Unidirectional mating-type switching confers self-fertility to *Thielaviopsis cerberus*, the only homothallic species in the genus

Daniella Krämer, Frances A. Lane, Emma T. Steenkamp, Brenda D. Wingfield, and P. Markus Wilken

Department of Biochemistry, Genetics and Microbiology; Forestry and Agricultural Biotechnology Institute, University of Pretoria, Private Bag X20, Hatfield, Pretoria, South Africa, 0028; corresponding author e-mail: markus.wilken@fabi.up.ac.za

Abstract

Sexual reproduction is ubiquitous in nature, and nowhere is this more so than in the fungi. Heterothallic behaviour is observed when there is a strict requirement of contact between two individuals of opposite mating-type for sexual reproduction to occur. In contrast, a homothallic species can complete the entire sexual cycle in isolation, although several genetic mechanisms underpin this self-fertility. The specific genetic mechanism that underpins homothallic behaviour can be inferred by characterising the structure and gene-content of the mating-type locus, which contains genes that are involved in the regulation of sexual reproduction. In this study, the genetic basis of homothallism in *Thielaviopsis cerberus* was investigated, the only known self-fertile species within this genus. Using genome sequencing and conventional molecular techniques, two versions of the mating-type locus were identified in this species. This is typical of species that have a unidirectional mating-type switching reproductive strategy. The first version was a self-fertile locus that contained four known mating-type genes, while the second was a self-sterile version with a single mating-type gene. The conversion from a self-fertile to a self-sterile locus is likely mediated by a homologous recombination event at two direct repeats present in the self-fertile locus, resulting in the deletion of three mating-type genes and one of the repeats. Both locus versions were present in isolates that were self-fertile, while self-sterility was caused by the presence of only a

switched locus. This study provides a clear example of the fluidity in the mating-type loci that is common among even closely related fungal species.

Keywords:

Homothallism, unidirectional mating-type switching, *Thielaviopsis*, mating-type locus, sexual reproduction

1. Introduction

Fungi represent fascinating systems of sexual reproduction (Heitman et al., 2013; Ni et al., 2011) because they employ a magnitude of reproductive strategies (Coppin et al., 1997; Glass et al., 1990; Haber, 2012; Turgeon and Yoder, 2000). Information gained from studying sexual processes in model fungi has facilitated significant progress in unravelling the mating systems of various non-model species. However, many non-model fungi have novel mating strategies, thus highlighting that much remains unknown about the diversity and associated mechanisms underlying sexual reproduction in fungi (Harrington et al., 1998; Kanematsu et al., 2007; Wilson et al., 2015b; Xu et al., 2016).

Fungal mating strategies are regarded as either heterothallic or homothallic. In heterothallism, initiation of the sexual cycle is defined by the need for contact between two isolates of opposite mating-type (Blakeslee, 1904). Mating-type identity is conferred by the genes present at the mating-type (*MAT1*) locus that controls many aspects of the sexual process (Yoder et al., 1986; Yun et al., 1999). In heterothallic species, each mating partner has either a *MAT1-1* or *MAT1-2* “allelic” or idiomorphic form of the locus (Metzenberg and Glass, 1990). Sexual reproduction takes place when the products of both these idiomorphs are present in a single cell, which is achieved in heterothallic species by fusion of the two interacting partners. At the genetic level, heterothallism is thus synonymous with obligate outcrossing (Glass and Nelson, 1994).

Homothallism is an umbrella term used to describe sexual reproduction in the absence of a mating partner (Blakeslee, 1904; Whitehouse, 1949; Wilson et al., 2015b). In other words, a homothallic individual is self-fertile and this condition may be achieved through various genetic mechanisms or homothallic behaviours (Wilson et al., 2015b). Most forms of homothallism are conferred by both *MAT1-1* and *MAT1-2* genes being present in a single cell (Aanen and Hoekstra, 2007). For example, in primary homothallic species, both these gene types are present at the *MAT1* locus (Blakeslee, 1904; Nasmyth, 1982). In the case of pseudohomothallic fungi, a single spore can contain two haploid nuclei of opposite mating-type and their subsequent germination produces heterokaryons capable of self-fertility (Lin and Heitman, 2007; Raju, 1992). In species capable of unidirectional mating-type switching, homothallism is mechanistically more complex (Harrington and McNew, 1997; Perkins, 1987; Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). The *MAT1* locus of these fungi is structured such that the *MAT1-2* genes are usually flanked by *MAT1-1* genes (Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). In certain cells, homologous recombination between direct repeats then eliminates the *MAT1-2* genes, thereby effectively restructuring their *MAT1* locus (Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). This confers *MAT1-1* identity to such cells, causing homokaryotic isolates harbouring them to be self-sterile (Yun et al., 2017). At the gene and genomic levels, these different mechanisms for achieving homothallism may be distinguished by analysis of the architecture of the *MAT1* locus, because they are each associated with a distinct architecture (Coppin et al., 1997; Nelson, 1996; Yun et al., 2000).

Here we studied the *MAT1* locus of *Thielaviopsis cerberus* (Ascomycota, Ceratocystidaceae), the only homothallic species within a genus of heterothallic fungi (De Beer et al., 2014; Mbenoun et al., 2014). *Thielaviopsis* species are known as pathogens on mainly monocotyledonous plants including sweet pea, pineapple and cacao tree (De Beer et al., 2014). Detailed analyses showed that all these fungi, except for *T. cerberus*, were obligately outcrossing in culture and their *MAT1* locus structure were typical of a heterothallic species, with two idiomorphs (*MAT1-1* and *MAT1-2*) present in the population (Wilken et al., 2018).

This previous study also showed the unusual position of a *MAT1-1* specific gene (i.e., *MAT1-1-2*) in the *MAT1-2* idiomorph (Wilken et al., 2018). In the case of *T. cerberus*, isolates typically showed self-fertility in culture by producing ascomata and ascospores in isolation (Mbenoun et al., 2014). Although nothing was known about the *MAT1* locus of this species, homothallism is widespread in the Ceratocystidaceae (Liu et al., 2018; Simpson et al., 2018; Wilken et al., 2018; Wilson et al., 2015b). For example, all species of *Ceratocystis* and *Endoconidiophora*, as well as *Davidsoniella virescens* achieve homothallism via unidirectional mating-type switching (Baker et al., 2003; Harrington and McNew, 1997; Wilken et al., 2014; Witthuhn et al., 2000).

The aim of the current study was to elucidate the genetic basis of homothallism in *T. cerberus*. To do this, a draft genome sequence was generated for this species to identify and annotate its *MAT1* locus. Single ascospore progeny were used to experimentally validate the structure and gene content of the locus, while RNA sequence data was generated and used to confirm the predicted gene models. This provides the basis for a model that explains the changes in locus configuration that accompany the transition between heterothallism and homothallism, as well as contributes to a larger effort to study sexual reproduction in the Ceratocystidaceae.

2. Materials and Methods

2.1. Thielaviopsis cerberus isolates and genome sequencing

For this study, two isolates (CMW36641 and CMW36653) of *T. cerberus* were obtained from the culture collection of the Forestry and Agricultural Biotechnology Institute (FABI) at the University of Pretoria. These were specifically selected for this study as they represent self-sterile and self-fertile individuals respectively, both of which were previously isolated from oil palm (*Elaeis guineensis*) in Cameroon (Mbenoun et al., 2014). The fungi were grown and maintained at 25°C on medium (MEA-TS) containing 20 g/L Malt Extract and 20 g/L Agar (Biolab, Merck, South Africa), as well as 100 mg/L Thiamine and 150 mg/L Streptomycin (Sigma, Steinheim, Germany).

To determine the full genome sequence for *T. cerberus*, the self-fertile isolate (CMW36653) was grown for 14 days at 25°C on MEA-TS before being subjected to genomic DNA isolation using the method of Plaza et al. (2004). The extracted DNA was sent to the Central Analytical Facility at the University of Stellenbosch (Stellenbosch, South Africa) where it was used to generate a single-read library of 400 base pair (bp) target length using the Ion S5 system with the Ion 530 Chip Kit (Thermo Fisher Scientific, Johannesburg, South Africa). The raw reads were subjected to error correction and then assembled with SPAdes v 3.14.0 using the default k-values but flagging the data as IonTorrent and using the –careful option to reduce the number of mismatches and short indels (Bankevich et al., 2012; Nurk et al., 2013).

General genome statistics (genome length, GC content, N50, L50, largest contig size and number of ambiguous nucleotides per 100 bp) for the assembled genome were calculated using QUAST v5.0.1 (Mikheenko et al., 2018), while completeness was assessed using the Benchmarking Universal Single Copy Orthologs tool (BUSCO v 2.0.1; Simão et al., 2015) with both the Fungi_odb9 (2017-02-13) and Ascomycota_odb9 (2017-02-13) databases. The number of protein coding genes in the genome was estimated using *de novo* prediction with AUGUSTUS based on *Fusarium graminearum* gene models (Keller et al., 2011; Stanke and Morgenstern, 2005).

2.2. Identification and annotation of the *MAT1* locus

CLC Main Workbench v. 20.0 (CLC Bio, Aarhus, Denmark) was used to screen the draft genome sequence of CMW36653 for the presence of mating-type genes using tBLASTn searches with the putative MAT1-1-1, MAT1-1-2, MAT1-2-1 and MAT1-2-7 protein sequences from *T. punctulata* (NCBI accession numbers: BK010318 and KX989056; Wilken et al., 2018). Contigs that produced matches with these proteins (expect [e]-values ≤ 0.01) were extracted and analysed with AUGUSTUS and FGENESH *de novo* Gene Prediction software (Solovyev et al., 2006; Stanke et al., 2006) to predict any genes that would be present, again using *F. graminearum* gene models. If any gene that produced a hit during the initial tBLASTn searches were not annotated during the *de novo* predictions, the relevant *T. punctulata* protein was

used as a guide for FGENESH+ prediction (Solovyev, 2007). All the identified gene models were then used to annotate the relevant contig.

