EVOLUTION OF THE MATING-TYPE LOCUS AND INSIGHTS INTO SEXUAL REPRODUCTION IN THE CRYPTOCOCCUS SPECIES COMPLEX

by

Keisha Monique Findley

Department of Molecular Genetics and Microbiology
Duke University

Date:_______________________
Approved:

Joseph Heitman, MD, PhD, Supervisor

___________________________
Fred Dietrich, PhD

___________________________
Daniel Lew, PhD

___________________________
John McCusker, PhD

___________________________
Rytas Vilgalys, PhD

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Molecular Genetics and Microbiology in the Graduate School of Duke University

2010
ABSTRACT

EVOLUTION OF THE MATING-TYPE LOCUS AND INSIGHTS INTO SEXUAL REPRODUCTION IN THE CRYPTOCOCCUS SPECIES COMPLEX

by

Keisha Monique Findley

Department of Molecular Genetics and Microbiology
Duke University

Date:_______________________
Approved:

Joseph Heitman, MD, PhD, Supervisor

Fred Dietrich, PhD

Daniel Lew, PhD

John McCusker, PhD

Rytas Vilgalys, PhD

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Molecular Genetics and Microbiology in the Graduate School of Duke University

2010
Abstract

Sexual reproduction in fungi is governed by a specialized genomic region called the mating-type locus (MAT). The ascomycetes, the largest phylum of fungi, primarily possess a bipolar mating system while the basidiomycetes, the second largest group, are mostly tetrapolar. The human fungal pathogen and basidiomycetous yeast Cryptococcus neoformans has evolved a bipolar mating system that encodes homeodomain (HD) and pheromone/receptor (P/R) genes. The MAT locus of C. neoformans is unusually large, spans greater than 100 kb, and encodes more than 20 genes. To understand how the pathogenic Cryptococcus species complex evolved this unique bipolar mating system, we investigated the evolution of MAT in closely and distantly related species and discovered an extant sexual cycle in Cryptococcus amylolentus.

Phylogenetic analysis using a six-gene multi-locus sequencing (MLS) approach identified the most closely related species to the pathogenic Cryptococcus species complex that are currently known. The two non-pathogenic sibling species, Tsuchiyaea wingfieldii and Cryptococcus amylolentus, and the more distantly related species Filobasidiella depauperata define the Filobasidiella clade. We also resolved the phylogeny of the species located in the sister clade, Kwoniella. A comprehensive tree dendrogram revealed that the 15 Tremellales species examined suggests a common saprobic ancestor. Moreover, the pathogenic Cryptococcus species have a saprobic origin but later emerged as pathogens. We further characterized the mating-type locus for T. wingfieldii and C. amylolentus by cloning and sequencing two unlinked genomic loci encoding the HD and P/R genes. Interestingly, linked and likely divergently transcribed
homologs for \textit{SXI1} and \textit{SXI2} are present in \textit{T. wingfieldii} and \textit{C. amylolentus}, while the P/R alleles contain many genes also found in the \textit{MAT} locus of the pathogenic \textit{Cryptococcus} species. Also, hypothetical genes present in \textit{C. neoformans} \textit{MAT} are also \textit{MAT}-linked in both species and indicate a possible translocation event between chromosomes 4 and 5 of \textit{C. neoformans}. Our analysis of \textit{MAT} in the sibling species indicates that \textit{T. wingfieldii} is likely tetrapolar, and the \textit{C. amylolentus} sequence comparison of the dimorphic \textit{SXI1} and \textit{SXI2} region and the pheromone receptor, \textit{STE3}, suggests that \textit{C. amylolentus} is also tetrapolar. The examination of \textit{MAT} in these sibling species confirms the model for \textit{MAT} evolution previously proposed in which this structure in \textit{C. neoformans} and \textit{C. gattii} evolved from an ancestral tetrapolar mating system. Moreover, the organization of \textit{MAT} in these sibling species mirrors key aspects of the proposed intermediates in the evolution of \textit{MAT} in the pathogenic \textit{Cryptococcus} species, and for sex chromosomes in plants, animals, and algae in general.

We discovered an extant sexual cycle for \textit{C. amylolentus}, a species previously thought to be asexual. Matings between two strains of opposite mating-types produce dikaryotic hyphae with fused clamp connections and uni- and bi-nucleate basidiospores. Genotyping of basidiospores using markers linked and unlinked to \textit{MAT} revealed that genetic exchange (recombination) occurs during the sexual cycle of \textit{C. amylolentus}, and it is likely that either aneuploids are generated during sex or more than one meiosis event occurs within each basidium. This is in contrast to \textit{C. neoformans}, where only one meiotic event per basidium has been observed. Uniparental mitochondrial inheritance has also been observed in \textit{C. amylolentus} progeny; similar to the pathogenic \textit{Cryptococcus}
species, mtDNA is inherited from the *C. amylolentus* MATa parent. Analysis of sex in *C. amylolentus* has provided insight into the mechanisms that phylogenetically related fungi employ in orchestrating sexual reproduction.

We also extended our analysis to include the distantly related tetrapolar basidiomycete *Tremella mesenterica*. We completed comparisons of MAT-specific genes between five strains of *T. mesenterica* and identified the regions that define its mating-type system. The HD locus is limited to the *SXII*- and *SXII*-like genes while the P/R locus is defined by *STE3, STE12, STE20*, and the pheromone gene, tremerogen a-13. Interestingly, many of the genes associated with the MAT locus of the pathogenic *Cryptococcus* species flank the HD and P/R locus and are not incorporated in MAT in *T. mesenterica*. The MAT region includes transposons and *C. neoformans* hypothetical genes also present in *T. wingfieldii* and *C. amylolentus*. The mating-type system in *T. mesenterica* reflects an ancestral intermediate in the evolution of the MAT locus in the pathogenic *Cryptococcus* species. In conclusion, this study provides an in-depth analysis on the structure, function, and evolution of an unusual mating-type locus with broader implications for the transitions in modes of sexual reproduction in fungi that impact gene flow in populations.
Dedication

This dissertation is dedicated to my mother, Maureen Findley, who instilled in me from an early age that there is no reward without labor. Because of you, I have made it this far. Additionally, I thank my family, friends, and most importantly, the Partners of Compassion for their unwavering support throughout this process.
Contents

Abstract.............................................................................................................................................. iv

List of Tables ........................................................................................................................................ xii

List of Figures ....................................................................................................................................... xiii

Acknowledgements ............................................................................................................................ xv

Chapter 1. Introduction ......................................................................................................................... 1

1.1 Cryptococcus as a pathogen .......................................................................................................... 1

1.2 The life cycle of C. neoformans ................................................................................................. 4

1.3 The evolution of the mating-type locus (MAT) in the Cryptococcus species complex .......... 5

1.4 Taxonomy of Cryptococcus species and the molecular phylogeny of the pathogenic Cryptococcus species ........................................................................................................ 9

1.5 Overview of this work ............................................................................................................... 11

Chapter 2. Phylogeny and phenotypic characterization of pathogenic Cryptococcus species and closely related saprobic taxa in the Tremellales ......................................................................... 14

Chapter 2 is modified from a published manuscript (38). ............................................................... 14

2.1 Introduction .................................................................................................................................. 14

2.2 Materials and methods .............................................................................................................. 19

2.2.1 Fungal isolates ....................................................................................................................... 19

2.2.2 DNA isolation, PCR, and sequencing .................................................................................... 19

2.2.3 Phylogenetic analysis ............................................................................................................ 20

2.2.4 Phenotypic analysis of isolates ............................................................................................ 21

2.2.5 Galleria mellonella survival assay ......................................................................................... 21

2.3 Results ......................................................................................................................................... 23

2.3.1 Cryptococcus amylolentus and Tsuchiyaea wingfieldii are the species most closely related to C. neoformans ............................................................................................................. 23
2.3.2 Phenotypic and morphological differences in close and distant relatives of C. neoformans .......................................................... 26

2.3.3 Virulence potential of Cryptococcus and neighboring taxa using G. mellonella. 29

2.4 Discussion ................................................................................................................. 32

Chapter 3. Evolution of MAT in the pathogenic Cryptococcus species and neighboring taxa, Tsuchiyaea wingfieldii and Cryptococcus amylolentus .............................................................. 40

3.1 Introduction ................................................................................................................. 40

3.2 Materials and methods .............................................................................................. 44

3.2.1 Strains and media ................................................................................................. 44

3.2.2 DNA extraction .................................................................................................... 44

3.2.3 Degenerate PCR .................................................................................................. 44

3.2.4 Fosmid library preparation and fosmid library screening .................................... 46

3.2.5 Sequencing and assembly ...................................................................................... 47

3.2.6 Fluorescence Activated Cell Sorting (FACS) analysis .......................................... 47

3.2.7 Pulsed-field gel electrophoresis (PFGE) and chromoblot analysis .................... 47

3.2.8 Bioinformatic analyses ......................................................................................... 48

3.2.9 Phylogenetic analysis .......................................................................................... 49

3.2.10 Southern blot analysis ....................................................................................... 49

3.3 Results ....................................................................................................................... 51

3.3.1 Characterizing MAT in T. wingfieldii and C. amylolentus .............................. 51

3.3.2 The HD and P/R loci are physically unlinked in the sibling species ............ 57

3.3.3 Identification of key genes that define MAT ..................................................... 58

3.3.4 Synteny across MAT in sibling species .............................................................. 63

3.3.5 Phylogenetic analysis of genes in C. amylolentus and T. wingfieldii ............. 65

3.4 Discussion ................................................................................................................. 69

Chapter 4. Discovery of an extant sexual cycle for Cryptococcus amylolentus ......... 79
4.1 Introduction ................................................................................................................................. 79
4.2 Materials and methods .............................................................................................................. 83
  4.2.1 Strains and media .................................................................................................................. 83
  4.2.2 Description of sexual cycle as Filobasidiella amylolenta .................................................... 83
  4.2.3 Microscopy .......................................................................................................................... 84
  4.2.4 Fluorescence Activated Cell Sorting (FACS) analysis ......................................................... 85
  4.2.5 Genotyping .......................................................................................................................... 85
  4.2.6 RAPD PCR and RFLP analysis ............................................................................................ 86
4.3 Results ........................................................................................................................................ 87
  4.3.1 Discovery of an extant sexual cycle in C. amylolentus ......................................................... 87
  4.3.2 Genotypic analysis of meiotic progeny ............................................................................... 93
  4.3.3 RAPD confirms that meiosis occurs during the C. amylolentus sexual cycle .................... 106
  4.3.4 Uniparental mitochondrial DNA inheritance ....................................................................... 108
4.4 Discussion .................................................................................................................................. 114

Chapter 5. Definition of the mating-type locus of the tetrapolar basidiomycete Tremella mesenterica ........................................................................................................................................ 123
  5.1 Introduction .............................................................................................................................. 123
  5.2 Materials and methods ........................................................................................................... 127
    5.2.1 Strains and media .............................................................................................................. 127
    5.2.2 DNA extraction ................................................................................................................. 128
    5.2.3 Primer design, PCR, and sequencing .............................................................................. 128
    5.2.4 Bioinformatic analyses .................................................................................................... 129
    5.2.5 Phylogenetic analysis ....................................................................................................... 129
  5.3 Results ...................................................................................................................................... 132
    5.3.1 Characterization of MAT in Tremella mesenterica ......................................................... 132
    5.3.2 Defining the boundaries of MAT in T. mesenterica .......................................................... 136
5.3.3 Synteny between T. mesenterica and C. neoformans ........................................ 141
5.3.4 Phylogenetic analysis of T. mesenterica genes ............................................. 142
5.3.5 Exploration of sex in T. mesenterica .......................................................... 144
5.4 Discussion ........................................................................................................ 150

Chapter 6. Conclusions and Future Directions ......................................................... 155

6.1 T. wingfieldii and C. amylolentus are sibling species and the closest relatives to the pathogenic Cryptococcus species cluster .................................................. 156
6.2 The MAT loci of T. wingfieldii and C. amylolentus are physically unlinked and lie on different chromosomes ................................................................. 157
6.3 Sexual reproduction in C. amylolentus – insights into mating in the pathogenic Cryptococcus species complex ................................................................. 160
6.4 The unlinked HD and P/R loci in T. mesenterica span ~3 kb and 18 kb, respectively ................................................................. 164

Appendix A ............................................................................................................. 167
Appendix B ............................................................................................................. 168
Appendix C ............................................................................................................. 169
Appendix D ............................................................................................................. 170
References ............................................................................................................. 171
Biography .............................................................................................................. 188
List of Tables

Table 1: Table of species in Chapter 1.................................................................................................................. 38

Table 2: Genbank accession numbers.............................................................................................................. 39

Table 3: Genbank accession numbers for T. wingfieldii and C. amylolentus ....................................................... 78

Table 4: List of genotyping markers (nuclear and mitochondrial). ..................................................................... 92

Table 5: Random spore dissection of set 1 F1 progeny and molecular analysis of the nuclear markers......................... 96

Table 6: Individual spore dissection of set 2 F1 and F2 progeny and molecular analysis of the nuclear markers.............................................. 97

Table 7: List of primers, RAPD and mitochondrial markers, used in the study................................................. 105

Table 8: RAPD analysis of set 1 F1 progeny..................................................................................................... 110

Table 9: RAPD analysis of set 2 F1 and F2 progeny including genotypes of individual spore chains.............................. 111

Table 10: Filamentous phenotype and mitochondrial DNA (mtDNA) inheritance identified in all 65 progeny......................................................... 112

Table 11: Description of species in Chapter 5.................................................................................................... 130

Table 12: List of primers used in Chapter 5 to identify the boundaries of MAT in T. mesenterica.......................... 131

Table 13: T. mesenterica Genbank accession numbers...................................................................................... 149
List of Figures

Figure 1: Phylogenetic relationship among members of the Tremellales......................... 24
Figure 2: Phylogram indicating the preferred habitats of the fungal species used in this study... 25
Figure 3: Phenotype and morphology of species grown under different conditions. .......... 28
Figure 4: Infection of the heterologous host *G. mellonella* with species closely related to *C. neoformans*.......................................................................................................................... 29
Figure 5: Infection of the heterologous host *G. mellonella* with species distantly related to *C. neoformans*.......................................................................................................................... 30
Figure 6: *T. wingfieldii* MAT loci and chromosomal locations.................................. 54
Figure 7: *C. amylolentus* MAT loci and chromosomal locations............................... 55
Figure 8: Alignment of the mating pheromone genes in the pathogenic *Cryptococcus* species and related taxa................................................................. 58
Figure 9: The *RPL22* gene is duplicated in the *C. amylolentus* genome. .................. 60
Figure 10: The homedomain genes, *SXII* and *SXII2*, define MAT.......................... 62
Figure 11: The pheromone receptor, *STE3*, defines MAT........................................ 63
Figure 12: Synteny exists across MAT in the sibling species. .................................... 64
Figure 13: Extensive inversions and gene rearrangements are present throughout MAT in the sibling species and *C. neoformans* strain JEC21. ................................................................. 65
Figure 14: Phylogenetic patterns of *C. amylolentus* genes........................................ 67
Figure 15: Phylogenetic patterns of additional *C. amylolentus* genes...................... 68
Figure 16: Model for the evolution of the mating-type locus in the pathogenic *Cryptococcus* species........................................................................................................................... 74
Figure 17: Sexual reproduction of *C. amylolentus*. .................................................. 90
Figure 18: Fluorescence microscopy of *C. amylolentus* mating structures................ 91
Figure 19: A model comparing the fungal bipolar and tetrapolar mating paradigms.... 94
Figure 20: Yeasts, blastospores, and basidiospores are produced during sex in *C. neoformans* and *C. amylolentus*........................................................................................................ 98
Figure 21: RFLP analysis of SXI1 and SXI2 in *C. amylolentus* progeny................................. 99

Figure 22: Set 2 F1 progeny #16 is recombinant and types as CBS6039 at SXI1 and as CBS6273 at SXI2. .................................................................................................................. 100

Figure 23: The SXI1 and SXI2 dimorphic region is similar in CBS6039 and set 2 F1 progeny #4. .................................................................................................................................. 101

Figure 24: The SXI1 and SXI2 dimorphic region differs between CBS6273 and progeny set 2 F1 progeny #16. ................................................................................................................................ 102

Figure 25: Nine unique, polymorphic RAPD bands identified....................................................... 107

Figure 26: RAPD primer, JOHE22656 – No. 1 produces a unique band in a subset of the progeny. ........................................................................................................................................... 109

Figure 27: *C. amylolentus* has a tetrapolar mating system. ............................................................ 113

Figure 28: *Tremella mesenterica* environmental isolate.............................................................. 132

Figure 29: Sequence assemblies spanning *MAT* in the sequenced strain *T. mesenterica* ATCC24925. ................................................................................................................................. 134

Figure 30: Alignment of an unlinked *MAT* gene, *RPO41*. ....................................................... 137

Figure 31: Alignment of an unlinked *MAT* gene, *GEF1*. ........................................................ 138

Figure 32: Comparison of *T. mesenterica* strain ATCC24925 (*a* allele) to CBS8219 (*A* allele) limits the P/R locus to ~18 kb................................................................. 139

Figure 33: Comparison of *T. mesenterica* strain ATCC24925 (*a* allele) to CBS8221 (*A* allele) limits the P/R locus to ~18 kb................................................................. 141

Figure 34: Comparison of *T. mesenterica* strain ATCC24925 (*BII* allele) and CBS8219 (*B* allele) restricts the HD locus to ~3 kb................................................................. 143

Figure 35: Synteny between *MAT* loci of *T. mesenterica* and *C. neoformans* var. *neoformans*. 144

Figure 36: Phylogenetic patterns of *T. mesenterica* genes. .......................................................... 146

Figure 37: Confrontation assay in CBS8219 and ATCC24925........................................................ 147

Figure 38: Confrontation assay in CBS8221 and ATCC24925........................................................ 148
Acknowledgements

I would like to thank everyone who helped to make this dissertation possible. I wish to thank my advisor, Dr. Joseph Heitman, for his guidance, support, and patience throughout the course of my graduate training. I would like to thank my committee members for their constructive criticism and advice in making this analysis comprehensive. I also acknowledge Dr. James Fraser for his initial training during the early part of my career. I thank Tim James for MLS primers and Marianela Rodriguez-Carres for advice and scientific discussion. I also thank Robyn Hicks, Anna Floyd Averette, Wenjun Li, Sheng Sun, and Yen-Ping Hsueh for technical assistance. Lastly, I wish to thank the Biological Sciences Sequencing Facility and the JGI for sequencing assistance and Dr. K. W. Gams for the Latin translation in Chapter 4.
Chapter 1. Introduction

1.1 Cryptococcus as a pathogen

The encapsulated basidiomycetous yeasts Cryptococcus neoformans and Cryptococcus gattii are fungal pathogens responsible for causing meningitis and pulmonary infections in immunocompromised and in a subset of immunocompetent patients (87). C. neoformans exhibits a diverse worldwide geographic distribution, while C. gattii tends to be tropical or sub-tropical. These species are speculated to have diverged from one another 40 million years ago (64). At least six cryptic species have been defined within the recognized pathogenic Cryptococcus species complex (16, 92). Population genetics studies and molecular analysis revealed the presence of two C. neoformans varieties that likely represent different cryptic species, var. grubii and var. neoformans. Eight molecular types comprise the pathogenic Cryptococcus species complex, including VNI/II/B and VNIIV for C. neoformans var. grubii and C. neoformans var. neoformans and the molecular types VGI, VGII, VGIII, and VGIV for C. gattii (16, 92).

The species and varieties of C. neoformans and C. gattii are divided into four serotypes (A, B, C, and D) (26, 85). Specifically, C. neoformans var. neoformans is serotype D, C. neoformans var. grubii is serotype A, and C. gattii species are serotypes B or C. The pathogenic Cryptococcus species are commonly found in association with some tree species, soil, and pigeon guano (26, 64). In the United States and Africa, C. neoformans infections occur in more than 95% of AIDS cases and are also common in patients with other forms of immunosuppression while C. gattii typically (~50%) infects
otherwise healthy individuals (26). In 1999, a *C. gattii* outbreak began in British Columbia, Canada (42, 47) and more recently population studies have reported the expansion of *C. gattii* from Canada into the Pacific Northwest United States (20-22, 33). This outbreak has impacted a wide range of hosts increasing the number of known susceptible hosts. If untreated, cryptococcal infections are 100% fatal. However, advances in antifungal drug therapy (114, 140), have decreased mortality rates though a significant threat still remains.

A robust set of tools developed for genetic, molecular, biochemical, signal transduction, and genomic studies have furthered our understanding of the virulence potential of the pathogenic *Cryptococcus* species complex (64). For example the deletion collection of the serotype A *C. grubii* strain H99 from the Madhani lab at UCSF has led to the discovery of novel melanin and capsule genes, infectivity factors, and a key regulator of virulence and capsule-independent anti-phagocytic mechanisms, Gat201 (30, 93, 94). Moreover, testing potential virulence factors in animal models (i.e. mouse or rat) and alternative hosts (i.e. amoeba, greater wax moth, or *Caenorhabditis elegans*) have identified numerous genes important in many biological processes (108, 109). Growth at high temperature (>37°C) is a feature of many pathogens for survival in the host, and *C. neoformans* and *C. gattii* can proliferate at high temperatures and calcineurin, a protein phosphatase controlled by intracellular calcium, is required for *C. neoformans* growth at high temperatures (26). The pathogenic species also produce capsule and melanin. Capsule is composed of a complex of sugars and forms a protective coating surrounding the yeast cells. Yeast cells that lack capsule are readily engulfed by macrophages while
the presence of capsule prevents uptake (3, 48). Melanin production is controlled by laccase, a copper-containing oxidase that produces pigment in fungal cells. The production of melanin shields *C. neoformans* from harmful UV radiation in the environment where the fungus encounters harsh conditions (26, 64).

The genomes of *C. neoformans* strains JEC21 and B-3501A (serotype D), *C. grubii* strain H99 (highly virulent, serotype A), and *C. gattii* strains WM276 (environmental, serotype B) and R265 (clinical, serotype B) have been sequenced (72, 97). The release of these genomes has provided insight into the emergence and evolution of the pathogenic species complex. The evolution of sex, virulence, and biology are under investigation in this group of genetically tractable fungi.

On average, the haploid genomes of the pathogenic *Cryptococcus* species are ~20 Mb, span 14 chromosomes, and encode ~6500 predicted genes (97). Genes are intron-rich and antisense transcripts and alternatively spliced messages are abundant (97). The genome is complex and rich in repetitive elements and transposons, and the predicted centromeres appear to be enriched for transposable elements. Interestingly, the sex-determining region (or mating-type locus) is also transposon-rich. Comparative studies of the content and architecture of the genomes of *C. neoformans* and *C. gattii* have provided a platform for additional genomic comparisons with other basidiomycetes and also more distantly related fungi such as the ascomycetous yeasts *Saccharomyces cerevisiae* and *Candida albicans.*
1.2 The life cycle of C. neoformans

June Kwon-Chung first described the sexual stage of C. neoformans in the mid 1970s (77-79). C. neoformans and C. gattii normally grow as budding yeasts in the environment and in the host, but during the Cryptococcus sexual cycle, a dimorphic switch from a yeast to a filamentous state occurs that ultimately culminates in the production of meiotic basidiospores. In C. neoformans, mating assays are performed on 5% V8 agar juice medium or more recently, the plant growth medium, Murashige and Skoog (MS) medium. Cultures are incubated in the dark for one week or longer, and in C. gattii, matings require an incubation period of four to six weeks (23, 24, 42, 47).

Heterothallic mating in the pathogenic Cryptococcus species requires cells of opposite mating-type for cell-cell fusion. Under nutrient limiting conditions, a and α cells secrete pheromones which promotes a-α cell fusion and the formation of dikaryotic filaments (77, 78). These dikaryons are connected by fused clamp cells through which the nuclei migrate. Nuclear fusion and meiosis occurs in the basidia producing four long chains of post-meiotic spores. Moreover, spores or dessicated yeasts are hypothesized to be the infectious propagules for C. neoformans infection (49, 164). Mating ability within the pathogenic Cryptococcus species differs. C. neoformans (serotype D) strains mate robustly, C. grubii (serotype A) displays mating strain specificity/preferences, while in C. gattii (serotypes B and C) only three of the four molecular types are fertile (VGI can mate with VGI, II, and III and VGII can mate with VGII) (64). Intervarietal matings also occur but the viability of progeny is reduced and many diploid and aneuploid progeny are generated (64).
In addition to heterothallic mating, homothallism (same-sex mating, $\alpha-\alpha$) has been observed in the pathogenic Cryptococcus species (88-90, 169). Same-sex mating in C. neoformans and C. gattii is also known as haploid or monokaryotic fruiting (89, 169). A previous study has shown that fruiting involves meiosis (89), and is not mitotic and asexual as was originally proposed (169). Moreover, the conditions that stimulate mating also promote monokaryotic fruiting resulting in filamentation and sporulation. Conditions that induce fruiting are nutrient limitation, darkness, and response to pheromones (89). In contrast to opposite-sex matings, fruiting produces monokaryotic (diploid $\alpha-\alpha$ nuclei) filaments with unfused clamp cells. Meiosis occurs during same-sex mating and can result in the production of recombinant haploid spores. Thus, genetic exchange occurs during this alternative mode of sexual reproduction. Over 95% of clinical and environmental isolates are of the $\alpha$ mating-type, and the $\alpha$ mating-type is more virulent than the $\alpha$ mating-type in some strain backgrounds. This bias could explain why the $\alpha$ alleles are able to undergo fruiting (85, 111, 113, 168), and C. neoformans may have evolved monokaryotic fruiting as an alternative form of sexual reproduction to promote inbreeding and to maintain favorable gene combinations (for more virulent strains).

1.3 The evolution of the mating-type locus (MAT) in the Cryptococcus species complex

In fungi, a specialized chromosomal region termed the mating-type (MAT) locus establishes cell identity and governs sexual reproduction. The MAT locus of fungi is organized into two structural paradigms: bipolar and tetrapolar. A bipolar system promotes inbreeding while a tetrapolar system promotes outbreeding. In a bipolar mating
system, a single locus encodes homeodomain transcription factors and pheromone/pheromone receptors, while in a tetrapolar mating system, unlinked HD and P/R loci define MAT. Typically, most ascomycetes are bipolar and basidiomycetes are primarily tetrapolar.

The bipolar mating system of the ascomycetous yeast, *S. cerevisiae* has been extensively examined and is a model system for other fungi. The opposite mating-type cells of *S. cerevisiae* are designated as a and α and its MAT locus is less than 1 kb (4). The yeast MAT locus encodes α-domain and homeodomain transcription factors; MATα encodes two proteins (α1, α-domain and α2, HD1) and MATα encodes only α1 (HD2). In haploid cells, α1 directs expression of α-specific genes and α2 represses the transcription of α-specific genes (69, 149). Haploid a or α cells respond to pheromone produced by cells of the opposite mating-type by forming an a/α diploid which undergoes meiosis to produce an ascus containing four haploid meiotic progeny. In the a/α diploid, an a1-α2 heterodimer forms and represses haploid-specific gene expression (35).

Among the basidiomycetes, the mating-type (MAT) locus of *Ustilago hordei*, *Coprinellus disseminatus*, *Schizophyllum commune*, *C. neoformans*, and *C. gattii* have been the target of numerous molecular studies (121, 133, 137, 153). The MAT loci of *S. commune* and *C. neoformans* are particularly interesting and present unique genetic arrangements (40). *S. commune*, a mushroom, is quite promiscuous and is predicted to have >15,000 mating-types. The MAT loci are multiallelic with an estimated 288 HD (A) and 81 P/R (B) locus alleles (147). The HD locus includes 9 Aα and 32 Aβ alleles and the P/R locus has 9 Bα and 9 Bβ alleles (147). The HD genes Aα and Aβ are an estimated
~650 kb apart and are highly similar to the divergently transcribed HD1/HD2 encoding genes in *Coprinellus cinerea* (27). The *Aβ* allele contains an HD motif of the HD2 gene and lacks an HD1 homolog (147). To date, only one HD2 gene has been identified in *S. commune*. The P/R locus is arranged in two loci in which the *Ba* locus encodes a single receptor and two to three pheromone genes while *Bβ* is 8 kb downstream and encodes eight different pheromones (27, 159). The *MAT* loci of *S. commune* control the switch from a non-fertile homokaryotic mycelium to a fertile dikaryotic mycelium ultimately giving rise to fruiting bodies (147). As with all other fungal species, meiosis and sporulation occurs in the fruiting bodies. For sex to occur and for a dikaryon to form in this heterothallic fungus, homokaryons must differ at *Aα* and/or *Aβ* as well as at *Ba* and/or *Bβ*. *S. commune* is also one of the oldest model systems employed for examining the genetics of mating in basidiomycetous fungi (40, 124).

The *MAT* locus of the pathogenic *Cryptococcus* species is unusually large, spanning >100 kb, and encoding ~20 linked genes. The bipolar system encodes a single non-recombining locus of linked HD and P/R genes in addition to many genes with unknown functions. In the pathogenic species, components associated with pheromone signaling (*STE11, STE12, and STE20*) are linked to *MAT* (18). Comparison of the *MAT* locus in *C. neoformans* species suggests that the locus has undergone extensive rearrangement and inversions, while in the *C. gattii* species, *MAT* is syntenic throughout the entire locus.

How did *C. neoformans* and *C. gattii* evolve a bipolar mating system? In 2005, Fraser et al. proposed that the ancestral form of *MAT* was a tetrapolar system (41). In this
model, two unlinked MAT defining regions (the HD and P/R loci) underwent sequential rounds of gene acquisition and expansion. Next, a chromosomal translocation event resulted in the formation of a transient tripolar intermediate, which collapsed to form the bipolar state. Gene rearrangements, gene conversions, and the acquisition of transposons yielded the extant bipolar alleles of *C. neoformans* and *C. gattii*. More recently, Metin et al. revisited the MAT evolutionary model and analyzed the MAT structure in the distantly related saprobic species, *Cryptococcus heveanensis* (38). The authors concluded the following: (1) the ancestral tetrapolar HD locus contains two unlinked and divergently transcribed HD genes and (2) either *SXI1* or *SXI2* were lost recently, possibly concomitantly, with the chromosomal translocation between chromosomes 4 and 5 of *C. neoformans* that gave rise to the unstable tripolar intermediate. Lastly, Hsueh et al. genetically engineered *C. neoformans* strains to have a tripolar and a tetrapolar mating system by moving *SXI1* and *SXI2* to a genomic region unlinked to MAT (60). Progeny from different crosses were dissected and the tripolar maters exhibited decreased fertility while the tetrapolar maters were able to complete sexual reproduction. The recapitulation of the ancestral tetrapolar and tripolar mating confirms that the evolution of the MAT locus of the pathogenic *Cryptococcus* species likely involved three stages: it first began as a tetrapolar system that formed an unstable tripolar intermediate and eventually collapsed to a bipolar mating system.

