Title

A genomics based approach to exploring the potentiality for a complete sexual cycle in the pathogenic fungus Candida albicans

Author

Tzung, Keh-Weei

Publication Date

2004

Peer reviewed|Thesis/dissertation

A Genomics Based Approach to Exploring the Potentiality for a Complete Sexual Cycle in the Pathogenic Fungus, *Candida albicans*

by

Keh-Weei Tzung

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

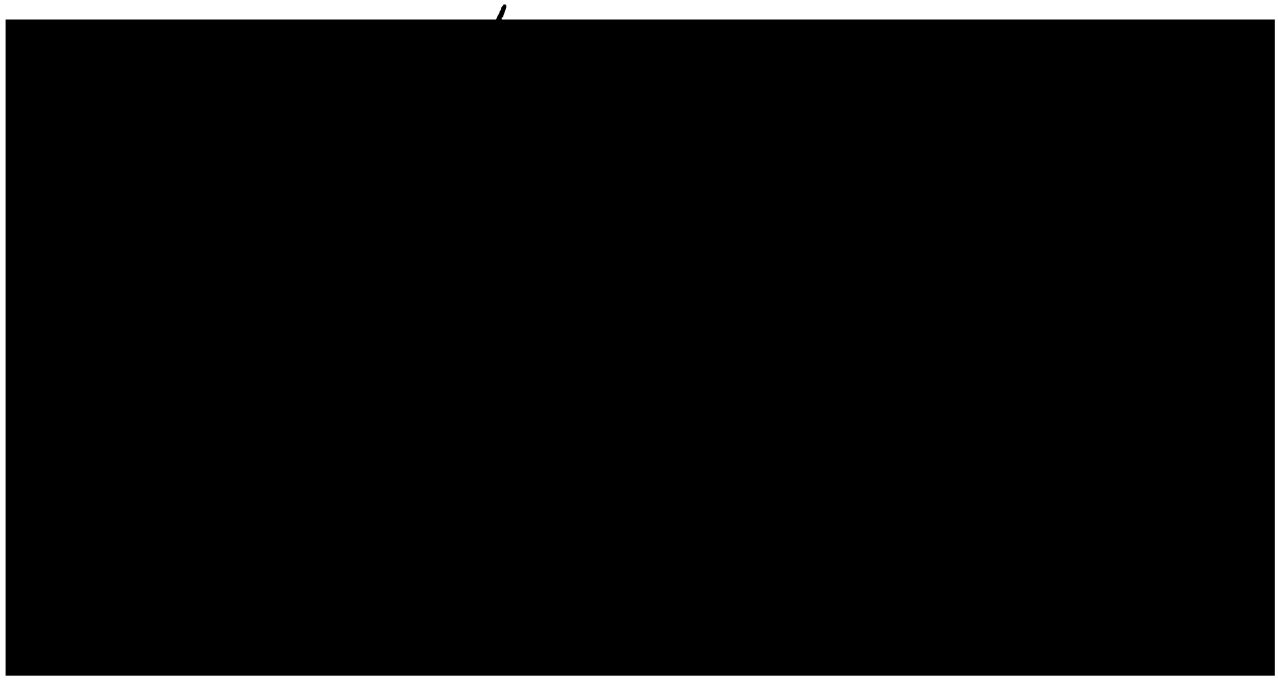
Oral and Craniofacial Sciences

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO



To my parents, 宗建成 and 黄澄江

for raising me, teaching me to be humble and confident and to respect others, for believing in me, for inspiring me to do my best, for encouraging me when I am in tears, to whom, I owe all my achievements.

Acknowledgements

"Most people are other people. Their thoughts are someone else's opinions, their lives a mimicry, their passion a quotation."—Oscar Wilde.

Graduate school is a place to champion ideas, nurture creativity, learn knowledge and explore the unknown. Independent thinkers are born and dreamers can cherish their academic freedom. I was fortunate enough to have spent the past several years at UCSF as a graduate student, and live in San Francisco to experience a variety of cultures and enjoy natural beauty. My experience in San Francisco was a journey of self-discovery and a "moveable feast".

First, I am grateful to my mentor, Nina Agabian, for her support and guidance. Nina encouraged me to pursue the questions which intrigued me most. She gave me complete freedom to learn science, which allowed me to mature and find my theme in science. She encouraged me to take the road less taken and to become fearlessly independent among the current scientific trend. She taught me how to write concisely and to present my ideas clearly. I cannot thank her enough for her scientific insight and for her faith in me.

I am also thankful for my thesis committee, Ira Herskowitz, Pat O'Farrell, Renee Reijo and Caroline Damsky. I was always amazed at Ira's ability to see the "big picture" in science and admired his relaxed attitude toward life. I remember how exciting and intimidating it was to discuss yeast meiosis with him and his follow up emails were always thoughtful and encouraging. I am indebted to the support he gave me at the onset of this project. Unfortunately, he died of pancreatic cancer last year and is deeply missed. To me, Pat is the ideal scientist both in his scientific depth and in his intellect. I

did my first lab rotation in his lab and it was an eye-opening and brainstorming experience as a foreign student. I enjoyed working with him and hope I can reach his intellectual caliber some day. I admire his quest for knowledge and his style of running a lab. Discussions with Pat have contributed significantly to my growth as a scientist and as a person. I also want to express my gratitude to Renee for her encouragement and input. I am delighted that I have gained new perspective on meiosis due to her suggestions. Her expertise was a boon to the completion of my project. Caroline joined my committee after Ira passed away. As the Oral and Craniofacial Sciences program director, she was always a cheerleader for my work. I would not have come to UCSF for graduate school if it were not for her support and encouragement. She guided me through some tough times in graduate school and her kindness is deeply appreciated.

I also want to thank the members of the Agabian lab, both current and past, for their friendship, support and tolerance toward my crazy love for meiosis. They have become a family for me in a foreign land. I am grateful to George Newport who always tried to answer my million questions, criticized my weird ideas without losing focus and corrected my writing with understanding. I really enjoyed conversation with him about films, which made late-night working much fun. I also want to thank Chung-Yu (Henry) Lan for his help on some tedious graphic work for my presentations, and for being a good listener who helps me weather through some dark days in the lab. I want to thank Nicholas Leaskou for his administrative help, and for numerous conversations we had regarding the arts, films and politics. My bench work would not be possible without the help from Samantha Donnelly, Jan Dungan, Gerwald Köhler, Alan Kuo, Gabriel Rodarte, Teresa Moeller, Luis Murillo, Isabel Cano, Silvio Favoreto, Kent Persson, Onix Adegbola, Cindy Lijima, Fang Feng, Jeff Roberts, Eva Harris, Jeff Holmes and Julie Johnson-Blake. We had fun discovering science together.

I must thank my collaborators who contributed to this work tremendously. Working with Stanford Genome Technology Center has been a privilege for me. I would like to thank Roy Williams, Ron Davis and Stew Scherer for their discussion and input for this project. I also deeply appreciate the help from CompleGen, especially John Swindle and David Weston, for their generosity in sharing protocols and technical help. In addition, I want to thank Brent Marsh for writing all the computer scripts for me, for his friendship and all the fun times we had together.

I also want to thank Tong-Seung Tseng, Shahid Chaudhry, Chun-Pyn Shen, JY Chen, for constantly thinking about me and cheering me up when I was down.

Finally, I want to thank my entire family for their love and support throughout the years. In particular, I am grateful to my parents, for loving me enough to let me go halfway across the world alone, and for supporting me during my endless academic pursuits, and for believing that I will continue to find the right path in life.

"To live and not to know why the cranes fly, why children are born, why the stars are in the sky. Either you know why you're alive or it's all nonsense, it's all dust in the wind"—Chekhov.

In the past several years, I have cried, stumbled, laughed, learned, fallen, and stood up. Finally, this long winding journey has come to an end. In the least, it is my wish that this thesis reflects my passion for knowledge and quest for the meaning of life.

UNIVERSITY OF CALIFORNIA, SAN FRANCISCO

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO SANTA BARBARA • SANTA CRUZ

MOLECULAR PATHOGENESIS

SCHOOL OF DENTISTRY
SCHOOL OF PHARMACY
SCHOOL OF MEDICINE

Box 0422, Clinical Sciences Bldg, C740
521 PARNASSUS AVENUE
SAN FRANCISCO, CA 94143-0422
(415) 476-6849 TEL
(415) 476-0864 FAX

FAXED 3.7.04

March 1, 2004

PNAS,
2101 Constitution Avenue,
NW, NAS 340,
Washington, DC 20418 USA

To whom it may concern:

I would like permission to include with my Ph.D. dissertation a copy of the following article cited below, which was published in *PNAS*. This dissertation will be microfilmed by University Microfilms Incorporated and they request permission to provide a single copy on demand.

Tzung KW, Williams RM, Scherer S, Federspiel N, Jones T, Hansen N, Bivolarevic V, Huizar L, Komp C, Surzycki R, Tamse R, Davis RW, Agabian N. Genomic evidence for a complete sexual cycle in Candida albicans. Proc Natl Acad Sci U S A. 2001 Mar 13;98(6):3249-53.

Please acknowledge receipt of this fax and respond by fax to 415-476-0864.

Thank you for your cooperation.

Sincerely,

Kuh-Wei Tzung
Kuh-Wei Tzung

Permission granted. Please cite full journal reference and "Copyright (year) National Academy of Sciences, U.S.A."

Diane M. Sullenberger, Executive Editor
03/05/04

A genomics based approach to exploring the potentiality for a complete sexual cycle in the pathogenic fungus, *Candida albicans*

Keh-Weei Tzung

Abstract

The role of sexual recombination in the evolution of eukaryotic pathogens is not well understood. For *C. albicans*, a fungal pathogen of immunocompromised individuals, the potential to undergo a complete sexual cycle is of both clinical and experimental importance. *C. albicans*, classically characterized as an asexual diploid, and thus imperfect fungus, has recently been shown to undergo mating without resolution of progeny; evidence suggests that some level of recombination might occur in nature. I have used comparative genomic and molecular genetic approaches to test the capacity of key genes in the meiotic pathway to support a complete sexual cycle in *C. albicans*. To explore the mechanism(s) governing the onset of meiosis, I have used a combined approach of functional complementation, genomic sequence comparison and gene expression analysis. These studies indicate that there is a fundamental difference between the putative meiotic signaling pathway of *C. albicans* and its *S. cerevisiae* counterpart. I have shown that the initiation complex for meiosis *IME1*/*UME6*/URS1, as defined in *S. cerevisiae,* does not exist in *C. albicans*. This finding is consistent with the notion that the master regulator for meiotic development varies among species, and represents a specific code that combines unique genetic and environmental signals. I have further demonstrated that the temporal expression of the *C. albicans* equivalent of the subset of genes in the *S. cerevisiae* meiotic pathway is not coordinately expressed. Although *C. albicans IME2* expression can be induced when *C. albicans* is grown in *S.*

cerevisiae sporulation medium, no ensuing meiosis occurs. In aggregate, these results indicate a differential wiring of the *S. cerevisiae* transcriptional circuitry in *C. albicans* for an as yet undetermined outcome/function. Either *C. albicans* has lost the capacity to execute a functional meiotic program, and putative meiotic genes have evolved to acquire new functions, or *C. albicans* might have developed an organism-specific meiotic machinery for its own evolutionary advantage—survival in a mammalian host. If the latter is the case, it provides a platform for understanding potential pathways whereby sexual recombination may take place in other imperfect fungi.

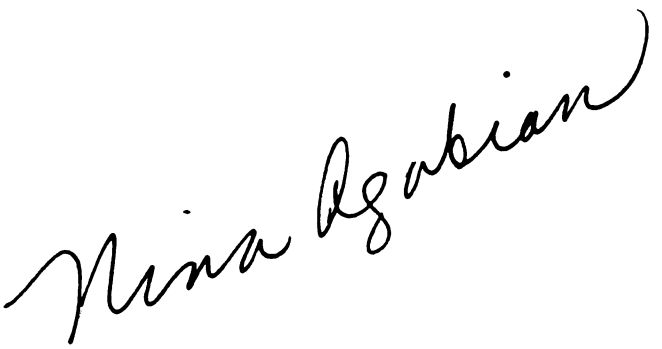


Table of Contents

List of Tables

Chapter 2

Chapter 3

List of Figures

Chapter 2

Chapter 3

Chapter 1

Introduction

Candida albicans as an opportunistic pathogen

C. albicans is the most important human fungal pathogen (1). It is present as a commensal in the gastrointestinal tract, oral cavity and vagina of many healthy individuals, and sometimes causes superficial and mucosal infection. In patients immunocompromised as a result of HIV infection, organ transplantation or cancer chemotherapy, *C. albicans* can cause a spectrum of diseases ranging from invasion of various cell and tissue types to life-threatening disseminated candidiasis (2).

Genomic variations in *C. albicans*

C. albicans is mainly a diploid organism. Traditionally it is classified as an imperfect fungus, because a sexual phase has never been observed (3). Despite its assumed asexual life cycle, *C. albicans* displays a broad spectrum of phenotypes which encompass morphological, karyotypical, biochemical and antigenic variations (4). In response to different environmental cues, *C. albicans* is able to proliferate in many forms—as budding yeast, pseudohyphae and filamentous hyphae (5). Under unfavorable conditions, it can also produce thick-walled, asexual chlamydospores, which have been suspected as vestigial sexual structures (6, 7). In addition to these morphogenetic transitions, *C. albicans* can undergo high-frequency phenotypic switching, a postulated strategy for survival in highly variable environments (8). Several switching systems are well-documented or characterized for *C. albicans*. The strain 3153 switching system manifests as variability of colony phenotypes including smooth, fuzzy, ring, irregular wrinkle and stipple (9). UV-irradiation experiments have linked DNA rearrangement to these differences in colony morphology. In the "white-opaque" switching system, identified in strain WO-1, *C. albicans* cells switch between ovoid, yeast-like " white cells" and elongated, larger, bean-shaped "opaque cells" (10). A whole-genome expression analysis of white and opaque cells indicates that there are

significant physiological differences between each phenotypic variant (11). The molecular mechanism of phenotypic switching is not clear, however several lines of evidence suggest that chromatin remodeling could be the primary determinant. The histone deacetylase genes *HDA1* and *RPD3*, and transcriptional regulators *TUP1* and *SIR2* have been shown to control phenotypic switching in *C. albicans* (12-14). It has also been proposed that there is a correlation between the diversity of karyotype and the variability of phenotype. For example, assimilation of sorbose is correlated with the loss of a homologue of chromosome 5, and the loss of a homologue of chromosome 4 is associated with fluconazole resistance (15, 16). These findings lead to the notion that *C. albicans* can deploy a rapid adaptative strategy to succeed as a pathogen through a variety of developmental programs including high- frequency phenotypic switching, the yeast-hyphae transition and the expression of virulence genes in various environmental/anatomical niches.

Sexual versus asexual reproduction

The sexual cycle has been crowned as "the masterpiece of nature" (17). Most species of plants and animals reproduce sexually, however a few organisms seem to lack a complete sexual cycle. What does this rarity of asexual reproduction among the eukaryotes suggest? One interpretation suggests that the loss of a sexual cycle might be a dead end of evolution, leading to early extinction (18). One exception is the Class Bdelloid of the Phylum Rotiferas which stands out as an "evolutionary scandal", as it represents the largest metazoan taxon where males, hermaphrodites and meiosis are unknown (19). There is much debate about why sexual recombination is dominant in nature. Two plausible arguments suggest that "sex is better" (20-22). The first argument indicates that sexual reproduction provides a milieu for the mixing of beneficial mutations from independent lineages. In an asexual organism, mutations arising from different

individuals will not be combined in one, and adaptive evolution occurs from successive accumulation of favorable mutations in individually evolving lineages. In a sexual species, assortment of chromosomes and recombination can bring favorable mutations together in one individual, providing a survival advantage in changing environments. Thereby sexual recombination becomes a more efficient way to spread beneficial genotypes. The second argument suggests that sexual recombination will facilitate DNA repair, or dilute deleterious mutations (23). In this postulated error correction mechanism the intact DNA of one haplotype can serve as a template for repairing the damaged DNA of the other. Furthermore, the unfavorable mutations on one haplotype will be compensated by intact DNA or dominant mutations on the other haplotype. As Muller's ratchet predicts that there is an irreversible accumulation of harmful mutations in a small asexual population, eventually there is no individual left without mutations (24, 25). In a sexual population, sex will purge the genome of unfavorable mutations because meiotic recombination will allow accumulation of harmful mutations in a few individuals which will be eliminated by selection. Taken together, this model indicates that sex helps remove harmful mutations.

However, sex is not always advantageous (26). Meiotic recombination disrupts co-adapted gene combinations, while advantageous genotypes will be broken through chromosome segregation and assortment. Meiosis and syngamy expend energy, and are not very efficient biological processes. For anisogamous sexual reproduction, there is a two-fold cost of sex, compared to asexual species. In addition, sexual species need to find a mate, which is a disadvantage in situations such as species with limited motility or in sparse populations. Moreover, some cytological types such as aneuploids can not be propagated successfully through meiosis.

Pathogenic fungi usually exhibit a complicated life history. Many have been shown to reproduce both sexually and asexually. Population structure analysis indicates that organisms such as *Mycosphaerella graminicola*, *Cryphonectria parasitica*, *Sclerotinia scleroticorum* possess both clonal and recombinational growth (27-30). Some fungal species were believed to be asexual; however evidence of genetic recombination is present in the wild, including the pathogens *Coccidioides immitis*, *C. albicans* and *Cryptococcus neoformans* (31-33).

What role the existence of a sexual cycle may play in the evolution of pathogens and their relationship with their hosts is not well understood. It has been suggested that the population structure of parasites has important evolutionary and public health implications. For example, Chagas disease caused by *Trypansoma cruzi* shows clinical pleotropism, and different strains affect heart and GI tract with varying degrees of virulence (34, 35). While *T. cruzi* displays clonal growth, the extant lineages diverged from one another before human colonization on earth. Therefore the genetic divergence among clonal lineages of *T. cruzi* may contribute to medical attributes such as drug and vaccine vulnerability and pathogenicity of this organism. For another parasite, *Toxoplasma gondii*, 3 clonal lineages have been formed globally without evidence of sexual recombination (36), however the completion of the sexual cycle in *T. gondii* produces recombinant progeny with enhanced virulence (37). The issue of sexuality in pathogenic fungi is more elusive. While mating has been demonstrated in *C. albicans*, meiosis has not been shown. A sexual cycle has been reported for *C. neoformans*, and most clinical isolates are α–mating strains, which are more virulent than **a** strains (38, 39). However, the function of the sexual cycle and its role in pathogenesis of *C. neoformans* remain unknown. In this thesis, I will use *C. albicans* as a case study to address some fundamental aspects in understanding reproduction in pathogenic yeasts.

Controversy about *Candida* sexuality

Since its classification in the "*Fungi imperfecti*", *C. albicans* has been the subject of controversy with respect to its obligate, asexual and diploid characteristics. In 1967, van der Walt reported that he had discovered a sexually active form of *C. albicans* (40). He suggested that the cells which he observed were sexual conidia, and that meiosis had occurred. However this was based on microscopic observation; there was no genetic proof for the existence of this sexual cycle. Later, in 1982 Whelan and Soll identified several balanced recessive lethals in UV-irradiated isolates, which suggested that temperature-sensitive or lethal heterozygosity might be frequent in natural strains of *C. albicans* (41, 42). Their results provided the first documentation of ploidy in *C. albicans*, and refuted the possibility of meiosis; as there would be low viability of meiotic progeny.

Since the 1980s, the study of the genetic structure of populations of microorganisms has provided new insight into understanding their epidemiology, pathogenesis and reproductivity. Several criteria have been proposed for assessing clonality of populations in microorganisms (43, 44). These include absence of segregating genotypes, fixed heterozygosity, the presence of over-represented genotypes, deviation from Hardy-Weinberg equilibrium, linkage disequilibrium and absence of recombinant genotypes. However, a clonal population structure does not imply that sex is totally absent, but only that it occurs too rarely to break the pattern of clonal propagation. The available data from several population genetics analyses for *C. albicans* are quite contradictory. Early studies have highlighted the key result and debate. Tibayrenc et al. (45, 46) first indicated that there was no evidence supporting clonality in *C. albicans*. Pujol et al. (47) have presented evidence of a lack of recombination and segregation in 21 gene loci in *C. albicans* strains, suggesting clonal growth. However, Grasser et al. (32) have identified both clonality and genetic exchange in the same *C. albicans* strains using

RAPD (random amplification of polymorphic DNA) and SSCP (single strand conformation polymorphism). Later, Forche et al. found evidence of clonality and genetic exchange in typical and atypical *C. albicans* strains from Africa (48). Taken together, these data indicate that *C. albicans* displays predominantly clonal reproduction, however some extent of recombination is likely to occur. So how much recombination has occurred in the *C. albicans* clinical isolates? What is the mechanism responsible for recombination? Is cryptic sex common in fungi? As atypical *C. albicans* strains and new *Candida* species emerge with increasing frequency, the question arises is *C. albicans* showing an "epidemic" population structure where new lineages arise from recombination of parental strains, and are expanded by clonal growth (48, 49)?

Transposon and Sexuality

Transposons are a major component of many eukaryotic genomes, and often constitute a significant fraction of the genome. For example, transposons occupy an estimated 50% of the human genome (50). There are two major classes of transposons: retrotransposons and DNA transposons. The retrotransposons encode reverse transcriptase and transpose via RNA intermediates. The DNA transposons usually encode a transposase, which excises the element from the host genome and reinserts it at a new site. Transposons are known to have impact on the evolution of genomes (51-53). For instance, recombination between transposable elements at different sites can lead to large-scale chromosome rearrangements; insertion of transposons within genes can cause mutations, while insertion in the promoter regions can change gene expression patterns. Host genomes usually have defensive strategies to suppress the potentially deleterious effects from transposon proliferation. For example, cytosine methylation limits the activity of transposable elements, and defends the mammalian genome against damage (54, 55). Alternatively, clustering of transposable elements or

targeting to specific locations in the genome can account for not only host tolerance of transposons to allow them to persist but also prevention of disruption of essential genes (56, 57).

What is the role of transposons in evolution (58)? The "helpful" transposon hypothesis suggests that transposable elements are important for the evolution of genomes by enhancing genome plasticity, accelerating large-scale genomic alteration and regulating gene expression. In this regard, transposons should provide a greater advantage to asexual organisms as compared with sexual organisms, which produce diversity through sexual recombination. Thus according to this hypothesis active transposon populations provide a vehicle for generating genotype diversity in asexual organisms.

The original "selfish DNA" hypothesis suggests that transposons replicate in their own interests without serious harm to the host (59). Further, Hickey proposed that transposons are "sexually-transmitted nuclear parasites", where sexual reproduction will allow transposons to proliferate even when they substantially endanger the host (60). In this parasite model, transposon activity is held to be deleterious. Therefore it predicts that sexually reproducing organisms will harbor more aggressive and abundant transposons, which in turn select for host suppressive mechanisms to limit transposon activity, while the genome of obligate asexual organisms should be free of these active elements (61).

Comparison between transposon populations in closely-related sexual and asexual species and their activity should resolve the relative merits of each of these two models. For example, it has been shown in *S. cerevisiae* that active Ty3 elements are more easily disseminated in sexual populations than in asexual populations (62). For the

obligate asexual organism, Bdelloid rotifers, there is no evidence of active transposons such as LINE-like and gypsy-like retrotransposons, which are prevalent among sexual organisms (63). It is noted that mitochondrial genomes in mammals are subject to maternal inheritance and are not sexual. They are devoid of transposons, while biparental mitochondrial genomes in sexually reproducing fungi frequently possess active transposable elements (64).

With the initiation and completion of several fungal genome projects, transposon populations have been characterized in *S. cerevisiae*, *C. albicans*, *C. neoformans*, *Schizosacchmromyces pombe*, and the hemiascomycetous yeasts (53, 65-68). These comparative analyses will certainly help to understand the role played by transposons in genome evolution.

Compared to *S. cerevisiae*, *C. albicans* has much more distinct retrotransposon families; these appear to be nonfunctional, and are present in low copy number. Moreover, *C. albicans* has non-LTR retrotransposons and DNA transposons, which are absent in *S. cerevisae*. In *C. neoformans*, Goodwin et al. have identified 15 distinct families of retrotranspons and several families of non-LTR retrotransposons (67). *C. neoformans* has fewer transposons than *C. albicans*, but has much more than either *S. cerevisiae* or *S. pombe*. These findings suggest that elimination of retrotransposons may occur more rapidly in *C. neoformans* than in *C. albicans* but not as rapidly as in *S.cerevisiae* and *S. pombe*. Why is the rate of loss of retrotransposons varied among different fungal species? How does the reproductive mode of the host contribute to the diversity of the transposons in the genome? As more genomes are sequenced and characterized, perhaps we will learn more about the genesis of transposon diversity in different species during evolution.

Identification of mating and meiosis gene homologues in *C. albicans*—pre-genomic era

Although *C. albicans* has long been classified as an asexual fungus, *Candida* homologues of genes associated with the sexual cycle in *S. cerevisiae* have been reported. Hicks laboratory first identified a gene encoding a Gα subunit in *C. albicans* (69). This *Candida* gene was able to complement similar deficiencies in the *S. cerevisiae* mating pathway, and was regulated by the *S. cerevisiae* **a**1-α2 transcriptional repressor. In *S. cerevisiae*, mating type switching involves transfer of a "silent cassette" to the resident *MAT* locus and thereby its expression (70). This model appeared to be useful for studying phenotypic switching in *C. albicans*, as reversibility is a main feature. In order to address the possible molecular basis for "white-opaque" switching in WO-1 strain, Soll's group first suggested that there might be a mating type locus present in *C. albicans* resembling the *MAT* cassette in *S. cerevisiae* (71). Southern hybridization with a cloned *MAT* ***a*** probe at medium stringency failed to identify any homologous sequences. The initial cloning of a G protein α subunit (*CAG1*) in *C. albicans* was significant in this regard. *CAG1* was shown not only to interact with the pheromone receptor, but was also subject to **a**1-α2 regulation in *S. cerevisiae*. This study provided the first clue that a master regulator similar to that found functioning in the sexual cycle of other fungi might also be present in *C. albicans*. Several *C. albicans* genes participating in the MAP kinase pathway were also identified by heterologous complementation or genomic library screening including *HST7* (*STE7*), *CPH1* (*STE12*), *CEK1*, *CST20* (*STE20*), and *CEK2* (72-76). *KEX2*, the processing enzyme for α–mating factor, and *HST6* (*STE6*), an ABC transporter for **a**-mating pheromone in *S. cerevisiae* have also been identified and functionally tested (77, 78). Further, *DLH1*, the homologue of meiosis-specific *DMC1* was identified (79). Taken together, these studies strongly

suggested that a potentially cryptic mating pathway might be present in *C. albicans*. Alternatively, these genes could have acquired additional or different functions during evolution.

Life in the genomic era

In 1996, the Stanford Genome Technology Center (SGTC) began the whole genome shotgun sequencing of *C. albicans* strain SC5314. The *C. albicans* genome project has provided the underlying information required to identify genes either unique to *C. albicans* or homologous to other organisms. One example of such gene discovery was the identification of a mating type-like locus (*MTL*) in *C. albicans*. Based on the sequence data from the publicly available *C. albicans* genome project, Hull and Johnson cloned the ***a1*** homologue from a *C. albicans* genomic library. By chromosome walking, they discovered a gene cluster of nearly 9 kilobases on two chromosome 5 homologues in *C. albicans* which resembled the mating type loci of other fungi (80). The entire mating type-like locus (*MTL*) encodes 9 ORFs including the *Candida* homologues of ***a1***,α*1* and α*2*; these appear to be arranged similarly to their counterparts in *S. cerevisiae*. There are six additional ORFs in the locus including two phosphatidylinositol kinases (*PIKs*), two oxysterol binding protein-like proteins (*OBPs*) and two poly (A) polymerases (*PAPs*). Gene disruption experiments suggest that *MTL**a*** or *MTL*α exists as a single copy; no *HMR* or *HML*-like silent cassettes have been identified as in *S. cerevisiae*. Moreover, no *HO* endo-nuclease, the enzyme which makes a double-stranded break at *MAT* locus, was found in the *C. albicans* genome (81, 82). Recently, the ORF annotated as *MTL**a2*** has been identified, which interestingly is more closely related to *MAT1-1-3* in *Cryphonectria parasitica*. The importance of this finding will be discussed in a later section.

Mating in *C. albicans*:

It has been reported that most clinical isolates of *C. albicans* are **a**/α cells by analogy with *S. cerevisiae*, however some naturally occurring **a** or α cells (nearly 3%) have also been found (83).Therefore, there are 3 different cell types in *C. albicans*, i.e. **a**, α and **a**/α cells. The marked imbalance in the *MTL***a**/ *MTL*α/*MTL***a**α ratio strongly suggests that *C. albicans* reproduces primarily through an asexual life cycle and that mating in nature must occur at low frequency.

The mating type locus is the master regulator of sexual morphogenesis and cell fate determination in yeasts (84). The identification of an *MTL* locus reopened the question of sexuality in *C. albicans*. Two different approaches have been taken to construct potentially mating-competent strains in *C. albicans*; both rely on creating **"a"** and "α" cell types similar to those in *S. cerevisiae*. Hull et al. (85) created **a** mating strains by deleting the α1 and α2 ORFs or by knocking out the entire *MTL*α locus; α mating strains were created by deleting the **a**1 ORF or by knocking out the whole *MTL***a** locus. Mating competent strains carried different auxotrophic markers, and mating was demonstrated in mice after tail vein injection with various ratios of **a** and α cells. The mating products were tetraploid or aneuploid as determined by FACS analysis, and showed a single nucleus by DAPI staining. No mating could be demonstrated *ex vivo*. They interpreted their results to suggest that the mammalian host provides the necessary environmental cues for mating.

Magee et al. (86) created mating competent strains using an alternative approach, based on the observation that culture of *C. albicans* on sorbose as sole carbon source will cause loss of a homologue of chromosome 5 (15); the *MTL* has been mapped to this

chromosome. Using this approach, mating strains of **a** and α cells were developed by growing *C. albicans* on sorbose, using strains which carried different auxotrophic markers on each homologue of chromosome 5. Mating was achieved *in vitro* by first cross-streaking opposite mating type cells on rich medium, then replica plating onto selective medium. The mating products were tetraploid, mononucleate, and contained both *MTL* alleles. Mating was better at room temperature than at 30 or 37^{0}C, suggesting that temperature might be a rate-limiting step in the process.

Phenotypic switching and mating:

C. albicans has the capacity to switch at high frequency and reversibly between different phenotypes. One of the well-characterized switching systems, first described in WO-1 strain, is the transition between "white and opaque" cell types (10). In this system, white cells and opaque cells grown at 25^{0}C can switch spontaneously at a low frequency of about 10^{-3}, and progeny cells can inherit the phenotype of their parents. Each phenotype is characterized by morphologically distinct forms. While white cells grow as budding yeast forms as most strains of *C. albicans*, opaque cells are larger, bean-shaped, and contain prominent protrusions called "pimples" on the cell surface. The molecular mechanism responsible for the white-opaque transition is not well understood, however it has been suggested to be regulated primarily at the level of transcription (87).

Most *C. albicans* strains do not display "white-opaque" switching, and appear to be locked in the white cell phase. Miller et al. (88) first observed that mating competent strains, **a** and α cells, constructed in a CAI4 (a SC5314 derivative) background could undergo "white-opaque" switching. Furthermore, when *MTL**a*** and *MTL*α cells are in the opaque cell phase, the mating frequency was dramatically increased by several orders

of magnitude as compared with both strains in the white cell phase. This indicates that a special developmental program (cell type or cell state) for mating (or sexual cycle) exists in *C. albicans*. Mating projections which mimic the well-characterized "shmoo" in *S. cerevisiae* have also been identified, however these seem to be restricted to **a** opaque cells; i.e. when **a** opaque cells and α opaque cells are mixed, only **a** cells will shmoo. This suggests that cell-cell signaling occurs in the process of mating, and further implies that a α–pheromone is expressed in *C. albicans*. In a large scale analysis of 220 clinical isolates of *C. albicans*, Lockhart et al. (83) have shown that nearly 97% of isolates were heterozygous at *MTL* locus, and 3% were homozygous at *MTL* locus. As observed in laboratory strains (CAI4 derivatives), most clinical strains are able to undergo white-opaque switching if homozygous at *MTL*. However, not all of the *MTL* homozygotes are capable of doing so, suggesting either that additional factors might contribute to this regulation or that there are variations among naturally occurring strains. Taken together, these results indicate that there is an interdependent relationship between mating and "white-opaque" phenotypic switching in *C. albicans* (89) (90).

Mating ritual in *C. albicans*

Players

So what makes opaque cells more competent than white cells for mating? Transcriptional profiling of the WO-1 strain (an α strain) reveals that several mating-associated genes, *STE3*, *MTLα1*, and mating factor-α are upregulated in opaque cells, suggesting that opaque cells are programmed for mating (11). There are also fundamental differences in the metabolic state of white cells and opaque cells. While opaque cells tend to utilize oxidative metabolism, white cells are fermentative (11). Whether these changes in metabolism are integral to the mating process is not known.

The opaque cell surface is also characteristically rough and covered with punctate protrusions called "pimples". In the EM the pimples appear as hollow cylinders on a pedestal of thickened cell wall. Opaque-specific antibody staining suggests that these pimples perform a secretory function, although the molecular nature of the secreted products is not known (91). Alternatively, these pimples might be involved in cell fusion or pheromone sensing during mating (90).

Pheromone

Genes which encode homologues of the *S. cerevisiae* pheromone receptor *STE2* and *STE3*, as well as pheromone processing enzymes such as *KEX2* have been identified in *C. albicans*, along with homologues of the MAP kinase pathway (92) (72-77, 93). Two independent groups have demonstrated that the MAP kinase pathway is essential for mating in *C. albicans*; deletion of *KEX2* also abrogates mating (72, 93). Using *in silico* methods Newport et al. analyzed the *C. albicans* genome for potential Kex2 substrates, finding among these a putative pheromone α-mating factor (94). However, an apparent **a**-mating pheromone homologue is not readily identified from sequence or structure analysis. When *C. albicans* opaque **a** cells are exposed to a synthetic α–pheromone, several key events of mating ensue. In the presence of the α–mating factor cell proliferation stops and "shmooing" is induced in **a** opaque cells, but not in **a** white cells, or in α white or opaque cells. The number of cells which undergo shmoo induction is dose dependent, and shmooing occurs only at 25 ^{0}C (95). Transcriptional analysis indicates that 62 genes are induced when **a** opaque cells are challenged with α– factor. Among these, 18 have close homologues in the mating pathway in *S. cerevisiae*. Eighteen genes are unique to *C. albicans*; seven genes encode proteins are either cell surface and secreted proteins, which have been implicated in virulence (96). These

results suggest an intimate relationship between mating and phenotypic switching, and also reveal a possible link between mating and pathogenesis in *C. albicans* (89, 90).

Details of mating (heterokaryon formation?)

Lockhart et al. (97) used continuous videomicroscopy, computer-assisted 3D re-construction and fluorescence microscopy to detail the mating process in *C. albicans*. While the initial steps of mating were very similar to those in *S. cerevisiae*, including shmooing, polarized cell growth (chemotropism) and cell fusion, the later events were quite different. Notably, after cell fusion the nuclei in *C. albicans* did not fuse, resulting in heterokaryon formation. This result is in distinction to that of Magee et al. (86) and Hull et al. (85) who showed that a single nucleus was formed after *C. albicans* mating using DAPI and FACS analysis. This result also contrasts with mating in *S. cerevisiae* where after plasma membrane fusion, the two parental nuclei move together and fusion of the nuclear envelops follows (98). In the Lockhart et al. study, the nuclei of the *C. albicans* mating cells were located in the tips of mating projections in the chemotropism stage. After cell fusion, the nuclei moved away from each other due to vacuole expansion. Asynchronous cell division of one nucleus followed, and the resulting zygotes contained 3 to 4 nuclei. After bud growth and maturation of daughter cells, the original fusion complex contained 2 or more nuclei. Southern blot analysis of daughter cells provided no evidence of mixing or exchange of mating type alleles or genetic markers in this study (97).

Heterokaryons are the product of cell fusion without subsequent chromosome loss or nuclear fusion. In *S. cerevisiae*, nuclear fusion during mating is a complex process, which is controlled by multiple genes such as *KAR1-5*, *CDC4* and *CDC37* (98). Mutation in any of these genes will cause the failure of karyogamy. When a *kar1-1* mutant and

KAR1 wild-type strain of *S. cerevisiae* are crossed, heterokaryons are formed and internuclear genetic exchange has been observed (99). Moreover, heterokaryons in yeasts are able to shuttle proteins and nucleic acids between nuclei (100). For filamentous ascomycetes, the heterokaryon provides the potential advantage of mitotic genetic exchange and functional diploidy within a parasexual cycle (101, 102). Interestingly, Tzung et al. first hypothesized hyphal fusion and heterokaryon formation in *C. albicans* from comparative genomic analysis. The identification of molecular components of the *het* loci suggests that these gene products might contribute to more complex sexual and asexual processes in *C. albicans* (92).

