Fungal Genome Initiative

A White Paper for Fungal Comparative Genomics

October 10, 2003

Submitted by The Fungal Genome Initiative Steering Committee
Corresponding author: Bruce Birren, Whitehead Institute / MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141 USA
Phone 617-258-0900; E-mail bwb@genome.wi.mit.edu

Introduction to Revised Fungal Genome Initiative White Paper

Our June 2003 Fungal Genome Initiative (FGI) white paper named 44 organisms that were chosen to address a wide range of biological questions. GRASPP's review of that proposal was justifiably critical of our failure to identify specific questions in human health and biology that would be served by sequencing these genomes. In response to these comments we now submit a list of seven fungi whose genome sequence will enable substantial advances in two research areas central to the mission of the NHGRI:

1. Comparative analysis of pathogenesis.

We have selected five fungi that will leverage existing sequence data to provide highly informative comparisons between aggressive human pathogens and nonpathogenic or less aggressive but phylogenetically closely related organisms. These fungi will directly contribute to our understanding of the pathogenicity of Candida, Coccidioides, Histoplasma and Cryptococcus.

2. Annotation of functional sequence elements in a model filamentous fungus.

A high-quality annotation for a reference filamentous fungus is sorely lacking. *Neurospora crassa* is the most well-understood and most widely studied filamentous fungus and the best choice as this reference. The ease of manipulation of Neurospora makes it especially valuable to those studying similar genes and pathways in other fungi, especially highly infectious BL3 organisms such as *Coccidioides immitis* and *Histoplasma capsulatum*. In addition, Neurospora physiology shares many processes with human cells that are lacking in yeast. We have identified two "helper" genomes that are at the appropriate evolutionary distance for identification of conserved functional elements in the sequence of Neurospora.

The sequence of these seven genomes, totaling 180 Mb, will stimulate the study of fungal pathogens and the use of filamentous fungi to study many basic aspects of human biology.

Background and Rationale

Not a single filamentous fungus had been sequenced by the year 2000. Since then, however, the Whitehead Institute/MIT Center for Genome Research (WICGR) has released genome sequences for seven filamentous fungi (www-genome.wi.mit.edu/annotation/fungi/fgi/). Nearly an equal number will have been publicly released by the end of this year from other sources. These sequences have given us a new appreciation for the complexity of the cellular processes these fungal organisms possess (Galagan, Calvo et al. 2003). Further sequencing of new fungal species will continue to reveal the biological diversity generated in the greater than 900 million years of divergence within the kingdom. However, our immediate goal is to better understand a small number of key species through comparative sequence analysis. First among the genomes we have chosen to focus on are the fungal pathogens that have a devastating impact on human health.

Pathogenesis

Fungal pathogens

Fungal diseases are a deadly threat to the growing number of patients receiving technically advanced and expensive treatments, such as surgical implants, organ transplants, and to the growing number of immunocompromised patients resulting from AIDS and cancer chemotherapy. Although the worldwide market for antifungals is estimated to reach $8.7 billion by 2007 (http://www.bccresearch.com), current

antifungals are notable for their inability to act rapidly, as well as their toxicity and limited spectrum. There is a critical need for the development of new antifungal agents. In addition, there is increasing incidence of resistance to the current antifungals such as fluconazole. The identification of fungal-specific genes and virulence factors would provide targets for new, more potent drugs.

Comparative genomic approaches

Comparing genomes from nonpathogenic species to related pathogenic organisms provides a comprehensive approach to identifying sequences that contribute to infection and disease. The value of comparative genomics to the study of pathogenesis is no longer hypothetical but has been validated in a number of studies, noted below. These studies have successfully identified:

- Proteins distinct to pathogenic strains of *E. coli*, some having known roles in pathogenesis (Perna, Plunkett et al. 2001).
- Pathogen-specific targets in *Neisseria meningitides*, where bactericidal antibodies have been developed against conserved surface-exposed antigens (Pizza, Scarlato et al. 2000).
- Novel antigens of meningococcus B that provide candidates for vaccine development. The same approach has been successfully applied to other important human pathogens, demonstrating the feasibility to develop vaccines against any infectious disease(Adu-Bobie, Capecchi et al. 2003).

