Accepted Manuscript

Transition from heterothallism to homothallism is hypothesised to have facilitated speciation among emerging *Botryosphaeriaceae* wheat-pathogens

Elisha Thynne, Megan C. McDonald, Peter S. Solomon

PII: S1087-1845(17)30150-0
DOI: <https://doi.org/10.1016/j.fgb.2017.10.005>
Reference: YFGBI 3087

To appear in: *Fungal Genetics and Biology*

Received Date: 18 November 2016
Revised Date: 14 October 2017
Accepted Date: 16 October 2017

Please cite this article as: Thynne, E., McDonald, M.C., Solomon, P.S., Transition from heterothallism to homothallism is hypothesised to have facilitated speciation among emerging *Botryosphaeriaceae* wheat-pathogens, *Fungal Genetics and Biology* (2017), doi: <https://doi.org/10.1016/j.fgb.2017.10.005>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Transition from heterothallism to homothallism is hypothesised to have facilitated speciation among emerging *Botryosphaeriaceae* wheat-pathogens

Elisha Thynne^a, Megan C. McDonald^{a,*}, Peter S. Solomon^{a,*}

* Co-corresponding authors

Email addresses: peter.solomon@anu.edu.au (Peter S. Solomon);

megan.mcdonald@anu.edu.au (Megan C. McDonald)

Keywords: Eutiarosporella, mating type, white grain disorder, speciation

ABSTRACT

White grain disorder (WGD) is a recently emerged wheat disease in Australia caused by three *Botryosphaeriaceae* fungi, from the genus *Eutiarosporella*. These species are *E. tritici-australis*, *E. darliae*, and *E. pseudodarliae*. Characterisation of the mating type genes for the WGD-species show that the genome sequence of a single *E. darliae* and *E. pseudodarliae* isolate both harbour *MAT1-2-1* and *MAT1-1-1*, which suggests that these species are homothallic. However, unlike most other characterised mating-type loci from other homothallic Dothideomycetes, these species' *MAT1-1-1* are located at a separate locus, inserted within the coding region of another gene. The sequenced strain of *E. tritici-australis* analysed did not harbour *MAT1-1-1*. Including the sequenced strain, we screened the mating type genes present in 16 *E. tritici-australis* individuals isolated from infected grain from fields in South Australia. Of these 16, 11 harbour *MAT1-1-1* and the other five harbour *MAT1-2-1*. The genome of a *MAT1-1-1* harbouring isolate was re-sequenced, which demonstrated that *MAT1-1-1* was present at the *MAT* locus. We examined non-coding DNA surrounding the *MAT1-1-1* gene in *E. pseudodarliae* and observed fragments of the *MAT* locus both up and downstream. These fragments and their orientation around *MAT1-1-1* is similar to characterised heterothallic *Botryosphaeriaceae*. Based on these gene arrangements, we conclude that the new *MAT1-1-1* containing locus likely originated from a cryptic DNA integration event between two heterothallic individuals. We hypothesise that this integration event led to the formation of a homothallic lineage, which is the common ancestor of *E. darliae* and *E. pseudodarliae*.

1. Introduction

Fungal phytopathogens are responsible for significant losses to the agricultural, horticultural and forestry industries globally. In addition to the established phytopathogens that these industries manage, there is also the threat of emerging pathogens. One such emerging disease was observed in 1999, infecting wheat in Queensland, Australia (Wildermuth *et al.*, 2001). This disease resulted in bleached and shrivelled grain-heads, and was described as white grain disorder (WGD) (Wildermuth *et al.*, 2001). This disease is now observed across the Eastern, and more recently Western wheat growing regions of Australia (Evans, 2013, Thomas and Jayasena, 2015). The causal agents of WGD were phenotypically classified as *Botryosphaeria zeae*, in the class Dothideomycetes (Wildermuth *et al.*, 2001). Using molecular markers and morphological comparisons, we recently demonstrated that the causal agents of WGD-fungi are in fact three separate species within the genus *Eutiarosporella*, classified as *Eutiarosporella darliae*, *E. pseudodarliae*, and *E. tritici-australis* (Thynne *et al.*, 2015). How or why three separate species emerged in a similar time frame to all cause an apparently novel disease in wheat is still a matter of speculation.

At the time when this disease was initially described there was little to no published sequence data available for this genus. Increased sampling, however, has led to the routine identification of new species from *Eutiarosporella* and sister genera internationally (Crous *et al.*, 2006, Jami *et al.*, 2012, Jami *et al.*, 2013, Thambugala *et al.*, 2014, Crous *et al.*, 2015). Although fungi of the family *Botryosphaeriaceae* are usually found in association with woody plants, some identified species were isolated from grasses (Sutton and Marasas, 1976, Thambugala *et al.*, 2014, Crous *et al.*, 2015). This indicates that *Eutiarosporella* fungi have a history of interacting with grasses.

Whilst many aspects of this disease remain unknown, field pathology studies have demonstrated that the WGD-*Eutiarospora* spp. produce ascospores in the field (Evans *et al.* 2015). This is significant as it is believed that a population of fungal phytopathogens has a greater adaptive potential if able to share and generate genetic diversity through sexual reproduction (McDonald and Linde, 2002). Abundant genetic diversity among pathogen populations can assist in overcoming strong selective pressure, such as fungicides or host-resistance. As such, it is important to understand how an emerging pathogen reproduces, in order to gain insight into its evolutionary potential (McDonald and Linde, 2002). This knowledge can assist in tactical disease management strategies (e.g. take steps to minimise the risk of resistance breakdown in host-cultivar breeding lines or overuse of single-mode of action fungicides) (McDonald and Linde, 2002).

Sexually reproducing Dothideomycetes are either homothallic (self-fertile) or heterothallic (non-self-fertile) (Yun *et al.*, 1999, Waalwijk *et al.*, 2002, Bennett, *et al.*, 2003, Yun *et al.* 2013, Bihon *et al.*, 2014, Bolton *et al.*, 2014, Marsberg *et al.*, 2016). Whether a species is homo- or heterothallic is determined genetically at the mating type locus (*MAT* locus) where the presence or absence of two genes, alpha-box encoding *MAT1-1-1* and HMG-box encoding *MAT1-2-1*, are known to regulate sexual reproduction. (Metzenberg *et al.*, 1990, Turgeon *et al.*, 2000). In heterothallic Dothideomycetes, individuals carry either *MAT1-1-1* or *MAT1-2-1* at their *MAT* locus, and require the alternate idiomorph in another individual in order to sexually reproduce (Yun *et al.*, 1999, Bennett *et al.*, 2003, Waalwijk *et al.*, 2002, Bihon *et al.*, 2014, Bolton *et al.*, 2014). On the other hand, homothallic fungi generally carry both *MAT1-1-1* and *MAT1-2-1* at the *MAT* locus (Yun *et al.*, 1999, Yun *et al.*, 2013, Marsberg *et al.*, 2016). In recent years, next generation genome sequencing has enabled the rapid identification of conserved *MAT* loci among fungal species (Bihon *et al.*, 2014, Bolton *et al.*, 2014, Marsberg *et al.*,

2013). As such, it is now much easier to identify and characterise the mating system of an emerging phytopathogen.

