Mating type and pheromone genes in the *Gibberella fujikuroi* species complex: an evolutionary perspective

by

SIMON HENRY MARTIN

Submitted in partial fulfilment of the requirements for the degree

MAGISTER SCIENTIAE

in the Faculty of Natural & Agricultural Sciences, Department of Genetics, University of Pretoria, Pretoria, South Africa

February 2011

SUPERVISOR: PROF. B.D. WINGFIELD

CO-SUPERVISORS: PROF. E.T. STEENKAMP AND PROF. M. J. WINGFIELD

© University of Pretoria

DECLARATION

I, the undersigned, hereby declare that the dissertation submitted herewith for the degree *Magister Scientae* to the University of Pretoria, contains my own independent work and has not been submitted hitherto for any degree at this or any other University.

Simon H. Martin
February 2011

For my father, Claude Martin, and my grandfather, Gustav Hoexter,
who have fuelled my passion for science, and equipped me (*via* biological inheritance and otherwise) with the tools necessary to pursue this dream.

ACKNOWLEDGEMENTS

The completion of this dissertation has only been possible thanks to the support and guidance of a number of remarkable people. My supervisor, Brenda Wingfield is to thank for the fact that this process began at all. Brenda's supervision has been a perfect combination of professionalism, passion and compassion. My co-supervisors, Emma Steenkamp and Mike Wingfield, have enabled me to produce a work of which I can be most proud. Through our long discussions, I have learnt more from Emma than anybody else during this Journey. If I have managed to assemble any ounce of logical reasoning in this dissertation, I owe it to Emma's guidance and Mike's critical evaluation of each and every paragraph.

I am grateful to the National Research Foundation (NRF), the University of Pretoria, the Tree Protection Cooperative Programme (TPCP) and the DST/NRF Centre of Excellence in Tree Health Biotechnology (CTHB), for financial support.

In FABI, the Forestry and Agricultural Biotechnology Institute, I have had access to extraordinary facilities and people. Several 'Fabians' have provided useful technical and academic insights, especially Quentin Santana and Martin Coetzee. However, most important has been Vivienne Clarence (my life-line from the lab to the "real world") whose contributions, big and small, toward my achieving this goal are innumerable.

My friends and family have offered unwavering support, despite my ineptitude in expressing just what it is that I do. In particular, I thank my lovely wife Nicci, whose eternal optimism and faith have made it easy to overcome even the worst disappointments.

Finally, for all things beyond my control (basically everything), I thank God, because

"*In his heart a man plans his course,*
but the Lord determines his steps."

The Proverbs of Solomon, 16:9

CONTENTS

PREFACE 1

CHAPTER ONE
Genetics and evolution of mate-recognition mechanisms in ascomycete fungi and their potential association with species boundaries 3

Introduction 4

Mating type 6
Genes encoded by the *MAT* idiomorphs and functions 7
Could *MAT* genes cause reproductive barriers? 8

MAT loci and mating systems 9
Various sources of self-fertility 10
Heterothallism to homothallism, or *vice versa*? 11
Could mating systems cause reproductive barriers? 12

Mating pheromones and receptors 13
Pheromones and receptors in heterothallic species 14
Pheromones and receptors in homothallic species 14
Could pheromones and receptors cause reproductive barriers? 16

Evolution of mate recognition and species boundaries 17
Ascomycete mate recognition genes evolve rapidly 17
Evolutionary forces and mate recognition mechanisms 17

Summary and conclusions 21

References 22

CHAPTER TWO

Structure and evolution of the *Fusarium* mating type locus: new insights from the *Gibberella fujikuroi* complex **27**

Abstract 28

Introduction 29

Materials and Methods 32

Fungal isolates 32

DNA sequencing 32

Structural organization and conserved regions in the *MAT* loci 33

Phylogenetic relationships 34

Sequence evolution in coding and non-coding regions 34

Results 36

Structural organization and conserved regions in the *MAT* loci 36

Phylogenetic relationships 38

Sequence evolution in coding and non-coding regions 38

Discussion 39

References 44

Tables 48

Figures 53

CHAPTER THREE

Causes and consequences of variability in peptide mating pheromones in ascomycete fungi **61**

Abstract 62

Introduction 63

Materials and Methods 65
Fungal isolates 65
Sequences and organization of pheromone precursor genes 66
Inter-specific variation and tests for positive selection 67
Relationships among tandemly-repeated α-factor-like pheromone sequences 68

Results 69
Sequences and organization of pheromone precursor genes 69
Inter-specific variation and tests for positive selection 72
Relationships among tandemly-repeated α-factor-like pheromone sequences 72

Discussion 73

References 77

Tables 82

Figures 84

SUMMARY 99

PREFACE

Genes involved in reproduction and mate recognition have become a major focus in evolutionary biology. This is largely due to the direct relationship between sexual isolation and speciation (the emergence of new species). In ascomycete fungi, the relative simplicity of the sexual compatibility system, coupled with the fact that these fungi display a variety of different reproductive strategies, offers a useful framework in which to test evolutionary hypotheses. Furthermore, sexual reproduction is of major relevance in fungal invasion and pathogenicity. Despite this appeal, the evolutionary biology of fungal reproductive genes has only recently received much attention. This dearth of knowledge has been the inspiration for this dissertation.

The first chapter of this dissertation is a review of the literature concerning the major molecular components of mate recognition in Ascomycetes. These are the mating type or *MAT* locus and the pheromone/receptor system. The function and evolution of these components is interrogated with the following question in mind: could the evolutionary modifications of these molecular mechanisms have been associated with the formation of reproductive boundaries between species? The review concludes with an exploration of the evolutionary forces that could cause Ascomycete reproductive genes to diverge over time.

The two research chapters describe experimental studies in which comparative methods were used to address questions regarding the evolution of Ascomycete reproductive genes. Both of these studies are centred on the *Gibberella fujikuroi* species complex, a group of morphologically similar filamentous Ascomycetes in the genus *Fusarium*. This complex has been of much interest to plant pathology as it includes many serious plant-pathogenic and mycotoxin producing species. Some species in the *G. fujikuroi* complex are not easily distinguished from one another, and were thus initially recognized as distinct "mating populations" based on sexual inter-compatibility (biological species recognition). Hence, the existing reproductive barriers among these species are well documented.

In the first study (Chapter Two), a high-throughput sequencing method was used to efficiently sequence the complete *MAT* loci from eleven *G. fujikuroi* species. These "heterothallic" (self-sterile) *MAT* loci were compared with the "homothallic" (self-fertile) *MAT* loci of species in the related *Fusarium graminearum* species complex. In addition to

the discovery of a previously unknown *MAT* gene, this study has shed light on the interesting evolutionary forces that affect *MAT* loci, and how they differ between heterothallic and homothallic species.

In the second study (Chapter Three), the mating pheromone genes and the pheromone peptides they produce were sequenced from twenty four isolates representing twelve species of the *G. fujikuroi* complex. These were compared with previously described pheromones and pheromone genes identified in available genome sequences, covering another fifty eight Ascomycetes. This study revealed the rich diversity of Ascomycete pheromones, which has resulted from both adaptive and non-adaptive evolution. Much of the observed inter-specific variation in pheromone peptides could directly contribute to reproductive barriers.

The data presented in this dissertation serve to answer a number of important questions regarding the evolution of Ascomycete mate recognition genes and its implications. These genes have complex evolutionary histories and their evolution has a direct influence on the biology of Ascomycetes. A number of avenues for future research in this field are also highlighted.

CHAPTER ONE

GENETICS AND EVOLUTION OF MATE RECOGNITION MECHANISMS IN ASCOMYCETE FUNGI AND THEIR POTENTIAL ASSOCIATION WITH SPECIES BOUNDARIES

Isn't it a task of science to detect fundamental similarities concealed by apparent unlikeness? A fundamental common property of species is the presence of isolating mechanisms. The very fact that isolating mechanisms are as diversified as they are is strong evidence for the prevention of interbreeding being an essential characteristic of the process of speciation. . . . Any gene that raises an effective barrier to the mingling of incipient species . . . may become the basis of speciation.

Theodosius Dobzhansky
Speciation as a Stage in Evolutionary Divergence (1940)

It appears certainly possible that an evolution of sexual preference . . . would establish an effective isolation between two differentiated parts of a species, even when geographical and other factors were least favourable to such separation.

Ronald A. Fisher
The Genetical Theory of Natural Selection (1930)

Introduction

Like most fungi, members of the phylum Ascomycota (Class: Ascomycetes) are defined by characteristics of their sexual reproductive biology. Species in this Phylum produce ascospores, the haploid products of meiosis, within sac-like structures known as asci. Ascomycetes are also capable of asexual reproduction, through budding in the Saccharomycotina or yeasts, and through the growth of vegetative hyphae and production of asexual spores such (conidia, mitospores) in the Pezizomycotina or filamentous species. Interestingly, many species that group phylogenetically within the Ascomycota, have never been observed in a sexual state (Taylor *et al.* 1999). This fact has complicated the definition of species (Taylor *et al.* 2000) and limited use of the biological species concept (Mayr 1963) to certain sexual genera such as *Neurospora* (e.g. Turner *et al.* 2001) and *Fusarium* (e.g. Kuhlman 1982). Nevertheless, when biological species are recognized, there is a strong correlation between biological and phylogenetic species. This pattern is exemplified by the *Gibberella fujikuroi* species complex (reviewed by Kvas *et al.* 2009). Indeed it has long been recognized that a loss of sexual compatibility and, therefore, the potential for gene flow will inevitably lead to speciation (Dobzhansky 1951).

The source of reproductive isolation in fungal speciation is a prevailing problem because many species are thought to have ubiquitous dispersal over an unlimited range (Finlay 2002), ruling out allopatry. Several recent studies have shown that allopatric speciation has occurred in some fungi; yet many cases of apparent sympatric speciation are yet to be fully understood (reviewed by Giraud *et al.* 2008). In sympatry, reproductive isolation can either be achieved via assortive mating or niche specialization, provided that mating occurs only in specific niches (Giraud *et al.* 2008). Le Gac and Giraud, (2008) proposed that host specificity, coupled with the fact that most Ascomycetes mate upon or within host tissue, makes specialization a likely source of isolation among sympatric species. This hypothesis was supported by the observation that most Ascomycete crosses display post-zygotic isolation, demonstrating a lack of species-specificity during pre-mating recognition (Le Gac and Giraud 2008). However, not all Ascomycetes exhibit strong or any host-specificity and many closely related but distinct sexual species share the same host as well as overlapping ranges. Furthermore, the typical mating experiments applied to Ascomycetes are of the kind referred to as “no choice” experiments (Coyne and Orr 2004), involving only two isolates, and always in an artificial setting. These factors could

mask the importance of any form of preference exhibited during mate-recognition or "behavioural" barriers that are present in nature.

The importance of behavioural (interaction between mates) and gametic (interation between gametes) isolating mechanisms in animal speciation is well recognised, as are the evolutionary forces that can give rise to such barriers (Coyne and Orr 2004 chapter 6). In fungi, our understanding of the interactions between mates has grown rapidly with the advent of molecular techniques (reviewed by Shiu and Glass 2000; Casselton 2002). However, we are only just beginning to explore the potential role of these mechanisms in species-specific mate recognition (e.g. Karlsson *et al.* 2008). There is huge diversity in mating systems among Ascomycetes, and genes involved in mating behaviour and mate recognition have been shown to diversify rapidly between species due to positive selection (Wik *et al.* 2008; Karlsson *et al.* 2008). This highlights the possibility that these mechanisms may be involved in reproductive barriers between species.

The term "mate recognition" usually refers to events that occur pre-mating with the interaction between gametes (post-mating but pre-zygotic) considered separately (Coyne and Orr 2004). In fungi, particularly unicellular yeasts, it is difficult to draw a distinction between mates and gametes. There is even evidence that recognition mechanisms occur between the nuclei after anastimosis or plasmogamy, the cytoplasmic fusion between gametes. Therefore, for the purpose of this review, "mate recognition" refers to all interactions that occur prior to karyogamy (nuclear fusion). Because mate recognition refers to both the detection of potential mates and the discrimination between potential and non-potential mates, the terms mate recognition and mate discrimination are used interchangeably.

In order to explore the potential role of mate-recognition mechanisms in reproductive isolation among Ascomycetes, a general understanding of mating in these species is relevant. In yeasts, such as *Saccharomyces cerevisiae*, mating occurs when haploid cells develop projections (schmoos), which fuse during fertilization (Lee *et al.* 2010). In most filamentous species, such as *Neurospora crassa*, male and female elements can be distinguished. The male (donor) element, or antheridium, can be either a conidium or a hypha. The female (acceptor) element is a coiled lateral hypha known as an ascogonium. In some species the ascogonium carries a receptive hypha called the trichogyne, which actively grows toward the male element and initiates fertilization (Bistis 1983). In yeasts, karyogamy immediately follows anastamosis (Lee *et al.* 2010). By contrast, in

filamentous species, karyogamy is delayed. Instead, the two nuclei proliferate through synchronized mitotic divisions, giving rise to multinucleate cells within a developing fruiting body (Coppin *et al.* 1997; Pöggeler *et al.* 2006). Nuclei from each parent pair up in a process termed inter-nuclear recognition and a single pair is packaged into an ascus mother cell where karyogamy occurs. This is the only diploid stage in the life cycle of filamentous species, and it is immediately followed by meiosis (Pöggeler *et al.* 2006). In yeasts, the diploid zygote is stable and can reproduce by mitosis. However, haploid progeny (ascospores) must be generated via meiosis before mating can again occur (Lee *et al.* 2010).

There are two major components to consider in Ascomycete mate recognition. The first is the mating type or *MAT* locus. In addition to regulating the expression of genes involved in sexual development, the *MAT* locus is intimately associated with the "mating system" of a species; in essence, whether it is self-fertile (homothallic) or self-sterile (heterothallic). The second major component in mate recognition is a pheromone/receptor system that governs the interaction between the cells and nuclei of the two mates (Kurjan 1993; Kim and Borkovich 2004, 2006). These pheromone and receptor-encoding genes are among those regulated by the *MAT* locus (Kurjan 1993).

This review serves to summarize the current knowledge of the molecular genetic mechanisms that control mate discrimination in Ascomycetes. The mating type and pheromone/receptor systems and their evolution are discussed. The potential involvement of each system in inter-specific reproductive barriers is explored. The review concludes with a discussion of the potential evolutionary forces that may drive species-specificity in mate recognition, and how these may affect the genes involved. Where appropriate, examples from the Ascomycetes are assessed.

Mating type

Almost all heterothallic Ascomycetes display bipolar heterothallism (Whitehouse 1949) with two mating types determined by the identity of a single genomic locus. Mating can occur only between haploid individuals dissimilar at the *MAT* locus. The two alternative sequences at the *MAT* locus share no real homology and sometimes contain more than one gene. These factors led to the use of the term "idiomorphs" as opposed to alleles (Metzenberg and Glass 1990). The borders of the *MAT* idiomorphs are defined by the points at which homology is restored between the genomes. To maintain the dissimilarity

between the two idiomorphs, recombination is suppressed in this region (Coppin *et al.* 1997; Kronstad and Staben 1997).

Mating types in many Ascomycetes were defined before the underlying molecular basis was understood. For this reason, there is little standardization in their nomenclature. For example, the two mating types are termed *MAT*a and *MAT*α in *S. cerevisiae*; *MAT*-A and *MAT*-a in *N. crassa*; *mat*- and *mat*+ in *Podospora anserina*; and *MAT*1-1 and *MAT*1-2 in *Cochliobolus heterostrophus* (Herskowitz 1988; Glass *et al.* 1988; Picard *et al.* 1991; Turgeon *et al.* 1993). This terminology has been complicated by the fact that homologous idiomorphs are not well conserved, and can differ markedly in gene composition.

Genes encoded by the *MAT* idimorphs and functions

Although *MAT* loci are ubiquitous among fungi, the gene composition of the *MAT* locus varies among species. Within the Ascomycota the most significant difference exists between the yeasts and filamentous species. A number of reviews have scrutinized the functions of the various *MAT* genes in both yeasts (Herskowitz 1988, 1989; Johnson 1995) and filamentous Ascomycetes (Kronstad and Staben 1997; Coppin *et al.* 1997). Here a brief summary of the genes and their putative functions is provided, focusing mainly on their relevance in the process of mate recognition.

***MAT* genes in yeasts:** The most important model organism used to describe yeast *MAT* genes is *S. cerevisiae*. The *MAT*a idiomorph contains a gene called a1; the *MAT*α idiomorph contains two genes, α1 and α2. The a1 and α2 genes encode homeodomain (HD) transcription factors. The α1 gene encodes a unique transcription factor with a motif known as the α-box. These transcription factors determine whether cells are the '**a**' or 'α' type by regulating the expression of multiple genes including those encoding specific pheromones and receptors.

The **a**-cells carry the *MAT*a idiomorph, however expression of the a1 gene, is not necessary to specify the '**a**' cell type. In fact, genes specific to this cell type are expressed constitutively while genes specific to the 'α' cell type are constitutively repressed. In α cells (which carry the *MAT*α idiomorph) the α1 and α2 genes are necessary to facilitate repression of the **a**-specific genes and concurrent expression of α-specific genes. The purpose of the a1 gene in the *MAT*a idiomorph becomes apparent only after fertilization, when yeasts form a stable diploid cell type refered to as **a**/α. In these cells, the homeodomain protein products of the a1 and α2 genes form a heterodimer that functions

to repress expression of all genes associated with the haploid cell types. The heterodimer also makes the diploid cell competent to undergo meiosis.

***MAT* genes in filamentous species:** The *MAT* genes of filamentous species are also putative transcription factors (Coppin *et al.* 1997). Homologs of the yeast homeodomain proteins are absent in filamentous Ascomycetes. Instead, one idiomorph contains a homologue of α1, the *S. cerevisiae* α-domain protein, while the other always carries a High Mobility Group (HMG)-domain protein. For the purpose of this review, and in accordance with the convention proposed by Turgeon and Yoder (2000), the idiomorph containing the α-domain protein is referred to as *MAT1-1* and isolates possessing this idiomorph as *MAT-1*s. Likewise, the idiomorph containing the HMG-domain protein is referred to as *MAT1*-2, and isolates possessing this idiomorph as *MAT-2*s.

In many filamentous species, including *N. crassa*, *P. anserina* and *Fusarium verticilioides,* the *MAT1-1* idiomorph contains two additional genes; an α-helical protein and a different HMG-domain protein (Ferreira *et al.* 1996; Debuchy and Coppin 1992; Yun *et al.* 2000). The latter HMG-domain protein has, therefore, been referred to as the HMG-1 and the one found in the *MAT1*-2 as HMG-2 (Shiu and Glass 2000). There are, therefore, four *MAT* gene types most commonly found in filamentous species: α-domain, α-helical and HMG-1 in *MAT1-1*, and HMG-2 in *MAT1*-2.

The α-domain protein (in *MAT-1* cells) and the HMG-2-domain protein (in *MAT*-2 cells) are termed "nuclear identity" genes. They are necessary for attraction and interaction between mates (reviewed by Kronstad and Staben 1997; Coppin *et al.* 1997) as well as the interaction between nuclei during internuclear recognition (reviewed by Shiu and Glass 2000). The HMG-1-domain protein, encoded by the *MAT1-1* idiomorph in some species, may form a heterodimer with the α-domain protein to perform the nuclear identity function. The third gene in the *MAT1-1* idiomorph of some species, encoding the α-helical protein, appears only to be necessary for later stages in mating. It may perform its function equally well if present at an ectopic site in the genome or in both genomes (Arnaise *et al.* 1997).

Could *MAT* genes cause reproductive barriers?

The *MAT* genes are known to be necessary for successful recognition between mates. However, as transcription factors, there is no evidence that the *MAT* proteins themselves are directly involved in the interaction between two individuals for mate discrimination.

One possible exception is the formation of the a1/α2 heterodimer after fertilization in yeasts. It is possible that differences in the homeodomain proteins that affect formation of the heterodimer could result in post-zygotic isolation; although this would not affect pre-mating recognition.

A more likely means by which *MAT* genes could cause isolation is indirectly, by orchestrating differential reproductive behaviour. Species can display differences in the environmental conditions that stimulate mating. It is well known that many Ascomycetes require unique and specific conditions before the sexual cycle is initiated (Poggeler *et al.* 2006). Another factor is the times at which mating takes place. Reproduction that occurs at different times may prevent gene flow in nature. In *S. cerevisiae* and *Saccharomyces paradoxis*, two sexually-compatible yeast species with overlapping distributions, differences in the rate at which mating occurs between the two species is also thought to reduce the probability of hybridization (Murphy *et al.* 2006).

Gene expression data for pheromone precursor genes in *Neurospora* have demonstrated the plausibility of the above arguments. The expression of both *N. crassa* pheromone precursors are dependant on circadian rhythms (Bobrowicz *et al.* 2002). In fact one precursor gene was originally described as ccg-4 (clock-controlled gene 4) (Bell-Pedersen *et al.* 1996). Although "allochronic" speciation due to temporal isolation is widely accepted, most theories have focused on seasonal differences in mating as opposed to time-of-day differences (White 1978, chapter 7; Coyne and Orr 2004, chapter 5). Quantitative differences in pheromone regulation have also been observed between successful and unsuccessful hybridizations in *Neurospora*, and were proposed as a possible cause of inter-sterility between certain species (see below) (Karlsson *et al.* 2008).

MAT loci and mating systems

A major complexity in Ascomycete biology is the existence of both homo- and heterothallic species in many genera. This has inspired endeavours to identify the mechanisms involved in determining thallism, as well which state is ancestral. The problem has now been largely solved, and it has become clear that the *MAT* locus is central to this phenomenon. However, the exact mechanism by which self-fertility is achieved varies among taxa.

Various sources of self-fertility

Some yeast species, such as *S. cerevisiae* display a homothallic lifestyle due to "mating type switching". In these fungi, a single isolate is capable of switching mating type, thereby becoming sexually compatible with "clonal" sibling cells derived asexually. This process is made possible by the presence of two silent loci in the genome carrying the genetic determinants of both *MAT*a and *MAT*α. Recombination between one of these silent loci and the active mating type locus enables mating type switching to occur (Herskowitz 1989). Interestingly, there are also many Saccharomycetes, such as *Candida albicans*, that are unable to switch mating type. Phylogenetic studies suggest that this unique ability arose in a step-wise manner in ancestors of *S. cerevisiae* (reviewed by Lee *et al.* 2010). It is intriguing that, *S. cerevisiae* isolates incapable of mating type switching due to a deleterious mutation, also occur in nature. The coexistence of isolates with and without the capacity for mating type switching (i.e. with or without self-fertility) suggests that stabilizing selection maintains both forms. This further implies that either heterothallism or homothallism may be optimal in different environmental conditions (reviewed by Lee *et al.* 2010).

In filamentous species, homothallism is achieved by very different means. Instead of carrying additional silent *MAT* loci, most homothallic filamentous species carry functional sequences of both *MAT1-1* and *MAT1*-2 idiomorphs in a single genome (Cisar *et al.* 1994; Glass *et al.* 1990; Glass and Smith 1994; Beatty *et al.* 1994; Yun *et al.* 2000). Independent studies on the *MAT* genes of the homothallic *Sordaria macrospora* and *Chochliobulus lutterellii* have demonstrated that the genes are able to confer a self-fertile phenotype when introduced into a heterothallic species (Pöggeler *et al.* 1997; Yun *et al.* 1999). The reverse of these experiments has been performed on the homothallic *Fusarium graminearum,* where the *MAT1-1* and *MAT1-2* genes were deleted independently. Deletion mutants lost the ability to self-mate but had become heterothallic in that they could still mate with the opposite deletion mutants. Apart from the ability to self, the process of sexual development is identical in homothallic and heterothallic species. This lead Coppin *et al.* (1997) to propose that alternate expression of one or the other mating type by different nuclei may facilitate fertilization, inter-nuclear recognition and karyogamy in homothallic species.

In a few exceptional homothallic *Neurospora* spp., only the *mat*-A (*MAT1-1*) idiomorph sequence appears to be present while the *mat*-a (*MAT1*-2) is absent (Glass *et al.* 1988; Glass and Smith 1994). Furthermore, in *Neurospora pannonica* and *Neurospora terricola*, both of which carry the HMG-2 gene, this gene has accumulated deleterious mutations through a lack of selective constraint (Wik *et al.* 2008). These results are in conflict with those for *F. graminearum*, as they suggest that only the *mat*-A (*MAT1-1*) determinants are necessary to complete the homothallic sexual cycle. It is possible that the *mat*-a (*MAT1-2*) function is substitute by alterations to downstream target genes (Pöggeler 1999).

A third route to the self-fertile life cycle has been achieved through what is termed pseudohomothallism. Like their heterothallic counterparts, pseudohomothallic species encode genes of only a single mating type in each nucleus, and karyogamy can occur only between haploid nuclei of different mating types. This obstacle may be overcome by various means. For example in *Neurospora tetrasperma*, ascospores can be heterokaryotic, carrying two haploid nuclei of different mating type, and can, therefore, give rise to hyphae and conidia that are sexually compatible (Kronstad and Staben 1997).

Heterothallism to homothallism, or *vice versa*?

Before the molecular determinants of thallism were understood, Nauta and Hoekstra, (1992) designed a model to examine the evolution of sexual strategies in Ascomycetes. They proposed that a homothallic strategy is more likely to "invade" a heterothallic population than *vice versa*, suggesting that heterothallism was the ancestral state. The abundance of homothallism in many lineages of filamentous Ascomycetes thus suggests that this system has emerged repeatedly and, therefore, has a strong selective advantage. This advantage may simply lie in the ability to self-mate, since this feature has also evolved independently in the form of pseudohomothalism (Coppin *et al.* 1997). Self mating, although it cannot give rise new genetic variation, allows recombination during meiosis, which may be important to avoid an accumulation of deleterious mutations through "Muller's Ratchet" (Muller 1964; Bidochka and De Koning 2001).

Molecular studies have improved our understanding of how transitions from heterothallism to homothallism might have occurred in filamentous species. In many homothallic species such as *S. macrospora*, *C. lutterellii*, and *F. graminearum* the *MAT1-1* and *MAT1*-2 genes are physically linked (Pöggeler *et al.* 1997; Yun *et al.* 1999; Yun *et al.* 2000), suggesting that homothallism may have arisen through fusion of *MAT*

idiomorphs. In *Chochliobolus*, Yun *et al.* (1999) identified evidence that a *MAT1-1*/*MAT1*-2 fusion resulted from an unequal crossover event between the *MAT1-1* and *MAT1*-2 idiomorphs. Furthermore, it was evident that multiple independent fusions had occurred in the genus, making homothallism a polyphyletic trait. Phylogenetic evidence suggests that, in *Neurospora*, homothallism may be ancestral, having potentially arisen in an earlier ancestor of both *Neurospora* and *Sordaria* (Pöggeler 1999). Determination of the ancestral state in *Aspergillus* is more complicated (as reviewed by Lee *et al.* 2010). Some heterothallic species appear to carry remnants of a linked *MAT* locus, suggesting that heterothallism could have arisen through alternative gene loss. On the other hand, some homothallic *Aspergillus* spp. carry sequences of both mating types at unlinked genomic loci, suggesting that a chromosomal translocation in a heterothallic species could have given rise to homothallism. It is certainly clear that transitions from one form of thallism to another have been numerous in the evolution of the Ascomycetes.

Could a change in mating system cause reproductive isolation?

There has been some debate as to whether mating systems themselves could cause reproductive isolation between species. Again, most ideas have focused on indirect mechanisms. However, in some plants there is evidence that genetic differences between self compatible (SC) and self incompatible (SI) species can directly lower the potential for hybridization (Lewis and Crowe 1958). Specifically, incompatibility occurs when pollen from an SC male is used to fertilize an SI female. This implies that the same mechanism that causes SI females to reject pollen from 'self', causes a rejection of pollen from SC species. Whether heterothallic Ascomycetes demonstrate the same level of attraction to spores from both heterothallic and homothallic relatives has not been investigated. However, there is evidence that pheromones and receptors may be less important, or even dispensible, for reproduction in some self-fertile species (see section 3). In fact, the homothallic *F. graminearum* may have adopted a previously unknown mechanism whereby contact between mycelia stimulates mating (Kim *et al.* 2008); this would severely affect its out-crossing ability.

Giraud *et al.*, (2008) proposed that, even if sexual compatibility is retained, high levels of selfing in a species can reduce gene flow between two species. It follows that a transition from heterothallism to homothallism could facilitate speciation. Coyne and Orr, (2004) argue that conversions to self-fertility or asexuality are not true isolating mechanisms because gene flow between individuals of the same taxon is reduced just as much as between individuals of different taxa. The authors further suggest that a set of completely

selfing individuals cannot be considered as a species but rather as a collection of "*microspecies*" with each individual "*propagating its own genetically isolated lineage*" (Coyne and Orr 2004). Despite this argument, Coyne and Orr, (2004) provide examples of how a change in mating system can indirectly lead to speciation. Selfing reduces the affective population size, increasing the effects of genetic drift, which can indirectly lead to post-zygotic incompatibilities such as chromosomal differences. By reducing gene flow from external populations, selfing or asexual reproduction could allow a single lineage to adapt more effectively to a specific habitat. Reproductive isolation could result as a secondary consequence of such adaptation. Selfing species might experience different selective pressures to those in out-crossing species. For example, as mentioned above and discussed in the next section, pheromone signalling may be less important or dispensable in homothallic fungi.

Mating type genes and mating systems certainly have the potential to decrease gene-flow between taxa. However, none of these mechanisms involves a deliberate preference of conspecific mates over heterospecific mates. If such preferences exist in Ascomycetes, they are more likely to be determined by the pheromone/receptor system that governs the attraction and interaction between mates.

Mating pheromones and receptors

Ascomycete pheromones and their associated response systems are best understood in the yeast *S. cerevisiae* (Kurjan 1993; Duntze *et al.* 1991; Caldwell *et al.* 1995; Chen *et al.* 1997): **a** cells release a-factor pheromones that bind to receptors on the surface of α cells. Similarly, α cells release α-factor pheromones that bind to receptors on the surface of **a** cells. The a-factor pheromone is a 12-amino-acid lipopeptide. It is processed from a larger precursor protein carrying a C-terminal prenylation signal, termed the "CaaX" motif, which is typical of pheromone precursors found in all fungi (Brake *et al.* 1985). Two genes encode precursors for the a-factor pheromone, *Mfa-1* and *Mfa-2*. The α-factor pheromone is a 13-amino-acid peptide that is also cleaved from a larger precursor. This precursor carries multiple repeats of the mature pheromone, each bordered by cleavage sites. Again, two genes encode precursors for the α-factor, *Mfα-1 and Mfα*-2 (Singh *et al.* 1983). The receptors for a-factor and α-factor are termed Ste2p and Ste3p, respectively. These proteins are both members of the large GPCR family of G-protein coupled receptors. Like all the members of this family, Ste2p and Ste3p contain seven transmembrane spanning (TM) domains (Nakayama *et al.* 1985). Both receptors also

activate the same downstream G-protein linked signal transduction pathway. Pheromone and pheromone receptor genes are unlinked from the *MAT* locus and both classes are present in the genomes of all individuals. They are, however, expressed in a mating type-specific manner (Herskowitz 1989).

