INTRODUCTION
I. INTRODUCTION

I.1. Fungi and their Diversity

Fungi are eukaryotic, heterotrophic organisms as they consume organic forms of carbon for energy. Fungi exist in three basic forms viz., unicellular yeasts, filamentous hyphae (molds) and, among the most basal groups, flagellated, swimming, unicellular organisms that encyst to form sporangia. The yeasts, hyphae, and sporangia have cell walls that contain polysaccharide chitin, along with a variety of glucans. Fungi represent a single eukaryotic kingdom characterized by an osmotrophic growth habitat in which extracellular enzymes are secreted to break down complex substrates, the resulting simple sugars and amino acids being taken up by the growing fungus. Fungi exist in two distinct morphological growth forms, the unicellular yeasts (which grow by budding or simple fission) and the filamentous fungi (which produce polarized hyphal strands that aggregate to form a network called a mycelium). The osmotrophic growth habit of fungi is extremely effective for colonizing diverse habitats and has made fungi the principal degraders of biomass in all terrestrial ecosystems and also important pathogens of both plants and animals (39).

I.2. Fungi and Pathogenesis

Fungi may be pathogens of plants, animals, insects, and even of other fungi, but in the context of fungal diversity, there are relatively few pathogens. In humans and animals, a potential restriction on their diversity and numbers is the requirement to grow at the body temperature of the host. In plants, pathogens may grow at a variety of temperatures, depending on the climate favored by the plant and the season in which the fungus grows, so the diversity bottleneck on plant pathogens is less severe. Ninety percent of plant pathogens are fungi. Fungi and the distantly related Oomycetes
represent a majority of plant pathogens and can have a dramatic economic impact on agricultural crops. Sixty percent of plant diseases are caused by fungi (217). There are three lifestyles of plant pathogenic fungi including the necrotrophs, biotrophs, and hemibiotrophs (217). Necrotrophs infect, invade, and kill cells to obtain nutrients. Examples of necrotrophs include mold species in the genuses *Alternaria*, *Cochliobolus*, and *Botrytis*. Biotrophs extract nutrients from living cells often using specialized structures called haustoria. Rusts and powdery mildews are all biotrophs and include species like *Ustilago maydis* and *Puccinia graminis*. Biotrophs need living tissue to obtain nutrients and cannot live saprophytically. The third class is the hemibiotrophs that initially obtain nutrients from the plant without killing it, but later secrete enzymes that kill the host and obtain nutrients saprophytically from dead cells. The tomato pathogen *Cladosporium fulvum*, wheat pathogen *Fusarium graminearum*, and the rice pathogen *Magnaporthe grisea* are examples of hemibiotrophic ascomycete fungi. All three forms of plant pathogenic fungi can cause severe damage to crops including rice, maize, wheat and other cereals, peanuts, oilseeds, and strawberries (75). Fungi have developed the ability to secrete low molecular weight secondary metabolites with a diverse array of functions. These are produced from specialized synthetases called non-ribosomal peptide synthetases (NRPS) and polyketide synthetases (PKS) (107). Many *Fusarium* species are hemibiotrophs that secrete the secondary metabolite and plant hormones auxin and gibberellin to induce plant overgrowth and eventual death (200). *Aspergillus flavus* and *Aspergillus parasiticus* produce aflatoxins, one of the most potent carcinogens known, and are found growing on nuts and maize (159). These mycotoxins have been implicated as carcinogens and kidney toxins in humans. In humans, fungi can cause diseases ranging from athlete’s foot to thrush to meningitis.
Of the roughly 1.5 million fungal species estimated to be alive today, only some 200 have been associated with the human body so far, either as pathogens or as commensals, and of these, a dozen or so represent the most common fungal pathogens (78). A healthy human immune system confers resistance to most of these fungi. For example, *Aspergillus fumigatus*, one of the most pervasive fungus, is the major cause of respiratory allergies, but rarely causes fatal infections except in immunocompromised individuals (21). Similarly, the basidiomycete fungus, *Cryptococcus neoformans* is an opportunistic pathogen that only infects individuals with immune system defects, although its cousin *Cryptococcus gattii* caused an outbreak of disease in Vancouver, British Columbia in Canada that infected healthy individuals (87). Some fungi are primary pathogens, those that can infect incompetent hosts, including *Coccidioides immitis*, the cause of “valley fever” (41) and *Histoplasma capsulatum* (214). *Candida albicans* is a ubiquitous fungal organism that often colonizes the skin and the mucosal surfaces of normal individuals, without causing disease. However, when the normal host defence mechanisms are impaired (for example, in patients who are undergoing chemotherapy for malignancies, receiving immunosuppressants after an organ transplant, or patients with AIDS), *C. albicans* can become a pathogen (29). The other *Candida* species *C. albicans*, *C. tropicalis*, *C. parapsilosis*, *C. guilliermondii*, *C. dubliiniensis*, and *C. glabrata*, which are components of the normal microbiota can also cause superficial and systemic infections (32). As animals and fungi are both eukaryotes, they share many cellular processes and essential genes, making it difficult to develop effective fungicides which are benign to human hosts. In contrast, antibacterials have been easier to discover and develop as they target cell wall structures or aspects of bacterial cell growth not found in eukaryotic cells. Identification of
important fungal genes and biological pathways that are different from the fungi's hosts will provide candidate targets for anti-fungal drugs (29).

