Fungal genomics and pathogenicity
OC Yoder* and B Gillian Turgeon†

The filamentous fungal genetics community has enthusiastically embraced the utilization of genomics technologies to resolve long-standing issues in fungal biology. For example, such technologies have been proposed to study the mechanics of tip growth, photoreception, gene silencing, the molecular basis of conidiation, the pathway leading to sexual reproduction, and mechanisms of pathogenesis. These studies have provided a refreshing change of pace in research on filamentous fungi, which has lagged behind that on other eukaryotes in the exploitation of genome-wide methodologies. Despite the late start, several fungal genome sequencing projects are underway. The resulting databases will allow the comprehensive analysis of developmental processes that are characteristic of fungi, including the molecular nature of pathogenicity. DNA databases underpin analyses of the fungal transcriptome, proteome, and metabolome. This combined information will contribute to our basic understanding of not only the mechanics of infection but also the evolution of pathogenicity.

Addresses
*†Torrey Mesa Research Institute, 3115 Merryfield Row, La Jolla, California 92121, USA
*e-mail: olen.yoder@syngenta.com
†e-mail: gillian.turgeon@syngenta.com
†Department of Plant Pathology, Cornell University, Ithaca, New York 14853, USA; e-mail: bgt1@cornell.edu


Introduction
Pathogenic fungi infect their hosts by multiple developmental processes requiring the activities of fungal molecules that either remain inside the fungus, are positioned on the fungal cell surface, or are secreted to allow direct physical interaction with receptors in the host. A few fungal molecules have already been proven critical to the outcome of infection. In general, these molecules can be classified as penetration effectors (e.g. melanin, glycerol and hydrophobins), toxins that poison host defenses or elicitors that induce them, enzymes that degrade host defenses, transporters that protect the fungus from host defenses, or components of signal transduction pathways that are required for fungal sensing of the host environment. It is not surprising that these classes of molecule were among the first discovered as fungal virulence factors. Toxins and enzymes often have readily assayable biochemical activities. The known functions of transporters and signal-transduction components suggest their involvement in fungal responses to the plant host. The group of known fungal virulence factors is, however, small and clearly incomplete [1–5] given the apparent complexity of the pathogenic lifestyle. Thus, we can anticipate that a systems approach to experimental investigation will reveal much about the molecular basis of fungal pathogenesis.

In the past, the most productive method to search for fungal virulence factors was forward genetics. Prior to the development of insertional-mutagenesis techniques, such as restriction-enzyme-mediated integration (REMI) [6–9] and the use of transposons [10–12], it was difficult to induce genetic variation in many filamentous fungi and even harder to clone genes that conferred pathological phenotypes. Gene-tagging methods eased the task but still permitted only ‘one-gene-at-a-time’ analyses. Genomics technologies [13**,14*,15–19] enable the identification of complete gene sets that control pathways such as those leading to pathogenicity. This review discusses factors responsible for the present state of fungal genomics, information recently gleaned from the first comparisons of fungal genomes, and approaches for the application of genomics technologies to the problem of fungal aggressiveness.

Current status of filamentous fungal genomics
Among fungal researchers, only the yeast (Saccharomyces cerevisiae) community has been in a position to develop, refine and exploit genomics technologies. Powerful genome-wide data-gathering methods have been used with yeast. These include genome sequencing; expression profiling using microarrays or serial analysis of gene expression (SAGE); protein identification by two-dimensional gels or chromatography coupled to tandem mass spectroscopy; and functional analysis using transposon insertions, targeted gene deletions, protein tagging or two-hybrid interactions [13**]. The yeast community also quickly recognized that the value of data collected using these technologies is severely limited without the construction of curated databases (e.g. the Yeast Protein Database [YPD] and the Stanford Genome Database [SGD]; Table 1).

Communities of researchers working with the filamentous fungi have been less well organized, and the fungi themselves have relatively undeveloped genetic systems compared to yeast. Despite that, genome-wide analyses of two genetic models, Aspergillus nidulans and Neurospora crassa, have been underway for several years [20–23]. The pace of fungal genome research has quickened recently with the completion of the N. crassa genome sequence (Table 1); another harbinger is the March 2001 issue of Genetics, which features a section on ‘Fungal Genomics Investigations’ that includes 12 original research reports [24]. Yet, by comparison with total efforts in genomics, those expended on the fungi pale. For example, in a collection of 379 genome-project websites
### Table 1

