MAT gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen

Wubetu Bihon[1,2*], Michael J. Wingfield[1], Bernard Slippers[1], Tuan A. Duong[1], Brenda D. Wingfield[1]

Department of Genetics, Forestry and Agricultural Biotechnology Institute, University of Pretoria, Private Bag X20, Hatfield, Pretoria 0028, South Africa[1];

Crop Protection Division, Vegetable and Ornamental Plant Institute, Agricultural Research Council, Private Bag X293, Pretoria 0001, South Africa[2]

[*]Corresponding author LegesseW@arc.agric.za

Phone: +27 795 481 639; Fax: +27 124 203 960

ABSTRACT

Diplodia pinea (= *Sphaeropsis sapinea*) is a well-known and economically important latent pathogen of *Pinus* spp. in many parts of the world. Despite intensive scrutiny, its sexual state has never been observed and the fungus has thus been considered exclusively asexual. It was, therefore, surprising that a recent population genetics study showed high genotypic diversity and random association of alleles in a number of populations, suggesting that the pathogen has a cryptic sexual stage. Using the genome sequence of two individual *D. pinea* isolates, we interrogated the structure of the *MAT* locus in this fungus. The results suggested that *D. pinea* is heterothallic (self-sterile) with complete and apparently functional copies of the *MAT* genes containing the α-1 and HMG domains present in different isolates. In addition to the *MAT1-2-1* and *MAT1-1-1* genes, we found a *MAT1-1-4* gene in the *MAT1-1* idiomorph and a novel *MAT1-2-5* gene in the *MAT1-2* idiomorph. Importantly, the frequencies of occurrence of both idiomorphs in populations examined were not significantly different from a 1:1 ratio, which would be expected in sexually reproducing populations. Although the sexual state has never been observed, the results strongly suggest that *D. pinea* has a cryptic, heterothallic sexual cycle.

Keywords: *Diplodia pinea*, endophyte, heterothallic, mating type, idiomorph, cryptic sex

1. Introduction

Diplodia pinea is an Ascomycete fungus known from many countries of the world where it is a latent pathogen that predominantly infects *Pinus* species (Swart et al., 1985; Palmer et al., 1987; Stanosz and Smith, 1996; Wingfield and Knox-Davies, 1980). It is an opportunistic pathogen that can result in significant economic losses in nurseries, plantations, natural forests and shelterbelts (Swart and Wingfield, 1991, Blodgett and Stanosz, 1997). The most significant impact has been in regions of the southern hemisphere where *Pinus* spp. are grown as non-natives in plantations (Swart et al., 1985; Zwolinski et al., 1990; Wingfield et al., 2001).

Pinus spp. infected with endophytic *D. pinea* display disease symptoms only when they are subjected to physiological or environmental stresses (Swart and Wingfield, 1991; Blodgett and Stanosz, 1999; Flowers et al., 2001). Asexual fruiting bodies of the fungus are commonly encountered on dead tissue and the conidia are easily transmitted by wind, rain splash and humans to new areas (Bihon et al., 2011; Swart and Wingfield, 1991). The asymptomatic persistence of *D. pinea* in the healthy tissues of host trees makes it particularly difficult to formulate or impose measures to control movement of the fungus to new environments.

A sexual state has never been observed in *D. pinea*. This is despite intensive studies on this pathogen since its discovery in France in 1842. The assumption has been that the fungus reproduces only asexually (Swart and Wingfield, 1991; Burgess et al., 2004; Bihon et al., 2011). However, Palmer et al. (1987) and Wingfield and Knox-Davies (1980) observed spermatial cells in *D. pinea*, which are often thought to be involved in the fertilization process and thus potentially in a cryptic sexual stage.

Although it does not provide unequivocal evidence for the presence of a sexual state, indirect tests can provide strong evidence for the existence of a sexual cycle in fungi for which sexual structures have not been encountered (Turgeon et al., 1993; Arie et al., 2000; Groenewald et al., 2006; Duong et al., 2013). Populations that undergo sexual recombination would typically have higher levels of genotypic diversity compared to those that reproduce asexually (Milgroom, 1996; McDonald and Linde, 2002). Furthermore, the alleles at unlinked loci in sexually

reproducing populations are expected to be recombining randomly as indicated by linkage disequilibrium, unlike the situation in exclusively asexual fungi where alleles are not reassorted (Smith et al., 1993). Population genetic studies using microsatellite markers have consistently shown these patterns expected for sexually reproducing species in *D. pinea*, namely high genotypic diversity in populations and linkage disequilibrium between alleles (Burgess et al., 2004; Bihon et al., 2012a, 2012b). Together with the observations of spermatia discussed above, these results demand a deeper investigation into the reproductive biology of *D. pinea*. Understanding the potential of this pathogen to undergo sexual recombination is important because it would provide an understanding of its evolutionary potential to adapt to different environments (McDonald and Linde, 2002) and thus also influence management strategies to reduce its impact.

Apart from population genetic studies using neutral markers, the study of mating type genes can also provide indirect evidence of sexual recombination. In Ascomycetes, sexual reproduction between individuals of the same species is governed by the genes found at a single mating type locus (*MAT-1*), which contains two core gene families (*MAT1-1* and *MAT1-2*) (Turgeon, 1998; Turgeon and Yoder, 2000; Lee et al., 2010; Debuchy et al., 2010; Ni et al., 2011). In heterothallic species, individuals contain only one of the two gene families at the *MAT-1* locus (idiomorphs), and require an individual of the opposite *MAT-1* idiomorph for sexual reproduction (Kronstad and Staben, 1997; Groenewald et al., 2008; Lee et al., 2010; Martin et al., 2011). The *MAT1-1* idiomorph contains at least a *MAT1-1-1* gene that encodes a protein with an α-1 box domain. The *MAT1-2* idiomorph contains at least a *MAT1-2-1* gene that encodes a regulatory protein with a DNA-binding domain known as high mobility group (HMG box). In contrast, homothallic fungus species are self-compatible and individuals have both the *MAT1-1* and the *MAT1-2* genes on the same idiomorph. Unfortunately the mating type loci of *D. pinea* have not been characterized, precluding their use in the study of population biology.

