Accepted Manuscript

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

Anabela Lopes, Benedetto T. Linaldeddu, Alan J.L. Phillips, Artur Alves

PII: S1878-6146(18)30054-0

DOI: 10.1016/j.funbio.2018.03.012

Reference: FUNBIO 914

To appear in: *Fungal Biology*

Received Date: 12 October 2017

Revised Date: 28 February 2018

Accepted Date: 27 March 2018

Please cite this article as: Lopes, A., Linaldeddu, B.T., Phillips, A.J.L., Alves, A., Mating type gene analyses in the genus Diplodia: from cryptic sex to cryptic species, *Fungal Biology* (2018), doi: 10.1016/j.funbio.2018.03.012.

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

Anabela LOPES[a], Benedetto T. LINALDEDDU[b], Alan J.L. PHILLIPS[c], Artur ALVES[a*]

[a]Departamento de Biologia, CESAM, Universidade de Aveiro, Campus Universitário de Santiago, 3810-193 Aveiro, Portugal

[b]Department of Land, Environment, Agriculture, and Forestry (TeSAF), University of Padova, Viale dell'Università 16, 35020 - Legnaro (PD), Italy

[c]University of Lisbon, Faculty of Sciences, Biosystems and Integrative Sciences Institute (BioISI), Campo Grande, 1749-016 Lisbon, Portugal

***Corresponding author:**

Artur Alves

e-mail: artur.alves@ua.pt

Phone: +351 234370766

Fax: +351 234 372587

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 amplification and sequencing of the *MAT1-1-1* and *MAT1-2-1* genes, and determination of 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 concatenated sequences of the ribosomal internal transcribed spacer (ITS), translation elongation factor 1-alpha (*tef1-α*) and beta-tubulin (*tub2*). Individual phylogenies based on *MAT* genes clearly differentiated all species analysed and agree 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 loci 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 revealed cryptic species in this genus (Phillips et al. 2012; Phillips et al. 2013; Alves et al. 2014; Linaldeddu et al. 2016). This feature is common in other genera of *Botryosphaeriaceae* and it is now widely accepted that morphology is a poor indicator of species limits. Thus, the 30 or more 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 the translation 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 of all species using 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, 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 and Heitman 2015). In ascomycetes, mating systems are generally assumed to be bipolar due to the existence in the single *MAT* locus of two alternate forms (idiomorphs) that define two mating types (Debuchy and

Turgeon 2006). 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 and Turgeon 2006; Lee et al. 2010; Sun and Heitman 2015). In ascomycetes, homothallic (self-fertile) species contain in their *MAT* locus both idiomorphs *MAT1-1* and *MAT1-2*, while heterothallic (self-sterile) species contain either *MAT1-1* or *MAT1-2* (Lee et al. 2010; Ni et al. 2011).

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 and Heitman 2015). Because of the evolutionary dynamism of *MAT* genes, using them to establish phylogenetic relationships, even in apparently asexual fungi, has resulted in clarification of species delimitation in numerous studies (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 induction of the sexual cycle *in vitro* has been unsuccessful (e.g. Bihon et al. 2014). Knowledge about *MAT* genes in the genus is currently limited to *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 compared 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

110 ***Fungal strains and culture conditions***

111 A total of 50 strains representing 20 species were studied and these are listed in
112 Table 1. All strains were grown and maintained on half-strength potato-dextrose agar
113 (PDA) (HIMEDIA, India) at room temperature.

114

115 ***Analyses of Diplodia sapinea genomes and design of MAT genes primers***

116 The genomes of two strains of *Diplodia sapinea* from each mating type were
117 available and the primers in the present study were firstly designed based on them
118 (PRJNA215898; PRJNA242796). Analysis of the *MAT* locus of *D. sapinea* had already
119 been done by Bihon et al. (2014), which facilitated our approach to *MAT* genes analyses
120 (KF551229, KF551228). Although the locus contained more than one gene only the main
121 *MAT1-1-1* and *MAT1-2-1* genes were used in this study.