Pheromone precursor and receptor genes, homologous to those of *S. cerevisiae*, have been characterized in a number of homothallic, heterothallic and even suspected asexual filamentous fungi. Homologs of the *Mfa* and *Mfα* pheromone precursor genes are often referred to as *ppg1* and *ppg2*, respectively in filamentous species (Mayrhofer *et al.* 2006; Lee *et al.* 2008). Genes encoding receptors homologous to *Ste3* and *Ste2* are termed *Pre1* and *Pre2*, respectively (Pöggeler and Kück 2001). Interestingly, homologs of the **a**-factor precursor gene (*ppg2*) have been identified only in species from one class of filamentous Ascomycetes, the Sordariomycetes, including *N. crassa*, *Magnaporthe grisea*, *P. anserina* and *Cryphonectria parasitica* (Bobrowicz *et al.* 2002; Shen *et al.* 1999; Zhang *et al.* 1998; Coppin *et al.* 2005). In the three species of the class Eurotiomycetes, *Aspergillus nidulans*, *Aspergillus fumigatus* and *Penecillium crysogenum*, in which a *ppg1* homolog was identified, a *ppg2* homolog may be absent (Dyer *et al.* 2003; Pöggeler 2002; Hoff *et al.* 2008). However, it is unclear how these species would compensate for its absence.

Pheromones and receptors in heterothallic species

Production of mating pheromones is essential for male fertility in filamentous, heterothallic species (Kim *et al.* 2002; Turina *et al.* 2003; Coppin *et al.* 2005). Kim and Borkovich, (2006) showed that pheromones confer male fertility by directing the polarized growth of trichogynes of the opposite mating type. Furthermore, mating pheromone expression is unnecessary for female fertility (Coppin *et al.* 2005; Kim and Borkovich 2006), as trichogynes need only be receptive to the pheromones released by the conidia. This is consistent with the observation that the pheromones of *P. anserina* are transcribed only at high levels in microconidia (Coppin *et al.* 2005). Pheromone receptor mutants are female sterile as they are unable to sense the presence of pheromone-secreting conidia (Kim and Borkovich 2004).

Pheromones and receptors in homothallic species

In the homothallic species *S. macrospora*, the absence of interacting conidia and trichogynes, suggests that pheromone signaling is not required to sense mating partners and initialise fertilization. Interestingly, both pheromone precursors as well as both

pheromone receptor genes are transcriptionally expressed in this fungus (Pöggeler 2000; Pöggeler and Kück 2001). In contrast to heterothallic species, *S. macrospora* individuals secrete both pheromone classes (Mayrhofer *et al.* 2006). This is probably due to the fact that this homothallic fungus carries *MAT1-1* and *MAT1-2* genes. Using a yeast expression-based assay, Mayrhofer and Pöggeler, (2005) showed that the *S. macrospora* α-factor-like pheromone interacts with and triggers a response in the *Pre2* receptor. The authors further demonstrated that *Pre2* from *S. macrospora* is capable of activating the *S. cerevisiae* pheromone response pathway. Despite these findings, mutants of *S. macrospora* lacking both pheromones are able to reproduce sexually by self-fertilization, although they do display reduced perithecial and ascospore production (Mayrhofer *et al.* 2006). In contrast, mutants lacking both receptors are completely unable to produce perithecia or ascospores (Mayrhofer *et al.* 2006). Mutants lacking only one of the pheromone/receptor pairs are unaffected, possibly because the other pair compensates for the loss (Mayrhofer and Pöggeler 2005, Mayrhofer *et al.* 2006). These results demonstrate that a pheromone/receptor interaction is partially dispensable in this species, but required for optimal sexual reproduction. At least one receptor is essential and probably involved in post-fertilization sexual development.

The pheromone/receptor system of *F. graminearum*, another homothallic species, is quite different to that in *S. macrospora* (Kim *et al.* 2008; Lee *et al.* 2008). Only the α-factor-like pheromone and *Pre2* form a functional pair, while the a-factor-like pheromone and *Pre1* may be completely non-functional (Lee *et al.* 2008; Kim *et al.* 2008). Unlike *S. macrospora*, mutants lacking both pheromone and receptor genes are completely self-fertile and capable of wild-type-level perithecia production (Lee *et al.* 2008; Kim *et al.* 2008). Kim *et al.*, (2008) proposed that contact between haploid structures (for example, mycelia and conidia) may be sufficient to initiate the mating type-directed signal transduction cascade in *F. graminearum*. This phenomenon may be similar to the *S. cerevisiae STE5* gain-of-function-mutant, in which this pathway is constitutively activated in the absence of pheromones (Pryciak and Huntress 1998; Sette *et al.* 2000). The results also imply that *F. graminearum* has no post-fertilization requirement for pheromones or pheromone receptors (Lee *et al.* 2008; Kim *et al.* 2008). G*ibberella zeae* may rather represent a more advanced stage in the evolution of homothallic Ascomycetes (Lee *et al.* 2008; Kim *et al.* 2008).

Could pheromones and receptors cause reproductive barriers?

In insects, there are several examples of reproductive isolation due to distinct pheromone signals (e.g. Roelofs *et al.* 2002; Nosil *et al.* 2007). These are however chemical compounds quite different from the peptide pheromones of fungi. Other than *N. crassa* and *S. macrospora* that share identical pheremones, the peptide pheromones characterized from filamentous Ascomycetes have all had unique sequences. It is certainly possible that these differences may be associated with barriers to mate-recognition between these species. In Basidiomycetes, minor mutations in the pheromone peptides or receptor proteins can drastically affect the success of pheromone reception (Olesnicky *et al.* 2000; Fowler *et al.* 2001).

There has been insufficient research to determine whether inter-specific differences in Ascomycete pheromones affect their functionality. In yeasts, there seems to be a negative correlation between pheromone similarity and hybridization ability (Burke *et al.* 1980) However, this would probably be true of any protein since more distant relatives are usually less reproductively compatible. Despite having distinct amino acid sequences, the a-factors of *S. cerevisiae*, *Saccharomyces kluyveri* and *Saccharomyces exiguus* have some inter-specific action (McCullough and Herskowitz 1979; Hisatomi *et al.* 1988). Potentially, only specific alterations may affect pheromone reception. In the *S. cerevisiae* α-factor peptide, the N-terminal residues have been shown to control binding to the receptor, while certain C-terminal residues, although unnecessary for receptor binding, are required to stimulate the receptor, initiating the mating response (Naider and Becker 2004). It is therefore possible that only certain inter-specific differences may be of importance while others may be inconsequential.

A clue to the significance of pheromone variation might be provided by a comparison of inter-specific variation with intra-specific similarity. Although this has not been performed, a notable point is the fact that the multiple tandem repeats of the α-factor-like pheromone are usually completely (or nearly completely) identical in amino acid sequence. If purifying selection acts to maintain this similarity, this would imply that pheromone function is sequence-sensitive and that inter-species differences could be significant. Alternatively, tandem repeats are often subject to homogenization through concerted evolution, so the conservation of multiple copies of the mature peptide is not clear evidence for functional constraint.

As mentioned previously, differential regulation of pheromones (spatially, temporally or quantitatively) could be another route to reduced mate-recognition between species. Karlsson *et al.* (2008) proposed this as a possible source of the observed inter-sterility between *N. crassa* and *N. intermedia* (Dettman *et al.* 2003). The authors demonstrated that pheromones were expressed in different quantities between compatible and incompatible mating events. However it is possible that the differential expression was merely a result of the failed mating rather than the cause. It is also unclear how, exactly, differential quantitative expression of pheromones could reduce mating success.

Evolution of mate-recognition and species boundaries

Ascomycete mate recognition genes evolve rapidly

Both the *MAT* idiomorphs and pheromone receptor genes have been found to be more divergent than other parts of the genome among certain species (Turgeon 1998, Wik *et al.* 2008; Pöggeler and Kück 2001). There is also evidence that positive diversifying selection might be responsible for this phenomenon, at least in some cases (Wik *et al.* 2008; Karlsson *et al.* 2008). These findings concur with the general trend of rapid evolution in sex-related genes, as reported in many species across the plant and animal kingdoms (Swanson and Vacquier 2002; Clark *et al.* 2006). A number of evolutionary forces have been proposed to explain this trend. A discussion of a few of these concepts and how they might apply to Ascomycete mate recognition follows. While only one of these forces, reinforcement, results directly from selection pressure to enhance reproductive isolation, the others could have the same effect as a byproduct of evolutionary divergence (Templeton 1989; Palumbi 2008). They are therefore all potentially important to species boundaries.

Evolutionary forces and mate recognition mechanisms

Reinforcement: Probably the most commonly invoked force that could drive the formation of pre-mating barriers is reinforcement. Dobzhansky (1951) proposed that if two species diverged in allopatry and returned to a sympatric distribution where hybrids between the two displayed reduced fitness; natural selection would favour alleles that enhance prezygotic isolation between the two species. Hence, reinforcement is thought to cause a pattern known as "reproductive character displacement", whereby sympatric individuals of two species are more strongly reproductively-isolated than allopatric individuals (Howard 1993). It is thought that such isolation would occur pre-mating, to avoid the physical cost of sexual reproduction (Coyne and Orr 2004). Reinforcement has

thus been proposed as a driving force behind the diversification of reproductive proteins (Howard 1993) and signalling molecules such as pheromones (Symonds and Elgar 2008).

Le Gac and Giraud, (2008) compiled data from numerous studies in which fungi had been crossed and, concluded that reproductive character displacement and premating isolation is rare in Ascomycetes and more common in Basidiomycetes. These authors proposed that this distinction could reflect the distinct life styles of the two groups: Basidiomycetes mate in the soil where there is a high probability of encounters with hetero-specifics. Ascomycetes generally mate on their specific host plants, so for species that display strong specificity for distinct hosts, the probability of such an encounter is reduced, even in sympatry. Hence, Ascomycetes may not experience the same selective pressures to develop species-specificity in mate recognition.

While the above model applies only to Ascomycetes that show strong host-specificity, the reality is that there is only a single reported case of reproductive character displacement in this Phylum. Dettman *et al.* (2003) demonstrated lower fitness in sympatric than allopatric crosses between *N. crassa* and *Neurospora intermedia*. Interestingly, this reduced compatibility was post-mating and quantitative, based on the number of ascospores produced. If Ascomycetes do encounter hetero-specific individuals in sympatry, the lack of examples of pre-mating isolation must be considered surprising. Turner *et al.* (2010) argued that reinforcement may favour post-mating barriers if selective constraint acting upon pheromone sequences is too great to allow divergence. Another interesting observation by Turner *et al.* (2010) is that in the case of *N. crassa* and *N. intermedia*, it appears that the female parent (spore receiver) is more likely to abort the production of ascospores. This is consistant with Coyne and Orr's (2004) prediction that reinforcement should be greater in the parent likely to incur the greater cost of reproduction. As discussed below, the fitness costs associated with playing the "female" role may be an important factor in the evolution of reproductive genes in filamentous Ascomycetes.

Sexual Conflict: It is unclear how many Ascomycete species have true male and female "sexes". By convention, "females" have a greater energy investment in the offspring and often respond to mating signals produced by males. This is probably the case for many filamentous fungi such as *N. crassa* in which males simply donate a haploid nucleus. Females on the other hand, must often locate and grow towards this nucleus, and, after plasmogamy, produce a complex ascogonial structure which protects the developing

ascospores (Pöggeler *et al.* 2006). Yeast species, in contrast, appear to have more equal investment as mating projections grow from both parents and fuse. Most Ascomycetes are also hermaphroditic, but there may be fitness benefits associated with performing the male function as opposed to the female function, a phenomenon called sexual conflict (Parker 1979). Consequently, individuals may mate preferentially as males, and be reluctant to mate as females. Interestingly, reduced female fertility and female-sterility has been reported in many Ascomycetes (Leslie and Klein 1996). Since these fungi can also reproduce asexually, female fertility could be entirely lost by a population, although any advantage associated with sexual reproduction may cause antagonistic selection to retain female fertility.

There is a direct connection between sexual conflict and mate recognition in filamentous Ascomycetes; pheromone receptors are required for female function while only pheromones are required for male function (Turina *et al.* 2003; Coppin *et al.* 2005; Kim and Borkovich 2006). Sexual conflict could result in a co-evolutionary chase, with receptors continually evolving to avoid reception by pheromones (to avoid performing the female role). Models have shown that such a phenomenon could cause reproductive signals to diversify in arbitrary directions in different species and could even cause reproductive isolation between two populations of the same species (Gavrilets 2000; Gavrilets *et al.* 2001). Sexual conflict has thus been proposed as a possible force driving the rapid evolution of receptors responsible for sperm-binding on the eggs of Abalone (*Haliotis rufescens*), which are thought to be associated with speciation in this genus (Galindo *et al.* 2003).

Sexual selection: Selection acting directly upon variation for reproductive success is known as sexual selection. This process is a common explanation for striking mating behaviour and display, but it may also act upon the interacting gametes (Eberhard 1996). In fungi, some individuals may be more successful in attraction or fertilization; either due to competition among gametes (analogous to sperm competition in animals) or due to a bias in female preference for fertilizing gametes (cryptic female choice). In animals, when there is polyandry (multiple-male mating), sperm competition can cause an "arms-race", acting on the reproductive system; particularly when it is coupled with sexual conflict (Clark *et al.* 2006). In line with this concept, Wik *et al.* (2008) postulated that competition among spores to attract trychogynes could drive the rapid evolution of mating type genes in heterothallic *Neurospora* species, perhaps to optimize spore attractiveness.

In Fisher's (1930) model, covariance between a characteristic male trait and female preference for that trait could result in "runaway sexual selection", whereby the male trait continually evolves to be ever more exaggerated. If females of different populations display differential preferences in the nature of the mating signal, the signal could diverge rapidly. In fact, several models predict that divergent sexual selection can occur within a population through assortive mating. This could consequently give rise to reproductive isolation (Lande 1981, West-Eberhard 1983)

Adaptive evolution: Paterson (1985) proposed that fertilization systems would diverge between species as they are optimized by natural selection to suit a specific habitat or lifestyle. A fungal example might be a transition from heterothallism to homothallism, to suit specific environmetal conditions. Templeton (1989) elaborated on the idea by suggesting that selection on reproductive systems could act not only to optimize reproduction *per se*, but also due to other environmental factors, for example to avoid predation. As pathogens, Ascomycete fungi may experience strong and unique selective forces imposed by the environment upon all aspects of their biology, including reproduction. One obvious example is pheromone signaling.

All signals that are intended for conspecific individuals are vulnerable to exploitation by unintended recipients (Haynes and Yeargan 1999). Reproductive signals could experience selective pressure to avoid detection by predators and parasites (Symonds and Elgar 2008) or microbial pathogens (Vacquier *et al.* 1997). Here, a related concept is proposed: that pheromones in Ascomycetes could experience selective pressure to avoid stimulating the specific defence systems of their plant or animal hosts. Ascomycete pathogens are unique in that spores that inoculate host tissue also form the male gametes and, therefore, produce pheromones abundantly. Hence, pheromone precursors are among the most highly expressed genes (Nelson *et al.* 1997; Kim *et al.* 2002). It is, therefore, plausible that pheromones could function as effectors of plant defence. This, in turn, could result in positive selection for changes in the nature of the pheromone signal, enabling pathogens to avoid detection. If this divergence leads to reproductive isolation as a side effect, this would be a rare example of adaptive evolution that indirectly causes a pre-mating barrier.

Genetic Drift: The ability of fungi to reproduce both sexually and asexually may make selective pressure on reproductive genes quite different from those in animals. The sexual

cycle may become dispensable when there is a limited advantage to be gained from the generation of genetic variation, for example in uniform environments such as crops and plantations. Relaxed purifying selection acting on reproductive genes could then cause them to be vulnerable to divergence due to drift.

Another situation under which sexual reproduction may be lost is during invasion, when founder populations of only a single mating type may arise, rendering sexual reproduction impossible (Paoletti *et al.* 2006). Under such conditions the "use it or lose it" principle may apply and reproductive genes could deteriorate. Homothallic species would presumably not suffer this fate as selfing maintains use of the sexual cycle. This ability to maintain an intact reproductive system might be one of the selective benefits of homothallism. Interestingly, Wik *et al.* (2008) reported evidence that contradicts this view. In some homothallic *Neurospora* spp. *MAT* genes contain deleterious mutations and premature stop-codons. The presence of genes from both mating types might in fact cause redundancy, and potentially even weaker constraint on reproductive genes.

Summary and Conclusions

The source of reproductive isolation in Ascomycete speciation, as well as the relative importance of pre- and post-mating isolating mechanisms, is not well understood. While mate and gamete recognition mechanisms in plants and animals often display species-specificity, we are only beginning to explore whether the relatively simple mate recognition mechanisms in Ascomycetes have a similar ability. The two major components of mate recognition in Ascomycetes are the *MAT* loci, which control mating systems and expression of reproductive genes, and the mating pheromone/receptor system that facilitates the attraction and interaction between mates. *MAT* loci might not control mate discrimination directly, but could cause reproductive isolation by directing distinct reproductive behaviour. Furthermore, transitions from heterothallism to homothallism, which involve the *MAT* loci, may lead to reduced gene-flow between groups. A more deliberate source of mate discrimination may be directed by the mating pheromone/receptor system. Minor differences in pheromone peptides or pheromone regulation could facilitate specificity in mate recognition.

In Ascomycetes, it is debatable whether reinforcement is a common force driving the formation of pre-mating barriers. Such barriers could also arise as a secondary consequence of divergence via other means, such as sexual selection, adaptive evolution

or genetic drift. Whatever its source, the rapid diversification of genes involved in mate recognition could be directly related to the great diversity of Ascomycetes.

References

Arnaise, S., Debuchy, R. and Picard, M. 1997. What is a bona fide mating-type gene? Internuclear complementation of mat mutants in Podospora anserina. *Molecular and General Genetics* **256**: 169-78.

Beatty, N. P., Smith, M. L. and Glass, N. L. 1994. Molecular characterization of mating-type loci in selected homothallic species of *Neurospora*, *Gelasinospora* and *Anixiella*. *Mycological Research* **98**: 1309-1316.

Bell-Pedersen, D., Shinohara, M. L., Loros, J. J. and Dunlap, J. C. 1996. Circadian clock-controlled genes isolated from *Neurospora crassa* are late night- to early morning-specific. *Proceedings of the National Academy of Sciences of the United States of America* **93**: 13096-13101.

Bidochka, M. J. and De Koning, J. 2001. Are teleomorphs really necessary: modelling the potential effects of Muller's Ratchet on deuteromycetous entomopathogenic fungi. *Mycological Research* **105**: 1014-1019.

Bistis, G. N. 1983. Evidence for diffusible, mating-type-specific trichogyne attractants in *Neurospora crassa*. *Experimental Mycology* **7**: 292-295.

Bobrowicz, P., Pawlak, R., Correa, A., Bell-Pedersen, D. and Ebbole, D. J. 2002. The *Neurospora crassa* pheromone precursor genes are regulated by the mating type locus and the circadian clock. *Molecular Microbiology* **45**: 795-804.

Brake, A. J., Brenner, C., Najarian, R., Laybourn, P. and Merrywheather, J. 1985 Structure of genes encoding precursors of the yeast peptide mating pheromone a-factor. In: *Protein Transport and Secretion*, (Gething, M. J., ed.). pp. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Burke, D., Mendonça-Previato, L. and Ballou, C. E. 1980. Cell-cell recognition in yeast: purification of *Hansenula wingei* 21-cell sexual agglutination factor and comparison of the factors from three genera. *Proceedings of the National Academy of Sciences of the United States of America* **77**: 318-322.

Caldwell, G. A., Naider, F. and Becker, J. M. 1995. Fungal lipopeptide mating pheromones: a model system for the study of protein prenylation. *Microbiological Reviews* **59**: 406-422.

Casselton, L. A. 2002. Mate recognition in fungi. *Heredity* **88**: 142-147.

Chen, P., Sapperstein, S. K., Choi, J. D. and Michaelis, S. 1997. Biogenesis of the *Saccharomyces cerevisiae* mating pheromone a-factor. *The Journal of Cell Biology* **136**: 251-269.

Cisar, C. R., TeBeest, D. O. and Spiegel, F. W. 1994. Sequence similarity of mating type idiomorphs: a method which detects similarity among the Sordariaceae fails to detect similar sequences in other filamentous ascomycetes. *Mycologia* **86**: 540-546.

Clark, N. L., Aagaard, J. E. and Swanson, W. J. 2006. Evolution of reproductive proteins from animals and plants. *Reproduction* **131**: 11-22.

Coppin, E., de Renty, C. and Debuchy, R. 2005. The function of the coding sequences for the putative pheromone precursors in *Podospora anserina* is restricted to fertilization. *Eukaryotic Cell* **4**: 407-420.

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

Coyne, J. A. and Orr, H. A. 2004. *Speciation*. Sinauer, Sunderland, Massachusetts.

Debuchy, R. and Coppin, E. 1992. The mating types of *Podospora anserina*: functional analysis and sequence of the fertilization domains. *Molecular and General Genetics* **233**: 113-121.

Dettman, J. R., Jacobson, D. J., Turner, E., Pringle, A. and Taylor, J. W. 2003. Reproductive isolation and phylogenetic divergence in *Neurospora*: comparing methods of species recognition in a model eukaryote. *Evolution* **57**: 2721-2741.

Dobzhansky, T. 1951. *Genetics and the origin of species*. Columbia University Press, New York, New York.

Duntze, W., Betz, R. and Nientiedt, M. 1991 Pheromones in yeasts. In: *The Mycota I: Growth, Differentiation and Sexuality*, (Wessels, J. M. F., ed.). pp. 381-399. Springer Verlag, Heidelberg.

Dyer, P. S., Paoletti, M. and Archer, D. B. 2003. Genomics reveals sexual secrets of *Aspergillus*. *Microbiology* **149**: 2301-2303.

Eberhard, W. G. 1996. *Female control: sexual selection by cryptic female choice*. Princeton University Press, New Jersey.

Ferreira, A. V. B., Saupe, S. and Glass, N. L. 1996. Transcriptional analysis of the *mtA* idiomorph of *Neurospora crassa* identifies two genes in addition to *mt A-1*. *Molecular and General Genetics* **250**: 767 - 774.

Finlay, B. J. 2002. Global dispersal of free-living microbial eukaryote species. *Science* **296**: 1061-1063.

Fisher, R. A. 1930. *The genetical theory of natural selection*. Oxford University Press, Oxford, UK.

Fowler, T. J., Mitton, M. F., Vaillancourt, L. J. and Raper, C. A. 2001. Changes in mate recognition through alterations of pheromones and receptors in the multisexual mushroom fungus *Schizophyllum commune*. *Genetics* **158**: 1491-1503.

Galindo, B. E., Vacquier, V. D. and Swanson, W. J. 2003. Positive selection in the egg receptor for abalone sperm lysin. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 4639-4643.

Gavrilets, S. 2000. Rapid evolution of reproductive barriers driven by sexual conflict. *Nature* **403**: 886-889.

Gavrilets, S., Arnqvist, G. and Friberg, U. 2001. The evolution of female mate choice by sexual conflict. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **268**: 531-539.

Giraud, T., Refrégier, G., Le Gac, M., de Vienne, D. M. and Hood, M. E. 2008. Speciation in fungi. *Fungal Genetics and Biology* **45**: 791-802.

Glass, N. L., Metzenberg, R. L. and Raju, N. B. 1990. Homothallic Sordariaceae from nature - the absence of strains containing only the a mating type sequence. *Experimental Mycology* **14**: 274 - 289.

Glass, N. L. and Smith, M. L. 1994. Structure and function of a mating-type gene from the homothallic species *Neurospora africana*. *Molecular and General Genetics* **244**: 401-409.

Glass, N. L., Vollmer, S. J., Staben, C., Grotelueschen, J., Metzenberg, R. L. and Yanofsky, C. 1988. DNAs of the two mating-type alleles of *Neurospora crassa* are highly dissimilar. *Science* **241**: 570-573.

Haynes, K. F. and Yeargan, K. V. 1999. Exploitation of intraspecific communication systems: illicit signalers and receivers. *Annals of the Entomological Society of America* **92**: 960-970.

Herskowitz 1988. Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Bicrobiology and Molecular Biology Reviews* **52**: 536-553.

Herskowitz, I. 1989. A regulatory hierarchy for cell specialization in yeast. *Nature* **342**: 749-757.

Hisatomi, T., Yanagishima, N., Sakurai, A. and Kobayashi, H. 1988. Interspecific actions of α mating pheromones on the a mating-type cells of three *Saccharomyces* yeasts. *Current Genetics* **13**: 25-27.

Hoff, B., Pöggeler, S. and Kuck, U. 2008. Eighty years after its discovery, Fleming's *Penicillium* strain discloses the secret of its sex. *Eukaryotic Cell* **7**: 465-470.

Howard, D. J. 1993 Reinforcement: origin, dynamics and fate of an evolutionary hypothesis. In: *Hybrid Zones and the Evolutionary Process*, (Harrison, R. G., ed.). pp. Oxford University Press, England.

Johnson, A. D. 1995. Molecular mechanisms of cell-type determination in budding yeast. *Current Opinion in Genetics and Development* **5**: 552-558.

Karlsson, M., Nygren, K. and Johannesson, H. 2008. The evolution of the pheromonal signal system and its potential role for reproductive isolation in heterothallic *Neurospora*. *Molecular Biology and Evolution* **25**: 168-178.

Kim, H.-K., Lee, T. and Yun, S.-H. 2008. A putative pheromone signaling pathway is dispensable for self-fertility in the homothallic ascomycete *Gibberella zeae*. *Fungal Genetics and Biology* **45**: 1188-1196.

Kim, H. and Borkovich, K. A. 2004. A pheromone receptor gene, *pre-1*, is essential for mating type-specific directional growth and fusion of trichogynes and female fertility in *Neurospora crassa*. *Molecular Microbiology* **52**: 1781-1798.

Kim, H. and Borkovich, K. A. 2006. Pheromones are essential for male fertility and sufficient to direct chemotropic polarized growth of trichogynes during mating in *Neurospora crassa*. *Eukaryotic Cell* **5**: 544-554.

Kim, H., Metzenberg, R. L. and Nelson, M. A. 2002. Multiple functions of *mfa-1*, a putative pheromone precursor gene of *Neurospora crassa*. *Eukaryotic Cell* **1**: 987-999.

Kronstad, J. W. and Staben, C. 1997. Mating type in filamentous fungi. *Annual Review of Genetics* **31**: 245-276.

Kuhlman, E. G. 1982. Varieties of *Gibberella fujikuroi* with Anamorphs in *Fusarium* Section *Mycologia* **74**: 759-768.

Kurjan, J. 1993. The pheromone response pathway in *Saccharomyces cerevisiae*. *Annual Review of Genetics* **27**: 147-179.

Kvas, M., Marasas, W. F. O., Wingfield, B. D., Wingfield, M. J. and Steenkamp, E. T. 2009. Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. *Fungal Diversity* **34**: 1-21.

Lande, R. 1981. Models of speciation by sexual selection on polygenic traits. *Proceedings of the National Academy of Sciences of the United States of America* **78**: 3721-3725.

Le Gac, M. and Giraud, T. 2008. Existence of a pattern of reproductive character displacement in *Homobasidiomycota* but not in *Ascomycota*. *Journal of Evolutionary Biology* **21**: 761-772.

Lee, J., Lee, T., Lee, Y. W., Yun, S. H. and Turgeon, B. G. 2003. Shifting fungal reproductive mode by manipulation of mating type genes: obligatory heterothallism of *Gibberella zeae*. *Molecular Microbiology* **50**: 145-152.

Lee, J., Leslie, J. F. and Bowden, R. L. 2008. Expression and function of sex pheromones and receptors in the homothallic ascomycete *Gibberella zeae*. *Eukaryotic Cell* **7**: 1211-1221.

Lee, S. C., Ni, M., Li, W., Shertz, C. and Heitman, J. 2010. The Evolution of Sex: a Perspective from the Fungal Kingdom. *Microbiology and Molecular Biology Reviews* **74**: 298-340.

Leslie, J. F. and Klein, K. K. 1996. Female fertility and mating type effects on effective population size and evolution in filamentous fungi. *Genetics* **144**: 557-567.

Lewis, D. and Crowe, L. K. 1958. Unilateral interspecific incompatibility in flowering plants. *Heredity* **12**: 233-256.

Mayr, E. 1963. *Animal species and evolution*. Harvard University Press, Cambridge, Massachusetts.

Mayrhofer, S. and Pöggeler, S. 2005. Functional characterization of an α-factor-like *Sordaria macrospora* peptide pheromone and analysis of its interaction with its cognate receptor in *Saccharomyces cerevisiae*. *Eukaryotic Cell* **4**: 661-672.

Mayrhofer, S., Weber, J. M. and Pöggeler, S. 2006. Pheromones and pheromone receptors are required for proper sexual development in the homothallic ascomycete *Sordaria macrospora*. *Genetics* **172**: 1521-1533.

McCullough, J. and Herskowitz, I. 1979. Mating pheromones of *Saccharomyces kluyveri*: pheromone interactions between *Saccharomyces kluyveri* and *Saccharomyces cerevisiae*. *Journal of Bacteriology* **138**: 146-154.

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

Muller, H. J. 1964. The relation of recombination to mutational advance. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **1**: 2-9.

Murphy, H. A., Kuehne, H. A., Francis, C. A. and Sniegowski, P. D. 2006. Mate choice assays and mating propensity differences in natural yeast populations. *Biology Letters* **2**: 553-556.

Naider, F. and Becker, J. M. 2004. The α-factor mating pheromone of *Saccharomyces cerevisiae*: a model for studying the interaction of peptide hormones and G protein-coupled receptors. *Peptides* **25**: 1441-1463.

Nakayama, N., Miyajima, A. and Arai, K. 1985. Nucleotide sequences of STE2 and STE3, cell type-specific sterile genes from *Saccharomyces cerevisiae*. *EMBO Journal* **4**: 2643.

Nauta, M. J. and Hoekstra, R. F. 1992. Evolution of reproductive systems in filamentous ascomycetes. I. Evolution of mating types. *Heredity* **68**: 405-410.

Nelson, M. A., Kang, S., Braun, E. L., Crawford, M. E., Dolan, P. L., Leonard, P. M., Mitchell, J., Armijo, A. M., Bean, L., Blueyes, E., Cushing, T., Errett, A., Fleharty, M., Gorman, M., Judson, K., Miller, R., Ortega, J., Pavlova, I., Perea, J., Todisco, S., Trujillo, R., Valentine, J., Wells, A., Werner-Washburne, M., Yazzie, S. and Natvig, D. O. 1997. Expressed sequences from conidial, mycelial, and sexual stages of *Neurospora crassa*. *Fungal Genetics and Biology* **21**: 348-363.

Nosil, P., Crespi, B., Gries, R. and Gries, G. 2007. Natural selection and divergence in mate preference during speciation. *Genetica* **129**: 309-327.

Olesnicky, N. S., Brown, A. J., Honda, Y., Dyos, S. L., Dowell, S. J. and Casselton, L. A. 2000. Self-compatible B mutants in *coprinus* with altered pheromone-receptor specificities. *Genetics* **156**: 1025-1033.

Palumbi, S. R. 2008. Speciation and the evolution of gamete recognition genes: pattern and process. *Heredity* **102**: 66-76.

Paoletti, M., Buck, K. W. and Brasier, C. M. 2006. Selective acquisition of novel mating type and vegetative incompatibility genes via interspecies gene transfer in the globally invading eukaryote *Ophiostoma novo-ulmi*. *Molecular Ecology* **15**: 249-262.

Parker, G. A. 1979 Sexual selection and sexual conflict. In: *Sexual selection and reproductive competition in insects*, (Blum, M. S. and Blum, N. A., eds.). Academic, New York, New York.

