

Genomic approaches to fungal pathogenicity

Michael C Lorenz

Within a few years, the genome sequences of a large number of medically and agriculturally important fungi will be known. With this resource come the promises of genomic approaches to study pathogenicity and host–fungus interactions. Genomics is particularly attractive for these questions, as conventional genetic and biochemical approaches are limited in many pathogenic fungi. Recent work has applied signature-tagged mutagenesis and DNA microarray analysis to virulence studies in several fungal species, and novel approaches, such as protein arrays and genomic deletion libraries, are being developed in *Saccharomyces cerevisiae* and have significant potential in other fungi. High-throughput gene-discovery approaches should greatly increase our understanding of fungal pathogenesis.

Addresses

Whitehead Institute for Biomedical Research, Nine Cambridge Center, Cambridge, Massachusetts 02142, USA; e-mail: lorenz@wi.mit.edu

Current Opinion in Microbiology 2002, 5:372–378

1369-5274/02/\$ – see front matter

© 2002 Elsevier Science Ltd. All rights reserved.

Published online 11 July 2002

Abbreviations

2D-PAGE	two-dimensional polyacrylamide gel electrophoresis
GST	glutathione-S-transferase
ICAT	isotope-coded affinity tag
ORF	open reading frame
PABA	<i>para</i> -aminobenzoic acid
SAGE	serial analysis of gene expression
STM	signature-tagged mutagenesis

Introduction

This is an exciting time for fungal research. The dramatic gap in research tools and resources that has long existed between the model yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe* and the rest of the mycology community is closing rapidly as the genomics revolution encompasses pathogenic fungi. The emerging genome sequence from many species allows the application of genomic technologies to the study of fungal pathogenicity. This is a particularly important development, given the historic difficulties that result from the limited genetic tools in these species — many are asexual, have poor efficiencies of transformation and/or homologous recombination, lack stable extrachromosomal elements, or grow poorly in the laboratory. Genomics promises to bypass some of these challenges and permit rapid identification of virulence genes. The experiences in *S. cerevisiae* demonstrate that genomics is a very powerful tool in the study of a wide variety of areas, including pathogenesis. In this review, I provide an overview of how these approaches are being, or can be, used in virulent species.

Genome sequencing efforts

The *S. cerevisiae* genome project was completed in 1996, making it the first eukaryotic organism to be sequenced

[1]. The *S. pombe* genome was finished recently, as well [2] (see Table 1). The genome sequence of *S. cerevisiae* has been a remarkable advance, both by speeding research using conventional techniques and by the development of novel tools. Fortunately, preliminary genome assemblies for the opportunistic pathogen *Candida albicans* and the mold *Neurospora crassa* have been released (Table 1) and there are ongoing public projects for several *Aspergillus* species, *Magnaporthe grisea*, *Candida glabrata*, *Cryptococcus neoformans*, *Pneumocystis carinii* and other fungal species. A list of fungal genomics resources on the World Wide Web can be found in a recent review [3] and at The Institute for Genome Research (Table 1).

In addition, a comprehensive Fungal Genome Initiative (FGI) has been proposed. Begun by Gerald Fink and colleagues at the Whitehead Institute and developed through community discussions, the FGI emphasizes that fungi represent both a terrible human cost in terms of disease and agricultural damage, and an unparalleled opportunity to study eukaryotic evolution. The FGI proposes to sequence, at a pace of one genome per month, 15 fungi representing a range of medical, agricultural and scientific interests, including both closely related species and widely divergent ones. Additional species will be considered after the initial phase is complete. This proposal, if fully funded, will provide an astonishing resource in less than two years [4••].

Mutagenic techniques

The cornerstone of genetics is, and will always be, the isolation and analysis of mutants. Genomic approaches can, however, make it easier to identify the mutated locus, or to create comprehensive pools of mutants. This is a significant advantage in pathogenesis studies.

Signature-tagged mutagenesis

Mutagenesis via insertion of a DNA fragment has been used in fungal pathogenesis studies for many years. Complete genome sequences greatly increase the power of insertional screens by allowing a single sequencing run to identify the tagged locus. Unfortunately, insertional mutagenesis is impractical for animal studies because each mutant must be tested individually. Signature-tagged mutagenesis (STM), developed by Holden for bacterial pathogens ([5]; reviewed in [6]), incorporates a 40 bp sequence tag into the insertional DNA element. Mutants, each with a unique tag, can be tested in the animal in pools of up to 96 strains. Persistence of individual mutants is detected by hybridizing PCR-amplified tags from organisms recovered from the animal to an array containing the tag sequences. Mutations disadvantageous *in vivo* will be underrepresented in the recovered population. The result is the ability to use the animal model itself as a virulence screen.