122 A first set of primers for each *MAT* gene was designed manually. Characteristics
123 such as length, GC content, melting temperature, potential hairpin formation,
124 complementarity and potential self-annealing sites were checked with OligoCalc
125 (Oligonucleotide Properties Calculator
126 http://biotools.nubic.northwestern.edu/OligoCalc.html) and Sequence Manipulation Suite
127 (http://www.bioinformatics.org/sms2/pcr_primer_stats.html). Primers were synthesized by
128 STAB Vida Lda (Lisbon, Portugal), reconstituted and diluted (10 pmol) in Tris-EDTA
129 buffer, and stored at -20°C. The first set of primers was then tested in *D. sapinea* and in all
130 the remaining species. Whenever primers gave rise to amplification, the amplicons were
131 sequenced and new sets of primers were designed based on consensus of the acquired *MAT*
132 sequences. The aim of this procedure was to reach a set of "universal" primers able to
133 amplify each gene in all the studied species. All the primers designed are listed in Table 2
134 (see also Fig. S1). The primers used in the amplification of the sequences used in
135 phylogenetic analyses of *MAT* genes are discriminated in the Table 3. Combinations of
136 primers and species that generated an amplicon are available in the supplementary Tables
137 S1 and S2.

138

139 ***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 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 and sequenced with primers T1 and Bt2b (Glass and Donaldson 1995; O'Donnell and Cigelnik 1997) with the cycling conditions previously described by Lopes et al. (2017). The histone gene (*his3*) was amplified and sequenced with primers CYLH3F and H3-1b (Glass and Donaldson 1995; Crous et al. 2004) according to Santos et al. (2017). For amplification and sequencing of actin (*act2*) the primers ACT-512F and ACT-783R (Carbone and Kohn 1999) were used 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.

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™ 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 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 were done with MEGA6 v. 6.06 (Tamura et al. 2013). All gaps were included in the analyses. The best model of DNA sequence evolution used for each dataset was determined automatically by the software (See Figs 3 and 4). Maximum likelihood (ML) analyses were performed on a Neighbor-Joining (NJ) starting tree generated by the software. A bootstrap analysis (1000 replicates) was used to estimate the robustness of nodes. Alignments were deposited in TreeBASE (http://purl.org/phylo/treebase/phylows/study/TB2:S21615).

Phylogenetic informativeness of the loci used in phylogenetic analyses

Profiles of phylogenetic informativeness for each locus were obtained using PhyDesign (López-Giráldez and Townsend 2011; http://phydesign.townsend.yale.edu/) as proposed by Townsend (2007). This allows a comparison of the different loci through calculation of the informativeness per base pair as well as net phylogenetic informativeness, which takes into account sequence length. Since the species are all

presumably 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 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) (Fig. 1, Table 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). Variability was also seen in the sequence alignments with the internal regions being more dissimilar than the regions closer to the ends of the genes. 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). In the species analysed one small intron (48–49 bp) located near the start of the gene was identified in each *MAT* gene using the FGENESH+ gene prediction program (Solovyev et al. 2006). 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 (Fig. 1).

No homothallic species could be detected. For *D. corticola*, *D. sapinea*, *D. scrobiculata* and *D. seriata* heterothallism was confirmed from the available sequenced genomes (PRJNA325745, PRJNA215898, PRJNA242796, PRJNA278001, PRJNA261773) through BLAST searches using the *MAT* genes of *D. sapinea* (Bihon et al. 2014; KF551229; KF551228). These confirmed the presence of only one mating type in the genomes. This was further confirmed by PCR amplification with the primers developed in this study. 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 3 and 4).

Phylogenetic analyses

All species of *Diplodia* were, in general, distinguished by the phylogenetic analyses of each *MAT* gene and by the combined dataset of ITS rDNA, *tef1-α* and *tub2* (Figs. 2 and 3). In the phylogenetic analysis of *MAT1-2-1* gene sequences of *D. mutila* and *D. subglobosa* grouped very closely with short branch lengths. *MAT1-2-1* sequences from both species are very similar and differ in only three nucleotide positions (*MAT1-2-1* alignment in supplementary data).

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. 2 and 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.

In the concatenated data analysis isolates of *D. corticola* are separated into two main groups as a result of fixed differences in the *tef1-α* sequences. In order to confirm that the separation of *D. corticola* isolates into two groups is stable and congruent, phylogenetic analyses were performed on each individual locus (*MAT1-1-1*, *MAT1-2-1*, ITS, *tef1-α* and tub2) as well as two additional loci (*act2* and *his3*). The two groups of *D. corticola* formed in the analysis of the three concatenated regions are not supported by either the *MAT* gene phylogenies or by the ITS and *his3* phylogenies (Fig. 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).

