Fungal Genetics and Biology 170 (2024) 103859

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

# Unidirectional mating-type switching is underpinned by a conserved *MAT1* locus architecture

P. Markus Wilken *, Frances A. Lane, Emma T. Steenkamp, Michael J. Wingfield, Brenda D. Wingfield

*Department of Biochemistry, Genetics and Microbiology, Forestry and Agricultural Biotechnology Institute (FABI), Faculty of Natural and Agricultural Sciences, University of Pretoria, Private Bag X20, Pretoria 0028, South Africa*

ARTICLE INFO

*Keywords:*
Sexual reproduction
*Ceratocystidaceae*
*Ceratocystis*
*Davidsoniella*
*Endoconidiophora*

ABSTRACT

Unidirectional mating-type switching is a form of homothallic reproduction known only in a small number of filamentous ascomycetes. Their ascospores can give rise to either self-sterile isolates that require compatible partners for subsequent sexual reproduction, or self-fertile individuals capable of completing this process in isolation. The limited studies previously conducted in these fungi suggest that the differences in mating specificity are determined by the architecture of the *MAT1* locus. In self-fertile isolates that have not undergone unidirectional mating-type switching, the locus contains both *MAT1-1* and *MAT1-2* mating-type genes, typical of primary homothallism. In the self-sterile isolates produced after a switching event, the *MAT1-2* genes are lacking from the locus, likely due to a recombination-mediated deletion of the *MAT1-2* gene information. To determine whether these arrangements of the *MAT1* locus support unidirectional mating-type switching in the *Ceratocystidaceae*, the largest known fungal assemblage capable of this reproduction strategy, a combination of genetic and genomic approaches were used. The *MAT1* locus was annotated in representative species of *Ceratocystis, Endoconidiophora,* and *Davidsoniella.* In all cases, *MAT1-2* genes interrupted the *MAT1-1–1* gene in self-fertile isolates. The *MAT1-2* genes were flanked by two copies of a direct repeat that accurately predicted the boundaries of the deletion event that would yield the *MAT1* locus of self-sterile isolates. Although the relative position of the *MAT1-2* gene region differed among species, it always disrupted the *MAT1-1–1* gene and/or its expression in the self-fertile *MAT1* locus. Following switching, this gene and/or its expression was restored in the self-sterile arrangement of the locus. This mirrors what has been reported in other species capable of unidirectional mating-type switching, providing the strongest support for a conserved *MAT1* locus structure that is associated with this process. This study contributes to our understanding of the evolution of unidirectional mating-type switching.

## 1. Introduction

Fungi in the *Ascomycota* have a diverse assemblage of sexual reproductive strategies coordinated by the genes located at the mating-type or *MAT1* locus (Dyer et al., 2016; Wilson et al., 2021b). Of these strategies, heterothallism and primary homothallism are the best known and considered as the archetypal fungal mating systems (Billiard et al., 2011; Wilson et al., 2015b). Single isolates of a heterothallic species are self-sterile and cannot produce sexual structures (ascomata) and sexual spores (ascospores) in pure culture (Ni et al., 2011; Wilson et al., 2015b). In these species, the sexual process can only be completed by two isolates of opposite mating-type that have different allelic versions or idiomorphs of the *MAT1* locus (Dyer et al., 2016; Metzenberg and Glass, 1990). As such, MAT-1 isolates harbour a *MAT1-1* idiomorph characterized by the presence of a *MAT1-1–1* gene, while the *MAT1-2–1* gene characterizes the *MAT1-2* idiomorph in MAT-2 isolates (Dyer et al., 2016; Wilken et al., 2017). Additional genes can be present on either idiomorph, but these tend to show a lineage-specific distribution (Turgeon and Debuchy, 2007; Wilson et al., 2021b). In contrast, primary homothallic species can complete the sexual cycle in the absence of another individual (Lin and Heitman, 2007). This is possible because the *MAT1* locus of these species contains the gene complement present in both idiomorphs of their heterothallic counterparts (Lin and Heitman, 2007; Wilson et al., 2015b), although in the case of unisexuality only a

* Corresponding author.
*E-mail address:* markus.wilken@fabi.up.ac.za (P.M. Wilken).

https://doi.org/10.1016/j.fgb.2023.103859
Received 9 October 2023; Received in revised form 11 December 2023; Accepted 14 December 2023
Available online 17 December 2023

single idiomorphic version is present (Roach et al., 2014; Wilson et al., 2021a; Wilson et al., 2015b).

Mating-type switching is a form of secondary homothallism, a reproductive strategy that resembles primary homothallism in appearance, but imitates heterothallism in function (Lin and Heitman, 2007; Wilson et al., 2015b). Two forms of this process have been described and are known as bidirectional and unidirectional mating-type switching (Wilson et al., 2015b). Bidirectional mating-type switching is well documented across various yeast species and is the most extensively studied instance of this distinct mating behaviour. It encompasses a controlled reorganization of the DNA at the mating-type locus, causing a change in the mating specificity of one of the daughter cells following a meiotic event (Wolfe and Butler, 2022). Since either of the two mating-types can be substituted, bidirectional switching is reversible. This contrasts with unidirectional mating-type switching where a section of the mating-type locus is permanently deleted, rendering the process irreversible (Wilken et al., 2014). Initially, this form of mating-type switching was identified in a few filamentous ascomycetes (Mathieson, 1952; Perkins, 1987; Uhm and Fujii, 1983a; Uhm and Fujii, 1983b; Wheeler, 1950) but has gained increasing attention due to advances in genome sequencing technologies and the availability of whole genome sequences (Krämer et al., 2021; Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017).

Species capable of unidirectional mating-type switching are known only in the *Pezizomycotina* and they have many conserved phenotypic and genetic characteristics (Krämer et al., 2021; Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). In all cases, two isolate types can be obtained from single ascospore isolations. The first is self-fertile isolates that are capable of producing ascomata and ascospores in isolation, a hallmark of homothallism (Dyer et al., 2016; Wilson et al., 2015b). Secondly, self-sterile isolates are produced following unidirectional mating-type switching, and these are only capable of sexual reproduction when paired with compatible partners in sexual crosses (Ferreira et al., 2010; Harrington and McNew, 1997; Yun et al., 2017). A single *MAT1* locus containing both *MAT1-1* and *MAT1-2* gene information are present in self-fertile isolates, while self-sterile isolates have an altered *MAT1* locus lacking the *MAT1-2* gene information (Krämer et al., 2021; Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). A switching event that is mediated by two copies of a direct repeat results in deletion of the *MAT1-2* gene information (Krämer et al., 2021; Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). In some cases, both self-fertile and self-sterile nuclei might be present together in a single culture or single cell (Xu et al., 2016; Yun et al., 2017).

The *Ceratocystidaceae* (*Microascales*) includes the largest number of related species known to undergo unidirectional mating-type switching. This unique behaviour has been reported in species of *Ceratocystis* and *Endoconidiophora*, as well as *Davidsoniella virescens* and *Thielaviopsis cerberus* (Andrus and Harter, 1933; Bakshi, 1951; Davidson, 1953; Harrington and McNew, 1997; Krämer et al., 2021; Lee et al., 2015; Wilken et al., 2014; Witthuhn et al., 2000). However, the switching mechanism has been extensively studied only in *Ceratocystis fimbriata* (Wilken et al., 2014) and *T. cerberus* (Krämer et al., 2021). In *C. fimbriata*, the *MAT1* locus of self-fertile isolates contains the genes *MAT1-1–1*, *MAT1-2–1*, and *MAT1-1–2* (in that order), with the *MAT1-2–1* gene flanked on either side by a 260 base pair (bp) direct repeat (Wilken et al., 2014). These motifs are thought to anchor a recombination event that deletes the *MAT1-2–1* gene, thereby producing the *MAT1* locus with a structure similar to those occurring in self-sterile isolates (Wilken et al., 2014). A similar situation likely occurs in *T. cerberus*, where the *MAT1* locus of a self-fertile isolate contains four genes (*MAT1-1–1*, *MAT1-1–2*, *MAT1-2–1,* and *MAT1-2–7*), in which three genes (*MAT1-1–2*, *MAT1-2–1* and *MAT1-2–7*) make up the *MAT1-2* gene region. *MAT1-1–2* being part of the *MAT1-2* region stems from heterothallic *Thielaviopsis* species where this gene is present only in the *MAT1-2* idiomorph as a secondary *MAT1-2* gene (Wilken et al., 2018). As such, the *MAT1-2* genes in *T. cerberus* are flanked on both sides by a 114 bp direct repeat that mediate deletion of all three of these genes during switching, leaving only the *MAT1-1–1* on the self-sterile *MAT1* locus (Krämer et al., 2021). Unidirectional mating-type switching associated deletion of the *MAT1-2–1* gene has been described in other *Ceratocystis* and *Endoconidiophora* species as well as in *D. virescens* (Baker Engelbrecht and Harrington, 2005; Harrington and McNew, 1997; Lee et al., 2015; Misse et al., 2017; Witthuhn et al., 2000), but the *MAT1* locus structure(s) and possible mechanism(s) of this deletion have not been resolved.