Table 1

Web sites of interest.

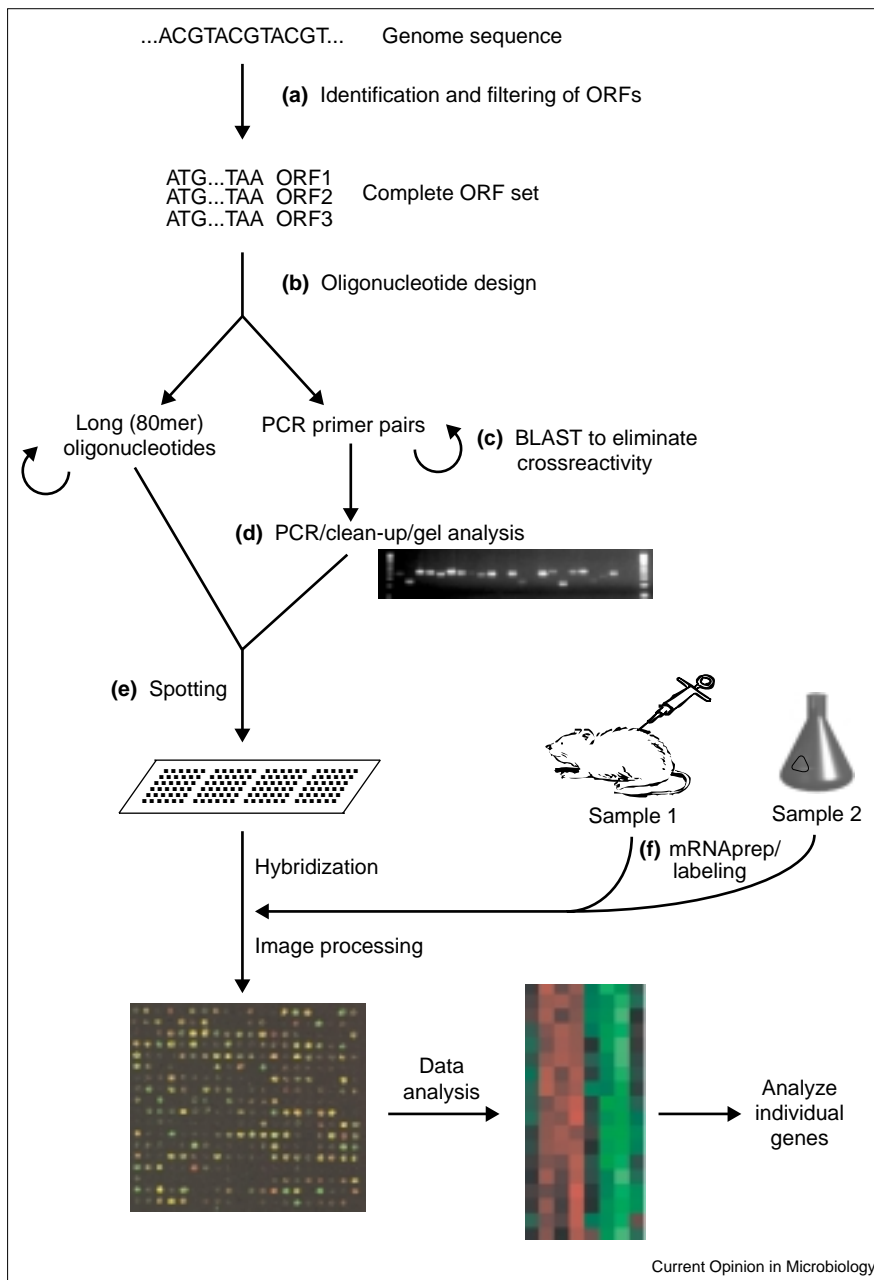
Web site name	Description	URL
Recent genome annotations/genome project information		
<i>Schizosaccharomyces pombe</i> GeneDB Prototype	A searchable annotation database for the complete genome	http://www.genedb.org/genedb/pombe/index.jsp
CandidaDB World Wide Web Server	A searchable partial annotation of the <i>C. albicans</i> genome (assembly 6)	http://genolist.pasteur.fr/CandidaDB
Chip DB	A searchable partial annotation of the <i>C. albicans</i> genome (assemblies 4–6)	http://staffa.wi.mit.edu/
<i>Neurospora crassa</i> Database	A searchable partial annotation (assembly 3)	http://www.genome.wi.mit.edu/annotation/fungi/neurospora/
<i>Saccharomyces</i> Genome Database	The king of databases; recently added <i>C. albicans</i> homology	http://genome-www.stanford.edu/Saccharomyces/
TIGR Microbial Database	Links/status of microbial genome projects at TIGR and elsewhere	http://www.tigr.org/tdb/mdb/mdbinprogress.html
Fungal Genome Initiative	Description and white paper for the FGI	http://www-genome.wi.mit.edu/seq/fgi/
Microarray protocols/software resources		
Microarrays.org	A good source for protocols, suppliers and software	http://microarrays.org
The MGuide	Do-it-yourself guide to microarray equipment	http://cmgm.stanford.edu/pbrown/mguide/index.html
The Brown Lab Protocols	Protocols developed and used in the Brown lab	http://brownlab.stanford.edu/protocols.html
Eisen Lab	Microarray software, including the popular 'TreeView' program	http://rana.lbl.gov/
Microarray Gene Expression Data Group	Developing standards for publication and evaluation of microarrays	http://www.mged.org
DeRisi Lab	Protocols, software and published datasets	http://derisilab.ucsf.edu/
BioConductor	Software and resources for bioinformatics	http://biowww.dfci.harvard.edu/~bioconductor/
Microarray resource list	Links to software for bioinformatics	http://linkage.rockefeller.edu/wli/microarray/soft.html
National Center for Genome Resources	Software and resources for bioinformatics	http://www.ncgr.org/software/
OligoArray	Application to design oligonucleotide-based arrays	http://berry.engin.umich.edu/oligoarray/
DNA probe database for DNA oligo microarray	Similar to OligoArray, has probe sets ready-made for four species	http://ural.wustl.edu/~lif/probe.pl
Primer3	Application to design PCR primers; can be used in bulk processing	http://www-genome.wi.mit.edu/cgi-bin/primer/primer3_www.cgi
Microarray datasets		
Stanford Microarray Database	14 published <i>S. cerevisiae</i> projects, plus many non-fungal projects	http://genome-www5.stanford.edu/MicroArray/SMD/
Yeast Microarray Global Viewer	>40 published datasets from multiple groups	http://www.transcriptome.ens.fr/ymgv/
SGD Expression Connection	11 published datasets from multiple groups	http://genome-www4.stanford.edu/cgi-bin/SGD/expression/expressionConnection.pl
NCBI Gene Expression Omnibus	The NCBI's database of gene expression data; not limited to fungi	http://www.ncbi.nlm.nih.gov/geo/
Fink Lab Genomics	Three published datasets from Fink lab work	http://staffa.wi.mit.edu/fink_public/index.html
Brown Lab Database	Four published datasets from Brown lab work	http://cmgm.stanford.edu/pbrown/explore/
2D-PAGE databases		
Yeast Protein Map	Spot identification of 282 <i>S. cerevisiae</i> proteins	http://www.ibgc.u-bordeaux2.fr/YPM/
Swiss 2DPAGE	Spot identification of 62 <i>S. cerevisiae</i> proteins	http://www.expasy.ch/ch2d/ch2d-top.html
COMPLUYEAST-2DPAGE	15 <i>C. albicans</i> and 37 <i>S. cerevisiae</i> proteins	http://babbage.csc.ucm.es/2d/2d.html
World-2DPAGE	Links to 2D-PAGE databases	http://www.expasy.ch/ch2d/2d-index.html