The comparison of these carefully selected groups of fungal genomes will help identify targets that are selective for the pathogens and, thus, less toxic to the host. This is especially important since fungi, being eukaryotic, are more difficult than bacteria to specifically target without harming the human host. Comparative studies are only the beginning of the work needed; experimental studies must follow. Functional studies based on the genome sequences are already underway for each of these fungi. But the design of these experimental studies will be vastly better when informed by comparative analysis.

We have identified fungi belonging to four key groups of pathogens — Candida, Histoplasma, Coccidioides and Cryptococcus — for which each new genome sequence will allow a novel comparative approach to define the genetic differences in the pathogens and the potential functions in the interactions with our immune system. In each case the new genome to be sequenced has been selected because of a specific difference in its pattern of infection of humans compared with previously sequenced genomes.

Candida

Candida species are the most common human fungal pathogens. Candida cause severe systemic disease in individuals who are immunocompromised, post-surgery, or taking broad-spectrum antibiotics. Candida species are the fourth most common blood culture isolate, accounting for 8% of all bacterial and fungal blood stream infections (Pfaller, Jones et al. 1998). The high prevalence of Candida compared with other fungi is likely because it is generally a benign commensal that resides on mucosal surfaces. Antifungal drugs such as azoles and amphotericin B are effective at suppressing most Candida infections, yet the attributed mortality rate from Candida infections is 38% (Wey, Mori et al. 1988). Not surprisingly, there is a large research effort focused on Candida, with over 200 investigators working on just *C. albicans*.

Although a single species, *C. albicans*, is responsible for about half of the Candida infections, a wide variety Candida species contribute to the remainder. The prevalence of these non-albicans infections is increasing (Hazen 1995). A recent study of candidemia in pediatric cancer patients reported 71% of the cases were non-albicans infections (Mullen, Abd El-Baki et al. 2003). Diagnosis and treatment of Candida infections must therefore be targeted to a range of Candida species. In addition, resistance to

current drugs is emerging. In 1998, approximately 10% of Candida infections were found to be resistant to azole antifungals (Pfaller, Jones et al. 1998). Increased use of antifungals is thought to be further driving the appearance of non-albicans candidiasis.

As is common with fungi, the shared name "Candida" belies the fact that the individual pathogens we are describing show tremendous evolutionary divergence. More than 300 millions years of divergence separate the different Candida described here. Hence, these Candida provide an excellent system to use comparative genomics to define the sequences being selected or lost as traits responsible for pathogenesis, disease association and prevalence over time.

Candida lusitaniae

The comparison of the genome sequence of *Candida lusitaniae* (aka *Clavispora lusitaniae*) to those of *C. albicans* and *C. tropicalis* will allow comprehensive analysis of protein conservation across very distantly related but infectious Candida species. Among the pathogenic Candida species, *C. lusitaniae* is the most distantly related to *C. albicans* (Figure 1). In fact, it has often been placed in phylogenetic trees among the Saccharomyces. The phylogenetic distance of the Candida species was analyzed in a comparison of 18S RNA; the average number of changes per nucleotide (times 100) are 0.32 for *C. albicans* vs *C. tropicalis*, 1.10 for *C. albicans* vs *C. guilliermondii* (see next section) and 3.96 for *C. albicans* vs *C. lusitaniae* (Barns 1991). Comparison of topoisomerase II genes yields similar numbers for distances (Kato 2001). The percent identity across topoisomerase II genes is 83% for *C. albicans* vs *C. tropicalis*; 74% with three alignment gaps for *C. albicans* vs *C. guilliermondii* and 71% with four alignment gaps for *C. albicans* vs *C. lusitaniae*. Hence, comparison of genomes of *C. albicans*, the closely related *C. tropicalis*, and the distantly related *C. lusitaniae*, will provide ideal distances to examine protein conservation among pathogenic Candida. Genes conserved among all pathogenic Candida, but not as highly conserved in humans, could be experimentally evaluated for their contribution to pathogenicity, and would provide ideal targets for antifungal drugs.