Not all fungi are harmful to humans and animals. Fungi have evolved many specialized ways to compete for resources in their local niche. This includes secretion of antibiotics or fungicides and the ability to grow filamentously to explore and forage more rapidly than single-celled competitors. Competition with bacteria led to the evolution of antibacterial products. The most famous of these, penicillin, is a secondary metabolite produced by *Penicillium* NRPS genes. Many other secondary metabolites are produced by fungi that have been harvested for human uses. Competition with other microorganisms probably leads to the evolution of fermentation capability and the evolution of the several alcohol dehydrogenase genes in the yeast *Saccharomyces cerevisiae* (197). Fungi play an important role in many bioindustrial processes. *Aspergillus niger* is one of the most efficient producers of citric acid (156), and is a well-known producer of extracellular fungal enzymes such as glucoamylase (192). The filamentous hemiascomycete *Ashbya gossypii* is used for industrial production of riboflavin (208). *Trichoderma reesei* is the main industrial source of cellulases and hemicellulases used to depolymerize biomass to simple sugars that are converted to chemical intermediates and biofuels, such as ethanol (130). The brewing yeast *S. cerevisiae* is used to make beer, bread, and wine while another fungus, *Aspergillus oryzae* is used in the brewing of sake and soy sauce (127). *Pichia stipitis* is the most efficient microbe for lignocellulose-bioconversion and xylose-fermentation (96).

**1.3. Fungal Phylogeny: Sub-kingdom Division**

Recently, a broad survey of the fungal kingdom using multilocus sequence typing analysis and evaluating the relationships among fungi and closely related organisms
based on the similarities in their genetic sequences provided evidence that the fungal kingdom contains at least eight distinct phyla (clusters of closely related organisms). According to the study (83), the phyla included in the kingdom Fungi include: Ascomycota, Basidiomycota, Zygomycota, Chytridiomycota, Glomeromycota, Blastocladiomycota, Neocallimastigomycota and Microsporidia. The Ascomycota and Basidiomycota are grouped under the sub-kingdom, the Dikarya (Fig. I.1.A).

The four major groups of fungal organism, i.e., Ascomycetes, Basidiomycetes, Zygomycetes, and Chytrids encompass 900 million years of evolutionary history leading to enormous biological diversity (65). Ascomycetes, under certain conditions, form ascospores inside the cell, while basidiomycetes develop external spores. Chytrids are unique among the true fungi in possessing zoospores, which move using posterior flagella, while Zygomycetes are fungi characterized by the formation of sexual spores (zygospores) and vegetative mycelium that lack septa (95). The Glomeromycota form a monophyletic group which probably diverged from the same ancestor as the Ascomycota and Basidiomycota. This group includes the ecologically and economically important Arbuscular Mycorrhizal fungi, crucial in the ecology and physiology of land plants (162). All species in this division are known to reproduce asexually form blastospores usually at the terminus of a hypha. Traditionally, the phylum Ascomycota is divided into three classes; the Euascomycetes (molds/filamentous fungi), the Archiascomycetes (of which S. pombe is the best-known representative) and the Hemiascomycetes (which groups all the ascomycetous yeasts) (149). In total, ~80 fungal genome sequences are currently publicly available with over 20 additional projects underway (NCBI Genome Project). These genomes represent important human pathogens, plant pathogens, saprophytes, and model organisms. They also encompass fungi that grow as yeasts, form mycelia or pseudo-hyphae, or are
Fig. 1.1. (A) Maximum likelihood phylogeny reconstructed using a concatenated alignment of 153 universally distributed fungal genes. The concatenated alignment contains 42 taxa and exactly 38,000 amino acid positions. Adapted from (62). (B) *Candida* clade Phylogeny. An average consensus supertree of CTG specific clade using 2146 gene families is shown. *Y. lipolytica* was chosen as an outgroup. Bootstrap scores are shown at all nodes. The alphabets A, B, C, D and F show splits in the alignment data. Adapted from (62).
capable of dimorphic (or polymorphic) growth. In addition, they include representatives of all four major fungal groups. i.e., Ascomycetes, Basidiomycetes, Zygomycetes and Chytrids. Importantly, the majority of available fungal genomes fall into clusters of related genomes that enable comparative analysis across a range of evolutionary distances (Fig. 1.A). These clusters also include related organisms that differ in terms of specific physiological traits (i.e. pathogenicity), thus allowing these traits to be explored through comparison.