Paterson, H. E. H. 1985 The recognition concept of species. In: *Species and Speciation*, (Vrba, E., ed.). pp. 21-29. Transvaal Museum, Pretoria, South Africa.

Picard, M., Debuchy, R. and Coppin, E. 1991. Cloning the mating types of the heterothallic fungus *podospora anserina*: developmental features of haploid transformants carrying both mating types. *Genetics* **128**: 539-547.

Pöggeler, S. 1999. Phylogenetic relationships between mating-type sequences from homothallic and heterothallic ascomycetes. *Current Genetics* **36**: 222-231.

Pöggeler, S. 2000. Two pheromone precursor genes are transcriptionally expressed in the homothallic ascomycete *Sordaria macrospora*. *Current Genetics* **37**: 403-411.

Pöggeler, S. 2002. Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. *Current Genetics* **42**: 153-160.

Pöggeler, S. and Kück, U. 2001. Identification of transcriptionally expressed pheromone receptor genes in filamentous ascomycetes. *Gene* **280**: 9-17.

Pöggeler, S., Nowrousian, M. and Kück, U. 2006 Fruiting-body development in Ascomycetes. In: *Growth, Differentiation and Sexuality*, (Kües, U. and Fischer, R., eds.). pp. 325-355. Springer, Berlin, Heidelberg.

Pöggeler, S., Risch, S., Kuck, U. and Osiewacz, H. D. 1997. Mating-type genes from the homothallic fungus *Sordaria macrospora* are functionally expressed in a heterothallic ascomycete. *Genetics* **147**: 567-580.

Pryciak, P. M. and Huntress, F. A. 1998. Membrane recruitment of the kinase cascade scaffold protein Ste5 by the Gβγ complex underlies activation of the yeast pheromone response pathway. *Genes & Development* **12**: 2684-2697.

Roelofs, W. L., Liu, W., Hao, G., Jiao, H., Rooney, A. P. and Linn, C. E. 2002. Evolution of moth sex pheromones via ancestral genes. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 13621-13626.

Sette, C., Inouye, C. J., Stroschein, S. L., Iaquinta, P. J. and Thorner, J. 2000. Mutational analysis suggests that activation of the yeast pheromone response mitogen-activated protein kinase pathway involves conformational changes in the Ste5 scaffold protein. *Molecular Biology of the Cell* **11**: 4033-4049.

Shen, W.-C., Bobrowicz, P. and Ebbole, D. J. 1999. Isolation of pheromone precursor genes of *Magnaporthe grisea*. *Fungal Genetics and Biology* **27**: 253-263.

Shiu, P. K. T. and Glass, N. L. 2000. Cell and nuclear recognition mechanisms mediated by mating type in filamentous ascomycetes. *Current Opinion in Microbiology* **3**: 183-188.

Singh, A., Chen, E. Y., Lugovoy, J. M., Chang, C. N., Hitzeman, R. A. and Seeburg, P. H. 1983. *Saccharomyces cerevisiae* contains two discrete genes coding for the α-factor pheromone. *Nucleic Acids Research* **11**: 4049-4063.

Swanson, W. J. and Vacquier, V. D. 2002. The rapid evolution of reproductive proteins. *Nature Reviews Genetics* **3**: 137-144.

Symonds, M. R. E. and Elgar, M. A. 2008. The evolution of pheromone diversity. *Trends in Ecology & Evolution* **23**: 220-228.

Taylor, J. W., Jacobson, D. J. and Fisher, M. C. 1999. The evolution of asexual fungi: reproduction, speciation and classification. *Annual Review of Phytopathology* **37**: 197-246.

Taylor, J. W., Jacobson, D. J., Kroken, S., Kasuga, T., Geiser, D. M., Hibbett, D. S. and Fisher, M. C. 2000. Phylogenetic species recognition and species concepts in fungi. *Fungal Genetics and Biology* **31**: 21-32.

Templeton, A. R. 1989 The meaning of species and speciation: a genetic perspective. In: *Speciation and its consequences*, (Otte, D. and Endler, J. A., eds.). pp. 3-27. Sinauer Associates, Sunderland, Massachusetts.

Turgeon, B. G. 1998. Application of mating type gene technology to problems in fungal biology. *Annual Review of Phytopathology* **36**: 115 - 137.

Turgeon, B. G., Bohlmann, H., Ciuffetti, L. M., Christiansen, S. K., Yang, G., Schafer, W. and Yoder, O. C. 1993. Cloning and analysis of the mating-type genes from *Cochliobolus heterostrophus*. *Molecular and General Genetics* **238**: 270-284.

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

Turina, M., Prodi, A. and Alfen, N. K. V. 2003. Role of the *Mf1-1* pheromone precursor gene of the filamentous ascomycete *Cryphonectria parasitica*. *Fungal Genetics and Biology* **40**: 242-251.

Turner, B. C., Perkins, D. D. and Fairfield, A. 2001. *Neurospora* from Natural Populations: A Global Study. *Fungal Genetics and Biology* **32**: 67-92.

Turner, E., Jacobson, D. J. and Taylor, J. W. 2010. Reinforced postmating reproductive isolation barriers in *Neurospora*, an Ascomycete microfungus. *Journal of Evolutionary Biology* **23**: 1642-1656.

Vacquier, V., Swanson, W. and Lee, Y.-H. 1997. Positive darwinian selection on two homologous fertilization proteins: what is the selective pressure driving their divergence? *Journal of Molecular Evolution* **44**: S15-S22.

West-Eberhard, M. J. 1983. Sexual selection, social competition and speciation. *The Quarterly Review of Biology* **58**: 155-183.

White, M. J. D. 1978. *Modes of speciation*. W. H. Freeman and Co., San Fransisco, California.

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

Wik, L., Karlsson, M. and Johannesson, H. 2008. The evolutionary trajectory of the mating-type (*mat*) genes in *Neurospora* relates to reproductive behavior of taxa. *BMC Evolutionary Biology* **8**: 109.

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

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

Zhang, L., Baasiri, R. A. and Van Alfen, N. K. 1998. Viral repression of fungal pheromone precursor gene expression. *Molecular and Cellular Biology* **18**: 953-959.

CHAPTER TWO

STRUCTURE AND EVOLUTION OF THE *FUSARIUM* MATING TYPE LOCUS:

NEW INSIGHTS FROM THE *GIBBERELLA FUJIKUROI* COMPLEX

Assemblons des faits pour nous donner des idées.
(Let us gather facts to give us some ideas)

Comte Georges-Louis Leclerc de Buffon
Histoire des Animaux (1749)

"Sequence First. Ask Questions Later."

Arend Sidow
Sequence First. Ask Questions Later. (2002)

Abstract

Mating type genes are central to sexual reproduction and compatibility in Ascomycete fungi. However the "*MAT*" loci experience unique evolutionary pressures that can result in rapid divergence and enhanced inter-specific gene-flow (lateral gene transfer). In this study, molecular evolution of *MAT* loci was considered using the genus *Fusarium* (Teleomorph: *Gibberella*) as a model. Both *MAT1-1* and *MAT1-2* "idiomorphs" from eleven species of the *Gibberella fujikuroi* species complex were sequenced. Molecular evolution of the *MAT* loci from these heterothallic (self-sterile) species was compared with that of the *MAT* loci from nine homothallic (self-fertile) species in the *Fusarium graminearum* species complex. Although *Fusarium* has previously been thought to have the same complement of four *MAT* genes that are found in *Neurospora*, we identified a novel gene, *MAT1-2-3*, that may be specific to the Hypocreales. All *MAT* genes share a similar set of cis-regulatory motifs, although homothallic species appear to have recruited novel regulatory elements to facilitate alternate expression of *MAT* genes. *MAT* phylogenies of the *G. fujikuroi* complex are inconsistent with species trees. Most notably, *Fusarium sacchari MAT1-1* appears to be unrelated to *MAT1-1* of any other *G. fujikuroi* relative. These findings could support theories that sex-related genes experience frequent inter-specific geneflow. *Fusarium MAT* genes are highly divergent compared to other nuclear genes as a result of both positive selection and, more importantly, relaxed selective constraint. However, *MAT* genes appear to have diverged more slowly among homothallic species in the *F. graminearum* complex, potentially due to frequent gene-flow.

Introduction

Ascomycete fungi exhibit either, or both, sexual and asexual modes of reproduction. Sexual reproduction involves fertilization between compatible haploid cells giving rise to a diploid zygote; followed by meiosis to generate recombinant, haploid, ascospores. Ascomycetes can be sexually self-fertile (homothallic) or self-sterile (heterothallic). In homothallic species, haploid cells derived by mitosis from the same haploid parent are sexually compatible. In heterothallic species, cells are compatible only with those derived from distinct haploid ascospores of different "mating type".

Heterothallic Ascomycetes have two mating types, determined by distinct allelic forms of the mating type (*MAT*) locus. The dissimilarity between the two "alleles" is maintained because this specific portion of the genome does not pair up and undergo homologous recombination during meiosis (Coppin *et al.* 1997; Kronstad and Staben 1997). Metzenberg and Glass (1990) introduced the term idiomorph to recognize the fact the two alleles of the *MAT* locus have distinct sequences that encode different proteins. By convention, the idiomorphs are referred to as *MAT1-1* and *MAT1-2* (Turgeon and Yoder 2000).

The genes encoded by the *MAT* locus are putative transcription factors thought to control sexual development and to regulate the expression of downstream, mating-type-specific genes (Coppin *et al.* 1997; Shiu and Glass 2000). While the gene composition of the locus varies dramatically among species, there are two *MAT* genes that are consistently found in filamentous Ascomycetes. *MAT1-1* always contains a gene called *MAT1-1-1*, which encodes a protein homologous to MATα1 of *Saccharomyces cerevisiae*. This protein has a unique motif called the α-box. *MAT1-2* always contains a gene called *MAT1-2-1* that encodes a protein with a high-mobility-group (HMG) DNA-binding domain.

Homothallic Ascomycetes carry homologues of both *MAT1-1-1* and *MAT1-1-2* genes in the same genome (with the possible exception of a few *Neurospora* species [Glass *et al.* 1988; Glass and Smith 1994]). Coppin *et al.* (1997) proposed that self-fertility may be achieved via alternate expression of genes for one or the other mating type in distinct cells. Homothallic species would thus require a mechanism for independent regulation of *MAT1-1-1* and *MAT1-1-2* (Coppin *et al.* 1997). In most Ascomycetes, homothallism is thought to be a derived state (e.g. Yun *et al.* 1999; O'Donnell *et al.* 2004). However, in

many genera, the ancestral state is difficult to determine and it is likely that multiple independent transitions from heterothallism to homothallism and *vice-versa* have occurred (reviewed by Lee *et al.* 2010).

Studies on model Ascomycetes have highlighted interesting patterns in *MAT* gene evolution. Turgeon (1998) noted that *MAT* loci appear to be highly divergent between species yet strongly conserved within species. Rapid diversification of sex-related genes is a trend that has been observed in many plants and animals, and the forces of sexual selection, sexual conflict and reinforcement of reproductive barriers have all been invoked to explain the phenomenon (Reviewed by Civetta and Singh 1998; Swanson and Vaquier 2002; Clark *et al.* 2006). In *Neurospora*, Wik *et al.* (2008) found that *MAT* genes were divergent in both heterothallic and homothallic species, but due to distinct causes. Homothallic *MAT* genes were evolving neutrally, under a lack of selective constraint, and some contained premature stop codons. It was proposed that *MAT* genes may, therefore, be dispensable in homothallic *Neurospora* spp. By contrast, heterothallic *MAT* genes were under positive diversifying selection. It was postulated that positive selection, limited only to the heterothallic species (which are facultative out-crossers) could be driven by competition in attracting mates (Wik *et al.* 2008), which is tantamount to sexual selection.

An important possibility that has been ignored in previous studies is that the distinct genetic orgainisation of the *MAT* loci could also shape their evolution. Unlike heterothallic *MAT* idiomorphs, homothallic *MAT* loci can presumably recombine freely during meiosis. The extent of recombination experienced by a locus could affect the efficiency of selection (Birky and Walsh 1988) as well as the actual nucleotide substitution rate (Vicoso *et al.* 2009). It is, therefore, important to distinguish between the effects of adaptive and non-adaptive forces on the diversity of *MAT* genes.

Recent research has shown that phylogenies constructed using *Neurospora MAT* genes are in conflict with species trees (Strandberg *et al.* 2010). It was proposed that reproductive genes may be more predisposed to inter-specific gene-flow than others. Transfer of sexual genes through hybridization could restore sexuality to species in which it has been lost, as was observed in *Ophiostoma* (Paoletti *et al.* 2006). It has not been established whether homothallic species display similar characteristics (Strandberg *et al.* 2010). In fact, it could be more difficult to detect lateral transfer in homothallic than heterothallic *MAT* loci. For the latter group, the lack of recombination within the *MAT*

locus ensures that the phylogenetic signal of hybridization or lateral transfer is not diluted by exchange of genetic material between introgressed and "native" alleles (Chaturvedi *et al.* 2002; Devier *et al.* 2010).

A useful model in which to study *MAT* locus evolution in fungi is found in the genus *Fusarium*. Heterothallic *Fusarium* spp. such as those in the well known *Gibberella fujikuroi* species complex, have three genes in the *MAT1-1* idiomorph (*MAT1-1-1*, *MAT1-1-2* and *MAT1-1-3*) and a single gene in the *MAT1-2* idiomorph (*MAT1-2-1*) (Arie *et al.* 1999; Yun *et al.* 2000). These four *MAT* genes are homologous to the four *MAT* genes of the model species *Neurospora crassa* (Yun *et al.* 2000), where they are found in the same arrangement (Glass *et al.* 1990; Staben and Yanofsky 1990). In the homothallic *Fusarium graminearum* complex, in which homothallism has a monophyletic origin (O'Donnell *et al.* 2004) all four *MAT* genes are adjacent within the *MAT* locus (Yun *et al.* 2000). Homothallism in this species is thought to have arisen through an unequal crossover between *MAT* idiomorphs in a heterothallic ancestor (O'Donnell *et al.* 2004). Unlike *Neurospora*, there is no evidence that *MAT* genes are dispensable in the homothallic *Fusarium graminearum*. Knock-out strains lacking either the *MAT1-1* or *MAT1-2* genes can only outcross in a heterothallic manner (Lee *et al.* 2003). Studies of *Fusarium MAT* loci have identified potential phylogenetic inconsistencies as well as findings in conflict with the selective trends in the *Neurospora* model (Steenkamp *et al.* 2000; O'Donnell *et al.* 2004). However, there has been no adequate comparison between the *F. graminearum* complex and an equivalent, monophyletic, heterothallic lineage.

In this study the *Gibberella fujikuroi* complex was used as a monophyletic, heterothallic lineage for comparison with species in the homothallic *F. graminearum* complex. The aims were to (1) identify and compare evolutionarily conserved regions and cis-regulatory motifs in the *MAT* loci of heterothallic and homothallic *Fusarium* spp.; (2) investigate whether phylogenetic conflicts exist between the *MAT* and species trees; and (3) compare patterns of molecular evolution in both coding and non-coding portions of heterothallic and homothallic loci. To achieve these aims, the complete *MAT1-1* and *MAT1-2* idiomorphs of eleven species in the *G. fujikuroi* complex, including nine mating populations (A to I), were sequenced. This allowed investigation of the structure and molecular evolution of the *MAT* locus, as well as a comparison between heterothallic and homothallic species.

Materials and Methods

Fungal isolates

Twenty two *Fusarium* isolates, a *MAT-1* and a *MAT-2* isolate for each of eleven species (Table 1), were used in this study. For the nine mating populations, the isolates used represent the standard *MAT-1* and *MAT-2* mating type tester strains. For two additional species, *F. subglutinans* group 1 and *F. mangiferae*, isolates characterized by Steenkamp *et al.* (2002) and Britz *et al.* (2002), respectively were used. DNA was extracted from all isolates using the procedure described by Steenkamp *et al.* (1999).

DNA sequencing

Using published *MAT1-1* and *MAT1-2* idiomorphs of *F. oxysporum* and *F. verticillioides* (Yun *et al.* 2000) and their genome sequences (GenBank accession numbers: AAXH01000548 and AAIM02000073, respectively), PCR primers were designed in conserved regions to amplify overlapping fragments covering the length of both idiomorphs (Fig. 1). Reaction mixtures were 25ul in volume and contained 4ng/µl template DNA, 1.5mM $MgCl_2$, 0.4µM of each primer, 1mM deoxynucleotide triphosphates (0.25 mM of each), and 0.05u/µl Super-Therm DNA Polymerase and reaction buffer (Southern Cross biotechnology [Pty.] Ltd., Cape Town, South Africa). The PCR cycling conditions consisted of an initial denaturation at 94°C for 60s; thirty cycles of denaturation at 94°C for 30s, annealing for 30s and extension at 70°C; followed by a final extension step at 70°C for 10 min. All PCR amplifications from genomic DNA were performed using the same protocol, except for annealing temperatures that depended on the primers used (Table 2) and extension times that depended on the expected length of the amplicons. Extension times were calculated by adding 1min for each kb of sequence amplified.

The complete idiomorphs were amplified using primers MFC-R1 and MFC-L1, located in the 5’ and 3’ flanking regions, respectively. In cases where this PCR failed, MFC-L1 was substituted with a primer just inside the 3’ end of the idiomorph (M1C-2 for *MAT1-1* and M2C-4 For *MAT1-2*). Products were analysed by electropheresis in a 1% agarose gel at 20V/cm to verify purity and fragment size. *MAT1-1* and *MAT1-2* products were pooled to be sequenced together in a single lane, thus reducing the cost of sequencing. Pools consisted of products from distinct species to avoid confusion in assembly arising from homology in the flanks. 454 sequencing was performed by Inqaba Biotech (Pretoria,

South Africa). The pools were prepared separately using adapters carrying distinct identifiers, and sequenced together in one lane on a Roche GS-FLX Platinum sequencer.

To fill in gaps as well as verify areas of low pyro-sequence coverage, PCR products were amplified using the appropriate primer pairs (Fig. 1 and Table 2). For *F. proliferatum* in which all PCRs incorporating the 3' MFC-L1 primer were unsuccessful, an additional primer, MFC-R2, was designed based on the 3' flank region of close relatives *F. fujikuroi* and *F. mangiferae*, to be used as a substitute for MFC-L1. Products were cloned in *Escherichia coli* using the pGEM®-T Easy Vector cloning system (Promega Corporation, Madison, WI). Cloned inserts were then amplified directly from colonies using the primers SP-6 (5'-ATTTAGGTGACACTATAG-3') and T-7 (5'-TAATACGACTCACTATAGGG-3'). This PCR reaction mixture was 25ul and contained 0.05u/µl FastStart Taq DNA Polymerase (Roche Diagnostics, Mannheim, Germany), 1x Faststart buffer with $MgCl_2$, 1mM deoxynucleotide triphosphates (0.25 mM of each), 0.4µM of each primer and one bacterial colony. The colony PCR reaction conditions were as follows: Denaturation at 94ºC for 5 min; followed by thirty cycles of denaturation at 94ºC for 30s, annealing at 50ºC for 30s and extension for 90s at 70ºC; followed by a final extension step at 70ºC for 10 min. Following purification with sodium acetate (pH 3.8; 0.1M) (Sambrook and Russell 2001) the PCR products were sequenced in both directions using the SP-6 and T-7 primers, the BigDye® terminator v3.1 cycle sequencing kit (Applied Biosystems, Foster City, CA) and an ABI PRISM 377 automated DNA sequencer (Perkin Elmer, Warrington, United Kingdom).

Structural organization and conserved regions in the *MAT* loci

Sequence assembly, annotation and *in silico* translation were performed using CLC Bio Genomics Workbench (CLC Bio, Aarhus, Denmark). Sequences were aligned using ClustalW (Thompson *et al.* 1994), except for coding sequences, which were aligned using the codon-based algorithm of SQUINT (Goode and Rodrigo 2007). The start of the flanking regions (i.e. the point at which homology is restored between *MAT-1* and *MAT-2* genomes) were identified manually by aligning the idiomorph sequences from the *MAT-1* and *MAT-2* individuals of the same species. Structural variation was investigated with the graphical alignment program GATA (Nix and Eisen 2005), based on the Bl2Seq algorithm (Tatusova and Madden 1999), which enables duplications and inversions to be visualized.

Evolutionarily conserved regions were identified using dnaSP Version 5 (Librado and Rozas 2009) by determining nucleotide diversity (π) over a sliding window of 100bp and sliding in increments of 25bp. Genes were identified using the online tools FGENESH and AUGUSTUS (Stanke *et al.* 2004). The 500bp region upstream of each start codon was investigated to identify putative cis-regulatory transcription factor binding sites using MOTIFSEARCH (http://motif.genome.jp), based on the TRANSFAC database (Wingender *et al.* 1996), using the default cut-off value of 85% similarity. Only 500bp was used because this is roughly the size of most intergenic regions in the *MAT* locus.

Phylogenetic relationships

Trees were generated for each of the *MAT* idiomorphs using the sequences obtained in this study. All flanking DNA was excluded to avoid any dilution of phylogenetic signal due to recombination. Most appropriate nucleotide substitution models were chosen using jModelTest (Posada 2008). Maximum likelihood phylogenetic trees were generated using PhyML 3.0 (Guindon *et al.* 2009) and internal branches were evaluated with 1000 bootstrap replicates. The *Fusarium oxysporum MAT* idiomorph sequences (AB011379 and AB011378) were used as outgroups.

The generated *MAT* trees were compared to one-another and to the species tree with respect to their fit to the sequence data using the Approximately Unbiased (AU) (Shimodaira 2002) and Shimodaira-Hasegawa (SH) (Shimodaira and Hasegawa 1999) tests, as implemented in CONSEL (Shimodaira and Hasegawa 2001). The species tree was inferred by identifying consistently well-supported nodes from previous studies (O'Donnel *et al.* 1998, O'Donnell *et al.* 2000, Kvas *et al.* 2009). Only nodes supported by bootstrap-values higher than 70% were retained in the *MAT* trees. Site likelihood values required by CONSEL were generated for each competing tree using PAUP ver. 4.0b10 (Swofford 2002).

Sequence evolution in coding and non-coding regions

Variability in the *Fusarium MAT* proteins was compared with an estimated genome-wide average. For this purpose, a set of amino acid sequences was generated using 100 randomly-selected protein-coding genes from each of four available *Fusarium* genomes: *F. verticillioides* (AAIM02000073), *F. oxysporum* (AAXH01000548), *F. circinatum* (unpublished) and *F. graminearum* (NZ_AACM00000000). To avoid variation due to differences in gene annotation, alignments for genes with large gaps or non-homologous regions were not included in the analysis. All amino acid sequence alignments are

provided in Appendix B. The percentage-identity values for the 100 proteins among the three species were compared with those for the five *MAT* proteins.

In order to consider the selective pressures affecting each *MAT* gene, the likelihoods of various models of codon evolution were calculated using the CODEML program of the PAML package, version 4.3 (Yang 2007). Each model imposed different constraints on ω, which is the ratio of non-synonymous substitutions per non-synonymous site to synonymous substitutions per synonymous site. We first investigated whether genes have experienced selective constraint significantly stronger than what could be expected under neutral evolution. A value of $\omega = 1$ indicates an equal rate of synonymous and non-synonymous substitutions, implying neutrality. A neutral model in which $\omega = 1$ was fixed, was compared to model M0 which allowed an estimated ω value. We then investigated whether positive selection drives *MAT* gene evolution. Positive selection is expected to cause more non-synonymous than synonymous substitutions, resulting in $\omega > 1$. However, since most proteins probably have some codons consistently under selective constraint, models that allowed codons to fall into various classes with different ω values were compared. Model M7 (the beta model) allowed a β distribution of codons in the range $0 \leq \omega \leq 1$, while model M8 (the beta&ω model) allowed an additional class of $\omega \geq 1$. The likelihoods of these two pairs of models were compared using the likelihood ratio test (LRT), as described by Yang *et al.* (2000). Identical analyses were performed on *MAT* genes from both species complexes. For the *F. graminearum* complex, *MAT* locus sequences obtained by O'Donnell *et al.* (2004) were used to represent nine *F. graminearum* species (one sequence for each species, accession numbers are listed in Appendix A).

The levels of nucleotide diversity in the heterothallic and homothallic *MAT* loci were also compared. Nucleotide diversity (π) and its variance was measured using dnaSP. *Fusarium sacchari* was excluded from this analysis to avoid any skewing of the result by the highly divergent *MAT1-1* sequence. Coding and non-coding portions of the locus were considered as two separate datasets. To eliminate variation due to different evolutionary time-scales, each measurement of π was divided by that generated for a sequence consisting of partial segments of house-keeping genes Translation Elongation Factor 1-α and Tubulin β-1 chain from the same set of species (accession numbers are listed in Appendix A). This scaling procedure was based on the assumption that these housekeeping genes would experience comparable selective pressures and evolutionary rates in both heterothallic and homothallic species.

Results

Structural organization and conserved regions in the *MAT* loci

Complete idiomorph sequences, including portions of both 5' and 3' flanks, were obtained for all 22 *Fusarium* isolates, except for the region at the 3' end of the *F. sacchari MAT1-2* idiomorph. Because the rest of the derived idiomorph sequence for the same isolate was identical to that obtained by Arie *et al.* (1999), the missing sequence was also assumed to be identical. All twenty two complete idiomorph sequences, with annotation, are provided in Appendix C. The length of the idiomorph (non-homologous region) ranged from 4.3kb to 4.6kb for *MAT1-1* and 3.5kb to 3.8kb for *MAT1-2*. This variation was largely due to various structural differences in the region at the 3' end of the idiomorph and extending into the 3' flank. There appeared to be three different categories of 3' flank structure, which corresponded to the major clades of the *G. fujikuroi* complex, designated the "African", "Asian" and "American" clades by O'Donnell *et al.* (1998). In the American and Asian clades, the final exon of *MAT1-1-1* extended slightly into the 3' flank (Fig. 2). However, the lack of an appropriate start codon precludes expression of this short ORF in *MAT-2* isolates. In the African clade, the non-homologous region extended about 250bp further, due to an inversion in the *MAT1-2* idiomorph (Fig. 2). In the American clade, the *MAT1-2* idiomorph was shorter than others due to a deletion of about 150bp at the 3' end (Fig. 2). In the Asian clade the position of the newly developed MFC-L1 primer was shifted 600bp downstream due to the insertion of two inverted fragments found further downstream in other clades. However, the 3' flanks of the two *F. sacchari* isolates were unlike those of other Asian clade species, and even somewhat dissimilar from each other. The structural events that gave rise to the unique organization of the *F. sacchari* flanks could not be resolved.

In general, non-coding portions of the *MAT* loci were found to be more variable between species than coding portions (Fig. 3). However, an area of low diversity was observed in the heterothallic *MAT1-2* idiomorph in a region where no gene had previously been described (Fig. 3). Gene prediction using both FGENESH and AUGUSTUS revealed a predicted gene in this region with three exons. Using the predicted protein product for tBLASTn analysis (protein query against translated nucleotide database) against the complete genomes of *F. verticillioides* and *F. oxysporum*, both of which represent *MAT-1* isolates, yielded no hits. In the homothallic *F. graminearum* genome however, there was a single significant hit to a predicted gene (FGSG 08894) located immediately adjacent to *MAT1-2-1* at the *MAT* locus. This location corresponded to the previously unknown

gene's position in the *MAT1-2* idiomorph of *F. verticillioides,* although it displayed an inverted orientation. GATA analysis provided evidence of an inversion of this region in *F. graminearum* (Fig. 4). The previously unknown gene was, therefore, recognized as a true *MAT* gene, and is designated *MAT1-2-3*. The name *MAT1-2-2* has already been used to describe a different gene (Kanamori *et al.* 2007), which is not homologous to *MAT1-2-3*.

To gain insight into the range of species in which *MAT1-2-3* may be present, a tBLASTn analysis against the GenBank database was performed. Apart from the *Fusarium* species mentioned herein, there were weak hits in two additional species in the order Hypocreales; *Cordyceps militaris* (accession, AB084257) (E = 0.026) and *Paecilomyces tenuipes* (AB084921) (E = 0.31). Both hits corresponded to a position immediately adjacent to the described *MAT1-2-1* gene, suggesting that these may be true orthologues.

All five *MAT* genes (*MAT1-1-1*, *MAT1-1-2*, *MAT1-1-3*, *MAT1-2-1*, and *MAT1-2-3*) had intact open reading frames (ORFs) of near-identical length among the eleven heterothallic *Fusarium* species, with two exceptions. In *F. sacchari*, *MAT1-1-1* had a stop-codon eleven codons upstream of that in the other species. This stop-codon is present in both the sequence produced in this study as well as that published by Arie *et al.* (1999). This result indicates that this species encodes a slightly truncated *MAT1-1-1* protein. In *F. circinatum*, the first exon of the newly-identified *MAT1-2-3* gene was predicted to begin 126nt downstream of the start codon predicted in the other species, thus reducing the predicted protein by 42 amino acids. The upstream start codon is present in *F. circinatum* but a stop codon follows downstream. The presence of this stop codon was confirmed by PCR and Sanger sequencing. It is, therefore, unclear whether the *F. circinatum* prediction represents the true start codon for all of these species or whether it has a different start codon to the others.

Eleven previously described transcription factor binding sites (most described in *S. cerevisia*e) were predicted in the upstream 500bp regions of *MAT* genes (Table 3). Multiple binding sites of a heat shock factor (HSF), an alcohol dehydrogenase gene regulator (ADR1) and an activator of nitrogen-regulated genes (NIT2) were predicted upstream of all genes, although the number of binding sites sometimes differed among species. Several binding sites were found upstream of a subset of the genes. For example, a stress response element (STRE) binding site was present upstream of all genes except *MAT1-2-3*. The most notable difference between the promoters of heterothallic and

homothallic species involved predicted binding sites of MAT-Mc and MATα2 (Fig. 5). In heterothallic species, all five genes carried a predicted MAT-Mc binding site and no MATα2 binding site. In homothallic species, *MAT1-1-1* had the same arrangement, however, the remaining four genes lacked MAT-Mc binding sites, and instead carried binding sites for MATα2 (Fig. 5).

Phylogenetic relationships

Phylogenetic analysis of both *MAT1-1* and *MAT1-2* sequences revealed three major clades for the G. *fujikuroi* isolates (Fig. 6). These correspond to the African, American and Asian clades described by O'Donnell *et al.* (1998). However the *MAT1-1* idiomorph of *F. sacchari* did not fall into any of the three major clades, while the *MAT1-2* idiomorph of this species grouped within the Asian clade as expected.

The *MAT* trees displayed several well-supported conflicts with the known species tree (Fig. 6). Comparison of competing topologies using the AU and SH tests revealed that neither idiomorph supported the known species tree topology over the topology of the ML tree (Table 4). Furthermore, neither *MAT1-1* nor *MAT1-2* supported the phylogeny representing the opposing idiomorph (Table 4). The unexpected grouping of the *F. sacchari MAT1-1* idiomorph was also addressed independently by comparing three trees that differed only in the position of *F. sacchari*. Placement of *F. sacchari* within the Asian clade could be rejected for the *MAT1-1* idiomorph but not *MAT1-2*. Placement of *F. sacchari* outside of the Asian clade but within the *G. fujikuroi* complex could be rejected for both idiomorphs. However, placement of *F. sacchari* outside of the complex could be rejected for *MAT1-2* but not for *MAT1-1*. This implies that the most likely phylogenetic placement of the *F. sacchari MAT1-1* idiomorph is outside of the *G. fujikuroi* complex.