The "step by step" cytological description of mating in *C. albicans* provided by Lockhart et al. (97) differs in both process and mating products from that originally reported, where karyogamy was observed, and mating products were primarily mononuclear tetraploids. There are several possibilities for these contradictory observations. The main differences between these two studies are the conditions and strains of *C. albicans* used for mating. Hull et al. (85) and Magee et al. (86) used laboratory engineered strains which were derived from CAI4, a common laboratory strain. Lockhart et al. (97) used the WO-1 strain and some clinical isolates. Heterokaryon formation observed by Lockhard et al. may be the result of mating between nonisogenic strains, or the strains and conditions used by Hull et al. and Magee et al. may otherwise be strongly conducive to karyogamy. The incompatibility in nuclear behavior after mating is unaccounted for.

Parasexual cycle in *C. albicans*

As I have discussed, while there is ample evidence supporting mating in *C. albicans*, there is a discrepancy in the description of the products of mating. As *C. albicans* is primarily diploid, mating products need to return to the diploid state, whether from a

heterokaryon or a zygote. Previous studies have shown that tetraploid *C. albicans* strains might be unstable (103). Tetraploid fusion products from protoplast fusions can be induced to become diploid through UV irradiation or heat shock treatment (104, 105).

Bennett et al. (106) created genetically marked tetraploid *C. albicans* from **a** and α mating competent cells, and then investigated the role of environmental conditions and possible mechanisms on chromosome loss. While the tetraploid was quite stable, chromosome loss was observed when these cells were grown at 37^0C in *S. cerevisiae* pre-sporulation medium (1% yeast extract, 0.8% peptone, 10% glucose). The resultant progeny showed a spectrum of ploidy, but nearly 33% were diploid. The **a** and α cells regenerated through chromosome loss were as capable of mating as their parental cells. Genetic marker analysis indicated that the pattern of chromosome loss was concerted (mitotic), and there was no evidence that reductive division had occurred. Moreover, chromosome loss was shown to be independent of several meiotic gene homologues such as *SPO11*, *HOP1* or *DLH1* (*DMC1*) as determined by gene deletion experiments. This study showed that *C. albicans* is able to complete a diploid-tetraploid life cycle through chromosome loss.

***MTLa2* and evolution**

Tsong et al. (107) have used genomic profiling of *MTL* mutants combined with functional assays of mating behavior to investigate the regulatory circuit of *MTL*. Seven genes were identified as **a**1-α2 targets on the basis of 2-fold or greater increase in gene expression in white cells of *MTL* mutants. Two hundred thirty-seven genes were upregulated, and 197 genes were downregulated in opaque cells irrespective of *MTL* configuration. The identified set of white and opaque-specific genes overlaps with the

published transcriptional profiling of the WO-1 strain (an α strain), even with the substantial difference in their genetic background (11). No differences in gene expression were observed between **a** and α white cells, and **a**- specific or α-specific genes were identified by comparing white and opaque strains in different *MTL* configurations. Amongst these differences, *STE3* and *MFα1*, were α-specific, *i.e. MTLα1* is essential for their expression. However, **a**-specific genes were identified only when opaque **a** cells were challenged with pheromone α–factor. Further analysis suggests that **a**2 is the sole transcriptional regulator responsible for the pheromone response including both transcriptional differences and morphological changes. These **a**-specific genes include *STE6*, *RAM2* and others. Taken together, these data suggest that the mating-type circuit in *C. albicans* is as follows: Mtl**a**2 and Mtlα1 are the transcriptional activators for **a** and α genes, respectively. The combined action of Mtl**a**1 and Mtlα2 represses genes required for mating and phenotypic switching, either directly or indirectly. For the mating type circuit in *S. cerevisiae*, Matα1 is required for transcription of α-specific genes. Matα2 is a transcriptional repressor for **a**-specific genes, and Mata1-Matα2 is a repressor of haploid specific genes (108).

Since *C. albicans* ***a2*** is widely conserved among ascomycetes, one of the more intriguing questions arising from this study is why does **a**2 retain its regulatory role in *C. albicans*, but loses it in *S. cerevisiae*? *MAT**a**2* encodes a protein of unknown function, and the *MAT**a**2* null mutant shows no growth, mating or sporulation defect in *S. cerevisiae*. Interestingly, *MAT**a**2* is one of the slowly evolving genes in *Saccharomyces* species, perfectly conserved in amino acid and nucleotide level among *S. cerevisiae*, *S. paradoxus*, *S. mikatae* and *S. bayanus* (109). Mtl**a**2 in *C. albicans* appears to be a HMG-box DNA-binding protein, which is most homologous to MAT1-1-3 of *Cryphonectria*

parasitica. Comparative analyses of the *MAT* locus or *MTL* in 9 yeast species have revealed that several landmark events occurred in the mating system during evolution. These include the appearance of *HMR*/*HML* silent cassettes, followed by acquisition of *HO* endonuclease, then loss of the HMG domain gene (*MAT**a**2*) (81). *C. albicans* diverged from *S. cerevisiae* 100-800 million years ago (110), the difference in their mating type circuits reflect not only their respective environmental niches, but also how sexual regulators evolve on a phylogenetic scale.

A river runs through it

I was introduced to *C. albicans* several years ago in the course of preparing a journal club presentation, which dealt with the regulatory role of *mei2*$^{+}$ in *S. pombe* meiosis. Mei2 is an RNA-binding protein which controls the cell cycle switch from mitosis to meiosis in fission yeast (111). I was fascinated by how master regulators specify cell fate in different organisms. Although I learned that *C. albicans* is an asexual obligate diploid yeast, I was puzzled by the finding that *C. albicans* possesses a homologue of the meiosis specific gene *DMC1* (*DLH1* in *C. albicans*).

Coincidently, the *C. albicans* genome project had just begun, and I realized that there was an opportunity to address the question of *C. albicans* sexuality using, in part, a genomic approach. Accordingly, I first used bioinformatics to search for and identify possible meiotic gene homologues present in the *C. albicans* genome. Surprisingly, I found many homologues of meiotic regulatory genes. Moreover, I noticed that *IME1*, the master regulator for meiotic differentiation in *S. cerevisiae*, appeared to be absent in *C. albicans* (112). While a homologue of *UME6*, the interacting partner of *IME1* in regulating the initiation of the meiotic pathway in *S. cerevisiae*, was present in the *C. albicans* genome. Several downstream targets of *IME1*/*UME6*, such as *IME2*, *HOP1*,

NDT80 were also identified. Therefore, I hypothesized that the lack of a sexual cycle in *C. albicans* is due to the functional absence of master regulators for entry into meiosis. This could arise either from the loss of positive activators, constitutively inhibitory signals or changes in the environment niche which would through selection alter the genome content.

I began my thesis work with *UME6*, the functional partner of *IME1*, because of its central role in the meiotic pathway in *S. cerevisiae*. My original goal was to clone *C. albicans UME6*, then to identify a possible *IME1* homologue by heterologous complementation in *S. cerevisiae*. I did not start with using *S. cerevisiae IME1* to complement meiosis in *C. albicans*, primarily because I had no basis to predict the potential nutritional conditions which might be required for *C. albicans* to undergo meiosis.

Focusing on *UME6*, I asked the following questions: First, if *C. albicans UME6* is genetically knocked out, would the downstream targets of *UME6* such as *C. albicans IME2* be upregulated even in the rich growth medium (YEPD)? Second, URS1 represents a promoter element which is needed for mitotic repression and meiotic activation of early meiotic genes in *S. cerevisiae*. If *C. albicans UME6* is expressed in *S. cerevisiae*, would *C. albicans UME6* recognize the regulatory element URS1, and would *C. albicans UME6* interact with *IME1* when grown in sporulation medium? Third, if Ime1 can not convert *C. albicans* Ume6 from a repressor to an activator it would indicate that either there is no interaction between these two proteins or *C. albicans* Ume6 needs an Ime1 analogue for the interaction. Would the possible candidates for *IME1* analogue be identified by heterologous complementation in *S. cerevisiae* with a *C. albicans* genomic library?

In the following chapters of this thesis, I will lead you through the long winding journey which I have traveled as a graduate student in the past several years. First, I will present a comparative genomic analysis between *C. albicans* and *S. cerevisiae*, with special emphasis on examining whether there is a basis for postulating a sexual life cycle in *C. albicans*. I hypothesize that *C. albicans* might have a complete sexual cycle in nature from this genomic comparison, which is the first evidence suggesting that this organism is equipped with the capacity to undergo meiosis. Then I will present the effort I have made to dissect the potential regulatory circuitry for entering meiosis in *C. albicans*. I have demonstrated how a gene circuitry can change during evolution. Finally I will give you some of my thoughts regarding *C. albicans* and its potential to undergo a complete sexual cycle, and the conservation of the meiotic pathway across species.

References

1. Calderone, R. A. (2002) *Candida and candidiasis* (ASM Press, Washington, D.C.).
2. Odds, F. C. (1988) *Candida and candidosis* (Bailliáere Tindall, London ; Philadelphia).
3. Kwon-Chung, K. J. & Bennett, J. E. (1992) *Medical mycology* (Lea & Febiger, Philadelphia).
4. Wickes, B. L. & Petter, R. (1996) *Curr Top Med Mycol* **7,** 71-86.
5. Liu, H. (2001) *Curr Opin Microbiol* **4,** 728-35.
6. Galgoczy, J., Daroczy, J. & Simon, G. (1988) *Mycoses* **31,** 515-22.
7. Raudonis, B. M. & Smith, A. G. (1982) *Mycopathologia* **78,** 87-91.
8. Soll, D. R. (2002) *Acta Trop* **81,** 101-10.
9. Slutsky, B., Buffo, J. & Soll, D. R. (1985) *Science* **230,** 666-9.
10. Slutsky, B., Staebell, M., Anderson, J., Risen, L., Pfaller, M. & Soll, D. R. (1987) *J Bacteriol* **169,** 189-97.
11. Lan, C. Y., Newport, G., Murillo, L. A., Jones, T., Scherer, S., Davis, R. W. & Agabian, N. (2002) *Proc Natl Acad Sci U S A* **99,** 14907-12.
12. Perez-Martin, J., Uria, J. A. & Johnson, A. D. (1999) *Embo J* **18,** 2580-92.
13. Srikantha, T., Tsai, L., Daniels, K., Klar, A. J. & Soll, D. R. (2001) *J Bacteriol* **183,** 4614-25.
14. Zhao, R., Lockhart, S. R., Daniels, K. & Soll, D. R. (2002) *Eukaryot Cell* **1,** 353-65.
15. Janbon, G., Sherman, F. & Rustchenko, E. (1998) *Proc Natl Acad Sci U S A* **95,** 5150-5.
16. Perepnikhatka, V., Fischer, F. J., Niimi, M., Baker, R. A., Cannon, R. D., Wang, Y. K., Sherman, F. & Rustchenko, E. (1999) *J Bacteriol* **181,** 4041-9.

17. Bell, G. (1982) *The masterpiece of nature : the evolution and genetics of sexuality* (University of California Press, Berkeley).

18. Maynard Smith, J. (1978) *The evolution of sex* (Cambridge University Press, Cambridge [Eng.] ; New York).

19. Judson, O. P. & Normark, B. B. (1996) *Trends in Ecology & Evolution* **11,** 41-46.

20. Barton, N. H. & Charlesworth, B. (1998) *Science* **281,** 1986-90.

21. Butlin, R. (2002) *Nat Rev Genet* **3,** 311-7.

22. Ayala, F. J. (1998) *Proc Natl Acad Sci U S A* **95,** 3346-8.

23. Bernstein, H., Hopf, F. A. & Michod, R. E. (1987) *Adv Genet* **24,** 323-70.

24. Haigh, J. (1978) *Theor Popul Biol* **14,** 251-67.

25. Muller, H. J. (1964) *Mutat Res* **106,** 2-9.

26. Crow, J. F. (1994) *Dev Genet* **15,** 205-13.

27. Kohli, Y. & Kohn, L. M. (1998) *Fungal Genet Biol* **23,** 139-49.

28. Chen, R. S. & McDonald, B. A. (1996) *Genetics* **142,** 1119-27.

29. Zhan, J., Pettway, R. E. & McDonald, B. A. (2003) *Fungal Genet Biol* **38,** 286-97.

30. Milgroom, M. G., Lipari, S. E. & Powell, W. A. (1992) *Genetics* **131,** 297-306.

31. Burt, A., Carter, D. A., Koenig, G. L., White, T. J. & Taylor, J. W. (1996) *Proc Natl Acad Sci U S A* **93,** 770-3.

32. Graser, Y., Volovsek, M., Arrington, J., Schonian, G., Presber, W., Mitchell, T. G. & Vilgalys, R. (1996) *Proc Natl Acad Sci U S A* **93,** 12473-7.

33. Litvintseva, A. P., Marra, R. E., Nielsen, K., Heitman, J., Vilgalys, R. & Mitchell, T. G. (2003) *Eukaryot Cell* **2,** 1162-8.

34. Andrade, L. O., Machado, C. R., Chiari, E., Pena, S. D. & Macedo, A. M. (2002) *Exp Parasitol* **100,** 269-75.

35. Prata, A. (2001) *Lancet Infect Dis* **1,** 92-100.

36. Su, C., Evans, D., Cole, R. H., Kissinger, J. C., Ajioka, J. W. & Sibley, L. D. (2003) *Science* **299,** 414-6.

37. Grigg, M. E., Bonnefoy, S., Hehl, A. B., Suzuki, Y. & Boothroyd, J. C. (2001) *Science* **294,** 161-5.

38. Wickes, B. L., Mayorga, M. E., Edman, U. & Edman, J. C. (1996) *Proc Natl Acad Sci U S A* **93,** 7327-31.

39. Nielsen, K., Cox, G. M., Wang, P., Toffaletti, D. L., Perfect, J. R. & Heitman, J. (2003) *Infect Immun* **71,** 4831-41.

40. van der Walt, J. P. (1967) *Antonie Van Leeuwenhoek* **33,** 246-56.

41. Whelan, W. L. & Magee, P. T. (1981) *J Bacteriol* **145,** 896-903.

42. Whelan, W. L. & Soll, D. R. (1982) *Mol Gen Genet* **187,** 477-85.

43. Tibayrenc, M., Kjellberg, F. & Ayala, F. J. (1990) *Proc Natl Acad Sci U S A* **87,** 2414-8.

44. Tibayrenc, M. & Ayala, F. J. (2002) *Trends Parasitol* **18,** 405-10.

45. Tibayrenc, M., Kjellberg, F., Arnaud, J., Oury, B., Breniere, S. F., Darde, M. L. & Ayala, F. J. (1991) *Proc Natl Acad Sci U S A* **88,** 5129-33.

46. Tibayrenc, M. (1997) *Trends Microbiol* **5,** 253-4; discussion 254-7.

47. Pujol, C., Reynes, J., Renaud, F., Raymond, M., Tibayrenc, M., Ayala, F. J., Janbon, F., Mallie, M. & Bastide, J. M. (1993) *Proc Natl Acad Sci U S A* **90,** 9456-9.

48. Forche, A., Schonian, G., Graser, Y., Vilgalys, R. & Mitchell, T. G. (1999) *Fungal Genet Biol* **28,** 107-25.

49. Smith, J. M., Smith, N. H., O'Rourke, M. & Spratt, B. G. (1993) *Proc Natl Acad Sci U S A* **90,** 4384-8.

50. Smit, A. F. (1999) *Curr Opin Genet Dev* **9,** 657-63.

51. Daboussi, M. J. & Capy, P. (2003) *Annu Rev Microbiol* **57,** 275-99.

52. Deragon, J. M. & Capy, P. (2000) *Ann Med* **32,** 264-73.

53. Kim, J. M., Vanguri, S., Boeke, J. D., Gabriel, A. & Voytas, D. F. (1998) *Genome Res* **8,** 464-78.

54. Bestor, T. H. (1998) *Novartis Found Symp* **214,** 187-95; discussion 195-9, 228-32.

55. Bestor, T. H. & Tycko, B. (1996) *Nat Genet* **12,** 363-7.

56. Sandmeyer, S. (1998) *Genome Res* **8,** 416-8.

57. Boeke, J. D. & Devine, S. E. (1998) *Cell* **93,** 1087-9.

58. Bestor, T. H. (2003) *Trends Genet* **19,** 185-90.

59. Doolittle, W. F. & Sapienza, C. (1980) *Nature* **284,** 601-3.

60. Hickey, D. A. (1982) *Genetics* **101,** 519-31.

61. Bestor, T. H. (1999) *Genetica* **107,** 289-95.

62. Zeyl, C., Bell, G. & Green, D. M. (1996) *Genetics* **143,** 1567-77.

63. Arkhipova, I. & Meselson, M. (2000) *Proc Natl Acad Sci U S A* **97,** 14473-7.

64. Moran, J. V., Zimmerly, S., Eskes, R., Kennell, J. C., Lambowitz, A. M., Butow, R. A. & Perlman, P. S. (1995) *Mol Cell Biol* **15,** 2828-38.

65. Bowen, N. J., Jordan, I. K., Epstein, J. A., Wood, V. & Levin, H. L. (2003) *Genome Res* **13,** 1984-97.

66. Goodwin, T. J. & Poulter, R. T. (2000) *Genome Res* **10,** 174-91.

67. Goodwin, T. J. & Poulter, R. T. (2001) *Yeast* **18,** 865-80.

68. Neuveglise, C., Feldmann, H., Bon, E., Gaillardin, C. & Casaregola, S. (2002) *Genome Res* **12,** 930-43.

69. Sadhu, C., Hoekstra, D., McEachern, M. J., Reed, S. I. & Hicks, J. B. (1992) *Mol Cell Biol* **12,** 1977-85.

70. Haber, J. E. (1992) *Trends Genet* **8,** 446-52.

71. Soll, D. R. (1990) in *The genetics of Candida*, eds. Kirsch, D. R., Kelly, R. & Kurtz, M. B. (CRC press, pp. 147-176.

72. Chen, J., Lane, S. & Liu, H. (2002) *Mol Microbiol* **46,** 1335-44.

73. Csank, C., Schroppel, K., Leberer, E., Harcus, D., Mohamed, O., Meloche, S., Thomas, D. Y. & Whiteway, M. (1998) *Infect Immun* **66,** 2713-21.

74. Kohler, J. R. & Fink, G. R. (1996) *Proc Natl Acad Sci U S A* **93,** 13223-8.

75. Leberer, E., Harcus, D., Broadbent, I. D., Clark, K. L., Dignard, D., Ziegelbauer, K., Schmidt, A., Gow, N. A., Brown, A. J. & Thomas, D. Y. (1996) *Proc Natl Acad Sci U S A* **93,** 13217-22.

76. Liu, H., Kohler, J. & Fink, G. R. (1994) *Science* **266,** 1723-6.

77. Newport, G. & Agabian, N. (1997) *J Biol Chem* **272,** 28954-61.

78. Raymond, M., Dignard, D., Alarco, A. M., Mainville, N., Magee, B. B. & Thomas, D. Y. (1998) *Mol Microbiol* **27,** 587-98.

79. Diener, A. C. & Fink, G. R. (1996) *Genetics* **143,** 769-76.

80. Hull, C. M. & Johnson, A. D. (1999) *Science* **285,** 1271-5.

81. Butler, G., Kenny, C., Fagan, A., Kurischko, C., Gaillardin, C. & Wolfe, K. H. (2004) *Proc Natl Acad Sci U S A* **101,** 1632-7.

82. Kostriken, R., Strathern, J. N., Klar, A. J., Hicks, J. B. & Heffron, F. (1983) *Cell* **35,** 167-74.

83. Lockhart, S. R., Pujol, C., Daniels, K. J., Miller, M. G., Johnson, A. D., Pfaller, M. A. & Soll, D. R. (2002) *Genetics* **162,** 737-45.

84. Herskowitz, I. (1989) *Nature* **342,** 749-57.

85. Hull, C. M., Raisner, R. M. & Johnson, A. D. (2000) *Science* **289,** 307-10.

86. Magee, B. B. & Magee, P. T. (2000) *Science* **289,** 310-3.

87. Soll, D. R. (1997) *Microbiology* **143 (Pt 2),** 279-88.

88. Miller, M. G. & Johnson, A. D. (2002) *Cell* **110,** 293-302.

89. Soll, D. R. (2004) *Bioessays* **26,** 10-20.

90. Johnson, A. D. (2003) *Nat Rev Microbiol* **1,** 106-116.

91. Anderson, J., Mihalik, R. & Soll, D. R. (1990) *J Bacteriol* **172,** 224-35.

92. Tzung, K. W., Williams, R. M., Scherer, S., Federspiel, N., Jones, T., Hansen, N., Bivolarevic, V., Huizar, L., Komp, C., Surzycki, R., Tamse, R., Davis, R. W. & Agabian, N. (2001) *Proc Natl Acad Sci U S A* **98,** 3249-53.

93. Magee, B. B., Legrand, M., Alarco, A. M., Raymond, M. & Magee, P. T. (2002) *Mol Microbiol* **46,** 1345-51.

94. Newport, G., Kuo, A., Flattery, A., Gill, C., Blake, J. J., Kurtz, M. B., Abruzzo, G. K. & Agabian, N. (2003) *J Biol Chem* **278,** 1713-20.

95. Lockhart, S. R., Zhao, R., Daniels, K. J. & Soll, D. R. (2003) *Eukaryot Cell* **2,** 847-55.

96. Bennett, R. J., Uhl, M. A., Miller, M. G. & Johnson, A. D. (2003) *Mol Cell Biol* **23,** 8189-201.

97. Lockhart, S. R., Daniels, K. J., Zhao, R., Wessels, D. & Soll, D. R. (2003) *Eukaryot Cell* **2,** 49-61.

98. Marsh, L. & Rose, M. D. (1997) in *The molecular and cellular biology of the yeast Saccharomyces ,cell cycle and cell biology*, eds. Pringle, J. R., Broach, J. R. & Jones, E. W. (Cold spring harbor laboratory press, Vol. 3, pp. 827-888.

99. Dutcher, S. K. (1981) *Mol Cell Biol* **1,** 245-53.

100. Olson, B. L. & Siliciano, P. G. (2003) *Yeast* **20,** 893-903.

101. Glass, N. L. & Kaneko, I. (2003) *Eukaryot Cell* **2,** 1-8.

102. Pontecorvo, G. (1956) *Annu Rev Microbiol* **10,** 393-400.

103. Whelan, W. L., Markie, D. M., Simpkin, K. G. & Poulter, R. M. (1985) *J Bacteriol* **161,** 1131-6.

104. Poulter, R. T. & Rikkerink, E. H. (1983) *J Bacteriol* **156,** 1066-77.

105. Hilton, C., Markie, D., Corner, B., Rikkerink, E. & Poulter, R. (1985) *Mol Gen Genet* **200,** 162-8.

106. Bennett, R. J. & Johnson, A. D. (2003) *Embo J* **22,** 2505-15.

107. Tsong, A. E., Miller, M. G., Raisner, R. M. & Johnson, A. D. (2003) *Cell* **115,** 389-99.

108. Johnson, A. D. (1995) *Curr Opin Genet Dev* **5,** 552-8.

109. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. (2003) *Nature* **423,** 241-54.

110. Heckman, D. S., Geiser, D. M., Eidell, B. R., Stauffer, R. L., Kardos, N. L. & Hedges, S. B. (2001) *Science* **293,** 1129-33.

111. Watanabe, Y., Shinozaki-Yabana, S., Chikashige, Y., Hiraoka, Y. & Yamamoto, M. (1997) *Nature* **386,** 187-90.

112. Kassir, Y., Granot, D. & Simchen, G. (1988) *Cell* **52,** 853-62.

Chapter 2

1. **The awesome power of comparative genomics**

2. **Genomic evidence for a complete sexual cycle in *Candida albicans***

The awesome power of comparative genomics

Understanding the book of life begins with comparative genomics

A complete genome sequence of an organism provides a window through which we may visualize the landscape of its genetic traits. With the number of sequenced genomes increasing daily, the prospect of comparative genomic analyses between phylogenetically disparate species promises to reveal invaluable information regarding evolutionary changes in genome size, contents and complexity, metabolic diversity and the adaptive ecology of organisms. Genomic comparisons between model organisms and other less studied species will facilitate annotation of DNA sequences in the latter and provide a genomic basis for predicting biological behavior. Already comparisons between the genomes of several *Saccaharomyces* species have identified conserved regulatory elements which are important for gene expression (1, 2). While *Caenorhabditis briggsae* and *Caenorhabditis elegans* share similar morphology and developmental processes, genomic analysis demonstrates that only about 60% of *C. elegans* genes have orthologues in *C. briggsae* (3). This suggests that the genomes of both organisms diverged significantly 80-110 million years ago, however their developmental programs remain conserved, implying that compensatory and coevolutionary changes in genes in the same regulatory pathway have occurred between these two organisms (3, 4). Comparing genome sequences from multiple closely-related species or distantly-related organisms should shed light on understanding functional and evolutionary aspects of organism biology. For instance, sequence comparisons in one targeted genomic region (orthologue of a 1.8 Mb segment on human chromosome 7) between 12 diverse species ranging from fish to human has disclosed functional conservation and mutational events such as transposon insertions and variation in genome expansion or compression among organisms (5). In particular, the identification of conserved non-coding segments in this region, not easily detectable

by pair-wise comparisons alone, suggests the power of comparison using more genomic sequences.

The genomes of several fungal species have now been fully sequenced and annotated, including *S. cerevisiae*, *Schizosaccharomyces pombe*, *Candida albicans* and *Neurospora crassa*. In addition, genomes of 13 hemiascomycetes have been partially sequenced. A Fungal Genome Initiative (FGI) launched in the Whitehead Center for Genome Research has targeted species of *Candida*, *Coccidioides*, *Histoplasma* and *Cryptococcus* for further sequencing. Genomic comparisons between these important agents of human diseases are now contributing to the emerging discipline of pathogenomics.

A whole genome shotgun sequencing project of *C. albicans* strain SC5314 was initiated at the Stanford Genome Technology Center (SGTC) in late 1996. The goal was not only to define a complete haploid genome for this apparently obligate diploid organism, but also to generate high quality DNA sequences to explore allelic variation. Adding further to the sequencing complexity and genome assembly, *C. albicans* displays considerable heterozygosity. The immediate challenge was to assemble the eight pairs of chromosomes where homologues have varying degrees of similarity. Based on a haploid assembly with 10.4X sequence coverage, a diploid *C. albicans* genome has been constructed which is in excellent agreement with available physical mapping data. The latest release of Assembly 19 provides a platform for studying *Candida* biology and pathogenesis (6).

What can we learn from comparative genomics?

S. cerevisiae is one of the best-studied experimental organisms. As such its genome database provides a rich resource, especially for comparative analysis with the *C. albicans* genome (7, 8). Several caveats have been suggested when gene discovery is the primary goal of genome comparisons. First, the absence of a particular gene in a completely sequenced genome does not necessarily suggest the loss of such function in that organism. For example, BLAST analysis between *C. albicans* and *S. cerevisiae* indicates that the *S. cerevisiae URA1* homologue is absent in *C. albicans*. Nevertheless, *C. albicans* does contain a *URA1* homologue, but one which is more closely related to that in *S. pombe*. The *URA1* in *S. cerevisiae* is cytoplasmic, and is necessary for biosynthesis of pyrimidine nucleotides, while *C. albicans* and *S. pombe URA1* are mitochondrial, indicating fundamental differences in the cellular pathway for performing this function in these yeasts (9, 10). Second, *C. albicans* diverged from *S. cervisiae* at least 100 million years ago, and some of its genes do not recapitulate its phylogenetic lineage as assigned through rRNA sequence. For instance, *C. albicans PHO85* is more related to that of *Aspergillus nidulans* than *S. cerevisiae* (11). The closest homologue of the *C. albicans MTLa2* is the *MAT1-1-3* of *Cryphonectria parasitica* (12), while other ORFs in the *MTL* of *C. albicans* are more closely related to those of *S. cerevisiae* (13), reinforcing the evolutionary distance between *S. cerevisiae* and *C. albicans*.

C. albicans has 8 pairs of chromosomes, and its haploid genome is about 16 Mb in size. The sequenced genome contains 6419 annotated ORFs (>100 a.a); amongst these are several large gene families some of which have been directly correctly with virulence (6). The *ALS* (agglutinin-like sequence) family contains 8 members, 10 genes comprise the *SAP* (secreted aspartyl proteinase) gene family, and 10 genes are in the lipase gene family. In addition, *C. albicans* also has 8 estrogen binding proteins, 4 acid

sphingomyelinases and 12 cyctochrome P450s. The *C. albicans* genome contains a higher percentage of ORFs of unassigned functions compared with *S. cerevisiae* or *S. pombe*, suggesting a repertoire of genes which may be important in supporting the pathogenic/ opportunistic natural history of *C. albicans* (14). Although *C. albicans* and *S. cerevisiae* are evolutionarily distant, 64% of the *C. albicans* ORFs have their best match in *S. cerevisiae*. Given that these two fungi differ in a wide range of characteristics such as habitats, pathogenicity, sexuality, hyphal formation and phenotypic switching, genomic comparisons which emphasize the differences between them will provide a platform for exploring the unique cellular process in each organism.

C. albicans in the era of comparative genomics

To date, most genome comparisons have by necessity focused on datasets derived from organisms separated by considerable phylogenetic distances. More recently, genomic comparisons between more closely related species of *Saccharomyes*, between mouse and human, and between *C. elegans* and *C. briggsae* underscore the importance of comparing sequence changes across smaller phylogenetic distances or comparing similar genomes which are under positive selection (15). Therefore the most powerful approach for understanding *Candida* biology will come from analyzing sequence differences between *C. albicans* strains and closely related *Candida* species especially those which are less or non-pathogenic.

When I began my thesis work, I was in the unique position of having access to both the *C. albicans* genome sequence in its earliest incarnations and the public database for *S. cerevisiae* (SGD). This allowed me to begin a comparative genomic analysis between these fungi with the objective of learning what might be the genetic basis for the apparent asexual nature of *C. albicans*, then thought to be an obligate diploid organism.

Surprisingly the results of my study revealed that *C. albicans* contains many potential homologues of genes which are found in the meiotic machinery defined in *S. cerevisiae*, albeit with notable differences and omissions. Moreover, I discovered that the *C. albicans* genome also encodes homologues of genes found in other organisms which are important for different aspects of meiosis, suggesting that these may play some compensatory roles in the process otherwise defined in *S. cerevisiae*. Taken together, I am convinced that these results suggest that *C. albicans* contains a genetic repertoire sufficient to support a complete sexual cycle (16).

Since *C. albicans* diverged from *S. cerevisiae* nearly 100-800 million years ago (Figure 1) (17, 18), gene order is substantially different between these two yeasts with significant loss of synteny (19, 20). One key caveat to be considered is that *S. cerevisiae* might not be the best candidate organism for genomic comparison with *C. albicans* because of their evolutionary distance and different biological habitats. Among the 21 *Candida* species classified as human pathogens, 13 have been reported to reproduce sexually (21). Several haploid species related to *C. albicans* including *C. lusitaniae*, *C. guilliermondii*, *C. krusei*, have a defined sexual cycle (Figure 2). In particular, genetic tools and yeast strains for studying the sexual cycle in *C. lusitaniae* have been established (22). Also, another two species closely related to *C. albicans* which have been proposed to be sequenced by the Fungal Genome Initiative (FGI) are *C. tropicalis* and *Lodderomyces elongisporus*. The asexual diploid *C. tropicalis* is usually associated with deep fungal infection, while *L. elongisporus* is the only known ascosporogenous species in the *Candida* clade (23, 24). Comparative genomic analysis between these various *Candida* species in the core meiotic machinery will lend insight into how *C. albicans* might complete its sexual cycle or explain how it may have lost the capacity to undergo meiosis.

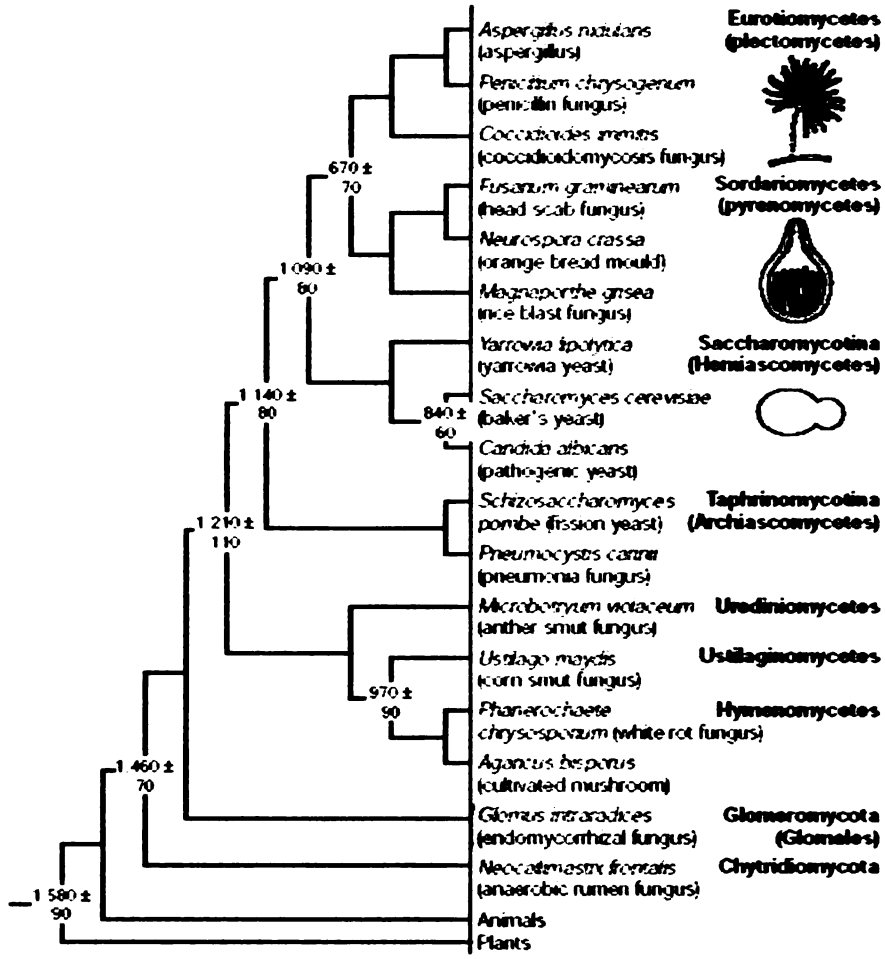


Figure 1. A phylogenetic tree of fungi.

The phylogeny of selected fungi are shown. The divergence times (millions of years ago) were estimated assuming an equal rate of substitution in fungi and animals (Hedges, 2002)

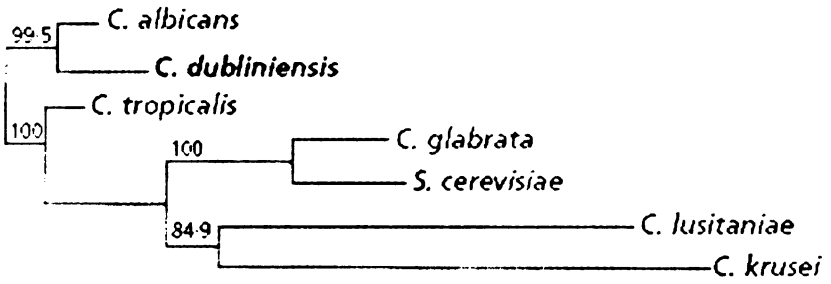


Figure 2. Phylogenetic relationship among *Candida* species
(Gilfillan et al.1998)

An unrooted phylogenetic neighbor joining tree was generated from the alignment of rRNA sequences from *C. albicans*, *S. cerevisiae* and other non-*Candida albicans Candida* species.

References

1. Kellis, M., Patterson, N., Endrizzi, M., Birren, B. & Lander, E. S. (2003) *Nature* **423,** 241-54.

2. Cliften, P., Sudarsanam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B. A. & Johnston, M. (2003) *Science* **301,** 71-6.

3. Stein, L. D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M. R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D. H., Fulton, L. A., Fulton, R. E., Griffiths-Jones, S., Harris, T. W., Hillier, L. W., Kamath, R., Kuwabara, P. E., Mardis, E. R., Marra, M. A., Miner, T. L., Minx, P., Mullikin, J. C., Plumb, R. W., Rogers, J., Schein, J. E., Sohrmann, M., Spieth, J., Stajich, J. E., Wei, C., Willey, D., Wilson, R. K., Durbin, R. & Waterston, R. H. (2003) *PLoS Biol* **1,** E45.

4. Blaxter, M. (2003) *Nature* **426,** 395-6.

5. Thomas, J. W., Touchman, J. W., Blakesley, R. W., Bouffard, G. G., Beckstrom-Sternberg, S. M., Margulies, E. H., Blanchette, M., Siepel, A. C., Thomas, P. J., McDowell, J. C., Maskeri, B., Hansen, N. F., Schwartz, M. S., Weber, R. J., Kent, W. J., Karolchik, D., Bruen, T. C., Bevan, R., Cutler, D. J., Schwartz, S., Elnitski, L., Idol, J. R., Prasad, A. B., Lee-Lin, S. Q., Maduro, V. V., Summers, T. J., Portnoy, M. E., Dietrich, N. L., Akhter, N., Ayele, K., Benjamin, B., Cariaga, K., Brinkley, C. P., Brooks, S. Y., Granite, S., Guan, X., Gupta, J., Haghighi, P., Ho, S. L., Huang, M. C., Karlins, E., Laric, P. L., Legaspi, R., Lim, M. J., Maduro, Q. L., Masiello, C. A., Mastrian, S. D., McCloskey, J. C., Pearson, R., Stantripop, S., Tiongson, E. E., Tran, J. T., Tsurgeon, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L. S., Young, A. C., Zhang, L. H., Osoegawa, K., Zhu, B., Zhao, B., Shu, C. L., De Jong, P. J., Lawrence, C. E., Smit, A. F., Chakravarti, A., Haussler, D., Green, P., Miller, W. & Green, E. D. (2003) *Nature* **424,** 788-93.

