Characterization of sex-pheromone receptor genes of *Fusarium* species and other *Sordariomycetes*

By

Tondani Kone

Submitted in partial fulfillment of the requirements for the degree

Magister Scientiae (Microbiology)

In the faculty of Natural and Agricultural Sciences

University of Pretoria

(30 June 2014)

Supervisor: Prof. E. T. Steenkamp

Co-supervisor: Mrs. G Fourie

Co-supervisor: Prof. N. A. van der Merwe

Co-supervisor: Prof. B. D. Wingfield

Declaration

I Tondani Kone declare that the thesis/dissertation, which I hereby submit for the degree Magister Scientiae (Microbiology) at the University of Pretoria, is my own work and has not been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE:……

DATE:……..

I dedicate this MSc degree to the people I have lost in my life:

My friend Nthuseni Nemukovhani (1986 – 2014)

Thank you for all the memories you have given me in the 27 years we shared together, the laughs, pranks and tough times. I will never know why you are gone however I know you were here to bring joy and laughter to those around. Your laughter and smile will remain one of the best gifts I have ever received, your love for your family is a lesson I will always keep and your desire to succeed will be shown through me.

The last day I saw you was when I accompanied you to finalize your MSc registration, you did not get a chance to complete the MSc yourself therefore I dedicate this to you, I would like this thesis to be my last gift to you “shaka langa”. I wish I had told you I love and appreciate you but I hope you knew it through my actions.
Rest in Peace Nthuseni “Thusi” Nemukovhani.

My grandfather Phophi William Kone Makhomu, a wise man with a big heart whose main teachings were love and unity.

Mrs Azwindini Emily Madzivhandila, the most gentle yet strong woman I have ever met, a real angel on earth.

Table of Contents

List of tables

Chapter 2:

Chapter 3:

List of figures

Chapter 2:

Chapter 3:

Acknowledgements

I would like to express my appreciation to the following people and funding bodies:

My supervisors, Prof. Emma Steenkamp, Prof. Brenda Wingfield and Prof. Albe van der Merve for the guidance, dedication and believing in me throughout my MSc degree. Your different characters and style of supervision were important in making my degree and thesis unique. You each offered different yet important roles in problem solving, the scientific and technological parts of this work.

Mrs Gerda Fourie, you were not just a co-supervisor but a friend. You did not only teach, assist and guide me with the MSc degree but also with my personal life. You were literally the “the woman behind a successful man” and I always know I can count on you. Thank you for never giving up on this project and the constant encouragement.

Mr Simon Martin and my six floor colleagues for all the assistance whenever I needed it, you made the work much easier and made me believe I too would complete my degree.

To my friends and family, thank you for the love, support and the laughter, each one of you made this journey and my life much easier.

To my mother, my biggest fan, thank you “mianga” for the love, encouragement, understanding and never giving up on me when the journey got tough. You have done

all these things expecting no payment, I will therefore like to paraphrase my favorite rap and say “there is no way I can pay you back, but this thesis is my way to show you that I understand why you do what you do for me. You are appreciated”.

To my partner, friend and confidant, the tough love and encouragement have finally paid off. You experienced the emotional roller coast of an MSc degree first hand and invested your time and emotions into this work as if it was your own. Now you finally get to see the rewards and feel the joy of completing an MSc degree, this is for us.

The University of Pretoria, the National Research Foundation (NRF), the Forestry and Agricultural Biotechnology Institute (FABI) and the Tree Protection Cooperative Program (TPCP) for the financial support for my MSc.

Preface

Ascomycete fungi consist of two pheromone receptor genes that function as pre-mating barriers. Gene sequence has been found to confer receptor specificity and thus the ability to mate between species. Little focus has however been put on studying and characterizing these receptors even though inter-fertility has been recorded between some species. In this study we look at these genes amongst various Ascomycetes species from the class Sordariomycetes which includes many pathogenic species such as *Fusarium* and *Neurospora.* their evolution is compared between inter-fertile and inter-sterile taxa in order to understand how inter-fertility persists and predict hybridization amongst species where the sexual stage has not observed.

In **chapter 1,** we provide an overview of available literature on the pheromone response system in fungi. The review starts by discussing sexual reproduction in fungi and then focuses specifically on the Ascomycota. These species were grouped into three sexual reproduction categories depending on whether they are self-fertile or not. The three categories, homothallism, heterothallism and pseudohomothallism are discussed. This is followed by a discussion of the two mating-type idiomorphs their roles in sexual reproduction. The pheromone response system, in terms of pheromones and corresponding receptors, as well as the downstream signaling pathways in various Ascomycota was considered. The chapter is concluded with a discussion of *Gibberella fujikuroi* species complex, the focal taxon of this study.

In Chapter 2, there was a comparison of the two pheromone receptor genes amongst *Fusarium spp* in the *Gibberella fujikuroi* species complex. Their phylogeny is first compared with the species phylogeny and to each other to determine if they evolved in a species specific manner and thus likely to confer mating-specificity. Due to DNA changes being either synonymous or non-synonymous, inter-specific differences in amino acid sequences were identified amongst the species in *Gibberella fujikuroi* complex in order to determine shared sequences that may be responsible for inter-fertility. An analysis of synonymous and non-synonymous nucleotide changes used to determine whether the two receptor genes are conserved or diverging amongst the species in the *GFC*. Furthermore regions under positive diversifying selection, usually associated with ligand binding, were identified to determine potential pheromone binging sites.

In **Chapter 3,** the receptors geneology of 35 Sordariomycetes is compared to the species phylogeny. In addition evolution of pheromone receptors amongst three taxa, two of which have inter-fertile species, was compared. Selection acting on the genes was analyzed to determine if the genes were conserved or diversifying in each taxa. Positive selected sites were also identified as well as lineages that experienced positive diversifying selection to determine its impact on inter-fertility in subsequent generations.

Chapter 1

Sexual reproduction in the Ascomycetes: A review

1. Introduction

Fungi are capable of both sexual and asexual reproduction. Sexual reproduction involves karyogamy followed by meiosis (Burnett 1968; Alexopoulos, Mims, Blackwell 1996), while asexual reproduction involves neither meiosis nor karyogamy but mitosis. Asexual reproduction in fungi, like with many other eukaryotes and prokaryotes, occurs by fragmentation, budding, or fission (Burnett 1968; Alexopoulos, Mims, Blackwell 1996). The mitotic spores that are formed through fungal asexual reproduction are known as conidiospores or conidia (singular = conidium). These are formed on hyphae or on conidiophores that are modified hyphal branches (Burnett 1968; Alexopoulos, Mims, Blackwell 1996; Leslie, Summerell, Bullock 2006).

Sexual and asexual reproduction not only differ in the processes and the structures involved but also play different roles within a population (Alexopoulos, Mims, Blackwell 1996). Though the benefits of sexual and asexual reproduction are still a matter to debate, most scientists agree that sexual reproduction provides the advantage of high genetic variation within a population (Heitman 2006), while asexual populations are considered clonal and less diverse (Burdon, Roelfs 1985). Variation is important for survival in changing environments and it is believed to be the reason why pathogens, such as *Candida albicans* and *Cryptococcus neoformans*, retain their sexual stage even though it is rarely observed in nature (Heitman 2006). Sexual recombination also allows for DNA repair and thus elimination of deleterious mutations within a population. In some fungi, sexual spores are more resilient and often produced for survival during

unfavorable conditions (Dyer, Ingram, Johnstone 1992; Milgroom 1996). Asexual reproduction on the other hand provides an energetically cheap and faster means to produce high numbers of progeny that are highly adapted to a specific environment compared to sexual spores which often requires more resources and are produced in lower numbers (Chamberlain, Ingram 1997; Fraser, Heitman 2004).

Fungal species often exhibit different morphologies during the sexual and asexual stages of their life cycle (Alexopoulos, Mims, Blackwell 1996; Leslie, Summerell, Bullock 2006). These morphologies have been termed the teleomorph and the anamorph, respectively, (Alexopoulos, Mims, Blackwell 1996; Leslie, Summerell, Bullock 2006) and are often given different species names. For example the teleomorph of *Fusarium proliferatum* is named *Gibberella intermedia* (Leslie, Summerell, Bullock 2006) and that of *Aspergillus nidulans* is named *Emaricella nidulans* (Seo, Han, Yu 2004). This dual nomenclature system has, however, recently been replaced with the one-fungus-one-name system (Wingfield et al. 2012). For most species the asexual stage is most commonly observed in nature compared to the sexual stage, in fact, the sexual stage of many fungi have not yet been described (e.g., *Fusarium oxysporum*, *F. andiyazi and F. dlamini*) (Leslie, Summerell, Bullock 2006). However the sexual stage of *Aspergillus parasiticus*, a species previously thought to not have the sexual stage, was recently observed and it is believed that most fungal species have a sexual stage which we have not been observed yet (Horn, Ramirez-Prado, Carbone 2009).

Mating of fungi in the dikaryotic phyla Ascomycota and Basidiomycota can be divided into various steps. The process of mating begins when two cells come into contact followed by cell wall fusion and migration of nuclei from the donor to recipient cell. This produces the dikaryon phase in which a cell contains two nuclei for a certain period before they fuse and undergo meiosis (Casselton, Olesnicky 1998; Billiard et al. 2012). In Ascomycota this phase is usually short lived while in the Basidiomycota, it is generally the predominant vegetative phase in nature (Casselton 2002; Billiard et al. 2012). After meiosis, mature sexual spores or meiospores are formed, in specialized cells called asci (singular = ascus) in Ascomycota. In the Basidiomycota these meiospores are formed outside on specialized cells called basidia (singular = basidium). These sexual spores are then released and germinate into colonies that are indistinguishable from those produced by asexual spores (Alexopoulos, Mims, Blackwell 1996; Bobrowicz et al. 2002).

The purpose of this review is to summarize current knowledge regarding sexual reproduction within the Ascomycota. This review will focus on sexual reproduction in terms of homothallism, heterothallism and pseudohomothallism, the two mating-type idiomorphs and their role in mating. The pheromone response system, in terms of pheromones and corresponding receptors, as well as the downstream signaling pathways will be considered in various model Ascomycota. The review concludes with an introduction of the focal taxon (the *Gibberella fujikuroi* species complex) of this

dissertation, a discussion of what is currently known regarding its mating and pheromone response and final mention of the objectives of this study.

2. Homothallism, heterothallism and pseudohomothallism

Sexual reproduction in fungi can be characterized as homothallic, pseudohomothallic or heterothallic (Kronstad, Staben 1997). Homothallic species have the ability to self-fertilize with their own mycelium, thus allowing them to complete the sexual cycle on their own without the need for an interacting partner (Glass, Smith 1994; Coppin et al. 1997). Pseudohomothallic fungus are also self-fertile, but this is due to the presence of two different mating-type nuclei in a single mycelium cell (Coppin et al. 1997; Poggeler 1999; Debuchy, Turgeon 2006). Heterothallic species cannot self-fertilize and mating can only occur between strains of opposite mating-type or strains with compatible mating specificities. Such strains do not differ morphologically but are both genetically (*i.e.*, they encode different sets of genes at their mating-type or *MAT* locus; see below) and physiologically (*e.g.*, genes encoding mating pheromones and pheromone receptors are differentially expressed) distinct (Alexopoulos, Mims, Blackwell 1996; Coppin et al. 1997).

3. Mating-type genes and their role in mating

Mating in Ascomycota is controlled by the mating-type (*MAT*) locus, which encodes an array of transcription factors that permit expression of genes required for mating including pheromones and their receptors (Pöggeler, Kück 2001; Casselton 2002). The genes and processes involved in mating of Ascomycota have been extensively studied in *Saccharomyces cerevisiae*, a heterothallic yeast species. The *MAT* locus can be occupied by one of two alternative idiomorphs, namely *MAT-**a** (MAT1-1)* and *MAT-**α** (MAT1-2)*, which have unrelated nucleotide sequences (Bölker, Kahmann 1993; Yun et al. 2000; Pöggeler, Kück 2001; Casselton 2002). These idiomorphs control the expression of genes that confer sexual identity of so-called **a**- and **α**-cells. The *MAT-**a*** idiomorph encodes for two proteins, *MATa1* and *MATa2* whereas the alternative idiomorph, *MAT-**α,*** encodes for *MATα1* and *MATα2* proteins (Bölker, Kahmann 1993; Pöggeler, Kück 2000; Yun et al. 2000; Casselton 2002).

The roles of *MATa1* and *MATa2* have not yet been identified in haploid cells where *MAT-**a*** specific genes are transcribed constitutively. However, in *MAT-**α*** cells transcription of *MAT-**α*** specific genes are controlled by the products of *MATα1* and *MATα2*. *MATα1* activates transcription of and **α**-specific gene set while *MATα2* represses transcription of the constitutively expressed **a**-specific gene set (Bölker, Kahmann 1993; Coppin et al. 1997; Yun et al. 2000; Casselton 2002). In these cells the *MATα1* protein and the minichromosomal maintenance protein, Mcm1p, function in activating the transcription of **α**-specific genes. *MATα2* encodes a repressor protein

which together with Mcm1p and other proteins suppresses **a**-cell specific genes (Bölker, Kahmann 1993; Casselton 2002). In diploid cells, *MATα2* forms a complex with the *MATa1* protein, which represses transcription of all genes that play vital roles during mating (Casselton 2002). The complex also has additional functions, including suppression of *MATα1* transcription and subsequent prevention of **α**-haploid cell-specific gene expression, as well as repressing the *rme1* gene whose product is an inhibitor of meiosis (Bender, Sprague-Jr 1989; Bölker, Kahmann 1993; Casselton 2002).

Homologs of *S. cerevisiae* mating-type idiomorphs have been identified in the model Ascomycota *Podospora anserina* and *Neurospora crassa* (Arie et al. 1997; Coppin et al. 1997; Kronstad, Staben 1997; Coppin, Debuchy 2000; Debuchy, Turgeon 2006). The *MAT* idiomorphs of these species and other filamentous Ascomycota are named differently (Bölker, Kahmann 1993; Coppin et al. 1997). Turgeon and Yoder (2000) proposed a standard *MAT* nomenclature for filamentous Ascomycota where *MAT1-1* and *MAT1-2* is equivalent to the yeast *MAT-**α*** and *MAT-**a*** idiomorphs, respectively.

The most studied and best understood mating system among filamentous Ascomycota is that of *Neurospora crassa* (Coppin et al. 1997; Dettman, Jacobson, Taylor 2003). Mating in this species is regulated by transcription factors *MAT1-1-1* and *MAT1-2-1*, expressed from *MAT1-1* and *MAT1-2*, respectively (Coppin et al. 1997; Turgeon, Yoder 2000). The *MAT1-1-1* gene encodes for an **α**-domain protein similar to the *S. cerevisiae* mating-type transcription activator *MAT*α1. This protein is the determinant of *MAT1-1*

mating-type specificity, but also plays a role post-fertilization such as ascospores formation (Chang, Staben 1994).

The *MAT1-2-1* polypeptide has a high mobility group (*HMG*) motif which allows this protein to bind to a specific DNA sequence and act as a transcription activator (Coppin et al. 1997; Turgeon, Yoder 2000). This gene was shown in various experiments to confer identity to *MAT1-2* cells. For example, deletion of the *MAT1-2-1* gene led to abolishment of *MAT1-2* identity (Coppin et al. 1997; Bobrowicz et al. 2002). Replacing the *MAT1-1* idiomorph with the *MAT1-2* idiomorph or only the *MAT1-2-1* Open Reading Frame (ORF), has also been shown to be enough to confer new mating-type identity from a *MAT1-1* to a *MAT1-2* cell (Chang, Staben 1994).

The *MAT1-1* idiomorph may also encode for two additional proteins, namely *MAT-1-2* and *MAT-1-3*, which respectively represent amphipathic **α**-helical and *HMG* box proteins (Coppin et al. 1997; Shiu, Glass 2000). Neither of these proteins are essential for mating and production of ascospores, although both have been found to increase the efficiency of mating and ascospores production (Coppin et al. 1997). A second ORF in the *MAT1-2* idiomorph was also identified in *Neurospora crassa* and has been named *MAT1-2-2,* but the putative polypeptide encoded by this gene has no known function in mating (Pöggeler, Kück 2000). Kanamori et al. (2007) identified an additional ORF within the *MAT1-2* locus of *Magnaporthe oryzae* (*MAT1-2-2*), although it is not known if this gene is homologous to *N. crassa MAT1-2-2* (Pöggeler 2001). Martin et al. (2011b)

identified yet another ORF within the *MAT1-2* locus (*MAT1-2-3*) of species within the *Gibberella fujikuroi* species complex. Additional MAT genes, MAT1-1-4, MAT1-1-5, MAT1-2-4 and MAT1-2-5 have also been identified in Ascomycetes species such as *Diplodia pinea* (Bihon et al. 2014) and *Botrytis cinerea* (Amselem et al. 2011).

4. The pheromone response system of Ascomycota

The pheromone response system involves the use of pheromones and corresponding receptors to initiate mating. This system allows Ascomycota to differentiate and grow towards cells of the opposite mating-type (Bölker, Kahmann 1993; Pöggeler, Kück 2001). Cells are able to do this because expression of the pheromones and receptors is under the control of the genes encoded at the mating-type locus (Coppin et al. 1997; Shiu, Glass 2000; Casselton 2002). The pheromones are excreted and bind to appropriate cognate receptors on the cell surface which leads to polarized growth and G1-cell cycle (the growth phase of interphase that occur before mitosis) arrest in *S. cerevisiae* (Bölker, Kahmann 1993; Naider, Becker 2004). The pheromone response pathway is also involved in transcriptional induction of a set of pheromone response genes and in *Schizosaccharomyces pombe* meiosis can only proceed when there is continued activation of this pathway (Bölker, Kahmann 1993; Casselton 2002).

4.1. Pheromone response in the model species *Saccharomyces cerevisiae*

Saccharomyces cerevisiae provides a framework for studying the pheromone response system (Kurjan 1993; Bradwell et al. 1994; Naider, Becker 2004). The two mating-types, *MAT-**a*** (*MAT1-2* in filamentous Ascomycota) and *MAT-**α*** (*MAT1-1* in filamentous Ascomycota), each confers for one type of pheromone and pheromone receptor (Turgeon 1998; Pöggeler, Kück 2001; Casselton 2002). *MAT-**a*** cells produce an **a**-factor pheromone encoded by two genes, *mfa1* and *mfa2* (mating factor **a**), which encode for a 34- and 36-amino acid precursor peptide, respectively (Bölker, Kahmann 1993; Casselton 2002). These precursors undergo post-translational modification to produce hydrophobic 12-amino-acid pheromones which differ at one position (Brake et al. 1985; Bölker, Kahmann 1993; Pöggeler, Kück 2001; Casselton 2002). The carboxyl terminals of these pheromone precursor peptides have CAAX motifs (where C stands for cysteine, A for any aliphatic amino acid; and X for any amino acid) which signals for prenylation and carboxymethylation of the cysteine as well as removal of the last three amino acids by proteolysis (Clarke 1992; Bölker, Kahmann 1993; Casselton 2002). Similarly, the **α**-factor pheromone is also encoded by a different set of alleles of the same genes (*mfα1* and *mfα2*), which encode for 165- and 120-amino acid precursor peptides, respectively (Kurjan, Herskowitz 1982; Singh et al. 1983; Pöggeler, Kück 2001; Bobrowicz et al. 2002). Each precursor contains tandem repeats of the **α**-factor

protein sequence bordered by protease processing sites that are cleaved to yield mature **α**-factor pheromones (Kurjan, Herskowitz 1982; Naider, Becker 2004).

The two pheromones are secreted by different pathways in *S. cerevisiae* cells. The **a**-factor is translocated by the *STE6* gene product, an ATP-binding cassette transporter (Kuchler, Sterne, Thorner 1989; Bölker, Kahmann 1993; Casselton 2002). The **α**-factor pheromone uses the classical yeast secretory pathway (Bölker, Kahmann 1993). Both pheromones however are known to activate the same transduction pathway (see below) (Pöggeler, Kück 2001; Bobrowicz et al. 2002).

Each pheromone binds to only one class of pheromone receptors which are found on the yeast cell surface. These receptors are named according to the pheromone they bind to whereas in filamentous species they are named based on the gene which they are expressed from (see below). The receptor for **a**-factor is expressed from the *ste3* (sterility) gene whereas *ste2* gene encodes an **α**-factor receptor (Mackay, Manney 1974; Hagen, McCaffrey, Sprague 1986; Pöggeler, Kück 2001; Bobrowicz et al. 2002). The primary sequences of the two receptors are unrelated although they have the same function and similar organization of functional domains.

The two receptors of *S. cerevisiae* are part of the family of G-protein coupled receptors (GPCRs) (Pöggeler, Kück 2001; Bobrowicz et al. 2002). The GPCRs are characterized by their extracellular N-terminal, intracellular C-terminal and typical 7 transmembrane

(TM) domains. These proteins are thus referred to as 7-TM receptor proteins. The 7-TM domains join three intra-cellular (IC) and extra-cellular (EC) loops with the third IC loop generally believed to bind to the G-protein (Wess 1997; Bockaert, Pin 1999). There has been evidence to suggest that the third EC loop is the target for **α**-factor in *S. cerevisiae* mating (Abel et al. 1998). The functions of the N-terminal and the C-terminal are not clearly understood although cells with truncated C-terminals show increased sensitivity of the receptor to the ligand for example pheromone peptides (Bradwell et al. 1994). Furthermore, the sixth (VI) TM segment contains a conserved proline residue which is involved in controlling the activity and trafficking of GPCRs (Konopka, Margarit, Dube 1996). The C-terminal of the 7-TM receptor contains multiple Ser-Thr residues which function as phosphoryl acceptors, and binding of appropriate pheromone ligands leads to an increased level of phosphorylation (Feng, Davis 2000).

The two pheromone receptors are capable of activating the same signal transduction pathway using the G-protein since they both belong to the GPCRs family of receptors (Bölker, Kahmann 1993; Pöggeler, Kück 2001). This pathway is activated by the secreted pheromones when they bind to the specific receptor, which leads to induction of the **α**-subunit of the G-protein to substitute the bound GDP with a GTP (Bölker, Kahmann 1993; Gustin et al. 1998; Bobrowicz et al. 2002; Casselton 2002). Binding by GTP causes the trimeric G-protein to breakup into the **α**-subunit and the βγ dimer (Bölker, Kahmann 1993). In most species the **α**-subunit bound to the Guanine triphosphate (GTP) interacts with a protein kinase to activate the transduction pathway,

after dissociation of the G-protein. However, in *S. cerevisiae* it has been shown that the βγ dimer plays that role and not the **α**-subunit (Leberer et al. 1992; Bölker, Kahmann 1993; Bobrowicz et al. 2002; Casselton 2002). Dissociation of G-protein is followed by a phosphorylation cascade involving various protein kinases and a transcription factor which eventually leads to physiological changes (Bölker, Kahmann 1993; Casselton 2002).

The first kinase to be activated is Fus3p, a mitogen activated protein (MAP), followed by three more kinases (*Ste5p*, *Ste11p* and *Ste7p*) which eventually transforms *Ste12p* transcription factor to an active state (Teague, Chaleff, Errede 1986; Leberer et al. 1992; Bölker, Kahmann 1993; Gustin et al. 1998; Casselton 2002; Kim, Borkovich 2004). This transcription factor is a member of the homeodomain proteins and it regulates the transcription of genes involved in mating by binding to a characteristic motif located on the promoter of these genes. *Ste12p* also enhances transcription in response to pheromone stimulation by binding Mcm1p on **a**- and **α**- gene-specific promoters which leads to pheromone signal augmentation (Gustin et al. 1998; Shiu, Glass 2000; Casselton 2002).

4.2. Pheromone response in the filamentous Ascomycota

A pheromone response system which is similar to that found in *S. cerevisiae*, has been identified in filamentous Ascomycota (Shiu, Glass 2000). Among these fungi, the best-studied system is that of *Neurospora crassa*, a heterothallic species commonly known as the red bread mould. Bistis (1981)observed polarized growth of trichogynes (specialized female receptive hyphae) towards conidia of the opposite mating-type. These results confirmed a previous suggestion that fusion between the “male cells” and the trichogyne was due to an attractant-receptor mechanism (Backus 1939). The attractants and the receptors have since been characterized in this species and others such as *Cryphonectria parasitica* and *Magnaporthe grisea*, as well as homothallic species such as *Sordaria macrospora* and *Gibberella zeae* (Zhang, Baasiri, Van Alfen 1998; Shen, Bobrowicz, Ebbole 1999; Pöggeler 2000; Bobrowicz et al. 2002; Kim, Borkovich 2006).

Like *S. cerevisiae*, heterothallic filamentous fungi have two classes of pheromones and receptors, and the genes for the two classes of pheromones and their receptors are present in strains of both mating-types and transcribed in a mating-type specific manner (Kim, Borkovich 2004; Mayrhofer, Weber, Pöggeler 2006). The pheromones and receptors also show homology and perform similar roles to those of *S. cerevisiae*,

showing that this system is important in heterothallic mating in the Ascomycota (Coppin et al. 1997; Shiu, Glass 2000; Casselton 2002).

4.2.1. Pheromones and receptors of the heterothallic *Neurospora crassa*

The pheromone response system in *N. crassa* involves two classes of pheromones whose expression is mating-type dependent. *MAT1-2* cells express only the *mfa-1* (mating factor expressed in *MAT-**a*** strain) pheromone gene whereas *MAT1-1* cells secrete only the *ccg-4* (clock controlled gene) pheromone gene (Bobrowicz et al. 2002). The polypeptides encoded by these genes show structural similarity to the **α**-factor and **a**-factor of *S. cerevisiae* (Kim, Borkovich 2006). The *N. crassa mfa-1* gene is predicted to transcribe for a polypeptide consisting of 24-amino acids and a CAAX motif at the C-terminal, which is also present in the *S. cerevisiae* **a**-factor. In yeast and other fungi, this motif signals for processing, resulting in a lipopeptide pheromone. In addition to the difference in sequence of Ccg-4 and Mfa-1 pheromone polypeptides, Ccg-4 is predicted to be synthesized from a precursor polypeptide made up of five identical repeats of the mature pheromone sequence. The precursor polypeptide is then cleaved by proteins into mature pheromones at specific processing sites similar to those in *S. cerevisiae*, with each repeat preceded by glutamine-alanine dipeptides. However, in *N. crassa* additional amino acids (Proline and Asparagine-Proline) are present in the fourth and fifth repeats between the glutamine-alanine dipeptide and the predicted pheromone peptide (Bobrowicz et al. 2002; Kim, Metzenberg, Nelson 2002; Kim, Borkovich 2006).