1.4. Hemiascomycetes Fungi

The largest numbers of known yeast species belong to the class Hemiascomycetes, also known as budding yeasts, because they multiply by the formation of buds that subsequently detach from their mother cell and grow independently into a new mother cell. The Hemiascomycetous fungi are thought to have diverged from the filamentous Ascomycetous fungi between 300 and 900 million years ago (79). The Archiascomycetes subdivision contains a few yeast species known as fission yeasts because they divide by equal partitioning of an elongated mother cell into two daughter cells. Phylogenetically, Archiascomycetous yeasts are extremely distant from Hemiascomycetous yeasts; accordingly, their genome organizations are markedly different. The Hemiascomycetes class comprises one large order, the Saccharomycetales (111). The members of this large order share important morphological features such as (i) Single or chained asci, (ii) cell walls devoid of chitin, (iii) no or rudimentary hyphae and (iv) proliferation of vegetative cells by budding or fission (45).

_ S. cerevisiae_, a model Hemiascomycete, was the first eukaryote to have a complete chromosome sequenced (151), and later to have its whole genome sequenced
S. cerevisiae has found its importance due to high economic value in brewing, baking and biotechnology, and has been the model eukaryote for fundamental research. It has one of the smallest genomes among well-studied eukaryotes (only 14 million basepairs), a high gene density (72% of the genome codes for protein), few introns and little repetitive DNA (47). These same features have more recently led to yeast species emerging as a powerful eukaryotic model system for comparative genomics and studies of genome evolution. The group of fungi that includes S. cerevisiae—the Hemiascomycetes—all have relatively small and non-repetitive genomes, and several are of biotechnological or medical interest. The result is that today we have genome sequences (either complete, or high-quality draft sequences) from 20 species of Hemiascomycetes. The genome of S. cerevisiae has now been sequenced three times: once from the laboratory strain S288C (70), once from the clinical isolate YJM789 (74), and once from the vineyard isolate RM11-1a (Broad Institute, unpublished; GenBank accession number AAEG01000000). Apart from the S. cerevisiae genome the sequenced Hemiascomycetes genomes include Saccharomyces bayanus, Saccharomyces castelli, Saccharomyces kluyveri, Klyveromyces lactis, Klyveromyces waltii, Saccharomyces paradoxus, Saccharomyces mikatae, Saccharomyces kudriavzevii, C. glabrata, Ashbya gossypii, Klyveromyces waltii, Yarrowia lipolytica, C. tropicalis, C. guilliermondii, C. lusitaniae, C. dubliniensis, Debaryomyces hansenii, C. albicans (SC5314 and W0-1 strains), Lodderomyces elongisporus and Pichia stipitis.

Phylogenetic analyses of ribosomal DNA (rDNA) and RNA polymerase II (RPB2) gene sequences support a single evolutionary origin of the Saccharomycetales and have been resolved by multigene sequence analysis into 14 clades (111). The Saccharomyces sensu stricto yeasts, including S. bayanus, S. cerevisiae,
S. kudriavzevii, S. mikatae and S. paradoxus, represent an isolated and well supported monophyletic group with overall phenotypic similarity (112). These species can mate with each other, but interspecific pairings result predominantly in sterile hybrids (147). While the post-zygotic barrier already exists among sensu stricto species, an efficient prezygotic barrier is still absent and therefore, this group of yeasts is still in the early stages of species formation. An ancestor of S. cerevisiae underwent whole genome duplication (WGD) after it diverged from non-WGD yeast lineages such as Kluyveromyces lactis, Kluyveromyces waltii, and Ashbya gossypii (Fig. I.1.A) (48, 213). The WGD had a major impact on the evolution of S. cerevisiae and its relatives, most notably by facilitating their adaptation to anaerobic growth (158) and contributing to their rapid speciation (169). In S. cerevisiae, 20% of genes are members of duplicated pairs that were formed in the WGD (26). The other loci became single-copy again during the sorting-out process (genome reduction), that occurred after the WGD. Similar large-scale loss of copies of duplicated genes from paleopolyploid genomes has occurred during the evolution of plants such as grasses and crucifers (155).

The Hemiascomycetes contain many species of practical importance and scientific interest, including several species that are pathogenic for humans, such as C. parapsilosis, C. tropicalis, and C. albicans, which is the most common human pathogenic fungus (28). During the translation of mRNA to polypeptides, several species of Candida exhibit alternative codon usage. The reassignment of the codon CUG from leucine to serine was first described by Kawaguchi et al. for C. cylindraceae (106). A phylogenetic analysis among Candida species having alternate codon usage showed that only 11 of 78 species of Candida used CUG as a codon for leucine (Fig. I.1.A.). The remaining 67 species translated CUG as serine, but they did not form a monophyletic group (190). The authors suggested a correlation between codon
reassignment and coenzyme Q9 (Co-Q9, mitochondrial electron carrier) as the predominant ubiquinone in these species based on chemotaxonomy. Co-Q has various numbers of isoprene units and the length of the isoprene chain is usually consistent within a monophyletic group and has therefore been used in yeast taxonomy. The 