The genome sequence of *C. lusitaniae* would also directly benefit research on this species. As a laboratory model, *C. lusitaniae* is experimentally tractable offering transformation, gene disruption by homologous recombination, and congenic a and α strains constructed for genetic studies. In a clinical setting, *C. lusitaniae* poses unique therapeutic challenges as strains resistant to amphotericin B are frequently isolated from patients (Minari, Hachem et al. 2001). Resistance of strains to azole drugs has also been reported.

Candida guilliermondii

Candida guilliermondii is a closely related nonpathogen to *C. albicans* and *C. tropicalis*. Comparison of the genome sequence of *C. guilliermondii* to *C. albicans* and *C. tropicalis* will allow identification of genes conserved among the two pathogens but not in the nonpathogen. These species are evenly spaced by multiple distance measures, and therefore are at an ideal distance to compare the conservation of proteins (see previous section). As *C. lusitaniae* is more distant from *C. albicans* and *C. tropicalis* than *C. guilliermondii*, the comparison of all four genomes will also allow identification of the genes conserved in the pathogenic Candidas but not the nonpathogenic species. This analysis will identify genetic differences correlated with pathogenesis and will inform our understanding of how species become or remain pathogenic through evolution.

An independent assembly of the sequence of *C. guilliermondii* will provide an additional important resource. A significant challenge to the assembly of the Candida genomes is the extent of polymorphism within the diploid sequences. The assembly of the genome sequence of strain SC5314 of *C. albicans* at

Stanford has been impeded by difficulties in joining contigs from the two homologous chromosomes. In the case of Candida, we must sequence the diploids to study the clinically relevant strains. However, we expect to face similar difficulties in assembling the diploid Candida species (*C. albicans*, *C. tropicalis*, and *Loddermyces elongisporus*) approved in our June 2003 white paper. The sequence of *C. guilliermondii*, a haploid very closely related to the three diploids being sequenced (Figure 1), will provide a robust assembly that can be used as a template to assemble the diploid genomes. Aligning consensus sequence from the diploids with the *C. guilliermondii* assembly will allow unambiguous placement of both copies (haplotypes) of the contigs that have not been assembled together.

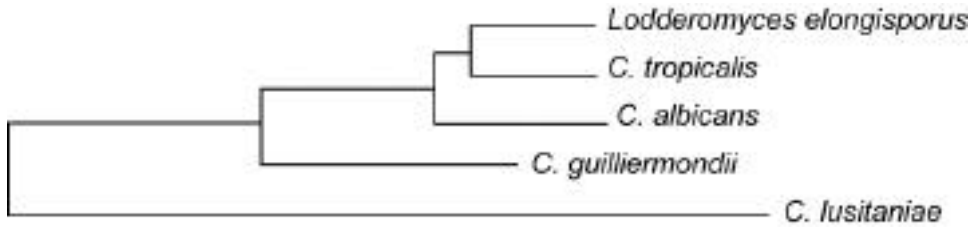

Figure 1. Phylogenetic relationship of Candida species based on the divergence of 26S RNA, after Kurtzman and Robnett (Kurtzman and Robnett 1997). Branch lengths are proportional to the number of nucleotide differences. For reference there were 39 differences in a region of 617 bases when comparing the *C. albicans* and the *C. tropicalis* sequences.

Unicinocarpus reesii

Coccidioides species cause serious and sometimes fatal disease (coccidiomycosis) in otherwise healthy people. The U.S. government regulates both *C. posadasii* and *C. immitis* under the Select Agent Program as potential bioterrorist threats (http://www.cdc.gov/od/sap/). For this reason, Coccidioides is an active target of molecular genetic and clinical investigation. The genome sequence of *C. posadasii* has been released (http://www.TIGR.org) and sequencing of *C. immitis* is underway at WICGR.

Unicinocarpus reesii is morphologically very similar to these Coccidioides species and sequence analysis indicates it is the closest known relative of Coccidioides. The sequence divergence of the 18S ribosomal DNA gene between *C. immitis* and *U. reesii* is approximately 0.7%, reflecting approximately 20–30 million years evolutionary distance (Bowman, White et al. 1996). However, unlike these Coccidioides species, *U. reesii* is nonpathogenic. Having the sequence from these three closely related fungi will allow a comparative approach to the study of pathogenesis in Coccidioides.