There have been very few studies on the pheromone receptors of filamentous Ascomycota and the majority of these studies were on homothallic species. Two putative pheromone receptor genes of *N. crassa* are *pre-1* and *pre-2* (Pöggeler, Kück 2001; Kim, Borkovich 2004; Kim, Borkovich 2006). These genes encode for receptors similar to the **a**-factor (*Ste3p*) and **α**-factor (*Ste2p*) receptors of *S. cerevisiae*, respectively. PRE-1 is expected to be the target for the hydrophobic pheromone MFA-1, while CCG-4 is the expected ligand of the PRE-2 receptor (Pöggeler, Kück 2001; Kim, Borkovich 2004). Both pheromone receptor genes, *pre-1* and *pre-2*, are present in the genome of *MAT1-1* and *MAT1-2* cells in single-copy and were found to be expressed in a mating-type independent manner (Pöggeler, Kück 2001). However, *MAT1-2* strains express the *pre-1* gene at much lower levels compared to *MAT1-1* strains, suggesting that *pre-1* may have additional functions to pheromone detection (Kim, Borkovich 2004).

The genes for the *N. crassa* pheromone receptors were first identified by Pöggeler and Kück (2001), who detected an ORF of 1265 base pairs (bp) that likely encodes a 402 amino acid protein and named it *pre-1*. Apart from being similar to the *S. cerevisiae Ste3p*, it also shares similarity with *MAP3* of *S. pombe* and the pheromone receptors of the Basidiomycota (Pöggeler, Kück 2001). A study by Kim, Borkovich (2004) however found that the *pre-1* ORF was instead 2307 bp long and contained two introns of 194 and 76 bp based on automatic gene annotation. When they performed RT-PCR analysis, they found similar results to those of Pöggeler, Kück (2001), with only one intron in the *pre-1* ORF but the transcript was approximately 3 kb in size, *i.e.*, almost

twice the size found by Pöggeler, Kück (2001). In that same study they showed that *pre-1* was critical in enabling *MAT1-1* females to identify and mate with *MAT1-2* male cells.

The second pheromone receptor of *N. crassa* was found to be encoded by a putative ORF of 1701 bp encoding a protein of 566 amino acids, which was named *pre-2* (Pöggeler, Kück 2001). Apart from being similar to Ste2p of *S. cerevisiae*, PRE-2 also shares similarity with the receptor of the homothallic species *Sordaria macrospora* (Pöggeler, Kück 2001). Interestingly, PRE-2 from *S. macrospora* has been transformed into *S. cerevisiae*, where it stimulated the pheromone response of *MAT-**a*** cells, with a mutated *ste2p* receptor gene, when activated by the *S. macrospora* peptide pheromone (Mayrhofer, Weber, Poggeler 2006). Since this receptor shows very high similarity to *N. crassa* PRE-2, it is expected that this protein is also activated by an appropriate pheromone and is involved in a G-protein transduction pathway in heterothallic species (Mayrhofer, Weber, Poggeler 2006).

Like Ste3p and Ste2p, PRE-1 and PRE-2 are predicted to be GPCR family proteins and thus have the characteristic 7 TM domains (Pöggeler, Kück 2001; Kim, Borkovich 2006). When amino acid sequences of *N. crassa* receptors and putative homologues from other species were aligned, they were found to be most similar at the TM helix regions (Pöggeler, Kück 2001). However, these comparisons showed that the sixth TM segment containing the conserved proline residue involved in controlling the activity and trafficking of GPCRs (Konopka, Margarit, Dube 1996) is present in both PRE-1 and

PRE-2 but its position is not conserved between the two proteins (Pöggeler, Kück 2001). Also, the Cytoplasmic C-terminal of the PRE-1 and PRE-2 receptors both have the multiple serine-threonine residues, which function as phosphoryl acceptors (Feng, Davis 2000). The pheromone receptor structures and functions thus seem to be conserved among Ascomycota (Pöggeler, Kück 2001; Kim, Borkovich 2004; Mayrhofer, Weber, Poggeler 2006; Lee, Leslie, Bowden 2008).

4.2.2. Pheromones and receptors of the homothallic *Gibberella zeae*

Among homothallic Ascomycota, pheromones and their receptors seem to play a limited role in mating but affect reproduction in various ways (Seo, Han, Yu 2004; Mayrhofer, Weber, Poggeler 2006). For example, *S. macrospora* with deleted pheromone and receptor genes were not able to produce ascospores (Mayrhofer, Weber, Poggeler 2006), while *Emaricella nidulans* (*Aspergillus nidulans*) double-knockout strains were not able to produce cleistothecia and ascospores (Seo, Han, Yu 2004). In *Gibberella zeae*, the pheromones-receptor pair (*ppg1/pre-2*) boost selfing and out-crossing whereas (*ppg2/pre-1*) pheromone-receptor pair does not show any immediate functions during sexual reproduction (Lee, Leslie, Bowden 2008).

Gibberella zeae (anamorph: *Fusarium graminearum*) is a homothallic species with all the mating-type genes clustered together (Yun et al. 2000; Lee, Leslie, Bowden 2008) and resemble the organization found in *S. macrospora* (Glass, Metzenberg, Raju 1990; Mayrhofer, Poggeler 2005). The pheromone and pheromone receptor genes of *G. zeae*

are similar to those characterized from other Ascomycota (Lee, Leslie, Bowden 2008). The putative pheromones are designated GzPPG1 and GzPPG2 whereas the receptors have been termed GzPRE1 and GzPRE2 (Kim, Lee, Yun 2008). Based on the amino acid sequence, GzPPG1 shows significant similarity to pheromones characterized from *S. macrosopra, N. crassa* and *C. parasitica* (Lee, Leslie, Bowden 2008). Expression analysis has shown that *Gzppg2* is not expressed, while *Gzpre1* is weakly expressed only at a certain growth stage. Therefore, the gene pair seems to be non-functional (Lee, Leslie, Bowden 2008). However *GzPR2* and *GzPPG1* may be involved in chemo-attraction of female cells by male cells in out-crossing mating (Lee, Leslie, Bowden 2008).

5. The focal group of this study – the *Gibberella fujikuroi* complex (GFC)

Fungal species previously accommodated in the *Fusarium* Section *Liseola* are now recognized as members of the *Gibberella fujikuroi* species complex (GFC) (Leslie, Summerell, Bullock 2006). Species in this complex have a wide geographic distribution and cause diseases in a wide variety of plant species (Leslie, Summerell, Bullock 2006; Kvas et al. 2009). They are well known and economically important because they are pathogenic to many crops such as maize, banana, mango, sorghum and sugarcane, but also forestry trees (Leslie, Summerell, Bullock 2006; Kvas et al. 2009). During growth, species in this complex may produce different secondary metabolites and mycotoxins

such as fumonisins (Rheeder, Marasas, Vismer 2002), gibberellic acid (Cerdá-Olmedo, Fernández-Martín, Ávalos 1994) and moniliformin (Marasas et al. 1986) that can cause cancer and birth defects if they are ingested with contaminated food or feed.

The GFC consists of species from 10 different biological species or mating populations (MP) that have been assigned letters A-J (*Fusarium verticillioides* (MP A), *F. sacchari* (MP B), *F. fujikuroi* (MP C), *F. intermedia* (MP D), *F. subglutinans* (MP E) subgroup 1 (now known as *F. temperatum* (Scauflaire, Gourgue, Munaut 2011)) and subgroup 2, *F. thapsinum* (MP F), *F. nygamai* (MP G), *F. circinatum* (MP H), *F. konza* (MP I) and *F. xylarioides* (MP J)) (Kuhlman 1982; Klittich, Leslie 1992; Klaasen, Nelson 1996; Klittich et al. 1997; Britz et al. 1999; Geiser et al. 2005; Lepoint, Munaut, Maraite 2005; Leslie et al. 2005;Kvas et al. 2009). Although these species are reproductively isolated (Taylor et al. 2000; Leslie, Summerell, Bullock 2006), there are exceptions where species from different mating populations are interfertile (Desjardins, Plattner, Gordon 2000; Steenkamp et al. 2001; Leslie et al. 2004). These include for example, *Gibberella fujikuroi* (*Fusarium fujikuroi*, MP C) and *Gibberella intermedia* (*Fusarium proliferatum*, MP D), which are morphologically very similar, but differ in some of the secondary metabolites they produce such as the levels of fumonosins and gibberellic acid (Tudzynski 1999; Rheeder, Marasas, Vismer 2002). Another example of such interspecies hybridization comes from *Fusarium subglutinans* (MP E) that are usually associated with maize and the pine pathogen *Fusarium circinatum* (MP H) (Desjardins, Plattner, Gordon 2000; Steenkamp et al. 2001). Such crosses between strains of

different mating populations under field conditions could result in progeny that produce unusual sets of secondary metabolites, or differences in their host range, which would complicate risk assessment of these strains in terms of mycotoxin production or pathogenicity (Leslie et al. 2004).

The GFC consists of many heterothallic species in which mating only occurs between strains of the opposite mating-types (Leslie, Summerell, Bullock 2006). Sexual reproduction in these species occurs by fusion between clearly defined male and female structures from different mating-types, similar to most heterothallic filamentous fungi (Leslie, Summerell, Bullock 2006). However, trichogynes have not been observed and thus, in these species the female structure are referred to as the protoperithecium, which is usually a differentiated cell that remains immature until fertilization by a male gamete. Sexual and asexual spores as well hyphal fragments can all act as the male structure to donate the male gametes (Leslie, Summerell, Bullock 2006).

Mating in species of the GFC is also controlled by the *MAT* locus. The two *MAT* idiomorphs, *MAT1-1* and *MAT1-2*, differ from *S. cerevisiae MAT* idiomorphs but they are very similar to those found in other filamentous species such as *Neurospora* and *Podospora*. *MAT1-1* encodes for three peptides, *MAT1-1-1*, *MAT1-1-2* and *MAT1-1-3*, whereas *MAT1-2* encodes for one polypeptide, *MAT1-2-1* and *MAT1-2-3* (Martin et al. 2011b). *MAT1-1-1* and *MAT1-2-1* are similar to *MATα1* and *MATa1* of *S. cerevisiae*, respectively (Arie et al. 1997; Coppin et al. 1997; Turgeon 1998). The other two

peptides encoded by *MAT1-1*, namely *MAT1-1-2* and *MAT1-1-3*, have homologous sequences in *N. crassa* and are expected to have the same functions (Coppin et al. 1997; Turgeon 1998), whereas *MAT1-2-3* could be unique to the Hypocreales (Martin et al. 2011b).

Many aspects of sexual reproduction are not well understood in the GFC, even though this process is essential for evolution and diversity. Understanding processes involved in sexual reproduction would be essential since species in this complex are found on all continents and they infect agricultural, forestry as well as ornamental plants (Leslie, Summerell, Bullock 2006; Kvas et al. 2009). One of the most vital processes in sexual reproduction of heterothallic species is the pheromone response system (Bistis 1981; Bistis 1983; Bender, Sprague-Jr 1989; Bölker, Kahmann 1993; Kim, Borkovich 2004). Therefore, characterization of the pheromone response system may provide novel insights into the initiation of mating, as well as species hybridization and the generation of genetic diversity in the GFC. Furthermore, sexual reproduction is also the basis of the biological species concept (Taylor et al. 2000; Kvas et al. 2009), which is often used in the grouping and identification of fungal species (Leslie, Summerell, Bullock 2006).

A recent study sequenced and characterized the two classes of mating pheromones in the GFC (Martin et al. 2011a). The overall objective of this study therefore to explore the notion that interactions between pheromones and pheromone receptors are responsible for the establishment of pre-zygotic species barrier during sexual recognition. Martin et

al. (2011a), however, showed that the genes of *Fusarium* species in the GFC encode the same set of mating peptide pheromones and did not detect significant differences in their amino acid sequences.

The aim of this study was to sequence, characterize and compare the pheromone receptor genes of *Fusarium* species from the GFC and of genes from other Sordariomycetes. This was done to study the forces driving the evolution of these genes and to evaluate whether the receptors could confer species-specificity amongst the different mating populations. It is hypothesized that genes governing sexual identity and mate detection show high degree of interspecies divergence, which is expected to play role in speciation (Swanson, Vacquier 2002a; Swanson, Vacquier 2002b; Dettman, Anderson, Kohn 2008). If the receptors are essential for mate recognition it could be expected that they are more conserved in species that can mate with each other than between non-mating species and the receptors could potentially also be more conserved between species that are inter-fertile than in species that are not.

In addition to the aims addressed in this study, future work should also seek to explore additional roles of the receptors in pre- and post-zygotic barriers in more detail. For example, effect of the pheromone repeats as well as expression levels of pheromone and receptor genes on mating. Finally, gene knock-outs studies of both pheromones and receptors should be conducted. Only by exploring these various avenues will it be

possible to fully understand the role that pheromones and pheromone receptors play in the biology of species in the GFC and Ascomycota as a whole.

6. References

Abel, MG, BK Lee, F Naider, JM Becker. 1998. Mutations affecting ligand specificity of the G-protein-coupled receptor for the *Saccharomyces cerevisiae* tridecapeptide pheromone. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1448:12-26.

Alexopoulos, CJ, CW Mims, M Blackwell. 1996. Introductory mycology. John Willey and Sons. Inc., New York 869.

Amselem, J, CA Cuomo, JAL van Kan, M Viaud, EP Benito, A Couloux, PM Coutinho, RP de Vries, PS Dyer, S Fillinger. 2011. Genomic analysis of the necrotrophic fungal pathogens Sclerotinia sclerotiorum and Botrytis cinerea. PLOS Genetics 7:1-27.

Arie, T, SK Christiansen, OC Yoder, BG Turgeon. 1997. Efficient cloning of Ascomycete mating-type genes by PCR amplification of the conserved MAT HMG box. Fungal Genetics and Biology 21:118-130.

Backus, MP. 1939. The mechanics of conidial fertilization in *Neurospora sitophila*. Bulletin of the Torrey Botanical Club 66:63-76.

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

Bihon, W, MJ Wingfield, B Slippers, TA Duong, BD Wingfield. 2014. MAT gene idiomorphs suggest a heterothallic sexual cycle in a predominantly asexual and important pine pathogen. Fungal Genetics and Biology 62:55-61.

Billiard, S, M López-Villavicencio, ME Hood, T Giraud. 2012. Sex, outcrossing and mating-types: unsolved questions in fungi and beyond. Journal of Evolutionary Biology 25:1020-1038.

Bistis, GN. 1981. Chemotropic interactions between trichogynes and conidia of opposite mating-type in *Neurospora crassa*. Mycologia 73:959-975.

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

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

Bockaert, J, JP Pin. 1999. Molecular tinkering of G protein-coupled receptors: an evolutionary success. The EMBO Journal 18:1723-1729.

Bölker, M, R Kahmann. 1993. Sexual pheromones and mating responses in fungi. The Plant Cell 5:1461-1469.

Bradwell, L, JG Cook, CJ Inouye, J Thorner. 1994. Signal propagation and regulation in the mating pheromone response pathway of the yeast *Saccharomyces cerevisiae*. Developmental Biology 166:363-379.

Brake, AJ, C Brenner, R Najarian, P Laybourn, J Merryweather. 1985. Structure of genes encoding precursors of the yeast peptide mating pheromone a-factor. Protein transport and secretion. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Britz, H, TA Coutinho, MJ Wingfield, WFO Marasas, TR Gordon, JF Leslie. 1999. *Fusarium subglutinans f. sp. pini* represents a distinct mating population in the *Gibberella fujikuroi* species complex. Applied and Environmental Microbiology 65:1198-1201.

Burdon, JJ, AP Roelfs. 1985. The effect of sexual and asexual reproduction on the isozyme structure of populations of *Puccinia graminis*. Phytopathology 75:1068-1073.

Burnett, JH. 1968. Fundamentals of mycology: Edward Arnold London.

Casselton, LA. 2002. Mate recognition in fungi. Heredity 88:142-147.

Casselton, LA, NS Olesnicky. 1998. Molecular genetics of mating recognition in Basidiomycete fungi. Microbiology and Molecular Biology Reviews 62:55-70.

Cerdá-Olmedo, E, R Fernández-Martín, J Ávalos. 1994. Genetics and gibberellin production in *Gibberella fujikuroi*. Antonie van Leeuwenhoek 65:217-225.

Chamberlain, M, DS Ingram. 1997. The balance and interplay between asexual and sexual reproduction in fungi. Advances in Botanical Research 24:71-87.

Chang, S, C Staben. 1994. Directed replacement of *mat A* by *mat a-1* effects *a* mating-type switch in *Neurospora crassa*. Genetics 138:75-81.

Clarke, S. 1992. Protein isoprenylation and methylation at carboxyl-terminal cysteine residues. Annual Review of Biochemistry 61:355-386.

Coppin, E, R Debuchy. 2000. Co-expression of the mating-type genes involved in internuclear recognition is lethal in *Podospora anserina*. Genetics 155:657-669.

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

Debuchy, R, BG Turgeon. 2006. Mating-type structure, evolution, and function in Euascomycetes. The Mycota: Growth, Differentiation and Sexuality:293-323.

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

Dettman, J, J Anderson, L Kohn. 2008. Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus *Neurospora*. BMC Evolutionary Biology 8:35-48.

Dettman, JR, DJ Jacobson, JW Taylor. 2003. A multilocus genealogical approach to phylogenetic species recognition in the model eukaryote *Neurospora*. Evolution 57:2703-2720.

Dyer, PS, DS Ingram, K Johnstone. 1992. The control of sexual morphogenesis in the Ascomycotina. Biological Reviews 67:421-458.

Feng, Y, NG Davis. 2000. Akr1p and the type I casein kinases act prior to the ubiquitination step of yeast endocytosis: Akr1p is required for kinase localization to the plasma membrane. Molecular and Cellular Biology 20:5350-5355.

Fraser, JA, J Heitman. 2004. Evolution of fungal sex chromosomes. Molecular Microbiology 51:299-306.

Geiser, DM, ML Lewis Ivey, G Hakiza, JH Juba, SA Miller. 2005. *Gibberella xylarioides* (anamorph: *Fusarium xylarioides*), a causative agent of coffee wilt disease in Africa, is a previously unrecognized member of the *G. fujikuroi* species complex. Mycologia 97:191-201.

Glass, NL, RL Metzenberg, NB Raju. 1990. Homothallic Sordariaceae from nature: The absence of strains containing only the *a* mating-type sequence. Experimental Mycology 14:274-289.

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

Gustin, MC, J Albertyn, M Alexander, K Davenport. 1998. MAP kinase pathways in the yeast *Saccharomyces cerevisiae*. Microbiology and Molecular Biology Reviews 62:1264-1300.

Hagen, DC, G McCaffrey, GF Sprague. 1986. Evidence the yeast *STE3* gene encodes a receptor for the peptide pheromone a factor: gene sequence and implications for the structure of the presumed receptor. Proceedings of the National Academy of Sciences 83:1418-1422.

Heitman, J. 2006. Sexual reproduction and the evolution of microbial pathogens. Current Biology 16:R711-R725.

Horn, BW, JH Ramirez-Prado, I Carbone. 2009. Sexual reproduction and recombination in the aflatoxin-producing fungus *Aspergillus parasiticus*. Fungal Genetics and Biology 46:169-175.

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

Kim, HK, T Lee, SH Yun. 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, KA Borkovich. 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, KA Borkovich. 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, RL Metzenberg, MA Nelson. 2002. Multiple functions of *mfa-1*, a putative pheromone precursor gene of *Neurospora crassa*. Eukaryotic Cell 1:987-999.

Klaasen, JA, PE Nelson. 1996. Identification of a mating population, *Gibberella nygamai* sp. nov., within the *Fusarium nygamai* anamorph. Mycologia 88:965-969.

Klittich, CJR, JF Leslie. 1992. Identification of a Second Mating Population within the *Fusarium moniliforme* Anamorph of *Gibberella fujikuroi*. Mycologia 84:541-547.

Klittich, CJR, JF Leslie, PE Nelson, WFO Marasas. 1997. *Fusarium thapsinum* (*Gibberella thapsina*): a new species in section Liseola from sorghum. Mycologia 89:643-652.

Konopka, JB, SM Margarit, P Dube. 1996. Mutation of Pro-258 in transmembrane domain 6 constitutively activates the G protein-coupled *alpha-factor* receptor. Proceedings of the National Academy of Sciences 93:6764-6769.

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

Kuchler, K, RE Sterne, J Thorner. 1989. *Saccharomyces cerevisiae* STE6 gene product: a novel pathway for protein export in eukaryotic cells. The EMBO Journal 8:3973-3984.

Kuhlman, EG. 1982. Varieties of *Gibberella fujikuroi* with anamorphs in *Fusarium* section Liseola. Mycologia 74:759-768.

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

Kurjan, J, I Herskowitz. 1982. Structure of a yeast pheromone gene (*MF alpha*): A putative *alpha-factor* precursor contains four tandem copies of mature *alpha-factor*. Cell 30:933-943.

Kvas, M, WFO Marasas, BD Wingfield, MJ Wingfield, ET Steenkamp. 2009. Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. Fungal Diversity 34:1-21.

Leberer, E, D Dignard, D Harcus, DY Thomas, M Whiteway. 1992. The protein kinase homologue *Ste20p* is required to link the yeast pheromone response G-protein beta gamma subunits to downstream signalling components. The EMBO Journal 11:4815-4824.

Lee, J, JF Leslie, RL Bowden. 2008. Expression and function of sex pheromones and receptors in the homothallic Ascomycete *Gibberella zeae*. Eukaryotic Cell 7:1211-1221.

Lepoint, PCE, FTJ Munaut, HMM Maraite. 2005. *Gibberella xylarioides sensu lato* from *Coffea canephora*: a new mating population in the *Gibberella fujikuroi* species complex. Applied and Environmental Microbiology 71:8466-8471.

Leslie, JF, BA Summerell, S Bullock. 2006. The *Fusarium* laboratory manual: Wiley Online Library.

Leslie, JF, BA Summerell, S Bullock, FJ Doe. 2005. Description of *Gibberella sacchari* and neotypification of its anamorph *Fusarium sacchari*. Mycologia 97:718-724.

Leslie, JF, K Zeller, M Wohler, B Summerell. 2004. Interfertility of two mating populations in the *Gibberella Fujikuroi* species complex. European Journal of Plant Pathology 110:611-618.

Mackay, V, TR Manney. 1974. Mutations affecting sexual conjugation and related processes in *Saccharomyces cerevisiae*. I. isolation and phenotypic characterization of nonmating mutants. Genetics 76:255-271.

Marasas, WFO, PG Thiel, CJ Rabie, PE Nelson, TA Toussoun. 1986. Moniliformin production in *Fusarium* section Liseola. Mycologia 78:242-247.

Martin, SH, BD Wingfield, MJ Wingfield, ET Steenkamp. 2011a. Causes and consequences of variability in peptide mating pheromones of ascomycete fungi. Molecular Biology and Evolution 28:1987-2003.

Martin, SH, BD Wingfield, MJ Wingfield, ET Steenkamp. 2011b. Structure and evolution of the *Fusarium* mating-type locus: New insights from the *Gibberella fujikuroi* complex. Fungal Genetics and Biology 48:731-740.

Mayrhofer, S, S Poggeler. 2005. Functional characterization of an alpha-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, JM Weber, S Poggeler. 2006. Pheromones and pheromone receptors are required for proper sexual development in the homothallic Ascomycete *Sordaria macrospora*. Genetics 172:1521-1533.

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

Poggeler, 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. 2001. Mating-type genes for classical strain improvements of Ascomycetes. Applied Microbiology and Biotechnology 56:589–601.

Pöggeler, S, U Kück. 2000. Comparative analysis of the mating-type loci from *Neurospora crassa* and *Sordaria macrospora*: identification of novel transcribed ORFs. Molecular and General Genetics 263:292-301.

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

Rheeder, JP, WFO Marasas, HF Vismer. 2002. Production of fumonisin analogs by *Fusarium* species. Applied and Environmental Microbiology 68:2101-2105.

Scauflaire, J, M Gourgue, F Munaut. 2011. *Fusarium temperatum sp. nov*. from maize, an emergent species closely related to *Fusarium subglutinans*. Mycologia 103:586-597.

Seo, JA, KH Han, JH Yu. 2004. The *gprA* and *gprB* genes encode putative G protein-coupled receptors required for self-fertilization in *Aspergillus nidulans*. Molecular Microbiology 53:1611-1623.

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

Shiu, PKT, NL Glass. 2000. Cell and nuclear recognition mechanisms mediated by mating-type in filamentous Ascomycetes. Current Opinion in Microbiology 3:183-188.

Singh, A, EY Chen, JM Lugovoy, CN Chang, RA Hitzeman, PH Seeburg. 1983. *Saccharomyces cerevisiae* contains two discrete genes coding for the alpha-factor pheromone. Nucleic Acids Research 11:4049-4063.

Steenkamp, ET, TA Coutinho, AE Desjardins, BD Wingfield, WFO Marasas, MJ Wingfield. 2001. *Gibberella fujikuroi* mating population E is associated with maize and teosinte. Molecular Plant Pathology 2:215-221.

Swanson, WJ, VD Vacquier. 2002a. The rapid evolution of reproductive proteins. Nature Reviews Genetics 3:137-144.

Swanson, WJ, VD Vacquier. 2002b. Reproductive protein evolution. Annual Review of Ecology and Systematics 33:161-179.

Taylor, JW, DJ Jacobson, S Kroken, T Kasuga, DM Geiser, DS Hibbett, MC Fisher. 2000. Phylogenetic species recognition and species concepts in fungi. Fungal Genetics and Biology 31:21-32.

Teague, MA, DT Chaleff, B Errede. 1986. Nucleotide sequence of the yeast regulatory gene *STE7* predicts a protein homologous to protein kinases. Proceedings of the National Academy of Sciences 83:7371-7375.

Tudzynski, B. 1999. Biosynthesis of gibberellins in *Gibberella fujikuroi*: biomolecular aspects. Applied Microbiology and Biotechnology 52:298-310.

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

Turgeon, BG, OC Yoder. 2000. Proposed nomenclature for mating type genes of filamentous Ascomycetes. Fungal Genetics and Biology 31:1-5.

Wess, J. 1997. G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition. The FASEB Journal 11:346-354.

Wingfield, MJ, ZW De Beer, B Slippers, BD Wingfield, JZ Groenewald, L Lombard, PW Crous. 2012. One fungus, one name promotes progressive plant pathology. Molecular Plant Pathology 13:604-613.

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

Zhang, L, RA Baasiri, NK Van Alfen. 1998. Viral repression of fungal pheromone precursor gene expression. Molecular and Cellular Biology 18:953.

Chapter 2

Phylogeny and evolution of pheromone receptor genes of *Fusarium* species in the *Gibberella fujikuroi* species complex

Introduction

The phylum Ascomycota contains many pathogenic fungal genera, including *Fusarium*. This genus includes agriculturally and economically important pathogens with a global distribution and mycotoxin producing abilities. Examples include *Fusarium circinatum*, a pine pathogen that causes pitch canker, and *F. verticillioides*, a maize pathogen that produces carcinogenic mycotoxins (D'mello, Placinta, Macdonald 1999; Placinta, D'mello, Macdonald 1999; Leslie, Summerell, Bullock 2006; Kvas et al. 2009). These species form part of the *Gibberella fujikuroi* species complex, which is a large group of *Fusarium* species such as *F. subglutinans*, *F. proliferatum* and *Fusarium intermedia* that have diverged relatively recently (O'Donnell, Cigelnik, Nirenberg 1998; Steenkamp et al. 1999; O'Donnell et al. 2000; Kvas et al. 2009). Many of the fungi in this complex can produce both sexual and asexual reproductive structures (Taylor et al. 2000; Kvas et al. 2009).