2D-PAGE, two-dimensional polyacrylamide gel electrophoresis.

Holden applied this system to *Aspergillus fumigatus*, testing 4648 STM strains in a mouse model of invasive aspergillosis. Two insertions were reproducibly defective, one of which was upstream of the *para*-aminobenzoic acid (PABA) synthetase gene (*pabaA*). This mutant had no *in vitro* phenotypes except for a growth defect in the absence of PABA, but was profoundly avirulent unless the diet of the mice was supplemented with PABA [7••]. It is important to note that availability of PABA had not previously been associated with virulence in fungi.

Cormack *et al.* [8] constructed 96 *C. glabrata* strains by integrating a signature tag into the disrupted *URA3* locus, and then used insertional mutagenesis to create the STM pool. In pools of 96, 4800 mutants were screened *in vitro* for adherence to epithelial cells; adherence was increased in five strains, reduced in 10 and eliminated in 16. Fourteen of the non-adherent mutants were integrations into the *EPA1* gene, a member of a large family of fungal cell-surface proteins. Heterologous expression of *EPA1* in *S. cerevisiae* cells allows this non-adherent species to bind to epithelial cells. A complete deletion of *CgEPA1* is

Figure 1



A schematic for microarray construction and application. Publicly available genome sequences for many species may not be completely assembled or annotated. (a) The first step in array design is to determine an ORF set, using criteria such as minimum size, coding probability and homology searches. The difficulty is increased in fungi with frequent introns. (b) Oligonucleotides are designed from this set. The type of oligonucleotide depends on its usage – in an oligonucleotide array (with 60–100mers) or a PCR product array, or to create fusion proteins. (c) The primers are compared to each other, and to the raw sequence to eliminate crossreactivity. (d) In a PCR-based array, products are amplified, purified and analyzed, then (e) spotted onto a solid support. (f) For pathogenesis studies, samples are taken from a relevant condition (for example, from an animal or plant, or in culture exposed to a stimulus of interest). It can be difficult to determine appropriate controls for these experiments, but the same strain grown *in vitro* can be used. Labeled samples are hybridized, processed and analyzed to identify differentially regulated genes.

non-adherent, but still fully virulent, so there are probably additional adhesins redundant with *EPA1 in vivo* [8].