Histoplasma capsulatum

Histoplasma capsulatum is the most common cause of fungal respiratory infections (histoplasmosis) in the world. While most infections are mild, 10% of cases result in life-threatening complications such as inflammation of the pericardium and fibrosis of major blood vessels (Durkin, Kohler et al. 2001). Once infected, a latent infection may be reactivated. Histoplasma poses a particular threat to the elderly and to immunocompromised patients (Rachid, Rezende et al. 2003).

Molecular characterization has shown that a number of fungi collectively known as *H. capsulatum* (=*Ajellomyces capsulatus*) actually comprises seven phylogenetic species. There are two discrete phylogenetic species in Northern America: Histoplasma class I (NAm I) and Histoplasma class II (NAm II) (Kasuga, White et al. 2003). Sequence analysis shows that the two Northern American phylogenetic species are as distant from each other as from any other Histoplasma clade, which radiated 3–30 million years ago (Kasuga, White et al. 2003). In addition to these sequence differences, the two North American *H. capsulatum* show a key difference in virulence. While the NAm I strains are chiefly isolated only from AIDS patients, NAm II strains infect otherwise healthy individuals.

H. capsulatum sequencing is underway at Washington University for NAm II. Sequencing the closely related NAm I strain will clarify the genomic basis for the difference in virulence and interactions with the host immune system.

***Cryptococcus neoformans* — serotype B**

Cryptococcus neoformans, a ubiquitous human pathogen, is the leading cause of infectious meningitis. The Cryptococcus species have been diverging over the past 40 million years into three distinct varieties and four serotypes: the predominantly opportunistic pathogens *C. neoformans* var. *neoformans* (serotype D) and *C. neoformans* var. *grubii* (serotype A) and the primary pathogen *C. neoformans* var. *gattii* (serotypes B and C). Ongoing genome projects at TIGR/Stanford and Whitehead/Duke are sequencing three strains [serotype A (strain H99), serotype D (JEC21 and B3501 strains)], and the University of British Colombia has produced ~5X coverage sequence for serotype B, Australian environmental strain WM276.

Serotype B isolates of *C. neoformans* var. *gattii* caused an outbreak on Vancouver Island in Canada that involved more than 50 infections in otherwise healthy human patients, with at least four fatalities (Fyfe, Black et al. 2002). Intervarietal mating tests revealed that the majority of the Vancouver clinical and environmental outbreak isolates were fertile compared with 100% of the environmental isolates from Australia that were sterile (Fraser, Subaran et al. 2003). The outbreak isolate R265 has a VGII molecular genotype by PCR fingerprinting, which differs from the VGI molecular genotype of strain WM276 (J. Kronstad, personal communication). Many mating-type-locus specific PCR reactions that work with a large variety of *gattii* isolates (both serotype B and C) do not work on these outbreak isolates (J. Fraser, personal communication). Thus, these data demonstrate sufficient genetic variations between this outbreak strain and the previously sequenced genome to warrant light coverage whole-genome sequencing.

The comparison of this outbreak strain to its closely related sister strain (WM276), based on comparisons among all three serotypes (40 million years divergence), will reveal recent evolutionary events leading to the change of virulence of this isolate.

Identifying Functional Elements in a Model Filamentous Fungus

Filamentous fungi

The vast majority of fungal pathogens are filamentous fungi. These include many species of Aspergillus, Cryptococcus, Coccidioides, Histoplasma, Fusarium and others. Filamentous fungi exhibit a number of cellular and developmental processes critical to infection that are absent in the laboratory models of budding or fission yeast. These include true hyphal growth, production of secondary metabolites, more diverse signaling pathways and secretion of extracellular enzymes. The great difference between the size and complexity of the genomes of the filamentous fungi compared with yeast highlights why direct study of filamentous fungi is required to fully understand their lifestyle and environmental interactions (Table 1). The genomes of Aspergillus, Neurospora, Magnaporthe and Fusarium are approximately three times as large as those of the sequenced yeasts, while the number of genes they contain is roughly double that of the yeasts.