Sexual species are grouped on the basis of mating assays, during which isolates that can mate and produce viable ascospores are recognized as members of a biological species, which is also known as a mating population (MP) (Leslie 1991; Leslie, Summerell, Bullock 2006). Ten such mating populations (MP A - J) currently exist within the GFC (Kvas et al. 2009), although a few laboratory studies have documented inter-fertility between *Fusarium fujikuroi* (MP C) and *Fusarium intermedia* (MP D) (Leslie et al. 2004), as well as between *Fusarium subglutinans* (MP E) and *Fusarium circinatum* (MP H) (Desjardins, Plattner, Gordon 2000). However, it is unclear whether

hybridization occurs in natural environments. The emergence of such inter-specific hybrids in nature is generally thought to be prevented by pre-zygotic and post-zygotic mating barriers between fungal species (Giraud et al. 2008). One possible pre-zygotic barrier in fungi is the pheromone response system, which is involved in inter- and intra-specific recognition during sexual reproduction (Bistis 1981; Bistis 1983; Bender, Sprague-Jr 1989; Chang, Staben 1994; Poggeler et al. 2000; Poggeler, Kuck 2001; Turner, Jacobson, Taylor 2010).

Ascomycete sexual reproduction is controlled by the mating-type idiomorphs (*MAT1-1* and *MAT1-2*), which encode for distinct regulatory proteins that control expression of pheromones and their associated receptor genes, as well as other genes involved in sexual reproduction (Chang, Staben 1994; Yun et al. 2000; Debuchy, Turgeon 2006). In heterothallic Ascomycetes, each haploid individual harbors one of the two idiomorphs and requires an opposite mating-type partner for sexual reproduction (Chang, Staben 1994; Kerényi et al. 1999; Turgeon, Yoder 2000; Yun et al. 2000). The pheromone response system is then used to locate a suitable mating partner, which is achieved by secreting pheromones into the environment that can be detected by compatible individuals through specific receptors on the cell surface (Bistis 1981; Bistis 1983; Bölker, Kahmann 1993; Kim, Borkovich 2006). This system has been linked to sexual incompatibility between *Ustilago maydis* and *Ustilago hordei* (Bakkeren, Kronstad 1996), as well as between heterothallic *Neurospora* species that are found in sympatry (Karlsson, Nygren, Johannesson 2008). In *Podospora anserina* and *Cryphonectria*

parasitica, mutations in the pheromone genes that affected expression levels lead to reduced fertility (Zhang et al. 1993; Coppin, Debuchy 2000; Coppin, De Renty, Debuchy 2005).

The pheromone and receptor genes have been extensively studied in the filamentous ascomycete *Neurospora crassa*. This species harbors two pheromone genes (*mfa-1* and *ccg-4*) and cognate receptor genes (*pre-1* and *pre-2*) in each genome. *MAT1-1* strains are sexually stimulated by *MFA-1* when it binds to PRE-1 receptors on their cell surface, while CCG-4 activates PRE-2 on the surface of *MAT1-2* cells (Bobrowicz et al. 2002; Kim, Borkovich 2004; Kim, Borkovich 2006). Homologous genes have been identified in yeast and other filamentous ascomycetes, including *Fusarium* species (Bölker, Kahmann 1993; Mayrhofer, Weber, Pöggeler 2006; Kim, Lee, Yun 2008; Lee, Leslie, Bowden 2008). Recently, the pheromone genes of species in the GFC were characterized and showed high levels of similarity among their amino acid sequences (Martin et al. 2011). The authors noticed that the pheromones were expressed in tandem repeats that differed in number between species. These genes were also shown to exhibit diversifying selection, which is common in reproductive genes involved in recognition (Yang, Swanson, Vacquier 2000; Swanson, Vacquier 2002a; Swanson, Vacquier 2002b; Karlsson, Nygren, Johannesson 2008).

Pheromone receptor genes have not yet been characterized in *Fusarium* species from the GFC, but they have previously been identified in *F. graminearum* and *F. oxysporum*

genomes (Lee, Leslie, Bowden 2008). These receptors are part of the G-protein coupled receptor (GPCRs) family which is characterized by 7 transmembrane (TM) spanning domains that are involved in specificity and activity as well as structural integrity and the helix-helix packaging of the receptor (Abel et al. 1998; Henry et al. 2002; Balasubramanian et al. 2005; Karlsson, Nygren, Johannesson 2008). The TM domains separate the intra-cellular (IC) and extra-cellular (EC) loops while the N-terminal and C-terminal are respectively located on the EC and IC side of the cell (Wess 1997; Bockaert, Pin 1999). The aim of this study was, therefore, to characterize the pheromone receptor genes of *Fusarium* species from the *Gibberella fujikuroi* species complex and to evaluate whether the receptors could confer species-specificity amongst the different mating populations.

Materials and methods

Fungal isolates and DNA extraction

A total of 24 *Fusarium* isolates were examined (Table 1). These included standard *MAT1-1* and *MAT1-2* mating-type tester strains from nine biological species (MP A to MP I). All mating populations corresponded to a single phylogenetic species, except for MP E, which consisted of two distinct phylogenetic species (referred to as groups 1 and 2) that are inter-fertile (Steenkamp et al. 2002). Note that *F. subglutinans* subgroup 1 has been named *F. temperatum* (Scauflaire, Gourgue, Munaut 2011) and this name is used to refer to *F. subglutinans* subgroup 1 in the current study. Moderate inter-fertility has also been shown between mating population C and D (*F. fujikuroi* and *F.*

proliferatum, respectively) (Leslie et al. 2004). Mating population E (*F. temperatum*) and mating population H (*F. circinatum*) have also been reported to be inter-fertile (Desjardins 2000). Two isolates from species for which the sexual state is not yet known, namely *F. mangiferae* and *F. sterilihyphosum*, were also included. All cultures are maintained in the culture collection (CMWF) of the Forestry and Agricultural Biotechnology Institute (FABI), University of Pretoria, South Africa. Genomic DNA was extracted from each isolate using a previously described protocol (Steenkamp et al. 1999).

PCR amplification and sequencing

Pheromone receptor genes *pre-1* and *pre-2* from species in the *Gibberella fujikuroi* complex were amplified using primers designed with Primer Designer 4.20 (Sci Ed Central, Cary, NC). For primer design the conserved regions flanking the pheromone receptor genes in *Fusarium verticilioides* and *Fusarium oxysporum* (GenBank accession numbers AAIM02000073 and AAXH01000548, respectively) were targeted. Due to the large size of the genes, both genes were amplified in two overlapping fragments, where both *pre-1* and *pre-2* were amplified using four primers (Table 2). Additional primers were used where amplification failed with the pre2-3/pre2-4 primer set (Table 3).

Polymerase chain reaction (PCR) mixtures contained ~5 ng/μl DNA, 0.3 μM of each primer, 250 μM dNTPs (Fermentas, Nunningen, Switzerland), 0.04 U/μl *Taq* DNA polymerase (Roche Molecular Biochemicals, Manheim, Germany) and PCR buffer with

$MgCl_2$ (Roche) in a final volume of 25 μl. The PCR cycling conditions consisted of initial denaturation at 94°C for 5 min followed by 35 cycles of denaturation at 94°C for 35 s, annealing for 35 s and extension at 72°C for 60 s, and concluding with a final extension step at 72°C for 5 min. All primer sets had an annealing temperature of 50°C except for pre2-3/pre2-4 and the species specific primers, which had an annealing temperature of 55°C. PCR products were purified by polyethylene glycol precipitation (Gargas, Taylor 1992) and sequenced in both directions with the original PCR primers using standard dye terminator cycle sequencing (Applied Biosystems, Foster City,CA) and an ABI PRISM 3700 DNA sequencer (Applied Biosystems). Raw sequences were visualized and edited with Chromas Lite 2.01 (Technelysium).

Phylogenetic analyses

Two phylogenetic trees (i.e., one tree for each pheromone receptor gene from *Fusarium* species) were constructed using the coding regions (exons). The sequences were aligned using MAFFT 5.85 (http://align.bmr.kyushu-u.ac.jp/mafft/online/server/) and included *F. oxysporum* sequences for out-group purposes. The best-fitting substitution model for the data was calculated using the Akaike Information Criterion in jModelTest 0.1.1 (Posada 2008). Phylogenetic relationships were inferred with PhyML version 3.0 (Guidon, Gascuel 2003) using maximum likelihood (ML) and the best-fit models obtained from jModelTest 0.1.1, and viewed using Mega 5.0 software (Tamura et al. 2007). Confidence in branches was obtained from 1000 bootstrap replicates.

Sequence variation and codon selection

Amino acid sequences from both genes were aligned using MAFFT 5.85 (http://align.bmr.kyushu-u.ac.jp/mafft/online/server/). This was done to evaluate the similarity among the products coded for the by the respective genes. For these analyses, the different domains of the receptors as identified by making use of TOPCONS (http://topcons.cbr.su.se/).

The ratio of non-synonymous and synonymous nucleotide substitutions was used to test for positive selection in the *pre-1* and *pre-2* genes of *Fusarium*. Selection would result in either amino acid changes (positive selection), or conservation of an amino acid at a specific position in a polypeptide (negative/purifying selection) (Swanson, Vacquier 2002b; Martin et al. 2011a).we utilized two set of analyses, Nielsen, Yang (1998) implemented in HyPhy and various methods of detecting positive selection implemented in Datamonkey. The forma method is widely employed while Data monkey makes use of the counting-based method, random effect models and fixed effect models. The three approaches have different upsides and downfalls for instance random effect methods have a high likelihood to give false positives in an analysis (Pond, Frost 2005b).

Positively selected sites within the pheromone receptor genes were identified using various codon-based maximum likelihood methods from Datamonkey (http://www.datamonkey.org/), namely Fixed Effect Likelihood (FEL), Random Effect Likelihood (REL) and Single Likelihood Ancestor Counting (SLAC), A Fast,

Unconstrained Bayesian AppRoximation for inferring selection (FUBAR) and Mixed Effect Model of Evolution (MEME) (Pond, Frost 2005a). The FEL method applies a likelihood ratio test to each codon site to determine the best-fitting substitution model, whereas REL searches for the best-fitting model for the entire gene, followed by a Bayesian approach to classify each site into a dN/dS class. The SLAC algorithm reconstructs ancestral codons using maximum likelihood and then detects selection based on observed vs expected substitutions. FUBAR is a faster alternative method that uses a site-to-site rate variation model and Bayesian Markov chain Monte Carlo (MCMC) to identify selected sites. MEME is a branch-site random effect model which detects episodic and widespread positive selection by allowing ω (dN/dS) to vary between sites as well as from branch to branch (Pond, Frost 2005a; Pond, Muse 2005; Delport et al. 2010). Default settings and the HKY85 nucleotide substitution model (Hasegawa, Kishino, Yano 1985) were used for all analyses, and results from all five methods were compared to obtain a consensus. For our study we only sites that could be consistently identified by at least three of the six methods (five from Datamonkey plus Nielsen-Yang) for further analyses.

The Nielsen-Yang maximum likelihood method was also performed using the HyPhy package (Pond, Muse 2005) to determine the type of selection acting on the *pre-1* and *pre-2* sequences, and to identify sites under positive selection. This method employs a likelihood ratio and the Bayes theorem to test for positive selection across a gene. These analyses involved hypothesis testing using LRTs, where the null model which

does not allow sites to be positively selected (ω>1) was compared with a broader model that does (Yang 2000; Pond, Muse 2005; Yang 2007). We compared three model pairs, M0 (single rate) vs M3 (discrete), M1 (neutral) vs M2 (selection), and M7 (β) vs M8 (β + ω) (Table 2) using the LRT statistics formula $2\Delta L=2$ ($\log H_1$ - $\log H_0$) with a χ^2 distribution. A Bayesian analysis was subsequently used to calculate the posterior probability of each site belonging to a certain selection class.

Results

Phylogenetic analyses

Phylogenetic analyses of the two pheromone receptor genes, *pre-1* and *pre-2*, separated 12 *Fusarium* species in the GFC into three clades (Figure 1, Figure 2), which were consistent with the so-called biogeographic clades (African, American and Asian) known for the complex (O'Donnell, Cigelnik, Nirenberg 1998). The African clade consisted of *F. verticilllioides* (MP A), *F. thapsinum* (MP F) and *F. nygamai* (MP G), whereas the Asian clade consisted of *F. sacchari* (MP B)*, F. fujikuroi* (MP C), *F. intermedia* (MP D) and *F. mangiferae. Fusarium subglutinans* and *F. temperatum* (MP E), *F. circinatum* (MP H), *F. konza* (MP I) as well as *F. sterilihyphosum* grouped together to form the American clade. The two mating-types from each biological species formed close sister lineages.

The tree constructed from *pre-2* sequences, however, was incongruent with *pre-1* gene tree. In the *pre-1* gene tree, the African clade was basal to the Asian and American clades, whereas the *pre-2* tree revealed the American clade as basal, while the African and Asian clades formed sister lineages. However, both trees had low bootstrap support at the basal nodes.

Sequence variation and codon selection

Amino acid alignments revealed sequence variation in predicted PRE-1 and PRE-2 proteins amongst *Fusarium* species belonging to the GFC (Figure 3). These polymorphisms were spread across and located within different domains of the predicted receptor peptide. The polymorphism was due to amino acid changes in a single isolate, single species or multiple species. However, there were substitutions that were conserved within clades as well as substitutions that were shared by species from different clades. The majority of the polymorphism were shared within clades, although the predicted protein sequence for PRE-1 included more polymorphic sites (38) compared to PRE-2 (28) (Figure 3).

All the domains of the PRE-1 receptor had at least one polymorphic amino acid residue, the majority of which were in the in the C-terminal followed by the N-terminal domain (Figure 3a - c). The PRE-1 receptor also contained amino acid variants that were limited to species from the same clade and some that were shared by species from different clades. These PRE-1 receptor sequence alignment also contained a high number of

Single amino acid polymorphism (SAAP) that were due to changes in only one species and a few in only one isolate (Figure 3a). The C-terminal had seven SAAPs due to changes in a single species while the N-terminal domain had six (Figure 3a & c). In addition both domains had variants conserved amongst species from the African clade, of which two were in the C-terminal and one was in the N-terminal. The amino acid variant in the N-terminal was however also observed in one of the Asian species, *F. mangiferae* at the same codon site (Figure 3a). Similar results were also observed in the first (I) EC loop (codon 100) and the third (III) IC loop (codon 219) where an amino acid variant conserved in the African clade was shared with a species from the Asian clade in this case *F. sacchari* (Figure 3a & b). The sixth (VI) TM domain (codon 256) and the third (III) EC loop (codon 297and 299) had variants that were shared by some but not all species from the Asian clade (Figure 3c). We also found variants in PRE-1 that were only shared by isolates of *F. fujikuroi* and *F. proliferatum* as well as *F. fujikuroi* with *F. sacchari* species (Figure 3a-c).

The PRE-2 alignment contained a Single amino acid polymorphism (SAAP) due to amino acid change in a single isolate (codon 83) that was located in the second (II) TM domain (Figure 4a). Several SAAPs were limited to specific species and these were found in the second (II) EC loop (codon 179), third (III) IC loop (codon 234) and the C-terminal domain (codon 293, 298, 306, 330, 345 and 354) (Figure 4a-c). Variation at codon sites 63 and 362, located in the first (I) IC loop and C-terminal respectively, were due to changes in sequences of species from different clades with species in the Asian

clade being the least conserved in both sites (Figure 4a & c). We identified SAAPs where an amino acid variant was limited to species from a single clade, i.e., Asian for codon 76 and African for codon 134 in the first (I) CL loop and third (III) TM respectively (Figure 4a). The fourth (IV) TM had an amino acid variant at codon 156 only present in *F. thapsinum* and *F. mangiferae* while in the third (III) IC loop all species in the Asian clade except *F. saccharin* shared a unique variant (Figure 4b).

The exon sequences for both the *pre-1* and *pre-2* genes from the *Fusarium* isolates (Table 1) were used to determine which selection forces were responsible for the observed amino acid changes. The consensus results of FEL, REL, FUBAR, MEME and SLAC revealed three sites under positive selection in both *pre-1* (codon 100, 219 and 292) and *pre-2* (codon 112, 181 and 309) genes (Table 4). Most of these codons were also identified by the Nielsen-Yang method ($P \leq 0.01$), namely codon 100 and 219 for *pre-1* and codon 112 for *pre-2* (Table 4). Additional sites in *pre-2* that were identified by two of the five methods from Datamonkey, namely codon 83 (identified by SLAC and MEME) and codon 156 (identified by REL and FEL), were also identified with the Nielsen-Yang method ($P \leq 0.01$) and considered for further analyses (Table 5). In total, *pre-1* had five sites under positive selection with three of these located in the third extracellular loop, while the remaining two were each located in the third intracellular and first extracellular loops (Figure 5b). Positively selected sites were scattered through the *pre-2* gene. Two of these sites were located in the second and fourth TM regions, while the first and second extracellular loop each contained a single site under positive

selection (Figure 5a). Two of the positive selected sites were found in the cytoplasmic side of the cell in the first loop and the C-terminal domain (Figure 5a).

Hypothesis testing of site models was performed on *pre-1* and *pre-2* genes using the Nielsen-Yang method implemented in the HyPhy package (Pond, Muse 2005). Results from the Nielsen-Yang method showed that *pre-1* and *pre-2* were under natural/purifying selection with dN/dS (ratio of nonsynonymous substitutions/ synonymous substitutions) values lower than one ($\omega < 1$) averaged over all codons (Table 5). Comparison of site models M1 (neutral) vs M2 (selection) and M7 (β) vs M8 ($\beta + \omega$) yielded values of 0.14 and 2.58 respectively, which did not allow for rejection of the null hypothesis of neutral selection in *pre-1*(Table 5). Similar results were also obtained when model M1 vs M2 were compared for the *pre-2* gene (Table 5). When we compared M7 vs M8, the results showed significant difference between the two methods, rejecting the null hypothesis in favor of positive selection acting on *pre-2* with LRT=20.9, which is significant at $P \leq 0.01$ (Table 5). The difference in results between the two pairs of hypothesis tests may be due to M1 vs M2 models being more conservative compared to M7 and M8. Positive selected sites were identified in both genes the method of Nielsen-Yang (1998), however the highest rate class was $\omega=1$ (neutral selection) in all selection models, with the exception of model M2 in *pre-1* where 6% of the sites had $\omega=1.22$ (positive selection) (Table 5).

Discussion

The results of this study showed that pheromone receptor genes have evolved in a mating-specific manner amongst *Gibberella fujikuroi* complex species. This is in contrast to the evolution of their ligands, the mating pheromones (Martin et al. 2011). In the pheromone receptor genes this evolution has led to clade- and species-specific amino acid changes with the majority of these located in the intracellular C-terminal domain of the receptors. The pheromone genes on the other hand do not show the same species and clade-specificity. Instead only one species *F. konzum* had species-specific amino acid changes in one of its pheromone genes, *ppg1*. The rest of the species had an identical sequence encoding for *ppg2* which consists of eight-residue repeats that make up the mature pheromone. These repeats are conserved within species through various methods and important for pheromone-receptor sequence specificity (Martin et al. 2011). Ppg1 on the other hand consisted of multiple copies of at least two variants that had minimal differences in amino acid sequence between each GFC species (Martin et al. 2011).

Phylogenies of *pre-1* and *pre-2* genes revealed three clades, namely the African, Asian and American clades. This finding is in agreement with previously determined species tree topologies for this group of fungi (O'Donnell, Cigelnik, Nirenberg 1998; Kvas et al. 2009). In the *pre-1* gene tree, the African clade represented the ancestral group, but in the *pre-2* phylogeny American sequences were basal and thus ancestral. Unfortunately, the basal branches of both trees did not have bootstrap support. These differences in

basal grouping are common for single gene trees of fungi in the GFC, but the generally accepted species tree has African isolates forming the basal and ancestral clade (O'Donnell, Cigelnik, Nirenberg 1998; Kvas et al. 2009).

The receptor genes from the two mating-types of the species examined in the *Gibberella fujikuroi* complex grouped closely together, even though each individual requires only one receptor type on its cell surface (Poggeler, Kuck 2001; Kim, Borkovich 2006). The low level of divergence between pheromone receptor genes within a single species may be due to their involvement in other functions, such as regulation of pheromone gene expression and post-fertilization activities such as nuclear fusion (Kim, Metzenberg, Nelson 2002; Karlsson, Nygren, Johannesson 2008). The involvement of the receptors in post-fertilization activities could also be supported by the conservation of these genes even in species with extensive asexual life cycles and those with no known sexual stages, such as *F. mangiferae* and *F. sterilihyphosum* (Britz et al. 1999), as well as the presence of these genes in the homothallic species, *F. graminearum* (Kim, Lee, Yun 2008; Lee, Leslie, Bowden 2008).

Comparisons of amino acid sequences at different domains of PRE-1 and PRE-2 receptors revealed variable sites in both extracellular and cytoplasmic domains as well as the TM. The majority of amino acid substitutions in the cytoplasmic domains were localized in the C-terminal. A high level of variation in this domain was also observed in a study that focused on pheromone receptors from *Neurospora* species (Karlsson,

Nygren, Johannesson 2008). This domain interacts with the G-protein, which activates the signal transduction pathway (Karlsson, Nygren, Johannesson 2008). The amino acid changes observed in PRE-1 and PRE-2 may therefore likely affect activation of the receptor in *Fusarium* species. Such an effect was observed in *Saccharomyces cerevisiae* and *Coprinellus disseminates* where single amino acid substitutions lead to constitutively activated and hypersensitive receptors (Boone, Davis, Sprague 1993) in the former and self-compatibility in the latter (James et al. 2006).

In this study, the inferred PRE-2 peptide had fewer amino acid substitutions in the extracellular domains compared to PRE-1, which had mutations in all extracellular domains including the N-terminal domain. Mutations in the extracellular domains affect the structure of the ligand-binding pockets, which are important for binding of the appropriate ligand and activating the receptor (van Rhee, Jacobson 1996; Lawson, Wheatley 2004; Wheatley et al. 2012). The third extracellular loop (3EL) which is specifically involved in ligand recognition and binding was conserved in the PRE-2, while PRE-1 had three amino acid substitutions in this domain with one of them conserved between isolates from *Fusarium fujikuroi* (MP C) and *Fusarium intermedia* (MP D), which are known to be capable of hybridization (Leslie et al. 2004). A similar change was also observed in the same region with isolates from *F. sacchari* (MP B) and *F. fujikuroi* (MP C), but there is no evidence of inter-fertility between these two species. Although these changes are limited to a few isolates, the relatively high number of

positively selected sites in this domain supports its important role in ligand binding and/or selection and this makes this region important for future studies.

An additional amino acid change from leucine to phenylalanine in the sixth TM domain was observed in *Fusarium fujikuroi* (MP C) and *Fusarium intermedia* (MP D). This domain has been shown to be important in ligand-binding as well as activation of the receptors (Abel et al. 1998; Henry et al. 2002). Single mutations in these regions can affect binding affinity and specificity as well as activation of the receptors (Abel et al. 1998; Balasubramanian et al. 2005). These changes may thus be responsible for the inter-fertility between *F. proliferatum* and *F. fujikuroi*, but experimental evidence for this possibility is still lacking. Variations in gene sequences arise from mutations and these mutation sites may experience positive or neutral (natural) selection. Reproductive proteins, such as the pheromone receptor genes, often evolve under positive selection which is identified by ω (dN/dS) > 1 when averaged over all codons (Swanson, Vacquier 2002a; Swanson, Vacquier 2002b). The pheromone receptor genes from *Fusarium* species in this study did not exhibit this signature, with the ω value of both genes less than one (Table 4). This was also the case in a study of reproductive proteins from *Arbacia* species and this may be due to differences in evolutionary rates of the gene between lineages. Thus, some lineages and/or codon sites may evolve rapidly while others are under conservation and therefore cancel each other out, leading to ω<1 (Swanson, Vacquier 2002b). Functional constraint may also result in ω<1 especially on genes such as the pheromone receptors which are involved in multiple

roles during sexual reproduction in fungi (Bobrowicz et al. 2002; Swanson, Vacquier 2002b; Kim, Borkovich 2006; Kim, Lee, Yun 2008).

Some codons appeared to be under positive selection in PRE-1 and PRE-2, and these codons were located in various domains of the receptors. Codon sites that evolve rapidly (ω>1), such as those found in these genes, may affect receptor-ligand compatibility, therefore are linked to functional adaptation, speciation and species reinforcement (Swanson, Vacquier 2002a; Swanson, Vacquier 2002b; Dettman, Anderson, Kohn 2008). The sites under positive selection in both genes had amino acid changes that were shared by species from the same clades and not limited to a single the species. Thus, these positive selected sites are not likely to confer the mating-specificity seen amongst *Fusarium* species. These amino acid changes however could be interpreted as a sign of active divergence in these gene regions between *Fusarium* species with different origins (clades).

The observed amino acid change in the sixth TM of *F. fujikuroi* and *F. proliferatum* was not due to positive selection. However this does not completely dispute its possible role in inter-fertility between these species. One possible explanation is that this change will not be under positive selection due to allopatry between these two species, with *F. fujikuroi* found on rice while *F. proliferatum* has a broad host range in nature (Leslie, Summerell, Bullock 2006; Kvas et al. 2009). Species in allopatry have a reduced likelihood of mating therefore reinforcement of pre-mating barriers such as pheromone

receptors is often minimal or absent (Turner, Jacobson, Taylor 2010). In such cases we often do not observe positive selection in those parts of the gene that maybe involved in mate-selection since there are geographical and/or niche barriers acting between the two species. The difference in reproductive isolation pressure due to reinforcement between species in allopatry and sympatry was also observed amongst *Neurospora crassa* and *N. intermedia* (Turner, Jacobson, Taylor 2010).

The results presented here indicate that pheromone receptor genes have evolved in a species-specific manner amongst *Fusarium* species in the GFC, which is in contrast to their corresponding pheromone peptides. Most of these differences are due to synonymous substitution, leading to very little amino acid sequence variation amongst species. However, there are sites that exhibit positive selection, thus divergence, in certain regions of both genes amongst this group of species. The changes observed between peptide sequences are not sufficient to allow receptors to confer mating-specificity amongst *Fusarium* species in the GFC. Therefore both the pheromones and the receptor peptides have not diverged enough to act as reproductive barriers amongst these species. Future studies should focus on pre- and post-zygotic barriers to hybridization and identifying the role of the receptors in sexual reproduction of *Fusarium* species.

References

Abel, MG, BK Lee, F Naider, JM Becker. 1998. Mutations affecting ligand specificity of the G-protein-coupled receptor for the *Saccharomyces cerevisiae* tridecapeptide pheromone. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research 1448:12-26.

Bakkeren, G, JW Kronstad. 1996. The pheromone cell signaling components of the *Ustilago a* mating-type loci determine intercompatibility between species. Genetics 143:1601.