Sequence evolution in coding and non-coding regions

Among the four species *F. oxysporum*, *F. verticillioides*, *F. circinatum*, and *F. graminearum*, all five *MAT* genes displayed levels of percentage identity well below the average of 100 randomly-selected nuclear genes (Fig. 7). The *MAT* genes shared most similarity between the two *G. fujikuroi* complex species *F. circinatum* and *F. verticillioides*, and were most divergent when the three heterothallic species were compared with *F. graminearum*.

The results of the CODEML analyses indicated that all five genes in heterothallic species deviated significantly from the neutral model of evolution. By contrast, in homothallic species, *MAT1-1-1* and *MAT1-1-3*, did not demonstrate a significant deviation from the neutral model. However, *MAT1-1-1* displayed evidence for the presence of certain codons under positive selection in both heterothallic and homothallic species (Table 5).

Scaled values of nucleotide diversity (π) demonstrated that the coding regions of the *MAT* loci were comparatively less divergent among homothallic species (Fig. 8). However, the ratio of π at non-synonymous sites to π at synonymous sites was similar in the heterothallic and homothallic species (0.253 and 0.243, respectively). This showed that the observed lower diversity in the coding regions of the homothallic *MAT* genes was reflected at both non-synonymous and synonymous codon positions. In contrast to the coding portion, the non-coding portion of the *MAT* locus was less divergent among heterothallic species (Fig. 8).

Discussion

Previous studies of *Fusarium* mating type loci have assumed that species in this genus carry the same complement of *MAT* genes as the model genus *Neurospora* (Arie *et al.* 1999; Yun *et al.* 2000; O'Donnell *et al.* 2004). In this study, a previously unknown gene was discovered in the *Fusarium MAT* locus and has been designated *MAT1-2-3*. The predicted protein sequence has no known functional domains, and further investigation is required to confirm that this gene is expressed and if so, to determine its ontology. This gene is absent from *MAT-1* genomes and is thus specific to individuals of the *MAT-2* mating type in heterothallic species. This unexpected discovery of a previously unknown *MAT* gene arose from the identification of a strongly conserved area in a region originally thought to constitute the non-coding portion of the *MAT1-2* idiomorph. This demonstrates the power of comparative sequence analysis for the identification of functionally conserved DNA sequences.

A homologue of the newly discovered *MAT1-2-3* is present in the *F. graminearum* genome and its location supports the view of O'Donnell *et al.* (2004) that homothallism arose through the fusion of heterothallic *MAT* loci. However, the gene's orientation is reversed, suggesting a segmental inversion of this region subsequent to the fusion. The portion of the homothallic genome referred to as the *MAT* locus (Yun *et al.* 2000) is, therefore, also larger than previously believed. It was possible to tentatively predict that

MAT1-2-3 may be specific to species in the Hypocreales. However, given the high variability of *MAT* genes, this distribution could represent an under-estimation of realty.

A similar set of transcription factors appears to be associated with all five *MAT* genes, including the newly identified *MAT1-2-3*, which further corroborates its authenticity. Predicted binding sites for transcription factors involved in coordination of cell growth and cell-cycle, mating and response to environmental stress were identified, the latter class being most abundant. Predicted binding sites of ADR1 and NIT2 were identified upstream of all genes. These transcription factors, along with MIG1, are involved in responding to the absence of certain nutrients. This is consistent with the fact that nutrient starvation is a well known stimulus of sexual reproduction in Ascomycetes (Coppin *et al.* 1997; Poggeler *et al.* 2006). The presence of multiple heat shock factor (HSF) binding sites suggests that heat shock may also be involved. In *Podospora anserina*, a heat shock protein *mod-E* is known to be necessary for sexual reproduction (Loubradou *et al.* 1997). Sexual development in *Fusarium* species is highly temperature-dependent (reviewed by Doohan *et al.* 2003), although heat-shock *per se* has never been implicated. The presence of predicted stress response element binding sites upstream of most *MAT* genes is not surprising, as stress response elements can be involved in the response to both nutrient starvation and heat-shock (Marchler *et al.* 1993).

The most significant difference in the regulatory motifs between heterothallic and homothallic species involved predicted binging sites of MAT-Mc and MATα2. MAT-Mc is an HMG-domain protein found in the *S. pombe MAT* locus (Kjaerulf *et al.* 1997) that may be orthologous to *MAT1-2-1* of filamentous species (Staben and Yanofsky 1990). Binding sites of MAT1-Mc were predicted upstream of all five *MAT* genes in heterothallic species. By contrast, in homothallic species, only *MAT1-1-1* carried a predicted MAT1-Mc binding site. The other four *MAT* genes instead carried predicted binding sites for MATα2. This is a homeodomain protein encoded by the *S. cerevisiae* MAT-α idiomorph (Johnson 1995), of which there is no known orthologue in filamentous species. Exactly how these mechanisms work to regulate the *Fusarium MAT* genes will require further investigation. Nevertheless, the significant implication of these findings is that the two most important *MAT* genes, *MAT1-1-1* and *MAT1-2-1*, are regulated by the same mechanism in heterothallic species but by two separate mechanisms in homothallic species. This supports Coppin's (1997) prediction that a transition from heterothallism to homothallism must involve recruitment of distinct regulatory elements to facilitate alternate expression of *MAT1-1-1* and *MAT1-2-1*. Following this reasoning, both *MAT1-*

1-2 and *MAT1-1-3* (which are *MAT-1*-associated genes in heterothallic species) appear to be co-regulated with the *MAT-2*-associated genes in homothallic species. This is plausible, as these two genes are probably not directly involved in mating type determination (Shiu and Glass 1997). In fact, the *P. anserina* homologue of *MAT1-1-2* is functional when present in either parental genome (Arnaise *et al.* 1997). To the best of our knowledge, this is the first evidence of an alteration in the regulatory mechanism of *MAT* genes associated with a shift in reproductive mode.

Both *MAT* idiomorphs of species in the *G. fujikuroi* complex displayed significant support for phylogenies in conflict with the recognized species tree. It is possible that some of the minor incongruities could be the result of incomplete lineage sorting. However, this process requires the existence of polymorphism prior to speciation, and intra-specific polymorphisms in heterothallic *MAT* loci are thought to be rare and short-lived (Turgeon 1998). Another possible cause of the phylogenetic irregularities is inter-specific gene flow (lateral gene transfer). Similar findings in *Neurospora* led Strandberg *et al.* (2010) to propose that reproductive genes might be more prone to lateral gene flow than others. This is because selection would favour lateral transfers that restore sexuality to populations in which it has been lost. Such lateral transfers between close relatives can simply occur through hybridization followed by back-crossing into the parent population (Paoletti *et al.* 2006). The potential to reproduce sexually can be lost through the rise of single-mating-type populations (in heterothallic species only) (e.g. Paolletti *et al.* 2006) or through genetic decay of reproductive genes (Strandberg *et al.* 2010; Wik *et al.* 2008). Indeed in *Fusarium*, many isolates display limited or no fertility (Kuhlman 1982), the causes of which remain unknown. It is, therefore, possible that selection may periodically favour lateral transfer of functional *MAT* genes among *Fusarium* spp.

The most intriguing phylogenetic aberration arising in this study was the grouping of the *F. sacchari MAT1-1* idiomorph outside the *G. fujikuroi* complex, while the *MAT1-2* idiomorph grouped in the Asian clade as expected. Further investigation including use of a larger dataset might improve phylogenetic resolution and rule out potential artefacts such as long-branch-attraction. The lateral transfer hypothesis appears doubtful because this would require a hybridization event between *F. sacchari* and a somewhat distant relative. Even within the *G. fujikuroi* complex, distinct mating populations (of which *F. sacchari* constitutes MP-B) are thought to be strongly reproductively isolated (Leslie 1991; Leslie *et al.* 2004). Interestingly, this opinion did not prevail originally; early studies on *F. sacchari* suggested that it may indeed be cross-fertile with several other

mating populations (Kuhlman 1982). However, Britz *et al.* (1999) demonstrated that the progeny produced in such crosses, rather than being hybrids, were uniparental. This implied that *F. sacchari* was in fact capable of homothallic reproduction. In light of the knowledge that *F. sacchari* has a heterothallic *MAT* locus and a *MAT1-1* idiomorph that is highly dissimilar from others in the complex, these earlier claims should be re-visited.

Unlike the phylogenies generated in this study, the *MAT* phylogeny of the homothallic *F. graminearum* complex generated by O'Donnell et al. (2004) contained few well-supported deviations from the species tree. This should not be construed as evidence against lateral gene-flow in homothallic species. On the contrary, given the presence of free recombination in homothallic *MAT* loci, rather than phylogenetic aberrations, one may expect to find that the *MAT* genes are simply under-diverged with low phylogenetic informativeness, a result that was indeed reported by O'Donnell et al. (2004).

Of the four available *Fusarium* genomes, only one was homothallic. A specific comparison between the rate of divergence among heterothallic species and the rate of divergence among homothallic species was not possible. However, in general *MAT* proteins were highly divergent compared to a genome-wide estimate. This is in agreement with the findings for *Cochliobolus* (Turgeon 1998) and *Neurospora* (Wik *et al.* 2008) and also with the trend of rapid evolution in reproductive proteins (Civetta and Singh 1998; Swanson and Vaquier 2002; Clark *et al.* 2006). This rapid divergence of *MAT* genes over short evolutionary distances helps to explain their dramatic lack of conservation among genera and the difficulty experienced by researchers in identifying *MAT* loci in new species (Cisar *et al.* 1994; Arie *et al.* 1997).

By scaling the level of nucleotide diversity at the *MAT* locus in the two species complexes to the level of diversity at two house-keeping genes, we were able to specifically compare the rate of divergence among homothallic species with the rate of divergence among heterothallic species. Using this measure, the coding portion of the homothallic locus indeed displayed lower relative nucleotide diversity. Nucleotide diversity was lower at both synonymous and non-synonymous sites, indicating that this was not the result of stronger purifying selection, which would only reduce diversity in the latter. This result is therefore consistent the idea that recombination combined with gene-flow might have a homogenising effect on *MAT* genes in the homothallic *F.*

graminearum complex. A lack of recombination in *MAT* loci of the heterothallic G. fujikuroi complex would thus result in more divergent *MAT* genes, even in the face of gene-flow.

In contrast with the coding sequences, non-coding portions of the *MAT* loci displayed greater relative inter-specific divergence in homothallic than heterothallic species. This result could possibly be explained by the difference in the extent of linkage. Heterothallic *MAT* loci do not recombine and all parts are therefore strongly linked. Non-coding regions of heterothallic *MAT* loci are therefore not independent, and could diverge more slowly due to functional constraint acting upon linked coding regions (Hudson 1994). Free recombination in homothallic species would mean that the non-coding portions are independent of the genes and liable to less constraint.

In both the heterothallic and homothallic complexes *MAT1-1-1* displayed evidence for the presence codons under positive diversifying selection. In homothallic species, *MAT1-1-1* and *MAT1-1-3* also exhibited evolutionary rates that were not significantly different from the neutral expectation. These findings are unlike those of O'Donnell *et al.* (2004) who reported *Fusarium MAT* genes to be under "*strong purifying selection*". This distinction could be attributed to the different datasets used. In the aforementioned study, all sequences were analysed together, while here a single sequence was used to represent each species and the distinct homothallic and heterothallic complexes were analysed separately. The results of this study are, however, similar to those of Wik *et al.* (2008), who reported a lack of selective constraint affecting *MAT* genes of homothallic *Neurospora* spp. While some *Neurospora MAT* genes might be dispensable for homothallic reproduction (Glass *et al.* 1990; Wik *et al.* 2008), this is not the case in *Fusarium* (Lee *et al.* 2003). It is thus likely that homothallic *Fusarium MAT* genes experience weaker purifying selection rather than a complete lack of selective constraint. In fact, even in the heterothallic *MAT* genes, the high level of protein diversity implies that a lower-than-average level of constraint is present. This might simply result from the fact the sexual reproduction is not an essential part of the life cycle. Because individuals can survive and reproduce with non-functional (or partially functional) *MAT* genes, fixation of some mildly deleterious mutations could be tolerated. This argument is not entirely, satisfactory because it is well recognised that the sexual spores of *F. graminearum* are the main mode of dispersal and infection, and therefore integral to

fitness. However, homothallic species retain the ability to outcross in a heterothallic manner, even in the absence of one functional *MAT* gene (Lee *et al* 2003). This redundancy might explain the observation of greater relaxation of selective constraint in homothallic species.

This study has provided new insights into the structure and evolution of the *Fusarium MAT* locus. *Fusarium* spp. have a novel *MAT* gene, *MAT1-2-3*, absent from many other genera, the functional relevance of which will be an important focus for future study. The presence of distinct regulatory motifs may facilitate alternate expression of *MAT1-1-1* and *MAT1-1-2* in homothallic species. However, further work is required to understand whether alternate expression occurs and also how the distinct regulatory motifs could have arisen *de novo* in homothallic species. *Fusarium MAT* proteins evolve rapidly, but in the *F. graminearum* complex *MAT* gene divergence could be limited by frequent inter-specific gene-flow. While *MAT1-1-1* is under positive selection, drift under relaxed selective constraint is probably a more important contributor to *MAT* gene divergence, particularly in homothallic species. Despite strong reproductive barriers in the *G. fujikuroi complex*, there is evidence that the *F. sacchari MAT1-1* sequence could have been acquired through lateral transfer from a somewhat distant relative, although further study is essential to verify this finding. Overall these findings demonstrate that *Fusarium MAT* loci offer a useful model in which study many different aspects of molecular evolution.

References

Arie, T., Christiansen, S. K., Yoder, O. C. and Turgeon, B. G. 1997. Efficient Cloning of Ascomycete Mating Type Genes by PCR Amplification of the ConservedMATHMG Box. *Fungal Genetics and Biology* **21**: 118-130.

Arie, T., Yoshida, T., Shimizu, T., Kawabe, M., Yoneyama, K. and Yamaguchi, I. 1999. Assessment of *Gibberella fujikuroi* mating type by PCR. *Mycoscience* **40**: 311-314.

Arnaise, S., Debuchy, R. and Picard, M. 1997. What is a bona fide mating-type gene? Internuclear complementation of mat mutants in Podospora anserina. *Molecular and General Genetics* **256**: 169-78.

Barbara, K., Haley, T., Willis, K. and Santangelo, G. 2007. The transcription factor Gcr1 stimulates cell growth by participating in nutrient-responsive gene expression on a global level. *Molecular Genetics and Genomics* **277**: 171-188.

Birky, C. W. and Walsh, J. B. 1988. Effects of linkage on rates of molecular evolution. *Proceedings of the National Academy of Sciences of the United States of America* **85**: 6414-6418.

Britz, H., Coutinho, T. A., Wingfield, M. J., Marasas, W. F. Â O., Gordon, T. R. and Leslie, J. F. 1999. *Fusarium subglutinans* f. sp. pini represents a distinct mating population in the *Gibberella fujikuroi* species complex. *Applied and Environmental Microbiology* **65**: 1198-1201.

Britz, H., Steenkamp, E. T., Coutinho, T. A., Wingfield, B. D., Marasas, W. F. O. and Wingfield, M. J. 2002. Two new species of *Fusarium* section *Liseola* associated with mango malformation. *Mycologia* **94**: 722-730.

Carlson, M. 1999. Glucose repression in yeast. *Current Opinion in Microbiology* **2**: 202-207.

Chaturvedi, V., Fan, J., Stein, B., Behr, M. J., Samsonoff, W. A., Wickes, B. L. and Chaturvedi, S. 2002. Molecular genetic analyses of mating pheromones reveal intervariety mating or hybridization in *Cryptococcus neoformans*. *Infection and Immunity* **70**: 5225-5235.

Cheng, C., Kacherovsky, N., Dombek, K. M., Camier, S., Thukral, S. K., Rhim, E. and Young, E. T. 1994. Identification of potential target genes for Adr1p through characterization of essential nucleotides in UAS1. *Mol. Cell. Biol.* **14**: 3842-3852.

Cisar, C. R., TeBeest, D. O. and Spiegel, F. W. 1994. Sequence similarity of mating type idiomorphs: a method which detects similarity among the Sordariaceae fails to detect similar sequences in other filamentous ascomycetes. *Mycologia* **86**: 540-546.

Civetta, A. and Singh, R. S. 1998. Sex-related genes, directional sexual selection, and speciation. *Molecular Biology and Evolution* **15**: 901-909.

Clark, N. L., Aagaard, J. E. and Swanson, W. J. 2006. Evolution of reproductive proteins from animals and plants. *Reproduction* **131**: 11-22.

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

Devier, B., Aguileta, G., Hood, M. E. and Giraud, T. 2010. Using phylogenies of pheromone receptor genes in the Microbotryum violaceum species complex to investigate possible speciation by hybridization. *Mycologia* **102**: 689-696.

Doohan, F. M., Brennan, J. and Cooke, B. M. 2003. Influence of climatic factors on *Fusarium* species pathogenic to cereals. *European Journal of Plant Pathology* **109**: 755-768.

Dutton, J. R., Johns, S. and Miller, B. L. 1997. StuAp is a sequence-specific transcription factor that regulates developmental complexity in Aspergillus nidulans. *EMBO J* **16**: 5710-5721.

Gasser, T., Muller, H. G. and Mammitzsch, V. 1985. Kernels for nonparametric curve estimation. *Journal of the Royal Statistical Society. Series B (Methodological)* **47**: 238-252.

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

Glass, N. L. and Smith, M. L. 1994. Structure and function of a mating-type gene from the homothallic species *Neurospora africana*. *Molecular and General Genetics* **244**: 401-409.

Glass, N. L., Vollmer, S. J., Staben, C., Grotelueschen, J., Metzenberg, R. L. and Yanofsky, C. 1988. DNAs of the two mating-type alleles of *Neurospora crassa* are highly dissimilar. *Science* **241**: 570-573.

Goode, M. G. and Rodrigo, A. G. 2007. SQUINT: a multiple alignment program and editor. *Bioinformatics* **23**: 1553-1555.

Guindon, S., Delsuc, F., Dufayard, J. F. and Gascuel, O. 2009. Estimating maximum likelihood phylogenies with PhyML. *Methods in Mollecular and Cellular Biology* **537**: 113-37.

Hudson, R. R. 1994. How can the low levels of DNA sequence variation in regions of the *Drosophila* genome with low recombination rates be explained? *Proceedings of the National Academy of Sciences of the United States of America* **91**: 6815-6818.

Ina, Y. 1996. Pattern of synonymous and nonsynonymous substitutions: An indicator of mechanisms of molecular evolution. *Journal of Genetics* **75**: 91-115.

Johnson, A. D. 1995. Molecular mechanisms of cell-type determination in budding yeast. *Current Opinion in Genetics & Development* **5**: 552-558.

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

Kjaerulff, S., Dooijes, D., Clevers, H. and Nielsen, O. 1997. Cell differentiation by interaction of two HMG-box proteins: *Mat1-Mc* activates M cell-specific genes in *S.pombe* by recruiting the ubiquitous transcription factor *Ste11* to weak binding sites. *EMBO Journal* **16**: 4021-4033.

Klaasen, J. A. and Nelson, P. E. 1996. Identification of a mating population, *Gibberella nygamai* sp. nov., within the *Fusarium nygamai* anamorph. *Mycologia* **88**: 965-969.

Klittich, C. J. R. and Leslie, J. F. 1992. Identification of a second mating population within the *Fusarium moniliforme* anamorph of *Gibberella fujikuroi*. *Mycologia* **84**: 541-547.

Klittich, C. J. R., Leslie, J. F., Nelson, P. E. and Marasas, W. F. O. 1997. *Fusarium thapsinum* (*Gibberella thapsina*): A new species in section *Liseola* from sorghum. *Mycologia* **89**: 643-652.

Kronstad, J. W. and Staben, C. 1997. Mating type in filamentous fungi. *Annual Review of Genetics* **31**: 245-276.

Kuhlman, E. G. 1982. Varieties of *Gibberella fujikuroi* with Anamorphs in *Fusarium* Section *Mycologia* **74**: 759-768.

Kurtz, S. and Shore, D. 1991. RAP1 protein activates and silences transcription of mating-type genes in yeast. *Genes & Development* **5**: 616-628.

Kvas, M., Marasas, W. F. O., Wingfield, B. D., Wingfield, M. J. and Steenkamp, E. T. 2009. Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. *Fungal Diversity* **34**: 1-21.

Lee, J., Lee, T., Lee, Y. W., Yun, S. H. and Turgeon, B. G. 2003. Shifting fungal reproductive mode by manipulation of mating type genes: obligatory heterothallism of *Gibberella zeae*. *Molecular Microbiology* **50**: 145-152.

Lee, S. C., Ni, M., Li, W., Shertz, C. and Heitman, J. 2010. The Evolution of Sex: a Perspective from the Fungal Kingdom. *Microbiology and Molecular Biology Reviews* **74**: 298-340.

Leslie, J., Zeller, K., Wohler, M. and Summerell, B. 2004. Interfertility of Two Mating Populations in the Gibberella Fujikuroi Species Complex. *European Journal of Plant Pathology* **110**: 611-618.

Leslie, J. F. 1991. Mating populations in *Gibberella fujikuroi* (*Fusarium* section Liseola). *Phytopathology* **81**: 1058-1060.

Leslie, J. F. 1995. *Gibberella fujikuroi*: available populations and variable traits. *Canadian Journal of Botany* **73**: 282-291.

Leslie, J. F., Summerell, B. A., Bullock, S. and Doe, F. J. 2005. Description of *Gibberella sacchari* and Neotypification of Its Anamorph *Fusarium sacchari*. *Mycologia* **97**: 718-724.

Loubradou, G., Begueret, J. and Turcq, B. 1997. A mutation in an HSP90 gene affects the sexual cycle and suppresses vegetative incompatibility in the fungus *Podospora anserina*. *Genetics* **147**: 581-588.

Marchler, G., Schüller, C., Adam, G. and Ruis, H. 1993. A *Saccharomyces cerevisiae* UAS element controlled by protein kinase A activates transcription in response to a variety of stress conditions. *EMBO Journal* **12**: 1997-2003.

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

Natvig, D. and May, G. 1996. Fungal evolution and speciation. *Journal of Genetics* **75**: 441-452.

Nix, D. and Eisen, M. 2005. GATA: a graphic alignment tool for comparative sequence analysis. *BMC Bioinformatics* **6**: 9.

O'Donnell, K., Cigelnik, E. and Nirenberg, H. I. 1998. Molecular systematic and phylogeography of the *Gibberella fujikuroi* species complex. *Mycologia* **90**: 465-493.

O'Donnell, K., Nirenberg, H. I., Aoki, T. and Cigelnik, E. 2000. A multigene phylogeny of the *Gibberella fujikuroi* species complex: Detection of additional phylogenetically distinct species. *Mycoscience* **41**: 61-78.

O'Donnell, K., Ward, T. J., Geiser, D. M., Corby Kistler, H. and Aoki, T. 2004. Genealogical concordance between the mating type locus and seven other nuclear genes supports formal recognition of nine phylogenetically distinct species within the *Fusarium graminearum* clade. *Fungal Genetics and Biology* **41**: 600-623.

Paoletti, M., Buck, K. W. and Brasier, C. M. 2006. Selective acquisition of novel mating type and vegetative incompatibility genes via interspecies gene transfer in the globally invading eukaryote *Ophiostoma novo-ulmi*. *Molecular Ecology* **15**: 249-262.

Posada, D. 2008. jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution* **25**: 1253-1256.

Sambrook, J. and Russell, D. 2001. *Molecular Cloning: A Laboratory Manual* 3rd ed. Cold Spring Harbor Laboratory Press, Plainview, NY.

Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology* **51**: 492-508.

Shimodaira, H. and Hasegawa, M. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular Biology and Evolution* **16**: 1114.

Shimodaira, H. and Hasegawa, M. 2001. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**: 1246-1247.

Shiu, P. K. T. and Glass, N. L. 2000. Cell and nuclear recognition mechanisms mediated by mating type in filamentous ascomycetes. *Current Opinion in Microbiology* **3**: 183-188.

Siliciano, P. G. and Tatchell, K. 1984. Transcription and regulatory signals at the mating type locus in yeast. *Cell* **37**: 969-978.

Staben, C. and Yanofsky, C. 1990. *Neurospora crassa* a mating-type region. *Proceedings of the National Academy of Sciences of the United States of America* **87**: 4917-4921.

Steenkamp, E. T., Wingfield, B. D., Coutinho, T. A., Wingfield, M. J. and Marasas, W. F. O. 1999. Differentiation of *Fusarium subglutinans* f. sp. *pini* by histone gene sequence data. *Applied and Environmental Microbiology* **65**: 3401-3406.

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

Steenkamp, E. T., Wingfield, B. D., Desjardins, A. E., Marasas, W. F. O. and Wingfield, M. J. 2002. Cryptic speciation in *Fusarium subglutinans*. *Mycologia* **94**: 1032-1043.

Strandberg, R., Nygren, K., Menkis, A., James, T. Y., Wik, L., Stajich, J. E. and Johannesson, H. Conflict between reproductive gene trees and species phylogeny among heterothallic and pseudohomothallic members of the filamentous ascomycete genus *Neurospora*. *Fungal Genetics and Biology* **47**: 869-878.

Swanson, W. J. and Vacquier, V. D. 2002. The rapid evolution of reproductive proteins. *Nature Reviews Genetics* **3**: 137-144.

Swofford, D. L. 2002. *PAUP*: phylogenetic analysis using parsimony (*and other methods), version 4.0 b10*. Sinauer Associates, Sunderland, MA.

Tao, Y. and Marzluf, G. A. 1999. The NIT2 nitrogen regulatory protein of *Neurospora*: expression and stability of nit-2 mRNA and protein. *Current Genetics* **36**: 153-158.

Tatusova, T. A. and Madden, T. L. 1999. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiology Letters* **174**: 247-50.

Turgeon, B. G. 1998. Application of mating type gene technology to problems in fungal biology. *Annual Review of Phytopathology* **36**: 115 - 137.

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

Vicoso, B. and Charlesworth, B. 2009. Recombination rates may affect the ratio of X to autosomal non-coding polymorphism in African populations of *Drosophila melanogaster*. *Genetics*: genetics.108.098004.

Wik, L., Karlsson, M. and Johannesson, H. 2008. The evolutionary trajectory of the mating-type (*mat*) genes in *Neurospora* relates to reproductive behavior of taxa. *BMC Evolutionary Biology* **8**: 109.

Yang, Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**: 1586-1591.

Yang, Z. H., Nielsen, R., Goldman, N. and Pedersen, A. M. K. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431 - 449.

Young, E. T., Dombek, K. M., Tachibana, C. and Ideker, T. 2003. Multiple pathways are co-regulated by the protein kinase Snf1 and the transcription factors Adr1 and Cat8. *Journal of Biological Chemistry* **278**: 26146-26158.

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

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

Zeller, K. A., Summerell, B. A., Bullock, S. and Leslie, J. F. 2003. *Gibberella konza* (*Fusarium konzum*) sp. nov. from prairie grasses, a new species in the *Gibberella fujikuroi* species complex. *Mycologia* **95**: 943-954.

Table 1. *Fusarium* isolates in the *Gibberella fujikuroi* species complex used in this study.

***Fusarium* spp.**	**isolate** [a]			
	Mating population	**Mating type**	**MRC** [b] **no.**	**KSU** [c] **no.**
F. verticillioides	A	*MAT-1*	8559	A-00149
		MAT-2	8560	A-00999
F. sacchari	B	*MAT-1*	8552	B-03853
		MAT-2	8551	B-03852
F. fujikuroi	C	*MAT-1*	8532	C-01993
		MAT-2	8534	C-01995
F. proliferatum	D	*MAT-1*	8549	D-04854
		MAT-2	8550	D-04853
F. subglutinans (group 2)	E	*MAT-1*	8553/6483	E-00990
		MAT-2	8554/6512	E-02192
F. subglutinans (group 1)		*MAT-1*	1084	
		MAT-2	7828	
F. thapsinum	F	*MAT-1*	8558	F-04094
		MAT-2	8557	F-04093
F. nygamai	G	*MAT-1*	8546	G-05111
		MAT-2	8547	G-05112
F. circinatum	H	*MAT-1*	7488	H-10847
		MAT-2	6213	H-10850
F. konzum	I	*MAT-1*	8545	I-11616
		MAT-2	8544	I-11615
F. mangiferae	N/A	*MAT-1*	8092/8093	X4382
		MAT-2	7559	11781

[a] Mating populations are represented by the standard *MAT-1* and *MAT-2* “tester” strains (Kuhlman 1982; Leslie 1991,1995; Klittich and Leslie 1992; Klaasen and Nelson 1996; Klittich *et al.* 1997; Britz *et al.* 1999; Britz *et al.* 2002; Zeller *et al.* 2003; Leslie *et al.* 2005)

[b] Medical Research Council (MRC), Tygerberg, South Africa

[c] Department of Plant Pathology, Kansas State University (KSU), Manhattan, Kansas, USA

Table 2. Sequences and annealing temperatures for primer pairs used in this study

Primer	Sequence	T_m
M1C-1 MFC-L1	ATGGCATCATGACCACTACC GCAGTGATTGGACTGGCTCAT	63
M1C-1 MFC-R2	ATGGCATCATGACCACTACC AGGCCACAGACATACATATC	58
M1C-3 M1C-2	TGACCTGTTCGGTCATGAGT GAGAGTGTTGCTCTCCATAG	60
M1C-5 M1C-13	CGACTTCCTCTACCATACTC AAGGCCATGAAGGCGTTA	60
M1C-7 M1C-6	CTCAACGGCATCCGAGCTAC AACCGCAGTCGACATGAAGG	64
M1C-9 M1C-8	GCGGTACTTGCTTCGTGCTT AGCTCTCGCGACTTAGAGAC	64
M2C-7 MFC-L1	TGTAGCAGCGTTCCAGGATG GCAGTGATTGGACTGGCTCAT	64
M2C-7 MFC-R2	TGTAGCAGCGTTCCAGGATG AGGCCACAGACATACATATC	58
M2C-9 M2C-4	ATCGCTTTCACCGGCCCTCT TTGTCCGTTCTGGCCGACAA	64
M2C-10 M2C-2	CTATTGACGCAAGAGAGTGG GCTCACAGCAAGTGAAGAAC	65
M2C-12 M2C-11	GCGCACTCTACAAGCAGATG AACAGTGCGTGGCAGGTGTC	64
MFC-R1 M1C-14	AGCATCGCCACGAGTCTGTC GTCGCAAGCCGATATTCACA	65
MFC-R1 M2C-13	AGCATCGCCACGAGTCTGTC GCGCACTCTACAAGCAGATG	65
MFC-R1 MFC-L1	AGCATCGCCACGAGTCTGTC GCAGTGATTGGACTGGCTCAT	65

Table 3. Predicted transcription factor binding sites in the upstream 500bp of *MAT* genes

TRANSFAC database entry [a]			Previously described function	Number of predicted binding sites [bc]									
				F. verticillioides					*F. graminearum*				
acc.	TF	sp.		1-1	1-2	1-3	2-1	2-3	1-1	1-2	1-3	2-1	2-3
M29	HSF	S. cerevisiae	heat shock factor	11	15	12	11	11	12	12	12	9	13
M31	MATα2	S. cerevisiae	activation of mating-type-specific genes (Johnson 1995)							1	1	1	1
M46	GCR1	S. cerevisiae	coordination of cell growth (Barbara et al. 2006)		2	2						1	1
M48	ADR1	S. cerevisiae	gene activation under glucose starvation (Young et al. 2003)	6	3	4	7	5	8	6	7	6	5
M61	MIG1	S. cerevisiae	gene repression umder glucose availability (Carlson 1999)	1									
M142	NIT2	N. crassa	gene repressor under nitrogen availability (Tao and Marzluff 1999)	4	5	5	2	5	1	4	4	2	2
M154	STRE	S. cerevisiae	stress response element	1	1	1	1		2	1	1	1	
M167	HSF	S. cerevisiae	heat shock factor				1						
M213	RAP1	S. cerevisiae	repression/activation of mating type genes (Kurtz and Shore 1991)		1	1							
M263	StuAp	E. nidulans	coordination of cell cycle (Dutton et al. 1997)	2			1						
M276	Mat1-Mc	S. pombe	activation of mating-type-specific genes (Kjaerulf et al. 1997)	2	2	1	1	3	3				

[a]Data obtained form the TRANSFAC database (Wingender *et al.* 1996). 'Acc.' Indicates the accession number of the binding motif, 'TF' the transcription factor name, and 'sp.' the species in which it was described (*E. nidulans* = *Emericella nidulans*)

[b]Number of predicted binding sites in the upstream region of *F. verticillioides* and *F. graminearum* genes (e.g. 11 indicates eleven different predicted binding sites for the transcription factor). Numbers were not necessarily identical in all other species.