Phylogenetic informativeness of the loci used in the phylogenetic analyses

As can be seen in Figure 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. 5).

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.

For a long time the lack of sequenced genomes of *Diplodia* species was an obstacle to the development of suitable primers to amplify and study *MAT* genes. 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. 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 *D. scrobiculata* CMW30223 (PRJNA278001; Wingfield et al. 2015). Similar to what has been reported for the genus *Neofusicoccum* (Lopes et al. 2017), *MAT* gene sequence alignments revealed that internal regions of the genes are more dissimilar than the regions closer to the ends 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 and 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 and Linde 2002).

Working with *MAT* genes in heterothallic species produces an additional challenge in comparison to homothallic species in that two idiomorphs are present in a population but only one idiomorph might have been sampled in a 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. 2 and 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. 3). They are clearly effective in resolving cryptic species in clade I. In the case of *D. mutila* and *D. subglobosa* (clade II) in the phylogenetic analysis of *MAT1-2-1* these species group very closely but still with some differences at sequence level (Fig. 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 to confirm if these differences

are stable within each species. Unfortunately, *MAT1-1* strains of *D. 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, which are a result of fixed differences in the *tef1-α* sequences, as different lineages within a single species. Although in our concatenated genes phylogeny isolates BL36 and BL37 appear to be more closely to *D. quercivora*, previous studies with more strains clearly indicate that these are two subgroups within *D. corticola* (Linadeddu et al. 2013, Smahi et al. 2017). Phylogenetic relationships of *D. 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). 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. 5). Thus, in the resolution of *Diplodia* species we strongly recommend the use of *MAT* genes due to their high resolving

power (Fig. 3) as well their high net and per-site phylogenetic informativeness (Fig. 5). In addition to *MAT* genes, the *tef1-α* locus seems to be useful in all species except *D. corticola*.

In conclusion, the 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* species 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.

Smahi H, Belhoucine-Guezouli L, Berraf-Tebbal A, Chouih S, Arkam M, Franceschini A, Linaldeddu BT, Phillips AJL, 2017. Molecular characterization and pathogenicity of *Diplodia corticola* and other Botryosphaeriaceae species associated with canker and dieback of *Quercus suber* in Algeria. *Mycosphere* **8**(2):1261–1272.

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.

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.

Townsend JP, 2007. Profiling phylogenetic informativeness. *Systematic Biology* **56**:222–231.

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.

Table 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	-

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	-
	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.

Table 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

Table 3 Primers and annealing temperatures used in the PCR

Species	Annealing Temperatures (°C)			
	MAT1-1-1		***MAT1-2-1***	
	Primer set A	**Primer set B**	**Primer set C**	**Primer set D**
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

Table 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	-
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.

Figure Captions

Fig 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).

Fig 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.

Fig 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.

Fig 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.

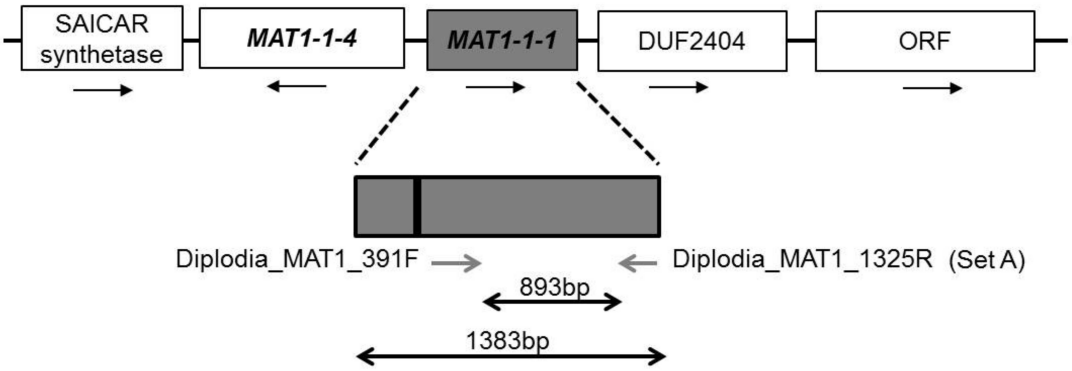
Fig 5 – Profiles of phylogenetic informativeness per gene through relative time periods (time units, TU). Net Phylogenetic informativeness for the loci ITS, *tef1-α*, *tub2*, *MAT1-1-1* (A) and ITS, *tef1-α*, *tub2*, *MAT1-1-2* (C); Phylogenetic informativeness per site for the loci ITS, *tef1-α*, *tub2*, *MAT1-1-1* (B) and ITS, *tef1-α*, *tub2*, *MAT1-1-2* (D).