<table>
<thead>
<tr>
<th>Fungus</th>
<th>Center</th>
<th>URL(s)</th>
<th>Strategy</th>
<th>Status</th>
<th>February 2001</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspergillus flavus</td>
<td>OU</td>
<td><a href="http://www.genome.ou.edu/fungal.html">http://www.genome.ou.edu/fungal.html</a></td>
<td>Cosmids, cDNA</td>
<td>~1250 ESTs</td>
<td></td>
</tr>
<tr>
<td>Aspergillus fumigatus</td>
<td>NIAID</td>
<td><a href="http://www.sanger.ac.uk/projects/A_fumigatus/">http://www.sanger.ac.uk/projects/A_fumigatus/</a></td>
<td>10 BACs</td>
<td>Library construction</td>
<td></td>
</tr>
<tr>
<td>Aspergillus fumigatus</td>
<td>TSC</td>
<td><a href="http://www.niaid.nih.gov/dmid/genomes/dharma">http://www.niaid.nih.gov/dmid/genomes/dharma</a> <a href="http://www.aspergillus.man.ac.uk">http://www.aspergillus.man.ac.uk</a></td>
<td>EST databases, Chromosome IV</td>
<td>~50,000 unique ESTs</td>
<td>~71X; 1503 contigs</td>
</tr>
<tr>
<td>Aspergillus nidulans</td>
<td>OU</td>
<td><a href="http://www.genome.ou.edu/fungal.html">http://www.genome.ou.edu/fungal.html</a></td>
<td>Cosmids, cDNA mixed vegetative and 24 h sexual development</td>
<td>~15,000 ESTs</td>
<td></td>
</tr>
<tr>
<td>Aspergillus oryzae</td>
<td>AIST</td>
<td><a href="http://www.aist.go.jp/RIODB/ffdb/index.html">http://www.aist.go.jp/RIODB/ffdb/index.html</a></td>
<td>cDNA</td>
<td>23,000 ESTs</td>
<td></td>
</tr>
<tr>
<td>Cryptococcus neoformans</td>
<td>SGTC</td>
<td><a href="http://www-sequence.stanford.edu/group/C.neoformans/">http://www-sequence.stanford.edu/group/C.neoformans/</a></td>
<td>5 shotgun</td>
<td>3X</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NUSM</td>
<td><a href="mailto:s-kohno@net.nagasaki-u.ac.jp">s-kohno@net.nagasaki-u.ac.jp</a></td>
<td>cDNA</td>
<td>~4000 ESTs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OU</td>
<td><a href="http://www.genome.ou.edu/fungal.html">http://www.genome.ou.edu/fungal.html</a></td>
<td>SAGE</td>
<td>42,000 unique SAGE tags</td>
<td></td>
</tr>
<tr>
<td></td>
<td>GSC</td>
<td><a href="mailto:kronstad@interchange.ubc.ca">kronstad@interchange.ubc.ca</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fusarium sporotrichioides</td>
<td>OU</td>
<td><a href="http://www.genome.ou.edu/fsporo.html">http://www.genome.ou.edu/fsporo.html</a></td>
<td>cDNA</td>
<td>~7500 ESTs</td>
<td></td>
</tr>
<tr>
<td>Magnaporthe grisea</td>
<td>CUGI</td>
<td><a href="http://www.genome.clemson.edu/projects/rice_blast/">http://www.genome.clemson.edu/projects/rice_blast/</a></td>
<td>BAC ends, ESTs</td>
<td>~17000 BAC ends</td>
<td>~5000 ESTs, 188 BAC contigs, 3 BACs complete</td>
</tr>
<tr>
<td></td>
<td>FGL</td>
<td><a href="http://www.fungalgenomics.ncsu.edu/">http://www.fungalgenomics.ncsu.edu/</a></td>
<td>Chromosome 7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Neurospora crassa</td>
<td>WICGR</td>
<td><a href="http://www-genome.wi.mit.edu/annotation/fungi/neurospora/">http://www-genome.wi.mit.edu/annotation/fungi/neurospora/</a></td>
<td>Shotgun, 4 kb</td>
<td>&gt;10X, ~1700 contigs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NGP</td>
<td><a href="http://www.unm.edu/~ngp/">http://www.unm.edu/~ngp/</a></td>
<td>cDNA, Chromosome II, V complete</td>
<td>~3500 ESTs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>MIPS</td>
<td><a href="http://www.mips.biochem.mpg.de/proj/neurospora/">http://www.mips.biochem.mpg.de/proj/neurospora/</a></td>
<td>Cosmids, cDNA</td>
<td>~3000 ESTs</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OU</td>
<td><a href="http://www.genome.ou.edu/fungal.html">http://www.genome.ou.edu/fungal.html</a></td>
<td>Fruiting bodies Evening Morning</td>
<td>~15,000 ESTs</td>
<td>~9100 ESTs, ~11,000 ESTs</td>
</tr>
<tr>
<td>Candida albicans</td>
<td>TSC</td>
<td><a href="http://www.sanger.ac.uk/projects/C_albicans/">http://www.sanger.ac.uk/projects/C_albicans/</a></td>
<td>10 cosmids</td>
<td>6 finished</td>
<td></td>
</tr>
<tr>
<td></td>
<td>SGTC</td>
<td><a href="http://www-sequence.stanford.edu/group/C.albicans/">http://www-sequence.stanford.edu/group/C.albicans/</a></td>
<td>Shotgun</td>
<td>~4X</td>
<td></td>
</tr>
<tr>
<td>Pneumocystis carinii</td>
<td>TSC</td>
<td><a href="http://www.sanger.ac.uk/projects/P_carinii/">http://www.sanger.ac.uk/projects/P_carinii/</a></td>
<td>10 telomeric cosmids</td>
<td>1 finished</td>
<td></td>
</tr>
<tr>
<td></td>
<td>UK</td>
<td><a href="http://biology.uky.edu/fc/">http://biology.uky.edu/fc/</a></td>
<td>cDNA</td>
<td>~4000 ESTs</td>
<td></td>
</tr>
<tr>
<td>Saccharomyces cerevisiae</td>
<td>SGD, WU</td>
<td><a href="http://wwwanford.edu/Saccharomyces/">http://wwwanford.edu/Saccharomyces/</a></td>
<td>Cosmid and lambda clones</td>
<td>Complete</td>
<td></td>
</tr>
<tr>
<td></td>
<td>TSC</td>
<td><a href="http://www.sanger.ac.uk/projects/S_cerevisiae/">http://www.sanger.ac.uk/projects/S_cerevisiae/</a></td>
<td>Cosmids</td>
<td>Almost complete</td>
<td></td>
</tr>
<tr>
<td>Schizosaccharomyces pombe</td>
<td>OU</td>
<td><a href="http://www.genome.ou.edu/fungal.html">http://www.genome.ou.edu/fungal.html</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>TSC</td>
<td><a href="http://www.sanger.ac.uk/projects/S_pombe/">http://www.sanger.ac.uk/projects/S_pombe/</a></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Phytophthora infestans</td>
<td>NCGR</td>
<td><a href="http://www.ncgr.org/research/pgi/index.html">http://www.ncgr.org/research/pgi/index.html</a></td>
<td>cDNA</td>
<td>3000 unigenes</td>
<td></td>
</tr>
</tbody>
</table>