Table 1. Genome size and gene number for yeast and representative filamentous fungi

Organism	Genome size (Mb)	Est. gene number
Saccharomyces cerevisiae	12	5,500
Schizosaccharomyces pombe	13.8	4,800
Neurospora crassa	38	10,000
Magnaporthe grisea	39	11,000
Fusarium graminearum	36	11,600
Aspergillus nidulans	30	9,500

Neurospora

The many decades of genetic and biochemical research using Neurospora have made it the most well-understood filamentous fungus. Because of its diverse biology, ease of culture, facile genetics and rapid growth rate, Neurospora remains the preeminent model for a filamentous fungus. It has the most identified genes and the densest and most accurate genetic map of any filamentous fungus (Davis 2000; Perkins, Radford et al. 2001; Galagan, Calvo et al. 2003). Currently in the United States there are about 70 research labs focusing on Neurospora. These U.S. labs are funded for more than 70 years of combined support, including over 500 graduate and postdoctoral training positions. The use of Neurospora is growing; the ease of genetic manipulation in Neurospora is inducing researchers working in other fungal systems to take advantage of Neurospora's tools to study the function of conserved genes and pathways. The value of a readily cultured ascomycete model like Neurospora cannot be overstated given the difficulties in experimental manipulation of BL3 organisms like Histoplasma and Coccidioides

In addition to the relevance of Neurospora to understanding fungal pathogens of humans, its study is directly applicable to understanding human biology in ways that studies of the yeasts are not. To take just one example, unlike the model yeasts, Neurospora shares with higher eukaryotes two mechanisms for epigenetic gene silencing. Its DNA is methylated and control of *N. crassa* DNA methylation requires a

histone methyltransferase (Tamaru and Selker 2001). This finding spurred study in higher eukaryotes, where this is also the case (Lehnertz, Ueda et al. 2003 and references therein). Structural analyses of the *N. crassa* histone H3 Lys-9 trimethyltransferase have established how this class of enzymes gains substrate specificity (Zhang, Yang et al. 2003). Also, RNA silencing occurs in *N. crassa* during vegetative growth, and silencing occurs through pathways similar to those found in animals and plants (Makeyev and Bamford 2002; Dykxhoorn, Novina et al. 2003 and references therein).

Why spend further sequencing resources on Neurospora?

To quote from the review of the previous white paper: "The reason for sequencing helper genomes is the characterization of functional units in the model". Despite its importance as a model, the current predicted gene set in Neurospora is far from comprehensive or accurate. Gene boundaries and intron positions remain uncertain and false-positive and -negative gene predictions are difficult to evaluate. Predicting genes in filamentous fungi is considerably more difficult than it is for yeasts because the gene structure is more complex. Introns are more frequent and larger in filamentous fungi than yeast (Galagan, Calvo et al. 2003) and evidence for alternative splicing is growing (C. Kodira and J. Galagan, personal communication). Manual annotation, so valuable for the yeasts (http://www.yeastgenome.org), is hampered in filamentous fungi by a lack of supporting evidence. For example, only 27% of Neurospora genes and 14% of Fusarium genes have supporting EST data.

Producing an accurate and strongly supported gene set for a reference filamentous fungus would provide a critical missing resource. The fact that Neurospora is one of the few filamentous fungi for which the genome sequence is being finished, the size of the user community and the advanced state of the genetic tools argues that Neurospora is the fungus for which this reference gene set would provide the greatest utility.

Finally, identification of noncoding functional elements in Neurospora will be especially valuable given the limited information about the regulation of gene expression in filamentous fungi. Multisequence alignments serve to identify noncoding functional elements, including transcription factor binding sites and other conserved elements.

Why *Podospora anserina* and *Chaetomium globosum*?

P. anserina and *C. globosum* have been selected as appropriate helper sequences based upon their divergence from the reference genome of *N. crassa*. Andrew Miller at the Field Museum recently completed a survey of 68 ascomycetes in which he performed sequence comparisons for a 1200 bp region between conserved motifs 5 and 7 within the *RPB2* gene (using the method in Liu, Whelen et al. 1999).