The existing evidence for the presence of a cryptic sexual cycle in *D. pinea* prompted the present study that sought to identify the *MAT* locus in this important tree pathogen. This was made possible by next generation sequencing technologies that enabled the sequencing of the entire *D. pinea* genome. In this study, we thus sequenced the genomes of two *D. pinea* isolates and

identified the *MAT* gene sequence using bioinformatics tools. The emerging data were then used to design specific primers to identify the mating types of isolates in populations of *D. pinea*.

2. Materials and methods

2.1. DNA isolation

Mycelium from a single conidial culture of *D. pinea* isolate CMW190 was grown on 2% malt extract agar (MEA) (Biolab, Merck). Three-day-old mycelium from a culture on MEA was then used to inoculate 2% malt extract broth containing 0.2% yeast, and incubated at 25 °C on a shaker. Mycelium grown in the broth was collected and transferred to 2 ml Eppendorf tubes. The mycelium was separated from the solution by centrifugation and then freeze dried. DNA was extracted using the salt extraction method of Aljanabi and Martinez (1997) and preserved at -20 °C.

2.2. Genome sequencing and MAT locus identification

The quality and quantity of the DNA extracted from *D. pinea* isolate CMW190 was measured using a NanoDrop spectrophotometer (Thermo Scientific) and on a 0.8% agarose gel after electrophoresis. More than 2 µg of gDNA was submitted to the Agricultural Research Council (ARC, South Africa) Sequencing Platform and paired end sequences were produced using Illumina HiSeq 2000 (Illumina, Inc.). The reads were assembled into contigs using the CLCbio Genomics Workbench 5.5 (CLCbio, Aarhus, Denmark) De novo assembly program for next generation sequences.

Analyses using the BLASTn algorithm on the CLCbio Genomics Workbench were conducted using the *Mycosphaerella graminicola MAT* idiomorphs sequences (Waalwijk et al. 2002) as a reference. Contig containing sequences having high levels of similarity were used in a BLAST analysis of the NCBI GenBank database with the BLASTn and BLASTx algorithms. *Diplodia pinea* genome sequences containing a putative *MAT* gene were also analyzed for the presence of other genes that have been previously associated with *MAT* idiomorphs, specifically DNA lyase/APN and *SLA* (Turgeon, 1998; Waalwijk et al., 2002).

2.3. MAT1-1-1 specific primer design and PCR amplification

A putative *MAT1-1-1* gene was identified from *D. pinea* isolate CMW190. Specific primers to amplify part of the *MAT1-1-1* locus were designed using Primer Premier v5 (PREMIER Biosoft International). Primer pairs were synthesized (Inqaba Biotech, Pretoria, South Africa) and tested on isolate CMW190 and other *D. pinea* isolates. The PCR reaction included: 0.5 µl DNA (30 ng/µl), 2.5 µl 10x PCR reaction buffer with $MgCl_2$, 200 µM of each dNTPs, 3.5 mM $MgCl_2$, 0.5 µM each primers, 1 U Faststart *Taq* polymerase (Roche) and sterilized distilled water to a total volume of 25 µl. The PCR conditions included a one cycle of 95 °C for 4 min, and 35 cycles of 95 °C for 30 seconds (s), 60 °C for 30 s, 72 °C for 30 s and one cycle of final extension at 72 °C for 7 min. Amplicons were sequenced using the same primers designed for PCR.

2.4. Identification of the MAT1-2-1 gene

The *MAT1-1-1* gene could not be amplified for all test isolates and we hypothesized that these isolates contained a *MAT1-2* idiomorph. One of these isolates (CMW39103) was then chosen for further characterization. DNA was extracted as described above and the genome of this isolate was sequenced using MiSeq, Illumina sequencer (Inqaba Biotech, Pretoria, South Africa). The single end reads were assembled and contigs constructed using CLCbio Genomics Workbench. A local BLASTn search targeting the *MAT1-2-1* gene using *M. graminicola MAT1-2-1* gene sequence as a reference was conducted on CLCbio Genomics Workbench. Contigs with sequences having high similarity to the *M. graminicola MAT1-2-1* sequence were analysed using FGENESH+ (http://linux1.softberry.com) and AUGUSTUS (http://augustus.gobics.de) gene prediction programs. Each of the ORFs determined in this analysis were subjected to a BLAST analysis in GenBank. Once the putative *MAT1-2-1* gene had been identified, a pair of primers was designed to amplify a 385 bp part of the *MAT1-2-1* gene. Designing primers, PCR and sequencing was conducted as described above for *MAT1-1* idiomorph.

2.5. Identification of genes flanking the MAT1-1-1 *and* MAT1-2-1 *genes*

Flanking regions of both of the idiomorphs were identified by aligning contigs containing the two *MAT* idiomorphs using an online version of MAFFT multiple sequence alignment for amino acid and nucleic acids (Katoh and Toh, 2008). The DNA sequence downstream of the *MAT1-1-1* gene was short to determine the flanking sequence. Primers were thus designed inside the *MAT1-1-1* gene and from the flanking region of the *MAT1-2-1* locus. The PCR amplicon generated using this primer pair was sequenced using Sanger sequencing (Sanger et al., 1977). This sequence was then used in an analysis with the DNA sequence of the *MAT1-2* and *MAT1-1* gene regions to extend the *MAT1-1* idiomorph sequence into the flanking region.

2.6. Mating type ratios and sequence similarity in D. pinea populations

Populations of *D. pinea* from Argentina, Australia, Ethiopia, United States, and South Africa (Table 1) were used to determine the occurrence and frequency of their mating type idiomorphs. These were the same populations of isolates that have been used in previous population genetic studies (Bihon et al. 2012a, 2012b). PCR was conducted using the two idiomorph primer sets described above and amplification verified using gel-electrophoresis. A chi-square test was used to compare the *MAT1-1/MAT1-2* ratio to a 1:1 ratio at a $P = 0.05$ significance level.

In order to evaluate sequence identity and similarity within and between populations, 96 isolates selected from all populations and representing 48 *MAT1-1-1* and 48 *MAT1-2-1* idiomorphs were sequenced.

2.7. Protein alignment and phylogenetic analysis

Predicted amino acid sequences of conserved portions of the *MAT1-1-1* (α-1 domain) and *MAT1-2-1* (HMGbox) genes of *D. pinea* were aligned with corresponding amino acid sequences from other Ascomycete *MAT* genes obtained from GenBank using the MAFFT multiple sequence alignment for amino acid and nucleic acids (Katoh and Toh, 2008). Shading of similar amino acid sequences (Figs. 2, 4) was applied using GeneDoc 2.7 (http://www.nrbsc.org).