6. Jones, T. F., Federspiel, N.A., Chibana, H. Dungan, J., Kalman, K., Magee, B.B. , Newport, G., Thorstenson, Y.R., Agabian, N., Magee, P.T., Davis, R.W. & Scherer, S. (2004) *Proc Natl Acad Sci U S A*, in press.

7. Goffeau, A., Barrell, B. G., Bussey, H., Davis, R. W., Dujon, B., Feldmann, H., Galibert, F., Hoheisel, J. D., Jacq, C., Johnston, M., Louis, E. J., Mewes, H. W., Murakami, Y., Philippsen, P., Tettelin, H. & Oliver, S. G. (1996) *Science* **274,** 546, 563-7.

8. Rubin, G. M., Yandell, M. D., Wortman, J. R., Gabor Miklos, G. L., Nelson, C. R., Hariharan, I. K., Fortini, M. E., Li, P. W., Apweiler, R., Fleischmann, W., Cherry, J. M., Henikoff, S., Skupski, M. P., Misra, S., Ashburner, M., Birney, E., Boguski, M. S., Brody, T., Brokstein, P., Celniker, S. E., Chervitz, S. A., Coates, D., Cravchik, A., Gabrielian, A., Galle, R. F., Gelbart, W. M., George, R. A., Goldstein, L. S., Gong, F., Guan, P., Harris, N. L., Hay, B. A., Hoskins, R. A., Li, J., Li, Z., Hynes, R. O., Jones, S. J., Kuehl, P. M., Lemaitre, B., Littleton, J. T., Morrison, D. K., Mungall, C., O'Farrell, P. H., Pickeral, O. K., Shue, C., Vosshall, L. B., Zhang, J., Zhao, Q., Zheng, X. H. & Lewis, S. (2000) *Science* **287,** 2204-15.

9. Scherer, S. (2002) in *Candida and Candidiasis*, ed. Calderone, R. A. (ASM press, Washington, D.C.), pp. 259-265.

10. Vorisek, J., Technikova, Z., Schwippel, J. & Benoist, P. (2002) *Yeast* **19,** 449-57.

11. Miyakawa, Y. (2000) *Yeast* **16,** 1045-51.

12. Tsong, A. E., Miller, M. G., Raisner, R. M. & Johnson, A. D. (2003) *Cell* **115,** 389-99.

13. Hull, C. M. & Johnson, A. D. (1999) *Science* **285,** 1271-5.

14. Herrero, E., de la Torre, M. A. & Valentin, E. (2003) *Int Microbiol* **6,** 183-90.

15. Hardison, R. C. (2003) *PLoS Biol* **1,** E58.

16. Tzung, K. W., Williams, R. M., Scherer, S., Federspiel, N., Jones, T., Hansen, N., Bivolarevic, V., Huizar, L., Komp, C., Surzycki, R., Tamse, R., Davis, R. W. & Agabian, N. (2001) *Proc Natl Acad Sci U S A* **98,** 3249-53.

17. Heckman, D. S., Geiser, D. M., Eidell, B. R., Stauffer, R. L., Kardos, N. L. & Hedges, S. B. (2001) *Science* **293,** 1129-33.

18. Gargas, A., DePriest, P. T., Grube, M. & Tehler, A. (1995) *Science* **268,** 1492-5.

19. Wong, S., Butler, G. & Wolfe, K. H. (2002) *Proc Natl Acad Sci U S A* **99,** 9272-7.

20. Seoighe, C., Federspiel, N., Jones, T., Hansen, N., Bivolarovic, V., Surzycki, R., Tamse, R., Komp, C., Huizar, L., Davis, R. W., Scherer, S., Tait, E., Shaw, D. J., Harris, D., Murphy, L., Oliver, K., Taylor, K., Rajandream, M. A., Barrell, B. G. & Wolfe, K. H. (2000) *Proc Natl Acad Sci U S A* **97,** 14433-7.

21. Hazen, K. C. (1995) *Clin Microbiol Rev* **8,** 462-78.

22. Young, L. Y., Lorenz, M. C. & Heitman, J. (2000) *Genetics* **155,** 17-29.

23. Munoz, P., Burillo, A. & Bouza, E. (2000) *Int J Antimicrob Agents* **15,** 83-90.

24. James, S. A., Collins, M. D. & Roberts, I. N. (1994) *Lett Appl Microbiol* **19,** 308-11.

Genomic evidence for a complete sexual cycle in *Candida albicans*

The original paper was published in *PNAS* as follows:

Tzung KW, Williams RM, Scherer S, Federspiel N, Jones T, Hansen N, Bivolarevic V, Huizar L, Komp C, Surzycki R, Tamse R, Davis RW, Agabian N. Genomic evidence for a complete sexual cycle in Candida albicans. Proc Natl Acad Sci U S A. 2001 Mar 13; 98(6):3249-53.

Keh-Weei Tzung was responsible for analyzing the genome information provided by Stanford Genome Technology Center and *S. cerevisiae* microarray data was provided by R.W. Davis lab at Stanford. Keh-Weei Tzung wrote the paper.

This section is reprinted with permission from *PNAS*.

ABSTRACT

Candida albicans is a diploid fungus which has become a medically important opportunistic pathogen in immunocompromised individuals. We have sequenced the *C. albicans* genome to 10.4-fold coverage and performed a comparative genomic analysis between *C. albicans* and *Saccharomyces cerevisiae* with the objective of assessing whether or not *Candida* possesses a genetic repertoire which could support a complete sexual cycle. Analyzing over 500 genes important for sexual differentiation in *S. cerevisiae*, we find many homologues of genes which are implicated in the initiation of meiosis, chromosome recombination and the formation of synaptonemal complexes. However, others are striking in their absence. *C. albicans* seems to have homologues of all the elements of a functional pheromone response pathway involved in mating in *S. cerevisiae*, but lacks many homologues of *S. cerevisiae* genes for meiosis. Other meiotic gene homologues in organisms ranging from filamentous fungi to *Drosophila melanogaster* and *Caenorhabditis elegans* were also found in the *C. albicans* genome, suggesting potential alternative mechanisms of genetic exchange.

INTRODUCTION

Meiosis represents a specialized cell division that is essential for sexual reproduction; it generates haploid germ cells from diploid parental cells (1). Because a sexual phase for *Candida albicans* has historically not been detected, it is classified among the *Fungi imperfecti* (2). However, the identification of a mating-type-like (*MTL*) locus and genes such as *CPH1*, *CAG1*, *DLH1*, *NDT80* and *HST6* in *C. albicans* (3), which participate in meiotic differentiation in *S. cerevisiae*, suggests that the classification of this diploid fungus belies the existence of a sexual cycle. Recently in fact, genetic manipulation of the *MTL* locus resulted in the demonstration that *C. albicans* strains can mate to produce triploid or tetraploid progeny at very low frequency either in culture or in experimental animals (4, 5). It thus appears that *Candida* can undergo cell fusion, dependent upon mating type. However, completion of a sexual cycle --i.e. meiosis and sporulation remains to be demonstrated.

C. albicans is an opportunistic pathogen which can cause disease in patients immunocompromised as a result of HIV infection, organ transplantation, and cancer chemotherapy (6). It is also a morphologically complex organism capable of proliferating either as a budding yeast or by the formation of pseudohyphae or filamentous hyphae. The inability to demonstrate a sexual cycle has significantly impeded conventional genetic analysis. Therefore the potential for its existence has both intrinsic and technical consequences.

Shotgun sequencing of the diploid *C. albicans* genome, undertaken by the Stanford Genome Technology Center, is complete with the sequencing of 10.4 haploid genome equivalents, which is sufficient to insure identification of all of the genes in this organism. A web page and database has been made available over the World Wide Web (http://sequence-www.stanford.edu/group/candida). Given the evolutionary proximity

between *C. albicans* and *S. cerevisiae* (7) and the differences in their virulence and habitat, genomic comparisons between these fungi are likely to illuminate aspects of the unique cell biology of both organisms. In this report, we address the potential for meiosis and sexual recombination in *C. albicans*.

MATERIALS and METHODS

Sequencing library construction, shotgun sequencing, assembly and analysis

C. albicans strain SC5314 was generously provided by Bristol-Meyers Squibb for use without restrictions. DNA from this strain was used for all M13 and plasmid sequencing library constructions. Electrocompetent *E. coli* DH12S and DH10B (Gibco/BRL Life Technologies) were used for transformation of M13 and plasmid libraries, respectively. *C. albicans* cultures were grown in YPD broth with shaking at 30°C. Spheroplasts were made, treated with SDS/proteinase K, and the DNA purified on a sucrose gradient. The purified *C. albicans* SC5314 genomic DNA was sheared by a point-sink shearing device (8) to a fragment size of 3-6 kb for cloning into the plasmid vector pUC119 and to 1-2 kb for cloning into M13mp18. Individual plasmid and M13 clones were picked, grown, and template DNA prepared using the automated instrumentation developed at Stanford (9). Sequencing reactions were performed on plasmid and M13 template DNA using BigDye Primer and BigDye Terminator kits from PE Applied Biosystems (Foster City, CA) according to manufacturer's specifications, with slight modifications. Sequencing analysis was performed on ABI377-XL Automated Sequencers at 96 lanes/gel. Sequence data were derived using the basecaller *phred* (10, 11), assembled with *phrap* (Phil Green, U. Washington), and viewed with *consed* (12).

Identification of the meiotic machinery in *C. albicans*

We have used several approaches to identify meiotic homologues in *C. albicans,* focusing mainly on genes critical for mating and meiosis in budding yeast. We also searched for homologues of genes which are regulators of cellular reproduction in other organisms. The *S. cerevisiae* and other sequences were used as query sequences in basic BLAST searches (13); using the blastp program, we searched for similar proteins in the *C. albicans* genome database, version 6 assembly. The blast output was sorted, and top hits ranked by blast scores. The E value cutoff used to assign homologues was $1e^{-6}$. However, each sequence required specific evaluation as there were exceptional instances where biological data indicated the presence of functional homologues although these lacked significant sequence homology. In Table 1 (which is published as supplemental data on the PNAS website, www.pnas.org) the functional groups used in the categories roughly correspond to: (1) mating differentiation (2) nutritional control (3) cell type control (4) initiation of meiosis (5) checkpoint control and progression through meiosis (6) recombination and the formation of synaptonemal complexes and (7) spore wall morphogenesis and ascus formation.

RESULTS and DISCUSSION

Using comprehensive genomic comparisons, we have assessed the repertoire of gene homologues in *C. albicans* which in *S. cerevisiae* are required in pathways leading to sexual differentiation. In a genome-wide transcription analysis of sporulation in *S. cerevisiae* (14), nearly 500 genes were expressed; those of particular interest are genes which are either meiosis-specific or have shown meiotic mutant phenotypes. Using these 500 genes as a reference point, *C. albicans* homologues were identified based on their sequence similarity with *S. cerevisiae* counterparts from 10.4x sequencing data (see materials & methods). Genes analyzed in these comparisons are listed in Table 1 and categorized based on their presumptive function in meiosis and sporulation. In this study, homologues of *S. cerevisiae*

genes involved in chromosome recombination and the formation of synaptonemal complexes (SC) were conspicuously absent from *C. albicans*, while other groups of genes important for meiosis, mating and sporulation contained many candidates in the *C. albicans* genome.

Mating differentiation:

In *S. cerevisiae*, mating between haploid cells is signaled by binding of pheromones to a cell type-specific receptor on cells of the opposite mating type (*STE2* expressed in **a** cells recognized by α–factor, and *STE3* expressed in α cells recognized by **a**-factor). The signal is transmitted by interaction of a heterotrimeric G protein complex composed of Gα(Gpa1), Gβ(Ste4), and Gγ(Ste18) through a downstream mitogen-activated protein (MAP) kinase cascade encoded by *STE20*, *STE11*, *STE7,* and *FUS3*. The resulting activation of the transcription factor Ste12p, is required for expression of mating type-specific genes, cell cycle arrest, fusion of mating partners and karyogamy (15). In *S. cerevisiae*, pseudohyphal growth and invasive growth respectively are initiated in diploid and haploid cells upon nutrient deprivation and signaled by genes shared with this MAPK pathway (16). As indicated in Table 1, *C. albicans* homologues for all of these genes in the MAP kinase cascade have been identified, as well as homologues of the pheromone receptor genes *STE2* and *STE3*. Genes involved in pheromone processing in *S. cerevisiae* such as *STE14*, *AXL1*, *STE23*, *RAM1, RAM2*, *STE24*, *RCE1*, *KEX2*, *KEX1*, and *STE13* (17) have homologues in the *Candida* genome, although there is no independent evidence that *C. albicans* can produce or respond to pheromones. We also have identified several pheromone-induced genes such as *FIG1*, *FIG3*, *FIG4* (Table 1), which appear to be important for different steps of mating cell differentiation in *S. cerevisiae* (18). Together, these data suggest that *C. albicans* may have preserved the ability to produce and respond

to mating pheromones. While extensive BLAST analysis failed to identify any mating-factor homologues, computer programs which take into account pheromone gene structure have provided us with several candidates for pheromone genes (unpublished results).

In *S. cerevisiae* α-agglutinin (*SAG1*) provides tight cell-cell adhesion during mating. It has been postulated that in *C. albicans* the adhesiveness of the homologous *ALS* gene family contributes to its pathogenesis (19). The presence of *FUS1* might suggest the ability of cell fusion for *C. albicans*. A *KAR1*-like sequence, which in *S. cerevisiae* is critical for nuclear fusion and spindle pole body (SPB) formation (20), was not found in *C. albicans.*

One of the major outcomes of *CPH1/STE12* activation through the MAP kinase pathway in *C. albicans* is to induce hyphal morphogenesis (21). Virulence of *cst20 /ste20* and *cph1 /ste12* disruption mutants is attenuated in the mouse model of systemic candidasis (22, 23), thereby establishing a potential link between pheromone signaling, filamentous growth and virulence as found in the pathogenic fungi *Cryptococcus neoformans* and *Ustilago maydis* (24, 25). In the evolution of pathogenesis, perhaps *C. albicans* has used the MAP kinase pathway to strictly control expression of its hyphal phenotype in response to changes in host environment. The identification of potential ligands and possible environmental cues which are either recognized by the homologues of *STE2*- and *STE3*-like receptors or which stimulate the MAP kinase pathway through alternate receptors could help us to understand hyphal induction and pathogenesis in *C. albicans*.

Nutritional control:

In *S. cerevisiae*, meiosis is initiated only by diploid cells deprived of glucose and nitrogen and grown in the presence of a nonfermentable carbon source, while other fungi have different and complex nutrient requirements for this process. For instance, nitrogen

starvation is required for mating and meiosis in *Schizosaccharomyces pombe* (26). The plant pathogen, *U. maydis,* enters meiosis only during growth in its host, *Zea mays* (27). *C. albicans* appears to contain homologues of *S. cerevisiae* genes involved in glucose repression and nitrogen metabolism: *MIG1*, *GAT1* and *UME6* as well as *RAS/cAMP*, *SNF1*, *MCK1*, which are involved in the nutrient sensing pathway (28).

The complexity and cross talk between nutritional and meiotic pathways suggest that, although similar genes may be present in both organisms, their participation in these pathways may have different biological consequences. For example, *SNF1* is essential for the viability of *C. albicans* (29) but is not essential in *S. cerevisiae,* where it coordinates glucose and acetate regulation of the early and late meiotic program (30). Another homologue found in *C. albicans*, *MCK1*, encodes a serine-threonine-tyrosine kinase, which functions as a positive regulator of meiotic gene expression in *S. cerevisiae* and is essential for ascus maturation; it governs centromere behavior in mitosis (28). Comparative genomic analyses also reveal the metabolic diversity of *C. albicans* (S. Scherer, personal communication), and suggest that the conditions traditionally used in the laboratory for its culture, may in part be responsible for the failure to detect a sexual cycle in this organism.

Cell type control:

In *S. cerevisiae*, only diploid cells which are heterozygous at the *MAT* locus can initiate meiosis and sporulation upon nutritional starvation. An important feature of **a**/α cells is the presence of a transcriptional repressor, **a**1-α2, which is a heterodimeric homeodomain protein (31). *RME1*, which encodes a negative regulator of meiosis, is one of the genes turned off by **a**1-α2 (32). The product of another gene, *IME4* mediates both cell type and

nutritional activation of *IME1* (see below) (33). No homologue of *IME1* was found in *C. albicans,* although putative counterparts of *RME1* and *IME4* have been identified.

Mating type loci in yeasts are master regulators of cell fate specification and sexual morphogenesis (34). In *C. neoformans*, there is an association between mating type, hyphal phase and infectivity (35). Recently, a mating-type-like (*MTL*) locus in *C. albicans* with homology to the *MAT* locus of *S. cerevisiae* was identified (36). The *C. albicans MTL* locus is large (approximately 8.8 kb), single copy and without silent cassettes such as *HML* or *HMR*. Whether there is a correspondence between mating type and virulence in *C. albicans* has not been established. However, most clinical isolates of *C. albicans* are heterozygous at *MTL* locus (P. Magee, personal communication), and by analogy, would be equivalent to **a**/α cells in *S. cerevisiae*.

Initiation of meiosis:

Initiation of meiosis and sporulation in diploid *S. cerevisiae* by nutritional limitation occurs through a transcriptional cascade with sequentially expressed distinct classes of meiosis-specific genes (37). *UME6* encodes a Zn_2Cys_6 DNA binding protein that functions as a developmental switch for mitotic repression and meiotic activation of early meiotic genes (38). Interaction between Ume6p and Ime1p, a transcriptional activator, is required for induction of early meiotic gene expression. In *S. cerevisiae, RIM11* and *RIM15* kinases are required for Ume6p-Ime1p interaction (39). *RIM101* defines a signaling pathway which activates *IME1*. As shown in Table 1, homologues of *UME6*, *RIM11*, *RIM15* and genes in the *RIM101* pathway have been identified in *C. albicans.* Experiments have shown that the *Candida RIM101* homologue participates in hyphal growth (40). In *S. cerevisiae*, Ime1p activates expression of *IME2*, a serine-threonine kinase, essential for pre-meiotic DNA replication (41). Despite extensive BLAST analysis we have not found counterparts of *IME1*

in *C. albicans*. Similarly efforts to functionally complement an *ime1* null mutant of *S. cerevisiae* with a *C. albicans* genomic library have failed (42), lending further support to the apparent absence of this gene in *C. albicans*. However, we have found downstream targets of Ime1p such as *IME2*. In *S. cerevisiae*, the Ime1p/Ume6p complex and the URS1 consensus binding site for early meiotic genes play a pivotal role in the cell's decision to enter meiosis. The absence of *IME1* suggests that the switch machinery used in *S. cerevisiae* to effect commitment to the meiotic pathway is missing. It remains possible that a functional analogue of *IME1* is present in *C. albicans*.

Checkpoint control and progression through meiosis:

Mitosis and meiosis, the two fundamental modes of cellular reproduction, have overlapping functions including DNA replication and chromosome segregation, as well as similar but distinct mechanisms to survey the progression of cell cycle events. We have identified *C. albicans* homologues of the mitotic DNA damage checkpoint genes *MEC1*, *RAD17*, *RAD24*, which are also required for meiotic progression (43), as well as meiosis-specific *MEK1* and the chromatin-silencing factors *SIR2* and *DOT1*, which are implicated in pachytene checkpoint control (44). We have also identified a set of gene homologues whose products in *S. cerevisiae* participate in DNA replication and chromosome segregation: *CDC5*, *CDC7*, *CDC14*, *CDC20*, *CDC25*, *CDC28*, *IPL1*, *CDH1* (45), several components of the anaphase-promoting complex (*APC1*, *APC2*, *APC11*, *CDC16*, *CDC23*, *CDC27*) and its meiotic activator *AMA1* (*SPO70*) (46). Overall, homologues for most of the genes which in *S. cerevisiae* participate both in mitosis and meiosis have been identified in *C. albicans*.

Meiosis I (MI) is a reductional division whereas meiosis II, like mitosis, is equational. The separation of sister chromatids must be tightly regulated during meiosis; homologues of *PDS5*, *SCC2*, *SMC1*, *SMC2*, *SMC3*, which in *S. cerevisiae* function in sister chromatid

cohesion, were also found in *C. albicans* (47). A homologue of *YPR007* (*REC8*), proposed to encode a protein mediating meiotic sister chromatid cohesion (48), was identified as was *NDT80*, a gene which encodes a meiosis-specific transcription factor. In *S. cerevisiae* Ndt80p is required for expression of middle meiotic genes, meiotic division and spore formation (49). Transcription of *NDT80* is dependent on *IME1*, which remains unidentified in the *C. albicans* genomic sequence. Another important gene apparently absent in *C. albicans,* is a homologue of *SPO13*. *spo13* null mutants in *S. cerevisiae* eliminate the reductional division (MI) during meiosis to produce diploid spores (50). The absence of *SPO13* might indicate that the *C. albicans* sexual cycle is one where cells undergo single-division meiosis.

Recombination and the formation of synaptonemal complexes:

Meiosis is fundamentally different from mitosis in the occurrence of high frequency recombination. Meiotic recombination, which is thought to proceed through a double-strand-break (DSB) repair mechanism (51), follows pre-meiotic DNA replication. More than ten genes are required to produce DSBs in meiotic cells in *S. cerevisiae;* strikingly, several of these are absent from *C. albicans*: *MER1*, *MER2*, *REC102*, *REC104*, *REC114*, and *MEI4*. The absence of homologues of approximately half of the genes required for the initiation of recombination in *S. cerevisiae* suggests that *C. albicans* may be greatly impaired in this initial step. However, we have identified *SPO11*, a type II topoisomerase implicated in nicking DNA to generate DSBs (52). Homologues of *SPO11* have also been identified in *Caenorhabditis elegans*, *Drosophila melanogaster, S. pombe,* mouse and human (53).

Homologues of most of the genes involved in strand invasion and Holliday junction formation have been identified as well: *RAD51*, *RAD52*, *RAD54*, *RAD57*, and *DMC1*(54). The *C. albicans DLH1* gene complements *DMC1* null mutants, its meiosis-specific

homologue in *S. cerevisiae* (55). Homologues of the meiosis-specific genes *SAE2* and *SAE3* are also absent from *C. albicans,* while those of genes which participate in both mitotic and meiotic mismatch repair *MSH2*, *MSH6*, *MSH3*, *MLH1* and *PMS1* are present (56).

Many unique characteristics of meiotic recombination are ensured through the formation of synaptonemal complexes (SC) (57). These structures not only exclude recombination between sister chromatids, but control the frequency of crossovers. Homologues of *HOP1* and *ZIP1*, molecular components of the SC, were found in *C. albicans* while those for other proteins, such as *ZIP2* and *HOP2*, involved in synapsis of chromosomes and formation of SC, are absent (58, 59). Homologues of meiosis-specific genes *MSH4* and *MSH5*, required to promote recombination between homologous chromosomes and Holliday junction resolution, were also identified. Together, these findings suggest that *C. albicans* is not competent to make mature SC. However, *C. albicans* may resemble other fungi such as *S. pombe* and *Aspergillus nidulans,* which exhibit meiotic recombination but do not form SC (60).

Spore wall morphogenesis and ascus formation:

In *S. cerevisiae*, the spore wall is formed *de novo* through a complex morphogenetic program dependent upon the initiation of meiosis. Although there are no detectable meiotic ascospores in *C. albicans*, homologues of genes involved in the sporulation pathway such as *SPS1*, a *STE20*-like protein kinase, and *SMK1*, a MAP kinase, were identified. We also found homologues of a set of mid/late sporulation-specific genes involved in spore wall maturation: *DIT1*, *DIT2*, *SPR3* and *YDR104*. *DIT1* and *DIT2* are important for the formation of the outer layer of the spore wall. *YDR104* is essential for spore wall formation in *S. cerevisiae* (14). The presence of these genes may reflect a possible MAP kinase pathway

which can transmit environmental signals to modify cell wall structures in *C. albicans*. The extent to which homologues of these genes function in yeast or hyphal wall synthesis is not known. The *C. albicans* cell wall contains variable amounts of dityrosine and its basic structural features are similar but not identical to those of *S. cerevisiae* (61).

Genomic comparison between *C. albicans* and other organisms

Given that gametogenesis in metazoans and sporulation in yeasts are evolutionary highly conserved processes, it is not surprising that a survey of available databases revealed several genes which function in meiosis in multicellular organisms and which have homologues in *C. albicans.* Some of the genes analyzed are listed in Table 2 (see Supplementary Material, www.pnas.org). For example, we have found homologues of *pelota* and *DES-1*, which are required for spermatogenesis in *Drosophila*. *DOM34* is the *S. cerevisiae* homologue of *pelota* and is important for meiosis, pseudohyphal growth and translation (62). *DES-1*, a transmembrane protein, is required for initiation of meiosis in *Drosophila* spermatogenesis. A *DES-1* homologue has been identified in mouse, *S. pombe*, *Arabidopsis thaliana* (63), but not in *S. cerevisiae*. The *Drosophila* protein is proposed to mediate the interaction between somatic cells and germ cells during development (64), raising the possibility that *DES-1* homologue might be required for communication between *C. albicans* and its environment.

We have also found several *S. pombe* gene homologues (in *C. albicans*) such as *RCD1* (*YNL288*), *NRD1* (*YPL184*), *STE20* (*YER093*) and *MEI4* (*FKH1*), which are involved in sexual differentiation in the fission yeast (26). Their counterparts in *S. cerevisiae* have been identified, but their functional role in meiosis has not been established. For example, *S. cerevisiae FKH1* regulates mitotic cell cycle progression and pseudohyphal growth (65) and is expressed early during sporulation (14).

The apparent dissimilarity in meiotic machinery between *S. cerevisiae* and *C. albicans* may reflect their relative phylogenetic distance (7). To investigate this further, we sought potential meiotic homologues in *C. albicans* which may resemble key meiotic regulators or cognates of genes which function in the sexual cycle in other fungi such as *C. neoformans*, *A. nidulans*, *Neurospora crassa* and *Podospora anserina*. From this analysis, we have identified several meiotic homologues including *SPO14*, *MCK1*, *MEK1* and a pH regulatory system with components of *palI*, *pacC*, *palA*, *palF* and *palB* shared by *C. albicans*, *S. cerevisiae* and *A. nidulans*. With respect to *pacC*, a homologue of *RIM101*, this gene is conserved in all three of these fungal species as are other elements of the *RIM101* pathway. In *S. cerevisiae* the *RIM101* pathway functions both in meiosis and invasive growth. In *C. albicans* it is required for pathogenesis (66).

Asexual sporulation is common among diverse groups of fungi. *C. albicans* can produce structures known as chlamydospores under special conditions (6); their biological role is unknown. We have identified a set of *C. albicans* genes which are homologues of those which participate in asexual sporulation in *A. nidulans* (67). These include *C. albicans* genes designated *SST2* (a homologue of *flbA*), *YPR013* (a homologue of *flbC*), *MYB1* (a homologue of *flbD*), *TEC1* (a homologue of *abaA*), *DOP1* (a homologue of *dopA*), and *EFG1* (a homologue of *stuA*). Several of these were also found in *S. cerevisiae*. Characterization of these gene homologues may help to define the role of chlamydospores in *C. albicans* cell biology.

In considering alternative reproductive pathways, the identification of *HET-C*, and *HET-E-1* homologues in *C. albicans* is particularly interesting. The *HET-C* gene, which encodes a glycolipid transfer protein, is proposed to function in vegetative incompatibility and ascospore formation in *Podospora anserina* (68). *HET-E-1* encodes a β–transducin-like

protein (69). No *HET-C* gene homologue was found in *S. cerevisiae*. Homologues of *MOD-D* and *MOD-E*, modifiers of the *het* locus, were also identified in *C. albicans*. *MOD-D* is a homologue of $G\alpha$ subunit, and *MOD-E* is an *HSP90* homologue (70, 71). As the *het* loci appear to regulate self/nonself-recognition during vegetative growth in filamentous fungi, heterokaryon formation could provide an opportunity for genetic exchange in imperfect fungi (72).

CONCLUSION

C. albicans can produce tetraploid progeny in animals and in culture when parental types homozygous at *MTL* locus are artificially created (4, 5). However, there is no evidence of reductive division from these matings, and it remains to be shown that this fungus has a meaningful sexual cycle. Most clinical isolates of *Candida* are heterozygous at the *MTL* locus and most proliferation of naturally occurring populations appears to be clonal (73), suggesting that mating is infrequent. Our comprehensive comparison and analysis of the *C. albicans* genome, focusing on the question of sexuality in *C. albicans,* suggests that while this organism suffers a multiple gene-defect in a *S. cerevisiae*-like meiotic machinery, it possesses a repertoire of genes composed of homologues found in sexual pathways of *S. cerevisiae*, filamentous fungi and metazoans, which in aggregate indicate that *C. albicans* has a sexual cycle in nature.

One of the more interesting results of our analysis is the finding that while *C. albicans* appears to lack a homologue of *IME1*, in *S. cerevisiae* the master switch for entry into the meiotic pathway, it does possess a number of homologues of the downstream target genes of the Ime1p/Ume6p regulatory complex in *S. cerevisiae*. As *IME1* is the integration point of genetic and nutritional signals for meiosis in *S. cerevisiae* (74), the regulation of initiation of meiosis might be achieved in *Candida* through an analogue of *IME1* which integrates a

different set of signals to effect commitment to meiosis. We have not been able to identify *IME1* homologues in other organisms despite extensive BLAST analysis, indicating that perhaps the commitment to meiosis effected by the interaction between Ime1p and Ume6p is unique to *S. cerevisiae.* Similarly absent in *C. albicans* and other organisms as well is *SPO13*, which in *S. cerevisiae* is essential for proper execution of meiosis I. Within the meiotic pathway itself, *C. albicans* is missing 6 of 10 homologues of genes which in *S. cerevisiae* are necessary for the initiation of DSBs and others which are involved in chromosome recombination and the formation of SC. Most of the missing genes are related to recombination and SC formation, although in each of these processes, *C. albicans* still possesses a number of homologues which may participate at various stages in the progression of meiosis.

The study of fungal developmental signalling pathways has revealed an intimate relationship between mating and filamentous growth (75) and suggests that different cell types can utilize shared signalling components to specify cell fate. *C. albicans* is able to proliferate in many forms--as yeast, pseudohyphae, and filamentous hyphae. It is also able to form chlamydospores and express different switch phenotypes in response to a variety of environmental cues. Any one of these forms may be capable of cell fusion, a possibility highlighted by the identification of several components of the *het* system as well as gene homologues which play crucial roles in cellular reproduction in various organisms. Our studies suggest the possibility for hyphal fusion and dikaryon formation in *C. albicans* and allude to more complex patterns of both asexual and sexual differentiation similar to those found in some filamentous ascomycetes and basidomycetes (76, 77).

ACKNOWLEDGEMENTS

K.W.T. would like to dedicate this paper to his parents. K.W.T. thanks G. Newport, C.Y. Lan and M.I. Cano for helpful discussions. Special thanks to members of the Agabian lab for their support and advice. We are grateful to I. Herskowitz, P. O'Farrell and R. Reijo for constructive comments on the manuscript. This work was supported by NIH grant 1R01DE12940-01 to N.A., and sequencing of *Candida albicans* genome was accomplished with the support of the NIDCR grant DE12302-02S2 and the Burroughs Wellcome Fund. K.W.T. is supported by NIH grant P01DE07946, and R.M.W. is supported by NIH grant HG01633 (awarded to R.W.D.).

REFERENCES

1. Kleckner, N. (1996) *Proc Natl Acad Sci U S A* **93,** 8167-74.
2. Kwon-Chung, K. J. & Bennett, J. E. (1992) *Medical mycology* (Lea & Febiger, Philadelphia).
3. See *Candida albicans* genome project information at http://alces.med.umn.edu/Candida.html and http://www-sequence.stanford.edu/group/candida
4. Hull, C. M., Raisner, R. M. & Johnson, A. D. (2000) *Science* **289,** 307-10.
5. Magee, B. B. & Magee, P. T. (2000) *Science* **289,** 310-3.
6. Odds, F. C. (1988) *Candida and candidosis* (Baillière Tindall, London ; Philadelphia).
7. Hendriks, L., Goris, A., Van de Peer, Y., Neefs, J. M., Vancanneyt, M., Kersters, K., Hennebert, G. L. & De Wachter, R. (1991) *J Gen Microbiol* **137,** 1223-30.
8. Thorstenson, Y. R., Hunicke-Smith, S. P., Oefner, P. J. & Davis, R. W. (1998) *Genome Res* **8,** 848-55.
9. Marziali, A., Willis, T. D., Federspiel, N. A. & Davis, R. W. (1999) *Genome Res* **9,** 457-62.
10. Ewing, B., Hillier, L., Wendl, M. C. & Green, P. (1998) *Genome Res* **8,** 175-85.

11. Ewing, B. & Green, P. (1998) *Genome Res* **8,** 186-94.

12. Gordon, D., Abajian, C. & Green, P. (1998) *Genome Res* **8,** 195-202.

13. Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) *Nucleic Acids Res* **25,** 3389-402.

14. Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O. & Herskowitz, I. (1998) *Science* **282,** 699-705.

15. Sprague, G. F.,Jr. & Thorner, J. W. (1992) in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, eds. Jones, E. W., Pringle, J. R. & Broach, J. R. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor), pp. 657-744.

16. Gustin, M. C., Albertyn, J., Alexander, M. & Davenport, K. (1998) *Microbiol Mol Biol Rev* **62,** 1264-300.

17. Davey, J., Davis, K., Hughes, M., Ladds, G. & Powner, D. (1998) *Semin Cell Dev Biol* **9,** 19-30.

18. Erdman, S., Lin, L., Malczynski, M. & Snyder, M. (1998) *J Cell Biol* **140,** 461-83.

19. Sundstrom, P. (1999) *Curr Opin Microbiol* **2,** 353-7.

20. Marsh, L. & Rose, M. D. (1997) in *The Molecular and Cellular Biology of the Yeast Saccharomyces : Cell Cycle and Cell Biology*, eds. Pringle, J. R., Broach, J. R. & Jones, E. W. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor), pp. 827-888.

21. Liu, H., Kohler, J. & Fink, G. R. (1994) *Science* **266,** 1723-6.

22. Leberer, E., Harcus, D., Broadbent, I. D., Clark, K. L., Dignard, D., Ziegelbauer, K., Schmidt, A., Gow, N. A., Brown, A. J. & Thomas, D. Y. (1996) *Proc Natl Acad Sci U S A* **93,** 13217-22.

23. Lo, H. J., Kohler, J. R., DiDomenico, B., Loebenberg, D., Cacciapuoti, A. & Fink, G. R. (1997) *Cell* **90,** 939-49.

24. Wang, P. & Heitman, J. (1999) *Curr Opin Microbiol* **2,** 358-62.

25. Banuett, F. & Herskowitz, I. (1994) *Genes Dev* **8,** 1367-78.

26. Davey, J. (1998) *Yeast* **14,** 1529-66.

27. Banuett, F. (1992) *Trends Genet* **8,** 174-80.

28. Kupiec, M., Byers, B., Esposito, R. E. & Mitchell, A. P. (1997) in *The Molecular and Cellular Biology of the Yeast Saccharomyces : Cell Cycle and Cell Biology*, eds. Pringle, J. R., Broach, J. R. & Jones, E. W. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor), pp. 889-1036.

29. Petter, R., Chang, Y. C. & Kwon-Chung, K. J. (1997) *Infect Immun* **65,** 4909-17.

30. Honigberg, S. M. & Lee, R. H. (1998) *Mol Cell Biol* **18,** 4548-55.

31. Dranginis, A. M. (1990) *Nature* **347,** 682-5.

32. Covitz, P. A., Herskowitz, I. & Mitchell, A. P. (1991) *Genes Dev* **5,** 1982-9.

33. Shah, J. C. & Clancy, M. J. (1992) *Mol Cell Biol* **12,** 1078-86.

34. Herskowitz, I. (1989) *Nature* **342,** 749-57.

35. Wickes, B. L., Mayorga, M. E., Edman, U. & Edman, J. C. (1996) *Proc Natl Acad Sci U S A* **93,** 7327-31.

36. Hull, C. M. & Johnson, A. D. (1999) *Science* **285,** 1271-5.

37. Mitchell, A. P. (1994) *Microbiol Rev* **58,** 56-70.

38. Steber, C. M. & Esposito, R. E. (1995) *Proc Natl Acad Sci U S A* **92,** 12490-4.

39. Rubin-Bejerano, I., Mandel, S., Robzyk, K. & Kassir, Y. (1996) *Mol Cell Biol* **16,** 2518-26.

40. Davis, D., Wilson, R. B. & Mitchell, A. P. (2000) *Mol Cell Biol* **20,** 971-8.

41. Dirick, L., Goetsch, L., Ammerer, G. & Byers, B. (1998) *Science* **281,** 1854-7.

42. Andaluz, E., Larriba, G. & Calderone, R. (1996) *Yeast* **12,** 893-8.

43. Lydall, D., Nikolsky, Y., Bishop, D. K. & Weinert, T. (1996) *Nature* **383,** 840-3.

44. Roeder, G. S. & Bailis, J. M. (2000) *Trends Genet* **16,** 395-403.

45. Lew, D. J., Weinert, T. & Pringle, J. R. (1997) in *The Molecular and Cellular Biology of the Yeast Saccharomyces : Cell Cycle and Cell Biology*, eds. Pringle, J. R., Broach, J. R. & Jones, E. W. (Cold Spring Harbor Laboratory, Cold Spring Harbor), pp. 607-695.