We sought to identify and describe the composition of the mating-type locus and associated genes in the three wheat infecting *Eutiarosporella* spp. to determine whether they are homo- or heterothallic. These analyses revealed that *E. tritici-australis* is heterothallic whilst both *E. darliae* and *E. pseudodarliae* were homothallic. Sequence analysis revealed that unlike in the *Mat1-1-1* *E. tritici-australis* isoaltes, the *Mat 1-1-1* gene in *E. darliae* and *E. pseudodarliae* was not present in the MAT locus but was embedded within another distal gene. Further analysis of the *Mat1-1-1* location in *E. darliae* and *E. pseudodarliae* uncovered a series of biologically relevant chromosomal rearrangements that prompted additional growth assays. These analyses shed light on the evolution of WGD-*Eutiarosporella* species.

2. Material and methods

2.1 Isolate sampling and DNA extractions.

Wheat grain, reportedly infected with *Eutiarosporella tritici-australis*, was provided by Marg Evans, South Australian Research and Development Institute (SARDI). Fungi were isolated from the surface sterilised grain on tap water agar (TWA). All isolates of fungi used were sub-cultured onto potato dextrose agar (PDA)(Difco, USA) with plugs of mycelia covered agar. These were incubated at 23°C with 12 hour light/dark cycles. Attempts to produce ascospores on wheat (with both pure and mixed isolate inoculum), under 12-hour light/dark cycles. Samples were examined and fruiting structures were screened periodically for over a month. No ascospores were observed from any fruiting

structures in *E. darliae* and *E. pseudodarliae*. DNA was extracted from mycelia scraped from the PDA plates using a QIAGEN DNeasy Plant Mini Kit (QIAGEN, Netherlands). For the MinIon sequencing, high molecular weight DNA was extracted using a modified C-TAB, phenol/chloroform method by Fulton *et al.* (1995). The full protocol detailing all steps and buffers used is publically available at [dx.doi.org/10.17504/protocols.io.hadb2a6](https://doi.org/10.17504/protocols.io.hadb2a6). Fungal mycelia used for the extraction was scraped from the surface of a PDA plate and freeze dried for 24 hrs. Approximately 100mg of dry weight tissue was used as the starting material. DNA was further purified using Agencourt AMPure XP magnetic beads (Beckman Coulter, AUS).

2.2 Genome sequencing and assembly

Illumina paired-end sequencing was performed using a HiSeq™ 2000 (Illumina, USA) in the Bimolecular Resource Facility at The Australian National University. Raw reads were deposited with the NCBI BioSample Database (Accessions: SAMN06075484, SAMN6075485, and SAMN06075486). These reads were quality trimmed using Trimmomatic v0.27 (Lohse *et al.*, 2012) and *de novo* assemblies were generated with SPAdes v2.5.0, with -k mer values 21, 33, 55, and 77 (NCBI BioProject PRJNA400815) (Bankevich *et al.*, 2012). The three Illumina *Eutiarospora* genomes are supplied in Files S1-S3 along with Table S1 which contains genome assembly statistics (Parra *et al.*, 2007). Gene sets for each species were predicted and annotated with the MAKER genome annotation pipeline (Cantarel *et al.*, 2008). Annotations were visualized in within Geneious v7.1.8 (Biomatters, New Zealand). The gene sets from *Botryosphaeria dothidea* v1.0 (http://genome.jgi.doe.gov/Botdo1_1/Botdo1_1.home.html) (Marsberg *et al.*, 2016) and *Parastagonospora nodorum* SN15 v2.0 (<http://genome.jgi.doe.gov/Stano2/Stano2.home.html>) were used for training genes.

PCRs were performed in 20µl reactions using Ex Taq (TaKaRa, Japan) polymerase with the program: (95°C, 2', 35×(90°C, 30'', 58°C, 30'', 72°C, 1')). Sequencing was performed using BigDye terminator (Applied Biosystems, USA) with the standard protocol. Sanger sequencing was performed at the Australian Genome Research Facility at The Australian National University.

For the MinIon sequencing, a high molecular weight DNA library for the Oxford Nanopore MinIon was prepared using 1D SQK-LSK108 Ligation sequencing Kit R9 Version (Oxford Nanopore, UK). Sequencing was performed on an R9.5 flow-cell with MinKNOWv1.6.11 (Oxford Nanopore, UK). Raw fast5 were base called with Albacore v1.2.1, with the following specific options; -c r95_450bps_linear.cfg -r -o fastq. Fastq files were corrected with Canu v1.5 and BLASTn was used to find the *MAT* locus amongst the fastq files (Altschul *et al.*, 1990, Koren *et al.* 2017). The *MAT* locus harbouring read is supplied in File S4 along with a gff file containing annotations.

2.3 Growth assay

The ability of *E. pseudodarliae* to metabolise glycosylated amino acids was assessed by comparing its growth to *E. darliae* on two different agar minimal media (MM) comprised of 10mls/L of 100× trace nutrient stock, 8.3g/L of glucose, 10mM of nitrogen source, and 1.2% of Bacto agar (Difco, USA). The first MM was made with glutamate as the nitrogen source. The second MM was made with lysine as the nitrogen source. We were unable to source pure glycosylated lysine or valine, so instead we boiled glucose+lysine MM to begin the Maillard reaction (indicated by browning of media) (Ajandouz *et al.*, 2001). It was assumed to be a crude mixture of glycosylated and unglycosylated lysine. The lysine MM was boiled for 30 minutes prior to autoclaving to ensure the resulting media is well caramelised. Both MMs were autoclaved to sterilise.

Plugs of mycelia covered agar from PDA plates were inoculated onto plates of each MM. Plates were incubated at 23°C with 12-hour light/dark cycles. The radial growth diameter of the fungi was measured at five days post inoculation (DPI).

2.4 Mating type analyses

The positions of the four putative *MAT* locus genes observed in the *Botryosphaeriaceae* species (*MAT1-1-1*, *MAT1-1-4*, *MAT1-2-1*, and *MAT1-2-5*) (as described by Bihon *et al.* (2014)) were located on the genomes of *M. phaseolina*, *B. dothidea*, *N. parvum*, *E. darliae*, *E. pseudodarliae* (*E. dactylidis*), and *E. tritici-australis* using tBLASTn (Altschul *et al.*, 1990) within Geneious v7.1.8 (Biomatters, New Zealand) with default settings. Annotated *MAT* amino acid sequences for the *Botryosphaeriaceae* spp., *Diplodia pinea* (Bihon *et al.*, 2014) were obtained from the GenBank database (*MAT1-1-1* (gb|AHA91692.1) and *MAT1-2-1* (gb|AHA91681.1)), were used as queries for the search. To compare the syntenic arrangement of the genes surrounding the *MAT* loci, predicted gene sequences from this locus within the *Eutiarosporella* spp.' genomes were used as queries for BLASTn and tBLASTx (no e-value cut-off) searches of each of the other genomes (Altschul *et al.*, 1990). Where appropriate, nucleotide alignments between genomes were performed using Mauve v2.3.1, with default settings as implemented in Geneious v7.1.8 (Darling *et al.* 2004). Mating type gene composition in the *E. tritici-australis* isolates were determined using primers for *MAT1-1-1* and *MAT1-2-1*, designed on the WGD-*Eutiarosporella* genomes (Table S2).