Ohno first proposed how animal sex chromosomes evolved (115). In his model, the original master sex-determining gene arose on an autosome, and then captured this chromosome to become a sex chromosome in conjunction with other genes with sex-
related functions. The evolution of distinct sex chromosomes was proposed to have occurred via suppression of recombination by mechanisms including inversions, leading to the divergence of large genomic regions and the emergence of diverse sex chromosomes. Gene capture and suppression of recombination therefore punctuate the evolution of sex chromosomes. These features of sex chromosome evolution in animals and plants mimic the evolution of MAT in the pathogenic Cryptococcus species complex (24, 37, 41, 46, 60). Fungi are haploid, and thus the degeneration of one of the sex determinants does not occur, unlike in animals (i.e. the Y chromosome).

1.4 Taxonomy of Cryptococcus species and the molecular phylogeny of the pathogenic Cryptococcus species

The Cryptococcus species are not a monophyletic clade. Instead, Cryptococcus is polyphyletic with the monophyletic pathogenic Cryptococcus species cluster embedded within a larger group of primarily non-pathogenic and saprobic yeast species (137). The genus Cryptococcus includes the Tremellales, Trichosporonales, Filobasidiales, and Cytofilobasidiales (137). Non-C. neoformans species such as Cryptococcus laurentii and Cryptococcus albidus also infect humans, though to a lesser extent than members of the pathogenic Cryptococcus species complex. The taxonomic classification of C. neoformans and C. gattii was assigned after the discovery of sex in this species in the 1970s (78, 79, 81-83). The teleomorphic genus, Filobasidiella, is based on morphological features of their sexual structures (81). Previous phylogenetic studies have focused only on the ribosomal DNA (rDNA) or the internal transcribed spacer (ITS) genes. In these
studies, the statistical support for the analysis was strong though lacking in the diversity of isolates analyzed.

The current understanding is that the *Filobasidiella* clade consists of the homothallic and obligately sexual filamentous fungus *Filobasidiella depauperata*, the pathogenic species *Filobasidiella neoformans* (teleomorph of *C. neoformans*) and *Filobasidiella bacillispora* (teleomorph of *C. gattii*), and the heterothallic fungus *Filobasidiella lutea* (141). These species are either human pathogens or mycoparasites found in association with dead insects or decaying insect frass. More recent studies have shown that *Cryptococcus amylolentus* and *Tsuchiyaea wingfieldii* are also phylogenetically related to the pathogenic *Cryptococcus* species (36, 54, 82, 137). Studies aimed at elucidating the phylogeny surrounding the pathogenic species are key to understanding how *C. neoformans* and *C. gattii* evolved mechanisms governing sex and the factors controlling virulence, especially in this monophyletic clade that contains primarily saprobic species.

*Kwoniella* is the next closest clade to the *Filobasidiella* species and is defined by the heterothallic fungus, *Kwoniella mangroviensis* (148). Analysis of the D1/D2 LSU rRNA gene revealed that *K. mangroviensis* is phylogenetically related to *Cryptococcus bestiolae, Cryptococcus dejecticola, Bullera dendrophila,* and *C. heveanensis*, and is also more distantly related to *Cryptococcus taiwaniana,* and *Cryptococcus cuniculi* (148). Members of the *Filobasidiella* clade produce aseptate basidia with four chains of spores emanating from the apex of the basidia while the basidia of *Kwoniella* species have cruciate septa and the spores typically surround the basidia, as in *C. heveanensis* (102,
Insight into the species that are phylogenetically related to the pathogenic *Cryptococcus* species provides a platform for future comparative genomics studies that examine fungal biology, genetics, and life cycles.

### 1.5 Overview of this work

The work described here focuses on the evolution of the mating-type locus and sexual reproduction in the closest relatives to the pathogenic *Cryptococcus* species complex. In Chapter 2, the phylogeny encompassing a subset of species in the Tremellales lineage that closely cluster with the pathogenic *Cryptococcus* species complex was resolved by employing a robust multi-locus sequencing approach for phylogenetic analysis (38). Furthermore, 15 basidiomycete species were subjected to virulence and phenotypic assays to evaluate their pathogenic potential. These studies reveal that *C. amylolentus* and *T. wingfieldii*, two nonpathogenic sibling species, are the most closely related taxa to the pathogens *C. neoformans* and *C. gattii*, and group with *Filobasidiella depauperata* to form the *Cryptococcus sensu stricto* group. Five other saprobic yeast species form the *Kwoniella* clade, which appears to comprise a more distantly related *sensu lato* group. This study establishes a foundation for future comparative genomics approaches that will provide insight into the structure, function, and evolution of the mating type locus, transitions in modes of sexual reproduction, and emergence of human pathogenic species from related or ancestral saprobic species.

We discuss the characterization of the *MAT* structure of the closest relatives of the pathogenic *Cryptococcus* species complex, the sibling species *C. amylolentus* and *T. wingfieldii* in Chapter 3. These studies determined that the *MAT* subloci (HD and P/R) in
the sibling species are unlinked and present on two different chromosomes, and that the mating system of *C. amylolentus* is tetrapolar. Interestingly, the characterization of *MAT* in the sibling species revealed that gene rearrangement, chromosomal translocation, gene conversion, and inversions shaped the evolution of the mating-type locus in *C. neoformans* and *C. gattii*.

In Chapter 4, we described the extant sexual cycle in *C. amylolentus*. Fluorescent microscopy revealed that dikaryotic hyphae are produced and four long chains of spores emanate from the basidia. Genotyping and RAPD analysis of microdissected spores from *C. amylolentus* suggest that meiosis and, thus, sexual reproduction occurs in this species. Interestingly, a majority of the microdissected progeny are sterile and only a subset are fertile isolates almost exclusively of the α mating-type. From our analysis, we conclude that sex in *C. amylolentus* is heterothallic and tetrapolar. Although some fungi have retained the appropriate machinery for sexual reproduction, sex has not yet been described for many species currently classified as asexual. Exploration of sex in these putatively asexual species will be key in elucidating the mechanisms employed by other species for sexual reproduction.

In Chapter 5, we identified the *MAT*-specific genes encoded in the tetrapolar mating system of the wood-rotting fungus *Tremella mesenterica* using comparative genomics. The Joint Genome Initiative sequenced the genome of ATCC24925 and we successfully identified *Cryptococcus MAT*-specific genes in *T. mesenterica* using BLAST analysis. We compared the sequenced isolate ATCC24925 to four *T. mesenterica* strains of opposite mating-types. From our analysis, we determined that the HD locus spans ~3
kb and is defined by $SXI1$ and $SXI2$-like genes while the P/R spans ~18 kb and is defined by $STE3$, $STE12$, $STE20$, the pheromone gene tremerogen a-13, and two hypothetical genes. The analysis of the more distantly related species $T. mesenterica$ is central to our understanding of the evolution of the $MAT$ locus in $Cryptococcus$. Taken together, our studies illustrate the evolutionary plasticity of both sex determinants and sexual reproduction in fungi with implications for model and pathogenic microbes as well as multicellular eukaryotes including alga, plants, and animals.
Chapter 2. Phylogeny and phenotypic characterization of pathogenic Cryptococcus species and closely related saprobic taxa in the Tremellales

Chapter 2 is modified from a published manuscript (38).

2.1 Introduction

Recent phylogenetic and genomic studies of the fungal kingdom have illustrated that analysis of both distantly and closely related species provides insight into the evolutionary trajectories for fungal species (67, 146). The Fungal Tree of Life Project (AFTOL) applied a high-resolution multi-locus sequencing (MLS) approach to 170 species, elucidating broad and specific evolutionary relationships among species (66). However the entire fungal kingdom encompasses an estimated 1.5 million species (57), many more than can be analyzed by this approach. More than 100 fungi have been subject to whole genome analysis, including distantly related but also, in some cases, closely aligned species (146). Taken together, these approaches illustrate that comparisons of closely related species, first by MLS and then by whole genome studies, can impact our understanding of how closely related pathogenic and saprobic fungi evolve to occupy specialized niches in nature.

The human fungal pathogen and basidiomycete Cryptococcus neoformans and its closest relative Cryptococcus gattii cause respiratory and neurological disease in immunocompromised and immunocompetent patients (26). The sexual states of those species, which develop on artificial culture media, were designated Filobasidiella neoformans and F. bacillispora, respectively (78, 79). C. neoformans and C. gattii are members of the pathogenic Cryptococcus species cluster, which contains two currently
recognized varieties, *C. neoformans* var. *neoformans* and *C. neoformans* var. *grubii*, and the sibling species *C. gattii* (26, 81). Recent studies based on MLST analysis provide evidence that *C. neoformans* var. *neoformans* and *C. neoformans* var. *grubii* are distinct species and a few sub-populations are found in var. *grubii* (VNI, VNII, VNB) (16, 91, 92). Furthermore, *C. gattii* can be further subdivided into at least four cryptic species (VGI, VGII, VGIII, VGIV) (16, 42, 103). Hence, as many as six species define the pathogenic *Cryptococcus* species complex. All of these pathogenic *Cryptococcus* species cluster together in the *Filobasidiella* clade, which appears to have emerged from the *Tremella* lineage (36, 137). Nonetheless, the genus *Cryptococcus* is polyphyletic and species cluster within the Tremellales, the Trichosporonales, the Filobasidiales, and the Cystofilobasidiales clades (36, 133, 137), including the less common human pathogens *Cryptococcus albidus*, *Cryptococcus laurentii*, and *Cryptococcus adeliensis* (126, 138), and the saprobe *Cryptotrichosporon anarcardii* (116), which shares similar phenotypic traits (melanin and capsule) with *C. neoformans*.

Previous taxonomic studies based on ultrastructural features and basidial morphology separated the hymenomycetous yeasts (now classified in the Tremellomycetes (59)) into two orders: Tremellales with cruciate-septate basidia and Filobasidiales with aseptate basidia (76). Because *Filobasidiella* species have aseptate holobasidia, they were included in the Filobasidiales together with *Filobasidium* and *Cystofilobasidium* species. However, molecular data later showed that *Filobasidiella* species are more closely related to Tremellales (137). Most of the species with defined
sexual cycles in the Tremellales have septate 2- to 4-celled basidia (14, 73, 76, 133, 135, 148). The holobasidia of *Filobasidiella* species are thus unique in the Tremellales.

Mating systems also differ among the Tremellales. All known heterothallic *Tremella* species have been reported to have a tetrapolar (bifactorial) mating system with a multiallelic *A* locus and a biallelic *B* locus involved in production and sensing of pheromones (56, 167). *Fibulobasidium inconspicuum* has also been shown to have the same type of mating system (8). However, other teleomorphic species in the Tremellales, such as *Auriculibuller fuscus* and *Kwoniella mangroviensis* are reported to have bipolar mating type systems (14, 134, 148).

Tremellales generally have haustorial branches, which are short branches of hypha with a basal clamp connection (101). Haustorial filaments of the mycoparasitic *Tremella* species were observed attached to host cells, enabling penetration and parasitic interactions (79, 148, 178). Haustorial branches have also been described in *C. neoformans* (77, 148), suggesting an ancestral mycoparasitism (7).

Previous studies examining the genetic relatedness of the species in the *Filobasidiella* clade, namely the homothallic filamentous fungus *Filobasidiella depauperata* and the heterothallic yeasts *C. neoformans* and *C. gattii* are based on ITS and/or ribosomal DNA sequence divergence and basidial morphology (42, 77, 82, 83, 153). Results from these studies strongly support that these three species form a monophyletic clade. Additional studies have shown that these fungi also align closely with *Cryptococcus amylophilus* and *Tsuchiyaea wingfieldii* (36, 54, 81, 106). Because of their proximity to the pathogenic *Cryptococcus* species cluster, these species constitute
the focus of our phylogenetic study. Several additional saprobic anamorphic yeasts including *Cryptococcus bestiolae* (157), *Cryptococcus dejecticola* (157), *Cryptococcus heveanensis*, and *Bullera dendrophila* (161, 163) also appear to be closely related to the *Filobasidiella* clade and form a monophyletic lineage related to the dimorphic, heterothallic basidiomycetous yeast *Kwoniella mangroviensis* (formerly *Cryptococcus* *spp.* CBS8507), recently discovered to have an extant sexual cycle (148). Therefore, for a more robust comparison several members of the *Kwoniella* lineage were also included in our analyses.

Many of the species nested in the Tremellales are mycoparasites and saprobes associated with wood, plants, soil, and arthropod frass (178). Similar habitats and environmental associations have been found for *C. neoformans* and *C. gattii*. These two species have been isolated from a variety of tree species, pigeon guano, and insects. For example, *C. neoformans* has been isolated from beetles (89) and *C. gattii* has been detected in honeybee hives and insect frass (87). This accumulating circumstantial evidence might imply a possible arthropod associated habitat for the pathogenic *Cryptococcus* species. Previous studies in several invertebrate model hosts such as *Acanthamoeba, Caenorhabditis elegans, Dictyostelium discoideum, Drosophila melanogaster*, and *Galleria mellonella* have been used to study bacterial or fungal infections in mammals (108). These host systems are in some cases genetically tractable, relatively simple to manipulate, and inexpensive (108). Interestingly, the virulence potential of the pathogenic *Cryptococcus* species correlates with the ability of the fungus to cause lethal infections in the heterologous insect host *G. mellonella* (wax moth) (109).
Moreover, *C. neoformans* pathogenicity has been studied extensively in *G. mellonella*, and genes required for infection of the wax moth larvae are necessary during *C. neoformans* and *C. gattii* infection of mammalian hosts (109). Results from these studies reveal insights into the ecological niche and the evolution of pathogenesis of *C. neoformans* and *C. gattii*.

The Tremellales lineage contains over 120 species, though many of their phylogenetic relationships are weakly supported due to the lack of multi-locus phylogenetic and phenotypic analyses. Here we molecularly and phenotypically examine the species surrounding the monophyletic pathogenic *Cryptococcus* complex clustered within the Tremellales. We employed a multi-locus approach encompassing six highly conserved genomic loci present in the fungal kingdom. *RPB1, RPB2, EF1α, mitSSU rRNA* gene, nuclear LSU rRNA (D1/D2 domains), and ITS were amplified and sequenced to determine the divergence among the 15 fungal isolates. The virulence potential of the isolates was also evaluated in *Galleria*. These studies provide insight into the genotypic and phenotypic trajectory of a highly successful pathogenic clade that likely emerged from saprobic fungi associated with insects in the environment.
2.2 Materials and methods

2.2.1 Fungal isolates

The isolates used in this study are listed in Table 1. All were grown and maintained on YPD medium at 24°C.

2.2.2 DNA isolation, PCR, and sequencing

To isolate fungal DNA, cells were harvested after shaking at 24°C in YPD liquid medium overnight followed by lyophilization. DNA was isolated using the CTAB method (165). PCR amplification was performed on six fungal specific, highly conserved genes encoding the following: the largest subunit of RNA polymerase II (RPB1), the second largest subunit of RNA polymerase II (RPB2), elongation factor 1 alpha (EF1α), the mitochondrial small subunit ribosomal RNA (mitSSU), the D1/D2 domains of the nuclear large subunit ribosomal RNA (nucLSU, D1/D2), and the internal transcribed spacer region (ITS) of the ribosomal DNA unit, which includes the ITS1 and ITS4 spacers and the 5.8S ribosomal RNA gene. Individual PCRs were performed for each of the six genes. Primer information can be found at http://www.aftol.org/primers.php (primers used in study: RPB1Af and RPB1Cr, RPB2-5f and RPB1-11bR, EF1α1F and EF1α1R, mitSSU1F and mitSSU3R, nucLSU LrDNA AND nucLSU LR3, and ITS1 and ITS4). PCR products were separated on an agarose gel and purified using the QIAquick PCR Purification Kit (Qiagen, Valencia, California, United States). Sequencing reactions were performed using Big Dye Chemistry v3.1 (Applied Biosystems, Foster City, California, United States) and analyzed on an Applied Biosystems 3730xl capillary
sequencer. Sequence reads were trimmed and assembled with the DNA sequence assembly software program Sequencher (Version 4.8, Gene Codes Corporation, Ann Arbor, Michigan, United States). Individual contigs were generated in Sequencher and BLAST (2) was performed to confirm the identity of the sequenced products. Genbank accession numbers are listed in Table 2.

2.2.3 Phylogenetic analysis

A total of 92 out of 96 DNA sequences (4% of the molecular data was missing) were aligned using ClustalW 1.81 (158). The FASTA alignment files for each of the six loci were imported into MacClade 4.08 (99) for manual editing and to identify and correct ambiguously aligned regions (732 characters). Edited files were concatenated resulting in 3,422 characters included for the phylogenetic analyses. Heuristic searches for maximum parsimony (MP, Figure 1A) and maximum likelihood (ML, Figure 1B) criteria were conducted using PAUP 4.0 (152). MP searches were also conducted on the concatenated data set using a user-defined step matrix generated in STEP3matrix for conserved regions (98) and a matrix generated by INAASE for ambiguously aligned regions (104). Model parameter estimates for ML analysis were obtained using MODELTEST (122). Statistical support was calculated using 1,000 bootstrap replicates under MP and ML, and all trees were rooted using *C. humicola* as the outgroup (a member of the Trichosporonales (36, 137), Table 1). MacClade software was also used to trace habitat preferences (Figure 2) of the fungi in the current study (99).
2.2.4 Phenotypic analysis of isolates

The phenotypic properties of each isolate were examined in several assays. First, the ability of each strain to grow on YPD and YNB medium at 24°C, 30°C, and 37°C was tested. Each isolate was analyzed for melanin production on Niger Seed, Rose Bengal, and low-glucose (0.1%) medium supplemented with the diphenolic molecule L-DOPA (100 mg/ml). To assay for capsule production, isolates were grown in capsule induction medium (5 g glucose, 5 g asparagine, 400 mg K₂HPO₄, 80 mg MgSO₄·7H₂O, and 250 mg CaCl₂·2H₂O and 100X vitamin-mineral mix (0.4 g thiamine, 0.057 g B(OH)₃, 0.004 g CuSO₄·5H₂O, 0.01 g MnCl₂·4H₂O, 2 g ZnSO₄·6H₂O, and 0.46 sodium molybdate in 1 L of water) with low iron (20 mg/L of the iron chelator EDDHA) at 24°C for 2 days followed by India ink staining. Isolates were also tested at 24°C for urease activity using Christensen’s agar (Beckton Dickinson, Cockeysville, MD). For these assays, all isolates were grown to exponential phase in liquid YPD, serially diluted in spots onto YPD, YNB, Niger Seed, Rose Bengal, L-DOPA, and Christensen’s agar, and incubated for 4 to 5 days. The phenotypic assays were performed in triplicate. We note that some images of growth patches on the same media were cropped to produce the composite image shown in Figure 3.

2.2.5 Galleria mellonella survival assay

To analyze virulence in the heterologous host G. mellonella, a previously described protocol (109) was followed with minor modifications. Wax moth larvae (~12) were injected in the last pseudopod with 5 x 10⁵ cells of each isolate and wax moths were incubated at 24°C. The number of surviving wax moths was monitored and recorded
daily. Survival curves were plotted using the Prism software (Version 4.0a, Prism Computational Sciences, Incorporated, Madison, Wisconsin, USA). A $p$ value of $<0.05$ was considered significant. $P$ values were calculated and compared to the reference PBS (or mock inoculation) control and to the H99 $C. neoformans$ var. $grubii$ pathogenic control (see graphs in Figures 4 and 5). The experiment was performed three times.
2.3 Results

2.3.1 Cryptococcus amylolentus and Tsuchiyaea wingfieldii are the species most closely related to C. neoformans

To establish evolutionary relationships of the pathogenic *Cryptococcus* species cluster and the related saprobic yeasts in the Tremellales, we performed a multi-locus phylogenetic analysis of the fifteen species listed in Table 1. The following six highly conserved, fungal specific genomic loci were sequenced and analyzed: *RPB1, RPB2, EF1α, mitSSU, nucLSU (D1/D2)*, and ITS. Phylogenetic trees from concatenated sequences were processed in PAUP (152) and the parsimony heuristic searches in PAUP (152) generated only one tree topology (Figure 1A). The same topology for the *Filobasidiella* clade was obtained under maximum likelihood searches (Figure 1B) when the substitution parameters estimated by the MODELTEST were used (122).

This MLS analysis resolved the phylogeny of the pathogenic *Cryptococcus* species complex and its closest relatives. In agreement with previous studies (153), two closely related monophyletic sister clades, the *Filobasidiella* and *Kwoniella* lineages, were identified and showed over 96% bootstrap support (Figure 1). The *Filobasidiella* clade represents a *sensu stricto* clade of *Cryptococcus* species (it contains the type species of the genus, *C. neoformans*) while the *Kwoniella* clade contains the majority of the more distantly related insect associated yeasts, named the *sensu lato* group (Figures 1 and 2). These two clades are reminiscent of the *sensu stricto* and the *sensu lato* groups used to describe the phylogeny of *Saccharomyces cerevisiae* and related species (162). Our results show that the two closest relatives of *C. neoformans* are the sibling species *C.
*amylolentus* and *T. wingfieldii*. These findings do not support previous phylogenetic studies (based on a single gene) that placed *F. depauperata* as the closest relative of the pathogenic *Cryptococcus* species (83, 116, 141). However, the current and previous results agree that *F. depauperata* is a member of the monophyletic *Filobasidiella* clade, which also includes *C. amylolentus*, *T. wingfieldii*, and the pathogenic species *C. neoformans* and *C. gattii*.

Figure 1: Phylogenetic relationship among members of the Tremellales.
A combined data set of concatenated gene sequences in a parsimony bootstrap phylogenetic tree representing six fungal specific genomic loci (*RPB1, RPB2, EF1α, mitSSU, nuclSU (D1/D2), and ITS*) reveals the phylogeny of the *Cryptococcus* pathogenic species complex with the insect associated species. Single topology for the most parsimonious tree shown in A, and maximum likelihood tree shown in B. Nucleotide sequences for each gene were aligned and analyzed using Mesquite and PAUP. Numbers on branches are bootstrap values from 1000 replicates. The outgroup in the analysis is *C. humicola*. 
The *Kwoniella* lineage includes several arthropod-associated species including *Bullera dendrophila*, *Cryptococcus dejecticola*, *Cryptococcus bestiolae*, and *C. heveanensis* (Note: the *C. heveanensis* strain used in this study was isolated from sheet rubber), and one isolated from mangrove areas, *Kwoniella mangroviensis* (Table 1 and Figure 2). Our findings confirm previous studies that suggested the *Kwoniella* clade is split into two monophyletic clusters, one lineage represented by *C. heveanensis*, and the other including *B. dendrophila*, *C. dejecticola*, *C. bestiolae*, and *K. mangroviensis* (Figures 1 and 2).

![Phylogram indicating the preferred habitats of the fungal species used in this study.](image)

Figure 2: Phylogram indicating the preferred habitats of the fungal species used in this study.
Tree diagram is based on data from Figure 1 and indicates evolutionary lineages and natural habitats of the isolates examined in the study. These species are possibly derived from a common insect-associated ancestor.

The three *Tremella* strains used in the analysis represent more distantly related taxa within the Tremellales. Interestingly, the two strains of *T. mesenterica* appear to be as divergent from each other as the two *C. neoformans* varieties, and also as diverged as the two *C. gattii* VG groups, suggesting the existence of sub-populations, varieties or cryptic species within *T. mesenterica* (Figures 1 and 2). The monophyletic clades defined based on this phylogenetic analyses were rooted with *C. humicola*, a member of the Trichosporonales (36, 137, 154).

### 2.3.2 Phenotypic and morphological differences in close and distant relatives of *C. neoformans*

The virulence potential of the studied isolates was assessed by examining three well-established virulence attributes: growth at 37°C, capsule, and melanin production. To assay for growth at different temperatures, serial dilutions of individual species were spotted at 24°C, 30°C, and 37°C on both nutrient rich and minimal medium. *C. neoformans* var. *grubii*, *C. neoformans* var. *neoformans*, and *C. gattii*, all of which belong to the pathogenic *Cryptococcus* species complex, grew at the highest temperature, 37°C (Figure 3). Additional species clustering in the *sensu stricto* Filobasidiella lineage, *C. amylolentus*, *T. wingfieldii*, and *F. depauperata*, failed to grow at 37°C. Five of the six species in the *sensu stricto* Filobasidiella lineage could grow at 30°C (Figure 3), and the only strain in this group that did not grow at 30°C was *T. wingfieldii*. Because we speculate that *T. wingfieldii* is a cryptic or sibling species of *C. amylolentus* (Figures 1
and 2), the lower thermotolerance of the former might be a distinct trait between the two species; alternatively, this property could vary between isolates of the same species. At present no other *T. wingfieldii* or *C. amylolentus* isolates are available to explore this further.

None of the species in the *sensu lato* *Kwoniella* clade, including *K. magroviensis*, *B. dendrophila*, *C. bestiolae*, *C. dejecticola*, and *C. heveanensis*, nor the species representing the paraphyletic lineages, *Tremella* spp. and *C. humicola*, grew at 37°C (Figure 3). Only two isolates *B. dendrophila* and *T. globispora*, could not grow at 30°C; a previous report showed that *B. dendrophila* displays little to no growth at 30°C or 35°C (13). In this case, lack of growth at 30°C seems to be a phenotypic difference that may aid in distinguishing *B. dendrophila* from the rest of the species in the *Kwoniella* lineage (Figure 3) because the majority of their close relatives are able to grow at 30°C (current study, (121)). In summary, the saprobic yeasts are distinguished from closely aligned pathogenic yeasts by their inability to grow at 37°C.

To examine capsule production, all strains were grown under conditions (low iron media) known to induce capsule formation in *C. neoformans* and *C. gattii*. Under the conditions tested, none of the saprobic yeasts produced capsules that could be visualized by india ink exclusion, in contrast to those observed for both *C. neoformans* and *C. gattii* (data not shown). Melanin production was tested using three different melanin-inducing media (DOPA, Rose Bengal, and Niger seed). The species in the pathogenic *Cryptococcus* clade all produced melanin (Figure 3), and *C. heveanensis* and *C. humicola* also showed faint coloration that might be attributable to melanin production (Figure 3).
The ability to hydrolyze urea was also tested on Christensen’s agar (117). All 15 species tested hydrolyzed urea confirming that they are all basidiomycetes. Based on these findings, we hypothesize that the environmental cues triggering capsule production might differ in *C. neoformans* and *C. gattii* compared to their saprobic relatives. Furthermore, both melanin production and growth at 37°C are characteristic of human pathogenic species, whereas other traits, such as the ability to hydrolyze urea, are pleisiomorphic.

![Figure 3: Phenotype and morphology of species grown under different conditions.](image)

Spot assays of individual isolates were grown for 4 to 5 days on different media at 24°C, 30°C, and 37°C. Each isolate was individually tested for the ability to grow on nutrient rich media (YPD) and minimal media (YNB). For the melanin assay, strains were grown for 4 to 5 days at 24°C on DOPA media. *F. depauperata* grows slower than any of the other isolates tested (approximately 10 days to achieve the same colony size shown in Figure 3), and is therefore not included here.
2.3.3 Virulence potential of *Cryptococcus* and neighboring taxa using *G. mellonella*

To further assess the virulence potential of each isolate in the heterologous host *G. mellonella*, larvae were inoculated with $5 \times 10^5$ yeasts, and incubated at 24°C for the duration of the experiment. Of the species in the *Filobasidiella* lineage, the two *C. neoformans* varieties and *C. gattii* exhibited the highest virulence potential ($p < 0.0001$, PBS) during infection assays in *G. mellonella* (sample survival curve is shown in Figure 4). Of the remaining species in the *Filobasidiella sensu stricto* lineage, *T. wingfieldii* ($p = 0.285$), *F. depauperata* ($p = 0.3458$), and *C. amyloleptus* ($p = 1.0$, Figure 4) were avirulent in *G. mellonella* compared to the negative control, PBS. Three of the five species clustering in the *Kwoniella* clade, *C. bestiolae* ($p = 0.0128$, PBS and $p = 0.0160$, H99), *B. dendrophila* ($p = 0.0382$, PBS and $p = 0.0128$, H99), and *C. heveanensis* ($p = 0.0181$, PBS and $p = 0.0160$, H99) displayed intermediate virulence (Figure 5).

![Survival curve](image)

**Figure 4**: Infection of the heterologous host *G. mellonella* with species closely related to *C. neoformans.* Survival of *G. mellonella* after inoculation with species in the *Filobasidiella* clade was assayed. At least twelve larvae were injected with $5 \times 10^5$ cells for each isolate. After inoculation larvae were incubated at 24°C and survival was monitored for 17 days post inoculation.
inoculation. The experiment was repeated three times and one representative experiment is presented here. In one of three replicates, *C. amylolentus* exhibited an intermediate virulence level (data not shown). The mock inoculation is injection of a PBS control.

The other two species in the *Kwoniella* clade, *C. dejecticola* (*p* = 0.1473, PBS and *p* < 0.0001, H99) and *K. mangroviensis* (*p* = 0.023, PBS and *p* < 0.0001, H99), were consistently severely attenuated or avirulent compared to the pathogenic H99 control. Furthermore, the *Tremella* outgroup species, *T. mesenterica* and *T. globispora* (*p* = 1, PBS and *p* < 0.0001, H99) were clearly avirulent in this assay. However, the *Trichosporon* outgroup species *C. humicola* displayed an intermediate level of virulence (*p* = 0.030, PBS and *p* = 0.02, H99). In summary, the pathogenic *Cryptococcus* species displayed the highest virulence in the greater wax moth whereas the majority of their close relatives were attenuated or avirulent.

![Figure 5](image)

Figure 5: Infection of the heterologous host *G. mellonella* with species distantly related to *C. neoformans*. Survival of *G. mellonella* after inoculation with species in the *Kwoniella* clade, and outgroup species representing *Trichosporon* and *Tremella* was assayed. At least twelve larvae were injected with 5 x 10^5 cells of each isolate. After inoculation, larvae were incubated at 24°C, survival was monitored daily, and the experiment was terminated 12 days post inoculation. The experiment was repeated three times and one representative is presented here. The mock inoculation is injection of a PBS control.
2.4 Discussion

Previous phylogenetic studies evaluating the phylogeny of the Filobasidiella clade used a single genomic locus, such as 5.8S, SSU, nucLSU (D1/D2), or the ITS region (36, 54, 133, 137). Results from these analyses were conflicting and failed to accurately resolve the phylogeny of the pathogenic Cryptococcus species within the Tremellales. To generate a robust data set to resolve the phylogenetic relationships surrounding this clade, fifteen species representing the Filobasidiella, Kwoniella, and Tremella lineages were examined by a multi-locus approach.