46. Zachariae, W. & Nasmyth, K. (1999) *Genes Dev* **13,** 2039-58.

47. Strunnikov, A. V. & Jessberger, R. (1999) *Eur J Biochem* **263,** 6-13.

48. Klein, F., Mahr, P., Galova, M., Buonomo, S. B., Michaelis, C., Nairz, K. & Nasmyth, K. (1999) *Cell* **98,** 91-103.

49. Chu, S. & Herskowitz, I. (1998) *Mol Cell* **1,** 685-96.

50. McCarroll, R. M. & Esposito, R. E. (1994) *Genetics* **138,** 47-60.

51. Szostak, J. W., Orr-Weaver, T. L., Rothstein, R. J. & Stahl, F. W. (1983) *Cell* **33,** 25-35.

52. Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P. C., Nicolas, A. & Forterre, P. (1997) *Nature* **386,** 414-7.

53. Romanienko, P. J. & Camerini-Otero, R. D. (1999) *Genomics* **61,** 156-69.

54. Smith, K. N. & Nicolas, A. (1998) *Curr Opin Genet Dev* **8,** 200-11.

55. Diener, A. C. & Fink, G. R. (1996) *Genetics* **143,** 769-76.

56. Kirkpatrick, D. T. (1999) *Cell Mol Life Sci* **55,** 437-49.

57. Roeder, G. S. (1997) *Genes Dev* **11,** 2600-21.

58. Chua, P. R. & Roeder, G. S. (1998) *Cell* **93,** 349-59.

59. Leu, J. Y., Chua, P. R. & Roeder, G. S. (1998) *Cell* **94,** 375-86.

60. Kohli, J. & Bahler, J. (1994) *Experientia* **50,** 295-306.

61. Chaffin, W. L., Lopez-Ribot, J. L., Casanova, M., Gozalbo, D. & Martinez, J. P. (1998) *Microbiol Mol Biol Rev* **62,** 130-80.

62. Davis, L. & Engebrecht, J. (1998) *Genetics* **149,** 45-56.

63. Endo, K., Matsuda, Y. & Kobayashi, S. (1997) *Dev Growth Differ* **39,** 399-403.

64. Endo, K., Akiyama, T., Kobayashi, S. & Okada, M. (1996) *Mol Gen Genet* **253,** 157-65.

65. Zhu, G., Spellman, P. T., Volpe, T., Brown, P. O., Botstein, D., Davis, T. N. & Futcher, B. (2000) *Nature* **406,** 90-4.

66. Davis, D., Edwards, J. E., Jr., Mitchell, A. P. & Ibrahim, A. S. (2000) *Infect Immun* **68,** 5953-5959.

67. Adams, T. H., Wieser, J. K. & Yu, J. H. (1998) *Microbiol Mol Biol Rev* **62,** 35-54.

68. Saupe, S., Descamps, C., Turcq, B. & Begueret, J. (1994) *Proc Natl Acad Sci U S A* **91,** 5927-31.

69. Espagne, E., Balhadere, P., Begueret, J. & Turcq, B. (1997) *Mol Gen Genet* **256,** 620-7.

70. Loubradou, G., Begueret, J. & Turcq, B. (1999) *Genetics* **152,** 519-28.

71. Loubradou, G., Begueret, J. & Turcq, B. (1997) *Genetics* **147,** 581-8.

72. Saupe, S. J. (2000) *Microbiol Mol Biol Rev* **64,** 489-502.

73. Graser, Y., Volovsek, M., Arrington, J., Schonian, G., Presber, W., Mitchell, T. G. & Vilgalys, R. (1996) *Proc Natl Acad Sci U S A* **93,** 12473-7.

74. Vershon, A. K. & Pierce, M. (2000) *Curr Opin Cell Biol* **12,** 334-9.

75. Madhani, H. D. & Fink, G. R. (1998) *Trends Cell Biol* **8,** 348-53.

76. Casselton, L. A. & Olesnicky, N. S. (1998) *Microbiol Mol Biol Rev* **62,** 55-70.

77. Coppin, E., Debuchy, R., Arnaise, S. & Picard, M. (1997) *Microbiol Mol Biol Rev* **61,** 411-28.

Table 1

Candida albicans genes homologous to *Saccharomyces cerevisiae* genes important for mating and meiosis. The functional groups used in the categories roughly correspond to : (1) mating differentiation; (2) nutritional control; (3) cell type control; (4) initiation of meiosis ; (5) checkpoint control and progression through meiosis; (6) recombination and the formation of synaptonemal complexes; and (7) spore wall morphogenesis and ascus formation. The functional description of each gene was obtained from YPD at proteome (http://www.proteome.com) and Saccharomyces Genome database (http://genome-www.stanford.edu/Saccharomyces/).

Sc ORF	Ca ORF	Ca Gene	Category #	Description
YLR389C	orf6.3250	STE23	1	Protease involved in a-factor processing; similar to Axl1p and human insulin-degrading enzymes
YJR117W	orf6.6894	STE24	1	Zinc metallo-protease; involved in processing of a-factor
YNL325C	orf6.4428	FIG4	1	Suppressor of sac1 mutation, induced by mating factor
YNL238W	orf6.4600	KEX2	1	Endoprotease of late Golgi compartment involved in processing of alpha-factor
YOR219C	orf6.4953	STE13	1	Dipeptidyl aminopeptidase A involved in maturation of alpha-factor
YBL016W	orf6.2854	FUS3	1	Serine/threonine protein kinase of the MAP kinase family required for cell cycle arrest and for cell fusion during mating
YPR141C	orf6.6836	KAR3	1	Kinesin-like protein involved in mitosis, essential for the congression step of karyogamy
YLR362W	orf6.3750	STE11	1	MAP kinase kinase kinase in the pheromone response pathway, the filamentous growth pathway and the high-osmolarity sensing pathway
YOR212W	orf6.1139	STE4	1	Guanine nucleotide-binding protein beta subunit of the pheromone response pathway, member of the WD-40 repeat family
YHR084W	orf6.695	CPH1	1	Transcription factor binds to pheromone response element to regulate genes required for mating; functions with Tec1p to regulate genes for filamentous growth
YLR452C	orf6.2767	SST2	1	Protein involved in desensitization to alpha-factor pheromone
YHR013C	orf6.8965	ARD1	1	Protein N-acetyltransferase subunit; mating functions are reduced in mutants due to derepression of silent mating type loci
YKL178C	orf6.6173	STE3	1	Pheromone a-factor receptor; seven transmembrane domains protein
YDL159W	orf6.3212	HST7	1	Serine/threonine/tyrosine protein kinase of MAP kinase kinase family, component of the pheromone pathway and a pathway regulating filamentous

				growth
YPR122W	orf6.8423	AXL1	1	Protease involved in proteolytic step of a-factor N-terminal processing
YLL021W	orf6.5791	SPA2	1	Protein involved in cell fusion and cell polarity during mating
YDR410C	orf6.1442	STE14	1	Farnesyl cysteine-carboxyl methyltransferase
YFL026W	orf6.4012	STE2	1	Pheromone alpha-factor receptor; has seven transmembrane segments
YJL157C	orf6.8448	FAR1	1	Inhibitor of CDK complexes involved in cell cycle arrest for mating
YCL027W	orf6.3131	FUS1	1	Protein with SH3 domain required for cell fusion during mating, located at the tip of the mating projection
YDR103W	no match	STE5	1	Scaffolding protein involved in a complex with Fus3p, Ste7p, and Ste11p in pheromone pathway
YOL051W	orf6.6345	GAL11	1	Component of RNA polymerase holoenzyme with positive and negative effects on transcription
YNL188W	no match	KAR1	1	Protein required for spindle pole body duplication and karyogamy
YKL209C	orf6.8600	HST6	1	Membrane transporter of ATP-binding cassette (ABC) superfamily responsible for export of a-factor mating pheromone
YHR005C	orf6.7428	CAG1	1	Guanine nucleotide-binding protein alpha subunit involved in pheromone response pathway
YJR004C	orf6.2112	ALS1	1	Alpha-agglutinin involved in cell-cell interactions during mating
YBR040W	orf6.3476	FIG1	1	Protein required for mating; induced by mating pheromone
YER068W	orf6.3869	SIG1	1	Zn finger transcriptional repressor, involved in pheromone signal transduction
YMR065W	orf6.2573	KAR5	1	Coiled-coil membrane protein, required for nuclear fusion, synonym: FIG3
YBR200W	orf6.1490	BEM1	1	Required for cell polarity, mating and budding; SH3 domain protein
YDR480W	no match	DIG2	1	MAP kinase-associated protein, negative regulator of mating and invasive growth
YNL279W	orf6.6562	PRM1	1	Required for membrane fusion during mating
YMR274C	orf6.7520	RCE1	1	Protease involved in a-factor C-terminal proteolysis
YKL019W	orf6.5223	RAM2	1	Required for a-factor prenylation
YDL090C	orf6.3323	RAM1	1	Required for Ras proteins and a-factor prenylation
YGL203C	orf6.7946	KEX1	1	Carboxypeptidase involved in processing of alpha-factor
YNL271C	orf6.5907	BNI1	1	Formin homolog; required for bipolar budding and default mating
YCL032W	orf6.6436	STE50	1	Protein involved in pheromone signal transduction; interacts with Ste11p and G protein
YGR238C	orf6.6249	KEL2	1	Kelch homolog; involved in cell fusion
YDR477W	orf6.5603	SNF1	2	Serine/threonine protein kinase for derepression of glucose-repressed genes

YER133W	orf6.5435	GLC7	2	Protein serine/threonine phosphatase PP1 required for glucose repression
YML115C	orf6.8893	VAN1	2	Vanadate resistance protein
YDL040C	orf6.4482	NAT1	2	Protein N-acetyltransferase subunit, mating reduced in mutants due to derepression of silent mating type loci
YOL081W	orf6.3680	IRA2	2	GTPase-activating protein (GAP) for Ras1p and Ras2p
YBR140C	orf6.3680	IRA1	2	GAP for Ras1p and Ras2p
YNL307C	orf6.3762	MCK1	2	Serine/threonine/tyrosine protein kinase, positive regulator of meiosis and spore formation
YMR139W	orf6.1071	RIM11	2	Serine/threonine protein kinase required for induction of IME2 by Ime1p
YGL115W	orf6.583	SNF4	2	Protein involved in derepression of glucose-repressed genes
YFL038C	orf6.7789	YPT1	2	GTP-binding protein of the rab family required for vesicle transport from ER to Golgi and within the Golgi stack
YIL099W	orf6.4841	SGA1	2	Glucoamylase, sporulation-specific
YIR026C	orf6.6306	YVH1	2	Dual-specificity protein phosphatase (PTPase) similar to vaccinia VH1
YPL084W	orf6.5745	BRO1	2	Protein interacting with the PKC1-MAP kinase pathway
YGL197W	orf6.8871	MDS3	2	Negative regulator of early meiotic genes
YOR101W	orf6.4089	RAS1	2	GTP-binding protein involved in regulation of cAMP pathway, homolog of proto-oncogene ras
YNL098C	orf6.4089	RAS2	2	GTP-binding protein involved in regulation of cAMP pathway
YGL035C	orf6.3170	MIG1	2	Zinc finger transcriptional repressor involved in glucose-repression
YFL021W	orf6.2898	GAT1	2	GATA zinc finger transcription factor that plays a supplemental role to Gln3p in the activation of genes to use non-preferred nitrogen sources (nitrogen catabolite repression)
YMR043W	orf6.7951	MCM1	3	Transcription factor of the MADS box family; recruits coregulatory proteins for both gene activation and repression
YGR044C	orf6.7251	RME1	3	Zinc finger transcription factor that represses meiosis in non-a/alpha cells
YCR097W	orf6.1884	MTLA1	3	Homeodomain protein MATa1p interacts with MATalpha2p to repress transcription of haploid-specific genes in diploid cells
YCR040W	orf6.4399	MTLAL1	3	Regulatory protein MATalpha1p interacts with Mcm1p activates alpha-specific genes
YCR096C	no match	MTLA2	3	Regulatory protein MATa2p (no known function)
YCR039C	orf6.4403	MTLAL2	3	Homeodomain regulatory protein MATalpha2p interacts with Mcm1p to turn off a-specific genes
YOL004W	orf6.5091	SIN3	4	Transcriptional regulator with negative and positive effects on individual gene expression
YNL330C	orf6.8830	RPD3	4	Component of histone deacetylase, transcriptional modifier

YCR084C	orf6.6232	TUP1	4	General repressor of transcription (with Ssn6p), member of WD-40 repeat family
YFL033C	orf6.7970	RIM15	4	Serine/threonine protein kinase, activator of IME2 expression and of sporulation
YGL192W	orf6.7097	IME4	4	Positive regulator for IME1 and IME2; controls meiosis by transmitting signals regarding mating type and nutritional status
YLR310C	orf6.7312	CDC25	4	Guanine-nucleotide exchange protein with SH3 domain for Ras1p and Ras2p
YLR071C	orf6.3992	RGR1	4	Component of RNA polymerase holoenzyme with positive and negative effects on transcription
YHL027W	orf6.8147	RIM101	4	Zinc finger protein involved in induction of IME1
YNL236W	orf6.2645	SIN4	4	Component of RNA polymerase holoenzyme with positive and negative effects on transcription
YMR154C	orf6.7408	RIM13	4	Protein involved in proteolytic processing of Rim101p
YDR207C	orf6.4701	UME6	4	Negative transcriptional regulator involved in nitrogen repression and induction of meiosis
YPL139C	no match	UME1	4	Negative regulator of meiosis
YMR063W	orf6.491	RIM9	4	Protein stimulating expression of IME1
YJL146W	no match	IDS2	4	Protein affecting the function of Ime2p
YJR094C	no match	IME1	4	Master switch for sporulation, positive regulator of IME2 and many early meiotic genes
YFR031C	orf6.1515	SMC2	5	Coiled-coil protein of the SMC family involved in chromosome condensation and segregation
YNL088W	orf6.4382	TOP2	5	DNA topoisomerase II; makes double-stranded breaks
YBL075C	orf6.2627	SSA3	5	Heat shock protein of HSP70 family, heat-induced cytoplasmic form
YJL074C	orf6.2013	SMC3	5	Coiled-coil protein, cohesin involved in chromosome condensation and segregation
YFL008W	orf6.2422	SMC1	5	Coiled-coil protein involved in chromosome condensation and segregation
YMR001C	orf6.5090	CDC5	5	Serine/threonine protein kinase required for exit from mitosis; involved in both meiotic divisions
YDL102W	orf6.546	POL3	5	DNA polymerase delta large subunit
YGR184C	orf6.5483	UBR1	5	Ubiquitin-protein ligase involved in selection of substrates for protein degradation
YBR112C	orf6.8833	SSN6	5	General repressor of transcription (with Tup1p), has tetratricopeptide (TPR) repeats
YPL154C	orf6.7029	PEP4	5	Proteinase A; aspartyl protease required for activation of various degradative enzymes
YKL166C	orf6.4228	TPK3	5	cAMP-dependent protein kinase 3
YNL172W	orf6.7590	APC1	5	Required for Clb2p degradation, metaphase-anaphase transition;

				component of APC
YDL008W	orf6.8992	APC11	5	Ubiquitin ligase, component of APC; required for metaphase-anaphase transition
YFR028C	orf6.2743	CDC14	5	Dual specificity protein phosphatase required for exit from mitosis
YBR160W	orf6.5246	CDC28	5	CDK essential for completion of START and for mitosis
YAL038W	orf6.5754	PYK1	5	Pyruvate kinase; required for glycolysis
YGL116W	orf6.2609	CDC20	5	Protein required for microtubule function at mitosis, member of WD-40 repeat family; activator of APC
YNL202W	orf6.2495	SPS19	5	Peroxisomal 2,4-dienyol-CoA reductase
YFL009W	orf6.3808	CDC4	5	Component of the SCF-Cdc4p complex, F box protein; required for ubiquitin-dependent degradation, has WD-40 repeats
YDL028C	orf6.8374	MPS1	5	Serine/threonine/tyrosine protein kinase involved in spindle pole body duplication and in checkpoint control in mitosis
YJL106W	orf6.3678	IME2	5	Serine/threonine protein kinase, positive regulator of early meiotic genes
YDL017W	orf6.5768	CDC7	5	Protein kinase interacts with Dbf4p for initiation of DNA synthesis; required for commitment to sporulation and for meiotic recombination
YBL084C	orf6.8231	CDC27	5	Component of anaphase-promoting complex; required for Clb2p degradation and for the metaphase-anaphase transition
YPL209C	orf6.4985	IPL1	5	Serine/threonine protein kinase involved in chromosome segregation
YDR364C	orf6.5776	CDC40	5	Protein required for mRNA splicing, member of WD-40 repeat family
YHR008C	orf6.4731	SOD2	5	Manganese superoxide dismutase in mitochondrial matrix
YKL022C	orf6.4371	CDC16	5	Component of APC; required for Clb2p degradation and for the metaphase-anaphase transition
YNL012W	orf6.553	SPO1	5	Transcriptional regulator in sporulation; similar to phospholipase B enzymes
YBL058W	orf6.1714	SHP1	5	Potential regulatory subunit for Glc7p; involved in carbohydrate metabolism and meiotic differentiation
YHR166C	orf6.7817	CDC23	5	Component of anaphase-promoting complex; required for Clb2p degradation and for the metaphase-anaphase transition
YMR076C	orf6.4261	PDS5	5	Protein involved in chromosome condensation and sister chromatid cohesion
YJR104C	orf6.7495	SOD1	5	Copper-zinc superoxide dismutase
YDR113C	no match	PDS1	5	Protein involved in sister chromatid separation
YJR057W	orf6.1573	CDC8	5	Thymidylate kinase; involved in dTTP biosynthesis pathway
YER012W	orf6.7438	PRE1	5	Proteasome subunit beta4_sc; involved in protein degradation
YAL009W	orf6.7149	SPO7	5	Protein required for meiosis

YDR440W	orf6.8562	DOT1	5	Protein required for pachytene checkpoint control and silencing at telomeres
YHR124W	orf6.4742	NDT80	5	Meiosis-specific transcription factor required for exit from pachytene
YLR127C	orf6.8810	APC2	5	Component of the APC; required for Clb2p degradation and for the metaphase-anaphase transition
YNL204C	orf6.2496	SPS18	5	Meiosis-specific zinc finger protein involved in activation of sporulation
YDR522C	orf6.1488	SPS2	5	Middle/late gene of meiosis
YDR108W	orf6.3955	GSG1	5	Protein required for sporulation and vesicular transport
YML031W	orf6.5018	NDC1	5	Protein involved in spindle pole body duplication; component of the nuclear envelope
YOR368W	orf6.2056	RAD17	5	Protein required for DNA damage checkpoint and meiotic recombination checkpoint
YDR285W	orf6.9102	ZIP1	5	Structural component of the synaptonemal complex central element, with coiled-coil domain
YPR007C	orf6.1672	SPO69	5	Required for sporulation; meiotic cohesin
YDL220C	no match	CDC13	5	Telomere-binding protein required for access of telomerase
YER180C	no match	ISC10	5	Meiosis-specific protein required for spore formation
YLR079W	no match	SIC1	5	Inhibitor of Cdc28p-Clb protein kinase complex; P40
YOR242C	orf6.2018	SSP2	5	Sporulation-specific protein
YCR086W	orf6.9011	SPO86	5	Protein involved in sporulation
YGR225W	orf6.9083	SPO70	5	Activator of meiotic anaphase-promoting complex
YGL173C	orf6.1895	KEM1	6	Nuclease with 5'-3' exonuclease activity; primarily degrades decapped mRNA
YDR097C	orf6.6045	MSH6	6	Component of mismatch repair complex; involved in single base mismatch repairs
YGL163C	orf6.8005	RAD54	6	DNA-dependent ATPase; required for recombination and repair of X-ray damage
YOL090W	orf6.6952	MSH2	6	DNA mismatch binding factor involved in repair of short insertions/deletions and single base mismatches; functions in both mitotic and meiotic repair
YMR167W	orf6.2626	MLH1	6	Mismatch repair protein; E. coli MutL homolog
YNL250W	orf6.2332	RAD50	6	Protein involved in recombinational DNA repair; required for resection at double-stranded breaks
YMR224C	orf6.7179	MRE11	6	Putative 5' to 3' exonuclease required for double–strand- break repair and meiotic recombination
YCR092C	orf6.2464	MSH3	6	DNA mismatch binding factor required for mismatch repair involving microsatellite sequences
YAR007C	orf6.7462	RFA1	6	DNA replication factor A; binds single-stranded DNA

YER095W	orf6.4214	RAD51	6	Protein stimulating pairing and strand-exchange between homologous single-stranded and double-stranded DNA, similar to E. coli RecA functionally
YPL008W	orf6.6360	CHL1	6	Protein of the DEAH box family; involved in chromosome transmission and cell cycle progression through G2/M
YFL003C	orf6.4680	MSH4	6	Protein homologous to E. coli MutS, meiosis-specific
YDL164C	orf6.5316	CDC9	6	DNA ligase; involved in DNA repair and replication
YER179W	orf6.2266	DLH1	6	Meiosis-specific recombination protein structurally related to Rad51p, Rad55p, Rad57p, and to E. coli RecA; required for repair of double-strand breaks during meiosis
YGR258C	orf6.4886	RAD2	6	Single-stranded DNA endonuclease required for nucleotide excision repair
YLR234W	orf6.9103	TOP3	6	DNA topoisomerase III; relaxes negatively supercoiled DNA
YLR399C	orf6.8227	BDF1	6	Protein with two bromodomain motifs and one ET domain; required for sporulation
YDL154W	orf6.2729	MSH5	6	Meiosis-specific protein involved in reciprocal recombination; component of late recombinational nodules
YDL042C	orf6.6367	SIR2	6	Protein involved in maintenance of silencing of HMR, HML and telomeres
YNL082W	orf6.2454	PMS1	6	Protein required for mismatch repair; E. coli MutL homolog
YPL153C	orf6.7322	RAD53	6	Serine/threonine/tyrosine protein kinase; functions as DNA damage checkpoints; required for viability of spore
YBR088C	orf6.7915	POL30	6	PCNA; required for DNA synthesis and DNA repair
YOR351C	orf6.7045	MEK1	6	Serine/threonine protein kinase required for meiotic recombination
YGL058W	orf6.8095	RAD6	6	Ubiquitin conjugating enzyme required for protein degradation and postreplication DNA repair
YML032C	orf6.5013	RAD52	6	Protein required for mitotic and meiotic recombination and repair of X-ray damage
YOR257W	orf6.7086	CDC31	6	Calmodulin-like calcium-binding protein of spindle pole body; involved in spindle pole body duplication
YER173W	orf6.2520	RAD24	6	Protein required for nucleotide excision repair and checkpoint control
YDR180W	orf6.3741	SCC2	6	Protein required for mitotic sister chromatid cohesion; cohesin
YDR004W	orf6.6089	RAD57	6	Protein involved in meiotic recombination and recombinational repair
YDR217C	orf6.3111	RAD9	6	DNA repair checkpoint required for cell cycle arrest following DNA damage
YIL072W	orf6.7201	HOP1	6	Component of the synaptonemal complex lateral elements; required for homologous chromosome synapsis and

				chiasmata formation
YNL216W	orf6.1679	RAP1	6	DNA-binding protein involved in silencing at telomeres and silent mating type loci; functions as activator or repressor
YHL022C	orf6.808	SPO11	6	Topoisomerase II-like protein; required for meiotic double- strand break formation
YLR263W	no match	RED1	6	Component of lateral elements of the synaptonemal complex; involved in homologous chromosome synapsis and chiasmata formation
YGL175C	no match	SAE2	6	Protein involved in meiotic recombination and chromosome metabolism
YDR369C	no match	XRS2	6	Protein required for meiotic recombination and DNA repair
YGL249W	no match	ZIP2	6	Protein involved in meiotic recombination and disjunction
YHR014W	no match	SPO13	6	Meiosis-specific protein required for meiosis I chromosome division
YMR133W	no match	REC114	6	Meiotic recombination protein required for double-strand break formation
YLR329W	no match	REC102	6	Meiotic recombination protein required for formation of the synaptonemal complex
YDR076W	no match	RAD55	6	Protein involved in meiotic recombination and recombinational repair
YHR157W	no match	REC104	6	Meiosis-specific protein required for synaptonemal complex formation and recombination
YNL210W	no match	MER1	6	Meiosis-specific protein required for splicing of MER2 mRNA and mRNA for other genes
YHR153C	no match	SPO16	6	Early meiotic protein required for spore formation
YER044C-A	no match	MEI4	6	Protein functions in early recombination and chromosome synapsis
YPL121C	no match	MEI5	6	Meiotic protein required for chromosome synapsis and meiotic recombination
YHR152W	no match	SPO12	6	Protein required for chromosome division in meiosis I
YHR079C-A	no match	SAE3	6	Protein involved in meiotic recombination and chromosome metabolism
YBR136W	orf6.4242	MEC1	6	Checkpoint protein required for DNA repair and mitotic and meiotic recombination
YJR021C	no match	MER2	6	Meiotic recombination protein
YOL104C	no match	NDJ1	6	Meiotic telomere protein involved in chromosomal synapsis and segregation
YGL033W	no match	HOP2	6	Component of synaptonemal complexes; required for pairing of homologous chromosomes in meiosis
YGL213C	orf6.6194	SKI8	6	Antiviral protein; required for meiotic recombination and spore viability
YGL195W	orf6.8322	GCN1	7	Protein required for activation of Gcn2p in response to starvation for amino

				acids or purines
YKR031C	orf6.1610	SPO14	7	Phospholipase D; required for meiosis and spore wall assembly
YOR190W	orf6.2395	SPR1	7	Sporulation-specific exo-beta-1,3-glucanase; involved in late sporulation
YDR403W	orf6.1352	DIT1	7	Enzyme required for biosynthesis of dityrosine in the outer layer of the spore wall
YDR523C	orf6.7786	SPS1	7	Serine/threonine protein kinase required for prospore formation; involved in middle/late stage of meiosis
YPR054W	orf6.8108	SMK1	7	Sporulation-specific MAP kinase; required for spore wall morphogenesis
YEL060C	orf6.1035	PRB1	7	Protease B; required for completion of sporulation
YGR059W	orf6.1204	SPR3	7	Sporulation-specific septin
YMR017W	orf6.1444	SPO20	7	DBF2-interacting protein; required for prospore membrane formation
YFL029C	orf6.2993	CAK1	7	Cdk-activating kinase involved in spore wall formation
YHR139C	no match	SPS100	7	Sporulation-specific protein involved in spore wall formation
YDR402C	orf6.1929	DIT2	7	Cytochrome P450 56, enzyme required for biosynthesis of dityrosine in the outer layer of the spore wall
YDR104C	orf6.5363	SPO71	7	Protein required for spore wall formation

Table 2
Candida albicans genes related to genes involved in cellular reproduction in various organisms. The abbreviations for organisms are Dm for *Drosophila melanogaster*, Sp for *Schizosaccharomyces pombe*, An for *Aspergillus nidulans* and Pa for *Podospora anserina*.

Organism	Gene	Ca ORF	Ca Gene	Function
Dm	DES1	orf6.4414	FAD3	Transmembrane protein; required for initiation of meiosis in fly, homologous to putative fatty acid desaturase in S. pombe
Dm	pelota	orf6.5174	DOM34	Required for spermatogenesis in fly; involved in meiosis, protein translation and pseudohyphal growth in S. cerevisiae
Sp	RCD1	orf6.8098	YNL288	Involved in sexual development in fission yeast
Sp	NRD1	orf6.2205	YPL184	RNA-binding protein, negative regulator of sexual differentiation in fission yeast
Sp	STE20	orf6.6015	YER093	Required for meiosis, amiloride sensitivity, pH regulation in S. pombe
Sp	MEI4	orf6.8625	FKH1	Meiosis-specific transcription factor with a forkhead domain
An	palI	orf6.491	RIM9	Aspergillus pH signal transduction component
An	pacC	orf6.8147	RIM101, PRR2	Zn finger protein, pH response regulator in Aspergillus
An	palA	orf6.3059	RIM20	Regulator of pH response in Aspergillus
An	palF	orf6.6250	RIM8	Regulator of pH response in Aspergillus
An	palB	orf6.7408	RIM13	Cysteine protease of calpain family
An	flbA	orf6.2767	SST2	Early regulator of asexual sporulation, normal colony development in A. nidulans
An	flbC	orf6.5156	YPR013	Putative Zn finger protein
An	flbD	orf6.7138	MYB1	Myb-like protein; initiates conidiophore development in Aspergillus
An	abaA	orf6.9129	TEC1	Developmental regulator with a TEF-1like DNA-binding domain and a potential leucine zipper domain
An	dopA	orf6.2672	DOP1	Required for correct cell morphology and sexual and asexual cycle development in A. nidulans; leucine zipper-like protein
An	stuA	orf6.4821	EFG1	Transcription factor with a DNA-binding domain related to Swi4, Mbp1, Phd1 in S. cerevisiae; involved in conidiophore development
Pa	HET-C	orf6.5520	HET1	Involved in vegetative incompatibility
Pa	HET-E-1	orf6.1133	YCR072	Protein with WD-40 domains; controls vegetative incompatibility
Pa	MOD-D	orf6.6451	GPA2	G alpha protein
Pa	MOD-E	orf6.7645	HSP90	Affecting sexual cycle and vegetative incompatibility

Chapter 3

Entry into meiosis in fungi: dissecting the potential regulatory circuit in *Candida albicans*

Introduction

In sexually reproducing eukaryotes, the meiotic cell cycle generates haploid gametes from diploid parental cells (1). During meiosis, cells undergo premeiotic DNA replication, chromosome pairing and recombination, followed by two consecutive rounds of chromosome segregation. Compared to mitosis, there are several features unique to meiotic chromosome behavior. In the first meiotic division (MI), homologous recombination joins maternal and paternal chromosomes, and sister centromeres cosegregate. Cells preserve cohesion of centromeric sister chromatids, and prevent DNA replication before the second meiotic division (MII) (2).

These fundamentals of meiosis appear to be highly conserved in nature. However, strategies for initiating the meiotic cell cycle can differ among organisms in order to allow reproduction to adapt to environmental cues. For example, in mammals spermatogenesis is controlled by an intrinsic genetic program, and regulated by interactive and extrinsic processes (3). The intrinsic genetic program coordinates germ cell development and neighboring cell functions. Sertoli cells and hormones such as follicle-stimulating hormone (FSH) and testosterone are instrumental in supporting germ cell development. In yeast species, the nutrient requirements for entering meiosis are quite diverse (4). For *Saccharomyces cerevisiae*, meiosis only occurs in diploid cells deprived of glucose and nitrogen, and grown in the presence of nonfermentable carbon sources (5). In *Schizosaccharomyces pombe*, nitrogen starvation is required for entering the sexual cycle (6). For the human pathogen, *Candida lusitaniae*, meiosis occurs upon depletion of ammonium and in the presence of glucose (7), while in the plant pathogen, *Ustilago maydis*, the sexual cycle is triggered when the fungus infects its host, *Zea mays* (8). Therefore, the nutritional conditions which initiate meiosis reflect the unique cell biology and habitat of each of these organisms.

In *S. cerevisiae*, sporulation involves a highly regulated transcriptional cascade in which distinct classes of meiosis-specific genes, driven by the interplay between transcription factors and protein kinases, are sequentially expressed (9, 10). In particular, *IME1* and *IME2* participate in the meiotic G1-S transition and *NDT80* in the G2-M transition (Figure 1, A). Ime1 is a transcriptional activator required for induction of early meiotic genes, and its expression is dependent on cell type and nutrients (11). Ume6, a Cys6Zn(II)2 DNA-binding protein, can recognize a URS1 element to act as a developmental switch to control vegetative repression and meiotic activation of meiosis-specific genes (12). The interaction of Ime1 with Ume6, essential for activating the transcription of meiosis-specific genes, requires both starvation and protein kinases Rim11 and Rim15 (13, 14). Ume6 represses early meiotic genes by recruitment of a Sin3-Rpd3 complex and targeted deacetylation (15) (Figure 1, B). *IME1* activates expression of *IME2*, which encodes a serine-threonine kinase (16). Ime2 is required for Ime1 destabilization, thus ensuring that Ime1 activity is restricted to a narrow window in early meiosis (17). *IME2* is a positive regulator of meiosis and is essential for premeiotic replication. The phosphorylation of Sic1, the primary CDK inhibitor, by Ime2 targets Sic1 degradation, thereby activating cyclin-dependent kinase (CDK) (18). *NDT80*, structurally related to *C. elegans* p53, is the key transcription factor responsible for expression of nearly 150 middle sporulation-specific genes (9, 19). In *S. cerevisiae*, phosporylation and expression of *NDT80* is dependent on Ime2, and is required for meiotic division and spore morphogenesis (20, 21).

C. albicans has been traditionally classified as a diploid, "imperfect" fungus, because a complete sexual cycle has never been discovered (22). Our comparative genomic analyses revealed that *C. albicans* lacked many of the genes that are either essential for, or closely associated with the *S. cerevisiae* meiotic machinery (summarized in Table

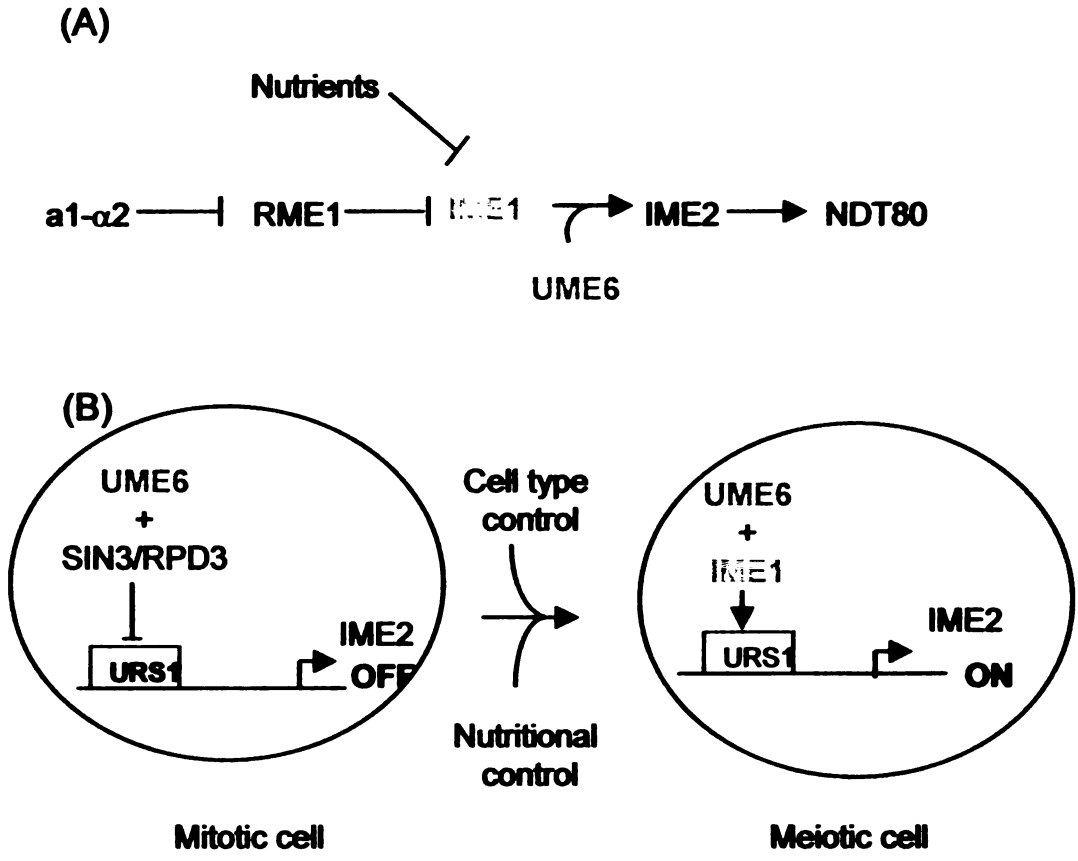


Figure 1. Transcriptional cascade of sporulation in *S. cerevisiae*.
A. Both nutrients and mating type control entry into meiosis by activating Ime1. Ime1 together with Ume6 activates the expression of early meiotic genes involved in DNA replication, chromosome synapsis and recombination. Ime2 is required for proper meiotic DNA replication, and for maximum induction of early, middle and late meiotic genes. *NDT80* expression is dependent on Ime1, and Ime2 can enhance Ndt80 activity by phosphorylation. Ndt80 turns on expression of the middle sporulation genes, which function in meiotic division and spore formation.
B. Ume6 binds to URS1, and recruits Sin3/Rpd3 histone deacetylase complexes to repress early meiotic genes in mitotic cells.

1). Possibly compensating for the missing *S. cerevisiae*-like functions, we identified homologues of meiotic genes found in other sexually reproducing organisms. The results of these genomic comparisons have led us to hypothesize that a complete sexual cycle can occur in *C. albicans* (23).