3. Results

3.1 The composition and location of MAT locus associated genes is varied among the Botryosphaeriaceae spp..

The WGD-*Eutiarosporella* spp. produce ascospores in the field (Evans *et al.*, 2015). To characterise the method of sexual reproduction used by these fungi (heterothallism or homothallism), the *MAT* locus, including the surrounding syntenic genes, were identified in the *de novo* assembled genomes of the WGD-*Eutiarosporella* spp. (Thynne *et al.* 2015). The composition of genes at the *MAT* locus in the *Botryosphaeriaceae* sp., *Diplodia pinea* was recently reported, including two additional “*MAT*” genes, *MAT1-1-4* and a novel gene, *MAT1-2-5* (Bihon *et al.*, 2014). These annotated nucleotide sequences were used as tBLASTx queries to identify homologues in the WGD-species. To compare the syntenic arrangement and conservation of this locus within the family, tBLASTx searches were also performed on *Botryosphaeria dothidea* (Marsberg *et al.*, 2016). The composition of genes at *B. dothidea*’s *MAT* locus were also recently reported, classifying this species as homothallic (Marsberg *et al.*, 2016) (Fig. 1). Eight genes, inclusive of, and surrounding the *MAT* locus were used for syntenic-reference among each species (Fig. 1, Fig. S1). These eight genes include a putative DNA lyase (*DNA lyase*), a gene encoding a protein with a COX6A domain (*COX6A*), a gene encoding a protein with an Apc5 domain (*Apc5*), a gene encoding a protein with a DUF 2404 domain (*DUF2404*), *MAT1-2-5*, *MAT1-1-4*,

MAT1-2-1 and *MAT1-1-1*. All eight of the genes discussed above were present on a single scaffold in *B. dothidea* (Fig. 1, Fig. S1). All of the *MAT* locus schematics displayed in Figure 1 are partial windows of the whole scaffold they are located on (aside from the *MAT* locus for *E. tritici-australis* isolate V16-1, described below). The full length scaffolds, with nucleotide lengths denoted, are displayed in Figure S1.

Within the three WGD-*Eutiarospora* spp. there were several additional rearrangements that differed from *B. dothidea*. Seven of the eight genes were present on the same scaffold in *E. tritici-australis* Isolate 153. The exception was *MAT1-1-1*, which was not found within this genome assembly. *E. darliae* also contained seven of the eight conserved genes on a single scaffold. *E. darliae*, however, had one additional predicted gene between *DNA lyase* and *COX6A*. No homologous proteins or gene sequences were identified on the NCBI-NR database when this 204 nucleotide predicted gene was used as a BLASTx and tBLASTx query. Without expression data we are unable to comment on whether this is a real gene or an incorrectly annotated sequence. Unlike *E. tritici-australis* Isolate 153, *MAT1-1-1* was identified in *E. darliae* but was located on a different scaffold to the other seven syntenic genes. In *E. pseudodarliae* these eight genes were dispersed across three different scaffolds (Fig. 1, Fig. S1). *DUF2404*, *MAT1-2-5*, and *MAT1-2-1* were all present in one scaffold and *MAT1-1-4*, *DNA lyase*, *COX6A*, and *Apc5* were all present on a second. *DUF2404* is located 14.5kb from the start of its containing scaffold, with 130kb following *MAT1-2-1* (Fig. S1). Therefore, the two separate sets of genes are located apart on *E. pseudodarliae*'s genome. Finally, *MAT1-1-1* was found on a third scaffold, orthologous to *E. pseudodarliae*'s sister species, *E. darliae*. The location of *MAT1-1-1* in these two species is discussed further below.

3.2 *E. tritici-australis* is a heterothallic species.

Unlike *E. darliae* and *E. pseudodarliae*, no *MAT1-1-1* gene was identified in the genome assembly for *E. tritici-australis* Isolate 153. We examined other isolates of *E. tritici-australis* to determine whether some individuals harbour *MAT1-2-1* and others *MAT1-1-1*. We isolated individuals of *E. tritici-australis* from WGD-infected grain harvested from South Australia. The ITS locus was amplified for each isolate (White *et al.*, 1990) and the product sequenced, confirming that each isolate is *E. tritici-australis* (Table S3). We screened these individuals for the presence of either *MAT1-1-1* or *MAT1-2-1* using primers designed in conserved regions of *E. darliae* and *E. pseudodarliae*'s *MAT* genes (Table S2).

In concordance with the genome assemblies for the species, *MAT1-2-1*, and *MAT1-1-1* were both amplified from *E. darliae* Isolate 2E2 (DAR 82491) and *E. pseudodarliae* Isolate HR599118 (DAR 82489) (Fig. 2A). Similarly, only *MAT1-2-1* was amplified from the genome sequenced isolate *E. tritici-australis* Isolate 153 (DAR 82485). We were unable to amplify *MAT1-1-1* from Isolate 153. Excluding Isolate 153, we used PCR to screen an additional 15 isolates of *E. tritici-australis* for their mating type genes. Four of these 15 isolates were identified as *MAT1-2-1* (Table S3). A representative *MAT1-2-1* amplification product from Isolate V6-1 was sequenced, confirming it as *MAT1-2-1* (GenBank accession: KY587365). The remaining 11 isolates of *E. tritici-australis* were identified harbouring *MAT1-1-1* (Table S3). These are represented in Fig. 2A by Isolates V2-1 and V16-1. The PCR amplification product for these two isolates resolves on the gel slightly higher than *E. darliae* and *E. pseudodarliae*'s *MAT1-1-1*, which have an expected product size of 501bp (Fig. 2A). Sequencing confirmed the product to be *MAT1-1-1* (GenBank accession: KY587366). Subsequent alignment of *E. tritici-australis* *MAT1-1-1* with *E. pseudodarliae* and *E. darliae*'s *MAT1-1-1* gene revealed it to be 63bp longer than

DNA that *E. pseudodarliae* and *E. darliae*'s sequence (Fig. 2B) thus explaining the band size discrepancy observed in Fig. 2A.

We next sought to confirm that the *MAT1-1-1* sequence was a true idiomorph. After several failed attempts at long-range PCR we re-sequenced *E. tritici-australis* Isolate V16-1 with the Oxford Nanopore MinIon. From the corrected-base called reads we were able to recover a single 14.8 kb read that harbours six of the eight *MAT* loci associated genes discussed for the other species (Fig. 1). The two missing genes were *MAT1-2-1* and *Apc5*. Within this read, *MAT1-1-1* is flanked by *MAT1-4-1* and *MAT1-2-5* (Fig. 1, Fig. S1). The arrangement of the additional *MAT* genes surrounding either *MAT1-2-1* or *MAT1-1-1* is similar to the predicted idiomorphs from the heterothallic relative *D. pinea* (Bihon *et al.*, 2014). These data demonstrate that *E. tritici-australis* is heterothallic with isolates harbouring either *MAT1-1-1* or *MAT1-2-1*, whereas *E. darliae* and *E. pseudodarliae* are homothallic.

3.3 Within the genomes of E. darliae and E. pseudodarliae, MAT1-1-1 is harboured within the middle of a transcription factor gene, distal to the MAT locus.