Mutant libraries

A more systematic mutagenesis approach is to construct a library of strains in which every (non-essential) gene is mutated. This also permits a saturating screen with a limited number of strains, an important consideration in virulence studies. A *S. cerevisiae* deletion set, incorporating signature tags, was created through a multilaboratory project [9], and is available from the American Type Culture Collection (ATCC), Research Genetics and EUROSCARF. The creation of similar libraries in other

fungi depends on genome sequence, of course, and is facilitated by efficient PCR-based disruption protocols. It will usually require a community project to share the labor and costs of primer design, amplification, transformation and genotyping, but this resource is well worth the effort.

Nucleotide-based approaches

It has been long appreciated that organisms express genes only in conditions in which they are needed. The induction of a gene in a particular condition can therefore provide clues about its function. Identification of expression changes between populations as a gene discovery tool underlies differential display (reviewed in [10]), which has

been invaluable in many fungal pathogens. Serial analysis of gene expression (SAGE) and microarrays, discussed below, rely on the same principles.

Microarray analysis

Microarray analysis is clearly the technology of choice for assaying expression changes between populations. It can be applied to dozens of conditions and can track thousands of genes simultaneously. The well-known obstacles to microarraying, mostly its cost and complexity, are being reduced as the technology becomes more widespread. Commercial and academic software is easing the task of data analysis, though ascribing biological meaning to array data remains a formidable challenge. A recent paper thoroughly reviews the theories and methods of microarray data analysis [11•]. An effort to create Minimum Information About Microarray Experiments (MIAME) standards to apply to publication of array work will hopefully ease the current problems encountered when comparing data from different laboratories [12].

In a microarray, thousands of DNA spots representing genomic features (usually open reading frames, or ORFs) are arrayed on a solid support, usually a glass microscope slide [13]. The DNA can be oligonucleotides, plasmids, genomic fragments or PCR products. PCR products are most common, but recent advances in technology and tools make long oligonucleotides (60–100mers) an attractive option [14•,15–17]. mRNA populations are enzymatically labeled (usually fluorescently) and hybridized to the array. Induction (or repression) of each gene is determined by the intensity ratio between the two conditions.

Table 1 lists links to some of the hundreds of *S. cerevisiae* datasets, plus additional genomic resources. Although little of this data is directly applicable to pathogenesis, one exception is recent work from our lab, in which we profiled *S. cerevisiae* cells that had been phagocytosed by cultured macrophages. The primary response in yeast isolated from the phagolysosome was induction of the glyoxylate cycle, a biochemical pathway involved in carbon utilization. This response is conserved in *C. albicans* (as determined by Northern analysis) and mutations in isocitrate lyase, an enzyme of this pathway, are less virulent in a mouse model of systemic candidiasis [18•].

To date, three projects have used partial *C. albicans* microarrays. In the first, exposure to the antifungal itraconazole induced expression changes in 296 genes, as detected on a glass slide array from Incyte Genomics containing 6600 ORFs, expressed sequence tags (ESTs) and genomic fragments. The upregulated genes included most of the ergosterol biosynthesis pathway [19], which is the target of the triazole compounds. This commercial microarray has since been withdrawn from the market.

Brown's laboratory, in collaboration with several other groups, constructed a 2000-gene filter-based array and

used it to study morphogenesis by profiling strains with mutations in CaTUP1 and CaNRG1. ScTUP1 is a global transcriptional repressor targeted to specific genes by ScNRG1. In *C. albicans*, mutations in *tup1* and *nrg1* confer constitutive filamentation [20,21,22•,23]. Consistent with this, CaTUP1 and CaNRG1 co-regulate a number of genes known to have a role in the yeast→hypha transition, including *HWPI*, *ECE1*, *ALS3* and *ALS8*. These genes also have promoter elements similar to sequences found to bind NRG1 *in vitro*. CaTUP1 is also targeted to promoters by the CaMIG1 protein, which together co-regulate a set of genes unrelated to morphogenesis [22•,23].