The most parsimonious and maximum likelihood tree concordantly defined two monophyletic groups: the sensu stricto Filobasidiella and the sensu lato Kwoniella lineage, as rooted with C. humicola (Figures 1 and 2). The monophyletic clades defined based on this phylogenetic analysis are analogous to the sensu stricto and the sensu lato groups in Saccharomycetaceae. The sensu stricto species include T. wingfieldii, C. amylolentus, and F. depauperata, which form a monophyletic cluster with the pathogenic Cryptococcus complex. The previously accepted Filobasidiella clade is composed of members with known sexual cycles. Therefore only Filobasidiella neoformans (anamorph C. neoformans), Filobasidiella bacillispora (anamorph C. gattii) and Filobasidiella depauperata were included (82, 141). Considerably less attention has focused on T. wingfieldii and C. amylolentus, although previous molecular evidence suggested they were closely related to the Filobasidiella clade (137). The holobasidium of the sexual taxa in the Filobasidiella clade (sensu stricto species) is clearly a
synapomorphy (shared derived character) for the clade. The phragmobasidium of the remaining taxa in the Tremellales seems to be a simplesiomorphy (ancestral character).

In the sister Kwoniella clade, the sensu lato species include B. dendrophila, C. bestiolae, C. dejecticola, C. heveanensis, and K. mangroviensis. The basidia of the only sexual species in the clade, K. mangroviensis, are similar to those of the remaining members of the Tremellales (148) and we could not identify a synapomorphy for the Kwoniella clade. Single gene phylogenies depicting the relationships between C. bestiolae, C. dejecticola, and neighboring species using nucLSU D1/D2 revealed conflicting views, especially when compared to the Neighbor-joining analyses of the nucLSU D1/D2 rRNA gene in Kwoniella and closely related species (148). Our multi-locus approach confirmed the relatedness among members of the Filobasidiella and Kwoniella clades as two distinct monophyletic clades, which had been previously suggested based on SSU (153) and ITS (137) sequence analyses. The six genomic loci employed in this study did not completely resolve the species within the Kwoniella clade given the lower bootstrap values (Figure 1). We attempted to add additional resolution to this clade by including ambiguously aligned regions in the analyses, and although MP (Figure 1) and Neighbor-joining trees (not shown) had identical topologies, results from bootstrapping and the ML topology strongly indicated that these relationships were not well supported. Future studies including additional species, such as recently discovered yeast species that cluster near these fungi or others in the Tremellales should increase the phylogenetic resolution of the Kwoniella clade. For example, Cryptococcus cuniculi (not included in this study), a previously described yeast isolated from rabbit feces in Korea is
phylogenetically related to *C. heveanensis* (139). Additionally, a novel *Cryptococcus* species, *Cryptococcus pinus*, was recently described (51). *C. pinus* was isolated from dead needles of *Pinus sylvestris* and rDNA sequence data reveal that *C. pinus* is also a member of the *Kwoniella* clade related to *C. dejecticola* (51). In summary, the robust data set generated in our study provides a platform for future studies exploring the biology, genetics, and genomics of defined species in the Tremellales and those that remain to be discovered.

Unlike many ascomycetous yeasts, which are commonly found in insect communities, basidiomycetes are less frequently isolated from living or dead insects (137). Moreover, these ascomycetes are not limited solely to insect communities, but can also be found in association with other habitats like plants and aquatic environments. Although basidiomycetes are rarely insect associated, most of the *sensu stricto* and *sensu lato* species in this study are saprobic yeasts frequently associated with decaying insects and arthropod frass. The data presented here suggest that the fungi in the *Filobasidiella* and *Kwoniella* lineages likely represent a group of phylogenetically related fungi that inhabit similar ecological niches (Figure 2). Moreover, another possibility is that the successful human pathogenic fungi emerged from an insect frass-associated ancestor.

Phenotypic assays were performed to identify features shared with or distinct from the extensively studied and well-characterized human pathogen *C. neoformans*. Melanin production and growth at 37°C was only observed in the pathogenic *Cryptococcus* species, whereas growth at 30°C appears to be fairly common in all lineages (Figure 3). Under the conditions we tested (low iron media to induce capsule
formation in the pathogenic *Cryptococcus*), visible capsule production was not observed in any of the close or more distant relatives. Previous results have provided evidence that *C. humicola* produces laccase, the enzyme required for melanin production, and a capsule composed of polysaccharides similar to those found in *Cryptococcus* (121). Additional studies have indicated the presence of a capsule in many of the *Tremella* species (39, 142). Like *C. neoformans*, the Trichosporonales *Cryptotrichosporon anacardii* (not included in this study) also produces both capsule and melanin. Biochemical studies have reported that *T. mesenterica* produces extracellular polysaccharides on YPD, such as those often found in the capsules of other Tremellales (39) and the pathogenic *Cryptococcus* (121). We speculate that these contradictory findings might result from different environmental cues triggering capsule production among the different lineages, or the sensitivity of the assays used to detect capsule production. Taken together, the results imply that shared ancestral traits (capsule, melanin, or growth at high temperatures), and also novel traits (or more recently evolved phenotypes) play a role in the pathogenic life styles of the *Cryptococcus* species.

*C. neoformans* and *C. gattii* displayed the highest virulence in the *G. mellonella* model (Figure 4). The species in the *Kwoniella* clade were also able to infect *G. mellonella* larvae and several exhibited intermediate virulence compared to the pathogenic *C. neoformans* and *C. gattii* species (Figure 5). The outgroup in this study, *C. humicola* has been previously isolated from immunocompromised patients (128) insect frass, and the gastrointestinal tracts of insects (150). Although the specific strain used in this study was isolated from soil (Table 1), it also displayed intermediate virulence in
Galleria. All of the Tremella isolates (T. globispora and the two T. mesenterica strains) were avirulent in G. mellonella (Figure 4). Thus, we hypothesize that ancestral characters in both the Kwoniella and Filobasidiella lineages may play a role in the pathogenic potential (Figure 3) of the pathogenic Cryptococcus species.

Moreover, several traits associated with C. neoformans pathogenic potential, for example virulence in G. mellonella and melanin production, appear to be present in other closely and distantly related species. Therefore, these shared traits might be examples of pleisiomorphic, ancestral traits that were retained in these species due to selection by the environment. Alternatively, some of these characters and their distribution, such as the faint melanin-like pigment produced by C. humicola, could be the product of convergent evolution. Current advances, such as the recently sequenced genome of T. mesenterica by the Joint Genome Institute (JGI) at the Department of Energy (isolate ATCC24925), in addition to the available sequenced genomes of several strains of C. neoformans and C. gattii (97), set the stage for future comparative genomic analyses to differentiate between these divergent evolutionary trajectories.

Within the Saccharomycotina the human associated and sometimes pathogenic Candida species are clustered within lineages of endosymbiotic and commensal Candida species found in the gastrointestinal tracts of insects (150). In a complementary study, four novel anamorphic yeast species were isolated from the gastrointestinal tracts of flower-visiting beetles in China (68). These novel insect-associated Candida species closely cluster with the human fungal pathogenic C. albicans/Lodderomyces elongisporus clade (68). The emergence of human pathogenic Cryptococcus species from an ancestral
lineage of mycoparasitic and insect frass associated fungi might be similar to these previously illustrated examples in the ascomycetes. Consequently, strong evolutionary selection might drive the emergence of successful human pathogens from ancestral yeasts associated with insects or other insect habitats or selection for isolates that survive at higher temperatures.
Table 1: Table of species in Chapter 1.

<table>
<thead>
<tr>
<th>Species</th>
<th>Strain Number</th>
<th>Original substrate and Location</th>
<th>Clade</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullera dendrophila</td>
<td>CBS6074</td>
<td>Frass of Buprestis larvae in South Africa</td>
<td>Kwniella</td>
</tr>
<tr>
<td>Cryptococcus amyloletrus</td>
<td>CBS6039</td>
<td>Frass of beetles in South Africa</td>
<td>Filobasidella</td>
</tr>
<tr>
<td>Cryptococcus bestiolae</td>
<td>CBS10118</td>
<td>Frass of the litchi fruit borer in Vietnam</td>
<td>Kwniella</td>
</tr>
<tr>
<td>Cryptococcus dejecticola</td>
<td>CBS10117</td>
<td>Frass of the litchi fruit borer in Vietnam</td>
<td>Kwniella</td>
</tr>
<tr>
<td>Cryptococcus heveanensis</td>
<td>CBS5569</td>
<td>Sheet rubber in Indonesia</td>
<td>Kwniella</td>
</tr>
<tr>
<td>Cryptococcus humicola</td>
<td>CBS571</td>
<td>Soil</td>
<td>Trichosporon</td>
</tr>
<tr>
<td>Cryptococcus gattii VGI</td>
<td>CBS10510</td>
<td>From debris of Eucalyptus tereticornis (WM276)</td>
<td>Filobasidella</td>
</tr>
<tr>
<td>Cryptococcus gattii VGI</td>
<td>CBS10514</td>
<td>Bronchial wash of male patient in Canada (R285)</td>
<td>Filobasidella</td>
</tr>
<tr>
<td>Cryptococcus neoformans var. grubii VNI</td>
<td>CBS8710</td>
<td>Patient with Hodgkin's lymphoma in North Carolina (H99)</td>
<td>Filobasidella</td>
</tr>
<tr>
<td>Cryptococcus neoformans var. neoformans VNI</td>
<td>ATCC MYA-565</td>
<td>Serotype D alpha lab congenic strain (JEC21)</td>
<td>Filobasidella</td>
</tr>
<tr>
<td>Filobasidella deapauperata</td>
<td>CBS7841</td>
<td>Dead spider in Canada</td>
<td>Filobasidella</td>
</tr>
<tr>
<td>Kwniella mangroviensis</td>
<td>CBS6507</td>
<td>Mangrove Cay in Bahamas</td>
<td>Kwniella</td>
</tr>
<tr>
<td>Tremella globispora</td>
<td>CBS6972</td>
<td>Diaporthie on Cornus nuttallii in Canada</td>
<td>Tremella</td>
</tr>
<tr>
<td>Tremella mesenterica</td>
<td>CBS6973</td>
<td>Fallen branch of Alnus rubra in Canada</td>
<td>Tremella</td>
</tr>
<tr>
<td>Tremella mesenterica</td>
<td>ATCC24925</td>
<td>Dead branch of Alnus rubra in Canada</td>
<td>Tremella</td>
</tr>
<tr>
<td>Tsuchiyae wingfieldii</td>
<td>CBS7118</td>
<td>Frass of Scolytid beetles in South Africa</td>
<td>Filobasidella</td>
</tr>
</tbody>
</table>

* Whole genome sequences available
Table 2: GenBank accession numbers.

<table>
<thead>
<tr>
<th>Species</th>
<th>EF1α</th>
<th>mitSSU</th>
<th>RPB1</th>
<th>RPB2</th>
<th>nuclSU</th>
<th>ITS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bullera dendrophila (CBS6074)</td>
<td>FJ534856</td>
<td>FJ534887</td>
<td>FJ534917</td>
<td>FJ534933</td>
<td>FJ534901</td>
<td>FJ534871</td>
</tr>
<tr>
<td>Cryptococcus amyloleotus (CBS6039)</td>
<td>FJ534857</td>
<td>FJ534888</td>
<td>FJ534918</td>
<td>FJ534934</td>
<td>FJ534902</td>
<td>FJ534872</td>
</tr>
<tr>
<td>Cryptococcus bestiae (CBS10118)</td>
<td>FJ534858</td>
<td>FJ534889</td>
<td>FJ534919</td>
<td>FJ534935</td>
<td>FJ534903</td>
<td>FJ534873</td>
</tr>
<tr>
<td>Cryptococcus dejecticola (CBS10117)</td>
<td>FJ534859</td>
<td>FJ534890</td>
<td>FJ534920</td>
<td>FJ534936</td>
<td>FJ534904</td>
<td>FJ534874</td>
</tr>
<tr>
<td>Cryptococcus heuanensis (CBS569)</td>
<td>FJ534860</td>
<td>FJ534891</td>
<td>FJ534921</td>
<td>FJ534937</td>
<td>FJ534905</td>
<td>FJ534875</td>
</tr>
<tr>
<td>Cryptococcus humicola (CBS571)</td>
<td>FJ534861</td>
<td>FJ534892</td>
<td>FJ534922</td>
<td>FJ534906</td>
<td>FJ534876</td>
<td></td>
</tr>
<tr>
<td>Cryptococcus gatti VGI (CBS10510)</td>
<td>FJ534862</td>
<td>FJ534924</td>
<td>FJ534939</td>
<td>FJ534908</td>
<td>FJ534878</td>
<td></td>
</tr>
<tr>
<td>Cryptococcus gatti VGI (CBS10514)</td>
<td>FJ534863</td>
<td>FJ534923</td>
<td>FJ534938</td>
<td>FJ534907</td>
<td>FJ534877</td>
<td></td>
</tr>
<tr>
<td>Cryptococcus neoformans var. grubii VN1 (CBS8710)</td>
<td>FJ534864</td>
<td>FJ534894</td>
<td>FJ534925</td>
<td>FJ534940</td>
<td>FJ534909</td>
<td>FJ534879</td>
</tr>
<tr>
<td>Cryptococcus neoformans var. neoformans VN9 (ATCC MYA-565)</td>
<td>FJ534865</td>
<td>FJ534895</td>
<td>FJ534926</td>
<td>FJ534941</td>
<td>FJ534910</td>
<td>FJ534880</td>
</tr>
<tr>
<td>Filobasidiella depauperata (CBS7841)</td>
<td>FJ534866</td>
<td>FJ534896</td>
<td>FJ534927</td>
<td>FJ534942</td>
<td>FJ534911</td>
<td>FJ534881</td>
</tr>
<tr>
<td>Kwoniealla mangroviensis (CBS8507)</td>
<td>FJ534866</td>
<td>FJ534897</td>
<td>FJ534928</td>
<td>FJ534943</td>
<td>FJ534912</td>
<td>FJ534882</td>
</tr>
<tr>
<td>Tremella globispora (CBS6972)</td>
<td>FJ534867</td>
<td>FJ534929</td>
<td>FJ534944</td>
<td>FJ534913</td>
<td>FJ534883</td>
<td></td>
</tr>
<tr>
<td>Tremella mesenterica (CBS6973)</td>
<td>FJ534868</td>
<td>FJ534898</td>
<td>FJ534930</td>
<td>FJ534945</td>
<td>FJ534914</td>
<td>FJ534884</td>
</tr>
<tr>
<td>Tremella mesenterica (ATCC24925)</td>
<td>FJ534869</td>
<td>FJ534899</td>
<td>FJ534931</td>
<td>FJ534946</td>
<td>FJ534915</td>
<td>FJ534885</td>
</tr>
<tr>
<td>Tsuchiyaea wingfieldii (CBS7118)</td>
<td>FJ534870</td>
<td>FJ534900</td>
<td>FJ534932</td>
<td>FJ534947</td>
<td>FJ534916</td>
<td>FJ534886</td>
</tr>
</tbody>
</table>
Chapter 3. Evolution of MAT in the pathogenic Cryptococcus species and neighboring taxa, Tsuchiyaea wingfieldii and Cryptococcus amylolentus

3.1 Introduction

There are two systems that define the genetic arrangement of fungal mating, the bipolar and tetrapolar mating systems (97). In a bipolar mating system, transcription factors that establish mating-type (MAT) are arranged in a single locus and in some examples genes encoding the pheromones and their receptors are also part of the MAT locus (40). For mating to occur compatible cells must have different MAT alleles (α and a), although there are examples of bipolar fungi that also undergo same-sex mating (i.e. Cryptococcus neoformans, α-α, (89)). The tetrapolar mating system is comprised of two unlinked genomic regions that also control and establish cell identity, these loci are often multiallelic, and both alleles of MAT must differ for sexual reproduction to occur. Bipolar mating systems also promote inbreeding while the tetrapolar systems promote outbreeding (74). Many ascomycetous yeast such as Saccharomyces cerevisiae and Candida albicans are bipolar while basidiomycetous yeasts are typically tetrapolar like Tremella mesenterica and Ustilago maydis (45). In dramatic contrast to most basidiomycetous species, Ustilago hordei, Coprinellus disseminatus, C. neoformans, and C. gattii all exhibit a bipolar mating system (5, 67, 86).

C. neoformans is a haploid, dimorphic fungus that has evolved a bipolar mating system, represented by two alleles, α and a (86). MAT spans 100 to 120 kilobases, and encodes more than 20 genes, many of which are not known to be involved in mating. Comparison of the MAT gene cluster among the members of the pathogenic
Cryptococcus species complex reveal that extensive rearrangements and gene conversions have occurred over time even though recombination in this gene cluster is generally suppressed (43, 44, 115). There are two interesting features of the α allele: 1) it is linked to virulence and α isolates can be more virulent than α isolates (84), and 2) it is more common in clinical and environmental isolates (80). For sex to occur in the pathogenic Cryptococcus species, both alleles of MAT do not have to differ (42, 89). Moreover, evidence for same-sex mating (also known as monokaryotic fruiting) suggests that in addition to its a-α heterothallic (outbreeding) lifestyle, C. neoformans and C. gattii also undergo an α-α homothallic life cycle (inbreeding). C. neoformans may have evolved the ability to undergo same-sex mating due to the absence of a compatible mating partner, or the evolution of same-sex mating may have resulted in the relative loss of the other mating-type.

Fraser et al. proposed that the ancestral form of MAT in the pathogenic Cryptococcus species was tetrapolar, with the homeodomain (HD) and pheromone/receptor (P/R) genes present in two unlinked sex-determining regions (41). Sequential rounds of gene acquisition led to the expansion of the ancestral tetrapolar MAT subloci. This expansion has also been described as a key hallmark in mammalian sex chromosome evolution (41, 46). A chromosomal translocation event fused the unlinked loci into a contiguous region resulting in the formation of a transient tripolar intermediate in which MAT is linked in one partner yet unlinked in the other. This unstable intermediate underwent gene conversion to link the other MAT locus alleles, one or the other homeodomain gene was lost, and MAT was subjected to multiple inversions.
and gene conversions events to yield the extant bipolar MAT locus of Cryptococcus. In sex-determining regions, recombination is suppressed and likely the result of several genomic events: inversions, gene conversions, gene rearrangements, and accumulation of repetitive elements and transposons (43).

A recent multi-locus sequence typing (MLST) approach and phylogenetic study highlighted the relationships of species in the pathogenic Cryptococcus species complex (38). The monophyletic Filobasidiella clade is comprised of not only the pathogenic species but also the sensu stricto or closest known relatives of C. neoformans and C. gattii, which are the saprobic and sibling species Tsuchiyaea wingfieldii and Cryptococcus amyloleucus (38). Additionally, Filobasidiella depauperata, a homothallic filamentous fungus, is also an integral member of the Filobasidiella clade and along with the two sibling species all are named according to the sexual structures produced during mating (81). The sensu lato, sister clade Kwoniella encompasses several saprobic and one aquatic-associated species: Bullera dendrophila, Cryptococcus heveanensis, Cryptococcus bestiolae, Cryptococcus dejecticola, and Kwoniella mangroviensis (38). Because of the close relationship of these species to the pathogenic Cryptococcus species complex, understanding the life cycle transitions of members of the sensu lato and particularly of the sensu stricto will provide key insights into the evolution of MAT in C. neoformans and C. gattii. MAT has been cloned and sequenced in C. heveanensis while in F. depauperata, only partial sequence for the MAT locus has been generated (127). The analysis of C. heveanensis revealed that MAT is tetrapolar with two unlinked gene clusters, a multiallelic HD locus and at least a biallelic P/R locus (102).
In this study, we characterized the life cycle of *T. wingfieldii* and *C. amylolentus* by sequencing MAT employing comparative genomic approaches and phylogenetics to predict the evolutionary intermediates that shaped the formation of this structure in the pathogenic *Cryptococcus* species (61). In both species, two unlinked gene clusters are present, one encodes the HD locus and the other encodes the P/R locus embedded within a dynamic >100 kb region containing homologs of many *Cryptococcus* MAT-associated genes. Pulsed-field gel electrophoresis of chromosomes (~8-10) isolated from *T. wingfieldii* and the two strains of *C. amylolentus* revealed that the HD and the P/R locus are unlinked and present on different chromosomes. Further characterization of the dimorphic SXI1 and SXI2 region in the two *C. amylolentus* strains suggest that the HD locus is restricted to the homeodomain genes and Southern blot analysis in both strains demonstrate that at least the pheromone receptor gene, *STE3*, displays allelic diversity (Sheng Sun, personal communication). Our studies (Chapter 4) provide evidence that *STE3* and the HD transcription factors, SXI1 and SXI2, govern opposite sexual identity in *C. amylolentus*. This study highlights a key evolutionary intermediate in the formation of MAT, which involves unlinked HD and P/R loci with MAT-associated genes and hypothetical genes from *C. neoformans* chromosomes 4 and 5 (61). We hypothesize that a chromosomal translocation occurred and that this event was integral in shaping the evolution of MAT. This study and similar analyses inform our understanding of sex-determining gene cluster evolution and formation in animals, plants, and fungi.
3.2 Materials and methods

3.2.1 Strains and media

The two strains of *C. amylolentus*, CBS6039 and CBS6273, and *T. wingfieldii*, CBS7118, were obtained from the Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre in the Netherlands. All species were grown and maintained on yeast extract-peptone-dextrose (YPD) medium at 24°C.

3.2.2 DNA extraction

To isolate genomic DNA from *T. wingfieldii* and *C. amylolentus*, cells were cultured in 50 mL of liquid YPD shaking overnight at 24°C. The pellets were then lyophilized overnight and the CTAB method of fungal DNA isolation was performed (165). Plasmid DNA from positive TOPO clones was extracted using the QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA), fosmid DNA was isolated using a modified miniprep protocol, and DNA from the shot-gun sequencing libraries was extracted using the DirectPrep96 Miniprep Kit (Qiagen, Valencia, CA). Additionally, progeny DNA was isolated using a modified miniprep protocol and colony lifts were performed to isolate DNA from individual colonies in each fosmid library according the protocol described in (127).

3.2.3 Degenerate PCR

We designed degenerate PCR primers using the online computer program, COnsensus-DEgenerate Hybrid Oligonucleotide Primer (CODEHOP, [http://blocks.fhcrc.org/codehop.html](http://blocks.fhcrc.org/codehop.html)) to identify *MAT* specific genes in *T. wingfieldii*. 
The primers consist of a relatively short 3’ degenerate core and a longer 5’ non-degenerate consensus clamp designed by multiple sequence alignments (129). We aligned sequences for two flanking genes, *FAO1* (JOHE14679-CCACCGGCACCACGNGAYCARAARG, JOHE14683-GGGAGGGTACTTCTTCTGGGNCCYSDAT) and *NOG2* (JOHE15835-TCCGACGCTCGTCAATCCAYGNTNGA, JOHE15840-GGGATCTTTGCGCGGWTVMARCTT), and two recently acquired *MAT* genes, *RPO41* (JOHE15850-CCGACAAGATGTGTCATHACNATHYT, JOHE15857-GCACTCGCCGAGCCARTCYWKDAT), and *LPD1* (JOHE15841-GCCTCAAGACCGCTGRTNGARARAG, JOHE15845-GGAGGGATGCGGCGCRTARTTNACRT) from *C. neoformans, C. grubii, C. gattii, U. maydis*, and *C. cinerea* to design the degenerate PCR primers. PCR was performed on genomic DNA isolated by the CTAB extraction method as template and products were separated by gel electrophoresis. Products with the strongest ethidium bromide-staining signal were then gel extracted using the QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) followed by transformation into *E. coli* using the TOPO-TA cloning Kit (Invitrogen, Carlsbad, CA). Transformants were then sequenced. For *C. amylolentus*, degenerate primers were not used. Instead, primers from *T. wingfieldii* were directly used to amplify *MYO2* (JOHE18132-TCTGTGACCGATGCGATGTTGTAATG, JOHE18133-CCCGTGATAGGTATGATGAAG), *LPD1* (JOHE20330-CAATGGCGACAAACGACATC, JOHE20331-GGCGAAACTCAAGTGCTGAAT), and *SXI1* (JOHE21462-
TCGGAGATGTCGGAGAAGG, JOHE21463-TCGTGAGTTCTTGCTTTGCTG) and SXI2 (JOHE21460-AGGTTTGTCTCGGCTGGAGA, JOHE21461-ACCAGATTATGGGCAGACTGCT) in both C. amylolentus strains.

3.2.4 Fosmid library preparation and fosmid library screening

We employed the CopyControl™ Fosmid Library Production Kit (Epicentre, Madison, WI) to generate fosmid libraries for T. wingfieldii and C. amylolentus strain CBS6039. At least 2.5 µg of CTAB isolated genomic DNA was randomly sheared using a 200 µl small bore pipette tip and end-repair sheared DNA was converted to blunt 5’ phosphorylated ends using End-Repair Enzyme Mix, dNTPs, and ATP. We then separated the end-repaired DNA overnight using contour-clamped homogenous electric field (CHEF) on a CHEF DR-II apparatus (Bio-Rad, Hercules, CA). The following conditions were used: 1- to 6-second switch time, 6 V/cm, 14°C for 14-15 hrs in 0.5X TBE. The size-fractionated DNA, 25 to 40 kb fragments, were recovered by gel extraction and the DNA was precipitated with sodium acetate and ethanol. The precipitated insert DNA was then ligated into the CopyControl pCC1FOS cloning-ready vector and incubated overnight at 24°C. The ligated DNA was packaged in phage particles and plated on E. coli phage-resistant cells overnight at 37°C. Approximately 16,000 fosmid clones were picked into 96-well plates and eventually transferred to 384-well plates for long-term storage at -80°C. The 384-well plates were replicated onto high-density filters and hybridizations using the MAT genes (T. wingfieldii: FAO1, NOG2, MYO2, LPD1, and SXI1 and C. amylolentus: MYO2, LPD1, SXI1, and RPL39) were conducted (127).
3.2.5 Sequencing and assembly

Positive fosmid clones were sequenced using the shot-gun sequencing method described in (102). Six fosmids were pooled and sequenced to generate the assembly for *T. wingfieldii* and three fosmids were individually sequenced to generate the assembly for *C. amylolentus* strain CBS6039. Sequencing reactions were performed using Big Dye chemistry v3.1 (Applied Biosystems, Foster City, California, United States) and analyzed on an Applied Biosystems 3730xl capillary sequencer in the Biological Sciences Sequencing Facility at Duke University. For each library, approximately ~1200 sequence reads were imported into UNIX using Phred and Phrap (34, 52, 53) to assemble the sequences into larger contigs of overlapping sequence. To close gaps in the assemblies, we designed primers from contig ends using Primer 3 (http://frodo.wi.mit.edu/primer3/). Genbank accession numbers are listed in Table 3.

3.2.6 Fluorescence Activated Cell Sorting (FACS) analysis

To determine the ploidy of the two *C. amylolentus* strains and *T. wingfieldii*, we cultured the species on YPD medium for 2 days at 24˚C. Each isolate was processed for flow cytometry as previously described (89, 156) and analyzed using the FL1 channel on a Becton-Dickinson FACScan. The ~20 Mb genome of *Cryptococcus neoformans/gattii* was used as a reference for ploidy determination (including haploid and diploid controls).

3.2.7 Pulsed-field gel electrophoresis (PFGE) and chromoblot analysis

To isolate chromosomes of *C. amylolentus* and *T. wingfieldii*, spheroplasts were generated following the spheroplasting protocol for *C. neoformans* and *C. gattii* (170).
The plugs containing spheroplasts were lysed at 55°C for at least 24 hrs in lysing solution (0.5 M EDTA/10 mM Tris-HCl (pH 10) and 1% Sarcosyl) and then loaded onto a PFGE apparatus and separated for approximately 5 days on a CHEF DR-II apparatus (Bio-Rad, Hercules, CA). The following conditions were used: Block 1: 75- to 150-second switch time, 4 V/cm, 13°C for 30 hrs and Block 2: 200 to 400-second switch time, 4 V/cm, 13°C for 60 hrs in 0.5X TBE. The gel was then stained in ethidium bromide for 15 minutes, destained for an hour, and visualized using a UV lamp. The chromosomes were blotted overnight onto Hybond (Amersham, Piscataway, NJ) membranes in 20X SSC using standard protocols with minor modifications (144). The membrane was then hybridized to MAT gene probes generated by PCR. We also performed Southern blot analysis on genomic DNA from C. amylolentus that was digested with EcoRV, PstI, BamHI, and NotI. The digested DNA was separated on an agarose gel and probed with the RPL22 gene probe amplified from C. amylolentus, but primers were designed from T. wingfieldii (JOHE23415-CCCAAGATCACTTCCTCCAA and JOHE23416-GTCCTTAGAGGTGGCAACGA).

3.2.8 Bioinformatic analyses

We compared sequences from the HD locus of T. wingfieldii to C. amylolentus by employing a matrix comparison (or dot plot) analysis. To generate each dot plot, we employed the Molecular Toolkit’s online nucleic acid dot plots program (http://www.vivo.colostate.edu/molkit/dnadot/). The parameters for the dot plot analyses are as follows: the window size is 51 and the mismatch limit is 6. We also employed the bioinformatic software, Artemis Comparison Tool Release 8.
(http://www.sanger.ac.uk/resources/software/) to generate comparison plots across MAT of *T. wingfieldii* to *C. amylolentus* and both sibling species to *C. neoformans* serotype D strain JEC21 (130). The input file was created using WebACT (http://www.webact.org/WebACT/home) with the Blastn algorithm (25).

### 3.2.9 Phylogenetic analysis

Phylogenetic analysis was performed on coding sequences using MEGA4 (155). To determine the phylogenetic relationship, the Neighbor-Joining method based on the Kimura 2-parameter model was employed (131). Statistical support was calculated from 500 replicates.

