

Functional Profiling of Human Fungal Pathogen Genomes

Alexi I. Goranov and Hiten D. Madhani

Department of Biochemistry and Biophysics, University of California, San Francisco, San Francisco, California 94158

Correspondence: hitenmadhani@gmail.com

Fungal infections are challenging to diagnose and often difficult to treat, with only a handful of drug classes existing. Understanding the molecular mechanisms by which pathogenic fungi cause human disease is imperative. Here, we discuss how the development and use of genome-scale genetic resources, such as whole-genome knockout collections, can address this unmet need. Using work in *Saccharomyces cerevisiae* as a guide, studies of *Cryptococcus neoformans* and *Candida albicans* have shown how the challenges of large-scale gene deletion can be overcome, and how such collections can be effectively used to obtain insights into mechanisms of pathogenesis. We conclude that, with concerted efforts, full genome-wide functional analysis of human fungal pathogen genomes is within reach.

The goal of this article is to describe the promise and challenges of systematic and comprehensive functional approaches to decipher the mechanisms of pathogenesis by human fungal pathogens. To set the stage, we begin with a description of the multifaceted roles of fungi in human affairs before introducing the two broad classes of fungi capable of infecting humans.

FUNGI IN HUMAN WELFARE

The fungal kingdom is arguably the most diverse and successful of the eukaryotic kingdoms. With more than 100,000 described and five million inferred fungal species inhabiting most niches in the biosphere, only bacteria match the diversity of fungi (O'Brien et al. 2005). Fungi are typically saprophytic organisms that extract energy by decomposing organ-

ic matter and, thus, play an essential role within the environment: recycling organic matter. Fungi also have many important industrially useful properties. Besides the familiar edible mushroom and brewer's and baker's yeast, other fungi produce a variety of bioactive molecules, many of which are well-known drugs, such as penicillin. A major threat to humans from fungi is the destruction of crops. Collectively, plant pathogenic fungi cause 10%–40% of crop loss worldwide (Fisher et al. 2012).

HUMAN FUNGAL PATHOGENS

Although many fungi are plant pathogens, relatively few (about 200 species) can infect animals or humans (Richardson and David 2003). It has been suggested that this is because most fungi cannot grow at mammalian body temperature

Editors: Arturo Casadevall, Aaron P. Mitchell, Judith Berman, Kyung J. Kwon-Chung, John R. Perfect, and Joseph Heitman
Additional Perspectives on Human Fungal Pathogens available at www.perspectivesinmedicine.org

Copyright © 2015 Cold Spring Harbor Laboratory Press; all rights reserved; doi: 10.1101/cshperspect.a019596
Cite this article as *Cold Spring Harb Perspect Med* 2015;5:a019596



(indeed, it has been suggested to be why mammals have been successful). Another likely reason is that the innate and adaptive immune systems of mammals provide an effective defense against infection by most fungi. A fungal infection, or mycosis, can be categorized as cutaneous, subcutaneous, or systemic. Cutaneous mycoses are quite common, for example, tinea (“ringworm”) infections caused by *Microsporum*, *Trichophyton*, and *Epidermophyton* fungi, but these are generally easily resolved. Systemic infections are a major clinical problem and challenging to diagnose and treat. They are caused either by fungi that are primary pathogens (e.g., *Histoplasma*, *Blastomyces*, and *Coccidioides* species [spp.]), which infect immunocompetent hosts, or opportunistic pathogens that require some degree of host compromise for infection (e.g., *Aspergillus*, *Candida*, *Cryptococcus*, *Mucor* spp.).

Presumably, the ability of these pathogens to cause disease involves the requirement for a specific set of properties, not shared with most fungi. Work over the last few decades has shed some light on the underlying molecular mechanisms. There are several simple requirements: proliferation at body temperature of 37°C, adaptation to hostile environments limited in essential nutrients, and specialization to overcome otherwise effective immune barriers, such as recognition and phagocytosis, and killing by macrophages (Levitz and DiBenedetto 1989). Fungi appear to have found a range of solutions for these problems, emphasizing the need for functional studies of a range of pathogens.

THE SYSTEMATIC PERSPECTIVE

Considering the question of how to most effectively approach the biology of the human pathogenic fungi, it is clear that comprehensive genetic and genomic resources represent an efficient set of tools that would enable the field to develop a detailed molecular understanding of fungal mechanisms of infection and drug sensitivities. In support of this perspective, we will briefly highlight the success of such approaches in baker’s yeast, as well as the initial applications of systematic genome-scale genetic studies in human fungal pathogens. Of the many impor-

tant human fungal pathogens, the application of the approaches described above is most advanced in *Cryptococcus neoformans* and *Candida albicans* and, thus, these will be the focus here. However, the approaches are broadly applicable once DNA transformation and homologous replacement strategies are developed in a given organism. The prospect for the latter is rapidly improving because of the development of CRISPR-based genome editing technology (Mali et al. 2013).

Saccharomyces cerevisiae AS A PROTOTYPE

Our perspective is informed by the recent history of genomics, which was driven by studies of *Saccharomyces cerevisiae*. We review the highlights below.