[c]Gene names are shortened, i.e. *1-1* represents *MAT1-1-1*; *2-3* represents *MAT1-2-3* etc.

Table 4. Results of the Approximately Unbiassed (AU) test of competing tree topologies

Trees Tested [a]	***P*** [b]		**Rejected** [c]	
	MAT1-1	***MAT1-2***	***MAT1-1***	***MAT1-2***
MAT trees vs species phylogeny				
Species tree	0.000	0.034	Yes	Yes
MAT1-1 ML tree	1.000	0.000		Yes
MAT1-2 ML tree	0.000	0.962	Yes	
Placement of *F. sacchari*				
Within Asian clade	0.009	1.000	Yes	
Outside Asian clade	0.016	0.001	Yes	Yes
Outside *G. fujikuroi* complex	0.993	0.000		Yes

[a] Alternative topologies being compared.

[b] Approximately Unbiased (AU) test statistic. Results of the Shimodaira-Hasegawa (SH) test are not shown as they echoed the AU results, but sometimes with lower significance.

[c] Likelihoods are significantly different at values of $P \leq 0.05$.

Table 5. Parameter estimates and likelihood values of the various models of codon evolution, including results of the LRT (likelihood ratio test).

Gene	Model [a]	Parameter estimates[b]	lnL[c]	LRT df[e]	LRT 2δ[d]
Gibberella fujikuroi species complex (heterothallic)					
MAT1-1-1	N	ω=1 (fixed)	-2805.04	1	61.802**
	M0	ω=0.299	-2774.14		
	M7	p=0.112 q=0.264	-2759.96	2	9.601**
	M8	p_0=0.96442 p=1.377 q= 5.044 ω=3.845	-2755.16		
MAT1-1-2	N	ω=1 (fixed)	-3726.44	1	71.514**
	M0	ω=0.284	-3690.68		
	M7	p=0.119 q=0.262	-3666.84	2	0.361
	M8	p_0=0.739 p=7.393 q=99.000 ω=1.034	-3666.66		
MAT1-1-3	N	ω=1 (fixed)	-1468.85	1	52.690**
	M0	ω=0.234	-1442.51		
	M7	p=0.066 q=0.196	-1435.65	2	0.006
	M8	p_0=0.843 p=0.081 q=0.630 ω=1.000	-1435.65		
MAT1-2-1	N	ω=1 (fixed)	-1493.2	1	71.806**
	M0	ω=0.147	-1457.29		
	M7	p=0.083 q=0.459	-1450.01	2	0.826
	M8	p_0=0.930 p=6.802 q=99.000 ω=1.423	-1449.6		
Fusarium graminearum species complex (homothallic)					
MAT1-1-1	N	ω=1 (fixed)	-1543.29	1	1.028
	M0	ω=0.29903	-1543.8		
	M7	p=0.002 q=0.008	-1539.72	2	6.957*
	M8	p_0=0.960 p=0.002 q=1.72 ω=7.888	-1536.24		
MAT1-1-2	N	ω=1 (fixed)	-2146.82	1	14.063**
	M0	ω=0.263	-2139.79		
	M7	p=35.475 q=99.000	-2139.79	2	0.000
	M8	p_0=0.999 p=35.474 q=99.000 ω=1.000	-2139.79		
MAT1-1-3	N	ω=1 (fixed)	-805.344	1	0.622
	M0	ω=0.42932	-805.655		
	M7	p=0.005 q=0.012	-804.883	2	1.282
	M8	p_0=0.887 p=0.005 q=1.400 ω=3.931	-804.242		
MAT1-2-1	N	ω=1 (fixed)	-1086.22	1	6.463*
	M0	ω=0.230	-1082.99		
	M7	p=29.649 q=99.000	-1082.99	2	0.000
	M8	p_0=0.999 p=29.642 q=99.000 ω= 1.000	-1082.99		
MAT1-2-3	N	ω=1 (fixed)	-2051.46	1	17.254**
	M0	ω=0.487	-2042.83		
	M7	p=0.535 q=0.548	-2040.62	2	0.000
	M8	p_0=0.999 p=0.535 q=0.548 ω=1.000	-2040.62		

[a] Model of codon evolution in CODEML, 'N' indicates the neutral model

[b] Parameter estimates from CODEML output

[c] ln likelihood of the model in question

[d] δ represents the difference between the likelihoods of the two models being compared

[e] degrees of freedom (the difference in the number of free parameters between the two models)

* and ** indicate significant difference at values of $P \leq 0.05$ and $P \leq 0.01$, respectively

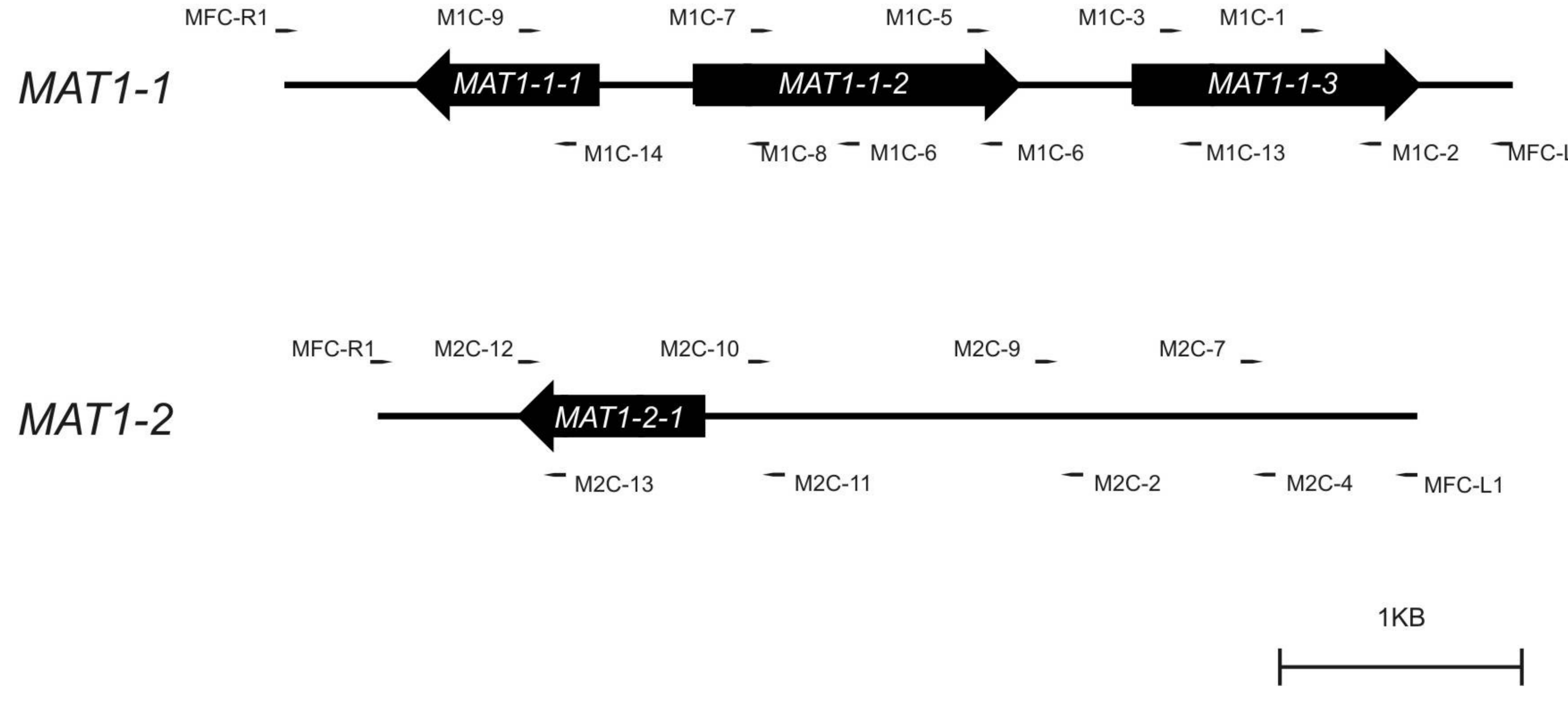


Fig 1. Binding sites of primers used to amplify *MAT1-1* and *MAT1-2* from *G. fujikuroi* species. Small arrows represent primer binding sites. Large block arrows represent *MAT* genes.

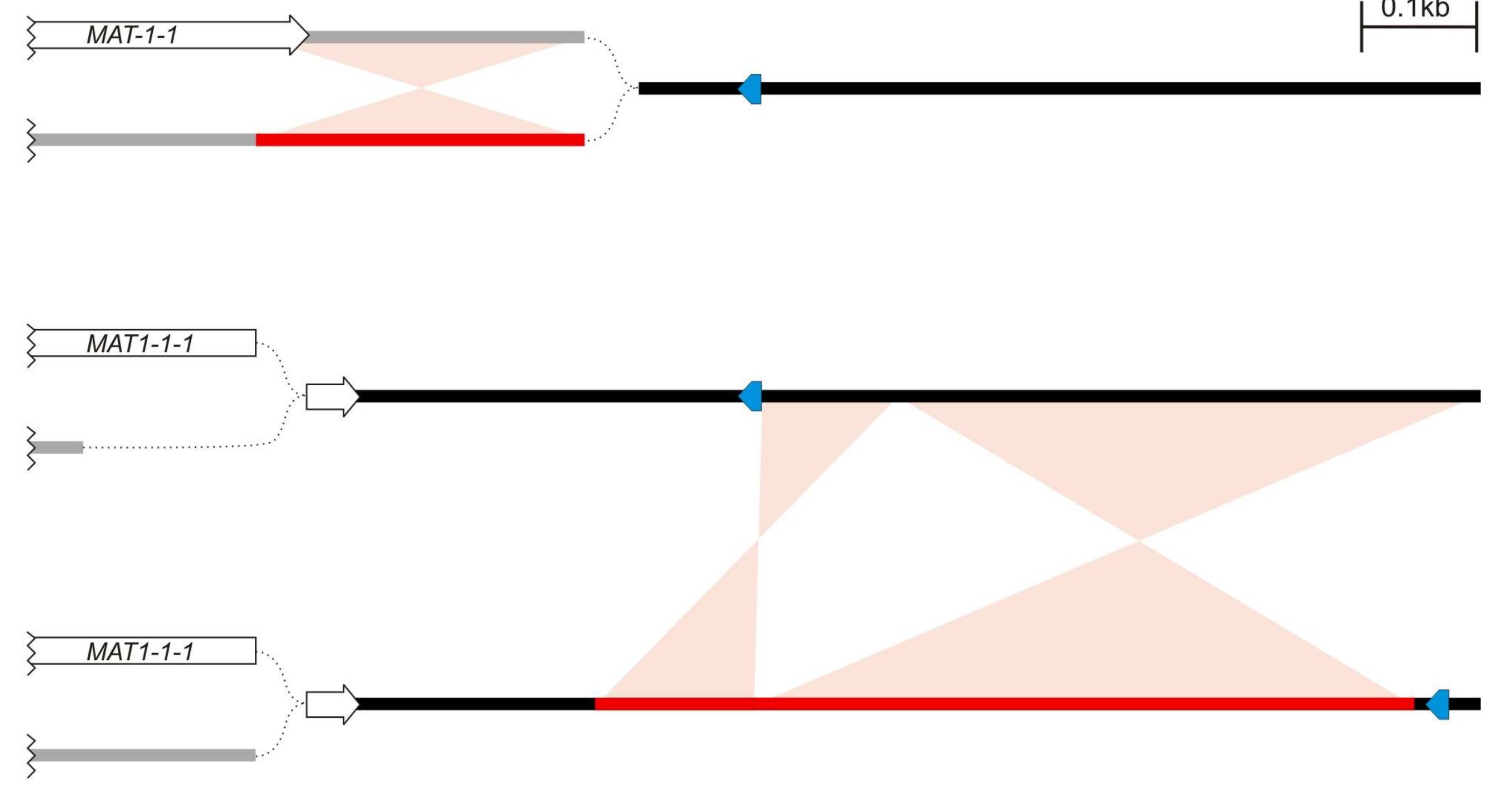
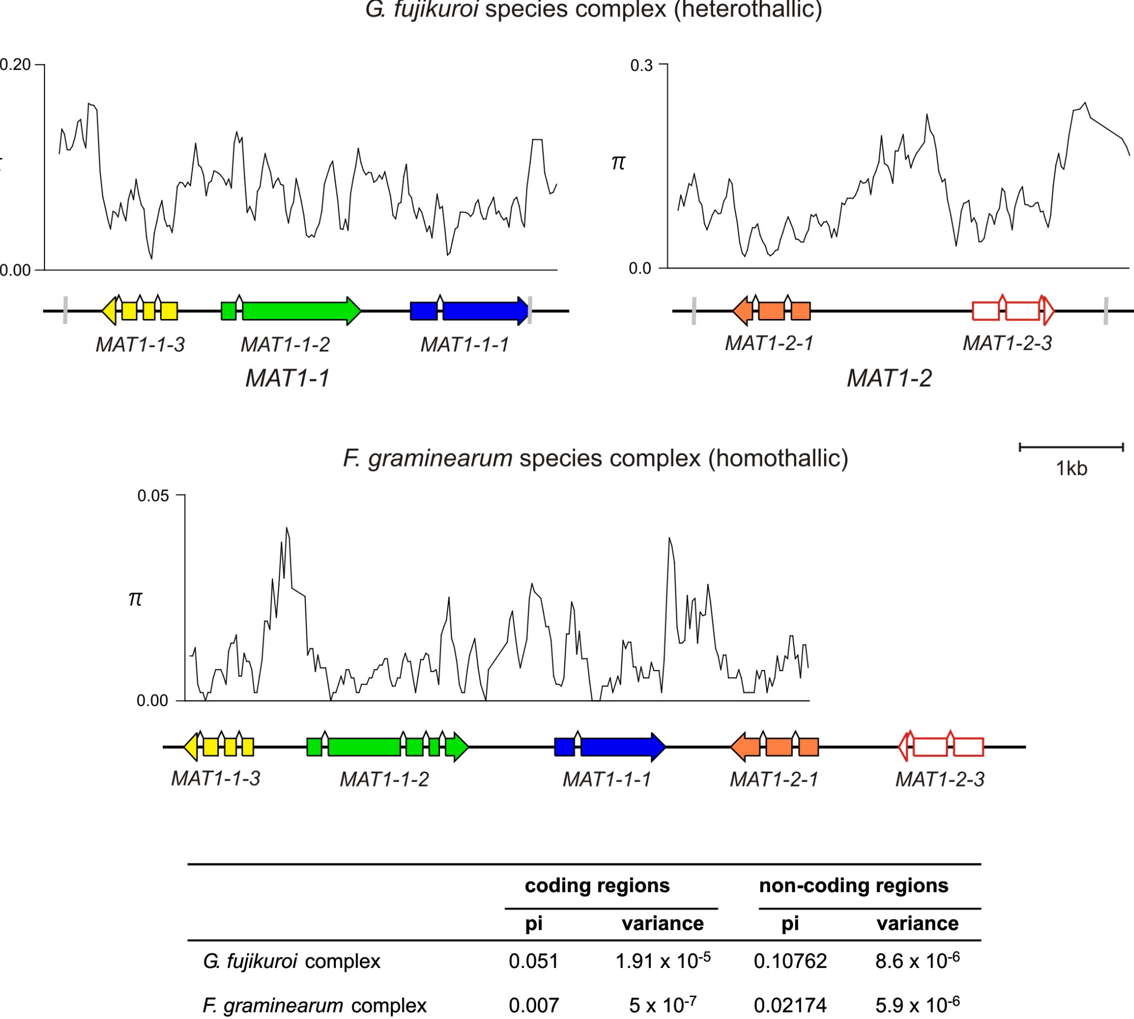


Fig. 2. Structural variation in the 3' flanking region. Major structural differences in the 3' flanking region between the African, American and Asian clades are displayed schematically. Inverted regions of homology, as interpreted from GATA analysis are shown in red. Grey lines represent idiomorphic (non-homologous regions) and black lines represent regions outside of the idiomorphs. Positions of the MFC-L1 primer binding sites are indicated in blue. The final exon of the MAT1-1-1 gene is represented by a white box arrow



	coding regions		non-coding regions	
	pi	variance	pi	variance
G. fujikuroi complex	0.051	1.91×10^{-5}	0.10762	8.6×10^{-6}
F. graminearum complex	0.007	5×10^{-7}	0.02174	5.9×10^{-6}

Fig. 3. Structures and nucleotide diversity of *Fusarium MAT* loci. Boxes represent exons with the direction of transcription indicated. Grey lines separate idiomorphs from flanking regions, although these points varied among species. Plots indicate nucleotide diversity (π), measured in a sliding window of 100bp. For species in the *G. fujikuroi* complex, π was measured using sequences generated in this study. For the *F. graminearum* species complex, sequences generated by O'Donnell *et al.* (2004) were used, hence the plot only covers the region previously thought to constitute *MAT* locus.

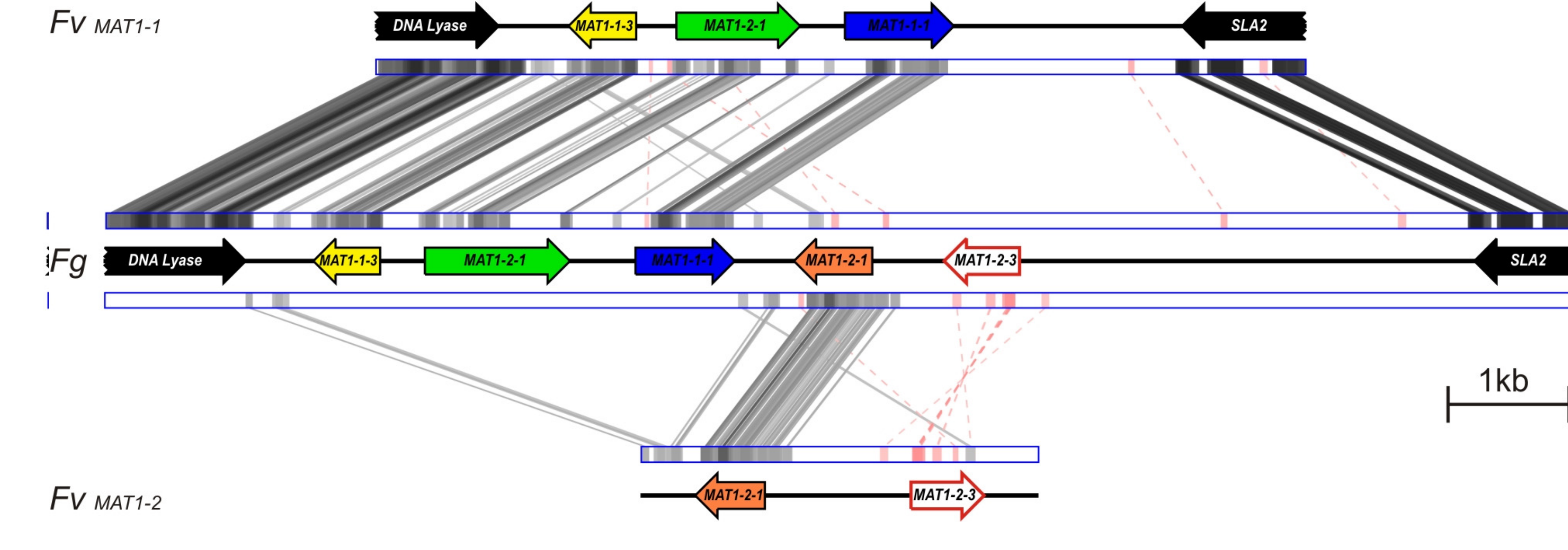


Fig. 4. Homology between the *F. graminearum* and *F. verticillioides MAT* loci. The diagram was prepared from the output of GATA. Regions of strong homology are shaded and connected by lines. The intensity of shading indicates the strength of homology. Red shading with dashed lines indicates inverted homology. Genes are represented by box arrows. The *F. graminearum* (*Fg*) and *F. verticillioides* (*Fv*) *MAT1-1* sequences were obtained from the genome sequences, while the *F. verticillioides MAT1-2* sequence was obtained in this study.

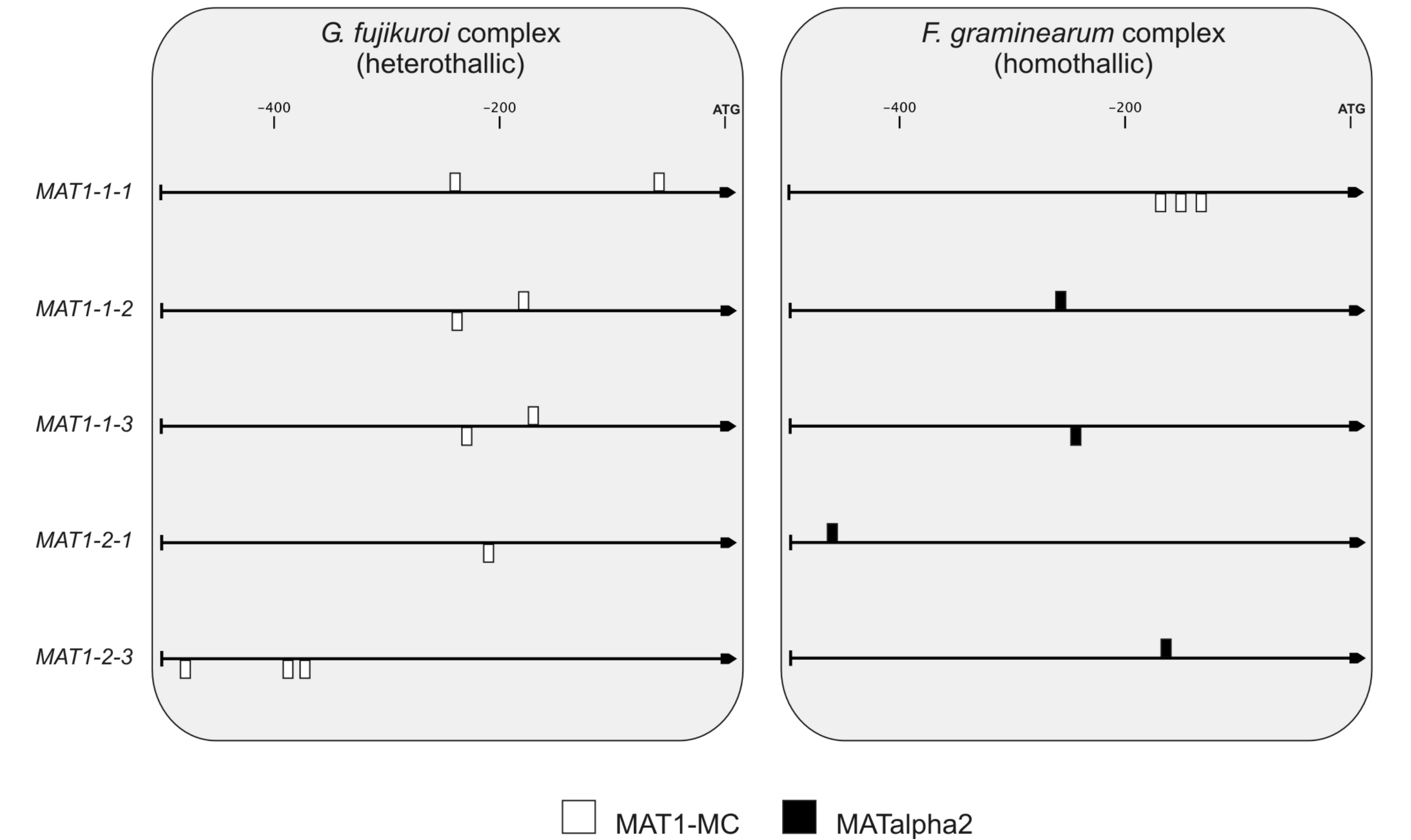


Fig. 5. Predicted binding sites of MAT1-Mc and MAT-α2. The 500bp upstream region of all five *MAT* genes is shown. Exact binding positions reflect those in *F. verticillioides* and *F. graminearum*, which were used as representatives of the two complexes. Coloured blocks indicate predicted binding sites as inferred by MOTIFSEARCH.

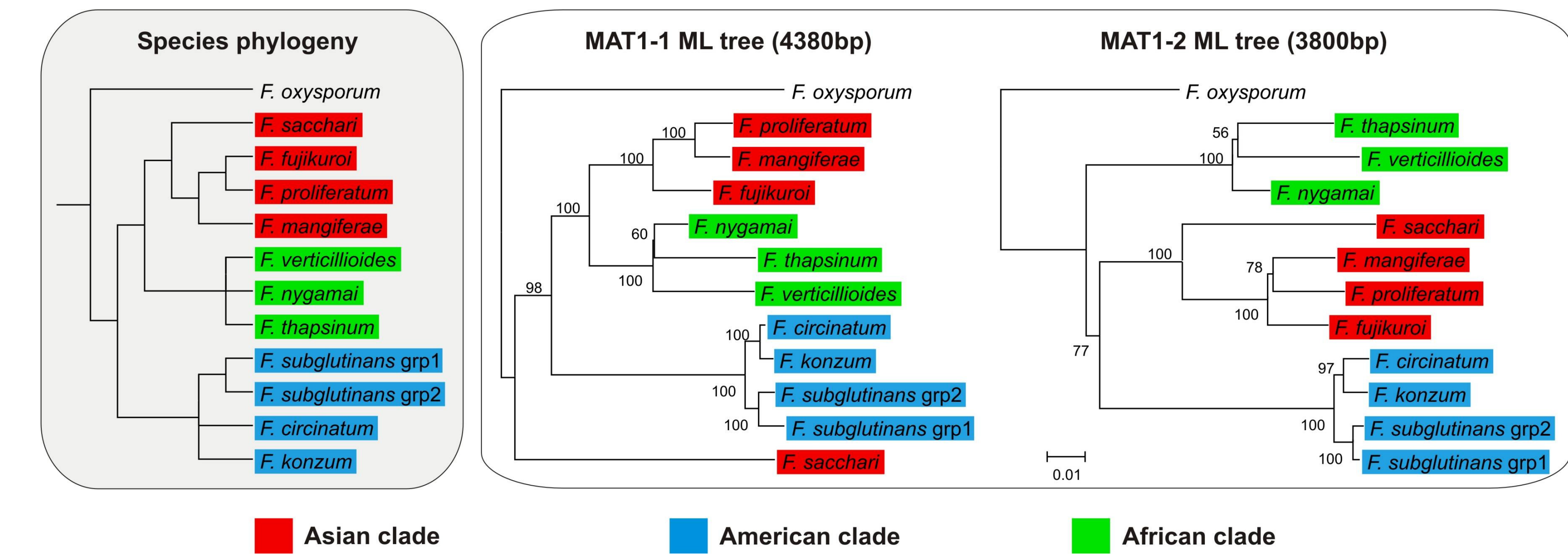


Fig. 6. Maximum likelihood phylograms for *G. fujikuroi* species complex *MAT* idiomorphs. Species trees represent consistently well supported relationships from previous studies (O'Donnell *et al.* 1998, O'Donnell *et al.* 2000, Kvas *et al.* 2009). The major clades of O'Donnell *et al.* (1998) are represented by colours. Numbers on branches indicate percentage bootstrap values from 1000 bootstraps. The *MAT1-1* and *MAT1-2* phylogenies were inferred using the general time reversible (GTR) model (Tavare 1986) and three-parameter (Kimura 1981) model with unequal base frequencies (TPM1uf; Posada 2008), respectively, both with gamma correction for among site rate variation.

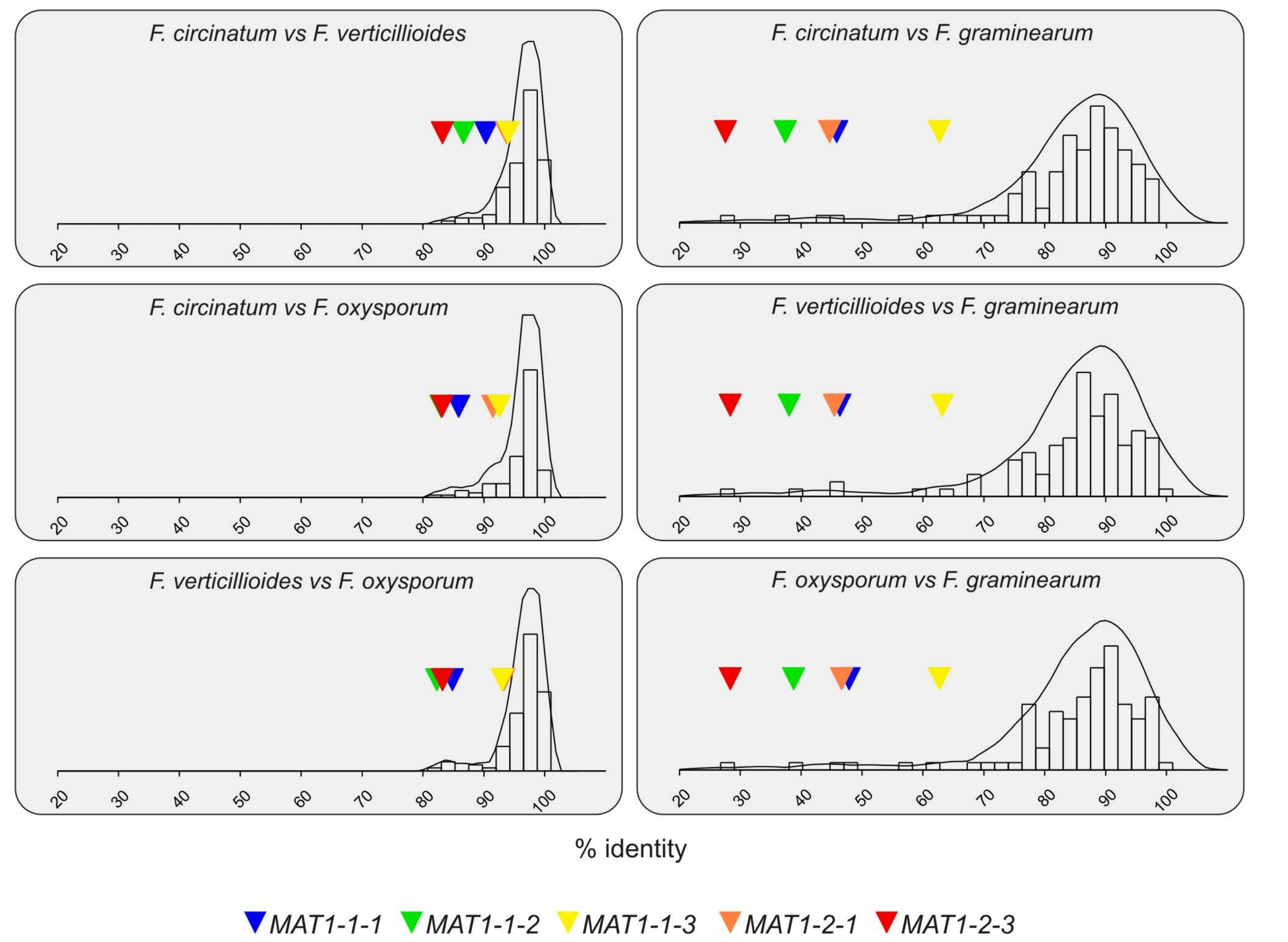


Fig. 7. Amino acid percentage-identity of *MAT* genes as compared to 100 nuclear genes from four genomes. Histograms indicate the distribution of percent identity for all 105 genes. Markers indicate the positions of the four *MAT* genes. The solid curve indicates an estimated non-parametric kernel curve (Gasser *et al.* 1985) of density.