He found the following levels of sequence similarity between these three taxa (A. Miller, personal Communication):

N. crassa – P. anserina = 68.1% sequence similarity
N. crassa – C. globosum = 68.1% sequence similarity
C. globosum – P. anserina = 68.4% sequence similarity

Sequence similarity between these genomes for the more slowly evolving beta-tubulin gene was found to be 81% (A. Miller, personal communication).

Thus, the sequences of these genomes conform to the guidelines for similarity defined by the comparative studies of yeast (Cliften, Sudarsanam et al. 2003; Kellis, Patterson et al. 2003) and mammals (Boffelli, McAuliffe et al. 2003). That is, these three genomes are similar enough to allow recognition and alignment of orthologous regions, but are sufficiently diverged to allow recognition of functional units through their conservation in multisequence alignments. Currently, the closest sequenced relative of *N. crassa* is *A. nidulans*. *N. crassa* and *A. nidulans* are thought to have diverged over 300 million years ago. As a result, only about half of the predicted Neurospora genes have a strong hit in the *A. nidulans* genome, and these show only approximately 48% sequence similarity (C. Cuomo and J. Galagan, unpublished). This degree of divergence has proven too great to permit identification of conserved noncoding elements in Neurospora, or even a comprehensive annotation of the Neurospora genes.

Additional value of *P. anserina* and *C. globosum* sequence

Podospora is a common model for a number of genetic, cellular and biochemical processes (Debuchy 2003 and references therein) that to date has been the target of only limited sequencing (Silar, Barreau et al. 2003). It has been an important model for understanding senescence in eukaryotic cells (Osiewacz 2002). Changes in the mitochondrial respiratory chain (respiration in Podospora more closely resembles mammals than does yeast) and in copper transport to respiratory enzyme complexes affect *P. anserina* longevity. Changes in mtDNA also affect longevity; this is particularly important in light of recent findings that specific changes in mtDNA sequences are correlated with human longevity and other aspects of human biology (Coskun, Ruiz-Pesini et al. 2003; Zhang, Asin-Cayuela et al. 2003). Another area of research directly relevant to human biology in which *P. anserina* can make an impact is in the study of amyloid formation by prions (Balguerie, Dos Reis et al. 2003; Nazabal, Dos Reis et al. 2003).

C. globosum is the most common clinical isolate of the Chaetomium, which ordinarily infect human skin and nails in healthy individuals. *C. globosum* is a rare, but deadly, systemic pathogen in immunocompromised patients for which there is no effective therapy (Serena, Ortoneda et al. 2003). In addition to its pathogenic capability, it is important to humans because it produces toxic mycotoxins when environmentally present as an infestation in moldy buildings (Nielsen, Gravesen et al. 1999; Pieckova 2003).

Sequencing Strategy and Coverage

Whole-genome shotgun sequence will be obtained as paired reads from both plasmid and Fosmid clones obtained by random shearing of genomic DNA. For each genome, 90–95% of the total sequence coverage will be produced from plasmid subclones (using a combination of 4-kb and 10-kb inserts), and 5–10% of the coverage will be produced from the Fosmid clones. Sequence reads from each genome will be assembled prior to alignment with the related genome sequences.

The coverage required for these organisms varies with the differing intended use of the data. For *C. guilliermondii*, we need an accurate reference assembly of this haploid upon which to build synteny maps of the other diploids being sequenced. *C. lusitaniae* is also haploid, but is sufficiently dissimilar from *C.*

guilliermondii that it requires a completely independent assembly. *P. anserina*, and *C. globosum* require draft assemblies of sufficient depth to use synteny information to ensure alignment of the correct orthologues. For each of these four genomes, 7X coverage will be produced. For *U. reesii*, and the strains of Cryptococcus and Histoplasma, 4X sequence coverage will be produced. The comparison of these genome sequences to the several fully sequenced close relatives does not require as complete representation of the genome.

The desired coverage and number of reads needed for each of these genomes is shown in Table 2. Current average read lengths (630 bp) and pass rates (88%) have been used for these calculations.