Neighbor joining (NJ) phylogenetic trees were generated following the approach of Arie et al. (2000) with 1000 bootstrap replicates using MEGA 5 (Tamura et al., 2011). To confirm the robustness of NJ analyses, a maximum parsimony analysis was also performed in PAUP 4.0 (Swofford, 2003), using a heuristic search and bootstrapping with 1000 replications. The protein sequences of the *MAT* genes of *Fusarium circinatum* (Wingfield et al. 2012) were used as an out-group.

2.8. Pairing of mating types in culture

Ten single conidial isolates of *D. pinea*, five each of which had been determined to be either *MAT1-1* or *MAT1-2* were crossed with each other in every possible combination in culture. Crosses were made in Petri dishes containing 1.5% water agar (w/v) (Biolab, Merck) and to which double sterilized pine needles had been added to the surface of the agar. The paired isolates were incubated at 25 °C for two weeks and until substantial mycelial growth had occurred. The plates were then transferred to a 4 °C incubator and they were monitored regularly for the development of sexual structures over a period of four months.

3. Results

3.1. Genome sequence, assembly and identification of MAT *loci*

Isolates CMW190 and CMW39103 of *D. pinea* each yielded a total of about 19 million paired and single reads. These reads were de-novo assembled using CLCbio Genomics Workbench 5.5 after trimming and removing reads shorter than 30 bp. After assembly, 2194 contigs with N50 equals 37.7 kb were obtained from isolate CMW190 and 7853 contigs with N50 of 16.2 kb obtained from isolate CMW39103. The data from the two assemblies were then merged and re-assembled that produced a combined assembly of 3843 contigs. The estimated genome size from this combined assembly of *D. pinea* was 37 Mb. This Whole Genome Shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession AXCF00000000. The version described in this paper is version AXCF01000000.

A contig of 17.8 kb containing a ~300 bp fragment of a gene similar to *M. graminicola MAT1-1-1* gene was identified using BLASTn in the assembled genome sequence of isolate CMW190. No genes with similarity to the *MAT1-2-1* gene were observed in the genome of this isolate. A contig, 33 kb in length with a putative *MAT1-2-1* gene encoding a protein containing an HMG box was identified in the genome of isolate CMW39103.

3.2. MAT1-1-1 *and* MAT1-2-1 *specific primers*

Primers (DipM1f: 5'-CAA GCC ATC GAC CGA AAC and DipM1r: 5'- GAA GAA GCG CAC CCT CAA T) were designed from the *MAT1-1-1* gene sequence from isolate CMW190. These primers were used to successfully amplify a fragment of about 428 bp of the target locus in a number of *D. pinea* isolates. Primer pairs (DipHMGf: 5'- ACA AAG TTC AGC GGA GCG and DipHMGr: 5'- CCT CCG CAG GTC ACT CAT) were designed for the identified *MAT1-2-1* gene from isolate CMW39103. Sequence and BLASTx analysis of the putative *MAT1-1-1* and *MAT1-2-1* gene sequences in the GenBank further confirmed the gene identities.

3.3. Structure of MAT *idiomorphs*

The total length of mRNA predicted using the FGENESH+ gene prediction program of *MAT1-1-1* gene was 1335 bp (444 aa) and that of *MAT1-2-1* was 1170 bp (389 aa) (Fig. 1). The MAT1-1 (accession no. KF551229) and MAT1-2 (accession no. KF551228) including other ORFs (Fig. 1) sequences have been deposited to GenBank nucleotide sequence databases. The predicted 1.3 kb *MAT1-1-1* open reading frames (ORF) had a single intron of 48 bp and it encoded a protein of 444 amino acids. The 3' end of the *MAT1-1* idiomorph also contained part of the *MAT1-2-5* gene described below. The *MAT1-2-1* gene was found to encode 398 amino acids and it contained a single 49 bp intron. As is the case with most Ascomycete *MAT1-2-1* genes, this intron was found to split a codon that codes for the amino acid Serine.

The 5' flanking region of both idiomorphs contained a DNA lyase gene. The DNA lyase gene had two introns of 47 bp and 86 bp, respectively. Five additional predicted ORFs were identified in both idiomorphs between the DNA lyase and the *MAT1-1-1* and *MAT1-2-1* genes (Fig. 1). In

the *MAT1-1* idiomorph, an additional ORF was identified that showed 25% and 27% amino acids sequence similarity to *MAT1-1-4* sequences of GenBank accession numbers XP003023854 and XP003014758, respectively. Similarly, a unique additional ORF was identified in the *MAT1-2* idiomorph. The gene observed in this idiomorph of *D. pinea* did not display similarity to the genes previously observed and is thus designated *MAT1-2-5*. This reflects the fact that it is contained in the *MAT1-2* idiomorph and is the fifth such gene to be identified in this region. The 5' end of the *MAT1-2* idiomorph also contains a portion of the *MAT1-1-4* gene.

3.4. Ratio of mating types and MAT *gene sequence similarity*

Of the 189 *D. pinea* isolates (Table 1) tested for mating type, the frequency of the *MAT1-1-1* to *MAT1-2-1* mating types ranged from 0.3 in Ethiopian to 1.6 in Argentinean *D. pinea* populations. The chi-square analysis in all of the populations indicated that the ratio of *MAT1-1-1* to *MAT1-2-1* was not significantly different from 1:1 except for the Ethiopian population (Table 1). The Australian population had a fairly low frequency of 0.53 but the difference was not significant at 5% level of probability.

Sequencing *MAT* amplicons (416 bp in *MAT1-1-1* and 390 bp in *MAT1-2-1*) from representative isolates showed (data not presented) no sequence differences in either the *MAT1-1-1* or *MAT1-2-1* idiomorphs.

3.5. Protein sequence and phylogenetic analysis

Amino acid sequences of the α-1 domain and HMG box of the *MAT1-1-1* and *MAT1-2-1* genes of *D. pinea* were most similar to other Dothideomycetes, even though that they had a very low similarity with all other species in this group (Fig 2, 3). There were 11 amino acids in the conserved α-1 domain, which were also found in all species used in the comparison and 15 amino acids were shared with nine or ten taxa (Fig. 2). In the aligned conserved HMG box region, only 11 of 100 amino acids were identical to all species used in the comparison (Fig. 3). Within the region designated in the HMG box, there was a maximum of 60% similarity across the compared species.