Balasubramanian, S, Y Xia, E Freinkman, M Gerstein. 2005. Sequence variation in G-protein-coupled receptors: analysis of single nucleotide polymorphisms. Nucleic Acids Research 33:1710-1721.

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

Bistis, G. 1981. Chemotropic interactions between trichogynes and conidia of opposite mating-type in *Neurospora crassa*. Mycologia 73:959-975.

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

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

Bockaert, J, JP Pin. 1999. Molecular tinkering of G protein-coupled receptors: an evolutionary success. The EMBO Journal 18:1723-1729.

Bölker, M, R Kahmann. 1993. Sexual pheromones and mating responses in fungi. The Plant Cell 5:1461-1469.

Boone, C, NG Davis, GF Sprague. 1993. Mutations that alter the third cytoplasmic loop of the a-factor receptor lead to a constitutive and hypersensitive phenotype. Proceedings of the National Academy of Sciences 90:9921-9925.

Britz, H, TA Coutinho, MJ Wingfield, WFO Marasas, TR Gordon, JF Leslie. 1999. *Fusarium subglutinans f. sp. pini* represents a distinct mating population in the *Gibberella fujikuroi* species complex. Applied and Environmental Microbiology 65:1198-1201.

Chang, S, C Staben. 1994. Directed replacement of *mat A* by *mat a-1* effects *a* mating-type switch in *Neurospora crassa*. Genetics 138:75-81.

Coppin, E, C De Renty, R Debuchy. 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, R Debuchy. 2000. Co-expression of the mating-type genes involved in internuclear recognition is lethal in *Podospora anserina*. Genetics 155:657-669.

D'mello, JPF, CM Placinta, AMC Macdonald. 1999. *Fusarium* mycotoxins: a review of global implications for animal health, welfare and productivity. Animal Feed Science and Technology 80:183-205.

Debuchy, R, BG Turgeon. 2006. Mating-type structure, evolution, and function in Euascomycetes. The Mycota: Growth, Differentiation and Sexuality 1:293-323.

Delport, W, AFY Poon, SDW Frost, SLK Pond. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26:2455-2457.

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

Dettman, J, J Anderson, L Kohn. 2008. Divergent adaptation promotes reproductive isolation among experimental populations of the filamentous fungus *Neurospora*. BMC Evolutionary Biology 8:35-48.

Gargas, A, JW Taylor. 1992. Polymerase chain reaction (PCR) primers for amplifying and sequencing nuclear 18S rDNA from lichenized fungi. Mycologia 84:589-592.

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

Guidon, S, O Gascuel. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52:696-704.

Hasegawa, M, H Kishino, T Yano. 1985. Dating of the human-ape splitting by a molecular clock of mitochondrial DNA. Journal of Molecular Evolution 22:160-174.

Henry, LK, S Khare, C Son, VVS Babu, F Naider, JM Becker. 2002. Identification of a contact region between the tridecapeptide *α-factor* mating pheromone of

Saccharomyces cerevisiae and its G protein-coupled receptor by photoaffinity labeling. Biochemistry 41:6128-6139.

James, TY, P Srivilai, U Kües, R Vilgalys. 2006. Evolution of the bipolar mating system of the mushroom *Coprinellus disseminatus* from its tetrapolar ancestors involves loss of mating-type-specific pheromone receptor function. Genetics 172:1877-1891.

Karlsson, M, K Nygren, H Johannesson. 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.

Kerényi, Z, K Zeller, L Hornok, JF Leslie. 1999. Molecular standardization of mating-type terminology in the *Gibberella fujikuroi* species complex. Applied and Environmental Microbiology 65:4071-4076.

Kim, H, KA Borkovich. 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, KA Borkovich. 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, RL Metzenberg, MA Nelson. 2002. Multiple functions of *mfa-1*, a putative pheromone precursor gene of *Neurospora crassa*. Eukaryotic Cell 1:987-999.

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

Kvas, M, WFO Marasas, BD Wingfield, MJ Wingfield, ET Steenkamp. 2009. Diversity and evolution of *Fusarium* species in the *Gibberella fujikuroi* complex. Fungal Diversity 34:1-21.

Lawson, Z, M Wheatley. 2004. The third extracellular loop of G-protein-coupled receptors: more than just a linker between two important transmembrane helices. Biochemical Society Transactions 32:1048-1050.

Lee, J, JF Leslie, RL Bowden. 2008. Expression and function of sex pheromones and receptors in the homothallic Ascomycete *Gibberella zeae*. Eukaryotic Cell 7:1211-1221.

Leslie, JF. 1991. Mating populations in *Gibberella fujikuro*i (*Fusarium* section Liseola). Phytopathology 81:1058-1060.

Leslie, JF, BA Summerell, S Bullock. 2006. The *Fusarium* laboratory manual: Wiley Online Library.

Leslie, JF, KA Zeller, M Wohler, BA Summerell. 2004. Interfertility of two mating populations in the *Gibberella fujikuroi* species complex. European Journal of Plant Pathology 110:611-618.

Martin, SH, BD Wingfield, MJ Wingfield, ET Steenkamp. 2011a. Causes and consequences of variability in peptide mating pheromones of ascomycete fungi. Molecular Biology and Evolution 28:1987-2003.

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

Nielsen, R, Z Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929-936.

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

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

Placinta, CM, JPF D'mello, AMC Macdonald. 1999. A review of worldwide contamination of cereal grains and animal feed with *Fusarium* mycotoxins. Animal Feed Science and Technology 78:21-37.

Poggeler, S, U Kuck. 2001. Identification of transcriptionally expressed pheromone receptor genes in filamentous ascomycetes. Gene 280:9-17.

Pond, SLK, SDW Frost. 2005a. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21:2531-2533.

Pond, SLK, SDW Frost. 2005b. Not so different after all: a comparison of methods for detecting amino acid sites under selection. Molecular Biology and Evolution 22:1208-1222.

Pond, SLK, SV Muse. 2005. HyPhy: hypothesis testing using phylogenies. Statistical Methods in Molecular Evolution:125-181.

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

Scauflaire, J, M Gourgue, F Munaut. 2011. *Fusarium temperatum sp. nov*. from maize, an emergent species closely related to *Fusarium subglutinans*. Mycologia 103:586-597.

Spyropoulos, IC, TD Liakopoulos, PG Bagos, SJ Hamodrakas. 2004. TMRPres2D: high quality visual representation of transmembrane protein models. Bioinformatics 20:3258-3260.

Steenkamp, ET, BD Wingfield, TA Coutinho, MJ Wingfield, WFO Marasas. 1999. Differentiation of *Fusarium subglutinans f. sp. pini* by histone gene sequence data. Applied and Environmental Microbiology 65:3401-3406.

Steenkamp, ET, BD Wingfield, AE Desjardins, WFO Marasas, MJ Wingfield. 2002. Cryptic speciation in *Fusarium subglutinans*. Mycologia 94:1032-1043.

Swanson, WJ, VD Vacquier. 2002a. The rapid evolution of reproductive proteins. Nature Reviews Genetics 3:137-144.

Swanson, WJ, VD Vacquier. 2002b. Reproductive protein evolution. Annual Review of Ecology and Systematics 33:161-179.

Tamura, K, J Dudley, M Nei, S Kumar. 2007. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. Molecular Biology and Evolution 24:1596-1599.

Taylor, JW, DJ Jacobson, S Kroken, T Kasuga, DM Geiser, DS Hibbett, MC Fisher. 2000. Phylogenetic species recognition and species concepts in fungi. Fungal Genetics and Biology 31:21-32.

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

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

van Rhee, AM, KA Jacobson. 1996. Molecular architecture of G protein-coupled receptors. Drug Development Research 37:1-38.

Wess, J. 1997. G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition. The FASEB Journal 11:346-354.

Wheatley, M, D Wootten, MT Conner, J Simms, R Kendrick, RT Logan, DR Poyner, J Barwell. 2012. Lifting the lid on GPCRs: the role of extracellular loops. British Journal of Pharmacology 165:1688-1703.

Yang, Z. 2000. Phylogenetic analysis by maximum likelihood (PAML). University College London 3.

Yang, Z. 2007. PAML 4: phylogenetic analysis by maximum likelihood. Molecular Biology and Evolution 24:1586-1591.

Yang, Z, WJ Swanson, VD Vacquier. 2000. 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.

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

Zhang, L, AC Churchill, P Kazmierczak, DH Kim, NK Van Alfen. 1993. Hypovirulence-associated traits induced by a mycovirus of *Cryphonectria parasitica* are mimicked by targeted inactivation of a host gene. Molecular and Cellular Biology 13:7782-7792.

Table 1. *Fusarium* isolates used in this study.

***Fusarium* species**	**Mating Population**	**Mating type**	**Isolate number**	
			MRC	**KSU**
Fusarium verticilllioides	A	*MAT1-1*	8559	A-00149
		MAT1-2	8560	A00999
Fusarium sacchari	B	*MAT1-1*	8552	B-03853
		MAT1-2	8551	B-03852
Fusarium fujikuroi	C	*MAT1-1*	8532	C-01993
		MAT1-2	8534	C-01995
Fusarium intermedia	D	*MAT1-1*	8549	D-04854
		MAT1-2	8550	D-04853
Fusarium subglutinans	E	*MAT1-1*	8553/6483	E-00990
		MAT1-2	8554/6512	E-02192
Fusarium temperatum		*MAT1-1*	1084	
		MAT1-2	7828	

Fusarium thapsinum	F	*MAT1-1*	8558	F-04094
		MAT1-2	8557	F-04093
Fusarium nygamai	G	*MAT1-1*	8546	G-05111
		MAT1-2	8547	G-05112
Fusarium circinatum	H	*MAT1-1*	7488	H-10847
		MAT1-2	6213	H-10850
Fusarium konza	I	*MAT1-1*	8545	I-11616
		MAT1-2	8544	I-11615
Fusarium mangiferae	N/A	*MAT1-1*	8092/8093	X4382
		MAT1-2	7559	11781
Fusarium sterilihyphosum	N/A	*MAT1-1*	2802	11783
		MAT1-2	8105	11782

Table 2. List of primers used in this study to amplify *pre-1* and *pre-2* genes.

Gene	Primer	Sequence 5′ → 3′
Pre 1	**Pre 1-1**	TAC GCC ATC GCC TGC AAT GA
	Pre 1-2	CAA CAG CAG CAG GTC TGC TA
	Pre 1-3	ATG ACT GCC GAC CAT GGA TA
	Pre 1-4	TGA TCA ACC GTT GAG CCA AG
Pre 2	**Pre 2-1**	CGA CAA GTA AGA GGC GGA TT
	Pre 2-2	TTC CAC AGA TTC GGC CAC AG
	Pre 2-3	TCG TCG TCC TCC TCG TCA TGG T
	Pre 2-4	GCC TGA GCG CTT ATC CAT CT
	Pre 2-3b	TCT GGA ACA TCT CCA AGT TCG
	Pre 2-4b	AGC AAC TTG TGC CGA TAC G

Table 3. Species specific primers used in this study to amplify *pre-2* gene

***Fusarium* species**	**Mating Population**	**Mating type**	**Isolate number MRC**	**KSU**	**Primers used 5' → 3'**
F. verticilllioides	A	*MAT1-1*	8559	A-00149	AF: 5' **GTCTCATGACTTTAGCCATATTC** AR: 5' **GCAGTCTTTTCGTGGAAGCT**
F. fujikuroi	C	*MAT1-1*	8532	C-01993	CF: 5' **TGCCGACTTTTACCCCTTCT** CR: 5' **GAAGAGGGTGGCCTAGATCC**
F. intermedia	D	*MAT1-2*	8550	D-04853	DF: 5' **TGCCGACTTTTACCCCTTCT** DR: 5' **TCTTCACGTTGCAAGAAGGC**
F. temperatum	E	*MAT1-1*	1084		EF: 5' **TCTGCTCGGTCAGAATGGTT** ER: 5' **CAGGAAGGCTTTGTCAACCC**
		MAT1-2	7828		
F. thapsinum	F	*MAT1-1*	8558	F-04094	GF: 5' **TCTGCTCGGTCAGAATGGTT** GR: 5' **AAGCCAGTGAAGGAGCTAGG**
		MAT1-2	8557	F-04093	
F. nygamai	G	*MAT1-1*	8546	G-05111	GF: 5' **TCTGCTCGGTCAGAATGGTT** GR: 5' **AAGCCAGTGAAGGAGCTAGG**

F. circinatum	H	*MAT1-1*	7488	H-10847	HF: 5' **CTCCAAGTTCGCCGACTTTT** HR: 5' **GCTGAAGTTGAAAGGGGTGG**
F. mangiferae	N/A	*MAT1-1*	8092/ 8093	X4382	MF: 5' **TGCCGACTTTTACCCCTTCT** MR: 5' **GAAGAGGGTGGCCTAGATCC**
F. sterilihyphosum	N/A	*MAT1-1*	2802	11783	HF: 5' **CTCCAAGTTCGCCGACTTTT** HR: 5' **GCTGAAGTTGAAAGGGGTGG**

Table 4. Results of the tests for selection conducted with datamonkey using *pre-1* and *pre-2* exon sequences showing codon sites identified by at least three of the five methods used.

	FEL	FUBAR	MEME	REL	SLAC
Pre-1	100	-	-	100	100
	219	-	-	219	219
	292	292	292	292	292
Pre-2	-	-	**83**	**-**	**83**
	112	112	112	112	112
	156	**-**	**-**	**156**	-
	181	-	181	181	-
	309	-	309	309	309

Codon sites in Bold were also identified by Nielsen, Yang (1998) method implemented in HyPhy (Pond, Muse 2005).

Table 5. Test results for positive selected sites based on the Nielsen-Yang (1998) models for the PRE-1and PRE-2 polypeptides.

Gene	Model	*p*	dN/dS	*ℓ*	2Δ*ℓ*	Estimated parameters	Positively selected
PRE-1	**M1** (Neutral)	59	0.17	-2961.73	**0.14**	ω1 = 0.10 *C2*= 0.91 ω2 = 1.00 *C3*= 0.09	Not allowed
	M2 (Selection)	61	0.17	-2961.66		ω1 = 0.11 *C1*= 0.94 ω2 = 1.22 *C2*= 0.06	100, 219, 292 and 297
	M7 (Neutral, β)	59	0.17	-2963.12	**2.58**	ω1 = 0.01 *C1*= 0.33 ω 2= 0.01 *C2*= 0.33 ω3 = 0.05 *C3*= 0.33	Not allowed
	M8 (Selection, β + ω)	61	0.17	-2961.83		ω1 = 0.09 *C1*= 0.62 ω2 = 0.18 *C2*= 0.30 ω3 = 1.00 *C3*= 0.08	100, 219, 292 and 297
PRE-2	**M1** (Neutral)	60	0.05	-3108.61	**0.00**	ω1 = 0.02 *C2*= 0.97 ω2 = 1.00 *C3*= 0.03	Not allowed
	M2 (Selection)	62	0.05	-3108.61		ω1 = 0.02 *C1*= 0.97 ω2 = 1.00 *C2*= 0.03	None
	M7 (Neutral, beta)	60	0.05	-3118.84	**20.9****	ω1 = 0.00 *C1*= 0.67 ω 2= 0.15 *C2*= 0.33	Not allowed
	M8 (Selection, beta + ω)	62	0.05	-3108.38		ω1 = 0.02 *C1*= 0.65 ω2 = 0.04 *C2*= 0.32 ω3 = 1.00 *C3*= 0.03	63, 83, 112 and 156

P is the number of free estimated parameters of **ω** generated by HyPhy (Pond, Muse 2005). ***ℓ*** is the log likelihood ratio scores for each model, as calculated by HyPhy. Significant difference between models indicated by **. ***C*** is the proportion of sites with a value equal to **ω**. **dN/dS** is the average ratio over the entire gene.

Figure 1. A maximum likelihood phylogenetic tree of *Fusarium* species within the *Gibberella fujikuroi* complex inferred from *pre-1* coding region (exons) sequences. Three major clades are indicated (African, American and Asian) as well as mating-type and mating population (MP A-I) for each species where applicable. Bootstrap values greater than 75% are indicated at the internodes and the tree is rooted with *F. oxysporum*.

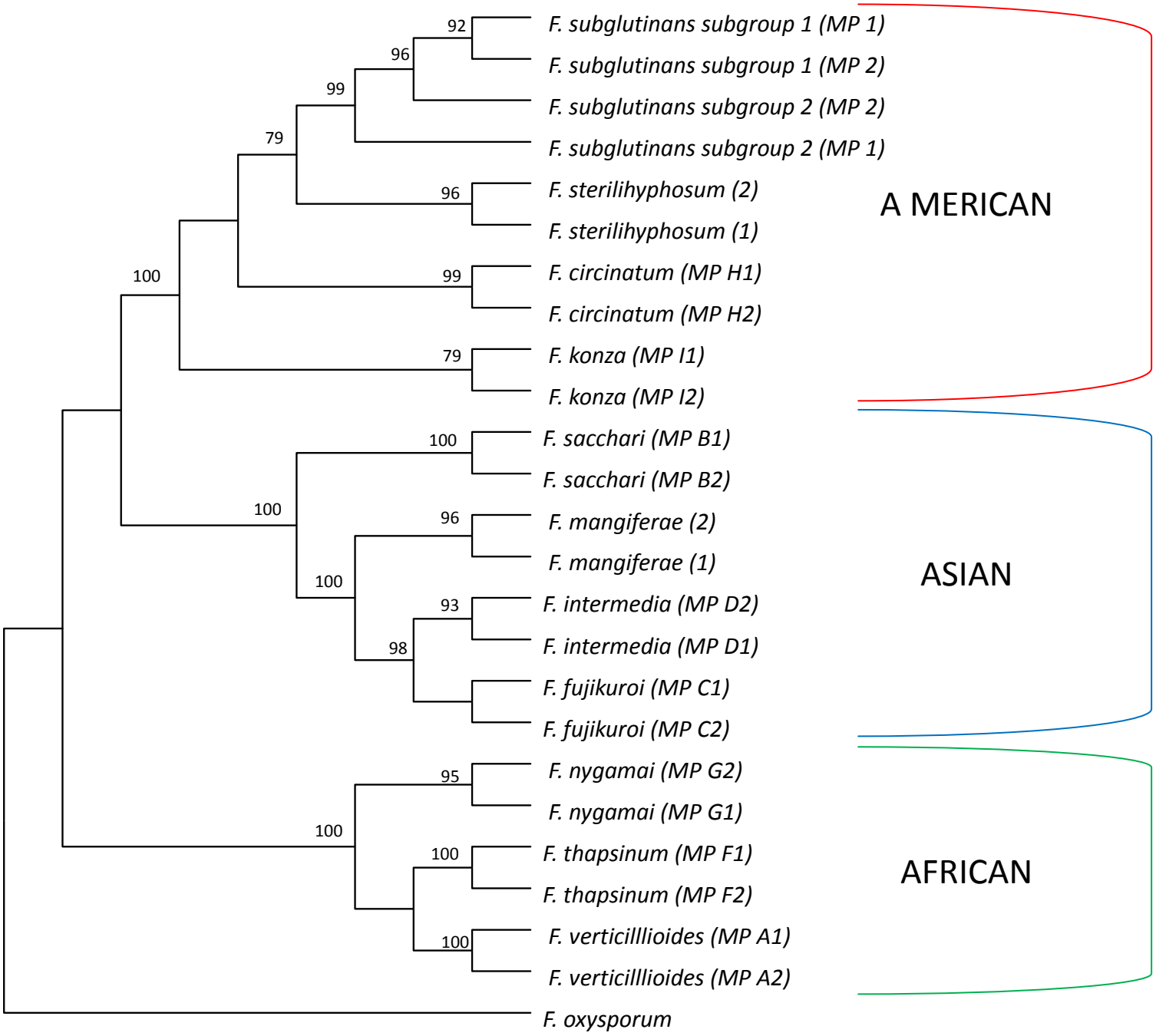


Figure 2. A maximum likelihood phylogenetic tree of *Fusarium* species within the *Gibberella fujikuroi* complex inferred from *pre-2* coding region (exons) sequences. Three major clades are indicated (African, American and Asian) as well as mating-type and mating population (MP A - I) for each species where applicable. Bootstrap values greater than 75% are indicated at the internodes and the tree is rooted with *F. oxysporum*.

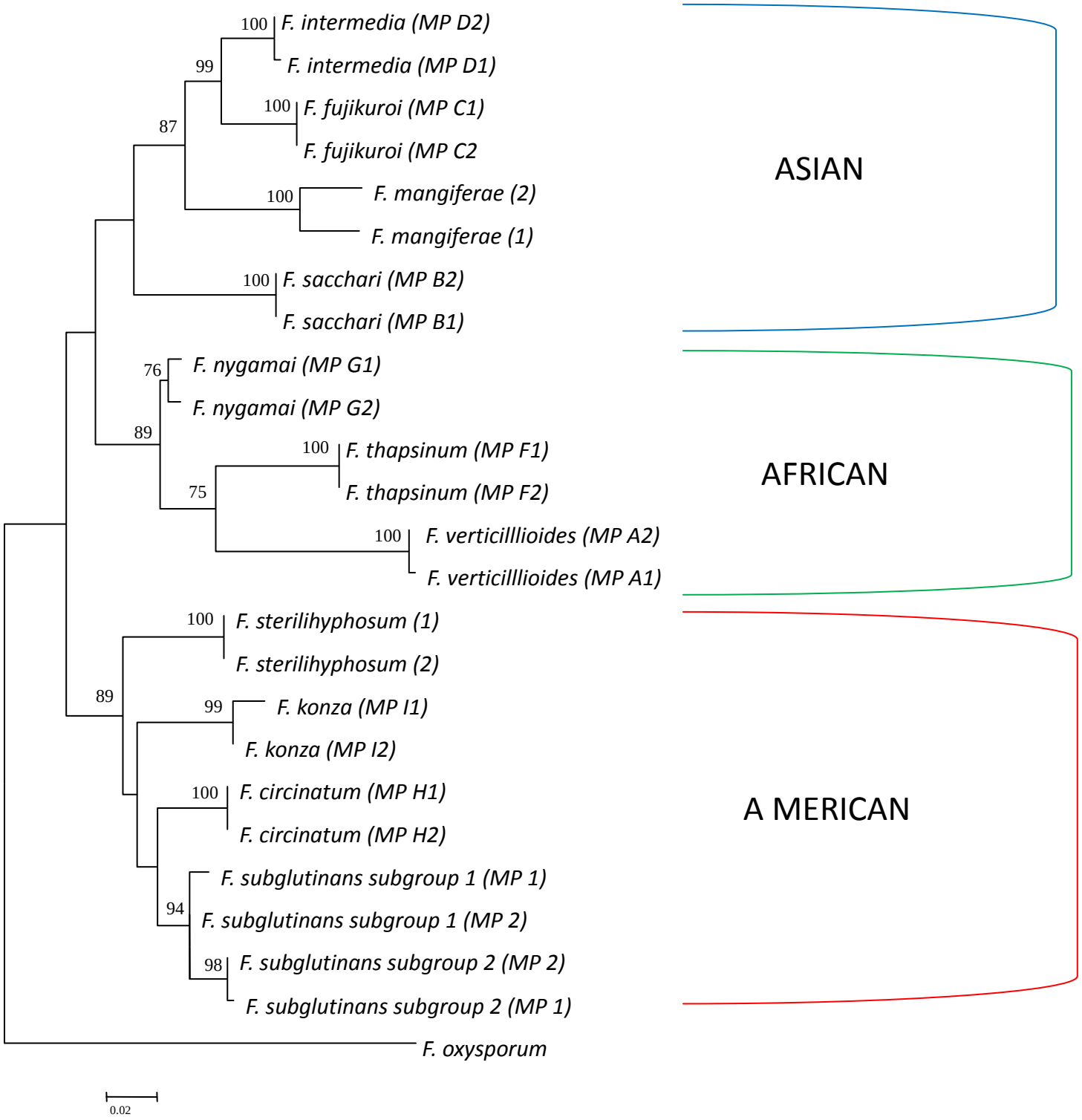


Figure 3 and 4. Comparison of PRE-1 (Fig. 3) and PRE-2 (Fig. 4) receptors sequences of *Fusarium* species from the *Gibberella fujikuroi* complex. Colored dots indicate identical residues and the alphabets indicate amino acid changes. Bars below the sequence alignment indicate receptor domains as predicted by TOPCONS (http://topcons.cbr.su.se/).