During the last decade, the genomes of many *Ceratocystidaceae* species have been sequenced (Molano et al., 2018; van der Nest et al., 2014a; Van der Nest et al., 2014b; Wilken et al., 2013; Wingfield et al., 2016a; Wingfield et al., 2015; Wingfield et al., 2018; Wingfield et al., 2016b). These include the genome sequences for seven species of *Ceratocystis* and two of *Endoconidiophora*, as well as the genome of *D. virescens* (Molano et al., 2018; van der Nest et al., 2014a; Van der Nest et al., 2014b; Wilken et al., 2013; Wingfield et al., 2016a; Wingfield et al., 2015; Wingfield et al., 2016b). These genome sequence data underpinned the primary aim of this study, which was to determine the structure of the *MAT1* locus in the Ceratocystidaceae species and to identify the changes that result in unidirectional mating-type switching. To do this, experimentally verified *MAT1* gene models were produced for *C. fimbriata* using RNA-seq analysis. These models were then utilised to identify and characterise the *MAT1* locus in the genomic sequences of the seven *Ceratocystis* species, two *Endoconidiophora* species, and *D. virescens*. A second objective was to determine the putative locus structure in self-sterile isolates (i.e., the structure of the locus after deletion) using *in silico* inference and the results were confirmed using conventional PCR and sequencing as well as RNA sequence analyses.

## 2. Materials and methods

### 2.1. Revision of the MAT1 gene models for Ceratocystis fimbriata

A sequence corresponding to the *MAT1* locus in a self-fertile *C. fimbriata* isolate (CMW 14799; Wilken et al., 2014) was obtained from GenBank (https://www.ncbi.nlm.nih.gov/genbank/; Benson et al., 2013) using accession number KF033902. This sequence was submitted to online versions of the AUGUSTUS (Stanke et al., 2006) and FGENESH (Solovyev et al., 2006) *de novo* gene prediction software. For AUGUSTUS, gene predictions were done using the *Fusarium graminearum* gene models, with the program parameters set to predict many alternative transcripts and any number of possibly partial genes on both strands. Similarly, the generic *Fusarium* gene models available on FGENESH were used for all predictions. All the gene models produced in this way were then used to annotate the *MAT1* locus of *C. fimbriata* using CLC Main Workbench 7.9.2 (QIAGEN Aarhus, Denmark).

The predicted gene models were confirmed using RNA-seq data for *C. fimbriata* isolate CMW 14799 obtained from the culture collection (CMW) of the Forestry and Agricultural Biotechnology Institute (FABI). The isolate was grown for ten days at 25 °C in the dark on medium overlaid with sterile cellophane. The growth medium (MEA-TS) contained 20 $g.L^{-1}$ Malt Extract (Biolab, Merck, South Africa), 20 $g.L^{-1}$ Agar (Biolab, Merck, South Africa), 100 $mg.L^{-1}$ Thiamine (SIGMA, Steinheim, Germany) and 150 $mg.L^{-1}$ Streptomycin (SIGMA, Steinheim, Germany). Mycelium was harvested from the plates, flash-frozen with liquid nitrogen, and ground to fine powder with a mortar and pestle. The ground material was used in RNA extractions with the Rneasy® Plant Mini Kit (Qiagen, Limburg, The Netherlands) following the manufacturer's instructions, apart from replacing the RLC buffer with the RLT buffer and including the optional Dnase-I (RNase-Free DNase Set, Qiagen, Limburg, The Netherlands) digestion step. The extracted RNA was enriched for mRNA using the Dynabead® mRNA purification kit (ThermoFisher Scientific, Waltham, USA) before being subjected to cDNA synthesis, library preparation, and sequencing at the Central Analytical Facilities (CAF), Stellenbosch University, South Africa, using

the Ion Proton Platform and PI$^{TM}$ Chip system (Life Technologies, Carlsbad, CA).

RNA-Seq reads were filtered for quality in CLC Genomics Workbench v8.5.1 (CLC Bio, Aarhus, Denmark) to remove reads with a Phred score of below 20 ($Q \leq 0.01$) or that were longer than 300 bases. From the remaining reads, up to two ambiguous terminal nucleotides were trimmed. The quality-filtered reads were then mapped onto the *C. fimbriata MAT1* region of isolate CMW 14799 using the "Map Reads to Reference" function in CLC Genomics Workbench v8.5.1 (CLC Bio, Aarhus, Denmark) using default parameters. The confirmed gene models were translated into putative proteins and subjected to BLASTp searches against the GenBank database as well as Pfam domain analyses using the full Pfam database (Finn et al., 2014) as implemented in CLC Main Workbench.

### 2.2. *MAT1* locus structure in self-fertile and self-sterile isolates of *Ceratocystis, Endoconidiophora, and Davidsoniella virescens*

The published genome sequences for *C. eucalypticola* (Wingfield et al., 2015), *C. manginecans* (Van der Nest et al., 2014b), *C. cacaofunesta* (Molano et al., 2018), *C. albifundus* (Van der Nest et al., 2019), *C. smalleyi* (Wingfield et al., 2018), *C. harringtonii* (Wingfield et al., 2016b), *E. polonica* (Wingfield et al., 2016a), *E. laricicola* (Wingfield et al., 2016a) and *D. virescens* (Wingfield et al., 2015), as well as the unpublished but publicly available genome sequence for *C. platani* (accession number LBBL01000000; Belbahri, 2015) were obtained from GenBank and imported into CLC Main Workbench. These genomes were screened for the presence of *MAT1* genes by making use of BLAST searches (Wilken et al., 2014). For this purpose, the inferred protein sequences for all the verified *C. fimbriata MAT1* gene models were used in tBLASTn searches against the respective genome sequences using CLC Main Workbench. Contigs with a BLAST expect-values of 0.01 or less were considered to contain *MAT1* genes.

In cases where the putative *MAT* genes produced results spread across multiple contigs, custom primers were designed with Primer3web (Koressaar and Remm, 2007; Untergasser et al., 2012) and used in PCR and sequencing reactions to reconstruct the full *MAT1* locus. To do this, representative isolates of each species were obtained from the CMW culture collection and grown on MEA-TS media at 25 °C for 10–14 days. Mycelium was harvested and used in DNA extractions based on the CTAB method of Damm et al. (2008) with modifications described by Krämer et al. (2021). Approximately 100 ng of the extracted DNA was used in 25 µl PCR reactions consisting of 1U KapaTaq DNA polymerase (Kapa Biosystems, USA), 1X KapaTaq Buffer A, 0.25 mM of each dNTP and 0.4 mM of each primer. These reactions were carried out in a GeneAmp 9700 PCR System (Applied Biosystems, Foster City, California, USA) using an initial denaturation at 95 °C for 3 min, followed by 35 cycles of 95 °C for 30 s, 50 °C for 30 s and 72 °C for 1 min, and a final extension at 72 °C for 7 min. All PCR products were subjected to agarose (1 %, w.v$^{-1}$, LE Agarose, SeaKem, USA) gel electrophoresis and visualised under UV light (Sambrook and Russell, 2001). PCR products of the expected sizes were then purified using Sephadex G50 columns (Sigma-Aldrich, USA) and sequenced in both directions using the original PCR primers, the BigDye Terminator Cycle Sequencing Kit v3.1 (Life Technologies, USA) and an ABI3500xL Genetic analyzer (Applied Biosystems, ThermoFisher Scientific, Waltham, MA, USA) at the Bioinformatics and Computational Biology Unit of the University of Pretoria. Electropherograms were curated in CLC Main Workbench and then used to complete the assemblies for the *MAT1* locus of the respective fungi.

Each of the full-length *MAT1* locus sequences were screened for the presence of repeat sequences using REPFIND (Betley et al., 2002). These analyses employed a P-value cut-off of 0.0001, a minimum repeat length of 50 bp (a maximum repeat length was not imposed) and a low complexity filter to eliminate the possible impact of nucleotide compositional biases on similarity searching (Morgulis et al., 2006). All the identified repeat motifs were manually annotated onto the respective *MAT1* sequences. Each of the repeats were also used in BLASTn searches of the GenBank database, as well as against each of the genomes using local BLAST searches in CLC Main Workbench. These repeats were also aligned with the online version of MAFFT v.7 (Katoh et al., 2002).

The "switched" version of *MAT1* locus for all the examined fungi was inferred using an *in silico* approach. To do this, the region flanked by the direct repeats, together with one of the repeats, were deleted from the sequence to mimic the changes that occur at the *MAT1* locus during unidirectional mating-type switching in *C. fimbriata* and *T. cerberus* (Krämer et al., 2021; Wilken et al., 2014). To experimentally verify the veracity of these predictions, a PCR approach as described by Wilken et al. (2014) was applied to DNAs obtained from representative isolates of these species (Table S1). In this case, two primers were designed with Primer3web to flank the putative deleted region for each species (Figure S1 and Table S2). A PCR product of defined size and sequence would then serve as proof that the locus configuration was present. For this purpose, DNA isolations, PCR, and sequencing were performed as described above. Comparisons involving the sequences for these products and those predicted for the "switched" locus were conducted in CLC Main Workbench.