Phylogenetic trees for the α-1 domain and HMG box conserved regions generated from NJ and MP analyses had similar topologies. In the tree obtained from α-1 domain sequence, *D. pinea* grouped most closely with *C. heterotrophus*, *A. alternata* and *P. nodorum*. In the tree obtained from the HMG box sequences, *D. pinea* grouped most closely to *M. graminicola*, *D. pini* and *C. zeina* (Fig. 4).

3.6. Pairing of mating types in culture

None of the pairings between isolates representing *MAT1-1-1* and *MAT1-2-1* mating types gave rise to sexual structures. They were, however, able to sporulate and all formed the asexual pycnidia with abundantly produced conidia. Spermatia were not observed in this study.

4. Discussion

This study provides evidence that *D. pinea* does not exclusively reproduce asexually as has previously been suggested but rather, most likely has a cryptic, heterothallic sexual cycle. This is evident from the identification of *MAT* idiomorphs containing *MAT1-1* and *MAT1-2* related genes. The study also showed that *D. pinea* has a novel *MAT* gene (*MAT1-2-5*). These idiomorphs occurred in ratios that are not significantly different from the expected 1:1 ratio, which is found in a randomly recombining species in at least four populations collected from various parts of the world. Sexual fruiting structures were, however, not observed, despite attempts to induce them using isolates of different mating types.

This study represents the first description of *MAT* gene sequences for any of the fungi in the Botryosphaeriales (Slippers and Wingfield, 2007) and will make a significant contribution to studying these genes in other members of the Order. There have been previous attempts to identify sexual structures in *D. pinea* but these have never been found (M. Wingfield, unpublished). Degenerate primers for both of the idiomorphs in other Ascomycetes (Groenewald et al., 2006, Arie et al., 1997), as well as our own attempts to design primers based on other Dothideomycetes *MAT* sequences failed to amplify any of the *MAT* genes in *D. pinea* (authors

unpublished). The high level of sequence dissimilarity of *MAT* genes in *D. pinea* with other known species in the Dothideomycetes revealed in this study now explains those failures.

The presence of *MAT1-1* or *MAT1-2* in different isolates is similar to that found in other sexually reproducing, heterothallic or self-incompatible Dothideomycetes and other Ascomycetes (Coppin et al., 1997; Paoletti et al., 2005). However, the organisation differed from other self-incompatible Dothideomycetes in which the idiomorphs consisted of a single gene encoding a *MAT1-1-1* or *MAT1-2-1* gene (containing α-1 domain or HMG box respectively) in close proximity to the DNA-lyase gene (Debuchy et al., 2010). In *D. pinea,* there are genes between the *MAT* locus and the DNA lyase gene, with approximately 13 kb between them. These genes include COX-VIa and APC5, which are rarely seen associated with this locus in other fungi (Debuchy and Turgeon, 2006; Tsui et al., 2013). While unusual, this organization is not unprecedented in Ascomycetes. For example, in *Grosmannia clavigera* the DNA-lyase gene was also found 15 kb from the *MAT* locus (Tsui et al., 2013). In *G. clavigera,* there were also five putative protein coding genes between *MAT* and DNA-lyase locus of unknown function, and including COX13. Even though these results support the hypothesis that *D. pinea* is heterothallic, they do not unequivocally prove sexuality, and neither do they preclude other forms of sexual reproduction in this fungus (e.g. Ni et al., 2011).

In addition to the *MAT1-1-1* and *MAT1-2-1* genes, we identified a *MAT1-1-4* gene in the *MAT1-1* idiomorph and a previously undescribed *MAT1-2-5* gene in the *MAT1-2* idiomorph. The *MAT1-1-4* gene has been previously characterized in *Trichophyton verrucosum* and *Anthroderma benhamiae* (http://www.ncbi.nlm.nih.gov). The novel ORF found alongside the *MAT1-2-1* gene in the *MAT1-2* idiomorph was designated as the *MAT1-2-5* gene being the fifth such gene described associated with the *MAT1-2* idiomorph. The function of these additional *MAT* genes in the *MAT* idiomorphs, which has also been observed in other fungal species (Kanamori et al., 2007; Martin et al., 2011; Chitrampalam et al., 2013), remain unknown.

Among 96 *D. pinea* isolates, half representing each of the *MAT1-1-1* and *MAT1-2-1* mating types, the sequences for representatives of both the genes were identical. The absence of within-

species variation of the *MAT* genes is consistent with studies on other fungi (Turgeon, 1998; Bennett et al., 2003). For example, in *C. heterostrophus* only three and four nucleotide differences were observed among 1405 bp of *MAT1-1* and 1280 bp of *MAT1-2* sequences (Turgeon, 1998). The sequenced region will thus not be informative for studying population diversity and structure. It might, however, be useful for studying cryptic species diversity in more closely related fungi such as *D. scrobiculata* (De Wet et al., 2003).

Amino acid sequences of both of the α-1 domain and HMG box of *D. pinea* aligned with sequences of other Dothidiomycetes species, but were not highly conserved. This was supported by the phylogenetic analysis of the conserved α-1 domain and HMG box regions that separated *D. pinea* from other genera in the class. The two trees were also not congruent, which is unlike other nuclear genes but has previously also been found for mating type genes (Goodwin et al., 2003; Yokoyama et al., 2006; Duong et al., 2013) and possibly relates to their different selection pressures, evolutionary rates and/or origins. Within the most conserved part of the HMG box in the *D. pinea MAT1-2-1*, gene the amino acid identity was 49% when compared with *C. zeina* and 48% in comparison to *D. pini*. The same was true in the α-1 domain of the *MAT1-1-1* gene where there was a 47% identity with *Phialocephala fortinii*. This lack of similarity highlights the phylogenetic distances within this group, which have also been shown using other gene regions (Schoch et al., 2006).