MAT1-1-1 is located within a syntenic location in the genomes of *E. darliae* and *E. pseudodarliae*, distal to the described *MAT* locus. In both species, *MAT1-1-1* is located in the middle of the coding sequence of a gene predicted to encode a protein with a fungal-specific transcription factor domain (*FSTF*, PFAM domain 11951). We are confident in the assembly placement of *MAT1-1-1* to this locus as paired-end reads tile across these genes (SAM files and corresponding fasta files supplied – Files S5.1-S5.2 and S6.1-S6.2). We identified orthologs of the *FSTF* gene, and adjacent genes, on a syntenic scaffold in two other species, *E. tritici-australis* Isolate 153, and *B. dothidea* (Fig. 3A). *MAT1-1-1* was not located at this locus in either of these two species.

In *E. pseudodarliae* there were 459bp of DNA upstream of the start codon for *MAT1-1-1* as well as 927bp downstream of the stop codon that were not part of either the mating type gene or the fungal transcription factor (Fig. 3B,C). We used this surrounding DNA as BLASTn and tBLASTx queries to identify homologous DNA in *E. tritici-australis* isolate 153 and at other loci in *E. pseudodarliae*. 135bp of the DNA upstream of *MAT1-1-1* was homologous to DNA from upstream of *MAT1-1-4* in both species (Fig. 3B). The BLAST-hit in *E. tritici-australis* isolate 153 had an e-value of $1.74e^{-16}$ and shares 72.1% pairwise identity. In *E. pseudodarliae*, the BLAST hit for this locus had an e-value of $7.29e^{-22}$ and a shared 91% pairwise identity with this *MAT1-1-4* locus DNA (Fig. 3B,C). The first 408bp downstream of the *MAT1-1-1* had no significant BLASTn hits (e-value $1e^{-5}$). The next 533bp of the DNA had a BLAST hit that matched DNA from *MAT1-2-5*, and its 5' border, in *E. tritici-australis* Isolate 153 (top e-value $224e^{-30}$). This 533bp region was used as a BLASTn query to search *E. pseudodarliae*'s genome. The two most closely aligned sequences were the 3' border of *MAT1-1-1*, from which the query originated, and *MAT1-2-5*, and its 5' border. The query DNA shares 89% pairwise identity with the latter (Fig. 3B,C). Therefore, the *MAT1-1-1* gene in *E. pseudodarliae* is surrounded by intergenic DNA upstream of *MAT1-1-4* and fragments of *MAT1-2-5*. These results describe a *MAT* locus with a similar gene arrangement to the *MAT1-1-1* idiomorph from *E. tritici-australis* isolate V16-1 (Fig. 3C). *E. pseudodarliae*'s *MAT 1-1-1* locus was used in Figure 3 B and C, and not *E. darlie*'s as the latter species has additional rearrangements at this locus separating the two *FSTF* fragments. We discuss the implications of these gene arrangements below.

3.4 *E. darliae* and *E. pseudodarliae* have highly plastic genomes surrounding their *MAT1-1-1*-loci

Aside from the insertion of *MAT1-1-1* within the *FSTF*, a number of significant gene rearrangements within this genome region were observed. In *E. darliae*, half of the *FSTF* gene and the neighbouring genes are present on a different scaffold (Fig. 3A). Because of the fragmented nature of the genome assembly, we could not determine whether these genes are linked to the scaffold containing the other half of the *FSTF* gene, which is interrupted by *MAT1-1-1*. However, neither half of the *FSTF* are at the end of a scaffold. The half adjacent to *MAT1-1-1* is approximately 12kb from the end of the scaffold (Fig. S1). Similarly, the other half is approximately 6kb into a scaffold. Therefore, even if the scaffolds were linked, the two gene fragments are separated. In *E. pseudodarliae*, the fragments of the *FSTF* gene are still present both up and downstream of the inserted *MAT1-1-1*. However, one of the predicted neighbouring genes was cleaved, with only a partial remnant of the coding sequence left at this original locus. The remainder of the gene was found within the genome assembly syntenic with the other neighbouring genes (Blue checkered gene (*) in Fig. 3A). Interestingly, these rearranged genes are still on the same scaffold, although over 38kb away, and in inverse orientation (Fig 3A, Fig. S1). This cleaved neighbouring gene remains intact in the genome of *E. tritici-australis* and *E. darliae*.

In order to determine the function of the cleaved gene, the predicted coding sequence from *E. tritici-australis* was used as a BLASTx query against the NCBI non-redundant database. Best-BLAST-hits included genes functionally annotated as encoding fructosyl amine oxidases (e-value of $5e^{-166}$ and 63% sequence identity to *Aspergillus oryzae* (Akazawa *et al.*, 2004). Fructosyl amine oxidases are involved in enzymatic deglycation (removal of the sugar) from glycosylated amino acids (Akazawa *et al.*, 2004). Glycosylated amino acids are a product of the Maillard reaction when reducing sugars react with amino acids (Akazawa *et al.*, 2004). As such, we hypothesised that *E. pseudodarliae*'s metabolic capabilities might be altered by the cleavage of this gene. This

gave us an opportunity to functionally test whether the genome rearrangements near the *MAT1-1-1* insertion were real or alternatively mis-assembled regions of the genome. To test this hypothesis the two sister-species, *E. darliae* and *E. pseudodarliae*, were grown on MM containing glycosylated amino acid (boiled glucose and lysine) as a nitrogen source (Ajandouz *et al.*, 2001). These fungi were also grown with glutamate as a nitrogen source as a control and radial growth was measured after five days post inoculation (DPI). There was no significant difference in radial growth of *E. pseudodarliae* and *E. darliae* on MM with glutamate (Fig. S2). However, radial growth was reduced for *E. pseudodarliae* on MM with boiled lysine suggesting a reduced ability to metabolise glycosylated amino acids. This suggests that the genome assembly in this region is correct.

4. Discussion

Much remains unknown about the emergence of WGD as a wheat disease in Australia. The aim of this study was to develop our understanding of the relationship among isolates of the WGD-*Eutiarosporrella* spp. and to determine the composition of mating type genes in these species. According to the composition of *MAT* genes we found that two of the species, *E. pseudodarliae* and *E. darliae* are homothallic, whereas the third species, *E. tritici-australis*, is heterothallic. Most homothallic Dothideomycetes harbour both mating type genes, *MAT1-1-1* and *MAT1-2-1*, at the *MAT* locus. In *E. darliae* and *E. pseudodarliae*, however, *MAT1-1-1* is located distally to the *MAT* locus.

There are other known Dothideomycete species whose *MAT* genes are not restricted to the *MAT* locus. The homothallic species, *Cochliobolus cymbopogonis* harbours both *MAT* genes but in unlinked loci (Yun *et al.*, 1999), similar to our observations with *E.*

darliae and *E. pseudodarliae*. How homothallism arose in *C. cymbopogonis* remains obscure (Yun *et al.*, 1999). Similarly, *MAT*-gene sequences are also observed removed from the *MAT* locus in the heterothallic species, *Cercospora beticola* (Bolton *et al.*, 2014). In this species each isolate harbours fragments of *MAT1-1-1* and *MAT1-2-1* throughout the genome, in addition to a complete *MAT* locus (Bolton *et al.*, 2014). These fragments of the *MAT* genes, however, did not contain introns, providing evidence that they were likely integrated from an RNA transcript (Bolton *et al.*, 2014). In our analysis, we showed that *MAT1-1-1* is distal to the *MAT* locus in *E. darliae* and *E. pseudodarliae*. However, unlike in *C. beticola*'s dispersed *MAT* genes, the two *Eutiarosporella* spp.'s *MAT1-1-1* retained their intron sequence. Non-coding DNA homologous to DNA from the *MAT* locus is also present at the new locus, as well as partial fragments for the secondary *MAT*-gene, *MAT1-2-5*, discussed further below.