Table 2. Genome size, desired coverage and required reads

Organism	Est. genome size (Mb)	Total coverage	Number of reads
C. guilliermondii	16	7X	202,000
C. lusitaniae	16	7X	201,800
U. reesii	30	4X	216,800
C. neoformans, outbreak strain	20	4X	144,000
H. capsulatum NAmI	28	4X	202,400
P. anserina	34	7X	429,000
C. globosum	36	7X	405,000
Total			1,826,000

Community and Resources

The existing connections between the fungal research community and WICGR provide access to the materials and information needed to perform these studies. The genomic DNA of *C. guilliermondii* and *C. lusitaniae* will be provided by Dr. David Soll at the University of Iowa. DNA for *U. reesii* will be supplied by John Taylor at the University of California, Berkeley. The Heitman lab will provide genomic DNA for the Cryptococcus strain R265. A BAC Library has been constructed from this strain at the Vancouver Genome Center and the clones will be end-sequenced and fingerprinted to generate a physical map. *P. anserina* DNA will be provided by R. Debuchy at the Institut de Génétique et Microbiology de l'Université d'Orsay.

Adu-Bobie, J., B. Capecchi, et al. (2003). Two years into reverse vaccinology. *Vaccine* **21**(7–8):605–610.

Balguerie, A., S. Dos Reis, et al. (2003). Domain organization and structure-function relationship of the HET-s prion protein of *Podospora anserina*. *EMBO J.* **22**(9):2071–2081.

Boffelli, D., J. McAuliffe, et al. (2003). Phylogenetic shadowing of primate sequences to find functional regions of the human genome. *Science* **299**(5611):1391–1394.

Bowman, B. H., T. J. White, et al. (1996). Human pathogeneic fungi and their close nonpathogenic relatives. *Mol. Phylogenet. Evol.* **6**(1):89–96.

Cliften, P., P. Sudarsanam, et al. (2003). Finding functional features in *Saccharomyces* genomes by phylogenetic footprinting. *Science* **301**(5629):71–76.

Coskun, P. E., E. Ruiz-Pesini, et al. (2003). Control region mtDNA variants: Longevity, climatic adaptation, and a forensic conundrum. *Proc. Natl. Acad. Sci. USA* **100**(5):2174–2176.

Davis, R. H. (2000). *Neurospora: Contributions of a Model Organism.* Oxford, Great Britain, Oxford University Press.

Debuchy, R. (2003). *Podospora anserina* bibliography no. 10 - Additions. *Fungal Genet. Newsl.* **in press**.

Durkin, M., S. Kohler, et al. (2001).Chronic infection and reactivation in a pulmonary challenge model of histoplasmosis. *J. Infect Dis.* **183**(12):1822–1824.

Dykxhoorn, D. M., C. D. Novina, et al. (2003). Killing the messenger: Short RNAs that silence gene expression. *Nat. Rev. Mol. Cell Biol.* **4**(6):457–467.

Fraser, J. A., R. L. Subaran, et al. (2003). Recapitulation of the sexual cycle of the primary fungal pathogen *Cryptococcus neoformans* var. *gattii*: Implications for an outbreak on Vancouver Island, Canada. *Eukaryotic Cell* **in press**.

Fyfe, M., W. Black, et al. (2002). *5th Int. Conf. Cryptococcus Cryptococcosis.* (1):1.

Galagan, J. E., S. E. Calvo, et al. (2003). The genome sequence of the filamentous fungus *Neurospora crassa*. *Nature* **422**(6934):859–868.

Hazen, K. C. (1995). New and emerging yeast pathogens. *Clin. Microbiol. Rev.* **8**(4):462–478.

Kasuga, T., T. J. White, et al. (2003). Phylogeography of the fungal pathogen *Histoplasma capsulatum*. *Molecular Ecology* **in press**.

Kato, M., M. Ozeki, et al. (2001). Phylogenetic relationship and mode of evolution of yeast DNA topoisomerase II gene in the pathogenic Candida species. *Gene* **272**(1-2):275–281.