As a final step, the complete sequences for the *MAT1* locus of the self-fertile isolates of *Ceratocystis*, *Endoconidiophora* and *Davidsoniella*, as well as those inferred and confirmed for self-sterile isolates were subjected to *de novo* gene predictions. This was done using AUGUSTUS and FGENESH as described above, and the results were used to annotate the respective contigs. In cases where the presence of a gene was expected from the original BLAST searches, but not predicted, the sequence was subjected to FGENESH+ (Solovyev, 2007) analyses using the generic *Fusarium* models and the corresponding *C. fimbriata MAT1* protein sequence as a guide. All predicted gene sequences were translated and subjected to BLASTp searches against the NCBI protein database and to Pfam domain analyses as described above. The annotated self-fertile and self-sterile loci for all species were submitted to the GenBank under the accession numbers OR922808 - OR922827.

### 2.3. *MAT1-1–1 gene expression from the self-sterile MAT1 locus*

An assay was designed to determine if the version of the *MAT1-1–1* gene produced after a switching event (the self-sterile locus) was expressed. To do this, a primer set for each species was designed that would detect expression from the locus in the self-sterile configuration only (Figure S2). For all *Ceratocystis* species (i.e., of *C. manginecans* CMW 17570, *C. eucalypticola* CMW 9998, *C. cacaofunesta* CMW 26375, *C. platani* 26380, *C. albifundus* CMW 4680, *C. harringtonii* CMW 14789, *C. smalleyi* CMW 14800), the first primer was positioned towards the 3′ end of the *MAT1-1–1* gene, while the second primer was targeted to the putative promoter region upstream of the start codon in the self-sterile locus. The fragment would also span a putative intron. For *D. virescens* (CMW 17339) and the two *Endoconidiophora* species (i.e., *E. polonica* CMW 20928, *E. laricicola* CMW 20930), the primer set were designed to span both the region that is deleted during switching, as well as at least one putative intron.

Total RNA was extracted from 5 to 7 day-old cultures grown on cellophane-covered MEA-TS medium as described above. Quality and quantity of the total RNA were assessed by agarose gel electrophoresis (Sambrook and Russell, 2001) and spectrophotometrically with NanoDrop ND1000 (ThermoScientific, Waltham, USA). Of the extracted RNA, ~5 µg was reverse transcribed using RevertAid H Minus First Strand cDNA Synthesis Kit (ThermoScientific, Waltham, USA) and oligo(dT)18 primers (ThermoScientific, Waltham, USA) following the manufacturer's instructions. The cDNA's produced were then used as template in PCRs with the primers designed to target the *MAT1-1–1* gene from the self-sterile locus. PCR products were sequenced and analysed in CLC Main Workbench as described above.

To detect expression of the *MAT1-1–1* gene in self-sterile isolates of *C. smalleyi* and *C. harringtonii*, RNA-seq libraries were prepared. This was done using total RNA (extracted as above) and the Illumina Stranded Total RNA Prep Ligation with Ribo-Zero Plus kit (Illumina, San Diego, USA), as well as the TruSeq Stranded Total RNA kit (Illumina, San Diego, USA). Each of the four libraries (two per species) were then pooled with unrelated samples and sequenced on an Illumina Hiseq X instrument (Illumina, San Diego, USA) to produce 150 bp pair-end sequences. RiboDetector (Deng et al., 2022) was used to remove reads with ribosomal RNA sequences, after which the remaining sequence reads were imported into CLC Genomics and mapped to the self-fertile and self-sterile *MAT1* locus sequences of *C. smalleyi* and *C. harringtonii* using the "Map reads to contigs" command with default settings. The data obtained were then visually inspected for read coverage across the respective *MAT1* gene models.

## 3. Results

### 3.1. Revision of the MAT1 gene models for Ceratocystis fimbriata

Using a combination of *in silico* predictions and RNA-seq data, the previously published *MAT1* locus structure for *C. fimbriata* was significantly updated (Wilken et al., 2014). AUGUSTUS and FGENESH respectively predicted three and four putative genes in the previously published *C. fimbriata MAT1* locus (Wilken et al., 2014), with AUGUSTUS also predicting multiple alternative models for the three genes. However, mapping of the RNA-seq data to the annotated contigs for isolate CMW 14799 supported the FGENESH prediction of four Open Reading Frames (ORFs) at the *MAT1* locus of this fungus (Fig. 1).

BLASTp searches with the translated protein sequences provided putative identities for all four genes. The first ORF corresponded to the *C. fimbriata MAT1-1–1* gene and spanned 1 937 bp including a single intron of 65 bases. It codes for a protein with 623 amino acids (aa) with the characteristic α-box DNA binding domain spanning the intron. The second ORF was not identified previously (Wilken et al., 2014), and was predicted only by FGENESH. It consisted of 653 bp with a single 50 bp

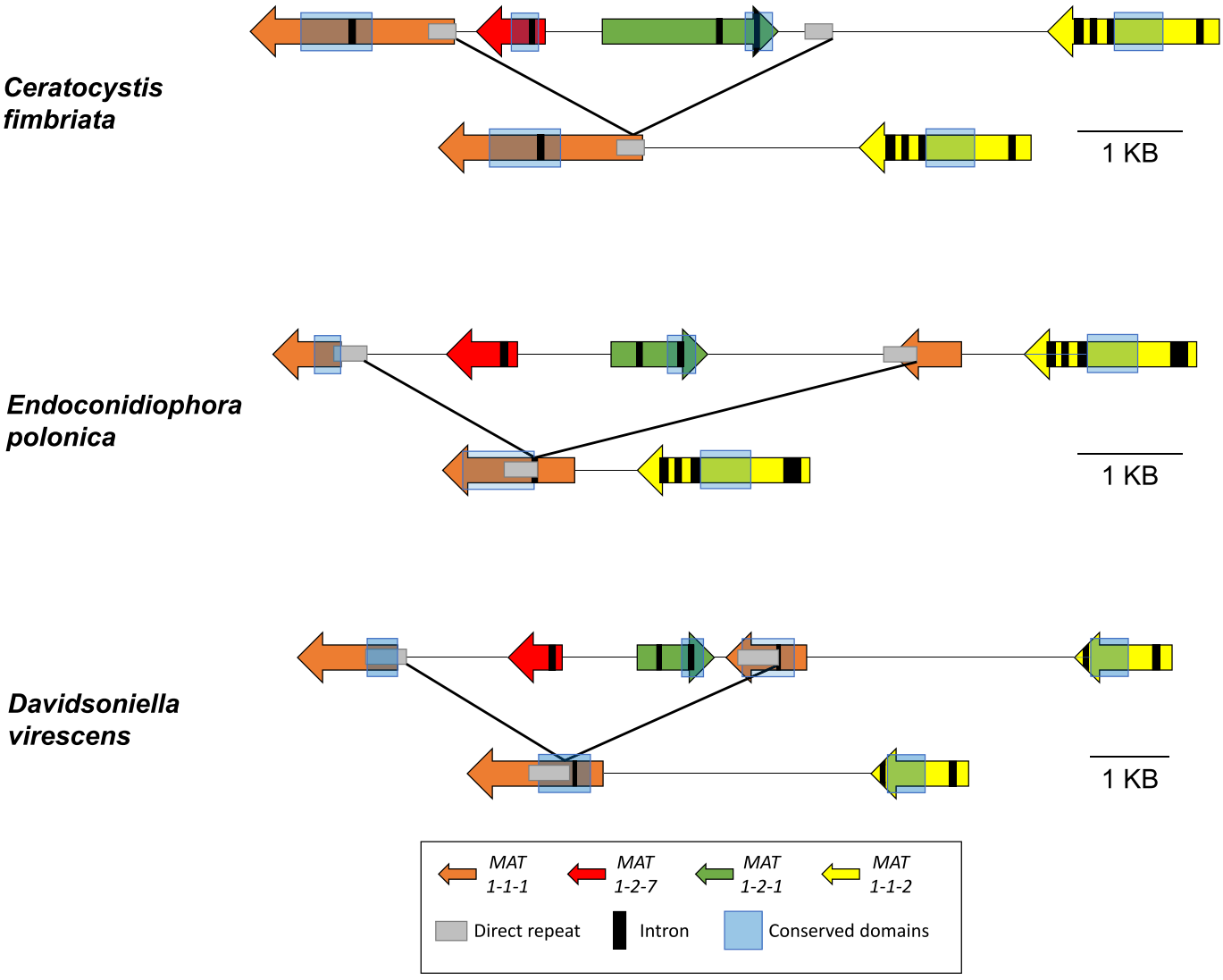


**Fig. 1.** The *MAT1* locus structure for *Ceratocystis fimbriata*, *Endoconidiophora polonica* and *Davidsoniella virescens*. In all three species, the self-fertile *MAT1* locus (shown at the top for each species) consists of four mating-type genes: *MAT1-2–7*, *MAT1-2–1*, *MAT1-1–2* and *MAT1-1–1*. In *E. polonica* and *D. virescens*, *MAT1-1–1* is interrupted by the *MAT1-2* gene information, which is flanked by two direct repeats. A putative recombination event involving these repeats deleted the *MAT1-2* information, leaving a self-sterile locus (shown at the bottom for each species) with only the *MAT1-1* related genes present. The conserved domain regions indicated corresponds to the α-box for *MAT1-1–1*, the HMG-box domain for *MAT1-2–1* and the conserved MAT1-1–2 domain for the *MAT1-1–2* gene. Figure is drawn to scale.