### 3.2.10 Southern blot analysis

We performed Southern blot analysis (Sheng Sun, pers. comm.) using standard protocols with minor modifications (144) on genomic DNA from *C. amylolentus* digested with BamHI, BglII, ClaI, EcoRI, and NcoI. The digested DNA was separated on an agarose gel and probed with the *STE3* gene (JOHE23993 - AACTTCCCTGGCCCTTGTTT and JOHE23994 – TTATACACCGATCGCTCGAC), stripped (0.1% SDS and 0.1X SSC in boiling water, 3X for 15 minutes each), and probed with the contig ends from the P/R assembly in *C. amylolentus* (*STE12* left forward primer – GCATGTAGTGAGTGAGGATGG, *STE12* left reverse primer – TTTTCCCCGTCCAGTTTCGCC, *LPD1* right forward primer – TCGGTTTCCACACCAATGGC, *LPD1* right reverse primer – TCAATCGGAGGCGGTATTCCT, *GEF1* left forward primer –
GCCCATGTGCTCTTGTGGAA, *GEF1* left reverse primer –
AGCGCAACCGGTATAAGGA, *MFA1* right forward primer -
GGTCTCTCGATCTGTCTTGCA, *MFA1* right reverse primer –
CCGATAAAATGCCAGGTTGCG, *MYO2* left forward primer –
AACACCCACCACCAACCAA, *MYO2* left reverse primer –
TTCAAGAGCGATGACGGGAC, *RPL39* right forward primer –
ATGCACCAGTTAGTGGGCCCT, and *RPL39* reverse primer –
TGAGTCAACAGCGGAAAGGC.
3.3 Results

3.3.1 Characterizing MAT in T. wingfieldii and C. amylolentus

To determine the MAT structure in the saprobic yeast-like species T. wingfieldii strain CBS7118, fosmid libraries were constructed and probed with several genes: MAT-linked MYO2, LPD1, and SXII and the flanking genes FAO1 and NOG2. The five probes were generated using degenerate PCR primers designed from aligning sequence from several basidiomycete species. Positive clones (3F11-3A15-5J15 (P/R locus), 2B23-2K10 (HD locus), and 4E07 (FAO1) see Appendix A) were pooled and sequenced resulting in the identification of two candidate MAT loci. The FAO1 gene was on a distinct fosmid and appears to be unlinked or distant from MAT. The region containing the P/R locus spans 70 kb and the region containing the HD locus spans 40 kb (Figure 6A).

We also cloned and sequenced MAT in the saprobic yeast-like, sibling species C. amylolentus strain (CBS6039) employing the same methods described for T. wingfieldii. Fosmid libraries were generated and probed with MYO2, LPD1, RPL39, and SXII. Primers specific for MAT genes in T. wingfieldii were used to generate probes for C. amylolentus and the identity of each probe was confirmed via cloning and sequencing. Positive clones (4E01 (SXII), 4E22 (MYO2), and 3H19 (LPD1), see Appendix B) were individually sequenced and assembled into two MAT loci. We later identified an additional fosmid (3N14, RPL39, see Appendix B) that was sequenced via primer walking. We experienced difficulty in closing a few relatively small sequence gaps in both the HD and P/R loci; the regions that were sequenced span ~20 kb and ~60 kb respectively, and each contain two small sequence gaps (~3 kb for each gap in the P/R
locus) we continue to work to close (Figure 7A). Because *T. wingfieldii* and *C. amylolentus* are closely related species (38), it is not surprising that the gene content in *MAT* appears to be largely conserved. However, we determined via Southern blot analysis that the order of the P/R locus genes in *C. amylolentus* differs from *T. wingfieldii* (Figures 6 and 7). Specifically, genomic DNA from CBS6039 and CBS6273 was digested with five restriction enzymes (BamHI, BglII, ClaI, EcoRI, and NcoI) and Southern blot analysis was performed using the ends of each contig in the P/R assembly of *C. amylolentus* as probes (Sheng Sun, pers. comm.). The analysis revealed that three inversion events have occurred between the P/R regions in *T. wingfieldii* compared to *C. amylolentus* suggesting that these species are distinct even though they are phylogenetically closely related (Figures 7 and 12).

The *MAT* loci of *T. wingfieldii* and *C. amylolentus* are organized similarly with the exception of the inversions and the defining features of this structure are highlighted in Figure 12. Analysis of the loci revealed that >20 *MAT*-associated genes have been identified and the gene content of the regions is similar to the *C. neoformans* and *C. gattii* *MAT* loci (86). In *T. wingfieldii* and *C. amylolentus*, only *FCY1* and *UAP1* flank the 5’ end of *MAT* while *FAO1* is unlinked and present elsewhere in the genome. We observed that *STE11* is not present in the P/R locus but based on PCR analysis it is located elsewhere in the genome in both *T. wingfieldii* and *C. amylolentus*. In the *MAT* locus of the pathogenic *Cryptococcus* species, *STE11* is present. In *C. heveanensis*, *STE11* is linked to but distant from and this may represent the ancestral configuration with retention in *C. neoformans* and *C. gattii* and translocation out of *MAT* in *C. amylolentus*.
and *T. wingfieldii* (102). In *T. wingfieldii*, the flanking gene at the 3’ end of MAT, NOG2, was used as a probe. It was present in a single contig within the larger fosmid assembly of *T. wingfieldii*, but has not been linked to either the HD or P/R loci contigs. This region proved difficult to sequence. PCR analysis (using gap closure) revealed a link to *LPD1* in the P/R locus, although this gap remains to be sequenced. Interestingly, *NCP1* and *NCP2* are duplicated genes in *T. wingfieldii* and *C. amylolentus* but not in the pathogenic *Cryptococcus* species. The *NCP1/2* genes are also duplicated in *C. heveanensis* (102), suggesting this might be ancestral.

We also identified several hypothetical genes (CND06020, CND06030, CND06040, CND01650, CNBE0480, CNE02690, and CNE02670) with *C. neoformans* genes as the most closely related homolog in other sequenced fungal genomes. A majority of these genes reside on chromosome 4 and a minority on chromosome 5 of *C. neoformans*, indicating that translocation (intra- and inter-chromosomal) events may have occurred between these two chromosomes during the evolution of *MAT* in the pathogenic *Cryptococcus* species (41, 127). In *C. heveanensis*, *F. depauperata*, and *T. mesenterica* there is additional evidence for such exchange between chromosomes (46, 102, 127). The arrangement of the *MAT* subloci in *T. wingfieldii* and *C. amylolentus* is a likely evolutionary intermediate in *MAT* evolution, the equivalent of a living fungal fossil/missing evolutionary link in which the loci (or their linked gene repertoire) are expanded but not yet fused.
Figure 6: *T. wingfieldii* MAT loci and chromosomal locations.
(A) Six fosmids were used to generate the assembly for *T. wingfieldii*. The MAT gene probes used to probe the *T. wingfieldii* library are indicated in blue. The HD (B) and P/R (A) loci span 40 and 70 kb respectively. Grey arrows indicate genes that either flank MAT or are hypothetical genes, black arrows are *Cryptococcus* MAT-specific genes, and yellow indicates the most recently acquired genes. Scale bar = 10 kb. (B) Chromosomes were isolated from *T. wingfieldii* and separated using PFGE. Three MAT-specific probes were used, two from the HD locus and one from the P/R locus.
Figure 7: *C. amylolentus* MAT loci and chromosomal locations.

(A) Four fosmids were used to generate the assembly for *C. amylolentus*. The MAT gene probes used to probe the *C. amylolentus* library are indicated in blue. The HD (B) and P/R (A) loci span 20 and 60 kb respectively. Grey arrows indicate genes that either flank MAT or are hypothetical genes, black arrows are *Cryptococcus* MAT-specific genes, and yellow indicates the most recently acquired genes. Several gaps remain in the MAT loci of *C. amylolentus*. Scale bar = 10 kb. Green bar under the assembly denotes gaps in sequence. (B) Chromosomes were isolated from *C. amylolentus* and separated using PFGE. Three MAT-specific probes were used, one from the HD locus and two from the P/R locus. The *RPL22* gene was also used as a probe.
In both sibling species, orthologs of both \textit{SXI1} and \textit{SXI2} are present in the HD locus implicating this as the ancestral configuration. The orientation of the homeodomain transcription factors mirrors the organization of the divergently transcribed genes, \textit{bE} and \textit{bW}, in the tetrapolar basidiomycete \textit{Ustilago maydis} (136). In contrast, in \textit{C. neoformans} and \textit{C. gattii}, only one HD gene is present and \textit{SXI1\alpha} is specific to the \textit{\alpha} allele while \textit{SXI2\alpha} is specific to the \textit{\alpha} allele; the partner gene has been lost in each extant \textit{MAT} allele.

The region corresponding to the P/R locus contains the mating pheromone genes, the pheromone receptor gene \textit{STE3}, and the five genes that were hypothesized to be those most recently acquired (\textit{LPD1}, \textit{RPO41}, \textit{BSP2}, \textit{CID1}, and \textit{GEF1}). In \textit{T. wingfieldii}, three pheromone genes (\textit{MF\alpha}1 and \textit{MF\alpha}3 are identical while \textit{MF\alpha}2 differs in only one amino acid) are present and share greater identity with the \textit{MF\alpha} genes of \textit{C. gattii} with an identity of 80\% compared to 70-75\% shared with the \textit{MF\alpha} pheromone gene (Figure 8). In \textit{C. amylolentus}, two pheromone genes (\textit{MF\alpha}1 and \textit{MF\alpha}2 differ in only two amino acids) have been identified and are slightly more diverged from the \textit{MF\alpha} protein product of \textit{C. neoformans} with an identity of 73\% and an identity of 65-70\% with the \textit{MF\alpha} pheromone gene (Figure 8). In summary, both \textit{SXI1} and \textit{SXI2} were present in the ancestral HD locus. We hypothesize that in the pathogenic \textit{Cryptococcus} species, loss of one or the other HD gene occurred during the evolution of \textit{MAT}. The five most recently acquired genes are linked to the ancestral P/R locus and not to the HD locus, in contrast to an earlier evolutionary model, suggesting a revision to the model (Figure 16) (41).
3.3.2 The HD and P/R loci are physically unlinked in the sibling species

To determine whether the HD and P/R loci are physically unlinked, we performed pulsed-field gel electrophoresis isolating the chromosomal DNA content in *T. wingfieldii* and both strains of *C. amylolentus* (CBS6039 and CBS6273). We succeeded in resolving approximately 8-10 chromosomes ranging in size from 800 kb to 2.2 Mb. Three genes were used to probe the *T. wingfieldii* chromosomes, two from the HD locus, *SXI1* and *RPL22*, and one from the P/R locus, *MYO2* (Figure 6B). For *C. amylolentus*, a total of four genes were used as probes: one from the P/R locus, *SXI1*, and two from the HD locus, *MYO2* and *ETF1* (Figure 7B). We also probed with *RPL22*, which was not present in the *MAT* assembly. From the chromoblot analysis, the two subloci are unlinked and located on separate chromosomes in the sibling species (~0.8 and 0.9 Mb). An unlinked HD and P/R loci indicate a tetrapolar-mating configuration for both *T. wingfieldii* and *C. amylolentus* in which the two loci lie on different chromosomes. Moreover, this orientation serves as the ancestral form of *MAT* in the formation of the bipolar mating system for the pathogenic *Cryptococcus* species. It appears that *RPL22* in *C. amylolentus* has been duplicated and it is present on the two chromosomes that correspond to the HD and P/R loci (Figure 7B). It will be interesting to elucidate how *RPL22* came to be duplicated in future studies. One model is that a telomeric-telomeric chromosomal fusion led to a dicentric chromosome whose resolution resulted in a segmental duplication containing the *RPL22* gene. In summary, the organization of *MAT* in these species resembles key aspects of the proposed intermediates in the evolution of *MAT* in the pathogenic *Cryptococcus* species.
3.3.3 Identification of key genes that define MAT

*MAT* is defined as a gene cluster (containing either HD and/or P/R genes) whose sequence is divergent between two strains of opposite mating-types. Based on the characterized structure of *MAT* in both species, we sought to determine which genes in each region govern and control sexual identity. The lack of additional *T. wingfieldii* strains has made it difficult to assess experimentally whether it has a sexual cycle and if so, which genes are involved.

Figure 8: Alignment of the mating pheromone genes in the pathogenic *Cryptococcus* species and related taxa. Sequence alignments of the pheromone gene in *C. neoformans* JEC21 MFα, *C. grubii* H99 MFα, *C. gattii* WM276 MFα, *C. neoformans* JEC20 MFα, *C. grubii* 125.91 MFα, *C. gattii* E566 MFα, *C. heveanensis* CBS569 MFα, *C. amylolentus* CBS6039 MFα, *T. wingfieldii* CBS7118 MFα1, and *T. mesenterica* ATCC24925 Tremerogen a-13. The black arrow denotes the predicted cleavage site.

Fortunately, in *C. amylolentus*, two strains are available and this enabled our analysis of *MAT* and sex in this species resulting in the discovery of an extant sexual cycle (see also Chapter 4).

The *SXI1* and *SXI2* dimorphic region defines *MAT* in *C. amylolentus*. We aligned the nucleotide sequences and performed a matrix comparison for the dimorphic region.
(~2 kb) spanning the SXI1 and SXI2 genes in CBS6039 and CBS6273. The diversity lies in the region between the two genes, and their divergently oriented 5’ regions span roughly 600 bp with a similarity score of 92% (Figure 10). This region encodes the N-terminal dimerization regions known to be variable and which also define alleles in other species. Moreover, the sequence length for CBS6273 is slightly shorter than CBS6039 at the 3’ end of the region we sequenced for the SXI2 gene and the dotted line in this region represents the difference. In summary, the SXI1 and SXI2 genes span ~3 kb and define the sex locus in *C. amylolentus*. Regions that define MAT typically display polymorphisms when comparing sequences from strains of opposite mating-type while the genes that flank MAT share a much higher level of identity (≥99%).
Figure 9: The RPL22 gene is duplicated in the C. amylolentus genome. Genomic DNA from the C. amylolentus parental strain CBS6039 was digested with EcoRV, PstI, BamHI, or NotI. The samples were resolved on an agarose gel and Southern blot analysis was performed using RPL22 (PCR amplified from C. amylolentus) as a probe. Two cross-hybridizing species are present in the PstI, BamHI, and NotI lanes.

To determine whether the pheromone receptor, STE3, is linked to mating-type, we performed Southern blot analysis using genomic DNA digested with BamHI, BglII, ClaI, EcoRI, and NcoI from the two strains of C. amylolentus, CBS6039 and CBS6273. The
STE3 PCR product derived from CBS6039 was used as a probe, and only hybridized to the lanes containing CBS6039 DNA while we observed no hybridization with CBS6273 (Figure 11) (Sheng Sun, pers. comm.). We are currently working to isolate the STE3 sequence for CBS6273. This analysis suggests that the STE3 gene differs between the two C. amylolentus strains and the pheromone receptor is also linked to mating-type. Furthermore, in basidiomycetes unlinked HD and P/R loci indicate a tetrapolar mating system and the nature of sex in C. amylolentus is likely a tetrapolar mating system in which the HD and the P/R loci define mating-type identity and both are responsible for contributing to sex.
Figure 10: The homedomain genes, *SXI1* and *SXI2*, define *MAT*. A percent identity plot of both *C. amylolentus* strains, CBS6039 and CBS6273 comparing the *SXI1* and *SXI2* dimorphic region in the HD locus. The red ellipsoid represents an EcoRV site, which only cleaves *SXI1* in CBS6039 while the blue ellipsoid represents Rsal, which only cleaves *SXI2* in CBS6039.
Figure 11: The pheromone receptor, *STE3*, defines *MAT*. Genomic DNA from the two *C. amylolentus* strains was digested with BamHI, BglII, ClaI, EcoRI, and Ncol and Southern blot analysis was performed using the *STE3* PCR product from CBS6039 as a probe against the parental strains CBS6039 and CBS6273. Data provided by Sheng Sun.

3.3.4 Synteny across *MAT* in sibling species

The multi-locus sequencing project revealed that *T. wingfieldii* and *C. amylolentus* are sibling species that share a close phylogenetic relationship to one another (38). This likely explains why we could successfully utilize *T. wingfieldii* specific primers to amplify genes in the genome of *C. amylolentus*. To determine if synteny is shared across *MAT* between these species, we employed the bioinformatic software, Artemis Comparison Tool Release 8, to compare *MAT* in *T. wingfieldii* and *C. amylolentus*. Unlike the synteny comparison between *C. neoformans* serotypes A and D (41) strains, synteny across the a and α alleles in the cryptic species *C. gattii* VGI, II, and
III is conserved. A major caveat in the assembly for C. amylolentus is that several gaps remain in the HD and P/R assemblies. However, we successfully completed Southern blot analysis and determined the order of the three contigs in the P/R assembly (Figure 7) (Sheng Sun, pers. comm.). From our analysis, a level of synteny exists across both MAT subloci in T. wingfieldii and C. amylolentus, but we also observed three inversion events that have occurred in C. amylolentus and are highlighted in blue in the P/R locus (Figure 12). We further compared each sibling species to the C. neoformans serotype D strain JEC21. Comparing JEC21 to each of the species revealed extensive gene rearrangements and inversions present throughout MAT (Figure 13) similar to the comparisons of MAT within the C. neoformans/C. gattii species complex.

Figure 12: Synteny exists across MAT in the sibling species. MAT sequences from T. wingfieldii and C. amylolentus were compared and a synteny analysis was performed. Red denotes conserved gene order while blue indicates inversion events. Green bars under the assembly denote gaps in sequence.
Figure 13: Extensive inversions and gene rearrangements are present throughout MAT in the sibling species and C. neoformans strain JEC21. MAT sequences from T. wingfieldii and C. amylolentus were compared to C. neoformans and a syntenic analysis was performed. Red denotes conserved gene order while blue indicates inversion events. Green bars under the assembly denote gaps in sequence.

3.3.5 Phylogenetic analysis of genes in C. amylolentus and T. wingfieldii

We conducted phylogenetic analysis of several unlinked and MAT-linked genes (CID1, ETF1, GEF1, LPD1, STE3, STE20, SXI1, and SXI2). In this analysis, we included C. neoformans var. neoformans, C. neoformans var. grubii, and C. gattii representatives from the pathogenic species cluster. We also included the sibling species and closest relatives C. amylolentus and T. wingfieldii, and the outgroup species C. heveanensis and
T. mesenterica. Interestingly, C. heveanensis and T. mesenterica displayed similar phylogenetic patterns, which were observed in T. wingfieldii and C. amylolentus, while the pathogenic species differed in their phylogenetic patterns. The sibling species exhibited three distinct phylogenetic patterns. Specifically, STE20 and ETF1 are very diverged from both the α and a mating-type specific alleles and display a mating-type specific phylogeny in C. neoformans and C. gattii, STE3 is a mating-type specific, and LPD1, GEF1, and CID1 demonstrate a species-specific pattern (Figures 14 and 15). Lastly, the alleles of SXI1 and SXI2 in C. amylolentus and T. wingfieldii are also very diverged from the pathogenic Cryptococcus species complex. In summary, the sibling species displayed similar phylogenetic patterns for the genes analyzed, while the pathogenic species differed in their phylogenetic profiles.
Figure 14: Phylogenetic patterns of *C. amylolentus* genes. The phylogenetic relationship of *C. amylolentus* to the pathogenic *Cryptococcus* species and neighboring taxa is highlighted and four representative genes *STE20*, *ETF1*, *STE3*, and *LPD1* are shown. *STE20* and *ETF1* exhibit a mating-type specific phylogeny in *C. neoformans* and *C. gattii*, but not in *T. wingfieldii* or *C. amylolentus*. *STE3* is a mating-type specific, and *LPD1* exhibits a species-specific phylogeny in *C. neoformans*. The trees were constructed using the Neighbor-Joining method and MEGA4. Bootstrap values on tree branches were calculated from 500 replicates. (α) indicates strains with the *MATα* locus, and (a) indicates strains with the *MATa* locus.
Figure 15: Phylogenetic patterns of additional *C. amylorentus* genes. The phylogenetic relationship of *C. amylorentus* to the pathogenic *Cryptococcus* species and neighboring taxa is highlighted and four additional genes *GEF1*, *CID1*, *SXI1*, and *SXI2* are shown. *GEF1* and *CID1* display a species-specific phylogeny and the *SXI1* and *SXI2* alleles are very diverged from the pathogenic *Cryptococcus* species. The trees were constructed using the Neighbor-Joining method and MEGA4. Bootstrap values on tree branches were calculated from 500 replicates. (α) indicates strains with the *MATα* locus, and (a) indicates strains with the *MATα* locus.
3.4 Discussion

The current study extends the previous analyses of the MAT locus in the pathogenic Cryptococcus species to the closest known relatives, T. wingfieldii and C. amylolentus. To determine the structure of MAT in both species, we cloned and sequenced the HD and P/R loci. We also identified and defined the sexual cycle of C. amylolentus (see Chapter 4). Due to their close phylogenetic relatedness (38), characterization of MAT has provided key insights into the evolution of MAT and revealed important aspects of the transition from an ancestral tetrapolar to a bipolar mating system in C. neoformans and C. gattii. Furthermore, sex chromosomes across all kingdoms share similar evolutionary events. Most importantly, these regions arose initially from autosomes and then underwent expansion, gene rearrangements and conversions, including suppression of recombination to form the sex-determining regions found in mammals, plants and fungi (37, 44, 46, 115).

A recent robust multi-locus sequence typing analysis resolved the phylogeny surrounding the pathogenic Cryptococcus species cluster and revealed that T. wingfieldii and C. amylolentus are sibling species, the closest relatives of the pathogenic species, and thus members of the Filobasidiella clade (38). The MAT structure for T. wingfieldii and C. amylolentus is syntenic and although three inversion events are present in the assembly for C. amylolentus, gene order is conserved (Figure 12). For this analysis, the type strain, the only isolate of T. wingfieldii available, was used. Two strains of C. amylolentus are available and we characterized MAT for the type strain CBS6039 and representative sequences for CBS6273. The two MAT subloci of T. wingfieldii and C.
*amylolentus* are physically unlinked and present on different chromosomes (Figure 6B and 7B). The *MAT* assembly for *C. amylolentus* is similar to *T. wingfieldii* in that both homeodomain transcription factors are present and divergently oriented, similar to *bE* and *bW* in *U. maydis*. Several other key genes (*SPO14, RPL22*, and *CAP1*) are present and these are also linked to *MAT* in *C. neoformans*. We were unable to clone and sequence the genes downstream of *CAP1* in *C. amylolentus* that are present in *T. wingfieldii*. However, we confirmed by PCR that *RPL22* is present in the genome of *C. amylolentus*. We also performed chromoblot analysis of the *RPL22* gene and the chromosomes for the HD and P/R loci both contain either partial remnants of *RPL22* or the entire gene has been duplicated. To verify our chromoblot data, we conducted Southern blot analysis and the hybridization pattern suggests that *RPL22* has been duplicated in the genome of *C. amylolentus*. We hypothesize that this duplication may have occurred via segmental or chromosomal duplication. Moreover, chromosomal translocation or transposition are also possible events that likely shaped the evolution of *RPL22*.

We also completed phylogenetic analysis on several *MAT*-linked and unlinked *MAT* markers. The analysis revealed that the phylogenetic patterns of these genes in *T. wingfieldii* and *C. amylolentus* are similar in that some genes are species-specific while *SXI2* and *STE3* are mating-type specific. The patterns for *T. mesenterica* and *C. heveanensis* were also identical to one another and to the sibling species and the profiles of the pathogenic *Cryptococcus* species changed depending on the gene that was analyzed. In *T. wingfieldii* and *C. amylolentus* species, the P/R locus and linked regions encode the pheromone receptor, *STE3*, the mating pheromone-sensing genes *STE12* and
STE20, the mating pheromone genes, and many other MAT-linked genes. STE11 is an exception in the sibling species; it is not linked to the P/R locus but instead is present elsewhere in the genome based on PCR (127). Although we elucidated the structure of the MAT loci, we cannot conclude whether the mating system of T. wingfieldii is tetrapolar or bipolar. In C. amylolentus, the organization of the loci appear tetrapolar and genotyping, RAPD, and mating analyses (see Chapter 4) support that it has a tetrapolar mating system.

Of the >20 genes identified in the HD (B) and P/R (A) loci of the sibling species, we determined which genes define MAT. Because only one strain of T. wingfieldii is available, we were unable to establish which of the genes in the B locus and the A locus are MAT-specific. Conversely, in C. amylolentus, we determined by comparing sequences from the two parental strains that the MAT-specific region in the HD locus is restricted to the ~3 kb SXI1 and SXI2 dimorphic region. Of the ~3 kb region, we sequenced ~2 kb and the divergence between the two parental strains CBS6039 and CBS6273 spans ~600 base pairs with a similarity score of 92%. The divergence is present in the 5’ regions of SXII and SXI2. Interestingly, the N-terminal regions of the homeodomain proteins in fungi are typically variable and heterodimerization only occurs when compatible (or different allelic versions) of the proteins are brought together promoting activation of genes required for sexual development (145).

We also sought to define MAT in the A locus by performing a Southern blot analysis of the pheromone receptor gene STE3 to determine how the STE3 sequences in the two strains of C. amylolentus are related. In mushroom fungi, the mating-type genes
display characteristic hyperpolymorphism (67). Our Southern blot findings revealed that the CBS6039 STE3-specific probe only hybridized to CBS6039 and not to CBS6273 (Sheng Sun, pers. comm.). We speculate that the STE3 sequences are therefore dissimilar in the two strains of *C. amylolentus*. Establishing a transformation and gene disruption system in *C. amylolentus* will advance our understanding of the role of not only STE3 and the mating pheromone genes, but also *SXI1* and *SXI2* during sexual reproduction in this species. We propose that the MAT locus of *C. amylolentus* possesses a tetrapolar mating system in which the B and A locus genes are linked to mating-type and therefore define MAT and we show this to be the case experimentally in the next chapter. In *C. amylolentus*, the B and A genes are currently biallelic and further characterization of the mating system in this species requires the identification of additional isolates to determine if either or both are multiallelic.

There is abundant evidence for translocation events in some members of the *Filobasidiella*, *Kwoniella*, and *Tremella* clades. The MAT-specific genes in *F. depauperata*, *C. heveanensis*, and *T. mesenterica* all contain a subset of hypothetical genes that correspond to chromosomes 4 and 5 in *C. neoformans*. The examples of chromosomal rearrangements (i.e. intrachromosomal events) are higher in the two chromosomes, suggesting that the arrangement of MAT in these species resembles the ancestral form of MAT (127). As the MAT region evolved these genes were lost in *C. neoformans* but maintained in the closely and more distantly related species (Figure 16). The predicted translocation events occurred near the telomeric regions of chromosome 4 in *C. neoformans* JEC21 (102, 127). Moreover, analysis of the two unlinked MAT loci in
*T. wingfieldii* and *C. amyloleunus* revealed many of the same hypothetical genes are also associated with *MAT*, indicating that translocation events were essential in shaping the evolution of the mating-type locus in the pathogenic *Cryptococcus* species.
Figure 16: Model for the evolution of the mating-type locus in the pathogenic *Cryptococcus* species.
The unlinked ancestral tetrapolar HD and P/R loci contained both homeodomain genes and the pheromone/receptor genes respectively. Additional genes were acquired into both loci, expanding the MAT-specific region. A translocation event occurred likely between chromosomes 4 and 5 of *Cryptococcus* resulting in the formation of a transient tripolar intermediate and one of the HD genes was lost. The hypothetical genes (grey arrows) relocated, likely through a translocation event to the telomeric ends of chromosome 4. The unstable tripolar intermediate later collapsed to a bipolar state. The fused loci were subjected to further gene rearrangement and gene conversion events, which led to the formation of the bipolar alleles of the pathogenic *Cryptococcus* species. White arrows
indicate the five most recently acquired genes and black arrows are MAT-specific genes present in the pathogenic Cryptococcus species.

This study, and those conducted on C. heveanensis and F. depauperata, reveals insights into the evolution of MAT in the pathogenic Cryptococcus species (102, 127). In T. wingfieldii and C. amylolentus, the two MAT subloci are unlinked based on chromoblot analysis. The two regions in which the HD and P/R loci are embedded span over 100 kb, encode >20 genes, and also contain hypothetical genes not linked to MAT in C. neoformans. A model for the evolution of MAT is described in Figure 16. The ancestral form of MAT was tetrapolar and represented by two unlinked loci. One locus encoded the homeodomain transcription factors (B) and the second locus encoded the pheromone genes and pheromone receptor (A). This region expanded to include additional MAT-specific genes. The loss of one of the HD genes likely occurred either simultaneously or prior to a chromosomal translocation event that resulted in the formation of a tripolar intermediate. The tripolar transition state later collapsed to form a bipolar mating system, which was subjected to further rearrangement and gene conversion to yield the current bipolar MAT locus of the pathogenic Cryptococcus species.

We have substantial evidence that suggests the ancestral form of the B locus contained both SXII and SXI2, similar to tetrapolar mating systems in other basidiomycetes, and that loss of one of the HD genes was a key step in the formation of a bipolar mating system. In S. cerevisiae species, the a2 HMG gene has also been lost from MAT and in Candida species the HD α2, or both the α2 and a1 HD genes have been lost (19, 125, 143). The bipolar system restricts outbreeding but promotes inbreeding and the
consequence is increased clonality. In pathogenic fungi, inbreeding may contribute to maintain gene configurations required for virulence (112).

From the characterization of $MAT$ in $T. wingfieldii$ and $C. amylolentus$, we identified the following proposed intermediates in the evolution of $MAT$: the presence of linked and likely divergently transcribed homeodomain genes $SXI1$ and $SXI2$, physically unlinked $B$ and $A$ loci, and chromosomal translocation events which include genes from the telomeric ends of chromosome 4 in $C. neoformans$. Several structural features define sex chromosomes and include expansion of the sex-determining region, suppression of recombination, and chromosomal rearrangements and translocation events. In fungi, plants, and animals these features are all common and conserved in evolutionarily divergent organisms and reflect the common evolutionary forces that shaped the sex-specific regions of their genomes (41).

We propose that $MAT$ in the pathogenic $Cryptococcus$ species evolved from an ancestral tetrapolar system with unlinked $B$ and $A$ loci and these loci fused into a large bipolar $MAT$ locus. In $C. amylolentus$, the structure of $MAT$ indicates a tetrapolar mating system with allelic diversity in both the $B$ and $A$ loci. Although evidence from $T. wingfieldii$, $C. amylolentus$, $C. heveanensis$ (102), and $C. disseminatus$ (67) suggests that $MAT$ evolved from a tetrapolar to a bipolar system, we cannot rule out the possibility that the ancestral form of $MAT$ was bipolar and instead evolved into a tetrapolar mating system. In such a scenario, a bipolar locus would have suffered a chromosomal break resulting in the formation of unlinked HD and P/R loci in a derived rather than ancestral tetrapolar fungal species. In this model, the tetrapolar state would then be ancestral in
some species and derived in others. Gene rearrangements and gene conversions may have occurred prior to the break in MAT. The model (Figure 16) we propose illustrates the evolution of the bipolar MAT in the pathogenic Cryptococcus species from an ancestral tetrapolar system and the evidence adduced for the sibling species supports our current model. Given that C. amylolentus, C. heveanensis, and Tremella mesenterica all have unlinked HD and P/R loci further suggest that tetrapolarity represents the ancestral configuration, and that the transition to bipolarity occurred recently and concomitantly with the emergence of the pathogenic C. neoformans/C. gattii species cluster.
Table 3: Genbank accession numbers for *T. wingfieldii* and *C. amylolentus*.