Fig S1 - 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.

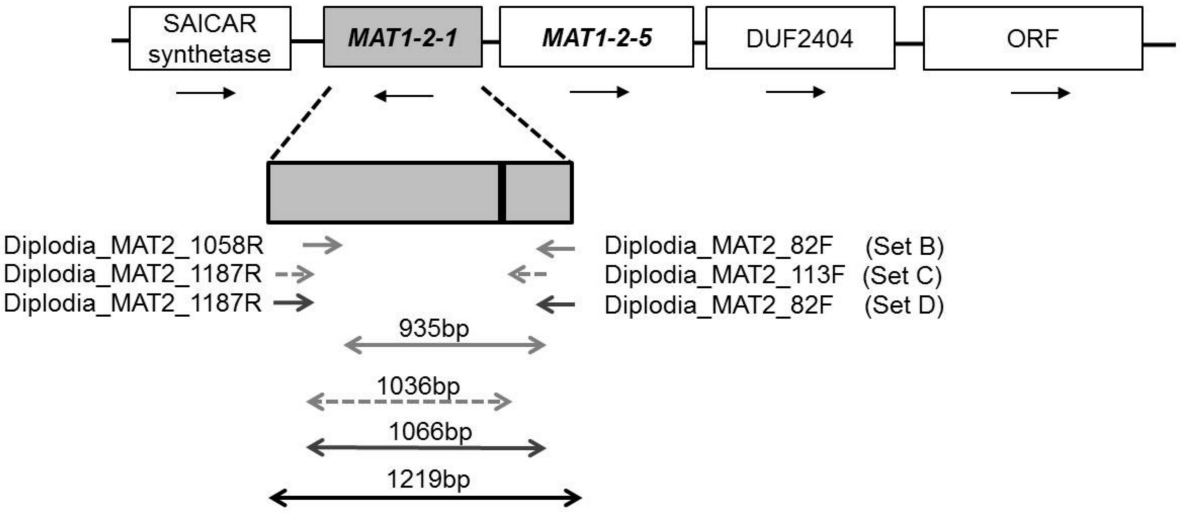
Table S1 - Possible primers combinations for *MAT1-1-1* gene