intron, and coded for a 200 aa protein lacking known conserved domains. However, BLASTp analysis identified this protein as MAT1-2–7, with the top BLAST hits corresponding those of *Thielaviopsis* (Wilken et al., 2018), as well as *Berkeleyomyces* (Nel et al., 2018), another member of *Ceratocystidaceae*. The third ORF was 1 670 bp in length, contained two introns (56 and 51 bp) and its predicted protein consisted of 520 aa. It also contained an HMG-box domain that was interrupted by the smaller intron across a serine codon. BLASTp searches identified this gene as *MAT1-2–1* that was previously described from *C. fimbriata* (Wilken et al., 2014). The fourth ORF had 1 631 bp and contained four introns (64, 68, 67 and 88 bp). Based on the BLASTp results, this 447 aa protein represented MAT1-1–2, and contained the conserved domain characteristic of this protein in other fungi (Dyer et al., 2016; Wilson et al., 2021b).

### 3.2. *MAT1* locus structure in self-fertile and self-sterile isolates of *Ceratocystis, Endoconidiophora and Davidsoniella virescens*

Using the protein sequences deduced from the confirmed *C. fimbriata MAT1* gene models, tBLASTn searches of the *Ceratocystis*, *Endoconidiophora* and *Davidsoniella* genome sequences identified one or two contigs containing *MAT1* gene sequences. These included matches to all four mating-type genes (i.e., *MAT1-1–1, MAT1-1–2, MAT1-2–1,* and *MAT1-2–7*) in all species apart from *D. virescens* that apparently lacked *MAT1-1–1*. However, a subsequent tBLASTn search of the *D. virescens* genome using the four *E. polonica* mating-type proteins (see below) identified an additional contig with similarity to the *MAT1-1–1* gene.

Conventional PCR and Sanger sequencing with species-specific primers (Table S3) were used to reconstruct a single, continuous sequence representing the self-fertile arrangement of the *MAT1* locus of each species. The size of the *MAT1* locus (defined as the region from the most distal to most proximal mating-type gene) ranged from 8 688 bp in *E. polonica* to 11 592 bp in *C. smalleyi* (Table 1). Subsequent annotation of the reconstructed *MAT1* locus sequences identified putative mating-type genes and repeat sequences for all *Ceratocystidaceae* species studied (Fig. 1 and Figure S3).

The *MAT1* locus of the *Ceratocystis* species contained homologs of the *MAT1-1–1*, *MAT1-2–1, MAT1-2–7,* and *MAT1-1–2* genes (Table 2). Overall, the position and structure of these genes closely matched those of *C. fimbriata*. In all these fungi, REPFIND also identified two copies of a direct repeat (Table 1). BLASTn searches of the repeat sequences against the respective genomes did not identify any additional copies of these sequences outside of the *MAT1* locus. BLASTn searches against the Genbank database only produced matches to the published *MAT1* locus of *C. fimbriata* (Wilken et al., 2014), apart from *C. albifundus* where no BLAST matches to any entry in the Genbank database were matched to the 84 bp repeat sequence. The start codon of the *MAT1-1–1* genes in all the *Ceratocystis* species were present within the first copy of the direct repeat, while the second copy was present in the intergenic sequence between the *MAT1-1–2* and *MAT1-2–1* genes. For example, in *C. fimbriata* most (256 bp) of the first repeat formed part of the coding region of the *MAT1-1–1* gene, with its start codon corresponding to the fifth base of the repeat. The second *C. fimbriata* repeat had an intergenic position between *MAT1-2–1* and *MAT1-1–2* (Fig. 1). A switching event would thus delete both *MAT1-2–1* and *MAT1-2–7*, while retaining the *MAT1-1–1* and *MAT1-1–2* genes (Wilken et al., 2014).

**Table 1**
Self-fertile and self-sterile *MAT1* locus size comparison between the species of *Ceratocystis*, *Endoconidiophora* and *Davidsoniella* used in this study.

| | *MAT1* locus | | | |
|---|---|---|---|---|
| Species | Self-fertile | Self-sterile | Deleted region | Direct repeat |
| | | *Ceratocystis* | | |
| *C. fimbriata* | 9205 bp | 5624 bp | 3581 bp | 260 bp |
| *C. manginecans* | 9215 bp | 5630 bp | 3585 bp | 260 bp |
| *C. eucalypticola* | 9178 bp | 5595 bp | 3583 bp | 260 bp |
| *C. cacaofunesta* | 9222 bp | 5637 bp | 3585 bp | 260 bp |
| *C. albifundus* | 9113 bp | 5765 bp | 3348 bp | 84 bp |
| *C. platani* | 9091 bp | 5116 bp | 3975 bp | 253 bp |
| *C. harringtonii* | 9772 bp | 6294 bp | 3478 bp | 126 bp |
| *C. smalleyi* | 11592 bp | 7739 bp | 3853 bp | 122 bp |
| | | *Endoconidiophora* | | |
| *E. polonica* | 8688 bp | 3477 bp | 5211 bp | 315 bp |
| *E. laricicola* | 11305 bp | 3490 bp | 7815 bp | 315 bp |
| | | *Davidsoniella* | | |
| *D. virescens* | 11030 bp | 6332 bp | 4698 bp | 517 bp |

The *MAT1-1–1*, *MAT1-2–1, MAT1-2–7* and *MAT1-1–2* genes were also annotated in the *MAT1* locus of both *Endoconidiophora* species (Fig. 1; Table 2). While the latter three genes closely resembled those of *Ceratocystis*, two partial *MAT1-1–1* genes were predicted in both *Endoconidiophora* species. As in *Ceratocystis*, two copies of a direct repeat were present in the *Endoconidiophora MAT1* locus, and BLAST searches with them produced matches to the *MAT1* locus of *Thielaviopsis* (Wilken et al., 2018), *Berkeleyomyces* (Nel et al., 2018) and *Huntiella* (Wilson et al., 2015a), but not against any other known repeat elements. The first repeat unit spanned the start codon of the predicted *MAT1-1-1L* gene, while the second copy was positioned between the *MAT1-2–1* and *MAT1-1–2* genes, in conjunction with *MAT1-1-1R* gene prediction (Fig. 1).

A single copy each of the *MAT1-1–2*, *MAT1-2–1* and *MAT1-2–7* genes and two genes with similarity to *MAT1-1–1* was annotated in the *D. virescens* self-fertile locus. One of the putative *MAT1-1–1* ORFs was in the same position as the *MAT1-1–1* gene of *Ceratocystis*, and its translated peptide contained only a partial α-box sequence. The other *MAT1-1–1* ORF was located between the *MAT1-2–1* and *MAT1-1–2* genes, and its predicted protein contained a complete α-box motif. The first *MAT1-1–1* gene was called *MAT1-1-1L* and the second *MAT1-1-1R*. Despite the presence of both the conserved domain and intervening intron, the translated *MAT1-1-1R* sequence was much shorter when compared to those of other *Ceratocystidaceae* (Table 2). Two direct repeats were identified only in the *MAT1* locus, which did not match any sequence in another part of the genome. A search of the Genbank database identified similarity to the *MAT1* loci of *Thielaviopsis* (Wilken et al., 2018), *Berkeleyomyces* (Nel et al., 2018) and *Huntiella* (Wilson et al., 2015a) only. Each of the predicted *MAT1-1–1* genes harboured a copy of the direct repeat.

Following what is known for *C. fimbriata* and *T. cerberus* (Krämer et al., 2021; Wilken et al., 2014), a self-sterile version of the *MAT1* locus for each species was produced *in-silico*. This was done by deleting of one copy of each repeat, together with the intervening sequence between the two repeat copies (Fig. 1). The validity of the self-sterile locus arrangement was confirmed in all species by amplification and sequencing with species-specific primers. *De novo* annotation and BLAST searches identified the *MAT1-1–1* and *MAT1-1–2* genes. In all cases, the former incorporated at least part of the remaining direct repeat into its predicted ORF. For *Ceratocystis*, the *MAT1-1–1* start codon was present in the repeat and did not differ from the one present in the self-fertile locus. In the two *Endoconidiophora* species, as well as *D. virescens*, the *MAT1-1–1* gene spanned the remaining repeat copy, which resulted in the predicted MAT1-1–1 protein from this locus to closely resemble that of *C. fimbriata* (Table 2). This is different to predictions for the self-fertile versions of the locus, where *MAT1-1–1* appeared to be truncated. Therefore, for all three species, the predicted protein contained a complete α-box domain, which was interrupted by an intron in the gene model.