Another filter-based array, containing 700 *C. albicans* genes, was used to study the regulation of filamentous growth by the transcription factors CPH1, CPH2 and EFG1. The *cpb1 efg1* mutant strain is non-filamentous in most *in vitro* conditions and is avirulent in mice [24], but these two factors are generally thought to respond to different inputs. These three factors co-regulate a set of hyphally induced genes, including *HWPI*, *ECE1*, *HYR1* and others, suggesting that these signaling pathways converge on a common set of response genes [25].

C. albicans arrays that are complete (or nearly so) have been developed by several groups, though none of this work has yet been published. Although microarray construction is certainly not simple (a schematic is shown in Figure 1), the multiple independent projects show that the widening accessibility of array technology has made it possible for collaborations, or even single laboratories, to undertake an array project from scratch. This is fortunate, as commercial array sources for species other than *S. cerevisiae* are unlikely in the near future. One group has undertaken a more comprehensive study of filamentation in *C. albicans*, using several mutants and about 20 different conditions, and glass slide arrays with >6000 genes. As well as confirming the complexity of hyphal induction, this work also provides hope that these regulatory networks, which have been associated with virulence [24], can be deciphered (M Whiteway, A Nantel, personal communication). Similar efforts can be expected in *C. albicans* and in other species within the next few years.

Serial analysis of gene expression

In SAGE, cellular mRNA is converted into cDNA, and then cleaved into precise 15 bp segments that are ligated together and sequenced [26]. The frequency of a particular sequence is proportional to the mRNA abundance. This is not as comprehensive as microarray analysis, but it does not require complete genome sequence and has proved to be an excellent method for identifying small ORFs missed during genome annotation, so-called non-annotated ORFs. Again, this has not been reported in fungi other than *S. cerevisiae*, but it is an RNA analysis method in species in which microarrays are not yet practical, and a tool for genome annotation.

Protein-based technologies

There are conflicting reports on whether or not mRNA levels generally correlate with protein abundance, with one report claiming that they do [27] and another disagreeing [28]. Regardless of the global relationship between protein and RNA, it is clear that post-translational regulation, through turnover or modifications, can affect protein abundance and activity. Detection of protein levels, modifications and protein–protein interactions is the goal of proteomics. None of these systematic approaches have been used to any significant degree in pathogenic fungi, so these descriptions will focus on *S. cerevisiae*. Again, the primary barriers to most of these are genome sequence and labor, so they should be possible in other species quickly. Graves presents a comprehensive review of many of these subjects, targeted to the molecular biologist [29•].

Two-hybrid system

First developed by Fields, the two-hybrid system has shown great utility in studying protein–protein interactions [30] and is well known (reviewed in [31]). The molecular biology of the system requires inframe fusions within ORFs, meaning that hundreds of thousands of library clones must be screened even for small fungal genomes. Rather than use random libraries, Uetz systematically cloned >6000 ORFs into both fusion constructs. They pooled the transformants and screened them, finding 692 interacting pairs. In an alternate approach, they arrayed 6000 strains, each expressing a single fusion, in microtiter format and screened against 192 selected ‘bait’ proteins, identifying 281 confirmed pairs. Although more systematic, this approach was noisy — only 20% of initial interactions were confirmed [32•]. Although some optimization is necessary, the genomic approach has obvious advantages.

Protein arrays

Snyder’s group created protein arrays with 5800 yeast proteins, expressed in yeast and purified as glutathione-*S*-transferase (GST) fusions, using a standard microarrayer. Control experiments using an anti-GST antibody found that 93.5% of the spots had a detectable signal. Probing this array with biotinylated calmodulin identified six known and 33 previously unknown calmodulin-binding proteins (six other known calmodulin-binding proteins were not present on the array). This array was also used to identify proteins that bind to liposomes containing various phosphatidylinositol compounds [33••]. A similar array was used to study protein kinase specificity [34]. It is not hard to see how this system could be applied to pathogenicity, for example, to find fungal proteins that bind host factors.

Expression analysis

Two-dimensional polyacrylamide gel electrophoresis (2D-PAGE) has been used to determine differential protein expression (or modifications, such as phosphorylation). 2D-PAGE is limited by the difficulty in ascribing

protein identities to the spots and by poor success with low-abundance proteins. Although microsequencing and mass spectroscopy can be used to identify proteins, fewer than 500 *S. cerevisiae* proteins can be identified on the basis of mobility alone. The use of 2D-PAGE in fungal pathogenesis is just beginning, but in *Mycobacterium*, 2D-PAGE identified a role for the glyoxylate cycle, similar to the *C. albicans* findings described above [35,36].