*Candida* clade organisms with available genome sequences include *C. tropicalis*, *C. albicans*, *C. dubliniensis*, *C. parapsilosis*, *C. lusitaniae*, *C. guilliermondii*, *Debaromyces hansenii* and *Lodderomyces elongisporus*. A phylogenetic analysis using Supertree (212) and alignment of multiple proteins, shows that *Candida* clade organisms form a monophyletic clade, with two sub-clades (62). The first clade contains *C. lusitaniae*, *C. guilliermondii* and *Debaromyces hansenii* and the second group is formed by *C. tropicalis*, *C. albicans*, *C. dubliniensis*, *C. parapsilosis* (Fig 1. B). *C. lusitaniae* and *C. guilliermondii* are haploid yeasts, and are apparently fully sexual (210, 221). *D. hansenii* is homothallic, with a fused mating locus (48, 58). In contrast, members of the second clade have at best a cryptic sexual cycle and have never been observed to undergo meiosis (88, 89). In the first clade, *D. hansenii* and *C. guilliermondii* form a sister group with the exclusion of *C. lusitaniae*. The industrially important *Candida* species, *Pichia stipitis* is known to be closely related to the *D. hansenii* and *C. guilliermondii* clade (96).

Phylogenetic distance between the pathogenic *Candida* species and the hierarchy of pathogenicity in *Candida* species do not correlate. Phylogenetic analysis of rRNA sequences has confirmed that *C. dubliniensis* and *C. albicans* are the two most closely related *Candida* species of clinical importance in humans. *C. albicans* is the most significant yeast pathogen responsible for superficial and deep-seated infections. *C. dubliniensis* is of lesser clinical importance as it is found in mucosal infections in non-HIV-infected patients, and is relatively insignificant in the case of bloodstream
infections (138). *C. dubliniensis* is less virulent than *C. albicans* in a murine model of systemic candidosis (203). The reason for the apparent difference in virulence between the two species is unknown as they are phenotypically very similar and seem to share many of the traits traditionally associated with virulence in *C. albicans*. In particular, both species have the ability to form true hyphae, to adhere to human epithelia and to produce secreted aspartyl proteinases. However, *C. dubliniensis* does not form hyphae as rapidly as *C. albicans* in response to shifts in pH/temperature or when incubated in serum (67). In contrast, when cultured on Staib agar or Pal's agar, *C. dubliniensis* forms abundant hyphae, pseudohyphae and chlamydospores, whereas *C. albicans* remains in the yeast phase (1, 184). *C. dubliniensis* also seems to be more sensitive to environmental stress, such as elevated temperature and NaCl concentration (138). Consistent with the fact that *C. dubliniensis* is generally less efficient than *C. albicans* at forming hyphae in response to serum, a comparative genomic hybridization study using *C. albicans* DNA microarrays revealed the absence and divergence of several genes and gene families in *C. dubliniensis* (203). These include putative virulence factors and many genes specific for or preferentially expressed in the hyphal phase, such as *SAP5*, *SAP6*, *HWP1* and *HYR1*. The array CGH data indicates the presence of genes homologous to many TFs involved in hypha formation in *C. albicans* (e.g. *EFG1*, *CPH1*, *TUP1*).

1.5. The **Candida albicans** Genome

*Candida* species are the most important opportunistic fungal pathogens of humans responsible for superficial and systemic infections (157). Among these species, *C. albicans* is responsible for the majority of infections, but other species are becoming increasingly common (157). Because of its predominance, *C. albicans* has been the
focus of genomic and molecular studies over the last 20 years, becoming a model organism for other pathogenic *Candida* species and fungal pathogens. The *C. albicans* genome was made publicly available by the Stanford Genome Technology Center at the end of the 1990s and different assemblies and annotations have been released since (24, 99, 202). This has been accompanied by the implementation of two main genomic databases: CandidaDB and the Candida Genome Database. A number of molecular techniques have been applied to understand the pathogenesis of *C. albicans* as well as to search for novel drug targets. However, *C. albicans* presents several difficulties for molecular biologists as it is diploid, only a part of its sexual cycle has been demonstrated, it has a very plastic genome, and it is highly heterozygous. Each of these properties is best investigated through a genomic approach. Hence, knowledge of the genome sequence has been an important goal for the past 10 years. More recently, genome structure and dynamics have become increasingly important in this organism as widespread aneuploidy (34, 172), the role of repeated DNA in chromosome loss (117), and chromosome rearrangements leading to drug resistance have been reported (173). A diploid assembly constructed from 10.9× coverage (Assembly 19) provided single contigs as well as allelic contigs. A finished sequence (Assembly 20) was recently published which assigned each contig to the eight different chromosomes. Assembly 21 was carefully constructed from a variety of resources including the sequences of *C. albicans* strain WO-1, the sister species *C. dubliniensis*, primary traces used to generate Assembly 4, the STS map and a whole-chromosome optical map (202). This assembly has eight linear DNA sequences including nine copies of the intermediate repeat called the Major Repeat Sequence (MRS), of which three have been completely sequenced.
I.6. *Candida albicans* Morphogenesis