Fig. 3a

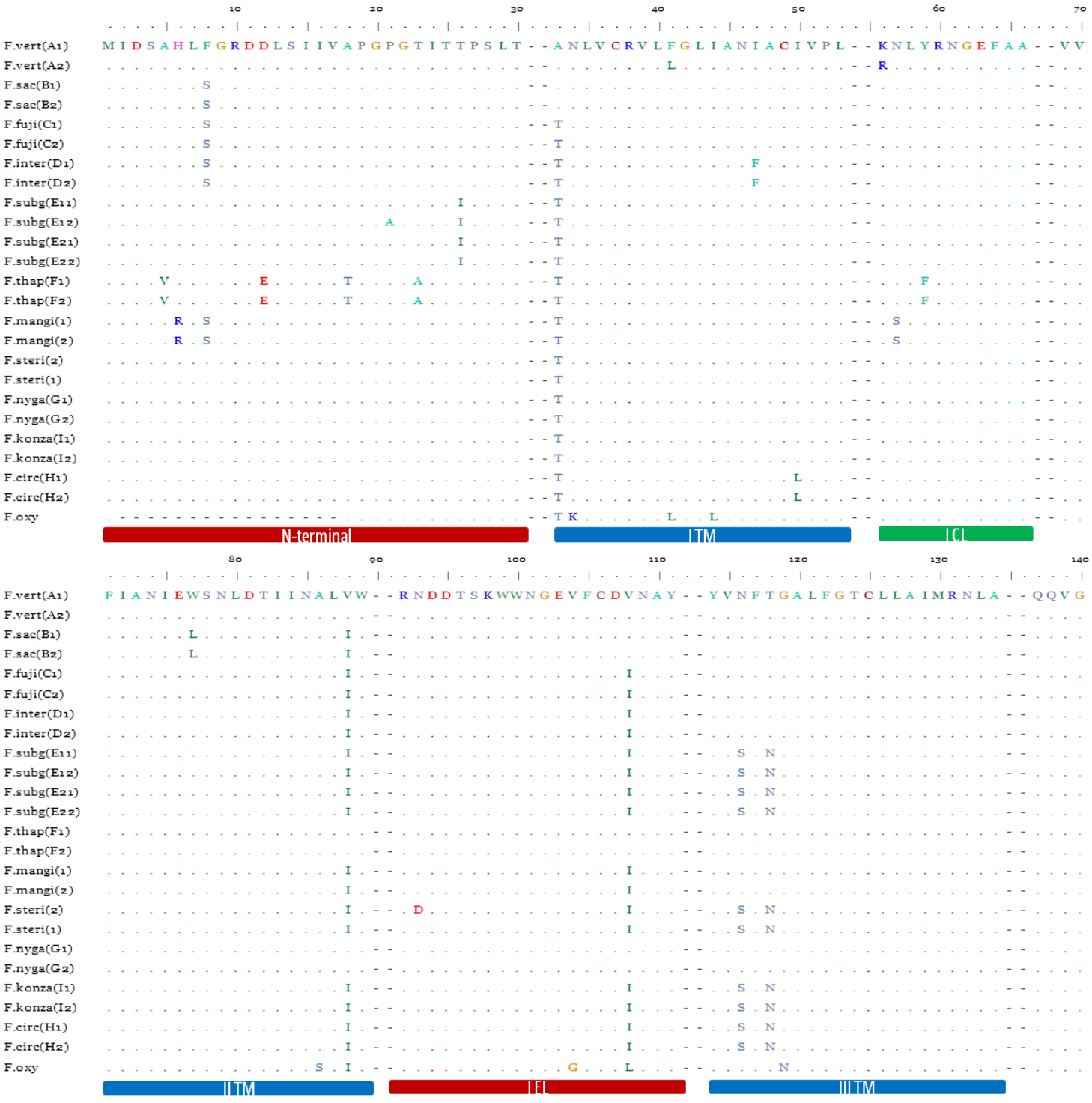


Fig. 3b

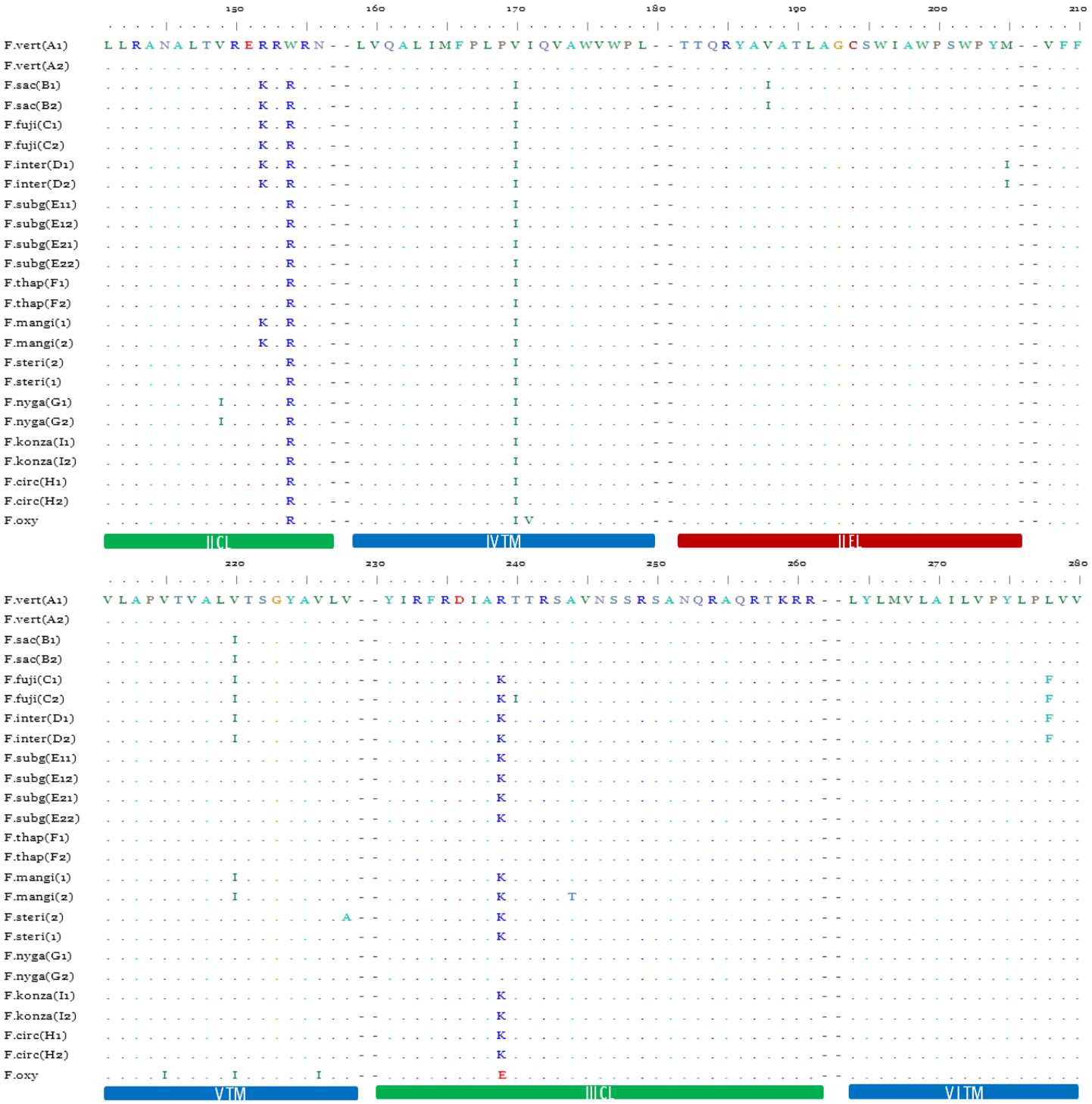


Fig. 3c

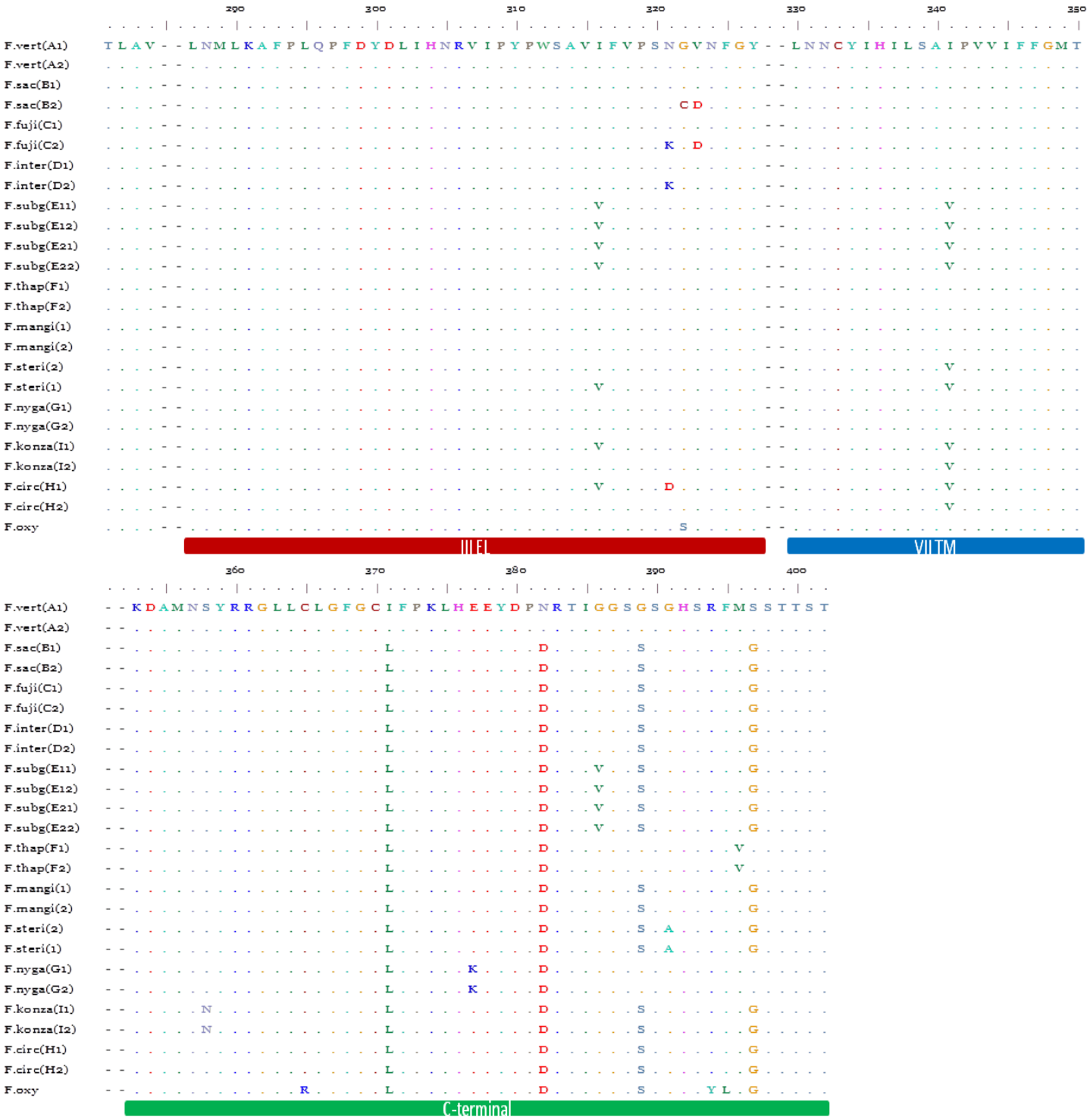


Fig. 4a

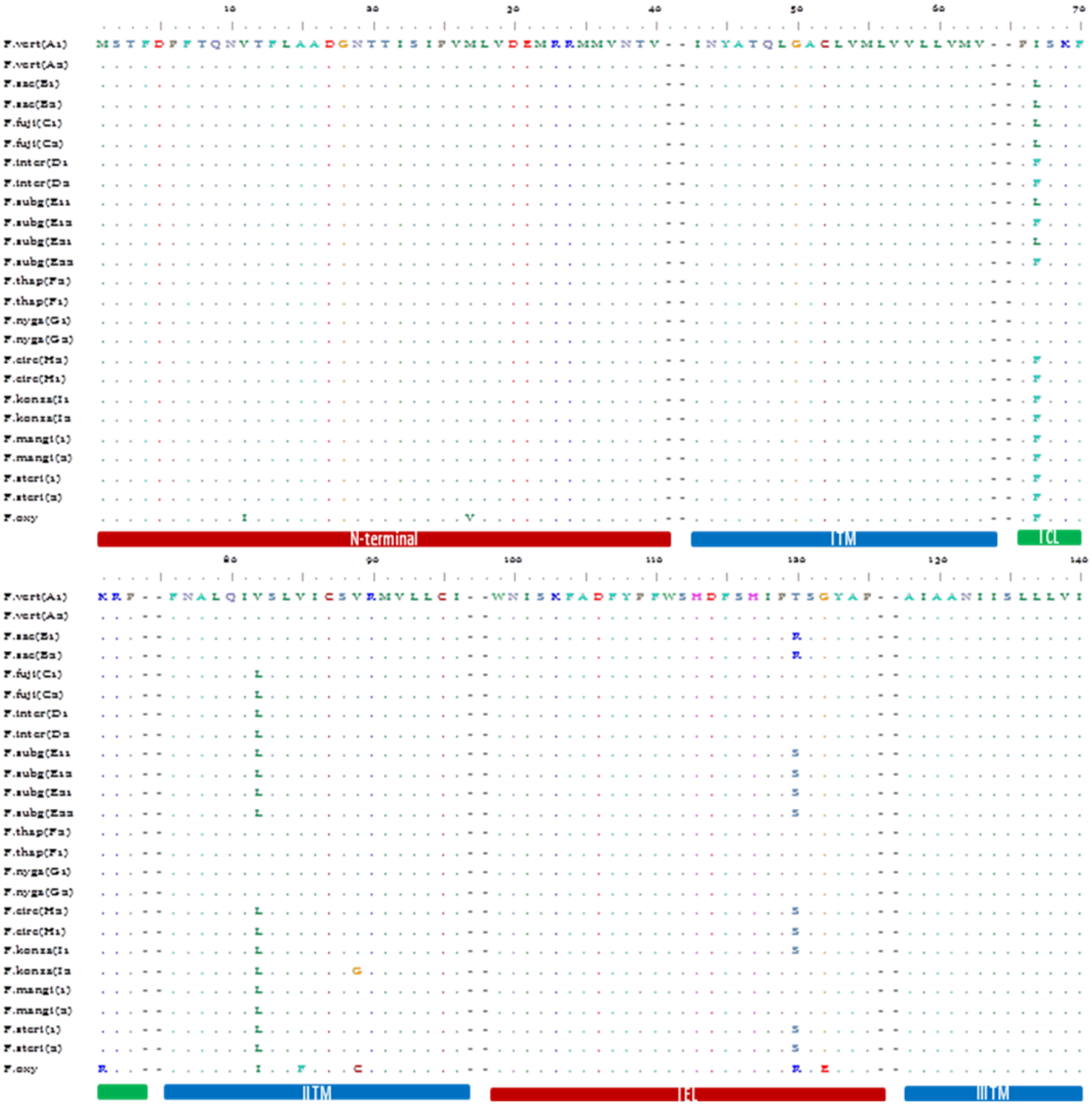


Fig. 4b

Fig. 4c

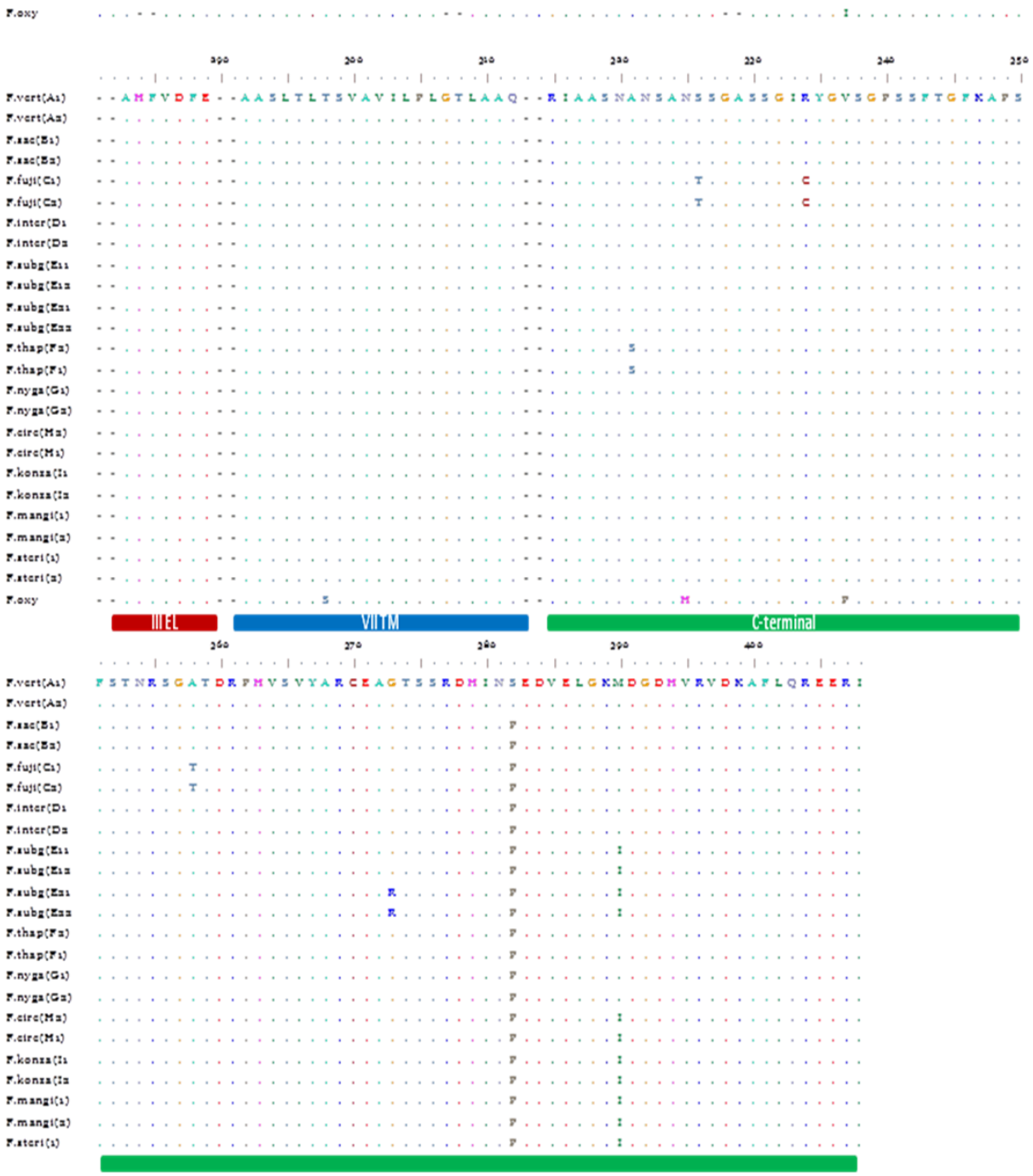
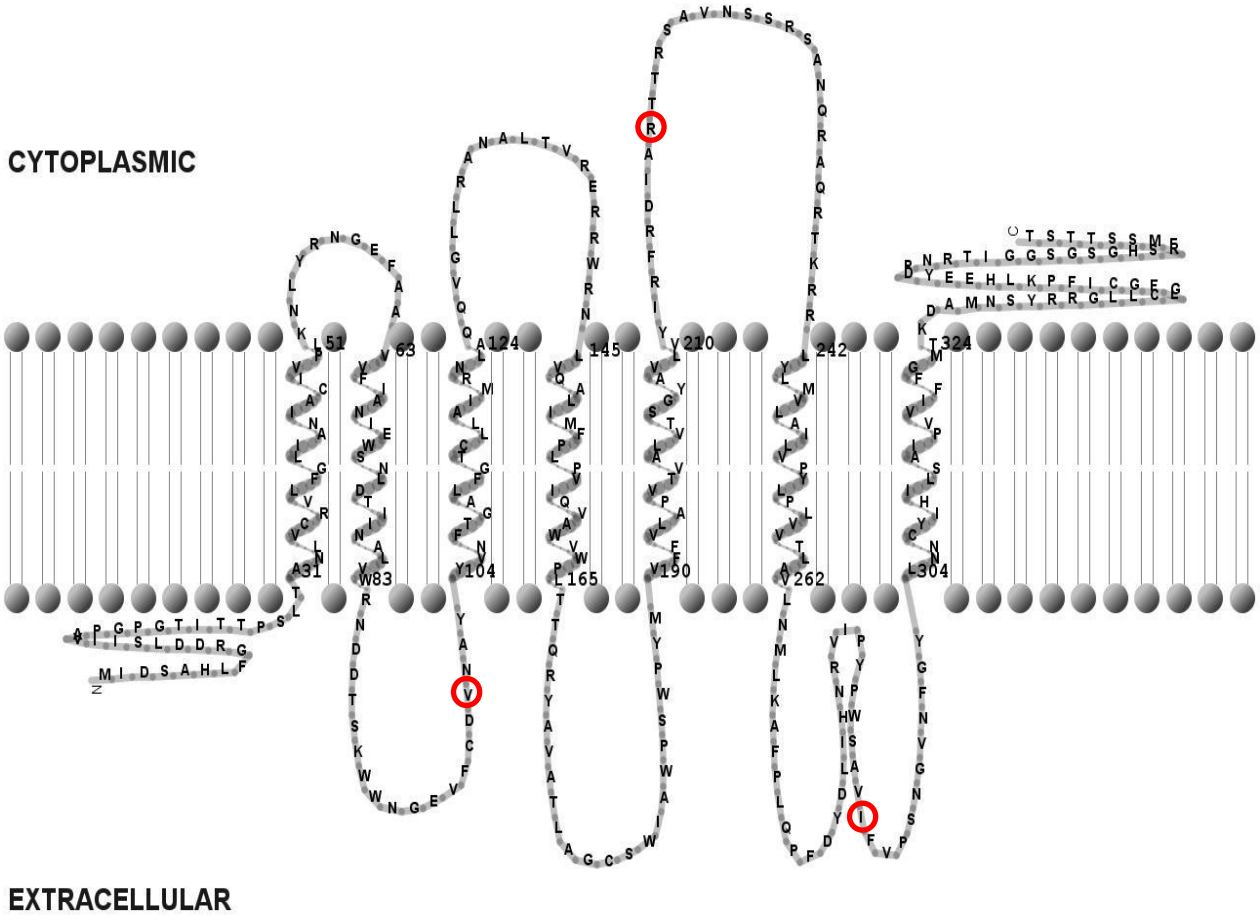
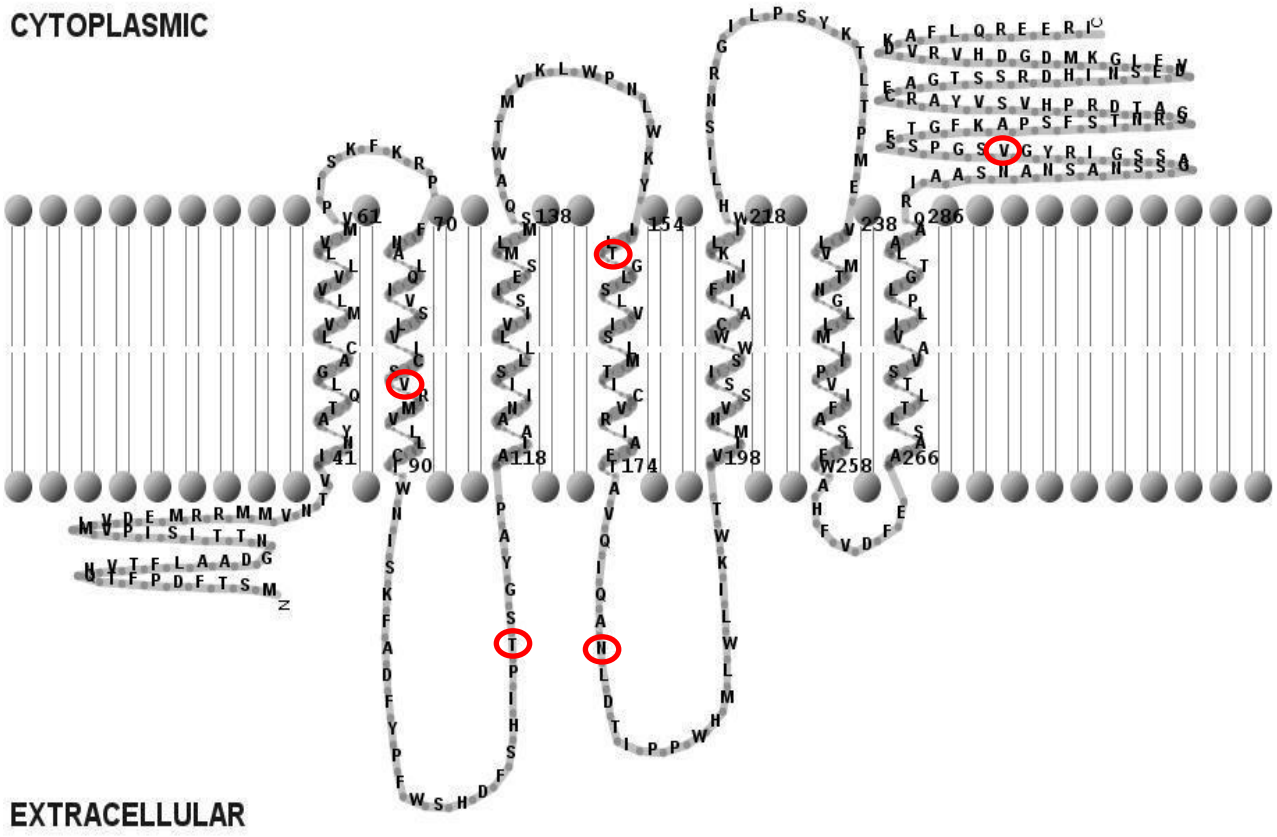


Figure 5a and b. The 2D structures of PRE-1 and PRE-2 constructed using TMRPres2D v0.91 (Spyropoulos et al. 2004) showing positive selected sites identified by at least three models highlighted in red.





Chapter 3

The evolution of pheromone receptor genes within the Sordariomycetes

Introduction

The Ascomycota consists of many species that have been isolated from diverse environments, including extreme environments such as Antarctica (Schoch et al. 2009). Species in this phylum are very diverse, not only in their niches but also their morphological structures and sizes (Zhang et al. 2006; Schoch et al. 2009). They play an important role in the decay of organic materials and are also found in mutualistic, pathogenic and parasitic relationships with other fungi, animals and plants (Schoch et al. 2009). The majority of the plant pathogenic Ascomycetes, such as *Cryphonectria parasitica*, *Magnaporthe grisea* and *Fusarium oxysporum* are found in the class Sordariomyceta which is subdivided into Leotiomycetes and Sordariomycetes (Zhang et al. 2006; Schoch et al. 2009; Wang et al. 2009). The latter includes at least 600 genera with more than 3000 species and three sub-classes, the Hypocreomycetidae, Sordariomycetidae and Xylariomycetidae (Zhang et al. 2006; Ainsworth 2008; Wang et al. 2009).

The Sordariomycetes include species which can reproduce sexually and asexually with the former involving three reproductive modes *i.e.* homothallism, pseudo-homothallism and heterothallism (Coppin et al. 1997; Kronstad, Staben 1997). The type of reproductive mode and the sexual reproduction processes are regulated by the mating type (*MAT*) locus which contains two MAT idiomorphs (*MAT1-1* and *MAT1-2*) (Coppin et al. 1997; Kronstad, Staben 1997). Genes found in the *MAT* idiomorphs control transcription of genes involved in mating, for example the pheromone and pheromone

receptor genes (Coppin et al. 1997; Kronstad, Staben 1997; Debuchy, Turgeon 2006; Martin et al. 2013). The pheromone genes encode short amino acid peptides that bind to target receptors, which initiate a signal transduction cascade that leads to activation of sexual reproduction processes (Bender, Sprague-Jr 1989; Bölker, Kahmann 1993; Shiu, Glass 2000; Kim, Borkovich 2006; Martin et al. 2011). Secretion and detection of pheromones thus form the initial stage of sexual reproduction, providing the first contact between the two potential mating partners (Casselton 2002; Martin et al. 2013).

The importance of the pheromone response system differs in the three reproductive modes (Coppin, De Renty, Debuchy 2005; Kim, Borkovich 2006; Mayrhofer, Weber, Poggeler 2006; Lee, Leslie, Bowden 2008; Nygren et al. 2012; Martin et al. 2013). This system is vital for heterothallic species that are unable to mate with themselves and thus use pheromones to attract compatible mating partners. For example, in heterothallic *Neurospora crassa, MAT1-1* cells predominantly express the *pre-1* pheromone receptor gene, and the resultant receptor will be activated by the *mfa-1* pheromone secreted by the *MAT1-2* cells. The reverse applies for the pheromone/receptor pair in *MAT1-2* cells (Pöggeler, Kück 2001).

Binding of pheromones to compatible receptors results in polarized growth of the cell towards the source of the stimulus, followed by cell fusion to initiate reproduction (Kim, Borkovich 2004; Kim et al. 2012). Self-fertile species (homothallic or peudo-homothallic) contain both *MAT1-1* and *MAT1-2* idiomorphs in each nucleus, and are thus able to

simultaneously express all genes required for sexual reproduction (Glass, Smith 1994; Coppin et al. 1997; Kronstad, Staben 1997; Yun et al. 2000; Kim, Lee, Yun 2008; Martin et al. 2013). These species do not need to use the pheromone response system to discriminate cells according to mating-type. This may lead to mutations and/or loss of gene function in one or both pheromone/receptor pairs (Pöggeler 2000; Turgeon, Yoder 2000; Coppin, De Renty, Debuchy 2005; Kim, Lee, Yun 2008). The retention of pheromone receptor genes, and the requirement for functional receptor genes in sexual reproduction by homothallic species, may be due to the multiple roles played by the pheromone receptors including ascospore production (Kim et al. 2012), internuclear recognition (Debuchy 1999) and meiosis (Coppin et al. 1997).