<table>
<thead>
<tr>
<th>Species</th>
<th>HD locus</th>
<th>P/R locus</th>
<th>Additional genes</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Tsuchiyayea wingfieldii</em> (CBS7118)</td>
<td>HM368525</td>
<td>HM368524</td>
<td></td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6039)</td>
<td>HM640220</td>
<td>HM640221</td>
<td></td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6039)</td>
<td></td>
<td></td>
<td>HM640223</td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6039)</td>
<td></td>
<td></td>
<td>HM640222</td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6273)</td>
<td></td>
<td></td>
<td>HM640224-SXII</td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6273)</td>
<td></td>
<td></td>
<td>HM640225-SXII2</td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6273)</td>
<td></td>
<td></td>
<td>HM640226-GEF1</td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6273)</td>
<td></td>
<td></td>
<td>HM640227-LPD1</td>
</tr>
<tr>
<td><em>Cryptococcus amylolentus</em> (CBS6273)</td>
<td></td>
<td></td>
<td>HM640228-ETF1</td>
</tr>
</tbody>
</table>
Chapter 4. Discovery of an extant sexual cycle for *Cryptococcus amylolentus*

4.1 Introduction

Sexual reproduction is ubiquitous throughout nature, generates population diversity, and has been described extensively in plants, animals, and microorganisms (105). The generation of diversity via genetic recombination in microbes, especially pathogenic species may be essential to maintain genes or factors required for virulence. Sex is costly and the benefits must outweigh the disadvantages or asexual species would arise more frequently. Some benefits of sex are accumulation of favorable mutation combinations, removal of deleterious mutations, and the ability thereby to adapt to changes in the host and environment by producing more fit recombinant offspring (32, 58, 105, 172). In sexually reproducing populations, outbreeding is common, but inbreeding forms of sex also occur that result in clonality. The benefits of clonal reproduction are similar to opposite-sex mating and same-sex mating may be an adapted virulence strategy for several microbial pathogens (58). Conversely, sexual reproduction is also disadvantageous (32). Sex is not always efficient and meiosis expends energy not required in asexual reproduction. Identifying a suitable mating partner also comes at a cost, especially in environments where opposite mating partners are scarce (32).

Modified forms of sex or sexual reproduction are present in many microbial pathogens such as viruses, bacteria, parasites, and fungi. In viruses, co-infection of pathogenic viruses results in recombination and antigenic shifts that can result in pandemics (105). Bacteria employ three distinct processes that lead to gene transfer:
transformation, conjugation, and transduction. Moreover, bacteria have exploited these
gene transfer mechanisms for horizontal genetic exchange between different bacterial
species and even across kingdoms (55, 172). Parasites also undergo sexual reproduction
and recombination, which has led to increased virulence throughout the evolution of
these eukaryotic microbes. Additionally, pathogenic fungi have exploited several modes
of sexual/parasexual reproduction to propagate the spread of their infectious propagules
(58). Some pathogenic microbes possess the ability to reproduce through same-sex
mating. This process promotes clonality and inbreeding, and may also impact virulence.

The pathogenic *Cryptococcus* species form a monophyletic cluster, which is
composed of *C. neoformans* var. *neoformans*, *C. neoformans* var. *grubii*, and the sibling
species *Cryptococcus gattii* (VGI, VGII, VGIII, VGIV) (155). The cluster encompasses
at least six species that have the potential to infect humans, animals, plants, and insects
(150, 151). Recent multi-locus sequence analysis resolved the phylogeny surrounding the
members of the monophyletic *Filobasidiella* clade, which includes the two sibling
species *C. amylolentus* and *T. wingfieldii* (both have unknown sexual cycles), and the
homothallic species *F. depauperata* (38, 109). The sister *Kwoniella* clade is more
extensive and includes notably *K. mangroviensis* and *C. heveanensis*. *T. mesenterica*
served as an outgroup for the study. All of the neighboring taxa are saprobic yeasts and
non-pathogenic, but a few display intermediate virulence in the greater wax moth

In the 1970s, sex was discovered in the pathogenic *Cryptococcus* species (61).
The sexual cycle and the structures governing sex (the mating-type locus) in the
pathogenic *Cryptococcus* species have been extensively examined and are well defined (27, 40). In a laboratory setting, *Cryptococcus* reproduces via either opposite-sex mating or same-sex mating (monokaryotic fruiting) (43, 78, 88, 89, 169). Because the mating system of *C. neoformans* and *C. gattii* is bipolar and defined by an a and α allele, mating (a-α) occurs via the following main steps: cell-cell fusion, production of a filamentous dikaryon with fused clamp cell connections, and nuclear fusion and meiosis both occurring in the basidia (27, 40). Specifically, meiosis gives rise to four haploid meiotic nuclei that undergo mitotic division to produce long chains of basidiospores that germinate into fertile yeasts that can mate with a partner/parent of the opposite mating-type. The major difference in monokaryotic (α-α) fruiting is that a monokaryon (instead of a dikaryon) forms, mating can involve two genetically distinct isolates (α1-α2) or two genetically identical genomes (α1-α1), and the resulting meiotic spore products are all α.

The pathogenic *Cryptococcus* and surrounding species are classified based on the morphological features of their sexual structures (81). These species all belong to the *Filobasidiella* clade and sex in *C. neoformans* and *C. gattii* generates hyphal filaments with fused clamp connections and basidia with four basipetal chains of basidiospores while hyphae generated by *F. depauperata* have unfused clamp connections (83, 127).

Of the species that are phylogenetically related to the pathogenic *Cryptococcus* species complex, sex has recently been described for *C. heveanensis* and *K. mangroviensis* (102, 148). Specifically, an extant heterothallic sexual cycle was observed in these two members of the *Kwoniella* clade and basidiospores associated with cruciate-septated basidia are produced during mating. Additionally in *F. depauperata* and *T.*
mesenterica, the nature of sex has also been revealed in previous studies and exemplifies homothallic and heterothallic life cycles, respectively (6, 9, 105, 148). The mating structures of F. depauperata resemble the basidia and basidiospores of C. neoformans and C. gattii while T. mesenterica mating products are similar to C. heveanensis and K. mangroviensis (6, 9, 148). The sexual cycles of the sibling species were unknown until recently. There is only one strain of T. wingfieldii known and therefore our ability to identify a sexual cycle for this isolate is limited. Fortunately, two strains of C. amylolentus are known and turn out to be of opposite mating-type. Here, we report the discovery of an extant heterothallic sexual cycle for C. amylolentus. We termed the teleomorph, Filobasidiella amylolenta. Successful matings between the two C. amylolentus strains produce dikaryotic hyphae with abundant basidia decorated with four long glorious spore chains composed of very round spores. Quite interestingly, the mating system of C. amylolentus is tetrapolar, in stark contrast to the bipolar sex-determining systems of the closely aligned pathogenic species C. neoformans and C. gattii. The C. amylolentus sexual cycle also generates many sterile progeny, illustrating a cost of sex, and also blastospores with the nuclear genome of one parent and the mitochondrial genome of the other indicating another potential benefit of sex.
4.2 Materials and methods

4.2.1 Strains and media

The two strains of *C. amylolentus*, CBS6039 and CBS6273, and *T. wingfieldii*, CBS7118, were obtained from the Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre in the Netherlands. All other isolates were obtained from the Heitman lab strain collection. All species were grown and maintained on yeast extract-peptone-dextrose (YPD) medium at 24°C and mating assays were performed on V8 medium pH=5 in the dark and also at 24°C. Microdissection was performed on YPD medium without (random spore dissection) and with zymolyase (to dissect individual spore chains) and incubated at 24°C.

4.2.2 Description of sexual cycle as *Filobasidiella amylolenta*

Standard description: *Filobasidiella amylolenta* Findley & Heitman sp. nov.

Etymol.: The epithet is chosen to be identical with that of *Cryptococcus amylolentus* (Van der Walt, D.B. Scott & Klift) Golubev (50). Paired cultures of *C. amylloentus* type strain, CBS6039T (A1B1), crossed to CBS6273 (A2B2) on V8 medium pH=5. Heterothallic fungus. Hyphae dikaryotic, clamped connections fused. Aseptate basidia, 3-5 µm diameter, and embedded in agar terminating in four chains of basidiospores. Basidiospores are aerial, round, and 2-2.5 µm in diameter.

Latin description: *Filobasidiella amylolenta* Findley & Heitman sp. nov.

4.2.3 Microscopy

Spores and yeast cells were cultured on slides coated with V8 pH=5 medium for one week to allow production of mating structures. The slide was first washed with phosphate buffered saline (PBS) followed by staining the cell wall using a solution of Calcofluor white (fluorescent brightener 28 F-3397; Sigma) for 15 minutes. Slides were rinsed with PBS and fixed for 15 minutes in fixing solution (3.7% formaldehyde and 1% Triton-X100 in PBS). After permeabilization of the fungal cells, nuclear content was examined by staining with Sytox green (Molecular Probes) for 30 minutes. Slides were washed with PBS and a cover slip was applied to the slide for observation. In addition to staining spores and yeasts, mating filaments were also stained. Agar pieces were removed from mating plates and washed several times with PBS. Calcofluor white was added directly to the agar piece for 30 minutes, followed by washing with PBS, and fixing for 45 minutes. After permeabilizing samples, filaments were washed with PBS and stained with 1 mg/ml Hoechst 33258 (Invitrogen, Carlsbad, CA) overnight at 4°C. The next day, samples were washed with PBS, a thin slice of the agar (containing the mating filaments) was removed using a razor blade and a mounting solution containing anti-fade (Invitrogen, Carlsbad, CA) was added to the agar slice on a slide. The slides were sealed with nail polish and stored at 4°C in the dark after microscopic evaluation. All staining was performed at 24°C, unless otherwise noted. SEM was performed on C. amylolentus
matings incubated on V8 pH=5 medium for 2 weeks. The specimen was prepared and analyzed as described in (127). Microscopy was performed with an Axioskop 2 plus upright microscope (Zeiss). Images were captured using an AxioCam MRm camera. Scanning electron microscopy was performed and viewed on a JEOL JSM 5900LV (JEOL U.S.A., Peabody, MA) SEM at 15 kV.

4.2.4 Fluorescence Activated Cell Sorting (FACS) analysis
To determine the ploidy of the microdissected progeny, we cultured the germinated spores on YPD medium for 2 days at 24°C. Each isolate was processed for flow cytometry as previously described (89, 156) and analyzed using the FL1 channel on a Becton-Dickinson FACScan. The ~20 Mb genome of Cryptococcus neoformans/gattii was used as a reference for ploidy determination (including haploid and diploid controls).

4.2.5 Genotyping
Microdissection of spores (random or individual spore chains using zymolyase (Zymo Research Corp., Orange, CA, USA)) was performed on YPD medium incubated at 24°C for two days to allow spores to germinate. Genomic DNA was isolated from each germinating spore and PCR was employed to amplify several MAT-specific genes: SXII (JOHE21462-TCGGAGATGTCGGAGAAGG, JOHE21463-TCGTGAGTTCTTGCTTTGCTG) and SXI2 (JOHE21460-AGGTTTTGTCTCGGCTGGAGA, JOHE21461-ACCAGATTATGGGCAGACTGCT), STE3 (JOHE23993-AACTTCCTGGCCCTTGTCTT, JOHE23994-TTATAACACCGATCGCTCGAC), GEF1 (JOHE24616-
AAGCGCAGAAGCTGAAAGAC, JOHE24618-GCGCTACCCAAATGTCAAC),

ETF1 (JOHE24001-TCCACAGACTCACCCCTCTC, JOHE24002-
TTTCGCAACCTCTTCAACC), and RPL39 (JOHE20326-
TCCGTGTTGCTACCATCTTC, JOHE20327-TGCCTGAGGATGAAAAGAAGA).

4.2.6 RAPD PCR and RFLP analysis

To detect the differences between the two strains of *C. amylolentus*, CBS6039 and CBS6273, and the 65 progeny, RAPD PCR was performed using a set of 25 random primers (Table 7). Primers were designed as previously described in (127). PCR programs were also identical to the conditions described in (127). PCR products were analyzed by agarose gel electrophoresis. The PCR reactions were repeated three times and the results were reproducible. The nine unique polymorphic bands identified were sequenced. RFLP was used to genotype several markers namely ETF1, SXI1, and SXI2. Products were PCR amplified with the genotyping primers and digested for 1 hr at 37°C. The ETF1 product was digested with DdeI, the SXI1 product was digested with EcoRV, and the SXI2 product was digested with RsaI. The digested samples were analyzed on an agarose gel. The same enzymes were used to digest the SXI1 and SXI2 gene products amplified in the parental strains including the set 2 F1 progeny #16. The products were analyzed following the same procedure for the RFLP analysis of the F1 progeny.
4.3 Results

4.3.1 Discovery of an extant sexual cycle in *C. amylolentus*

It was previously thought that the haploid, sibling species *T. wingfieldii* and *C. amylolentus* were asexual (13); however, our recent findings have revealed an extant heterothallic sexual cycle for *C. amylolentus*. We conducted mating assays and found the following optimal conditions: V8 pH=5 solid medium with incubation for one week or longer at room temperature in the dark. The cross between strains CBS6039 and CBS6273 produced hyphae with fused clamp connections and aseptated basidia terminating in four long individual spore chains (see description in Materials and methods section), similar to matings in *C. neoformans* and *C. gattii*. Sterigmata were not observed (Figure 17A-F). A marked, obvious difference in the meiotic progeny is the shape of the spores. In particular, the spores of the pathogenic species are ellipsoid (short or long) (164) whereas the spores of *C. amylolentus* are round and similar in size to yeast cells. Crosses of either *C. amylolentus* strain with *T. wingfieldii* did not result in the production of mating hyphae. Because there is only one strain of *T. wingfieldii* available, *T. wingfieldii* might be fertile in the presence of a suitable partner, similar to the two interfertile *C. amylolentus* strains, or it could be a sterile isolate.

In *C. amylolentus*, we observed that the periphery of some mating patches contain a mixture of both monokaryotic hyphae and sectors in which mating occurs to produce dikaryotic hyphae indicative of sexual reproduction. This is an interesting feature of *C. amylolentus* and the dikaryotic sectoring phenotype is present in most mating patches (Figure 17G) and also serves as a visual assay to determine if mating has occurred. We
performed microscopy on the mating patches to visualize the structures produced during the sexual cycle of \textit{C. amylolentus} in greater detail. The four spore chains are each very long consisting of >15 (quantified by counting 10 individual basidia) spores per chain and clamp cell connections are visible by light microscopy and SEM (Figure 17A-F). We also employed fluorescence microscopy to examine the nuclear content in spores and in the mating hyphae. Although, the mating hyphae were embedded in the agar and thus difficult to stain, we were able to utilize the nuclear stain, Hoechst 33258. We visualized dikaryotic hyphae and both uni-and bi-nucleated spores (Figure 18A-D). Additionally, we employed Calcofluor white and Sytox green to visualize clamp cells and the nuclear content of spores (Figure 18E-H). In the \textit{Filobasidiella} lineage, \textit{C. neoformans} and \textit{C. gattii} produce both dikaryotic (heterothallic) and monokaryotic (homothallic) hyphae while \textit{F. depauperata} produces only monokaryotic hyphae. The presence of dikaryotic hyphae in \textit{C. amylolentus} provides evidence that opposite-sex mating occurs during the sexual cycle (78, 79). Additionally, the presence of two nuclei in some basidiospores is an interesting phenomenon and could result from either a post-meiotic nuclear division in the spore or packaging of two nuclei into some spores as is known to occur in pseudo-homothallic species (123).

Interestingly, the cap of the spore chain represents a quartet of basidiospores. These spores are the oldest in the spore chain and remain tightly attached to each other. Younger spores in the four spore chains remain attached to the preceding and following spores in the chain but often not to their meiotic siblings in the other three spore chains (Figures 17C-D). Thus, the quartet spore cap appears to tether the ends of the spore
chains together. This feature has not been described in the pathogenic *Cryptococcus* species. Microscopic examination of mating structures in *C. amylolentus* has revealed both shared hallmarks with sexual reproduction in the pathogenic *Cryptococcus* species and novel features.

We microdissected a total of 65 basidiospores and examined the progeny produced by *C. amylolentus* sexual reproduction for any obvious phenotypic differences. We observed that a subset of the progeny are self-filamentous on YPD medium and this phenotype does not appear to segregate with either mating-type locus (Table 16). Both parents appear filamentous on YPD but CBS6273 is more filamentous on YPD than the CBS6039 parent. By light microscopy, growth of CBS6273 on YPD results in a mixture of hyphae (including pseudohyphae) and yeast-like cells not immediately apparent when CBS6039 is initially grown on YPD solid medium. Fluorescence microscopy of the hyphal cells reveals monokaryotic filaments while the mating filaments are dikaryotic. In the first F1 progeny set, 6 out of 28 progeny are filamentous on YPD and only one of the filamentous isolates is fertile (set 1 F1 #18) (Table 10). In the second set, fertility and filamentous growth are correlated. Of the 13 fertile progeny out of 31, 12 are also filamentous on YPD. It is possible that filamentous growth indicates fertility although several of the filamentous progeny are sterile. In the F2 progeny set, all of the progeny are filamentous and all are fertile, either with the parental strain or interfertile with a subset of the progeny (Table 10).
Figure 17: Sexual reproduction of *C. amylolentus*.
(A) SEM of basidiospores attached to basidia. Scale bar represents 10 µm. (B) SEM of fused and unfused clamp connections. (C and D) Light microscopy at a magnification of 20X of hyphal filaments, basidia, and basidiospores, scale bar = 10 µm. (E) Basidium with youngest spores attached and associated detached spore chains, scale bar = 1 µm. (F) A cluster of basidiospores and basidia. Scale bar = 10 µm. (G) Mating patch on V8 pH=5 medium. Basida and basidiospores are present in the white fuzzy area extending from the patch.
Figure 18: Fluorescence microscopy of *C. amylolentus* mating structures. (A and B) Staining of basidiospores, scale bars = 2 µm. (A) Differential Interference Contrast (DIC) image, and (B) fluorescence image of basidiospores, nuclei stained with Hoechst 33258. (C and D) Staining of a dikaryotic mating filament. (C) DIC image, and (D) fluorescence image of mating filaments, nuclei stained with Hoechst 33258. (E-F) Staining of mating filament and clamp cell. (E) Nuclear content of filament stained with Sytox green and (F) Calcofluor white for cell wall visualization. (G-H) Staining of basidiospores, scale bar = 5 µm. (G) Nuclear content of basidiospores stained with Sytox green and (H) cell wall with Calcofluor white.
Table 4: List of genotyping markers (nuclear and mitochondrial).

<table>
<thead>
<tr>
<th>MARKERS:</th>
<th>NUCLEAR</th>
<th>MITOCHONDRIAL*</th>
</tr>
</thead>
<tbody>
<tr>
<td>HD LOCUS</td>
<td>SXI1</td>
<td>NAD4</td>
</tr>
<tr>
<td></td>
<td>SXI2</td>
<td>NAD5</td>
</tr>
<tr>
<td>PR LOCUS</td>
<td>RPL39</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GEF1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ETF1</td>
<td></td>
</tr>
<tr>
<td></td>
<td>STE3</td>
<td></td>
</tr>
</tbody>
</table>
4.3.2 Genotypic analysis of meiotic progeny

Because *C. amylolentus* sexual spores are large and not completely embedded in the agar, unlike the mating hyphae, we were able to employ a microdissection technique to isolate spores, as commonly used for the pathogenic *Cryptococcus* species. To determine if recombination occurs, and to further assess whether the mating system of *C. amylolentus* is tetrapolar or bipolar (Figure 19), we performed microdissection of random progeny and individual spore chains followed by molecular genotyping analysis (Table 4). We designate the CBS6039 parent as A1B1 and the CBS6273 parent as A2B2 according to the designation used for a tetrapolar mating system from our findings in Chapter 3 and assigning *A* as the P/R locus and *B* as the HD locus (as in *T. mesenterica* and *C. heveanensis* [6, 102]). First, we randomly dissected set 1 F1 progeny from the *C. amylolentus* matings. The spore germination frequency was 70% (28/40) (Table 5). The progeny were all found to be haploid based on FACS analysis using *C. neoformans* as a reference genome (data not shown).
Figure 19: A model comparing the fungal bipolar and tetrapolar mating paradigms. In a bipolar mating system, haploid a and α cells fuse to form a diploid cell, a/α. Sex culminates in meiosis, which gives rise to four meiotic progeny, 2 a and 2 α. The a progeny can mate with the α parent (50%) while the α progeny can mate with the a parent (50%). In a tetrapolar mating system, haploid A1B1 and A2B2 cells fuse to form a dikaryon/diploid A1B1/A2B2. Meiosis also results in the production of four haploid meiotic progeny: A1B1 can mate with the A2B2 parent (25%), A2B2 can mate with the A1B1 parent (25%), and A1B2 and A2B1 are recombinants (50%) that are sterile with either parent but interfertile with one another.

Remarkably, 78% (22/28) of the progeny are sterile and unable to undergo sexual reproduction with either parent or their F1 siblings. A MAT recombinant progeny from set 1 F1 #18 (A2B1) mates with another MAT recombinant progeny from set 2 F1#10 (A1B2) but does not mate with either parent. Molecular analysis of the following genes: SXII, SXI2, RPL39, GEF1, STE3, and ETF1 revealed that the majority of the progeny are closely related to the A1B1 parent (Table 5). For all progeny sets, SXII, SXI2, and ETF1 were scored using PCR-RFLP while the remaining genes were PCR amplified and scored for the presence of absence of a PCR product (see Appendix C). Recombination is scored
according to marker exchange at the P/R and/or HD locus and is infrequent among this progeny set, 3/28 (11\%) are recombinant (#17, 27, and 28) and 1/28 (3.5\%) has independent assortment of the MAT locus chromosomes (#18) and we hypothesize that this is likely due to the dissection of a mixture of yeast cells, blastospores (mitotic pre-meiotic cells produced by budding from the hyphae or clamp cells), and basidiospores (meiotic sexual spores) (Figure 20). This is supported by analysis of the mitochondrial genome segregation shown below.
Table 5: Random spore dissection of set 1 F1 progeny and molecular analysis of the nuclear markers.

<table>
<thead>
<tr>
<th>F1 Set 1</th>
<th>Mating as</th>
<th>SXI2</th>
<th>SXI1</th>
<th>RPL39</th>
<th>GEF1</th>
<th>ETF1</th>
<th>STE3</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>2</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>3</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>4</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>5</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>6</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>7</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>8</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>9</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>10</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>11</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>12</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>13</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>14</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>15</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>16</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>17</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td>18</td>
<td>A2B2</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>19</td>
<td>alpha</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>20</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>21</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>22</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>23</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>24</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>25</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>26</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td>27</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td>28</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
</tbody>
</table>

Yellow = fertile w/ parents  
Blue = fertile w/ siblings from set 2 F1 progeny  
Red = recombinants  

a = CBS6039 A1B1 parent (a=A1 (P/R), a=B1 (HD))  
b = CBS6273 A2B2 parent (b=A2 (P/R), b=B2 (HD))
Table 6: Individual spore dissection of set 2 F1 and F2 progeny and molecular analysis of the nuclear markers.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th>SX12</th>
<th>SX11</th>
<th>RPL39</th>
<th>GEF1</th>
<th>ETF1</th>
<th>STE3</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>1</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>A1B1×A2B2</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>A1B2</td>
<td>b</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>A2B2</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>A1B2</td>
<td>b</td>
<td>b</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>21</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>22</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>23</td>
<td>sterile</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>24</td>
<td>A1B1</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
<td>a</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>27</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>28</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>29</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
<tr>
<td></td>
<td>31</td>
<td>sterile</td>
<td>a</td>
<td>a</td>
<td>b</td>
<td>b</td>
<td>b</td>
</tr>
</tbody>
</table>

Yellow = fertile w/ parents
Blue = fertile w/ siblings
Red = recombinants
from set 1 F1 progeny

<table>
<thead>
<tr>
<th>B locus</th>
<th>A locus</th>
</tr>
</thead>
<tbody>
<tr>
<td>a = CBS6039 A1B1 parent (a=A1 (P/R), a=B1 (HD))</td>
<td></td>
</tr>
<tr>
<td>b = CBS6273 A2B2 parent (b=A2 (P/R), b=B2 (HD))</td>
<td></td>
</tr>
</tbody>
</table>

97
Figure 20: Yeasts, blastospores, and basidiospores are produced during sex in *C. neoformans* and *C. amylolentus*.

To isolate meiotic progeny, dissections are commonly performed. There are three possible products that are dissected during our genotyping and RAPD analysis: yeasts, blastospores (yeast cells produced via mitotic budding from the hyphae/clamp cell), and basidiospores (sexual spores, generated by meiosis). In the genotyping analysis, our interpretation is that set 1 F1 is a mixture of yeasts, blastospores, and basidiospores while set 2 F1 and F2 are solely derived from basidiospores. Adapted from Lin et al., *Nature* **434**, 1017-1021.

To analyze basidiospores specifically, we dissected a second F1 progeny set of four individual spore chains (one chain from four different basidia) using zymolyase to isolate single spores. FACS was also performed on this progeny set and all were haploid (data not shown). The germination frequency was 90% (28/31) and 58% (18/31) of the progeny were sterile (Table 6). Molecular analysis of this set using the same six genes revealed that 64.5% (20/31) of the progeny are closely related to one or the other parent while the other >30% exhibit evidence of independent chromosome reassortment (A1B2 or A2B1) or recombination within the MAT A or B locus. The data differs from the first
progeny set in that more assorted/recombinant progeny and fewer parental genotypes are apparent; consistent with our interpretation that the first set consisted of yeast cells, blastospores, and basidiospores. The second F1 set is likely an accurate representation of meiotic products in *C. amylolentus*. RFLP analysis of *SXI1* products digested with EcoRV and *SXI2* products digested RsaI from set 1 and 2 F1 progeny served as a second approach to confirm the molecular typing data (Figure 21).

**Figure 21**: RFLP analysis of *SXI1* and *SXI2* in *C. amylolentus* progeny. (A) The *SXI1* gene from set 1 F1 progeny and the two parental strains CBS6039 and CBS6273 digested with EcoRV. (B) The *SXI2* gene from set 2 F1 progeny and the two parental strains CBS6039 and CBS6273 digested with RsaI. D = digested and U = undigested. L = 1 kb ladder.

In the RFLP analysis, there is an extra band in progeny #10, #15, and #16 although they appear to type as the CBS6273 parent based on PCR sequence analysis. We hypothesize that the products were not completely digested. In Figure 21, progeny
#16 has a double band and in Figure 22, we repeated the RFLP analysis on #16. The RFLP analysis revealed that progeny #16 types as the CBS6273 parent.

![Image of RFLP analysis](image)

Figure 22: Set 2 F1 progeny #16 is recombinant and types as CBS6039 at SXII and as CBS6273 at SXI2.

RFLP analysis of the recombinant set 2 F1 progeny #16. The progeny and both parental strains were digested with EcoRV and Rsal at SXI1 and SXI2, respectively. These enzymes only cleave the PCR product in the parental strain CBS6039. L = 1 kb ladder.

Furthermore, the molecular data for these progeny indicate that they are identical to CBS6273 and the progeny are not diploid but are haploid by FACS analysis. Interesting features of the second F1 set are: (1) F1 #4 is a bi-mater and mates with both parental strains (see Appendix D) and (2) F1 #16 is recombinant at the HD locus and types as the A1B1 parent at one base pair in SXI1 and as the A2B2 parent at all other
positions in \textit{SXI2} (Figure 24 and Table 6). Also, the spores in chain D are completely identical to each other in both the genotyping and RAPD analysis but differ from both parents (Table 6).

![Figure 23: The \textit{SXI1} and \textit{SXI2} dimorphic region is similar in CBS6039 and set 2 F1 progeny #4. Percent identity plots comparing the ~2 kb region containing the HD genes in the \textit{C. amylolentus} parental strain CBS6039 compared to set 2 F1 progeny #4.](image)

We performed dot plot analysis on the \textit{SXI1} and \textit{SXI2} (~2 kb analyzed) region comparing both parental strains to progeny #4 (bi-mater) and #16 (recombinant at HD locus) (Figures 23 and 24). We sequenced the dimorphic \textit{SXI1} and \textit{SXI2} region in the bi-mater, progeny #4 to determine if any SNPs, deletions, or insertions could explain the ability of this progeny to mate with both parental strains. Interestingly, #4 was 100% identical to CBS6039 (A1B1). We determined that the percent identity between #4 to CBS6039 is 100% and #4 to CBS6273 is 92%, which is in agreement with the genotyping data. For the \textit{MAT} recombinant progeny #16 the percent identity to CBS6039
Figure 24: The \textit{SXI1}-\textit{SXI2} dimorphic region differs between CBS6273 and progeny set 2 F1 progeny #16. Percent identity plots comparing the \(-2\) kb region containing the HD genes in the \textit{C. amylolentus} parental strain CBS6273 compared to set 2 F1 progeny #16. An example of the region of crossover or gene conversion in the HD locus of the recombinant progeny #16 is highlighted with an orange box.

is 92% and to CBS6273 it is 99.9%. The divergence in #16 is in the \textit{SXI1} gene, which types as the CBS6039 parent and the \textit{SXI2} gene is completely identical to CBS6273. Specifically the divergence in \textit{SXI1} in progeny #16 spans just one nucleotide (at the EcoRV site). This is likely the result of a very short gene conversion track; if a crossover occurred, it flanked (on either side) the region immediately surrounding the SNP (Figure 24). We determined that the SNP present in progeny #16 is synonymous with CBS6273 even though the progeny types as CBS6039 (thus, no change in amino acid). We also
completed RFLP analysis of progeny #16 and the results illustrate that this progeny is recombinant at SXI1 and SXI2 at this single position (Figure 22). Further microdissection and analysis of progeny will provide insight into how meiosis controls sex in *C. amylolentus*, in particular the ability of progeny to acquire a recombinant pattern at the SXI1 and SXI2 loci, which may be one mechanism by which new mating-types arise. In summary, the majority of the fertile progeny are fertile with one of the parents.