E. darliae and *E. pseudodarliae* are sister species that share very similar morphological characteristics (Thynne *et al.*, 2015). Herein, we show that these species carry *MAT1-1-1* in a syntenic genome location, within a fungal specific transcription factor. Therefore, the insertion of *MAT1-1-1* to its new locus most likely occurred in a common ancestor. Another *Eutiarosporella* spp., *E. dactylidis*, an isolate identified from dead grass in (Thambugala *et al.*, 2014) resolves in a monophyletic clade with *E. pseudodarliae* (comparing the ITS and EF1- α) (Dissanayake, 2016). Due to their close relationship, it is likely that *E. dactylidis* also shares *MAT1-1-1* insertion, but with no genome sequence for *E. dactylidis* currently available, this cannot be confirmed.

However, the sexual state has been observed for *E. dactylidis* in single-spored culture, indicating that this species is functionally homothallic, which provides circumstantial evidence to support our hypothesis.

In comparison with *E. darliae* and *E. pseudodarliae*, the third WGD species, *E. tritici-australis* is heterothallic. Heterothallism has been demonstrated to be the ancestral state

in other Dothideomycete genera where both heterothallic and homothallic species are observed. Yun *et al.* (1999) described the *MAT* loci of a number of homothallic and heterothallic *Cochliobolus* spp.. In this study they show that the organisation of genes at the *MAT* loci within heterothallic spp. was conserved, whereas the organisations of genes at the *MAT* loci of the homothallic spp. were each unique. Based on these results, they proposed that homothallism arose from an unequal crossing-over event, resulting in both *MAT* genes present within a single individual at the same *MAT* locus (Yun *et al.*, 1999). Such events are considered to be infrequent as sex determination genes across eukaryotic organisms are generally in regions of the genome which have suppressed recombination (Idnurm, 2011). Unlike the homothallic species reported by Yun *et al.* (1999), the homothallic-like species *E. darliae* and *E. pseudodarliae* do not appear to carry their *MAT1-1-1* genes near the *MAT* locus. Below, we consider an alternative model whereby *MAT1-1-1* and surrounding DNA originated from a heterothallic *MAT1-1-1* *MAT* locus (Fig. 3).

In *B. dothidea*, a homothallic *Botryosphaeriaceae* spp. (Marsberg *et al.*, 2016), *MAT1-2-1* is downstream from *MAT1-1-1* (Fig. 1). The gene arrangement at the *MAT* locus in each of *E. darliae* and *E. tritici-australis* is syntenic with the *MAT* locus genes in *B. dothidea*, aside from lacking *MAT1-1-1*. Hypothetically, if *MAT1-1-1* was once present in the WGD-species' *MAT* loci, it would likely have originated from downstream of *MAT1-2-1*. However, there is no fragment of *MAT1-2-1* at the new *MAT1-1-1* locus, only fragments of *MAT1-1-4* and *MAT1-2-5* (Fig. 3B,C). The gene arrangement of *MAT1-1-4* followed by *MAT1-1-1*, and followed by *MAT1-2-5* is strikingly similar to that observed in the *MAT* loci of the heterothallic WGD species, *E. tritici-australis* (Fig. 3C). The gene arrangement at the *MAT1-1-1* idiomorph in *E. tritici-australis* isolate V16-1 was demonstrated as *MAT1-1-4* followed by *MAT1-1-1* followed by *MAT1-2-5* (Fig. 1, Fig. 3C). The *MAT1-2-1* idiomorph in isolate 153 contained *MAT1-1-4* followed by *MAT1-2-1*, followed by *MAT1-*

2-5 (Fig. 1, Fig. 3C). The *MAT1-1-1* orientation in isolate V16-1 is similar to the gene arrangement observed at the *MAT1-1-1* locus of *E. pseudodarliae* (Fig. 3C). Similarly, the *MAT1-2-1*-locus is similar between *E. darliae* and *E. tritici-australis* (Fig. 1, Fig. 3C). As such, the two *MAT* locus variants from the homothallic WGD-species (Fig. 3C; *MAT1-2-1* represented by *E. darliae* and *MAT1-1-1* represented by *E. pseudodarliae*) resemble the *MAT* loci of two alternative heterothallic individuals. We hypothesise that the insertion of *MAT1-1-1* (and surrounding DNA) in *E. darliae* and *E. pseudodarliae* arose via the integration of DNA from a *MAT1-1-1* heterothallic individual. This is very similar to other Ascomycete fungi such as *Neosartorya fischeri* and *Eupenicillium crustaceum* (Rydholm *et al.*, 2007, Poggeler *et al.*, 2011). In both these species, the two *MAT* idiomorphs are not linked, and it was proposed that homothallism arose in each from recombination between two heterothallic individuals (Rydholm *et al.*, 2007, Poggeler *et al.*, 2011). The proposed model is that recombination during meiosis resulted in the transition from heterothallism to homothallism, citing similarity to the event proposed in the generation of homothallic *Cochliobolus* species (Yun *et al.*, 1999, Rydholm *et al.*, 2007). However, instead of an unequal crossing-over resulting in linked or fused *MAT* genes (Yun *et al.*, 1999), during meiosis DNA from one *MAT* idiomorph was translocated to a new, distal locus (Rydholm *et al.*, 2007). We hypothesise that this is likely how the homothallic *Eutiarosporella* lineages were generated.

Because *E. tritici-australis* remains a heterothallic species, credence is lent to the hypothesis that all of these species might have once been heterothallic. Figure 4 presents a model as to how the three WGD species may have arisen from a single ancestral heterothallic population. After the integration of the *MAT1-1-1* DNA, *E. darliae* and *E. pseudodarliae*'s ancestral species underwent sympatric speciation through selfing. In comparison, *E. tritici-australis* remained heterothallic. Further examination of the mating type systems in closely related *Eutiarosporella* spp. could shed light on the

evolution of mating type systems in the genus. It will also be interesting to determine whether transition to homothallism has provided a tangible lifestyle benefit to the species of this lineage. Perhaps a benefit is due to comparative ease of spread and sexual propagation to new locations, as homothallic species have the potential to generate a new, sexual population from a single spore.

While the insertion point of the *MAT1-1-1* gene is conserved between *E. darliae* and *E. pseudodarliae* there are a number of significant within scaffold rearrangements surrounding the new *MAT1-1-1*-locus. The inversions and gene rearrangements observed among these species are similar to the intra-chromosomal rearrangements described as mesosyteny (Hane *et al.*, 2011). Mesosyntenic variation is frequently observed among Dothideomycete species (Hane *et al.*, 2011, Goodwin *et al.*, 2011, Ohm *et al.*, 2012) and so it is consistent that mesosyteny would be observed in *Botryosphaeriaceae* spp.. The mechanisms that lead to mesosytenic gene rearrangement remain obscure although it is speculated that the process occurs during meiosis (Hane *et al.*, 2011). The lineage specific scaffold rearrangements in *E. darliae* and *E. pseudodarliae* at this locus bear the clear hallmarks of mesosyteny, with clear intra-chromosomal rearrangements.