From Genome to Public Resources

S. cerevisiae was the first eukaryotic organism whose genome was sequenced and annotated (Goffeau et al. 1996). The available genomic data paved the way for the undertaking of the construction of a deletion library for every non-essential gene, as well as a diploid heterozygous for essential genes (Giaever et al. 2002). Critical for the success of this effort was the fact that homologous recombination is unusually efficient in *S. cerevisiae*, with 50 bp being sufficient for targeting homology (Bahler et al. 1998; Longtine et al. 1998). This allows cloning-free, polymerase chain reaction (PCR)-based construction of targeting vectors. Moreover, transformation using lithium chloride and polyethylene glycol has been highly optimized, as has the ability to rapidly screen colonies for homologous replacement events (Fink and Guthrie 2002). In addition to the well-known gene-deletion library, a number of useful ordered whole-genome libraries have been created. Among these is a library that allows for the overexpression and affinity purification of every yeast protein, which has allowed for efficient activity-based biochemical screens (Martzen et al. 1999), as well as a comprehensive library of genes tagged with green fluorescent protein, which

allowed for the determination of protein localization at an unprecedented scale (Ghaemmaghani et al. 2003)

The Efficiency of the Systematic Approach

The ease of genetic manipulations, which include sexual crosses, combined with the ability to systematically assay the behavior of strains lacking every nonessential gene in a controlled, parallel manner under any desired laboratory condition or in combination with other mutations (synthetic genetic array), rapidly enhanced the understanding of the biology of *S. cerevisiae* (Botstein and Fink 2011). Two additional advances enhanced the utility of the deletion collection. First, robotics for moving arrays of colonies between agar plates was developed. Second, strains and methods were developed that allowed for the selection of the products of a sexual cross in a high-throughput manner (Tong and Boone 2006). The most impressive example of the application of these tools was the generation and analysis of 5.4 million double mutants (Costanzo et al. 2010). Such double mutant analysis (also termed genetic interaction mapping) has been shown to be very useful for linking genes to functional categories or physiological processes. Deletions of single genes often show no easily detectible phenotypes because other genes can perform similar functions inside the cell and can, therefore, “cover” for the absence of the deleted gene. This is the notion of redundant function and, at one extreme, only the double mutant displays a phenotype, producing a so-called synthetic genetic interaction. In reality, nearly all single-gene deletions produce a measurable phenotype (Breslow et al. 2008), and it is in cases in which the double mutant produces more than a multiplicative defect that it is considered a synthetic interaction.

With the availability of the ordered genetic arrays of deletion mutants, the rapid screening for genes whose absence causes sensitivity or resistance to biological compounds of interest (chemogenomics) has become possible (Wuster and Madan Babu 2008). Such studies have been useful for determining drug mode-of-action and gene annotation (as gene deletions with

similar phenotypic signature tend to specify factors that function in the same pathway). Because the individual deletion strains of the yeast deletion collection each carry a unique DNA bar code, experiments in which the entire deletion collection is pooled and subject to chemical challenge have also been performed. Custom bar code microarrays or, more recently, bar code sequencing, have been used to assess the fitness of individual strains in the pool after treatment. Such systematic chemical-genetic screens are now routinely used (Lopez et al. 2008; Wuster and Madan Babu 2008).

In addition to allowing for the systematic generation of deletions, the availability of the annotated genomic sequence for budding yeast led to the generation of the first whole-genome DNA microarrays, which allowed for the assessment of messenger RNA accumulation for all genes at the same time (expression profiling) (DeRisi et al. 1997). This ability to assay the expression of every gene at once transformed how questions are framed and investigated (Botstein and Fink 2011). By determining what sets of genes are induced or repressed by a given condition, one can quickly formulate hypotheses as to how the condition affects the physiology of the cell. Although microarray technology is being superseded by RNA-seq, expression profiling is a mainstay in genetic analysis. Likewise, systematic whole-genome chromatin immunoprecipitation of all annotated DNA-binding regulatory proteins has impacted studies of the structure, function, and evolution of transcription circuits. The latter is only beginning to be applied to fungal pathogens, whereas transcriptomics is widespread in its application attributable to the ease of isolating RNA from cells.

The Functionally Annotated Genome

As a result of such large-scale studies, most *S. cerevisiae* genes have been assigned a function or role in a particular physiological process. The assignments produced by the combination of systematic approaches range in specificity, from the fine-grained and mechanistic to high-level and diffuse. Nonetheless, the current state of genome annotation vastly exceeds what is

available in other systems (Dolinski and Botstein 2005). An important aspect of this recent history is the availability of an effective annotation database curated by PhD-level scientists (*Saccharomyces* genome database, www.yeastgenome.org). In addition to allowing for rapid analysis, the availability of these systematic resources also allows for comprehensive analysis such that entire pathways can be defined. Importantly, a scientific culture of sharing and dissemination of resources has also accelerated progress.

Can We Systematize the Analysis of Human Fungal Pathogens?

S. cerevisiae is not a human pathogen, and it is known that fungal pathogens harbor mechanisms and pathways required for virulence that either do not exist in model yeasts or have been substantially repurposed during their evolution. Thus, the ability to dissect biology of infection and drug responses in these organisms in an efficient and comprehensive manner requires the development of genetic resources and, equally important, the effective systematic analysis of such resources. However, accomplishing this task in a human fungal pathogen is far more challenging than in *S. cerevisiae* for a variety of technical reasons discussed below. Despite these challenges, substantial beginnings of such efforts have been accomplished in two human fungal pathogens, *C. neoformans* and *C. albicans*. Although random insertional mutagenesis methods are effective tools for identifying mutants, they have disadvantages for the application of systematic approaches. Thus, we will focus on what can be accomplished with ordered collections. For *Cryptococcus* and *Candida*, we will describe the types of technologies and approaches that allowed the challenges to be addressed successfully, and insights gained from the systematic mindset (Table 1).