To provide support for our gene models, the predicted protein sequences of all the *MAT1* genes were compared to a list of published *Thielaviopsis MAT1* proteins (Wilken et al., 2018). The sequences for seven self-fertile and self-sterile loci, representing four *Thielaviopsis* species (supplementary table 2), were obtained from the GenBank database (Sayers et al., 2019). The annotated gene models were translated into proteins in CLC Main Workbench v. 20.0, producing a dataset of known *Thielaviopsis* proteins. A best-hit reciprocal BLASTp analysis was done where the predicted *T. cerberus* proteins were compared to the *Thielaviopsis* data set, and vice versa (Altschul et al., 1990; States and Gish, 1994). Identified *MAT1* genes were screened for the presence of putative conserved domains using the online interphase of the InterPro v 83.0 tool (Blum et al., 2020).

The *MAT1* locus of species capable of unidirectional mating-type switching are characterized by direct repeats that allow recombination-based elimination of *MAT1-2* gene elements (Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017), and therefore all contigs with putative *MAT1* sequences were subjected to analysis using RepFind (Betley et al., 2002). The following parameters were used: The P-value cut-off was 0.0001, the minimum and maximum repeat length were 3 bp and infinity respectively, the low complexity filter was on and the order of background Markov model was set to 1. These contigs were also subjected to analysis with REPuter (Kurtz et al., 2001) with settings as follows: maximum number of repeats was set to 50 and the minimum size of the repeats were 30 bp. Furthermore, the repeats in all directions were selected (forward (direct), reverse, complement and palindromic). The E-values produced by the programme were used to evaluate the statistical support for the repeat regions (Kurtz et al., 2001).

2.3. Preparation of cultures established from single ascospores and hyphal tips

To experimentally verify the *MAT1* locus structure and gene content of *T. cerberus*, individual ascospores or hyphal tips were isolated and cultured to obtain genetically distinct individuals. For this purpose, isolates CMW36641 and CMW36653 were inspected routinely for the formation of ascomata with ascospore drops indicative of self-fertility. With a sterile needle, 2-3 ascospore droplets were then inoculated onto fresh MEA-TS and incubated for 14 days at 25°C. When no ascomata were formed, a small block of medium overgrown with mycelium was transferred with a sterile scalpel blade onto fresh MEA-TS and incubated for 14 days at 25°C. These cultures were then used to prepare sub-cultures originating from single ascospores or hyphal tips.

For self-fertile isolates, an end-point spore dilution method was used. The spore droplets were collected from the tops of two ascomata using a sterile needle and suspended in 50 µl of an isoparaffin solvent, Soltrol 130 (Chemfit, Gauteng, South Africa). The sample was diluted in a series ranging from 10^{-1} through to 10^{-4}. The undiluted sample and each dilution were spread evenly onto MEA-TS and incubated for 48 hours at 30°C, the optimum growth temperature for *T. cerberus* (Mbenoun et al., 2014). Fifty single colonies that emerged were isolated and sub-cultured on MEA-TS and monitored for ascomata formation. The segregation of self-sterile to self-fertile progeny are expected to occur at a ratio of 1:1 (Mathieson, 1952; Uhm and Fujii, 1983a, b). To evaluate if the results from the current study differed significantly from this expectation, a chi-square test with a null hypothesis of a 1:1 ratio was employed (Lee et al., 2018; Lee et al., 2015).

In cases where no ascomata were present, hyphal tip isolations were performed to produce 20 single hyphal tip cultures. For this, a small mycelial block was cut from the culture using a sterile scalpel and placed on fresh MEA-TS. After 24 hours at 30°C, a sterile needle was used to remove single hyphal strands growing into the medium with the aid of a stereo microscope. Each strand was placed onto fresh MEA-TS, incubated at 30°C and monitored for ascomata formation.

2.4. Experimental verification of the *T. cerberus MAT1* locus structure and gene content

To evaluate the predicted structure and genic content of the *MAT1* locus, an approach based on conventional PCR and Sanger sequencing was used. Genomic DNA was extracted from *T. cerberus* isolates CMW36641 and CMW36653, as well as from all subcultures prepared from single ascospores and hyphal tips. For this purpose, the hexadecyltrimethyl ammonium bromide (CTAB) extraction protocol modified from the one described by Damm *et al*. (2008) was used. Briefly, this entailed harvesting of mycelium and disrupting the cell walls with glass beads in CTAB (0.2 M Tris, 1.4 M NaCl, 20 mM EDTA, 0.2 g/L CTAB). Following incubation at 100°C for 3 min and 10 mins on ice, the supernatant was collected via centrifugation (5 min; 18 000 rcf) and subjected to chloroform: isoamylalcohol (24:1) extraction. Nucleic acids in the aqueous phase were then precipitated using 0.5 volumes of 7.5 M ammonium acetate and 2 volumes of ice-cold isopropanol and incubated at -20°C overnight. Precipitated nucleic acids were then pelleted via centrifugation (18 000 rcf at 4°C for 10 min), washed with 70% ethanol and resuspended in double distilled water. DNA quality and concentration were estimated using a NanoDrop spectrophotometer (ThermoFisher Scientific, Wilmington, USA).

To confirm the predicted structure of the *MAT1* locus, PCR primers were designed using Primer3Plus (Rozen and Skaletsky, 1999; Untergasser et al., 2007). These primers (supplementary table 1) were used in PCR reactions with the extracted DNA from both isolate CMW36653 and CMW36641 as template. Each 25 µl volume reaction included 1 U of KAPA *Taq* DNA polymerase (Kapa Biosystems, USA), 1 X KAPA Taq Buffer A, 0.4 mM of each primer, 0.25 mM of each of the dNTPs and 90-150 ng of template DNA. The PCR cycling conditions were as follows: initial denaturation was carried out for 3 min at 96°C, followed by 30 cycles of 30 sec at 96°C, 55°C for 30 sec and then 30 sec at 72°C, and a final extension step at 72°C for 7 min. The PCR products were mixed with GelRed (Biotium, Hayward, US) and visualised on 1% agarose (LE Agarose, SeaKem, USA) by gel electrophoresis.

To map the PCR products back to the annotated *MAT1* locus predicted from the assembled genome, representative PCR products from each gene were subjected to Sanger sequencing. The selected PCR products were purified using Sephadex G50 Columns (Sigma-Aldrich, USA). Amplicons were then sequenced in both directions using the original PCR primers, the BigDye Terminator Cycle Sequencing Kit v3.1 (Life Technologies, USA) and an ABI PRISM 3300 Genetic Analyser (Applied Biosystems, USA). Chromatograms were analysed in CLC Main Workbench and their sequences were compared to those of the sequenced *T. cerberus* genome using nucleotide BLAST searches.

To confirm the predicted *MAT1* gene models, RNA was extracted from isolate CMW36653 and CMW36641 using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). The RNase-free DNase kit (Qiagen, Hilden, Germany) was used to eliminate residual DNA by applying it to an on-column DNA digestion at room temperature for 15 min. A total of 2.5 µg of the extracted RNA, quantified using a NanoDrop spectrophotometer, was used as template in a cDNA synthesis reaction with the RevertAid H Minus First Strand cDNA Synthesis kit (ThermoFisher Scientific, Waltham, USA). The reaction was performed in a thermocycler set at 42°C for 60 min, and it was terminated by incubation at 70°C for 5 min. The synthesized cDNA was then used in PCR and sequencing reactions as described above, using primers specifically designed to target the coding regions of the predicted genes (supplementary table 1). Chromatograms were analysed using CLC Main Workbench to determine the intron-exon boundaries for all the genes.

A multiplex PCR was designed for rapid screening of the *MAT1* locus structure of the sub-cultures prepared from single ascospores or individual hyphal strands. These were also conducted in 25 µl reaction volumes that included 1 U of KAPA *Taq* DNA polymerase, 1 X KAPA *Taq* Buffer A, 0.25 mM of each dNTP, 90-150 ng of respective DNA templates, 0.2-0.4 µM of each primer. For amplification, an initial denaturation was carried out for 3 min at 96°C, followed by 30 cycles of 30 sec at 96°C, 55°C for 30 sec and then 30 sec at 72°C. This was

followed by a final extension step at 72°C for 7 min. The PCR products were visualised as described above.

3. Results

3.1. Genome sequence of *Thielaviopsis cerberus* isolate CMW36653

A total of 3 024 097 single-reads with an average length of 274.362 bases were produced. These were assembled into a draft genome sequence of 29,1 Mbp contained in 9 307 contigs. Of these, 4 089 contigs were larger than 1000 bp, with the largest being 58 730 bp in length. The genome had a GC-content of 49.7%, N50 and L50 of 10 338 bp and 821 contigs, respectively, and 255.89 ambiguous nucleotides per 100 000 bp. AUGUSTUS predicted that the genome contained 8 308 protein coding genes. The results of the BUSCO analysis showed that the genome had a high level of completeness (i.e., 80% using the Ascomycota dataset and 88.6% using the Fungi dataset). The genome was deposited as a Whole Genome Shotgun project to DDBJ/ENA/GenBank under accession JACYXV000000000. The version described in this paper is version JACYXV010000000.