Gels can be eliminated using isotope-coded affinity tags (ICATs) of slightly different molecular weights to label two protein extracts. Mass spectroscopy can determine the ratio of the labels, hence the relative abundance of each protein. Peaks of interest can be easily sequenced using mass spectroscopy. This technique permits identification of low-abundance proteins (a drawback of 2D-PAGE), but only works in proteins with cysteines on conveniently sized peptides. A test using ICATs to assess proteomic changes in response to carbon source in *S. cerevisiae* was consistent with previous data [37].

Combinatorial approaches

In *S. cerevisiae*, we are now seeing the combination of multiple genomic technologies to develop detailed models of biological behavior. By combining a few directed experiments with publicly available microarray, 2D-PAGE and two-hybrid studies, Ideker *et al.* [38•] reconstructed the galactose-utilization pathway via computer modeling. Young and colleagues [39] combined chromatin immunoprecipitation, which allowed them to find global transcription-factor-binding sites for all nine known cell-cycle regulators, with expression analysis to develop a model of the orderly and continual progression of the cell cycle. Multiple genomic approaches have also been used to study the co-regulation of protein complexes [40]. The application of these data-intense modeling approaches to other fungi is still well into the future, but they have great potential to make predictions about the behavior of biological systems.

Conclusions

Because of the ability of genomics to bypass traditional challenges in fungal pathogens and to provide a comprehensive view of the system, it represents a tremendous opportunity to study fungal virulence. As is clear from the above discussion, few of these technologies have yet been applied to fungal pathogens but, of course, significant genome sequencing is just now appearing for fungal pathogens. *C. albicans* is the most complete of these projects [41]. As this changes, we should see the rapid and productive application of genomics to this field. When used in combination with traditional approaches, genomics promises to revolutionize our understanding of fungal pathogenesis.