An interesting parallel observation was that in *E. pseudodarliae*, but not *E. darliae*, these rearrangements cleaved a putative fructosyl amine oxidase. In *E. darliae*, intrachromosomal rearrangements moved the 5' *FSTF* fragment along with the entire fructosyl amine oxidase to a new locus, represented on a different scaffold to the *MAT1-1-1* and 3' *FSTF* fragment harbouring scaffold. (Fig. S1(*)). In contrast, within the genome of *E. pseudodarliae*, both *FSTF* fragments remained localised with *MAT1-1-1*. A 5' fragment of the fructosyl amine oxidase also remained at this locus, whereas the remainder of the gene was rearranged. Both *FSTF* fragments are represented on the

same scaffold (Fig. S1(**)). A functionally characterised homologue of this gene in *Aspergillus oryzae* is required for metabolism of fructosyl-valine and fructosyl-lysine (Akazawa *et al.*, 2004). These are products of the Maillard reaction, when reducing sugars glycosylate amino acids (Akazawa *et al.*, 2004). Our growth experiment on caramelised glucose and lysine MM demonstrated that *E. pseudodarliae* has reduced growth on this medium. Although this data is preliminary, it provides initial evidence that these gene rearrangements altered *E. pseudodarliae*'s metabolic capabilities. We have not observed *E. pseudodarliae*-specific growth impairment under other conditions. It is often assumed that species with larger and/or more equipped genomes may be better suited to survive under altered conditions or have greater adaptive potential (Raffaele and Kamoun, 2012). Gene rearrangements could have unknown consequences and may have reduced the ability of these *Eutiarosporella* spp. to overcome certain evolutionary challenges. However, it remains unknown if the disruption of the fructosyl amine oxidase has affected *E. pseudodarliae*'s fecundity or adaptive potential.

5. Conclusions

We have shown that both *E. darliae* and *E. pseudodarliae* each possess *MAT1-1-1* and *MAT1-2-1* genes and conclude that these species are homothallic. This is despite the fact that *MAT1-1-1* is not localised within the *MAT* locus. In contrast, the third WGD-species, *E. tritici-australis* is likely heterothallic, with isolates harbouring either *MAT1-1-1* or *MAT1-2-1* at the *MAT* locus. We also present evidence that *E. pseudodarliae* and *E. darliae* potentially transitioned to homothallism via a cryptic DNA integration event between two heterothallic individuals. This opens interesting new questions about

fungus speciation via genome adaptation events between closely related species or between individuals of the same species, as well as the genetic events that precede them.

Acknowledgments

The authors thank Joseph Spatafora and Bernard Slippers for allowing use of the *B. dothidea* genome prior to publication (Marsberg *et al.*, 2016). The *B. dothidea* genome was sequenced by the NSF-Funded AFTOL2 project (DEB-0732993). The authors also acknowledge Marg Evans (SARDI) for supplying WGD infected grain for analysis and for personal communications relating to sexual reproduction within the WGD-*Eutiarosporella* spp.. ET is funded through a Grains Research and Development Corporation Grains Research Scholarship.

References

- Akazawa, S.-I., Karino, T. Yoshida, N., Katsuragi, T., Tani, Y., 2004. Functional analysis of fructosyl-amino acid oxidases of *Aspergillus oryzae*. Appl. Environ. Microbiol. 70, 5882-5890.
- Ajandouz, E.H., Tchiakpe, L.S., Dalle Ore, F., Benajiba, A., Puigserver, A., 2001. Effects of pH on caramelization and caillard reaction kinetics in fructose-lysine model systems. J. Food. Sci. 66, 926-931.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–10.
- 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. *J. Comp. Biol.* 19, 455-477.
- Bennett, R.S., Yun, S.-H., Lee, T.Y., Turgeon, B.G., Arseniuk, E., Cunfer, B.M., Bergstrom, G.C., 2003. Identity and conservation of mating type genes in geographically diverse isolates of *Phaeosphaeria nodorum*. *Fungal Genet. Biol.* 40, 25-37.
- Bihon, W., Wingfield, M.J., Slippers B., Duong, T.A., Wingfield, B.D., 2014. *MAT* gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. *Fungal Genet. Biol.* 62, 55-61.
- Blanco-Ulate, B., Rolshausen, P., Cantu, D., 2013. Draft genome sequence of *Neofusicoccum parvum* isolate UCR-NP2, a fungal vascular pathogen associated with grapevine cankers, *Genome Announc.* 1: e00339-13.
- Bolton, M.D, de Jonge, R., Inderbitzin, P., Liu, Z., Birla, K., Van de Peer, Subbarao, K.V., Thomma, B. P., Secor, G.A., 2014. The heterothallic sugarbeet pathogen *Cercospora beticola* contains exon fragments of both *MAT* genes that are homogenized by concerted evolution. *Fungal Genet. Biol.* 62, 43-54.
- Cantarel, B.L., Korf, I., Robb, S.M., Parra, G., Ross, E., Moore, B., Holt, C., Alvarado, A.S., Yandell, M., 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* 18, 188-196.
- Crous, P.W., Slippers, B., Wingfield, M.J., Rheeder, J., Marasas, W.F., Philips, A.J., Alves, A., Burgess, T., Barber, P., Groenewald, J.Z., 2006. Phylogenetic lineages in the *Botryosphaeriaceae*. *Stud. Mycol.* 55, 235-253.
- Crous, P.W., Muller, M.M., Sanchez, R., Giordano, L., Bianchinotti, M.V., Anderson, F.E., Groenewald, J.Z., 2015. Resolving *Tiarosporella* spp. allied to *Botryosphaeriaceae* and *Phacidiaceae*, *Phytotaxa* 202, 73-93.

- Darling, A.C., Mau, B., Blattner, F.R. and Perna, N.T., 2004. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394-1403.
- Dissanayake, A.J., Phillips, A.J.L., Li, X.H., Hyde, K.D., (2016) *Botryosphaeriaceae*: current status of genera and species. *Mycosphere.* 7, 1001-1073.
- Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792-1797.
- Evans, M., 2013. White grain rejection spreading. *Ground Cover* 102
- Evans, M., Walter, J., Wallwork, H., 2015. Pathogen of the month: *Eutiarosporella* Crous species. *Australas. Plant Pathol.* October, 15
<https://www.appsnet.org/Publications/potm/pdf/oct15.pdf>.
- Fulton, T.M., Chunwongse, J., Tanksley, S.D., 1995. Microprep protocol for extraction of DNA from tomato and other herbaceous plants. *Plant Mol. Biol.* 13, 207-209.
- Goodwin, S.B., M'Barek, S.B., Dhillon, B., Wittenberg, A.H., Crane, C.F., Hane, J.K., Foster, A.J., Van der Lee, T.A., Grimwood, J., Aerts, A., Antoniw, J., Bailey, A., Bluhm, B., Bowler, J., Bristow, J., van der Burgt, A., Canto-Canche, B., Churchill, A.C., Conde-Ferraez, L., Cools, H.J., Coutinho, P.M., Csukai, M., Dehal, P., De Wit, P., Donzelli, B., van de Geest, H.C., van Ham, R.C., Hammond-Kosack, K.E., Henrissat, B., Kilian, A., Kobayashi, A.K., Koopmann, E., Kourmpetis, Y., Kuzniar, A., Lindquist, E., Lombard, V., Maliepaard, C., Martins, N., Mehrabi, R., Nap, J.P., Ponomarenko, A., Rudd, J.J., Salamov, A., Schmutz, J., Schouten, H.J., Shapiro, H., Stergiopoulos, I., Torriani, S.F., Tu, H., de Vries, R.P., Waalwijk, C., Ware, S.B., Wiebenga, A., Zwiers, L.H., Oliver, R.P., Grigoriev, I.V., Kema, G.H., 2011. Finished genome of the fungal wheat pathogen *Mycosphaerella graminicola* reveals dispensome structure, chromosome plasticity, and stealth pathogenesis. *PLoS Genet.* 7, e10020.

