was amplified with the primers EF1-688F and EF1-1251R (Alves et al. 2008) with the conditions described by Phillips et al. (2005). Part of the β-tubulin gene (*tub2*) was amplified with primers T1 (O'Donnell & Cigelnik 1997) and Bt2b (Glass & Donaldson 1995) using the following conditions: 95°C for 3 min; 35 cycles at 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min; final extension at 72°C for 10 min.

PCR amplicons were purified with the DNA Clean & Concentrator™-5 kit (Zymo Research, CA, USA) before DNA sequencing. Both strands of the PCR products were sequenced at GATC Biotech (Cologne, Germany). The nucleotide sequences were read with FinchTV v.1.4.0 (Geospiza Inc.). All sequences were checked manually, and nucleotide arrangements at ambiguous positions were clarified using both primer direction sequences. Sequences were deposited in GenBank (Table 2.1) and the alignment in TreeBase (reviewer link: http://purl.org/phylo/treebase/phylows/study/TB2:S19901?x-access-code=9ab4bf7e4b620fcdbcc31cb6d30a1726&format=html).

Available ITS, *tef1-α* and *tub2* sequences from known and well-characterized *Neofusicoccum* species were retrieved from GenBank and also included in the phylogenetic analyses. Sequences of *Dothiorella sarmentorum* and *Do. iberica* were used as outgroup (Table 2.1).

Sequences were aligned with ClustalX v. 2.1 (Thompson et al. 1997), using the following parameters: pairwise alignment parameters (gap opening = 10, gap extension = 0.1) and multiple alignment parameters (gap opening = 10, gap extension = 0.2, transition weight = 0.5, delay divergent sequences = 25 %). Alignments were checked and edited with BioEdit Alignment Editor v. 7.2.5 (Hall 1999). Phylogenetic analyses of sequence data were done with MEGA6 v. 6.06 (Tamura et al. 2013). All gaps were included in the analyses. The model of DNA sequence evolution used for each dataset was determined by the software. Maximum likelihood (ML) analyses were performed on a Neighbour-Joining (NJ) starting tree automatically generated by the software. Bootstrap analyses with 1000 replicates were used to estimate the consistency of each node of the trees. Trees were visualized and edited with Interactive tree of life (iTOL) v3 (Letunic & Bork 2016).

Results

Fungal isolation

A collection of 351 isolates from different hosts was established. These isolates were initially selected based on culture characteristics typical of the *Botryosphaeriaceae*, namely fast-growing fluffy white aerial mycelium becoming grey to black with grey to indigo-grey or black pigment visible from the reverse side of Petri dishes. Most isolates sporulated within 2–3 weeks on ¼ strength PDA supplemented with pine needles. Micromorphology of conidia and conidiogenesis assigned them to *Neofusicoccum*. Host ranges of the species were determined from the SMML Fungus-Host Distribution Database as well as available literature. Several new fungus-host associations were established (Table 2.2).

Table 2.1 - Identity of the isolates studied and GenBank accession numbers of sequences used in phylogenetic analyses.

				GenBank Accession[b]		
Species	**Isolate[a]**	**Origin**	**Host**	**ITS**	***tef1-α***	***tub2***
Dothiorella iberica	**CBS115041**	Spain	*Quercus ilex*	*AY573202*	*AY573222*	*EU673096*
D. sarmentorum	**IMI63581b**	United Kingdom	*Ulmus* sp.	*AY573212*	*AY573235*	*EU673102*
N. algeriense	**CBS137504**	Algeria	*Vitis vinifera*	*KJ657702*	*KX505893*	*KX505915*
	CAA322	Portugal	*Malus domestica*	*KX505906*	*KX505894*	*KX505916*
	CAA366	Portugal	*Eucalyptus globulus*	*KT440951*	*KT441011*	KX871764
	PE32	Portugal	*Eucalyptus globulus*	*KT440952*	*KT441012*	KX871765
N. andinum	**CBS117453**	Venezuela	*Eucalyptus* sp.	*GU251155*	*GU251287*	*GU251815*
N. arbuti	**CBS116131**	USA	*Arbutus menziesii*	*AY819720*	*KF531792*	*KF531793*
	CBS117090	USA	*Arbutus menziesii*	*AY819724*	*KF531791*	*KF531794*
N. australe	**CMW6837**	Australia	*Acacia* sp.	*AY339262*	*AY339270*	*AY339254*
	CMW6853	Australia	*Sequoiadendron giganteum*	*AY339263*	*AY339271*	*AY339255*
	CAA178	Portugal	*Ferula communis*	KX871844	KX871800	KX871709
	CAA184	Portugal	*Ferula communis*	KX871845	KX871801	KX871710
	CAA191	Portugal	*Ferula communis*	KX871846	KX871802	KX871711
	CAA195	Portugal	*Ferula communis*	KX871847	KX871803	KX871712
	CAA197	Portugal	*Ferula communis*	KX871848	KX871804	KX871713
	CAA202	Portugal	*Melia azederach*	KX871849	KX871805	KX871714
	CAA231	Portugal	*Hydrangea macrophylla*	KX871850	KX871806	KX871715
	CAA233	Portugal	*Hydrangea macrophylla*	KX871851	KX871807	KX871716
	CAA242	Portugal	*Hydrangea macrophylla*	KX871852	KX871808	KX871717
	CAA319	Portugal	*Eucalyptus globulus*	*KT440900*	*KT440960*	KX871718
	CAA320	Portugal	*Eucalyptus globulus*	*KT440901*	*KT440961*	KX871719
	CAA326	Portugal	*Pyracantha coccinea*	KX871853	KX871809	KX871720
	CAA327	Portugal	*Pyracantha coccinea*	KX871854	KX871810	KX871721
	CAA332	Portugal	*Eucalyptus globulus*	*KT440902*	*KT440962*	KX871722
	CAA341	Portugal	*Eucalyptus globulus*	*KT440903*	*KT440963*	KX871723
	CAA344	Portugal	*Eucalyptus globulus*	*KT440904*	*KT440964*	KX871724
	CAA351	Portugal	*Eucalyptus globulus*	*KT440905*	*KT440965*	KX871725
	CAA357	Portugal	*Eucalyptus globulus*	*KT440906*	*KT440966*	KX871726
	CAA359	Portugal	*Eucalyptus globulus*	*KT440907*	*KT440967*	KX871727
	CAA392	Portugal	*Quercus robur*	KX871855	KX871811	KX871728
	CAA398	Portugal	*Eucalyptus globulus*	KX871856	KX871812	KX871729
	CAA400	Portugal	*Eucalyptus globulus*	*KT440908*	*KT440968*	KX871730

				GenBank Accession[b]		
Species	**Isolate[a]**	**Origin**	**Host**	**ITS**	***tef1-α***	***tub2***
	CAA401	Portugal	*Eucalyptus globulus*	*KT440909*	*KT440969*	KX871731
	CAA406	Portugal	*Eucalyptus globulus*	*KT440910*	*KT440970*	KX871732
	CAA420	Portugal	*Eucalyptus globulus*	*KT440911*	*KT440971*	KX871733
	CAA427	Portugal	*Eucalyptus globulus*	*KT440912*	*KT440972*	KX871734
	CAA434	Portugal	*Eucalyptus globulus*	*KT440913*	*KT440973*	*KX505927*
	CAA441	Portugal	*Eucalyptus globulus*	*KT440914*	*KT440974*	KX871735
	CAA455	Portugal	*Eucalyptus globulus*	*KT440915*	*KT440975*	*KX505928*
	CAA464	Portugal	*Eucalyptus globulus*	*KT440916*	*KT440976*	KX871736
	CAA466	Portugal	*Eucalyptus globulus*	*KT440917*	*KT440977*	KX871737
	CAA468	Portugal	*Olea europaea*	KX871857	KX871813	KX871738
	CAA475	Portugal	*Olea europaea*	KX871858	KX871814	KX871739
	CAA546	Portugal	*Eucalyptus globulus*	*KT440918*	*KT440978*	KX871740
	CAA549	Portugal	*Eucalyptus globulus*	*KT440919*	*KT440979*	KX871741
	CAA550	Portugal	*Eucalyptus globulus*	KX871859	KX871815	KX871742
	CAA571	Portugal	*Eucalyptus globulus*	KX871860	KX871816	KX871743
	CAA647	Portugal	*Eucalyptus globulus*	*KT440920*	*KT440980*	KX871744
	CAA648	Portugal	*Eucalyptus globulus*	*KT440921*	*KT440981*	KX871745
	CAA649	Portugal	*Eucalyptus globulus*	KX871861	KX871817	KX871746
	CAA723	Portugal	*Tilia platyphyllos*	KX871862	KX871818	KX871747
	CAA741	Portugal	*Acacia longifolia*	KX871863	KX871819	KX871748
	CAA743	Portugal	*Acacia longifolia*	KX871864	KX871820	KX871749
	CAA747	Portugal	*Acacia longifolia*	KX871865	KX871821	KX871750
	CAA749	Portugal	*Acacia longifolia*	KX871866	KX871822	KX871751
	CAA750	Portugal	*Acacia longifolia*	KX871867	KX871823	KX871752
	CAA751	Portugal	*Acacia longifolia*	KX871868	KX871824	KX871753
N. batangarum	**CBS124924**	Cameroon	*Terminalia catappa*	*FJ900607*	*FJ900653*	*FJ900634*
	CBS124923	Cameroon	*Terminalia catappa*	*FJ900608*	*FJ900654*	*FJ900635*
N. brasiliense	**CMM1338**	Brazil	*Mangifera indica*	*JX513630*	*JX513610*	*KC794031*
	CMM1285	Brazil	*Mangifera indica*	*JX513628*	*JX513608*	*KC794030*
N. cordaticola	**CBS123634**	South Africa	*Syzygium cordatum*	*EU821898*	*EU821868*	*EU821838*
	CBS123635	South Africa	*Syzygium cordatum*	*EU821903*	*EU821873*	*EU821843*
N. cryptoaustrale	CMW23785	South Africa	*Eucalyptus* sp.	*FJ752742*	*FJ752713*	*FJ752756*
	CMW20738	South Africa	*Eucalyptus citriodora*	*FJ752740*	*FJ752710*	*FJ752754*
N. eucalypticola	**CBS115679**	Australia	*Eucalyptus grandis*	*AY615141*	*AY615133*	*AY615125*

				GenBank Accession[b]		
Species	**Isolate[a]**	**Origin**	**Host**	**ITS**	***tef1-α***	***tub2***
	CBS115766	Australia	*Eucalyptus rossi*	*AY615143*	*AY615135*	*AY615127*
N. eucalyptorum	**CBS115791**	South Africa	*Eucalyptus grandis*	*AF283686*	*AY236891*	*AY236920*
	CAA369	Portugal	*Eucalyptus globulus*	*KT440922*	*KT440982*	KX871773
	CAA450	Portugal	*Eucalyptus globulus*	*KT440923*	*KT440983*	KX871774
	CAA517	Portugal	*Eucalyptus globulus*	*KT440924*	*KT440984*	KX871775
	CAA518	Portugal	*Eucalyptus globulus*	KX871883	KX871839	KX871776
	CAA520	Portugal	*Eucalyptus globulus*	*KT440925*	*KT440985*	KX871777
	CAA522	Portugal	*Eucalyptus globulus*	*KT440926*	*KT440986*	KX871778
	CAA528	Portugal	*Eucalyptus globulus*	*KT440927*	*KT440987*	KX871779
	CAA532	Portugal	*Eucalyptus globulus*	*KT440928*	*KT440988*	KX871780
	CAA535	Portugal	*Eucalyptus globulus*	*KT440929*	*KT440989*	KX871781
	CAA536	Portugal	*Eucalyptus globulus*	*KT440930*	*KT440990*	KX871782
	CAA539	Portugal	*Eucalyptus globulus*	KX871884	KX871840	KX871783
	CAA542	Portugal	*Eucalyptus globulus*	*KT440931*	*KT440991*	KX871784
	CAA558	Portugal	*Eucalyptus globulus*	*KT440932*	*KT440992*	KX871785
	CAA561	Portugal	*Fraxinus excelsior*	KX871885	KX871841	KX871786
	CAA601	Portugal	*Eucalyptus globulus*	*KT440933*	*KT440993*	KX871787
	CAA604	Portugal	*Eucalyptus globulus*	*KT440934*	*KT440994*	KX871788
	CAA618	Portugal	*Eucalyptus globulus*	*KT440935*	*KT440995*	KX871789
	CAA624	Portugal	*Eucalyptus globulus*	*KT440936*	*KT440996*	KX871790
	CAA651	Portugal	*Eucalyptus globulus*	*KT440937*	*KT440997*	KX871791
	CAA680	Portugal	*Eucalyptus globulus*	*KT440938*	*KT440998*	KX871792
	CAA683	Portugal	*Eucalyptus globulus*	*KT440939*	*KT440999*	KX871793
	CAA695	Portugal	*Eucalyptus globulus*	*KT440940*	*KT441000*	KX871794
	CAA709	Portugal	*Eucalyptus globulus*	*KT440941*	*KT441001*	*KX505920*
	CAA712	Portugal	*Eucalyptus globulus*	*KT440942*	*KT441002*	KX871795
	CAA713	Portugal	*Eucalyptus globulus*	*KT440943*	*KT441003*	*KX505921*
	CAA714	Portugal	*Eucalyptus globulus*	KX871886	KX871842	KX871796
	PE20	Portugal	*Eucalyptus globulus*	*KT440944*	*KT441004*	KX871797
	PE21	Portugal	*Eucalyptus globulus*	*KT440945*	*KT441005*	KX871798
	PE23	Portugal	*Eucalyptus globulus*	KX871887	KX871843	KX871799
N. hellenicum	**CERC1947**	Greece	*Pistacia vera*	*KP217053*	*KP217061*	*KP217069*
	CERC1948	Greece	*Pistacia vera*	*KP217054*	*KP217062*	*KP217070*
N. kwambonambiense	**CBS123639**	South Africa	*Syzygium cordatum*	*EU821900*	*EU821870*	*EU821840*

				GenBank Accession[b]		
Species	**Isolate[a]**	**Origin**	**Host**	**ITS**	***tef1-α***	***tub2***
	CBS123641	South Africa	*Syzygium cordatum*	*EU821919*	*EU821889*	*EU821859*
N. luteum	**CBS110299**	Portugal	*Vitis vinifera*	*AY259091*	*AY573217*	*DQ458848*
	CBS110497	Portugal	*Vitis vinifera*	*EU673311*	*EU673277*	*EU673092*
	CAA200	Portugal	*Melia azederach*	KX871869	KX871825	KX871754
	CAA203	Portugal	*Melia azederach*	KX871870	KX871826	KX871755
	CAA352	Portugal	*Quercus robur*	KX871871	KX871827	KX871756
	CAA360	Portugal	*Fraxinus ornus*	KX871872	KX871828	KX871757
	CAA362	Portugal	*Fraxinus ornus*	KX871873	KX871829	KX871758
	CAA365	Portugal	*Quercus robur*	KX871874	KX871830	KX871759
	CAA379	Portugal	*Melia azederach*	KX871875	KX871831	KX871760
	CAA412	Portugal	*Populus alba*	KX871876	KX871832	KX871761
	CAA505	Portugal	*Fraxinus ornus*	KX871877	KX871833	KX871762
	CAA628	Portugal	*Fraxinus excelsior*	*KX505911*	*KX505902*	*KX505929*
	CAA720	Portugal	*Tilia platyphyllos*	KX871878	KX871834	KX871763
N. macroclavatum	**CBS118223**	Australia	*Eucalyptus globulus*	*DQ093196*	*DQ093217*	*DQ093206*
	WAC12445	Australia	*Eucalyptus globulus*	*DQ093197*	*DQ093218*	*DQ093208*
N. mangiferae	CBS118531	Australia	*Mangifera indica*	*AY615185*	*DQ093221*	*AY615172*
	CBS118532	Australia	*Mangifera indica*	*AY615186*	*DQ093220*	*AY615173*
N. mediterraneum	**CBS121718**	Greece	*Eucalyptus* sp.	*GU251176*	*GU251308*	*GU251836*
	CBS121558	USA	*Vitis vinifera*	*GU799463*	*GU799462*	*GU799461*
N. nonquaesitum	**CBS126655**	USA	*Umbellularia californica*	*GU251163*	*GU251295*	*GU251823*
N. nonquaesitum	PD301	Chile	*Vaccinium corymbosum*	*GU251164*	*GU251296*	*GU251824*
N. occulatum	**CBS128008**	Australia	*Eucalyptus grandis* hybrid	*EU301030*	*EU339509*	*EU339472*
	MUCC286	Australia	*Eucalyptus pellita*	*EU736947*	*EU339511*	*EU339474*
N. parvum	**CMW9081**	New Zealand	*Populus nigra*	*AY236943*	*AY236888*	*AY236917*
	UCR-NP2	USA	*Vitis vinifera*	*AORE01001444*	*AORE01000046*	*AORE01001255*
	CBS110301	Portugal	*Vitis vinifera*	*AY259098*	*AY573221*	*EU673095*
	CAA189	Portugal	*Ferula communis*	KX871879	KX871835	KX871766
	CAA192	Portugal	*Ferula communis*	*KX505905*	*KX505892*	*KX505913*
	CAA384	Portugal	*Rosa* sp.	KX871880	KX871836	KX871767
	CAA386	Portugal	*Rosa* sp.	KX871881	KX871837	KX871768
	CAA608	Portugal	*Aesculus hippocastanum*	KX871882	KX871838	KX871769
	CAA692	Portugal	*Eucalyptus globulus*	*KT440950*	*KT441010*	KX871770

				GenBank Accession[b]		
Species	**Isolate[a]**	**Origin**	**Host**	**ITS**	***tef1-α***	***tub2***
	PE17	Portugal	*Eucalyptus globulus*	*KT440948*	*KT441008*	KX871771
	PE18	Portugal	*Eucalyptus globulus*	*KT440949*	*KT441009*	KX871772
N. pennatisporum	**MUCC510**	Australia	*Allocasuarina fraseriana*	*EF591925*	*EF591976*	*EF591959*
N. protearum	MUCC497	Australia	*Santalum acuminatum*	*EF591912*	*EF591965*	*EF591948*
N. ribis	**CBS115475**	USA	*Ribes* sp.	*AY236935*	*AY236877*	*AY236906*
	CBS121.26	Unknown	*Ribes rubrum*	*AF241177*	*AY236879*	*AY236908*
N. umdonicola	**CBS123645**	South Africa	*Syzygium cordatum*	*EU821904*	*EU821874*	*EU821844*
	CBS123646	South Africa	*Syzygium cordatum*	*EU821905*	*EU821875*	*EU821845*
N. vitifusiforme	5H022	California	*Juglans regia*	*KF778869*	*KF779059*	*KF778964*
	B8	Italy	*Vitis vinifera*	*KC469638*	*KX505897*	*KX505922*
	B9	Italy	*Vitis vinifera*	*KX505908*	*KX505898*	*KX505923*

[a]Acronyms of culture collections: **CAA** – Personal culture collection Artur Alves, Universidade de Aveiro, Portugal; **CBS** – Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands; **CERC** – China Eucalypt Research Center, Beijing, China; **CMM** – Coleção de culturas de fungos fitopatogénicos Prof. Maria Menezes, Universidade Federal Rural de Pernambuco, Brazil; **CMW** – Tree Pathology Co-operative Program, Forestry and Agricultural Biotechnology Institute, University of Pretoria, South Africa; **IMI** - International Mycological Institute, CBI-Bioscience, Egham, Bakeham Lane, UK; **MUCC** – Murdoch University Culture Collection, Perth, Australia; **PD** - University of California, Davis, USA; **UCR** – College of Natural and Agricultural Sciences, Riverside, California, USA; **WAC** - Department of Agriculture, Western Australia Plant Pathogen Collection, South Perth, Western Australia.

[b]Sequence numbers in italics were retrieved from GenBank. All others were determined in the present study.

Isolates in bold are ex-type cultures.

Table 2.2 - *Neofusicoccum* species isolated in this study and their respective hosts.

Species	Host
N. australe	*Acacia longifolia*[a,b]
	Castanea sativa[a,b]
	Eucalyptus globulus
	Ferula communis[a,b]
	Fraxinus excelsior[a,b]
	Hydrangea macrophylla[a,b]
	Melia azedarach[a,b]
	Olea europaea[b]
	Populus alba[a,b]
	Pyracantha coccinea[a,b]
	Quercus robur
	Tilia platyphyllos[a,b]
	Ulmus minor[a,b]
N. eucalyptorum	*Eucalyptus globulus*
	Fraxinus excelsior[a,b]
N. kwambonambiense	*Eucalyptus globulus*
N. luteum	*Fraxinus excelsior*[a,b]
	Fraxinus ornus[a,b]
	Melia azedarach[a,b]
	Populus alba[a,b]
	Populus tremula[a,b]
	Quercus robur
	Tilia platyphyllos[a,b]
N. parvum	*Aesculus hippocastanum*[b]
	Eucalyptus globulus
	Ferula communis[a,b]
	Malus domestica
	Melia azedarach[a,b]
	Rosa sp.[a,b]

[a]new host reported for the species; [b]first report from Portugal

Molecular characterization

BOX-PCR fingerprinting analysis divided the 351 isolates into 7 distinct clusters, which were presumed to represent distinct species. A total of 99 isolates representative of each group were selected for further molecular characterization (Table 2.1). The 7 clusters formed by the BOX-PCR fingerprinting analysis were resolved into 5 clades by multilocus (ITS, *tef1-α* and *tub2*) phylogenetic analysis (Figure 2.2).

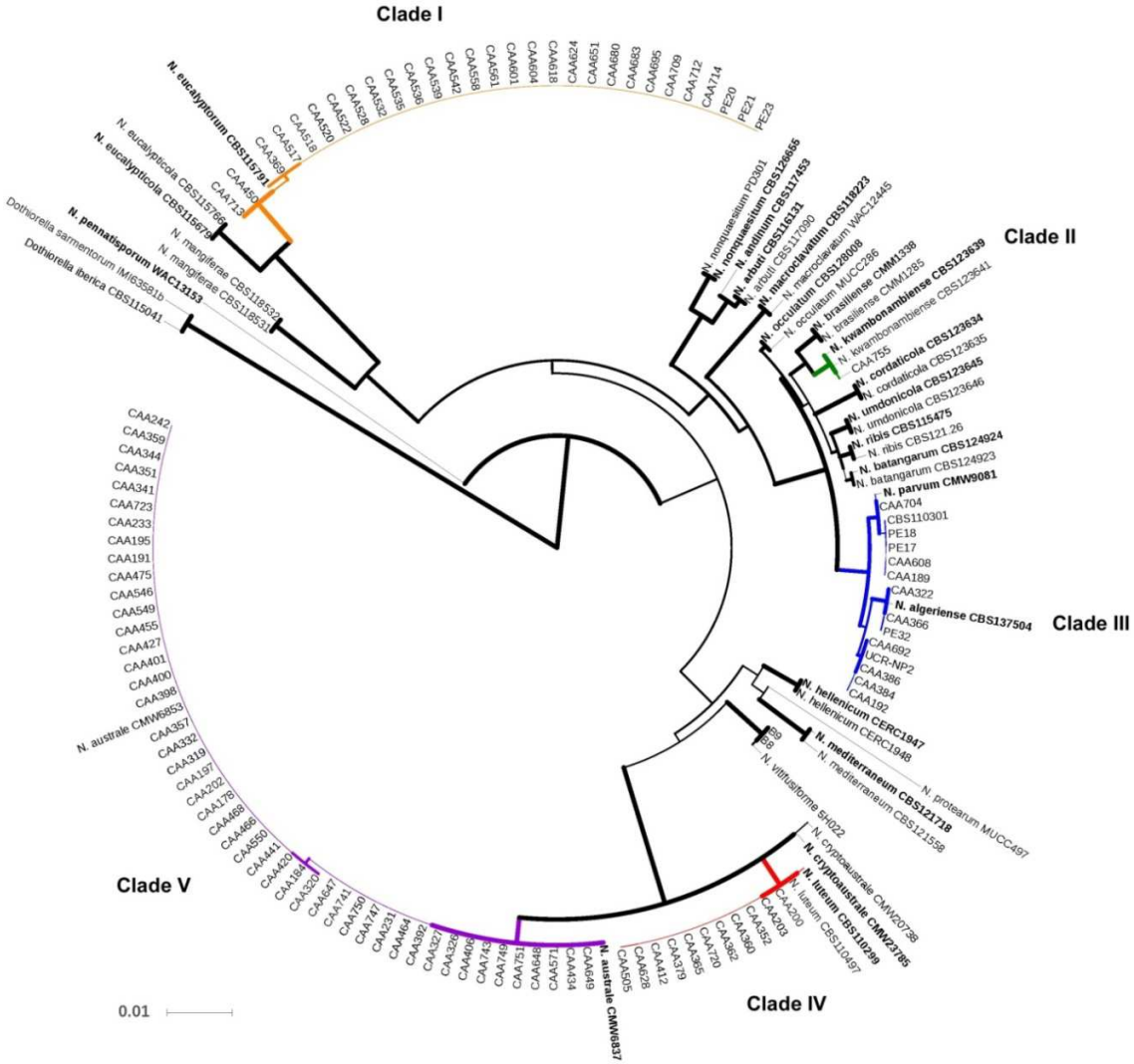

Figure 2.2 - Combined ITS, *tef1-α* and *tub2* maximum likelihood tree based on the Tamura 3-parameter model. A discrete Gamma distribution was used to model evolutionary rate differences among sites. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The thickness of branches is proportional to bootstrap support values.

Clades I, II, IV and V were clearly resolved and represent the species *N. eucalyptorum*, *N. kwambonambiense*, *N. luteum* and *N. australe* respectively. Clade III contained isolates belonging to two species (*N. parvum* and *N. algeriense*) and was further divided into 3 subclades. However, these showed incongruence between phylogenetic analyses results obtained from each individual locus (data not shown) and from the combined dataset (Figure 2.2). Moreover, there were no fixed alleles between the different subclades.

Within the clades formed by *N. eucalyptorum* and *N. australe* two subclades were also noticeable. However, a comparison of sequences of the three loci from members of each subclade showed minor differences between them. Thus, only 1 bp difference in the *tef1-α* of *N. eucalyptorum* isolates and 1 bp in the *tub2* sequence of *N. australe* isolates.

Discussion

In this study a collection of 351 isolates retrieved from a large diversity of plant hosts was characterised by morphological and PCR typing analysis. Selected representative isolates of each PCR typing group were further characterised by multilocus phylogenetic analyses. The isolates studied grouped into five clades, four of which clearly represented distinct species (Figure 2.2).

The clade containing *N. parvum* and *N. algeriense* (Clade III) was not clearly resolved, exhibiting incongruence between phylogenetic analysis results obtained from each individual locus and the combined dataset. A similar inconsistency was seen in phylogenetic analyses based on *MAT* genes (Lopes et al. 2017). By applying the principle of Phylogenetic Species Recognition (Taylor et al. 2000) where the transition from concordance to conflict determines the limits of species Lopes et al. (2017) considered that this clade represented a single species and synonymized *N. algeriense* with *N. parvum*. This study is in agreement with this previous finding.

Neofusicoccum australe and *N. eucalyptorum* were the most common species found. *Neofusicoccum australe* was originally regarded as native to Australia but since then it has been shown to have a widespread distribution occurring on a broad range of hosts (Sakalidis et al. 2011; Phillips et al. 2013). In Portugal, *N. australe* was found on *Rubus* sp. (Phillips et al. 2006), *Quercus robur* (Barradas et al. 2013), *Eucalyptus globulus* (Barradas et al. 2016),

Robinia pseudoacacia (van Niekerk et al. 2004) and several species of conifers (Alves et al. 2013). To our knowledge, this study is the first to report *N. australe* occurring on *A. longifolia*, *C. sativa*, *F. communis*, *F. excelsior*, *H. macrophylla*, *M. azederach*, *P. alba*, *P. coccinea*, *T. platyphyllos* and *U. minor*. It is also the first time that this species is found on *O. europaea* in Portugal. Another interesting finding was the isolation of *N. australe* from *A. longifolia*, being the first report of this species colonizing *Acacia* spp. outside of Australia. This could have serious repercussions on the dissemination of *N. australe* in Portugal since *Acacia* spp. are introduced exotic species that have spread rapidly to several new areas, from the coast to inland forests. Colonization of the invasive species *A. longifolia* will allow *N. australe* to be rapidly introduced into new geographic areas, possibly infecting new hosts.

Neofusicoccum eucalyptorum was first found on diseased *Eucalyptus grandis* and *E. nitens* in South Africa (Smith et al. 2001). Later, the species was isolated from cankers on native and planted eucalypts in eastern Australia (Slippers et al. 2004). Based on the dominance and wide distribution in eastern Australia, the authors suggested that the pathogen is probably native to this area (Slippers et al. 2004). Meanwhile, the presence of *N. eucalyptorum* was also detected on eucalypt species in other countries including Portugal (Barradas et al. 2016). Several authors suggested that the occurrence of the species on *Eucalyptus* in others parts of the world is a consequence of anthropogenic actions due to the large amounts of germplasm traded (Pérez et al. 2009; Barradas et al. 2016). Although this species is apparently specialized in the infection of *Eucalyptus* spp., it has also been associated with other genera in the Myrtaceae (Pérez et al. 2009; Pérez et al. 2010). In our study we report the occurrence of *N. eucalyptorum* in *Fraxinus excelsior* (Oleaceae) planted as ornamental. This is the first time that *N. eucalyptorum* is associated with a host outside of the family Myrtaceae. However, it is important to note that the *F. excelsior* tree from which the fungus was isolated was surrounded by a large number of eucalypts. Thus, it is possible that *F. excelsior* was colonized due to the high pressure of the surrounding inoculum or the fungus used it as a transition host. Further studies should be carried out to test the pathogenicity of *N. eucalyptorum* to this host and evaluate the impact that host jumps may have on the fungus host expansion and pathogenicity.