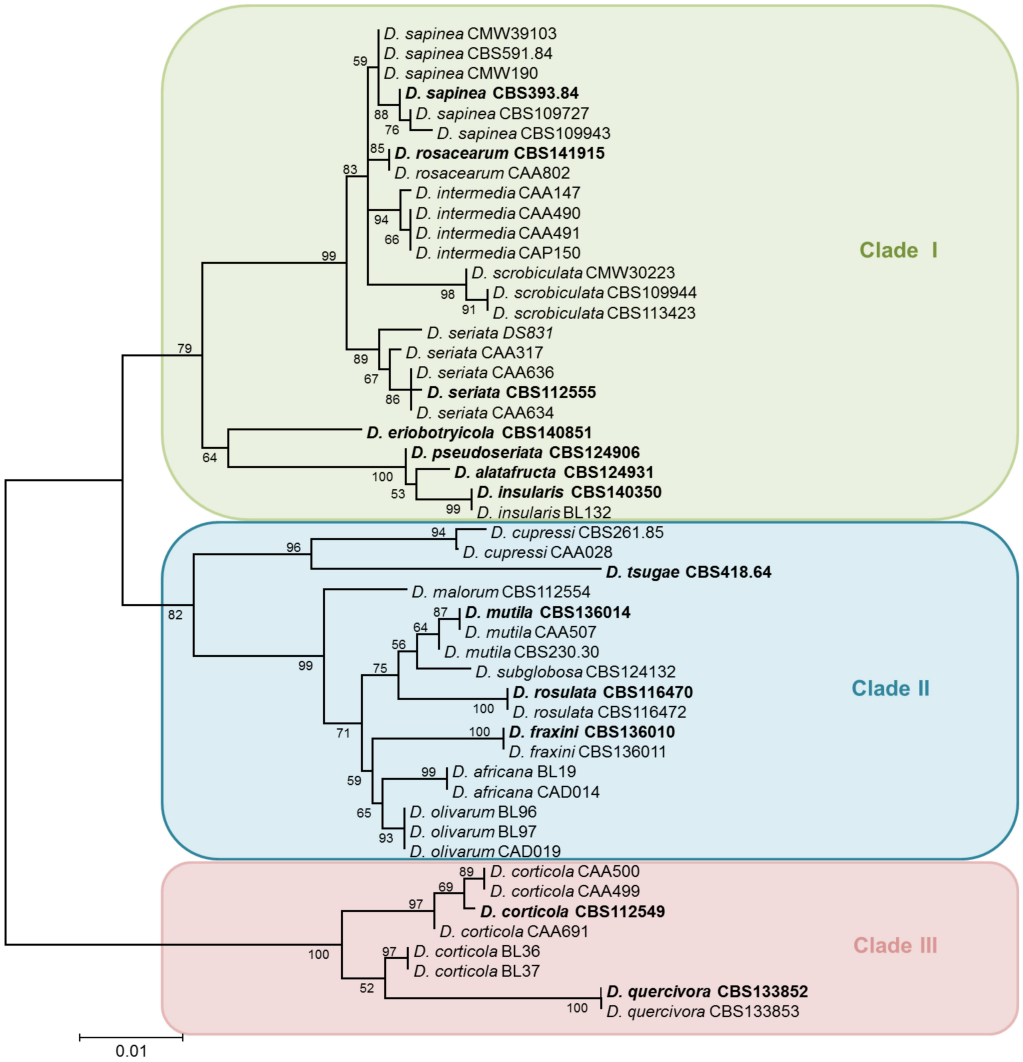
Table S2 - Possible primers combinations for *MAT1-2-1* gene