In this study, I have tried to characterize the potential regulatory network for meiotic initiation in *C. albicans*. Using the combined approach of genomic sequence analysis and functional heterologous complementation, I have shown that *IME1*, the master regulator for initiating meiosis in *S. cerevisiae*, is absent from *C. albicans*. I have further demonstrated that the key developmental switch for transition from mitosis to meiosis in *S. cerevisiae*, namely the interaction of Ume6 with the URS1 regulatory element does not exist in *C. albicans*. I developed a *C. albicans* strain in which *IME2* is conditionally expressed in order to test whether downstream targets of a potential meiotic pathway could be activated by *IME2* expression. *C. albicans IME2* expression is induced in *S. cerevisiae* sporulation medium (SPM), however no ensuing meiosis was observed. The phenomenon of "flocculation" was observed when *C. albicans* was transferred into SPM, indicating that growth under this condition induces cell-cell adhesion/interactions. These results suggest that there are significant regulatory differences in the expression and function of meiotic homologues between *S. cerevisiae* and *C. albicans*.

Materials and Methods

Yeast strains

C. albicans strains: SC5314; CAI4(Δ*ura3*::λ*imm434*/Δ*ura3*::λ*imm434*); *C. albicans* Robin Berkhout SGY243 (*ade2*Δ*ura3*::*ADE2*/*ade2* Δ*ura3*::*ADE2*); KWY100 (Δ*ura3*::λ*imm434*

/Δ*ura3*::λ*imm434*, *RP10*::p*MET3-IME2*) KWY101 (Δ*ura3*::λ*imm434*/Δ*ura3*::λ*imm434*, *RP10*::*pMET3*)

S. cerevisiae strains: Y442 (*MATa/MATα ade2-1/ade2-R8*); Y752 (*MATa/MATα ura3-52/ura3-52 ade2-1/ade2-R8 ime2::LEU2/ime2::LEU2*); W303 (*MATa/MATα ade2/ade2 can1-100/can1-100 his3-11,15/his3-11,15 leu2-3,112/leu2-3,112 trp1-1/trp1-1 UME6/UME6 ura3-1/ura3-1*); W303 *ume6Δ* (*MATa/MATα ade2/ade2 ade6/ADE6 can1-100/can1ADE2:CAN1 his3-11,15/his3-11,15 leu2-3,112/leu2-3,112 trp1-1/trp1-1 ume6D1/ume6D1 ura3-1/ura3-1*).

Media

All media were prepared as described (24), unless stated otherwise. Sporulation medium (SPM): 3 g potassium acetate and 0.2 g raffinose in 1 liter distilled water. Minimal acetate medium (PSP2): 6.7 g yeast nitrogen base (YNB) without amino acids, 1 g yeast extract and 10 g potassium acetate in 1 liter of 50 mM potassium phthalate buffer (pH 5.0). SPII: 2% potassium acetate. Synthetic minimum medium (SD): 0.67% YNB without amino acids, 2% Dextrose in 1 liter water. Synthetic complete medium (SCM) contains SD with complete supplement mixture (CSM). As *C. albicans* utilizes uridine not uracil, all media were supplemented with 0.097mg/ml uridine.

Initial cloning and gene disruption of a putative *C. albicans UME6*

A lambda zap library of *C. albicans* SC5314 strain genomic fragments (kind gift from B. Brun) was screened for the homologue of *S. cerevisiae UME6*, using a probe designed from the trace sequences provided by the *C. albicans* genome project. Probe hybridizations, plaque lifts, phage purifications, and plasmid excisions were carried out

according to standard protocols (25). For each screening, ~10000 plaques were screened and estimated to represent 2.5 to 5 genomic equivalents. Several independent clones were identified, and genomic inserts were sequenced at UCSF, Biomolecular Resource Center.

Disruption constructs for this putative *UME6* were constructed with the URA- blaster construct pMB-7 (26). The 5'-end of deletion construct primers 5'-CGGAAGCTT TGATGGTACACCTTAACCAAGAAA-3' and 5'-CGCCTGCAGAA GAGAACGGGAACAA AGTGA-3' yield a PCR fragment with *Hind* III and *Pst* I ends, and the 3' end of the deletion construct primers 5'-CGCAGATCTCGAATCAAAACAAAC GAGCC-3' and 5'-CCCGGTACCGCCGTGACTAACACAAAAACA-3' yield a PCR fragment with *Bgl* II and *Kpn* I ends. The fragments were cut and ligated sequentially into pMB-7. The deletion construct was released by cleavage with *Hind* III and *Kpn* I.

The deletion construct was transformed into *C. albicans* strain CAI4 using electroporation (27, 28). Correct targeted deletion strains were identified by Southern blot analysis. Transformed deletion strains were counter-selected to eliminate the *URA3* marker gene by growing on yeast synthetic plates containing 0.08% 5-fluoroorotic acid (5-FOA) and 0.097 mg/ml uridine.

C. albicans genomic library construction

Genomic DNA from *C. albicans* (Strain SGY243) was prepared by standard methods. The vector used in library construction was p16, a derivative of pYC2/CT (Invitrogen, Carlsbad, CA) but without the V5-6xHis C-terminal epitope tag. *Sau*3A I partially digested DNA was size-fractionated by agarose gel electrophoresis. Fragments between 1-6 Kbp were purified and ligated into the *Bam*H I site of the vector. Inserts

cloned into p16 are under the control of the *GAL1* promoter and can be propagated in yeast by taking advantage of the *URA3* marker. A CEN6/ARSH4 sequence allows non-integrative centromeric maintenance and low copy replication in *S. cerevisiae*. Transformation of the library into *E. coli* DH10B resulted in 1.2 x 10^6 primary transformants, 92% with inserts which averaged 2.2 Kbp in size; this represents approximately 150 genome equivalents.

Southern blot and Northern blot analyses

General molecular biology techniques were carried out according to standard protocols (29). Briefly, for Southern blot analysis, genomic DNA was isolated as described (30). Five μg of restriction enzyme digested genomic DNA sample was fractionated in a 0.8% agarose gel, blotted, and probed with labeled DNA fragments. For Northern blot analysis, total RNA was extracted as published (31). RNA samples were quantitated, and 10 μg of each sample was electrophoresed on a 1% formaldehyde gel, transferred to a nylon membrane, and probed with random- primed ^{32}P-PCR fragments. Blot hybridization and washing were according to Church and Gilbert (32).

Functional complementation for *S. cerevisiae ime1* deletion mutant strain

Diploid *S. cerevisiae* strain CGY52b was transformed with 40μg of the *C. albicans* library according to standard methods (33). Transformed cells were amplified overnight in the absence of uracil followed by a 5 day incubation in sporulation buffer (1% KOAc, 0.005% ZnOAc, 5mg/L adenine). On day 6 the transformation culture was treated with zymolyase (ICN Aurora, OH) to eliminate diploid cells, and then plated on germination medium (Synthetic Complete -URA, -ARG, 2% galactose, 60 μg/ml canavanine, 540 μg/ml G418) and allowed to incubate at 30°C for 4-5 days.

Construction of an inducible *C. albicans IME2* strain

For methionine-regulated expression of *IME2*, the entire ORF of *CaIME2* (ORF6.3678) was amplified by PCR. The PCR fragment was cloned into pCaEXP (34). The resulting plasmid was linearized at its *Stu*1 site, and transformed into the CAI4 strain using electroporation. Transformants were selected for uracil prototrophy. Correct integrants were identified by Southern analysis. The growth characteristics of several independent clones were determined on medium supplemented with a mixture of 2.5 mM methionine and 0.5 mM cysteine to repress the *MET3* promoter (34).

Sporulation assay

Growth and sporulation of *S. cerevisiae* were conducted at 30^0C as described (24). Briefly, late logarithmic phase diploid cells (~$1x10^7$ cells per ml) from PSP2 (minimum acetate medium) were washed, suspended in SPM or SPII supplemented with required amino acids, and sampled various times; samples were used for fluorescence-activated cell sorter (FACS) analysis and 4',6-diamidino-2-phenylindole (DAPI) staining.

***In silico* analysis**

Assembly 6 and Assembly 19 of *C. albicans* genome were constructed by Stanford Genome Technology Center (SGTC), and we compiled a database of intergenic sequences by extracting 1000 bp upstream of the designated start codon and 1000 bp downstream of the designated stop codon of each ORF, using a script which was written in Perl. From this intergenic database a string search for a particular motif present in regulatory regions was performed in Visual Basic 6. Moreover, we have performed a motif search in the translated open reading frames (ORFs, >100 amino acids) encoding predicted proteins from Assembly 6 by algorithms written in Visual Basic .net.

Reverse transcription-PCR

The RT-PCR reactions were performed using the Qiagen Onestep RT-PCR system. 200 ng DNase I-treated total RNA was used as template in each reaction. The primers for each gene are listed (Table 5). The reaction was carried out as follows: reverse transcription, 45 min 50^0C; initial PCR activation, 15 min 95^0C; 3-step cycling, 45sec 94^0C, 45sec 55-60^0C, 1min 72^0C, 30 cycles; 10 min 72^0C.

Results

Functional complementation of *S. cerevisiae ime1* mutant

Based on our comparative genomic analysis of meiosis-related genes, I concluded that no *IME1* homologue could be identified in *C. albicans* by *in silico* analysis alone. Therefore, we entered into a collaboration with Dr. John Swindle of CompleGen to see if it was possible to complement an *ime1* deletion mutant of *S. cerevisiae* using high throughput screening and selection methods (Figure 2). A *C. albicans* genomic expression library was transformed into the *S. cerevisiae ime1* deletion strain resulting in $1.4x10^7$ primary transformants, representing more than 10-fold coverage of the library. After 6 days of growth under sporulation conditions, no spores or tetrads were observed by microscopic examination. Transformed cells were then treated with zymolase to remove diploid cells and subsequently plated on germination medium. No colonies were observed following 4-5 days' incubation. Usually between 5 and 20 complementing plasmids for other gene deletions originating in organisms ranging from yeasts to human are detected in the CompleGen screening method. Considering the complexity of the library, it is reasonable to assert that *C. albicans* does not contain either a functional homologue or an analogue of *S. cerevisiae IME1*.

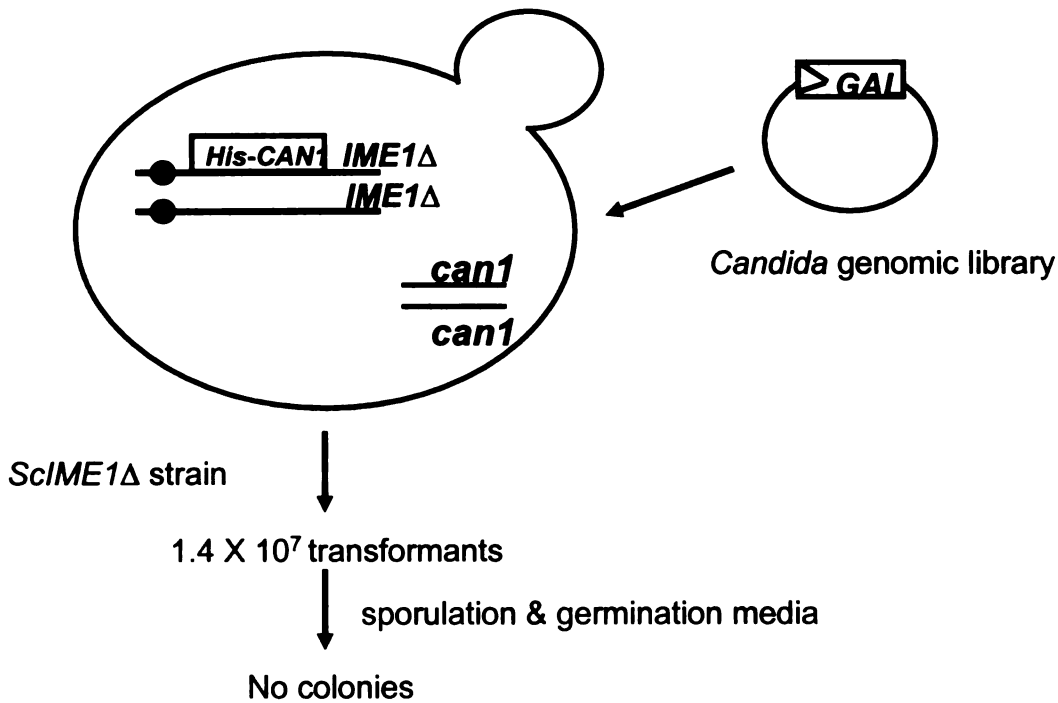


Figure 2. The strategy for functional complementation of *S. cerevisiae ime1* null mutant.

Diploid *S. cerevisiae* strain CGY52a (*ime1Δ/ime1Δ*) was modified by inserting a *HIS3-CAN1* suicide cassette between the YJR001W gene and centromere on chromosome X. The resulting strain, CGY52b was used in the complementation experiement. A *C. albicans* genomic library driven by a *GAL1* promoter was transformed into the CGY52b strain, was screened by diploid negative selection. No spores were identified after 5 days incubation in sporulation medium. Further, there was no colony formation when the zymolase-treated transformed cells were grown on germination plates for another 5 days. Unless otherwise stated, growth of yeast took place at 30°C, sporulation took place at 23°C.

In silico, expression and functional analysis of potential *UME6* homologues

In silico analysis, supported by genetic studies, failed to identify an *IME1* counterpart in *C. albicans*. However, when I began this study, early in the genome sequencing effort (2.5 x coverage), a *UME6* homologue had been assigned. Given the apparent absence of an *IME1* homologue from *C. albicans*, I was especially interested in characterizing *UME6* function. Therefore I went on to clone and characterize the candidate *UME6* gene from *C. albicans*. Knockout mutants (heterozygous and homozygous) were generated (Figure 3) and characterized using a variety of assays including a mating assay, morphological analysis, a reporter assay for testing *UME6* transcriptional repression activity and Northern blot analysis (Figure 4); no particular phenotype was observed.

As I completed the construction and characterization of the deletion strains for this particular *UME6* gene candidate, the genome project for *C. albicans* released its genome sequence of 10.4 X coverage. I therefore reexamined the interim genome Assembly 6 for ORFs containing the signature motif for Cys6Zn(II)2 binuclear cluster proteins [$CX_2CX_6CX_{5\text{-}16}CX_2CX_{6\text{-}8}C$] (35, 36). A total of 86 putative ORFs were identified; several of these were annotated as potential *UME6* homologues based on the presence of this sequence motif (Table 2). Proteins containing a Cys6Zn(II)2 domain are generally transcriptional regulators, and have only so far been identified in fungi. Overall the homology is relatively poor between *S. cerevisiae UME6* and all the candidate *UME6* genes containing this motif. The most highly homologous region lies in the Cys6Zn(II)2 DNA binding domain (Figure 5 A and B). I identified ORF6.887, 6.1326 and 6.4701 as the most likely candidate homologues of *UME6* in *C. albicans* because of additional sequence homology between these and *S. cerevisiae UME6* in regions outside of the Cys6Zn(II)2 domain. Of these, ORF6.887 and ORF6.4701 are allelic, and were the *UME6* candidate gene of my initial cloning and characterization.

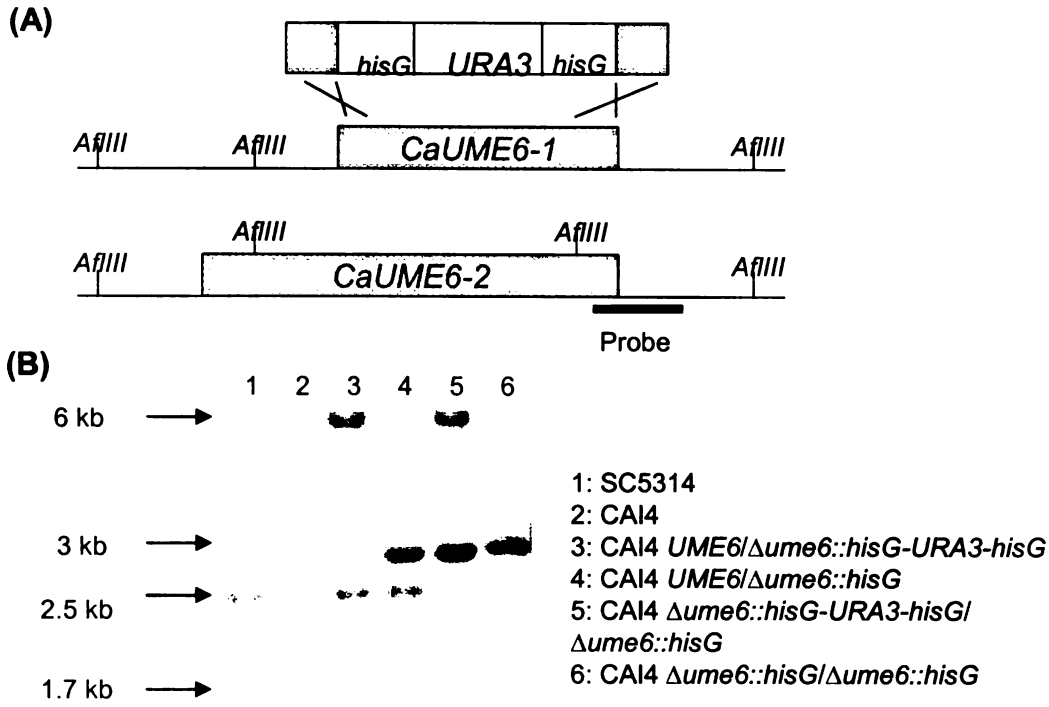


Figure 3. Sequential disruption of a putative *UME6* gene in *C. albicans*.

A. Schematic of a putative *CaUME6* (alleles: ORF6.887 & ORF6.4701) and disruption targets which were initially cloned in this work. Gene disruption of this *UME6* gene was achieved by URA-blaster strategy.
B. Correct targeted strains were confirmed by Southern analysis using *Afl*II-digested genomic DNA from *C. albicans* wild type strain SC5314 (lane1), CAI4 (lane2) and *UME6* heterozygous (lane3, lane4) and *ume6* homozygous (lane5, lane6) mutants, with a *UME6* fragment as identified probe shown in the figure 3A.

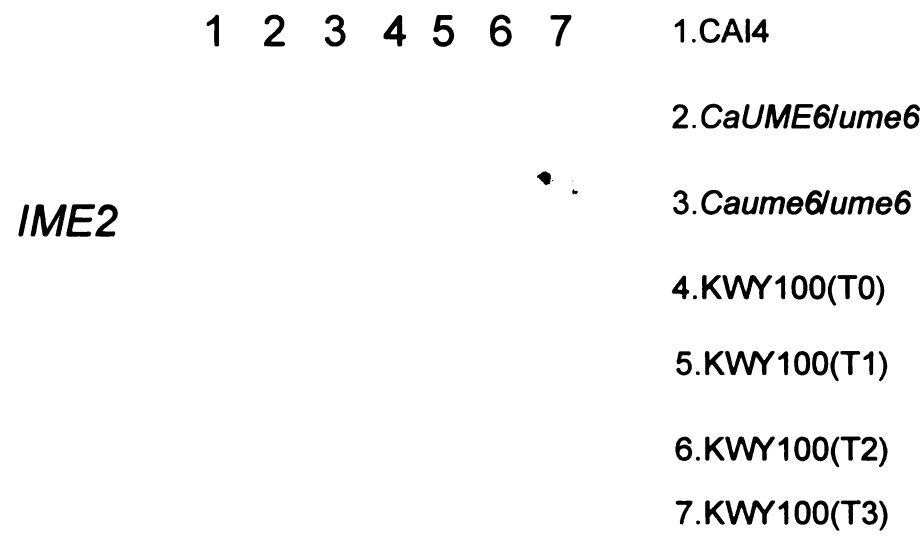


Figure 4. Northern blot analysis of *C. albicans IME2* expression.

Northern blot analysis was performed using RNA isolated from *C.albicans* CAI4, *UME6* heterozygous (*UME6/ume6*) and homozygous (*ume6/ume6*) mutant strains grown in YEPD during vegetative growth (OD_{600}1.0). Similarly, Northern blot was analyzed using RNA harvested from KWY100 (RP10::p*MET3-IME2*) grown in SCM to OD_{600}0.8 (T0) in the presence of 2.5 mM methionine and 0.5 mM cysteine, washed, and transferred to Met- and Cys-free medium. Samples were collected at various times post shift (T1:2hrs,T2:4hrs,T3:6hrs). The Northern blot was probed with *CaIME2*. In *S. cerevisiae*, *IME2* expression is upregulated in *ume6* null mutants. There was no *IME2* expression in CAI4 (lane1), *UME6* heterozygous mutant (lane 2) and homozygous mutant (lane 3). *CaIME2* was expressed after shifted to Met- and Cys-free medium (lane5-7) (see also Figure 10).

A

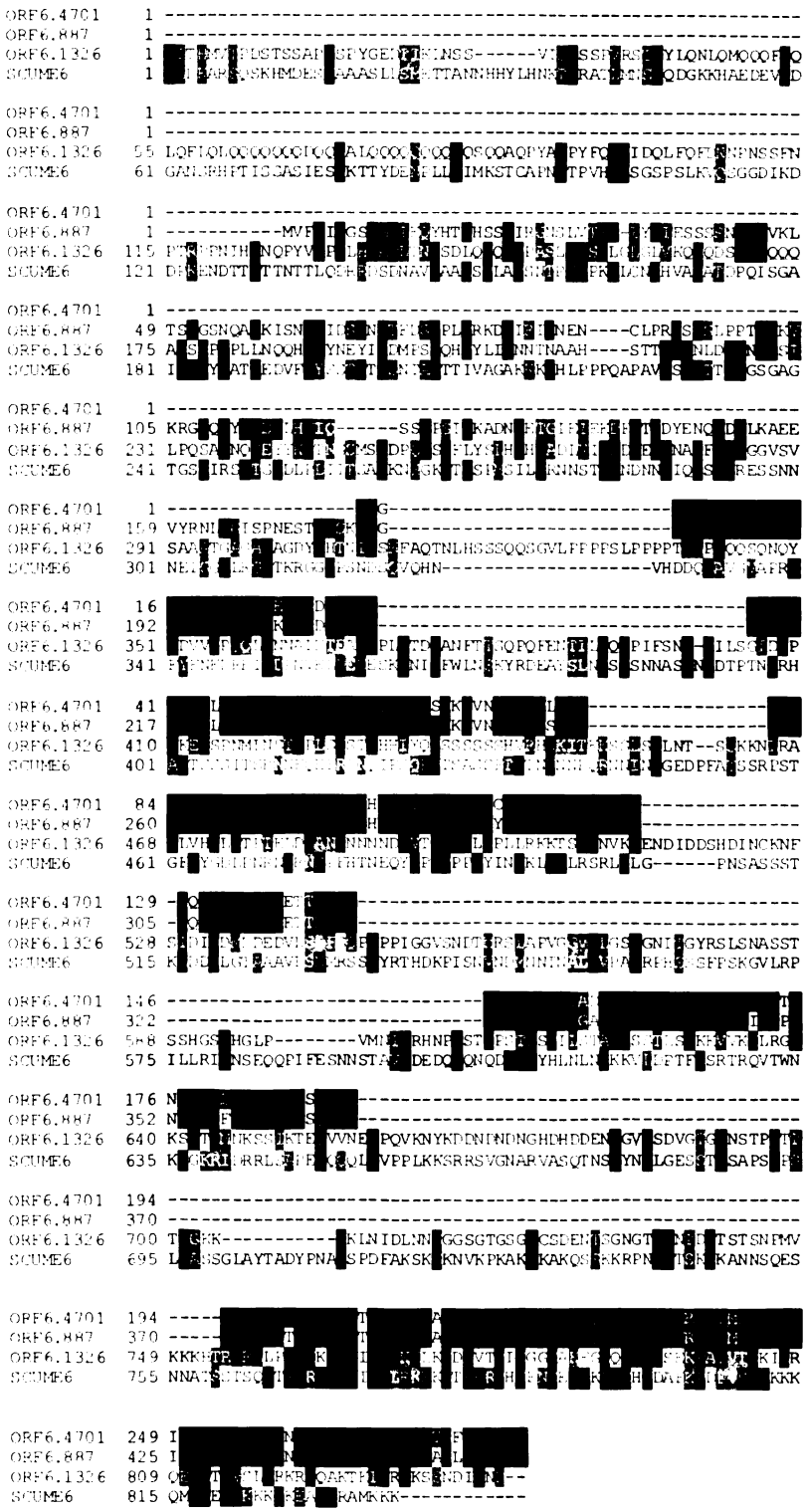


Figure 5 A. Sequence comparison between putative *UME6* homologues in *C. albicans*.

B.

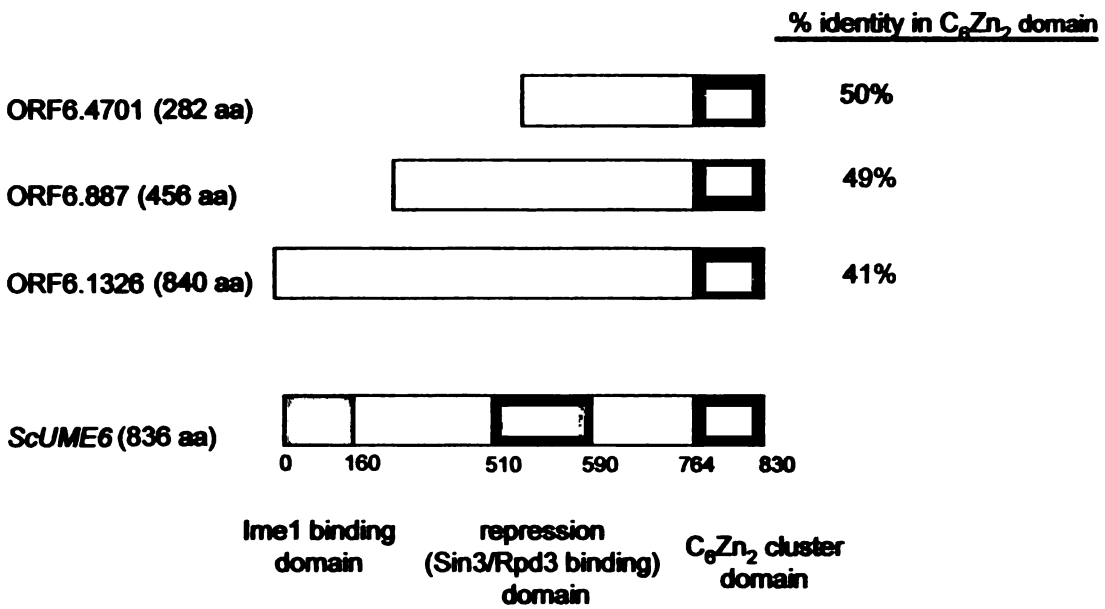


Figure 5 B. Schematic depiction of potential *UME6* homologues in *C. albicans*.

In *S. cerevisiae*, Ume6 contains three major domains including the Ime1 binding domain, Sin3 binding domain and C6Zn2 DNA binding domain. Blast analyses suggest there are three putative *UME6* candidates in *C. albicans*; these are ORF6.887, 6.1326 and 6.4701. However, sequence similarity between these ORFs and *S. cerevisiae* Ume6 lies mainly in the C6Zn2 DNA binding domain.

To better understand the characteristics of the potential *UME6* ORFs, I first used DNA microarray data available from our laboratory to examine the respective expression patterns of *UME6* candidate RNAs in different growth conditions. In summary, ORF6.887, while transcriptionally silent in a variety of conditions, was expressed when iron concentrations in the medium were low (5-10 μM). ORF6.887 and 6.4701 were transcribed when cells were grown in rich medium (YEPD); as assessed by RT-PCR. ORF6.1326 was expressed in every stage of biofilm formation, and was down-regulated in *C. albicans kex2* mutant which was grown in F12 medium in the presence of CO_2. ORF6.1326 was also up-regulated during the yeast to hyphae transition and in a *C. albicans* phospholipase B (*plb3*) deletion mutant. None of these data were particularly informative in assigning a more likely *UME6* candidate among these ORFs.

To further determine which potential *UME6* ORF might be a functional homologue of the *S. cerevisiae UME6*, I assessed the ability of each of these *UME6* candidates to functionally complement the *S. cerevisiae ume6* deletion mutant using sporulation as the assay (12). I generated expression plasmids from ORF6.4701, 6.887, 6.5864, 6.6594 and 6.1326; each was placed under the control of *S. cerevisiae IME2* promoter, and transformed into a *S. cerevisiae ume6* deletion mutant. After growth in sporulation medium for 7 days, the transformed cells were still arrested in G2 phase; the cells were mononucleate as determined by DAPI staining. The positive control, complementation of the *S. cerevisiae* deletion mutant with *S. cerevisiae UME6* expressed from the *IME2* promoter showed a multinucleate staining pattern (Figure 6). RT-PCR analyses showed that these putative *UME6* ORFs and *S. cerevisiae UME6* were expressed in sporulation conditions (data not shown). Accordingly I concluded that none of the potential *UME6* ORFs can complement the *S. cerevisiae ume6*Δ defect. Taken together, the results

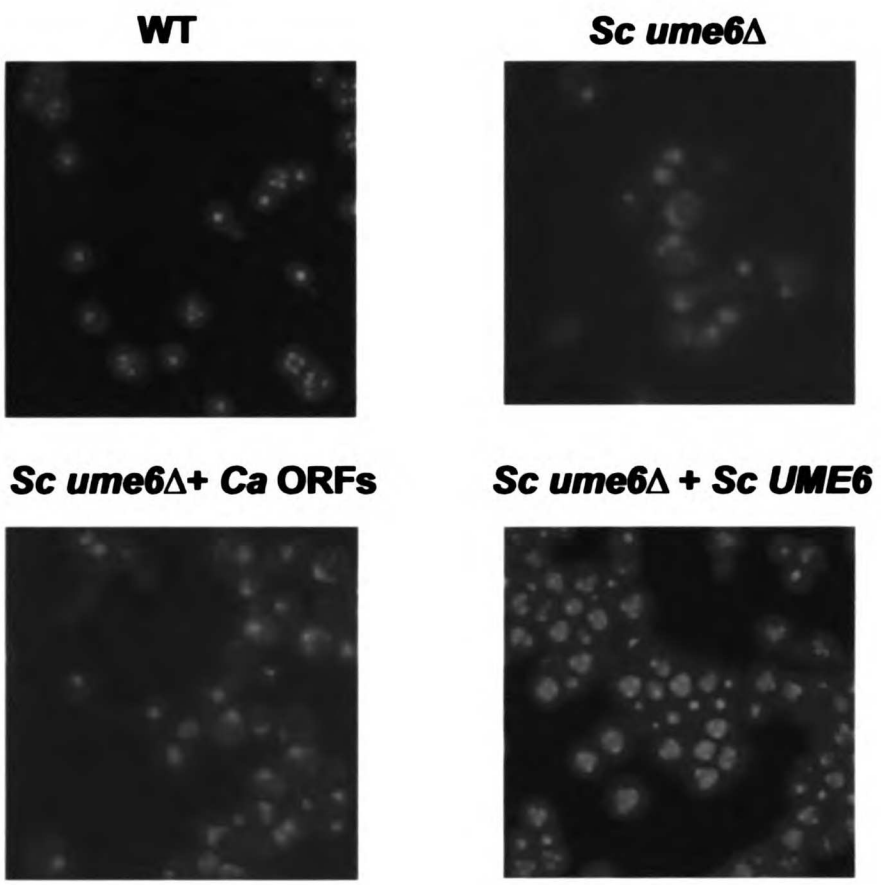


Figure 6. Functional complementation of *S. cerevisiae ume6* null mutants.
Growth and sporulation were conducted at 30^{0}C as described. Late logarithmic phase cells from *S. cerevisiae* strains were washed, transferred to SPII (2% potassium acetate) and samples were collected at various time points to assay several meiotic events. *S. cerevisiae UME6* is essential for MI and MII. Analysis of DAPI-stained nuclei after 7 days of sporulation suggests that the majority of cells in the *ume6* mutant were arrested in G2 with one staining nucleus. The *S.cerevisiae ume6* mutant transformed with various *C. albicans* putative *UME6* ORFs showed a very similar staining pattern to the *ume6* mutant.In the positive control, the *ume6* mutant transformed with *S. cerevisiae UME6* displayed the wild type phenotype with the majority of cells showing 3 or 4 staining nuclei.

suggest that there is no *UME6* homologue in *C. albicans* which can substitute *S. cerevisiae UME6* for its meiotic function.

Genome-wide search for URS1 sites in *C. albicans* intergenic regions

In *S. cerevisiae*, Ume6 can recognize the URS1 site to function as a developmental switch for vegetative repression and meiotic induction of early meiotic genes. In order to understand the regulatory mechanism of expression of potential meiotic homologues in *C. albicans*, I searched the entire genome sequence of *C. albicans* for potential regulatory elements. An intergenic database was created by extracting 1000 bp of sequence both upstream and downstream of each ORF assigned in Assembly 6 of the *C. albicans* genome. First, a string search was used to locate the URS1 site "5'-GGCGGCNAN-3" in the intergenic database (37). Second, I also compiled a list of potential meiotic gene homologues and then searched upstream sequences of this subset by using MEME to identify possible consensus motifs (38). In both instances, I was not able to identify either a conserved URS1-like binding site or shared motifs common to the promoters of these putative meiotic homologues.

Sequence, expression and functional analysis of *C. albicans IME2*-like protein

In *S. cerevisiae IME2* encodes a meiosis-specific kinase which is essential for premeiotic DNA replication, nuclear division and ascus formation. Sequence analysis (BLAST) of the *C. albicans* genome identified a potential *IME2* homologue which shares significant sequence similarity with *S. cerevisiae IME2*, *S. pombe mde3+*, *crk1* of *U. maydis*, *Neurospora crassa IME2* and human MAK. For *S. pombe*, *mde3+* is expressed in meiosis, and is involved in spore formation (39). In *U. maydis*, *crk1* is required for morphogenesis (40). In humans, MAK (male germ cell-associated kinase) is strictly expressed in testicular germ cells; however mouse knockout experiments suggest it is

not essential for spermatogenesis (41). The sequence similarities between *C. albicans IME2* and those *IME2* homologues are found mainly in the N- terminal region encoding the kinase domain (Figure 7). Thus *C. albicans IME2* appears to belong to a conserved family of protein kinases which function in determining cell fate.

To further explore the biological function of *C. albicans IME2*, I performed a functional complementation in *S. cerevisiae*. An *ime2* deletion mutant or a yeast transformant with a centromeric plasmid carrying the *C. albicans IME2* ORF driven by the *S. cerevisiae IME2* promoter was grown under sporulation conditions. Samples were taken at various times to evaluate two major meiotic events: (1) DNA replication, which was determined by FACS analysis, and (2) completion of meiosis and tetrad formation, which was determined by DAPI staining and microscopic observation (42). Based on FACS analysis (Figure 8), after 2 -12 hours of incubation in SPM, the *ime2* mutant was arrested in G1 phase with a DNA content primarily of 2N. The yeast transformant carrying p*ScIME2-CaIME2* plasmid displayed a very similar DNA content profile to that of the *S. cerevisiae ime2* deletion strain, suggesting that *C. albicans IME2* cannot rescue premeiotic DNA replication in the *S. cerevisiae ime2* mutant. No evidence of spore formation was observed, even after 48 hours incubation in SPM, in either the *ime2* deletion mutant or the yeast transformant carrying the *C. albicans IME2* plasmid. Taken together, these results indicate that despite significant sequence homology, *C. albicans IME2* is not a functional homologue of *S. cerevisiae IME2*.

In order to shed some light onto the possible function of *IME2*, I have also examined DNA microarray data available in our laboratory from cells grown in a variety of conditions. *C. albicans IME2* was not expressed in most conditions which suggests that there might be a special nutrient requirement for its transcription. To determine how

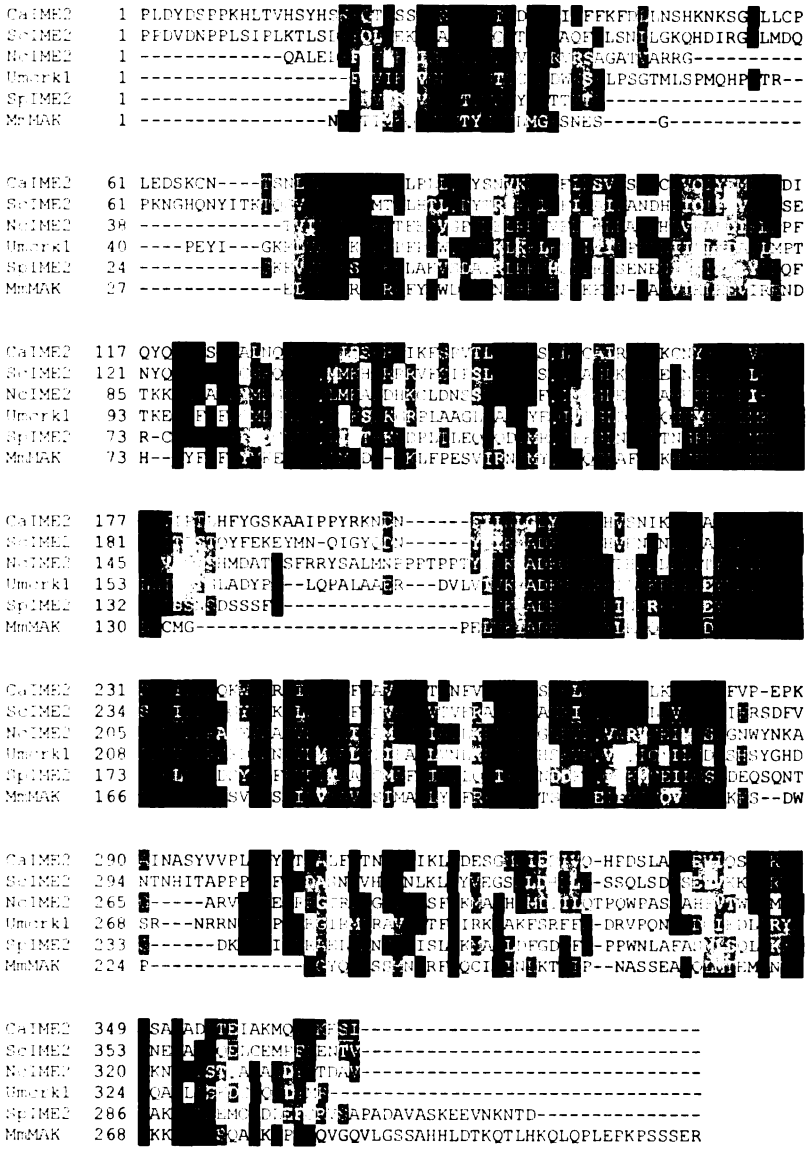


Figure 7. *C. albicans* Ime2 is a serine/threonine kinase.