The species *N. luteum* and *N. parvum* were also found in this study although the number of isolates was lower. Both species are known to occur on a wide range of hosts worldwide (Phillips et al. 2013). *Neofusicoccum luteum* has been associated with dieback and canker mostly on crops (e.g. Phillips 2002; Úrbez-Torres et al. 2013) but also on ornamentals (Marincowitz et al. 2008; Varela et al. 2011). In Portugal, *N. luteum* has been found to infect conifers (Alves et al. 2013), *Quercus robur* (Barradas et al. 2013), grapevines, *Fraxinus angustifolia* and *Sophora japonica* (Phillips et al. 2002). In our study we found new host associations namely with *M. azederach, F. ornus, F. excelsior, P. alba, P. tremula* and *T. platyphyllos*, all of them planted as ornamentals.

Neofusicoccum parvum is probably the species within the genus with the widest geographic distribution, host range and proven ability to cause disease (Phillips et al. 2013; Sakalidis et al. 2013). It has been found associated with many forest species (Iturritxa et al. 2011), fruit trees (Ismail et al. 2013) and ornamental plants (Marincowitz et al. 2008, Zlatković et al. 2016). In Portugal, *N. parvum* was found associated with *Protea cynaroides* and *P. repens* (Crous et al. 2013), grapevines (Phillips 2002), conifers (Alves et al. 2013) and *E. globulus* (Barradas et al. 2016). To our knowledge, this study is the first to report the association of *N. parvum* with *Rosa* spp., *F. communis*, *M. azederach* and also the first occurrence of *N. parvum* on *A. hippocastanum* in Portugal. The fungus was only recently associated for the first time with *A. hippocastanum* in the Western Balkans, showing symptoms of canker and dieback (Zlatković et al. 2016). In our study it was isolated from severely affected trees with trunk cankers (Figure 2.1) and planted as ornamentals on streetscapes. However, since no pathogenicity tests were carried out we cannot conclude that *N. parvum* was the cause of the observed symptoms. This aspect should be addressed in future studies.

The presence, in this study, of species in such a wide diversity of hosts confirms that *Neofusicoccum* species are opportunistic fungi that can potentially colonize most plant hosts that it comes into contact with and represents a threat to vulnerable plants. This study reinforces the urgent need to understand the routes of introduction and dissemination of these fungi, not only in natural environments but also in the less studied urban environments where many potential hosts are planted as ornamentals.

Acknowledgements

This work was financed by European Funds through COMPETE and National Funds through the Portuguese Foundation for Science and Technology (FCT) to project ALIEN (PTDC/AGR-PRO/2183/2014), CESAM (UID/AMB/50017/2013 – POCI-01- 0145-FEDER-007638), Artur Alves (FCT Investigator Programme – IF/00835/2013), Anabela Lopes (PhD grant – SFRH/BD/85615/2012) and Carla Barradas (PhD grant – SFRH/BD/77939/2011).

References

Alves A, Correia A, Luque J, Phillips AJL, 2004. *Botryosphaeria corticola* sp. nov. on *Quercus* species, with notes and description of *Botryosphaeria stevensii* and its anamorph *Diplodia mutila*. *Mycologia* **96**:598–613.

Alves A, Phillips AJL, Henriques I, Correia A, 2007. Rapid differentiation of species of *Botryosphaeriaceae* by PCR fingerprinting. *Research in Microbiology* **158**:112–121.

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

Alves A, Barradas C, Phillips AJL, Correia A, 2013. Diversity of *Botryosphaeriaceae* species associated with conifers in Portugal. *European Journal of Plant Pathology* **135**:791-804.

Barradas C, Correia A, Alves A, 2013. First Report of *Neofusicoccum australe* and *Neofusicoccum luteum* Associated with Canker and Dieback of *Quercus robur* in Portugal. *Plant Disease* **97**:560.

Barradas C, Phillips AJL, Correia A, Diogo E, Bragança H, Alves A, 2016. Diversity and potential impact of *Botryosphaeriaceae* species associated with *Eucalyptus globulus* plantations in Portugal. *European Journal of Plant Pathology* **146**:245–257.

Berraf-Tebbal A, Guerreiro MA, Phillips AJL, 2014. Phylogeny of *Neofusicoccum* species associated with grapevine trunk diseases in Algeria, with description of *Neofusicoccum algeriense* sp. nov. *Phytopathologia Mediterranea* **53**:416−427.

Burgess TI, Barber PA, Hardy GEStJ, 2005. *Botryosphaeria* spp. associated with eucalypts in Western Australia, including the description of *Fusicoccum macroclavatum* sp. nov. *Australasian Plant Pathology* **34**:557–567.

Crous PW, Slippers B, Wingfield MJ, Rheeder J, Marasas WFO, Phillips AJL, Alves A, Burgess T, Barber P, Groenewald JZ, 2006. Phylogenetic lineages in the *Botryosphaeriaceae*. *Studies in Mycology* **55**:235–253.

Crous PW, Denman S, Taylor JE, Swart L, Bezuidenhout CM, Hoffman L, Palm ME, Groenewald JZ, 2013. Cultivation and Disease of Proteaceae: Leucadendron, Leucospermum, and Protea: Second Edition. CBS Biodiversity Series.

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

Golzar H, Burgess TI, 2011. *Neofusicoccum parvum*, a causal agent associated with cankers and decline of Norfolk Island pine in Australia. *Australasian Plant Pathology* **40**:484–489.

Hall TA, 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**:95–98.

ICNF, 2013. IFN6 – Áreas dos usos do solo e das espécies florestais de Portugal continental. Resultados preliminares. [pdf], 34 pp. Instituto da Conservação da Natureza e das Florestas. Lisboa (accessed April 2016).

Inderbitzin P, Bostock RM, Trouillas FP, Michailides TJ, 2010. A six locus phylogeny reveals high species diversity in *Botryosphaeriaceae* from California almond. *Mycologia* **102**:1350–1368.

Ismail AM, Cirvilleri G, Lombard L, Crous PW, Groenewald JZ, Polizzi G, 2013. Characterization of *Neofusicoccum* species causing mango dieback in Italy. *Journal of Plant Pathology* **95**:549-557.

Iturritxa E, Slippers B, Mesanza N, Wingfield MJ, 2011. First report of *Neofusicoccum parvum* causing canker and die-back of *Eucalyptus* in Spain. *Australasian Plant Disease Notes* **6**:57–59.

Letunic I, Bork P, 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research* **44**:W242-245.

Lopes A, Phillips AJL, Alves A, 2017. Mating type genes in the genus *Neofusicoccum*: mating strategies and usefulness in species delimitation. *Fungal Biology* **121**:394-404.

Marincowitz S, Groenewald JZ, Wingfield MJ, Crous PW, 2008. Species of *Botryosphaeriaceae* occurring on Proteaceae. *Persoonia* **21**:111–118.

McDonald V, Eskalen A, 2011. *Botryosphaeriaceae* Species Associated with Avocado Branch Cankers in California. *Plant Disease* **95**:1465-1473.

Mondello V, Picolo SL, Conigliaro G, Alfonzo A, Torta L, Burruano S, 2013. First report of *Neofusiccoccum vitifusiforme* and presence of other *Botryosphaeriaceae* species associated with Botryosphaeria dieback of grapevine in Sicily (Italy). *Phytopathologia Mediterranea* **52**:388−396.

Ngobisa AICN, Abidin MAZ, Wong MY, Noordin MWDW, 2013. *Neofusicoccum ribis* Associated with Leaf Blight on Rubber (*Hevea brasiliensis*) in Peninsular Malaysia. *The Plant Pathology Journal* **29**:10-16.

O'Donnell K, Cigelnik E, 1997. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. *Molecular and Phylogenetics Evolution* **7**:103-116.

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

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ, 2009b. 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.

Pérez CA, Wingfield MJ, Slippers B, Altierd NA, Blanchette RA, 2009. *Neofusicoccum eucalyptorum*, a *Eucalyptus* pathogen, on native Myrtaceae in Uruguay. *Plant Pathology* **58**:964–970.

Pérez CA, Wingfield MJ, Slippers B, Altier NA, Blanchette RA, 2010. Endophytic and canker-associated *Botryosphaeriaceae* occurring on non-native *Eucalyptus* and native Myrtaceae trees in Uruguay. *Fungal Diversity* **41**:53–69.

Pérez SF, Meriño-Gergichevich C, Guerrero JC, 2014. Detection of *Neofusicoccum nonquaesitum* causing dieback and canker in highbush blueberry from Southern Chile. *Journal of Soil Science and Plant Nutrition* **14**:581-588.

Phillips AJL, 2002. *Botryosphaeria* species associated with diseases of grapevines in Portugal. *Phytopathologia Mediterranea* **41**:3-18.

Phillips AJL, Fonseca F, Povoa V, Castilho R, Nolasco G, 2002. A reassessment of the anamorphic fungus *Fusicoccum luteum* and description of its teleomorph *Botryosphaeria lutea* sp. nov. *Sydowia* **54**:59-77.

Phillips AJL, Alves A, Correia A, Luque J, 2005. Two new species of *Botryosphaeria* with brown, 1-septate ascospores and *Dothiorella* anamorphs. *Mycologia* **97**:513−529.

Phillips AJL, Oudemans PV, Correia A, Alves A, 2006. Characterisation and epitypification of *Botryosphaeria corticis*, the cause of blueberry cane canker. *Fungal Diversity* **21**:141-155.

Phillips AJL, Alves A, Abdollahzadeh J, Slippers B, Wingfield MJ, Groenewald JZ, Crous PW, 2013. The *Botryosphaeriaceae*: genera and species known from culture. *Studies in Mycology* **76**:51-167.

Sakalidis ML, Hardy GEStJ, Burgess TI, 2011. Class III endophytes, clandestine movement amongst hosts and habitats and their potential for disease; a focus on *Neofusicoccum australe*. *Australasian Plant Pathology* **40**:510–521.

Sakalidis ML, Slippers B, Wingfield BD, Hardy GE St J, Burguess TI, 2013. The challenge of understanding the origin, pathways and extent of fungal invasions: global populations of the *Neofusicoccum parvum – N. ribis* species complex. *Diversity and Distributions* **19**:873-883.

Slippers B, Fourie G, Crous PW, Coutinho TA, Wingfield BD, Carnegie AJ, Wingfield MJ, 2004. Speciation and distribution of *Botryosphaeria* spp. on native and introduced *Eucalyptus* trees in Australia and South Africa. *Studies in Mycology* **50**:343–58.

Slippers B, Summerell BA, Crous PW, Coutinho TA, Wingfield BD, Wingfield MJ, 2005. Preliminary studies on *Botryosphaeria* species from Southern Hemisphere conifers in Australasia and South Africa. *Australasian Plant Pathology* **34**:213–220.

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

Slippers B, Boissin E, Phillips AJL, Groenewald JZ, Lombard L, Wingfield MJ, Postma A, Burgess T, Crous PW, 2013. Phylogenetic lineages in the Botryosphaeriales: a systematic and evolutionary framework. *Studies in Mycology* **76**:31–49.

Smith H, Crous PW, Wingfield MJ, Coutinho TA, Wingfield BD, 2001. *Botryosphaeria eucalyptorum* sp. nov., a new species in the *B. dothidea*-complex on *Eucalyptus* in South Africa. Mycologia **93**:277–285.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**:2725-2729.

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.

Thomidis T, Michailides TJ, Exadaktylou E, 2011. *Neofusicoccum parvum* associated with fruit rot and shoot blight of peaches in Greece. *European Journal of Plant Pathology* **131**:661-668.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG, 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**:4876–4882.

Triki MA, HadjTaieb SK, Cheffi M, Gharbi Y, Rhouma A, 2015. First Report of Dieback of Olive Trees caused by *Neofusicoccum australe* in Tunisia. *Journal of Plant Pathology* **97**:209-220.

Úrbez-Torres JR, Peduto F, Vossen PM, Krueger WH, Gubler WD, 2013. Olive twig and branch dieback: etiology, incidence and distribution in California. *Plant Disease* **97**:231–244.

van Niekerk JM, Crous PW, Groenewald JZE, Fourie PH, Halleen F, 2004. DNA phylogeny, morphology and pathogenicity of *Botryosphaeria* species on grapevines. *Mycologia* **96**:781–798.

Varela CP, Fernández VR, Vásquez JPM, Casal OA, 2011. First Report of Dieback on Hybrid Rhododendrons Caused by *Neofusicoccum luteum* and *N. parvum* in Spain. *Plant Disease* **95**:221.

White TJ, Bruns T, Lee S, Taylor J, 1990. Amplification and direct sequencing of fungal genes for phylogenies. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, (eds) PCR protocols: a guide to methods and applications. Academic Press, California, 315−322.

Yu Z, Tang G, Peng S, Chen H, Zhai M, 2015. *Neofusicoccum parvum* causing canker of seedlings of Juglans regia in China. *Journal of Forestry Research* **26**:1019–1024.

Zlatković M, Keča N, Wingfield MJ, Jami F, Slippers B, 2016. *Botryosphaeriaceae* associated with the die-back of ornamental trees in the Western Balkans. *Antonie van Leeuwenhoek* **109**:543-564.

CHAPTER 3

Mating type genes in the genus *Neofusicoccum*: Mating strategies and usefulness in species delimitation

Lopes A, Phillips AJL, Alves A, 2017. Mating type genes in the genus *Neofusicoccum*: Mating strategies and usefulness in species delimitation. *Fungal Biology* **121**:394-404, Doi 10.1016/j.funbio.2016.08.011

Abstract

The genus *Neofusicoccum* includes species with wide geographical and plant host distribution, some of them of economic importance. The genus currently comprises 27 species that are difficult to identify based on morphological features alone. Thus, species differentiation is based on phylogenetic species recognition using multigene genealogies. In this study, we characterised the mating type genes of *Neofusicoccum* species. Specific primers were designed to amplify and sequence *MAT* genes in several species and a PCR-based mating type diagnostic assay was developed. Homothallism was the predominant mating strategy among the species tested. Furthermore, the potential of mating type gene sequences for species delimitation was evaluated. Phylogenetic analyses were performed on both *MAT* genes and compared with multigene genealogies using sequences of the ribosomal internal transcribed spacer region, translation elongation factor 1-alpha and beta-tubulin. Phylogenies based on mating type genes could discriminate between the species analysed and are in concordance with the results obtained with the more conventional multilocus phylogenetic analysis approach. Thus, *MAT* genes represent a powerful tool to delimit cryptic species in the genus *Neofusicoccum*.

Keywords – *Botryosphaeriaceae*, *MAT* genes, PCR-based assay, Phylogeny, Reproductive strategies

Introduction

Mating type (*MAT*) genes play important roles in the biology and evolution of fungi. They are responsible for determining mating compatibility and regulation of sexual reproduction (Coppin et al. 1997; Kronstad & Staben 1997; Lee et al. 2010). In Ascomycota the mating type system is bipolar due to the existence of a single *MAT* locus (*MAT1*) with two alternate forms (or idiomorphs), *MAT1-1* and *MAT1-2*. *MAT* genes encode transcriptional regulators that mediate the expression of genes required for sexual development (Coppin et al.

1997; Kronstad& Staben 1997; Lee et al. 2010). Although more genes may be present in the loci, the idiomorphs are usually characterised by the presence in *MAT1-1* of an open reading frame (ORF) encoding a protein with an alpha box motif (*MAT1-1-1* gene), and in *MAT1-2* of an ORF encoding a regulatory DNA-binding protein containing a high mobility group (HMG) motif (*MAT1-2-1* gene) (Coppin et al. 1997; Turgeon & Yoder 2000; Yun et al. 2000; McGuire et al. 2001; Lee et al. 2010). *MAT* gene sequences have been used to study evolutionary trends of mating systems (Fraser et al. 2007; Wik et al. 2008), to understand and predict population genetics and dynamics (Zhan et al. 2002; Groenewald et al. 2007), and to establish phylogenetic relationships and species boundaries (Steenkamp et al. 2000; Ueng et al. 2003; O'Donnell et al. 2004; Pöggeler et al. 2011; Kashyap et al. 2014). Several studies showed that phylogenetic analyses based on *MAT* sequences could have better resolution than those obtained with conventional molecular markers such as the ribosomal internal transcribed spacer (ITS rDNA) and some protein coding genes (Pöggeler 1999; Witthuhn et al. 2000; Du et al. 2005; Yokoyama et al. 2006). High evolution rates, with low levels of variation within species but high levels between species could be the reason for such good results (Turgeon 1998; Pöggeler 1999). Even in apparently asexual fungi, the analysis of *MAT* genes sequences proved to be useful in phylogenetic resolution (Groenewald et al. 2006). Studies on *MAT* genes also provide important knowledge of mode and frequency of reproduction in plant pathogens which is relevant to the epidemiology and control of plant diseases. Sexual populations may have evolutionary advantages over asexual populations since more fit genotypes may arise through recombination (Bennett et al. 2003).

The genus *Neofusicoccum* (Ascomycetes, Botryosphaeriales, *Botryosphaeriaceae*) includes species with a very wide range of geographical and host distribution, including woody plants of agricultural, forestry and economic importance. *Neofusicoccum* species are typically endophytes but under stress conditions some species can cause decline and dieback symptoms (Crous et al. 2006; Slippers &Wingfield 2007). This genus comprises 27 species that are difficult to identify based on morphological features alone. Species defined solely on morphological characteristics often encompass at least two, but often more cryptic species, giving an underestimation of the real diversity. For example, the *Neofusicoccum parvum-Neofusicoccum ribis* complex has been shown to include at least seven cryptic species when

molecular phylogenetics was applied (Sakalidis et al. 2011). To avoid unreliable identifications Taylor et al. (2000) encouraged the use of multiple concordant gene genealogies to establish species boundaries. Since the ITS region alone proved to be insufficient to resolve certain species complexes in *Neofusicoccum*, such as the *N. parvum-N. ribis* complex and *Neofusicoccum luteum-Neofusicoccum australe* complex (Slippers et al. 2004; Pavlic et al. 2009a,b), additional DNA sequences of the elongation factor 1-alpha (*tef1-α*), β-tubulin (*tub2*) and the RNA polymerase II subunit (*RPB2*) genes are often used (Pavlic et al. 2009a,b; Sakalidis et al. 2011).

Since there is no knowledge about *MAT* genes in the genus *Neofusicoccum* the aims of this study were (i) to analyse the *MAT* locus on the available sequenced genome of *N. parvum*, (ii) to develop a PCR-based assay for *MAT* genes amplification and sequencing, (iii) to evaluate the usefulness of *MAT* genes as phylogenetic markers, compared to conventional phylogenetic markers, and (iv) to gather information about sexual reproductive strategies in the genus.

Materials and methods

Fungal strains and culture conditions

The fungal strains used in this study are listed in Table 3.1. A total of 28 strains representing 13 species were used. Cultures were grown and maintained on half-strength potato-dextrose agar (PDA) (HIMEDIA, India).

Table 3.1 - Identity of the *Neofusicoccum* isolates studied and GenBank accession numbers of the sequences used in phylogenetic analyses.

			GenBank Accession Numbers[b]				
Species	**Isolate No.[a]**	**Host**	**ITS**	***tef1-α***	***tub2***	***MAT1-1-1***	***MAT1-2-1***
N. algeriense	**CBS137504**	*Vitis vinifera*	*KJ657702*	KX505893	KX505915	KX505936	KX505876
	CAA322	*Eucalyptus globulus*	KX505906	KX505894	KX505916	KX505937	KX505877
N. arbuti	**CBS116131**	*Arbutus menziesii*	*AY819720*	*KF531792*	*KF531793*	KX505942	-
	CBS117090	*Arbutus menziesii*	*AY819724*	*KF531791*	*KF531794*	KX505943	-
N. australe	CAA434	*Eucalyptus globulus*	*KT440913*	*KT440973*	KX505927	KX505951	KX505885
	CAA455	*Eucalyptus globulus*	*KT440915*	*KT440975*	KX505928	KX505952	KX505886
N. cryptoaustrale	LM03	*Pistacia lentiscus*	KX505912	KX505903	KX505930	KX505955	KX505890
	BL34	*Vitis vinifera*	*KJ638328*	KX505904	KX505931	KX505956	KX505891
N. eucalyptorum	CAA511	*Eucalyptus globulus*	KX505907	KX505896	KX505919	KX505944	KX505881
	CAA709	*Eucalyptus globulus*	*KT440941*	*KT441001*	KX505920	KX505945	KX505882
	CAA713	*Eucalyptus globulus*	*KT440943*	*KT441003*	KX505921	KX505946	KX505883
N. kwambonambiense	CAA755	*Eucalyptus globulus*	*KT440946*	*KT441006*	KX505917	KX505938	KX505878
N. luteum	**CBS110299**	*Vitis vinifera*	*AY259091*	*AY573217*	*DQ458848*	KX505953	KX505887
	CAA628	*Fraxinus excelsior*	KX505911	KX505902	KX505929	KX505954	KX505888
N. mangiferae	CBS118531	*Mangifera indica*	*AY615185*	*DQ093221*	*AY615172*	-	KX505889
N. mediterraneum	CAA001	*Pistacia vera*	KX505909	KX505899	KX505924	-	KX505884
	CAA002	*Pistacia vera*	*EU017537*	KX505900	KX505925	KX505949	-
	SPA9	*Pistacia lentiscus*	KX505910	KX505901	KX505926	KX505950	-
N. nonquaesitum	IMI500168	*Vaccinium corymbosum*	*JX217819*	KX505895	KX505918	KX505941	-
N. parvum	**CMW9081**	*Populus nigra*	*AY236943*	*AY236888*	*AY236917*	KX505932	KX505872
	UCR-NP2	*Vitis vinifera*	*AORE01001444*	*AORE01000046*	*AORE01001255*	*KB915846*	*KB916244*
	CBS110301	*Vitis vinifera*	*AY259098*	*AY573221*	*EU673095*	KX505933	KX505873

			GenBank Accession Numbers[b]				
Species	**Isolate No.[a]**	**Host**	**ITS**	***tef1-α***	***tub2***	***MAT1-1-1***	***MAT1-2-1***
N. ribis	**CBS115475**	*Ribes* sp.	*AY236935*	*AY236877*	*AY236906*	KX505939	KX505879
	CBS121.26	*Ribes rubrum*	*AF241177*	*AY236879*	*AY236908*	KX505940	KX505880
N. vitifusiforme	B8	*Vitis vinifera*	*KC469638*	KX505897	KX505922	KX505947	-
	B9	*Vitis vinifera*	KX505908	KX505898	KX505923	KX505948	-
***Neofusicoccum* sp.**	CAA192	*Ferula communis*	KX505905	KX505892	KX505913	KX505934	KX505874
	CAA704	*Eucalyptus globulus*	*KT440947*	*KT441007*	KX505914	KX505935	KX505875

[a] Acronyms of culture collections: **CAA** – Personal culture collection Artur Alves, Universidade de Aveiro, Portugal; **CAP** - Personal culture collection Alan Phillips, Universidade Nova de Lisboa, Portugal; **CBS** – Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands; **CMW** – Tree Pathology Co-operative Program, Forestry and Agricultural Biotechnology Institute, University of Pretoria, South Africa; **IMI** - International Mycological Institute, CABI-Bioscience, Egham, Bakeham Lane, UK; **UCR** – College of Natural and Agricultural Sciences, Riverside, California, USA.
[b] Sequence numbers in italics were retrieved from GenBank. All others were determined in the present study.
Isolates in bold are ex-type cultures.

Analyses of Neofusicoccum parvum genome and MAT genes primers design

The study of *MAT* genes and further primers design were based on the only sequenced genome of *N. parvum* UCR-NP2 available in GenBank Database (Blanco-Ulate et al. 2013; Accession number PRJNA187491). The genome was checked for the presence of one or both *MAT* genes using the genome annotation and BLAST searches with *MAT* genes of *Diplodia sapinea* (Bihon et al. 2014; KF551229; KF551228), a species in the same family. Once they were found the locus was carefully studied and the first set of primers was manually designed for the flanking regions of *MAT* genes. Characteristics such as length, GC content, melting temperature, potential hairpin formation, complementarity and potential self-annealing sites were then automatically checked with the free software OligoCalc: Oligonucleotide Properties Calculator (http://bio-tools.nubic.northwestern.edu/OligoCalc.html) and Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/pcr_primer_stats.html). Primers were synthesized by STAB Vida Lda (Lisbon, Portugal), reconstituted and diluted (10 pmol) in Tris-EDTA buffer, and stored at -20°C. The first set of primers was first tested on *N. parvum* and then on the remaining 12 species that were studied in this paper. Whenever primers gave rise to amplification, new sets of primers were designed based on consensus of already acquired *MAT* sequences. To obtain sequences for all the studied species and to develop a set of 'universal' primers for each *MAT* gene that would work for all *Neofusicoccum* species, a large number of primers was designed and tested. All the primers designed for amplification of the *MAT* genes are listed in Table 3.2. Combinations of primers and species in which they have worked are available in the Supplementary Tables S3.1 and S3.2. Phylogenetic analyses of *MAT* genes were based on sequences retrieved with the 'universal' primers (accessions numbers are given in Table 3.1).

DNA extraction, PCR amplification, and sequencing

Genomic DNA was extracted from fresh mycelium grown on half-strength PDA plates for 5 d at approximately 23°C, according to Alves et al. (2004). All PCR reactions were carried out in 25 µL reaction mixtures with NZYTaq 2x Green Master Mix (2.5 mM $MgCl_2$; 200 mM dNTPs; 0.2 U/µL DNA polymerase) (Lisbon, Portugal), in a Bio-Rad C-1000

Touch™ Thermal Cycler (Hercules, CA, USA). Negative controls with sterile water instead of the template DNA were used in every PCR reaction. Amplification of *MAT1-1-1* gene was performed with the primers Neo_MAT1_113F and Neo_MAT1_1215R. Thermal conditions were denaturation at 95°C for 3 min; 35 cycles at 94°C for 30 s, 48°C for 30 s, and 72°C for 1min 10 s; final extension at 72°C for 10 min. For *MAT1-2-1* amplification primers Neo_MAT2_-156F and Neo_MAT2_1070R were used with the following thermal conditions: denaturation at 95°C for 3 min; 35 cycles at 94°C for 30 s, 52°C for 30 s, and 72°C for 1 min 15 s; final extension at 72°C for 10 min.

Primers ITS1 and ITS4 (White et al. 1990) were used for amplification and sequencing of the ITS region of the ribosomal RNA as described by Alves et al. (2004). Part of the translation elongation factor 1-alpha (*tef1-α*) was amplified with the primers EF1-688F and EF1-1251R (Alves et al. 2008) with the conditions described by Phillips et al. (2005). Beta-tubulin gene was amplified with T1 and Bt2b primers (Glass & Donaldson 1995; O'Donnell & Cigelnik 1997) with the thermal conditions: 95°C for 3min; 35 cycles at 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min; final extension at 72°C for 10 min.

After amplification, 2 µL of each PCR product were separated by electrophoresis in 1.5 % agarose gels at 90 V for 1 h in 1x TAE buffer (40 mM Tris, 40 mM acetate, 2 mM EDTA, pH 8.0). A GeneRuler DNA Ladder Mix (Thermo Scientific, USA) was also included. Gels were stained with ethidium bromide and visualised on a BioRad Molecular Imager Gel Doc™ XR$^+$ to assess PCR amplification. The amplified PCR fragments were purified with the DNA Clean & Concentrator™-5 kit (Zymo Research, CA, USA) before sequencing.

Both strands of the PCR products were sequenced at GATC Biotech (Cologne, Germany). The nucleotide sequences were read with FinchTV v.1.4.0 (Geospiza http://www.geospiza.- com/finchtv). All sequences were checked manually, and nucleotide arrangements at ambiguous positions were clarified using both primer direction sequences. Sequences were deposited in GenBank (see Table 3.1 for accession numbers). The sequences generated in this study were deposited in GenBank.

Table 3.2 - Primers used for amplification of *MAT* genes.