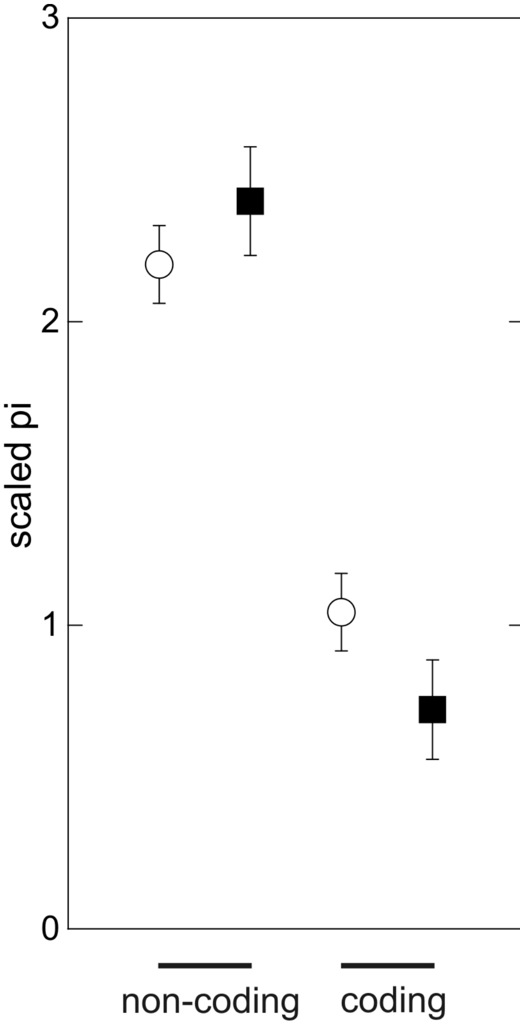


Fig. 8. Scaled nucleotide diversity (π) in coding and non-coding portions of *MAT* loci. Values of π for from *MAT* loci were scaled based on the π values of EF1-α and BT. Bars indicate the length of one standard deviation from the mean.

CHAPTER THREE

CAUSES AND CONSEQUENCES OF VARIABILITY IN PEPTIDE MATING PHEROMONES OF ASCOMYCETE FUNGI

In Press as:
Causes and consequences of variability in peptide mating pheromones of ascomycete fungi.
Simon H. Martin, Brenda D. Wingfield, Michael J. Wingfield and Emma T. Steenkamp**.**
Molecular Biology and Evolution
doi: 10.1093/molbev/msr022

It is no use for a species to have individuals of two sexes, supplied with reproductive apparatus mutually adapted for their propagation, if the individuals aren't endowed with suitable genetically determined impulses. . . . prompting them to desire one another . . .

Cyril Dean Darlington
Genetics and Man (1964)

Abstract

The reproductive genes of fungi, like those of many other organisms, are thought to diversify rapidly. This phenomenon could be associated with the formation of reproductive barriers and speciation. Heterothallic Ascomycetes, produce two classes of mating-type-specific peptide pheromones, which are essential for recognition between the mating types. Little is known regarding the diversity or the extent of species-specificity in pheromone peptides among these fungi. We compared the putative protein-coding DNA sequences of the two pheromone classes from 70 species of Ascomycetes. The dataset included previously-described pheromones and putative pheromones identified from genomic sequences. In addition, pheromone genes from twelve *Fusarium* species in the *Gibberella fujikuro*i complex were amplified and sequenced. Pheromones were largely conserved among species in this complex and, therefore, cannot alone account for the reproductive barriers observed between these species. In contrast, pheromone peptides were highly diverse among many other Ascomycetes, and occasionally under positive diversifying selection. Furthermore, repeats of the α-factor-like pheromone, which occur in tandem arrays of variable copy number, were conserved through selective constraint and not concerted evolution. This implies that sequence-specificity is important. Our findings suggest that frequent duplication and loss causes the tandem-repeats to evolve through "birth-and-death" evolution, which could contribute to their diversity among species.

Introduction

Sexual reproduction in ascomycete fungi (phylum Ascomycota) commences with the interaction between compatible cells, which is facilitated by pheromones and receptors (Kurjan 1993; Kim and Borkovich 2004, 2006). In heterothallic Ascomycetes sexual reproduction can occur only between two individuals of different "mating type". Pheromones are essential for successful mating and are produced in a mating type-specific manner in heterothallic species (Bender and Sprague-Jr 1989; Kim and Borkovich 2004; Coppin *et al.* 2005; Kim and Borkovich 2006; Mayrhofer *et al.* 2006). In homothallic species, which can reproduce sexually through self-fertilization, mating pheromones might be somewhat dispensable (Mayrhofer and Pöggeler 2005; Kim *et al.* 2008; Lee *et al.* 2008).

Ascomycete fungi produce two classes of mating pheromones. These were first comprehensively studied in the yeast *Saccharomyces cerevisiae* and characterized as mating factors a and α (or a-factor and α-factor) (Stötzler and Duntze 1976; Betz *et al.* 1977; Betz *et al.* 1987). Both pheromone classes are cleaved from larger precursor proteins (Kurjan and Herskowitz 1982; Singh *et al.* 1983; Brake *et al.* 1985). Homologous pheromones and precursor genes have been described from numerous yeasts and filamentous Ascomycetes (all references and GenBank accession numbers are provided as in Appendix D).

In *S. cerevisiae*, two genes (*Mfa1* and *Mfa2*) each encode a Prepro a-factor precursor polypeptide, from which the eleven amino acid mature a-factor pheromone is cleaved. The precursor is characterized by the presence of a carboxyl (C)-terminal prenylation signal or so-called CaaX motif, typical of fungal pheromone precursors. Prenylation is required for transport of the precursor to the plasma membrane where cleavage occurs (Brake *et al.* 1985). There are also two genes (*MFα1* and *MFα2*) that encode Prepro α-factor polypeptides, which respectively contain four and two repeats of the thirteen-residue mature α-factor (Singh *et al.* 1983). Each mature peptide repeat is bordered by signals for cleavage by Ste13p and Kex2p on its N and C-terminal borders, respectively (Julius *et al.* 1983; Julius *et al.* 1984). The amino acid sequences of the multiple mature peptide repeats are usually identical or near-identical. It is thought that all the copies of the mature α-factor are cleaved and released by the cell (Caplan and Kurjan 1991).

Observations in a wide range of organisms have demonstrated that reproductive proteins, particularly those under sex-biased expression such as gamete recognition proteins, evolve rapidly, often under positive, diversifying selection (reviewed by Swanson and Vacquier 2002; Clark *et al.* 2006). Whether ascomyete pheromones display this trend has not been ascertained, although this appears to be the case for fungal mating-type (*MAT*) loci (Turgeon 1998; Brown and Casselton 2001; Wik *et al.* 2008). The *MAT* loci encode putative transcription factors that determine mating type and control sexual development (including pheromone production) (Coppin *et al.* 1997). A number of processes have been proposed to explain the phenomenon of rapid evolution in reproductive proteins. These include reinforcement (Howard 1993), and sexual selection (Palumbi 1999; Galindo *et al.* 2003). Regardless of its source, a fundamental consequence of the rapid evolution of interacting reproductive proteins is the potential to generate reproductive isolation between populations or species (Clark *et al.* 2006; Palumbi 2008).

The importance of pheromone peptide sequence for successful mate recognition is unclear. In yeasts, pheromones appear to have some species-specificity; although this may be weak among members of the same genus (McCullough and Herskowitz 1979; Burke *et al.* 1980; Hisatomi *et al.* 1988). Naider and Becker (2004) have demonstrated that only certain residues in the α-factor are essential for receptor stimulation while others might be less important. For the pheromones of Basidiomycetes, alteration of certain residues can drastically affect the success of pheromone reception (Olesnicky *et al.* 2000; Fowler *et al.* 2001). It is, therefore, conceivable that reproductive barriers between some ascomycete species could stem from differences in their pheromone peptide sequences.

The potential for reproductive isolation as a result of species-specific mating cues, including chemical signals, is well known (Coyne and Orr 2004, p. 214-215). However, large scale phylogenetic studies have been principally focussed on visual and acoustic mating signals of animals, with no thorough, equivalent studies on chemical signals such as pheromones (Symonds and Elgar 2008). In fungi in particular, there has only recently been exploration into the potential role of the pheromone/receptor system in the development of species boundaries (e.g., Karlsson *et al.* 2008). In this study we considered the diversity of ascomycete pheromones and whether these peptides could potentially play a role in the generation of reproductive isolation and speciation. To achieve this goal, we considered (i) species-specificity in pheromone peptides, (ii) the inter-specific diversity and evolution of ascomycete pheromone peptides and (iii) the

mechanism(s) of sequence and copy-number evolution of multiple tandem repeats of the α-factor-like pheromone.

Nucleotide sequences of previously described pheromones, as well as putative pheromones identified from genomic sequences available in public domain databases were used. In addition, we determined the pheromone gene sequences for twelve *Fusarium* spp. in the *Gibberella fujikuroi* complex. Species in this group display varying levels of sexual inter-compatibility, ranging from complete reproductive isolation to complete inter-fertility (Desjardins *et al.* 2000; Leslie *et al.* 2004; Leslie and Summerall 2006; Kvas *et al.* 2009). To understand diversity in ascomycete pheromones, the rate of pheromone evolution was compared with that of house-keeping genes and tests for positive diversifying selection were performed. Furthermore, to test the hypothesis that concerted evolution could act to homogenize multiple tandem repeats of the α-factor-like pheromone, evolutionary relationships among repeats within and between species were examined. The structural organization of the α-class precursor gene was also compared among the twelve *G. fujikuroi* complex species.

Materials and Methods

Fungal isolates

A total of 24 *Fusarium* isolates were included in this study (Table 1). These isolates represent the standard *MAT-1* and *MAT-2* mating type tester strains for the nine biological species or reproductively-isolated mating populations (MP-A to MP-I) of the *G. fujikuroi* complex (reviewed by Kvas *et al.* 2009). All but one of these mating populations corresponds to a single phylogenetic species (O'Donnell *et al.* 1998; Kvas *et al.* 2009). The exception is MP-E, which consists of two distinct phylogenetic species that are inter-fertile (referred to here as *F. subglutinans* groups '1' and '2') (Steenkamp *et al.* 2002). Mating populations C and D (*F. fujikuroi* and *F. proliferatum*, respectively) have been shown to share a moderate level of inter-fertility (Leslie *et al.* 2004). A single inter-specific cross has also been observed between a pair of isolates from MP-H (*F. circinatum*) and MP-E (*F. subglutinans* group 1) (Desjardins *et al.* 2000). We also included two isolates from each of two species (*F. mangiferae* and *F. sterilihyphosum*) for which no sexual stage is known (Britz *et al.* 2002). DNA was extracted from all isolates following a protocol based on that of Steenkamp *et al.* (1999).

Sequences and organization of pheromone precursor genes

Putative pheromone precursor genes were identified by BLASTp and tBLASTn using, as queries, all currently known homologues of these genes in Ascomycetes (References and GenBank accession numbers are provided in Appendix D). In addition, a-class precursors were recognized by the presence of the characteristic C-terminal 'CaaX' motif (Brake *et al.* 1985). Likewise, α-class precursors were recognized by the presence of multiple repeats of an α-factor-like peptide bordered by the characteristic cleavage signals for homologues of Ste13p (Julius *et al.* 1983; Pöggeler 2000; Bobrowicz *et al.* 2002) and Kex2p (Julius *et al.* 1984; Darby and Smyth 1990).

To amplify the putative a- and α-class precursor genes (designated *ppg1* and *ppg2*, respectively) from species of the *G. fujikuroi* complex, Primer Designer 4.20 (Sci Ed Central, Cary, NC) was used to design primers in the conserved regions flanking the genes in *Fusarium verticillioides* and *Fusarium oxysporum* (GenBank accession numbers AAIM02000073 and AAXH01000548). To amplify *ppg2*, primers ppg2-2 (5'-TGTCTGGCAGCAACACCATC-3') and ppg2-3 (5'-CCGTCACTCAGAGCCAGGTA-3') were used (Fig.1). The *ppg1* gene was amplified in three overlapping fragments (Fig. 1) with primer set ppg1-7c (5'-ATATCACCGACGTACTGTAA-3') and ppg1-8c (5'-TACGAGTACCACTCACTT-3'), primer set ppg1-S1 (5'CTGCAACCTCGAYTAYAA-3') and ppg1-T1 (5'-ACCAGTAGCACCRTCRTC-3'), and primer set ppg1-S1 and ppg1-3 (5'-ATGGAGCGCTTGGCCTTGTG-3'). PCR reaction mixtures were 25µl in volume and contained 4ng/µl template DNA, 1.5mM $MgCl_2$, 0.4µM of each primer, 1mM deoxynucleotide triphosphates (0.25 mM of each), and 0.05U/µl Super-Therm DNA Polymerase and reaction buffer (Southern Cross biotechnology [Pty.] Ltd., Cape Town, South Africa). The PCR cycling conditions consisted of an initial denaturation at 94°C for 60s; thirty cycles of denaturation at 94°C for 30s, annealing for 30s and extension at 70°C for 90s; followed by a final extension step at 70°C for 10 min. An annealing temperature of 60°C was used for all primer combinations, except ppg1-7c and ppg1-8c, where an annealing temperature of 58°C was used.

For sequencing, PCR products were precipitated overnight in ethanol containing 0.1M Sodium Acetate (pH 3.8) at 4°C followed by centrifugation at 16,000 *rcf* for 30 min at 4°C. The pellets were washed in 70% ethanol, air dried following centrifugation (16,000 *rcf*; 10 min; 4°C) and removal of the supernatant, and resuspended in 50µl sterile distilled water. These purified products were then subjected to automated Sanger sequencing in both directions using the original PCR primers, the BigDye® terminator v3.1 cycle

sequencing kit (Applied Biosystems, Foster City, CA) and an ABI PRISM® 377 DNA sequencer (Applied Biosystems). Sequence assembly, annotation and *in silico* translation was performed using Vector NTI 9.0 (Invitrogen Life Technologies, Carlsbad, CA, USA) and CLC Bio Main Workbench (CLC Bio, Aarhus, Denmark).

PCR products that did not yield clean chromatograms when sequenced were purified by cloning in *Escherichia coli* using the pGEM®-T Easy Vector cloning system (Promega Corporation, Madison, WI). Cloned inserts were then amplified directly from colonies using the primers SP-6 (5'-ATTTAGGTGACACTATAG-3') and T-7 (5'-TAATACGACTCACTATAGGG-3'). This PCR reaction mixture was 25µl and contained 1mM deoxynucleotide triphosphates (0.25 mM of each), 0.4µM of each primer, 0.05U/µl FastStart *Taq* DNA Polymerase and FastStart reaction buffer with $MgCl_2$ (Roche Diagnostics, Mannheim, Germany). The PCR reaction conditions were as follows: denaturation at 94°C for 5 min; followed by thirty cycles of denaturation at 94°C for 30s, annealing at 50°C for 30s and extension for 90s at 72°C; followed by a final extension step at 72°C for 10 min. Sequencing was performed as described above, but using the SP-6 and T-7 primers.

Inter-specific variation and tests for positive selection

For evolutionary analyses, only the regions of the precursor genes encoding the predicted mature peptides were considered. As these peptides were divergent over large phylogenetic distances, species were only compared with other members of the same fungal Class (Fig. 2 and 3). Coding DNA sequences were aligned using the codon-based algorithm implemented in SQUINT (Goode and Rodrigo 2007). As a measure of the rate of protein sequence evolution, pair-wise d_N/d_S ratios were used, where d_N is the number of non-synonymous substitutions per non-synonymous site and d_S the number of synonymous substitutions per synonymous site. This is an effective measure of evolutionary rate as it controls for differential phylogenetic distances and differential mutation rates (Ina 1996). To calculate d_N/d_S the Nei and Gojobori (1986) method was used as implemented by the PAML package, version 4.3 (Yang 2007). All sequence pairs that were inferred by the program to be saturated for synonymous differences were excluded. For the multiple-copy peptides, all possible pairs between two species were averaged to give a single pair-wise value. Pair-wise d_N/d_S values were also determined for house-keeping genes using coding sequences for the β-tubulin and translation elongation factor 1-α genes (when available on GenBank; accession numbers are provided in Appendix A).

Statistical analyses were performed using Bill Miller's OpenStat software (http://statpages.org/miller/openstat/). To determine whether d_N/d_S values differed significantly between pheromone and house-keeping genes, the Wilcoxon Signed Ranks test (Wilcoxon 1945) was used. This test was implemented because the data were not normally distributed according to the Shapiro-Wilk test (Shapiro and Wilk 1965).

To test for positive diversifying selection acting at specific sites, two "site-models" (Yang *et al.* 2000a) were tested using the maximum-likelihood approach implemented by the CODEML program in the PAML package. Model M7 (the beta model) assumes a β distribution of ω with $0 \leq \omega \leq 1$, while model M8 (the beta&ω model) adds a class of codons with $\omega \geq 1$. The two models were compared with respect to their fit to the data using a likelihood ratio test (LRT), as described by Yang *et al.* (2000a). The Bayes empirical Bayes (BEB) output of CODEML was used to identify specific codons likely to fall into codon class $\omega \geq 1$ in model M8. For the multiple-repeat α-factor pheromone, the first repeat in each module was used to represent each species. This was done to avoid the risk that purifying selection acting at the within-species level could mask diversifying selection at the between-species level.

Relationships among tandemly-repeated α-factor-like pheromone sequences

To test the hypothesis that concerted evolution acts to homogenize DNA sequences among tandemly repeated copies of the α-factor-like pheromone, the relationships among DNA sequences encoding the mature pheromone peptides were examined. All repeats from all species in the Saccharomycetes and Sordariomycetes datasets (see Appendix D for sequences) were considered, as these two datasets contained a large number of species including both close and distant relatives. DNA sequences were aligned using CLUSTAL W (Thompson *et al.* 1994). The proportion of nucleotide sites at which each pair of sequences differed (*i.e.*, p-distances) (Nei and Kumar 2000) was used to construct neighbor-joining trees using MEGA version 4 (Tamura *et al.* 2007). This simple distance-based model was selected because the mode of evolution of repeats was unclear, and repeats experiencing concerted evolution would not necessarily follow any particular nucleotide-substitution model. Hence, the focus was not on evolutionary relationships *per se* but simply on whether the DNA sequence of each repeat shared more similarity with adjacent tandem repeats or with a corresponding repeat from a related species.

Another, more sensitive test proposed by Swanson and Vacquier (1998) to distinguish between concerted evolution and purifying selection was applied by comparing synonymous distance (d_S) with non-synonymous distance (d_N). Purifying selection should theoretically act only at the protein level, consequently maintaining a low d_N while d_S is allowed to increase. In contrast, concerted evolution acts at the DNA level, causing both d_N and d_S to be low among tandem repeats. These values were calculated using the Yang and Nielsen method, as implemented in the program yn00 in the PAML package. Pair-wise d_S and d_N values were generated for every possible peptide pair and then averaged such that each species had a single within-species value and each species pair had a single between-species value. Pairs saturated for synonymous or non-synonymous substitutions were discarded. The covariance and distribution of d_S and d_N values from within-species and between-species pairs were compared statistically.

To investigate the structural events that give rise to gain or loss of repeats, the *ppg1* gene sequence from an isolate of each of the twelve *Fusarium* species in the *G. fujikuroi* complex were used. The programme GATA (graphic alignment tool for comparative sequence analysis) (Nix and Eisen 2005) was used to visualise the DNA sequence relationships between each pair of species. Given two input sequences, this application uses the NCBI BLASTn (Altschul *et al.* 1990) and Bl2seq (Tatusova and Madden 1999) algorithms to generate all possible local alignments between two sequences. In this way the application is able to identify duplications, deletions and inversions. A word size of 20 was used, along with a lower cut-off bit-score of 24.9 to ensure that only regions of high homology were considered.

Results

Sequences and organization of pheromone precursor genes

For the a-class precursors, published sequences were available for nine species (three yeast and six filamentous Ascomycetes) (Fig. 2). For the α-class precursors, published sequences were available for seventeen species (nine yeast and eight filamentous Ascomycetes) (Fig. 3) (see Appendix D for references and GenBank accession numbers). Putative a-class and α-class precursor genes in the genomes of another 20 and 42 species respectively, including *F. verticillioides* were identified. In addition, the putative a-class and α-class precursor genes (*ppg2* and *ppg1*, respectively) from the 24 *Fusarium* isolates, representing 12 species of the *G. fujikuroi* complex, including *F. verticillioides* (Table 1) were sequenced (Sequences are provided in Appendix D). Sequences were obtained for

both genes from all twelve species, although one isolate of *F. sterilihyphosum* (MRC2802), failed to yield a fragment corresponding to *ppg1*. Thus a-class precursor sequences were obtained from 40 species, covering three major classes of Ascomycetes: Saccharomycetes, Sordariomycetes and Taphrinomycetes (Hibbett *et al.* 2007) (Fig. 2). For the α-class precursor, sequences were obtained from 69 species, covering six major classes: Saccharomycetes, Sordariomycetes, Taphrinomycetes, Leotiomycetes, Dothidiomycetes and Eurotiomycetes (Hibbett *et al.* 2007) (Fig. 3). Most of the yeasts carried two or more copies of one or both precursor genes, while among the filamentous species, only *Cryphonectria parasitica* carried two copies of the a-class gene (*Mf2/1* and *Mf2/2*), as previously reported by Zhang *et al.* (1998) (Fig. 2). There was only one species, *Ashbya gosypii*, for which the a-class but no α-class precursor was obtained. In contrast, there were 31 species for which only the α-class precursor was obtained.

The putative a-class precursors were all small, ranging from 21 to 66 amino acids in length (Fig. 2). All carried a C-terminal CaaX motif (containing a Cystein [C] residue, followed by two aliphatic (a) residues and terminating in an arbitrary [X] residue), although in some species, one or more additional putative CaaX motifs were present in a repeated nature, and were usually preceded by a conserved peptide sequence. This was true for the a-class gene *ppg2* for all twelve *Fusarium* species, where *F. nygamai* carried two repeats while the other eleven species each carried three (Fig. 2). Among the other Sordariomycetes, similar repeats, each carrying a CaaX motif, were found in the a-class precursors of *F. oxysporum* and both *Verticillium* species studied. Multiple CaaX motifs were also identified in two of the three *Trichoderma* species, although these were not preceded by a conserved peptide repeat. The remaining Sordariomycetes, including *G. zeae* (*Fusarium graminearum*) and *Trichoderma atroviridae* had the typical arrangement with a single, C-terminal CaaX motif (Fig. 2).

Based on sequence similarity to described a-factor-like peptides (Brake *et al.* 1985; Davey 1992), it was possible to identify predicted mature a-factor-like peptides from the remaining Saccharomycetes species (Fig. 2). Similar inferences could not be made for the 25 sordariomycete species examined because the exact sequence of the mature a-factor-like peptide has not been described from any filamentous Ascomycetes. Further analysis of a-factor peptide evolution was, therefore, based only on the yeast a-factors.

In the *ppg2* gene of the twelve *Fusarium* species representing the *G. fujikuroi* complex, there was a conserved eight-residue peptide located immediately upstream of each CaaX

motif (Fig. 2). This could represent the mature a-factor-like pheromone, implying that these species encode multiple repeats of the mature peptide. There was complete identity in this proposed pheromone sequence across all twelve *Fusarium* species. Outside of the eight residue motifs, most of the variation was found in the aliphatic residues of the putative CaaX motifs, which presumably do not form part of the mature peptides (Brake *et al.* 1985). Furthermore, there was only one example of a species-specific polymorphism, at residue 26 in both of the *F. konzum* isolates (Fig. 2).

Among all the fungal sequences examined, the putative α-class precursor proteins were larger, ranging from 102 to 711 amino acids in length. All α-class precursors carried copies of a motif bordered by the Ste13p and Kex2p protease-like cleavage signals: an N-terminal run of X-Alanine or X-Proline dipeptides (Julius *et al.* 1983; Pöggeler 2000; Bobrowicz *et al.* 2002) and a C-terminal KR (Lysine-Arginine) or RR dipeptide (Julius *et al.* 1984; Darby and Smyth 1990). The motifs occurred in a tandem arrangement of between one (e.g. *Saccharomyces castellii*) and sixteen (e.g. *F. subglutinans* group 1) repeats, with an average of five (Fig. 3).

The predicted sequences of the mature α-factor-like peptides were inferred on the basis of sequence similarity to previously described homologues, and the presence of the Ste13p and Kex2p protease-like cleavage signals. Some predicted mature peptides, such as the last peptide repeat in all of the *Fusarium* spp., had Kex2p protease-like cleavage sites at both ends of the peptide. In several species, one or more copies of the predicted mature peptide were not bordered by one or either of the signature cleavage motifs. For example, in 20% of the putative peptides among the 12 *Fusarium* species examined, the Kex2p protease-like cleavage site (usually KR or RR) was replaced by TR. However, these were still regarded as pheromones for evolutionary analyses, unless they were highly dissimilar from other repeats in the precursor. In 29 species, there was complete identity in amino acid sequence among all encoded repeats of the mature peptide, while the remaining species each encoded between two and six similar yet distinct "variants" of the pheromone (Fig. 3).

The number of repeats in the precursor gene varied considerably among the species and sometimes between the two isolates of the same species; from a minimum of nine to a maximum of sixteen (Fig. 3). Every one of the isolates encoded between four and eleven copies (average = 8) of a peptide with the sequence WCTWRGQPCW. Each isolate also encoded between two and five copies (average = 4) of a second peptide, in which the

Threonine residue was substituted with Methionine. A single repeat of two additional variants of the peptide (one in which the Threonine was replaced by Leucine, and one in which the third Tryptophan residue was replaced with Cysteine) were each found in both isolates of *F. konzum*. A single copy of a fifth variant of the peptide in which the second Tryptophan residue is replaced with an Arginine residue, occurred in one of the two *F. circinatum* isolates (Fig. 3).

Inter-specific variation and tests for positive selection

Four datasets (a-factor-like pheromones of the Saccharomycetes and α-factor-like pheromones of the Saccharomycetes, Sordariomycetes and Eurotiomycetes) included a sufficient number of species pairs for sound statistical analyses. For most species pairs, pheromones displayed a far greater evolutionary rate than house-keeping genes, although a few species had pheromones that were strongly conserved, with even slower evolutionary divergence than the house-keeping genes. However, average pairwise d_N/d_S values were found to be significantly greater for pheromones than house-keeping genes in all four datasets ($P \leq 0.05$) (Fig. 4).

Results of the CODEML analyses and LRTs are displayed in Table 2. In three of the four datasets, there was no significant difference in the likelihood values of each pair of models. Hence, the α-factor-like pheromone of Sordariomycetes and both the a- and α-factor-like pheromones of Saccharomycetes appear not to have diversified under positive selection. In contrast, the α-factor-like pheromone of the Eurotiomycetes showed evidence of having diversified under positive selection. The CODEML analyses further indicated that this selection is likely to only have affected the codon encoding the third residue of this nine-residue peptide (see Fig. 3).

Relationships among tandemly-repeated α-factor-like pheromone sequences

Distance analysis of multiple tandem repeats of the α-factor-like pheromone yielded two conflicting patterns (Fig. 5). In some species, the multiple tandem repeats were more closely related in DNA sequence to one another than to repeats from other related species. In contrast, some closely related species, including the members of the *G. fujikuroi* complex and *Saccharomyces sensu stricto* clade, displayed a different pattern in which repeats from several different species grouped together. In these instances there was some positional bias in the clustering, where clusters usually consisted of repeats from a similar region in the precursor protein. For example, repeat 5 of *F. sacchari* clustered with repeat 5 of *F. mangiferae* while repeats 13 and 14 of the same two species

clustered together. However, this pattern only appeared in species that shared close relatives within the dataset.

Broken down into synonymous and non-synonymous distances, within-species comparisons among peptide repeats had, on average, a significantly lower d_N than the between-species comparisons. However d_S values from within- and between-species comparisons were similar. Some within-species d_S values were greater than many between species values (Fig. 6).

Results of the GATA analysis were compiled (Fig. 7) so that all changes in repeat copy number in the *G. fujikuroi* complex could be considered simultaneously. The results demonstrated that multiple independent changes in repeat copy number have occurred within and among these species. In particular, it appeared that the more central repeats have been more frequently involved in these structural events.

Discussion

This study has substantially increased the number of a- and α-class pheromone precursor genes that have been described, providing new insights into the distribution, sequence diversity and structural organization of these genes. Some acomycete pheromones are strongly conserved between species but most are highly divergent. The rich diversity in peptide sequence has probably resulted from both adaptive and non-adaptive evolutionary forces. Both pheromone precursor classes can have a modular nature, a characteristic that could contribute to the inter-specific divergence of the pheromone peptides. Much of this variation might also have functional relevance, implying a potential role of pheromones in species-specific mate recognition. The rapid evolution of pheromones could, therefore, contribute to speciation in Ascomycetes.

The α-class precursor was distributed across species from all four filamentous classes (Sordariomycetes, Eurotiomycetes, Dothidiomycetes and Leotiomycetes) as well as both yeast classes (Saccharomycetes and Taphrinomycetes) included in this study. By contrast, the a-class precursor was found in far fewer species, spanning the two yeast classes and a single filamentous class, Sordariomycetes. Several previous studies have been unable to identify the a-class precursor in various genomes (Pöggeler 2002; Dyer *et al.* 2003; Hoff *et al.* 2008; Butler *et al.* 2009). This small gene might have been altered to a point where it is undetectable by BLAST analysis in these fungi. If it is indeed absent, this would

require that some other element has assumed the role of the a-factor-like pheromone, at least in heterothallic species. In some homothallic species, only one pheromone-receptor pair might be necessary, while the other could be lost (Kim *et al.* 2008; Lee *et al.* 2008).

All the pheromone precursor genes that were found in this study appeared to have intact open reading frames. These include those in species for which no sexual stage is known such as *F. mangiferae*, *F. sterilihyphosum* and *F. oxysporum*. While this could be interpreted as evidence that these species might have a cryptic sexual cycle, these genes could be retained in asexual species if they perform additional functions outside of sexual reproduction. For example, work on *N. crassa* pheromone precursor *Mfa-1* has suggested a possible role in "conglutination", the cementing of hyphae during perithecium formation (Kim *et al.* 2002). In *Candida* spp., the α-factor-like pheromone appears to be necessary for inter-cellular signalling during biofilm formation (Daniels *et al.* 2006; Sahni *et al.* 2010). Further work is, therefore, necessary to distinguish between the reproductive and non-reproductive functions of these pheromones.