Kellis, M., N. Patterson, et al. (2003). Sequencing and comparison of yeast species to identify genes and regulatory elements. *Nature* **423**(6937):241–254.

Kurtzman, C. P. and C. J. Robnett (1997). Identification of clinically important ascomycetous yeasts based on nucleotide divergence in the 5' end of the large-subunit (26S) ribosomal DNA gene. *J. Clin. Microbiol.* **35**(5):1216–1223.

Lehnertz, B., Y. Ueda, et al. (2003). Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr. Biol.* **13**(14):1192–1200.

Liu, Y. J., S. Whelen, et al. (1999). Phylogenetic relationships among ascomycetes: Evidence from an RNA polymerase II subunit. *Mol. Biol. Evol.* **16**(12):1799–1808.

Makeyev, E. V. and D. H. Bamford (2002). Cellular RNA-dependent RNA polymerase involved in posttranscriptional gene silencing has two distinct activity modes. *Mol. Cell* **10**(6):1417–1427.

Minari, A., R. Hachem, et al. (2001). *Candida lusitaniae*: A cause of breakthrough fungemia in cancer patients. *Clin. Infect. Dis.* **32**(2):186–190.

Mullen, C. A., H. Abd El-Baki, et al. (2003). Non-albicans Candida is the most common cause of candidemia in pediatric cancer patients. *Support Care Cancer* **11**(5):321–325.

Nazabal, A., S. Dos Reis, et al. (2003). Conformational transition occurring upon amyloid aggregation of the HET-s prion protein of *Podospora anserina* analyzed by hydrogen/deuterium exchange and mass spectrometry. *Biochemistry* **42**(29):8852–8861.

Nielsen, K. F., S. Gravesen, et al. (1999). Production of mycotoxins on artificially and naturally infested building materials. *Mycopathologia* **145**(1):43–56.

Osiewacz, H. D. (2002). Aging in fungi: Role of mitochondria in *Podospora anserina*. *Mech. Ageing Dev.* **123**(7):755–764.

Perkins, D. D., A. Radford, et al. (2001). *The Neurospora Compendium: Chromosomal loci.*. San Diego, CA, Academic Press.

Perna, N. T., G. Plunkett, 3rd, et al. (2001). Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7. *Nature* **409**(6819):529–533.

Pfaller, M. A., R. N. Jones, et al. (1998). National surveillance of nosocomial blood stream infection due to *Candida albicans*: Frequency of occurrence and antifungal susceptibility in the SCOPE Program. *Diagn. Microbiol. Infect. Dis.* **31**(1):327–332.

Pieckova, E. (2003). In vitro toxicity of indoor Chaetomium Kunze ex Fr. *Ann. Agric. Environ. Med.* **10**(1):9–14.

Pizza, M., V. Scarlato, et al. (2000). Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science* **287**(5459):1816–1820.

Rachid, A., L.S. Rezende et al. (2003). A case study of disseminated histoplasmosis linked to common variable immunodeficiency. *Braz J Infect Dis.* **7**(4):268–272.

Serena, C., M. Ortoneda, et al. (2003). In vitro activities of new antifungal agents against Chaetomium spp. and inoculum standardization. *Antimicrob. Agents Chemother.* **47**(10):3161–3164.

Silar, P., C. Barreau, et al. (2003). Characterization of the genomic organization of the region bordering the centromere of chromosome V of *Podospora anserina* by direct sequencing. *Fungal Genet. Biol.* **39**(3):250–263.

Tamaru, H. and E. U. Selker (2001). A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* **414**(6861):277–283.

Wey, S. B., M. Mori, et al. (1988). Hospital-acquired candidemia. The attributable mortality and excess length of stay. *Arch. Intern. Med.* **148**(12):2642–2645.

Zhang, J., J. Asin-Cayuela, et al. (2003). Strikingly higher frequency in centenarians and twins of mtDNA mutation causing remodeling of replication origin in leukocytes. *Proc. Natl. Acad. Sci. USA* **100**(3):1116–1121.

Zhang, X., Z. Yang, et al. (2003). Structural basis for the product specificity of histone lysine methyltransferases. *Mol. Cell* **12**(1):177–185.