Locus	Primer name	Nucleotide Sequence (5' – 3') [a]	Location	Tm (°C)
MAT1-1-1	Neo_MAT1_+92F	TTGGCGCCACATACGCC	Downstream Flank Region	59.2
	Neo_MAT1_+227F	GTCAATGCAGTATCCCCAGC		56.0
	Neo_MAT1_+470F	CTTTGCTTCTGTGCTGTGCC		57.3
	Neo_MAT1_113F[b]	CACTCTCAACTGCTTCGTCG	*MAT1-1-1* gene	55.8
	Neo_MAT1_226F	CAGAAGGACAGGTCGGGC		58.2
	Neo_MAT1_240F	GACCTGTCCTTCTGTTGATGC		55.8
	Neo_MAT1_771F	TGCTGGGCATTCTGAGCAGC		60.8
	Neo_MAT1_1215R[b]	CGAAGGTCCGAGTA**N**TTG		51.2
	Neo_MAT1_1301R	CTTGATCGGACTGTCCAACC		55.5
	Neo_MAT1_1511R	CATTGTCAAAGTGGTCGGCG		57.0
	Neo_MAT1_-84F	GTGCAGTCCTACACGATTCC	Upstream Flank Region	55.5
	Neo_MAT1_-273R	GCATAAGTACTCGCCCAAGC		56.1
	Neo_MAT1_-283R	CGCTTGTGGCGCATAAGTACTCG		60.4
	Neo_MAT1_-1154R	GTTCATCTGCATCTGAGGATCG		55.4
	Neo_MAT1_-1892R	TACGATGTCGTGCATTGGGA	Upstream ORF	56.5
	Neo_MAT1_-2002R	TGGATTGGGTGGGGAATTGG		57.7
MAT1-2-1	Neo_MAT2_-359F	GGAAATACATACGCTCCTGTGG	Upstream ORF	55.4
	Neo_MAT2_-156F[b]	TATCGTTCTTGGAGCGACTCAGC	Upstream Flank Region	58.7
	Neo_MAT2_-77F	TCACTGCTTTGGCTGCACACC		60.9
	Neo_MAT2_119F	CTACGAGCAACAAATGCCATTGC	*MAT1-2-1* gene	57.4
	Neo_MAT2_268F	CTCAGCCTCTCATGAACCAG		55.1
	Neo_MAT2_1070R[b]	GCATTGTCAGGATAGTCCGC		55.9
	Neo_MAT2_1405R	CAAGCGAAGTGAAGTCGAAGC		56.5
	Neo_MAT2_+116R	AGGCAGTGGTCTTTCGTTCC	Downstream Flank Region	57.5
	Neo_MAT2_+162R	CCTTGATCGAAAGACGCAGAGTG		57.8
	Neo_MAT2_+975R	TGGGTGCGTCGTTGTTAGAGG	Downstream ORF	59.5

[a]N = A, C, G or T.

[b]Universal' primers.

Phylogenetic analyses

Sequences were aligned with ClustalX v. 2.1 (Thompson et al. 1997), using the following parameters: pairwise alignment parameters (gap opening = 10, gap extension = 0.1) and multiple alignment parameters (gap opening = 10, gap extension = 0.2, transition weight = 0.5, delay divergent sequences = 25 %). Alignments were checked and edited with BioEdit Alignment Editor v. 7.2.5 (Hall 1999). Phylogenetic analyses of sequence data were done with MEGA6 v. 6.06 (Tamura et al. 2013). All gaps were included in the analyses. The model of DNA sequence evolution used for each data set was determined by the software (see Figures 3.3 and 3.4). Maximum Likelihood (ML) analyses were performed on a Neighbour-Joining (NJ) starting tree automatically generated by the software. A bootstrap analysis (1000 replicates) was used to estimate the consistency of each node of the trees. The alignments of all datasets were deposited in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S19498).

Results

Analyses of Neofusicoccum parvum MAT loci and MAT genes sequencing

Analyses of the *N. parvum* UCR-NP2 genome revealed the presence of both *MAT1-1-1* and *MAT1-2-1* genes, thus showing it to be a homothallic species. A detailed analysis of the *MAT* loci structure showed that apart from *MAT1-1-1* and *MAT1-2-1* other *MAT* genes were present, namely *MAT1-1-4* and *MAT1-2-5*. *MAT* genes (*MAT1-1-1* and *MAT1-2-1*) are not located close to each other or fused, and it was not possible to determine if they are located in the same chromosome. Moreover, in the analysed scaffolds *MAT* genes are not flanked by the frequently found genes DNA lyase (APN1/2) and/or cytoskeleton protein (SLA2) (Figure 3.1).

Neofusicoccum parvum MAT1-1-1 gene is 1516 base pairs (bp) long with one intron whereas *MAT1-2-1* gene is slightly smaller with 1447 bp but has three introns (Figure 3.1). The knowledge gathered from the identification of both *MAT* genes allowed the development of specific primers for *MAT1-1-1* and *MAT1-2-1* genes, which were then tested with other species of the genus. Due to the great sequence variability between species it was necessary to

develop several primers until a 'universal' set of primers was obtained for both *MAT* genes (Table 3.2 and Figure S3.1).

The large number of possible primer combinations enabled sequencing the complete *MAT1-1-1* gene for all the studied species and complete *MAT1-2-1* for most of them. The length of the fragments obtained with the 'universal' set of primers was different for each *MAT* gene and between species (Table 3.3). Sequence alignments showed that the internal regions are more dissimilar than the regions closer to the end of the genes corroborating the idea that flanking regions of *MAT* loci are usually more conserved. On the other hand among isolates of the same species the sequences of both genes were identical (Supplementary data).

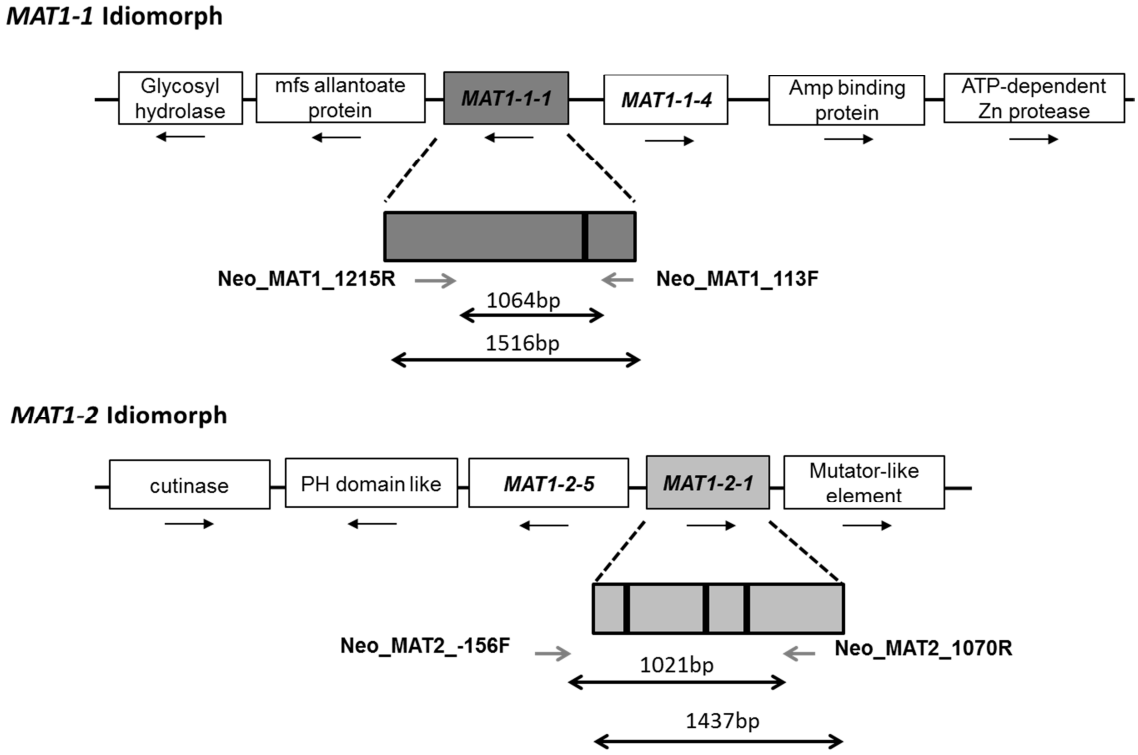

Figure 3.1 - Structure of the mating type locus of the homothallic species *Neofusicoccum parvum* UCR-NP2. Arrows below the genes indicate gene orientation (5'-3'). Grey arrows represent the binding sites of the 'Universal primers'. Introns are also represented inside the *MAT* genes by vertical black lines. Distances and sizes are not drawn to a scale

Mating strategies

This study showed that *Neofusicoccum* comprises both homothallic and heterothallic species (Table 3.3). From the 13 species studied the majority (eight) are homothallic and one is heterothallic. The remaining species are putative heterothallic. Thus, *Neofusicoccum algeriense, Neofusicoccum australe, Neofusicoccum cryptoaustrale, Neofusicoccum eucalyptorum, Neofusicoccum luteum, Neofusicoccum parvum, Neofusicoccum ribis,* and *Neofusicoccum vitifusiforme* strains possess both *MAT* genes and are therefore homothallic. For *Neofusicoccum mediterraneum* two strains contained the gene *MAT1-1-1* and the third strain the gene *MAT1-2-1*, thus confirming that the species is heterothallic (Figure 3.2). On the contrary, for *Neofusicoccum nonquaesitum, Neofusicoccum arbuti*, *Neofusicoccum vitifusiforme*, and *Neofusicoccum mangiferae* only one of the genes (either *MAT1-1-1* or *MAT1-2-1*) was amplified indicating that these are putative heterothallic species.

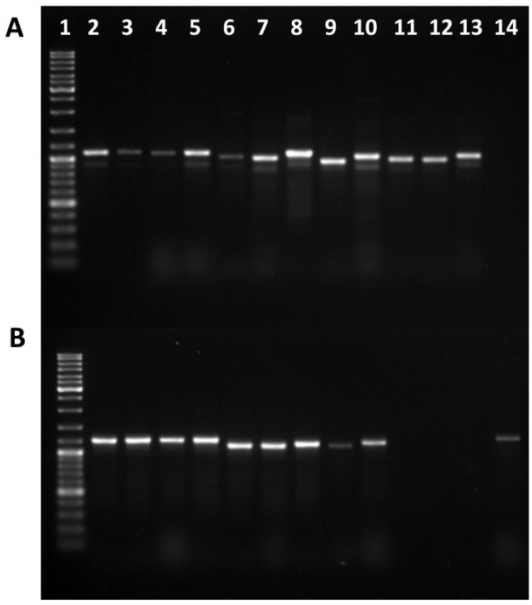

Figure 3.2 - PCR amplicons of partial *MAT1-1-1* (A) and *MAT1-2-1* (B) genes obtained with 'universal' primers. (1) DNA ladder; (2) *N. parvum*; (3) *N. ribis*; (4) *N. kwambonambiense*; (5) *N. algeriense*; (6) *N. luteum*; (7) *N. australe*; (8A) *N. mediterraneum*, CAA002; (8B) *N. mediterraneum*, CAA001; (9) *N. eucalyptorum*; (10) *N. cryptoaustrale*; (11A) *N. nonquaesitum*; (12A) *N. arbuti*; (13A) *N. vitifusiforme*; (14B) *N. mangiferae*.

Table 3.3 - Mating strategies and fragment lengths obtained for each *Neofusicoccum* species.

			Fragment Length (bp)	
Species	**Mating Strategy**	**Isolate**[a]	***MAT1-1-1***	***MAT1-2-1***
N. algeriense	Homothallic	**CBS137504** CAA322	1064	1021
N. arbuti	Heterothallic (?)	**CBS116131** CBS117090	950	-
N. australe	Homothallic	CAA434 CAA455	980	904
N. cryptoaustrale	Homothallic	LM03 BL34	1000	938
N. eucalyptorum	Homothallic	CAA511 CAA709 CAA713	950	901
N. kwambonambiense	Homothallic	CAA755	1052	1021
N. luteum	Homothallic	**CBS110299** CAA628	1001	904
N. mangiferae	Heterothallic (?)	CBS118531	-	984
N. mediterraneum	Heterothallic	CAA001 CAA002 SPA9	- 1044 1044	937 - -
N. nonquaesitum	Heterothallic (?)	IMI500168	956	-
N. parvum	Homothallic	UCR-NP2 **CMW9081** CBS110301	1064	1021
N. ribis	Homothallic	**CBS115475** CBS121.26	1064	1021
N. vitifusiforme	Heterothallic (?)	B8 B9	1004	-
***Neofusicoccum* sp.**	Homothallic	CAA192 CAA704	1064	1021

[a]Isolates in bold are ex-type cultures.

Phylogenetic analyses

Phylogenetic analyses of *MAT* genes and the three combined loci ITS, *tef1-α* and *tub2* were generally in concordance (Figures 3.3 and 3.4). *Neofusicoccum eucalyptorum, Neofusicoccum mangiferae, Neofusicoccum cryptoaustrale, Neofusicoccum luteum, Neofusicoccum australe, Neofusicoccum vitifusiforme, Neofusicoccum mediterraneum, Neofusicoccum arbuti, Neofusicoccum nonquaesitum, Neofusicoccum kwambonambiense* and *Neofusicoccum ribis* were clearly delimited in the phylogeny resulting from combined loci ITS, *tef1-α* and *tub2*, as well as *MAT1-1-1* and/or *MAT1-2-1* phylogenies. The only exception was the clade containing *Neofusicoccum parvum-Neofusicoccum algeriense* isolates that showed some incongruence between the *MAT1-1-1*, *MAT1-2-1* and combined ITS, *tef1-α*, *tub2* phylogenies (Figures 3.3 and 3.4). The isolates in this group formed three well-supported clades in the combined ITS, *tef1-α*, *tub2* phylogeny, corresponding to *N. parvum*, *N. algeriense* and a third one containing isolates previously identified as *N. parvum* including isolate UCR-NP2 whose genome has been sequenced. In contrast, in both *MAT1-1-1* and *MAT1-2-1* phylogenies these isolates were not resolved into the same three clades and formed a single, well-supported clade. Within this clade there were sub-clades that received poor or low bootstrap support. Also, the topologies of the *MAT1-1-1* and *MAT1-2-1* as well as individual *tef1-α* and *tub2* gene trees denote some conflict in the position of isolates within this group.

Taxonomy

Neofusicoccum parvum (Pennycook & Samuels) Crous, Slippers & A.J.L. Phillips, *Stud. Mycol.* **55**: 248 (2006).

MycoBank: MB500879.

= *Neofusicoccum algeriense* A. Berraf-Tebbal & A.J.L. Phillips, *Phytopathol. Mediterr.* **53**: 423 (2014).

MycoBank: MB808496.

Notes: Berraf-Tebbal et al. (2014) described *N. algeriense* from grapevines in Algeria as a distinct species based on DNA sequence data. According to the authors two unique, fixed alleles in ITS and one in *tef1-α* separate *N. algeriense* from all other *Neofusicoccum* species. However, we show in this study that *N. algeriense* is phylogenetically indistinguishable from *N. parvum*.

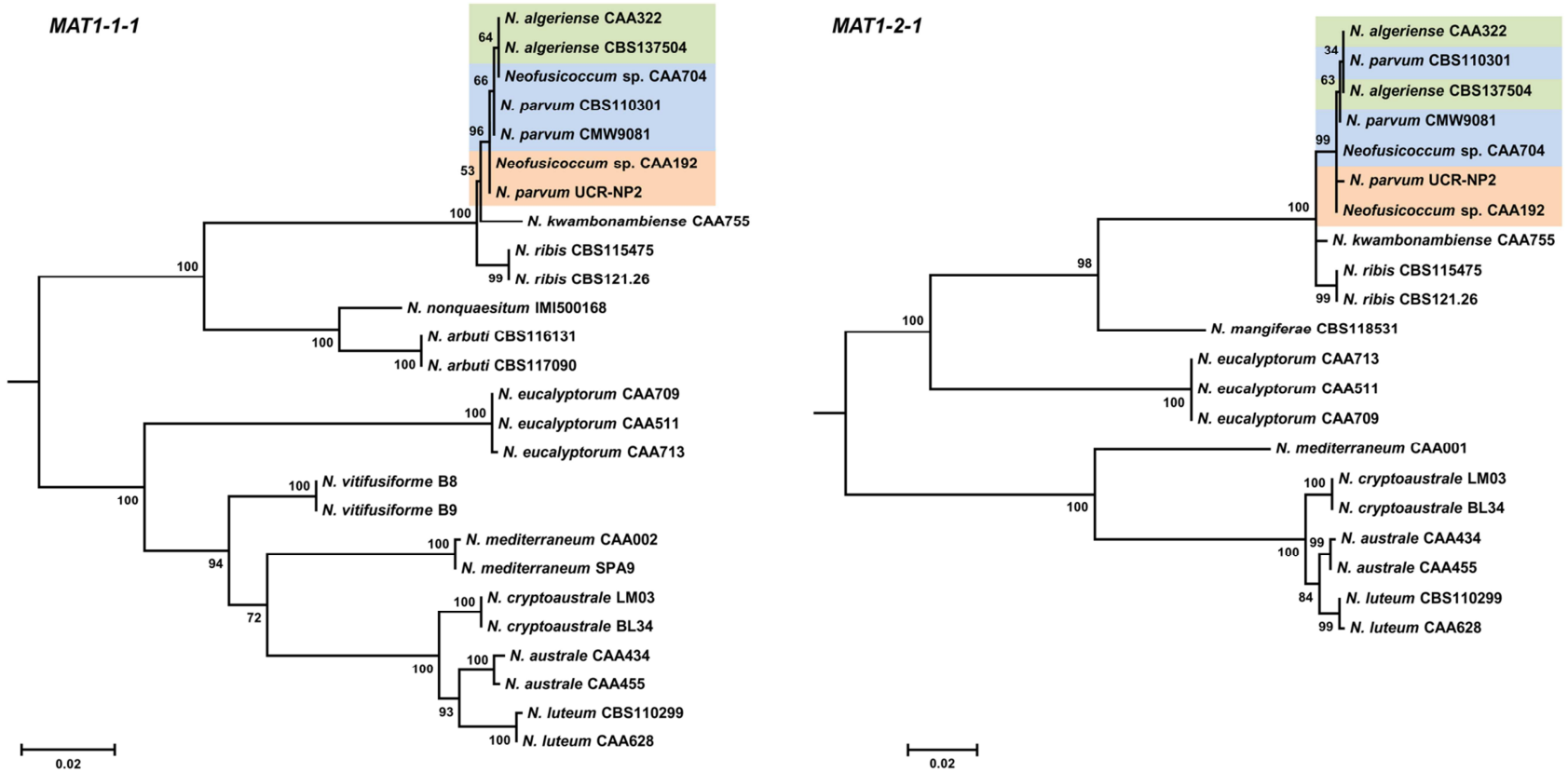

Figure 3.3 - Phylogenetic relationships of *Neofusicoccum* species based on sequence data from *MAT1-1-1* and *MAT1-2-1* genes. There were a total of 1123 positions (*MAT1-1-1*) and 1031 positions (*MAT1-2-1*) in the final datasets. Phylogenies were inferred using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model. The tree with the highest log likelihood is shown. There were a total of 1210 positions in the final dataset. A discrete Gamma distribution was used to model evolutionary rate differences among sites. Bootstrap values are given at the nodes. The tree is drawn to scale, with branch length measured in the number of substitutions per site.

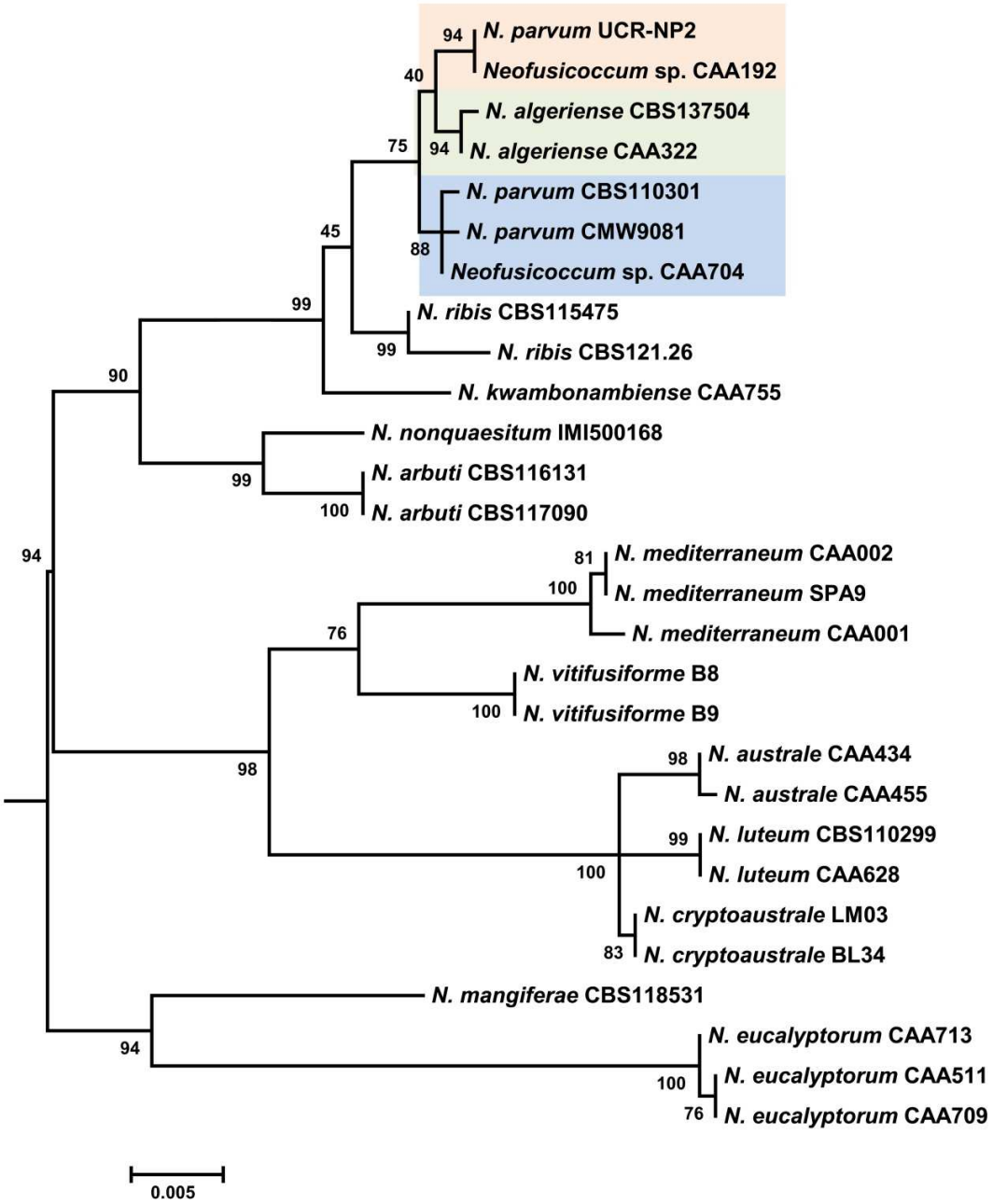

Figure 3.4 - Phylogenetic relationships of *Neofusicoccum* species based on combined sequence data from ITS, *tef1-α* and *tub2* regions. There were a total of 1210 positions in the final dataset. The phylogeny was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The tree with the highest log likelihood is shown. A discrete Gamma distribution was used to model evolutionary rate differences among sites. Bootstrap values are given at the nodes. The tree is drawn to scale, with branch length measured in the number of substitutions per site.

Discussion

The availability of the genome sequence of *Neofusicoccum parvum* UCR-NP2 created the opportunity to unveil *MAT* loci structure and mating strategies in the genus *Neofusicoccum*. In comparison to *MAT* loci structure of other Ascomycetes, the *N. parvum MAT* loci structure seems different in some aspects. *MAT* genes in *Neofusicoccum* are not fused or located adjacent to each other as frequently observed in most known *MAT* loci (Debuchy & Turgeon 2006). Also, the synteny of the flanking genes DNA lyase (APN1/2) and/or cytoskeleton protein (SLA2), usually conserved in members of the Ascomycota (Debuchy & Turgeon 2006; Yokoyama et al. 2006; Zaffarano et al. 2010; Bihon et al. 2014; Wey et al. 2016) was not seen in *N. parvum*. However, considering that there are gaps in the genome sequence we cannot rule out their presence upstream or downstream from the *MAT* loci. A reason for such unusual structure is not yet clear but some hypotheses have been proposed. In the genera *Neurospora* and *Sclerotinia*, for example, the *MAT* loci structure was explained as a possible result of the insertion of transposable elements (Gioti et al. 2012; Putman et al. 2015).

Apart from the *MAT1-1-1* and *MAT1-2-1* genes additional *MAT1-1-4* and *MAT1-2-5* genes were observed in the *MAT* loci of *Neofusicoccum*. In *Diplodia*, a genus from the same family, a similar structure of the *MAT* loci has been reported (Bihon et al. 2014). However, in Ascomycetes the absence of these additional mating genes may also occur (Yun et al. 1999; Galagan et al. 2005; Rydholmet al. 2007; Pöggeler et al. 2011).

Knowledge of *MAT* genes structure in *N. parvum* allowed us to develop 'universal' primers that can be used for *MAT* genes amplification and sequencing of other species of the genus. The sequence alignments showed that the inner regions of the genes are highly dissimilar between species. Some of this variability could possibly be explained by the presence of introns. In the *MAT1-1-1* gene most species had a single intron that occupied the same position. The sequence was the same for closely related species such as *N. parvum*, *Neofusicoccum kwambonambiense* and *Neofusicoccum ribis* and more divergent for the others. A second intron was present only in *Neofusicoccum arbuti*. On the other hand, the *MAT1-2-1* gene showed much larger variation. Almost all species had three introns, except *Neofusicoccum eucalyptorum* that had only two. Typically the first intron was more conserved

in position but with the exception again of the group *N. parvum, N. kwambonambiense*, and *N. ribis* all species had great variation in length, position and sequence in the other two introns. The biological significance and mechanisms of the presence of introns at specific positions in one species but absence in closely related taxa are not yet clear but it has been postulated that introns can be gained and lost in different genomes in response to strong selective forces and that could be a significant driving force in the evolution of fungal genes (Stergiopoulos et al. 2007). For that reason and the variability introduced by them we decided to include the introns in the phylogenetic analyses.

Concerning reproductive strategies, the genus *Neofusicoccum* comprises homothallic and heterothallic species, with homothallism being the most common mating strategy within the species studied (Table 3.3). The uncertainty about which reproductive strategy was adopted remains for the species *Neofusicoccum nonquaesitum*, *Neofusicoccum arbuti*, *Neofusicoccum vitifusiforme*, and *Neofusicoccum mangiferae* where only one *MAT* gene was amplified. We cannot be entirely sure if these species actually have only one of the *MAT* genes (being heterothallic) or if the primers did not amplify the other one. Until more isolates are tested it is premature to draw any conclusions about their mating strategy. In fact, one of the biggest problems encountered when using *MAT* genes in phylogenetic analyses is that sometimes only one mating type is known, or only one isolate of a species is available, and this isolate carries only one of the two mating type genes. In these cases, it is not possible to compare and/or combine phylogenies obtained by each *MAT* gene and other molecular markers without missing information, or conclude about their reproductive strategies.

Phylogenetic analyses based on mating type gene sequences have been done in different ascomycetes (O'Donnell et al. 2004; Inderbitzin et al. 2005; Wik et al. 2008; Strandberg et al. 2010). Several studies revealed concordant topologies between the *MAT* genes and the most common combined molecular markers ITS, *tef1-α*, *tub2*, *RPB2* and glyceraldehyde-3-phosphate dehydrogenase (GPD) (O'Donnell et al. 2004; Inderbitzin et al. 2005) but in others topologies conflicted (Wik et al. 2008; Strandberg et al. 2010). In our study, we compared phylogenies based on *MAT* genes with those based on the commonly used molecular markers in the genus *Neofusicoccum*, ITS, *tef1-α* and *tub2* (Phillips et al. 2013) to evaluate the usefulness of *MAT* genes for phylogenetic analyses and delimitation of species.

Apart from the clear resolution of most species obtained with the dataset of combined genes ITS, *tef1-α* and *tub2*, also individual *MAT* gene genealogies clearly resolved the same species, proving their usefulness as phylogenetic markers for species delimitation in *Neofusicoccum* (Figures 3.3 and 3.4). However, the main problem in working with *MAT* genes highlighted above is reflected here. From the 11 well delimited species there was no *MAT1-1-1* sequence for *N. mangiferae* or *MAT1-2-1* sequence for *N. vitifusiforme*, *N. arbuti*, and *N. nonquaesitum*. This limits our approach since it hampers combination of *MAT* genes with each other or with the other three genes without losing information about some species.

The *Neofusicoccum parvum-Neofusicoccum algeriense* clade was the only one that denoted incongruences between phylogenetic analyses obtained with *MAT* genes and the combined genes ITS, *tef1-α* and *tub2*. As stated by Taylor et al. (2000) conflict among gene trees is possibly the result of recombination among individuals within a species, and the transition from concordance to conflict determines the limits of species and this is the principle of phylogenetic species recognition widely applied to fungi and specifically to *Neofusicoccum*. Thus, it is likely that the *N. parvum-N. algeriense* clade represents in fact a single species and not three species as suggested by combined ITS, *tef1-α*, *tub2* phylogenies. One could argue that *MAT* genes are too conserved and do not discriminate between the three phylogenetic species in this group. However, *MAT* genes are notoriously highly divergent and known to evolve quickly (Turgeon 1998). In fact, our data shows that these genes are highly divergent between species (even cryptic species such as *Neofusicoccum australe*, *Neofusicoccum cryptoaustrale*, and *Neofusicoccum luteum*) and quite conserved within isolates of the same species. *MAT1-1-1* and *MAT1-2-1* sequences of *N. parvum-N. algeriense* isolates have only minor differences between the three lineages and these are within the range seen for intraspecific variability in other *Neofusicoccum* species. Also, examining both *MAT* genes sequences as well as ITS, *tef1-α*, and *tub2* sequences revealed no fixed alleles in these three lineages (Table 3.4).