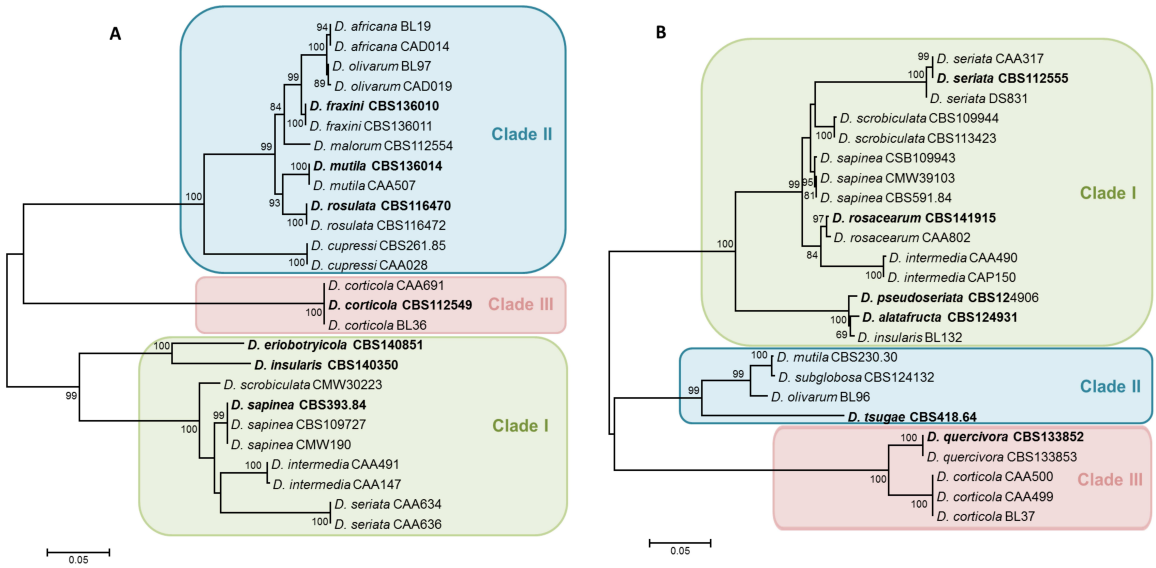
***MAT1-1* Idiomorph (Strain CMW190)**

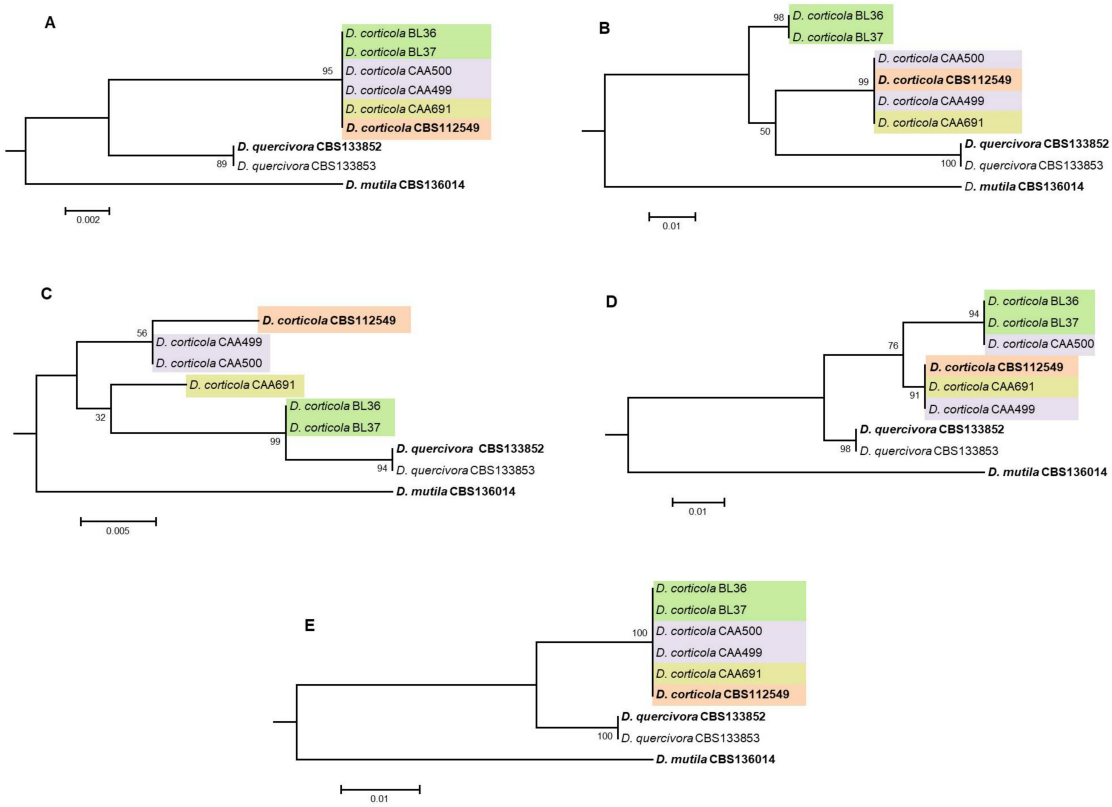


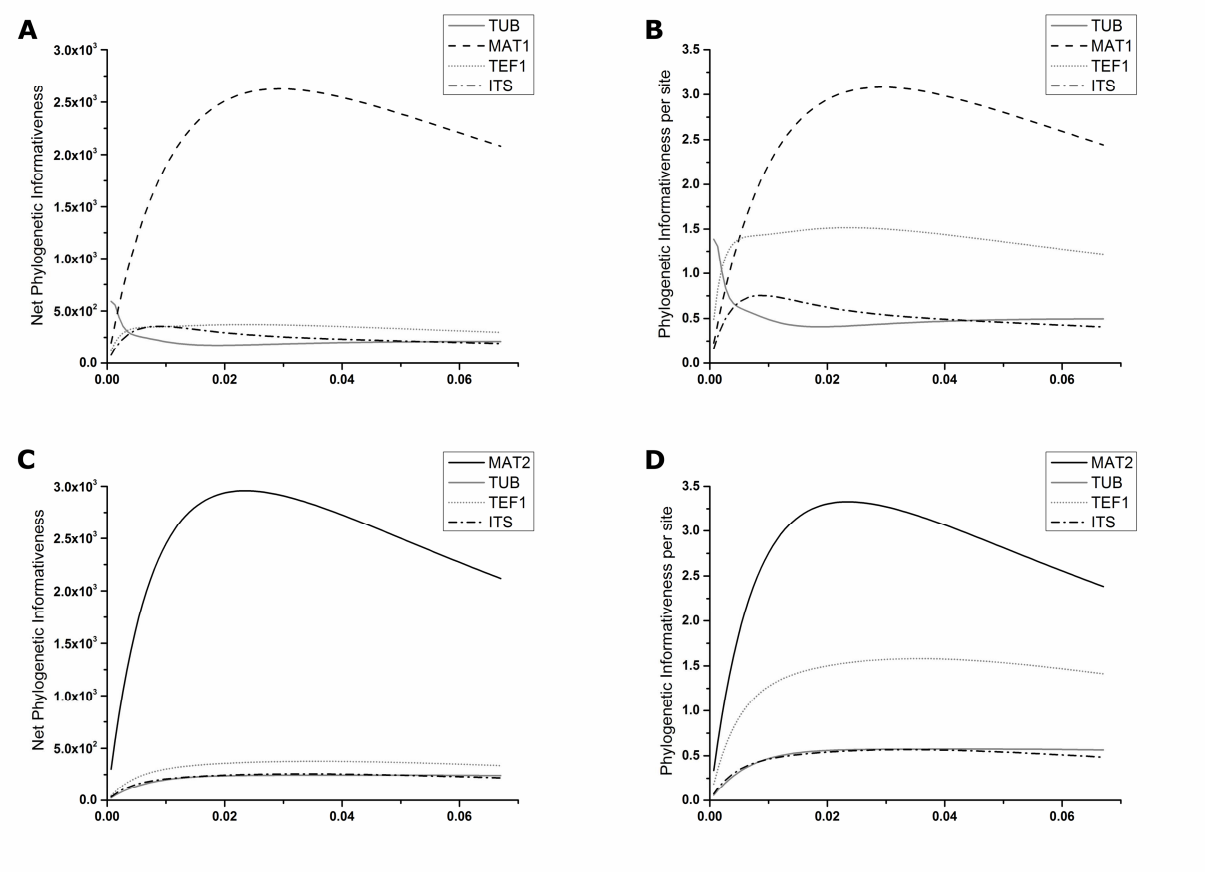
***MAT1-2* Idiomorph (Strain CMW39103)**




D. sapinea CMW39103
D. sapinea CBS591.84
D. sapinea CMW190
D. sapinea CBS393.84
D. sapinea CBS109727
D. sapinea CBS109943
D. rosacearum CBS141915
D. rosacearum CAA802
D. intermedia CAA147
D. intermedia CAA490
D. intermedia CAA491
D. intermedia CAP150
D. scrobiculata CMW30223
D. scrobiculata CBS109944
D. scrobiculata CBS113423
D. seriata DS831
D. seriata CAA317
D. seriata CAA636
D. seriata CBS112555
D. seriata CAA634
D. eriobotryicola CBS140851
D. pseudoseriata CBS124906
D. alatafructa CBS124931
D. insularis CBS140350
D. insularis BL132
Clade I
D. cupressi CBS261.85
D. cupressi CAA028
D. tsugae CBS418.64
D. malorum CBS112554
D. mutila CBS136014
D. mutila CAA507
D. mutila CBS230.30
D. subglobosa CBS124132
D. rosulata CBS116470
D. rosulata CBS116472
D. fraxini CBS136010
D. fraxini CBS136011
D. africana BL19
D. africana CAD014
D. olivarum BL96
D. olivarum BL97
D. olivarum CAD019
Clade II
D. corticola CAA500
D. corticola CAA499
D. corticola CBS112549
D. corticola CAA691
D. corticola BL36
D. corticola BL37
D. quercivora CBS133852
D. quercivora CBS133853
Clade III
0.01







Research highlights

- Revealing *MAT* loci organization in *Diplodia* species
- All *Diplodia* species studied are heterothallic
- Development of PCR-based assays for mating type determination
- *MAT* genes phylogenies distinguish species in *Diplodia*