Amplification of mating type loci in *D. pinea* isolates collected in four different countries on four continents showed that the frequency of the *MAT1-1-1* to *MAT1-2-1* does not differ from the 1:1 ratio. This is expected in randomly mating populations (Milgroom, 1996; Zhan et al., 2002; Paoletti et al., 2005; Groenewald et al., 2006). The results were unusually consistent given that all of the populations studied were for countries where the pathogen has been introduced. They might thus have been expected to be slanted towards one or other mating type if it was reproducing asexually, and given the bottleneck that usually accompanies introduced populations (Taylor et al., 1999). The only exception was found in the Ethiopian population and this result could have been influenced by sampling.

Previous reports of spermatia (Wingfield and Knox-Davies, 1980; Palmer et al., 1987), as well as the random association of loci and high genetic diversity in *D. pinea* suggested the presence of an active sexual state in *D. pinea* (Bihon et al. 2012a, 2012b). The observation of an intact *MAT* locus and the even distribution of mating genes (Coppin et al., 1997; Hull et al., 2000; Lengeler et al., 2002; Poggeler, 2002) strengthen these previous indications of the presence of a cryptic sexual cycle in *D. pinea*. Despite this evidence for the presence of a sexual cycle, it remains clear that the asexual state underpins an important reproductive and dispersal strategy for the fungus, given its common occurrence on infected trees. Gene expression and other functional studies would strengthen the case for cryptic sexuality (Paoletti et al., 2005) but the presence of two idiomorphs illustrated in this study and previous population genetic data (Burgess et al. 2004; Bihon et al. 2012b) make a compelling case for a cryptic sexual cycle.

We were not able to induce a sexual state for *D. pinea* in this study despite the fact that isolates of known opposite mating type were used in our tests. Failure of fungal species to produce sexual structures either in nature or laboratory could be due to the loss of function of genes other than *MAT* genes (Arie et al., 2000; Linde et al., 2003), geographic barriers, climate and growing medium (Kerenyi et al., 2004; Debuchy et al., 2010) and many other factors. The population genetic studies conducted on *D. pinea* (Burgess et al., 1994; Bihon et al., 2012b) however, suggest that this state exists in nature, and is not due to one of these factors. This is also the case for many of fungal species where it is known that the sexual structures are rare in nature and difficult or impossible to obtain in culture.

Molecular genetics tools and especially next generation sequencing increasingly provide evidence of sexual reproduction in Ascomycete fungi that were previously thought to be asexual (e.g. Paoletti et al., 2005; Debuchy and Turgeon, 2006; O'Gorman et al., 2009; and others). *Diplodia pinea* is now added to that list. These findings have significance not only for understanding the biology of *D. pinea*, but they also have significant applied relevance when considering the evolutionary potential of this fungus to overcome various control strategies.

Table 1. Summary of *MAT* gene occurrence and frequency in *D. pinea* populations

Population	N	*MAT1-1-1*	*MAT1-2-1*	Ratio[a]	Chi-Sq[b]	*P*-value[c]
Argentina	34	21	13	1.62	1.88	0.1701[NS]
Australia	29	10	19	0.53	2.79	0.0947[NS]
Ethiopia	36	8	28	0.29	11.11	0.0009*
South Africa	67	37	30	1.23	0.73	0.3924[NS]
USA	23	10	13	0.77	0.39	0.5316[NS]

N Number of isolates per each population

[a] Ratio of *MAT1-1-1* to *MAT1-2-1*

[b] Chi-square value for relative to the expected ratio of 1:1

[c] Probability of chi-square values at $P \leq 0.05$ at 1 degree of freedom

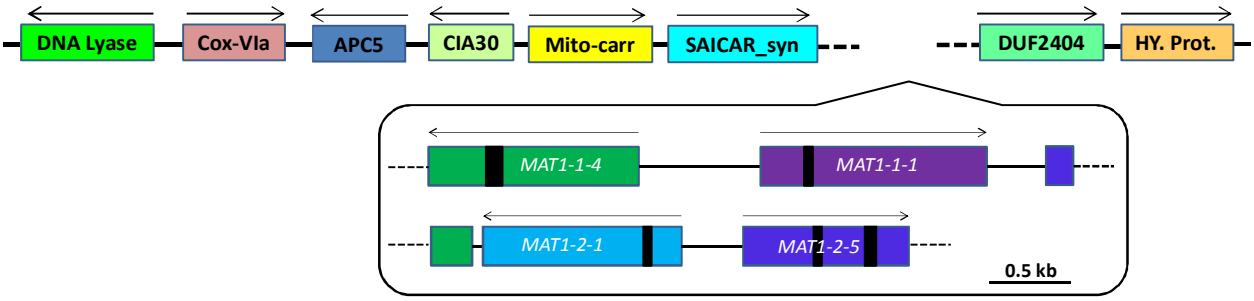

Fig. 1. Organization of *MAT* locus in heterothallic *D. pinea*. The scale bar indicates sizes of only the *MAT* idiomorphs, the distance and sizes of other genes are not to a scale. Introns in the *MAT* genes are indicated by black boxes. Arrows indicate the orientation of the *MAT* genes. Regions of similar DNA sequence are indicated by the same colour, thus the *MAT1-2* idiomorph contains a partial *MAT1-1-4* gene sequence. The *MAT1-1* idiomorph contains a partial *MAT 1-2-5* gene sequence. Gene or superfamily names refer to a gene encoding a protein showing the highest percentage identity to the putative gene product following BLASTp analysis. *Cox-Via* = Cytochrome c oxidase subunit VIa (*Macrophomina phaseolina* MS6); *APC5* = Anaphase-promoting complex subunit 5; *CIA30* = Complex I intermediate-associated protein 30; *Mito-carr* = Mitochondrian carrier protein; *SAICAR*Synthetase; *DUF2404* = Putative integral membrane protein conserved region; and Hy. Prot. = hypothetical protein.