Table 3.4 - Single nucleotide polymorphisms (SNPs) from sequence data of ITS rDNA, *tef1-α*, *tub2*, *MAT1-1-1* and *MAT1-2-1*.

Species	**Isolate**[a]	**ITS**					***tef1-α***		***tub2***		***MAT1-1-1***		***MAT1-2-1***			
		49	108	158	376	377	226	230	369	387	418	717	41	170	231	732
N. parvum	**CMW9081**	T	G	T	C	C	A	A	T	T	G	C	A	A	T	T
N. parvum	CBS110301	A	A	T	C	C	A	A	T	T	G	C	A	G	T	T
***Neofusicoccum* sp.**	CAA704	A	G	T	C	C	A	A	T	T	A	C	A	A	C	T
***Neofusicoccum* sp.**	CAA192	A	G	T	C	C	G	C	C	C	G	T	A	A	C	T
N. parvum	UCRNP2	A	G	T	C	C	G	C	C	C	G	T	G	A	C	C
N. algeriense	**CBS137504**	A	G	A	T	T	G	C	T	T	A	C	A	G	T	T
N. algeriense	CAA322	A	G	T	T	T	G	C	T	T	A	C	A	G	T	T

[a]Isolates in bold are ex-type cultures.

An alternative hypothesis could be that this group of fungi is genetically isolated in nature, but retained the ancestral character of interbreeding. Thus although they could be seen as a biological species due to mating ability, or in this case by having identical *MAT* gene sequences as a proxy for mating, they would in fact represent distinct phylogenetic species. However, this seems unlikely because isolates from the different lineages were obtained from the same region (Portugal) and therefore would not be geographically isolated. Also, although these isolates came from different hosts this would not be sufficient to prevent gene flow because *Neofusicoccum* species are not host specific.

Although *MAT1-2-1* has been considered a better phylogenetic marker than *MAT1-1-1* in some studies (e.g. Martin et al. 2010), here we cannot draw the same conclusion because, as already mentioned, it was impossible to obtain both *MAT* genes for all species analysed and therefore a full comparison of *MAT* phylogenies could not be accomplished. However, based on the results presented *MAT1-2-1* and *MAT1-1-1* appear to be equally efficient in resolving species. Despite this, the fact that *MAT1-2-1* contains a larger number and less conserved introns than *MAT1-1-1* and the better PCR amplification (Figure 3.2) suggest that *MAT1-2-1* analysis is a good approach to accurate and reliable species differentiation within the genus *Neofusicoccum*.

Results from our study represent an advance on the knowledge of *MAT* genes and reproductive strategies in the filamentous ascomycete *Neofusicoccum*. We also developed a PCR- based mating type assay that allows the scoring of the mating type within populations of *Neofusicoccum* in a fast, robust and reliable way.

In future studies it would be important to gain further knowledge about the *MAT* genes flanking regions and analyse the entire structure of *MAT* loci in other species, to study more isolates from the hypothetical heterothallic species in order to confirm their mating strategies, and to test the effectiveness of primers in all known species of *Neofusicoccum*. Additionally, it would be relevant to evaluate the distribution of the mating type genes at the population level as well as to confirm if these genes are expressed and fully functional.

Acknowledgements

This work was financed by European Funds through COMPETE and National Funds through the Portuguese Foundation for Science and Technology (FCT) to project ALIEN (PTDC/AGR-PRO/2183/2014 – POCI-01-0145-FEDER-016788), CESAM (UID/AMB/50017/2013 – POCI-01- 0145-FEDER-007638), Artur Alves (FCT Investigator Programme – IF/00835/2013) and Anabela Lopes (PhD grant – SFRH/BD/85615/2012). The authors are thankful to Benedetto Linaldeddu (University of Sassari, Italy), Vincenzo Mondello (University of Palermo, Italy) and Akila Berraf-Tebbal (University of Blida, Algeria) for sharing some of their isolates.

References

Alves A, Correia A, Luque J, Phillips AJL, 2004. *Botryosphaeria corticol*a sp. nov. on *Quercus* species, with notes and description of *Botryosphaeria stevensii* and its anamorph *Diplodia mutila*. *Mycologia* **96**: 598–613.

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

Bennett RS, Yun S-H, Lee TY, Turgeon BG, Arseniuk E, Cunfer BM, Bergstrom GC, 2003. Identity and conservation of mating type genes in geographically diverse isolates of *Phaeosphaeria nodorum*. *Fungal Genetics and Biology* **40**: 25−37.

Berraf-Tebbal A., Guerreiro MA, Phillips AJL, 2014. Phylogeny of *Neofusicoccum* species associated with grapevine trunk diseases in Algeria, with description of *Neofusicoccum algeriense* sp. nov. *Phytopathologia Mediterranea* **53:** 416−427.

Bihon W, Wingfield MJ, Slippers B, Duong TA, Wingfield BD, 2014. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. *Fungal Genetics and Biology* **62**: 55−61.

Blanco-Ulate B, Rolshausen P, Cantu D, 2013. Draft Genome Sequence of *Neofusicoccum parvum* Isolate UCR-NP2, a Fungal Vascular Pathogen Associated with Grapevine Cankers. *Genome Announcements* **1**(3): e00339-13.

Coppin E, Debuchy R, Arnaise S, Picard M, 1997. Mating Types and Sexual Development in Filamentous Ascomycetes. *Microbiology and Molecular Biology Reviews* **61**(4): 411−428.

Crous PW, Slippers B, Wingfield MJ, Reeder J, Marasas WFO, Phillips AJL, Alves A, Burgess T, Barber P, Groenewald JZ, 2006. Phylogenetic lineages in the Botryosphaeriaceae. *Studies in Mycology* **55**: 235−253.

Debuchy R, Turgeon BG, 2006. Mating-Type Structure, Evolution, and Function in Euascomycetes. In: Kües/Fisher (eds.) *The Mycota I - Growth, Differentiation and Sexuality*. Springer-Verlag, Berlin, Heidelberg, pp 293−323.

Du M, Schardl CL, Nuckles EM, Vaillancourt LJ, 2005. Using mating-type gene sequences for improved phylogenetic resolution of *Collectotrichum* species complexes. *Mycologia* **97**(3): 641−658.

Fraser JA, Stajich JE, Tarcha EJ, Cole GT, Inglis DO, Sil A, Heitman J, 2007. Evolution of the Mating Type Locus: Insights Gained from the Dimorphic Primary Fungal Pathogens *Histoplasma capsulatum*, *Coccidioides immitis*, and *Coccidioides posadasii*. *Eukaryotic Cell* **6**(4): 622−629.

Galagan JE, Calvo SE, Cuomo C, Ma L-J, Wortman JR, Batzoglou S, Lee S-I, Bastürkmen M., Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scazzocchio C, Farman M, Butler J, Purcell S, Harris S, Braus GH, Draht O, Busch S, D'Enfert C, Bouchier C, Goldman GH, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Peñalva MA, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW, 2005.

Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* **438**: 1105–1115.

Gioti A, Mushegian AA, Strandberg R, Stajich JE, Johannesson H, 2012. Unidirectional evolutionary transitions in fungal mating systems and the role of transposable elements. *Molecular Biology and Evolution* **2**: 3215–3226.

Glass NL, Donaldson G, 1995. Development of primer sets designed for use with PCR to amplify conserved genes from filamentous ascomycetes. *Applied and Environmental Microbiology* **61**: 1323−1330.

Groenewald M, Groenewald JZ, Harrington TC, Abeln ECA, Crous PW, 2006. Mating type gene analysis in apparently asexual *Cercospora* species is suggestive of cryptic sex. *Fungal Genetics and Biology* **43**: 813−825.

Groenewald M, Barnes I, Bradshaw RE, Brown AV, Dale A, Groenewald JZ, Lewis KJ, Wingfield BD, Wingfield MJ, Crous PW, 2007. Characterization and Distribution of Mating Type Genes in the Dothistroma Needle Blight Pathogens. *Phytopathology* **97**(7): 825−834.

Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95–98.

Inderbitzin P, Harkness J, Turgeon BG, Berbee ML, 2005. Lateral transfer of mating system in *Stemphylium*. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 11390–11395.

Kashyap PL, Rai S, Kumar S, Srivastava AK, Anandaraj M, Sharma AK, 2014. Mating type genes and genetic markers to decipher intraspecific variability among *Fusarium udum* isolates from pigeon pea. *Journal of Basic Microbiology* **54**: 1−11.

Kronstad JW, Staben C, 1997. Mating Type in Filamentous Fungi. *Annual Review of Genetics* **31**: 245−276.

Lee SC, Ni M, Li W, Shertz C, Heitman J, 2010. The Evolution of Sex: a Perspective from the Fungal Kingdom. *Microbiology and Molecular Biology Reviews* **74**(2): 298–340.

Martin T, Lu SW, van Tilbeurgh H, Ripoll DR, Dixelius C, Turgeon BG, Debuchy R, 2010. Tracing the origin of the fungal alpha 1 domain places its ancestor in the HMG-box superfamily: implication for fungal mating-type evolution. *PLoS ONE* **5**: e15199.

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.

O'Donnell K, Cigelnik E, 1997. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. *Molecular Phylogenetics and Evolution* **7**: 103−116.

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

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ, 2009a. Molecular and phenotypic characterization of three phylogenetic species discovered within the *Neofusicoccum parvum*/*N. ribis* complex. *Mycologia* **101**(5): 636–647.

Pavlic D, Slippers B, Coutinho TA, Wingfield MJ, 2009b. 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.

Phillips AJL, Alves A, Correia A, Luque J, 2005. Two new species of *Botryosphaeria* with brown, 1-septate ascospores and *Dothiorella* anamorphs. *Mycologia* **97**: 513−529.

Phillips AJL, Alves A, Abdollahzadeh J, Slippers B, Wingfield MJ, Groenewald JZ, Crous PW, 2013. The Botryosphaeriaceae: genera and species known from culture. *Studies in Mycology* **76**: 51−167.

Pöggeler S, 1999. Phylogenetic relationships between mating-type sequences from homothallic and heterothallic ascomycetes. *Current Genetics* **36**: 222−231.

Pöggeler S, O'Gorman CM, Hoff B, Kück U, 2011. Molecular organization of the mating-type loci in the homothallic Ascomycete *Eupenicillium crustaceum*. *Fungal Biology* **115**: 615−624.

Putman AI, Tredway LP, Carbone I, 2015. Characterization and distribution of mating-type genes of the turfgrass pathogen *Sclerotinia homoeocarpa* on a global scale. *Fungal Genetics and Biology* **81**: 25−40.

Rydholm C, Dyer PS, Lutzoni F, 2007. DNA Sequence Characterization and Molecular Evolution of MAT1 and MAT2 Mating-Type Loci of the Self-Compatible Ascomycete Mold *Neosartorya fischeri*. *Eukaryotic Cell* **6**(5): 868−874.

Sakalidis ML, Hardy GEStJ, Burgess TI, 2011. Use of the Genealogical Sorting Index (GSI) to delineate species boundaries in the *Neofusicoccum parvum-Neofusicoccum ribis* species complex. *Molecular Phylogenetics and Evolution* **60**(3): 333–344.

Slippers B, Fourie G, Crous PW, Coutinho TA, Wingfield BD, Wingfield MJ, 2004. Multiple gene sequences delimit *Botryosphaeria australis* sp. nov. from *B. lutea*. *Mycologia* **96**: 1028–1039.

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

Steenkamp ET, Wingfield BD, Coutinho TA, Zeller KA, Wingfield MJ, Marasas WFO, Leslie JF, 2000. PCR-Based Identification of MAT-1 and MAT-2 in the *Gibberella fujikuroi* Species Complex. *Applied Environmental Microbiology* **66**(10): 4378-4382.

Stergiopoulos I, Groenewald M, Staats M, Lindhout P, Crous PW, De Wit PJGM, 2007. Mating-type genes and the genetic structure of a world-wide collection of the tomato pathogen *Cladosporium fulvum*. *Fungal Genetics and Biology* **44**: 415−429.

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.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**(12): 2725−2729.

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.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG, 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**: 4876–4882.

Turgeon BG, 1998. Application of mating type gene technology to problems in fungal biology. *Annual Review of Phytopathology* **36**: 115−137.

Turgeon BG, Yoder OC, 2000. Proposed Nomenclature for Mating Type Genes of Filamentous Ascomycetes. *Fungal Genetics and Biology* **31**: 1−5.

Ueng PP, Dai Q, Cui K-R, Czembor PC, Cunfer BM, Tsang H, Arseniuk E, Bergstrom GC, 2003. Sequence diversity of mating-type genes in *Phaeosphaeria avenaria*. *Current Genetics* **43**: 121−130.

Wey T, Schlegel M, Stroheker S, Gross A, 2016. MAT – gene structure and mating behavior of *Hymenoscyphus fraxineus* and *Hymenoscyphus albidus*. *Fungal Genetics and Biology* **87**: 54−63.

White TJ, Bruns T, Lee S, Taylor J, 1990. Amplification and direct sequencing of fungal genes for phylogenies. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, (eds) *PCR protocols: a guide to methods and applications*. Academic Press, California, pp 315−322.

Wik L, Karlsson M, Johannesson H, 2008. The evolutionary trajectory of the mating-type (*mat*) genes in *Neurospora* relates to reproductive behavior of taxa. *BMC Evolutionary Biology* **8**: 109−120.

Witthuhn RC, Harrington TC, Steimel JP, Wingfield BD, Wingfield MJ, 2000. Comparison of isozymes, rDNA spacer regions and MAT-2 DNA sequences as phylogenetic characters in the analysis of the *Ceratocystis coerulescens* complex. *Mycologia* **92**: 447–452.

Yokoyama E, Arakawa M, Yamagishi K, Hara A, 2006. Phylogenetic and structural analyses of the mating-type loci in Clavicipitaceae. *FEMS Microbiology Letters* **264**: 182−191.

Yun SH, Berbee ML, Yoder OC, Turgeon BG, 1999. Evolution of the fungal self-fertile reproductive life style from self-sterile ancestors. *Proceedings of the National Academy of Sciences* **96**: 5592−5597.

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.

Zaffarano PL, Duò A, Grünig CR, 2010. Characterization of the mating type (MAT) locus in the *Phialocephala fortinii* s.l. – *Acephala applanata* species complex. *Fungal Genetics and Biology* **47**: 761−772.

Zhan J, Kema GHJ, Waalwijk C, McDonald BA, 2002. Distribution of mating type alleles in the wheat pathogen *Mycosphaerella graminicola* over spatial scales from lesions to continents. *Fungal Genetics and Biology* **36**: 128−136.

Supplementary data

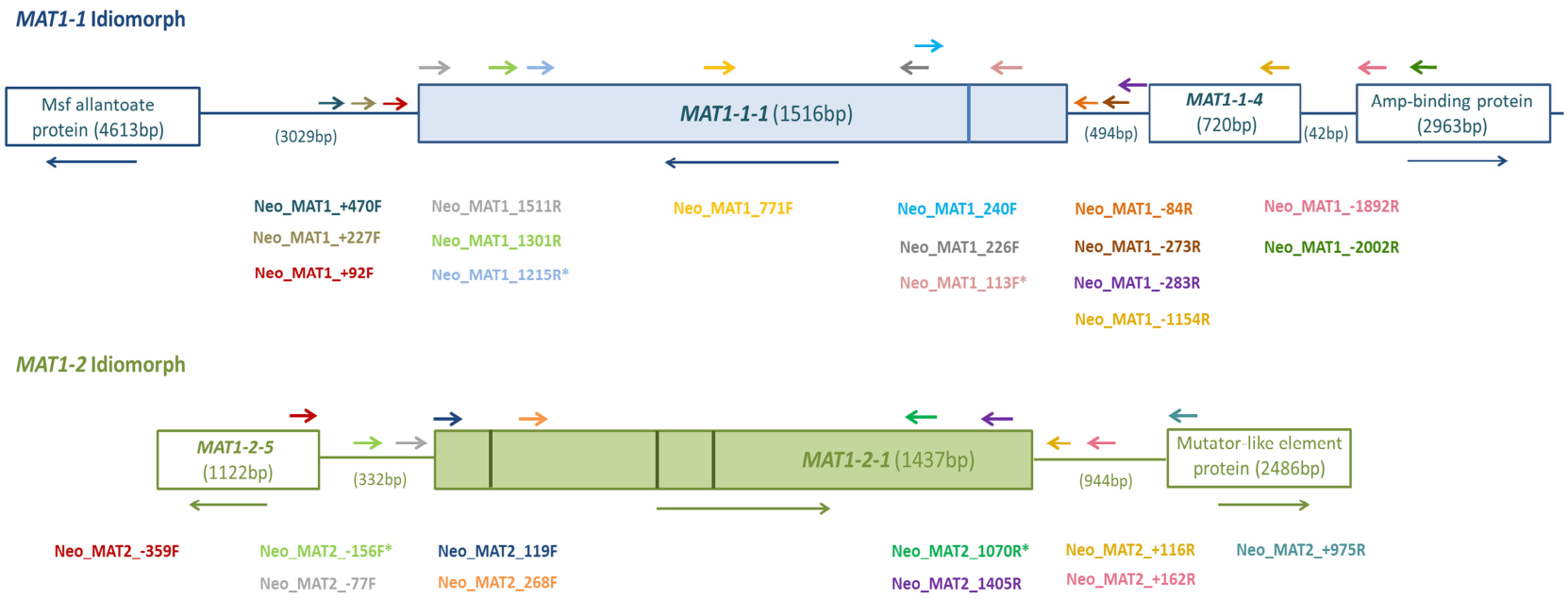

Figure S3.1 - Binding sites scheme of all the primers designed and tested for *MAT* loci. The lines inside the genes indicate introns. Arrows below the genes give the gene orientation (5'-3'). Distances and sizes are not drawn to a scale. *Universal primers

Table S3.1 - Possible primers combinations for *MAT1-1* idiomorph.

Primers \ Species	*N. parvum*	*N. ribis*	*N. kwamb.*	*N. luteum*	*N. australe*	*N. nonq.*	*N. medit.*	*N. eucal.*	*N. vitif.*	*N. arbuti*	*N. alge.*	*N. crypto.*
Neo_MAT1_+92F/Neo_MAT1_-1892R (3463bp)	+	+	NT	-	-	-	-	-	NT	-	NT	NT
Neo_MAT1_+92F/Neo_MAT1_-2002R (3573bp)	+	NT	NT	NT	NT	NT	-	-	NT	NT	NT	NT
Neo_MAT1_+92F/Neo_MAT1_-1154R (2723bp)	NT	NT	NT	-	-	NT	-	-	NT	NT	NT	NT
Neo_MAT1_+92F/Neo_MAT1_-273R (1844bp)	+	NT	+	+	+	-	-	-	-	-	NT	NT
Neo_MAT1_+92F/Neo_MAT1_-283R (1851bp)	NT	+	NT	NT	NT	NT	-	+	NT	-	-	NT
Neo_MAT1_+92F/Neo_MAT1_-84F (1655bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_+92F/Neo_MAT1_113F (1458bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_+92F/Neo_MAT1_226F (1347bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_+470F/Neo_MAT1_-2002R (3948bp)	+	NT	NT	-	-	-	-	-	NT	-	NT	NT
Neo_MAT1_+470F/Neo_MAT1_-1892R (3838bp)	+	NT	NT	NT	NT	NT	-	-	NT	NT	NT	NT
Neo_MAT1_+470F/Neo_MAT1_-1154R (3098bp)	-	NT	NT	NT	NT	NT	-	-	NT	NT	NT	NT
Neo_MAT1_+470F/Neo_MAT1_-273R (2219bp)	-	-	NT	-	-	-	-	-	-	-	NT	NT
Neo_MAT1_+470F/Neo_MAT1_-84F (2030bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_+470F/Neo_MAT1_113F (1833bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_+470F/Neo_MAT1_226F (1722bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_-84F/Neo_MAT1_1511R (1555bp)	+	NT	NT	-	-	+	-	-	-	+	NT	NT
Neo_MAT1_-84F/Neo_MAT1_1301R (1345bp)	NT	NT	NT	-	-	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_-84F/Neo_MAT1_1215R (1261bp)	NT	NT	-	NT	NT	NT	NT	+	-	+	NT	NT
Neo_MAT1_1511R/Neo_MAT1_-273R (1744bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1511R/Neo_MAT1_-1154R (2623bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1511R/Neo_MAT1_-1892R (3363bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1511R/Neo_MAT1_-2002R (3473bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_113F/Neo_MAT1_1301R (1148bp)	+	NT	NT	+	+	NT	+	-	NT	-	NT	NT
Neo_MAT1_113F/Neo_MAT1_1511R (1358bp)	NT	NT	NT	-	-	NT	NT	-	-	NT	NT	NT
Neo_MAT1_113F/Neo_MAT1_1215R (1064bp)	+	+	+	+	+	+	+	+	+	+	+	+
Neo_MAT1_1301R/Neo_MAT1_-273R (1534bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	-	NT	NT
Neo_MAT1_1301R/Neo_MAT1_-1154R (2413bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1301R/Neo_MAT1_-1892R (3153bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1301R/Neo_MAT1_-2002R (3263bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_226F/Neo_MAT1_1215R (953bp)	+	NT	NT	NT	NT	NT	NT	-	NT	-	NT	NT

Primers \ Species	*N. parvum*	*N. ribis*	*N. kwamb.*	*N. luteum*	*N. australe*	*N. nonq.*	*N. medit.*	*N. eucal.*	*N. vitif.*	*N. arbuti*	*N. alge.*	*N. crypto.*
Neo_MAT1_226F/Neo_MAT1_1301R (1037bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_226F/Neo_MAT1_1511R (1247bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1215R/Neo_MAT1_-273R (1450bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1215R/Neo_MAT1_-1154R (2329bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1215R/Neo_MAT1_-1892R (3069bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_1215R/Neo_MAT1_-2002R (3179bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_771F/Neo_MAT1_-1154R (1883bp)	-	-	NT	-	-	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_771F/Neo_MAT1_226F (507bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_771F/Neo_MAT1_113F (618bp)	NT	NT	NT	-	-	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_771F/Neo_MAT1_-84F (815bp)	NT	NT	NT	-	-	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_771F/Neo_MAT1_-273R (1004bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_771F/Neo_MAT1_-1892R (2623bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_771F/Neo_MAT1_-2002R (2733bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_240F/Neo_MAT1_-1154R (1351bp)	NT	NT	NT	-	-	NT	NT	+	NT	+	NT	NT
Neo_MAT1_240F/Neo_MAT1_-273R (472bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_240F/Neo_MAT1_-1892R (2091bp)	NT	NT	+	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_240F/Neo_MAT1_-2002R (2201bp)	NT	NT	NT	NT	NT	NT	NT	-	NT	NT	NT	NT
Neo_MAT1_+227F/Neo_MAT1_-283R (1983bp)	+	+	+	-	-	+	+	-	+	-	NT	NT
Neo_MAT1_+227F/Neo_MAT1_-1892R (3595bp)	+	NT	NT	NT	NT	NT	-	NT	NT	NT	NT	NT
Neo_MAT1_+227F/Neo_MAT1_113F (1590bp)	NT	NT	NT	NT	NT	NT	NT	NT	NT	+	NT	NT
Neo_MAT1_+227F/Neo_MAT1_-1154R (2855bp)	+	NT	NT	NT	NT	NT	-	NT	NT	NT	NT	NT

Legend:

(+) - Specific amplification
(-) - Nonspecific or without amplification
NT - Not tested
bp - Base pair

"Universal combination" is shaded in blue

Expected sizes of the fragments based on *N.parvum*

Table S3.2 - Possible primers combinations for *MAT1-2* idiomorph.

Primers \ Species	*N. parvum*	*N. ribis*	*N. mangiferae*	*N. kwamb.*	*N. algeriense*	*N. luteum*	*N. australe*	*N. medit.*	*N. eucal.*	*N. crypto.*
Neo_MAT2_-359F/Neo_MAT2_+975R (2738bp)	+	-	+	NT	NT	-	-	-	-	NT
Neo_MAT2_-359F/Neo_MAT2_1070R (1397bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-359F/Neo_MAT2_1405R (1731bp)	NT	+	NT	+	NT	-	-	-	-	NT
Neo_MAT2_-359F/Neo_MAT2_+116R (1880bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-359F/Neo_MAT2_+162R (1923bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-156F/Neo_MAT2_1070R (1193bp)	+	+	+	+	+	+	+	+	+	+
Neo_MAT2_-156F/Neo_MAT2_1405R (1527bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-156F/Neo_MAT2_+116R (1676bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-156F/Neo_MAT2_+162R (1719bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-156F/Neo_MAT2_+975R (2534bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-77F/Neo_MAT2_1070R (1116bp)	NT	NT	NT	NT	NT	NT	NT	NT	-	NT
Neo_MAT2_-77F/Neo_MAT2_1405R (1450bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-77F/Neo_MAT2_+116R (1599bp)	NT	+	NT	NT	+	-	-	-	-	NT
Neo_MAT2_-77F/Neo_MAT2_+162R (1642bp)	NT	NT	NT	NT	NT	-	-	-	-	NT
Neo_MAT2_-77F/Neo_MAT2_+975R (2457bp)	NT	-	-	-	NT	-	-	-	-	NT
Neo_MAT2_119F/Neo_MAT2_1070R (919bp)	NT	NT	NT	NT	NT	NT	NT	NT	-	NT
Neo_MAT2_119F/Neo_MAT2_1405R (1253bp)	+	+	+	+	NT	+	-	-	-	-
Neo_MAT2_119F/Neo_MAT2_+116R (1402bp)	+	+	+	+	NT	-	-	-	-	NT
Neo_MAT2_119F/Neo_MAT2_+975R (2260bp)	+	-	NT	-	NT	-	-	-	-	NT
Neo_MAT2_268F/Neo_MAT2_1070R (773bp)	+	+	+	+	+	+	+	+	+	-
Neo_MAT2_268F/Neo_MAT2_1405R (1107bp)	NT	NT	NT	NT	NT	NT	NT	NT	-	NT
Neo_MAT2_268F/Neo_MAT2_+116R (1256bp)	NT	NT	NT	NT	NT	NT	NT	NT	-	NT
Neo_MAT2_268F/Neo_MAT2_+162R (1299bp)	NT	NT	NT	NT	NT	-	-	+	+	-
Neo_MAT2_268F/Neo_MAT2_+975R (2114bp)	NT	NT	NT	NT	NT	NT	NT	NT	-	NT

Legend:

(+) - Specific amplification
(-) - Nonspecific or without amplification
NT - Not tested
bp - Base pair

"Universal combination" is shaded in green

Expected sizes of the fragments based on *N.parvum*

CHAPTER 4

Mating type gene analyses in the genus *Diplodia*: from cryptic sex to cryptic species

Lopes A, Linaldeddu BT, Phillips AJL, Alves A, 2017. Mating type gene analyses in the genus *Diplodia*: from cryptic sex to cryptic species. Submitted to *Fungal Biology*.

Abstract

Cryptic species are common in *Diplodia*, a genus that includes some well-known and economically important plant pathogens. Thus, species delimitation has been based on the phylogenetic species recognition approach using multigene genealogies. We assessed the potential of mating type (*MAT*) genes sequences as phylogenetic markers for species delimitation in the genus *Diplodia*. A PCR-based mating type diagnostic assay was developed that allowed amplifying and sequencing *MAT1-1-1* and *MAT1-2-1* genes, and determining the mating strategies used by different species. All species tested were shown to be heterothallic. Phylogenetic analyses were performed on both *MAT* genes and also, for comparative purposes, on combined sequences of the ribosomal internal transcribed spacer (ITS), translation elongation factor 1-alpha (*tef1-α*) and beta-tubulin (*tub2*). *MAT* genes individual phylogenies clearly differentiated all species analysed and are in agreement with the results obtained with the commonly used multilocus phylogenetic analysis approach. However, *MAT* genes genealogies were superior to multigene genealogies in resolving closely related cryptic species. The phylogenetic informativeness of each locus was evaluated revealing that *MAT* genes were the most informative locus followed by *tef1-α*. Hence, *MAT* genes can be successfully used to establish species boundaries in the genus *Diplodia*.

Keywords: *Botryosphaeriaceae*, Heterothallism, PCR-based assay, Phylogenetic Informativeness

Introduction

The genus *Diplodia* (*Ascomycetes*, *Botryosphaeriales*, *Botryosphaeriaceae*) includes species that are pathogens, endophytes and saprobes of mostly woody hosts (Phillips et al. 2013). Several studies have shown that cryptic species are frequent in this genus (Phillips et al. 2012; Phillips et al. 2013; Alves et al. 2014; Linaldeddu et al. 2016), a feature that is common to other genera in the *Botryosphaeriaceae*, and it is now widely accepted that morphology is a

poor indicator of species limits. Thus, the over 30 species of *Diplodia* currently known from culture (Phillips et al. 2013; Dissanayake et al. 2016; González-Domínguez et al. 2017; Yang et al. 2017) have been delimited on the basis of the Genealogical Concordance Phylogenetic Species Recognition (GCPSR) proposed by Taylor et al. (2000). This has been based largely on sequence data of two loci, the ribosomal internal transcribed spacer (ITS rDNA) and elongation factor 1-alpha (*tef1-α*) (Lazzizera et al. 2008; Jami et al. 2012; Phillips et al. 2012; Linaldeddu et al. 2013; Lynch et al. 2013; Phillips et al. 2013; Alves et al. 2014).