C. neoformans: SYSTEMATIC GENETICS IN A HAPLOID FUNGUS

C. neoformans Biology, Life Cycle, and Genome

C. neoformans is a basidiomycetous fungus that is an opportunistic human pathogen, infect-

ing, predominantly, individuals with a compromised immune system. *C. neoformans* is the leading cause of fungal meningitis and causes more than 600,000 deaths per year in sub-Saharan Africa alone. *Cryptococcus* is found growing in the environment, typically on decaying wood matter or bird guano (Lin and Heitman 2006). Inhalation of desiccated yeast is likely the major route of infection. *Cryptococcus* species that infect humans belong to four different serotypes (A–D): Serotype A is termed *C. neoformans* var. *grubii*, serotype B and C are called *Cryptococcus gatii*, and serotype D is called *C. neoformans* var. *neoformans*. The different serotypes of *C. neoformans* display difference in phenotypes and genotypes (Lin et al. 2005; Lee et al. 2009). Serotypes A and, to a lesser extent, D account for most clinical isolates from immunocompromised patients, whereas the *C. neoformans* var. *gatii* serotype B/C isolates tend to be isolated from more immunocompetent individuals (Lin and Heitman 2006). Virulence factors of *C. neoformans* include the ability to grow at 37°C and produce melanin and a polysaccharide capsule (Idnurm et al. 2005; Lin and Heitman 2006).

Because of its medical importance, clinical *C. neoformans* isolates belonging to serotype A and D have been sequenced and their genomes have been annotated. Draft sequence of a serotype B isolate is also available (Broad Institute Serotype B Genome Project, www.broadinstitute.org/annotation/genome/cryptococcus_neoformans_b). *C. neoformans* var. *neoformans* has a 19-Mb genome containing 14 chromosomes encoding 6962 predicted proteins (Loftus et al. 2005). The genome of *C. neoformans* var. *grubii* is very similar in size. *C. neoformans* has a complete sexual cycle with two mating types, thus allowing for standard genetic crosses (Hull and Heitman 2002; Lin and Heitman 2006; Idnurm 2010). A current deficit in *C. neoformans* genomics is that there is no publicly funded annotation database akin to the *Saccharomyces* genome database or the *Candida* genome database. However, Broad Institute maintains the sequence database for the H99 serotype A reference strain genome (Broad Institute Serotype A Genome Project, www.broadinstitute.org/annotation/genome/cryptococcus_neoformans).



Table 1. Summary of genomic resources available for *C. neoformans* and *C. albicans* and some of the insights they have provided

Organism	Genomic resource	Screens	Major findings
<i>C. neoformans</i>	1201 gene-deletion library	In vivo infectivity (Liu et al. 2008)	Identified 197 genes with increased or decreased virulence
		Phagocytosis resistance (Liu et al. 2008)	Identified critical role for <i>GAT201</i> in preventing engulfment by macrophages
		Melanin production (Liu et al. 2008)	Identified 38 genes with defective melanin production; many of the genes also had virulence defects
		Growth at 37°C (Liu et al. 2008)	Identified 92 genes important for proliferation at body temperature
		Capsule production (Liu et al. 2007, 2008)	Identified <i>CPL1</i> , <i>GAT201</i> , <i>PBX1</i> , and <i>PBX2</i> as regulating capsule production or assembly
		Growth under hypoxia (Chang et al. 2007; Chun et al. 2007)	SREBP pathway controls growth under hypoxia
<i>C. albicans</i>	~700 homozygous gene-deletion library	In vivo infectivity (Noble et al. 2010)	Identified 115 mutants defective in infectivity; 46 of those had no defects in morphogenesis
		Growth at 37°C (Noble et al. 2010)	Identified 68 mutants defective in proliferation
		Colony morphology (Noble et al. 2010)	Identified 133 mutants defective in morphogenesis; 66 of those had no defect in infectivity
	143 homozygous transcription factor deletion library	Tested 55 conditions (Homann et al. 2009)	Most mutants affect growth under at least one condition
		Surface adherence (Finkel et al. 2012)	Identified the transcription factor <i>SNF5</i> as a key regulator of adhesion
GRACE collection (~1100 conditional mutants)	Growth requirement (Roemer et al. 2003)	Identified 567 essential genes	

See text for details. GRACE, gene replacement and conditional expression; SREBP, sterol regulatory element-binding protein.

Gene-Deletion Methods *C. neoformans*

C. neoformans can be transformed by electroporation or bombardment by DNA-coated particles (biolistic transformation) (Toffaletti et al. 1993). The latter is the mainstay of work in the most common clinical serotype, serotype A (*C. neoformans* var. *grubii*). Early studies showed that substantial homology was required for homologous gene replacement, with replacement frequencies increasing steadily until homology reached ~1 kb on either side of a selectable marker (Davidson et al. 2000; Nelson et al. 2003). Cloning-free fusion PCR methods were

found to be suitable for the addition of long homology arms for gene targeting, which aided in the rapid synthesis of targeting constructs. These advances, together with the development of methods for rapid screening of colonies (which was nontrivial for an encapsulated organism) and optimization of transformation parameters, made it feasible to construct large numbers of gene deletions in *C. neoformans* (Chun and Madhani 2010). In addition, the development of multiple dominant drug selectable marker systems was also critical. This molecular toolbox, together with a complete sexual cycle, make *C. neoformans* a particularly exper-

imentally tractable pathogen, perhaps the most tractable eukaryotic pathogen of humans.