The different morphological forms associated with *C. albicans* makes it unique among the human pathogens. Due to the presence of distinct morphological forms, governed by different environmental cues, this important pathogen can be considered as a polymorphic or pleomorphic organism. *C. albicans* can exist in three forms that have distinct shapes: yeast cells (also known as blastospores), pseudohyphal cells and true hyphal cells (18, 188). The elongated hyphae evade or escape phagocytic cells and yeast cells disseminate in the bloodstream suggesting that morphology contributes to the survival of *C. albicans* (167). The ability to switch rapidly from yeast-form growth to hyphal growth is a defining characteristic of *C. albicans* cells, and it is believed that each form of growth provides critical functions required for the pathogenic lifestyle (188). Yeast cells are round to ovoid in shape and separate readily from each other. Pseudohyphae resemble elongated, ellipsoid yeast cells that remain attached to one another at the constricted septation site and usually grow in a branching pattern that is thought to facilitate foraging for nutrients away from the parental cell and colony (19). True hyphal cells are long and highly polarized, with parallel sides and no obvious constrictions between cells (19, 188).

These different morphologies are often treated as different developmental states. Laboratory grown cultures at low temperature and low pH contain mostly ellipsoid yeast cells. Long, narrow hyphae develop from yeast cells grown at 37°C and neutral pH; external stimuli such as serum, N-acetyl-glucosamine, spider medium also induce hyphae. Elongated pseudohyphal cells develop at intermediate temperatures and pH, and in nitrogen deficient medium. Furthermore, pseudohyphal cultures always contain some yeast and/or some hyphal cells.
Yeast cells grow by asymmetric budding, forming smooth, round colonies. Septin rings appear before bud emergence and nuclei divide across the mother bud neck (206). Yeast cells separate after cytokinesis, when daughter cells have not yet reached the size of their mother cells. Yeast and pseudohyphae of *C. albicans* are similar to those of *S. cerevisiae* in shape, size and follow similar cell cycle steps. Both in *S. cerevisiae* and *C. albicans* polarized growth at the tip, polarized deposition of cell wall material (that is required for septation) and changes in actin patch distribution signify filament formation. However isotropic growth occurs throughout the bud. This switch is an early cell cycle event in the yeast form and occurs later in the pseudohyphal form (18). *C. albicans* pseudohyphal cells bud in a unipolar pattern. The cells remain attached after cytokinesis, forming branched chains of elongated buds and colonies that are fibrous or rough. Filaments invade the agar below the colony and extend across the agar from the colony edge. As in yeast form cells, septin rings form before bud emergence, and nuclei divide across the neck (189). As with *S. cerevisiae* pseudohyphae (110), *C. albicans* pseudohyphal cells spend more time growing in a polarized manner and remain in G2 longer than do yeast cells. Hyphae are narrower than pseudohyphal cells (~2 μm) and have parallel walls with no obvious constriction at the site of septation (188).

In addition to the presence of the septin ring and nuclear division, other characteristics distinguish a *C. albicans* cell initiating a bud from those initiating a germ tube that elaborate into hyphae an important difference is the presence of a structure called a Spitzenkorper near the tip of the growing cell that is implicated in polarized growth (204). Spitzenkorper is present in filamentous fungi such as *Aspergillus* and *Neurospora* near the tip of the growing cell. Spitzenkorper is believed to direct secretion to the growing tip and thus has a function similar to that of the
polarisome of the budding yeast (204). However the growing tips of all three 
*C. albicans* morphological forms have another structure called polarisome.

Much of the analysis of the yeast-hyphal transition has been concerned with transcriptional control and the relevant transcription factors. The initial observation that the combined loss of the Cph1p and Efg1p transcription regulators blocked the hyphal transition in most tested conditions, and led to reduced virulence, focused interest on these two transcription factors and their controlling networks (124). Subsequently, a large number of transcription regulators have been reported to play roles, either positive or negative, in the yeast-to-hyphal transition. The APSES protein Efg1p plays a central role in the general control of morphogenesis, in the yeast-hyphal transition the regulatory circuit leading through Efg1p is linked to cAMP metabolism. The role of the cAMP-mediated signaling pathway has been well established in the yeast-to-hyphal switch (124), as a variety of mutations that affect the pathway influence hyphal development (9, 37, 100, 136).