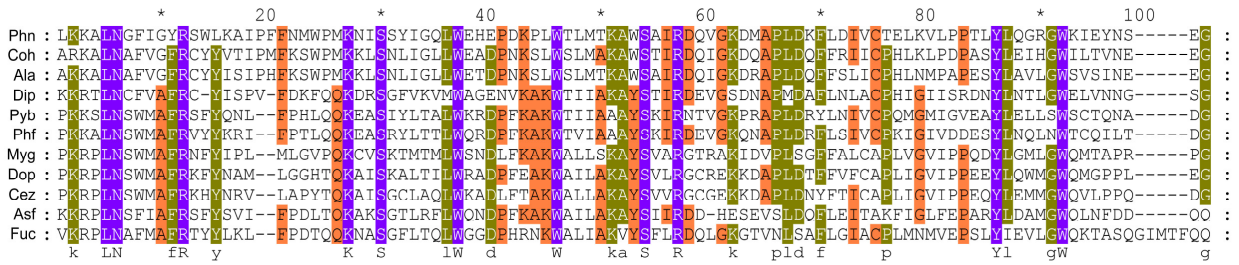

Fig. 2. Comparison of amino acid sequences of α-1 domain of *D. pinea* with selected sequences from GenBank collections. Abbreviations of species names and GenBank accession numbers in brackets: Phn = *Phaeosphaeria nodorum* (AA031742), Coh = *Cochliobolus heterostrophus* (Q02991), Ala = *Alternaria alternata* (BAA75908), Dip = *Diplodia pinea* (KF551229), Pyb = *Pyrenopeziza brassicae* (CAA06843), Phf = *Phialocephala cf. fortinii* (ADJ38491), Myg = *Mycosphaerella graminicola* (AAL30836), Dop = *Dothistroma pini* (ABK91353), Cez = *Cercospora zeina* (ABB83720), Asf = *Aspergillus fumigatus* (XP751590), Fuc = *Fusarium circinatum* (AAQ18158). Those with consensus sequences are identical and those light shaded are similar among the sequences.

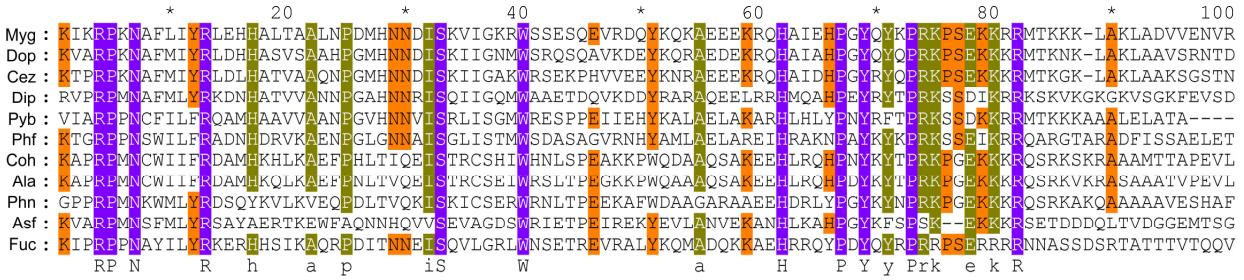

Fig. 3. Comparison of amino acid sequences of HMG box of *D. pinea* with selected sequences from the GenBank collections. Abbreviations of species name and GenBank accession numbers in brackets: Myg = *Mycosphaerella graminicola* (XP003847598), Dop = *Dothistroma pini* (ABK91355), Cez = *Cercospora zeina* (ABB83706), Dip = *Diplodia pinea* (KF551228), Pyb = *Pyrenopeziza brassicae* (CAA06844), Phf = *Phialocephala cf. fortinii* (ADJ38450), Coh = *Cochliobolus heterostrophus* (Q02990), Ala = *Alternaria alternata* (AB009451), Phn = *Phaeosphaeria nodorum* (XP001791062), Asf = *Aspergillus fumigatus* (AAX83122), Fuc = *Fursarium circinatum* (AEP03825). Those with consensus sequences are identical and those light shaded are similar among the sequences.

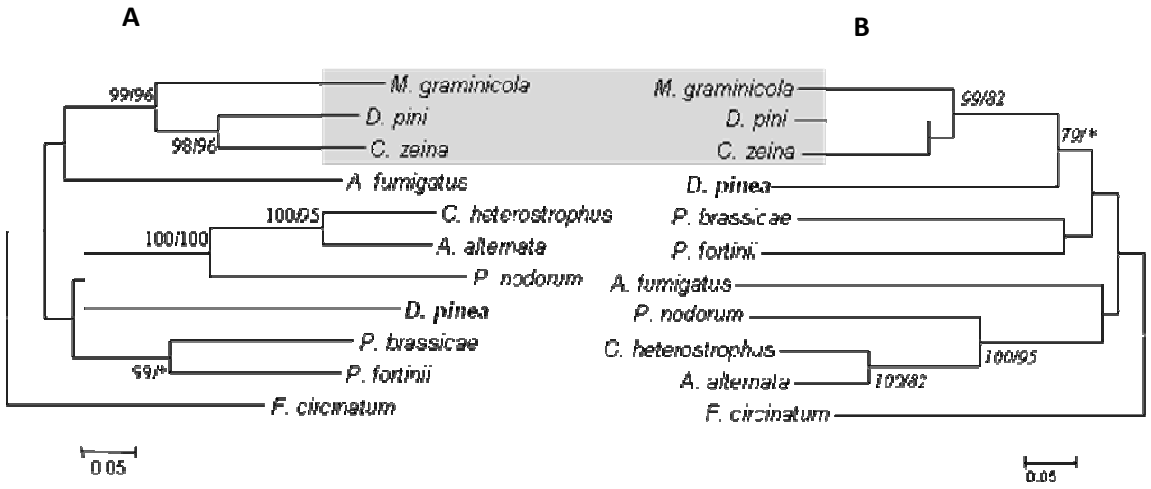

Fig. 4. Neighbor joining trees constructed using the highly conserved region of (A) α-1 box region of the *MAT1-1-1* gene and (B) HMG box region of the *MAT1-2-1* gene with *F. circinatum* as an out group. Numbers on the nodes are bootstrap values over 75% are from NJ/Maximum parsimony analysis. * Indicate bootstrap values less than 75%.

Acknowledgments

We are grateful to the Department of Science and Technology (DST)/National Research Foundation (NRF) Centre of Excellence in Tree Health Biotechnology, members of the Tree Protection Cooperative Program (TPCP) and Genomics Research Institute (University of Pretoria Institutional Research Theme) for financial support that made this study possible. The first author also acknowledges the Claude Leon Foundation and the University of Pretoria for providing post-doctoral support.

This work is based on the research supported in part by a number of grants from the National Research Foundation of South Africa (includes Grant specific unique reference number (UID) 83924). The Grant holders acknowledge that opinions, findings and conclusions or recommendations expressed in any publication generated by the NRF supported research are that of the author(s), and that the NRF accepts no liability whatsoever in this regard.