3.2. Identification and annotation of the *MAT1* locus of *T. cerberus*

With the aid of tBLASTn searches, we identified mating-type genes on a single fragment of contiguous sequence in the *T. cerberus* genome assembly (i.e., contig 1 379). The *T. punctulata* mating-type genes showed BLAST identity to this contig with E-values ranging from 4.56E-112 to 2.91E-26. Analysis with AUGUSTUS and FGENESH+ showed that it encoded three complete genes and two partial gene sequences. The inferred *T. cerberus* MAT1-1-2 MAT1-2-1 and MAT1-2-7 proteins, as well as two different portions of MAT1-1-1, were considered as homologous to the published *Thielaviopsis* mating-type proteins based on a reciprocal best-hit BLASTp search, confirming the identities of the different mating-type genes.

InterPro domain analyses found no conserved domains in the *MAT1-2-7* gene, while the *MAT1-1-2* gene encoded the known MAT1-1-2/MatA-2/Smr1 domain (IPR031472; PF17043). The *MAT1-2-1* gene showed three signatures for the HMG_box domain (IPR009071;

PF00505). However, a characteristic of this gene in the Pezizomycotina is the presence of a conserved serine residue in the HMG domain whose codon is split by an intron (Wilken et al., 2017), which was also the case for the *MAT1-2-1* gene of *T. cerberus*. For the two fragments of the *MAT1-1-1* gene, the piece upstream of *MAT1-1-2* encoded an identifiable MATalpha_HMGbox domain (IPR006856; PF04769) across a predicted intron, although no stop codon was present. No conserved domain was predicted from the second *MAT1-1-1* fragment. These predicted *T. cerberus* genes also corresponded in terms of size to those known from the *MAT1* locus of Ceratocystidaceae species (Wilken et al., 2018; Wilken et al., 2014; Wilson et al., 2015a). The sizes of the predicted *MAT1-1-2*, *MAT1-2-1* and *MAT1-2-7* genes were 1 231 bp, 890 bp and 651 bp, respectively, while the two fragments of the *MAT1-1-1* gene were 962 bp and 513 bp respectively.

The sequenced cDNA was used to confirm the intron-exon boundaries in the gene models predicted from the genome sequence. The cDNA was used to ensure that the predicted gene models had the correct intron-exon boundaries. In terms of the overall architecture of the genes on the contig containing the *MAT1* genes, *MAT1-1-2*, *MAT1-2-1* and *MAT1-2-7* were located together (and in this order). The group was flanked on the one side by the 962 bp fragment of *MAT1-1-1* and on the other side by the 513 bp fragment of *MAT1-1-1*. Additionally, analysis with RepFind and REPuter identified a direct repeat of 114 bp in each of the *MAT1-1-1* coding regions (supplementary figure 2). REPuter calculated an output e-value of 1.94E-62 for these 114 bp direct repeats (Kurtz et al., 2001). One of these was located towards the 3' end in the 962 bp *MAT1-1-1* fragment, while the other was located towards the 5' end in the 513 bp fragment.

3.3. ***MAT1*** **locus structure in self-fertile and self-sterile individuals**

Because of structural similarities to the *MAT1* locus of fungi capable of unidirectional mating-type switching (Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017), we generated the expected "self-sterile" version of the locus *in silico* by deleting the region between the two repeats and retaining only one of the repeats. This produced a single open reading frame of

1 361 bp in length that was predicted to encode a complete copy of *MAT1-1-1*, harbouring the expected MATalpha_HMGbox (IPR006856; PF04769) domain across the direct repeat and the predicted intron (figure 1a), as expected (Wilken et al., 2017). We accordingly used PCR and sequencing to verify the gene content in self-fertile isolate CMW36653 and self-sterile isolate CMW36641 (supplementary figure 3). Both the nucleotide sequences corresponding to the self-fertile and self-sterile loci of *T. cerberus* were submitted to GenBank with accession numbers MW057848 and MW057849, respectively.

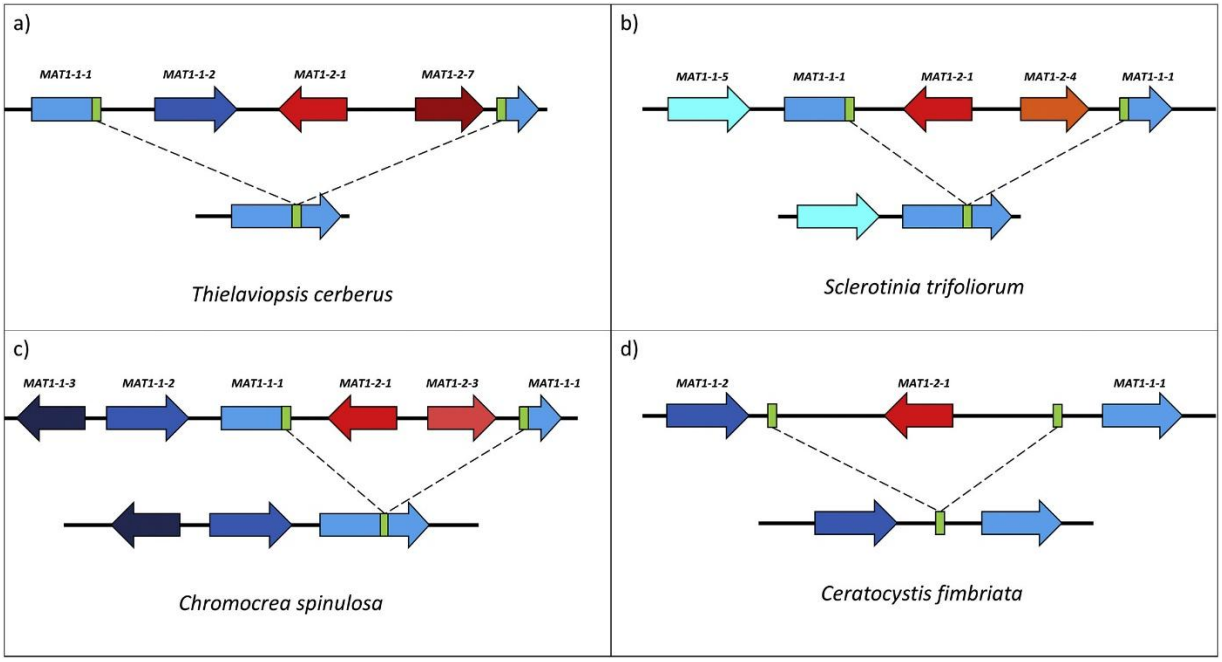


Fig. 1. Models of the mating-type loci of four Sordariomycetes that utilise unidirectional mating-type switching. a-d) The self-fertile loci (top) and self-sterile loci (bottom) of *T. cerberus*, *Sclerotinia trifoliorum*, *Chromocrea spinulosa* and *Ceratocystis fimbriata,* respectively. The *MAT1-1* information is in shades of blue and the *MAT1-2* information is in shades of red in the self-fertile versions of the *MAT1* loci. The *MAT1-2* information is located within the *MAT1-1* information and is flanked by two direct repeats (shown by green rectangles) and results in a loss of the *MAT1-2* information in the self-sterile version of the loci. The dotted lines show the loss of the intervening DNA region between the direct repeats, resulting in a single repeat left. In a, b and c, the self-fertile loci contain two fragments of the *MAT1-1-1* gene, and the self-sterile loci contain a single *MAT1-1-1* gene consisting of the joined fragments present in the self-fertile locus.

Within isolate CMW36653, the expected fragments for the *MAT1-1-2*, *MAT1-2-1* and *MAT1-2-7* genes were amplified, as well as for a truncated portion of *MAT1-1-1*. By contrast, a fragment corresponding to the complete *MAT1-1-1* gene was amplified only for isolate CMW36641, and with the primer sets utilized, none of the other genes could be amplified from it. Similar results were found when the PCR analysis was repeated using reverse transcription PCR (RT-PCR). Most strikingly, this allowed amplification of *MAT1-1-1* (660 bp) from isolate CMW36641, while none of the other three genes were detected. In the self-fertile individual (CMW36653), RT-PCR generated amplicons of the expected size for *MAT1-1-2*, *MAT1-2-1,* and *MAT1-2-7.* An amplicon was also produced for *MAT1-1-1* from this individual, but sequencing confirmed that the fragment was generated from the self-sterile locus.

3.4. Unidirectional mating-type switching in *T. cerberus*

Comparison of the *MAT1* locus structure in isolate CMW36641 and CMW36653 suggested that self-fertility in the latter is due to unidirectional mating-type switching. We therefore tested whether the inheritance of this phenotype follows standard Mendelian genetics. This was achieved by examining the fertility of sexual progeny collected from isolate CMW36653. Of the 50 single ascospore isolates prepared for this fungus, 29 were self-fertile producing abundant ascomata and ascospores, while 21 were self-sterile (supplementary figure 4). A chi-square test showed that the ratio of self-fertile to self-sterile isolates did not deviate significantly (p = 0.2579) from the expected 1:1 ratio (Mathieson, 1952; Uhm and Fujii, 1983a, b). By comparison, isolate CMW36641 never produced any ascomata, nor did any of the 20 isolates prepared from hyphal tips of this culture.