Most Sordariomycete species have overlapping niches with other species of the same class, yet hybrid species are rarely observed. Species therefore rely on the specificity of pheromones and/or receptors to prevent hybridization. This is achieved through the use of Guanine (G) protein coupled receptors (GPCRs) on cell membranes, which are also found in animals and plant cells (Bockaert, Pin 1999; Li et al. 2007). The GPCRs are identified by their characteristic seven trans-membrane (TM) domains and the G-protein attached to the cytoplasmic domains of the receptor (Bockaert, Pin 1999; Li et al. 2007). The GPCRs consist of three extracellular and three cytoplasmic loops that link the TM spanning domains as well as the N- and C-terminal regions of the protein, which are located outside and inside the cell, respectively (Bockaert, Pin 1999; Li et al. 2007). Domains found on the extracellular side are usually responsible for ligand binding, while

cytoplasmic domains interact with the G-protein (van Rhee, Jacobson 1996; Li et al. 2007). The TM domains give the receptor its tertiary structure, which is critical for efficient binding and interaction with the ligand (van Rhee, Jacobson 1996; Bockaert, Pin 1999).

Since the GPCRs share a common 7TM protein structure, specificity is most likely provided in the ligand binding domains by species-specific amino acid sequences. For example, the *pre-2* gene from *Sordaria macrospora* was able to activate the yeast pheromone response system of *Saccharomyces cerevisiae* when it was exposed to *S. macrospora* PPG1 pheromone (Mayrhofer, Poggeler 2005). The pheromone response system is also involved in reproductive reinforcement amongst *Neurospora* species, where a correlation exists between the number of viable spores produced and pheromone/receptor gene expression levels during inter-specific sympatric crosses (Dettman et al. 2003).

Reproductive reinforcement may be driven by either positive or negative selection, which involves the selection for or against sequence changes that are beneficial or detrimental to reproduction (Swanson, Vacquier 2002a). Very little information on these evolutionary processes in the Sordariomycetes is available, even though determining their involvement in the evolution of pheromone receptor genes may help us understand reproductive reinforcement and inter-fertility seen amongst some Sordariomycetes. The aim of this study was thus to determine how pheromone receptor genes evolved

amongst a sample of Sordariomycete species with different levels of inter-fertility and to determine the role of positive selection in the divergence of these genes.

Materials and methods

Phylogenetic analyses

A maximum likelihood gene tree was constructed for homologs of the GPCRs PRE-1 and PRE-2, which encode the mating pheromone receptors of Sordariomycetes. Homologous receptor sequences were identified from 35 Sordariomycetes (Table 1) available from the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/), Broad institute of MIT and Harvard (http://www.broadinstitute.org/) and Joint Genome Institute of the US Department of Energy (JGI; http://genome.jgi.doe.gov/) using a protein blast (blastp) search. Amino acid sequences were downloaded and aligned using ClustalW version 1.4 (Thompson, Higgins, Gibson 1994) as implemented in BioEdit version 7.0.9.0 (Hall 2004). Sequences were subsequently analyzed using the Phylogeny.fr pipeline (http://www.phylogeny.fr/version2_cgi/), which makes use of various programs and have high computational power (Dereeper et al. 2008). The “Advanced analysis” mode was used to allow for manual adjustment of parameters associated with each step in the workflow. Sequences were aligned using MUSCLE 3.7 (Edgar 2004) before inference of maximum likelihood phylogenies using PhyML 3.0 (Guindon et al. 2010) while implementing the best-fit model (WAG; Whelan, Goldman (2001)) of substitution.

Confidence in branches was assessed using 100 bootstrap replicates. Both trees were rooted using homologous sequences from *Botrytis cinerea* and *Sclerotinia sclerotiorum*.

Selection analyses

The pheromone receptor genes from *Trichoderma* and *Neurospora* were subjected to tests for selection in order to investigate whether these genes were under positive selection. Subsequently, these results could be compared to results for *Fusarium* species in the *Gibberella fujikuroi* complex (Chapter 2). Protein coding DNA sequences of *Trichoderma* and *Neurospora* pheromone receptor genes were downloaded from JGI and NCBI (Table 1). Sequences were aligned in BioEdit (Hall 2004) and to infer maximum likelihood trees based on the best-fit substitution model according to jModelTest 0.1.1 (Posada 2008). Confidence in branches was tested using 1000 bootstrap replicates in PhyML version 3.0 (Guindon et al. 2010). The inferred tree topologies were subsequently used in the codon-based maximum likelihood method (Nielsen, Yang 1998) implemented in the HyPhy package (Pond, Muse 2005) to detect evidence of positive selection. This method involves hypothesis testing, during which a LRT is used to determine the best-fit model for the data between neutral evolution and the alternative positive selection model (Nielsen, Yang 1998). In this study two pairs of models were tested, namely M1 (neutral) vs M2 (positive selection) and M7 (β) vs M8 (β + ω). The neutral model (M1) restricts sites into two classes, where they are either conserved ($\omega<1$) or neutral ($\omega=1$), whereas the M2 model allows for and extra class of sites under positive selection ($\omega>1$) (Nielsen, Yang 1998). The more complex M7 model

allows ω to vary across sites in a beta (β) distribution manner, while M8 allows for an additional class of ω>1 (Llopart, Comeron 2008). The posterior probability that a codon site belongs to one of the three classes is then calculated using Bayes Empirical Bayes (BEB) (Nielsen, Yang 1998).

To improve accuracy, additional methods implemented in the Datamonkey webserver (http://www.datamonkey.org/) were used to identify sites under positive selection (Delport et al. 2010). These methods, namely Fixed Effect Likelihood (FEL), Random Effect Likelihood (REL) and Single Likelihood Ancestor Counting (SLAC), FUBAR and Mixed Effect Model of Evolution (MEME), were previously described in chapter 2. The TM domains of each receptor were identified using TOPCONS (http://topcons.cbr.su.se/) in order to identify the domain in which each positive selected site was located (Bernsel et al. 2009).

Lineages do not experience the same selection pressure at the same time and this usually affects the mean ω value of the gene. Thus, it is necessary to identify which lineages experienced episodic positive selection. In this study we applied the Branch-Site REL (BS-REL) method (Pond et al. 2011) available in Datamonkey. Unlike other similar methods, BS-REL uses the random effect likelihood (REL) and allows ω to vary in each site and branch. Using coding DNA sequences and the user's phylogenetic trees constructed with PhyML version 3.0 (Guidon, Gascuel 2003) using maximum likelihood and the best-fit models obtained from jModelTest 0.1.1, we identified *pre-1*

and *pre-2* lineages under episodic positive selection amongst *Fusarium*, *Trichoderma* and *Neurospora* species.

Results

Phylogenetic analyses

The evolution of pheromone receptor genes (*pre-1* and *pre-2*) amongst 35 Sordariomycete species was investigated based on the phylogeny of *pre-1* and *pre-2* amino acid sequences. Each species belonged to one of six sister taxa, namely the Diaporthales, Glomerellales, Hypocreales, Magnaporthales, Sordariales and Xylariales. PRE-1 and PRE-2 genealogies (Figure 1 & 2) were compared to the species tree topology (Supplementary Figure 1) and only branches with significant support (Bootstrap value >75) are discussed and labeled in this study.

The topology of the receptor gene phylogenies showed conflicts in the ancestral taxa when they were compared to each other and to the species Tree (Supplementary Figure 1). The Xylariales, represented by *Apiospora montagnei*, formed the basal group of the Sordariomycetes (Supplementary Figure 1), but in the pheromone receptor genes, Magnaporthales and Hypocreales formed the basal groups (Figure 1 & 2). The PRE-1 genealogy also placed *Glomerella cingulata* in a paraphyletic group with other Glomerellales and Hypocreales with a significant bootstrap support value of 78 (Figure 1). Other conflicts with significant support were observed in the PRE-1 gene tree, where *Podospora anserina* clustered with the Magnaporthales (bootstrap =100), sharing

identical amino acid sequences. A major topological conflict was also observed in the PRE-2 genealogy where *Cryphonectria parasitica* (Diaporthales) clustered within the Sordariales (Figure 2) instead of as a sister group such as in the species tree (Supplementary Figure 1). The two species from the genus *Magnaporthe*, namely *M. grisea* and *M. oryzae*, shared amino acid sequence identity for both PRE-1 and PRE-2 (Figure 2). In contrast, *Verticillium dahliae* and *V. albo-atrum* PRE-1 sequences were identical (Figure 1) while there were minor polymorphisms between their PRE-2 sequences (Figure 2). *Neurospora* and *Trichoderma* included multiple species (8 and 7 respectively), and the highest level of amino acid sequence divergence existed between the *Trichoderma* clade compared to *Neurospora* (Figure 1 & 2).

Selection analyses

The evolutionary histories of the pheromone receptor genes, *pre-1* and *pre-2*, were compared among the *Fusarium*, *Neurospora* and *Trichoderma* clades. We investigated positive selective pressure acting on these genes as well as lineages and codons that experienced positive selection. Our results showed that none of the pheromone receptor genes evolved under positive selection (mean ω >1), but there were individual sites and lineages that experienced positive selection. The mean dN/dS (ω) of both *pre-1* and *pre-2* genes showed strong purifying selection with mean ω (dN/dS) < 1. The *Trichoderma* and *Neurospora* genes also had means ω < 1 (Table 2). LRTs were used to determine the best-fitting model that explains diversity within *pre* genes between the neutral models (M1 and M7) and positive selection models (M2 and M8). Results

showed that the neutral and positive selection models were not significantly different (*P* = 0.05) for *Fusarium* genes, suggesting that sequence diversity in both genes was likely due to neutral selection. In contrast, the two selection models fitted the data better than their nested neutral models for both genes from *Trichoderma* species (*P* = 0.05) with the exception of M1 vs M2 in *pre-2*, where the null hypothesis could not be rejected (Table 2). Purifying selection, or conservation, in the *Trichoderma* genes was not as high as that observed in the *Fusarium*, with mean ω values as high as 0.81 for M2 in *pre-1* and 0.45 for M1 in *pre-2* (Table 2). Comparison of model M1 to M2 for *Neurospora*, allowed rejection of the null hypothesis of neutral evolution (*P* = 0.05) for *pre-1* (Table 2). However the results differed for the more complex M7 and M8 models, showing insignificant difference between the two models (*P* = 0.05) (Table 2). The *Neurospora pre-2* gene showed similar results to those observed for the *Fusarium* genes, where the null hypothesis could not be rejected (Table 2). The average dN/dS for *pre-1* was found to be 0.37 (M2) and 0.35 (M7) while *pre-2* experienced even stronger purifying selection with values as low as 0.13 (M1) and 0.12 (M7) according to the best-fitting models (Table 2).

Different methods used in Datamonkey (Pond, Frost 2005; Delport et al. 2010) and the *Nielsen-Yang* method implemented in HyPhy (Pond, Muse 2005) were applied to the data to identify sites evolving under positive selection. Seven methods were implemented to improve accuracy, and only sites that were identified by at least four of the methods were considered further. Both pheromone receptor genes from the three

taxa investigated in this study contained positive selected sites located in various domains of the receptor protein (Figures 3-6). Sites under positive selection in the *Fusarium* taxa are available in the previous chapter (Chapter 2) and will not be repeated in this chapter but will be included in the discussion section.

Consensus results identified three sites (172, 559 and 703) under positive selection, ω (dN/dS) >1, in the *Neurospora pre-1* gene (Figure 3). Codon 172 was located in the fourth TM domain in close proximity to the extracellular side of the membrane, whereas the remaining two sites were both situated in the cytoplasmic C-terminal domain (Figure 3). The *pre-2* gene had a relatively high number of positive selected sites. The extracellular N-terminal domain of this gene contained three such sites (27, 30 and 55), followed by two sites (79 and 86) in the adjacent first TM domain (Figure 4). Another site (195) was identified in the fourth TM while five additional sites occupied the C-terminal domain (Figure 4). None of these sites were identified by the *Nielsen-Yang* method in HyPhy. *Trichoderma pre-1* and *pre-2* genes both contained two sites under positive selection. These were located in the N-terminal in *pre-2*, while *pre-1* had one site in the N-terminal and the other in the third intracellular loop (Figures 5 & 6). None of the sites detected in this study were under selection in all three taxa.

Branch-site REL (Pond et al. 2011) was used to identify lineages that experienced episodic positive selection in the evolution of pheromone receptor genes of representative *Fusarium, Neurospora* and *Trichoderma* species. Episodic positive

selection appeared to be absent in all *Fusarium* lineages for both of the pheromone receptor genes. In *Neurospora*, positive selection could be detected in multiple branches of the *pre-1* and *pre-2* gene trees ($P \leq 0.05$). Both *pre-1* and *pre-2* identified node 1 and node 2 to be under positive selection in *Neurospora* (Figures 7 & 8). Multiple branches were under episodic positive selection in *Trichoderma pre-1* and *pre-2* gene trees ($P \leq 0.05$). Four of those branches, namely node 1, branch 1 (B1), 2 (B2) and 4 (B4), were identified in both *pre* genes (Figures 9 & 10).

Discussion

The receptors function at the amino acid level, implying that they should differ at this level to provide species-specificity. In the current chapter we investigated whether amino acid sequence phylogenies could be resolved in other taxa of the Sordariomycetes. These results showed separation of receptor genes similar to the species phylogeny, with a few exceptions where species shared amino acid sequences or could not be separated due to insignificant differences. In chapter 2 the nucleotide phylogeny of pheromone receptor genes belonging to *Fusarium* species in the *Gibberella fujikuroi* complex (GFC) was inferred, and it was shown that the DNA sequences of these genes separated in a species-specific manner. However, when amino acid sequences were used, the phylogeny could not be resolved. Similar results were observed with the *a*- and α-class pheromone genes, where amino acid sequences were unique to some species including those in the *Trichoderma* lineage, but not in the GFC (Martin et al. 2011).

Sequence diversity can often be linked to the probability that species can hybridize (Martin et al. 2013). Relatively low amino acid sequence variability in *pre*-genes was observed in the *Fusarium* and *Neurospora* clades, where inter-fertility has been reported (Desjardins, Plattner, Gordon 2000; Dettman et al. 2003; Leslie et al. 2004; Karlsson, Nygren, Johannesson 2008). In contrast, inter-species hybridization had not yet been observed in the *Trichoderma* clade, although *Fusarium* and *Neurospora* species show different levels of inter-fertility, ranging from initiation of mating between species, which may be aborted, to the production of viable spores. The low diversity in *pre-1* and *pre-2* resulted in unresolved genealogies, with short internal tree branches that showed a low level of divergence in the receptors amongst these species. Pheromone receptors have been shown to act as a pre-mating barrier (Martin et al. 2013), but our results suggest that in these taxa the receptors have not diverged enough to play that role. However, it is still not clear how many amino acid changes are required for the receptors to become a pre-mating barrier and this remains to be explored experimentally.

Pheromones and their receptors show signs of co-evolution in Sordariomycete species. For example, *Magnaporthe grisea* and *M. oryzae* display identical amino acid sequences for both their pheromones and their pheromone receptors. In contrast, *Verticillium albo-atrum and V. dahlia* shared sequence identity for both the *a*-class pheromone and its cognate PRE-1 receptor, but different *α*-class and PRE-2 amino acid

sequences. Relatively high sequence diversity of pheromone and receptor sequences could also be detected amongst *Trichoderma* species, while few differences existed amongst *Fusarium* species for the same genes (Chapter 2). Sexual reproduction genes have also been identified in many fungal species that were classified as asexual, such as *Aspergillus nidulans* (Anderson et al. 2005; Dyer, Paoletti 2005). This observation suggests that these genes have been conserved in many fungal species. Two potential explanations exist for the conservation and co-evolution of these genes in both sexual and "asexual" species. First, sexual reproduction may be happening more often than previously expected, even in species that are presumed to be asexual, and second, the pheromones and their receptors may have additional functions (Casselton 2002; Martin et al. 2013). These two potential explanations could be tested, because it would be expected that different levels and types of selection would act on the genes, depending on which explanation is most likely.

We determined whether positive selection was involved on the evolutionary pattern of pheromone receptor genes in species of *Fusarium, Neurospora* and *Trichoderma*. In both *Fusarium* and *Neurospora*, there have been reports of interfetility between different species of the same genus (Dettman et al. 2003; Leslie et al. 2004). Inter-fertility has not been reported amongst *Trichoderma* species included in this study (Druzhinina, Kubicek 2005). Results showed that receptor genes are under purifying selection in all three taxa. These results support the notion that pheromones and their receptors may have additional functions. This is because most reproductive genes, whose main

function is mate detection, have been found to evolve rapidly in many species including humans and plants (Yang, Swanson, Vacquier 2000; Swanson, Vacquier 2002a; Swanson, Vacquier 2002b). Also the receptors show little sequence variation between sexual and asexual species such as those in the *Fusarium* and *Neurospora taxa*, even with homothallic species where at least one of the receptors is not crucial for initiation of sexual reproduction (Glass, Smith 1994; Kim, Lee, Yun 2008; Lee, Leslie, Bowden 2008). In contrast if the receptors were being conserved in both sexual and asexual species because of their importance in sex initiation and occurring frequently even in "asexual" species, we would expect the amount of mutations to be greater in pseudo- and homothallic compared to heterothallic and "asexual" species.

Trichoderma had the lowest levels of purifying selection, with values approaching ω=1 (neutral selection) in *pre-1*, while *pre-2* showed mild purifying selection at ω=0.45. In addition, both positive selection models (M2 and M8) were the best-fitting for *pre-1* data while only one of them (M8) was best-fitting for *pre-2* data in the hypothesis testing analysis. *Fusarium* and *Neurospora pre-1* averaged 0.17 and 0.36 respectively, which can be interpreted as a sign of high conservation of this gene due to selection with the strongest selection amongst the examined *Fusarium* species compared to the other two taxa. This conservation could be a result of strong purifying selection or due to relatively recent divergence, and thus fewer mutations amongst the relatively recently diverged *Fusarium* species. This was also the case in the *pre-2* gene, where *Fusarium* displayed the lowest ω value of 0.05, followed by *Neurospora* with 0.13. However, it is important

to note that *Neurospora pre-1* and *Fusarium pre-2* showed significant differences between neutral and positive selection models, thus positive selection models could not be rejected (M2 and M8 respectively), which suggests that these genes had previously experienced positive selection.

Overall, the results show that positive selection was involved in divergence of at least one of the pheromone receptor genes amongst the three taxa. This finding is in accordance with previous studies that showed that at least one functional pheromone receptor is required for sexual reproduction (Kim, Borkovich 2004). It will be worth exploring if there are any functional differences between positive selected *pre*-genes and those that are not in *Fusarium* and *Neurospora*. These results may also suggest that one pheromone-receptor pair is involved in reproduction more often than the other, and thus it experiences selection due to reinforcement (Giraud et al. 2008; Martin et al. 2013) or other processes that lead to divergence. For instance, studies have found that there was abundance of one mating type in a population of *Fusarium* (Leslie, Summerell, Bullock 2006), such situations would lead to the pheromone/receptor pair associated with that mating type to be expressed more often, thus exposed to stronger selection than the non-abundant pair.

One might also predict that for complete inter-sterility to occur, pheromone receptor amino acid sequences must be different in both PRE-1 and PRE-2. This could explain positive selection acting on both pheromone receptor genes of *Trichoderma* species

that have no record of inter-fertility. Because Fungi are hermaphrodite, where each cell act as both male and female and thus produces one of each receptor and pheromone type (that are not compatible), we expect that both the receptor and pheromone genes need to have diverged between any two species to prevent inter-fertility. For example, *Verticillium dahliae* and *Verticillium albo-atrum* differ in one pheromone/receptor pair but not the other, and are believed to be the parents of a hybrid species, namely *V. longisporum* (Pantou et al. 2005). Therefore, in the case of *Fusarium* and *Neurospora* we can predict that the pheromone receptor under positive selection may be more involved as a reproductive barrier whereas the other genes may sometimes, though rarely, allow for inter-fertility due to low levels of sequence divergence between species.

GPCRs have a characteristic structure with different domains that are involved in various functions within the receptor. Sites under positive selection are usually involved in protein functioning, specifically during binding of a ligand (Wess 1997; Bockaert, Pin 1999; Li et al. 2007). We identified positive selected sites in both genes from all three taxa and determined if they were located in the predicted binding sites, since previous studies showed that the TM and the extracellular loops are involved in ligand recognition and binding (Wess 1997; Bockaert, Pin 1999). *Fusarium* PRE-1 had four sites in the first and third extracellular loops, while PRE-2 had two sites in the first and second extracellular loop and one site in the fourth TM domain (Chapter 2). Positive selected sites in the fourth TM domain were also found in both PRE-1 and PRE-2 proteins in *Neurospora* with an additional site in the first TM of PRE-2. In contrast,

Trichoderma had no selected sites in the predicted ligand binding domains in either of the receptors, but this may be because these proteins have already diversified amongst these species. The C-terminal domain has been shown to be important for receptor sensitivity, and both *Neurospora* proteins contained multiple sites under positive selection in this domain. *Fusarium* PRE-2 had one site in this domain but none in PRE-1 while none were found in either of the two *Trichoderma* receptors' C-Terminal. This further support that positive selection is vital in the early stages of receptors evolution amongst closely related species specifically in domains involved in specificity and sensitivity which are likely to lead to speciation.

Because positive selection is often limited to only certain lineages of the phylogenetic tree, we also examined whether specific lineages were more prone to experiencing positive selection, it has been shown that only a few lineages diverge through positive selection (Pond et al. 2011). We were able to identify positive selection on pheromone receptor genes in some lineages of *Trichoderma* and *Neurospora* but not in *Fusarium*. Both PRE-1 and PRE-2 experienced positive selection at node 1 (Figure 9 & 10). One can speculate that this was the stage when divergence due to reproductive isolation occurred between the two clades because of changes in the pheromone receptor gene sequences. The *Neurospora* phylogeny also showed similar results with node 1 (Figure 7 & 8), which separates *N. discreta* from the other species in both *pre-1* and *pre-2*. These results indicate the importance reproductive isolation, specifically due to pheromone receptor genes changes, during species divergence. However, it is

important to note that this is not the only process that leads to speciation, since not all the divergences in these phylogenies showed signs of positive selection.

Another interesting result was that *Trichoderma*, which showed greater sequence divergence and are not inter-fertile, had more positive selected lineages compared to *Neurospora* (with just a few in each gene) and *Fusarium* (which had none in either gene). It is thus possible that the high number of positive selected lineages can be linked to inter-fertility and how this process has influenced inter-sterility amongst *Trichoderma.* The *Neurospora* phylogeny showed differences in the number of positive selected lineages between the two genes, which show that *pre-1* and *pre-2* genes do not always experience the same selection pressure. This observation also supports the notion that one pheromone/receptor pair may be dormant within a population of heterothallic species, which could affect reproductive barriers if the dormant pair was to be reactivated.

This study showed that pheromone receptor genes have evolved in a species-specific manner. In addition, phylogenetic and selection studies showed that these genes experience positive selection pressure and in some cases this may coincide with species divergence and speciation. Future research should determine how much change and at which domains is required to prevent inter-fertility between species.

References

Ainsworth, GC. 2008. Ainsworth and Bisby's Dictionary of the Fungi: CABI.

Anderson, MJ, G Szakacs, J-P Latgé, DW Denning. 2005. Evidence for sexuality in the opportunistic fungal pathogen *Aspergillus fumigatus*. Current Biology 15:1242-1248.

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

Bernsel, A, H Viklund, A Hennerdal, A Elofsson. 2009. TOPCONS: consensus prediction of membrane protein topology. Nucleic Acids Research 37:W465-W468.

Bockaert, J, JP Pin. 1999. Molecular tinkering of G protein-coupled receptors: an evolutionary success. The EMBO Journal 18:1723-1729.

Bölker, M, R Kahmann. 1993. Sexual pheromones and mating responses in fungi. The Plant Cell 5:1461-1469.

Casselton, LA. 2002. Mate recognition in fungi. Heredity 88:142-147.

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

Debuchy, R. 1999. Internuclear recognition: a possible connection between Euascomycetes and Homobasidiomycetes. Fungal Genetics and Biology 27:218-223.

Debuchy, R, BG Turgeon. 2006. Mating-type structure, evolution, and function in Euascomycetes. The Mycota: Growth, Differentiation and Sexuality 1:293-323.

Delport, W, AFY Poon, SDW Frost, SLK Pond. 2010. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. Bioinformatics 26:2455-2457.

Dereeper, A, V Guignon, G Blanc, et al. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Research 36:W465-W469.

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

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

Druzhinina, I, CP Kubicek. 2005. Species concepts and biodiversity in *Trichoderma* and *Hypocrea*: from aggregate species to species clusters? Journal of Zhejiang University Science B 6:100-112.

Dyer, PS, M Paoletti. 2005. Reproduction in *Aspergillus fumigatus*: sexuality in a supposedly asexual species? Medical Mycology 43:S7-S14.

Edgar, RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Research 32:1792-1797.

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

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

Guidon, S, O Gascuel. 2003. A simple, fast and accurate algorithm to estimate large phylogenies by maximum likelihood. Systematic Biology 52:696-704.

Guindon, S, JF Dufayard, V Lefort, M Anisimova, W Hordijk, O Gascuel. 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. Systematic Biology 59:307-321.

Hall, T. 2004. BioEdit version 7.0. 0. Distributed by the author, website: www. mbio. ncsu. edu/BioEdit/bioedit. html.

Karlsson, M, K Nygren, H Johannesson. 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, HK, T Lee, SH Yun. 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, KA Borkovich. 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, KA Borkovich. 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, SJ Wright, G Park, S Ouyang, S Krystofova, KA Borkovich. 2012. Roles for receptors, pheromones, G proteins, and mating-type genes during sexual reproduction in *Neurospora crassa*. Genetics 190:1389-1404.

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

Lee, J, JF Leslie, RL Bowden. 2008. Expression and function of sex pheromones and receptors in the homothallic Ascomycete *Gibberella zeae*. Eukaryotic Cell 7:1211-1221.

Leslie, JF, BA Summerell, S Bullock. 2006. The *Fusarium* laboratory manual: Wiley Online Library.

Leslie, JF, K Zeller, M Wohler, B Summerell. 2004. Interfertility of two mating populations in the *Gibberella Fujikuroi* species complex. European Journal of Plant Pathology 110:611-618.

Li, L, SJ Wright, S Krystofova, G Park, KA Borkovich. 2007. Heterotrimeric G protein signaling in filamentous fungi. Annual Review of Microbiology 61:423-452.

Llopart, A, JM Comeron. 2008. Recurrent events of positive selection in independent *Drosophila* lineages at the spermatogenesis gene roughex. Genetics 179:1009-1020.