References

Aljanabi, S.M., Martinez, I., 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. Nucleic Acid Res. 25, 4692 – 4693.

Arie, T., Christiansen, S.K., Yoder, O.C., Turgeon, B.G., 1997. Efficient cloning of ascomycete mating type genes by PCR amplification of the conserved *MAT*-HMG box. Fungal Genet. Biol. 21, 118 – 130.

Arie, T., Kaneko, I., Yoshida, T., Noguchi, M., Nomura, Y., Yamaguchi, I., 2000. Mating type genes from asexual phytopathogenic ascomycetes *Fusarium oxysporum* and *Alternaria alternata*. Mol. Plant-Microbe Interact. 13, 1330 – 1339.

Bennett, R.S., Yun, S.H., Lee, T.Y., Turgeon, B.G., Arseniuk, E., Cunfer, B.M., Bergstrom, G.C., 2003. Identity and conservation of mating type genes in geographically diverse isolates of *Phaeosphaeria nodorum*. Fungal Genet. Biol. 40, 25 – 37.

Bihon, W., Burgess, T., Slippers, B., Wingfield, M.J., Wingfield, B.D., 2011. Distribution of *Diplodia pinea* and its genotypic diversity within asymptomatic *Pinus patula* trees. Australas. Plant Path. 40, 540 – 548.

Bihon, W., Burgess, T., Slippers, B., Wingfield, M.J., Wingfield, B.D., 2012a. High levels of genetic diversity and cryptic recombination is widespread in introduced *Diplodia pinea* populations. Australas. Plant Path. 41, 41 – 46.

Bihon, W., Slippers, B., Burgess, T., Wingfield, M.J., Wingfield, B.D., 2012b. Diverse sources of infection and cryptic recombination revealed in South African *Diplodia pinea* populations. Fungal Biol. 116, 112 – 120.

Blodgett, J.T., Stanosz, G.R., 1997. *Sphaeropsis sapinea* morphotypes differ in aggressiveness, but both infect on wounded red or jack pines. Plant Dis. 81, 143 – 147.

Blodgett, J.T., Stanosz, G.R., 1999. Differences in aggressiveness of *Sphaeropsis sapinea* RAPD marker group isolates on several conifers. Plant Dis. 83, 853 – 856.

Burgess, T.I., Wingfield, M.J., Wingfield, B.D., 2004. Global distribution of *Diplodia pinea* genotypes revealed using simple sequence repeat (SSR) markers. Australas. Plant Path. 33, 513 – 519.

Chitrampalam, P., Inderbitzin, P., Maruthachalam, K., Wu, B-M., Subbarao, K.V., 2013. The *Sclerotinia sclerotiorum* mating type locus (MAT) contains a 3.6-kb region that is

inverted in every meiotic generation. *PLOS ONE* 8, e56895. doi,10.1371/journal.pone.0056895.

Coppin, E., Debuchy, R., Srnaise, S., Picard, M., 1997. Mating types and sexual development in filamentous ascomycetes. Microbiol. Mol. Boil. Rev. 16, 411 – 428.

De Wet, J., Burgess, T., Slippers, B., Preisig, O., Wingfield, B.D., Wingfield, M.J., 2003. Multiple gene genealogies and microsatellite markers reflect relationships between morphotypes of *Sphaeropsis sapinea* and distinguish a new species of *Diplodia*. Mycol. Res. 107, 557 – 566.

Debuchy, R., Turgeon, B.G., 2006. Mating-type structure, evolution, and function in Euascomycetes. The Mycota1, 293 – 323.

Debuchy, R., Berteaux-Lecellier, V., Silar, P., 2010. Mating systems and sexual morphogenesis in Ascomycetes. In, Cellular and Molecular biology of filamentous fungi. (eds. Borkovich and Ebbole), ASM Press, Washington DC, pp501 – 535.

Duong, T.A., De Beer, Z. W., Wingfield, B.D., Wingfield, M.J., 2013. Characterization of the mating-type genes in *Leptographium procerum* and *Leptographium profanum*. Fungal Biol. 117, 411 – 421.

Flowers, J., Nuckles, E., Hartman, J., Vaillancourt, L., 2001. Latent infection of Austrian and Scots pine tissues by *Sphaeropsis sapinea*. Plant Dis. 85, 1107 – 1112.

Goodwin, S.B., Waalwijk, C., Kema, G.H.J., Cavaletto, J.R., Zhang, G., 2003. Clonning and analysis of the mating-type idiomorphs from the barely pathogen *Septorai passerinii*. Mol. Gen. Genomics 269, 1 – 12.

Groenewald, M., Groenewald, J.Z., Harrington, T.C., Abeln, E.C.A., Crous, P.W., 2006. Mating type gene analysis in apparently asexual *Cercospora* species is suggestive of cryptic sex. Fungal Genet. Biol. 43, 813 – 825.

Groenewald, M., Linde,C.C., Groenewald, J.Z., Crous, P.W., 2008. Indirect evidence for sexual reproduction in *Cercospora beticola* populations from sugar beet. Plant Pathol. 57, 25 – 32.

Hull, C.M., Raisner, R.M., Johnson A.D., 2000. Evidence for mating of the "asexual" yeast *Candida albicans* in a mammalian host. Science 289, 307 – 310.

Kanamori, M., Kato, H., Yasuda, N., Koizumi, S., Peever, T.L., Kamakura, T., Teraoka, T., Arie, T., 2007. Novel mating type-dependent transcripts at the mating type locus in *Magnaporthe oryzae*. Gene 403, 6–17.

Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. 9, 286 – 298.

Kerenyi, Z., Moretti, A., Waalwiji, C., Olah, B., Hornok, L., 2004. Mating type sequences in asexually reproducing *Fusarium* species. Appl. Environ. Microbiol. 70, 4419 – 4423.

Kronstad, J.W., Staben, C., 1997. Mating type in filamentous fungi. Annu. Rev. Genet. 31, 245 – 276.

Lee, S.C., Ni, M., Li, W., Shertz, C., Heitman, J., 2010. The evolution of sex, a perspective from the fungal kingdom. Microbiol. Mol. Biol. Rev. 74, 298 – 340.