The self-fertility and self-sterility observed in the sexual progeny of isolate CMW36653 were confirmed using a multiplex PCR designed in this study (supplementary figures 5 and 6). This method utilized one forward primer and two reverse primers. Products amplified from the forward primer and first reverse primer were expected to be 692 bp in size (supplementary figure 6a), because the latter primer binds to a region of the sequence that is lost during the switch. Products amplified from the forward primer and the second reverse primer were

expected to be 469 bp in size after the switch (supplementary figure 6b), because this third primer was designed to bind to the *MAT1-1-1* gene, downstream of the region of DNA that is lost (i.e., switching brings the primer site closer to the forward primer). Application of this multiplex PCR on DNAs for the self-fertile parental isolate (CMW36653) and all its 29 self-fertile progeny, generated both amplicons that respectively corresponded to "unswitched" and "switched" versions of the *MAT1-1-1* gene. For self-sterile isolate CMW36641 and the 21 self-sterile progeny of isolate CMW36653, only the smaller amplicon was produced, which indicated that the *MAT1* locus was restructured during the switching event to produce an intact *MAT1-1-1* gene. Also, in all the sub-cultures prepared from single hyphal strands of CMW36641, only the smaller amplicon was detected.

4. Discussion

The findings presented here showed that *T. cerberus* and its *MAT1* locus bear all of the hallmarks associated with species that achieve homothallism via unidirectional mating-type switching (Harrington and McNew, 1997; Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). Similar to what was observed previously (Mbenoun et al., 2014), some isolates of *T. cerberus* were self-fertile and capable of producing ascomata and ascospores in the absence of a mating partner. Also, about half of the isolates grown from individual ascospores retained this phenotype, while the remainder were self-sterile having lost the ability to produce ascomata. This difference in reproductive ability was reflected by the overall architecture of the *MAT1* locus of the fungus. In self-sterile cells, the *MAT1* locus was restructured via a deletion event to produce a stretch of sequence encoding only the *MAT1-1-1* gene. Prior to deletion, this gene was interrupted by a stretch of DNA encoding three genes (*MAT1-1-2*, *MAT1-2-1,* and *MAT1-2-7*). The latter three-gene version of the locus confers *MAT1-2* mating specificity to the cell, while the one-gene version confers *MAT1-1* specificity. As a consequence, sexual interactions between cells harbouring the different versions of the locus are fertile and needed for the production of meiotic offspring. Therefore, as in other fungi utilizing unidirectional mating-type switching, homothallism is achieved by changing the mating-type specificity of

certain cells through a process involving modification to the structure and gene content of the *MAT1* locus (Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017).

The gene content of the *T. cerberus MAT1* locus closely resembles that of other members of the Microascales. Among the Ceratocystidaceae studied so far, *T. cerberus* and other *Thielaviopsis* species share their complement of *MAT1* genes with *Huntiella* species (Liu et al., 2018; Wilson et al., 2015a; Wilson et al., 2020). The *Knoxdaviesia* species (family Gondwanamycetaceae) also have the same gene complement (Aylward et al., 2016). Although the *MAT1-1-1*, *MAT1-1-2* and *MAT1-2-1* genes are present across many Sordariomycetes, the *MAT1-2-7* gene appears to be unique to the Microascales (Dyer et al., 2016; Wilken et al., 2017; Wilson et al., 2020). The presence of lineage-specific mating-type genes is not uncommon (Wilken et al., 2017), and functional analysis has shown that these are often necessary for the sexual process (Rodenburg et al., 2018; Wilson et al., 2020). In *Huntiella*, the *MAT1-2-7* gene is essential for ascomatal maturation (Wilson et al., 2020), a role likely conserved across the Microascales.

Species capable of unidirectional mating-type switching show significant similarity in the organisation of their *MAT1* locus (see figure 1; Harrington and McNew, 1997; Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). The “unswitched” version of the *T. cerberus MAT1* locus is similar to these other fungi, in that its *MAT1-1* information is split by a repeat-flanked region containing *MAT1-2* information (Yun et al., 2017). The mating-type “switch” then restructures the locus, producing a *MAT1* locus encoding only *MAT1-1-1* and one of the direct repeats (Wilken et al., 2014). In the case of *T. cerberus*, the 114 bp repeats are within the *MAT1-1-1* gene, and after the “switch”, the gene is completed. Although the exact mechanism remains to be elucidated, this restructuring of the locus is believed to be facilitated by a recombination event across the direct repeats that flank the sequence splitting the *MAT1-1* information (Xu et al., 2016). Therefore, the placement of the direct repeats in the locus varies between species that undergo unidirectional mating-type switching, but they always flank the

region that is deleted and this results in the permanent loss of *MAT1-2* information from the genome (Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017).

In this study, a self-fertile to self-sterile ratio of 1:1 was obtained when using an end-point spore dilution method, which suggests that unidirectional mating-type switching is governed by Mendelian genetics. However, this ratio is not always recovered for other species from the Ceratocystidaceae that are capable of unidirectional mating-type switching. A study of isolates, previously recognised as part of the *Ceratocystis fimbriata* species complex, reported a range of self-fertile to self-sterile ratios up to a ratio of 9:1, with 3:2 being the most common (Webster and Butler, 1967). The variation can in part be attributed to a lack of spore size dimorphism, where in other species capable of this switching behaviour colonies from large spores produce self-fertile cultures while small ascospores result in self-sterile cultures (Mathieson, 1952), making the selection of self-fertile and self-sterile isolates straightforward. Additionally, in at least one member of the Ceratocystidaceae, *Ceratocystis albifundus*, self-sterility carries a fitness cost in the form of reduced growth rate (Lee et al., 2015) which in turn can also influence the unbiased selection of single-ascospores when using a germling isolation approach (Harrington and McNew, 1997; Oliveira et al., 2015). Instead, an end-point spore dilution method used here does not rely on the growth-rate of the germlings, a factor that could introduce a bias during isolation, and could be useful to future studies relying on the production of self-fertile and self-sterile progeny.

The unique structure of the mating-type loci from *Thielaviopsis* species allowed for the development of a putative model explaining the evolution of unidirectional mating-type switching in *T. cerberus* (see figure 2). The model suggests that a translocation of the *MAT1-1* specific gene, *MAT1-1-2*, occurred in the heterothallic ancestor of the Ceratocystidaceae (Wilken et al., 2018). Therefore, in heterothallic members of *Thielaviopsis* the *MAT1-1-2* gene is part of the *MAT1-2* idiomorph, a unique organization that appears to be specific to these species within the Ceratocystidaceae (Wilken et al., 2018). In *T. cerberus*, the complete *Thielaviopsis MAT1-2* idiomorph (including the *MAT1-1-2* gene) was transferred to the *MAT1-*

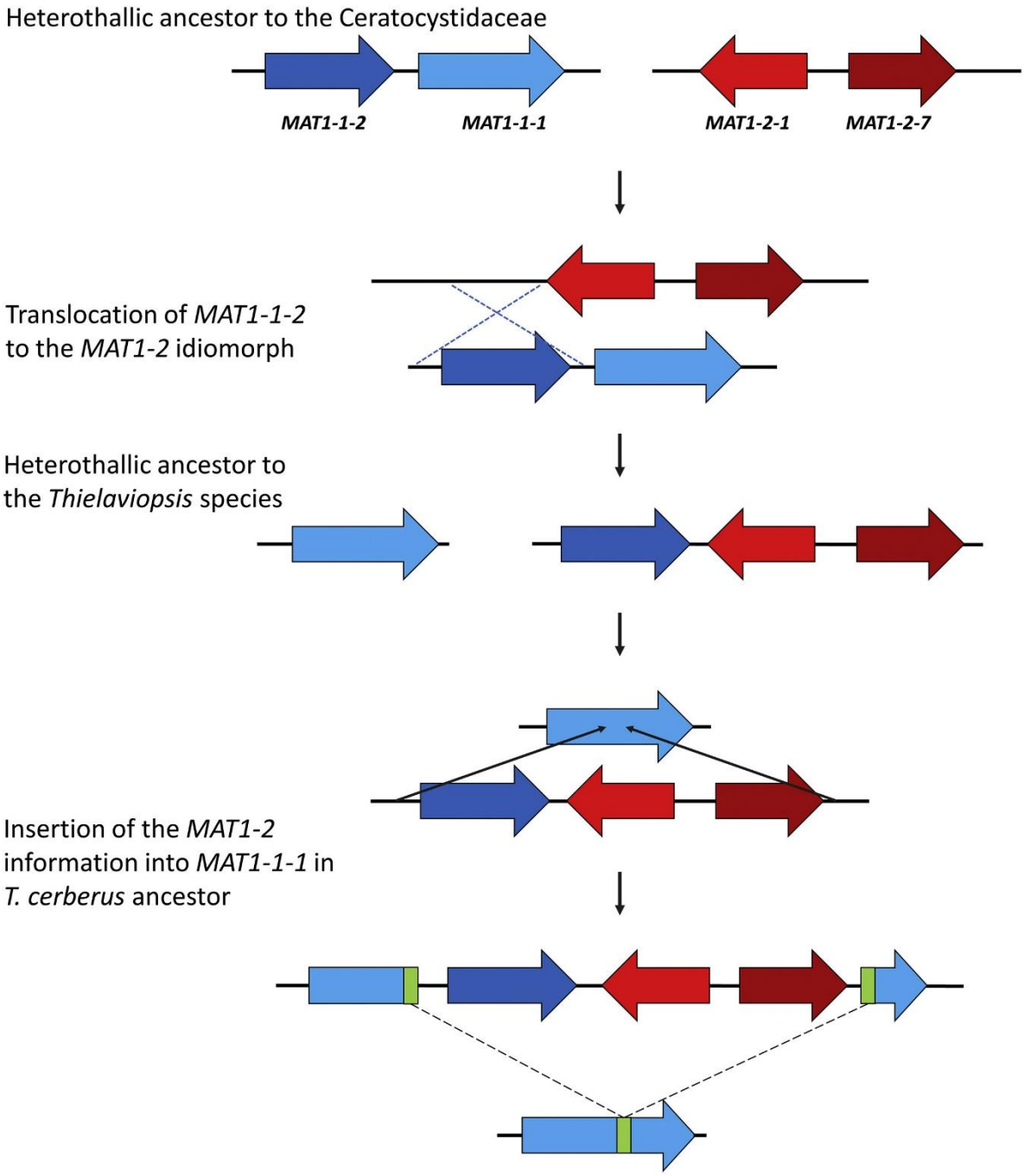


Fig. 2. A putative model for the development of unidirectional mating-type switching in *T. cerberus*. An unequal cross over event resulted in the insertion of the *MAT1-1-2* gene, typically a *MAT1-1* gene, into the *MAT1-2* idiomorph in heterothallic species of *Thielaviopsis*. In the ancestor of *T. cerberus*, this *MAT1-2* idiomorph was inserted into the *MAT1-1-1* gene, splitting it in two.