Although the use of combined ITS and *tef1-α* sequences has helped to resolve a number of cryptic species (e.g. Phillips et al. 2012; Alves et al. 2014) it has some shortcomings. For example, in the clade containing species with brown, aseptate conidia that occasionally develop one or two septa (e.g. *D. seriata*, *D. sapinea*), resolution all species with these two loci can be difficult and some species clades tend to have low phylogenetic support (Phillips et al. 2012; Linaldeddu et al. 2016; Giambra et al. 2016). Finding alternative loci with higher resolving power would be important for adequate species delimitation within this clade and the genus in general.

Mating type (*MAT*) genes have been used as molecular markers to establish phylogenetic relationships and species boundaries in several fungi (Steenkamp et al. 2000; Ueng et al. 2003; O'Donnell et al. 2004; Pöggeler et al. 2011; Martin et al. 2013; Kashyap et al. 2015), including the genus *Neofusicocccum*, a member of the *Botryosphaeriaceae* (Lopes et al. 2017). *MAT* genes are responsible for determining mating compatibility and for the regulation of the sexual cycle, and are located in a specialized region of the genome, the mating type locus (Lee et al. 2010; Ni et al. 2011; Sun & Heitman 2015). In ascomycetes, mating systems are bipolar due to the existence in the single *MAT* locus of two alternate forms (idiomorphs), defining two mating types. Both idiomorphs encode transcriptional regulators. The *MAT1-1-1* idiomorph encodes a protein with an alpha box motif, and the *MAT1-2-1* idiomorph encodes a regulatory DNA-binding protein with a high mobility group (HMG) motif (Debuchy & Turgeon 2006; Lee et al. 2010; Sun & Heitman 2015).

MAT genes are known to be evolutionarily dynamic, undergoing expansions and contractions, as well as chromosomal rearrangements such as translocations and inversions, resulting in high evolutionary rates (Gioti et al. 2012; Martin et al. 2013; Sun & Heitman

2015). This could be the reason for the good results obtained by several studies where *MAT* genes were used to establish phylogenetic relationships, even in apparently asexual fungi (Groenewald et al. 2006; Yokoyama et al. 2006; Lopes et al. 2017). Apart from the usefulness in determining phylogenetic relationships between species and determining species boundaries, the study of *MAT* genes provides information about the mode of reproduction, which is of relevance for population genetic analyses of plant pathogens.

Despite the importance of the genus *Diplodia* very little is known about the genetics of their mating systems as well as the mating strategies employed by each species. For the majority of the species there is no known sexual stage and its induction in vitro has been unsuccessful. Knowledge about *MAT* genes in the genus is currently limited to the species *D. sapinea*, a species that has been considered strictly asexual, but for which a cryptic heterothallic sexual cycle has been proposed (Bihon et al. 2014). Considering this, the goals of the present study were (i) to expand knowledge about *MAT* genes and sexual reproductive strategies in *Diplodia* species and (ii) to evaluate the usefulness of *MAT* genes as phylogenetic markers to delimit species in the genus, in comparison to the more commonly used loci. To accomplish this, a PCR-based assay was developed to amplify and sequence *MAT* genes from *Diplodia* species.

Materials and Methods

Fungal strains and culture conditions

A total of 50 strains representing 20 species were studied and these are listed in Table 4.1. All strains were grown and maintained on half-strength potato-dextrose agar (PDA) (HIMEDIA, India) at room temperature.

Table 4.1 - Identity of the *Diplodia* isolates studied and GenBank accession numbers of the sequences used in phylogenetic analyses.

			GenBank Accession Numbers[b]				
Species	**Isolate No.[a]**	**Host**	**ITS**	***tef1-α***	***tub2***	***MAT1-1-1***	***MAT1-2-1***
D. africana	BL19	*Juniperus phoenicea*	*JF302648*	*JN157807*	MG015797	MG015751	-
	CAD014	*Vitis vinifera*	*KJ638326*	*KJ638344*	MG015798	MG015752	-
D. alatafructa	**CBS124931**	*Pterocarpus angolensis*	*FJ888460*	*FJ888444*	MG015799	-	MG015775
D. corticola	**CBS112549**	*Quercus suber*	*AY259100*	*AY573227*	*DQ458853*	MG015753	-
	CAA499	*Eucalyptus globulus*	MG015741	MG015723	MG015800	-	MG015776
	CAA500	*Eucalyptus globulus*	*KT440895*	*KT440958*	MG015801	-	MG015777
	CAA691	*Eucalyptus globulus*	*KT440896*	*KT440959*	MG015802	MG015754	-
	BL36	*Quercus canariensis*	*JX894196*	*JX894215*	MG015803	MG015755	-
	BL37	*Quercus suber*	*JX894197*	*JX894216*	MG015804	-	MG015778
D. cupressi	CBS261.85	*Cupressus sempervirens*	*DQ458894*	*DQ458879*	*DQ458862*	MG015756	-
	CAA028	*Juniperus scopulorum*	MG015742	MG015724	MG015805	MG015757	-
D. eriobotryicola	**CBS140851**	*Eriobotrya japonica*	*KT240355*	*KT240193*	MG015806	MG015758	-
D. fraxini	**CBS136010**	*Fraxinus angustifolia*	*KF307700*	*KF318747*	MG015807	MG015759	-
	CBS136011	*Fraxinus angustifolia*	*KF307711*	*KF318758*	MG015808	MG015760	-
D. insularis	**CBS140350**	*Pistacia lentiscus*	*KX833072*	*KX833073*	MG015809	MG015761	-
	BL132	*Fraxinus angustifolia*	*KF307720*	*KF318767*	MG015810	-	MG015779
D. intermedia	CAA147	*Malus domestica* (fruit rot)	*GQ923857*	*GQ923825*	MG015811	MG015762	-
	CAA490	*Pyracantha coccinea*	MG015744	MG015726	MG015812	-	MG015780
	CAA491	*Pyracantha coccinea*	MG015745	MG015727	MG015813	MG015763	-
	CAP150	Quince fruit	MG015743	MG015725	MG015814	-	MG015781
D. malorum	CBS112554	*Malus sylvestris*	*AY259095*	*DQ458870*	*DQ458851*	MG015764	-

			GenBank Accession Numbers[b]				
Species	**Isolate No.[a]**	**Host**	**ITS**	***tef1-α***	***tub2***	***MAT1-1-1***	***MAT1-2-1***
D. mutila	**CBS136014**	*Populus alba*	*KJ361837*	*KJ361829*	MG015815	MG015765	-
	CAA507	*Fraxinus ornus*	MG015746	MG015728	MG015816	MG015766	-
	CBS230.30	*Phoenix dactylifera*	*DQ458886*	*DQ458869*	*DQ458849*	-	MG015782
D. olivarum	BL96	*Pistacia lentiscus*	*KX833078*	*KX833079*	MG015817	-	MG015783
	BL97	*Quercus coccifera*	*KF307719*	*KF318766*	MG015818	MG015767	-
	CAD019	*Vitis vinifera*	*KJ638323*	*KJ638341*	MG015819	MG015768	-
D. pseudoseriata	**CBS124906**	*Blepharocalyx salicifolius*	*EU080927*	*EU863181*	MG015820	-	MG015784
D. quercivora	**CBS133852**	*Quercus canariensis*	*JX894205*	*JX894229*	MG015821	-	MG015785
	CBS133853	*Quercus canariensis*	*JX894206*	*JX894230*	MG015822	-	MG015786
D. rosacearum	**CBS141915**	*Eriobotrya japonica*	*KT956270*	*KU378605*	MG015823	-	MG015787
	CAA802	*Sorbus intermedia*	MG015747	MG015729	MG015824	-	MG015788
D. rosulata	**CBS116470**	*Prunus africana*	*EU430265*	*EU430267*	*EU673132*	MG015769	-
	CBS116472	*Prunus africana*	*EU430266*	*EU430268*	*EU673131*	MG015770	-
D. sapinea	**CBS393.84**	*Pinus nigra*	*DQ458895*	*DQ458880*	*DQ458863*	MG015771	-
	CBS109727	*Pinus radiata*	*DQ458897*	*DQ458882*	*DQ458865*	MG015772	-
	CBS591.84	*Pinus radiata*	MG015748	MG015730	MG015825	-	MG015789
	CBS109943	*Pinus patula*	*DQ458898*	*DQ458883*	*DQ458866*	-	MG015790
	CMW190	*Pinus radiata*	*KF766159*	*PRJNA215898*	*PRJNA215898*	*KF551229*	-
	CMW39103	*Pinus radiata*	*PRJNA242796*	*PRJNA242796*	*PRJNA242796*	-	*KF551228*
D. seriata	**CBS112555**	*Vitis vinifera*	*AY259094*	*AY573220*	*DQ458856*	-	MG015793
	CAA317	*Eucalyptus globulus*	*KT440897*	*KT440955*	MG015826	-	MG015794
	CAA634	*Fraxinus ornus*	MG015749	MG015731	MG015827	MG015773	-
	CAA636	*Fraxinus ornus*	MG015750	MG015732	MG015828	MG015774	-

			GenBank Accession Numbers[b]				
Species	**Isolate No.[a]**	**Host**	**ITS**	***tef1-α***	***tub2***	***MAT1-1-1***	***MAT1-2-1***
	DS831	*Vitis vinifera*	*KP296243*	*PRJNA261773*	*PRJNA261773*	-	*PRJNA261773*
D. scrobiculata	CBS109944	*Pinus greggii*	*DQ458899*	*DQ458884*	*DQ458867*	-	MG015791
	CBS113423	*Pinus greggii*	*DQ458900*	*DQ458885*	*DQ458868*	-	MG015792
	CMW30223	*Pinus patula*	*PRJNA278001*	*PRJNA278001*	*PRJNA278001*	*PRJNA278001*	-
D. subglobosa	CBS124132	*Fraxinus excelsior*	*DQ458887*	*DQ458871*	*DQ458852*	-	MG015795
D. tsugae	**CBS418.64**	*Tsuga heterophylla*	*DQ458888*	*DQ458873*	*DQ458855*	-	MG015796

[a] Acronyms of culture collections: **BL** - B.T. Linaldeddu, Università degli Studi di Sassari, Italy; **CAA** – Personal culture collection Artur Alves, Universidade de Aveiro, Portugal; **CAD** - A. Deidda, Università degli Studi di Sassari, Italy; **CAP** - Personal culture collection Alan Phillips, Universidade de Lisboa, Portugal; **CBS** – Centraalbureau voor Schimmelcultures, Utrecht, The Netherlands; **CMW** – Tree Patholgy Co-operative Program, Forestry and Agricultural Biotechnology Institute, University of Pretoria, South Africa.
[b] Sequence numbers in italics were retrieved from GenBank. All others were determined in the present study.
Isolates in bold are ex-type cultures.

Analyses of Diplodia sapinea genomes and design of MAT genes primers

Currently genomes of six *Diplodia* species have been sequenced and are available in GenBank public database. However, when this study began the genomes of only two strains of *Diplodia sapinea* were ready to use and the primers in the present study were firstly designed based on them (PRJNA215898; PRJNA242796). Analysis of the *MAT* locus had already been done by Bihon et al. (2014), which facilitated our approach to *MAT* genes analyses (KF551229; KF551228). Although the locus contained more than one gene only the main *MAT1-1-1* and *MAT1-2-1* genes were used in this study.

A first set of primers for each *MAT* gene was designed manually inside the genes. Characteristics such as length, GC content, melting temperature, potential hairpin formation, complementarity and potential self-annealing sites were then checked with the free software OligoCalc: Oligonucleotide Properties Calculator (http://biotools.nubic.northwestern.edu/OligoCalc.html) and Sequence Manipulation Suite (http://www.bioinformatics.org/sms2/pcr_primer_stats.html). Primers were synthesized by STAB Vida Lda (Lisbon, Portugal), reconstituted and diluted (10 pmol) in Tris-EDTA buffer, and stored at -20°C. The first set of primers was then tested in *D. sapinea* and in all the remaining species. Whenever primers gave rise to amplification, the amplicons were sequenced and new sets of primers were designed based on consensus of the acquired *MAT* sequences. The aim of this procedure was to reach a set of "universal" primers able to amplify each gene in all the studied species. All the primers designed are listed in Table 4.2 (see also Fig. S4.1). The primers used in the amplification of the sequences used in phylogenetic analyses of *MAT* genes are discriminated in the Table 4.3. Combinations of primers and species that generated an amplicon are available in the supplementary Tables S4.1 and S4.2.

Table 4.2 - Primers designed in this study.

Target	Name	Nucleotide Sequence (5'⟶3')	Tm (°C)
***MAT1-1-1* gene**	Diplodia_MAT1_292F	CTCAGCTGACACTACGCAGG	57.7
	Diplodia_MAT1_391F	GTCAAGGCCAAATGGACCATC	56.5
	Diplodia_MAT1_1159R	CCATCGTGCCAGACTTCTC	55.5
	Diplodia_MAT1_1174R	CCTTCTCACCAACTTCCATCG	55.5
	Diplodia_MAT1_1325R	GCGAGACGGTGCATGTCGAAT	60.2
***MAT1-2-1* gene**	Diplodia_MAT2_82F	GTCGCACTTCAGCAACTGAAG	56.6
	Diplodia_MAT2_113F	CCTCGATCGATTTGCCTCAC	55.8
	Diplodia_MAT2_978R	GTGGCATCAGCATTGGCTTTAG	56.9
	Diplodia_MAT2_1058R	CGTTGAGCTGGAAGCCACCAT	60.2
	Diplodia_MAT2_1187R	GGTCGAAGTTGGCCTCACG	58.6

DNA extraction, PCR amplification and sequencing

Genomic DNA was extracted from fresh mycelium of cultures grown on half-strength PDA plates for 5 d at approximately 23°C, according to Alves et al. (2004). All PCR reactions were carried out in 25 µL reaction mixtures with NZYTaq 2× Green Master Mix (2.5 mM MgCl2; 200 µM dNTPs; 0.2 U/µL DNA polymerase) (Lisbon, Portugal), in a Bio-Rad C-1000 Touch^{TM} Thermal Cycler (Hercules, CA, USA). Negative controls with sterile water instead of template DNA were used in every PCR reaction. Amplification of *MAT1-1-1* gene was performed with the primers Diplodia_MAT1_391F and Diplodia_MAT1_1325R (primers set A). Thermal conditions were denaturation at 95°C for 3 min; 35 cycles at 94°C for 30 s, 50-56°C for 30 s, and 72°C for 1 min; final extension at 72°C for 10 min. For *MAT1-2-1* amplification the primers Diplodia_MAT2_82F and Diplodia_MAT2_1058R (primers set B), Diplodia_MAT2_113F and Diplodia_MAT2_1187R (primers set C) or Diplodia_MAT2_82F and Diplodia_MAT2_1187R (primers set D) were used with the following thermal conditions: denaturation at 95°C for 3 min; 35 cycles at 94°C for 30 s, 50-52°C for 30 s, and 72°C for 1 min 10s; final extension at 72°C for 10 min (see Table 4.3 for annealing temperatures specification).

Primers ITS1 and ITS4 (White et al. 1990) were used for amplification and sequencing of the ITS region of the ribosomal DNA as described by Alves et al. (2004). Part of the translation elongation factor 1-alpha was amplified and sequenced with primers EF1-688F and

EF1-1251R (Alves et al. 2008). Beta-tubulin (*tub2*) gene was amplified with T1 and Bt2b primers (Glass & Donaldson 1995; O'Donnell & Cigelnik 1997) with the cycling conditions previously described by Lopes et al. (2017). The histone gene was amplified and sequenced with primers CYLH3F and H3-1b (Glass & Donaldson 1995; Crous et al. 2004) according to Santos et al. (2017). For amplification and sequencing of actin were used the primers ACT-512F and ACT-783R (Carbone & Kohn 1999) with an initial denaturation at 95°C for 5 min; followed by 35 cycles at 94°C for 30 s, 55°C for 30 s, and 72°C for 1 min; and then a final extension at 72°C for 10 min.

Table 4.3 - Primers and annealing temperatures used in the PCR.

		Annealing Temperatures (°C)			
		MAT1-1-1		***MAT1-2-1***	
		Primer set A	**Primer set B**	**Primer set C**	**Primer set D**
Species	***D. alatafructa***	a	-	-	52
	D. africana	52	b	b	b
	D. corticola	52	-	52	-
	D. cupressi	50	b	b	b
	D. eriobotryicola	52	b	b	b
	D. fraxini	52	b	b	b
	D. insularis	52	-	-	52
	D. intermedia	52	52	-	NT
	D. malorum	56	b	b	b
	D. mutila	56	-	52	-
	D. olivarum	52	NT	52	NT
	D. pseudoseriata	a	-	-	52
	D. quercivora	a	-	52	+
	D. rosacearum	a	52	-	NT
	D. rosulata	50	b	b	b
	D. sapinea	52	52	-	NT
	D. seriata	52	52	52	NT
	D. scrobiculata	a	50	-	NT
	D. subglobosa	a	-	52	-
	D. tsugae	a	-	52	-

(a) – No *MAT1-1* isolates available
(b) – No *MAT1-2* isolates available
(-) – No amplification
NT – Not tested

After amplification, 2 µL of each PCR product were separated by electrophoresis in 1.5% agarose gels at 90 volts for 1 h in 1× TAE buffer (40 mM Tris, 40 mM acetate, 2 mM EDTA, pH 8.0). A GeneRuler DNA Ladder Mix (Thermo Scientific, USA) was also included. Gels were stained with ethidium bromide and visualized on a BioRad Molecular Imager Gel Doc^{TM} XR^{+} to assess PCR amplification. The amplified PCR fragments were purified with the DNA Clean & Concentrator™-5 kit (Zymo Research, CA, USA) before sequencing.

Both strands of the PCR products were sequenced at GATC Biotech (Cologne, Germany). The nucleotide sequences were read with FinchTV v.1.4.0 (Geospiza Inc. http://www.geospiza.com/finchtv). All sequences were checked manually, and nucleotide arrangements at ambiguous positions were clarified using both primer direction sequences. Sequences generated in this study were deposited in GenBank (see Table 4.1; accession numbers for actin and histone are MG015715-MGO15722, MG015829 and MG015733-MG015740, MG015830, respectively).

Phylogenetic analyses

Sequences were aligned with ClustalX v. 2.1 (Thompson et al. 1997), using the following parameters: pairwise alignment parameters (gap opening = 10, gap extension = 0.1) and multiple alignment parameters (gap opening = 10, gap extension = 0.2, transition weight = 0.5, delay divergent sequences = 25 %). Alignments were checked and edited with BioEdit Alignment Editor v. 7.2.5 (Hall 1999). Phylogenetic analyses of sequence data were done with MEGA6 v. 6.06 (Tamura et al. 2013). All gaps were included in the analyses. The model of DNA sequence evolution used for each dataset was determined by the software (See Figs 4.3 and 4.4). Maximum likelihood (ML) analyses were performed on a Neighbor-Joining (NJ) starting tree automatically generated by the software. A bootstrap analysis (1000 replicates) was used to estimate the consistency of each node of the trees. The alignments of all datasets were deposited in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S21615).

Phylogenetic informativeness of the loci used in the phylogenetic analyses

Profiles of phylogenetic informativeness for each locus were obtained using PhyDesign (López-Giráldez & Townsend 2011; http://phydesign.townsend.yale.edu/). This allows a comparison of different loci through calculation of the informativeness per base pair. Since the species are heterothallic, comparisons between each *MAT* gene and the other loci used in phylogenetic analyses were done separately and comparisons between *MAT1-1-1* and *MAT1-2-1* were therefore not possible.

Results

MAT genes sequencing and mating strategies

To amplify and sequence *MAT* genes in all studied species a large number of primers were designed (Table 4.2). For *MAT1-1-1* a single combination of primers (set A) worked for all species but in the case of *MAT1-2-1* three sets of primers were needed (set B, C and D) (Figure 4.1, Table 4.3). The four sets of primers annealed inside the genes amplifying partial fragments of *MAT1-1-1* and *MAT1-2-1* with variable lengths for each *MAT* gene and for each species (Table 4.4). Variability was also seen in the sequence alignments with the internal regions of the genes being more dissimilar than the regions closer to the end of the genes. It was also noticed that among isolates of the same species the sequences of both genes were mostly identical (*MAT1-1-1* and *MAT1-2-1* alignments in supplementary data). All analysed species have only one small intron (48–49 bp) in each *MAT* gene, located near the start of the gene. For *MAT1-1-1* this intron was not included in the region amplified and thus was not included in the phylogenetic analyses. In *MAT1-2-1* the intron was included in the amplification and phylogenetic analyses (Figure 4.1).

No homothallic species could be detected. For *D. corticola*, *D. sapinea*, *D. scrobiculata* and *D. seriata* heterothallism was confirmed from the sequenced genomes (PRJNA325745, PRJNA215898, PRJNA242796, PRJNA278001, PRJNA261773) in which only one mating type is present. Also, in this study strains of each mating type for *D. insularis*, *D. intermedia*, *D. mutila* and *D. olivarum* were analysed confirming the presence of only one *MAT* gene in each strain. On the other hand, *D. africana*, *D. alatafructa*, *D. cupressi*, *D. eriobotryicola*, *D. fraxini*, *D. malorum*, *D. pseudoseriata*, *D. quercivora*, *D. rosacearum*, *D.*

rosulata, D. subglobosa and *D. tsugae* are putative heterothallic species since only one of the genes (either *MAT1-1-1* or *MAT1-2-1*) was detected by PCR amplification (Table 4.3 and 4.4).

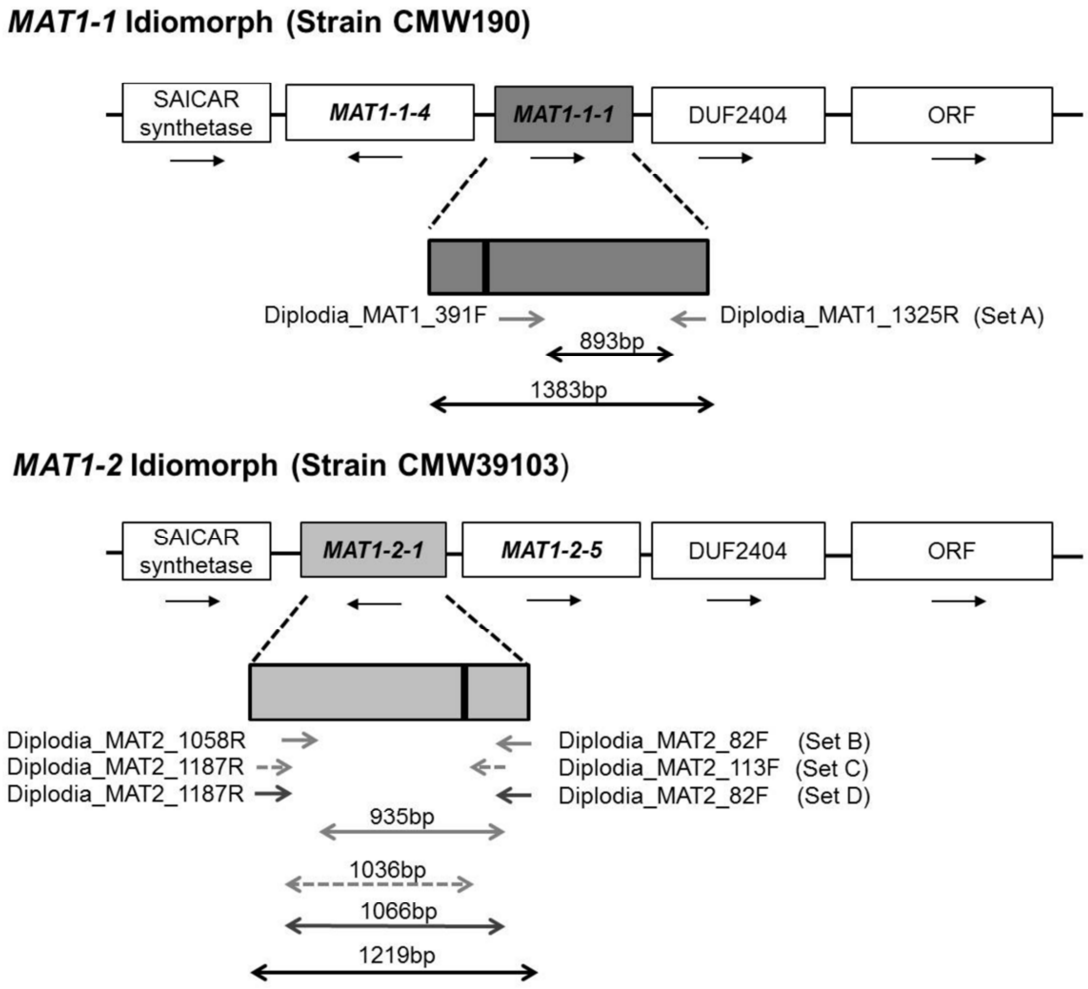

Figure 4.1 - Structure of the mating type locus of the heterothallic fungus *Diplodia sapinea*. Arrows below the genes indicate genes orientation (5'-3'). Introns are represented inside the *MAT* genes by vertical black lines. Primers used to amplify partial *MAT* genes are noted with respective fragment lengths. Distances and sizes are not drawn to a scale. Adapted from Bihon et al. (2014).

Table 4.4 - Fragment lengths obtained for each *Diplodia* species and mating strategies.

Species	Isolate No.[a]	Mating strategy	Fragment length (bp)	
			MAT1-1-1	*MAT1-2-1*
D. africana	BL19	Heterothallic (?)	671	-
	CAD014		671	-
D. alatafructa	**CBS124931**	Heterothallic (?)	-	797
D. corticola	**CBS112549**	Heterothallic	782	-
	CAA499		-	851
	CAA500		-	851
	CAA691		782	-
	BL36		782	-
	BL37		-	851
D. cupressi	CBS261.85	Heterothallic (?)	737	-
	CAA028		737	-
D. eriobotryicola	**CBS140851**	Heterothallic (?)	668	-
D. fraxini	**CBS136010**	Heterothallic (?)	632	-
	CBS136011		632	-
D. insularis	**CBS140350**	Heterothallic	671	-
	BL132		-	797
D. intermedia	CAA147	Heterothallic	668	-
	CAA490		-	792
	CAA491		668	-
	CAP150		-	792
D. malorum	CBS112554	Heterothallic (?)	671	-
D. mutila	**CBS136014**	Heterothallic	671	-
	CAA507		671	-
	CBS230.30		-	871
D. olivarum	BL96	Heterothallic	-	871
	BL97		671	-
	CAD019		671	-

Species	Isolate No.[a]	Mating strategy	Fragment length (bp)	
			MAT1-1-1	*MAT1-2-1*
D. pseudoseriata	**CBS124906**	Heterothallic (?)	-	797
D. quercivora	**CBS133852**	Heterothallic (?)	-	851
	CBS133853		-	851
D. rosacearum	**CBS141915**	Heterothallic (?)	-	792
	CAA802		-	792
D. rosulata	**CBS116470**	Heterothallic (?)	671	-
	CBS116472		671	-
D. sapinea	**CBS393.84**	Heterothallic	659	-
	CBS109727		659	-
	CBS591.84		-	792
	CBS109943		-	792
	CMW190		659	-
	CMW39103		-	792
D. seriata	**CBS112555**	Heterothallic	-	786
	CAA317		-	786
	CAA634		668	-
	CAA636		668	-
	DS831		-	786
D. scrobiculata	CBS109944	Heterothallic	-	792
	CBS113423		-	792
	CMW30223		668	-
D. subglobosa	CBS124132	Heterothallic (?)	-	871
D. tsugae	**CBS418.64**	Heterothallic (?)	-	877

[a]Isolates in bold are ex-type cultures.

Phylogenetic analyses

All species of *Diplodia* were, in general, distinguished by the phylogenetic analyses of each *MAT* gene and by the combined dataset of the regions ITS rDNA, *tef1-α* and *tub2* (Figs. 4.2 and 4.3). The only exception was found in the phylogenetic analysis obtained with *MAT1-2-1* sequences where *D. mutila* and *D. subglobosa* were not clearly separated. *MAT1-2-1* sequences from both species are very similar and differ in only three nucleotide positions (*MAT1-2-1* alignment in supplementary data).