Acknowledgements

I am grateful to K Richards, R Wheeler, R Prusty, T Reynolds, G Fink and K Haynes for helping to make this manuscript a reality.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M *et al.*: **Life with 6000 genes.** *Science* 1996, **274**:546-563.
2. Wood V, Gwilliam R, Rajandream MA, Lyne M, Lyne R, Stewart A, Sgouros J, Peat N, Hayles J, Baker S *et al.*: **The genome sequence of *Schizosaccharomyces pombe*.** *Nature* 2002, **415**:871-880.
3. Yoder OC, Turgeon BG: **Fungal genomics and pathogenicity.** *Curr Opin Plant Biol* 2001, **4**:315-321.
4. **Fungal Genome Initiative: A White Paper developed by the Fungal Research Community.** URL: <http://www-genome.wi.mit.edu/seq/fgi>
This white paper outlines the Fungal Genome Initiative (FGI), which was developed through conversations within the fungal research community. It presents the logical arguments for the FGI, including the agricultural and medical importance of fungal diseases. The fungi are also a diverse kingdom and can tell us much about the evolution of eukaryotes.
5. Hensel M, Shea JE, Gleeson C, Jones MD, Dalton E, Holden DW: **Simultaneous identification of bacterial virulence genes by negative selection.** *Science* 1995, **269**:400-403.
6. Meccas J: **Use of signature-tagged mutagenesis in pathogenesis studies.** *Curr Opin Microbiol* 2002, **5**:33-37.
7. Brown JS, Aufauvre-Brown A, Brown J, Jennings JM, Arst H Jr, Holden DW: **Signature-tagged and directed mutagenesis identify PABA synthetase as essential for *Aspergillus fumigatus* pathogenicity.** *Mol Microbiol* 2000, **36**:1371-1380.
These authors apply signature-tagged mutagenesis to *A. fumigatus* via methodology very similar to that used previously on bacteria. A mouse model identified an enzyme, PABA synthetase, which is required for virulence. Amazingly, supplementation of the animal diet with PABA restores virulence in this mutant.
8. Cormack BP, Ghori N, Falkow S: **An adhesin of the yeast pathogen *Candida glabrata* mediating adherence to human epithelial cells.** *Science* 1999, **285**:578-582.
9. Winzeler EA, Shoemaker DD, Astromoff A, Liang H, Anderson K, Andre B, Bangham R, Benito R, Boeke JD, Bussey H *et al.*: **Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis.** *Science* 1999, **285**:901-906.
10. Sturtevant J: **Applications of differential-display reverse transcription-PCR to molecular pathogenesis and medical mycology.** *Clin Microbiol Rev* 2000, **13**:408-427.
11. Quackenbush J: **Computational analysis of microarray data.** *Nat Rev Genet* 2001, **2**:418-427.
This review succinctly explains the computational issues regarding microarray analysis, including the benefits and appropriate usage of normalization techniques, distance metrics, clustering algorithms and other statistical tools. The review is a must-read for biologists getting started with microarray work.
12. Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball CA, Causton HC *et al.*: **Minimum information about a microarray experiment (MIAME)-toward standards for microarray data.** *Nat Genet* 2001, **29**:365-371.
13. Schena M, Shalon D, Davis RW, Brown PO: **Quantitative monitoring of gene expression patterns with a complementary DNA microarray.** *Science* 1995, **270**:467-470.
14. Hughes TR, Mao M, Jones AR, Burchard J, Marton MJ, Shannon KW, Lefkowitz SM, Ziman M, Schelter JM, Meyer MR *et al.*: **Expression profiling using microarrays fabricated by an ink-jet oligonucleotide synthesizer.** *Nat Biotechnol* 2001, **19**:342-347.
This paper describes the development of an 'ink-jet' printing system that synthesizes oligonucleotides directly on a slide surface. They found that 60mers provide a good combination of specificity and sensitivity. The arrays were compatible with existing labeling and amplification protocols. This represents an option for *de novo* array construction with the need for tedious PCR steps.
15. Kane MD, Jatke TA, Stumpf CR, Lu J, Thomas JD, Madore SJ: **Assessment of the sensitivity and specificity of oligonucleotide (50mer) microarrays.** *Nucleic Acids Res* 2000, **28**:4552-4557.
16. Li F, Stormo GD: **Selection of optimal DNA oligos for gene expression arrays.** *Bioinformatics* 2001, **17**:1067-1076.
17. Rouillard JM, Herbert CJ, Zuker M: **OligoArray: genome-scale oligonucleotide design for microarrays.** *Bioinformatics* 2002, **18**:486-487.
18. Lorenz MC, Fink GR: **The glyoxylate cycle is required for fungal virulence.** *Nature* 2001, **412**:83-86.
This paper uses *S. cerevisiae* as a model to understand the interaction between fungal cells and the immune system, particularly macrophages. Transcriptional profiles of phagocytosed yeast identified the primary response as induction of the glyoxylate cycle, a biochemical pathway. A similar response occurs in *C. albicans*, and mutation of this pathway reduces virulence in this organism.
19. De Backer MD, Ilyina T, Ma XJ, Vandoninck S, Luyten WH, Vanden Bossche H: **Genomic profiling of the response of *Candida albicans* to itraconazole treatment using a DNA microarray.** *Antimicrob Agents Chemother* 2001, **45**:1660-1670.
20. Braun BR, Johnson AD: **Control of filament formation in *Candida albicans* by the transcriptional repressor TUP1.** *Science* 1997, **277**:105-109.
21. Braun BR, Kadosh D, Johnson AD: **NRG1, a repressor of filamentous growth in *C. albicans*, is down-regulated during filament induction.** *EMBO J* 2001, **20**:4753-4761.
22. Murad AM, d'Enfert C, Gaillardin C, Tournu H, Tekaija F, Talibi D, Marechal D, Marchais V, Cottin J, Brown AJ: **Transcript profiling in *Candida albicans* reveals new cellular functions for the transcriptional repressors CaTup1, CaMig1 and CaNrg1.** *Mol Microbiol* 2001, **42**:981-993.
One of the first reports of *C. albicans* microarray work, this paper investigates the transcriptional control of the yeast→hypha transition. Mutations in both the DNA-binding proteins TUP1 and NRG1 confer constitutive filamentation, co-regulate a set of hypha-specific genes, and physically interact. Though done with an incomplete array (~2000 genes), this provides a harbinger of future microarray studies in *C. albicans*.
23. Murad AM, Leng P, Straffon M, Wishart J, Macaskill S, MacCallum D, Schnell N, Talibi D, Marechal D, Tekaija F *et al.*: **NRG1 represses yeast-hypha morphogenesis and hypha-specific gene expression in *Candida albicans*.** *EMBO J* 2001, **20**:4742-4752.
24. Lo HJ, Kohler JR, DiDomenico B, Loeberberg D, Cacciapuoti A, Fink GR: **Nonfilamentous *C. albicans* mutants are avirulent.** *Cell* 1997, **90**:939-949.
25. Lane S, Birse C, Zhou S, Matson R, Liu H: **DNA array studies demonstrate convergent regulation of virulence factors by Cph1, Cph2, and Efg1 in *Candida albicans*.** *J Biol Chem* 2001, **276**:48988-48996.
26. Velculescu VE, Zhang L, Zhou W, Vogelstein J, Basrai MA, Bassett DE Jr, Hieter P, Vogelstein B, Kinzler KW: **Characterization of the yeast transcriptome.** *Cell* 1997, **88**:243-251.
27. Fitcher B, Latter GI, Monardo P, McLaughlin CS, Garrels JL: **A sampling of the yeast proteome.** *Mol Cell Biol* 1999, **19**:7357-7368.
28. Gygi SP, Rochon Y, Franz BR, Aebersold R: **Correlation between protein and mRNA abundance in yeast.** *Mol Cell Biol* 1999, **19**:1720-1730.
29. Graves PR, Haystead TA: **Molecular biologist's guide to proteomics.** *Microbiol Mol Biol Rev* 2002, **66**:39-63.
This review describes the value of proteomics and explains the basic technologies, including mass spectroscopy, protein arrays, 2D-PAGE and ICATs. It is written for the molecular biology audience and provides an outstanding introduction to proteomic methodology and its potential uses.
30. Fields S, Song O: **A novel genetic system to detect protein-protein interactions.** *Nature* 1989, **340**:245-246.
31. Drees BL: **Progress and variations in two-hybrid and three-hybrid technologies.** *Curr Opin Chem Biol* 1999, **3**:64-70.
32. Uetz P, Giot L, Cagney G, Mansfield TA, Judson RS, Knight JR, Lockshon D, Narayan V, Srinivasan M, Pochart P *et al.*: **A comprehensive analysis of protein-protein interactions in *Saccharomyces cerevisiae*.** *Nature* 2000, **403**:623-627.
Most (>5000) *S. cerevisiae* ORFs were cloned into two-hybrid fusion vectors. The authors performed screens in two ways, first by pooling transformants expressing these fusions and screening in a traditional manner, albeit with a much smaller library size. Secondly, these transformants were arrayed in microtiter format and screened against 192 selected bait fusions.