Martin, SH, ET Steenkamp, MJ Wingfield, BD Wingfield. 2013. Mate-recognition and species boundaries in the Ascomycetes. Fungal Diversity 58:1-12.

Martin, SH, BD Wingfield, MJ Wingfield, ET Steenkamp. 2011. Causes and consequences of variability in peptide mating pheromones of Ascomycete fungi. Molecular Biology and Evolution 28:1987-2003.

Mayrhofer, S, S Poggeler. 2005. Functional characterization of an alpha-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, JM Weber, S Pöggeler. 2006. Pheromones and pheromone receptors are required for proper sexual development in the homothallic ascomycete *Sordaria macrospora*. Genetics 172:1521-1523.

Nielsen, R, Z Yang. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148:929-936.

Nygren, K, R Strandberg, A Gioti, M Karlsson, H Johannesson. 2012. Deciphering the relationship between mating system and the molecular evolution of the pheromone and receptor genes in *Neurospora*. Molecular Biology and Evolution 29:3827-3842.

Pantou, MP, OK Strunnikova, VY Shakhnazarova, NA Vishnevskaya, VG Papalouka, MA Typas. 2005. Molecular and immunochemical phylogeny of *Verticillium* species. Mycological Research 109:889-902.

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

Poggeler, S, U Kuck. 2001. Identification of transcriptionally expressed pheromone receptor genes in filamentous ascomycetes. Gene 280:9-17.

Pond, SLK, SDW Frost. 2005. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics 21:2531-2533.

Pond, SLK, B Murrell, M Fourment, SDW Frost, W Delport, K Scheffler. 2011. A random effects branch-site model for detecting episodic diversifying selection. Molecular Biology and Evolution 28:3033-3043.

Pond, SLK, SV Muse. 2005. HyPhy: hypothesis testing using phylogenies. Statistical Methods in Molecular Evolution:125-181.

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

Schoch, CL, G-H Sung, F López-Giráldez, JP Townsend, J Miadlikowska, V Hofstetter, B Robbertse, PB Matheny, F Kauff, Z Wang. 2009. The Ascomycota tree of life: a phylum-wide phylogeny clarifies the origin and evolution of fundamental reproductive and ecological traits. Systematic Biology 58:224-239.

Shiu, PKT, NL Glass. 2000. Cell and nuclear recognition mechanisms mediated by mating-type in filamentous Ascomycetes. Current Opinion in Microbiology 3:183-188.

Swanson, WJ, VD Vacquier. 2002a. The rapid evolution of reproductive proteins. Nature Reviews Genetics 3:137-144.

Swanson, WJ, VD Vacquier. 2002b. Reproductive protein evolution. Annual Review of Ecology and Systematics 33:161-179.

Thompson, JD, DG Higgins, TJ Gibson. 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, BG, OC Yoder. 2000. Proposed nomenclature for mating-type genes of filamentous Ascomycetes. Fungal Genetics and Biology 31:1-5.

van Rhee, AM, KA Jacobson. 1996. Molecular architecture of G protein-coupled receptors. Drug Development Research 37:1-38.

Wang, H, Z Xu, L Gao, B Hao. 2009. A fungal phylogeny based on 82 complete genomes using the composition vector method. BMC Evolutionary Biology 9:195.

Wess, J. 1997. G-protein-coupled receptors: molecular mechanisms involved in receptor activation and selectivity of G-protein recognition. The FASEB Journal 11:346.

Whelan, S, N Goldman. 2001. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Molecular Biology and Evolution 18:691-699.

Yang, Z, WJ Swanson, VD Vacquier. 2000. 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.

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

Zhang, N, LA Castlebury, AN Miller, SM Huhndorf, CL Schoch, KA Seifert, AY Rossman, JD Rogers, J Kohlmeyer, B Volkmann-Kohlmeyer. 2006. An overview of the systematics of the Sordariomycetes based on a four-gene phylogeny. Mycologia 98:1076-1087.

Table 1. Accession/taxon numbers, protein ID and names of species used in this study.

Species	Accession number	Pheromone receptors protein ID Pre-1	Pre-2
Acremonium alcalophilum	591952	1081599 [a]	1062940 [a]
Apiospora montagnei	1149870	167715 [a]	437444 [a]
Botrytis cinerea	332648	BC1G_07387 [b]	BC1G_13582 [b]
Chaetomium globosum	306901	CHGG_05819 [b]	CHGG_08469 [b]
Chaetomium thermophilum	759272	340959538	340939445
Cryphonectria parasitica	5116	247536 [a]	268940 [a]
Fusarium verticilliodes	NRRL 20956	FVEG_05310 [b]	FVEG_09280 [b]
Fusarium oxysporum	NRRL 54002	FOZG_05084 [b]	FOZG_10125 [b]
Fusarium graminearum	NRRL 31084	FGSG_07270 [b]	FGSG_02655 [b]
Glomerella cingulata	1314770	1720719 [a]	766503 [a]
Magnaporthe grisea	242507	MGG_06452 [b]	MGG_04711 [b]
Magnaporthe oryzae	242507	440478310	351650355
Myceliophthora thermophila	573729	2138966 [a]	2111571 [a]
Nectria haematococca	660122	302916955	256725842
Neurospora crassa	5141	283468630	283468458
Neurospora hispaniol	588809	8815_pre1 [c]	8815_pre2 [c]
Neurospora intermedia	5142	83468644	283468468

Neurospora metzenbergii	588798	8853_pre1 [c]	8853_pre2 [c]
Neurospora perkinsi	588810	8835_pre1 [c]	8835_pre2 [c]
Neurospora discrete	29879	283468688	8827_pre2 [c]
Neurospora tetrasperma	40127	283468664	283468482
Neurospora sitophila	40126	283468658	283468480
Podospora anserina	515849	170943293	171677600
sordario macrospora	5147	380090235	83318576
Sclerotina sclerotiorum	665079	SS1G_07464 [b]	SS1G_10310 [b]
Thielavia terrestris	578455	2130979 [a]	2127001 [a]
Trichoderma asperellum	1042311	22510 [a]	453452 [a]
Trichoderma atroviride	452589	267561 [a]	275682 [a]
Trichoderma citrinoviride *			
Trichoderma harzianum	983964	13779 [a]	265909 [a]
Trichoderma longibrachiatum	983965	1009536 [a]	1441218 [a]
Trichoderma reesei	1344414	32877 [a]	36457 [a]
Trichoderma virens	413071	227543 [a]	147400 [a]
Verticillium albo-atrum	526221	261360139	302417054
Verticillium dahlia	498257	346971396	346971006

Homologous *pre-1* and *pre-2* protein sequences were downloaded from [a] *JGI: DOE Joint Genome Institute* (http://genome.jgi.doe.gov/), [b] *Broad institute of MIT and Harvard* (http://www.broadinstitute.org/) and [c](Strandberg et al. 2010). The rest were downloaded from NCBI*: National Center for Biotechnology Information* (http://www.ncbi.nlm.nih.gov/). * Sequences were downloaded from *JGI* but subsequently not available anymore.

Table 2. log-likelihood values and parameter estimates for *Fusarium, Neurospora* and *Trichoderma* based on *Nielsen-Yang* (2000) models for *pre-1* and *pre-2* genes.

Taxon	Model	*p*	dN/dS	*ℓ*	VS	Model	*p*	dN/dS	*ℓ*	2Δ*ℓ*
Pre-1										
Fusarium	**M1** (Neutral)	59	0.17	-2961.73		**M2** (Selection)	61	0.17	-2961.66	**0.14**
	M7 (Neutral, β)	59	0.17	-2963.12		**M8** (Selection, β + ω)	61	0.17	-2961.83	**2.58**
Neurospora	**M1** (Neutral)	39	0.33	-5475.81		**M2** (Selection)	41	0.37	-5467.76	**16.10****
	M7 (Neutral, β)	59	0.35	-5478.78		**M8** (Selection, β + ω)	61	0.17	-5478.92	- **0.28**
Trichoderma	**M1** (Neutral)	24	1.00	-7934.13		**M2** (Selection)	26	0.89	-7584.85	**12.50****
	M7 (Neutral, β)	24	0.35	-7566.74		**M8** (Selection, β + ω)	26	0.35	-7566.74	**6.94****
Pre-2										
Fusarium	**M1** (Neutral)	60	0.05	-3108.61		**M2** (Selection)	62	0.05	-3108.61	**0.00**
	M7 (Neutral, beta)	60	0.05	-3118.84		**M8** (Selection, beta + ω)	62	0.05	-3108.38	**20.9****
Neurospora	**M1** (Neutral)	41	0.13	-4789.99		**M2** (Selection)	43	0.13	-4789.99	**0.00**
	M7 (Neutral, β)	41	0.12	-4789.40		**M8** (Selection, β + ω)	43	0.12	-4789.28	**0.24**
Trichoderma	**M1** (Neutral)	26	0.45	-6871.47		**M2** (Selection)	28	0.51	-6870.34	**2.26**
	M7 (Neutral, β)	26	0.43	-6872.72		**M8** (Selection, β + ω)	28	0.43	-6861.66	**22.12****

P is the number of free estimated parameters of **ω** generated by HyPhy (Pond, Muse 2005). ***ℓ*** is the log likelihood ratio scores for each model, as calculated by HyPhy. Significant difference between models indicated by **. **dN/dS** is the average ratio over the entire gene.

Figure 1. A maximum likelihood phylogenetic tree of Sordariomycetes inferred from *pre-1* amino acid sequences. Vertical bars indicate the different orders and bootstrap values greater than 75% are indicated at the internodes. The tree is rooted with *Botrytis cinerea* and *Sclerotina sclerotiorum*.

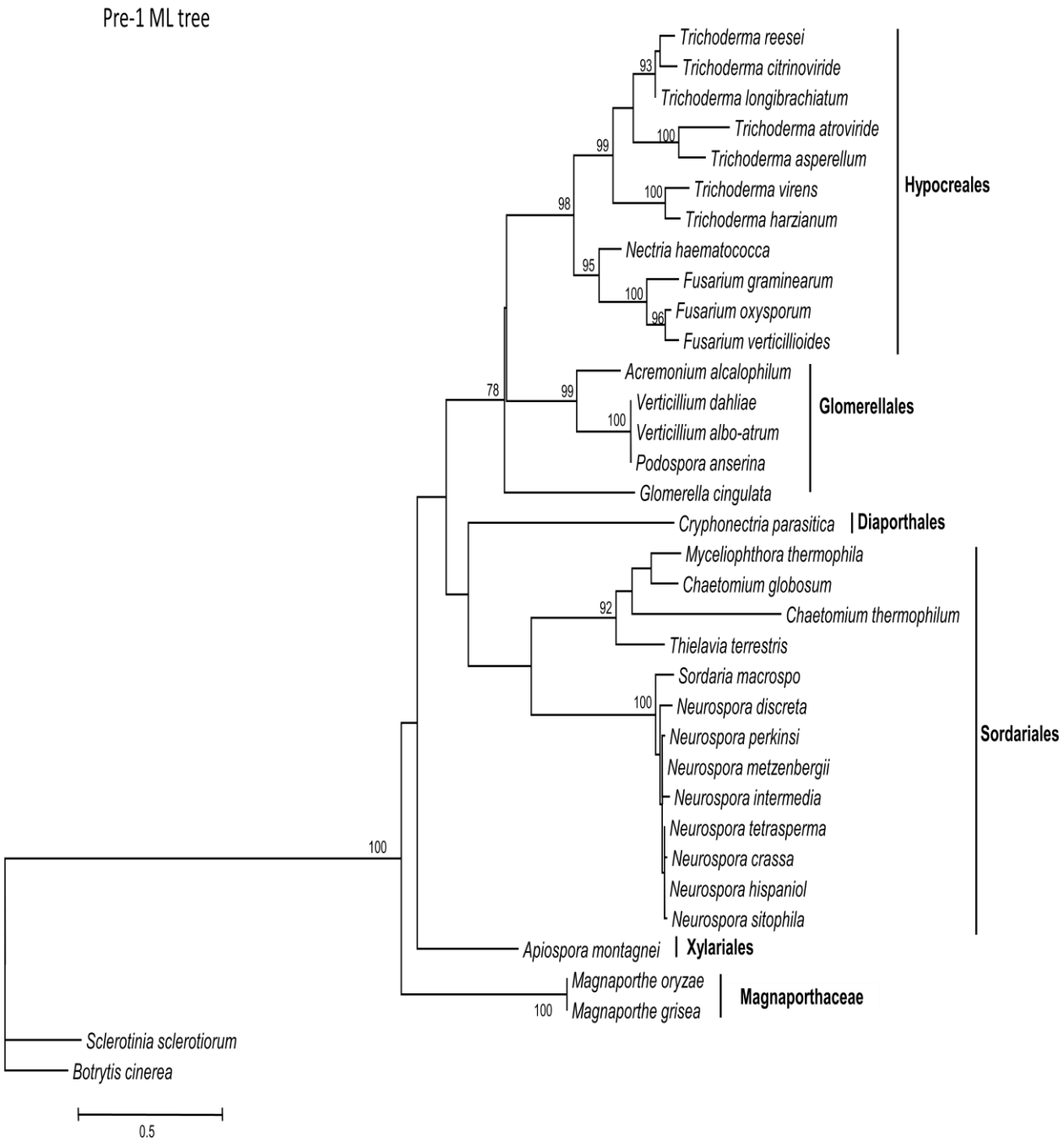


Figure 2. A maximum likelihood phylogenetic tree of Sordariomycetes inferred from *pre-2* amino acid sequences. Vertical bars indicate the different orders and bootstrap values greater than 75% are indicated at the internodes. The tree is rooted with *Botrytis cinerea* and *Sclerotina sclerotiorum*.

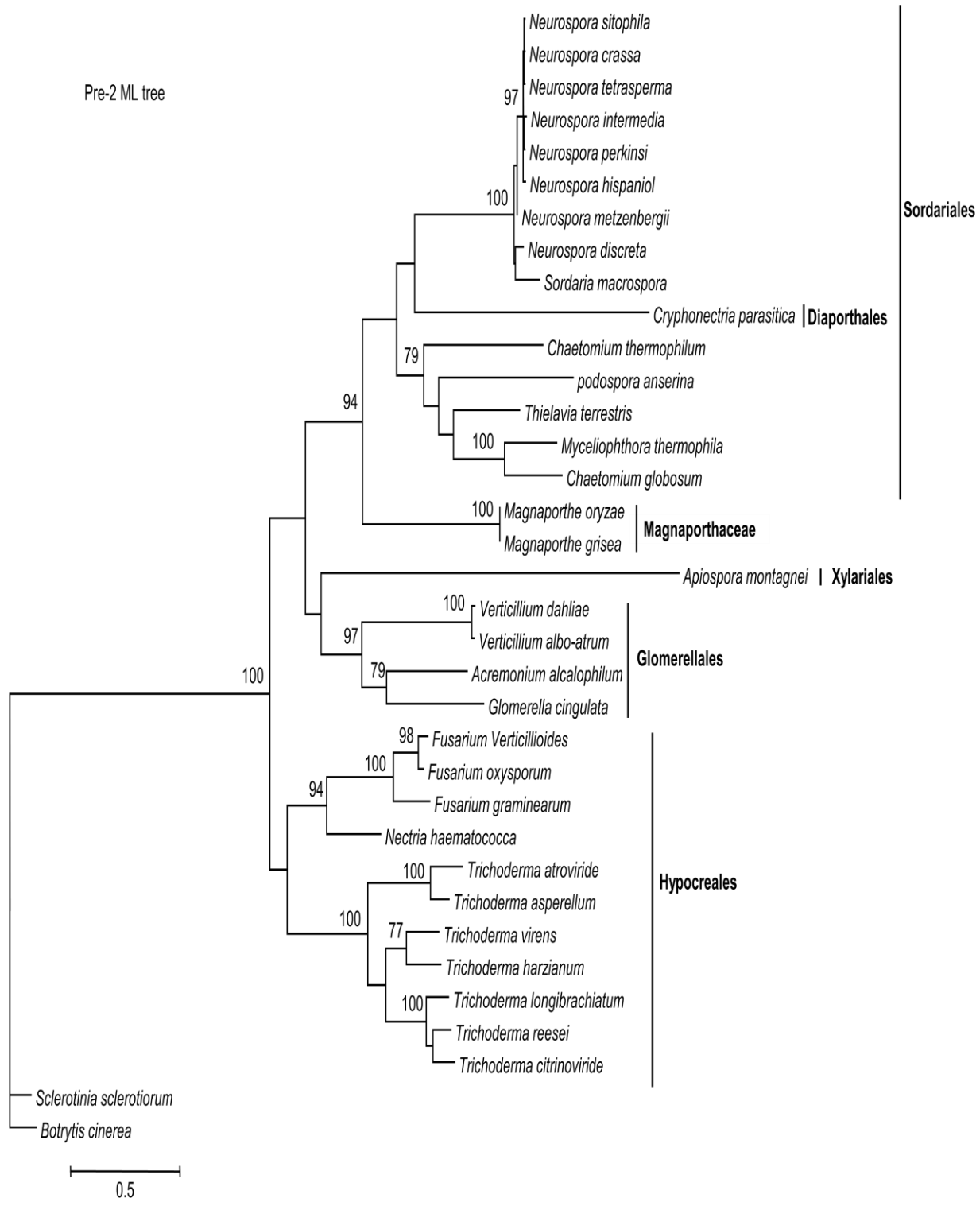


Figure 3. The 2D structures of PRE-1 receptor constructed using TMRPres2D v0.91 (Spyropoulos et al. 2004) showing positive selected sites and the different domains of the protein. The domains were predicted using amino acid sequences with TOPCONS (Bernsel et al. 2009) an online Consensus prediction membrane protein topology program.

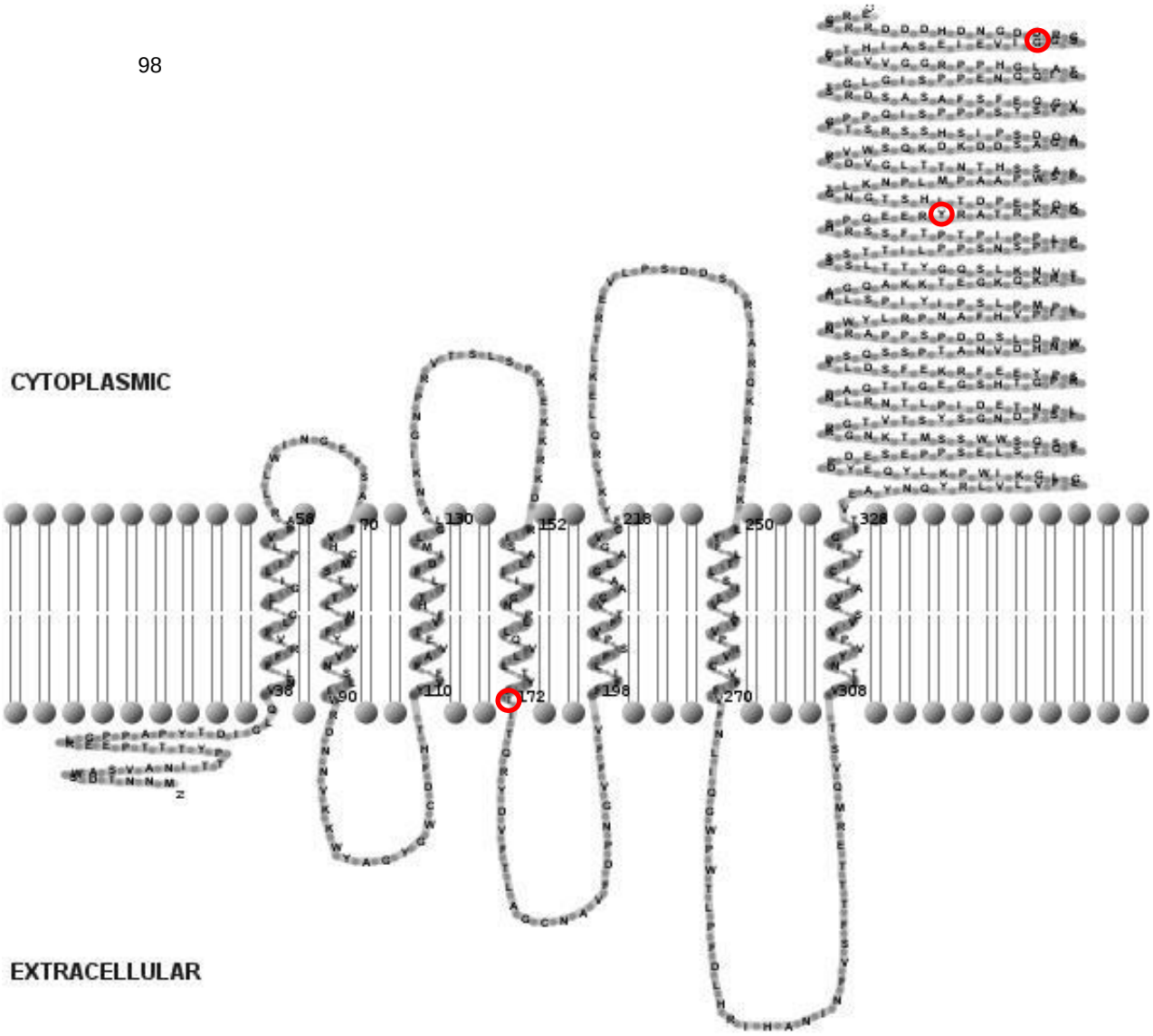


Figure 4. The 2D structures of PRE-2 receptor constructed using TMRPres2D v0.91 (Spyropoulos et al. 2004) showing positive selected sites and the different domains of the protein. The domains were predicted using *Neurospora* amino acid sequences with TOPCONS (Bernsel et al. 2009) an online Consensus prediction membrane protein topology program.

Glomerellales

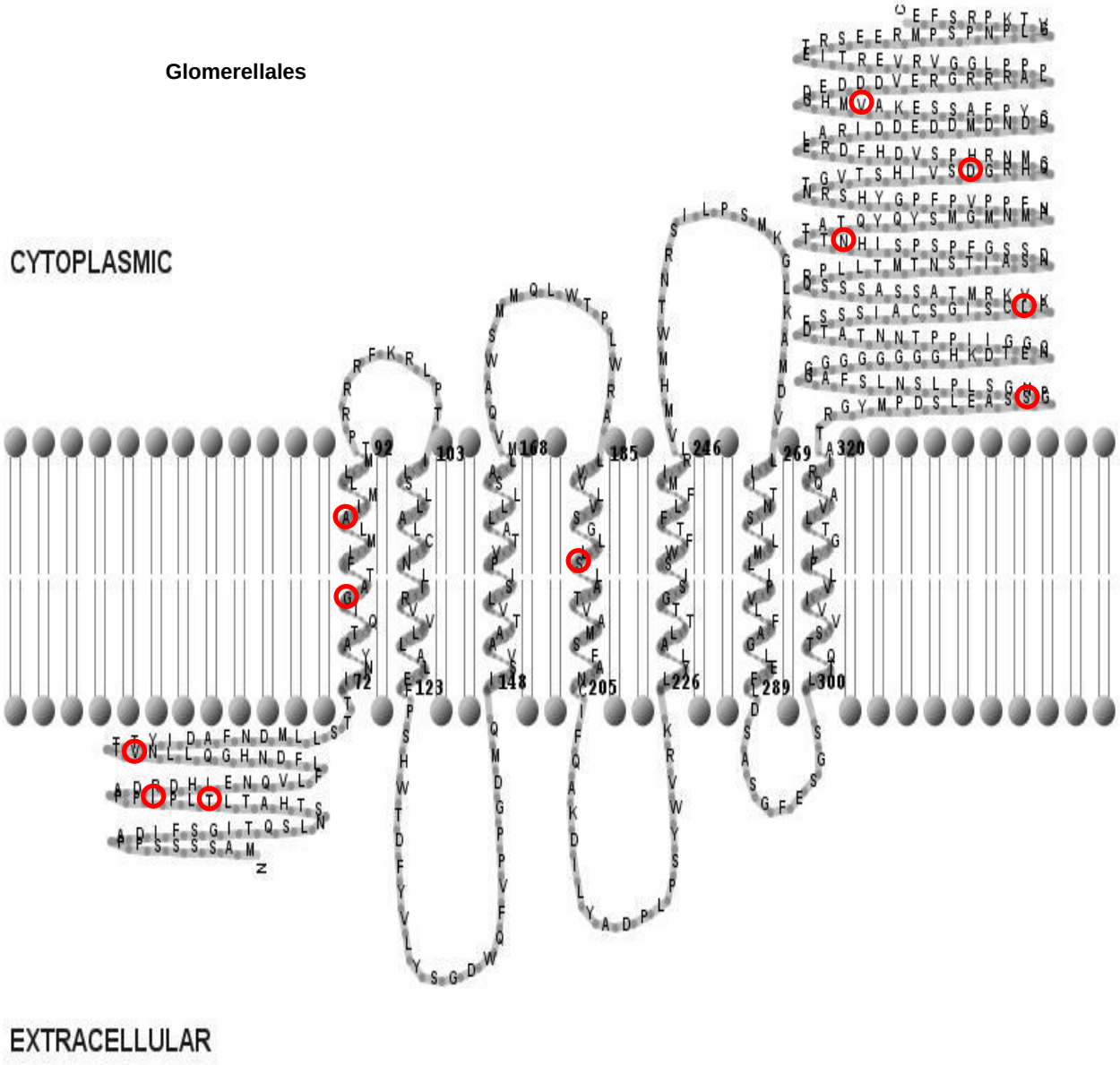


Figure 5. The 2D structures of PRE-1 receptor constructed using TMRPres2D v0.91 (Spyropoulos et al. 2004) showing positive selected sites and the different domains of the protein. The domains were predicted using *Trichoderma* amino acid sequences with TOPCONS (Bernsel et al. 2009) an online Consensus prediction membrane protein topology program.