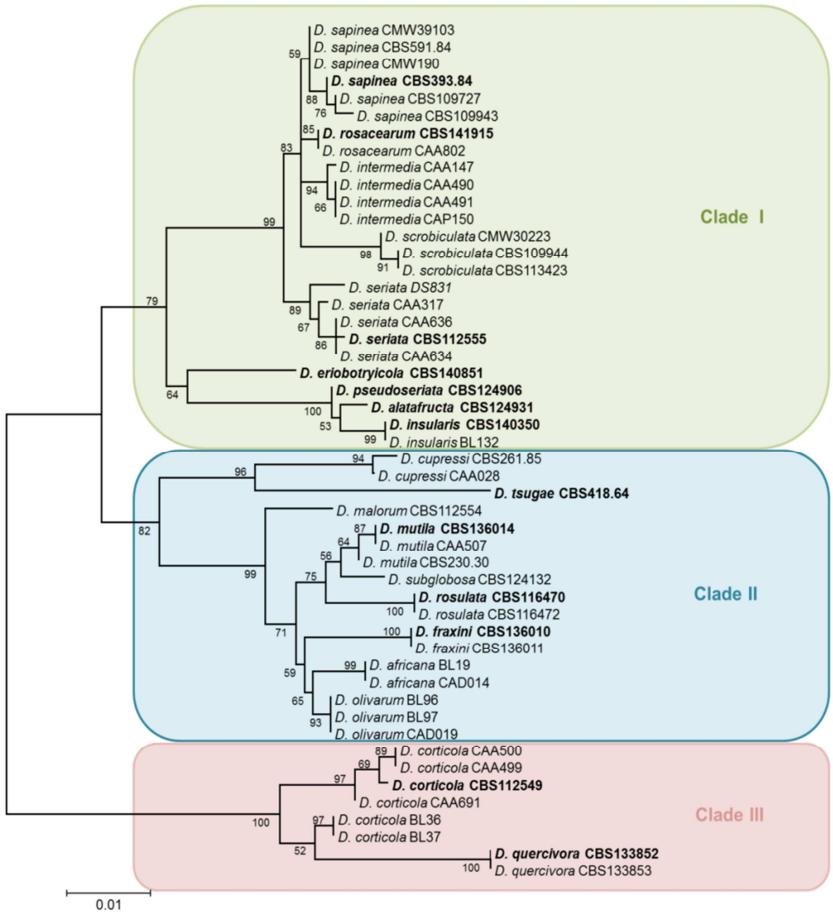

Figure 4.2 - Phylogenetic relationships of *Diplodia* species based on the combined sequence data from ITS, *tef1-α* and *tub2* regions. There were a total of 1157 positions in the final dataset. The phylogeny was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model. The tree with the highest log likelihood is shown. A discrete Gamma distribution was used to model evolutionary rate differences among sites. Bootstrap values (>50%) are given at the nodes. The tree is drawn to scale, with branch length measured in the number of substitutions per site.

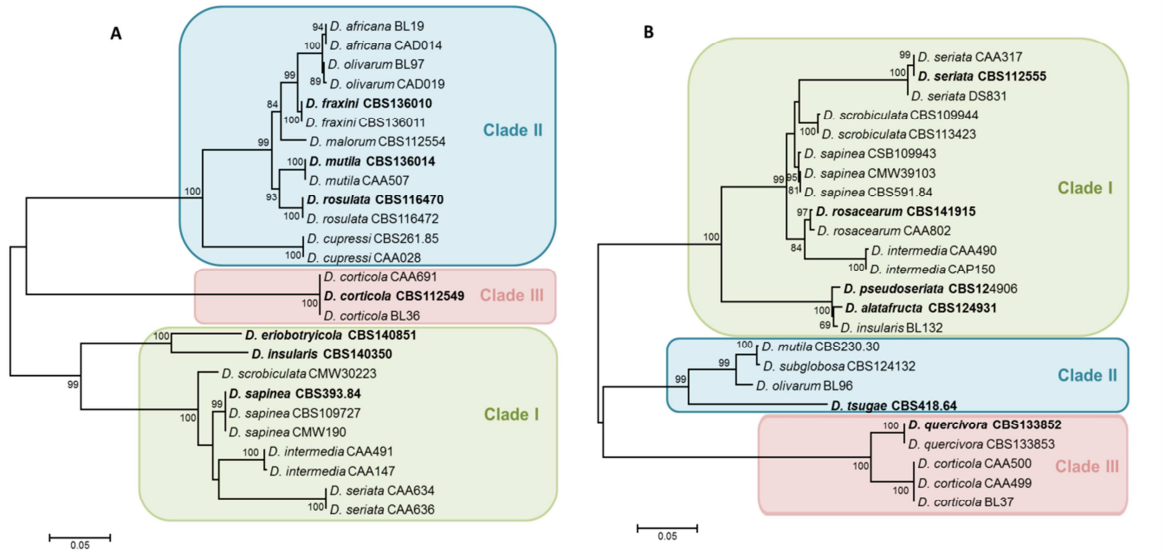

Figure 4.3 - Phylogenetic relationships of *Diplodia* species based on the sequence data from *MAT1-1-1* (A) and *MAT1-2-1* (B) genes. For *MAT1-1-1* gene a total of 864 positions were used. Phylogeny was inferred using the Maximum Likelihood method based on the Hasegawa-Kishino-Yano model. In the *MAT1-2-1* phylogeny were used a total of 904 positions in the final dataset and the Maximum Likelihood method based on the Tamura-3 parameter model was performed. A discrete Gamma distribution was used to model evolutionary rate differences among sites for both phylogenies. The trees with the highest log likelihood are shown. Bootstrap values (>50%) are given at the nodes. The trees are drawn to scale, with branch length measured in the number of substitutions per site.

Topology of the tree resulting from the analysis of the three concatenated regions was, overall, in concordance with that obtained with *MAT1-1-1* and *MAT1-2-1* sequences. All analyses grouped the species in three major clades (I, II and III) (Figs. 4.2 and 4.3). In the concatenated data analysis, which includes all the species studied, clade I grouped the closely related species *D. alatafructa*, *D. eriobotryicola*, *D. insularis*, *D. intermedia*, *D. pseudoseriata*, *D. sapinea*, *D. scrobiculata*, *D. seriata* and *D. rosacearum*. Clade II is represented by *D. africana*, *D. cupressi*, *D. fraxini*, *D. malorum*, *D. mutila*, *D. olivarum*, *D. rosulata*, *D. subglobosa* and *D. tsugae*. Clade III constitutes only *D. corticola* and *D. quercivora*. The phylogenetic analyses also revealed subgroups within some species. For example, in the concatenated data analysis isolates of *D. corticola* are clearly separated into two main groups as a result of fixed differences in the *tef1-α* sequences. The two main

subgroups of *D. corticola* formed in the analysis of the three concatenated regions are not supported by either of the *MAT* gene phylogenies or by *his3* phylogeny (Fig. 4.4). Phylogenies generated for *tub2* and *act2* were not congruent but both grouped the isolates of *D. corticola* representing the two *tef1-α* lineages in other separate but closely related subclades (Fig. 4.4).

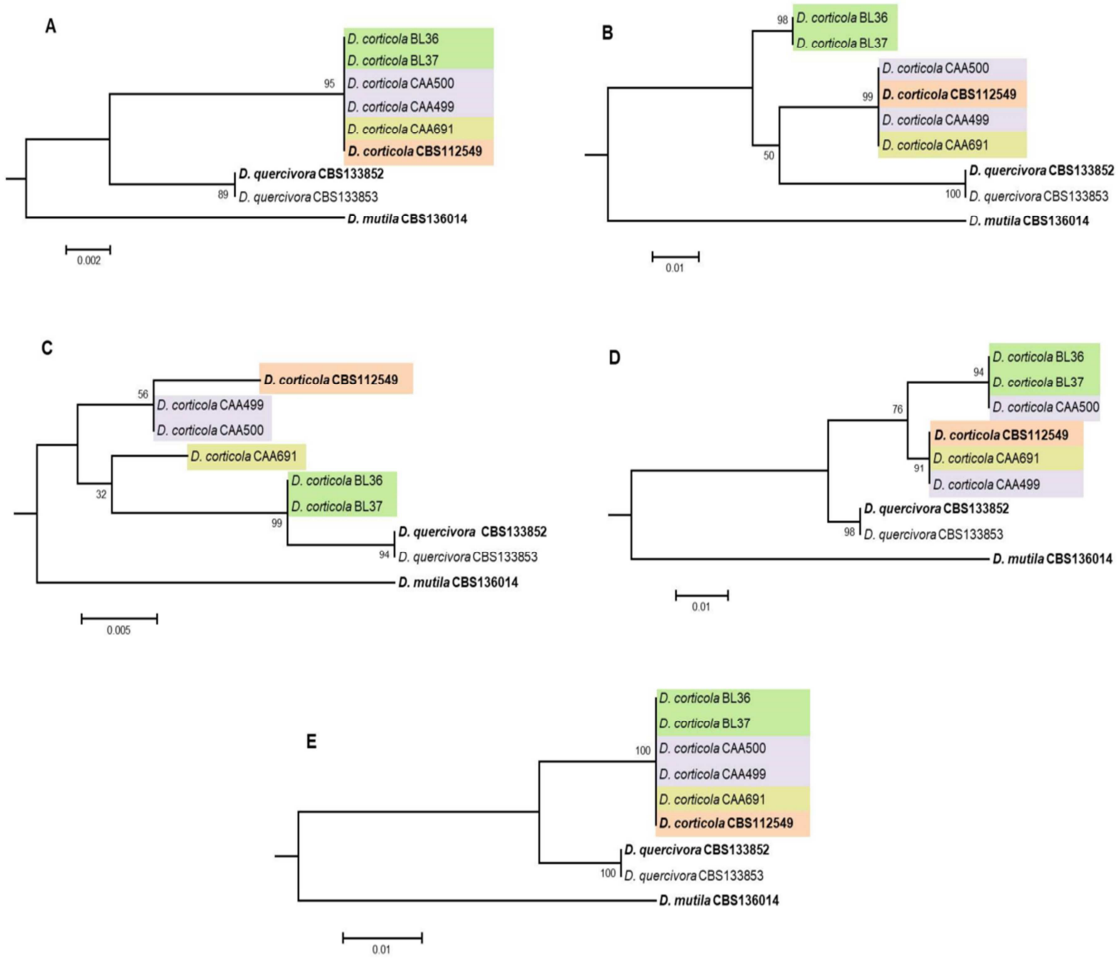

Figure 4.4 - Phylogenetic relationships of *Diplodia corticola* isolates based on the sequence data from ITS (A), *tef1-α* (B), *tub2* (C), *act2* (D) and *his3* (E) regions. The phylogeny was inferred using the Maximum Likelihood method based on the Tamura 3-parameter model except for the ITS region that was inferred based on the Jukes-Cantor model. The tree with the highest log likelihood is shown. A uniform distribution was used to model evolutionary rate differences among sites. Bootstrap values are given at the nodes. The tree is drawn to scale, with branch length measured in the number of substitutions per site.

With less variability between isolates are the species *D. intermedia*, *D. mutila*, *D. sapinea*, *D. scrobiculata* and *D. seriata*. Some of this variability is reinforced in the phylogenetic analyses of *MAT1-2-1* gene (*D. sapinea* and *D. seriata*). On the contrary, the phylogenetic analysis of *MAT1-1-1* gene denotes sequences more conserved within isolates of the same species.

Phylogenetic informativeness of the loci used in the phylogenetic analyses

As can be seen in Figure 4.5, *MAT* genes are the highest contributors to both net and per-site phylogenetic informativeness from all loci studied. Furthermore, ITS and *tub2* sequences are the least informative to resolve *Diplodia* species (Fig. 4.5).

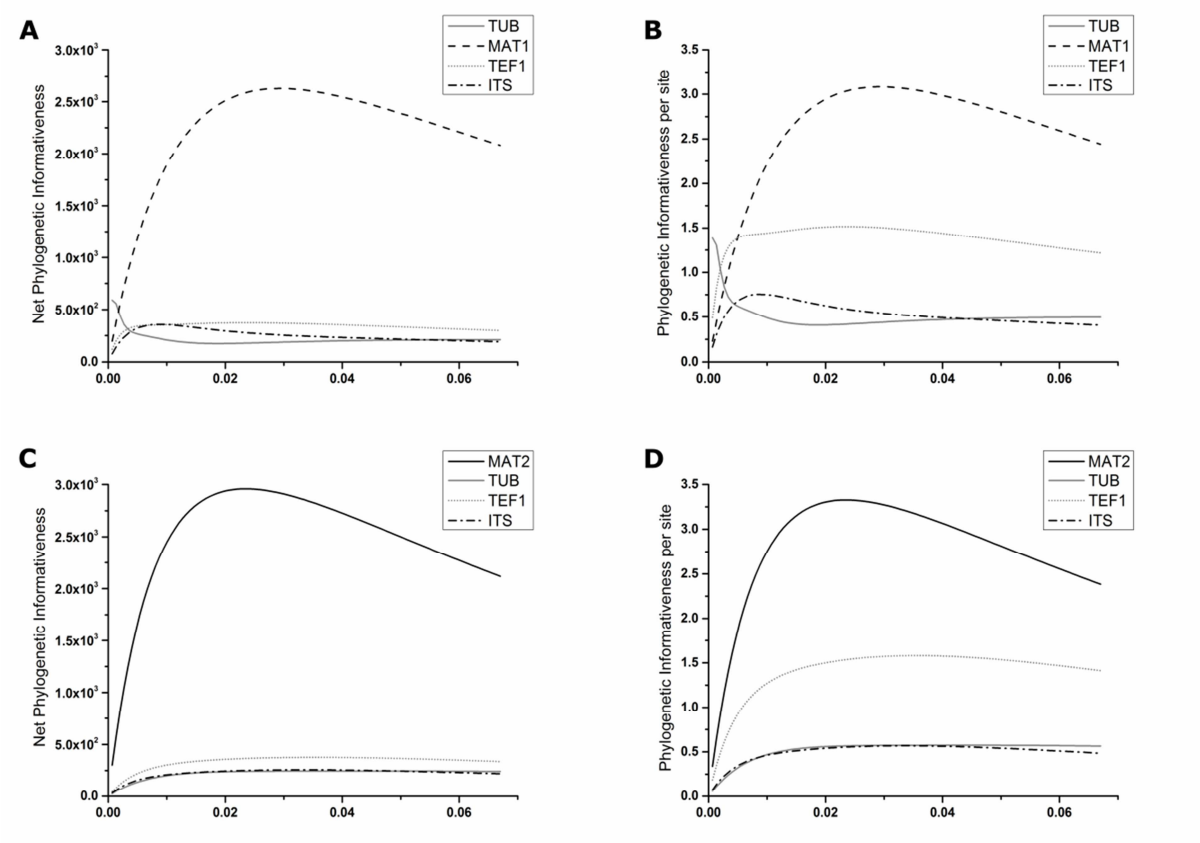

Figure 4.5 - Profiles of phylogenetic informativeness. Net Phylogenetic informativeness for the loci ITS, *tef1-α*, *tub2*, *MAT1-1-1* (A) and ITS, *tef1-α*, *tub2*, *MAT1-1-2* (C); Phylogenetic informativeness per site for the loci ITS, *tef1-α*, *tub2*, *MAT1-1-1* (B) and ITS, *tef1-α*, *tub2*, *MAT1-1-2* (D).

Discussion

In this study primers were developed to amplify and sequence *MAT* loci from several species of *Diplodia*. Phylogenies resulting from these *MAT* sequences were compared with those obtained from the combined analyses of ITS, *tef1-α* and *tub2* regions revealing a general concordance between them, but with higher support in *MAT* phylogenies.

The availability of genomes from both mating types of *D. sapinea* was fundamental to the development of the first primers. Despite the presence of additional *MAT* genes (*MAT1-1-4* and *MAT1-2-5* in the *MAT1-1* and *MAT1-2* locus respectively) the targets of this study were the main *MAT1-1-1* and *MAT1-2-1* genes (Fig. 4.1). Five primers per idiomorph were necessary to reach the right combinations that allowed the amplification of the genes in all the studied species. Since we do not have in our collection any strain of *D. scrobiculata* with the *MAT1-1-1* idiomorph the primers could not be tested and we cannot draw any conclusion about their performance in this species. The sequence used in the phylogenetic analysis was retrieved from the genome of strain CMW30223 (PRJNA278001; Wingfield et al. 2015). Similar to what has been reported for the genus *Neofusicoccum*, *MAT* gene sequence alignments revealed that internal regions of the genes are more dissimilar than the regions closer to the end of the genes and that among isolates of the same species sequences of both genes are almost identical (Supplementary data). The great variability of the *MAT1-2-1* gene between species was also reflected by the need for more than one set of primers to amplify it in all the studied species.

As with *D. sapinea* all other species studied are apparently heterothallic. This is very different from what has been found in the genus *Neofusicoccum* where the majority of the species are homothallic (Lopes et al. 2017). The prevalence of heterothallism as mating strategy in fungal plant pathogens such as *Diplodia* species is an important feature with implications on the genetic variability and evolutionary potential of a species. Obligate outcrossing species are able to rapidly generate new genotypes with differing infection capacity and virulence, with the more favourable genotypes maintained through clonal reproduction (McDonald & Linde 2002). Thus, knowledge about the mating strategy adopted by a fungal plant pathogen is relevant for the development and implementation of adequate disease management strategies (McDonald & Linde 2002).

Working with *MAT* genes of heterothallic species is however quite challenging since we could not analyse the same idiomorph for all the species in study. In our collection, for example, only isolates with the idiomorph *MAT1-1* were available for *D. africana* and *D. rosulata*, while for *D. subglobosa* and *D. tsugae* only isolates with *MAT1-2* idiomorph were present. In these cases, comparison of phylogenies is possible between only those species for which both genes are available. Furthermore, comparison of phylogenetic informativeness of *MAT1-1-1* with *MAT1-2-1* could not be tested. Despite this limitation, the results obtained using the concatenated data (ITS, *tef1-α*, *tub2*) and *MAT* genes are in concordance, with each analysis resolving three major clades (I, II, III) (Fig. 4.2 and 4.3). These results are also in concordance with the analyses of Phillips et al. (2013), where the same clades were represented. The three clades are constituted by closely related species that are sometimes difficult to distinguish. Specifically in clade I where cryptic species (e.g. *D. intermedia*/*D. sapinea*/*D. seriata* and *D. alatafructa*/*D. pseudoseriata*) are very common the resolution of species is not well supported in phylogenies based on ITS and *tef1-α* (Phillips et al. 2012, 2013). In our combined analyses, where partial *tub2* sequences were included, the delimitation of the cryptic species *D. intermedia* and *D. seriata* was improved but bootstrap support for *D. sapinea*, *D. eriobotryicola* and *D. alatafructa* is low. On the contrary, phylogenetic analyses based on *MAT1-1-1* and/or *MAT1-2-1* genes discriminate all species with high bootstrap support values (Fig.4. 3). They are clearly effective in resolving cryptic species in clade I. The only exception was for *Diplodia mutila* and *Diplodia subglobosa* (clade II) in the phylogenetic analysis of *MAT1-2-1* where they are almost indistinguishable but still with some differences at sequence level (Fig. 4.3 and *MAT1-2-1* alignment in supplementary data). This could be a result of a recent divergence of these two species. Thus, it would be important in the future to study more isolates of *MAT1-2* idiomorph for both species. Unfortunately, *MAT1-1* strains of *Diplodia subglobosa* were not available and thus it was not possible to determine how this species is placed in the *MAT1-1-1* phylogenetic analysis.

In all phylogenetic analyses minor genetic variations between isolates of the same species are usually due to intraspecific variability. However, for *D. corticola*, *D. sapinea* and *D. seriata*, multiple phylogenetic subgroups were revealed in the analyses of the three concatenated genes (ITS, *tef1-α* and *tub2*) raising questions about the putative existence of

distinct species. In the case of *D. corticola*, Linaldeddu et al. (2013) considered these subgroups as different lineages within a single species. Phylogenetic relationships of *Diplodia corticola* isolates based on the sequence data from ITS, *tef1-α*, *tub2*, *act2* and *his3* regions were not concordant revealing the existence of gene flow between lineages (Fig. 4.4). Considering this and the analyses of *MAT* genes presented here, where no differences were found between these lineages, which reinforces the possibility of crossing between them, we agree with Linaldeddu et al. (2013) that these lineages represent a single species, *D. corticola*. In the case of *D. sapinea* and *D. seriata* only minor differences within a species were noticed in the analysis of *MAT1-2-1* gene suggesting that these variations probably reflect intraspecific variability rather than interspecific variation.

As mentioned above it was impossible to obtain both *MAT* genes for all species analysed and therefore a full comparison of *MAT* phylogenies could not be accomplished, precluding us from drawing any definite conclusion about which of the *MAT* genes is more efficient in resolving *Diplodia* species. The *MAT1-1-1* gene seemed to be more useful than *MAT1-2-1* gene in PCR amplification since for *MAT1-1-1* gene one set of primers was enough to amplify the gene in all species, reducing the workload, cost and time. However, in the phylogenetic analysis both genes appeared to be useful in delimiting species. Also, the test for phylogenetic informativeness showed that both *MAT* genes are highly effective as opposed to ITS and *tub2* which are least effective (Fig. 4.5). Thus, in the resolution of *Diplodia* species we strongly recommend the use of *MAT* genes due to their high resolving power (Fig. 4.3) as well their high net and per-site phylogenetic informativeness (Fig. 4.5). In addition to *MAT* genes, the *tef1-α* locus seems to be useful with the exception of *D. corticola*.

In conclusion, results from our study represent an advance in the current knowledge about *MAT* genes in the genus *Diplodia*. The PCR-based assay developed here will be a valuable tool to assess and score the mating types within populations of *Diplodia* in a fast, robust and reliable way, thus helping to unravel cryptic sex and cryptic sexual strategies. Our results also suggest that *MAT* gene analysis is a good approach for accurate and reliable species differentiation within the genus *Diplodia*, either alone or in combination with other loci such as the *tef1-α*. Future studies should expand these analyses to all currently known

Diplodia species; analyse the structure and organization of the *MAT* locus within different species in the genus; and undertake a functional characterization of both *MAT* idiomorphs.

Acknowledgements

This work was financed by European Funds through COMPETE and National Funds through the Portuguese Foundation for Science and Technology (FCT) to project ALIEN (PTDC/AGR-PRO/2183/2014 – POCI-01-0145-FEDER-016788), CESAM (UID/AMB/50017/2013 – POCI-01- 0145-FEDER-007638), Artur Alves (FCT Investigator Programme – IF/00835/2013) and Anabela Lopes (PhD grant – SFRH/BD/85615/2012). Alan JL Phillips acknowledges the support from Biosystems and Integrative Sciences Institute (BioISI, FCT/UID/ Multi/04046/2013).

References

Alves A, Correia A, Luque J, Phillips AJL, 2004. *Botryosphaeria corticola* sp. nov. on *Quercus* species, with notes and description of *Botryosphaeria stevensii* and its anamorph *Diplodia mutila*. *Mycologia* **96**:598–613.

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

Alves A, Linaldeddu BT, Deidda A, Scanu B, Phillips AJL, 2014. The complex of *Diplodia* species associated with *Fraxinus* and some other woody hosts in Italy and Portugal. *Fungal Diversity* **67**:143–156.

Bihon W, Wingfield MJ, Slippers B, Duong TA, Wingfield BD, 2014. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. *Fungal Genetics and Biology* **62**:55–61.

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

Crous PW, Groenewald JZ, Risede JM, Hywel-Jones NL, 2004. *Calonectria* species and their *Cylindrocladium* anamorphs: species with sphaeropedunculate vesicles. *Studies in Mycology* **50**:415–430.

Debuchy R, Turgeon BG, 2006. Mating-Type Structure, Evolution, and Function in Euascomycetes. In: Kües/Fisher (eds.) *The Mycota I - Growth, Differentiation and Sexuality*. Springer-Verlag, Berlin, Heidelberg, pp 293-323.

Dissanayake AJ, Phillips AJL, Li XH, Hyde KD, 2016. *Botryosphaeriaceae*: Current status of genera and species. *Mycosphere* **7**:1001–1073.

Giambra S, Piazza G, Alves A, Mondello V, Berbegal M, Armengol J, Burruano S, 2016. Botryosphaeriaceae species associated with diseased loquat trees in Italy and description of *Diplodia rosacearum* sp. nov. *Mycosphere* **7**: 978–989.

Gioti A, Mushegian AA, Strandberg R, Stajich JE, Johannesson H, 2012. Unidirectional evolutionary transitions in fungal mating systems and the role of transposable elements. *Molecular Biology and Evolution* **29**, 3215–3226.

Glass NL, Donaldson G, 1995. Development of primer sets designed for use with PCR to amplify conserved genes from filamentous ascomycetes. *Applied and Environmental Microbiology* **61**:1323–1330.

González-Domínguez E, Alves A, León M, Armengol J, 2017. Characterization of Botryosphaeriaceae species associated with diseased loquat (*Eriobotrya japonica*) in Spain. *Plant Pathology* **66**:77–89.

Groenewald M, Groenewald JZ, Harrington TC, Abeln ECA, Crous PW, 2006. Mating type gene analysis in apparently asexual *Cercospora* species is suggestive of cryptic sex. *Fungal Genetics and Biology* **43**:813–825.

Hall TA, 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series* **41**:95–98.

Jami F, Slippers B, Wingfield MJ, Gryzenhout M, 2012. Five new species of the *Botryosphaeriaceae* from *Acacia karroo* in South Africa. *Cryptogamie, Mycologie* **33**:245–266.

Kashyap PL, Rai S, Kumar S, Srivastava AK, Anandaraj M, Sharma AK, 2015. Mating type genes and genetic markers to decipher intraspecific variability among *Fusarium udum* isolates from pigeonpea. *Journal of Basic Microbiology* **55**:846–856.

Lazzizera C, Frisullo S, Alves A, Lopes J, Phillips AJL, 2008. Phylogeny and morphology of *Diplodia* species on olives in southern Italy and description of *Diplodia olivarum*. *Fungal Diversity* **31**:63–71.

Lee SC, Ni M, Li W, Shertz C, Heitman J, 2010. The Evolution of Sex: a Perspective from the Fungal Kingdom. *Microbiology and Molecular Biology Reviews* **74**: 298–340.

Linaldeddu BT, Franceschini A, Alves A, Phillips AJL, 2013. *Diplodia quercivora* sp. nov.: a new species of *Diplodia* found on declining *Quercus canariensis* trees in Tunisia. *Mycologia* **105**:1266–1274.

Linaldeddu BT, Maddau L, Franceschini A, Alves A, Phillips AJL, 2016. *Botryosphaeriaceae* species associated with lentisk dieback in Italy and description of *Diplodia insularis* sp. nov. *Mycosphere* **7**:962–977.

Lopes A, Phillips AJL, Alves A, 2017. Mating type genes in the genus *Neofusicoccum*: Mating strategies and usefulness in species delimitation. *Fungal Biology* **121**:394–404.

López-Giráldez F, Townsend JP, 2011. PhyDesign: an online application for profiling phylogenetic informativeness. *BMC Evolutionary Biology* **11**:152.

Lynch SC, Eskalen A, Zambino PJ, Mayorquin JS, Wang DH, 2013. Identification and pathogenicity of *Botryosphaeriaceae* species associated with coast live oak (*Quercus agrifolia*) decline in southern California. *Mycologia* **105**:124–140.

Martin SH, Steenkamp ET, Wingfield MJ, Wingfield BD, 2013. Mate-recognition and species boundaries in the ascomycetes. *Fungal Diversity* **58**:1–12.

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

Ni M, Feretzaki M, Sun S, Wang X, Heitman J, 2011. Sex in fungi. *Annual Review of Genetics* **45**:405–430.

O'Donnell K, Cigelnik E, 1997. Two divergent intragenomic rDNA ITS2 types within a monophyletic lineage of the fungus *Fusarium* are nonorthologous. *Molecular Phylogenetic and Evolution* **7**:103–116.

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

Phillips AJL, Lopes J, Abdollahzadeh J, Bobev S, Alves A, 2012. Resolving the *Diplodia* complex on apple and other *Rosaceae* hosts. *Persoonia* **29**:29–38.

Phillips AJL, Alves A, Abdollahzadeh J, Slippers B, Wingfield MJ, Groenewald JZ, Crous PW, 2013. The *Botryosphaeriaceae*: genera and species known from culture. *Studies in Mycology* **76**:51–167.

Pöggeler S, O'Gorman CM, Hoff B, Kück U, 2011. Molecular organization of the mating-type loci in the homothallic Ascomycete *Eupenicillium crustaceum*. *Fungal Biology* **115**:615–624.

Santos L, Phillips AJL, Crous PW, Alves A, 2017. *Diaporthe* species on Rosaceae with descriptions of *D. pyracanthae* sp. nov. and *D. malorum* sp. nov. *Mycosphere* **8**(5):485-511.

Steenkamp ET, Wingfield BD, Coutinho TA, Zeller KA, Wingfield MJ, Marasas WFO, Leslie JF, 2000. PCR-Based Identification of *MAT-1* and *MAT-2* in the *Gibberella fujikuroi* Species Complex. *Applied and Environmental Microbiology* **66**:4378–4382.

Sun S, Heitman J, 2015. From two to one: Unipolar sexual reproduction. *Fungal Biology Reviews* **29**:118–125.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S, 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* **30**:2725–2729.

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.

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, and Higgins DG, 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research* **25**:4876–4882.

Ueng PP, Dai Q, Cui K-R, Czembor PC, Cunfer BM, Tsang H, Arseniuk E, Bergstrom GC, 2003. Sequence diversity of mating-type genes in *Phaeosphaeria avenaria*. *Current Genetics* **43**:121–130.

White TJ, Bruns T, Lee S, Taylor J, 1990. Amplification and direct sequencing of fungal genes for phylogenies. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, (eds) *PCR Protocols: A Guide to Methods and Applications*. Academic Press, California, pp 315–322.