- Hane, J.K., Rouxel, T., Howlett, B.J., Kema, G.H., Goodwin, S.B., Oliver, R.P., 2011. A novel mode of chromosomal evolution peculiar to filamentous *Ascomycete* fungi. *Genome Biol.* 12, R45.
- Idnurm, A., 2011. Sex and speciation: the paradox that non-recombining DNA promotes recombination. *Fungal Biol. Rev.* 25, 121-127.
- Islam, M.S., Haque, M.S., Islam, M.M., Emdad, E.M., Halim, A., Hossen, Q.M., Hossain, M.Z., Ahmed, B., et al., 2012. Tools to kill: genome of one of the most destructive plant pathogenic fungi *Macrophomina phaseolina*. *BMC Genomics* 13, e493.
- Jami, F., Slippers, B., Wingfield, M.J., Gryzenhout, M., (2012). Five new species of the *Botryosphaeriaceae* from Acacia Karroo in South Africa. *Cryptogam., Mycol.* 33, 245-266.
- Jami, F., Slippers, B., Wingfield, M.J., Gryzenhout, M., 2013. Greater *Botryosphaeriaceae* diversity in healthy than associated diseased Acacia karroo tree tissues. *Australas. Plant Pathol.* 42, 421-430.
- Jami, F., Slippers, B., Wingfield, M.J., Gryzenhout, M., 2014. *Botryosphaeriaceae* species overlap on four unrelated, native South African hosts. *Fungal Biol.* 118, 168-179.
- Lohse, M., Bolger, A.M., Nagel, A., Fernie, A.R., Lunn, J.E., Stitt, M., Usadel, B., 2012. RobiNA: a user-friendly, integrated software solution for RNA-seq-based transcriptomics. *Nucleic Acids Res.* 40, W622-W627.
- Koren, S., Walenz, B.P., Berlin, K., Miller J.R., Bergman N.H., Phillippy A.M., 2017. Canu: scalable and accurate long-read assembly via adaptive *k*-mer weighting and repeat separation. *Genome Res.* 27, 722-736.
- Leigh, J.W., Susko, E., Baumgartner, M., Roger, A.J., 2008. Testing congruence in phylogenomic analysis. *Syst. Biol.* 57, 104-115.
- Marsberg, A., Kemler, M., Jami, F., Nagel, J.H., Postma-Smidt, A., Naidoo, S., Wingfield, M.J., Crous, P.W., Spatafora, J., Hesse, C.N., Robbertse, B., Slippers, B., 2016.

- Botryosphaeria dothidea*: A latent pathogen of global importance to woody plant health. *Mol. Plant Pathol.* 18, 477-488.
- McDonald, B.A., Linde, C., 2002. Pathogen population genetics, evolutionary potential, and durable resistance. *Annu. Rev. Phytopathol.* 40, 349–379
- Metzenberg, R.L., Glass, L.N., 1990. Mating type and mating strategies in *Neurospora*. *Bioessays* 12, 53–59.
- Ohm, R.A., Feau, N., Henrissat, B., Schoch, C.L., Horwitz, B.A., Barry, K.W., Condon, B.J., Copeland, A.C., Dhillon, B., Glaser, F., Hesse, C.N., Kosti, I., LaButti, K., Lindquist, E.A., Lucas, S., Salamov, A.A., Bradshaw, R.E., Ciuffetti, L., Hamelin, R.C., Kema, G.H., Lawrence, C., Scott, J.A., Spatafora, J.W., Turgeon, B.G., de Wit, P.J., Zhong, S., Goodwin, S.B., Grigoriev, I.V., 2012. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes fungi. *PLoS Pathog.* 8, e1003037.
- Parra, G., Bradnam, K., Korf, I., 2007. CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics.* 23, 1061-1067.
- Poggeler, S., O’Gorman C.M., Hoff, B., Kuck, U., 2011. Molecular organization of the mating-type loci in the homothallic Ascomycete *Eupenicillium crustaceum*. *Fungal Biol.* 115, 615-624.
- Raffaele, S., Kamoun, S., 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nat. Rev. Microbiol.* 10, 417-430.
- Rydholm, C., Dyer, P.S., Lutzoni, F., 2007. DNA sequence characterisation and molecular evolution of *MAT1* and *MAT2* mating-type loci of the self-fertile compatible ascomycete mold *Neosartorya fischeri*. *Eukaryotic Cell* 6, 868-874.
- Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688-90.

- Sutton, B.C., Marasas, W.F.O., 1976. Observations on *Neottiosporina* and *Tiarosporella*.
Trans. Br. Mycol. Soc. 67, 69-76.
- Thambugala, K.M., Daranagama, D.A., Camporesi, E., Singtripop, C., Liu, Z.-Y, Hyde, K.D.,
2014. Multi-locus phylogeny reveals the sexual state of *Tiarosporella* in
Botryosphaeriaceae. Cryptogam., Mycol. 35, 359-367.
- Thomas, G., Jayasena, K., 2015. White grain disorder of wheat in Western Australia.
https://www.giwa.org.au/pdfs/CR_2015/SORT/EOI80_Thomas_Geoff_White_grain_disorder-paper-revised_CU15_EOI80_.pdf
- Thynne, E., McDonald, M.C., Evans, M., Wallwork, H., Neate, S., Solomon, P.S., 2015. Re-classification of the causal agent of white grain disorder on wheat as three separate species of *Eutiarosporella*. Austral. Plant Pathol. 44, 527-539.
- Thynne, E., Saur, I.M., Simbaqueba, J., Ogilvie, H.A., Gonazalez-Cendales, Y., Mead, O., Taranto, A., Catanzariti, A.M., McDonald M.C., Schwessinger, B., Jones, D.A., Rathjen, J.P., Solomon, P.S., 2016. Fungal phytopathogens encode functional homologues of plant rapid alkalisation factor (RALF) peptides. Mol. Plant. Pathol. 18, 811-824.
- Turgeon, B.G., Yoder, O.C., 2000. Proposed nomenclature for mating type genes in filamentous ascomycetes. Fungal Genet. Biol. 31, 1-5.
- White, T.J., Bruns, T., Lee, S., Taylor, J.W., 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis, M.A., Gelfand, D.H., Sninsky, J.J., White, T.J. (Eds.), PCR protocols: a guide to methods and applications. Academic Press, Inc., New York, pp. 315-322.
- Waalwijk, C., Mendes, O., Verstappen, E.C.P., de Waard, M.A., Kema, G.H.J., 2002. Isolation and characterization of the mating type idiomorphs from the wheat septoria leaf blotch fungus *Mycosphaerella graminicola*. Fungal Genet. Biol. 35, 277-286.