1-1 gene splitting it into two fragments that effectively disrupted the coding sequence, possibly through a recombination or insertion event. This resulted in a homothallic *T. cerberus* capable of producing a self-fertile and a self-sterile locus through unidirectional mating-type switching. The order and orientation of the genes in *T. cerberus* is identical to that of the heterothallic species, supporting a model where movement of the *MAT1-1-2* gene into the *MAT1-2* idiomorph preceded the insertion event that shaped the *T. cerberus MAT1* region. These results indicate that unidirectional mating-type switching could be derived from heterothallism through the insertion of the *MAT1-2* idiomorph into the opposite mating type and appears to be both required and sufficient to establish unidirectional mating-type switching as the reproductive strategy.

Finally, the genome sequence information generated in this study was complemented by mRNA-based gene models and the development of a multiplex PCR for diagnosing the fertility status of *T. cerberus* isolates. Through cDNA amplification and sequencing, all predicted gene models were confirmed. We thus provided the first set of verified mating-type genes for *Thielaviopsis* species, and in any species within the Ceratocystidaceae that makes use of unidirectional mating-type switching. Such gene models are crucial for the accurate description of the *MAT1* locus (Wilken et al., 2017), and will be useful to guide annotations in future projects. Furthermore, the development of a multiplex PCR that could distinguish between the self-fertile and self-sterile versions of the *MAT1* locus proved valuable. It allowed us to screen many isolates conveniently, saving on time and cost as compared to gene-specific PCRs. Multiplex PCRs have been used in studies on mating identity in the past, with much success (Dyer et al., 2001; Steenkamp et al., 2000). It is possible that primers developed for mating-type genes in a specific species could be used across a range of species (Li et al., 2020). It may be useful to test the primers developed here for *T. cerberus* in other species as well, as mating-type gene content is very conserved across species (Li et al., 2020; Nel et al., 2018; Wilken et al., 2014; Wilson et al., 2018).

Conflict of interest

The authors declare no conflicts of interest.

Acknowledgements

This project was funded by the University of Pretoria, as well as the Department of Science and Technology (DST)/ National Research Foundation (NRF) Centre of Excellence in Plant Healthy Biotechnology and the South African Research Chairs Initiative (SARChI) in Fungal Genomics. 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

Aanen, D.K., Hoekstra, R.F., 2007. Why sex is good: On fungi and beyond. In: Heitman, J., Kronstad, J.W., Taylor, J.W., Casselton, L.A. (eds), *Sex in fungi: Molecular determination and evolutionary implications*. ASM Press, Washington, DC, pp. 527-534.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal of Molecular Biology 215, 403-410.

Aylward, J., Steenkamp, E.T., Dreyer, L.L., Roets, F., Wingfield, M.J., Wingfield, B.D., 2016. Genetic basis for high population diversity in *Protea*-associated *Knoxdaviesia*. Fungal Genetics and Biology 96, 47-57.

Baker, C.J., Harrington, T.C., Krauss, U., Alfenas, A.C., 2003. Genetic variability and host specialization in the Latin American clade of *Ceratocystis fimbriata*. Phytopathology 93, 1274-1284.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A.V., Sirotkin, A.V., Vyahhi, N., Tesler, G.,

Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19, 455-477.

Betley, J.N., Frith, M.C., Graber, J.H., Choo, S., Deshler, J.O., 2002. A ubiquitous and conserved signal for RNA localization in chordates. Current Biology 12, 1756-1761.

Blakeslee, A.F., 1904. Zygospore formation a sexual process. Science 19, 864-866.

Blum, M., Chang, H.Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., Nuka, G., Paysan-Lafosse, T., Qureshi, M., Raj, S., Richardson, L., Salazar, G.A., Williams, L., Bork, P., Bridge, A., Gough, J., Haft, D.H., Letunic, I., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Orengo, C.A., Pandurangan, A.P., Rivoire, C., Sigrist, C.J.A., Sillitoe, I., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Wu, C.H., Bateman, A., Finn, R.D., 2020. The InterPro protein families and domains database: 20 years on. Nucleic Acids Research, DOI: 10.1093/nar/gkaa977.

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

Damm, U., Mostert, L., Crous, P.W., Fourie, P.H., 2008. Novel *Phaeoacremonium* species associated with necrotic wood of *Prunus* trees. Persoonia 20, 87-102.

De Beer, Z.W., Duong, T.A., Barnes, I., Wingfield, B.D., Wingfield, M.J., 2014. Redefining *Ceratocystis* and allied genera. Studies in Mycology 79, 187-219.

Dyer, P.S., Furneaux, P.A., Douhan, G., Murray, T.D., 2001. A multiplex PCR test for determination of mating type applied to the plant pathogens *Tapesia yallundae* and *Tapesia acuformis*. Fungal Genetics and Biology 33, 173-180.

Dyer, P.S., Inderbitzin, P., Debuchy, R., 2016. Mating-type structure, function, regulation and evolution in Pezizomycotina. In: Wendland, J. (Ed.), *The Mycota: Growth, Differentiation and Sexuality,* 3rd edn. Springer International Publishing, Switzerland, pp. 351-385.

Glass, N.L., Grotelueschen, J., Metzenberg, R.L., 1990. *Neurospora crassa* A mating-type region. Proceedings of the National Academy of Sciences of the United States of America 87, 4912-4916.

Glass, N.L., Nelson, M.A., 1994. Mating-type genes in mycelial ascomycetes. In: Wessels, J.G.H., Meinhardt, F. (Eds.), *The Mycota. I. Growth, Differentiation and Sexuality,* 1st edn. Springer-Verlag, Berlin, Germany, pp. 295-306.

Haber, J.E., 2012. Mating-type genes and *MAT* switching in *Saccharomyces cerevisiae*. Genetics 191, 33-64.

Harrington, T.C., McNew, D.L., 1997. Self-fertility and uni-directional mating-type switching in *Ceratocystis coerulescens*, a filamentous ascomycete. Current Genetics 32, 52-59.

Harrington, T.C., Steimel, J., Kile, G., 1998. Genetic variation in three *Ceratocystis* species with outcrossing, selfing and asexual reproductive strategies. European Journal of Forest Pathology 28, 217-226.

Heitman, J., Sun, S., James, T.Y., 2013. Evolution of fungal sexual reproduction. Mycologia 105, 1-27.

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

Keller, O., Kollmar, M., Stanke, M., Waack, S., 2011. A novel hybrid gene prediction method employing protein multiple sequence alignments. Bioinformatics 27, 757-763.

Kurtz, S., Choudhuri, J.V., Ohlebusch, E., Schleiermacher, C., Stoye, J., Giegerich, R., 2001. REPuter: The manifold applications of repeat analysis on a genomic scale. Nucleic Acids Research 29, 4633-4642.

Lee, D.-H., Roux, J., Wingfield, B.D., Wingfield, M.J., 2018. Non-Mendelian segregation influences the infection biology and genetic structure of the African tree pathogen *Ceratocystis albifundus*. Fungal Biology 122, 222-230.

Lee, D.H., Roux, J., Wingfield, B.D., Wingfield, M.J., 2015. Variation in growth rates and aggressiveness of naturally occurring self-fertile and self-sterile isolates of the wilt pathogen *Ceratocystis albifundus*. Plant Pathology 64, 1103-1109.

Li, J.Q., Wingfield, B.D., Wingfield, M.J., Barnes, I., Fourie, A., Crous, P.W., Chen, S.F., 2020. Mating genes in *Calonectria* and evidence for a heterothallic ancestral state. Persoonia 45, 163-176.

Lin, X., Heitman, J., 2007. Mechanisms of homothallism in fungi and transitions between heterothallism and homothallism. In: Heitman, J., Kroonstad, J.W., Taylor, J.W. (Eds.), *Sex in Fungi: Molecular Determination and Evolutionary Implications*. ASM Press, Washington, DC, pp. 35-57.

Liu, F., Li, G., Roux, J., Barnes, I., Wilson, A.M., Wingfield, M.J., Chen, S., 2018. Nine novel species of *Huntiella* from southern China with three distinct mating strategies and variable levels of pathogenicity. Mycologia 110, 1145-1171.

Mathieson, M.J., 1952. Ascospore dimorphism and mating type in *Chromocrea spinulosa* (Fuckel) Petch n. comb. Annals of Botany 16, 449-468.

Mbenoun, M., de Beer, Z.W., Wingfield, M.J., Wingfield, B.D., Roux, J., 2014. Reconsidering species boundaries in the *Ceratocystis paradoxa* complex, including a new species from oil palm and cacao in Cameroon. Mycologia 106, 757-784.

Metzenberg, R.L., Glass, N.L., 1990. Mating type and mating strategies in *Neurospora*. BioEssays 12, 53-59.