Of the 59 F1 progeny, there are 14 A1B1 F1 progeny that are fertile with the A2B2 parent, one A2B2 progeny fertile with the A1B1 parent, and one that is fertile with both parents. From the second F1 set, the following progeny are interfertile: set 2 F1 #3 x 13 (A1B1 x A2B2) and #13 x 24 (A2B2 x A1B1). We also examined the mating ability of A1B2 progeny crossed to A2B1 progeny in all three sets. Surprisingly, only set 2 F1 #10 and #16 were interfertile with F2 #1, 2, and 5, and set 2 F1 #10 with set 1 F1 #18. Light microscopic examination of the intercross matings revealed the presence of basidia and basidiospores. Some of the basidia are barren and the relative lengths of the spore chains vary significantly, some chains contain >10 spores/chain while others contain <5 spores/chain (see Appendix D). These are examples of A1B2 crossed to A2B1 interfertile progeny that are sterile with both parents, consistent with a tetrapolar mating system (Figure 27A). All other combinations of A1B2 crossed to A2B1 progeny tested were sterile. The high rate of sterile progeny illustrates a risk to sexual reproduction in generating progeny and could involve sex induced silencing of repetitive elements or transposon insertion(s) within *MAT*. 
Because only 32% (19/59) of the progeny are fertile in both progeny sets, we assessed whether fertility increases with an additional sexual cross or mitotic passage. In all progeny sets, we tested sterility after passaging the progeny. Even after several passages on YPD, the sterile phenotype remained stable. We crossed set 2 F1 #3 and CBS6273 to generate an “F2” progeny set and dissected spores from two individual spore chains using zymolyase. The germination rate was 54% (6/11) and all of the progeny are fertile (50% with the CBS6039 parent and the remaining A2B1 progeny are interfertile with the set 2 F1 A1B2 progeny #10 and #16 (Table 6). We also genotyped the F2 progeny at the $\text{SXI1, SXI2, RPL39, GEF1, STE3, and ETF1}$ loci. Unlike the first F1 set, 50% of the progeny in the F2 generation type as CBS6273 (Table 6). The F2 progeny set is a backcross of set 2 F1 #3 to CBS6273 and therefore ~75% of the markers type as the CBS6273 marker (“b”) rather than 50% for the F1 progeny (Tables 6 and 9). Interestingly, most of the basidia in the cross were barren and if spore chains were present, the number of spores per chain was significantly reduced when compared to matings between the parental strains (light microscopy to inspect mating sector). All of the progeny examined are haploid with the exception of F2 #4, which is diploid by FACS (data not shown). We also tested F2 #4 for self-fertility on V8 but it was not self-fertile (data not shown). The genotyping data indicates that recombinants are present among the progeny and our evidence is that an extant sexual cycle of *C. amylolentus* conforms to a classical tetrapolar mating system with the caveat that sterile progeny are also produced.
Table 7: List of primers, RAPD and mitochondrial markers, used in the study.

<table>
<thead>
<tr>
<th>Primers</th>
<th>Primer Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAD4 Forward (mito)</td>
<td>GGGAGAATTTGATTCAAGTGCAAC</td>
</tr>
<tr>
<td>NAD4 Reverse (mito)</td>
<td>ATGAGTTGCAATCTGGCATCATACT</td>
</tr>
<tr>
<td>NAD5 Forward (mito)</td>
<td>CTATTGGTGTTACAGGAAGCTCAC</td>
</tr>
<tr>
<td>NAD5 Reverse (mito)</td>
<td>GAGCCCTCATACCTGCCCTATTTGAG</td>
</tr>
<tr>
<td>OPA4 Modified</td>
<td>AATCGGGCTGAGG</td>
</tr>
<tr>
<td>OPA5 Modified</td>
<td>AGGGGTCCTTGGAG</td>
</tr>
<tr>
<td>Pi Random 5</td>
<td>AATGGGGCATCCTTCG</td>
</tr>
<tr>
<td>Pi Random 8</td>
<td>TAAGTCTGCTTCGAG</td>
</tr>
<tr>
<td>Pi Random 9</td>
<td>CCCCTCAAGAGACTTGG</td>
</tr>
<tr>
<td>Pi Random 15</td>
<td>CCTATCCGATCCTGG</td>
</tr>
<tr>
<td>Pi Random 20</td>
<td>AACATGGGGTCACGCG</td>
</tr>
<tr>
<td>Pi Random 21</td>
<td>TTCTGAAACCAGCCC</td>
</tr>
<tr>
<td>Pi Random 24</td>
<td>CGTGCAAGGGAGCACCC</td>
</tr>
<tr>
<td>JOHE22492</td>
<td>CTCTCAGACGCAGAAG</td>
</tr>
<tr>
<td>JOHE22621</td>
<td>GACTCACTCCAGCCATG</td>
</tr>
<tr>
<td>JOHE22631</td>
<td>CGGCAGTGCTCTGGCG</td>
</tr>
<tr>
<td>JOHE22643</td>
<td>ATCTGGTCTGGAGAAGT</td>
</tr>
<tr>
<td>JOHE22655</td>
<td>CTTGCGAAGGGTTCCGG</td>
</tr>
<tr>
<td>JOHE22656</td>
<td>ACTATCTGGGGCCGTC</td>
</tr>
<tr>
<td>JOHE22660</td>
<td>GCTCCATAACGGGCGAG</td>
</tr>
</tbody>
</table>
4.3.3 RAPD confirms that meiosis occurs during the *C. amylolentus* sexual cycle

In a complementary, independent approach to determine if meiosis occurs during sexual reproduction, we employed Random Amplification of Polymorphic DNA (RAPD) PCR (103) using markers unlinked to *MAT*. We designed 20 primers with arbitrary sequences and also included several standard RAPD primers (Table 7). A total of five primers yielded nine unique, polymorphic bands (Figure 25). An example gel illustrating a primer, OPA5 modified, that yielded a unique band is highlighted in Figure 26. These products were cloned and sequenced followed by BLAST (2) analysis, which revealed that the nine polymorphic sequences are unlinked to *MAT* and also do not appear to match any sequences in the *C. neoformans* and/or *C. gattii* genomes and are therefore unique to *C. amylolentus*. Scoring these unique regions identified as polymorphic in the parental strains and in their progeny provides robust evidence that recombination occurs during sexual reproduction in *C. amylolentus* and this is also apparent in the second F1 set in which all progeny are clearly recombinant (Tables 8 and 9). In set 1 F1, there are a few progeny that appear to be recombinants either through meiosis or gene conversion (Table 8). In set 2 F1, the observed recombination frequency is much higher than typical mitotic gene conversion frequency. In addition, analysis of the markers implemented in this study revealed that each parental marker was inherited roughly 50:50 across the entire progeny set, and each progeny inherited roughly half of its marker from one parent and the other half from the other parent, as would be expected from meiosis. The observed high level of recombination and the 1:1 segregation of the parental alleles support the conclusion that meiosis has occurred in *C. amylolentus*. Moreover, we
observed that meiotic recombination in *C. amylolentus* resulted in the generation of new combinations of alleles in the progeny given the multiple genotypes present in the different spore chains analyzed. In summary, RAPD analysis provides corroborating evidence that meiosis occurs in *C. amylolentus*.

![Figure 25](image)

Figure 25: Nine unique, polymorphic RAPD bands identified. Representative gels of the polymorphic bands in the parental strains of *C. amylolentus* isolated using RAPD primers JOHE22621, JOHE22643, JOHE22656, and JOHE22655. A = CBS6039 and B = CBS6273. L = 100 bp or 1 kb ladder. Data provided by Sheng Sun.
4.3.4 Uniparental mitochondrial DNA inheritance

We determined the inheritance pattern of mitochondrial DNA (mtDNA) in the progeny produced by *C. amylolentus* sexual reproduction. In *C. neoformans*, mtDNA is inherited uniparentally, from the a parent (173-176). Therefore, we tested the mitochondrial inheritance pattern resulting from mating of the two *C. amylolentus* strains. We first compared the sequence for both mitochondrial markers in the parental strains and identified 12/438 SNPs for *NAD4* in CBS6039 compared to CBS6273 and 6/345 SNPs for *NAD5* in CBS6039 vs. CBS6273. We also conducted PCR and sequencing on the 65 progeny using the SNPs identified for both markers in the parental strains (Tables 4 and 7). All of the progeny with the exception of two from the first F1 set typed as the A2B2 parent, CBS6273 (Table 10). The two progeny (set 1 F1 #13 and 16) that contain the A1B1 mitochondrial genome may descend from dissected parental yeast cells (or progeny in which uniparental mitochondrial inheritance from the A2B2 parent did not occur) (Table 10). For the other non-recombinant progeny that type as the A1B1 parent, the fact that they have the A2B2 mitochondrial genome suggests that they descend from blastospores following cell-cell fusion (Table 10). These results demonstrate that mitochondria are uniparentally inherited from the A2B2 parent during *C. amylolentus* sexual reproduction. This is in agreement with the mitochondrial uniparental inheritance pattern studies in *C. neoformans* and indicates that this phenomenon is evolutionarily conserved in *C. amylolentus*. It also illustrates that the A1B1 progeny we interpret as being derived from blastospores have inherited the mitochondrial genome from the A2B2 parent as a result of cytoduction.
Figure 26: RAPD primer, JOHE22656 – No. 1 produces a unique band in a subset of the progeny.
Representative RAPD gel using primer JOHE22656 comparing progeny to parental strains identified a product that is present in only the CBS6273 parent (A= CBS6039 and B= CBS6273). Data provided by Sheng Sun. Note: Several progeny are missing from the analysis.
Table 8: RAPD analysis of set 1 F1 progeny.
Twenty RAPD primers were examined, a = CBS6039 (in grey) and b = CBS6273. Data provided by Sheng Sun.

| OPA4 MODI | OPA5 MODI | PI Random_5 | PI Random_8 | PI Random_15 | PI Random_20 | PI Random_21 | PI Random_24 No.1 | PI Random_24 No.2 | JOHE22643 | JOHE22655 No.1 | JOHE22655 No.2 | JOHE22656 No.1 | JOHE22656 No.2 | JOHE22656 No.3 |
|-----------|-----------|-------------|-------------|-------------|-------------|-------------|------------------|------------------|-------------|----------------|----------------|----------------|----------------|----------------|----------------|
| F1S1_1    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_2    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_3    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_4    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_5    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_6    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_7    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_8    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_9    | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_10   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_11   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_12   | a a a b a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_13   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_14   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_15   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_16   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_17   | a a b a b a b b b a a a a a a a a b a b       |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_18   | b b b b a b b b b a a a a a a a a b a b       |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_19   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_20   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_21   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_22   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_23   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_24   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_25   | b b b a b a a b b b b a a b b b a b b a b     |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_26   | a a a a a a a a a a a a a a a a a a a a a a a |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_27   | a a b a b a b b b a b a b b b b a b a b     |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
| F1S1_28   | a a b a b a b b b a b a b b b b a b a b     |             |             |             |             |             |                  |                  |             |                |                |                |                |                |                |
Table 9: RAPD analysis of set 2 F1 and F2 progeny including genotypes of individual spore chains.
Twenty RAPD primers were examined, a = CBS6039 (in grey) and b = CBS6273. Genotypes for each spore chain indicated to the right of the table. Data provided by Sheng Sun.

| F1S2_1 | a | a | a | b | a | a | b | a | a | a | a | a | b | b | b | A |
| F1S2_2 | b | b | b | a | a | b | b | a | a | a | b | b | a | a | a | a | B |
| F1S2_3 | b | b | b | a | b | a | b | b | b | a | a | b | a | b | a | a | C |
| F1S2_4 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | D |
| F1S2_5 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | E |
| F1S2_6 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | F |
| F1S2_7 | b | b | b | b | a | b | a | a | b | a | a | a | a | a | a | a | G |
| F1S2_8 | b | b | b | b | a | a | a | a | a | a | a | a | a | a | a | a | H |
| F1S2_9 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | I |
| F1S2_10 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | J |
| F1S2_11 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | K |
| F1S2_12 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | L |
| F1S2_13 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | M |
| F1S2_14 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | N |
| F1S2_15 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | O |
| F1S2_16 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | P |
| F1S2_17 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | Q |
| F1S2_18 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | R |
| F1S2_19 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | S |
| F1S2_20 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | T |
| F1S2_21 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | U |
| F1S2_22 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | V |
| F1S2_23 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | W |
| F1S2_24 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | X |
| F1S2_25 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | Y |
| F1S2_26 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | Z |
| F1S2_27 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | |
| F1S2_28 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | |
| F1S2_29 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | |
| F1S2_30 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | |
| F1S2_31 | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | a | |
| F2_1 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | |
| F2_2 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | |
| F2_3 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | |
| F2_4 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | |
| F2_5 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | |
| F2_6 | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | b | |

Genotypes:
Table 10: Filamentous phenotype and mitochondrial DNA (mtDNA) inheritance identified in all 65 progeny.
The filamentous phenotype was assayed on YPD medium and the mtDNA markers were amplified and sequenced (a = CBS6039 and b = CBS6273). Circled progeny in the set 1 F1 are recombinants, all others in this set are blastospores, and #13 and #16 are possible parental, CBS6039 yeasts).
Figure 27: *C. amylostentus* has a tetrapolar mating system.
A. An example of the tetrapolar mating system that defines *C. amylostentus* sex in which the fertile progeny are sterile with both parents, but are interfertile with one another, and the parental strains are fertile with one another. (“−” indicates lack of sexual reproduction and “+” indicates sexual reproduction occurs.) B. Heterothallic sex occurs in both *C. amylostentus* and *C. neoformans*. Haploid cells of opposite mating-type fuse to form a diploid nucleus, which undergoes meiosis producing four meiotic progeny. The major differences between the two are that 1) *C. amylostentus* is tetrapolar and *C. neoformans* is bipolar and 2) >1 meiotic event may occur in the basidium while in *C. neoformans* there is only one meiotic event per basidium or aneuploids are generated during the sexual cycle of *C. amylostentus*. 
4.4 Discussion

This study is the first description of sexual reproduction in *C. amylolentus*, which was heretofore thought to be strictly asexual. Optimal conditions for mating are similar to *C. neoformans* and *C. gattii* and include V8 pH=5 medium with incubation in the dark and at room temperature for one week or longer with unbagged/non-parafilmed plates (78). Matings between the two parental strains (CBS6039 and CBS6273) result in the production of basidia, dikaryotic hyphae, and fused clamp cell connections. The presence of dikaryotic hyphae is indicative of opposite-sex mating (or some hyphae could contain two α nuclei).

In a bipolar mating system, 100% of the progeny are mating competent with one or the other parental strain (50% with each). In a tetrapolar mating system, 25% of the progeny are fertile with one parent, 25% with the other parent, and 50% of the progeny are recombinant and cannot mate with either parent but are interfertile with siblings. Mating in *C. amylolentus* is distinct in that many of the progeny are sterile. This demonstrates that sexual reproduction may pose a risk in which not all of the progeny produced are fertile. Additionally, meiosis or sporulation-specific lethal mutations may be present in the sterile progeny. We do not think it is due to chromosomal translocations based on CHEF gel analysis and the robust spore germination (70-80%). It is possible that sterile aneuploids (1N+1) are generated during meiosis. FACS analysis of the examined progeny revealed that all of the progeny are haploid with the exception of a single diploid (F2 progeny #4), but FACS is not sensitive enough to detect 1N+1 aneuploids. Employing comparative genomic hybridization (CGH) of the *C. amylolentus*
parental strains with the sterile progeny will be necessary to address the issue of aneuploidy. Sexual reproduction may also increase transposition in the genome and the insertion of transposons in MAT or elsewhere might result in sterility. An additional explanation for increased sterility among progeny is sex induced silencing of repetitive elements within MAT and linked fertility genes. MAT may also be damaged as a result of gene conversion. These models provide testable hypotheses to define the mechanism(s) causing sterility in the C. amylolentus progeny that remain to be elucidated. We also further examined the filamentous phenotype observed in the parental strains. We determined that a majority of the fertile progeny also displayed the filamentous phenotype and none of the afilamentous progeny were fertile. Some were filamentous and sterile with both parents including set 1 F1 #18, set 2 F1 #10 and #16, and three of the F2 progeny (Table 6). These progeny are filamentous and interfertile, mating as A1B2 x A2B1. We conclude that the observed filamentation in the progeny and parental strains may in some cases indicate mating competent strains.

The structures produced during sex resemble matings in the pathogenic Cryptococcus species with a few exceptions (164). The spores are rounder and tightly associated with one another. They remain attached even after the fixing process for SEM preparation. Fixing spores in the pathogenic species usually results in shorter spore chains that tend to fall apart upon fixation. C. amylolentus spores appear to have a tight connection to each other throughout the entire spore chain but not between spore chains except at the very tip whereas in C. neoformans and C. gattii, spores in an individual spore chain are attached, but not attached to other spore chains at the tip. The fact that the
spore chains stay together until digested with enzymes facilitated the microdissection analysis. We also observed that the cap of the spore chains (or the first born spores) form a quartet at the very top of the chain. This quartet structure resembles the tetrads that form in the asci of *S. cerevisiae*. Further analysis and microdissection of this structure would be of interest. We speculate that β-glucans are likely candidates mediating spore-spore attachment.

Micromanipulation of spores in *C. amylolentus* is facilitated due to the larger size of spores (2-2.5 µm wide in *C. amylolentus* and ~1-2 µm wide in *C. neoformans*) (63, 164). In total, 65 spores were dissected, 59 represent sets 1 and 2 F1 progeny and the remaining 6 are F2 progeny. We performed genotyping analysis on all 65 progeny using six *MAT*-linked markers (*SXII*, *SXI2*, *RPL39*, *GEF1*, *STE3*, and *ETF1*). The majority of the first set of F1 progeny typed as the A1B1 parent, CBS6039. The first set, a random progeny microdissection, contained very few recombinants (11%) and we likely dissected a mixture of yeasts, blastospores, and basidiospores (89). Moreover, the recombinant progeny #18 is fertile with set 2 F1 #10 and all other recombinants from set 1 are sterile.

We therefore completed a second round of microdissections dissecting individual spore chains. The progeny typed as either parent at the six markers and only 2/31 (6.5%) are recombinant. In the F2 generation (set 2 F1 #3 X CBS6273), half of the progeny type as A2B2 and 54% of the progeny germinated. This is in contrast to the first two sets and the most plausible explanation is that in this backcross, ~75% of the genome resembles the A2B2 parent while in the F1 progeny population, 50% of the genome is A1B1 and the remaining 50% is A2B2. In set 2 A1B2 F1 #10 and 16 are interfertile with A2B1 fertile
F2 progeny #1, 2, and 5. The discovery of sex in *C. amylolentus* has increased the number of species in the *Filobasidiella* clade that have defined sexual cycles adding to our foundation of knowledge acquired through similar studies in both closely and distantly related species (81, 141, 148).

The microdissection of the six individual spore chains representing set 2 F1 and the F2 progeny generated several genotypes that typify a tetrapolar mating system. These include A1B1, A2B2, A1B2 and A2B1. The A1B2 genotype was underrepresented and set 2 F1 #10 and #16 were the only examples present in this analysis. Of the spore chains analyzed, the fourth chain (progeny 25-31) is sterile and all one recombinant class, A2B1. The simplest explanation is that the basidium underwent a meiotic event in which one of the genotypes is aneuploid and inviable and only the viable genotype, A2B1, survived meiosis. Similar examples of aneuploidy associated with loss of one genotypic class were reported for *C. neoformans* (63). If aneuploids are present, we are currently unable to score the heterozygous state with our markers and FACS cannot discriminate differences such as 1N+1 (viable) or 1N-1 (inviable).

Our studies provide evidence that meiosis occurs during the sexual cycle of *C. amylolentus*. From the genotyping analysis, it is evident that recombinants are present among the progeny population, but a second confirmatory approach to address this question was implemented. We employed RAPD (127) using 20 unlinked and RAPD-specific markers identifying nine unique, polymorphic bands in the parental strains. We also conducted the analysis on the 65 progeny and the number of recombinants detected overall increased based on this analysis. Further examination of the markers employed in
this RAPD study revealed a 50:50 (1:1) segregation pattern of the two parental alleles in the progeny population. This segregation data and the high level of genomic reshuffling (or genetic exchange) in the set 2 F1 progeny is indicative of meiotic recombination and not gene conversion, which would yield biased allelic frequencies. Additionally, RAPD analysis revealed that in the different spore chains from set 2 F1 and the F2 progeny, more than four genotypes are present in a single chain. Several possibilities are considered. In *C. neoformans*, meiosis typically gives rise to four meiotic products and it was recently shown that a single meiotic event occurs in each basidium (63). In *C. amylolentus*, more than one meiotic event could occur in the basidium. This would involve post meiotic nuclear fusion and a second round of meiosis. We speculate that if for example two meiotic events occurred, at most eight genotypes are produced. If a mixture of same-sex and opposite-sex mating occurs, this would generate at least five genotypes (four from the A2B2+A1B1 cross and one from the A1B1 cross). Also, high gene conversion events favoring some alleles over others could result in a non-Mendelian inheritance pattern and skew the resulting genotypes in each individual spore chain gives rise to greater diversity. If sex in this species were highly clonal (A1B1+A1B1, same-sex mating), the generation of multiple genotypes would result in more genetic exchange and also an increase in diversity.

We also speculate that the presence of aneuploids in the progeny population could explain the multiple genotypes (>4) observed in the RAPD analysis. The RAPD markers employed in this analysis differentiate the two parental strains by the presence or absence of a PCR product. If progeny are aneuploid for one or some of the chromosomes, they
could appear unique and differ from the two parental strains. We expect the basidiospores from one basidium to share four common genotypes with the exception of a few rare genotypes (or potential aneuploids). This is also consistent with what we observed from our RAPD data (Table 9) in which some spore chains contain four distinct and several anomalous genotypes that are not completely unrelated to other genotypes in a given spore chain. Moreover, we are currently unable to score the heterozygous state of the aneuploids. In addition to markers that can be scored by PCR-RFLP, CGH analysis can also be applied in future studies to detect aneuploidy. In conclusion sex in *C. amylolentus* involves meiosis but differs from *C. neoformans* in that the pattern of segregation from a single basidium is more complex.

Taken together, the sexual cycle of *C. amylolentus* is complex. The segregation patterns of the DNA markers linked to the mating-type genes fit a tetrapolar mating system. Extending the genotyping analysis to RAPD PCR advanced our current understanding of sex in this species. We conclude that meiosis is an integral component of the *C. amylolentus* sexual cycle. Most importantly, meiosis gives rise to multiple genotypes, aneuploids, and/or sterile progeny. In *Armillaria gallica*, genetic mosaicism is found in the basidiome stage and this novel mechanism (118-120) could also be responsible for the multiple genotypes observed in the progeny of *C. amylolentus* within an individual spore chain, a consequence of repeated rounds of meiosis.

Cytoplasmic exchange has also been observed. The mitochondrial inheritance pattern in the progeny is uniparental and, of the 65 progeny, only two (#13 and 16) from set 1 F1 typed as the CBS6039 parent for both nuclear and mitochondrial markers and are
possible parental yeasts. The recombinants in set 1 F1 (random spore dissection) have undergone cytoplasmic and nuclear exchange while the non-recombining population with the mitochondrion from the other parent are derived from blastospores. Moreover, the blastospores have not undergone nuclear exchange based on our analysis with one possible exception (set 1 F1 progeny #12) that may be explained by chromoduction of a single introgressed chromosome from one nucleus to the other. The remaining progeny from set 2 F1 and the F2 (individual spore chain dissection) progeny typed as the CBS6273 parent. All of the progeny in these sets are recombinants based on RAPD analysis and have completed both cytoplasmic and nuclear exchange. The uniparental mtDNA inheritance detected in *C. amylolentus* is similar to that in *C. neoformans* (176).

In Chapter 3, we determined that *C. amylolentus* has a tetrapolar mating system organization in which the homeodomain transcription factors *SXI1* and *SXI2* and the pheromone receptor *STE3* are dimorphic, govern and control mating, and the two subloci are unlinked and present on different chromosomes. To elucidate the mating system of *C. amylolentus*, we performed dot plot analysis and sequence alignments comparing the *SXI1* and *SXI2* genes and Southern blot analysis using *STE3* as a probe generated from the CBS6039 strain in both parents, CBS6039 and CBS6273. From our analysis of the HD and P/R key genes, we determined that *MAT* divergence in *SXI1* and *SXI2* is 92% (see Chapter 3, Figure 10). Moreover, *STE3* segregates in the progeny and Southern blot analysis reveals that in the parental strains *STE3* differs significantly (see Chapter 3, Figure 11). In *C. amylolentus* the HD and P/R genes are both linked to mating-type and define sex in this species.
Dot plot analysis of the dimorphic HD genes was extended to two key progeny from the second set F1 progeny (Figures 23 and 24). From the analysis, #4 is completely identical across the entire region to CBS6039, #4 and CBS6039 share 92% similarity with #16 and CBS6273, and #16 and shares 99.9% similarity with CBS6273 (Figure 23 and 24). We expected to identify differences between CBS6039 and progeny #4 although we observed no polymorphisms in #4. If differences were observed we would speculate that these nucleotide changes were responsible for the bi-mater phenotype. The bi-mater may have acquired a new mating-type that likely involved the acquisition of an additional homeodomain gene that is unlinked to the HD locus or self-compatible mutations in the HD genes could have generated the bi-mater. We have not observed any self-compatible mutations and speculate that a mutation linked or unlinked to MAT or an epigenetic event could have given rise to the bi-mater. Progeny #16 is sterile and recombinant at the HD locus. SXI1 and SXI2 are adjacent and the tight linkage association of these genes suggests that crossover events are less likely to occur. The dot plot analysis reveals that in SXI1, one nucleotide that differs between CBS6273 and #16 is shared between CBS6039 and #16. While in the remainder of the SXI1 gene and SXI2 gene, the identity between CBS6273 and #16 is completely identical. We hypothesize that gene conversion occurred from CBS6039 and is responsible for the genotyping pattern at the HD locus of progeny #16.

Progeny microdissection and genotyping using MAT-linked markers revealed that sex in C. amylolentus produces largely progeny that resemble either parental strain. Matings of meiotic progeny to parental strains of opposite mating-type was successful
(Tables 5 and 6) while in others no mating structures were observed. The progeny primarily type as A2B1 (14/65) while 3/65 type as A1B2 at the \textit{MAT} markers. Some of the matings between these two combinations were also successful (namely set 2 F1 #10 with F2 #1, 2, and 5 and set 1 F1 #18, and set 2 F1 #16 with F2 #1, 2, and 5). Sex produced basidia terminating in four spore chains, similar to the matings in the parental strains, CBS6039 and CBS6273, though not as efficient. Additionally, the A1B1 x A2B2 and A1B2 x A2B1 crosses yielded mating products and it is possible that longer incubation (>1 week to 1 month) is necessary for these strains to engage in sexual reproduction. The presence of \textit{MAT} recombinant progeny and their ability in some cases when co-cultured on V8 mating media to produce basidia and spores is strong evidence that sex in \textit{C. amylolentus} is heterothallic and tetrapolar (Figure 27A-B).

Although questions remain, progress in understanding the sexual cycle of another species in the \textit{Filobasidiella} clade has been achieved. We have known for several years that \textit{F. depauperata} (127) is homothallic and may be obligately sexual. In the sister \textit{Kwoniella} clade, we discovered that \textit{C. heveanensis} is sexual and tetrapolar, and in \textit{C. amylolentus} we discovered an extant sexual cycle that is governed by a tetrapolar mating system (102). It is quite possible that the species in the \textit{Filobasidiella} clade are mostly bipolar with the possible exception of \textit{C. amylolentus}, which is tetrapolar and \textit{T. wingfieldii}, which is currently asexual. The species in the more distant clades are also mostly tetrapolar. Uncovering sex in the species surrounding the pathogenic species will shed light into how sex evolved specifically in basidiomycetes and other fungi.
Chapter 5. Definition of the mating-type locus of the tetrapolar basidiomycete *Tremella mesenterica*

5.1 Introduction

*Tremella* species are jelly fungi that belong to the Hymenomycetes lineage, Tremellales order, a large and diverse group of fungi. Specifically, three orders comprise jelly fungi: Tremellales, Auriculariales, and the Dacrymycetales (10). All jelly fungi are heterobasidiomycetes and are named according to their fruiting body structures, the consistency of which are similar to jelly. These species are also brightly colored (orange, yellow, or white). They are mostly saprotrophic, commonly associated with dead and decaying wood in temperate and tropical regions, and some are mycoparasitic (*Tremella mesenterica, Tremella encephala, and Tremella mycophaga*) on plants and insects (177). The presence of specialized cells termed tremelloid haustorial cells establishes mycoparasitism.

Of the estimated 40 *Tremella* species, the dimorphic isolate *T. mesenterica* (common name, Witch’s Butter) has garnered the most attention. *T. mesenterica* is a wood rotting fungus associated primarily with dead trees and rarely with living trees. The primary function of wood rotting fungi is to break down woody plants and trees into their basic elements, enriching the soil for plant growth (17). Although *T. mesenterica* is non-pathogenic, recent studies have investigated the medicinal properties of *T. mesenterica*. Specifically, the glucuronoxylomannan (GXM) polysaccharide present in its fruiting body provides protection against radiation and has antidiabetic and hypocholesteremic properties (28, 29, 75, 95, 96). A second interest in *T. mesenterica* is the potential to
derive fuel ethanol from lignocellulosic biomass (17). Identifying the factors responsible for degrading biomass is central to understanding the biological processes that will eventually advance bio-energy research.

Mycologists have studied mating in *Tremella* for several decades. In 1963, Bandoni described sex in *T. mesenterica* as an incompatibility system (also described in *Tremella globospora, Tremella fuciformis, and Tremella foliacea*) (6, 171). The *A* (P/R) locus is biallelic and encodes the diffusible pheromones, tremerogens *A*-10 and *a*-13, also identified by Bandoni (6, 9, 132). Both pheromones are farnesylated; *A*-10 has a farnesylated and carboxymethylated cysteine and *a*-13 is a farnesylated 13-amino acid lipopeptide (6, 9, 132). In *T. mesenterica* and other fungi, peptide pheromones are essential for specific cell recognition and fusion (70, 71), components of sexual reproduction. The *B* (HD) locus is multiallelic (~146 B factors) and encodes the homeodomain transcription factors. The mating system is therefore tetrapolar, like the majority of basidiomycetes, in which the homeodomain and pheromone/receptor loci are physically unlinked and present on different chromosomes (45). Moreover, the *A* locus controls production and fusion of conjugation tubes while the *B* locus controls clamp cell formation and the production of dikaryotic mycelia during the sexual cycle of *Tremella* (171). This is in dramatic contrast to ascomycetous yeasts and the basidiomycetes species, *Cryptococcus neoformans/C. gattii* and *Coprinellus disseminatus*, which have evolved a bipolar mating system (67, 86).