Systematic Genetic Analysis of Virulence in *C. neoformans*

To systematically identify virulence factors in *Cryptococcus*, our laboratory previously constructed a collection of approximately 1200 deletion mutants and screened them in vitro and in vivo for virulence phenotypes (Liu et al. 2008). Each deletion harbored one of 48 DNA oligonucleotide bar codes (Gerik et al. 2005) that enabled quantitative detection of each mutant in a pooled sample. Mutants were pooled in groups of 48 and used to infect mice with an intranasal inhalation model that is thought to mimic the natural route of infection. After allowing for infection to be established, mice were killed, and *C. neoformans* cells were isolated from the lungs. The relative representation of each of the 48 mutants could be assessed by quantitative PCR using 48 bar-code-specific primer pairs on genomic DNA prepared from cells obtained from the mouse versus those derived from a portion of the input pool. Mutants that were significantly reduced in abundance were reconstructed to verify linkage between genotype and phenotype. Importantly, nearly all mutants that were reconstructed maintained their original phenotypes, indicating that unlinked mutations rarely caused phenotypes in the deletion strains. Moreover, all mutants that displayed a defect in a pooled setting also displayed the anticipated impact on virulence in a single infection experiment. Many of the identified mutants were defective in the production of the well-characterized virulence factors: capsule, melanin or growth at 37°C. However, as many as 40 virulence mutants without clear defects in the well-established virulence markers were also isolated (Liu et al. 2008). These are outstanding candidates for factors that participate in novel virulence pathways. These studies validated the use of large-scale signature-tagged mutagenesis screens for *C. neoformans* using comprehensive collections of mutants. Two additional screens of this gene-deletion collection (described below) revealed new aspects of cryptococcal disease.

A Major Role for Capsule-Independent Phagocytosis Inhibition in Cryptococcal Disease

A great advantage of a deletion collection is that the same collection can be screened for any phenotype of interest. One such study asked the question of which genes mediate the ability of *C. neoformans* to resist phagocytosis. *C. neoformans* can be phagocytosed by activated macrophages in vitro (Levitz and DiBenedetto 1989; Liu et al. 2008). The expectation would be that genes required for the synthesis of the *C. neoformans* capsule would be the critical factors. Remarkably, screening of the 1200 deletion library on RAW264.7 macrophage cells revealed that a deletion of the gene coding for the GATA family transcription factor Gat201 had caused an increase in phagocytosis that was far greater than that caused by lack of capsule (Liu et al. 2008). Additional genetic studies showed that this phenotype did not require the major capsular polysaccharide (Liu et al. 2008). In subsequent work, the first whole-genome chromatin immunoprecipitation studies of *C. neoformans* were reported that identified the direct targets of Gat201 (Chun et al. 2011). Systematic deletion of the Gat201 target genes identified two genes, Gat204, encoding another GATA family transcription factor, and Blp1, a member of a family of *C. neoformans*-specific proteins that each harbors a putative amino-terminal signal sequence, which played a key role in mediating phagocytosis inhibition and pathogen fitness during infection. Further evidence indicated that this was independent of the production of polysaccharide capsule, underlining the conclusion that phagocytosis avoidance and virulence in *Cryptococcus* requires mechanisms that have yet to be identified (Chun et al. 2011).

A Role for the SREBP Pathway in Virulence and the Discovery of Novel Virulence Factors

Another independent screen of the deletion collection asked the question of whether *C. neoformans* experienced hypoxia during infection and whether this was important for survival in the host environment. This screen revealed two

pathways required for hypoxic adaptation. Both played a role in *C. neoformans* fitness, implying that the ability to respond to hypoxia is critical for infection (Chun et al. 2007). One of the two pathways encodes a fungal analog of the sterol-sensing SREBP (sterol regulatory element-binding protein) system of mammals. In the mammalian system, SREBP is processed from a membrane-bound precursor in response to sterol depletion in the membrane, yielding an active transcription factor that translocates into the nucleus to activate cholesterol biosynthesis genes. This system had been shown to function as hypoxia sensor in *Schizosaccharomyces pombe* because of the requirement for molecular oxygen in fungal ergosterol biogenesis (Hughes et al. 2005). A similar mechanism was shown to operate in *C. neoformans* (Chun et al. 2007). Directed studies of the SREBP pathway came to the same conclusions (Chang et al. 2007).

Although the polysaccharide capsule has been the subject of genetic screen for decades, the systematic approach identified a number of novel players in capsule production. Two of these genes, *PBX1* and *PBX2*, were initially identified on the basis of their dry colony phenotype, which hinted at a capsule defect (Liu et al. 2007). These proteins encode paired β -helix proteins. Many proteins with this structure play a role in carbohydrate biochemistry. Indeed, the two mutants, although capable of producing capsule monomers, were defective in properly assembling the capsule (Liu et al. 2007). The biochemical function of Pbx1 and Pbx2 is at present unclear, but likely plays a key role in capsule assembly, a process that is poorly understood.

The examples described above show the utility and efficiency of screening a large deletion collection, and emphasize the additional utility of construction and analysis of a comprehensive collection.

Cross-Species Genetic Interaction Mapping: Leveraging *S. cerevisiae* to Study Pathogen Proteins

Because many *C. neoformans* genes that appear relevant for virulence lack assigned biological

function, the genetic tools available for *Saccharomyces* have been used to uncover the function of *C. neoformans* genes (Brown and Madhani 2013). The cross-species genetic interaction approach consists of introducing genes of interest from *Cryptococcus* into the *S. cerevisiae* deletion collection and measuring their genetic interactions, assessed as growth phenotypes, with any of the *S. cerevisiae* gene deletions. Based on the plethora of existing data sets of genetic interactions of the identified *S. cerevisiae* genes, one can predict the function of the tested *C. neoformans* gene. Six *C. neoformans* genes with unknown function, but important for virulence, were initially tested. One of the tested genes, *LIV7*, had a genetic interaction profile similar to that of *S. cerevisiae* mutants defective in Golgi transport, indicating that Liv7 may be involved in vesicle trafficking, a hypothesis that was tested and supported by further evidence (Brown and Madhani 2013). Predictions about the function of the remaining five tested *C. neoformans* genes could also be made (endosome vesicle fusion, proteolysis, cell cycle progression, etc.), and preliminary experiments indicate these predictions to be correct (Brown and Madhani 2013).