Mikheenko, A., Prjibelski, A., Saveliev, V., Antipov, D., Gurevich, A., 2018. Versatile genome assembly evaluation with QUAST-LG. Bioinformatics 34, i142-i150.

Nasmyth, K.A., 1982. Molecular genetics of yeast mating type. Annual Review of Genetics 16, 439-500.

Nel, W.J., Duong, T.A., Wingfield, M.J., Wingfield, B.D., Hammerbacher, A., De Beer, Z.W., 2018. Heterothallism revealed in the root rot fungi *Berkeleyomyces basicola* and *B. rouxiae*. Fungal Biology 122, 1031-1040.

Nelson, M.A., 1996. Mating systems in ascomycetes: A romp in the sac. Trends in Genetics 12, 69-74.

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

Nurk, S., Bankevich, A., Antipov, D., Gurevich, A., Korobeynikov, A., Lapidus, A., Prjibelsky, A., Pyshkin, A., Sirotkin, A., Sirotkin, Y., 2013. Assembling genomes and mini-metagenomes from highly chimeric reads. Lecture Notes in Computer Science 7821, 158-170.

Oliveira, L.S.S., Guimarães, L.M.S., Ferreira, M.A., Nunes, A.S., Pimenta, L.V.A., Alfenas, A.C., 2015. Aggressiveness, cultural characteristics and genetic variation of *Ceratocystis fimbriata* on *Eucalyptus* spp. Forest Pathology 45, 505-514.

Perkins, D.D., 1987. Mating-type switching in filamentous ascomycetes. Genetics 115, 215-216.

Plaza, G.A., Upchurch, R., Brigmon, R.L., Whitman, W.B., Ulfig, K., 2004. Rapid DNA extraction for screening soil filamentous fungi using PCR amplification. Polish Journal of Environmental Studies 13, 315-318.

Raju, N.B., 1992. Functional heterothallism resulting from homokaryotic conidia and ascospores in *Neurospora tetrasperma*. Mycological Research 96, 103-116.

Rodenburg, S.Y.A., Terhem, R.B., Veloso, J., Stassen, J.H.M., van Kan, J.A.L., 2018. Functional analysis of mating type genes and transcriptome analysis during fruiting body development of *Botrytis cinerea*. mBio 9, e01939-01917.

Rozen, S., Skaletsky, H., 1999. Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology 132, 365-386.

Sayers, E.W., Agarwala, R., Bolton, E.E., Brister, J.R., Canese, K., Clark, K., Connor, R., Fiorini, N., Funk, K., Hefferon, T., Holmes, J.B., Kim, S., Kimchi, A., Kitts, P.A., Lathrop, S., Lu, Z., Madden, T.L., Marchler-Bauer, A., Phan, L., Schneider, V.A., Schoch, C.L., Pruitt, K.D., Ostell, J., 2019. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research 47(D1), D23-D28.

Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: Assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31, 3210–3212.

Simpson, M.C., Coetzee, M.P.A., van der Nest, M.A., Wingfield, M.J., Wingfield, B.D., 2018. *Ceratocystidaceae* exhibit high levels of recombination at the mating-type (*MAT*) locus. Fungal Biology 122, 1184-1191.

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

Solovyev, V.V., 2007. Statistical approaches in Eukaryotic gene prediction. In: Balding, D., Cannings, C., Bishop, M. (Eds.), *Handbook of Statistical Genetics*, 3rd ed. John Wiley & Sons, Ltd, New York, pp. 83-127.

Stanke, M., Morgenstern, B., 2005. AUGUSTUS: A web server for gene prediction in Eukaryotes that allows user-defined constraints. Nucleic Acids Research 33, 465-467.

Stanke, M., Tzvetkova, A., Morgenstern, B., 2006. AUGUSTUS at EGASP: using EST, protein and genomic alignments for improved gene prediction in the human genome. Genome Biology 7, 1-8.

States, D.J., Gish, W., 1994. Combined use of sequence similarity and codon bias for coding region identification. Journal of Computational Biology : A Journal of Computational Molecular Cell Biology 1, 39-50.

Steenkamp, E.T., Wingfield, B.D., Coutinho, T.A., Zeller, K.A., Wingfield, M.J., Marasas, W.F., Leslie, J.F., 2000. PCR-based identification of *MAT-1* and *MAT-2* in the *Gibberella fujikuroi* species complex. Applied and Environmental Microbiology 66, 4378-4382.

Turgeon, B.G., Yoder, O.C., 2000. Proposed nomenclature for mating type genes of filamentous Ascomycetes. Fungal Genetics and Biology 31, 1-5.

Uhm, J.Y., Fujii, H., 1983a. Ascospore dimorphism in *Sclerotinia trifoliorum* and cultural characters of strains from different-sized spores. Phytopathology 73, 565-569.

Uhm, J.Y., Fujii, H., 1983b. Heterothallism and mating type mutation in *Sclerotinia trifoliorum*. Phytopathology 73, 569-572.

Untergasser, A., Nijveen, H., Rao, X., Bisseling, T., Geurts, R., Leunissen, J.A.M., 2007. Primer3Plus, an enhanced web interface to Primer3. Nucleic Acids Research 35, W71-W74.

Webster, R.K., Butler, E.E., 1967. The origin of self-sterile, cross-fertile strains and culture sterility in *Ceratocystis fimbriata*. Mycologia 59, 212-221.

Whitehouse, H.L.K., 1949. Heterothallism and sex in the fungi. Biological Reviews 24, 411-447.

Wilken, P.M., Steenkamp, E.T., van der Nest, M.A., Wingfield, M.J., De Beer, Z.W., Wingfield, B.D., 2018. Unexpected placement of the *MAT1-1-2* gene in the *MAT1-2* idiomorph of *Thielaviopsis*. Fungal Genetics and Biology 113, 32-41.

Wilken, P.M., Steenkamp, E.T., Wingfield, M.J., De Beer, Z.W., Wingfield, B.D., 2014. DNA loss at the *Ceratocystis fimbriata* mating locus results in self-sterility. PLoS One 9, e92180.

Wilken, P.M., Steenkamp, E.T., Wingfield, M.J., De Beer, Z.W., Wingfield, B.D., 2017. Which *MAT* gene? Pezizomycotina (Ascomycota) mating-type gene nomenclature reconsidered. Fungal Biology Reviews 31, 199-211.

Wilson, A.M., Godlonton, T., van der Nest, M.A., Wilken, P.M., Wingfield, M.J., Wingfield, B.D., 2015a. Unisexual reproduction in *Huntiella moniliformis*. Fungal Genetics and Biology 80, 1-9.

Wilson, A.M., van der Nest, M.A., Wilken, P.M., Wingfield, M.J., Wingfield, B.D., 2018. Pheromone expression reveals putative mechanism of unisexuality in a saprobic ascomycete fungus. PLoS One 13, e0192517.

Wilson, A.M., Wilken, P.M., van der Nest, M.A., Steenkamp, E.T., Wingfield, M.J., Wingfield, B.D., 2015b. Homothallism: An umbrella term for describing diverse sexual behaviours. IMA Fungus 6, 207-214.

Wilson, A.M., Wilken, P.M., van der Nest, M.A., Wingfield, M.J., Wingfield, B.D., 2020. The novel *Huntiella omanensis* mating gene, *MAT1-2-7*, is essential for ascomatal maturation. Fungal Genetics and Biology 137, 103335.

Witthuhn, R.C., Harrington, T.C., Wingfield, B.D., Steimel, J.P., Wingfield, M.J., 2000. Deletion of the *MAT- 2* mating-type gene during uni-directional mating-type switching in *Ceratocystis*. Current Genetics 38, 48-52.

Xu, L., Jardini, T.M., Chen, W., 2016. Direct repeat-mediated DNA deletion of the mating type *MAT1-2* genes results in unidirectional mating type switching in *Sclerotinia trifoliorum*. Scientific Reports 6, 27083.

Yoder, O., Valent, B., Chumley, F., 1986. Genetic nomenclature and practice for plant pathogenic fungi. Phytopathology 76, 383-385.

Yun, S.-H., Arie, T., Kaneko, I., Yoder, O.C., Turgeon, B.G., 2000. Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella/Fusarium* species. Fungal Genetics and Biology 31, 7-20.

Yun, S.-H., Berbee, M.L., Yoder, O.C., Turgeon, B.G., 1999. Evolution of the fungal self-fertile reproductive life style from self-sterile ancestors. Proceedings of the National Academy of Sciences of the United States of America 96, 5592-5597.

Yun, S.H., Kim, H.K., Lee, T., Turgeon, B.G., 2017. Self-fertility in *Chromocrea spinulosa* is a consequence of direct repeat-mediated loss of *MAT1-2*, subsequent imbalance of nuclei differing in mating type, and recognition between unlike nuclei in a common cytoplasm. PLoS Genetics 13, e1006981.