The life cycle of *T. mesenterica* includes both yeast-like and mycelial phases (6, 9, 171). The yeast phase is restricted to mitotic reproduction via budding and the mycelial
phase is dikaryotic, only occurring when compatible monokaryons differing at the $A$ and $B$ loci are crossed. During the initial stages of sexual reproduction in *T. mesenterica*, budding ceases and conjugation tube formation is induced ($A$ locus) (6, 9, 11). Cell-cell fusion results in the production of dikaryotic hyphae ($B$ locus) and meiosis ensues in the basidium (site of spore production). Upon maturation of the basidia, four elongated sterigmata form that bears spores.

The basidia of jelly fungi vary. Members of the Tremellales have vertically septate basidia, the Auriculariales have transversely septate basidia, and the Dacrymycetales display tuning fork basidia (11, 107). *T. mesenterica* has large cruciate, septate basidia attached to fruiting bodies (or basidiocarps) that associate with wood bark. In a laboratory setting, crosses of a co-mixture of opposite mating partners are conducted on Conjugation Medium Agar (CJM-1) under nutrient starved conditions at low temperatures (12). Importantly, spores are exposed on the surface of the basidiocarp while basidia are usually embedded within the mating medium by five days. We attempted to establish matings in the laboratory, but were unsuccessful. However, we were able to examine conjugation tube formation.

The human fungal pathogens *C. neoformans* and *C. gattii* also belong to the Tremellales order and the genomes of both *Cryptococcus* species have been sequenced (36). *Tremella* is a close relative of *C. neoformans* and comparative genomic studies between these species will advance our current appreciation of basidiomycete biology (141). A recent study compared *MAT* from *T. mesenterica* to *Cryptococcus heveanensis*, a more distant relative to the pathogenic *Cryptococcus* species complex (36, 38, 141).
This study revealed that the structure of MAT in C. heveanensis is similar to T. mesenterica in that the homeodomain locus is multiallelic and the pheromone/receptor locus is at least biallelic (102). This feature is also shared in Ustilago maydis, an extensively examined tetrapolar basidiomycete and plant pathogen (160). The STE11 gene in Filobasidiella depauperata, Tsuchiyaea wingfieldii, and Cryptococcus amylolentus is present outside of the A locus and in pathogenic Cryptococcus sp. STE11 is included. Additionally, the mating structures produced in T. mesenterica and C. heveanensis are similar in that the basidia are globose and have cruciate septa. These advances in comparative genomics among a sampling of species in the Tremellales are pivotal to identifying both shared and distinct characteristics that led to the formation of MAT in this fungal order.

We submitted a fungal whitepaper to the Department of Energy’s Joint Genome Initiative (JGI) program advocating for the sequencing of T. mesenterica in 2007 that was approved as part of the Community Genomics Consortium. We supplied the JGI with well-validated genomic DNA (50 µg) and RNA (for EST sequencing). To sequence the genome of T. mesenterica strain ATCC24925, the JGI applied a whole genome shotgun approach resulting in 7.44X coverage, producing a high-quality draft assembly. The haploid genome of T. mesenterica is approximately 28.6 Mb, somewhat larger than the genome of the pathogenic Cryptococcus species (~20 Mb), and a total of 8,313 genes are predicted. Automated annotations have been performed and we are in the process of completing manual annotations on distinct gene families in collaboration with colleagues around the world for a global analysis of the T. mesenterica genome.
In the current study, we performed BLAST analyses identifying *Cryptococcus* MAT-specific genes in *T. mesenterica* strain ATCC24925 (2). The HD locus spans ~3 kb and is embedded in a ~640 kb contig, while the P/R locus spans ~18 kb and is present on a ~350 kb contig containing many of the genes including hypothetical ones found in the MAT locus of *C. neoformans* and *C. gattii*. We determined through PCR employing five *T. mesenterica* strains, primer walking, sequencing, and sequence comparisons that the core BII locus is limited to *SXI1* and *SXI2* with *RPL22* and a retrotransposon flanking the locus while the a allele is limited to *STE3*, *STE12*, *STE20*, the pheromone gene, tremerogen *a13*, and several hypothetical genes flanked by CND01570 and CNG04540. We successfully elucidated the key sex-determinants responsible for controlling sexual reproduction in this tetrapolar basidiomycetous fungus using comparative genomics. This study has provided further insight into the evolutionary events that shaped MAT in the pathogenic *Cryptococcus* species complex.

5.2 Materials and methods

5.2.1 Strains and media

The species examined in this study are listed in Table 11. Three strains of *T. mesenterica*, CBS8214, CBS8219, and CBS8221, were obtained from the Centraalbureau voor Schimmelcultures (CBS) Fungal Biodiversity Centre in the Netherlands and two strains, ATCC24925 and ATCC42219, were obtained from the American Type Culture Collection (ATCC): The Global Bioresource Center in Manassas, Virginia. All other isolates were obtained from the Heitman lab strain collection. Mating and confrontation assays were performed on Conjugation Medium Agar (CJM-1: 2 g glucose; 2 g soytone;
15 g agar; and 1 L distilled water) for two weeks at 20°C and for one week at 24°C in the dark, respectively. All species were grown and maintained on yeast extract-peptone-dextrose (YPD) medium at 24°C.

5.2.2 DNA extraction

To isolate genomic DNA from *T. mesenterica*, cells were cultured in 50 mL of liquid YPD shaking overnight at 24°C. The pellets were then lyophilized overnight and the CTAB method of fungal DNA isolation was performed (165).

5.2.3 Primer design, PCR, and sequencing

To identify the boundaries of the HD and P/R loci in the sequenced *T. mesenterica* strain ATCC24925, we designed primers to flanking genes using the online Primer 3 program (http://frodo.wi.mit.edu/primer3/). Primers used in this study are listed in Table 12. Standard PCR conditions were used. We used Takara ExTaq for PCR (TAKARA BIO INC., Otsu, Shiga, Japan) and Phire Hot Start DNA Polymerase (Finnzymes, Woburn, Massachusetts, USA) for reactions that were more difficult to amplify. The PCR products were either gel purified using QIAquick Gel Extraction Kit (Qiagen, Valencia, CA) or PCR purified using ExoSAP-IT (GE Healthcare Bio-Sciences Corp., Piscataway, New Jersey, USA). Sequencing reactions were performed with Big Dye chemistry v3.1 (Applied Biosystems, Foster City, California, USA) and analyzed on an Applied Biosystems 3730xl capillary sequencer in the Biological Sciences Sequencing Facility at Duke University.
5.2.4 Bioinformatic analyses

We compared sequences from the successful PCR reactions on the flanking genes of the P/R and HD loci using a matrix comparison (or dot plot) of two sequences (the sequenced strain ATCC24925 versus CBS8214, CBS8219, or CBS8221). To generate each of the dot plots, we employed the Molecular Toolkit’s online nucleic acid dot plots program (http://www.vivo.colostate.edu/molkit/dnadot/). The parameters for the dot plot analyses were a window size of 51 and the mismatch limit of 6. We employed the bioinformatics software Artemis Comparison Tool Release 8 (http://www.sanger.ac.uk/resources/software/) to generate comparison plots across MAT of T. mesenterica and C. neoformans/C. gattii (130). The input file was created using WebACT (http://www.webact.org/WebACT/home) with the tBlastx algorithm (25). GenBank accession numbers are included in Table 13.

5.2.5 Phylogenetic analysis

Phylogenetic analysis was performed on coding sequences using MEGA4 (155). To determine the phylogenetic relationship, the Neighbor-Joining method based on the Kimura 2-parameter model was used (131). Statistical support was calculated from 500 replicates.
Table 11: Description of species in Chapter 5.

<table>
<thead>
<tr>
<th>Strain Number</th>
<th>Source</th>
<th>Mating designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATCC24925</td>
<td>Fallen <em>Alnus rubra</em>, Canada</td>
<td>aBII</td>
</tr>
<tr>
<td>ATCC42219</td>
<td>Fallen <em>Alnus rubra</em>, Canada</td>
<td>unknown</td>
</tr>
<tr>
<td>CBS8214</td>
<td>Dead branch of <em>Alnus rubra</em>, Canada</td>
<td>ab</td>
</tr>
<tr>
<td>CBS8219</td>
<td>Dead branch of <em>Alnus rubra</em>, Canada</td>
<td>AB</td>
</tr>
<tr>
<td>CBS8221</td>
<td>Dead branch of <em>Alnus rubura</em>, Canada</td>
<td>ABI</td>
</tr>
</tbody>
</table>
Table 12: List of primers used in Chapter 5 to identify the boundaries of MAT in *T. mesenterica*

<table>
<thead>
<tr>
<th>Primers</th>
<th>Primer Sequence</th>
<th>Region Amplified</th>
</tr>
</thead>
<tbody>
<tr>
<td>17110-17111</td>
<td>TTCTCSCCGGGGCAAAAGT, AGGATCGAGTTGGGCTGTT</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA7 IA10</td>
<td>TTTCCCATCTCGCCATATCTC, TTCTAATCGGGGGAGAGCA</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA11-IA18</td>
<td>AGAACTCCCTCTCCACTCT, GTGTTGGACCATATTCCG</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA20-IA21</td>
<td>ACTGGTTCTCGCTGTTT, TGTTCAAATCAGCCAAAGCTGG</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA21-IA44</td>
<td>TGCCAATACGCAAGCTGG, CCCAGGTTCATTTTCTTGA</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA21-IA77</td>
<td>TGCCAATACGCAAGCTGG, TACTGATCGCTTTCAGGACT</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA25-17111</td>
<td>CCTTTTACAGCGGCTCAAGTC, AGGATCGAGTTGGGCTGTT</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>SX1F-SX1R</td>
<td>CAACCAACTGCTTCTTCTTCTCT, CGATTTCCAGTTTATGTCAGC</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>SX1F-SX1R</td>
<td>CCAGGTAGCCTTTCAGACCA, ATAGGGCGACCTCGAGAACCT</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA55-17111</td>
<td>GAAATGAAGCTTGGGAGCACG, AGGATCGAGTTGGGCTGTT</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA10-IA55</td>
<td>TTTCTAATCGGGGGAGACGA, GAAATGAAGCTTGGGAGCAG</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA55-SX1R</td>
<td>GAAATGAAGCTTGGGAGCACG, ATAGGGCGACCTCGAGAACCT</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>SX1F-SX1R</td>
<td>CCAGGTAGCCTTTCAGACCA, ATAGGGCGACCTCGAGAACCT</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA62-IA63</td>
<td>CCTCGGTCGTCATACAAA, AGACCGTGTAAGTGAGGTA</td>
<td>HD allele - CBS 8214</td>
</tr>
<tr>
<td>IA63-IA59</td>
<td>TGCAACTCTCCCCATCCACTGT, ATGGTCGCGTTTTTCTTTCCC</td>
<td>P/R allele - CBS 8219</td>
</tr>
<tr>
<td>IA58-IA61</td>
<td>TGCAACTCTCCCCATCCACTGT, ATGGTCGCGTTTTTCTTTCCC</td>
<td>P/R allele - CBS 8219</td>
</tr>
<tr>
<td>IA64-IA66</td>
<td>GTCGCCGTGATGTCGCGAAT, GACCCCGAAGAGCATCAAA</td>
<td>P/R allele - CBS 8219</td>
</tr>
<tr>
<td>IA38-17113</td>
<td>GGGTTCTTCCTCGCTGATT, GGTGTCCTCAGCTCGCTT</td>
<td>P/R allele - CBS 8219</td>
</tr>
<tr>
<td>IA50-2082</td>
<td>TGAGAGCTCTTTCGCGTGA, TTTCTGCTGCTCGCTCTCT</td>
<td>P/R allele - CBS 8219</td>
</tr>
<tr>
<td>IA52-IA53</td>
<td>CCTCTACAGGATGACCGAGT, CAACATCGGTGACACCAAGG</td>
<td>P/R allele - CBS 8219</td>
</tr>
<tr>
<td>IA66-IA71</td>
<td>GACCCCGAAGAGCATCAAA, GACCCCGAAGAGCATCAAA</td>
<td>P/R allele - CBS 8219</td>
</tr>
<tr>
<td>IA58-IA59</td>
<td>TGCAACTCTCCCCATCCACTGT, ATGGTCGCGTTTTTCTTTCCC</td>
<td>P/R allele - CBS 8221</td>
</tr>
<tr>
<td>IA58-IA61</td>
<td>TGCAACTCTCCCCATCCACTGT, ATGGTCGCGTTTTTCTTTCCC</td>
<td>P/R allele - CBS 8221</td>
</tr>
<tr>
<td>IA64-IA66</td>
<td>GTCGCCGTGATGTCGCGAAT, GACCCCGAAGAGCATCAAA</td>
<td>P/R allele - CBS 8221</td>
</tr>
<tr>
<td>IA38-17113</td>
<td>GGGTTCTTCCTCGCTGATT, GGTGTCCTCAGCTCGCTT</td>
<td>P/R allele - CBS 8271</td>
</tr>
<tr>
<td>IA46-IA69</td>
<td>GACCCCGAAGAGCATCAAA, AGCATGTCGCTGACTATC</td>
<td>P/R allele - CBS 8221</td>
</tr>
<tr>
<td>IA46-IA71</td>
<td>GACCCCGAAGAGCATCAAA, AGCATGTCGCTGACTATC</td>
<td>P/R allele - CBS 8221</td>
</tr>
<tr>
<td>IA46-IA67</td>
<td>GACCCCGAAGAGCATCAAA, CGAAAACGACAGCAGACTT</td>
<td>P/R allele - CBS 8221</td>
</tr>
<tr>
<td>JH021555</td>
<td>AACTTGTTCTGCTGATGGAT</td>
<td>GEF1 Forward</td>
</tr>
<tr>
<td>JH02156</td>
<td>CAGAAAATGGACTGGGTAAGC</td>
<td>GEF1 Reverse</td>
</tr>
<tr>
<td>JH02149</td>
<td>TGGTGACTGACTTCTGCTGGACG</td>
<td>RPO41 Forward</td>
</tr>
<tr>
<td>JH02150</td>
<td>TCACCTTGGCGAAGTTT</td>
<td>RPO41 Reverse</td>
</tr>
</tbody>
</table>
5.3 Results

5.3.1 Characterization of MAT in *Tremella mesenterica*

In 2007, we submitted a fungal whitepaper to the DOE’s JGI program advocating for the sequencing of the wood-rotting fungus *Tremella mesenterica*, and the proposal was approved. We also established a consortium of researchers to assist with the annotation of the completed sequenced genome of *T. mesenterica*. With the recent release of the haploid 28.6 Mb *T. mesenterica* genome, we performed reciprocal BLAST analysis identifying putative orthologs of >20 MAT-linked genes found in *C. neoformans* and *C. gattii*. The P/R and HD alleles of the tetrapolar basidiomycete *T. mesenterica* strain ATCC24925 (Figure 28) have undergone extensive expansion, yet the sex-specific regions containing the P/R and HD genes are restricted to ~18 kb and 3 kb, respectively.

![Tremella mesenterica environmental isolate. Rehydrated sample of *T. mesenterica* on a tree branch.](image)
A majority of the MAT-linked genes identified are also present but are randomly located throughout the 640 kb contig containing the P/R locus (Figure 29A). Most importantly, the pheromone gene tremerogen a-13, STE3, STE12, and STE20 are linked and within the region delimited as mating-type specific. The five most recently acquired genes are present, and RPO41 and BSP2 are linked at one distant end of the locus, while LPD1, CID1, and GEF1 are linked and located on the 5’ end of the P/R locus. Although, the C. neoformans 3’ MAT flanking gene, NOG2, is located at the very end of the contig containing the P/R locus containing contig in T. mesenterica, it remains closely associated with the genes that define MAT in the pathogenic Cryptococcus species.

Several hypothetical genes are also found throughout the contig containing the P/R locus and correspond to chromosome 4, the location of MAT in C. neoformans. In particular, many of the hypothetical genes in T. mesenterica (and the other fungal species integral to the current study) correspond to the telomeric region of JEC21 and intrachromosomal rearrangements may have also occurred during the evolution of MAT. We have also identified genes (or sequences) that only have homologs in fungi unrelated to Cryptococcus. For example, several U. maydis hypothetical gene (UM02602 and UM04613) homologs are located in the P/R locus, as well as a hypothetical gene most closely related to a gene from Laccaria proxima (Figure 29A). The presence of such genes has also been observed in C. heveanensis (102). We hypothesize that these genes have been lost from MAT and relocated to the telomeric ends of chromosome 4 in C. neoformans, as previously described in Metin et al. (102).
The arrangement of the genes in the HD locus revealed that the \textit{SXI1} and \textit{SXI2} orthologs are present, linked, and divergently oriented, similar to \textit{bE} and \textit{bW} in \textit{U. maydis} (Figure 29B). Two additional HD linked genes, \textit{RPL22} and \textit{SPO14}, are also present. Interestingly the \textit{SPO14} gene is not adjacent to the homeodomain transcription factor genes. Instead it is >120 kb upstream of the key sex determining genes. Similar to the P/R locus, hypothetical genes are numerous and retrotransposons, retroelements, and transposases have also been identified and linked to the HD genes, including candidate genes from \textit{Laccaria} and \textit{Coprinopsis} (Figure 29B). The genomes of \textit{C. neoformans} and \textit{C. gattii} are rich in transposons, and many cluster near centromeric regions (97). The \textit{MAT} locus of the pathogenic species contains many transposons and their presence suppresses recombination within \textit{MAT}, a key hallmark of sex-determining regions in animals, plants, and fungi (43, 46, 115).

![Contig containing the \textit{a} (biallelic) P/R locus ~640 kb](image)

![Contig containing the \textit{Bll} (multiallelic) HD locus ~240 kb](image)

Figure 29: Sequence assemblies spanning \textit{MAT} in the sequenced strain \textit{T. mesenterica} ATCC24925. We performed BLAST analysis and the contig containing the embedded biallelic \textit{a} (P/R) locus spans ~640 kb of sequence (A) while the contig containing the embedded
multiallelic B (HD) locus spans ~240 kb (B). Black arrows indicate genes with homologs in *C. neoformans* and yellow indicates hypothetical genes. *MAT*-linked genes are underlined in red and sequenced genes outside of the *MAT* region are underlined in green. Scale bar = 10 kb (A) and 20 kb (B).

Several interesting features of the *MAT* loci were identified. A common theme previously described in *F. depauperata*, *C. heveanensis*, *T. wingfieldii*, *C. amylolentus* and more recently in *T. mesenterica* is the absence of the *STE11* gene in *MAT*. *STE11* is located outside of the *STE* gene cluster in *T. mesenterica* but is present in the larger ~640 kb contig containing the P/R locus. In the other species, *STE11* is unlinked to *MAT* and found elsewhere in the genome (Figure 29A). Additionally, the *C. neoformans* *SXII* and *SXI2* flanking genes *FAO1, UAP1*, and/or *FCY1* are not linked to the HD scaffold and are likely present elsewhere in the genome of *T. mesenterica*. Two copies of the *NCP1* gene exist in *T. mesenterica* (Figure 29A). This duplication has also been observed in *C. heveanensis, T. wingfieldii*, and *C. amylolentus*, and we speculate that this organization represents the ancestral arrangement of these genes in which two copies were present. In *C. neoformans*, however, one of the *NCP1* genes has been lost or translocated to another region in the genome (102). The current analysis of the *MAT* loci in *T. mesenterica* has revealed key evolutionary steps, including the presence of transposons, unlinked HD and P/R loci, and several chromosomal translocation events that punctuated the evolution of *MAT* in the pathogenic *Cryptococcus* species.
5.3.2 Defining the boundaries of $MAT$ in $T.~mesenterica$

In strains of opposite mating-types, the only difference in the genome between the two lies within $MAT$. Therefore, the boundaries (or flanking regions) surrounding $MAT$ are typically syntenic with highly conserved sequences whereas the sequence of the $MAT$-specific genes are diverged and often rearranged. In $T.~mesenterica$, the P/R and HD allele specific genes span >28 kb. We hypothesized that the core regions defining sex in this species would be restricted to $STE3$, $STE12$, $STE20$, and tremerogen a13 in the P/R locus and $SXI1$ and $SXI2$ in the HD locus. To determine the boundaries, we PCR amplified across the loci or flanking regions in all five strains, including the sequenced strain ATCC24925 as a control. Because the expected size of the fragments for the HD (~3 kb) and P/R (~18 kb) loci are >2 kb, we completed several rounds of primer walking to identify or delimit the boundaries of each locus.

We have successfully identified the key genes encoding the tetrapolar mating system of $T.~mesenterica$. The boundaries of the P/R locus are the hypothetical genes CNG04540 and CND01570 (Figures 32 and 33). We identified the boundaries by performing a dot plot comparison of CBS8219 and CBS8221 to the sequenced strain ATCC24925. The percent identity for CND01570 (~900 bp analyzed) in CBS8219 compared to ATCC24925 is 98% and CNG04540 (~1 kb analyzed) in CBS8219 compared to ATCC24925 is 97%. Additionally, CND01570 (~900 bp analyzed) in CBS8221 compared to ATCC24925 is 98% and CNG04540 (~1 kb analyzed) in CBS8221 compared to ATCC24925 is 96%. The high percent identities suggest that these genes lie outside the sex-divergent region of $MAT$. The $STE$ genes (except $STE11$),
tremerogen a-13, CNB00600, and CNB00610 define the P/R allele. We amplified the 5’ and 3’ flanking genes in the P/R locus of CBS8219 and CBS8221 (both strains have the A allele) but were unable to amplify the individual genes that define the locus (Figures 32 and 33). Unfortunately, the sequence divergence is very high and primers designed for the sequenced strain did not yield products in the two A strains. The regions that flank the P/R allele share greater than 96% identity between CBS8219 (Figure 32), CBS8221 (Figure 33) and the sequenced strain ATCC24925.

Figure 30: Alignment of an unlinked MAT gene, RPO41. Sequence alignment of RPO41 in T. mesenterica strains CBS8214, CBS8219, CBS8221, ATCC24925, and ATCC42219. The sequences share ~96% identity.

A caveat of our analysis is that there could be diverged regions (or genes) beyond the boundaries we identified as MAT-specific in the HD and P/R locus of ATCC24925.
To address this caveat, we amplified two unlinked P/R genes, RPO41 and GEF1, from the *T. mesenterica* strains. Only partial fragments were sequenced and the alignment suggests that the two genes share roughly 96% identity between each strain and the sequenced isolate ATCC24925 (Figures 30 and 31). Based on this analysis, we propose that MAT is restricted to the 3 and 18 kb regions defined. However, given the structure of the MAT locus of the pathogenic *Cryptococcus* species the conserved genes that we interpret as flanking MAT could reside within, and additional divergent regions may remain currently unidentified beyond these borders.

![Figure 31: Alignment of an unlinked MAT gene, GEF1.](image)

Sequence alignment of GEF1 in *T. mesenterica* strains CBS8214, CBS8219, CBS8221, ATCC24925, and ATCC42219. The sequences share ~96% identity.
Figure 32: Comparison of *T. mesenterica* strain ATCC24925 (*a* allele) to CBS8219 (*A* allele) limits the P/R locus to ~18 kb. A percent identity plot spanning the flanking genes in the P/R locus identified the presumptive boundaries of *MAT*. The red arrow indicates the genes (CND01570 and CNG04540) that flank the 5’ and 3’ ends of the locus respectively. Black arrows indicate genes with homologs in *C. neoformans* and yellow indicates hypothetical genes. Window size = 35 and mismatch limit = 4. Scale bar = 5 kb.

Conversely, the HD locus includes only *SXII* and *SXI2* and is flanked by *RPL22* and a retrotransposon (Figure 34). Amplification of the *B* allele in *T. mesenterica* strain CBS8219 was successful. In comparison to the *A* allele, this region is much smaller. Therefore, amplifying an ~6 kb region for the HD locus, which spans *SXII*, *SXI2*, and the two flanking genes was feasible and the percent identity for this region is 85% between CBS8219 compared to the sequenced strain ATCC24925. A sequence alignment of
RPL22 in the sequenced strain and CBS8219 displays 99% nucleotide identity while the 3’ flanking retrotransposon shares 100% identity between the sequenced strain and CBS8219. In summary, we determined from both PCR and dot plot comparison that the HD locus spans ~3 kb and the P/R locus spans ~18 kb. This analysis and similar studies will provide insight into how mating-type loci evolved in the pathogenic Cryptococcus species by focusing on closely related species. Extending these studies to more distantly related isolates may reveal further insights into the transition of the ancestral form of MAT in Cryptococcus from tetrapolar to bipolar mating systems.
5.3.3 Synteny between *T. mesenterica* and *C. neoformans*

We successfully defined the *MAT* boundaries in the tetrapolar basidiomycetous fungus *T. mesenterica* employing a PCR based approach. *MAT* is restricted to ~3 kb in the HD locus and ~18 kb in the P/R locus. We recognize that there could be other *MAT*-specific regions beyond the borders we identified. For example, *MAT* could have two parts (or subloci) with conserved genes. To further characterize *MAT*, we conducted a
synteny analysis using Artemis (tBlastn algorithm) to compare the MAT loci of *T. mesenterica* to MAT in the *C. neoformans* serotype D strain JEC20. In the synteny analysis of the sibling species (see Chapter 3), we used Blastn for the analysis. Because *T. mesenterica* is a more distant relative of the pathogenic *Cryptococcus* species, tBlastn was employed. More importantly, sequence divergence between *C. neoformans* and *T. mesenterica* is high across the regions encoding MAT (Figure 35). Similar to the sibling species *T. wingfieldii* and *C. amylolentus*, MAT in *T. mesenterica* has undergone extensive rearrangement, gene inversions, multiple transposon insertions, and chromosomal translocation events. The inclusion of transposons in MAT has been described in the pathogenic *Cryptococcus* species and has now been identified in the *T. mesenterica* HD locus (97). Many of these events are also evident in the synteny analysis for the *C. neoformans* species although the degree of divergence is much higher in *T. wingfieldii*, *C. amylolentus*, and *T. mesenterica*.

### 5.3.4 Phylogenetic analysis of *T. mesenterica* genes

We conducted phylogenetic analysis on several unlinked and MAT-linked genes (*CID1, ETF1, GEF1, LPD1, STE3, SXI1, SXI2, and STE20*, see Chapter 3 for additional gene trees). Four representative genes (*STE20, ETF1, STE3, and LPD1*) highlight different phylogenetic patterns: mating-type specific (*STE3, ETF1, and STE20*), species-specific (*GEF1, CID1, and LPD1*), and *SXI1 and SXI2* are very diverged from *C. neoformans* and *C. gattii* (Figure 36). In this analysis, we included *C. neoformans* var. *neoformans*, *C. neoformans* var. *grubii*, and *C. gattii* representatives from the pathogenic species cluster. We also included species residing outside of the complex, *C. heveanensis*.
and the sibling species *C. amylolentus* and *T. wingfieldii*. Interestingly, *C. heveanensis* and *T. mesenterica* displayed similar phylogenetic patterns, which were also observed in *T. wingfieldii* and *C. amylolentus*, while the pathogenic species differed in their phylogenetic patterns.

**Figure 34**: Comparison of *T. mesenterica* strain ATCC24925 (BII allele) and CBS8219 (B allele) restricts the HD locus to ~3 kb.
A percent identity plot spanning the flanking genes in the HD locus identified the boundaries of MAT. Black arrows indicate flanking genes. The parameters for dot plot generation: window size = 51 and mismatch limit = 6.

Figure 35: Synteny between MAT loci of *T. mesenterica* and *C. neoformans* var. *neoformans*.

Black arrows indicate genes with homologs in *C. neoformans* and yellow indicates hypothetical genes. The grey arrows are genes that flank the MAT locus of *C. neoformans*. Synteny is depicted in blue or red. Red indicates that gene order is conserved while blue depicts gene inversions or stretches of sequences have been inverted between the two species. The top bars represent *T. mesenterica* and the bottom bar represents *C. neoformans*.

### 5.3.5 Exploration of sex in *T. mesenterica*

We performed mating assays in pairwise combinations (CBS8214, CBS8219, CBS8221, and ATCC42219 X ATCC24925) employing five *T. mesenterica* strains on Conjugation Medium Agar (CJM-1). The matings were incubated at 20°C for two weeks in the dark. The matings were unsuccessful and only CBS8219 X ATCC24925 produced a small number of conjugation tubes. In the other crosses, conjugation tubes and/or clamp
cells were not produced. Because the matings were fairly unsuccessful, we conducted confrontation assays with *T. mesenterica* strains and confronted CBS8219 and CBS8221 with ATCC24925. These confrontation assays produced what appeared to be conjugation tubes and the results are shown in Figures 37 and 38. Importantly, confrontations of either of the strains alone did not produce conjugation tubes. In *T. mesenterica*, the *A* locus controls the production and fusion of conjugation tubes while the *B* locus establishes clamp cell and dikaryon formation. We speculate that the confrontation assays we performed have shown that conjugation tube formation is induced by the *A* locus and clamp cell formation is not initiated. This is likely due to an *Aon, Boff* mating system. The *a* cell produces pheromone that *A* responds to and we only see conjugation tube formation in the *A* allele. Although the mating designation of the strains indicates that the *B* alleles are compatible, the designation may be incorrect. Alternatively these strains may have lost fertility with storage/passage, or the conditions may not be fully conducive for mating. Additional studies are required to recapitulate the sexual cycle of *T. mesenterica* originally described by Bandoni.
Figure 36: Phylogenetic patterns of T. mesenterica genes.

The phylogenetic relationship of T. mesenterica to the pathogenic Cryptococcus species and neighboring taxa is highlighted and four representative genes STE20, ETF1, STE3, and LPD1 are shown. STE20 and ETF1 exhibit mating-type specific phylogeny, STE3 displays an a mating-type specific phylogenetic pattern, and LPD1 is species-specific. The trees were constructed using the Neighbor-Joining method and MEGA4. Bootstrap values on tree branches were calculated from 500 replicates. (α) indicates strains with the MATα locus, and (a) indicates strains with the MATα locus.
Figure 37: Confrontation assay in CBS8219 and ATCC24925. Strains were confronted on CJM-1 agar at 24°C for 2 weeks in the dark. The inset reveals conjugation tubes that resulted from a CBS8219 X ATCC24925 cross.
Figure 38: Confrontation assay in CBS8221 and ATCC24925. Strains were confronted on CJM-1 agar at 24°C for 2 weeks in the dark. Conjugation tubes are produced only when the $A$ allele is present.
Table 13: *T. mesenterica* Genbank accession numbers.