Wingfield BD, Ades PK, Al-Naemi FA, Beirn LA, Bihon W, Crouch JA, Beer ZW, De Vos L, Duong TA, Fields CJ, Fourie G, Kanzil AM, Malapi-Wight M, Pethybridge SJ, Radwan O, Rendon G, Slippers B, Santana QC, Steenkamp ET, Taylor PWJ, Vaghefi N, van der Merwe N, Veltri D, Wingfield MJ, 2015. Draft genome sequences of *Chrysoporthe austroafricana*, *Diplodia scrobiculata*, *Fusarium nygamai*,

Leptographium lundbergii, *Limonomyces culmigenus*, *Stagonosporopsis tanaceti*, and *Thielaviopsis punctulata*. *IMA Fungus* **6**: 233–248.

Yang T, Groenewald JZ, Ratchadawan C, Fahimeh J, Abdollahzadeh J, Lombard L, Crous PW, 2017. Families, genera, and species of Botryosphaeriales. *Fungal Biology* 121: 322–346.

Yokoyama E, Arakawa M, Yamagishi K, Hara A, 2006. Phylogenetic and structural analyses of the mating-type loci in *Clavicipitaceae*. *FEMS Microbiology Letters* **264**:182–191.

Supplementary Data

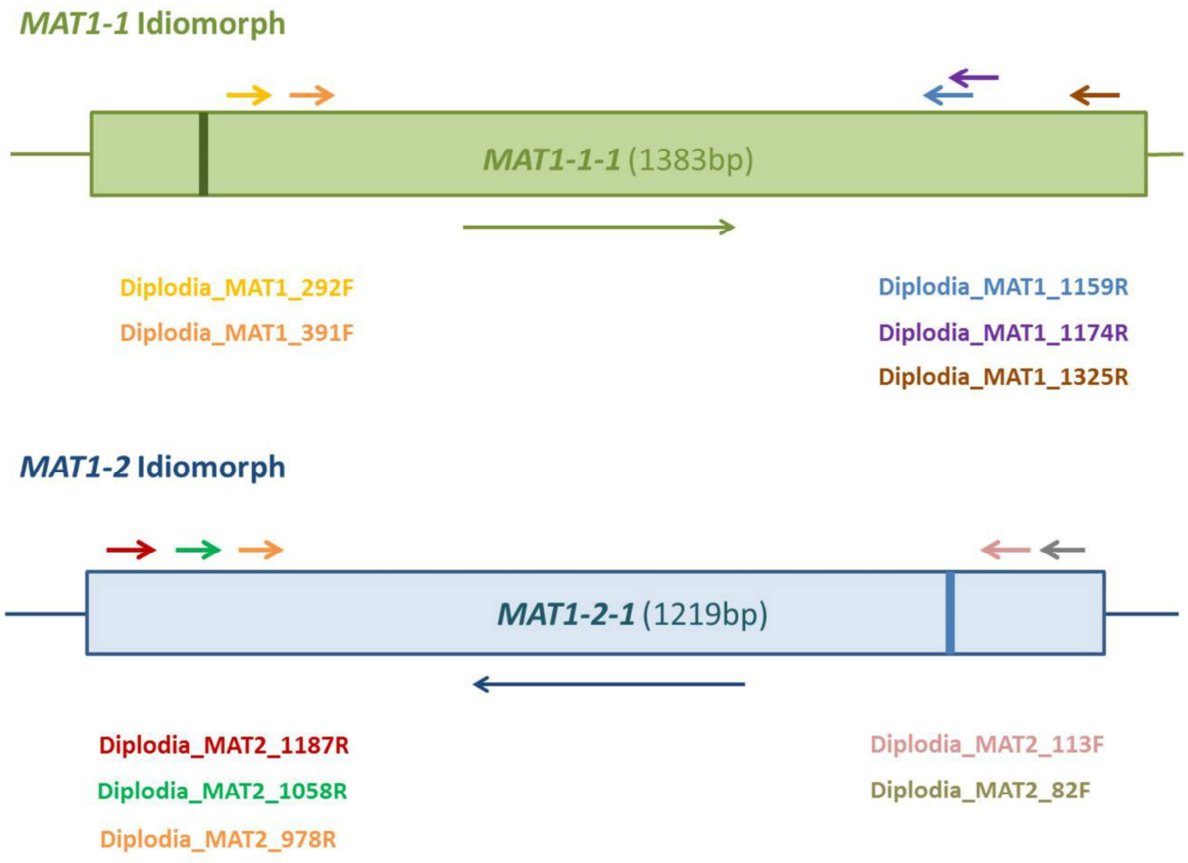

Figure S4.1 - Binding sites scheme of all the primers designed and tested for *MAT* loci. Introns are indicated by the vertical lines inside the genes. Arrows below the genes give the gene orientation (5'-3'). Distances and sizes are not drawn to a scale.

Table S4.1 - Possible primers combinations for *MAT1-1-1* gene.

Primers \ Species	*D. africana*	*D. corticola*	*D. cupressi*	*D. eriobotryicola*	*D. fraxini*	*D. insularis*	*D. intermedia*
Diplodia_MAT1_292F/Diplodia_MAT1_1159R (829bp)	NT	NT	-	-	-	-	NT
Diplodia_MAT1_292F/Diplodia_MAT1_1174R (842bp)	+	+	-	-	-	-	+
Diplodia_MAT1_292F/Diplodia_MAT1_1325R (993bp)	NT	NT	-	-	-	NT	NT
Diplodia_MAT1_391F/Diplodia_MAT1_1159R (729bp)	NT	NT	-	+	-	+	NT
Diplodia_MAT1_391F/Diplodia_MAT1_1174R (742bp)	NT	NT	-	+	-	+	NT
Diplodia_MAT1_391F/Diplodia_MAT1_1325R (893bp)	+	+	+	+	+	+	+

Primers \ Species	*D. malorum*	*D. mutila*	*D. olivarum*	*D. rosulata*	*D. sapinea*	*D. seriata*
Diplodia_MAT1_292F/Diplodia_MAT1_1159R (829bp)	-	NT	NT	-	NT	NT
Diplodia_MAT1_292F/Diplodia_MAT1_1174R (842bp)	-	-	+	-	+	+
Diplodia_MAT1_292F/Diplodia_MAT1_1325R (993bp)	-	NT	NT	NT	NT	NT
Diplodia_MAT1_391F/Diplodia_MAT1_1159R (729bp)	-	+	NT	+	NT	NT
Diplodia_MAT1_391F/Diplodia_MAT1_1174R (742bp)	-	NT	NT	-	NT	NT
Diplodia_MAT1_391F/Diplodia_MAT1_1325R (893bp)	+	+	+	+	+	+

Legend:
(+) - Specific amplification
(-) - Nonspecific or without amplification
NT - Not tested
bp - Base pair

Primers Set A
Expected sizes of the fragments based on *D. sapinea*

Table S4.2 - Possible primers combinations for *MAT1-2-1* gene.

Primers \ Species	*D. alatafructa*	*D. corticola*	*D. insularis*	*D. intermedia*	*D. mutila*	*D. olivarum*	*D. pseudoseriata*
Diplodia_MAT2_82F/Diplodia_MAT2_978R (854bp)	+	-	+	NT	-	NT	+
Diplodia_MAT2_82F/Diplodia_MAT2_1058R (935bp)	-	-	-	+	-	NT	-
Diplodia_MAT2_82F/Diplodia_MAT2_1187R (1066bp)	+	-	+	NT	-	NT	+
Diplodia_MAT2_113F/Diplodia_MAT2_978R (824bp)	-	-	-	NT	-	NT	-
Diplodia_MAT2_113F/Diplodia_MAT2_1058R (905bp)	-	-	-	NT	-	NT	-
Diplodia_MAT2_113F/Diplodia_MAT2_1187R (1036bp)	-	+	-	-	+	+	-

Primers \ Species	*D. quercivora*	*D. rosacearum*	*D. sapinea*	*D. scrobiculata*	*D. seriata*	*D. subglobosa*	*D. tsugae*
Diplodia_MAT2_82F/Diplodia_MAT2_978R (854bp)	-	NT	NT	NT	+	-	-
Diplodia_MAT2_82F/Diplodia_MAT2_1058R (935bp)	-	+	+	+	+	-	-
Diplodia_MAT2_82F/Diplodia_MAT2_1187R (1066bp)	+	NT	NT	NT	NT	-	-
Diplodia_MAT2_113F/Diplodia_MAT2_978R (824bp)	-	NT	NT	NT	+	-	-
Diplodia_MAT2_113F/Diplodia_MAT2_1058R (905bp)	-	NT	NT	NT	+	-	-
Diplodia_MAT2_113F/Diplodia_MAT2_1187R (1036bp)	+	-	-	-	+	+	+

Legend:
(+) - Specific amplification
(-) - Nonspecific or without amplification
NT - Not tested
bp - Base pair

Primers Set B **Primers Set D**

Primers Set C

Expected sizes of the fragments based on *D. sapinea*

CHAPTER 5

Genomic organization of mating type loci in the order *Botryosphaeriales*

Abstract

The study of *MAT* loci and is of relevance address diverse aspects of fungal biology other than their sexual reproduction. However, due to their high interspecific diversity it is sometimes quite difficult to study these genomic regions through conventional DNA cloning strategies or PCR-based strategies. However, the recent increase of fungal genome sequencing projects greatly facilitated this task. Here analyses of the genomic organization of *MAT* loci and their flanking regions were performed on the available sequenced genomes of species belonging to the order *Botryosphaeriales*. The main *MAT* genes (*MAT1-1-1* and *MAT1-2-1*), as well as additional *MAT* genes and genes present in the flanking regions, were identified and properly annotated when needed. Results showed that heterothallism is the most prevalent reproductive strategy among the species analysed. With few exceptions, genomic organization of the *MAT* loci revealed some degree of gene synteny between the different families. From all species analyzed, *Neofusicoccum parvum* showed the most divergent structure of the locus. Future studies are needed to clarify the evolutionary trajectory of the *MAT* locus in the order *Botryosphaeriales*.

Keywords: *Aplosporellaceae*, *Botryosphaeriaceae*, *MAT* genes, *Phyllostictaceae*, sexual strategies, *Saccharataceae*

Introduction

The order *Botryosphaeriales* (*Ascomycota*, *Dothideomycetes*) harbours nine families with several genera of endophytes or pathogens from a wide range of woody hosts (Slippers et al. 2017; Yang et al. 2017). For many years the fungal taxa were defined based on morphology alone which resulted in an underestimation of the real diversity. To solve this problem molecular techniques based on DNA sequence data has been used in the last years. In the order *Botryosphaeriales*, for instance, the most complete phylogenies of representative groups have been based on a dataset of six molecular markers (Slippers et al. 2013). Apart from these molecular markers *MAT* genes have also been used to solve phylogenetic issues (Steenkamp et al. 2000; Ueng et al. 2003; O'Donnell et al. 2004; Pöggeler et al. 2011; Martin et al. 2013; Kashyap et al. 2015), including in the *Botryosphaeriales* (Amorim et al. 2017; Lopes et al.

2017a,b). Besides phylogenetic studies, the knowledge of *MAT* genes and *MAT* locus organization can provide insights into several aspects of fungal biology such as evolution, sexual reproduction, population genetics, and epidemiology, among others (Debuchy et al. 2010).

MAT genes are responsible for determining sexual identity and controlling sexual reproduction and are located, in the ascomycetes, in a single locus (*MAT1* locus) under two alternate forms (idiomorphs), defining the mating types *MAT1-1* and *MAT1-2* (Lee et al. 2010; Ni et al. 2011; Sun & Heitman 2015). The *MAT1-1* idiomorph contains the *MAT1-1-1* gene that encodes an alpha-box protein, and the *MAT1-2* idiomorph is characterized by the presence of *MAT1-2-1* gene that encodes a high mobility group (HMG)-domain protein (Lee et al. 2010; Ni et al. 2011). In addition to these master genes, other genes may be present in the *MAT* locus (Debuchy et al. 2010). Heterothallic species have only one mating type and the sexual reproduction occurs when opposite mating type partners interact (Debuchy et al. 2010). On the contrary, both *MAT1-1-1* and *MAT1-2-1* genes are present in homothallic fungi either in a fused *MAT* locus or in proximity. For these species the initiation of sexual reproduction is not dependent on the interaction with an isolate of opposite mating type (Debuchy et al. 2010).

Information about the structure and organization of *MAT* genes allows the development of molecular markers that can be used to determine the mating types of isolates, without the laborious and time consuming process of developing and crossing mating tester strains, to infer about the reproductive mode and mating systems of a fungal species (Duong et al. 2016; Wang et al. 2016; Wey et al. 2016; Amorim et al. 2017) and to study patterns of genetic diversity (Duong et al. 2015; Kashyap et al. 2015). Such information is particularly important for fungal pathogens as sexual and asexual reproduction have significantly different effects on the population structures of the pathogens, which in practical situations require different disease management strategies (MacDonald & Linde 2002). Evolutionary advantages could be inherent to sexual populations since more fit genotypes may arise through recombination which could be reflected in the mode and distance of dispersal as well as longevity (MacDonald & Linde 2002).

Because of their relevance *MAT* loci have been cloned and characterized from several filamentous ascomycetes (Belfiori et al. 2013; Geng et al. 2014; Hughes et al. 2014; Putman et

al. 2015). However, from all those only three belong to species of the order *Botryosphaeriales* (Bihon et al. 2014; Wang et al. 2016; Amorim et al. 2017). For that reason in this work our goals were: (i) to search in the public databases for available genomes of species belonging to the order *Botryosphaeriales*, (ii) to identify the *MAT* loci in each genome, (iii) to characterize the global organization of each locus and flanking regions, (iv) to compare the genomic architecture of *MAT* loci from members of different families in the *Botryosphaeriales*.

Materials and Methods

The draft genomes used for the identification of *MAT* locus are listed in table 5.1. The known idiomorphs of *Diplodia sapinea* (Bihon et al. 2014), *Phyllosticta citricarpa* (Wang et al. 2016; Amorim et al. 2017) and *Phyllosticta capitalensis* (Wang et al. 2016) were the references to search for orthologs in the remaining genomes using BLASTn and BLASTp algorithms against the National Center for Biotechnology Information (NCBI) database or against the Joint Genome Institute (JGI) Genome Portal database. Once there was a hit, the relevant contig was selected for further analysis. For those genomes that were already annotated the identification, position, orientation and length of the genes were assumed as the same as in the annotation with some exceptions that are discussed in the results/discussion section. In the non-annotated genomes, the putative open reading frames (ORFs) in the idiomorphs and flanking regions were predicted with the FGENESH tool (Solovyev et al. 2006) in the MOLQUEST software package (www.softberry.com) using the *Diplodia seriata* or *Phyllosticta citricarpa* data set as reference. The schematic organization of the loci was created using the free software SnapGene Viewer 4.0.1 (SnapGene.com).

Table 5.1 - General information about the genomes and *MAT* loci analysed in this study.

Family	Species[a]	Genome Accession	MAT Locus Location	Mating Strategy	Mating type	References
Aplosporellaceae	*Aplosporella prunicola* CBS121167	JGI Fungal Genome Portal Project ID: 1006427	Scaffold 3:898312-944247	Heterothallic	MAT1-1	-
Botryosphaeriaceae	*Botryosphaeria dothidea* **CBS115476**	JGI Fungal Genome Portal	Scaffold 547:757-28535	Homothallic	MAT1-1 and MAT1-2	Marsberg et al. 2017
	Diplodia corticola **CBS112549**	NCBI PRJNA325745	MNUE01000007 (Scaffold 7: 103897-130618)	Heterothallic	MAT1-1	-
	Diplodia sapinea CMW190	NCBI PRJNA215898	KF551229	Heterothallic	MAT1-1	Bihon et al. 2014; van der Nest et al. 2014
	Diplodia sapinea CMW39103	NCBI PRJNA242796	KF551228	Heterothallic	MAT1-2	Bihon et al. 2014; van der Nest et al. 2014
	Diplodia scrobiculata CMW30223	NCBI PRJNA278001	LAEG01000830 (Contig 831:15582-26956) LAEG01002648 (Contig 2649: 1993-2751) LAEG01002687 (Contig 2688:106-6497)	Heterothallic	MAT1-1	Wingfield et al. 2015

Family	Species[a]	Genome Accession	MAT Locus Location	Mating Strategy	Mating type	References
	Diplodia seriata F98.1	NCBI PRJNA350273	MSZU01000074 (Scaffold 1: 1073334-1100549)	Heterothallic	MAT1-1	-
	Diplodia seriata DS831	NCBI PRJNA261773	LAQI01000001 (Scaffold v01.1: 216499-241489)	Heterothallic	MAT1-2	-
	Lasiodiplodia theobromae CSS-01s	NCBI PRJNA339237	MDYX01000006 (Contig 6: 791528-831439)	Heterothallic	MAT1-2	-
	Macrophomina phaseolina MS6	NCBI PRJNA78845	AHHD01000467 (Contig 00467: 114094-141461)	Heterothallic	MAT1-1	Islam et al. 2012
	Macrophomina phaseolina MP00003	NCBI PRJNA291855	LHTM01000070 (Scaffold 28:432-20110) LHTM01002900 (Scaffold 942:1370-7153)	Heterothallic	MAT1-2	-
	Neofusicoccum parvum UCR-NP2	NCBI PRJNA187491	AORE01000290 (Contig 290:38915-54033) AORE01000875 (Contig 875:4952-13811) AORE01000876 (Contig 876:1-4459)	Homothallic	MAT1-1 and MAT1-2	Blanco-Ulate et al. 2013
	Neoscytalidium dimidiatum UM 880	NCBI PRJEB14660	FLVB01000333 (Contig 333:45410-74434)	Heterothallic	MAT1-1-1	-

Family	Species[a]	Genome Accession	MAT Locus Location	Mating Strategy	Mating type	References
Phyllostictaceae	*Phyllosticta capitalensis* Gm33	NCBI PRJNA301361	LOEO01000513 (Contig 513:10643-28784) LOEO01001011 (Contig 01011:754-8653)	Homothallic	MAT1-1 and MAT1-2	Wang et al. 2016
	Phyllosticta citriasiana CBS 120486	JGI Fungal Genome Portal Project ID: 1011301	Scaffold 34:93636-116407	Heterothallic	MAT1-1	-
	Phyllosticta citricarpa CGMCC3.14348	NCBI PRJNA188924	AOTE01003968 (Contig 4070:6463-19369) AOTE01003969 (Contig 4071:1-9497)	Heterothallic	MAT1-1	Amorim et al. 2016
	Phyllosticta citricarpa Gc12	NCBI PRJNA301361	LOEN01001345 (Contig 01345:1-1743) LOEN01002477 (Contig 02477:7300-12388) LOEN01003787 (Contig 03787:1-16464)	Heterothallic	MAT1-2	Wang et al. 2016
Saccharataceae	*Saccharata protea* CBS121410	JGI Fungal Genome Portal Project ID: 1011317	Scaffold 3:603071-633886	Heterothallic	MAT1-2	-

[a]Isolates in bold are ex-type cultures.

Phylogenetic analyses

The sequences used in the phylogenetic analyses are listed in table 5.1 (*MAT1-1-1* and *MAT1-2-1*) and table 5.2 (ITS and *tef1-α*). Sequences were aligned with ClustalX v. 2.1 (Thompson et al. 1997), using the following parameters: pairwise alignment parameters (gap opening = 10, gap extension = 0.1) and multiple alignment parameters (gap opening = 10, gap extension = 0.2, transition weight = 0.5, delay divergent sequences = 25 %). Alignments were checked and edited with BioEdit Alignment Editor v. 7.2.5 (Hall 1999). Phylogenetic analyses of sequence data were done with MEGA6 v. 6.06 (Tamura et al. 2013). Maximum likelihood (ML) analyses were performed on a Neighbor-Joining (NJ) starting tree automatically generated by the software. All gaps were included in the analyses. Mating type genes phylogenies were inferred based on the Hasegawa-Kishino-Yano model and the concatenated data of ITS-*tef1*-α was based on the Kimura 2-parameter model. A discrete Gamma distribution was used to model evolutionary rate differences among sites for concatenated data phylogeny while for both *MAT* genes phylogenies was applied a Gamma distribution with invariant sites. A bootstrap analysis (1000 replicates) was used to estimate the consistency of each node of the trees.

Table 5.2 - Accession numbers/locations of the ITS and *tef1*-α sequences used in the phylogenetic tree.

	Accession numbers/Location	
Species	**ITS**	***tef1*-α**
Aplosporella prunicola CBS121167	KF766147	Scaffold 19:144554-146768
Botryosphaeria dothidea CBS115476	AY236949	AY236898
Diplodia corticola CBS112549	AY259100	AY573227
Diplodia sapinea CMW190	KF766159	AY624251
Diplodia sapinea CMW39103	JHUM01000052 (Contig 52:2714-3031)	JHUM01000933 (Contig 941:10016-10280)
Diplodia scrobiculata CMW30223	HM100278	HM100269
Diplodia seriata DS831	KP296243	LAQI01000236 (Scaffold 236:20351-20615)
Diplodia seriata F98.1	KU568530	MSZU01000076 (Scaffold 11: 377015-376506)
Lasiodiplodia theobromae **CSS-01s**	MDYX01000025 (Contig 25:4483-4805)	MDYX01000018 (Contig 18:10913-11178)
Macrophomina phaseolina MS6	AHHD01000808 (Contig 00813:1-187) AHHD01000980 (Contig 00996:1-136)	AHHD01000250 (Contig 00250:98970-99200)
Macrophomina phaseolina MP00003	LHTM01002602 (Scaffold 1571:2803-3125)	LHTM01002443 (Scaffold 449:595-825)
Neofusicoccum parvum UCR-NP2	AORE01001444 (Contig 1444:5389-5707)	AORE01000046 (Contig 46:95433-95656)
Neoscytalidium dimidiatum UM 880	FLVB01001845 (Contig 1845:3682-4000)	FLVB01001781 (Contig 1781:10710-10952)

	Accession numbers/Location	
Species	**ITS**	***tef1-α***
Phyllosticta capitalensis Gm33	LOEO01000224 (Contig 00224:370-710)	LOEO01000452 (Contig 00452:17829-18062)
Phyllosticta citriasiana CBS120486	FJ538360	FJ538418
Phyllosticta citricarpa CGMCC3.14348	JN791637	JN791482
Phyllosticta citricarpa Gc12	LOEN01000348 (Contig 00348:1-276)	LOEN01002819 (Contig 02819:10381-10614)
Saccharata protea CBS121410	KX464238	KX464775

Results and Discussion

The increasing number of fungal genomes now available facilitates the study of the structure, function, and evolution of *MAT* loci. It is known, for instance, that the structure of the mating type locus in most heterothallic ascomycetes is relatively conserved and usually flanked by the DNA lyase (APN1/2) and/or cytoskeleton assembly control (SLA2) genes (Debuchy et al. 2010). However, the *MAT* locus structure of each fungal species is unique in terms of gene number and arrangement, and transcriptional orientation (Butler 2007). Also, the molecular organization of the mating type locus in different homothallic species is often singular. The two idiomorphs can exist within the same fungal strain either as a single fused *MAT1-1*/*MAT1-2* locus or in separate loci located in different regions of the genome (Debuchy et al. 2010). Finally, it is known that both homothallic and heterothallic mating systems co-exist and are closely aligned throughout the Ascomycota, but with a higher prevalence of heterothallism in some lineages (Butler 2007). Bearing in mind all these informations, the purpose of this study was to take advantage of the available sequenced genomes from species of *Botryosphaeriales* and gain knowledge about the structure and organization of the *MAT* loci, as well as the sexual reproductive strategies used by member of different families.

In the public databases NCBI and JGI there are available sequenced genomes of species belonging to four families of the order *Botryosphaeriales*, namely the families

Aplosporellaceae (*Aplosporella prunicola*), *Botryosphaeriaceae* (*Botryosphaeria dothidea*, *Diplodia corticola*, *Diplodia sapinea*, *Diplodia scrobiculata*, *Diplodia seriata*, *Lasiodiplodia theobromae*, *Macrophomina phaseolina*, *Neofusicoccum parvum*, *Neoscytalidium dimidiatum*), *Phyllostictaceae* (*Phyllosticta capitalensis*, *Phyllosticta citriasiana*, *Phyllosticta citricarpa*) and *Saccharataceae* (*Saccharata protea*) (Table 5.1). Figure 5.1 shows the structure and organization of each *MAT* locus and their respective flanking regions and also the phylogenetic relationships between the different species and families. Analysis of the *MAT* loci revealed some degree of gene synteny even between different families. With the exception of *N. parvum*, the flanking regions of all species analyzed present the SAICAR, Mitochondrial carrier, CIA30, APC5, Cox-VIa and DNA lyase genes. Except for *L. theobromae*, these genes are always in the same order and are found downstream of *MAT* genes. In some cases the DNA lyase is near the *MAT* genes (*A. prunicola*, *B. dothidea*, *M. phaseolina*, *Phyllosticta* species and *S. proteae*) but in the others is the SAICAR gene that is closest to *MAT* genes (*Diplodia* species). In the particular case of *L. theobromae* it is observed a switch between the DNA lyase and Cox-VIa position being the last one closest to *MAT1-2-1* gene. The presence of this set of genes was previously reported in the species of *D. sapinea* (Bihon et al. 2014) and *P. citricarpa* (Amorim et al. 2017) but also in species belonging to the order *Pleosporales* (Vaghefi et al. 2015), which also belong in the Dothideomycetes. Individually, Cox-VIa and APC5 genes have also been found in some *MAT* loci (Cozijnsen & Howlett 2003; Li et al. 2010; Tsui et al 2013). In the flanking regions, including in the *MAT* locus of *N. parvum*, a precursor of a protein with a pleckstrin homology (PH) domain was always found in the opposite side of the genes mentioned above. In the families *Aplosporellaceae*, *Phyllostictaceae* and *Saccharataceae* the PH-domain like gene is accompanied by a precursor of a ribosomal protein while in some species of the family *Botryosphaeriaceae* the PH-domain like gene is located near to a cutinase precursor. This cutinase precursor is also present in the families *Aplosporellaceae* and *Saccharataceae* in the opposite side of PH-domain like gene but absent in the family *Phyllostictaceae*.

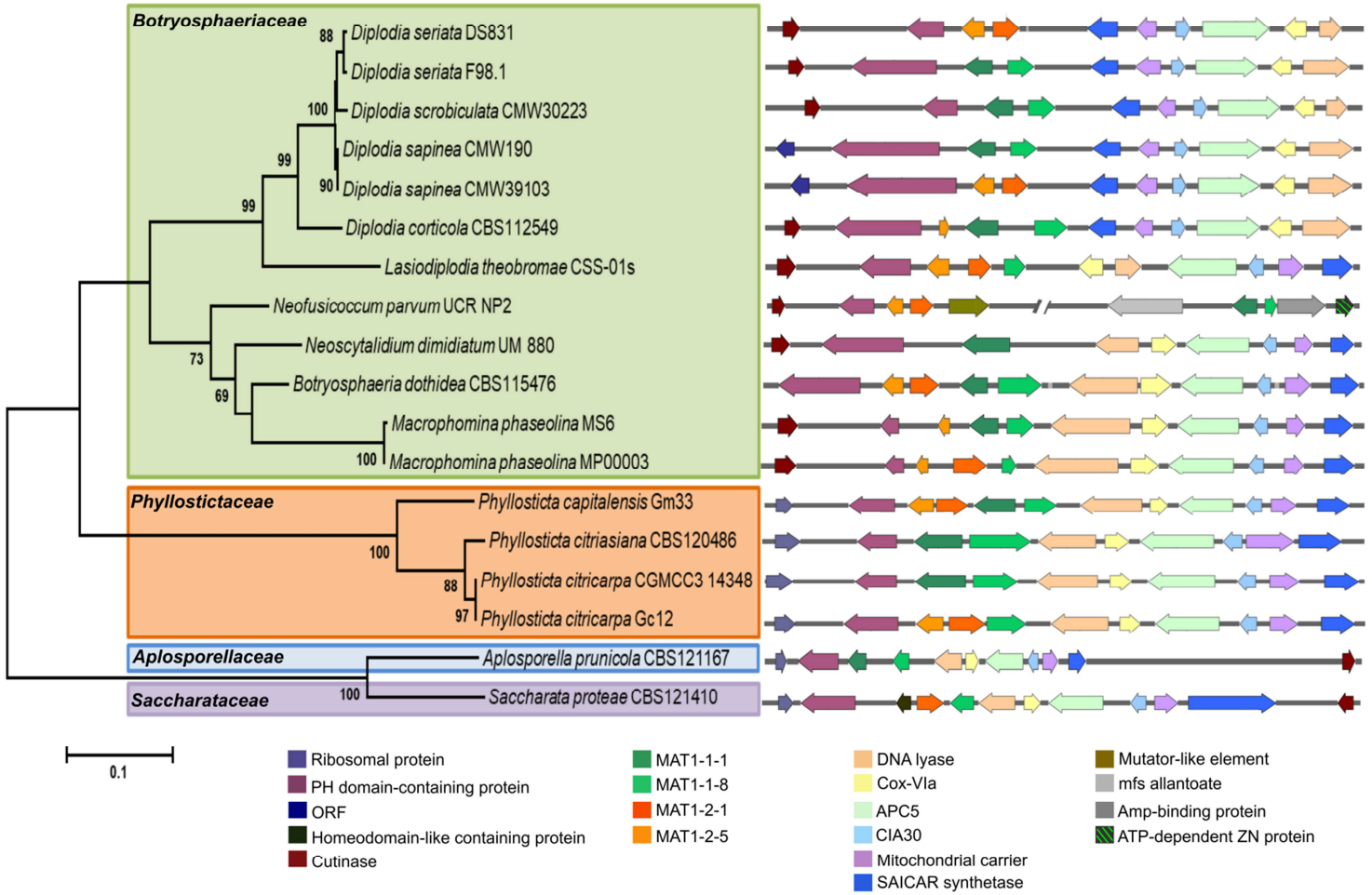

Figure 5.1 – Structure of the *MAT* locus (including flanking regions) and phylogenetic relationships between species and families based on ITS-*tef1-α* concatenated sequence data. Bootstrap values (>70%) are given at the nodes. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site.