Supplementary information

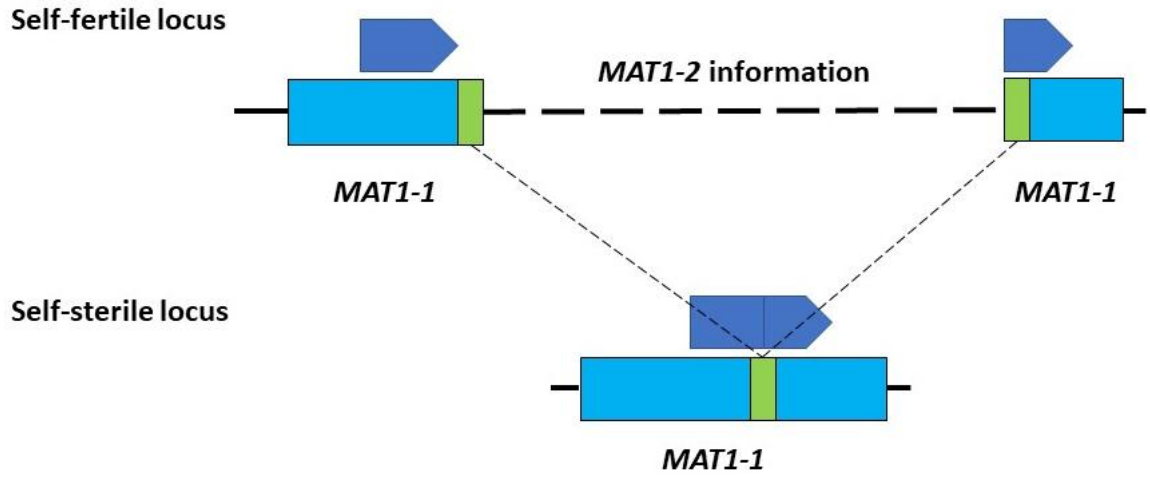
Supplementary table 1: Primers designed in this study to amplify fragments of mating-type genes in *T. cerberus*

Primer name	Primer sequence (5' to 3')	*MAT1* gene targeted
T_cer_111_aL	ATCGTTCCTCTCCTTTCATG	*MAT1-1-1*
T_cer_111_bR	TTGATATAGCTCGGGGTTTG	
T_cer_111_cR	ATTGTAAATGCCCGGAAGAT	
T_cer_111L	AGTTTAAGGGGGAAACTGTC	
T_cer_111R	CGGTAGTACGCTGATAAAGT	
T_cer_c111L	GGAGCCTGAAGACAACGAGT	
T_cer_c111R	CATGAGGAGCGACGAGGAC	
T_cer_c111/112L	GGTCATTTGTTAGCCACACT	
T_cer_c111/112R	ATCGACAATGAAAGGTATGG	
T_cer_112L	TCTGTTGCGGAAAAACTTTC	*MAT1-1-2*
T_cer_112R	TCTGTTGCGGAAAAACTTTC	
T_cer_c112R1	CTTACAGCTTGACCGATACC	
T_cer_c112R2	ATGCCATGATGGTCTCTATG	
T_cer_121L	AAAATTGCGAACTCGTGTTT	*MAT1-2-1*
T_cer_121R	CAGTTCCTGGCTATGATACC	
T_cer_c121L	TCCTCCGAAATTCTTCTGTA	
T_cer_c121R	ACACCAGGTGATGGATATTG	
T_cer_127L	GCTACTAGCTCTTGTTACCC	*MAT1-2-7*
T_cer_127R	AATCCTGTCGATAAATGGCA	

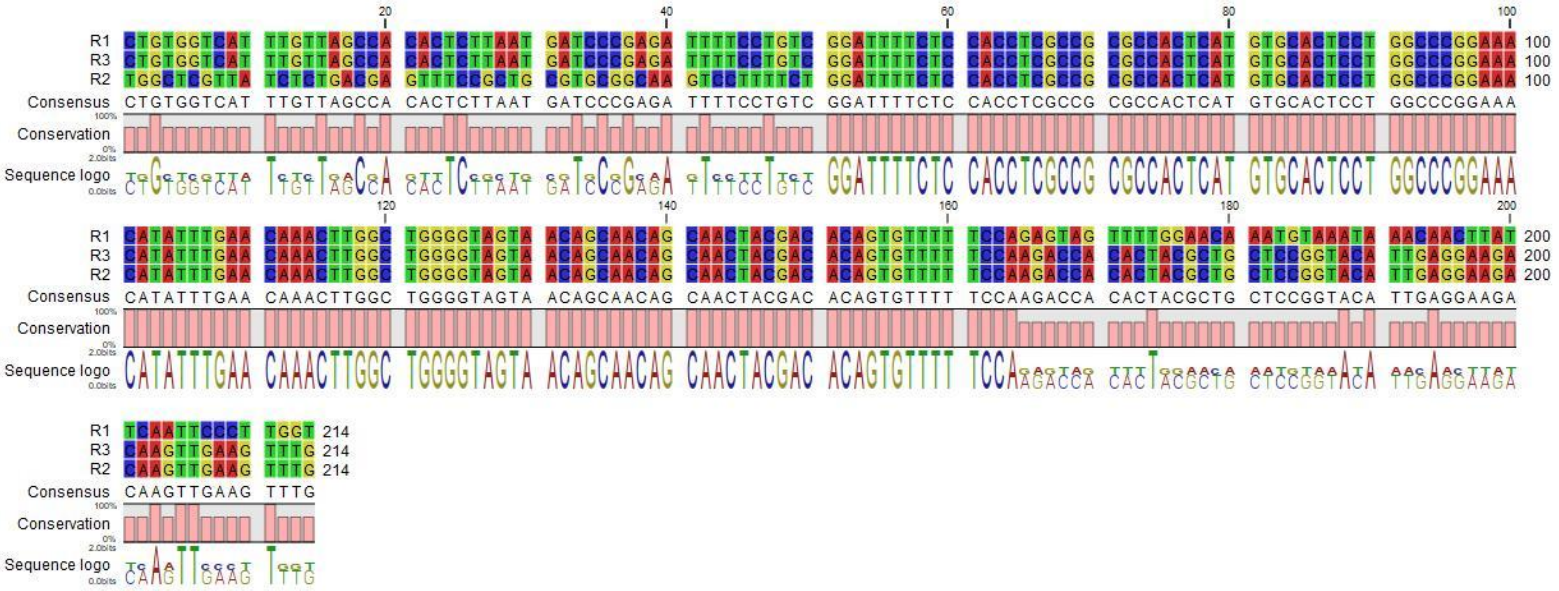
Supplementary table 2: The heterothallic *Thielaviopsis* species mating genes used in this study with their Genbank accession numbers

***Thielaviopsis* species**	**Accession number***	**Mating gene**
T. paradoxa	MF476807	*MAT1-1-2*
		MAT1-2-1
		MAT1-2-7
	BK010319	*MAT1-1-1*
	BK010320	*MAT1-1-1*
T. punctulata	KX989056	*MAT1-1-1*
	BK010318	*MAT1-1-2*
		MAT1-2-1
		MAT1-2-7
T. euricoi	MF476808	*MAT1-1-1*
T. musarum	BK010321	*MAT1-1-1*

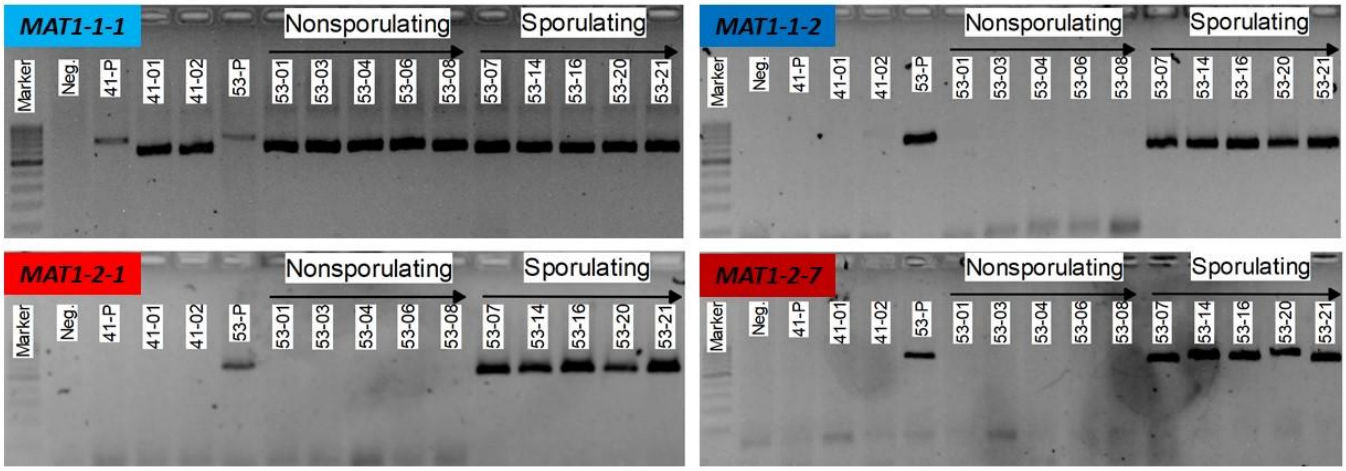
* accession numbers taken from (Wilken et al., 2017).