<table>
<thead>
<tr>
<th>Species</th>
<th>HD Locus</th>
<th>P/R locus</th>
<th>RPO41</th>
<th>GEF1</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Tremella mesenterica</em> (ATCC24925)</td>
<td>HM440937</td>
<td>HM440938</td>
<td>HM440942</td>
<td>HM440947</td>
</tr>
<tr>
<td><em>Tremella mesenterica</em> (CBS8214)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tremella mesenterica</em> (CBS8219)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tremella mesenterica</em> (CBS8221)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Tremella mesenterica</em> (ATCC42219)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

5.4 Discussion

Bandoni first described *T. mesenterica* in the 1960s. Later Bandoni also discovered the sexual cycle and isolated the pheromone, tremerogen a-13 (6, 9). In recent years, interest in investigating the sexual cycle of *T. mesenterica* has waned. Significant advances have been made in employing *T. mesenterica* for medicinal purposes due to its anticancer, antiviral, and antidiabetic properties (28, 29, 75, 95, 96). The haploid genome of *T. mesenterica* was recently sequenced and a genome wide analysis is under way. Although studies on the healing properties of *T. mesenterica* are important, we focused our attention on the evolution of sex in the pathogenic *Cryptococcus* species using *Tremella* as an outgroup. We also were successful in generating conjugation tubes upon co-culture of potential compatible strains of *T. mesenterica*. Analysis of MAT from closely and distantly related species has revealed shared aspects of MAT locus evolution in *Cryptococcus* and more generally, fungal species.

The mating system of the tetrapolar basidiomycete *T. mesenterica* consists of an unlinked biallelic locus that encodes the pheromone and pheromone receptor and a multiallelic B locus that encodes the homeodomain transcription factors. BLAST was conducted to identify MAT-specific genes in both loci. The scaffold containing the P/R locus spans ~350 kb while the scaffold containing the HD locus spans ~640 kb in the sequenced strain ATCC24925. The scaffolds containing MAT genes representing the HD and P/R loci are large and include many hypothetical genes also present in the genomes of *C. neoformans* and *C. gattii*. The inclusion of several additional strains of *T. mesenterica* (CBS8214, CBS8219, CBS8221, and ATCC42219) assisted in determining
which key genes in these larger regions define MAT. MAT in most basidiomycetes is relatively small, as in *Rhodosporidium toruloides*, *U. maydis* and *C. disseminatus*, but in the pathogenic *Cryptococcus* species, it is unusually large spanning >100 kb (31). We determined from sequence comparisons of the HD and P/R loci in strains of opposite mating-type that the MAT alleles in *T. mesenterica* span ~28 kb total. Specifically, the HD locus spans ~3 kb and is restricted to **SXI1** and **SXI2**. **RPL22** and a retrotransposon flank this region (>99% identity). We also determined that **STE3**, **STE12**, **STE20**, tremerogen a-13, **CNB0600**, and **CNB0610** define the ~18 kb P/R locus and **CND01570** and **CNG04540** flank the P/R locus. We typed two genes that lie on the P/R scaffold, but far from MAT >30 kb from the MAT locus, **RPO41** and **GEF1**, and the alignment revealed that the genes share a high level of identity in the *T. mesenterica* strains (>96% identity). It is possible that other regions specific to MAT could be present outside of the regions we have defined as the HD and P/R locus. Additional analyses that focus on examining adjacent genes will address whether MAT is larger in *T. mesenterica*. We speculate that the larger scaffolds that comprise the HD and P/R loci include numerous flanking genes that are not mating-type specific in *T. mesenterica* but were recruited into MAT in the pathogenic *Cryptococcus* species cluster.

Similar to *C. neoformans*, **NOG2** flanks the P/R locus of *T. mesenterica* while the genes that flank the HD locus in *C. neoformans* are missing in *T. mesenterica*. We speculate that these genes (**FCY1**, **FAO1**, and **UAPI**) are found elsewhere in the genome, and were translocated to the regions near the HD locus in the pathogenic species lineage. Moreover, in *T. wingfieldii*, *C. amylolentus*, and *F. depauperata*, **FAO1** is absent from the
HD locus. The *STE11* gene is also missing from the *MAT* regions in the aforementioned species and in *T. mesenterica*, this gene is present in the larger P/R contig scaffold, but is absent from the *STE3, STE12, and STE20* gene cluster. In the *MAT* locus of *C. neoformans* and *C. gattii, STE11* is present. These differences in the organization of *MAT* may represent the ancestral form in which several of the flanking genes and *STE11* are not linked to *MAT*. Two HD genes are present in the HD locus of *T. mesenterica* (HD1 and HD2) while in *C. neoformans, SXI1* (HD1) is associated with the α allele while *SXI2* (HD2) is associated with the a allele. Additionally, the presence of both homeodomain transcription factors in *T. mesenterica* and its distant relationship to the pathogenic *Cryptococcus* species as well as their presence in *C. heveanensis*, suggest that the arrangement of *MAT* in *T. mesenterica* is similar to the ancestral form of *MAT*. The inclusion of these genes did not occur until the pathogenic species complex formed.

Characterizing *MAT* in the sibling species and *C. heveanensis* and *F. depauperata* revealed a number of hypothetical genes related to those present in the *C. neoformans* genome, a majority of which are located on chromosome 4. In *T. mesenterica*, a majority of the hypothetical genes identified by BLAST correspond to those from *C. neoformans* chromosome 4. Interestingly, chromosome 4 also harbors the *MAT* locus. In addition, several orthologs from *Laccaria, Ustilago,* and *Coprinopsis* are randomly distributed among the hypothetical genes in the HD and P/R loci. We speculate that some of these genes have been lost from the regions flanking *MAT* and are now present in other areas of the *Cryptococcus* genome while a minority corresponds to the telomeric end of *C. neoformans* chromosome 4. Taken together, these results suggest that rearrangements and
intrachromosomal translocations have shaped the evolution of the extant MAT structure in the pathogenic Cryptococcus species complex (41, 45, 46, 86).

The tetrapolar mating system of the basidiomycete Laccaria bicolor shares features that have been described in the evolution of MAT in C. neoformans and sex chromosome evolution in animals and plants (43, 44, 47, 110). In L. bicolor, the A (HD) locus encodes 1 pair of HD genes with 45 mating-types (100, 110). The B (P/R) locus is more complex and consists of three linked subloci ordered in a tandem duplication of P/R genes. Similar to the species we have examined, the B locus has been subjected to gene duplications, translocations, and transposon insertions. These features suggest that the B mating-type evolved faster than the A locus. We have identified examples of these key features in MAT evolution in T. wingfieldii, C. amylolentus, C. heveanensis (102), and T. mesenterica. Specifically, NCP1 has been duplicated and is located within the contig that defines the P/R locus. Further, the presence of hypothetical genes from different chromosomes of the pathogenic Cryptococcus species suggests chromosomal translocation events have transpired, and transposons are found in the HD locus of T. mesenterica. Moreover, recombination is also suppressed and gene conversion events occur within the regions that define MAT in L. bicolor and the Cryptococcus species.

The MAT features discussed in detail in this chapter are similar to those identified in T. wingfieldii and C. amylolentus, including the evidence for chromosomal translocation events, the presence of two linked and divergently oriented HD genes, and two unlinked loci that encode HD and P/R genes (60). With the addition of the tetrapolar basidiomycete and outgroup T. mesenterica, the evolution of MAT (see Chapter 3, Figure
10) analysis has been extended to both closely and distantly related species of the pathogenic Cryptococcus species cluster. Moreover, the evolution of sex determinants in not only fungi, but also plants and animals, involves expansion of the sex-determining region, gene rearrangements and inversions, chromosomal translocations, and suppression of recombination in the regions that will eventually define MAT or a sex chromosome (43, 44, 46, 115). From this analysis, it is apparent that a series of evolutionary events are implicated in MAT formation. These events include: the tetrapolar to tripolar to bipolar transition, a multiallelic HD that evolved into a biallelic locus, and paired HD genes that underwent a loss event resulting in only one HD associated with each allele. Therefore, T. mesenterica serves as an intermediate in the formation of the bipolar mating system in C. neoformans and C. gattii, with a tetrapolar MAT configuration involving two unlinked loci, and thus an evolutionary window descended from a lost common ancestor prior to the MAT fusion event that occurred in the lineage leading to the pathogenic Cryptococcus species cluster.
Chapter 6. Conclusions and Future Directions

Sex is found in a myriad of microorganisms, plants, and animals and is both beneficial and costly. In microbial pathogens, sex generates diversity by reshuffling the genome for favorable and more virulent gene combinations, and deleterious gene combinations are readily expunged from the genome. But sex can also reassert favorable allele configurations and confer a cost. Sexual reproduction in fungi is governed by either a bipolar or a tetrapolar mating system; a bipolar system promotes inbreeding while the tetrapolar system involves numerous mating-types and thus promotes outbreeding. Sexual reproduction in some pathogenic fungi had been unknown (15, 18), but several recent studies have reported the discovery of heterothallic and/or homothallic sex in *Candida albicans*, *Candida dubliniensis*, and *Aspergillus fumigatus* (1, 19, 23, 58). The human pathogen and basidiomycete *Cryptococcus neoformans* has evolved a bipolar mating system, which governs and controls sex. Its *MAT* locus is unusually large, spanning >100 kb and encoding >20 genes (41, 86). The bipolar mating system in *C. neoformans* and *C. gattii* evolved from an ancestrally unlinked *MAT*, which has experienced fusion, extensive rearrangement, and gene conversion events (41).

Most basidiomycetes have a tetrapolar mating system and thus, the *C. neoformans* *MAT* locus is unique in its bipolar organization. In this dissertation, we investigated the evolution of the unusual *MAT* structure in the pathogenic *Cryptococcus* species using phylogenetics and comparative genomics approaches. Prior contributions of our lab to this field focused on the pathogenic *Cryptococcus* complex (46, 86), while the current study extended the analysis to both closely and distantly related basidiomycetes in the
Tremellales lineage. We also report the discovery and characterization of an extant sexual cycle in *C. amylolentus*.

**6.1 *T. wingfieldii* and *C. amylolentus* are sibling species and the closest relatives to the pathogenic *Cryptococcus* species cluster**

Previous phylogenetic studies of the pathogenic *Cryptococcus* species complex were limited in the number and diversity of species analyzed, and in many cases only the rDNA and ITS markers were examined (54, 82, 137, 153). In Chapter 2, we resolved the phylogeny surrounding the pathogenic *Cryptococcus* species using a robust analysis of six fungal specific markers in a multi-locus sequencing (MLS) approach (38, 66). This study revealed that *C. amylolentus* and *T. wingfieldii*, two nonpathogenic sibling species, are the most closely related taxa to the pathogens *C. neoformans* and *C. gattii*, and group with *F. depauperata* to form the *Cryptococcus sensu stricto* group. Additionally, the five saprobic yeast species in the *Kwoniella* clade (*Bullera dendrophila*, *Cryptococcus heveanensis*, *Kwoniella mangroviensis*, *Cryptococcus bestiolae*, and *Cryptococcus dejecticola*) appear to be part of a more distantly related *sensu lato* group. The majority of the isolates in the *Filobasidiella* and *Kwoniella* clades are saprobic yeasts that are commonly associated with dead and decaying insect frass in the environment. Moreover, this phylogenetic study revealed that members of the Tremellales lineage likely evolved from a common insect-associated (or saprobic) ancestor that inhabited similar ecological niches.

None of the species (*T. wingfieldii*, *C. amylolentus*, *F. depauperata*, *K. mangroviensis*, *C. bestiolae*, *C. dejecticola*, *C. heveanensis*, *B. dendrophila*, *Tremella*...
mesenterica, and Cryptococcus humicola) displayed any of the virulence attributes commonly associated with C. neoformans and C. gattii. Specifically, we tested melanin and capsule production as well as the growth at high temperature, and we examined the virulence potential of each isolate in the wax moth model of infection. Surprisingly, the saprobic Cryptococcus species in the Filobasidiella clade and T. mesenterica were avirulent while C. humicola and members of the Kwoniella clade displayed an intermediate virulence. Taken together, these studies provide insight into the genotypic and phenotypic profiles of the successful pathogenic Cryptococcus species. These species likely evolved from a common saprobic ancestor, later emerging as a pathogen. An example of another human fungal pathogen with a predicted saprobic ancestor is Candida (150, 151). Finally, if newly identified isolates of Cryptococcus species become available, these species should be included in future phylogenomic studies with the goal of generating a broad and robust analysis of the species that are phylogenetically related to the pathogenic Cryptococcus species.

6.2 The MAT loci of T. wingfieldii and C. amylolentus are physically unlinked and lie on different chromosomes

Having successfully identified taxa most closely related to C. neoformans and C. gattii, we investigated the evolution of the mating-type locus in the pathogenic Cryptococcus species by characterizing the MAT locus in the sibling species T. wingfieldii and C. amylolentus. Chromoblot analysis revealed that the HD and P/R loci are physically unlinked and located on different chromosomes. This suggests that the ancestral form of MAT is defined by two unlinked loci, HD (B) and P/R (A), and further
supports the amended *MAT* evolution model (see Chapter 4, Figure 10). An interesting question worth addressing in future studies is whether a link exists between the unique organization of *MAT* in *Cryptococcus* and its emergence as a pathogen.

Previous identification and characterization of the *MAT* loci in the tetrapolar basidiomycete *L. bicolor* revealed that sex chromosomes in plants, animals, and the *MAT* locus in fungi underwent similar genomic events that shaped the formation of sex-determining regions in these species (43-45, 110, 115). Based on our studies of *C. amylolentus* and *T. wingfieldii*, we conclude that *MAT* in the sibling species contains (1) two linked and likely divergently transcribed homeodomain genes as in other tetrapolar basidiomycetes and (2) a chromosomal translocation event that likely occurred between chromosomes 4 and 5 of *C. neoformans*.

In addition to the *T. wingfieldii* and *C. amylolentus* fosmid libraries we used to clone *MAT*, we also generated libraries for members of the *Kwoniella* clade: *K. mangroviensis* (CBS8507), *B. dendrophila* (CBS6459, CBS6460, and CBS6074), and *Cryptococcus flavus* (CBS14551). Future characterization of the *MAT* locus from these species will provide insight into the evolution of this structure within the pathogenic *Cryptococcus* species cluster and more broadly in the Tremellales. A fungal whitepaper was recently submitted to the Joint Genome Initiative advocating for whole genome sequencing of *T. wingfieldii*, *C. amylolentus*, and *F. depauwerata*. Once complete, data from these genome-sequencing projects will provide the opportunity to identify *MAT*-linked genes in the genomes of the fungi in the monophyletic *Filobasidiella* clade. Whole genome sequencing will be helpful in analyzing difficult regions of the genome,
including repetitive sequences that surround the pheromone genes and will assist in
closing gaps in the $MAT$ assemblies for $C.\ amylolentus$.

We were unable to determine whether the $T.\ wingfieldii$ mating system is bipolar
or tetrapolar because only one isolate is currently available. Because the HD and P/R loci
are unlinked, the organization of the mating system appears to be tetrapolar.
Identification of additional strains of $T.\ wingfieldii$ will further advance our
understanding of $MAT$ and will allow us to explore sex in this species. Sampling diverse
environmental niches may assist in the recovery of more $T.\ wingfieldii$ strains.

We successfully identified the boundaries of $MAT$ in $C.\ amylolentus$. The
homeodomain genes $SXI1$ and $SXI$, are polymorphic primarily in the 5’ ends of both
genes, share ~92% identity, and define the HD locus. The $SXI1$ and $SXI2$ dimorphic
regions of $C.\ amylolentus$ are similar to the HD genes of basidiomycetes. In particular,
the N-terminal regions are more variable and define the mating-type alleles. The
pheromone receptor gene, $STE3$, is also linked to $MAT$ in $C.\ amylolentus$. Southern blot
analysis using a CBS6039 $STE3$-specific PCR product as a probe revealed that $STE3$ is
highly divergent in the two $C.\ amylolentus$ strains. Therefore, the HD and P/R subloci
define mating-type and are biallelic. Upon successful identification of additional $C.\ amylolentus$
isolates, future studies will investigate whether the loci are multiallelic or
remain biallelic.

While we were successful in determining which genes in the HD and P/R loci
define $MAT$ in $C.\ amylolentus$, information on their roles during sexual reproduction is
still lacking. To determine the function of the $MAT$-linked HD genes, $STE3$, and the
pheromone genes, it will be useful to establish a transformation system in \textit{C. amylolentus}. Gene disruptions of \textit{SXI1, SXI2, STE3}, and the pheromone genes will be instrumental in determining the function of each gene during sex in \textit{C. amylolentus}. Additionally, a transformation system is needed for \textit{T. wingfieldii} to perform gene disruptions that will help elucidate the role of the HD and the P/R genes. In \textit{C. neoformans}, \textit{SXI1} and \textit{SXI2}, are upregulated during sexual development (62), and it will be interesting to investigate the expression of \textit{C. amylolentus} HD genes under mating-stimulating (V8) and nutrient-rich (YPD) conditions. In summary, \textit{C. amylolentus} has a tetrapolar mating system that is defined by the HD and P/R locus, and its \textit{MAT} configuration serves as an evolutionary intermediate in the formation of the bipolar mating system in the pathogenic \textit{Cryptococcus} species.

\section*{6.3 Sexual reproduction in \textit{C. amylolentus} – insights into mating in the pathogenic \textit{Cryptococcus} species complex}

In Chapter 4, we described an extant sexual cycle for \textit{C. amylolentus}, a fungus that was reported to be asexual. Surprisingly, matings between the only two strains of \textit{C. amylolentus} that are available and that have been stored frozen at the CBS for several decades were successful and resulted in the production of basidia terminating in four long individual chains of basidiospores. Fluorescent staining of mating filaments revealed dikaryotic hyphae and uni- and dinucleate spores. Sex in \textit{C. amylolentus} is reminiscent of opposite-sex mating in the pathogenic \textit{Cryptococcus} species. The steps involved in \textit{C. amylolentus} sexual reproduction are cell fusion, clamp cell formation and fusion, and nuclear fusion, culminating in meiosis and basidiospore production. The discovery of sex
in *C. amylolentus* suggests that the reported asexuality of certain fungal species may be a reflection of limited knowledge of methods for detecting mating (i.e. length of incubation or proper conditions). We have at our disposal a wide range of mating media, which can be utilized to determine if, under the appropriate conditions, mating is stimulated in an otherwise asexual species.

We microdissected 65 meiotic progeny from *C. amylolentus* matings and performed genotyping and RAPD analysis of markers linked and unlinked to *MAT*. This analysis revealed that meiosis occurs during sexual reproduction in this species as evidenced by the 50:50 marker segregation pattern observed in the progeny population and the high level of genetic exchange in the set 2 F1 progeny. Meiosis typically results in the production of 4 genotypes, and in *C. neoformans* a single meiotic event takes place in any given basidium (63). Similar to a recent study by Idnurm (63), we genotyped spores from a single chain of spores derived from a single basidium of *C. amylolentus* to determine if more than one meiotic event occurred per basidium. Our analysis of individual spore chains resulted in >4 genotypes produced per spore chain, providing evidence that more than one meiosis event occurs in a given basidia. A similar event has been described in *Armillaria gallica*, (118-120) in which single basidiomes (fruiting structures) give rise to multiple (>4) genotypes. This phenomenon is referred to as genetic mosaicism (118-120). The two-diploidization-two-haploidization life cycle of *A. gallica* produces genetic mosaicism, and this cell-line genetic variation occurs in response to different environmental conditions (i.e. pH and temperature). Additionally, we cannot rule out the possibility that sex in *C. amylolentus* results in the production of
aneuploids, which would also give rise to multiple genotypes observed in the individual spore chains. Genotyping additional spores of *C. amylolentus* will contribute to elucidating how and why some spore chains generate multiple genotypes. In conclusion, although the mechanisms governing sex in *C. amylolentus* appear to be similar to those in the pathogenic *Cryptococcus* species, the large number of multiple genotypes observed per spore chain indicates that the output of meiosis does not result in the typical four genotypes per basidium.

mtDNA is uniparentally inherited from the A2B2 parent in *C. amylolentus* like in *C. neoformans* and *C. gattii* (174), though a small percentage inherit mtDNA from the A1B1 parent (possible parental yeasts). Analysis of the mtDNA markers *NAD4* and *NAD5* revealed cytoplasmic exchange of mitochondria in *C. amylolentus*. In *C. neoformans*, mitotic/asexual spores or blastospores bud directly off hyphal filaments during sexual reproduction. We speculate that in *C. amylolentus* a second type of product (basidiospores) of sex inherit the A2B2 mitochondria and are produced primarily in set 1 F1 progeny. Surprisingly, one of the progeny from set 2 F1 is a bimater, mating with both CBS6039 and CBS6273. We hypothesize that sexual reproduction in *C. amylolentus* can produce a new mating-type in which additional unlinked HD genes may also contribute to sex. Further investigation of the mating-type specific genes will be necessary to understand how a bimater forms during sex in *C. amylolentus*.

Interestingly, most of the microdissected fertile progeny are α-mating-type and only mate with the CBS6273 parent, and >60% of the microdissected progeny are sterile and unable to mate with either of the parental strains or the other progeny. Why are so
many of the progeny sterile? Meiosis may damage MAT alleles in the sterile progeny, inducing breaks in the regions that define MAT, or reshuffling of the genome during sex may increase chromosomal rearrangements (or aneuploids) that impair fertility. Whole genome sequencing of C. amylolentus and comparative genomic hybridization (CGH) could be used to investigate the possibility of aneuploidy in the sterile progeny that could cause sterility. Analysis of the expression of SXII and SXI2 in the sterile versus the fertile progeny and the parental strains might also be informative. Further, expression of the HD genes in the sterile progeny might be extremely low, even under nutrient rich conditions (on YPD), relative to levels in the fertile progeny and in the parental strains.

During the analysis of the meiotic progeny, we noted that a subset of the progeny appear ruffled and filamentous on YPD like the CBS6273 parent, while other progeny are smooth and mucoid, resembling the CBS6039 parent. The filamentous phenotype appears to be linked to mating ability. This phenotype can be further examined by identifying factor(s) that control filamentation in C. amylolentus. In future studies, Agrobacterium-mediated transformation, which has been used with much success in the pathogenic Cryptococcus species (65, 166), of the CBS6039 and CBS6273 parental strains could be used to isolate hyperfilamentous or afilamentous mutants. In Chapter 3, we characterized the tetrapolar mating system of C. amylolentus and in Chapter 4 we described the discovery of its extant sexual cycle and sex in the sibling species shares features with reproduction in C. neoformans and C. gattii.
6.4 The unlinked HD and P/R loci in *T. mesenterica* span ~3 kb and 18 kb, respectively

The wood rotting fungus and tetrapolar basidiomycete *T. mesenterica* served as an outgroup for the studies on the evolution of the mating-type locus in the pathogenic *Cryptococcus* species complex. The recent release of the genome sequenced at the JGI/DOE allowed the identification of *MAT* genes in the *T. mesenterica* genome using BLAST to query the database. Comparative genomics of *T. mesenterica* strains with opposite mating types revealed that the *SXI1-* and *SXI2*-homolog genes define the HD locus and span ~3 kb while the P/R locus is defined by *STE3, STE12, STE20, tremerogen a-13,* and several hypothetical genes spanning ~18 kb. Although these regions are relatively small in comparison to the unusually large *MAT* locus of the pathogenic *Cryptococcus* species, a vast majority of the *Cryptococcus MAT*-linked genes flank the regions that define *MAT* in *T. mesenterica*. The flanking genes are not yet incorporated into the *MAT* loci of *T. mesenterica* but are in close proximity to the *MAT*-defining regions, and inversions driven by transposons are likely responsible for recruiting these genes into the *MAT* locus of the pathogenic *Cryptococcus* species. To determine if *MAT*-specific regions are located outside of the regions we defined as *MAT*, we analyzed two genes (*GEF1* and *RPO41*) from the P/R locus in the *T. mesenterica* strains. Their identity is >96%, suggesting that they may lie outside of the P/R defining region. This analysis is not comprehensive and should be extended to additional genes residing in regions unlinked to the HD or P/R locus in *T. mesenterica*.

Many of the defining characteristics of *MAT* evolution discussed in Chapter 3 are also present in *T. mesenterica*. We identified retrotransposons flanking *MAT* and
unlinked HD and P/R loci, with two homeodomain transcription factors linked to the HD locus while the pheromone receptor and pheromone genes are linked to the P/R locus. Furthermore, the HD loci of all of the species examined in this dissertation contain two divergently transcribed genes. We also identified hypothetical genes in MAT of *T. mesenterica* and in the sibling species that are present on the telomeric ends of chromosome 4 in *C. neoformans* strain JEC21 (102, 127). It is apparent from our analyses in Chapters 3 and 5 that sex chromosome and MAT evolution in fungi, animals, and plants are characterized by similar events such as transposon insertion, chromosomal translocations, gene conversion, and gene rearrangements, all of which have shaped the evolution of MAT in *T. mesenterica*, *T. wingfieldii*, *C. amylolentus*, *C. heveanensis* (102), and the pathogenic *Cryptococcus* species complex.

In conclusion, the previous MAT evolution model proposed by Fraser et al. can be amended based on this current study (41) (see Chapter 3, Figure 16). We confirmed that the presence of unlinked HD and P/R loci preceded the fusion of the subloci into a bipolar mating system. First, the HD locus encodes two divergently transcribed homeodomain transcription factors and the P/R locus encodes the pheromone receptor and pheromone genes. Next, both loci underwent gene acquisition and further expansion, accumulating additional MAT-specific genes. A translocation event occurred between chromosomes 4 and 5 of *C. neoformans* resulting in the formation of a transient tripolar intermediate and loss of one of the HD genes. Loss of one HD gene restricts outbreeding and increases inbreeding (124). Finally, the tripolar intermediate collapsed to form a
bipolar transition state that undergoes further gene rearrangement and inversions resulting in the bipolar mating system of *Cryptococcus*.

Hsueh et al. recapitulated the intermediates in the MAT evolutionary model by genetically engineering *C. neoformans* to mirror a tetrapolar and tripolar mating system. Surprisingly, *C. neoformans* is able to complete sex in a tetrapolar mating system, although the tripolar system is hypothesized to have been detrimental and inbreeding was further restricted (60). These observations support the model that the bipolar mating system in the pathogenic *Cryptococcus* species evolved from a tetrapolar system. In this dissertation, comprehensive analysis of the mating-type locus in the closely related sibling species and the more distantly related species *T. mesenterica* provides insight into the evolution of MAT and sexual reproduction in the *Cryptococcus* species complex. In a broader sense, our analysis of sex chromosome evolution in phylogenetically closely and more distantly related fungi shows that strong selective pressures have driven evolution of the overall structure and organization of the plastic sex-determining regions in these species. These studies have broad implications for the analysis of the evolutionary events that helped shape the formation of the MAT locus and ultimately control sexual reproduction in fungal pathogens.
Appendix A

*Tsuchiyaea wingfieldii* – *MAT* loci

- **Fosmid 2B23**
  - Homeodomain Locus – 40 kb

- **Fosmid 2K10**

- **Fosmid 4E07**
  - 31 kb

- **Fosmid 3F11**

- **Fosmid 3A15**

- **Fosmid 5J15**

Pheromone/Receptor Locus – 70 kb
Appendix B

Cryptococcus amyloleotus – MAT loci

Homeodomain Locus – 20 kb

Pheromone/Receptor Locus 60 kb
Appendix C

Appendix C: Genotyping analysis of 65 C. amyloentus progeny. 
ETF1 PCR-RFLP of products digested with DdeI. RPL39, GEF1, and STE3 were scored for the absence or presence of a single PCR product. Sample order: Row 1 - set 1 F1 # 1-24, Row 2 - 4- set 1 # 25-28 and set 2 F1 # 1-20; Row 3 - set 2 F1 # 21-31, set 2 F1 #8, 9, 10, 23, and F2 #1-6. With the exception of ETF1, CBS6039, CBS6273, and water are included as controls and 1 kb ladder is located in the first and last lanes of each row. (Refer to Tables 5 and 6.) (Data provided by Sheng Sun.)
Appendix D: Mating assays between MAT recombinant progeny and the bi-mater set 1 F1 #4 crossed to the parental strains. Light microscopy of mating structures produced in the interfertile progeny (top row, Left panels – 2 F1 #10 x F2 #1 and right panels – Set 2 F1 #16 x F2 #1 and Set 2 F1 #16 x F2 #5) and backcrosses of progeny #4 to CBS6039 and CBS6273 (bottom row). Scale bar = 10 μm.
References


standardize techniques on which to base a detailed epidemiological survey. Electrophoresis 20:1790-9.


149. **Strathern, J., B. Shafer, J. Hicks, and C. McGill.** 1988. a/α-specific repression by MATα2. Genetics 120:75-81.


161. **van der Walt, J. P.** 1973. \textit{Aessosporon dendrophilum} sp. nov., the perfect state of \textit{Bullera dendrophila}. Antonie Van Leeuwenhoek \textbf{39}:455-60.


175. **Yan, Z., C. M. Hull, S. Sun, J. Heitman, and J. Xu.** 2007. The mating type-specific homeodomain genes *SXI1alpha* and *SXI2a* coordinately control uniparental mitochondrial inheritance in *Cryptococcus neoformans*. Curr Genet **51**:187-95.

176. **Yan, Z., and J. Xu.** 2003. Mitochondria are inherited from the *MATa* parent in crosses of the basidiomycete fungus *Cryptococcus neoformans*. Genetics **163**:1315-25.


Biography

Keisha Monique Findley

Born
February 27, 1982 in Washington, D.C.

Education
Bachelor of Science  May 2004
Department of Biochemistry and Molecular Biology, University of Maryland Baltimore County

Publications

Research papers


Book chapters

**Conference Poster Presentations**

American Society of Microbiology General Meeting (May 2009)
Poster: “Evolution of MAT and Sex in the *Cryptococcus* Species Complex”

North Carolina Branch of the American Society for Microbiology Annual Meeting (October 2008)
Poster: “Evolution of MAT and Sex in *Cryptococcus*”

Gordon Conference: Cellular & Molecular Fungal Biology (June/July 2008)
Poster: “Evolution of MAT and Sex in *Cryptococcus*”

Fungal Genetics Conference (March 2007)

**Oral Presentations**

Genome Biology Seminar (September 2008)
Talk Title: “Evolution of MAT and Sex in *Cryptococcus*”

Departmental Meeting (February 2007)
Talk Title: “Evolution of the *Cryptococcus* Mating-type Locus: a Comparative Genomics Approach”

**Awards**

Carl Storm Fellowship Recipient to attend Gordon Conference (Summer 2008)
Minority Supplement funded by NIH (2006-2010)

**Professional Affiliations**

Student Member of the Mycological Society of America (2008-present)
Student Member of American Society of Microbiology (2004-present)
Student Member of American Association of the Advancement of Science (2004-present)