This study is the first to scrutinize similarities and differences in the pheromones of a group of closely related filamentous Ascomycetes with varying degrees of sexual compatibility. Most species pairs in the *G. fujikuroi* complex are inter-sterile. However, we have found a lack of significant differences among the putative pheromone peptides in species of this complex. We hypothesise that the repeated eight-residue motif encoded by *ppg2* represents the mature a-factor like pheromone (Schrnoll *et al.* 2010). If so, there is complete identity in this pheromone among all twelve species. In *ppg1* each species carried multiple copies of two or more distinct variants of the pheromone that differed slightly in amino acid sequence. Nevertheless, only one species (*F. konzum*) carried variants that were unique and potentially species-specific. Unless these pheromones undergo differential post-translational modifications (an avenue that has not been explored in ascomycete pheromones), our results imply that these species should recognize one-another as potential mates. A mechanism further downstream must, therefore, be responsible for the observed reproductive barriers. Indeed, inter-specific mating experiments in *Fusarium* can lead to the formation of barren perithecia (Leslie *et al.* 2004), suggesting that hyphal fusion could have occurred. Such post-mating barriers are common in the genus *Neurospora*, which led Turner *et al.* (2010) to propose that pheromone divergence could be prevented by strong selective constraint.

On the whole, strong conservation of pheromones was not the norm; in fact pheromones were significantly more divergent than house-keeping genes in all four datasets considered. These findings are in agreement with the rapid evolution of reproductive proteins observed in the plant and animal kingdoms (as reviewed by Swanson and Vacquier 2002; Clark *et al.* 2006). This phenomenon has been proposed as a driving force in speciation, particularly when associated with mate-recognition proteins. A comparable example is the rapid evolution of interacting sperm and egg surface proteins in Abalone (*Haliotis rufescens*), which could cause reproductive isolation between populations (Yang *et al.* 2000b; Galindo *et al.* 2003). If pheromone peptide sequence is indeed crucial for successful reception, the observed inter-specific differences and rapid evolution of pheromones in this study might have major implications for reproductive isolation and speciation in Ascomycetes. An understanding of the forces that drive pheromone diversification could shed light on the importance of these modifications for mate recognition. Alterations that come about through relaxed selective constraint are probably inconsequential, while those driven by positive selection are likely to have functional relevance.

Our analyses identified evidence for positive diversifying selection acting in the dataset of α-factor-like pheromones from Eurotiomycetes, but not in the larger datasets of the Sordariomycetes and Saccharomycetes. This implies that the variation among species in the latter groups might have accumulated over time through relaxed constraint. However, the test performed to detect positive selection in this study addresses each dataset as a whole and is most sensitive when positive selection acts in all lineages. Yang and Nielsen (2002) demonstrated that selection acting only along certain lineages of the phylogeny might go undetected by the “site models”. It is plausible that pheromone peptides follow this sort of evolutionary pattern given the observation that the pheromones of some species pairs are highly divergent, while those of a few remain even more conserved than house-keeping genes. Another factor that could have masked the signal of selection is the inclusion of homothallic species. Recent work has shown that some reproductive genes are obsolete in homothallic species and hence evolve via neutral evolution (Wik *et al.* 2008). For example, in *G. zeae* one pheromone appears to have become redundant (Kim *et al.* 2008; Lee *et al.* 2008). The observed pheromone diversity might therefore result from both positive selection and relaxed constraint, acting in different lineages.

In the α-class precursors described here and also in those previously described, amino-acid sequences of the multiple tandem repeats of the mature peptide were usually

identical or nearly-identical. We found that this similarity was largely reflected at the DNA level, in that tandem repeats were often more closely related to one another in DNA sequence than to repeats from any other species. Such a pattern has been interpreted as the signature of concerted evolution, (e.g., Swanson and Vacquier 1998). However, the repeats of some species did not form independent clusters, and instead clustered with corresponding repeats from closely related species. For example, all repeats in each *Fusarium* species had their closest relative in another *Fusarium* species. This pattern has been termed trans-specific-polymorphism (Klein 1987), and occurs when duplication predates speciation. Thus, paralogues of a certain sequence are more similar to their corresponding copy in a related species than to the sister paralog in the same genome. Accordingly, concerted evolution in the conventional sense, which is thought to rapidly homogenise DNA sequences among tandem repeats (Elder and Turner 1995), appears not to occur in the α-class precursor.

The trans-specific polymorphism was absent in all the species that did not have close relatives in our dataset, indicating that it degrades over larger evolutionary distances. Rather than concerted evolution, this pattern is probably an inevitable result of continuous duplication and deletion of repeats, as evidenced by the huge variation in repeat copy number. Evolution through duplication and deletion has been termed "birth-and-death" evolution and is common in tandem repeat sequences (Nei and Rooney 2005). Characteristic examples of genes experiencing birth-and-death are the major histocompatibility complex (MHC) and imunoglobin (Ig) genes of mammals (Nei *et al.* 1997) as well as the ribosomal RNA genes in fungi (Rooney and Ward 2005). Concerted evolution has been rejected in these cases because, while some members of the multigene families are highly similar, others have their closest relatives in distantly related species. Fig. 8 illustrates how continuous duplications and deletions acting randomly on a set of repeats could create such a phylogenetic scenario, eventually creating exclusive monophyletic clusters.

Birth-and death is not the only mechanism acting to maintain similarity among peptide repeats. We have shown that while repeats from the same species always have low non-synonymous distances, some have high synonymous distances. This indicates that not all repeats are similar only due to recent duplication, but that the sequences of the encoded peptides are also conserved by purifying selection (Ina 1996). We, therefore, propose that the similarity observed among repeats of the mature peptide is adaptive, and reflects a

requirement for sequence specificity in pheromone reception. Differences between species could therefore indeed reflect functional divergence.

The unique structure and evolutionary system of the α-class precursor could further contribute to the rapid evolution of the α-factor-like pheromones. Random events of duplication and loss could cause the "fixation" of a certain pheromone variant in the precursor gene of one species and the simultaneous fixation of a different variant in another species. This could be facilitated by positive selection, but might even occur by chance in the absence of selection (Nei and Rooney 2005). The gradual process of succession by which a new repeat variant can spread through the precursor, replacing the old one, might also facilitate the co-evolution of pheromones and receptors. A new compatible pheromone/receptor pair could spread in a population without a dramatic fitness cost, as individuals carrying both pheromone variants maintain the ability to be recognized by the "wild-type" receptor. To illustrate this point, species B in Fig. 8 would remain compatible with individuals carrying a receptor for either the new or the wild-type pheromone up until step 4, after which the new variant completely succeeds the old. Lastly, the ability to harbor multiple repeat variants without a significant fitness cost could increase the ability to evolve by providing a constant source of variation, the raw material for natural selection.

References

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

Bender, A. and Sprague-Jr, G. F. 1989. Pheromones and pheromone receptors are the primary determinants of mating specificity in the yeast *Saccharomyces cerevisiae*. *Genetics* **121**: 463-476.

Betz, R., Crabb, J. W., Meyer, H. E., Wittig, R. and Duntze, W. 1987. Amino acid sequences of a-factor mating peptides from *Saccharomyces cerevisiae*. *Journal of Biological Chemistry* **262**: 546-548.

Betz, R., MacKay, V. L. and Duntze, W. 1977. a-Factor from *Saccharomyces cerevisiae*: partial characterization of a mating hormone produced by cells of mating type a. *Journal of Bacteriology* **132**: 462-472.

Bobrowicz, P., Pawlak, R., Correa, A., Bell-Pedersen, D. and Ebbole, D. J. 2002. The *Neurospora crassa* pheromone precursor genes are regulated by the mating type locus and the circadian clock. *Molecular Microbiology* **45**: 795-804.

Brake, A. J., Brenner, C., Najarian, R., Laybourn, P. and Merrywheather, J. 1985. Structure of genes encoding precursors of the yeast peptide mating pheromone a-factor. In: *Protein Transport and Secretion*, (Gething, M. J., ed.). pp. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.

Britz, H., Steenkamp, E. T., Coutinho, T. A., Wingfield, B. D., Marasas, W. F. O. and Wingfield, M. J. 2002. Two new species of *Fusarium* section *Liseola* associated with mango malformation. *Mycologia* **94**: 722-730.

Brown, A. J. and Casselton, L. A. 2001. Mating in mushrooms: increasing the chances but prolonging the affair. *Trends in Genetics* **17**: 393-400.

Burke, D., Mendonça-Previato, L. and Ballou, C. E. 1980. Cell-cell recognition in yeast: purification of *Hansenula wingei* 21-cell sexual agglutination factor and comparison of the factors from three genera. *Proceedings of the National Academy of Sciences of the United States of America* **77**: 318-322.

Butler, G., Rasmussen, M. D., Lin, M. F., Santos, M. A. S., Sakthikumar, S., Munro, C. A., Rheinbay, E., Grabherr, M., Forche, A., Reedy, J. L., Agrafioti, I., Arnaud, M. B., Bates, S., Brown, A. J. P., Brunke, S., Costanzo, M. C., Fitzpatrick, D. A., de Groot, P. W. J., Harris, D., Hoyer, L. L., Hube, B., Klis, F. M., Kodira, C., Lennard, N., Logue, M. E., Martin, R., Neiman, A. M., Nikolaou, E., Quail, M. A., Quinn, J., Santos, M. C., Schmitzberger, F. F., Sherlock, G., Shah, P., Silverstein, K. A. T., Skrzypek, M. S., Soll, D., Staggs, R., Stansfield, I., Stumpf, M. P. H., Sudbery, P. E., Srikantha, T., Zeng, Q., Berman, J., Berriman, M., Heitman, J., Gow, N. A. R., Lorenz, M. C., Birren, B. W., Kellis, M. and Cuomo, C. A. 2009. Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature* **459**: 657-662.

Caplan, S. and Kurjan, J. 1991. Role of α-factor and the MFα1 α-factor precursor in mating in yeast. *Genetics* **127**: 299-307.

Clark, N. L., Aagaard, J. E. and Swanson, W. J. 2006. Evolution of reproductive proteins from animals and plants. *Reproduction* **131**: 11-22.

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

Coppin, E., de Renty, C. and Debuchy, R. 2005. The function of the coding sequences for the putative pheromone precursors in *Podospora anserina* is restricted to fertilization. *Eukaryotic Cell* **4**: 407-420.

Coyne, J. A. and Orr, H. A. 2004. *Speciation*. Sinauer, Sunderland, Massachusetts.

Daniels, K. J., Srikantha, T., Lockhart, S. R., Pujol, C. and Soll, D. R. 2006. Opaque cells signal white cells to form biofilms in *Candida albicans*. *EMBO Journal* **25**: 2240-2252.

Darby, N. J. and Smyth, D. G. 1990. Endopeptidases and prohormone processing. *Bioscience Reports* **10**: 1-13.

Davey, J. 1992. Mating pheromones of the fission yeast *Schizosaccharomyces pombe*: purification and structural characterization of M-factor and isolation and analysis of two genes encoding the pheromone. *EMBO Journal* **11**: 951-960.

Desjardins, A. E., Plattner, R. D. and Gordon, T. R. 2000. *Gibberella fujikuroi* mating population A and *Fusarium subglutinans* from teosinte species and maize from Mexico and Central America. *Mycological Research* **104**: 865-872.

Dyer, P. S., Paoletti, M. and Archer, D. B. 2003. Genomics reveals sexual secrets of *Aspergillus*. *Microbiology* **149**: 2301-2303.

Elder, J. F., Jr. and Turner, B. J. 1995. Concerted evolution of repetitive DNA sequences in eukaryotes. *The Quarterly Review of Biology* **70**: 297-320.

Fowler, T. J., Mitton, M. F., Vaillancourt, L. J. and Raper, C. A. 2001. Changes in mate recognition through alterations of pheromones and receptors in the multisexual mushroom fungus *Schizophyllum commune*. *Genetics* **158**: 1491-1503.

Galindo, B. E., Vacquier, V. D. and Swanson, W. J. 2003. Positive selection in the egg receptor for abalone sperm lysin. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 4639-4643.

Goode, M. G. and Rodrigo, A. G. 2007. SQUINT: a multiple alignment program and editor. *Bioinformatics* **23**: 1553-1555.

Hall, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symposium Series* **41**: 95 - 98.

Hibbett, D. S., Binder, M., Bischoff, J. F., Blackwell, M., Cannon, P. F., Eriksson, O. E., Huhndorf, S., James, T., Kirk, P. M., Lücking, R., Thorsten Lumbsch, H., Lutzoni, F., Matheny, P. B., McLaughlin, D. J., Powell, M. J., Redhead, S., Schoch, C. L., Spatafora, J. W., Stalpers, J. A., Vilgalys, R., Aime, M. C., Aptroot, A., Bauer, R., Begerow, D., Benny, G. L., Castlebury, L. A., Crous, P. W., Dai, Y.-C., Gams, W., Geiser, D. M., Griffith, G. W., Gueidan, C., Hawksworth, D. L., Hestmark, G., Hosaka, K., Humber, R. A., Hyde, K. D., Ironside, J. E., Kõljalg, U., Kurtzman, C. P., Larsson, K.-H., Lichtwardt, R., Longcore, J., Miadlikowska, J., Miller, A., Moncalvo, J.-M., Mozley-Standridge, S., Oberwinkler, F., Parmasto, E., Reeb, V., Rogers, J. D., Roux, C., Ryvarden, L., Sampaio, J. P., Schüßler, A., Sugiyama, J., Thorn, R. G., Tibell, L., Untereiner, W.

A., Walker, C., Wang, Z., Weir, A., Weiss, M., White, M. M., Winka, K., Yao, Y.-J. and Zhang, N. 2007. A higher-level phylogenetic classification of the Fungi. *Mycological Research* **111**: 509-547.

Hisatomi, T., Yanagishima, N., Sakurai, A. and Kobayashi, H. 1988. Interspecific actions of α mating pheromones on the a mating-type cells of three *Saccharomyces* yeasts. *Current Genetics* **13**: 25-27.

Hoff, B., Pöggeler, S. and Kuck, U. 2008. Eighty years after its discovery, Fleming's *Penicillium* strain discloses the secret of its sex. *Eukaryotic Cell* **7**: 465-470.

Howard, D. J. 1993. Reinforcement: origin, dynamics and fate of an evolutionary hypothesis. In: *Hybrid Zones and the Evolutionary Process*, (Harrison, R. G., ed.). pp. Oxford University Press, England.

Ina, Y. 1996. Pattern of synonymous and nonsynonymous substitutions: An indicator of mechanisms of molecular evolution. *Journal of Genetics* **75**: 91-115.

Julius, D., Blair, L., Brake, A., Sprague, G. and Thorner, J. 1983. Yeast α factor is processed from a larger precursor polypeptide: The essential role of a membrane-bound dipeptidyl aminopeptidase. *Cell* **32**: 839-852.

Julius, D., Schekman, R. and Thorner, J. 1984. Glycosylation and processing of prepro-α-factor through the yeast secretory pathway. *Cell* **36**: 309-318.

Karlsson, M., Nygren, K. and Johannesson, H. 2008. The evolution of the pheromonal signal system and its potential role for reproductive isolation in heterothallic *Neurospora*. *Molecular Biology and Evolution* **25**: 168-178.

Kim, H.-K., Lee, T. and Yun, S.-H. 2008. A putative pheromone signaling pathway is dispensable for self-fertility in the homothallic ascomycete *Gibberella zeae*. *Fungal Genetics and Biology* **45**: 1188-1196.

Kim, H. and Borkovich, K. A. 2004. A pheromone receptor gene, *pre-1*, is essential for mating type-specific directional growth and fusion of trichogynes and female fertility in *Neurospora crassa*. *Molecular Microbiology* **52**: 1781-1798.

Kim, H. and Borkovich, K. A. 2006. Pheromones are essential for male fertility and sufficient to direct chemotropic polarized growth of trichogynes during mating in *Neurospora crassa*. *Eukaryotic Cell* **5**: 544-554.

Kim, H., Metzenberg, R. L. and Nelson, M. A. 2002. Multiple functions of *mfa-1*, a putative pheromone precursor gene of *Neurospora crassa*. *Eukaryotic Cell* **1**: 987-999.

Klein, J. 1987. Origin of major histocompatibility complex polymorphism: The trans-species hypothesis. *Human Immunology* **19**: 155-162.

Kurjan, J. 1993. The pheromone response pathway in *Saccharomyces cerevisiae*. *Annual Review of Genetics* **27**: 147-179.

Kvas, M., Marasas, W. F. O., Wingfield, B. D., Wingfield, M. J. and Steenkamp, E. T. 2009. Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. *Fungal Diversity* **34**: 1-21.

Lee, J., Leslie, J. F. and Bowden, R. L. 2008. Expression and function of sex pheromones and receptors in the homothallic ascomycete *Gibberella zeae*. *Eukaryotic Cell* **7**: 1211-1221.

Leslie, J., Zeller, K., Wohler, M. and Summerell, B. 2004. Interfertility of Two Mating Populations in the Gibberella Fujikuroi Species Complex. *European Journal of Plant Pathology* **110**: 611-618.

Leslie, J. F. and Summerall, B. A. 2006. *The Fusarium laboratory manual*. Blackwell Professional, Ames, Iowa.

Mayrhofer, S. and Pöggeler, S. 2005. Functional characterization of an α-factor-like *Sordaria macrospora* peptide pheromone and analysis of its interaction with its cognate receptor in *Saccharomyces cerevisiae*. *Eukaryotic Cell* **4**: 661-672.

Mayrhofer, S., Weber, J. M. and Pöggeler, S. 2006. Pheromones and pheromone receptors are required for proper sexual development in the homothallic ascomycete *Sordaria macrospora*. *Genetics* **172**: 1521-1533.

McCullough, J. and Herskowitz, I. 1979. Mating pheromones of *Saccharomyces kluyveri*: pheromone interactions between *Saccharomyces kluyveri* and *Saccharomyces cerevisiae*. *Journal of Bacteriology* **138**: 146-154.

Naider, F. and Becker, J. M. 2004. The α-factor mating pheromone of *Saccharomyces cerevisiae*: a model for studying the interaction of peptide hormones and G protein-coupled receptors. *Peptides* **25**: 1441-1463.

Nei, M. and Gojobori, T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**: 418-426.

Nei, M., Gu, X. and Sitnikova, T. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune-system. *Proceedings of the National Academy of Sciences of the United States of America* **94**: 7799-7806.

Nei, M. and Kumar, S. 2000. *Molecular evolution and phylogenetics*. Oxford University Press, New York, New York.

Nei, M. and Rooney, A. P. 2005. Concerted and birth-and-death evolution of multigene families. *Annual Review of Genetics* **39**: 121-152.

Nix, D. and Eisen, M. 2005. GATA: a graphic alignment tool for comparative sequence analysis. *BMC Bioinformatics* **6**: 9.

O'Donnell, K., Cigelnik, E. and Nirenberg, H. I. 1998. Molecular systematic and phylogeography of the *Gibberella fujikuroi* species complex. *Mycologia* **90**: 465-493.

Olesnicky, N. S., Brown, A. J., Honda, Y., Dyos, S. L., Dowell, S. J. and Casselton, L. A. 2000. Self-compatible B mutants in *coprinus* with altered pheromone-receptor specificities. *Genetics* **156**: 1025-1033.

Palumbi, S. R. 1999. All males are not created equal: Fertility differences depend on gamete recognition polymorphisms in sea urchins. *Proceedings of the National Academy of Sciences of the United States of America* **96**: 12632-12637.

Palumbi, S. R. 2008. Speciation and the evolution of gamete recognition genes: pattern and process. *Heredity* **102**: 66-76.

Pöggeler, S. 2000. Two pheromone precursor genes are transcriptionally expressed in the homothallic ascomycete *Sordaria macrospora*. *Current Genetics* **37**: 403-411.

Pöggeler, S. 2002. Genomic evidence for mating abilities in the asexual pathogen *Aspergillus fumigatus*. *Current Genetics* **42**: 153-160.

Rooney, A. P. and Ward, T. J. 2005. Evolution of a large ribosomal RNA multigene family in filamentous fungi: Birth and death of a concerted evolution paradigm. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 5084-5089.

Sahni, N., Yi, S., Daniels, K. J., Huang, G., Srikantha, T. and Soll, D. R. 2010. Tec1 mediates the pheromone response of the white phenotype of *Candida albicans*: insights into the evolution of new signal transduction pathways. *PLoS Biology* **8**.

Shapiro, S. S. and Wilk, M. B. 1965. An analysis of variance test for normality (complete samples). *Biometrika* **52**: 591-611.

Singh, A., Chen, E. Y., Lugovoy, J. M., Chang, C. N., Hitzeman, R. A. and Seeburg, P. H. 1983. *Saccharomyces cerevisiae* contains two discrete genes coding for the α-factor pheromone. *Nucleic Acids Research* **11**: 4049-4063.

Steenkamp, E. T., Wingfield, B. D., Coutinho, T. A., Wingfield, M. J. and Marasas, W. F. O. 1999. Differentiation of *Fusarium subglutinans* f. sp. *pini* by histone gene sequence data. *Applied and Environmental Microbiology* **65**: 3401-3406.

Steenkamp, E. T., Wingfield, B. D., Desjardins, A. E., Marasas, W. F. O. and Wingfield, M. J. 2002. Cryptic speciation in *Fusarium subglutinans*. *Mycologia* **94**: 1032-1043.

Stötzler, D. and Duntze, W. 1976. Isolation and characterization of four related peptides exhibiting alpha factor activity from *Saccharomyces cerevisiae*. *European Journal of Biochemistry* **65**: 257-262.

Swanson, W. J. and Vacquier, V. D. 2002. The rapid evolution of reproductive proteins. *Nature Reviews Genetics* **3**: 137-144.

Swanson, W. J. and Vacquier, V. D. 1998. Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. *Science* **281**: 710-712.

Symonds, M. R. E. and Elgar, M. A. 2008. The evolution of pheromone diversity. *Trends in Ecology and Evolution* **23**: 220-228.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0. *Molecular Biology and Evolution* **24**: 1596-1599.

Tatusova, T. A. and Madden, T. L. 1999. BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences. *FEMS Microbiology Letters* **174**: 247-50.

Thompson, J. D., Higgins, D. G. and Gibson, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**: 4673-4680.

Turgeon, B. G. 1998. Application of mating type gene technology to problems in fungal biology. *Annual Review of Phytopathology* **36**: 115 - 137.

Turner, E., Jacobson, D. J. and Taylor, J. W. 2010. Reinforced postmating reproductive isolation barriers in *Neurospora*, an Ascomycete microfungus. *Journal of Evolutionary Biology* **23**: 1642-1656.

Wik, L., Karlsson, M. and Johannesson, H. 2008. The evolutionary trajectory of the mating-type (*mat*) genes in *Neurospora* relates to reproductive behavior of taxa. *BMC Evolutionary Biology* **8**: 109.

Wilcoxon, F. 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* **1**: 80-83.

Yang, Z. 2007. PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**: 1586-1591.

Yang, Z. and Nielsen, R. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution* **19**: 908-917.

Yang, Z. H., Nielsen, R., Goldman, N. and Pedersen, A. M. K. 2000a. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431 - 449.

Yang, Z. H., Swanson, W. J. and Vacquier, V. D. 2000b. Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Molecular Biology and Evolution* **17**: 1446 - 1455.

Zhang, L., Baasiri, R. A. and Van Alfen, N. K. 1998. Viral repression of fungal pheromone precursor gene expression. *Molecular and Cellular Biology* **18**: 953-959.

Table 1. *Fusarium* isolates from the *Gibberella fujikuroi* complex used in this study

Isolate number [a]		*Fusarium* spp.	Mating population	Mating type
MRC	KSU			
8559	A-00149	*F. verticillioides*	A	*MAT-1*
8560	A-00999			*MAT-2*
8552	B-03853	*F. sacchari*	B	*MAT-1*
8551	B-03852			*MAT-2*
8532	C-01993	*F. fujikuroi*	C	*MAT-1*
8534	C-01995			*MAT-2*
8549	D-04854	*F. proliferatum*	D	*MAT-1*
8550	D-04853			*MAT-2*
8553/6483	E-00990	*F. subglutinans* (group 2)	E	*MAT-1*
8554/6512	E-02192			*MAT-2*
1084		*F. subglutinans* (group 1)		*MAT-1*
7828				*MAT-2*
8558	F-04094	*F. thapsinum*	F	*MAT-1*
8557	F-04093			*MAT-2*
8546	G-05111	*F. nygamai*	G	*MAT-1*
8547	G-05112			*MAT-2*
7488	H-10847	*F. circinatum*	H	*MAT-1*
6213	H-10850			*MAT-2*
8545	I-11616	*F. konzum*	I	*MAT-1*
8544	I-11615			*MAT-2*
8092/8093	X4382	*F. mangiferae*	N/A	*MAT-1*
7559	11781			*MAT-2*
2802	11783	*F. sterilihyphosum*	N/A	*MAT-1*
8105	11782			*?*

[a] Isolates may be obtained from PROMEC, Medical Research Council (MRC), Tygerberg, South Africa or the Department of Plant Pathology, Kansas State University (KSU), Manhattan, Kansas, USA;

Table 2. Results of tests for positive selection

Dataset	Model [a]	Parameter Estimates [b]	lnL [c]	2δ[d]
Saccharomycetes α	M7	p=0.839 q=2.764	-566.745	0.046
	M8	p_0=0.964 p=0.906 q=3.339 ω=1.000	-566.722	
Saccharomycetes a	M7	p=0.371 q=3.837	-368.243	0.000
	M8	p_0=1.000 p=0.372 q=3.837 ω=1.000	-368.243	
Sordariomycetes α	M7	p=0.332 q=22.689	-273.156	0.000
	M8	p_0=0.999 p=0.332 q=22.689 ω=1.000	-273.156	
Eurotiomycetes α	M7	p=0.164 q=1.281	-148.346	6.230*
	M8	p_0=0.933 p=0.394 q=11.128 ω=9.923	-145.231	

[a] "Site" models in CODEML (Yang *et al.* 2000a)

[b] Parameter estimates generated by CODEML

[c] ln likelihood score for each model, as calculated by CODEML

[d] Likelihood Ratio Test (LRT) statistic, δ is the differences between the ln likelihood values

* Indicates a significant difference at a value of $P \leq 0.05$

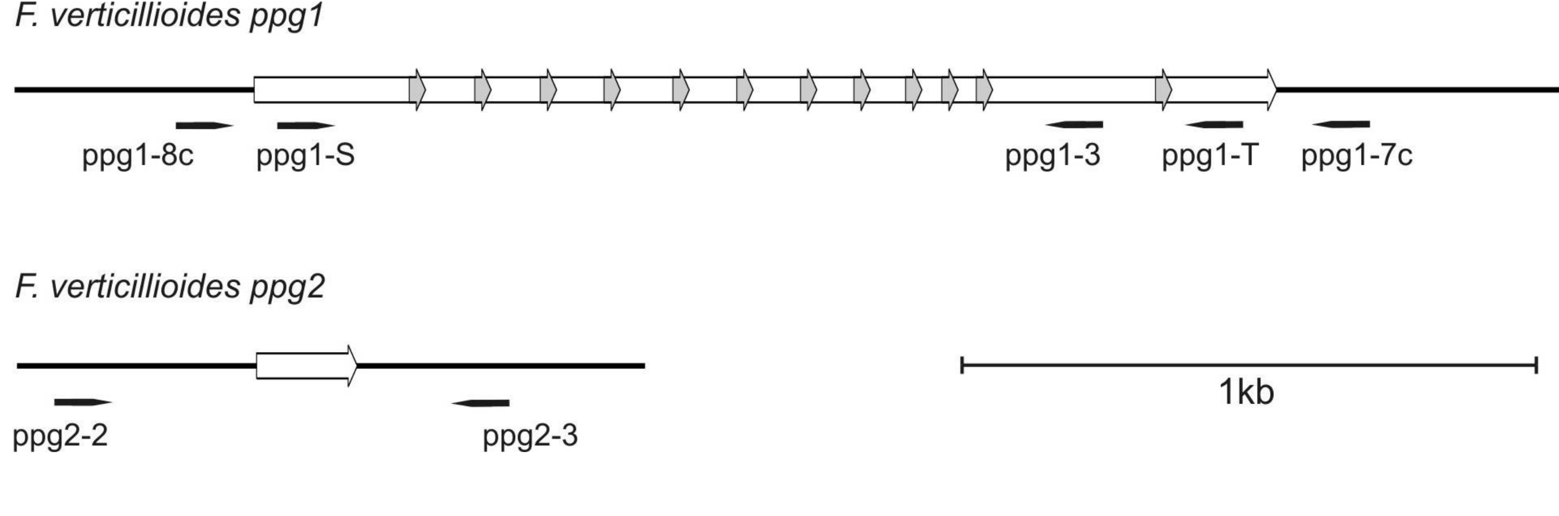


Fig. 1. Predicted *ppg1* and *ppg2* genes from *F. verticillioides*. (GenBank accession number AAIM02000073). Clear block arrows indicate coding regions. Grey block arrows indicate repeats of the putative mature α-factor-like peptide. Black arrows indicate primer binding sites and direction.

Fig. 2. a-class pheromone precursor proteins and predicted mature peptides. See Appendix D for complete DNA sequences, GenBank accession numbers and references. CaaX and CpaX motifs are shaded yellow. Mature a-factor peptides from yeasts are underlined. The hypothetical mature peptide repeats of the *G. fujikuroi* species are boxed. Multiple genes encoded in the same genome are indicated by numbers in brackets. The default amino acid coloration of BioEdit (PAM 250) (Hall, 1999) is used.