Comparison of the predicted amino acid sequence encoded by *C. albicans IME2* with *S. cerevisae* Ime2, *S. pombe* Mde3, *N. crassa* Ime2- like protein, *U. maydis* Ime2 and mouse Mak shows high similarity among these proteins. Only the kinase domains were used for the alignment, which was generated by Clustal W1.8.

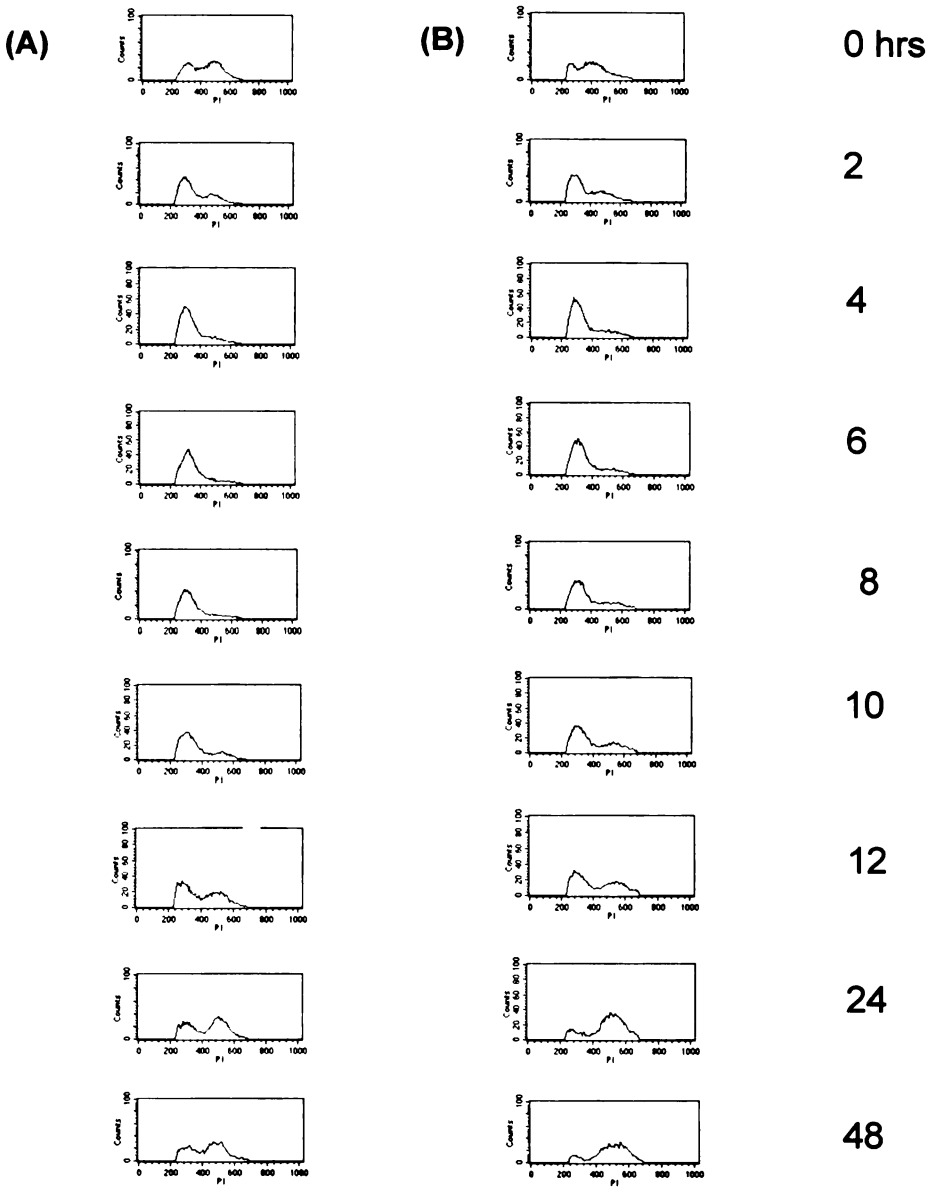


Figure 8. *C. albicans IME2* is not a functional homologue of *S. cerevisiae IME2*.

Expression of *C. albicans IME2* in early meiosis is not able to rescue a premeiotic DNA replication defect in *S. cerevisiae ime2*Δ diploid cells. *S. cerevisiae ime2* deletion mutant (Y752) and Y752 carrying a pSc*IME2-CaIME2* on a centromeric plasmid were grown at 30^{0}C in PSP2 to stationary phase (1X10^7 cells/ml). Cells were washed, suspended in SPM, and incubated at 30^{0}C. Samples were taken at the indicated times for FACS analysis. Both scenarios (A: *ime2*Δ mutant, B: *ime2*Δ mutant+pSc*IME2-CaIME2*) displayed a very similar DNA content profile.

C. albicans IME2 expression is regulated, I compared the promoter of the *C. albicans IME2* with the promoters of *S. cerevisiae IME2* and *S. pombe med3+*. There are several regulatory sites found in the promoter of *S. cerevisiae IME2*, including UAS, URS1 and T4C (43). There is a FLEX- like element (-5'-GTAAACAAACA-3') present in *S. pombe mde3+* promoter (39). I have not identified a potential URS1 site or FLEX-like element in the promoter of *C. albicans IME2*, however a potential T4C site was found. These results suggest the difference in transcriptional regulation of *IME2* between *C. albicans* and other yeasts.

Conditional expression of *C. albicans IME2*

In *S. cerevisiae*, increased gene dosage of *IME2* enables meiosis in an *ime1* deletion mutant. To test the hypothesis that *C. albicans IME2* can function as a key transcriptional regulator of a potential meiotic pathway, *C. albicans* strains were constructed in which an additional copy of *C. albicans IME2* was placed under the control of the inducible *MET3* promoter, and integrated at the *RP10* locus (34) (Figure 9). To validate the expression of *IME2* from the *MET3* promoter, I also created a strain where the *MET3* vector alone was targeted at the *RP10* locus. Correct integrants were identified by Southern analysis. As *C. albicans IME2* is not expressed in most conditions, I first attempted to modulate its expression during vegetative growth. I cultured the strains in synthetic complete medium in the presence of methionine (Met) and cysteine (Cys), and then shifted the strains into Met- and Cys-free medium in exponentially growing cells. RNA was harvested from cells at various times, and RT-PCR was used to assay the expression of several key meiotic homologues.

In agreement with DNA microarray experiments, I found that *C. albicans IME2* was not expressed in vegetative cells in the strain containing the integrated vector alone. As

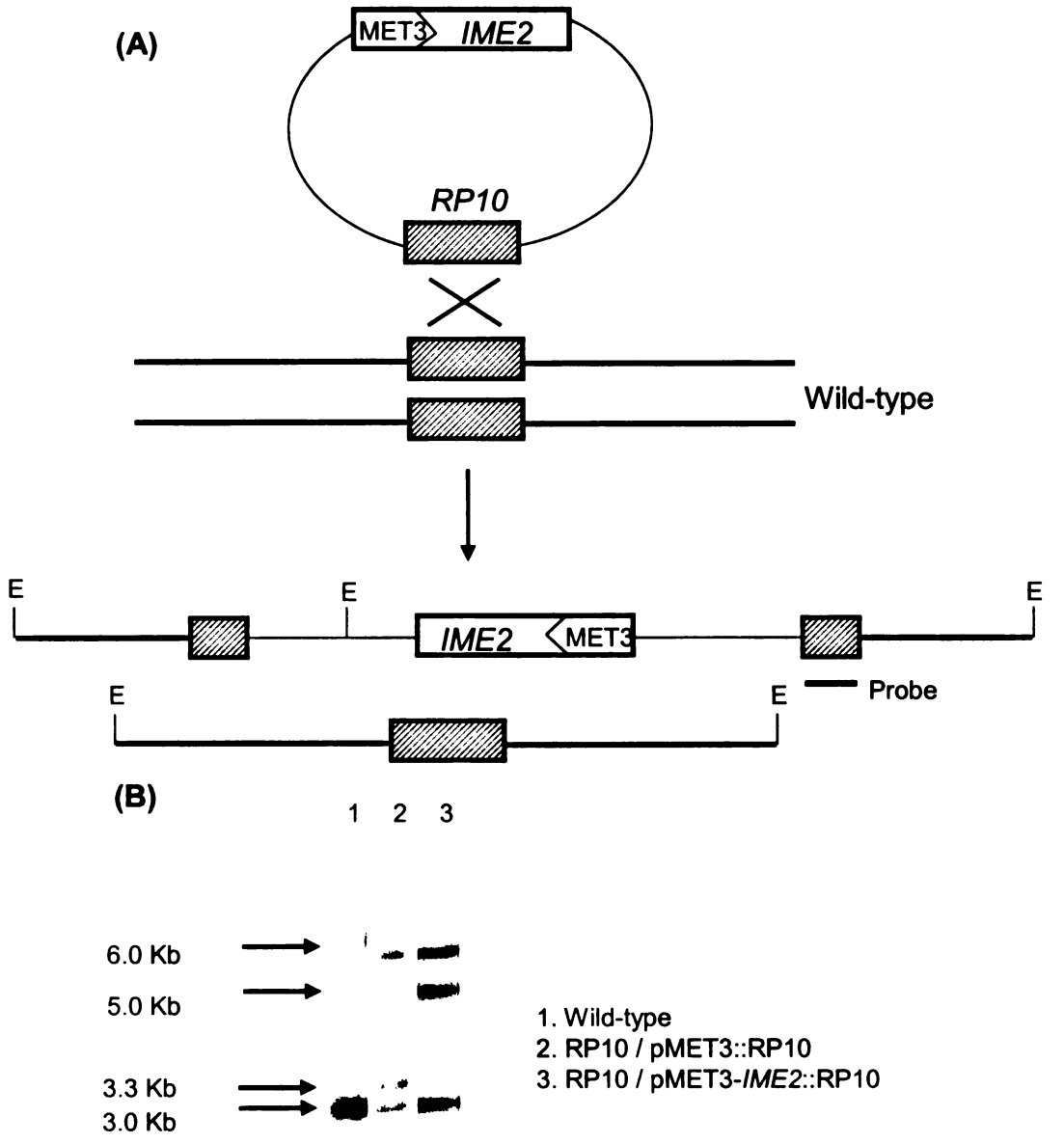


Figure 9. Construction of inducible *C. albicans IME2* strains

A. Scheme depicts conditional expression *IME2* strain constructions where inducible *IME2* was targeted to the *RP10* locus. Correct integrants were identified by Southern blot analysis.
B. Five micrograms each of total genomic DNA from the indicated strains was digested with *EcoR* I and analyzed by Southern blotting. The blot was probed with a PCR-amplified *RP10* fragment.

expected, *IME2* was expressed after cells were transferred into Met-and Cys- free medium (Figure 10). In *S. cerevisiae*, *NDT80* is the key transcriptional activator for middle sporulation genes. *IME2* is essential for expression and phosphorylation of Ndt80. Contrary to the situation in *S. cerevisiae*, where there is a single *NDT80* gene, the *C. albicans* genome sequence revealed two potential *NDT80* homologues, ORF6.4742 and 6.6725 (Figure 11). To test the possibility that conditional expression of *C. albicans IME2* might identify those potential *NDT80* homologues which may be associated with the meiotic pathway, I examined *C. albicans NDT80* expression by RT-PCR. I detected expression of both potential *NDT80* ORFs using this assay (Figure 12). In available DNA microarray data, ORF6.4742 was constitutively expressed, and ORF6.6725 was not expressed under most conditions. Similar analyses were performed to assess expression of middle sporulation gene homologues *SPS1* and *SPS4*; these are *NDT80*- regulated genes in *S. cerevisiae* (21). While *SPS1* was expressed in low abundance; no *SPS4* expression was detected in this assay (Figure 13). Taken together, these results suggest that *NDT80* expression is not dependent on *IME2*, and expression of *SPS1* and *SPS4* is not dependent on *NDT80* in *C. albicans*. I further compared the upstream promoter regions of ORF6.4742 and ORF6.6725 with that of *S. cerevisiae NDT80*. There are URS1 sites and middle sporulation elements (MSE; gNCRCAAAA/T) present within the upstream sequence of *S. cerevisiae NDT80* (44). I have located a URS1 site in the upstream region of ORF6.6725, however no potential MSE was identified in the promoter regions of either potential *NDT80* ORF. A similar analysis was performed to compare promoters of putative *SPS1* and *SPS4* with their *S. cerevisiae* counterparts. I could not identify MSE sequences within the promoter region of *C. albicans SPS1* or *SPS4*. In aggregate, results from these "phylogenetic footprinting" analyses bolster the notion that transcriptional regulation of the meiotic pathway is different in *C. albicans* and *S. cerevisiae*.

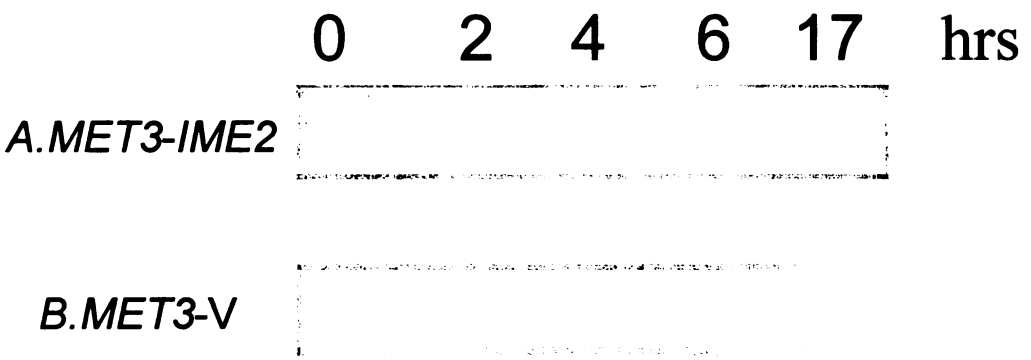


Figure 10. Induction of *IME2* expression during vegetative growth.

RT-PCR analysis was performed using RNA from *C. albicans* strains A.KWY100 (*RP10*::p*MET3-IME2*) and B.KWY101(*RP10::pMET3*) grown to OD_{600} 0.8 (time 0)in SCM in the presence of 2.5 mM Met and 0.5 mM Cys to O.D. 0.8, washed, and transferred to Met- and Cys-free SCM. Samples were collected at the indicated times post shift. As shown, *C. albicans IME2* was expressed after shifted to Met- and Cys free medium. Very low levels of expression was detected in the control strain KWY101, indicating basal level of *IME2* expression. Absence of DNA contamination in RNA samples was confirmed by RT-PCR(data not shown).

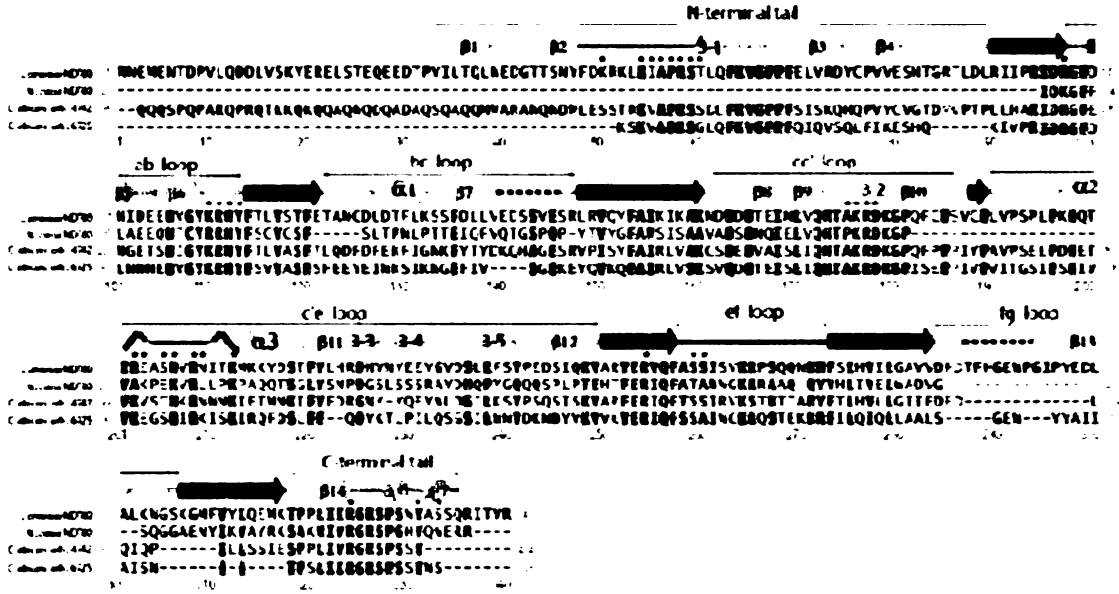


Figure 11. Sequence alignment of the DNA-binding domain of Ndt80 (Lamoureux et al. 2002)

There are two ORFs annotated as *NDT80* in the *C. albicans* genome, ORF6.4742 and 6.6725. Here, these two *C. albicans* ORFs are aligned with *S. cerevisiae* Ndt80 and *Neurospora crassa* Ndt80 that share high sequence similarity with one another. The regions highlighted in green share greater than 75% sequence identity. Amino acid residues with stars above them are involved in DNA contacts.

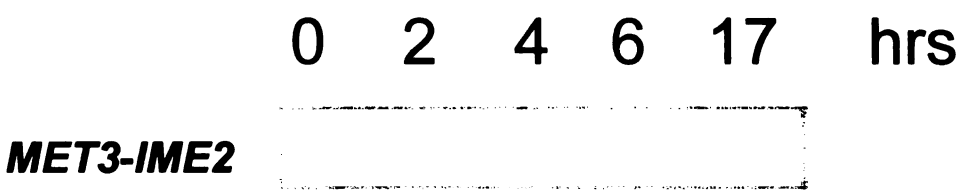


Figure 12. Expression of *C. albicans NDT80* during vegetative growth.

RT-PCR was performed to detect the expression of the *NDT80* homologues in *C. albicans*. RNA was isolated from strain KWY100 grown in SCM to OD_{600}0.8 (time 0) in the presence of 2.5 mM Met and 0.5 mM Cys. Cells were washed, transferred to Met- and Cys-free SCM, and samples were collected at the indicated times. Expression of both ORFs was detected by this assay.

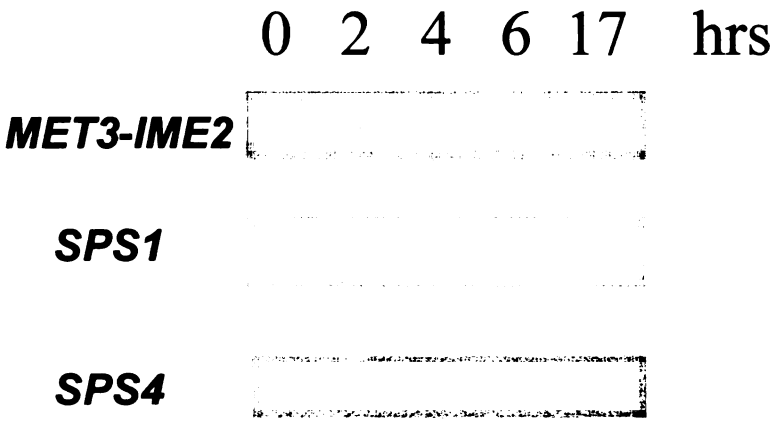


Figure 13. Expression of potential *NDT80*-regulated genes during vegetative growth.

RT-PCR analysis was performed to detect the expression of putative *SPS1* and *SPS4* homologues in *C. albicans* under conditions of *C. albicans IME2* expression. RNA was isolated and samples were taken as described in legend of Figure 12. As shown, *SPS1* was expressed in very low abundance, and *SPS4* was not expressed.

Nutritional regulation of *C. albicans IME2*

In *S. cerevisiae*, glucose inhibits meiosis by turning off *IME1* and *IME2* transcription, while acetate will activate *IME1* and *IME2* expression through *SNF1* kinase (45). In *C. albicans IME2* was not expressed in vegetative cells grown in glucose medium, therefore I tested the possibility that either cell cycle arrest or acetate might activate *C. albicans IME2* expression. *C. albicans* wild type SC5314 cells growing exponentially in glucose-based rich medium (YEPD) were allowed to reach a high cell density and accumulated in stationary phase. On transfer to SPM, cells were harvested at various times and RNA prepared for RT-PCR assay. In stationary phase, there was no *IME2* expression which suggests nutritional starvation and cell cycle arrest are not sufficient to stimulate *IME2* expression. *IME2* was expressed after shifting cultures to SPM; maximum levels of expression were induced after 6 to 9 hours (Figure 14). This observation implies that nonfermentable carbon sources can activate *IME2* transcription in *C. albicans*.

Transcriptional analysis of *C. albicans* grown in sporulation medium

The nutritional environment is a primary determinant of cellular behavior (46). In SPM, *C. albicans* activates *IME2* expression, however *C. albicans* does not initiate meiosis. To further characterize how Ime2 might coordinate transcription in SPM, I performed RT-PCR assays to detect the possible expression of several key meiotic homologues which are dependent on *IME2* expression in the *S. cerevisiae* meiotic cascade. These include candidate genes involved in cell type and nutritional control, the core meiotic recombination machinery, the synaptonemal complex and meiotic chromosome behaviors.

I first checked *IME4* expression (Figure14). Ime4, an RNA methyltransferase, mediates cell type and nutritional regulation for meiosis, and is expressed in SPM in *S. cerevisiae*.

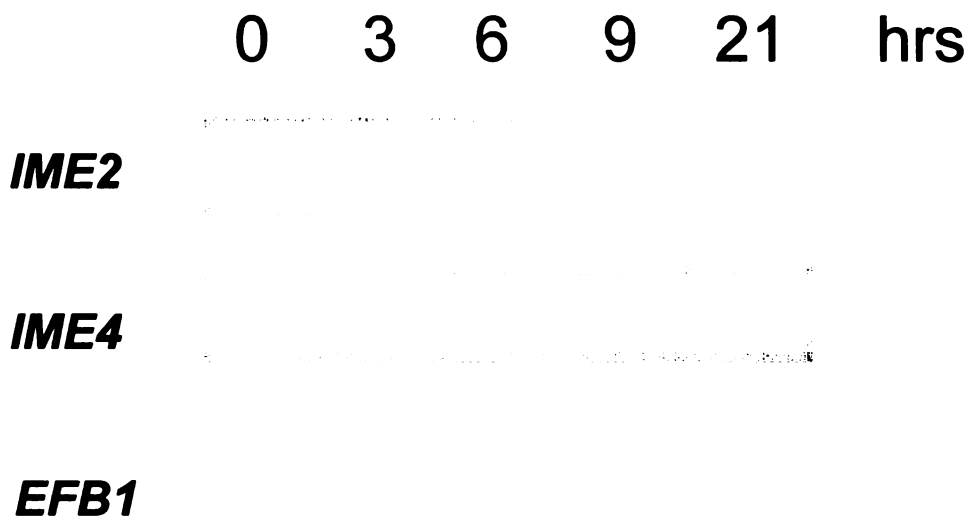


Figure 14. Expression of *C. albicans IME2* and *IME4* in SPM.

RT-PCR was used to assay expression of putative *IME2* and *IME4* homologues when *C. albicans* was grown in SPM. RNA was harvested from *C. albicans* SC5314 strain grown in YEPD to reach stationary phase (time 0), washed, then transferred to SPM, and samples were collected at the indicated times. As shown, *C. albicans IME2* was expressed in SPM, and no *IME4* expression was detected.

It has been hypothesized that *IME4* might methylate a subset of early meiotic genes to activate meiosis (47). I did not detect any *IME4* expression in SPM or in vegetative cells, an observation which serve to highlight one of the regulatory differences in the signaling network between *C. albicans* and *S. cerevisiae*.

Recombination occurs at a much higher frequency during meiosis than during mitosis. I therefore asked whether candidate genes in the core meiotic recombination machinery would be activated as they are know to be in *S. cerevisiae* in response to *IME2* expression. The genes which I analyzed included *SPO11*, *RAD50*, *DLH1*, *MRE11*, *RAD51*, *MSH4*, *MSH5* and *MLH1* (48). In summary, I detected the expression of *SPO11*, *MLH1*, *MRE11*, *RAD50*, *RAD51* and *MSH4*. *MSH5* was expressed at very low levels, and *DLH1* (*DMC1* homologue) expression was barely detectable (Figure 15). I have also detected *SPO11* and *MSH4* expression in vegetative cells, and their expression does not appear to be controlled by *IME2*. No *MSH5* or *DLH1* expression was detected during vegetative growth (Figure 16). This result suggests that genes which encode the core meiotic recombination machinery in *C. albicans* are expressed in SPM, however SPM does not activate these genes.

In the initial comparative genomic analysis (23), I concluded that *C. albicans* is not able to make mature synaptonemal complexes (SC) due to the lack of several molecular components. I have identified putative *HOP1* and *ZIP1* homologues, which are important for homologous chromosome synapsis and SC formation in *S. cerevisiae* (1). RT-PCR showed that *HOP1* was not expressed in SPM or in vegetative cells. On the contrary, *ZIP1* was constitutively expressed through the mitotic cell cycle and in SPM (Figure 17). Further BLAST analysis of *C. albicans ZIP1* showed that it bears very weak homology to

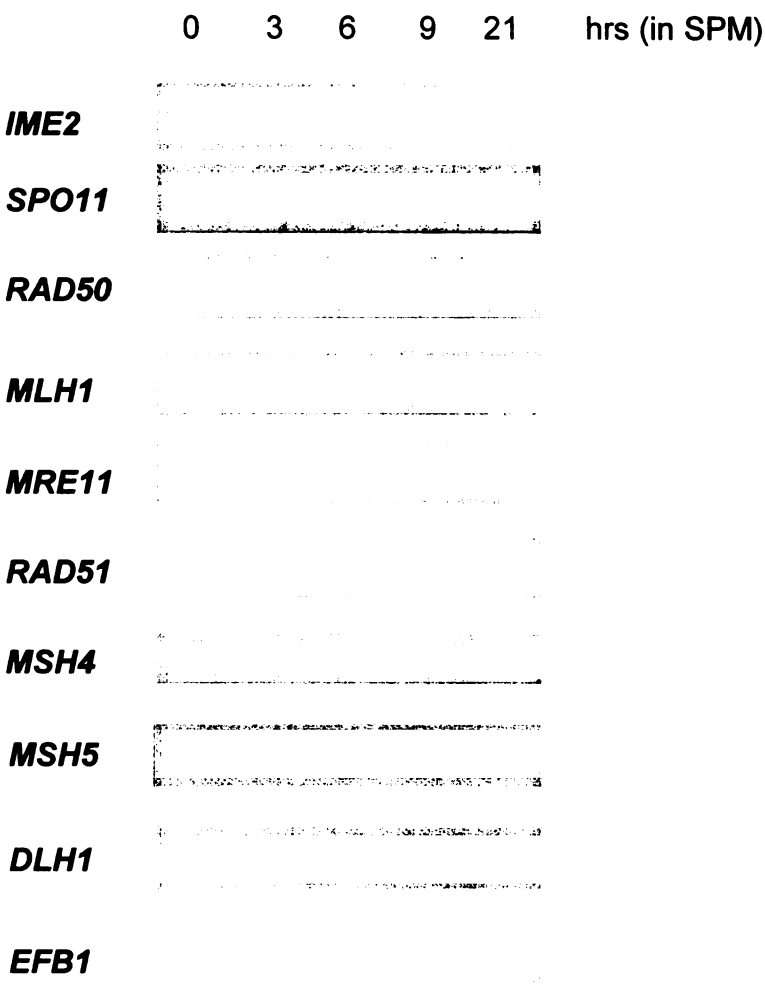


Figure 15. Expression of the core meiotic recombination machinery in SPM.

RT-PCR was used to assess expression in SPM of gene homologues associated with the core meiotic recombination machinery. RNA was isolated from SC5314 strain grown in YEPD to reach stationary phase, washed, and transferred to SPM. Samples were collected at various times. As shown, most of genes were expressed, except for *MSH5* and *DLH1*. These results suggest that *C. albicans* homologues of *S. cerevisiae* core meiotic machinery are active in SPM.

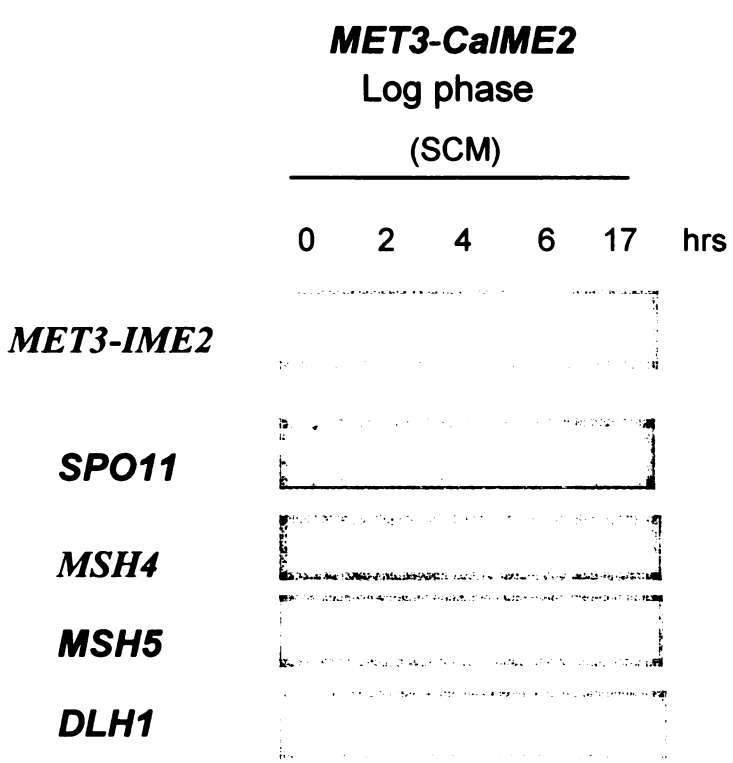


Figure 16. Expression of *C. albicans* putative meiotic homologues during vegetative growth.

RT-PCR was performed to detect expression of several putative meiotic homologues in vegetative cells. RNA was isolated from strain KWY100 grown to OD_{600} 0.8 (time 0) in SCM in the presence of 2.5 mM Met and 0.5 mM Cys, washed and transferred to Met- and Cys-free SCM. Samples were collected at the indicated times. Both *SPO11* and *MSH4* were expressed during vegetative growth.

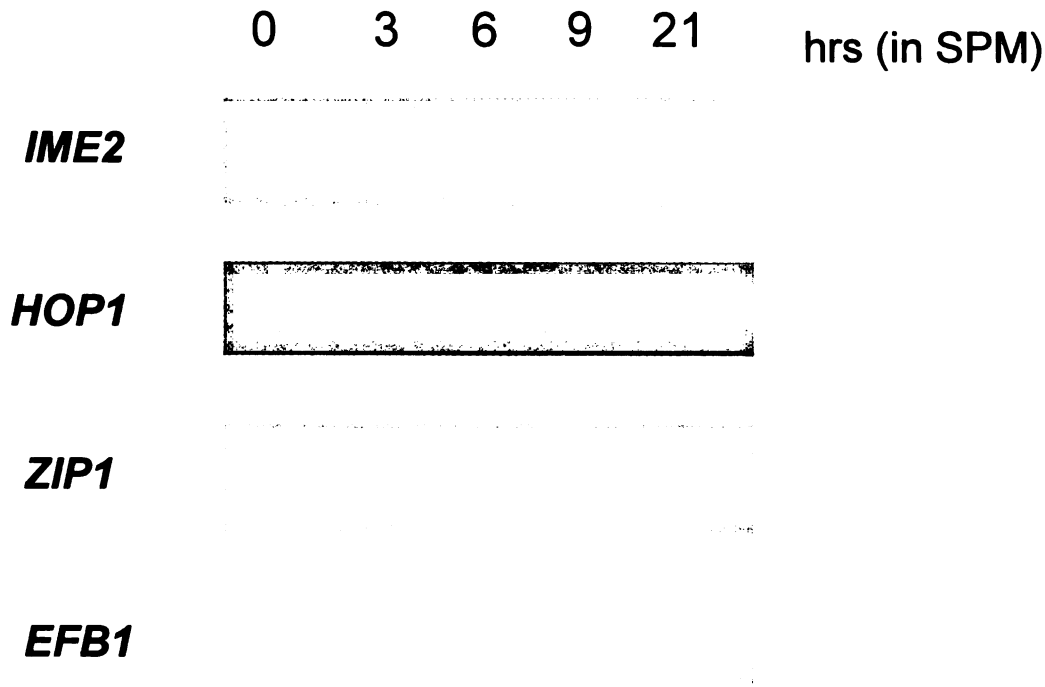


Figure 17. Expression of components of synaptonemal complexes in SPM.

RT-PCR was performed to detect the expression of putative *HOP1* and *ZIP1* homologues in SPM. RNA was harvested and samples were taken as described in the legend of Figure 14. No *HOP1* expression was detected; *ZIP1* was expressed.

S. cerevisiae ZIP1. These findings indicate that there is no SC formation in *C. albicans* under conditions which support sporulation in *S. cerevisiae*.

One of the key differences between mitosis and meiosis is the behavior of sister chromatids. *MEK1*, a meiosis specific kinase, is required for proper SC morphogenesis and sister-chromatid cohesion (49). *SPO70*, the activator for meiotic anaphase-promoting complex, is essential for meiosis and spore development (50). I could not detect either *MEK1* or *SPO70* expression in *C. albicans* cells grown either in SPM or vegetatively. This suggests that key regulators for meiotic chromosome behavior are not expressed in SPM (Figure 18).

Flocculation and expression of virulence genes

While *C. albicans* does not initiate meiosis in SPM, I did observe that yeast cell cultures begin to aggregate and form large flocculants when grown in SPM as compared with cells grown in YEPD as individual yeast forms (Figure 19). It has been suggested that flocculation is important for fungal pathogenesis due to the expression of adhesins which contributes to *C. albicans*'s interaction with its mammalian host cells (51). I tried to assess the possible relationship between flocculation and virulence by assaying expression of genes involved in virulence or cell-cell adhesion. These included *FLO1*, *SAP7* and *MUC1* (ORF6.2933). The best known flocculation gene in *S. cerevisiae* is *FLO1*. *FLO1*, a cell surface protein, is regulated by *SSK1* which in *C. albicans* is involved in the oxidative stress response (52). I detected the expression of *FLO1* in SPM, but I also found that *FLO1* is constitutively expressed during vegetative growth, indicating that expression of this gene is not associated with the flocculation phenotype I observed in SPM. Expression of *MUC1* (ORF6.2993), another member of the flocculin gene family, was barely detectable in either condition. *SAP7* is a unique member in the

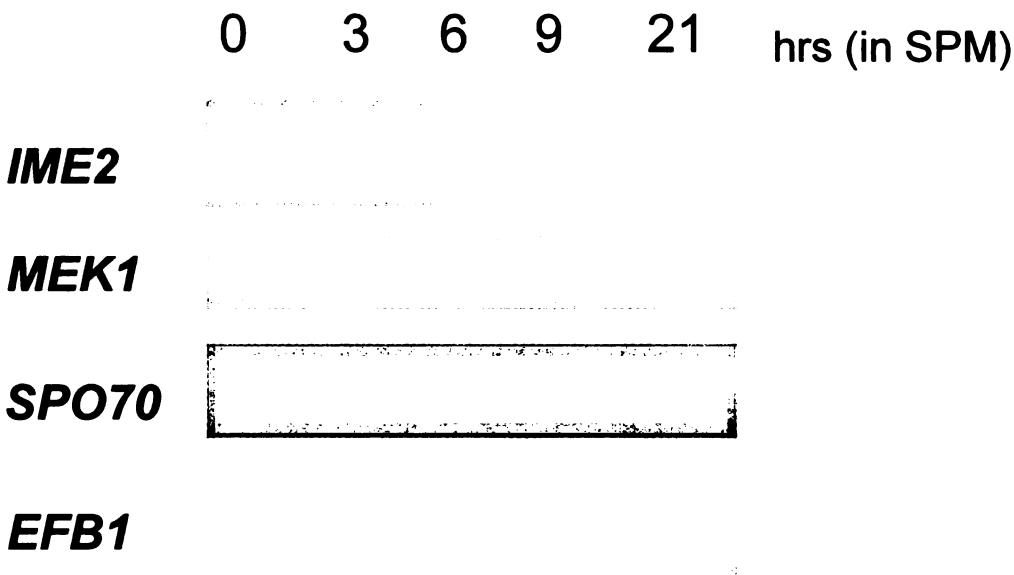


Figure 18. Expression of genes dictating meiotic chromosome behavior in SPM.