### 3.3. *MAT1-1–1* gene expression from the self-sterile *MAT1* locus

After 5 to 7 days of growth on MEA-TS medium, the presence of both ascomata and mRNA from the self-sterile *MAT1* locus was evaluated. Ascomata were present only in *C. manginecans*, *C. eucalypticola*, *C. cacaofunesta*, *C. platani* and *C. albifundus*, and in all these species, *MAT1-1–1* gene transcripts from the self-sterile version of the *MAT1*

**Table 2**
In-silico predictions of the genes and their introns identified at the self-fertile (SF) and self-sterile (SS) versions of the *MAT1* locus of the *Ceratocystis*, *Endoconidiophora* and *Davidsoniella* species examined in this study.

| | MAT1-1–1 (SF) | | MAT1-1–1 (SS) | | MAT1-2–7 (SF, SS) | | MAT1-2–1 (SF, SS) | | MAT1-1–2 (SF, SS) | |
|---|---|---|---|---|---|---|---|---|---|---|
| Species | Gene length (polypeptide length) | No. of introns (size) | Gene length (polypeptide length) | No. of introns (size) | Gene length (polypeptide length) | No. of introns (size) | Gene length (polypeptide length) | No. of introns (size) | Gene length (polypeptide length) | No. of introns (size) |
| *Ceratocystis fimbriata* | 1937 bp (623 aa) | 1 (65 bp) | 1937 bp (623 aa) | 1 (65 bp) | 653 bp (200 aa) | 1 (50 bp) | 1670 bp (520 aa) | 2 (56, 51 bp) | 1631 (447 aa) | 4 (64, 68, 67, 88 bp) |
| *C. manginecans* | 2078 bp (670 aa) | 1 (65 bp) | 2078 bp (670 aa) | 1 (65 bp) | 653 bp (200 aa) | 1 (50 bp) | 1670 bp (520 aa) | 2 (56, 51 bp) | 1631 bp (447 aa) | 4 (64, 68, 67, 88 bp) |
| *C. eucalypticola* | 2096 bp (676 aa) | 1 (65 bp) | 2096 bp (676 aa) | 1 (65 bp) | 653 bp (200 aa) | 1 (50 bp) | 1670 bp (520 aa) | 2 (56, 51 bp) | 1631 bp (447 aa) | 4 (64, 68, 67, 88 bp) |
| *C. cacaofunesta* | 1763 bp (565 aa) | 1 (65 bp) | 1763 bp (565 aa) | 1 (65 bp) | 653 bp (200 aa) | 1 (50 bp) | 1670 bp (520 aa) | 2 (56, 51 bp) | 1631 bp (447 aa) | 4 (64, 68, 67, 88 bp) |
| *C. albifundus* | 1813 bp (579 aa) | 1 (73 bp) | 1813 bp (579 aa) | 1 (73 bp) | 653 bp (200 aa) | 1 (50 bp) | 1602 bp (498 aa) | 2 (54, 51 bp) | 1632 bp (447 aa) | 4 (65, 68, 67, 88 bp) |
| *C. platani* | 1892 bp (608 aa) | 1 (65 bp) | 1892 bp (608 aa) | 1 (65 bp) | 653 bp (200 aa) | 1 (50 bp) | 1670 bp (520 aa) | 2 (56, 51 bp) | 1631 bp (447 aa) | 4 (64, 68, 67, 88 bp) |
| *C. harringtonii* | 2062 bp (638 aa) | 2 (78, 67 bp) | 2062 bp (638 aa) | 2 (78, 67 bp) | 653 bp (200 aa) | 1 (50 bp) | 1665 bp (519 aa) | 2 (54, 51 bp) | 1408 bp (435 aa) | 3 (68, 67, 76 bp) |
| *C. smalleyi* | 1664 bp (505 aa) | 2 (78, 68 bp) | 1664 bp (417 aa) | 2 (78, 68 bp) | 653 bp (200 aa) | 1 (50 bp) | 2012 bp (603 aa) | 3 (93, 55, 52 bp) | 1408 bp (434 aa) | 3 (68, 67, 76 bp) |
| *Endoconidiophora polonica* | 648 bp (215 aa) | 0 (- bp) | 1255 bp (399 aa) | 1 (58 bp) | 676 bp (200 aa) | 1 (73 bp) | 918 bp (267 aa) | 2 (56, 61 bp) | 1638 bp (412 aa) | 4 (166, 87, 66, 80 bp) |
| *E. laricicola* | 639 bp (212 aa) | 0 (- bp) | 1246 bp (395 aa) | 1 (58 bp) | 677 bp (200 aa) | 1 (74 bp) | 903 bp (261 aa) | 2 (56, 61 bp) | 1662 bp (412 aa) | 4 (165, 87, 90, 81 bp) |
| *Davidsoniella virescens*[1] | 1266 & 1019 bp (421 & 321 aa) | 0 & 1 (53 bp) | 1718 bp (554 aa) | 1 (53 bp) | 684 bp (200 aa) | 1 (81 bp) | 974 bp (281 aa) | 2 (58, 70 bp) | 1234 bp (357 aa) | 2 (91, 69 bp) |

[1] Two partial *MAT1-1-1* genes were predicted for the self-fertile *D. virescens MAT1* locus.

locus were detected using cDNA-based PCR analysis. Ascomata were not observed in cultures of *C. harringtonii*, *C. smalleyi*, *E. polonica*, *E. laricicola* and *D. virescens*, although cDNA-based PCR analysis suggested that mRNA from the self-sterile version of the locus was encoded. In *Ceratocystis*, the primer position was such that amplification was only possible when a region likely corresponding to the promoter is positioned next to the start codon for *MAT1-1–1*, which would only be the case following a switching event (Figure S2). This confirmed that the amplicon was derived from the *MAT1-1–1* gene version present in the self-sterile isolate, while the absence of the intron in the cDNA sequence confirmed that the amplicon was derived from mRNA. Similarly, primer design for *E. laricicola, E. polonica* and *D. virescens* was such that amplification of the target size was only possible from the *MAT1-1–1* gene of the self-sterile locus, while the absence of the predicted intron confirmed that the fragment originated from mRNA (Figure S2).

No amplicons corresponding to the *MAT1-1–1* transcripts produced from the switched locus were amplified from the cDNA of *C. harringtonii* or *C. smalleyi*, despite multiple attempts. The RNA-seq analysis for the *MAT1* locus of both these species also yielded no transcripts that could unequivocally be linked to the self-sterile version of the *MAT1* locus (Figure S4). No RNA reads were mapped to either of the *MAT1-1–1* or *MAT1-1–2* genes in the self-fertile or self-sterile locus versions of *C. harringtonii*, although expression was apparent for the *MAT1-2–1* and *MAT1-2–7* genes. The RNA-seq dataset for *C. smalleyi* contained reads mapping to all four mating-type genes, but the number of reads associated with the *MAT1-2–1* (up to 600 reads) and *MAT1-2–7* (up to 21 reads) genes was somewhat higher than the 9 reads mapping to the self-sterile version of *MAT1-1–1* and the single read mapping to the self-fertile version of *MAT1-1–1*.

## 4. Discussion

This study highlights a conserved *MAT1* locus structure that underpins unidirectional mating-type switching across diverse species and it also provides clues regarding how this unique reproductive strategy functions. This unique architecture becomes apparent when comparing the *MAT1* locus among *Ceratocystidaceae* with those published for *Chromocrea spinulosa* (Yun et al., 2017) and *Sclerotinia trifoliorum* (Xu et al., 2016). The *MAT1* locus has distinct organisational structures in the respective self-fertile and self-sterile isolates of all these species. In the self-fertile isolates, the locus has all the expected *MAT1-1* genes, but the *MAT1-1–1* gene is disrupted by a region containing the *MAT1-2* genes. The latter region is located between two direct repeats that mediate a recombination that deletes the entire stretch of DNA between

them as suggested in earlier studies (Wilken et al., 2014; Xu et al., 2016; Yun et al., 2017). The *MAT1* locus in the self-sterile isolates contains only *MAT1-1* gene information, with one copy of the repeat being part of the *MAT1-1–1* gene. Unlike yeasts that harbour additional loci encoding mating-type genes (Wolfe and Butler, 2022), recombination between the two repeats in the studied *Ceratocystidaceae* completely deletes the *MAT1-2* gene-containing region of the genome. Therefore, the change in locus architecture forms the basis of unidirectional mating-type switching in these fungi as it results in an irreversible switch from self-fertility to self-sterility. The presence of these two architectures linked to the *MAT1* locus serves as a distinguishing feature in fungi exhibiting unidirectional mating-type switching. Utilizing this understanding should facilitate the identification of additional fungi displaying this unique mating behaviour.