The systematic and high-throughput nature opens new opportunities for two-hybrid screens.

33. Zhu H, Bilgin M, Bangham R, Hall D, Casamayor A, Bertone P, Lan N, Jansen R, Bidlingmaier S, Houfek T *et al.*: **Global analysis of protein activities using proteome chips.** *Science* 2001, **293**:2101-2105.
 This paper reports the production of a nearly complete protein array for *S. cerevisiae*. 5800 ORFs were cloned into a standard galactose-inducible vector fused to GST and a hexahistidine tag. Expressed and purified from *S. cerevisiae*, the proteins were spotted to nickel-coated glass slides. This array was tested for binding to calmodulin and several types of liposomes, both labeled biotin. The protein array approach was successful in identifying binding proteins in both cases.
34. Zhu H, Klemic JF, Chang S, Bertone P, Casamayor A, Klemic KG, Smith D, Gerstein M, Reed MA, Snyder M: **Analysis of yeast protein kinases using protein chips.** *Nat Genet* 2000, **26**:283-289.
35. Honer Zu Bentrup K, Miczak A, Swenson DL, Russell DG: **Characterization of activity and expression of isocitrate lyase in *Mycobacterium avium* and *Mycobacterium tuberculosis*.** *J Bacteriol* 1999, **181**:7161-7167.
36. McKinney JD, Honer zu Bentrup K, Munoz-Elias EJ, Miczak A, Chen B, Chan WT, Swenson D, Sacchettini JC, Jacobs WR Jr, Russell DG: **Persistence of *Mycobacterium tuberculosis* in macrophages and mice requires the glyoxylate shunt enzyme isocitrate lyase.** *Nature* 2000, **406**:735-738.
37. Griffin TJ, Gygi SP, Rist B, Aebersold R, Loboda A, Jilkine A, Ens W, Standing KG: **Quantitative proteomic analysis using a MALDI quadrupole time-of-flight mass spectrometer.** *Anal Chem* 2001, **73**:978-986.
38. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, Eng JK, Bumgarner R, Goodlett DR, Aebersold R, Hood L: **Integrated genomic and proteomic analyses of a systematically perturbed metabolic network.** *Science* 2001, **292**:929-934.
 This work inputs a few directed microarray experiments and a large database of previously published protein-protein interactions and microarray datasets into a modeling system that was able to 'predict' the galactose-utilization pathway in yeast. This combination of multiple genomic approaches with tailored experiments and computer simulation is a powerful predictive technology.
39. Simon I, Barnett J, Hannett N, Harbison CT, Rinaldi NJ, Volkert TL, Wyrick JJ, Zeitlinger J, Gifford DK, Jaakkola TS *et al.*: **Serial regulation of transcriptional regulators in the yeast cell cycle.** *Cell* 2001, **106**:697-708.
40. Jansen R, Greenbaum D, Gerstein M: **Relating whole-genome expression data with protein-protein interactions.** *Genome Res* 2002, **12**:37-46.
41. Tzung KW, Williams RM, Scherer S, Federspiel N, Jones T, Hansen N, Bivolarevic V, Huizar L, Komp C, Surzycki R *et al.*: **Genomic evidence for a complete sexual cycle in *Candida albicans*.** *Proc Natl Acad Sci USA* 2001, **98**:3249-3253.