The identification of the main *MAT* genes (*MAT1-1-1* and/or *MAT1-2-1*) in all the *MAT* loci allowed us to infer about the mating strategy adopted by each species. The majority of the studied species are heterothallic which means that only one mating type (*MAT1-1* or *MAT1-2*) is present in the locus. For the species *D. sapinea*, *D. seriata*, *M. phaseolina* and *P. citricarpa* the sequencing of both idiomorphs confirms that they differ only in the *MAT* genes content having identical flanking regions. The species *B. dothidea*, *N. parvum* and *P. capitalensis* have both *MAT1-1-1* and *MAT1-2-1* genes thus confirming homothallism as the mating strategy. Previous studies showed that homothallism is the prevalent mating strategy among *Neofusicoccum* species (Lopes et al. 2017a) while heterothallism was the unique strategy observed so far within the genus *Diplodia* (Lopes et al. 2017b). However, further studies are necessary to conclude about the prevalence of each mating strategy in the remaining genera as well as to infer about the evolution of *MAT* loci.

Beyond the main *MAT* genes additional genes were found in the idiomorphs. These genes are often not annotated and identified only as open reading frame (e.g.in the species *B. dothidea*) or not identified at all (e.g. in strain DS831 of *D. seriata*). In these cases the genes were predicted and/or identified using the FGNESH software and BLASTp algorithms respectively. With the exception of *Neoscytalidium dimidiatum*, where no additional *MAT* gene was found, all the *MAT1-1* loci analysed contain the gene *MAT1-1-8*. Until recently this gene was named as *MAT1-1-4* in *D. sapinea* and *P. citricarpa* (Bihon et al. 2014; Amorim et al. 2017) but Wilken et al. (2017) argued that the low similarity between the original *MAT1-1-4* protein (a metallothionein) and the protein of *D. sapinea* justify a different designation (*MAT1-1-8*). In the *MAT1-2* loci the additional gene *MAT1-2-5* is always present with the exception of *S. proteae*. This gene was denominated by Wang et al. (2016) as *MAT1-2-9* in the species *P. citricarpa* and *P. capitalensis* and later as *MAT1-2-5* by Amorim et al. (2017) in *P. citricarpa*. In this study we adopted the designation *MAT1-2-5*, according to the nomenclature recommendation of Wilken et al. (2017) for the class *Dothideomycetes*. Curiously, in the *MAT1-2* locus of *S. proteae* an open reading frame coding for a protein belonging to a homeodomain superfamily was found. This gene has not an evident homology with *MAT1-2-5* gene found in the remaining families. It would be interesting to understand if the gene in *S. proteae* belongs or not to the *MAT* locus. Further analysis of a *MAT1-1* locus of *S. proteae*

could give some insights about this question. *MAT* genes with homeodomains are typical in the basidiomycetes and also in the subphylum of *Saccharomycotina* but not in the subphylum *Pezizomycotin*a (Dyer 2008).

The analysis of *MAT* loci revealed also a frequent presence of additional genes belonging to one mating type in the opposite mating type. For example, the gene *MAT1-1-8* was found in the *MAT1-2* locus of the species *M. phaseolina*, *D. sapinea* (Bihon et al. 2014), *S. proteae* and *P. citricarpa*. On the other hand, in the *MAT1-1* locus of the species *M. phaseolina, D. corticola* and *D. sapinea* (Bihon et al. 2014) the *MAT1-2-5* gene was identified. The genes found in the opposite mating type are usually truncated pseudogenes because part of the original sequence is missing due to possible deletions, translocations or unequal crossovers (Gioti et al. 2012). Although not typical this situation is not unusual and has been used to study the origin and evolution of mating systems (Tsui et al. 2013).

A completely different *MAT* locus structure was found in the species *N. parvum*. Although homothallic species could present a unique structure, we cannot find an explanation for such difference between the locus of *N. parvum* and the remaining homothallic species here analysed. Apart from the distance between the two mating types, the flanking regions are completely different except upstream of the *MAT1-2* idiomorph were cutinase and PH domain genes were identified. It would be interesting to analyse the *MAT* locus of other homothallic species as well as of heterothallic species of the genus *Neofusicoccum* in order to properly address this issue and study the evolution of the *MAT* locus in the genus.

The phylogenies inferred by ITS-*tef1-α, MAT1-1-1* and *MAT1-2-1* genes clearly grouped all species in the respective family (Fig 5.2). However, the organization and composition of all *MAT* loci are overall very similar which probably points for a recent divergence. The family *Botryosphaeriaceae* exhibits more heterogeneity between *MAT* loci, but this is most likely a reflection of the larger number of genera analysed. Further analyses with more families and genera will surely allow a better understanding of this issue.

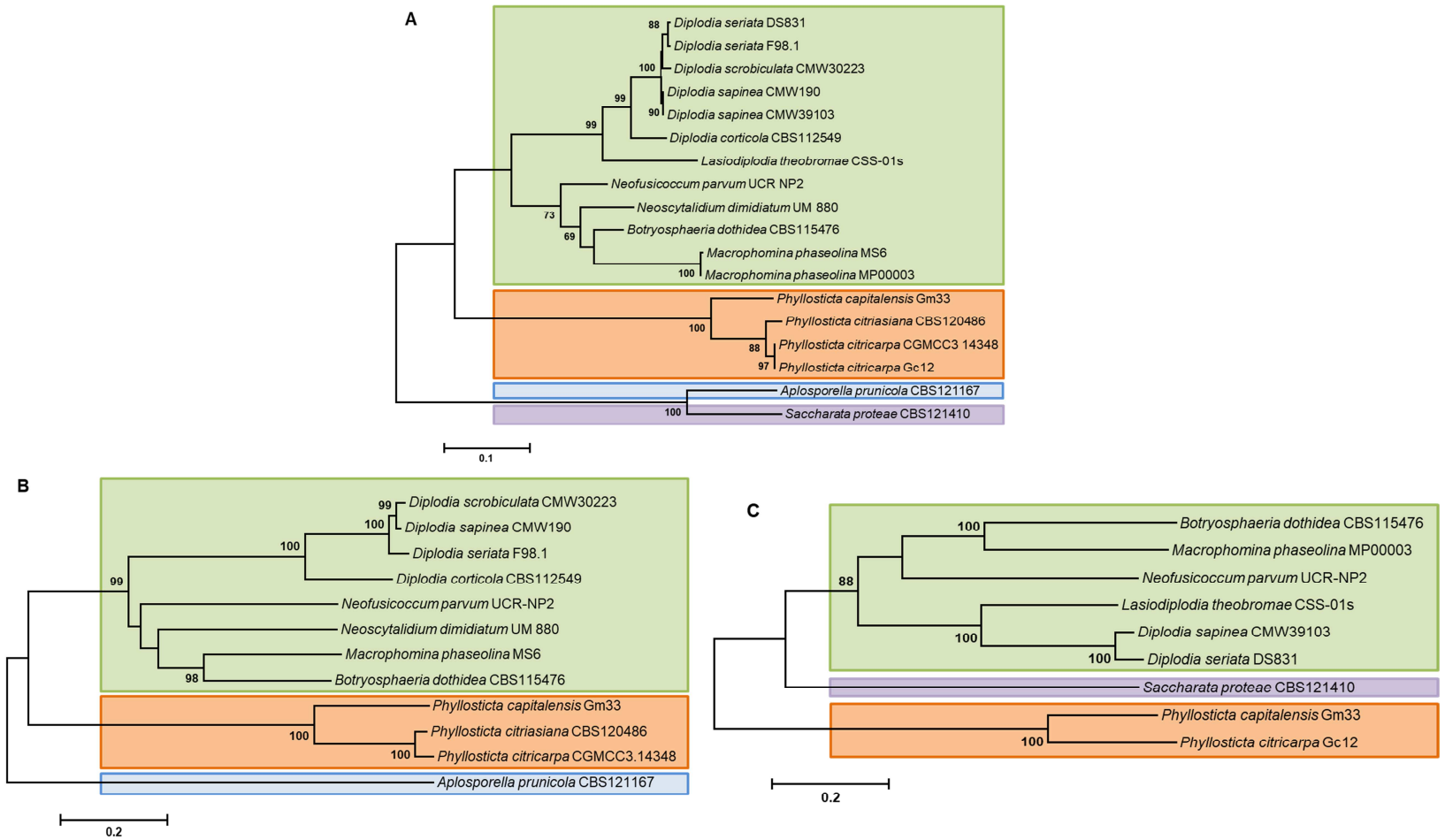

Figure 5.2 - Phylogenetic relationships of species belonging to the families *Botryosphaeriaceae*, *Phyllostictaceae*, *Aplosporellaceae* and *Saccharataceae* based on the sequence data of ITS-*tef1-α* (A) *MAT1-1-1* (B) and *MAT1-2-1* (C) regions. Bootstrap values (>70%) are given at the nodes. The trees are drawn to scale, with branch length measured in the number of substitutions per site.

These preliminary analyses lay the foundations for future studies addressing the structure and evolution of *MAT* loci in the *Botryosphaeriales*. The few studies about *MAT* genes in the order (Bihon et al. 2014; Lopes et al. 2017a,b) proved that they are very different and difficult to access without genus specific primers. On the other hand it was shown here that although the genes present in flanking regions are conserved, the order and direction could be different. This is important to have in mind if the objective is to access the *MAT* locus by the design of primers for the flanking regions. To get around these issues we expect that more genomes will be available soon in order to get a more complete knowledge about *MAT* loci in this relevant group of fungi.

References

Amorim R, Savi DC, Ferreira-Maba L, Aluizio R, Goulin EH, Takita MA, Machado MA, Glienke C, 2017. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in the citrus pathogen *Phyllosticta citricarpa*. *European Journal of Plant Pathology* **147**:325-337.

Belfiori B, Riccioni C, Paolocci F, Rubini A, 2013. Mating type locus of chinese black truffles reveals heterothallism and the presence of cryptic species within the *T. indicum* species complex. *PLoS ONE* **8**:e82353.

Bihon W, Wingfield MJ, Slippers B, Duong TA, Wingfield BD, 2014. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. *Fungal Genetics and Biology* **62**:55-61.

Blanco-Ulate B, Rolshausen P, Cantu D, 2013. Draft Genome Sequence of *Neofusicoccum parvum* Isolate UCR-NP2, a Fungal Vascular Pathogen Associated with Grapevine Cankers. *Genome Announcements* **1**:e00339-13.

Butler G, 2007. The evolution of MAT: the ascomycetes. In: Heitman J, Kronstad JW, Taylor J, Casselton L (eds), Sex in fungi: molecular determination and evolutionary implications. *ASM Press*, Washington, DC, pp. 3-18.

Cozijnsen AJ, Howlett BJ, 2003. Characterisation of the mating-type locus of the plant pathogenic ascomycete *Leptosphaeria maculans*. *Current Genetics* **43**:351–357.

Debuchy R, Berteaux-Lecellier V, Silar P, 2010. Mating systems and sexual morphogenesis in Ascomycetes. In: Borkovich KA, Ebbole DJ (eds.), Cellular and Molecular Biology of Filamentous Fungi. *ASM Press*, Washington, DC, pp. 501–535.

Duong TA, de Beer ZW, Wingfield BD, Eckhardt LG, Wingfield MJ, 2015. Microsatellite and mating type markers reveal unexpected patterns of genetic diversity in the pine root-infecting fungus *Grosmannia alacris*. *Plant Pathology* **64**, 235–242.

Duong TA, de Beer ZW, Wingfield BD, Wingfield MJ, 2016. Mating type markers reveal high levels of heterothallism in *Leptographium sensu* lato. *Fungal Biology* **120**:538-546.

Dyer PS, 2008. Evolutionary Biology: Genomic Clues to Original Sex in Fungi. *Current Biology* **18**:207-209.

Geng Y, Li Z, Xia L-Y, Wang Q, Hu X-M, Zhang X-G, 2014. Characterization and phylogenetic analysis of the mating-type loci in the asexual ascomycete genus *Ulocladium*. *Mycologia* **106**:649–665.

Gioti A, Mushegian AA, Strandberg R, Stajich JE, Johannesson H, 2012. Unidirectional Evolutionary Transitions in Fungal Mating Systems and the Role of Transposable Elements. *Molecular Biology and Evolution* **29**:3215–3226.

Hughes TJ, O'Donnell K, Sink S, Rooney AP, Scandiani MM, Luque A, Bhattacharyya MK, Huang X, 2014. Genetic architecture and evolution of the mating type locus in fusaria that cause soybean sudden death syndrome and bean root rot. *Mycologia* **106**:686–697.

Islam MS, Haque MS, Islam MM, Emdad EM, Halim A, Hossen QMM, Hossain MZ, Ahmed B, Rahim S, Rahman MS, Alam MM, Hou S, Wan X, Saito JA, Alam M, 2012. Tools

to kill: Genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. *BMC Genomics* **13**:493.

Kashyap PL, Rai S, Kumar S, Srivastava AK, Anandaraj M, Sharma AK, 2015. Mating type genes and genetic markers to decipher intraspecific variability among *Fusarium udum* isolates from pigeonpea. *Journal of Basic Microbiology* **55**:846-856.

Lee SC, Ni M, Li W, Shertz C, Heitman J, 2010. The Evolution of Sex: a Perspective from the Fungal Kingdom. *Microbiology and Molecular Biology Reviews* **74**: 298–340.

Li W, Metin B, White TC, Heitman J, 2010. Organization and Evolutionary Trajectory of the Mating Type (MAT) Locus in Dermatophyte and Dimorphic Fungal Pathogens. *Eukaryotic Cell* **9**:46-58.

Lopes A, Phillips AJL, Alves A, 2017a. Mating type genes in the genus *Neofusicoccum*: Mating strategies and usefulness in species delimitation. *Fungal Biology* **121**:394-404.

Lopes A, Linaldeddu B, Phillips AJL, Alves A, 2017b. Mating type gene analyses in the genus Diplodia: from cryptic sex to cryptic species. Submitted to *Fungal Biology*.

Marsberg A, Kemler M, Jami F, Nagel JH, Postma-Smidt A, Naidoo S, Wingfield MJ, Crous PW, Spatafora JW, Hesse CN, Robbertse B, Slippers B, 2017. *Botryosphaeria dothidea*: a latent pathogen of global importance to woody plant health. *Molecular Plant Pathology* **18**:477-488.

Martin SH, Steenkamp ET, Wingfield MJ, Wingfield BD, 2013. Mate-recognition and species boundaries in the ascomycetes. *Fungal Diversity* **58**:1–12.

McDonald BA, Linde C, 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annual Reviews of Phytopathology* **40**:349–79.

Ni M, Feretzaki M, Sun S, Wang X, Heitman J, 2011. Sex in fungi. *Annual Review of Genetics* **45**:405-430.

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

Pöggeler S, O'Gorman CM, Hoff B, Kück U, 2011. Molecular organization of the mating-type loci in the homothallic Ascomycete *Eupenicillium crustaceum*. *Fungal Biology* **115**:615-624.

Putman AI, Tredway LP, Carbone I, 2015. Characterization and distribution of mating-type genes of the turfgrass pathogen *Sclerotinia homoeocarpa* on a global scale. *Fungal Genetics and Biology* **81**:25-40.

Slippers B, Boissin E, Phillips AJL, Groenewald JZ, Lombard L, Wingfield MJ, Postma A, Burgess T, Crous PW, 2013. Phylogenetic lineages in the *Botryosphaeriales*: a systematic and evolutionary framework. *Studies in Mycology* **76**:31–49.

Slippers B, Crous PW, Jami F, Groenewald JZ, Wingfield MJ, 2017. Diversity in the *Botryosphaeriales*: Looking back, looking forward. *Fungal Biology* **121**:307-321.

Solovyev V, Kosarev P, Seledsov I, Vorobyev D, 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biology* **7**:1-12.

Steenkamp ET, Wingfield BD, Coutinho TA, Zeller KA, Wingfield MJ, Marasas WFO, Leslie JF, 2000. PCR-Based Identification of *MAT-1* and *MAT-2* in the *Gibberella fujikuroi* Species Complex. *Applied and Environmental Microbiology* **66**:4378-4382.

Sun S, Heitman J, 2015. From two to one: Unipolar sexual reproduction. *Fungal Biology Reviews* **29**:118-125.

Tsui CK-M, DiGuistini S, Wang Y, Feau N, Dhillon B, Bohlmann J, Hamelin RC, 2013. Unequal Recombination and Evolution of the Mating-Type (MAT) Loci in the Pathogenic Fungus *Grosmannia clavigera* and Relatives. *G3* **3**:465-480.

Ueng PP, Dai Q, Cui K-R, Czembor PC, Cunfer BM, Tsang H, Arseniuk E, Bergstrom GC, 2003. Sequence diversity of mating-type genes in *Phaeosphaeria avenaria*. *Current Genetics* **43**:121-130.

Vaghefi N, Ades PK, Hay FS, Pethybridge SJ, Ford R, Taylor PWJ, 2015. Identification of the MAT1 locus in *Stagonosporopsis tanaceti*, and exploring its potential for sexual reproduction in Australian pyrethrum fields. *Fungal Biology* **119**:408-419.

van der Nest MA, Bihon W, De Vos L, Naidoo K, Roodt D, Rubagotti E, Slippers B, Steenkamp ET, Wilken PM, Wilson A, Wingfield MJ, Wingfield BD, 2014. IMA Genome-F 2: *Ceratocystis manginecans*, *Ceratocystis moniliformis*, *Diplodia sapinea*: Draft genome sequences of *Diplodia sapinea*, *Ceratocystis manginecans*, and *Ceratocystis moniliformis*. *IMA Fungus* **5**:135-140.

Wang N-Y, Zhang K, Huguet-Tapia JC, Rollins JA, Dewdney MM, 2016. Mating type and simple sequence repeat markers indicate a clonal population of *Phyllosticta citricarpa* in Florida. *Phytopathology* **106**:1300-1310.

Wey T, Schlegel M, Stroheker S, Gross A, 2016. MAT – gene structure and mating behavior of Hymenoscyphus fraxineus and Hymenoscyphus albidus. Fungal Genetics and Biology 87:54-63.

Wilken PM, Steenkamp ET, Wingfield MJ, De Beer ZW, Wingfield BD, 2017. Which MAT gene? Pezizomycotina (Ascomycota) mating-type gene nomenclature reconsidered. *Fungal Biology Reviews*, *In press*. Doi: 10.1016/j.fbr.2017.05.003

Wingfield BD, Ades PK, Al-Naemi FA, Beirn LA, Bihon W, Crouch JA, de Beer ZW, De Vos L, Duong TA, Fields CJ, Fourie G, Kanzi AM, Malapi-Wight M, Pethybridge SJ, Radwan O, Rendon G, Slippers B, Santana QC, Steenkamp ET, Taylor PW, Vaghefi N, van der Merwe NA, Veltri D, Wingfield MJ, 2015. IMA Genome-F 4: Draft genome sequences of *Chrysoporthe austroafricana*, *Diplodia scrobiculata*, *Fusarium*

nygamai, *Leptographium lundbergii*, *Limonomyces culmigenus*, *Stagonosporopsis tanaceti*, and *Thielaviopsis punctulata*. *IMA Fungus* **6**:233-248.

Yang T, Groenewald JZ, Ratchadawan C, Fahimeh J, Abdollahzadeh J, Lombard L, Crous PW, 2017. Families, genera, and species of Botryosphaeriales. *Fungal Biology* 121: 322–346.

CHAPTER 6

Concluding Remarks

Final considerations

The ecological and economic relevance of plant pathogenic fungi belonging to the order *Botryosphaeriales* triggered this study. Particularly in the family *Botryosphaeriaceae*, the main concerns about this group of fungi are their potential to cause diseases in a wide range of hosts and geographical areas and consequently the adoption of measures that could prevent or minimize the damages. However, to do that, basic knowledge such as which fungi we are dealing with is fundamental but not always easy. To improve the knowledge about this fungal group we proposed to explore its diversity in Portugal, to develop new tools to delimit cryptic species and to unveil sexual reproductive strategies.

Considering that *Neofusicoccum* species are widespread and have the capacity to cause disease, several important forest, crop and ornamental plants were sampled in order to gain further knowledge about the diversity of these fungi in Portugal. Despite the low diversity of species found, the number of new host associations was high, attesting their ability for host jumping. Host jumps are usually associated to anthropogenic actions that inadvertently spread fungi all over the world allowing the contact with unrelated hosts that have not coevolved defences against the fungi, which often causes disease epidemics (Slippers et al. 2005). Also, host jumps have been pointed as the main force in speciation events of several fungi meaning that they are evolving in order to improve their capacity of infection (Van der Merwe et al. 2008; Kepler et al. 2012; McTaggart et al. 2016). For these reasons, host jumping is an issue that should not be minimized but considered as enough reason to implement quarantine measures to prevent the dissemination in new areas (Slippers et al. 2005).

Beyond the high capacity to colonize new hosts several genera of the family *Botryosphaeriaceae* are rich in cryptic species resulting in misidentifications. In this study we showed that the use of the GCPRS criterion is fundamental to establish solid species boundaries in the genera *Neofusicoccum* and *Diplodia*. Apart from the use of more conventional molecular markers, *MAT* genes were also used. Similar to the good results obtained in several previous reports (Barve et al. 2003; Du et al. 2005; Yokoyama et al. 2006; Rau et al. 2007) here *MAT* genes also proved to be a powerful molecular marker to delimit species with levels of informativeness higher than the other regions used. However, despite the low discriminatory power of ITS, this region has been accepted as the DNA barcode for

fungi. The availability of universal primers and the ease of amplification and sequencing are enough reasons to provide the continuity of its use but more important than that is the large amount of sequences available in databases. As good as one molecular marker may be it will be useless without sequences to compare with in the databases. As noticed during this study, for the family *Botryosphaeriaceae* the regions ITS and *tef1-α* have the largest number of sequences deposited while for the regions *tub2*, *act2* or *his3*, for example, the number decreases sharply. In the case of *MAT* genes, despite the ease of amplification and sequencing associated with an excellent species resolution power, the lack of deposited sequences unable the regular use of this genes as molecular markers for a while. However, with the knowledge gathered in studies like this the scenario may become different soon.

Several methods have been used to identify *MAT* genes. Among them the amplification of the conserved DNA-binding domains using degenerate primers, followed by chromosome walking using Thermal Asymmetric Interlaced (TAIL) PCR (Arie et al. 1997), inverse PCR (Kerenyi et al. 2004), commercial kits (Conde-Ferraez et al. 2007) or a combination of these methods (Barve et al. 2003; Phan et al. 2003) are the most common. Such methods, although have been proven useful in isolation of *MAT* genes in multiple species, are time-consuming and laborious, and cannot be applied to taxa for which *MAT* sequences of closely related species are not available (Bihon et al. 2014). On the other hand, partial genome assembly for identifying genes of interest, is a relatively simple technique, and also provides a valuable source of additional data for future research. Due to their small genomes compared to other eukaryotes, fungi are very well suited for de novo assembly of the whole genome to address questions regarding their biology and evolution (Vaghefi et al. 2015). The obvious advantages of this tool are the reason for the increased number of genomes deposited in the databases. Among *Botryosphaeriaceae* species the number of available genomes increased from two to nine in a very short time period. In this work we exploited the *MAT* loci and flanking regions of the genomes of species belonging to the order *Botryosphaeriales* which allowed us to identify the content and organization of each locus and point out the major differences and similarities. Also the sexual strategies adopted were readily detected.

In conclusion, the objectives proposed to this work were accomplished. Now, we have more information about the diversity of the genus *Neofusicoccum* in Portugal and a new

powerful tool to delimit their species in a quick and reliable way. The usefulness of *MAT* genes in resolving species boundaries was proved not only for the genus *Neofusicoccum* but also for the genus *Diplodia*. The PCR-based assays developed here will be useful to access *MAT* genes for this or other purposes. Also, the similarities and differences pointed out in the analysis of different *MAT* loci among species in the *Botryosphaeriales* will be useful for further analyses. Although we have focused in the potential of *MAT* genes in resolving cryptic species overall the results obtained in this work will be valuable in other studies such as the evolution of *MAT* genes and mating strategies, population structure, mating tests, gene expression and functions.

Future perspectives

In science the questions and the search for new knowledge never end and it was felt during this work. Despite the important results obtained here we immediately realized that it is just the beginning of a long road. Regarding to the diversity of *Neofusicoccum* species and *Botryosphaeriaceae* in general in Portugal, several new studies could be done. In this work the sampling was made in several hosts but in random geographic areas. In the future, studies may be carried out with new samplings taking in account, for instance, all the possible hosts in each geographic/climatic Portuguese region. Also, it could be interesting to perform a survey within the preserved national forests, which harbours in some cases species not seen anywhere else.

If questions emerged regarding biodiversity, the same happened in the next subject. We were aware that working with *MAT* genes in taxa with almost no previous knowledge would be a hard challenge, as it was indeed. However, the excellent quality of *MAT* genes as molecular markers proved here was worth all the effort and made us advance our knowledge on this group of fungi. Thus, in the future we would like to evaluate the functionality of the PCR-based assays here developed in all *Neofusicoccum* and *Diplodia* species in order to complete the information needed to make full phylogenetic comparisons and also to conclude about the mating strategies present in each genus and their evolution. This kind of study can also be extended to other genera of the *Botryosphaeriaceae* allowing the use of *MAT* genes as molecular markers in the whole family.

Another interesting quest will be to explain the difference found in the *MAT* loci and flanking regions of *N. parvum*. In order to understand if this structure is typical from the genus *Neofusicoccum* sequencing genomes of other species will be crucial.

References

Arie T, Chrisiansen SK, Yoder OC, Turgeon BG, 1997. Efficient cloning of Ascomycete mating type genes by PCR amplification of the conserved MAT HMG box. *Fungal Genetics and Biology* **21**:118-130.

Barve MP, Arie T, Salimath SS, Muehlbauer FJ, Peever TL, 2003. Cloning and characterization of the mating type (*MAT*) locus from *Ascochyta rabiei* (teleomorph: *Didymella rabiei*) and a *MAT* phylogeny of legume-associated *Ascochyta* spp. *Fungal Genetics and Biology* **39**:151-167.

Bihon W, Wingfield MJ, Slippers B, Duong TA, Wingfield BD, 2014. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. *Fungal Genetics and Biology* **62**:55-61.

Du M, Schardl CL, Nuckles EM, Vaillancourt LJ, 2005. Using mating- type gene sequences for improved phylogenetic resolution of *Colletotrichum* species complexes. *Mycologia* **97**:641-658.

Kepler RM, Sung G-H, Harada Y, Tanaka K, Tanaka E, Hosoya T, Bischoff JF, Spatafora JW, 2012. Host jumping onto close relatives and across kingdoms by *Tyrannicordyceps* (Clavicipitaceae) gen. nov. and *Ustilaginoidea* (Clavicipitaceae). *American Journal of Botany* **99**:552-561.

Kerényi Z, Antonio M, Waalwijk C, Oláh B, Hornok L, 2004. Mating type sequences in asexually reproducing *Fusarium* species. *Applied and Environmental Microbiology* **70**(8): 4419-4423.

McTaggart AR, Shivas RG, van der Nest MA, Roux J, Wingfield B, Wingfield MJ, 2016. Host jumps shaped the diversity of extant rust fungi (Pucciniales). *New Phytologist* **209**:1149-1158.

Phan HTT, Ford R, Taylor PWJ, 2003. Mapping the mating type locus of *Ascochyta rabiei*, the causal agent of ascochyta blight of chickpea. *Molecular Plant Pathology* **4**:373-381.

Rau D, Attene G, Brown AHD, Nanni L, Maier FJ, Balmas V, Saba E, Schäfer W, Papa R. 2007. Phylogeny and evolution of mating-type genes from *Pyrenophora teres*, the causal agent of barley ‘‘net blotch’’disease. *Current Genetics* **51**:377-392.

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

Vaghefi N, Ades PK, Hay FS, Pethybridge SJ, Ford R, Taylor PW, 2015. Identification of the *MAT1* locus in *Stagonosporopsis tanaceti*, and exploring its potential for sexual reproduction in Australian pyrethrum fields. *Fungal Biology* **119**:408-419.

Van der Merwe MM, Walker J, Ericson L, Burdon JJ, 2008. Coevolution with higher taxonomic host groups within the *Puccinia*/*Uromyces* rust lineage obscured by host jumps. *Mycological Research* **112**:1387–1408.

Yokoyama E, Arakawa M, Yamagishi K, Hara A, 2006. Phylogenetic and structural analyses of the mating-type loci in *Clavicipitaceae*. *FEMS Microbiology Letters* **264**:182-191.