Organism		Sequence	
SORDARYOMYCETES			
F. verticillioides		MPSTKNTTAQTPGYPLTCSGMAKPTKDNQTPGYPLTCNVMKKPAANGQTPGYPLTCTVM	(MRC8559)
		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8560)
F. sacchari		MPSTKNTTAQTPGYPLTCTVMAKPIKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8552)
		MPSTKNTTAQTPGYPLTCTVMAKPIKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8551)
F. fujikuroi		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8532)
		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8534)
F. proliferatum		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8549)
		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8550)
F. subglutinans 1		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPVANGQTPGYPLTCTVM	(MRC1084)
		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPVANGQTPGYPLTCTVM	(MRC7828)
F. subglutinans 2		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8553)
		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8554)
F. thapsinum		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8558)
		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8557)
F. nygamai		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCTVM	(MRC8546)
		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCTVM	(MRC8547)
F. circinatum		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCSVMKKPVANGQTPGYPLTCTVM	(MRC7488)
		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCSVMKKPVANGQTPGYPLTCTVM	(MRC6213)
F. konzum		MPSTKNTTAQTPGYPLTCTIMAKPTQDNQTPGYPLTCSVMKKPVANGQTPGYPLTCTVM	(MRC8545)
		MPSTKNTTAQTPGYPLTCTIMAKPTQDNQTPGYPLTCSVMKKPVANGQTPGYPLTCTVM	(MRC8544)
F. mangiferae		MPSTKNTTAQTPGYPLTCTVMAKPIKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8092)
		MPSTKNTTAQTPGYPLTCSVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC7559)
F. sterilihyphosum		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC8105)
		MPSTKNTTAQTPGYPLTCTVMAKPTKDNQTPGYPLTCSVMKKPAANGQTPGYPLTCTVM	(MRC2802)
F. oxysporum		MPSTKNTTAQTPGYPLSCSVMAKPTKDNQTPGYPLSCSVMKKPAANGQTPGYPLSCTVM	
G. zeae		MPSTKPTSSQKPGYPLSCTVM	
V. dahliae		MPSYTQKNGGGHSGCSIMKNGGGHSGCSIMKNGGGHSGCSIMKNGGGHSGCSVMKNGGGHSGCAIM	
V. albo-atrum		MPSYTQKNGGGHSGCSIMKNGGGHSGCSIMKNGGGHSGCSIMKNGGGHSGCSVMKNGGGHSGCAIM	
T. atroviridae		MASEGVQSFSAVQAKKGQSPQNSPATSQFIGYLGCTIM	
T. reesii		MAQTGNLGCTVMAKPQSVERKRLIGCSVMTKPAANDKKFTGLLGCTVM	
T. virens		MAAIRTTTTIGCTVMAKPKPTTTVGCNVMAKPHGNNKSIFDKYHTTIGCTIM	
M. grisea		MSPSTKNIPAPVAGARAGPIHYCVIM	
C. parasitica	(1)	MPSNTQTSNSSMGVNGYSYCVVM	
	(2)	MPSNTQTSNSSMGVNGYSYCVVM	
P. anserina		MPSTTAQTKVPQTSTNFNSYCVVM	
N. crassa		MPSTAASTKVPQTTMNFNGYCVVM	
S. macrospora		MPSTAASTKVPQTTMNFNGYCVVM	
C. globosum		MPSTTTQTKVPQTSTNFNGYCVVM	
SACCHAROMYCETES			
S. cerevisiae	(1)	MQPSTATAAPKEKTSSEKKDNYIIKGVFWDPACVIA	
	(2)	MQPITTASTQATQKDKSSEKKDNYIIKGLFWDPACVIA	
S. bayanus	(1)	MQPITTVSAAPKDKTSTEKKDNYIIKGVFWDPACIIV	
	(2)	MQPVATVSAQASQKDKSSEKKDNYIIKGLFWDPACVIA	
S. kudriavzevii	(1)	MQPSTITAAPKDKTSSEKKDNYIIKGVFWDPACVIA	
	(2)	MQPTATVSAQASQKDRSSEKKDNFIIKGLFWDPACVIA	
S. mikatae	(1)	MQPSTITAAPKDKTNTENKDNYIIKGVFWDPACVIA	
	(2)	MQPITTTSAQAAQKNKSSEKKDNYIIKGLFWDPACVIA	
S. paradoxis		MQPSTVTAAPKDKTSAEKKDNYIIKGVFWDPACVIA	
S. pastorianus	(1)	MQPSTATAAPKEKTSSEKKDNYIIKGVFWDPACVIA	
	(2)	MQPITTASTQATQKDKSSEKKDNYIIKGLFWDPACVIA	
	(3)	MQPVTTISAQASQKDKSSEKKDNYIIKGLFWDPACVIA	
S. kluyveri	(1)	MQPKSNATQKDSAENKDNWIIEGLAWNPQCVII	
	(2)	MKAATHATQKGSTEDKENWIIKGFAWDPQCIIA	
K. lactis		MQPTQQASQNESAENKENWIIPGFVWVPQCVVA	
K. waltii		MQPIAQATQNDSSDNKDNWIHKGLAWDPQCVIA	
C. glabrata		MQPTIEATQKDNTQEKRDNYIVKGFFWSPDCVIA	
V. polyspora	(1)	MQSTTYAAQKNSSEKKDNYIVKGWFWDPECVIA	
	(2)	MQSTTYAAQKNSSEKKDNYIVKGWFWDPECVIV	
	(3)	MQSTTYAAQKNSSEKKDNYIVKGWFWDPECVIV	
K. delphensis		MEPAQATQKDNSQDKKDNYVWRGRFWYPECVVV	
A. gossypii		MQLTNNTNKDESTENKDNWIAKGYMWTPQCVIV	
S. castellii	(1)	MQPTTQATHKDNSAEKQDNYIVKGLFWDPACVIA	
	(2)	MQPSAQASQKDNTAENKDNYIVKGLFWDPACVIA	
TAPHRINOMYCETES			
S. pombe	(1)	MDSMANSVSSSSVVNAGNKPAETLNKTVKNYTPKVPYMCVIA	
	(2)	MDSIATNTHSSSIVNAYNNNPTDVVKTQNIKNYTPKVPYMCVIA	
	(3)	MDSMANTVSSSVVNTGNKPSETLNKTVKNYTPKVPYMCVIA	

Fig. 3. α-class precursor protein structures and mature pheromone sequences. See Appendix D for complete DNA sequences, GenBank accession numbers and references. Lines with blocks schematically represent the number of repeats in each precursor protein. Multiple genes present within the same genome are displayed beneath one another and numbered. Repeat colours indicate different amino acid sequence "variants" of the pheromone within a single species. For each species, the sequence of all putative pheromone repeats are displayed on the right, with dots indicate residues that are identical to those in the first sequence. The default amino acid coloration of BioEdit (PAM 250) (Hall 1999) is used.

Species	Strain	Pheromone sequence
SORDARYOMYCETES		
Fusarium verticillioides	MRC8559 MRC8560	WCTWRGQPCW ..M.......
Fusarium sacchari	MRC8552 MRC8551	WCTWRGQPCW ..M.......
Fusarium fujikuroi	MRC8532 MRC8534	WCTWRGQPCW ..M.......
Fusarium proliferatum	MRC8549 MRC8550	WCTWRGQPCW ..M.......
Fusarium subglutinans (1)	MRC1084 MRC7828	WCTWRGQPCW ..M.......
Fusarium subglutinans (2)	MRC8553 MRC8554	WCTWRGQPCW ..M.......
Fusarium thapsinum	MRC8558 MRC8557	WCTWRGQPCW ..M.......
Fusarium nygamai	MRC8546 MRC8547	WCTWRGQPCW ..M.......
Fusarium circinatum	MRC7488 MRC6213	WCTWRGQPCW ..M....... ...R......
Fusarium konzum	MRC8545 MRC8545	WCTWRGQPCW ..M....... ....C ..L.......
Fusarium mangiferae	MRC8092 MRC7559	WCTWRGQPCW ..M.......
Fusarium sterilihyphosum	MRC8105	WCTWRGQPCW ..M.......
Fusarium oxysporum		WCTWRGQPCW ..L....... ..M.......
Gibberella zeae		WCTWKGQPCW ..W.......
Trichoderma reesei		WCYRIGEPCW ..WIL.GK..
Trichoderma atroviridae		WCWRVGESCW
Trichoderma virens		WCYRVGMTCGW ...QT.QP.-.
Verticillium dahliae		GKPCPRPGQGCW D........... Q...........
Verticillium albo-atrum		GKPCPRPGQGCW Q...........
Cryphonectria parasitica		WCLFHGEGCW
Magnaporthe grisea		QWCPRRGQPCW
Neurospora crassa		QWCRIHGQSCW
Sordaria macrospora		QWCRIHGQSCW
Podospora anserina		QWCLRFVGQSCW
Chaetomium globosum		WCKQFLGMPCW ..T....Q... ..R.......-
LEOTIOMYCETES		
Botryotinia fuckeliana		WCGRPGQPC
Sclerotinia sclerotiorum		WCGRPGQPC
EUROTIOMYCETES		
Penicillium chrysogenum		WCGHIGQGC
Aspergillus nidulans		WCRFAGRIC ....R.QV.
Aspergillus fumigatus		WCHLPGQGC
Neosartorya fischeri		WCHLPGQGC
Aspergillus clavatus		WCELPGQGC
Aspergillus terreus		WCWLPGQGC
Aspergillus niger		WCVLPGQPC
Aspergillus flavus		WCSLPAQGC ..A..G...
Aspergillus oryzae		WCSLPAQGC ..A..G...
DOTHIDIOMYCETES		
Alternaria brassicicola		WSFTQKRPYGLPIG ......H.......
Phaeosphaeria nodorum		WKYNGWRYRPYGLPVG .R.-----........ -............... -R..............
Pyrenophora tritici-repentis		WTQKRPYGMPVG ...H........ .R..........

Species	Genes	Pheromone sequences
SACCHAROMYCETES		
Saccharomyces cerevisiae	(1) (2)	WHWLQLKPGQPMY ....N.R......
Saccharomyces bayanus	(1) (2)	WHWLQLKPGQPMY ......R...... ....K.R......
Saccharomyces kudriavzevii	(1) (2)	WHWLQLKPGQPMY ......R......
Saccharomyces mikatae	(1) (2)	WHWLQLKPGQPMY ....S.R......
Saccharomyces paradoxus	(1) (2)	WHWLQLKPGQPMY ...N.R......
Saccharomyces Pastorianus	(1) (2) (3) (4)	WHWLQLKPGQPMY .....F....... ....K.R...... ....T.R...... ......R...... ....N.R......
Saccharomyces kluyveri		WHWLSFSKGEPMY
Saccharomyces naganishii		WHWLRLSYGQPIY
Saccharomyces uvarum		WHWLQLKPGQPMY
Saccharomyces castellii	(1) (2) (3)	WHWLRLDPGQPLY ....S..A.....
Vanderwaltozyma polyspora	(1) (2)	WHWLELDNGQPIY ....R.RY.E...
Candida glabrata		WHWVKIRKGQGLF ....RL.......
Pichia stipitis		WHWTSYGVFEPG
Lodderomyces elongisporus	(1) (2)	WMWTRYGRFSPV
Kluyveromyces delphensis		WHWLSVRPGQPIY
Kluyveromyces lactis		WSWITLRPGQPIF
Kluyveromyces waltii		WRWLSLARGQPMY
Lachanceae thermotolerans		WRWLSLSRGQPMY ......A......
Clavispora lusitaniae		WGWIHFLNTDVIG .K....R...... .K..K.R...... .R..N.R......
Yarrowia lipolytica	(1) (2)	WRWFWLPGYGEPNW
Candida albicans		GFRLTNFGYFEPG
Candida dubliniensis		RFRLTNFGYFEPG K.K.......... G............
Candida tropicalis		KFKFRLTRYGWFSPN
Candida parapsilosis		KPHWTTYGYYEPQ
Pichia guilliermondi		KKNSRFLTYWFFQPIM
Deberomyces hansenii		KFHWMTYRFFQPNL
Zygosaccharomyces rouxii		HFIELDPGQPMF ..V.........
TAPHRINOMYCETES		
Schizosaccharomyces japonicus		VSDRVKQMLSHWWNFRNPDTANL
Schizosaccharomyces pombe		KSYADFLRVYQSWNTFANPDRPNL .T......A.......V....... ........A.H.....V.......
Schizosaccharomyces octosporus		KTYEDFLRVYKNWQTFQNPDRPDL ....WS......... ..........Q..E.......... ......E...Q..E..........

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

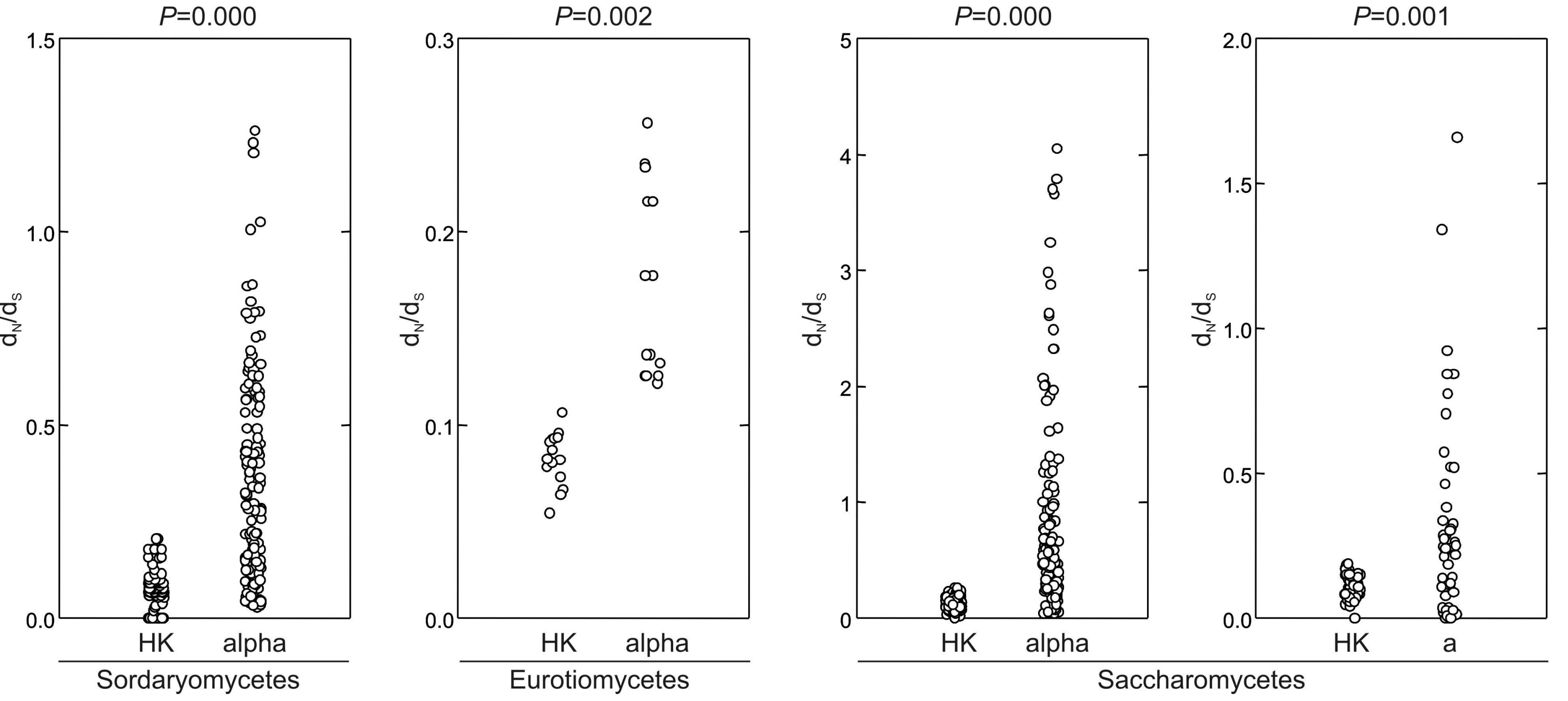


Fig. 4. Dot plots of pairwise d_N/d_S values for pheromones versus house-keeping genes. Each point represents a pairwise d_N/d_S value between a pair of species. Datasets are displayed independently, with values for house-keeping (HK) genes on the left and pheromones on the right. "*P*" values represent significance values from the Wilcoxon signed ranks test. $P \leq 0.05$ indicates a significant difference between the two samples. Sample sizes for the four datasets are 156, 15, 145, and 55 species pairs, respectively.

Fig. 5. Neighbour joining distance trees of DNA sequences encoding mature α-factor-like peptide repeats. (A) Saccharomycetes dataset, excluding species with repeats too divergent to be aligned. (B) Sordariomycetes dataset with the subtree containing all repeats from *Fusarium* species collapsed. (C) Subtree containing all repeats from *Fusarium* species. Colors indicate repeats from the same species. All the sordariomycete taxon names are followed by a single number, representing repeat position starting at the N-terminal. The Saccharomycetes each have two numbers because some species carry multiple genes. The first number represents the gene; the second represents the repeat position within the precursor, again starting at the N-terminal. Bootstrap values greater than 50 are displayed on branches.

0.05

A

S. paradoxis 1 1
S. kudriavzevii 1 4
S. bayanus 1 2
S. kudriavzevii 1 1
S. kudriavzevii 1 2
S. kudriavzevii 1 3
S. bayanus 1 1
S. paradoxis 1 2
S. paradoxis 1 3
S. paradoxis 1 4
S. paradoxis 2 2
S. cerevisiae 1 1
S. uvarum 1 2
S. pastorianus 1 2
S. mikatae 2 2
S. cerevisiae 2 2
S. pastorianus 4 2
S. pastorianus 3 3
S. bayanus 2 3
S. kudriavzevii 2 2
S. cerevisiae 1 2
S. cerevisiae 1 3
S. pastorianus 1 3
S. pastorianus 1 1
S. uvarum 1 1
S. mikatae 1 1
S. mikatae 1 2
S. mikatae 1 3
S. cerevisiae 1 4
S. uvarum 1 3
S. pastorianus 1 5
S. pastorianus 1 4
S. pastorianus 3 2
S. paradoxis 2 1
S. pastorianus 2 1
S. pastorianus 2 3
S. pastorianus 2 2
S. pastorianus 3 1
S. bayanus 2 1
S. bayanus 1 3
S. bayanus 2 2
S. kudriavzevii 2 1
S. mikatae 2 1
S. pastorianus 4 1
S. cerevisiae 2 1
K. delphensis 1 1
K. delphensis 1 2
K. delphensis 1 3
K. delphensis 1 4
V. polyspora 1 1
V. polyspora 1 2
V. polyspora 1 3
V. polyspora 1 4
V. polyspora 1 5
Z. rouxii 1 1
Z. rouxii 1 3
Z. rouxii 1 2
Z. rouxii 1 4
Z. rouxii 1 6
Z. rouxii 1 5
Z. rouxii 1 8
Z. rouxii 1 7
S. castellii 1 1
S. castellii 1 3
S. castellii 1 2
S. castellii 1 4
S. castellii 1 5
S. castellii 2 3
S. castellii 2 2
S. castellii 2 1
S. castellii 3 1
C. glabrata 1 1
C. glabrata 1 2
C. glabrata 1 3
K. lactis 1 1
K. lactis 1 2
K. lactis 1 3
K. lactis 1 4
D. hansenii 1 1
D. hansenii 1 3
D. hansenii 1 4
D. hansenii 1 2
D. hansenii 1 5
V. polyspora 2 1
V. polyspora 2 2
V. polyspora 2 4
V. polyspora 2 3
S. naganishi 1 1
S. naganishi 1 3
S. naganishi 1 2
L. thermotolerans 1 1
L. thermotolerans 1 2
L. thermotolerans 1 3
K. waltii 1 1
K. waltii 1 2
K. waltii 1 3
K. waltii 1 4
K. waltii 1 5
L. thermotolerans 1 4
Y. lypolytica 2 3
Y. lypolytica 2 1
Y. lypolytica 2 2
Y. lypolytica 1 1
Y. lypolytica 1 2
Y. lypolytica 1 4
Y. lypolytica 1 3
S. kluyveri 1 1
S. kluyveri 1 2

B

C

P. anserina 1
P. anserina 2
C.globosum 3
C.globosum 1
C.globosum 2
T. reesii 1
T. reesii 3
T. reesii 2
T. reesii 5
T. virens 4
T. virens 2
T. virens 1
T. virens 3
T. atroviridae 1
T. atroviridae 3
T. atroviridae 2
T. atroviridae 4
M. grisea 1
M. grisea 2
M. grisea 4
M. grisea 3
V. albo-atrum 1
V. albo-atrum 3
V. albo-atrum 4
V. dahliae 4
V. dahliae 6
V. dahliae 5
V. albo-atrum 2
V. dahliae 2
V. dahliae 1
V. dahliae 3
C. parasitica 1
C. parasitica 7
C. parasitica 3
C. parasitica 5
C. parasitica 2
C. parasitica 4
C. parasitica 6
T. reesii 4
N. crassa 4
S. macrospora 5
N. crassa 5
N. crassa 2
N. crassa 1
S. macrospora 4
S. macrospora 1
N. crassa 3
S. macrospora 2
S. macrospora 3

C

F. sterilihyphosum 6
F. subglutinans (grp1) 6
F. fujikuroi 6
F. sacchari 7
G. zeae 5
G. zeae 6
G. zeae 8
G. zeae 7
G. zeae 9
G. zeae 3
G. zeae 1
G. zeae 2
G. zeae 4
F. sacchari 6
F. mangiferae 6
F. nygamai 2
F. fujikuroi 5
F. proliferatum 4
F. konzum 5
F. sterilihyphosum 5
F. subglutinans (grp1) 5
F. nygamai 4
F. thapsinum 4
F. circinatum 5
F. oxysporum 5
F. subglutinans (grp2) 3
F. oxysporum 9
F. konzum 10
F. nygamai 10
F. subglutinans (grp1) 14
F. circinatum 12
F. subglutinans (grp2) 8
F. sterilihyphosum 13
F. verticillioides 12
F. proliferatum 14
F. thapsinum 11
F. mangiferae 14
F. sacchari 13
F. fujikuroi 10
F. proliferatum 15
F. thapsinum 12
F. subglutinans (grp1) 15
F. konzum 11
F. subglutinans (grp2) 9
F. sterilihyphosum 14
F. oxysporum 10
F. nygamai 11
F. mangiferae 15
F. sacchari 15
F. fujikuroi 11
F. verticillioides 13
F. mangiferae 12
F. circinatum 10
F. nygamai 8
F. subglutinans (grp2) 6
F. sterilihyphosum 12
F. sacchari 12
F. subglutinans (grp1) 12
F. oxysporum 7
F. thapsinum 9
F. proliferatum 12
F. verticillioides 10
F. fujikuroi 8
F. nygamai 9
F. verticillioides 11
F. oxysporum 8
F. subglutinans (grp2) 7
F. konzum 9
F. circinatum 11
F. thapsinum 10
F. subglutinans (grp1) 13
F. konzum 8
F. subglutinans (grp2) 5
F. sterilihyphosum 11
F. proliferatum 11
F. circinatum 9
F. subglutinans (grp1) 11
F. circinatum 13
F. thapsinum 7
F. fujikuroi 9
F. verticillioides 9
F. mangiferae 13
F. proliferatum 13
F. sacchari 14
F. verticillioides 7
F. nygamai 7
F. subglutinans (grp2) 4
F. circinatum 8
F. subglutinans (grp1) 10
F. subglutinans (grp1) 7
F. thapsinum 8
F. proliferatum 10
F. fujikuroi 1
F. subglutinans (grp1) 9
F. oxysporum 6
F. mangiferae 11
F. sacchari 11
F. verticillioides 6
F. sterilihyphosum 10
F. verticillioides 8
F. sterilihyphosum 9
F. konzum 7
F. nygamai 6
F. mangiferae 10
F. sacchari 10
F. oxysporum 4
F. thapsinum 3
F. circinatum 4
F. subglutinans (grp1) 4
F. sterilihyphosum 4
F. fujikuroi 4
F. konzum 4
F. proliferatum 8
F. proliferatum 3
F. nygamai 3
F. verticillioides 5
F. mangiferae 5
F. sacchari 5
F. oxysporum 3
F. thapsinum 2
F. fujikuroi 3
F. verticillioides 4
F. mangiferae 4
F. sacchari 4
F. konzum 3
F. proliferatum 2
F. subglutinans (grp1) 3
F. sterilihyphosum 3
F. circinatum 3
F. mangiferae 3
F. konzum 2
F. subglutinans (grp2) 2
F. sacchari 3
F. oxysporum 2
F. circinatum 2
F. fujikuroi 2
F. sterilihyphosum 2
F. verticillioides 3
F. mangiferae 2
F. sacchari 2
F. subglutinans (grp1) 2
F. verticillioides 2
F. circinatum 7
F. sterilihyphosum 1
F. circinatum 1
F. subglutinans (grp1) 8
F. sacchari 9
F. fujikuroi 7
F. sterilihyphosum 8
F. mangiferae 9
F. proliferatum 9
F. thapsinum 6
F. proliferatum 7
F. mangiferae 1
F. sacchari 1
F. konzum 1
F. nygamai 1
F. oxysporum 1
F. thapsinum 1
F. verticillioides 1
F. proliferatum 1
F. subglutinans (grp2) 1
F. subglutinans (grp1) 1
F. sterilihyphosum 7
F. proliferatum 6
F. sacchari 8
F. mangiferae 8
F. proliferatum 5
F. konzum 6
F. thapsinum 5
F. mangiferae 7
F. nygamai 5
F. circinatum 6

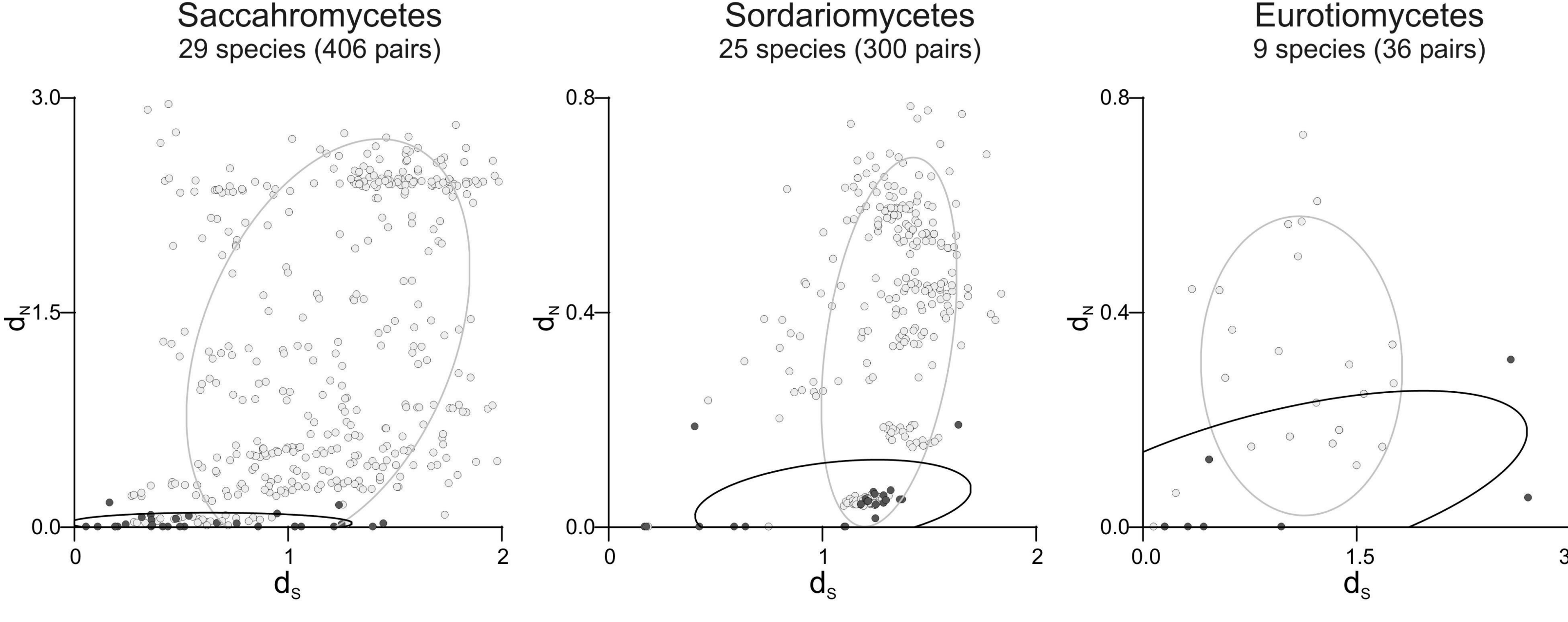


Fig. 6. Scatter plots of average pairwise d_N values for α-factor-like pheromones plotted against the corresponding d_S values. Each light (between-species) point represents the averaged d_N and d_S values for pair-wise comparisons between all possible pairs of repeats for a single pair of species. Each dark (within-species) point represents averaged d_N and d_S values for pair-wise comparisons between all possible pairs of tandem repeats from within the same precursor gene. Light points outnumber dark points because the light represent all possible species pairs while the dark each represent a single species. Gaussian bivariate ellipses are drawn for each group. Ellipses are centerd on sample means, with the major axes determined by the unbiased sample standard deviations ($P = 0.6827$) and orientation by the sample covariance between d_N and d_S.

F. verticillioides (MRC8560)
F. nygamai (MRC8546)
F. thapsinum (MRC8559)
F. fujikuroi (MRC8532)
F. proliferatum (MRC8550)
F. mangiferae (MRC8092)
F. sacchari (MRC8552)
F. konzum (MRC8544)
F. sterilihyphosum (MRC8105)
F. subglutinans grp1 (MRC1084)
F. subglutinans grp2 (MRC8554)
F. circinatum (MRC7488)

Fig. 7. GATA alignments of *ppg1* among species in the *Giberella fujikuroi* species complex. Gray boxes indicate positions of the predicted α-factor-like mature peptides (see figure 3 for amino acid sequences). Lines connect portions of sequence with high homology (lower cut-off bit-score of 24.9) and the darker the shading the stronger the homology. The accompanying reference phylogeny represents a consensus that was inferred from the results of previous studies (O'Donnell *et al.* 1998; Kvas *et al.* 2009).

○ synonymous substitution: no change to encoded encoded peptide
● non-synonymous substitution: resultes in alteration of the encoded peptide

Fig. 8. Theoretical model showing birth and death evolution of tandem pheromone repeats. Blocks indicate repeats of the mature α-factor-like peptide. Fill patterns indicate distinct amino acid sequence variants. Accompanying phylogenies represent relationships among underlying DNA sequences at each step.

SUMMARY

Reproductive isolation is an essential stage in speciation. In Ascomycetes, the ubiquitous distribution of many species suggests that sympatric speciation through assertive mating should be an important factor. The *MAT* locus and the pheromone/receptor system could both potentially contribute to the development of such sexual isolation. Alterations at the *MAT* loci could lead to distinct reproductive habits or a change in mating system, both of which can reduce gene-flow between species. However, if deliberate pre-mating sexual preferences exist, they are more likely to be determined by the pheromone/receptor system. This study of *Fusarium* reproductive genes, and comparisons with other Ascomycetes, has yielded numerous interesting findings regarding the evolution of these mate-recognition mechanisms and the implications thereof.

The *G. fujikuroi* and *F. graminearum* species complexes have offered an interesting comparison between heterothallic and homothallic *MAT* locus evolution. The value of comparative sequence analysis has been demonstrated in the discovery of a previously unknown gene, *MAT1-2-3*, which may be specific to members of the Order Hypocreales. While all *MAT* genes share similar regulatory elements, this is the first report of evidence that a transition to homothallism can be accompanied by the recruitment of distinct elements that could facilitate alternate expression of *MAT* genes. The *MAT* genes are also highly divergent between *Fusarium* spp., largely due to relaxed selective constraint, particularly in homothallic species. However, inter-specific gene-flow could curb *MAT* gene divergence among homothallic species. Despite strong reproductive barriers in the *G. fujikuroi complex*, the *F. sacchari MAT1-1* sequence appears to have been acquired through lateral transfer from a distant relative. Analytical analysis of the *MAT* locus novelties reported here, including the new *MAT* gene, will be necessary to determine their biological significance.

To investigate the extent of pheromone diversity in the Ascomycetes, and to gain clues as to its biological importance, pheromone peptides from seventy ascomycete species were compared. A number of reproductively incompatible species, such as those in the *G.*

fujikuroi complex, share identical pheromones; which implies that another mechanism must be responsible for the observed reproductive barriers. However, on the whole, pheromones are highly divergent among species. Both adaptive and non-adaptive evolution could have contributed to this pattern. In fact the structure of the α-class pheromone precursor gene, which consists of multiple repeats of the pheromone module, could facilitate rapid diversification through "birth-and-death" evolution. Within species, selection maintains pheromone peptides, implying that much of the inter-specific variation is functionally relevant. This further suggests that pheromone evolution could contribute to the generation of reproductive isolation between species.

The most general trend in the findings of this study is that ascomycete reproductive genes are highly divergent. This is in agreement with findings in other Kingdoms. A number of evolutionary forces are probably involved but weaker selective constraint, resulting from the fact that reproduction is not essential in these fungi, appears to be a common factor. This reproductive gene variability could be directly linked to speciation and, therefore, the great diversity in Ascomycetes.