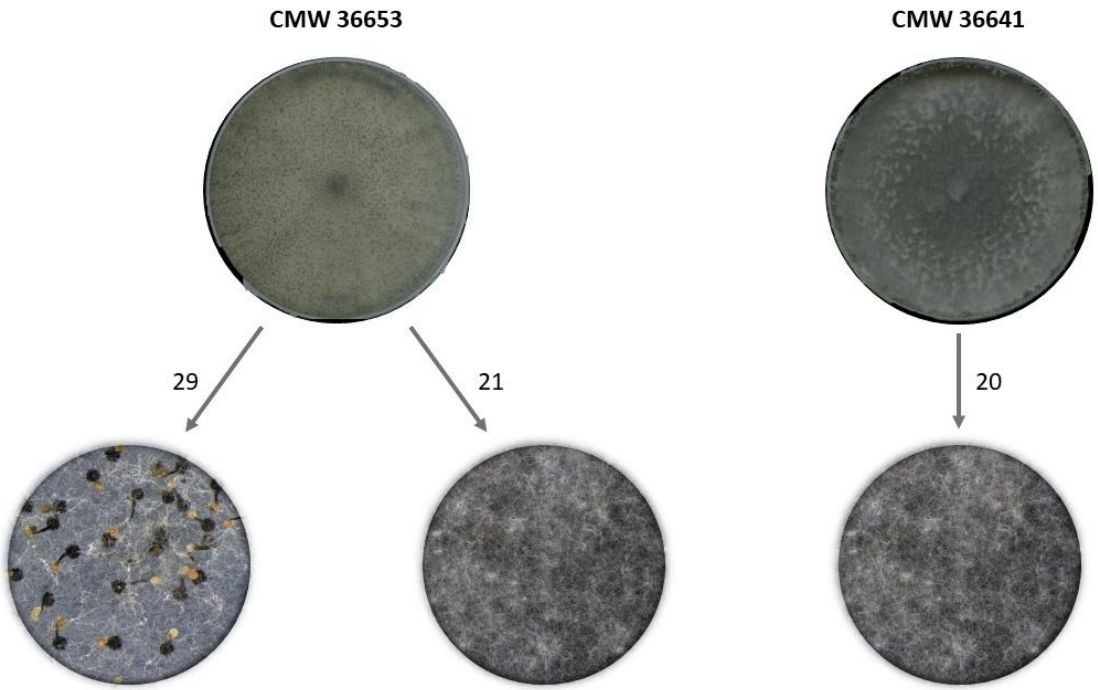
Supplementary figure 1: A model showing the fragments of the *MAT1-1-1* gene in the self-fertile locus of *Thielaviopsis cerberus*. The dark blue pentagon shapes indicate the conserved MATalpha_HMGbox domain in the self-fertile and self-sterile loci. In the self-fertile locus, the *MAT1-2* information interrupts the *MAT1-1-1* gene and associated MATalpha_HMGbox domain as shown. After the switch (dotted black lines), the fragments of the *MAT1-1-1* gene is brought together to complete the MATalpha_HMGbox.



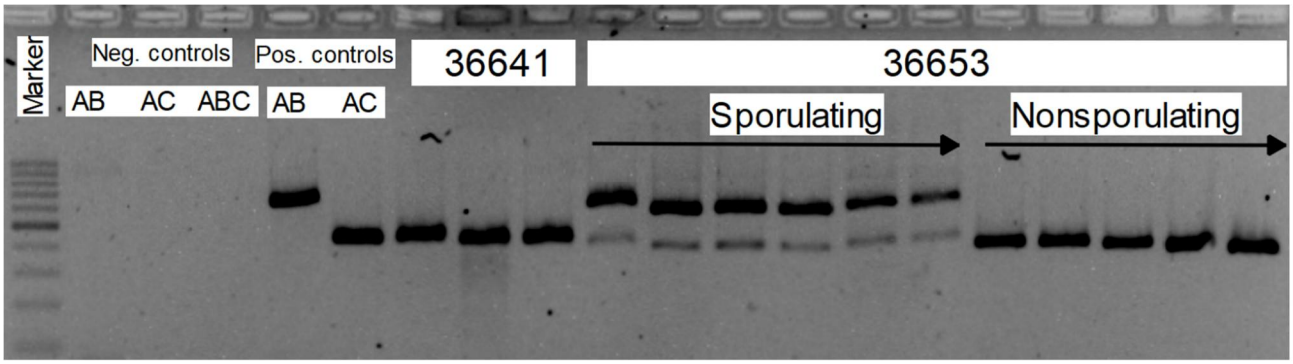
Supplementary figure 2: A sequence alignment of the 114 bp direct repeats in *Thielaviopsis cerberus*. R1 is the direct repeat from the portion of the *MAT1-1-1* gene containing the start codon, while sequence R2 is from the other portion of the *MAT1-1-1* gene in the self-fertile version of the locus. Sequence R3 is the single direct repeat in the self-sterile version of the locus, from the single *MAT1-1-1* gene. The 114 bp repeats have been included as well as 50 bp upstream and downstream of each sequence.



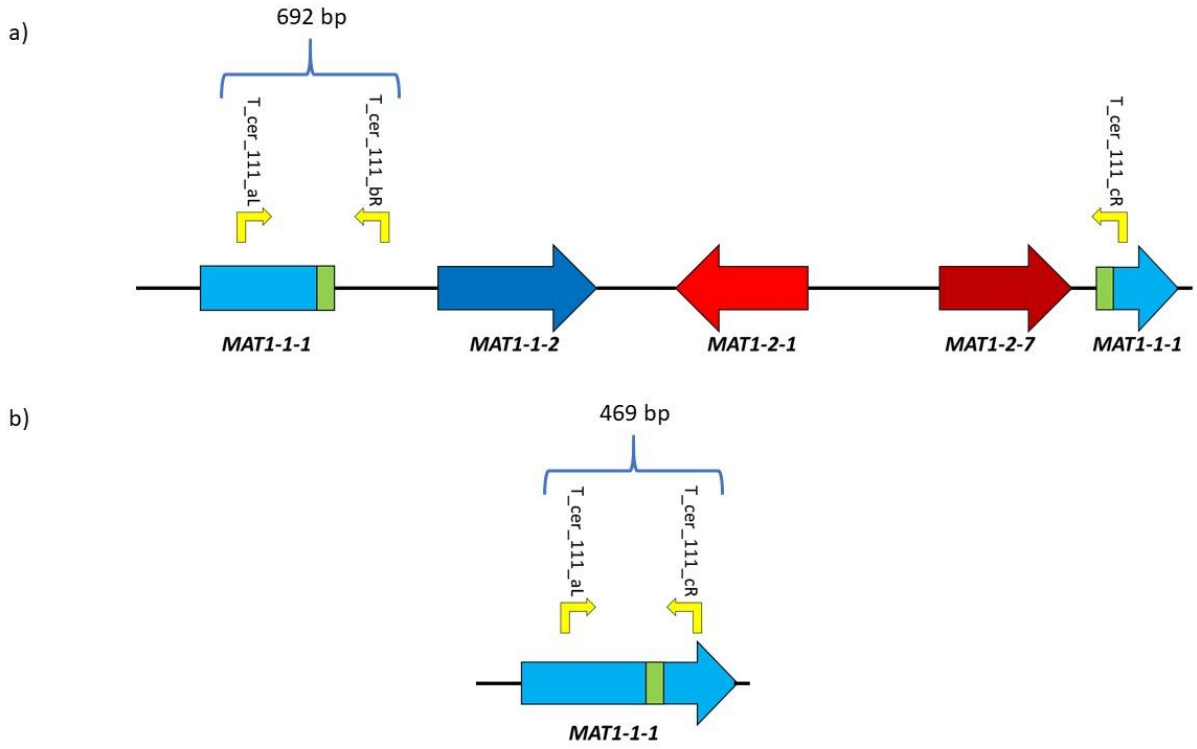
Supplementary figure 3: The agarose gels of the PCR products with gene-specific primers on *Thielaviopsis cerberus* cultures and subcultures. PCR products of representatives of 36641, as well as of nonsporulating and sporulating isolates of 36653 were randomly selected for this figure. The first lane is a 100 bp molecular marker. The second lane is the negative control of each primer set. Lanes that start with a label of 41 are from isolate 36641. None of these were sporulating and so only three isolates are shown (the original isolate, 41-P, and two subcultures). All cultures of 36641 contain only the *MAT1-1-1* gene and none of the other three genes. Lanes starting with a label of 53 are from isolate 36653, and the original sporulating isolate is indicated in lane 53-P. Progeny produced from this isolate were either sporulating or non-sporulating. Five representative progeny showing each phenotype are shown. The *MAT1-1-1* gene is found in every single isolate tested, regardless of sporulation capability. The *MAT1-1-2*, *MAT1-2-1* and *MAT1-2-7* genes are only present in sporulating isolates.



Supplementary figure 4: Isolates CMW36641 and CMW36653 of *T. cerberus*. The figure shows the MEA plates of CMW36653, with the original isolate sporulating and producing ascomata while CMW36641 had only mycelia and no ascomata produced. The fifty single spore isolations of CMW36653 yielded twenty-nine self-fertile progeny, with the development of ascomata, while twenty-one were self-sterile, with so ascomata like CMW36641. The twenty hyphal tip isolations of CMW36641 remained self-sterile.



Supplementary figure 5: The agarose gel of the PCR products using the multiplex primers after visualisation by electrophoresis. All isolates that were nonsporulating produced a single band (+-700 bp) which corresponded to the self-sterile mating locus. All isolates that were capable of sporulating showed two bands (+-700 and +-470), one that corresponded to the self-sterile locus and one that corresponded to the self-fertile locus (indicated by the larger band). This shows that sporulating isolates are made up of a mixed population of self-sterile and self-fertile individuals.



Supplementary figure 6: The two versions of the mating-type loci of *Thielaviopsis cerberus*. a) The self-fertile locus with primers designed to bind in three places. Primers T_cer_111_aL and T_cer_111_cR bind within the coding regions of the portions of the *MAT1-1-1* genes while the third primer, T_cer_111_bR binds to the non-coding region after the direct repeat (green rectangle). Only primers T_cer_111_aL and T_cer_111_bR can produce an amplicon of 692 bp in size. b) The self-sterile locus of *T. cerberus* after the loss of the intervening region between the direct repeats. Only T_cer_111_aL and T_cer_111_cR can bind to this version of the locus and produce and amplicon of 469 bp in size, as primer T_cer_111_bR has nowhere to bind after the switch.