Various commonalities in the gene content of the *MAT1* locus are known in species capable of unidirectional mating-type switching. In all cases, both the self-fertile and self-sterile arrangement of the locus contained intact copies of the so-called "secondary" *MAT1-1* genes, e.g., *MAT1-1–2* in the *Ceratocystidaceae* (Krämer et al., 2021; Wilken et al., 2014), *MAT1-1–2* and *MAT1-1–3* in *Ch. spinulosa* (Yun et al., 2017) and *MAT1-1–5* in *S. trifoliorum* (Xu et al., 2016). This was also seen in the current study where the *MAT1-1–2* gene was present in all the *Ceratocystis*, *Endoconidiophora* and *Davidsoniella* species considered. The self-sterile arrangement of the locus is further defined by the presence of a single, intact *MAT1-1–1* gene with all the expected features, including the intron-containing motif that codes for the protein's characteristic α-box domain (Wilken et al., 2017). In the self-fertile arrangement, *MAT1-2* genes are present as a repeat-flanked group that include *MAT1-2–1* together with so-called "secondary" *MAT1-2* genes, e.g., *MAT1-2–7* in the *Ceratocystidaceae*, *MAT1-2–4* in *S. trifoliorum* (Xu et al., 2016) and *MAT1-2–3 Ch. spinulosa* (Yun et al., 2017). This group of *MAT1-2* genes split the *MAT1-1–1* gene, either between the coding region and the putative promoter (as in *Ceratocystis*) or entirely split the coding region as seen in *Endoconidiophora* species, *D. virescens*, *T. cerberus*, *Ch. spinulosa* and *S. trifoliorum.*

The results of this study suggest that the *MAT1-2* intervening sequence in the *MAT1* locus of *Ceratocystis* species eliminates the expression of *MAT1-1–1*. The self-fertile version of the locus contains an apparently intact version of *MAT1-1–1*, but its start codon and first part of the gene lies within one of the direct repeats. In this position, we hypothesize that the *MAT1-2* intervening sequence separates *MAT1-1–1* from its promoter, which then places transcriptional regulation of the *MAT1-1–1* gene under the control of the switching event (Fig. 2). Deletion of the intervening *MAT1-2* sequence is required to bring together the *MAT1-1–1* gene and its promoter element located after the second repeat. Therefore, despite there being no detectable differences in *MAT1-1–1* gene sequences between self-fertile and self-sterile isolates, the gene would not be expressed in the self-fertile isolates. This was confirmed for *C. albifundus*, *C. cacaofunesta*, *C. eucalypticola*, *C. fimbriata*, *C. manginecans* and *C. platani* where expression analysis showed the presence of transcripts containing the full *MAT1-1–1* gene and putative promoter, brought together after the switching event. In *C. harringtonii* and *C. smalleyi*, only expression of the *MAT1-2* genes was detected, suggesting that only the self-fertile arrangement for the *MAT1* locus was

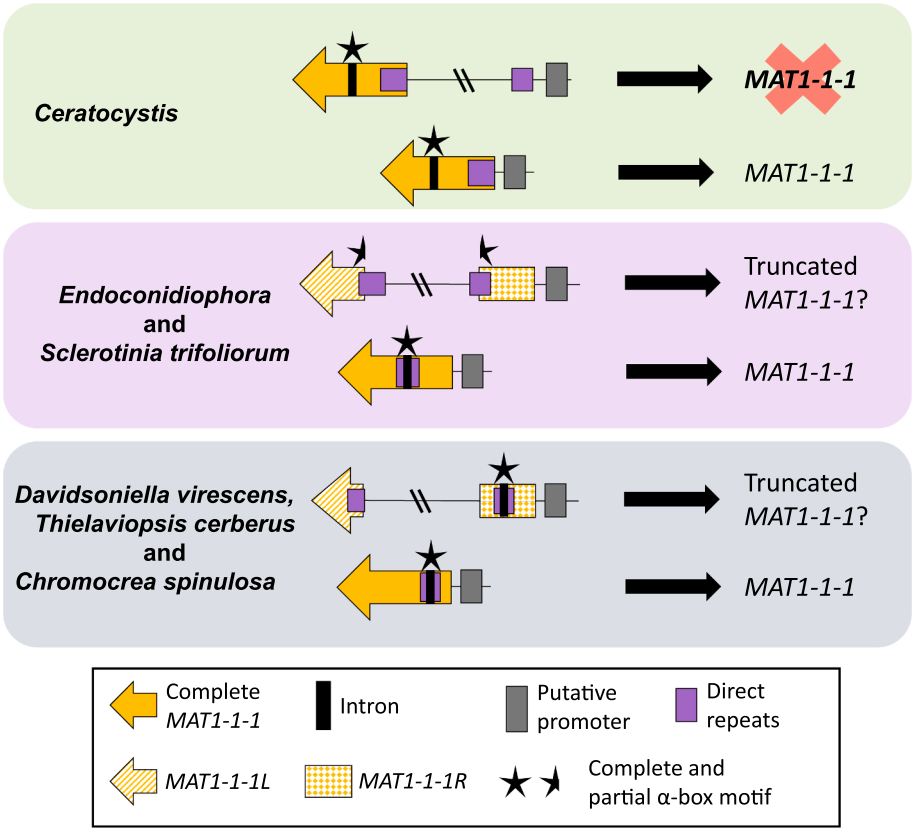


**Fig. 2.** Models illustrating control of *MAT1-1–1* gene expression through unidirectional mating-type switching. In *Ceratocystis* species, the hypothetical promoter of the *MAT1-1–1* gene is separated from the coding region through the presence of the *MAT1-2* gene region. Switching brings the *MAT1-1–1* coding region and promoter into close proximity, resulting in expression of the protein-coding mRNA. In *Endoconidiophora* species and *Sclerotinia trifoliorum,* the putative promoter region as well as the 5′ region of *MAT1-1–1* is separated from the remaining part of the gene. Deletion of the intervening *MAT1-2* gene information produces a complete *MAT1-1–1* gene, bringing together the two regions each containing a fragment of the α-box. For *Davidsoniella virescens, Thielaviopsis cerberus* and *Chromocrea spinulosa,* the putative promoter region as well as the 5′ end of *MAT1-1–1* containing the α-box is separated from the remaining part of the gene by the *MAT1-2* gene information. Deletion of the intervening sequence produces a mature *MAT1-1–1* gene by joining the two fragments. Figure is not drawn to scale.

present. This was supported by the absence of ascomata or ascospore drops in cultures of these species, which together indicates that switching had not occurred in these fungi. Our future research will thus seek to identify and characterize the predicted promoter in these fungi, as well as to explore *MAT* gene expression in sporulating cultures of these species.

In *Endoconidiophora*, *D. virescens* and *T. cerberus*, parts of the *MAT1-1–1* gene disrupted by the *MAT1-2* intervening region might be associated with additional functions. Such disruption of the *C. spinulosa MAT1-1–1* gene has been shown to generate gene fragments that were termed *MAT1-1-1L*, containing approximately 83 % of the gene including the alpha-box, and *MAT1-1-1S*, containing the remainder of the gene sequence (Yun et al., 2017). Using a range of transformants, coupled with extensive molecular analysis, Yun et al. (2017) showed that the *MAT1-1-1L* gene likely functions as an autonomous mating gene that, together with the complete *MAT1-1–1* gene, controls karyogamy during sexual reproduction in *C. spinulosa*. Whether one of the gene fragments of *Endoconidiophora*, *D. virescens* (both considered in this study) and *T. cerberus* (Krämer et al., 2021) could also have functional roles is not known. However, in the case of *D. virescens*, the *MAT1-1-1L* fragment also encodes for a full α-box domain, suggesting that its possible involvement in the mating process requires further scrutiny (Yun et al., 2017). In *T. cerberus* and *Endoconidiophora*, it is less likely that *MAT1-1-1L* and *MAT1-1-1R* fragments function as autonomous mating genes as the α-box encoding motif is split by the *MAT1-2* intervening region. Nevertheless, given that both self-fertile and self-sterile arrangements of the *MAT1* locus are needed for sexual reproduction (Krämer et al., 2021; Yun et al., 2017), our findings provide a valuable framework within which to employ transformation protocols in development (Lane et al., 2021; Sayari et al., 2019) to functionally characterize all of the genes involved in *Ceratocystidaceae* sexual reproduction.