Other pathways that transmit external signals to the transcription machinery, including a MAP kinase signaling pathway, an external pH sensing pathway and a matrix-sensing pathway, have also been identified as playing roles in hyphal development under specific conditions (109, 115). The target of the MAP kinase cascade is proposed to be the Cph1p transcription factor on the basis of comparison with *S. cerevisiae*, and Cph1p provided a residual level of virulence in the efg1 mutant background. Apart from above two transcription factors, a number of TFs have been identified which regulate yeast to hyphae transition in one or other conditions. This include alkali induced hyphal regulation by Rim101p (44), matrix induced hyphae regulation by Czflp (25), serum induced hyphae by Flo8p and Tec1p (31, 171), Ace2

15
regulated hypoxia induced hyphae (108) and Hap5p and Ash1p controlled filamentation in Spider medium (92, 98).

The central cell cycle regulatory kinase, as well as a group of regulatory cyclin molecules, has been identified in *C. albicans*, and several of these molecules appear to have specific roles in morphogenetic control. In *C. albicans* a single cyclin dependent kinase (Cdc28p) is the key regulator of cell cycle progression, and this kinase interacts with a variety of cyclins to control the G1, S, and mitotic transitions, as well as aspects of morphogenesis (201). Several G1 cyclins like Ccn1p, Hgc1p and Cln3p are required for the hyphal development under variety of conditions but not for the initiation of hyphae formation (7, 125, 223). B-type cyclins and the cyclin-dependent protein kinase Cdc28p are also implicated in control of morphogenesis, but not directly in control of the yeast-to-hyphal transition. Loss of the B-type cyclin Clb4p caused pseudohyphal growth, whereas shut-off of the essential cyclin Clb2p resulted in more highly polarized filaments called elongated buds (209). Similar filamentation has been seen in mutants that lack function in other essential cell cycle components (6).

I.7. Pseudohyphal Growth

*C. albicans* cells exhibit several patterns of growth that involve extensive cell elongation but do not involve the formation of true hyphae. The pseudohyphal pattern is the best studied of these filamentous growth modes. Pseudohyphal growth involves elongated cells that remain in chains but are separated by true constrictions, and not by septa. It has been controversial whether pseudohyphal growth is an intermediate stage in the yeast-to-hypha transition, but current work suggests that pseudohyphal cells represent a discrete cell growth pattern (188). The absence of a Spitzenkorper, and the positioning of nuclear division across the mother-daughter junction, suggest that
pseudohyphal growth is closer to yeast growth than to hyphal growth (209). The major distinctions from the yeast form are the extended period of polarized growth exhibited by the pseudohyphal cells and the tendency for the cells to remain attached even though proper septa have formed between them. Some environmental conditions, such as medium rich in phosphate (85) or in alkanes (183), have been reported to stimulate the pseudohyphal pattern of division. In addition, a number of mutants lead to constitutive pseudohyphal growth. Several of these mutations are in genes involved in cell cycle regulation, implicating this process in the control of this morphogenetic state. This involvement is logical because the switch from yeast growth to pseudohyphal growth arises from a subtle modulation in the length of time the cell spends in the polarized growth mode relative to the isotropic growth mode. Inactivation of the Fkh2p transcription factor implicated in the regulation of mitotic cyclins, or of the Fkh2p target cyclin Clb4p, generates cells that remain pseudohyphal under both hyphae-inducing and yeast-growth conditions (17). Loss of Grr1p, an F box protein involved in the degradation of the G1 cyclins Ccn1p and Cln3p, leads to a pseudohyphal growth state (120). Perhaps extension of G1 phase by stabilizing the G1 cyclins or reducing the G2 cyclins can trigger the pseudohyphal state. In C. albicans, the Nim1p-related kinases Gin4p and Hsl1p function in the regulation of the pseudohyphal state; mutation of either kinase leads to constitutive pseudohyphae (211). In addition to cell cycle perturbations that lead to constitutive pseudohyphae, defects in the Tup1p regulatory circuit lead to this phenotype. Tup1p and the associated protein Ssn6p play important regulatory roles in S. cerevisiae; they interact with different DNA binding partners to provide a generic gene repression function (181).

For specific transcription modules, the functional consequences of loss of Tup1p or Ssn6p can be different; therefore Tup1p and Ssn6p do not provide completely
overlapping functions. In *C. albicans*, Tup1p inactivation causes cells to proliferate in a pseudohyphal growth mode (22). However, the ssn6Δ mutant does not trigger equivalent pseudohyphal development in *C. albicans* (66), although normal morphology is perturbed. The global transcriptional consequences of the *tup1Δ* and *ssn6Δ* mutations are different (66), suggesting that the Ssn6p corepressor is not a critical component of the Tup1p regulatory circuit that represses the pseudohyphal growth pattern. An alternative tetratricopeptide repeat protein, Tcc1p, may act as a corepressor with Tup1p in the repression of pseudohyphal growth, as loss of Tcc1p creates similar phenotypes and affects patterns of gene expression similar to the loss of Tup1p (103). It appears the Nrg1p DNA binding protein targets the Tup1p (Tcc1p) corepressor to a variety of promoters involved in regulation of morphogenesis, as loss of Nrg1p function leads to constitutive pseudohyphal growth (66). Microarray analysis suggests that Tup1p and Nrg1p influence distinct but overlapping gene sets. A second DNA binding protein, Rfg1p (101), appears to function as a targeting element for Tup1p (Tcc1p); loss of Rfg1p leads to a similar constitutive pseudohyphal phenotype, but the networks of gene expression modulated by Nrg1p and Rfg1p are different. *NRG1* is repressed by hyphal inducing conditions, and Rfg1p may be shut off posttranslationally (101, 140). Thus inactivation of this negative regulatory circuit is a critical component of morphogenetic control in *C. albicans*.