- Wildermuth, G., Williamson, P., Shivas, R., McNamara, R., 2001. Premature head blight and white grain – a new disease of wheat. Proc. of 13th Biennial Austral. Plant Pathol. Conference
- Yoshida, K., Burbano, H.A., Krause, J., Thines, M., Weigel, D., Kamoun, S., 2014. Mining herbaria for plant pathogen genomes: back to the future. PLoS pathog. 10, e1004028.
- Yun, S.-H., Berbee, M.L., Yoder, O., 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., Arie, T., Kaneko, I., Yoder, O., Turgeon, B.G., 2000. Molecular organization of mating type loci in heterothallic, homothallic, and asexual *Gibberella/Fusarium* species. Fungal Genet. Biol. 31, 7-20.
- Yun, S.H., Yoder, O.C., Turgeon, B.G., 2013. Structure and function of the mating-type locus in the homothallic ascomycete, *Didymella zea-maydis*. J. Microbiol. 51, 814-820.

Fig. 1. Graphical representation of syntenic genes shared at the *MAT* locus among a selection of *Botryosphaeriaceae* spp. Eight genes were used for syntenic reference among these species, all eight of which are present in *B. dothidea*'s *MAT* locus. These genes are in the legend at the bottom of the figure. Each of these genes is shown on the partial scaffold/s where they were found in the WGD species. Scaffolds are shown in black, arrows on the end of the contigs indicate there is more sequence data available, which is not shown. Blunt lines denote contig ends.

Fig. 2. *E. tritici-australis* is heterothallic. A) An isolate of each of *E. darliae* and *E. pseudodarliae* were PCR screened for *MAT1-1-1* and *MAT1-2-1*. Expected product sizes for *MAT1-1-1* are 501bp in size (PCR band A) and 901bp for *MAT1-2-1* (PCR band B). Isolates of *E. tritici-australis* represented in this figure by Isolate 153 (*MAT1-2-1*), Isolate V2-1 (*MAT1-1-1*, PCR band B), and Isolate V16-1 (*MAT1-1-1*, PCR band A). B) A 448 bp alignment of partial *MAT1-1-1* nucleotide sequences obtained from sanger sequencing of *E. tritici-australis* (Isolate V6-1) and the genome sequences of *MAT1-1-1* *E. pseudodarliae* (Isolate V4B6), and *E. darliae* (Isolate 2G6). *E. tritici-australis* has an additional 63bp of DNA when compared to *E. pseudodarliae* and *E. darliae*.

Fig. 3. *MAT1-1-1* in *E. darliae* and *E. pseudodarliae* are located within the middle of a fungal specific transcription factor (*FSTF*). A) The location of *MAT1-1-1* (red) in *E. darliae* and *E. pseudodarliae* and surrounding genes. *B. dothidea* and *E. tritici-australis* share three syntenic genes at this locus. The partial scaffolds where the genes are located are shown in black, arrows on the end of the contigs indicate there is more sequence data, which is not shown. The grey annotation overlaying the fructosyl amine oxidase and neighbouring gene indicates that these two genes shared a single annotation in the initial *in silico* gene prediction. B) DNA upstream and downstream of

MAT1-1-1 in *E. pseudodarliae* is homologous to *MAT* locus DNA. C) Schematic representation of the *MAT* idomorphs of heterothallic *E. tritici-australis* compared against the *MAT 1-2-1* locus of *E. darliae* and the *MAT1-1-1*-locus of *E. pseudodarliae*. *E. darliae* and *E. pseudodarliae* each possess both idiomorphs. However, there are rearrangements in each species' alternate idiomorph locus (*E. darliae*'s *MAT1-1-1* locus, and *E. pseudodarliae*'s *MAT 1-2-1* locus) that makes them unsuitable for comparisons of gene synteny. As such they are excluded from this figure. Square blocks represent partial sequences. Blue square block in *E. pseudodarliae* represents partial sequence upstream of *MAT1-1-4*, not coding sequence.

Fig. 4 Diverging mating type systems generated a new lineage of *Eutiarosporella*

spp. A) Schematic representation of the DNA integration of *MAT1-1-1* resulting in a homothallic offspring (concept adapted from Rydholm *et al.*, 2007). (1) Sexual recombination occurred between a *MAT1-1-1* harbouring individual and a *MAT1-2-1* harbouring individual. (2) A DNA breakage or transposition event results in excision of DNA containing *MAT1-1-1* and surrounding. (3) Excised *MAT1-1-1* is integrated within the coding region of the *FSTF* gene, (4) which ends up within the same offspring genome as the *MAT1-2-1* harbouring *MAT* locus. Square blocks represent partial sequences. Blue square block represents partial sequence upstream of *MAT1-1-4*, not coding sequence. Dotted line denotes DNA from the same genome, but not necessarily on the same chromosome. B) Schematic representation of an ancestral heterothallic *Eutiarosporella* spp. from which modern WGD spp. arose. *E. tritici-australis* represents a modern heterothallic *Eutiarosporella* spp. from this ancestral population. A cryptic DNA integration event between two heterothallic individuals with alternate *MAT* idiomorphs is predicted to be the event that generated a homothallic lineage. *E. darliae* and *E. pseudodarliae* represent modern homothallic *Eutiarosporella* spp. from this lineage.

Supplementary Figure 1. Whole scaffolds harbouring *MAT* and associated genes from *B. dothidea*, *E. tritici-australis*, *E. darliae*, and *E. pseudodarliae*. Each of the eight genes discussed in *Fig. 1* are labelled on these scaffolds (if present). Gene blocks from this locus are highlighted in red. In addition, scaffolds harbouring genes associated with *E. darliae* and *E. pseudodarliae*'s *MAT1-1-1* locus (the fungal specific transcription factor gene and the fructosyl amine oxidase gene), are labelled. Gene blocks from this locus are highlighted in purple. The * highlights rearrangement of 5' *FSTF* fragment and fructosyl amine oxidase from the *MAT1-1-1* locus, to different scaffold in the genome assembly of *E. darliae*. The ** highlights rearrangement of the 3' fragment of the fructosyl amine oxidase from the *MAT1-1-1* locus, to a new position on the same scaffold in *E. pseudodarliae*.

Supplementary Figure 2. *E. pseudodarliae* displays more reduced radial growth on caramelised glucose/lysine minimal media (MM), than *E. darliae*. A) Growth of *E. darliae* and *E. pseudodarliae* five days post inoculation (DPI) on two different MM; one with glucose and glutamate as the respective carbon and nitrogen source, and the second with glucose and lysine. The four plates for each of the species represent replicates of the same isolate (*E. pseudodarliae* isolate HR599118 and *E. darliae* isolate 2E2). B) Radial growth diameter measurements for both species, on both media, quantifies the reduction of growth of *E. pseudodarliae* on glucose+lysine MM. Each growth assay consisted of four replicates. Standard errors bars are shown.

Figure 2

B

Figure 3

A

B

E. pseudodarliae

C

Figure 4

A

B

Highlights

- *E. darliae* and *E. pseudodarliae* harbour both mating type genes and are homothallic. However, *MAT1-1-1* is located distal to the MAT-locus.
- Individuals of *E. tritici-australis* harbour either *MAT1-1-1* or *MAT1-2-1* and are therefore heterothallic.
- Switch to homothallism in a common ancestor of *E. darliae* and *E. pseudodarliae* likely facilitated further speciation.