Experimental confirmation of the predicted gene models in *C. fimbriata* was crucial for the accurate identification and annotation of the various *MAT1* genes. Reannotation of the *MAT1* locus using various software options also allowed identification of the *MAT1-2–7* gene that was not detected in the previous annotation of the *C. fimbriata MAT1* locus (Wilken et al., 2014). This gene was first described in the *Ceratocystidaceae* species *Huntiella omanensis* and *H. moniliformis* (Wilson et al., 2015a), but has subsequently been identified in other *Microascales* species (Aylward et al., 2016; Krämer et al., 2021; Nel et al., 2018; Wilken et al., 2018). The experimentally confirmed models for the *C. fimbriata MAT1-1–1*, *MAT1-2–1* and *MAT1-1–2* genes in this study differed significantly from those originally predicted for the fungus by (Wilken et al., 2014). However, the identification of the conserved domains for the respective proteins as well as the presence of conserved introns in the former two genes (Dyer et al., 2016; Wilken et al., 2017) provided support for our annotations. In addition, the incorporation of the direct repeat sequences into the *MAT1-1–1* gene aligned well with the *MAT1-1–1* genes of *Ch. spinulosa* and *Sclerotinia trifoliorum* (Xu et al., 2016; Yun et al., 2017). Consequently, our results highlight the importance of obtaining experimental support for predicted mating-type genes, as these genes are known to be highly diverse, even between closely related species (Butler, 2007; Debuchy et al., 2010; Wilken et al., 2017). The availability of the corrected models emerging from the present study will support future work that aim to elucidate the role of unidirectional mating-type switching in the sexual cycles of these fungi.

An intriguing but unresolved question pertains to the evolutionary origin of unidirectional mating-type switching as a reproductive strategy. One hypothesis is that mating-type switching evolved for reproductive assurance (reviewed in (Nieuwenhuis and Immler, 2016), where the apparent "self-compatibility" of a fungus ensures that sexual reproduction (and the associated formation of sexual structures) can occur in the absence of a suitable mating partner. Consistent with that view, *Ceratocystidaceae* capable of unidirectional mating-type switching are closely related to heterothallic species in the genera *Thielaviopsis* (Mbenoun et al., 2014; Wilken et al., 2018), *Huntiella* (Liu et al., 2018) and *Berkeleyomyces* (Nel et al., 2018). This relatedness is even more pronounced for *D. virescens* that has a heterothallic sister-species in *D. eucalypti* (Kile et al., 1996), while *T. cerberus* is the only *Thielaviopsis* species capable of unidirectional mating-type switching while all other species in this genus are heterothallic (Krämer et al., 2021; Wilken et al., 2018). In light of these phylogenetic relationships, the evolutionary roots of unidirectional mating-type switching may lie in abnormal crossing over events that occurs during sexual reproduction. Such events around regions of identity in the highly diverse *MAT* locus are known to shape transitions between reproductive strategies (e.g., Gioti et al., 2012; Yun et al., 1999; Yun et al., 2017).

Sexual reproduction is essential in the biology of many *Ceratocystidaceae* species where ascospores are produced in sticky drops on elongated ascomatal necks that facilitate dispersal by insects (Wingfield et al., 2013). As such, the formation of ascomata is an important component of the ecology of these species and supports the evolution of a reproductive strategy that ensures sexual reproduction under all environmental conditions. Information on the sexual cycles and mating-type locus of species in ambrosial-associated *Ceratocystidaceae* genera such as *Ambrosiella* (De Beer et al., 2014), *Meredithiella* (Mayers et al., 2015), *Toshionella* (Mayers et al., 2020b), *Phialophoropsis* (Mayers et al., 2015) and *Wolfgangiella* (Mayers et al., 2020a) remain lacking. Many of these species do not have described sexual states, although modified ascomata that lack long necks and sticky ascospore drops have been described for a small number of these species (Mayers et al., 2020a; Mayers et al., 2017; Mayers et al., 2015). These structures do not appear to have a role in the dispersal of ascospores via adhesion to the insect body, and may rather be linked to the generation of genetic variation through meiosis (Mayers et al., 2020a). An analysis of the *MAT1* locus in these ambrosial-associated species could provide further information on the role of sexual reproduction in their biology.

## Funding

Markus Wilken was supported in part by a post-doctoral research fellowship from the Claude Leon Foundation. We thank the University of Pretoria, the South African Department of Science and Technology (DST) and National Research Foundation (NRF) for funding via the Centres of Excellence programme (Centre of Excellence in Tree Heath Biotechnology) and the South African Research Chairs Initiative (SARChI; SARChI Chair in Fungal Genomics). The Grant holders acknowledge that opinions, findings and conclusions or recommendations expressed in publications generated by NRF supported research are that of the authors, and that the NRF accepts no liability whatsoever in this regard.

## CRediT authorship contribution statement

**P. Markus Wilken:** Conceptualization, Methodology, Formal analysis, Investigation, Data curation, Writing – original draft, Writing – review & editing, Visualization, Project administration, Funding acquisition. **Frances A. Lane:** Data curation, Investigation, Writing – review & editing. **Emma T. Steenkamp:** Conceptualization, Funding acquisition, Writing – review & editing. **Michael J. Wingfield:** Conceptualization, Writing – review & editing, Funding acquisition. **Brenda D. Wingfield:** Conceptualization, Writing – review & editing, Funding acquisition.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

Markus Wilken was supported in part by a post-doctoral research fellowship from the Claude Leon Foundation. We thank the University of Pretoria, the South African Department of Science and Technology (DST) and National Research Foundation (NRF) for funding via the Centres of Excellence programme (Centre of Excellence in Tree Heath Biotechnology) and the South African Research Chairs Initiative (SARChI; SARChI Chair in Fungal Genomics). The Grant holders acknowledge that opinions, findings and conclusions or recommendations expressed in publications generated by NRF supported research are that of the authors, and that the NRF accepts no liability whatsoever in this regard.

## Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fgb.2023.103859.

## References

Andrus, C.F., Harter, L., 1933. Morphology of reproduction in *Ceratostomella fimbriata*. J. Agric. Res. 46, 1059–1078.

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 Genet. Biol. 96, 47–57.

Baker Engelbrecht, C.J., Harrington, T.C., 2005. Intersterility, morphology and taxonomy of *Ceratocystis fimbriata* on sweet potato, cacao and sycamore. Mycologia 97, 57–69.

Bakshi, B.K., 1951. Studies on four species of *Ceratocystis*, with a discussion on fungi causing sap-stain in Britain. Mycological Papers 35, 1–16.

Belbahri, L., 2015. Genome sequence of *Ceratocystis platani*, a major pathogen of plane trees. https://www.ncbi.nlm.nih.gov/nuccore/LBBL00000000.1.

Benson, D.A., Cavanaugh, M., Clark, K., Karsch-Mizrachi, I., Lipman, D.J., et al., 2013. GenBank. Nucleic Acids Res. 41, D36–D42.

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. Curr. Biol. 12, 1756–1761.

Billiard, S., Lopez-Villavicencio, M., Devier, B., Hood, M.E., Fairhead, C., Giraud, T., 2011. Having sex, yes, but with whom? Inferences from fungi on the evolution of anisogamy and mating types. Biol. Rev. Camb. Philos. Soc. 86, 421–442.

Butler, G., 2007. The evolution of *MAT*: the Ascomycetes. In: Heitman, J., Kronstad, J.W., Taylor, J.W., Casselton, L.A. (Eds.), Sex in Fungi: Molecular Determination and Evolutionary Implications. American Society of Microbiology, Washington, DC, pp. 3–18.

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.

Davidson, R.W., 1953. Two common lumber-staining fungi in the western United States. Mycologia 45, 579–586.

De Beer, Z.W., Duong, T.A., Barnes, I., Wingfield, B.D., Wingfield, M.J., 2014. Redefining *Ceratocystis* and allied genera. Stud. Mycol. 79, 187–219.

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

Deng, Z.-L., Münch, P.C., Mreches, R., McHardy, A.C., 2022. Rapid and accurate identification of ribosomal RNA sequences via deep learning. Nucleic Acids Res. 50, e60–e.

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

Ferreira, E.M., Harrington, T.C., Thorpe, D.J., Alfenas, A.C., 2010. Genetic diversity and interfertility among highly differentiated populations of *Ceratocystis fimbriata* in Brazil. Plant Pathol. 59, 721–735.

Finn, R. D., Miller, B. L., Clements, J., Bateman, A., 2014. iPfam: a database of protein family and domain interactions found in the Protein Data Bank. Nucleic Acids Res. 42, D364-D373.

Gioti, A., Mushegian, A.A., Strandberg, R., Stajich, J.E., Johannesson, H., 2012. Unidirectional evolutionary transitions in fungal mating systems and the role of transposable elements. Mol. Biol. Evol. 29, 3215–3226.

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

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kile, G.A., Harrington, T.C., Yuan, Z.Q., Dudzinski, M.J., Old, K.M., 1996. *Ceratocystis eucalypti* sp. nov., a vascular stain fungus from eucalypts in Australia. Mycol. Res. 100, 571–579.

Koressaar, T., Remm, M., 2007. Enhancements and modifications of primer design program Primer3. Bioinformatics 23, 1289–1291.