In addition to this yeast-hyphal transition that defines *C. albicans* as a dimorphic fungus, a number of other naturally occurring morphological forms are characteristic of specific cellular functions. These distinct morphologies include the opaque form, characteristic of mating-competent cells (182); the chlamydospore, a currently enigmatic thick-walled cell formed typically under suboptimal growth conditions (134) and the hyperpolarized growth under certain cellular stresses or in the
absence of certain gene products that influence the cell cycle (e.g. HU and nocadazole treatment). The elongated buds resemble true hyphae in that they maintain polarized growth at the tip and do not show periodic constricted growth as do pseudohyphae. They also demonstrate nuclear movement out of the mother yeast cell and into the filament. Similar to pseudohyphae, a constriction is present at the junction between the yeast cell and elongated bud, which is consistent with the bud forming prior to the cell cycle arrest and the elongated growth mode. The elongated buds eventually die, which is not unexpected given that they cannot continue the cell cycle.

1.8. Role of Spt10p-Spt21p in histone gene transcription in S. cerevisiae

The regulation of histone synthesis is a critical aspect of eukaryotic cell division because histones play essential roles in all aspects of chromosome function. In S. cerevisiae, altered histone levels have been shown to impair chromosome segregation, transcription, and other processes (153). The regulation of histone gene transcription in S. cerevisiae is complex. Two sets of divergently transcribed gene pairs encode histones H2A and H2B (HTA1-HTB1, HTA2-HTB2) while two other loci encode histones H3 and H4 (HHT1-HHF1 and HHT2-HHF2). While transcription of all four histone loci is cell-cycle regulated, with peak expression during S phase, the promoters at the four loci are highly divergent and are dependent on overlapping but distinct sets of transcription factors (80, 216). The insertion of Ty retrotransposons or Ty long terminal repeat sequences (LTRs or δ-elements) into or near the promoters of genes can cause transcriptional defects (20). The suppressors of these transcriptional defects have helped identify a large set of genes referred to as the SPT genes. Two functionally related genes, SPT10 and SPT21, were isolated in two different selections for spt mutants (59, 146). Studies of spt10 and spt21 mutants have shown that they share
several related mutant phenotypes in addition to their Spf$^-$ phenotype i.e., impaired transcription of particular histone genes, suppression of the transcriptional defects caused by loss of upstream activating sequences or activators at particular genes, and defects in chromatin structure (46, 104, 133, 146, 160, 220).

Spt10p and Spt21p play important roles in the transcription of histone genes. It has been shown that Spt10p and Spt21p are necessary for the transcription of two of the four histone loci and that they bind to the promoters of all four histone loci (46, 80). Studies on Spt10p and Spt21p have suggested a possible mechanism for their activation of histone gene transcription. First, Spt10p has motifs found in acetyltransferases, including the histone acetyltransferase (HAT) Gcn5 (148). Acetyltransferases have been shown to play critical roles in transcription through their ability to acetylate both histones and other transcription factors (187). Consistent with a role for Spt10p in acetylation of histones, recent work has shown that the transcriptional induction of the \textit{CUP1} gene is defective in \textit{spt10} mutants and that this defect correlates with a decreased level of acetylation of histones H3 and H4 at the \textit{CUP1} promoter (176). Second, \textit{SPT21} mRNA levels are cell cycle controlled, with a peak at the G1/S transition, consistent with a role in controlling histone gene transcription (35). Thus, at the time when histone genes are transcribed, Spt21p may play a role in activating Spt10p to enable it to activate transcription of particular histone genes through its putative acetyltransferase activity. Moreover it has been shown that, the activation by histone genes by Spt10p is dependent on the Spt10p acetyltransferase domain. Both Spt10p and Spt21p are physically associated with the promoters of histone genes, strongly suggesting a direct role for these factors in transcriptional control. Furthermore, Spt21p is required for the association of Spt10p with histone promoters at S phase and that Spt21p directly interacts with Spt10p (80). Recently it has been shown that Spt10p acts
as a sequence specific DNA binding protein that recognizes histone upstream activating sequences (UAS) element with consensus of (G/A)TTCC(N)₆TTCNC (56).