Lessons Learned

Given its genetic tractability, *Cryptococcus* represents an obvious test bed for the application of genomics. The concerted investment into methods development by a small number of laboratories was critical in producing the conditions conducive to the construction of the first large-scale genetic resource and its use in studies of pathogenesis. Ongoing efforts are aimed at constructing a whole-genome deletion collection, which we anticipate to have a significant impact on studies of this important human pathogen.

C. albicans: SYSTEMATIC GENETIC ANALYSIS IN A DIPLOID COMMENSAL PATHOGEN WITHOUT A COMPLETE SEXUAL CYCLE

C. albicans Biology, Life Cycle, and Genome

C. albicans, an ascomyceteous fungus, is a commensal of humans found in the gut, skin, and



mucous membranes. A common cause of mucosal infections, *C. albicans* can also cause serious disseminated infection in immunocompromised individuals. According to the Centers for Disease Control, *C. albicans* is the fourth most common cause for hospital acquired bloodstream infections. In response to environmental stimuli, such as pH, hypoxia, and starvation, *C. albicans* can switch from yeast to hyphal, invasive growth (Whiteway and Bachewich 2007). The switch between yeast and hyphal forms is crucial for pathogenesis (Calderone and Fonzi 2001). The potential sexual cycle of *Candida* has been the subject of much interest. Although haploid strains can be selected in the laboratory (Hickman et al. 2013), naturally occurring strains are diploid or close to diploid. *Candida* is thought not to undergo meiosis, a specialized cell cycle during which diploid cells produce haploid gametes, as it lacks many of the key genes that regulate meiosis that are conserved from *Saccharomyces* to humans. Nonetheless, *C. albicans* encodes a mating-type locus encoding DNA-binding regulators of cell type (Hull and Johnson 1999). Derivatives capable of mating can be constructed (Hull et al. 2000). These can be used to derive tetraploids, which then lose chromosomes, presumably via errors in chromosome segregation, to produce diploids (Bennett and Johnson 2005).

The *C. albicans* genome encodes more than 6200 genes, the majority of which (more than 4500 genes) are considered uncharacterized according to the *C. albicans* genome database (www.candidagenome.org). Aneuploidy is common, as is heterozygosity (Selmecki et al. 2006). These factors and the diploid nature of *C. albicans* represent enormous challenges for functional analysis. Nonetheless, steady advances in genetic and genomic tool production have led to the development and use of several large-scale genetic resources.

Gene-Deletion Methods in *C. albicans*: Technical Challenges and Solutions

Because most laboratory *C. albicans* stains are diploid, introducing recessive mutations in the genome requires two successive transformation

steps to modify both copies of the locus of interest (Fonzi and Irwin 1993). Initially, the strategy applied to introduce mutations used *ura3* deficient mutants and replaced the gene of interest with a *URA3* cassette (Fonzi and Irwin 1993). Flanking of the *URA3* gene with repeats allowed for the recycling of the marker using the negative selective agent 5-fluororotic acid. The sequential deletion of the two alleles was challenging for two reasons. First, the marker had to be recycled, and, second, there was no way to select for homologous recombination at the unaltered (vs. deleted) allele. A clever genetic trick was developed by Mitchell and colleagues that used an *ARG4* gene flanked by overlapping fragments of the *URA3* gene (the *UAU1* marker) (Nobile and Mitchell 2009). This construct allows for a double selection, thus, greatly increasing the likelihood for cells lacking both wild-type alleles. This was used to generate the first-ordered homozygous mutant libraries for *C. albicans* and discovery of the first regulators of biofilm formation (see below). Biofilms are thought to be important clinically as a significant number of *Candida* infections are associated with indwelling catheters. This early application of an unbiased approach is what led to the discovery of what we now know to be a complex regulatory circuitry that controls biofilms (Nobile et al. 2012). Although clearly powerful, the *UAU1* method suffers from the complication that it relies on secondary genetic events that are difficult to control and can result in loss-of-heterozygosity for large regions (Nobile and Mitchell 2009).

Furthermore, *URA3* has turned out to not be the ideal marker for studies of *C. albicans*. *ura3* null mutants are defective in both morphogenetic switching and pathogenesis (Lay et al. 1998; Cheng et al. 2003; Brand et al. 2004). Moreover, strains complemented with the wild-type *URA3* gene can suffer from position effects that impact the expression of *URA3* and the associated phenotypes it controls (Brand et al. 2004). These findings raised the possibility that studies investigating the virulence of *C. albicans* strains in the *ura3* genetic background may have missed important aspects of *C. albicans* virulence. These findings and other challenges spurred the development of strains better suited to genetic inves-

tigations of pathogenesis mechanisms. These strains used the *arg4*, *leu2*, and *his1* loci, which have no detectable virulence or commensalism phenotypes (Noble and Johnson 2005). Moreover, new heterologous complementing auxotrophic markers and cloning-free methods for the generation of homologous targeting constructs containing several hundred base pairs of targeting homolog were also developed. Finally, the ability to select for the first deletion allele using one marker, and the second using a distinct marker, was used. Together, these advances allowed for the rapid generation of deletions. These strains and approaches are widely used in the field. Dominant drug selectable markers and conditional promoter-based systems have also been developed, and the *C. albicans* toolbox continues to expand rapidly.