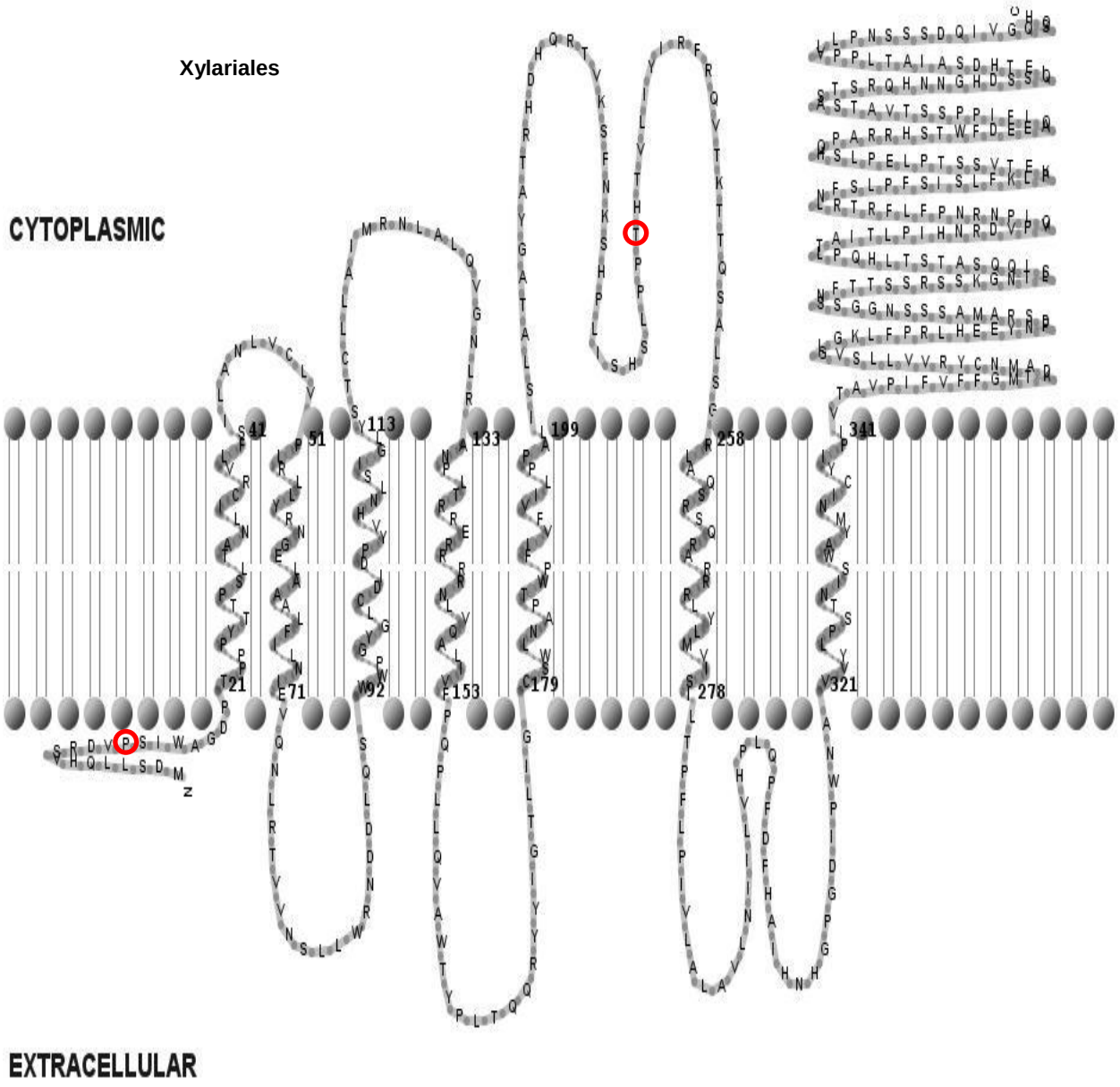


Figure 6. The 2D structures of PRE-2 receptor constructed using TMRPres2D v0.91 (Spyropoulos et al. 2004) showing positive selected sites and the different domains of the protein. The domains were predicted using *Trichoderma* amino acid sequences with TOPCONS (Bernsel et al. 2009) an online Consensus prediction membrane protein topology program.

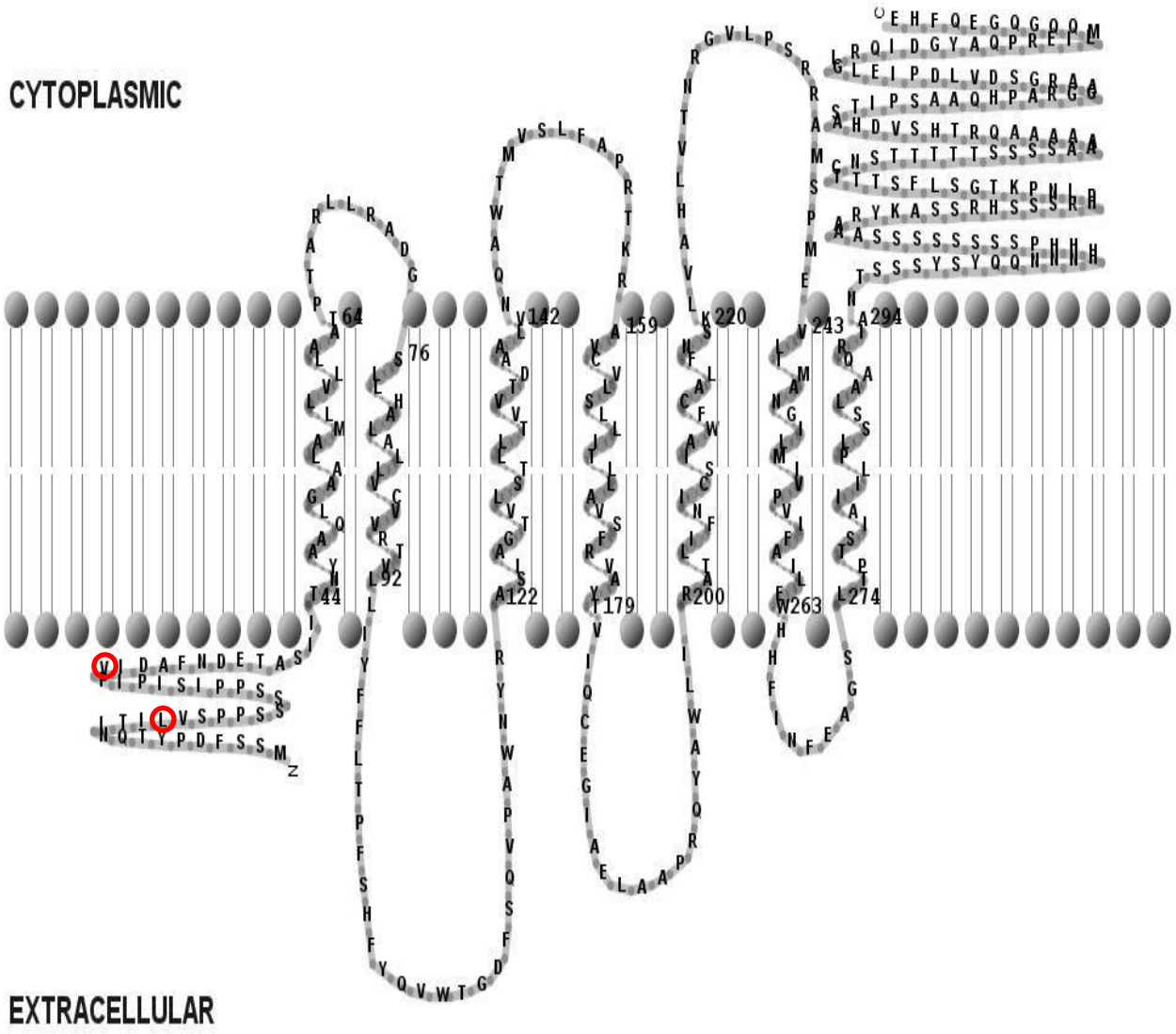


Figure 7. Phylogenetic tree of *Neurospora* species showing lineages that experienced episodic selection during the evolution of *pre-1* genes. Branches that are coloured red showed evidence of positive selection based on Branch-Site REL method (Pond et al. 2011).

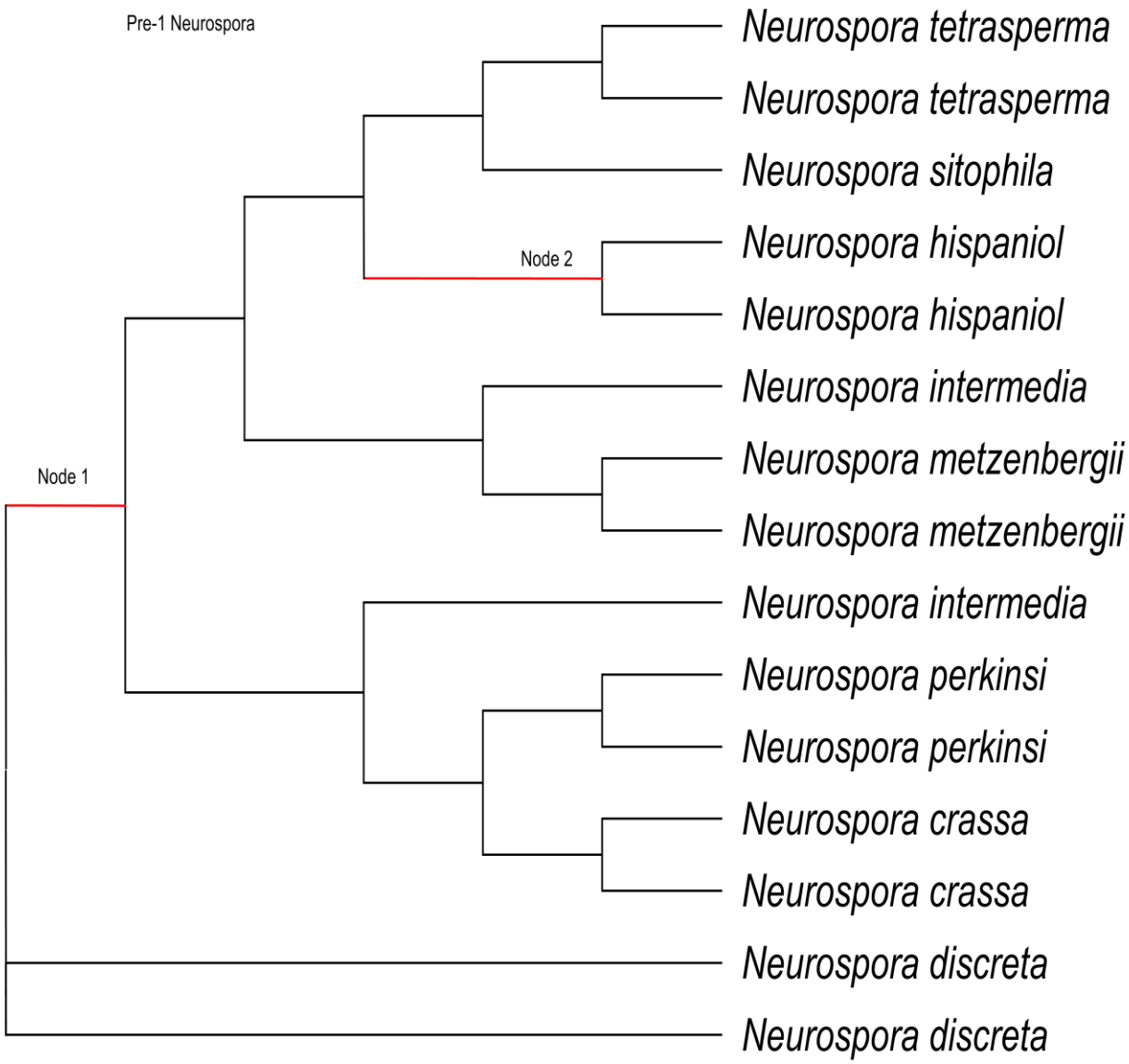


Figure 8. Phylogenetic tree of *Neurospora* species showing lineages that experienced episodic selection during the evolution of *pre-2* genes. Branches that are colored red showed evidence of positive selection based on Branch-Site REL method (Pond et al. 2011).

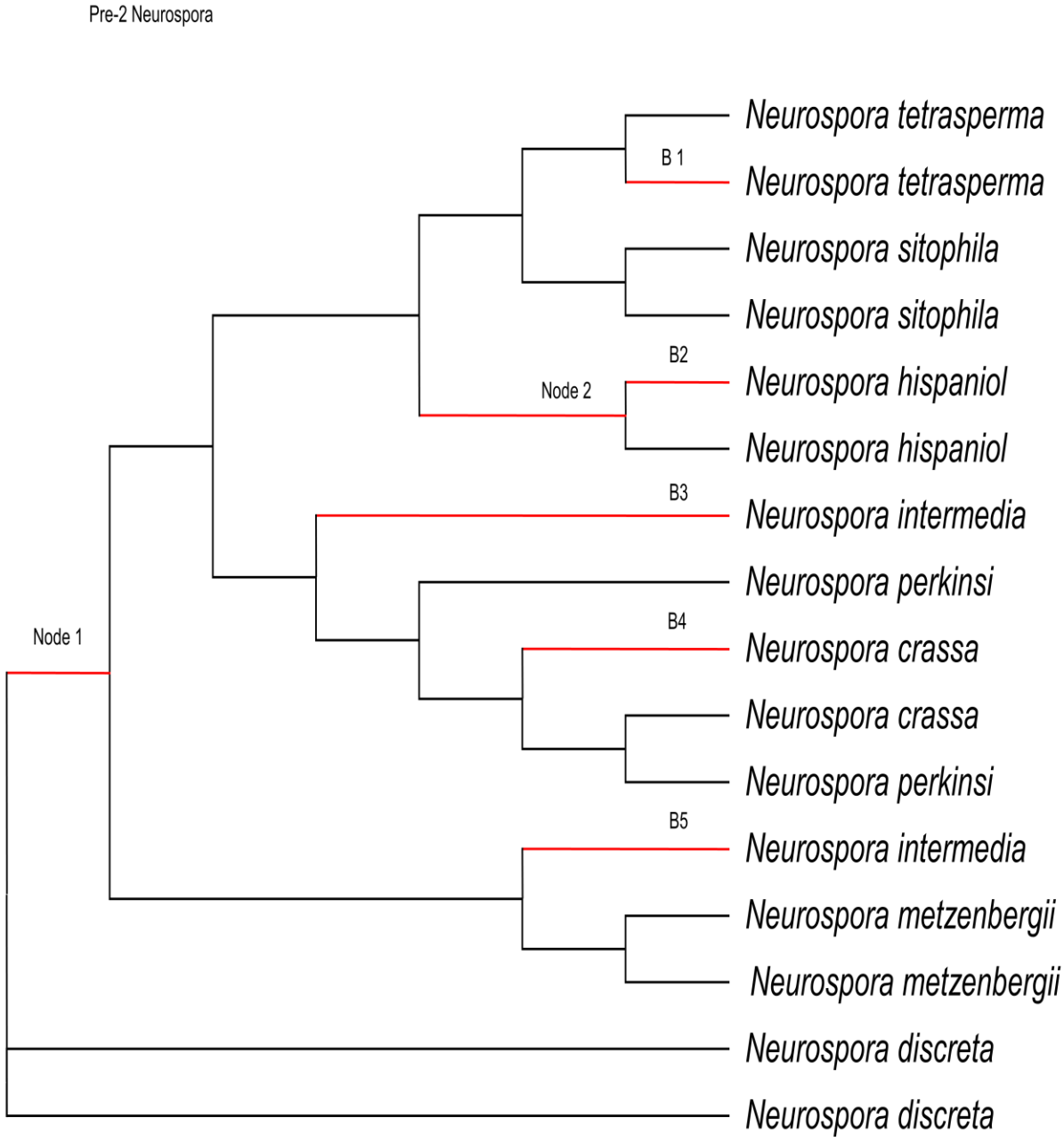


Figure 9. Phylogenetic tree of *Trichoderma* species showing lineages that experienced episodic selection during the evolution of *pre-1* genes. Branches that are colored red showed evidence of positive selection based on Branch-Site REL method (Pond et al. 2011).

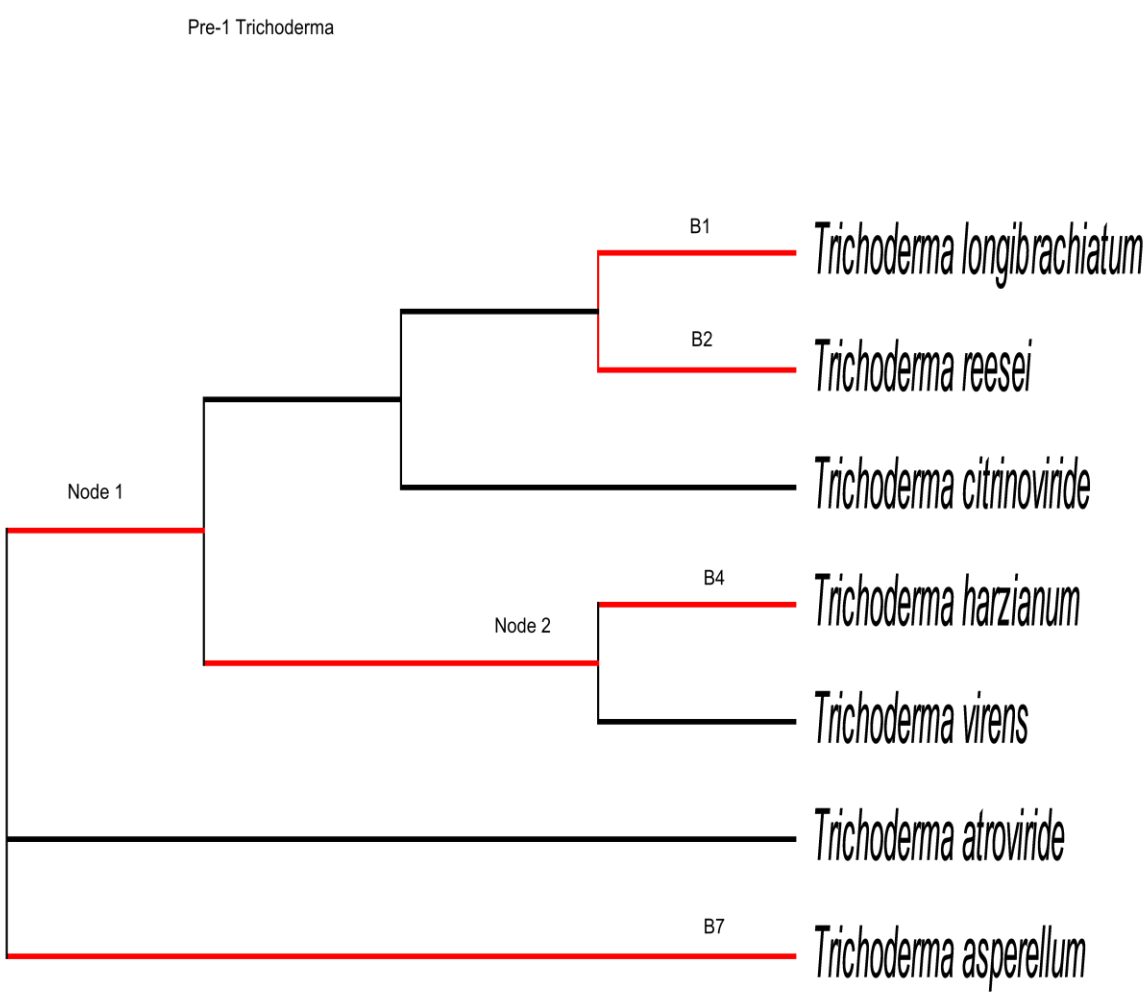
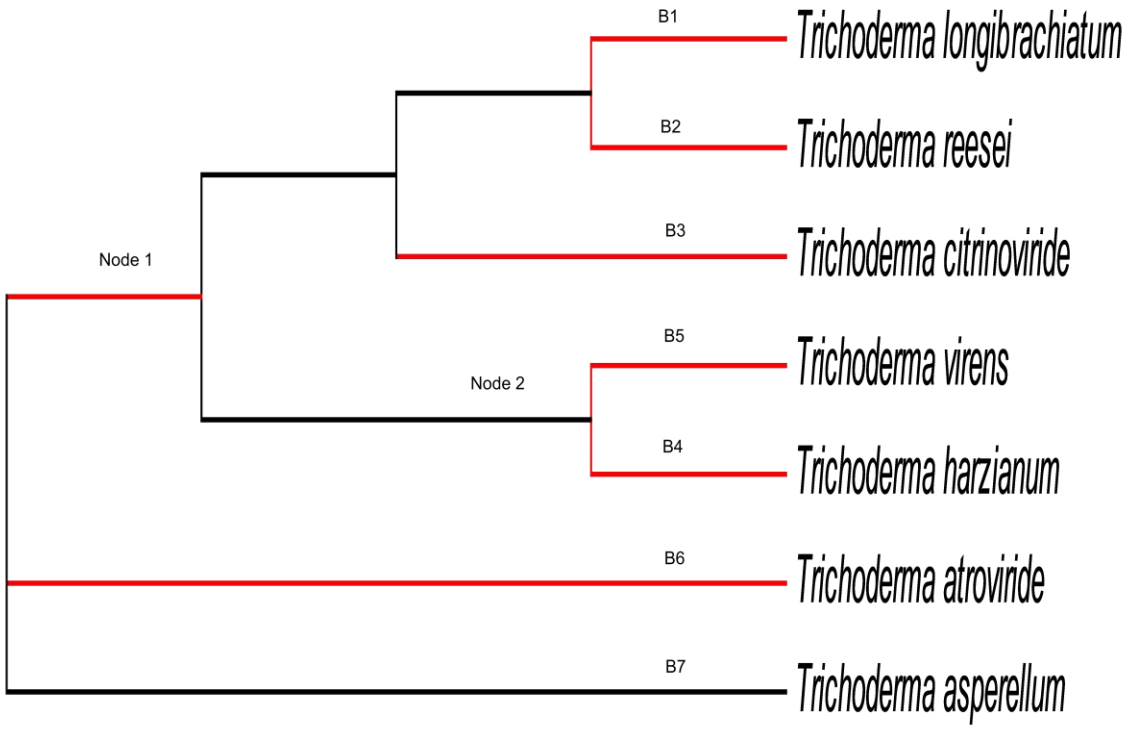
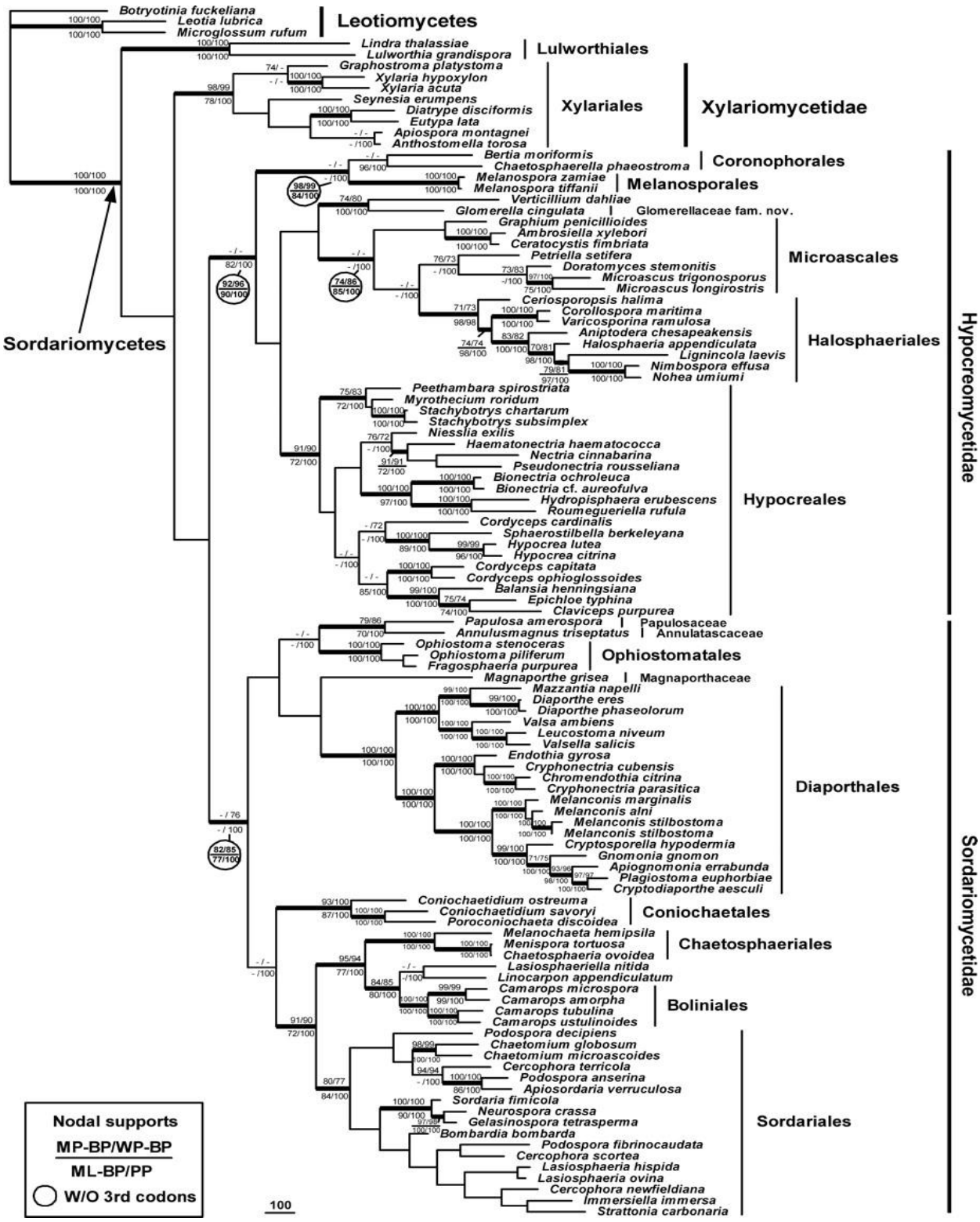


Figure 10. Phylogenetic tree of *Trichoderma* species showing lineages that experienced episodic selection during the evolution of *pre-2* genes. Branches that are colored red showed evidence of positive selection based on Branch-Site REL method (Pond et al. 2011).

Pre-2 Trichoderma



Supplementary Figure 1. Phylogenetic tree of Sordariomycetes constructed by Zhang et al. (2006) from single most parsimonious tree from weighted parsimony analysis. Bootstrap proportions (MP-BP, WP-BP and ML-BP) of ≥ 70% and posterior probabilities (PP) of ≥ 95% are provided for the corresponding internodes. Bootstrap proportions and posterior probabilities in circles were calculated with the exclusion of the third codon positions in TEF and RPB2. The internodes that are supported by all bootstrap proportions and posterior probabilities are shown as a thicker line. The name of taxa (e.g. subclasses, orders, families) is provided to the right of species names.



Summary

Inter-fertility in fungi is prevented by either pre-mating or post-mating barriers. The Ascomycota have two classes of pheromone-receptors, PRE-1 and PRE-2 that form part of the pre-mating barrier. These receptors bind to specific pheromones secreted by fungal cells to initiate mating. Pheromones diversity of Ascomycete species was recently explored and this study aimed to characterize the pheromone receptor genes from the Ascomycota with main focus on the Sordariomyctes and to investigate the evolution of these genes. Some of these Sordariomycetes are pathogenic affecting with a vast range of hosts and therefore of economic importance. PRE-1 and PRE-2 receptors from three taxa, *Fusarium*, *Neurospora* and *Trichoderma* were utilized in the selection analysis studies. Results showed differences in selection profiles in each genus and this seemed to be related to inter-fertility and this is likely to be true for all Ascomycetes taxa. Using previously characterized receptors from other ascomycetes, the phylogeny of the receptor genes was inferred and compared with the species phylogeny. The receptors phylogeny was congruent with the species phylogeny and based on the genetic variation, there receptors are likely to confer mating-specificity. This study was not able to solve the question of how inter-fertility is limited to only a few species however it improved our understanding of the pheromone-receptor systems in Ascomycete species.