RT-PCR was used to assess expression of *C. albicans MEK1* and *SPO70* homologues in SPM. RNA was harvested from SC5314 cells grown to stationary phase in YEPD, washed and transferred to SPM. Samples were collected at the indicated times. Neither gene homologue was expressed in SPM.

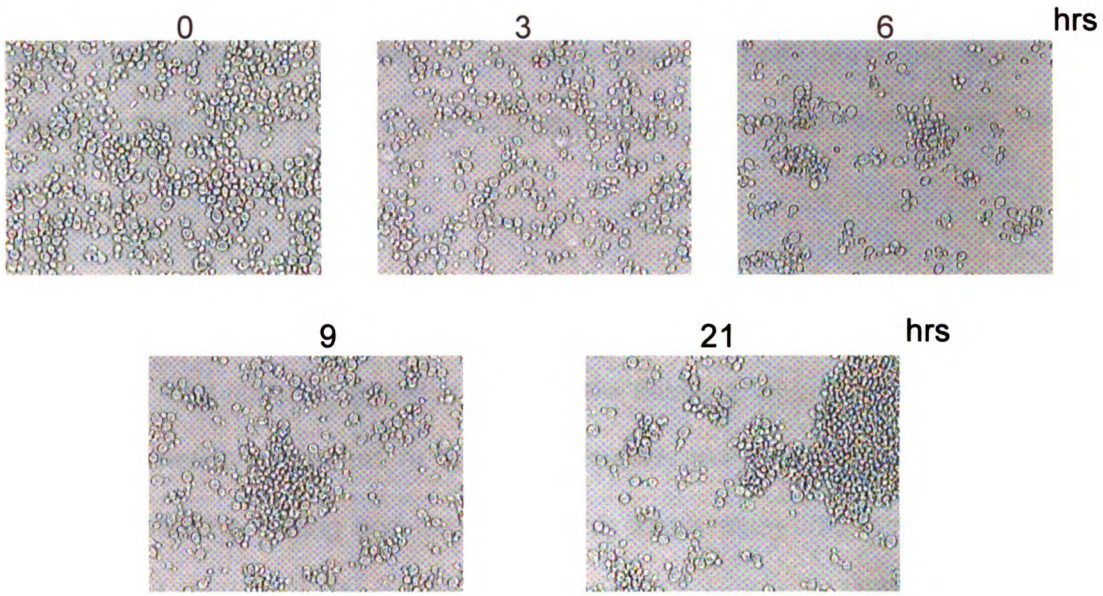


Figure 19. *C. albicans* shows flocculation in SPM.

When *C. albicans* SC5314 strain is transferred to SPM after being grown to stationary phase, it begins to aggregate and displays a flocculation phenotype. The majority of stationary phase cells remain primarily individual round shaped yeast form with prominent vacuoles.

secreted aspartyl proteinase (*SAP*) gene family, thought to be involved in mucosal infection based on its expression pattern (53). I did not observe the expression of *SAP7* during the flocculation process, although I did detect the expression of *SAP7* when the *C. albicans* strain was grown vegetatively (Figure 20). In summary, these results do not suggest that any of the genes assayed is involved in flocculation. There are, however, several factors affecting *C. albicans* self-aggregation, in both yeast and hyphal forms which also play pivotal roles in microbial adhesion, aggregation and biofilm formation (54).

Discussion:

It has been a long standing question whether the failure to observe a sexual cycle in *C. albicans* is genetically inherent, or due to our inability to define the correct conditions or strains for detecting mating and meiosis. In this thesis study I have used comparative genomics and molecular genetic approaches to explore the regulatory circuitry which might control initiation of meiosis in *C. albicans*. The results of this work provide clues regarding how developmental pathways evolve under different selective pressures. I have shown that *IME1*/*UME6*/URS1, the initiation complex as defined for meiosis in *S. cerevisiae*, does not exist in *C. albicans*. In *S. cerevisiae*, nonfermentatable carbon sources increase *IME1* and *IME2* transcript levels, leading to DNA replication and recombination (45). While I have demonstrated that acetate can activate *IME2* expression in *C. albicans*, growth in this carbon source nevertheless fails to result in the transcription of several gene homologues which in *S. cerevisiae* are key downstream targets of *IME2*. My results demonstrate that there are fundamental differences between the nutrient utilization and signaling networks that affect sexual reproduction in *C. albicans* and *S. cerevisiae*.

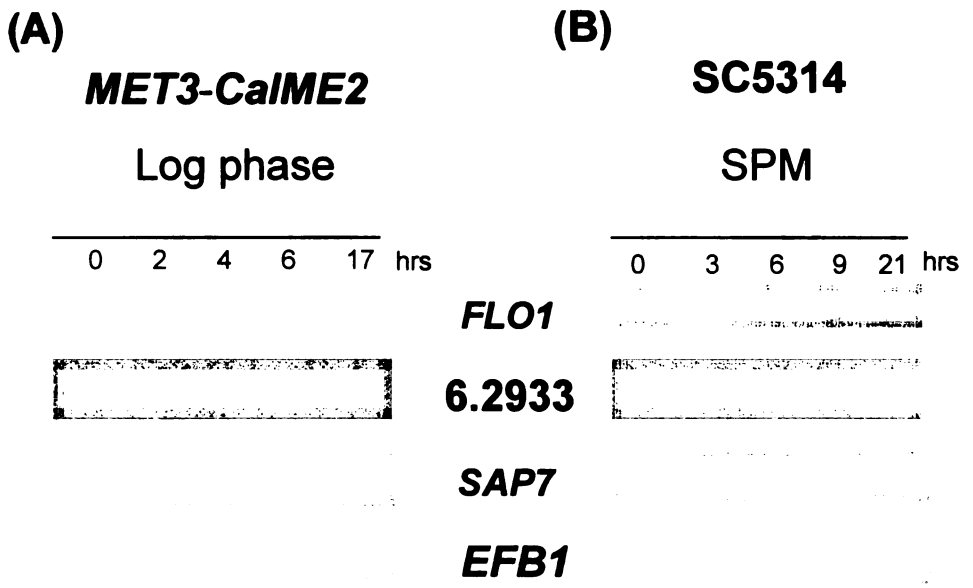


Figure 20. Expression of potential virulence-associated genes in *C. albicans*.

RT-PCR was used to assay expression of *FLO1*, *MUC1*(ORF6.2933) and *SAP7* under conditions (SPM) where *C. albicans* displays a flocculation phenotype. RNA was harvested from SC5314 cells grown in YEPD to stationary phase(time 0, B), washed, transferred to SPM, and samples were collected at the indicated times. Similarly, expression of these genes was assessed in RNA from strain KWY100 grown to OD_{600}0.8 (time 0, A) in SCM in the presence of 2.5 mM Met and 0.5 mM Cys, washed, transferred to Met- and Cys-free SCM; samples were collected at the indicated times. As shown, *FLO1* was constitutively expressed, and *MUC1* was not expressed in both instances. Expression of *SAP7* was detected during vegetative growth.

Defining an initiation complex for meiosis in *C. albicans*

I interpret the results of my studies to conclude that *C. albicans* lacks an *IME1* homologue or analogue. This conclusion is based on the following observations. First, BLAST analysis fails to identify any homologous sequences in the genome. Second, we performed saturated functional complementation assays in an *ime1* deficient mutant of *S. cerevisiae* and were unable to identify any complementing *C. albicans* genes. These results indicate that the master regulator as defined in *S. cerevisiae* for entering meiosis is absent in *C. albicans*. It further suggests an underlying regulatory difference for meiotic initiation in *C. albicans* vs. *S. cerevisiae*. Clearly, the absence of this gene explains why *C. albicans* does not undergo meiosis in response to nitrogen starvation and nonfermentable carbon sources which are required for meiotic initiation in *S. cerevisiae*. As an obligate parasite of mammals with no known free-living state, an appropriate cue for *C. albicans* to initiate meiosis is most likely to come from its host. It has been suggested that temperature differences at different anatomical sites may play a role in mating in *C. albicans*. It is thus possible that there has been no selective pressure in *C. albicans* to maintain an intact *IME1* gene, and an alternative gene has evolved in *C. albicans* which senses the environmental signals to regulate entry into meiosis. Consistent with this hypothesis, there is no *IME1* homologue in *S. pombe*, and it has been shown that either *mei2+* encoding an RNA-binding protein, or *ste11+* encoding a transcription factor, play an analogous role as *IME1* in this organism (55, 56).

For *S. cerevisiae*, Ume6 functions as a key component of the regulatory complex which activates or represses early meiotic genes transcription through its association with Sin3 or Ime1 respectively. Since the precondition for meiosis in *S. cerevisiae* differs markedly from that in *C. albicans* (if meiosis does occur), *IME1* may have been supplanted by a different functional analogue. Thus, one would expect that the sequence divergence in

UME6 between *S. cerevisiae* and *C. albicans*, as well as a possible change in the nature of the *UME6*/URS1 partnership might occur during evolution.

My studies provide several lines of evidence which indicate that the relationship between *UME6* and its DNA binding site URS1, as established in *S. cerevisiae*, is not conserved in *C. albicans*. First, *UME6* appears to have diverged significantly from its *S. cerevisiae* counterpart. Sequence comparisons between the potential *UME6* ORFs in *C. albicans* and *S. cerevisiae UME6* using each of the three domains of *S. cerevisiae UME6* including the Zn cluster domain, repression domain (Sin3-Rpd3 binding) and Ime1 binding domain have all failed to identify a clear *UME6* homologue. Second, genomic expression analysis of *C. albicans* is not consistent with the pattern of *UME6* expression as deduced from its role in *S. cerevisiae* and finally, functional complementation of *S. cerevisiae ume6* mutants with potential *UME6* ORFs from *C. albicans*, does not rescue the meiotic defect in *S. cerevisiae*.

My studies also support genome-wide changes in URS1 distribution and reassortment of this transcriptional regulatory element and its target genes. Genomic comparisons show that URS1 is not a conserved regulatory element common to the putative meiotic homologues in *C. albicans*. Interestingly motif searches (MEME) of the upstream sequences of putative meiotic genes in *C. albicans* also did not identify any conserved motifs, suggesting that no equivalent of a URS1 regulatory element functions to coordinate transcription of this gene subset. While in *S. cerevisiae* many early meiotic genes contain a URS1 site in their promoters, implying that they are coregulated by Ume6, Ume6 also recognizes different regulatory elements (57, 58). Moreover, the URS1 element is also found in the promoters of non-meiotic genes, and proteins other than Ume6 have been demonstrated to bind at the URS1 site (59, 60). To determine if

any pattern might emerge from understanding the distribution of URS1 binding sites in *C. albicans*, I catalogued the potential ORFs in *C. albicans* which possess upstream URS1 sequences (Table 3). The resulting gene set is quite different from that defined using the same parameters in *S. cerevisiae*. As mentioned above, I have shown that the URS1 site is not shared by the putative early meiotic homologues in *C. albicans*. Taken together, these results support the view that *C. albicans* does not possess a *UME6* homologue and further suggest that alternative transcriptional regulators will orchestrate expression of the putative meiotic homologues through different regulatory sites. This also indicates that a new regulatory mechanism may have emerged due to changing environmental cues for sexual development in *C. albicans*. Of note, it has been shown that *C. albicans SIN3* does not complement a *S. cerevisiae sin3* mutant (61). One caveat to these complementation studies however is that the difference in decoding CUG in these two yeasts may explain why potential *UME6* ORFs could not complement a *ume6* mutant in *S. cerevisiae*. CUG is an ambiguous codon in *C. albicans*, which can be decoded as leucine or serine (62).

Acetate activates *IME2* expression in *C. albicans*

Despite their overall sequence similarity, I have shown by both complementation and conditional expression analysis that *C. albicans IME2* is not a functional homologue of *S. cerevisiae IME2*. I further demonstrated that *C. albicans IME2* transcription is activated in the presence of acetate as a carbon source however, unlike the case in *S. cerevisiae*, this did not induce meiosis. *IME2* and *NDT80* are two key regulators of the transcriptional cascade of sporulation in *S. cerevisiae*. Transcriptional analysis in *C. albicans* reveals that the expression of either ORF of *NDT80* is not dependent on *IME2*, and *NDT80* cannot activate the expression of *SPS1* or *SPS4* which are Ndt80-dependent middle sporulation genes in *S. cerevisiae*, further supporting the notion of

underlying differences in signaling networks between *C. albicans* and *S. cerevisiae*. I have shown that the core meiotic recombination machinery is active when *C. albicans* grows in SPM as assessed by the transcription of the putative meiotic homologues including *SPO11*, *RAD50*, *MLH1*, *MRE11*, *RAD51* and *MSH4*. The genes dictating meiotic chromosome behavior are not expressed, and we predict that synaptonemal complexes (SC) fail to form (23), providing ample rationale for the absence of *C. albicans* meiosis in SPM.

Expression of *C. albicans* putative meiotic homologues in SPM

It has been suggested that meiosis has evolved from mitosis (18). A failure in the mitotic cell cycle or accidental nuclear fusion will result in polyploid nuclei. As one of the key functions in meiosis is ploidy reduction, meiosis might have evolved from repairing such mistakes (63). Ploidy reduction in meiosis is achieved by several concerted mechanisms, including (1) homologous chromosome paring; (2) a delay in sister centromere splitting until MII; (3) a lack of DNA replication between MI and MII; (4) the reorientation of sister centromeres to the same pole (2). Genes important for each of these aspects have been identified and characterized in *S. cerevisiae*. Notably, *SPO11* functions as the catalyst of double-stranded breaks, which in turn initiates meiotic recombination (64). *MSH4/MSH5* is important for reciprocal recombination, while meiosis specific *DMC1* is required for repairing of double-stranded breaks and chromosome pairing (65-67). The timing for cleavage of *REC8* (meiotic cohesion), and the activation of *AMA1* (the meiotic activator of APC), are essential for the reductional behavior of meiotic division and sister chromatid separation (50, 68, 69). *MAM1* (monopolin) is important for suppressing sister kinetochore bi-orientation, however this gene appears to be unique to *S. cerevisiae* (70).

I have shown that *C. albicans SPO11* is expressed in SPM and in the mitotic cell cycle. There is no detectable expression of *DLH1* (*DMC1* homologue), and *MSH5* is expressed in very low levels in SPM. I have noted the expression of *MSH4* in mitosis as well as in SPM. No expression of *MEK1* or *SPO70* was observed in SPM. These results provide information about two important aspects of cellular differentiation in *C. albicans*. First, there might be an unidentified environmental cue different from that for *S. cerevisiae*, which activates evolutionarily important genes for meiosis such as *REC8* and *DLH1* in *C. albicans*. Second, sporulation is a transcriptional cascade in *S. cerevisiae*, with one set of genes sequentially triggering the expression of another. Our previous study suggests that *C. albicans* possesses a potential sexual cycle similar to *S. cerevisiae*. However, transcriptional analysis in *C. albicans* reveals that the expression of similar sets of genes is not coordinated, indicating a rewiring of the regulatory circuitry in this pathogenic fungus.

Flocculation and fungal pathogenesis

When stationary phase *C. albicans* cells are transferred to SPM, cells start to aggregate and form large clumps, indicating that yeast cell-yeast cell flocculation has occurred. In brewer's yeast, flocculation has practical industrial applications (71). For fungal pathogens, expression of cell-surface proteins contributing to the flocculation phenotype might change microbial adherence to host cells, and facilitate the expression of virulent traits. It has been hypothesized that flocculation involves expression of lectin- like proteins (flocculins) on the surface of yeast cell walls, which selectively bind to cell-wall mannose of neighboring cells (72). The *FLO* gene family has been implicated in flocculation due to expression of these flocculin genes (73). For example, Flo1 is a cell wall protein in *S. cerevisiae* and plays a role in cell-cell interaction, indicating its structural role in flocculation (74). Many cell wall related proteins in *C. albicans* have

been assigned adhesin functions which may be important for interactions between *C. albicans* and host epithelial or endothelial cells. It has been shown that the expression of *C. albicans CAD1/AAF1* gene in *S. cerevisiae* induces flocculation and enhances adherence to endothelial cells, mimicking the Flo1 phenotype (75). Other *C. albicans* genes which might be involved in flocculation include *ALS1* (agglutinin-like sequence), *CHK1* (histidine kinase) and *SSK1* (two-component response regulator) (52, 76, 77). The RT-PCR analysis I performed does not support the hypothesis that any reported gene associated with flocculation contributes to the aggregation phenotype in our system. Since yeast flocculation is governed by multiple factors such as nutrient availability, temperature, pH, oxygen and genetic background of strains (72), our results suggest that multiple signaling pathways will be involved in flocculation.

Is there meiosis after mating in *C. albicans*?

While ample evidence indicates that *C. albicans* can be made to mate, a body of other evidence in particular population studies suggest that mating is infrequent. Clinical isolates of *C. albicans* are primarily heterozygous (**a**/α cells) resulting in a marked imbalance in the ratio of mating type strains found in natural circumstances. Moreover, mating is linked to phenotypic switching, the opaque cells, which only grow at 25-30^{0}C being the active mating type, an observation which is at odds with the lack of any recognized ex vivo reservoirs of *C. albicans*. Taken together, these observations suggest that mating is a rare event in the natural history of this pathogen. Several scenarios arise from these findings. It is possible that *C. albicans* is primarily asexual, and genes involved in mating have acquired additional functions during evolution. Alternatively *C. albicans* may utilize mating for genetic exchange despite the fact that its mating potential may be limited in nature; after mating, the organism would return to a diploid state through either meiosis or a parasexual cycle. It has been reported that most

commonly a single *C. albicans* strain is found in a given anatomical/ infection site in the host, and more than one *C. albcans* strain can be found in patients with mixed infections (78, 79). Hence, one intriguing possibility is that *C. albicans* can self-fertilize at anatomical sites where the identical or closely related strains are found (80, 81). To date, meiosis has not been observed in *C. albicans*; most clinical strains are diploid or polyploid. So, is *C. albicans* able to undergo meiosis? Our previous study suggests that *C. albicans* has a genetic repertoire sufficient to support a complete sexual cycle; however it might manifest very rarely, explaining why no haploid progeny have yet been found. Available data indicate that *C. albicans* is able to have a diploid-tetraploid life cycle (82); therefore it might be advantageous for *C. albicans* to be an obligate diploid organism, which can overcome recessive lethal mutations leading to haploid lethality (83-85). Perhaps *C. albicans* undergoes single-division meiosis, since *SPO13* is missing in the *C. albicans* genome. In *S. cerevisiae*, *spo13* mutants produce diploid spores due to the elimination of MI (86). Among the 21 *Candida* species, 13 have been reported to reproduce sexually (87). Several species related to *C. albicans* including *C. lusitaniae*, *C. guilliermondii* and *C. krusei* have a defined haploid-diploid sexual cycle (88). Comparative genomic analysis between *C. albicans* and these relatives in the core meiotic machinery is essential for understanding how *C. albicans* might complete its sexual cycle or explain how it lost the capacity to perform meiosis. In particular, population studies have revealed that there are five geographically distinct clades in *C. albicans* (78). It has been shown that flucytosine resistance of *C. albicans* is clade specific, suggesting the possibility of prevalence of phenotypic differences between *C. albicans* strains in various geographic locations (89). For another fungal pathogen, *Crytococcus neoformans*, evidence indicates that clinical isolates from different regions possess differing capacities for mating (90). Perhaps meiosis in *C. albicans* is restricted to specific geographic isolates.

Genetic conservation in meiosis among organisms?

In this study, I have demonstrated that a master regulatory complex (*UME6*/*IME1*/URS1) essential for initiation of meiosis in *S. cerevisiae* is absent in *C. albicans*. Evidence from comparing meiosis in different species reveals that the master regulator can change dramatically as new species evolve. As the master regulator integrates genetic and environmental signals to regulate the onset of meiosis, our results reinforce the notion that there is an intimate relationship between the master regulator for meiotic initiation and environmental niches of organisms. I further provide evidence for the loss of conservation of function for the key genes (*IME2*, *NDT80*) in *C. albicans* which are important for the progression of meiosis in *S. cerevisiae*. Moreover, there is no ensuing meiosis when *C. albicans* grows in conditions known to induce sporulation in *S. cerevisiae*. In particular, several gene homologues important for meiosis such as *DLH1* (*DMC1*) and *SPO70* are not expressed in SPM as defined for *S. cerevisae*. In any case, the possibility that *C. albicans* is able to undergo meiosis is still open. These results suggest that an environmental trigger for meiosis is likely to be different from that for *S. cerevisiae*, and a different signaling pathway will be executed if meiosis occurs in *C. albicans*. Furthermore our results also raise the question whether meiotic transcriptome is conserved among organisms.

It appears that the general features for germ cell development are well conserved among a wide variety of species. Following premeiotic DNA replication, chromosome pairing (recombination) and two successive rounds of meiotic divisions without intervening DNA replication highlight the unique characteristics of meiosis. In particular, MI behaves reductionally and MII resembles mitosis. For example, it has been shown that sporulation in yeasts and spermatogenesis in mammals are analogous differentiation processes where meiotic divisions are coupled with gamete

morphogenesis (21). However, it is not well understood whether the genes participating in the meiotic events are also conserved among eukaryotes.

Based on induced gene expression patterns in yeast sporulation and their functional relevance to mammal meiosis, Hwang et al. (91) have chosen 164 yeast genes, and searched for their counterparts in mouse and/or human by BLAST analysis. Over 70% (116/164) of these yeast genes have homologues in *C. elegans*, *Drosophila* and mammals, suggesting conserved genetic information in meiosis among organisms through evolution. Moreover there are 49 genes unique to *S. cerevisiae*, which indicates the fundamental difference in the meiotic machinery between unicellular fungi and multicelluar metazoans. Su et al. (92) have conducted a large scale of analysis between the human and mouse transcriptome. Nearly half of 799 putative orthologue pairs displayed similar expression patterns between human and mouse. Among them, 104 (24%) gene pairs are expressed in testis, which provides evidence of differential gene expression in specific tissue. Furthermore, genomic comparisons in mouse and human suggest that genes involved in reproduction show a high degree of variability (93), which supports the notion that reproductive traits such as placental structures and estrous cycles will shape the genetic content of the meiotic machinery. Comparing the meiotic transcriptome between *S. cerevisiae* and *S. pombe* reveals a small set of (~75) genes is shared by these two yeasts (94). These include cell cycle regulators (B-type cyclins, *plo1/CDC5, flp1/CDC14*), APC components, APC regulators (*CDC20*), recombination proteins (*rec7/REC114*, *dmc1/DMC1*, *meu13/HOP2*) and cohesin (*rec8/REC8*). As *S. pombe* diverged from *S. cerevisiae* nearly ~330-420 million years ago (95), the small shared transcriptome indicates that meiosis has evolved independently in these two yeast species. Moreover, transcriptional profiling comparisons among *S. cerevisiae*, *S. pombe* and *C. elegans* show that genes shared by these three organisms are highly

expressed in vegetative cells, while *S. pombe*-specific genes are notably expressed in meiotic prophase (96). The absence of SC formation for sexual development in *S. pombe* is consistent with the concept that organism-specific genes generally function in specialized cellular processes, or in fulfilling particular needs for cellular specifications.

In order to understand the genes that are coexpressed in meiosis across multiple species, I searched the core meiotic machinery as defined in our previous study for their counterparts in other organisms using a web-based platform (97). Over 62% (84/134) of *C. albicans* or *S. cerevisiae* genes appear to have homologues in fly, worm and human (Table 4). In agreement with genomic comparisons performed by other investigators (91), this implies that there is a core transcriptome of meiosis shared by eukaryotic organisms. It also suggests that the number of genes that constitute the genetic core for gametogenesis is in this range (~80) (98); these include genes important for general features of meiosis such as DNA replication, recombination and chromosome cohesion. Hence, in addition to the core genetic repertoire shared by all eukaryotes during meiosis, individual species will deploy different strategies such as recruiting new proteins, changing regulatory networks or specifying alternative signaling pathways to constitute an organism-specific transcriptome for meiosis. For example, Mos is important for maintaining the metaphase arrest of oocytes and preventing DNA replication between MI and MII in vertebrates (99). The *Drosophila* homologue of Mos possesses cytostatic activity and activates a MAPK pathway as its metazoan counterpart does, however it is not essential for meiosis (100). This suggests that possible redundant pathways operate in the fly for the completion of meiosis. My work does not rule out a meiotic cell cycle in *C. albicans*, however our results suggest that *C. albicans* will execute an alternative signaling pathway if meiosis occurs. For *C. albicans*, we hypothesize that the meiotic machinery should be composed of the shared genetic core and other *C. albicans*-

specific genes. In particular, there are nearly 22% of *C. albicans* ORFs in its genome without homologues in other organisms (101). We speculate that some of these *C. albicans*-specific genes will play some role in meioisis. It has been suggested that the enrichment of organism-specific genes in defined cellular processes may drive the separation of species. Therefore, an emerging paradigm from available analyses supports the view of divergence of regulatory mechanisms for meiosis in diverse organisms.

References

1. Roeder, G. S. (1997) *Genes Dev* 11, 2600-21.
2. Petronczki, M., Siomos, M. F. & Nasmyth, K. (2003) *Cell* 112, 423-40.
3. Eddy, E. M. (2002) *Recent Prog Horm Res* 57, 103-28.
4. Honigberg, S. M. & Purnapatre, K. (2003) *J Cell Sci* 116, 2137-47.
5. Kassir, Y., Adir, N., Boger-Nadjar, E., Raviv, N. G., Rubin-Bejerano, I., Sagee, S. & Shenhar, G. (2003) *Int Rev Cytol* 224, 111-71.
6. Yamamoto, M. (1996) *Trends Biochem Sci* 21, 18-22.
7. Francois, F., Noel, T., Pepin, R., Brulfert, A., Chastin, C., Favel, A. & Villard, J. (2001) *J Clin Microbiol* 39, 3906-14.
8. Banuett, F. (1992) *Trends Genet* 8, 174-80.
9. Chu, S., DeRisi, J., Eisen, M., Mulholland, J., Botstein, D., Brown, P. O. & Herskowitz, I. (1998) *Science* 282, 699-705.
10. Primig, M., Williams, R. M., Winzeler, E. A., Tevzadze, G. G., Conway, A. R., Hwang, S. Y., Davis, R. W. & Esposito, R. E. (2000) *Nat Genet* 26, 415-23.
11. Smith, H. E., Su, S. S., Neigeborn, L., Driscoll, S. E. & Mitchell, A. P. (1990) *Mol Cell Biol* 10, 6103-13.
12. Steber, C. M. & Esposito, R. E. (1995) *Proc Natl Acad Sci U S A* 92, 12490-4.
13. Vidan, S. & Mitchell, A. P. (1997) *Mol Cell Biol* 17, 2688-97.
14. Rubin-Bejerano, I., Mandel, S., Robzyk, K. & Kassir, Y. (1996) *Mol Cell Biol* 16, 2518-26.
15. Kadosh, D. & Struhl, K. (1997) *Cell* 89, 365-71.
16. Smith, H. E. & Mitchell, A. P. (1989) *Mol Cell Biol* 9, 2142-52.
17. Guttmann-Raviv, N., Martin, S. & Kassir, Y. (2002) *Mol Cell Biol* 22, 2047-56.
18. Dirick, L., Goetsch, L., Ammerer, G. & Byers, B. (1998) *Science* 281, 1854-7.

19. Lamoureux, J. S., Stuart, D., Tsang, R., Wu, C. & Glover, J. N. (2002) *Embo J* 21, 5721-32.

20. Sopko, R., Raithatha, S. & Stuart, D. (2002) *Mol Cell Biol* 22, 7024-40.

21. Chu, S. & Herskowitz, I. (1998) *Mol Cell* 1, 685-96.

22. Kwon-Chung, K. J. & Bennett, J. E. (1992) *Medical mycology* (Lea & Febiger, Philadelphia).

23. Tzung, K. W., Williams, R. M., Scherer, S., Federspiel, N., Jones, T., Hansen, N., Bivolarevic, V., Huizar, L., Komp, C., Surzycki, R., Tamse, R., Davis, R. W. & Agabian, N. (2001) *Proc Natl Acad Sci U S A* 98, 3249-53.

24. Guthrie, C. & Fink, G. R. (1991) *Guide to yeast genetics and molecular biology* (Academic Press, San Diego).

25. Ausubel, F. M. (1994) *Current protocols in molecular biology* (John Wiley & Sons, New York).

26. Fonzi, W. A. & Irwin, M. Y. (1993) *Genetics* 134, 717-28.

27. De Backer, M. D., Maes, D., Vandoninck, S., Logghe, M., Contreras, R. & Luyten, W. H. (1999) *Yeast* 15, 1609-18.

28. Thompson, J. R., Register, E., Curotto, J., Kurtz, M. & Kelly, R. (1998) *Yeast* 14, 565-71.

29. Sambrook, J. & Russell, D. W. (2001) *Molecular cloning : a laboratory manual* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.).

30. Hoffman, C. S. & Winston, F. (1987) *Gene* 57, 267-72.

31. Johnston, J. R. (1994) *Molecular genetics of yeast : a practical approach* (IRL Press at Oxford University Press, Oxford ; New York).

32. Church, G. M. & Gilbert, W. (1984) *Proc Natl Acad Sci U S A* 81, 1991-5.

33. Burke, D., Dawson, D., Stearns, T. & Cold Spring Harbor Laboratory. (2000) *Methods in yeast genetics : a Cold Spring Harbor Laboratory course manual* (Cold Spring Harbor Laboratory Press, Plainview, N.Y.).

34. Care, R. S., Trevethick, J., Binley, K. M. & Sudbery, P. E. (1999) *Mol Microbiol* 34, 792-8.

35. Todd, R. B. & Andrianopoulos, A. (1997) *Fungal Genet Biol* 21, 388-405.

36. Schjerling, P. & Holmberg, S. (1996) *Nucleic Acids Res* 24, 4599-607.

37. Bernstein, B. E., Tong, J. K. & Schreiber, S. L. (2000) *Proc Natl Acad Sci U S A* 97, 13708-13.

38. Bailey, T. L. & Elkan, C. (1995) *Proc Int Conf Intell Syst Mol Biol* 3, 21-9.

39. Abe, H. & Shimoda, C. (2000) *Genetics* 154, 1497-508.

40. Garrido, E. & Perez-Martin, J. (2003) *Mol Microbiol* 47, 729-43.

41. Shinkai, Y., Satoh, H., Takeda, N., Fukuda, M., Chiba, E., Kato, T., Kuramochi, T. & Araki, Y. (2002) *Mol Cell Biol* 22, 3276-80.

42. Guttmann-Raviv, N., Boger-Nadjar, E., Edri, I. & Kassir, Y. (2001) *Genetics* 159, 1547-58.

43. Bowdish, K. S. & Mitchell, A. P. (1993) *Mol Cell Biol* 13, 2172-81.

44. Pierce, M., Benjamin, K. R., Montano, S. P., Georgiadis, M. M., Winter, E. & Vershon, A. K. (2003) *Mol Cell Biol* 23, 4814-25.

45. Honigberg, S. M. & Lee, R. H. (1998) *Mol Cell Biol* 18, 4548-55.

46. Wilson, W. A. & Roach, P. J. (2002) *Cell* 111, 155-8.

47. Clancy, M. J., Shambaugh, M. E., Timpte, C. S. & Bokar, J. A. (2002) *Nucleic Acids Res* 30, 4509-18.

48. Villeneuve, A. M. & Hillers, K. J. (2001) *Cell* 106, 647-50.

49. Bailis, J. M. & Roeder, G. S. (1998) *Genes Dev* 12, 3551-63.

50. Cooper, K. F., Mallory, M. J., Egeland, D. B., Jarnik, M. & Strich, R. (2000) *Proc Natl Acad Sci U S A* 97, 14548-53.

51. Sundstrom, P. (2002) *Cell Microbiol* 4, 461-9.

52. Chauhan, N., Inglis, D., Roman, E., Pla, J., Li, D., Calera, J. A. & Calderone, R. (2003) *Eukaryot Cell* 2, 1018-24.

53. Naglik, J. R., Challacombe, S. J. & Hube, B. (2003) *Microbiol Mol Biol Rev* 67, 400-28, table of contents.

54. Douglas, L. J. (2003) *Trends Microbiol* 11, 30-6.

55. Sugiyama, A., Tanaka, K., Okazaki, K., Nojima, H. & Okayama, H. (1994) *Embo J* 13, 1881-7.

56. Watanabe, Y., Shinozaki-Yabana, S., Chikashige, Y., Hiraoka, Y. & Yamamoto, M. (1997) *Nature* 386, 187-90.

57. Sweet, D. H., Jang, Y. K. & Sancar, G. B. (1997) *Mol Cell Biol* 17, 6223-35.

58. Gailus-Durner, V., Chintamaneni, C., Wilson, R., Brill, S. J. & Vershon, A. K. (1997) *Mol Cell Biol* 17, 3536-46.

59. Kaadige, M. R. & Lopes, J. M. (2003) *Mol Microbiol* 48, 823-32.

60. Strich, R., Surosky, R. T., Steber, C., Dubois, E., Messenguy, F. & Esposito, R. E. (1994) *Genes Dev* 8, 796-810.

61. Heyken, W. T., Wagner, C., Wittmann, J., Albrecht, A. & Schuller, H. J. (2003) *Yeast* 20, 1177-88.

62. Santos, M. A., Moura, G., Massey, S. E. & Tuite, M. F. (2004) *Trends Genet* 20, 95-102.

63. Cavalier-Smith, T. (2002) *Heredity* 88, 125-41.

64. Celerin, M., Merino, S. T., Stone, J. E., Menzie, A. M. & Zolan, M. E. (2000) *Embo J* 19, 2739-50.

65. Hollingsworth, N. M., Ponte, L. & Halsey, C. (1995) *Genes Dev* 9, 1728-39.

66. Novak, J. E., Ross-Macdonald, P. B. & Roeder, G. S. (2001) *Genetics* 158, 1013-25.

67. Bishop, D. K., Park, D., Xu, L. & Kleckner, N. (1992) *Cell* 69, 439-56.

68. Buonomo, S. B., Clyne, R. K., Fuchs, J., Loidl, J., Uhlmann, F. & Nasmyth, K. (2000) *Cell* 103, 387-98.

69. Watanabe, Y. & Nurse, P. (1999) *Nature* 400, 461-4.

70. Toth, A., Rabitsch, K. P., Galova, M., Schleiffer, A., Buonomo, S. B. & Nasmyth, K. (2000) *Cell* 103, 1155-68.

71. Maraz, A. (2002) *Acta Microbiol Immunol Hung* 49, 483-91.

72. Verstrepen, K. J., Derdelinckx, G., Verachtert, H. & Delvaux, F. R. (2003) *Appl Microbiol Biotechnol* 61, 197-205.

73. Stratford, M. (1992) *Yeast* 8, 25-38.

74. Bony, M., Thines-Sempoux, D., Barre, P. & Blondin, B. (1997) *J Bacteriol* 179, 4929-36.

75. Fu, Y., Filler, S. G., Spellberg, B. J., Fonzi, W., Ibrahim, A. S., Kanbe, T., Ghannoum, M. A. & Edwards, J. E., Jr. (1998) *Infect Immun* 66, 2078-84.

76. Calera, J. A. & Calderone, R. (1999) *Microbiology* 145 (Pt 6), 1431-42.

77. Fu, Y., Ibrahim, A. S., Sheppard, D. C., Chen, Y. C., French, S. W., Cutler, J. E., Filler, S. G. & Edwards, J. E., Jr. (2002) *Mol Microbiol* 44, 61-72.

78. Soll, D. R. & Pujol, C. (2003) *FEMS Immunol Med Microbiol* 39, 1-7.

79. Soll, D. R. (2000) *Clin Microbiol Rev* 13, 332-70.

80. Gow, N. A. (2002) *Mol Cell* 10, 217-8.

81. Murtagh, G. J., Dyer, P. S. & Crittenden, P. D. (2000) *Nature* 404, 564.

82. Bennett, R. J. & Johnson, A. D. (2003) *Embo J* 22, 2505-15.

83. Whelan, W. L. & Soll, D. R. (1982) *Mol Gen Genet* 187, 477-85.

84. Whelan, W. L. & Magee, P. T. (1981) *J Bacteriol* 145, 896-903.

85. Whelan, W. L., Partridge, R. M. & Magee, P. T. (1980) *Mol Gen Genet* 180, 107-13.

86. Wang, H. T., Frackman, S., Kowalisyn, J., Esposito, R. E. & Elder, R. (1987) *Mol Cell Biol* 7, 1425-35.

87. Hazen, K. C. (1995) *Clin Microbiol Rev* 8, 462-78.

88. Calderone, R. A. (2002) *Candida and candidiasis* (ASM Press, Washington, D.C.).

89. Pujol, C., Pfaller, M. A. & Soll, D. R. (2004) *Antimicrob Agents Chemother* 48, 262-6.

90. Fraser, J. A., Subaran, R. L., Nichols, C. B. & Heitman, J. (2003) *Eukaryot Cell* 2, 1036-45.

91. Hwang, S. Y., Oh, B., Knowles, B. B., Solter, D. & Lee, J. S. (2001) *Mol Reprod Dev* 59, 144-58.

92. Su, A. I., Cooke, M. P., Ching, K. A., Hakak, Y., Walker, J. R., Wiltshire, T., Orth, A. P., Vega, R. G., Sapinoso, L. M., Moqrich, A., Patapoutian, A., Hampton, G. M., Schultz, P. G. & Hogenesch, J. B. (2002) *Proc Natl Acad Sci U S A* 99, 4465-70.