I.9. The Ams2 GATA factor in *S. pombe*

Like other eukaryotes, the fission yeast, *Schizosaccharomyces pombe* has reiterated histone genes, the genomic organization of which consists of a single H2Aβ (*hta2*⁺), a pair of H2A-H2B (*hta1⁺-htb1⁺*), and three pairs of H3-H4 (*hht1⁺-hhf1⁺*, *hht2⁺-hhf2⁺* and *hht3⁺-hhf3⁺*) histone genes (132). In *S. pombe*, Ams2 is a cell cycle-regulated GATA-type transcription factor that acts as an activator of core histone genes at S-phase (33). Ams2 accumulates on the nuclear chromatin when chromosomes are duplicated during S-phase, whereas little accumulation of this protein was detected in mid-to-late G2 or early M-phase (194). The three histone gene pairs transcriptional regulation is controlled by a combination of Ams2-dependent transcriptional activation at S-phase and Ams2-independent constitutive transcription throughout the cell cycle. Ams2 binds directly to the promoter regions of all the core histone genes in vivo in an AACCCT-box-dependent manner (194). The transcription factor Ams2 was originally identified as one of four multicopy suppressors in the SpCENP-A (Centromeric Histone H3 variant) ts mutant *cnpl-1* . CENP-A is an evolutionarily conserved histone H3 variant (154, 191). CENP-A is essential for chromosome segregation in all organisms so far investigated. It has been shown to be specifically localized only at active centromeres and CENP-A-containing nucleosomes possibly serve to form the primary kinetochore-building scaffold (193). Ams2 is dispensable for cell viability, but Ams2-depletion results in growth retardation with a high frequency of chromosome missegregation (33). Ams2 is shown to be a nuclear protein in *S. pombe* that is...
enriched in the central centromeres and at promoter regions of presumed target genes directly to the GATA-like DNA sequences in vitro (33).

I.10. C. albicans Histone Genes and Centromere Organization

The genome of C. albicans has two sets of divergently transcribed gene pairs encode histones H2A and H2B (HTA1-HTB1, HTA2-HTB2) and two other sets of loci encoding histones H3 and H4 (HHT1-HHF1 and HHT2-HHF2). This organization is similar to S. cerevisiae histones. However the genome of C. albicans contains an extra copy each of histone H2A (HTA3) and H3 (HHTI). The regulation of histone genes in C. albicans has not been studied. Expression profiling showed that histone genes (orf19.1051, orf19.6924 orf19.1853, orf19.1061, orf19.1059) are repressed under the hyphal growth conditions (YPD, 37°C) and up-regulated at 30°C (54). The significance of the repression of histone expression during hyphal growth is not understood, however, a special role for reduced histone dosage during morphogenesis is an interesting possibility. In mammalian cells, hydroxyurea creates a rapid decrease of the histone mRNA pool (14), while in C. albicans, hydroxyurea is able to induce filament formation (6). It is therefore possible that a decrease in the amount of histone message plays a role in hyphal development.

The centromere (CEN) is essential for the precise segregation of chromosomes during the cell cycle. Although CEN function is conserved across eukaryotes, the CEN DNA sequence varies between and even within species. For example, CEN DNAs in budding yeasts such as S. cerevisiae are only ~125 bp long and are known as point centromeres. In contrast, CEN DNAs in most organisms are relatively large (40–4,000 kb). Termed as regional centromeres, they may contain tandem repeated DNA, inverted repeats, retroelements, and unique DNAs (53). Despite such significant sequence
variation, CEN DNA is associated with specialized nucleosomes containing a unique histone H3 variant (CENP-A/Cse4p family) in every eukaryotic system studied to date. Mechanisms that determine the targeting of specialized nucleosomes containing CENP-A to the CEN DNA sequences, however, remain largely unknown. Such a specialized functional association between CENP-A and divergent CEN DNA sequences is especially intriguing when chromosomes within a species harbor unique and different CEN DNAs.

*C. albicans* centromere on each of the eight chromosomes are uniform in size (~3-4.5 kb) and do not contain any conserved DNA sequence motifs (13, 168). The centromeric regions in *C. albicans* and the 4–7 kb central core regions (cnt) of *S. pombe* centromeres are similar in size, DNA sequence heterogeneity, and in sharing an unusual chromatin structure (13). Most of the *C. albicans* centromere regions completely lack IR sequences and thus are analogous to the central core (cnt) regions of fission yeast centromeres. Those with IR sequences (such as CEN5) are structurally analogous to the cnt and inner-most inverted repeat (imr) regions of fission yeast centromeres (135, 168). Despite this structural analogy, the localization of Cse4p differs somewhat in these two yeast species. In fission yeast, both the cnt and imr regions are bound to Cnp1p (Cse4p homolog), while in *C. albicans*, CaCse4p binds only to core sequences but not to any flanking inverted repeats. Thus the *C. albicans* centromeres are analogous to the *S. pombe* CEN central cores, and do not activate de novo because of the lack of activator regions similar to the outer repeat region. It has been proposed that unknown epigenetic processes could mark centromere formation and function in *C. albicans*. 