Systematic Analysis of a Homozygous Knockout Collection Reveals Multiple Insights

In the largest such academic effort to date, Noble et al. (2010) generated more than 3000 knockout strains covering nearly 700 genes. Each homozygous deletion was constructed twice independently to enhance the probability that a given genotype and phenotype pair would be linked. The mutants were systematically assayed for growth at 37°C, colony and cell morphology, and fitness in a pooled infection of mice. Prior studies had shown that nearly all mutants defective in virulence that had been described in the field also had defects in the morphology switch from yeast to hyphae (Calderone and Fonzi 2001). However, unbiased screens of the new library showed that many mutants defective in virulence had no detectable defects in morphogenetic switching or growth (Noble et al. 2010). Thus, the first insight from this unbiased approach was that there is evidently much more to *C. albicans* pathogenesis than morphogenetic switching.

Further analysis of mutants defective in infection, but not in growth or morphogenetic switching, led to studies that revealed the importance of glucosylceramide in infection (Noble et al. 2010). The investigators note that related yeasts, like *S. cerevisiae* and *S. pombe*, have lost

the glucosylceramide synthesis genes, indicating that this glycolipid may be directly relevant for survival in the host. Additional studies of a transcription factor required for virulence, but not growth or morphology, led to discovery of a new iron regulatory network whose activation is critical for virulence and repression is important for gut commensalism (Noble et al. 2010). Genetic screens of this library in the context of gastrointestinal tract led to the identification of a specialized cell type required for *C. albicans* commensalism in the gut (Pande et al. 2013). We emphasize that these three insights came directly out of the systematic, unbiased, forward genetic approach enabled by the development of new genetic tools and resources for *C. albicans*.

Transcription Factor Knockout Collection: A Systematic Analysis

In a related study, the 143 deletion mutants of nonessential *C. albicans* transcription factors were constructed (Homann et al. 2009). Initial characterization of the mutants revealed most mutants affected growth under at least one condition. In addition, about a quarter of the mutants affected sensitivity to commonly used antifungal drugs, and a quarter of the mutants affected colony morphology (Homann et al. 2009). Other groups used the transcription factor library to investigate the adherence to surfaces by growing *Candida* cells (biofilm formation) (Finkel et al. 2012). As mentioned above, biofilms are thought to allow *C. albicans* cells to colonize the surfaces of medical hardware and withstand drug treatment and immune responses (Lynch and Robertson 2008). The study identified multiple factors regulating biofilms but the chromatin-remodeling factor Snf5 had one of the most pronounced phenotypes. The investigators also showed that the effects of Snf5 are largely through regulating another transcription factor, Ace2 (Finkel et al. 2012).

Collections Constructed in Industry: Insights and Challenges

To recognize possible drug targets, several large-scale studies have attempted to identify new

essential or haploinsufficient genes in *Candida* via the systematic construction of mutants. A company constructed a library of more than 1100 mutants, which carry a deletion of one copy of a gene and the second copy is placed under a conditional, doxycycline-repressible promoter (Roemer et al. 2003). The methodology used was termed GRACE (gene replacement and conditional expression), and has the benefit of allowing the investigation of essential genes, as one copy can be kept expressed for the purposes of maintaining the strain. Using this resource, the 567 essential genes were identified and prioritized as potential drug targets. One interesting finding was the discordance of essentiality among homologous genes in *C. albicans*, *S. cerevisiae*, and *S. pombe*. Only 61% of the tested *C. albicans* genes that are homologous to *S. cerevisiae* essential genes showed significant reduction in growth. This resource was not made available for academic research until recently. Current use requires negotiation of a materials transfer agreement with Merck Sharp & Dohme (Whitehouse Station, NJ) for academic use. The same is true of a larger heterozygous mutant collection. The latter was successfully used to profile the effects of 35 small molecules and identify their targets (Xu et al. 2007).

Lessons Learned

As an important human fungal pathogen, *C. albicans* remains the focus of research for molecular mycologists worldwide. As a diploid pathogen without a complete sexual cycle, the technical challenges it poses are substantial. Nonetheless, the patient development of methods for systematic gene deletions has shown that the construction of comprehensive resources is feasible, even in an organism without traditional genetics. Moreover, insights obtained to date from unbiased functional approaches indicate that whole-genome resources, if created and made available without restriction, would have an enormous impact. Importantly, an extremely well-curated genome database (www.candida-genome.org) is in place and widely used in the field. Together, with advances in genetic tools,

the field as a whole is poised to make the next move.

CONCLUDING REMARKS

This work makes it evident that the construction of whole-genome deletion collections, as well as other resources (tag libraries, overexpression libraries, etc.), could have a transformative impact on our ability to rapidly and efficiently develop insights into the mechanism of pathogenesis of human fungal pathogens. Although such efforts require substantial investment of time and resources and are not without substantial challenges, there seems little doubt that the impact on the field would be strong and sustained.

ACKNOWLEDGMENTS

We thank Madhani laboratory members for reading and commenting on this manuscript. *C. neoformans* deletion resource development in our laboratory is supported by R01 AI100272.

REFERENCES

- Bahler J, Wu JQ, Longtine MS, Shah NG, McKenzie A 3rd, Steever AB, Wach A, Philippsen P, Pringle JR. 1998. Heterologous modules for efficient and versatile PCR-based gene targeting in *Schizosaccharomyces pombe*. *Yeast* **14**: 943–951.
- Bennett RJ, Johnson AD. 2005. Mating in *Candida albicans* and the search for a sexual cycle. *Annu Rev Microbiol* **59**: 233–255.
- Botstein D, Fink GR. 2011. Yeast: An experimental organism for 21st Century biology. *Genetics* **189**: 695–704.
- Brand A, MacCallum DM, Brown AJ, Gow NA, Odds FC. 2004. Ectopic expression of URA3 can influence the virulence phenotypes and proteome of *Candida albicans* but can be overcome by targeted reintegration of URA3 at the RPS10 locus. *Eukaryot Cell* **3**: 900–909.
- Breslow DK, Cameron DM, Collins SR, Schuldiner M, Stewart-Ornstein J, Newman HW, Braun S, Madhani HD, Krogan NJ, Weissman JS. 2008. A comprehensive strategy enabling high-resolution functional analysis of the yeast genome. *Nat Methods* **5**: 711–718.
- Brown JC, Madhani HD. 2013. Approaching the functional annotation of fungal virulence factors using cross-species genetic interaction profiling. *PLoS Genet* **8**: e1003168.
- Calderone RA, Fonzi WA. 2001. Virulence factors of *Candida albicans*. *Trends Microbiol* **9**: 327–335.
- Chang YC, Bien CM, Lee H, Espenshade PJ, Kwon-Chung KJ. 2007. Sre1p, a regulator of oxygen sensing and sterol