Krämer, D., Lane, F.A., Steenkamp, E.T., Wingfield, B.D., Wilken, P.M., 2021. Unidirectional mating-type switching confers self-fertility to *Thielaviopsis cerberus*, the only homothallic species in the genus. Fungal Biol. 125, 427–434.

Lane, F.A., Du Plessis, D., Wingfield, B.D., Wilken, P.M., 2021. Transferring an *Agrobacterium*-mediated transformation protocol across eight genera in the Ceratocystidaceae. For. Pathol. 51, e12688.

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 Pathol. 64, 1103–1109.

Lin, X., Heitman, J., 2007. Mechanisms of homothallism in fungi and transitions between heterothallism and homothallism. 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. 35–57.

Liu, F., Li, G., Roux, J., Barnes, I., Wilson, A.M., et al., 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. Ann. Bot. 16, 449–468.

Mayers, C.G., McNew, D.L., Harrington, T.C., Roeper, R.A., Fraedrich, S.W., et al., 2015. Three genera in the *Ceratocystidaceae* are the respective symbionts of three independent lineages of ambrosia beetles with large, complex mycangia. Fungal Biol. 119, 1075–1092.

Mayers, C.G., Harrington, T.C., Ranger, C.M., 2017. First report of a sexual state in an ambrosia fungus: *Ambrosiella cleistominuta* sp. nov. associated with the ambrosia beetle *Anisandrus maiche*. Botany 95, 503–512.

Mayers, C.G., Harrington, T.C., Masuya, H., Jordal, B.H., McNew, D.L., et al., 2020a. Patterns of coevolution between ambrosia beetle mycangia and the *Ceratocystidaceae*, with five new fungal genera and seven new species. Persoonia 44, 41–66.

Mayers, C.G., Harrington, T.C., McNew, D.L., Roeper, R.A., Biedermann, P.H.W., et al., 2020b. Four mycangium types and four genera of ambrosia fungi suggest a complex history of fungus farming in the ambrosia beetle tribe Xyloterini. Mycologia 112, 1104–1137.

Mbenoun, M., De Beer, Z.W., 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.

Misse, A.C., Barnes, I., Roets, F., Mbenoun, M., Wingfield, M.J., Roux, J., 2017. Ecology and population structure of a tree wound-infecting fungus in a native South African forest environment. Fungal Biol. 121, 69–81.

Molano, E.P.L., Cabrera, O.G., Jose, J., do Nascimento, L.C., Carazzolle, M.F., et al., 2018. Ceratocystis cacaofunesta genome analysis reveals a large expansion of extracellular phosphatidylinositol-specific phospholipase-C genes (PI-PLC). BMC Genomics 19, 58.

Morgulis, A., Gertz, E.M., Schäffer, A.A., Agarwala, R., 2006. A fast and symmetric DUST implementation to mask low-complexity DNA sequences. J. Comput. Biol. 13, 1028–1040.

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 Biol. 122, 1031–1040.

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

Nieuwenhuis, B.P., Immler, S., 2016. The evolution of mating-type switching for reproductive assurance. Bioessays 38, 1141–1149.

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

Roach, K.C., Feretzaki, M., Sun, S., Heitman, J., 2014. Unisexual reproduction. Adv. Genet. 85, 255–305.

Sambrook, J., Russell, D.W., 2001. Molecular Cloning - A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Sayari, M., Van der Nest, M.A., Steenkamp, E.T., Adegeye, O., Marincowitz, S., Wingfield, B.D., 2019. *Agrobacterium*-mediated transformation of *Ceratocystis albifundus*. Microbiol. Res. 226, 55–64.

Solovyev, V., Kosarev, P., Seledsov, I., Vorobyev, D., 2006. Automatic annotation of eukaryotic genes, pseudogenes and promoters. Genome Biol. 7 Suppl. 1, S10.1-S10.12.

Solovyev, V. V., 2007. Statistical approaches in Eukaryotic gene prediction. In: D. Balding, C. Cannings, M. Bishop (Eds.), Handbook of Statistical Genetics. Wiley-Interscience, pp. 1616.

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 Biol. 7 Suppl. 1, S11.1–S11.8.

Turgeon, B.G., Debuchy, R., 2007. *Cochliobolus* and *Podospora*: mechanisms of sex determination and the evolution of reproductive lifestyle. 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. 93–121.

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., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., et al., 2012. Primer3-new capabilities and interfaces. Nucleic Acids Res. 40, e115.

van der Nest, M.A., Beirn, L.A., Crouch, J.A., Demers, J.E., de Beer, Z.W., et al., 2014a. IMA Genome-F 3: Draft genomes of *Amanita jacksonii*, *Ceratocystis albifundus*,

*Fusarium circinatum*, *Huntiella omanensis*, *Leptographium procerum*, *Rutstroemia sydowiana*, and *Sclerotinia echinophila*. IMA Fungus 5, 472–485.
Van der Nest, M.A., Bihon, W., De Vos, L., Naidoo, K., Roodt, D., et al., 2014b. IMA Genome-F 2: Draft genome sequences of *Diplodia sapinea*, *Ceratocystis manginecans*, and *Ceratocystis moniliformis*. IMA Fungus 5, 135–140.
Van der Nest, M.A., Steenkamp, E.T., Roodt, D., Soal, N.C., Palmer, M., et al., 2019. Genomic analysis of the aggressive tree pathogen *Ceratocystis albifundus*. Fungal Biol. 123, 351–363.
Wheeler, H., 1950. Genetics of *Glomerella*. VIII. A genetic basis for the occurrence of minus mutants. Am. J. Bot. 37, 304–312.
Wilken, P.M., Steenkamp, E.T., De Beer, Z.W., Wingfield, M.J., Wingfield, B.D., 2013. IMA Genome-F1: Draft nuclear genome sequence for the plant pathogen, *Ceratocystis fimbriata*. IMA Fungus 4, 357–358.
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 Biol. Rev. 31, 199–211.
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 Genet. Biol. 113, 32–41.
Wilson, A. M., Wilken, P. M., Wingfield, M. J., Wingfield, B. D., 2021b. Genetic networks that govern sexual reproduction in the Pezizomycotina. Microbiol. Mol. Biol. Rev. 85, e00020-21.
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 Genet. Biol. 80, 1–9.
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., Gabriel, R., Singer, S.W., Schuerg, T., Wilken, P.M., et al., 2021a. Doing it alone: Unisexual reproduction in filamentous ascomycete fungi. Fungal Biol. 35, 1–13.
Wingfield, B.D., Barnes, I., De Beer, Z.W., De Vos, L., Duong, T.A., et al., 2015. IMA Genome-F 5: Draft genome sequences of *Ceratocystis eucalypticola*, *Chrysoporthe cubensis*, *C. deuterocubensis*, *Davidsoniella virescens*, *Fusarium temperatum*, *Graphilbum fragrans*, *Penicillium nordicum*, and *Thielaviopsis musarum*. IMA Fungus 6, 493–506.
Wingfield, B.D., Ambler, J.M., Coetzee, M., De Beer, Z.W., Duong, T.A., et al., 2016a. IMA Genome-F 6: Draft genome sequences of *Armillaria fuscipes*, *Ceratocystiopsis minuta*, *Ceratocystis adiposa*, *Endoconidiophora laricicola*, *E. polonica* and *Penicillium freii* DAOMC 242723. IMA Fungus 7, 217–227.
Wingfield, B.D., Duong, T.A., Hammerbacher, A., van der Nest, M.A., Wilson, A., et al., 2016b. IMA Genome-F 7: Draft genome sequences for *Ceratocystis fagacearum*, *C. harringtonii*, *Grosmannia penicillata*, and *Huntiella bhutanensis*. IMA Fungus 7, 317–323.
Wingfield, B.D., Bills, G.F., Dong, Y., Huang, W., Nel, W.J., et al., 2018. IMA Genome-F 9: Draft genome sequence of Annulohypoxylon stygium, Aspergillus mulundensis, Berkeleyomyces basicola (syn. Thielaviopsis basicola), Ceratocystis smalleyi, two Cercospora beticola strains, Coleophoma cylindrospora, Fusarium fracticaudum, Phialophora cf. hyalina, and Morchella septimelata. IMA Fungus 9, 199–223.
Wingfield, M.J., Roux, J., Wingfield, B.D., Slippers, B., 2013. *Ceratocystis* and *Ophiostoma*: international spread, new associations and plant health. In: Seifert, K.A., De Beer, Z.W., Wingfield, M.J. (Eds.), The Ophiostomatoid Fungi: Expanding Frontiers. CBS-KNAW Fungal Biodiversity Centre, CBS, Utrecht, The Netherlands, pp. 191–200.
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*. Curr. Genet. 38, 48–52.
Wolfe, K.H., Butler, G., 2022. Mating-type switching in budding yeasts, from flip/flop inversion to cassette mechanisms. Microbiol. Mol. Biol. Rev. 86, e00007–e00021.
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*. Sci. Rep. 6, 27083.
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. Proc. Natl. Acad. Sci. u.s.a. 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 Genet. 13, e1006981.