Lengeler, K.B., Fox, D.S., Fraser, J.A., Allen, A., Forrester, K., Dietrich, F.S., Heitman, J., 2002. Mating-type locus of *Cryptococcus neoformans*: a step in the evolution of sex chromosomes. Eukaryot. Cell 1, 704 – 718.

Linde, C.C., Zala, M., Ceccarelli, S., McDonald, B.A., 2003. Further evidence for sexual reproduction in *Rhynchosporium secalis* based on distribution and frequency of mating-type alleles. Fungal Genet. Biol. 40, 115 – 125.

Martin, S.H., Wingfield, B.D, Wingfield, M.J., Steenkamp, E.T., 2011. Structure and evolution of the *Fusarium* mating type locus, new insights from the *Gibberella fujikuroi* complex. Fungal Genet. Biol. 48, 731 – 740.

McDonald, B.A., Linde, C., 2002. Pathogen population genetics, evolutionary potential and durable resistance. Annu. Rev. Phytopathol. 40, 349 – 379.

Milgroom, M.G., 1996. Recombination and the multilocus structure of fungal populations. Annu. Rev. Phytopathol. 34, 457 – 477.

Ni, M., Feretzaki, M., Sun, S., Wang, X., Heitman, J., 2011. Sex in fungi. Annu. Rev. Genet. 45, 405 – 430.

O'Gorman, C. M., Fuller, H.T., Dyer, P.S. 2009. Discovery of a sexual cycle in the opportunistic fungal pathogen *Aspergillus fumigatus*. Nature 457, 471 – 474.

Palmer, M.A., Stewart, E.L., Wingfield, M.J., 1987. Variation among isolates of *Sphaeropsis sapinea* in the north central United States. Phytopathol. 77, 944 – 948.

Paoletti, M., Rydholm, C., Schwier, E. U., Anderson, M. J., Szakacs, G., Lutzoni, F., Debeaupuis, J.-P., Latgé, J.-P., Denning, D.W., Dyer, P.S., 2005. Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. Current Biol. 15, 1242 – 1248.

Poggeler, S., 2002. Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. Curr. Genet. 42, 153 – 160.

Sanger, F., Nicklen, S., Coulson, A.R., 1977. DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. USA 74, 5463 – 5467.

Schoch, C.L., Shoemaker, R.A., Seifert, K.A., Spatafora, J.W., Crous, P.W., 2006. A multigene phylogeny of the Dothideomycetes using four nuclear loci. Mycologia 98, 1041 - 1052.

Slippers, B., Wingfield, M.J., 2007. Botryosphaeriaceae as endophytes and latent pathogens of woody plants: diversity, ecology and impact. Fugal Biol. Rev. 21, 90 – 106.

Smith, J.M., Smith, N.H., O'Rourke, M., Spratt, B.G., 1993. How clonal are bacteria? Proc. Natl. Acad. Sci. USA 90, 4384 – 4388.

Stanosz, G.R., Smith, D.R., 1996. Evaluation of fungicides for control of *Sphearopsis* shoot blight or red pine nursery seedlings. Can. J. For. Res. 26, 492 – 497.

Swart, W.J., Wingfield, M.J., 1991. Biology and control of *Sphaeropsis sapinea* on *Pinus* species in South Africa. Plant Dis. 75, 761 – 766.

Swart, W.J., Knox-Davies, P.S., Wingfield, M.J., 1985. *Sphaeropsis sapinea*, with special reference to its occurrence on *Pinus* spp. in South Africa. South Afr. Forest. J. 35, 1 – 8.

Swofford, D.L., 2003. PAUP*, Phylogenetic analysis using parsimony (* and other methods). Version 4.0b 10. Sinauer Associates, Sunderland, Massachusetts, Sinauer Associates.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA 5**,** Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. Mole. Biol. Evol. 28, 2731 – 2739.

Taylor, J.W., Geiser, D.M., Burt, A., Koufopanou, V., 1999. The evolutionary biology and population genetics underlying fungal strain typing. Clin. Microbiol. Rev. 12, 126 – 146.

Tsui, C.K.M., DiGuistini, S., Wang, Y., Feau, N., Dhillon, B., Bohlmann, J., Hamelin, R.C., 2013. Unequal recombination and evolution of the mating type (*MAT*) loci in the pathogenic fungus *Grosmannia clavigera* and relatives. G3 Genes|Genomics|Genetics 3, 465 – 480.

Turgeon, B.G., Bohlmann, H., Ciuffetti, L.M., Christiansen, S.K., Yang, G., Schafer, W., Yoder, O, C., 1993. Cloning and analysis of the mating type genes from *Cochliobolus heterostrophus*. Mol. Gen. Genet. 238, 270 – 284.

Turgeon, B.G., 1998. Application of mating type gene technology to problems in fungal biology. Annu. Rev. Phytopathol. 36, 115 – 137.

Turgeon, B.G., Yoder, O.C., 2000. Proposed nomenclature for mating type genes of filamentous ascomycetes. Fungal Genet. Biol. 31, 1 – 5.

Waalwijk, C., Mendes, O., Verstappen, E.C.P., de Waard, M.A., Kema, H.J., 2002. Isolation and characterization of the mating-type idiomorphs from the wheat septoria blotch fungus *Mycosphaerella graminicola*. Fungal Genet. Biol. 35, 277 – 286.

Wingfield, B.D., Steenkamp, E.T., Santana Q.C., et al., 2012. First fungal genome sequence from Africa: A preliminary analysis. S. Afr. J. Sci. 108, 104 – 112.

Wingfield, M.J., Knox-Davies, P.S., 1980. Association of *Diplodia pinea* with a root disease of pines in South Africa. Plant Dis. 64, 221 – 223.

Wingfield, M.J., Slippers, B., Roux, J., Wingfield, B.D. 2001. Worldwide movement of exotic forest fungi, especially in the tropics and the southern hemisphere. Biosci. 51, 134 – 140.

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

Zhan, J., Kema, G.H.J., Waalwijk, C., McDonald, B.A., 2002. Distribution of mating type alleles in the wheat pathogen *Mycosphaerella graminicola* over spatial scales from lesions to continents. Fungal Genet. Biol. 36, 128 – 136.

Zwolinski, J.B., Swart, W.J., Wingfield, M.J., 1990. Intensity of dieback induced by *Sphaeropsis sapinea* in relation to site conditions. Eur. J. For. Pathol. 20, 167 – 174.