- homeostasis, is required for virulence in *Cryptococcus neoformans*. *Mol Microbiol* **64**: 614–629.
- Cheng S, Nguyen MH, Zhang Z, Jia H, Handfield M, Clancy CJ. 2003. Evaluation of the roles of four *Candida albicans* genes in virulence by using gene disruption strains that express URA3 from the native locus. *Infect Immun* **71**: 6101–6103.
- Chun CD, Madhani HD. 2010. Applying genetics and molecular biology to the study of the human pathogen *Cryptococcus neoformans*. *Methods Enzymol* **470**: 797–831.
- Chun CD, Liu OW, Madhani HD. 2007. A link between virulence and homeostatic responses to hypoxia during infection by the human fungal pathogen *Cryptococcus neoformans*. *PLoS Pathog* **3**: e22.
- Chun CD, Brown JC, Madhani HD. 2011. A major role for capsule-independent phagocytosis-inhibitory mechanisms in mammalian infection by *Cryptococcus neoformans*. *Cell Host Microbe* **9**: 243–251.
- Costanzo M, Baryshnikova A, Bellay J, Kim Y, Spear ED, Sevier CS, Ding H, Koh JL, Toufighi K, Mostafavi S, et al. 2010. The genetic landscape of a cell. *Science* **327**: 425–431.
- Davidson RC, Cruz MC, Sia RA, Allen B, Alspaugh JA, Heitman J. 2000. Gene disruption by biolistic transformation in serotype D strains of *Cryptococcus neoformans*. *Fungal Genet Biol* **29**: 38–48.
- DeRisi JL, Iyer VR, Brown PO. 1997. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**: 680–686.
- Dolinski K, Botstein D. 2005. Changing perspectives in yeast research nearly a decade after the genome sequence. *Genome Res* **15**: 1611–1619.
- Fink GR, Guthrie C. 2002. *Guide to yeast genetics and molecular and cell biology, Part B*. Academic, New York.
- Finkel JS, Xu W, Huang D, Hill EM, Desai JV, Woolford CA, Nett JE, Taff H, Norice CT, Andes DR, et al. 2012. Portrait of *Candida albicans* adherence regulators. *PLoS Pathog* **8**: e1002525.
- Fisher MC, Henk DA, Briggs CJ, Brownstein JS, Madoff LC, McCraw SL, Gurr SJ. 2012. Emerging fungal threats to animal, plant and ecosystem health. *Nature* **484**: 186–194.
- Fonzi WA, Irwin MY. 1993. Isogenic strain construction and gene mapping in *Candida albicans*. *Genetics* **134**: 717–728.
- Gerik KJ, Donlin MJ, Soto CE, Banks AM, Banks IR, Maligie MA, Selitrennikoff CP, Lodge JK. 2005. Cell wall integrity is dependent on the PKC1 signal transduction pathway in *Cryptococcus neoformans*. *Mol Microbiol* **58**: 393–408.
- Ghaemmaghami S, Huh WK, Bower K, Howson RW, Belle A, Dephoure N, O’Shea EK, Weissman JS. 2003. Global analysis of protein expression in yeast. *Nature* **425**: 737–741.
- Giaever G, Chu AM, Ni L, Connelly C, Riles L, Veronneau S, Dow S, Lucau-Danila A, Anderson K, Andre B, et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**: 387–391.
- Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, et al. 1996. Life with 6000 genes. *Science* **274**: 546, 563–547.
- Hickman MA, Zeng G, Forche A, Hirakawa MP, Abbey D, Harrison BD, Wang YM, Su CH, Bennett RJ, Wang Y, et al. 2013. The “obligate diploid” *Candida albicans* forms mating-competent haploids. *Nature* **494**: 55–59.
- Homann OR, Dea J, Noble SM, Johnson AD. 2009. A phenotypic profile of the *Candida albicans* regulatory network. *PLoS Genet* **5**: e1000783.
- Hughes AL, Todd BL, Espenshade PJ. 2005. SREBP pathway responds to sterols and functions as an oxygen sensor in fission yeast. *Cell* **120**: 831–842.
- Hull CM, Heitman J. 2002. Genetics of *Cryptococcus neoformans*. *Annu Rev Genet* **36**: 557–615.
- Hull CM, Johnson AD. 1999. Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans*. *Science* **285**: 1271–1275.
- Hull CM, Raisner RM, Johnson AD. 2000. Evidence for mating of the “asexual” yeast *Candida albicans* in a mammalian host. *Science* **289**: 307–310.
- Idnurm A. 2010. A tetrad analysis of the basidiomycete fungus *Cryptococcus neoformans*. *Genetics* **185**: 153–163.
- Idnurm A, Bahn YS, Nielsen K, Lin X, Fraser JA, Heitman J. 2005. Deciphering the model pathogenic fungus *Cryptococcus neoformans*. *Nat Rev Microbiol* **3**: 753–764.
- Lay J, Henry LK, Clifford J, Koltin Y, Bulawa CE, Becker JM. 1998. Altered expression of selectable marker URA3 in gene-disrupted *Candida albicans* strains complicates interpretation of virulence studies. *Infect Immun* **66**: 5301–5306.
- Lee H, Chang YC, Varma A, Kwon-Chung KJ. 2009. Regulatory diversity of TUP1 in *Cryptococcus neoformans*. *Eukaryot Cell* **8**: 1901–1908.
- Levitz SM, DiBenedetto DJ. 1989. Paradoxical role of capsule in murine bronchoalveolar macrophage-mediated killing of *Cryptococcus neoformans*. *J Immunol* **142**: 659–665.
- Lin X, Heitman J. 2006. The biology of the *Cryptococcus neoformans* species complex. *Annu Rev Microbiol* **60**: 69–105.
- Lin X, Hull CM, Heitman J. 2005. Sexual reproduction between partners of the same mating type in *Cryptococcus neoformans*. *Nature* **434**: 1017–1021.
- Liu OW, Kelly MJ, Chow ED, Madhani HD. 2007. Parallel β -helix proteins required for accurate capsule polysaccharide synthesis and virulence in the yeast *Cryptococcus neoformans*. *Eukaryot Cell* **6**: 630–640.
- Liu OW, Chun CD, Chow ED, Chen C, Madhani HD, Noble SM. 2008. Systematic genetic analysis of virulence in the human fungal pathogen *Cryptococcus neoformans*. *Cell* **135**: 174–188.
- Loftus BJ, Fung E, Roncaglia P, Rowley D, Amedeo P, Bruno D, Vamathevan J, Miranda M, Anderson IJ, Fraser JA, et al. 2005. The genome of the basidiomycetous yeast and human pathogen *Cryptococcus neoformans*. *Science* **307**: 1321–1324.
- Longtine MS, McKenzie A 3rd, Demarini DJ, Shah NG, Wach A, Brachat A, Philippsen P, Pringle JR. 1998. Additional modules for versatile and economical PCR-based gene deletion and modification in *Saccharomyces cerevisiae*. *Yeast* **14**: 953–961.
- Lopez A, Parsons AB, Nislow C, Giaever G, Boone C. 2008. Chemical-genetic approaches for exploring the mode of