93. Waterston, R. H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J. F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., Antonarakis, S. E., Attwood, J., Baertsch, R., Bailey, J., Barlow, K., Beck, S., Berry, E., Birren, B., Bloom, T., Bork, P., Botcherby, M., Bray, N., Brent, M. R., Brown, D. G., Brown, S. D., Bult, C., Burton, J., Butler, J., Campbell, R. D., Carninci, P., Cawley, S., Chiaromonte, F., Chinwalla, A. T., Church, D. M., Clamp, M., Clee, C., Collins, F. S., Cook, L. L., Copley, R. R., Coulson, A., Couronne, O., Cuff, J., Curwen, V., Cutts, T., Daly, M., David, R., Davies, J., Delehaunty, K. D., Deri, J., Dermitzakis, E. T., Dewey, C., Dickens, N. J., Diekhans, M., Dodge, S., Dubchak, I., Dunn, D.

M., Eddy, S. R., Elnitski, L., Emes, R. D., Eswara, P., Eyras, E., Felsenfeld, A., Fewell, G. A., Flicek, P., Foley, K., Frankel, W. N., Fulton, L. A., Fulton, R. S., Furey, T. S., Gage, D., Gibbs, R. A., Glusman, G., Gnerre, S., Goldman, N., Goodstadt, L., Grafham, D., Graves, T. A., Green, E. D., Gregory, S., Guigo, R., Guyer, M., Hardison, R. C., Haussler, D., Hayashizaki, Y., Hillier, L. W., Hinrichs, A., Hlavina, W., Holzer, T., Hsu, F., Hua, A., Hubbard, T., Hunt, A., Jackson, I., Jaffe, D. B., Johnson, L. S., Jones, M., Jones, T. A., Joy, A., Kamal, M., Karlsson, E. K., et al. (2002) *Nature* 420, 520-62.

94. Mata, J., Lyne, R., Burns, G. & Bahler, J. (2002) *Nat Genet* 32, 143-7.

95. Sipiczki, M. (2000) *Genome Biol* 1, REVIEWS1011.

96. Mata, J. & Bahler, J. (2003) *Genome Res* 13, 2686-90.

97. Stuart, J. M., Segal, E., Koller, D. & Kim, S. K. (2003) *Science* 302, 249-55.

98. Schlecht, U. & Primig, M. (2003) *Reproduction* 125, 447-56.

99. Kishimoto, T. (2003) *Curr Opin Cell Biol* 15, 654-63.

100. Ivanovska, I., Lee, E., Kwan, K. M., Fenger, D. D. & Orr-Weaver, T. L. (2004) *Curr Biol* 14, 75-80.

101. Jones, T. F., Federspiel, N.A., Chibana, H., Dungan, J. Kalman, K. Magee, B.B. , Newport, G., Thorstenson, Y.R., Agabian, N., Magee, P.T., Davis, R.W. & Scherer, S. (2004) *Proc Natl Acad Sci U S A*, in press.

Table 1. Comparison in the core meiotic machinery between *C. albicans* and *S. cerevisiae*

Category	Gene # in *S. cerevisiae*	Gene # in *C. albicans*	Genes missing in *C. albicans*
Mating	39	36	*STE5,KAR1,DIG2*
Nutritional Control	18	18	
Cell Type Control	6	5	*MTAA2*
Initiation of Meiosis	15	12	*UME1,IDS2,IME1*
Checkpoint Control & Progression through Meiosis	52	48	*PDS1,CDC13, ISC10,SIC1*
Recombination & SC Formation	53	35	*RED1,SAE2,XRS2 ZIP2,SPO13,MER2 REC114,REC102 RAD55,REC104, HOP2MER1,SPO16, SPO12MEI4,MEI5, SAE3,NDJ1*
Spore Wall & Ascus formation	13	12	*SPS100*

Genomic comparison indicates that *C. albicans* appears to have an intact mating pathway as defined in *S. cerevisiae*. While *C. albicans* possesses most genes associated with the *S. cerevisiae* meiotic machinery, several crucial genes are missing. In particular, *IME1* is important for initiation of meiosis, and *SPO13* is essential for regulation of MI. Interestingly, the major set of genes absent form the *C. albicans* genome is related to chromosome recombination and synaptonemal complexes formation.

Table 2 *C. albicans* proteins containing the potential C_6Zn_2 domain

ORF name	Protein length (no. of amino acids)	Location of C_6Zn_2 domain within the protein	Potential C_6Zn_2 binuclear cluster domain $C(X)_2C(X)_6C(X)_{5-16}C(X)_2C(X)_{6-8}C$
orf6.236	124	38	CLACKKRKVKCDKGKPACGGCVRNGVGHLC
orf6.265	276	54	CKSCHSLKVKCTPSDPNNPSAPCVRCINANRIC
orf6.484	709	82	CVECRQQKSRCDAFEKQPDPCTRCAKKGLHC
orf6.536	400	68	CIRCRTRHIRCPGGDPCRKCQLAKSKC
orf6.692	838	10	CISCRKRKIKCNRIKPCDQCIKRHLPC
orf6.852	126	26	CDYCRKRKSKCNGEQPCSKCLDKNRNC
orf6.854	526	40	CDSCRRKKIKCNGSYPCGNCIQAKNTSNC
orf6.867	196	18	CLTCRKRKKKCDENKPKCNSCIHLNKEC
orf6.887	458	381	CWTCRIRHKACPEEKPSCSQCIRLQLDC
orf6.999	921	28	CDFCRSKKAKCDGKSPVCSNCFANKEEC
orf6.1023	1009	66	CEYCRKKKVKCNGCQPCLNCLQSNNGDC
orf6.1106	253	28	CKECKRRKIKCDEGKPSCWQCIRLRKDC
orf6.1277	388	29	CTRCRRHKTKCDAQDTNPYPCSHCFKRNLDC
orf6.1295	592	40	CRECKRRKIRCPEEKPYCSTCVRLGKQC
orf6.1323	119	23	CSNCDGKCPICDSFVKPTEQVRICQDCSQGHLSNKC
orf6.1326	840	765	CWICRIKHLKCDEVTPICGGCAKFGLQC
orf6.1405	992	95	CDTCRQKKVKCDGKQPCIHCTVYSYKC
orf6.1414	860	116	CANCKERRVKCLENLPSCTNCIKHRVKC
orf6.1797	101	7	CFLCCFFCFLCFLCLLGLLCLLCFLCFLCFLYFLCLLC
orf6.1907	689	13	CKECKRRKIRCPEDKPSCATCVRLGKVC
orf6.2248	340	18	CDFCRSKKAKCDGKSPVCSNCFANKEEC
orf6.2322	947	54	CKSCHSLKVKCTPSDPNNPSAPCVRCINANRIC
orf6.2403	830	15	CDYCKKRKFKCSGVSPCELCTKKGIQC
orf6.2536	974	12	CWTCRSRKVKCDLTKPQCNRCLKSNRIC
orf6.2537	1029	90	CDPCRKKKIKCNAEYSELEKKVTKICTSCAKNKEIC
orf6.2541	246	194	CNNCRLKKLKCDEQKPQCSRCSKKGYDC
orf6.2542	1100	48	CFTCRSRKIKCDLTKPQCEKCTRAGLIC
orf6.3100	805	17	CDYCKKRKYKCSGIAPCNLCSKKQIDC
orf6.3346	1196	239	CDTCRLKKIKCDNVRPRCGSCIKNGNLNC
orf6.3356	528	33	CFNCKRKKIKCNRQQPCLSCIKSNLESKC
orf6.3429	311	246	CWICRIKHLKCDEVTPICGGCAKFGLQC
orf6.3441	195	55	CLTCRKRKKKCDEDKVNGKCQACTRNFLDCC
orf6.3543	1144	84	CTVCRSRKVRCDAEIHIPCTNCITFGCDC
orf6.3545	871	31	CDLCKRHKRKCNGDNPCSYCQEKSLQC
orf6.3585	944	15	CDYCKKRKFKCSGVSPCELCTKKNIQC
orf6.3816	123	33	CFNCKRKKIKCNRQQPCLSCIKSNLESKC
orf6.3817	1242	33	CSVCRKRKSKCDRARPCGTCIKKSIVHLC
orf6.3987	405	28	CKECKRRKIKCDEGKPSCWQCIRLRKDC
orf6.4032	1056	54	CDRCRAKKTKCDGQNPCSTCQSVGLEC
orf6.4118	122	40	CRECKRRKIRCPEEKPYCSTCVRLGKQC
orf6.4142	861	44	CAECRRRKIKCDELKPYCHNCTRLNKLC
orf6.4154	449	115	CDTCAAKKVKCDANRPCGQCRTNNLEC
orf6.4213	917	90	CTFCRQHKIKCNASDNYPNPCERCKKMGLKC
orf6.4484	922	26	CDYCRKRKSKCNGEQPCSKCLDKNRNC
orf6.4487	1010	66	CEYCRKKKVKCNGCQPCLNCLQSNNGNC
orf6.4585	1078	32	CDRCRIKKIKCDGQSPCRNCQKVSVEC
orf6.4701	282	205	CWTCRIRHKACPEEKPSCSQCIRLQLDC
orf6.4842	915	69	CGPCKLLKIKCDLQIPCTACKKFNRVNRC
orf6.4908	654	13	CDNCRKRRRKCNGELPCHYCSGKNKPC
orf6.4951	578	51	CVYCRRSHMICDESRPCQRCIKRGIAHLC
orf6.5198	890	175	CDRCRSHKIKCSGECPCATCLKQAKEC
orf6.5234	1085	38	CLACKKRKVKCDKGKPACGGCVRNGVGHLC
orf6.5403	710	11	CDACRSRKIKCNRQTPCASCHKSKRDC
orf6.5421	978	68	CIRCRTRHIRCPGGDPCRKCQLAKSKC
orf6.5722	796	48	CVYCHSKKIKCDRQKPCSQCTKLGMEC
orf6.5847	619	44	CELCRKQKTRCFRSPENPNSCLRCRFLSKTC
orf6.5864	385	315	CLTCRQRKKRCCETRPRCTECTRLRLNC
orf6.5965	1105	12	CDFCKRRKVKCDLGNPCSMCVKYKKSPC
orf6.6060	629	128	CGTCKKRRIKCDETLPACLNCLKGKLHC
orf6.6273	712	54	CSTCKKRRVKCDEQRPVCGNCTKLKLDC
orf6.6500	775	21	CLECRRRHFKCDGKQPVCDRCQKSNKPC

Table 2. Continued

ORF name	Protein length (no. of amino acids)	Location of C_6Zn_2 domain within the protein	Potential C_6Zn_2 binuclear cluster domain $C(X)_2C(X)_6C(X)_{5-16}C(X)_2C(X)_{6-8}C$
orf6.6594	699	42	CFTCRKRKKKCDEHQPNCENCIRNKLKC
orf6.6620	984	82	CVECRQQKSRCDAFEKRPDPCTRCAKKGLHC
orf6.6789	787	31	CDNCKKRKFKCSGEKPCFECSKKGHDC
orf6.7200	738	18	CLNCKRKKVRCDESLPECKNCVKGKKETC
orf6.7203	640	39	CLRCRKLKKKCDKSTPHCLNCENANEEC
orf6.7263	566	20	CVFCHQKHLQCSNERPCKNCVKRNIAHGC
orf6.7399	860	35	CKRCRTKKIKCDQKFPQCGKCEVNGVEC
orf6.7478	834	38	CLTCKKKRLKCDETKPNCLNCTKKNIEC
orf6.7582	583	51	CLQCRARKKKCNEEHPVCGSCKRRKVNC
orf6.7691	799	15	CCNCKRSKVKCVYTSSLPCERCVKTGQAATC
orf6.7731	900	42	CQRCRTRKIKCNYELPCFNCKRDGSQC
orf6.7749	1065	45	CLNCRLRKVKCDLGPVDNPHDGKCARCLRERKDC
orf6.8252	878	120	CDQCRKRKIKCVLVPNTNNCVQCEAKQVTC
orf6.8327	261	24	CDSCRKRKLKCSKEFPRCSKCIQHEWCC
orf6.8399	214	17	CKNCRLRKRKCDELHPTCTFCHTRDLIC
orf6.8400	501	13	CDSCSFRKVKCDMKTPCSRCVLNNLKC
orf6.8531	1102	23	CLSCRKKKVKCDKGKPCSNCNKSQPDKC
orf6.8532	1108	31	CTNCRKRKIRCDRQHPCNNCIKSKKHNAC
orf6.8541	624	19	CVTCRDRHIKCDEQQPVCKNCQKSNRKC
orf6.8770	582	14	CNLCRSSKIKCINNSDSTRCHRCASLDLEC
orf6.8814	517	26	CDSCRIKKTKCDGKKPCNRCTLDNKIC
orf6.8918	803	74	CLTCKVRKKKCSEEKPVCNDCSRFGKNC
orf6.8931	681	25	CTRCRRFKKKCSFENPSFKSCARCFKNGYEC
orf6.9045	446	305	CLTCRKRRIKCDERKPTCFNCERSKKSC
orf6.9097	716	14	CHTCKRSKIKCDENKPTCSYCSKTKAIC

Table 3. *C. albicans* genes containing the URS1 site in the regulatory regions

ORF	Gene annotation
6.1569	no good homologies
6.2385	YBL051, weak similarity to RNA-binding proteins
6.3	no good homologies
6.3052	no good homologies
6.3068	AGP3, amino acid permease
6.3114	GEA2, ARF GTP/GDP exchange factor
6.3115	no good homologies
6.3305	COX15, cytochrome oxidase assembly
6.3667	ERG3, C5, 6 desaturase
6.3743	VID21
6.3821	sim to flagellar calmodulin
6.3899	sim to Hypothetical protein M151.4
6.4091	no good homologies
6.4415	no good homologies
6.4537	PMT99, Similar to SPAC3A12.06c membrane transport protein
6.4908	YNA2, sim to Pichia Yeast Nitrate Assimilation factor; potential fungal Zn(2)-Cys(6) binuclear cluster domain
6.5119	CDC42, RHO small monomeric GTPase; signal transduction
6.5260	no good homologies
6.5986	YOR289, homolog of YOR289w
6.6044	no good homologies
6.6072	no good homologies
6.6073	no good homologies
6.6117	no good homologies
6.6217	YOR292, similar to to human glomerulosclerosis protein
6.6540	MCT99, malonyl CoA:acyl carrier transacylase
6.6570	CHO1, phosphotidylcholine synthase
6.6585	YLR110, weakly similar to YLR110c
6.6725	sim to Ndt80p
6.6785	DPH5, diphthamide synthesis
6.6900	PEX10, similar to Pichia PEX10 Zn-binding protein
6.6908	no good homologies
6.7281	CHK1, histidine kinase 1
6.7537	no good homologies
6.7608	YFR021, homolog of YFR021W
6.7702	no good homologies
6.7858	YML059, homolog of YML059C
6.7902	ARL1, signal transduction
6.799	YJL171, sim to YJL171
6.8046	no good homologies
6.8200	no good homologies

Table 3. *C. albicans* genes containing the URS1 site in the regulatory regions

ORF	Gene annotation
6.8333	no good homologies
6.8373	ARP2, endocytosis and membrane growth and polarity
6.8496	YGL245, simillar to YGL245W
6.8525	FYV10, weak sim to FYV10
6.8532	HAP1, sim to Heme Activated Protein; Co-ordinate control of synthesis of mitochondrial and non-mitochondrial hemoproteins; potential fungal Zn(2)-Cys(6) binuclear cluster domain
6.8574	FLO2, agglutinin-like protein 6
6.8689	no good homologies
6.8699	SEC6, fusion of post-Golgi vesicles with plasma membrane
6.8719	no good homologies
6.8732	no good homologies
6.8789	BNI1, cytoskeletal regulatory protein binding
6.8893	VAN1, Vanadate resistance protein
6.8959	TRX1, Thioredoxin
6.8967	YHR194, homolog of YHR194W
6.8986	PRO1, gamma-glutamyl kinase
6.926	no good homologies
6.932	TOM37, translocase of outer mitochondrial membrane

Table 4. The core meiotic machinery[a]

Gene Name	Meta genes[b]
YAL009W	No meta-gene
YAL038W	MEG1823
YAR007C	MEG171
YBL058W	MEG1132
YBL075C	No meta-gene
YBL084C	MEG1395
YBR088C	MEG2077
YBR112C	MEG194
YBR136W	MEG804
YBR140C	MEG56
YBR160W	MEG806
YCR039C	No meta-gene
YCR040W	No meta-gene
YCR084C	No meta-gene
YCR086W	No meta-gene
YCR092C	MEG1663
YCR096C	No meta-gene
YCR097W-A	No meta-gene
YCR097W	No meta-gene
YDL008W	MEG1599
YDL017W	MEG2189
YDL028C	MEG592
YDL040C	MEG738
YDL042C	No meta-gene
YDL102W	MEG474
YDL154W	MEG63
YDL164C	MEG348
YDR004W	No meta-gene
YDR097C	MEG1566
YDR104C	No meta-gene
YDR108W	No meta-gene
YDR180W	MEG2003
YDR207C	No meta-gene
YDR217C	No meta-gene
YDR285W	No meta-gene
YDR364C	MEG865
YDR402C	MEG2
YDR403W	No meta-gene
YDR440W	MEG1872
YDR477W	MEG2036
YDR522C	No meta-gene
YDR523C	MEG1523
YEL060C	MEG54
YER012W	MEG2013
YER095W	MEG756
YER133W	MEG2319
YER173W	MEG2032
YER179W	MEG611
YFL003C	MEG1634
YFL008W	MEG536
YFL009W	MEG2346
YFL021W	MEG1221

Table 4. Continued

Gene Name	Meta genes[b]
YFL029C	No meta-gene
YFL033C	MEG238
YFL038C	MEG670
YFR028C	MEG336
YFR031C-A	No meta-gene
YFR031C	MEG2321
YGL035C	No meta-gene
YGL058W	MEG862
YGL115W	MEG1708
YGL116W	No meta-gene
YGL163C	MEG564
YGL173C	MEG1705
YGL192W	MEG1874
YGL195W	MEG1143
YGL197W	No meta-gene
YGL213C	No meta-gene
YGR044C	No meta-gene
YGR059W	No meta-gene
YGR184C	MEG1813
YGR225W	No meta-gene
YGR258C	MEG533
YHL022C	MEG2283
YHL027W	No meta-gene
YHR008C	MEG1808
YHR124W	No meta-gene
YHR139C-A	No meta-gene
YHR166C	MEG130
YIL072W	MEG1032
YIL099W	No meta-gene
YIR026C	MEG2433
YJL074C	MEG798
YJL106W	MEG1940
YJR057W	MEG1898
YJR104C	MEG723
YKL022C	MEG659
YKL166C	No meta-gene
YKR031C	MEG2022
YLR071C	No meta-gene
YLR127C	MEG1208
YLR234W	MEG1365
YLR310C	MEG5090
YLR399C	MEG426
YML031W	No meta-gene
YML032C	MEG721
YML115C	No meta-gene
YMR001C	MEG2373
YMR017W	No meta-gene
YMR043W	MEG1692
YMR063W	No meta-gene
YMR076C	MEG1106
YMR139W	MEG1409
YMR154C	MEG2403

Table 4. Continued

Gene Name	Meta genes[b]
YMR167W	MEG1269
YMR224C	MEG1235
YNL012W	No meta-gene
YNL082W	MEG1928
YNL088W	MEG518
YNL098C	No meta-gene
YNL172W	MEG647
YNL202W	MEG2204
YNL204C	No meta-gene
YNL216W	No meta-gene
YNL236W	No meta-gene
YNL250W	MEG1926
YNL307C	No meta-gene
YNL330C	MEG1026
YOL004W	MEG640
YOL081W	MEG56
YOL090W	MEG1366
YOR101W	MEG2018
YOR190W	No meta-gene
YOR242C	No meta-gene
YOR257W	MEG1333
YOR351C	No meta-gene
YOR368W	No meta-gene
YPL008W	MEG2416
YPL084W	MEG1994
YPL153C	No meta-gene
YPL154C	MEG2417
YPL209C	MEG1828
YPR007C	No meta-gene
YPR054W	No meta-gene

[a]The core meiotic machinery is defined as gene homologues important for meiosis in *S. cerevisiae*, which have counterparts in *C. albicans*, human, fly and worm.
[b]A metagene is defined as a set of genes across multiple organisms whose protein sequences are one another's best reciprocal BLAST hits.

Table 5. Primers used in reverse transcription PCR

Genes	Sequence (5' to 3')
DLH1-a-L	CAATTAGGTAATGAATTAGCAGAAGG
DLH1-a-R	CAAACAATGCACTGGCTCC
CaMSH5-a-L	ATACCACAGCTAGGGTTCTTGG
CaMSH5-a-R	AACTCTTTGTTGCAAGTCCTGG
CaMRE11-a-L	ATTAGCTGAAGGTGAAGTTGCC
CaMRE11-a-R	GATGGTGGTATTTTTCTTAGGGG
CaMLH1-a-L	CATCAAACACCAATGTTGAACC
CaMLH1-a-R	TCTTCTTTTTCGTTTTCTTCGG
CaRAD50-a-L	ACAACACCATGAGTATGCAACC
CaRAD50-a-R	CAACTCTTTCAATTTGCATTCG
CaRAD51-a-L	TAGTTCCCTTGGGTTTCACC
CaRAD51-a-R	AGATTCTGCCATCATTTCAGC
FLO1-a-L	AGTGCATTATTTGCTGCTACC
FLO1-a-R	AGCATGACAAGCATCATTATGA
ALS1-a-L	CAGGATACCCAACTTGGAAT
ALS1-a-R	CCAGTATTCAGTAGTAGTGA
ORF6.2933-a-L	TGAATCAAGTCATGCTACCACC
ORF6.2933-a-R	TCCATTGTAGCACTTGGTTTTG
CaIME4-a-L	GTTGAAAACACCAGAAAAACCC
CaIME4-a-R	GATTTCCAATTGTAAGCCAACC
CaMSH4-a-L	ATGCATCGACTATTGTGTTTCG
CaMSH4-a-R	ATTGGTGTGTAGCTGTTGATGC
CaMEK1-a-L	AACTAAAAATTGTCCACCGGG
CaMEK1-a-R	GAGGCTTGATTGTCTGTCTCC
CaSPO70-a-L	AAGTAACCGTTTCATTCCTTCG
CaSPO70-a-R	CTTCTCCAAAGTCATCACCAGC
SAP7-F	GAAATGCAAAGAGTATTAGAGTTATTAC
SAP7-R	GAATGATTTGGTTTACATCATCTTCAACTG
ACT1-F	GGCTGGTAGAGACTTGACCAACCATTTG
ACT1-R	GGAGTTGAAAGTGGTTTGGTCAATAC
ORF6.4742-F	CAACAAGATCGAAACTATGAATAAAATC
ORF6.4742-R	TTACTGTGGAGGAGTAGGGGTTG
ORF6.6725-F	GCTATGATAGTATGGATTCTTATACTATAAACAC
ORF6.6725-R	CTAAGCTGTCATTACATCAGAAGGTGTCAC
SAP7-a-L	TCTTCTTCACTGGAAGCTGC
SAP7-a-R	AGGAACAACGGCATGGTTATC
CaSPS1-a-L	CTCATCATCTTCACCACTGAGC
CaSPS1-a-R	CCTTGCCATTGTCTAATTTTCC
CaSPS4-a-L	CAATTCCTGAACTTCTTTCTTTCC
CaSPS4-a-R	TCCAGAGACTCTTTCTCTTTTTCC
EFB1-a-L	AGTCATTGAACGAATTCTTGGCTG
EFB1-a-R	TTCTTCAACAGCAGCTTGTAAGTC
CaSPO11-a-L	ATTAAGGAAATTAGGAAGTTATTGCG
CaSPO11-a-R	TTATTTCGTATAGCTAGGCGTTCC
CaHOP1-a-L	TTCTTATCAAAGGAACATCCGC
CaHOP1-a-R	TAGTTGTTTACCATCATTGGCG

Table 6. Oligonucleotides used in gene complementation

Oligonucleotides for *UME6*	Sequence
ScUME6f-1	attcccgggatgctagacaaggcgcgctctc
ScUME6r-1	cccgcggccgcttatttttttttcattgctcttc
orf6.887f	attcccgggatggtattttcgatatacgggtctagtg
orf6.887r	taagcggccgcttattgagtattactagttctcaatcctgc
orf6.4701f	attcccgggatgacaggaatatcactttgttcccgttc
orf6.4701r	taagcggccgcttattgagtattactagttctaaatcctgt
orf6.1326f	ttcccgggatgattacccatatggttacac
orf6.1326r	aagcggccgctcaatcattggttatatc
orf6.867f	ttcccgggatgaaactatctactgttggg
orf6.867r	aagcggccgctcatggtgaattattcagttg
orf6.5864f	ttcccgggatgagttcaatacccaatatc
orf6.5864r	aagcggccgcttatttacttctgtattcaac
orf6.6594f	aacccgggatgaattcagaaatatcaacg
orf6.6594r	ttgcggccgctcaaaaggcgacactttcg
oligonucleotides for *IME2*	
orf6.3678f	ttcccgggatgtcacgacaccattctttc
orf6.3678r	aagcggccgctcataaattcacaacttcctc

Chapter 4

Conclusion

One of the puzzles in *Candida* biology is why *C. albicans* appears to be an obligate, diploid organism. Traditionally *C. albicans* is classified as an imperfect fungus due to a failure to demonstrate or identify a sexual phase in this organism. Supporting this contention, population studies suggest that reproduction in *C. albicans* is primarily clonal, but do not rule out the possibility for low levels of recombination. Recently, it has been shown that engineered *C. albicans* strains are capable of mating either in a mouse host or under laboratory conditions. Most *C. albicans* strains found in nature however are heterozygous at their *MTL* locus (i.e. **a**/α cells), and mating is linked to a complex phenomenon of phenotypic switching. Hence, all evidence suggests that mating is a rare event in the wild for *C. albicans*.

It has been demonstrated that tetraploid mating products of *C. albicans* can return to a diploid state through a parasexaul cycle; no evidence of reductional division has ever been observed. Comparative analyses of the *C. albicans* genome have revealed a hidden commitment to sexual reproduction. In agreement with mating demonstrated in the laboratory, genomic analysis indicates that the mating pathway, the gateway toward a sexual cycle, is intact in *C. albicans*. Moreover, *C. albicans* possesses homologues of most of the genes associated with the meiotic machinery in *S. cerevisiae*, however several crucial genes are missing. In particular, *IME1*, the master regulator for triggering meiosis in *S. cerevisiae*, is absent in *C. albicans*. Ime1 coordinates genetic and nutritional signals which culminate in the initiation of the sporulation cascade in *S. cerevisiae*. The apparent lack of this homologue in *C. albicans* suggests a fundamental difference in the regulation of meiotic initiation between these two yeasts. Another missing gene is *SPO13*; *spo13* mutants of *S. cerevisiae* undergo single-division meiosis to produce diploid spores. This suggests that possibly *C. albicans* has an alternative reductional mechanism to complete meiosis, since most *C. albicans* strains are diploid.

Conspicuously, the major subset of genes which are absent in *C. albicans* is related to chromosome recombination and synaptonemal complexes formation. Furthermore, homologues of genes found in a variety of other organisms, which are associated with different aspects of meiosis are present in the *C. albicans* genome. In aggregate, comparative genomics indicates that *C. albicans* is equipped with a genetic repertoire to support a complete sexual cycle.

So, when and where, does *C. albicans* undergo meiosis? As an obligate vertebrate pathogen with no known *ex vivo* reservoir, one can assume that the environmental signals to initiate meiosis in *C. albicans* must come from its host. This is clearly different from *S. cerevisiae* where entry into meiosis is dependent on nitrogen starvation and the presence of nonfermentable carbon sources. It has been shown that temperature and skin are important factors for mating in *C. albicans* (1), suggesting that a requirement for special environmental conditions for initiating meiosis is reflected genetically in this organism. Thus it seems likely that *C. albicans* has lost *IME1* due to different selective pressures, indicating that the environmental cues which activate *IME1* in *S. cerevisiae* are not relevant to meiotic initiation in *C. albicans*.

Similarly, I show here that there is significant sequence divergence between all of the *C. albicans UME6* candidates and *S. cerevisiae UME6*, and further that there is a change in the relationship between *UME6* and its target URS1 binding site in *C. albicans*. Taken together, these results strongly suggest that *IME1*/*UME6*/URS1, the initiation complex for meiosis as defined in *S. cerevisiae* does not exist in *C. albicans*. These data also indicate that alternative transcriptional regulators must be identified to fulfill an analogous role as *IME1*/*UME6* in the controlling switch from mitosis to meiosis in *C. albicans*.

We have shown that while acetate activates *IME2* expression in *C. albicans*, no meiosis ensues. Furthermore, *IME2* does not activate transcription of either of the candidate *NDT80* genes, and there is no dependency upon the expression of *SPS1* and *SPS4* with *NDT80* as there is in *S. cerevisiae*. In particular, the temporal expression of different sets of genes in a putative meiotic pathway is not coordinated in *C. albicans* as observed in *S. cerevisiae*. These results imply that the same nutrient elicits a fundamentally different effect in putatively the same regulatory and signaling pathways between *C. albicans* and *S. cerevisiae*, indicating a rewiring of transcriptional circuitry in *C. albicans*.

The generation of new transcriptional patterns from existing gene products to create evolutionary novelty is well documented (2, 3). This rewiring of regulatory circuitry can be achieved by several mechanisms such as mutations in promoter sequences, recombination between regulatory regions, or divergence in coding sequences. Accordingly, changes in transcriptional regulators will change their activity, binding of regulatory elements, or change the nature of associated regulons. For example, the sex determining pathway is generally less conserved across phyla in metazoans. *Sex-lethal* (*Sxl*), the master regulator for sex determination in *Drosophila melanogaster*, controls the alternative splicing of a downstream gene *transformer* (*tra*), that acts with tra2 to control the alternative splicing of *doublesex* (*dsx*) which encodes both male and female-specific transcription factors. Moreover, *Drosophila Sxl* controls its own splicing by an autoregulatory feedback loop to ensure *Sxl* expression in females, but not in males (4). In *Ceratitis capitata*, *tra* is regulated by alternative splicing and controls the splicing of *dsx*. However, *tra* is not regulated by *Sxl* in *Ceratitis capitata*; it appears to be autoregulated as Sxl in *Drosophila* (5). Hence, *Ceratitis* tra is taking over the master regulator role held by *Drosophila* Sxl, and the regulatory pathway in sex determination seems to be conserved downstream in *Ceratitis capitata* (6).

For single-cell fungi, it has been shown that the evolution of the mating-type locus (i.e. sex determining gene) has been driven by several landmark events including gain of silent cassettes, gain of the *HO* gene and Ho cleavage site in *MAT*α*1*, and loss of MAT**a**2. Comparing the transcriptional hierarchy governed by the mating-type locus between *C. albicans* and *S. cerevisiae*, highlights the association between *MTL* and phenotypic switching in *C. albicans*, and reveals a different regulatory role of the mating-type locus in each respective organism. Taken together, these observations, as discussed in Tsong et al., suggest that regulatory circuits can change over evolutionary timescales (7).

Is meiosis-specific transcription conserved in all eukaryotes? If the primary signals triggering meiosis are varied among organisms, then what about regulatory genes responding to these signals? Our data can be extrapolated to indicate that each organism has a unique master regulator to control the initiation of meiosis, suggesting that the analogue of the *IME1*/*UME6* complex awaits to be discovered in *C. albicans*. Further, genomic analysis suggests that *C. albicans* possesses a *S. cerevisiae*-like meiotic pathway; however functional study indicates that there are dissimilarities in gene function and expression between these two yeasts. My data implies that the meiotic pathway is not well conserved in different species. It has been shown that gene expression and transcription regulation correlate well with gene conservation. While conserved genes are expressed at high levels, poorly expressed genes, generally organism-specific, are expressed in highly defined cellular processes. Our data support the idea that there is a minimal number of gene sets in the meiotic transcriptome which is shared between organisms; these include genes important for DNA replication, chromosome recombination and cohesion. Each species will develop its own meiotic transcriptome tailored to its own evolutionary advantage. Compared with *S. cerevisiae* or

S. pombe, *C. albicans* possesses a higher percentage (~22%) of ORFs with no known homologues in other organisms, suggesting that there is a genetic repertoire which might be associated with its adaptation to biological habitats and opportunistic nature as a pathogen (8, 9). I hypothesize that some of these *C. albicans*-specific genes will play an important role in meiosis; thereby the *C. albicans*-specific meiotic machinery will include the shared genetic core among eukaryotes and *Candida*-specific gene sets.

In this work, my goal was to understand why *C. albicans* has historically been classified as an imperfect fungus, i.e. a fungus without a sexual cycle. From genomic comparisons, my work suggests that *C. albicans* may have a complete sexual cycle in nature. This conclusion is based on identifying a number of genes homologous to those in the meiotic pathway of *S. cerevisiae* and other sexually reproducing organisms. To test the possibility that these gene homologues support meiosis in *C. albicans*, I used a combined approach of sequence analysis, functional complementation and gene expression to study the key gene homologues *IME2*, *NDT80* and their potential transcriptional targets in *C. albicans*. My results suggest that while these genes retain considerable sequence similarity with their *S. cerevisiae* counterparts, they do not appear to function in the meiotic regulatory cascade, as defined in *S. cerevisiae*. Thus, if *C. abicans* can undergo meiosis, it does so through an alternative pathway. I wish to emphasize that this work does not rule out the possibility that *C. albicans* can not execute a functional meiotic program, but rather introduces a new set of questions to be addressed. For example, if *C. albicans* can undergo meiosis, what is environmental signal for the initiation of the process and what signaling pathway is involved? On the other hand, if *C. albicans* has lost the capacity to undergo meiois, what is the advantage of this mode of evolution? The role of reproduction in the evolution of parasites has been controversial. For *Trypanosoma cruzi*, the causative agent of Chagas disease, clonal

growth appears to be advantageous. Indeed, it has been argued that asexual reproduction is advantageous to the parasitic mode of life (10). On the other hand, while *Toxoplasma gondii* has a clonal population structure, sexual recombination can enhance its virulence (11), although one must question whether this is advantageous as it fundamentally leads to destruction of its obligate habitat. In another fungal pathogen, *Cryptococcus neoformans*, a sexual cycle has been defined, however, the functional role of the sexual cycle and when and how it occurs is unknown (12). If *C. albicans* is capable of undergoing meiosis, discovery of how to induce this phenomenon in the laboratory will greatly facilitate research in understanding of the biology of this organism. In a broader evolutionary context, a deeper understanding of reproduction in *C. albicans*, will provide an additional basis for understanding the intricate affair between pathogens and their hosts.

References

1. Lachke, S. A., Lockhart, S. R., Daniels, K. J. & Soll, D. R. (2003) *Infect Immun* **71,** 4970-6.
2. Gerhart, J. & Kirschner, M. (1997) *Cells, embryos, and evolution : toward a cellular and developmental understanding of phenotypic variation and evolutionary adaptability* (Blackwell Science, Malden, Mass.).
3. Carroll, S. B., Grenier, J. K. & Weatherbee, S. D. (2001) *From DNA to diversity : molecular genetics and the evolution of animal design* (Blackwell Science, Oxford ; Malden, Mass.).
4. Graham, P., Penn, J. K. & Schedl, P. (2003) *Bioessays* **25,** 1-4.
5. Pane, A., Salvemini, M., Delli Bovi, P., Polito, C. & Saccone, G. (2002) *Development* **129,** 3715-25.
6. Saccone, G., Pane, A. & Polito, L. C. (2002) *Genetica* **116,** 15-23.
7. Tsong, A. E., Miller, M. G., Raisner, R. M. & Johnson, A. D. (2003) *Cell* **115,** 389-99.
8. Herrero, E., de la Torre, M. A. & Valentin, E. (2003) *Int Microbiol* **6,** 183-90.
9. Jones, T. F., Federspiel, N.A., Chibana, H., Dungan, J., Kalman, K., Magee, B.B. , Newport, G., Thorstenson, Y.R., Agabian, N., Magee, P.T., Davis, R.W. & Scherer, S. (2004) *Proc Natl Acad Sci U S A*, in press.
10. Ayala, F. J. (1998) *Proc Natl Acad Sci U S A* **95,** 3346-8.
11. Grigg, M. E., Bonnefoy, S., Hehl, A. B., Suzuki, Y. & Boothroyd, J. C. (2001) *Science* **294,** 161-5.
12. Nielsen, K., Cox, G. M., Wang, P., Toffaletti, D. L., Perfect, J. R. & Heitman, J. (2003) *Infect Immun* **71,** 4831-41.