- action of natural products. *Prog Drug Res* **66**: v237, 239–271.
- Lynch AS, Robertson GT. 2008. Bacterial and fungal biofilm infections. *Annu Rev Med* **59**: 415–428.
- Mali P, Esvelt KM, Church GM. 2013. Cas9 as a versatile tool for engineering biology. *Nat Methods* **10**: 957–963.
- Martzen MR, McCraith SM, Spinelli SL, Torres FM, Fields S, Grayhack EJ, Phizicky EM. 1999. A biochemical genomics approach for identifying genes by the activity of their products. *Science* **286**: 1153–1155.
- Nelson RT, Pryor BA, Lodge JK. 2003. Sequence length required for homologous recombination in *Cryptococcus neoformans*. *Fungal Genet Biol* **38**: 1–9.
- Nobile CJ, Mitchell AP. 2009. Large-scale gene disruption using the UAU1 cassette. *Methods Mol Biol* **499**: 175–194.
- Nobile CJ, Fox EP, Nett JE, Sorrells TR, Mitrovich QM, Hernday AD, Tuch BB, Andes DR, Johnson AD. 2012. A recently evolved transcriptional network controls biofilm development in *Candida albicans*. *Cell* **148**: 126–138.
- Noble SM, Johnson AD. 2005. Strains and strategies for large-scale gene deletion studies of the diploid human fungal pathogen *Candida albicans*. *Eukaryot Cell* **4**: 298–309.
- Noble SM, French S, Kohn LA, Chen V, Johnson AD. 2010. Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nat Genet* **42**: 590–598.
- O'Brien HE, Parrent JL, Jackson JA, Moncalvo JM, Vilgalys R. 2005. Fungal community analysis by large-scale sequencing of environmental samples. *Appl Environ Microbiol* **71**: 5544–5550.
- Pande K, Chen C, Noble SM. 2013. Passage through the mammalian gut triggers a phenotypic switch that promotes *Candida albicans* commensalism. *Nat Genet* **45**: 1088–1091.
- Richardson MDW, David W. 2003. *Fungal infection: Diagnosis and management*. Blackwell, Oxford.
- Roemer T, Jiang B, Davison J, Ketela T, Veillette K, Breton A, Tandia F, Linteau A, Sillaots S, Marta C, et al. 2003. Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery. *Mol Microbiol* **50**: 167–181.
- Selmecki A, Forche A, Berman J. 2006. Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* **313**: 367–370.
- Toffaletti DL, Rude TH, Johnston SA, Durack DT, Perfect JR. 1993. Gene transfer in *Cryptococcus neoformans* by use of biolistic delivery of DNA. *J Bacteriol* **175**: 1405–1411.
- Tong AH, Boone C. 2006. Synthetic genetic array analysis in *Saccharomyces cerevisiae*. *Methods Mol Biol* **313**: 171–192.
- Whiteway M, Bachewich C. 2007. Morphogenesis in *Candida albicans*. *Annu Rev Microbiol* **61**: 529–553.
- Wuster A, Madan Babu M. 2008. Chemogenomics and biotechnology. *Trends Biotechnol* **26**: 252–258.
- Xu D, Jiang B, Ketela T, Lemieux S, Veillette K, Martel N, Davison J, Sillaots S, Trosok S, Bachewich C, et al. 2007. Genome-wide fitness test and mechanism-of-action studies of inhibitory compounds in *Candida albicans*. *PLoS Pathog* **3**: e92.