AIST, National Institute of Advanced Industrial Science and Technology; CalpD, Candida albicans Protein Database; CUGI, Clemson University Genomics Institute; FGL, Fungal Genomics Laboratory, NC State University; GSC, Genome Sequence Centre, B.C. Cancer Agency; MIPS, Munich Information Center for Protein Sequences; MycoPathPD, Human Pathogenic Fungi Protein Database; NCGR, National Center for Genomics Research; NGP, Neurospora Genome Project University of New Mexico; NIAID, National Institute of Allergy and Infectious Diseases; NUSC, Nagasaki University School of Medicine; OSU, Oklahoma State University; OU, Oklahoma University; PompeBED, Schizophyllum commune Protein Database; SGD, Stanford Genome Database; SGTC, Stanford Genome Technology Center; TSC, The Sanger Centre; UK, University of Kentucky; WICGR, Whitehead Institute; WU, Washington University; YPD, Yeast Protein Database.
55 projects are complete and 324 are in progress. Only 16 of these websites (i.e. 4% of the total) involve fungi, and only 12 species of fungi are represented, of which seven are pathogens of either plants or humans. Details of the publicly accessible fungal and oomycete (*Phytophthora*) websites are shown in Table 1. The list includes projects on plant pathogens, human pathogens [25,26], and saprophytes. The genome sequencing of *N. crassa* is near completion at approximately 10X shotgun coverage and 1700 contigs (Table 1). Although the majority of fungal genome projects are publicly supported, some are funded by ‘for-profit’ organizations, which can constrain public access. There are cases, however, in which private fungal genome data have been made available to the public. A case in point is a partial sequence of the *A. nidulans* genome that is accessible (subject to certain restrictions) on the Cereon/Monsanto website (URL www.cereon.com).

A substantial impediment to the sequencing of fungal genomes, especially those of pathogens, has been lack of adequate financial support in the public sector. Traditional research grants usually are too small to defray the costs of major genomics efforts, especially of whole-genome sequencing, and government funding agencies have needed time to adjust to this reality. The sequencing of the *N. crassa* genome, which was supported by the National Science Foundation, is a good example of how partnership among academic groups, a research institute, and government can work. Although a consortium of academic researchers developed the proposal for the *N. crassa* project, the contract for sequencing was awarded to the Whitehead Institute, which has both the expertise and capacity to do the job efficiently. The project was completed, under budget, in a few weeks, and is now publicly available on the Whitehead website (Table 1). We propose that the *N. crassa* model be followed for the public support of genomics projects with other fungi. Indeed, the Whitehead has suggested a filamentous fungal genome initiative that would aim to sequence a genome per week for as long as funding is available (announced at the XXI Fungal Genetics Conference, Asilomar, March 2001).

**Genome comparisons**

In Figure 1a, the genome sizes of selected filamentous fungi are compared with those of various model organisms. Note that the filamentous fungal genomes are about three times the size of the yeast genome (except *Ashbya*, which...
is smaller than yeast), one-third the size of the *Arabidopsis thaliana* genome, and an order of magnitude smaller than the genome of rice, the model cereal. The small size of many fungal genomes compared to those of most eukaryotes, combined with their high gene density and relatively low amount of repetitive DNA, make the sequencing of multiple fungal genomes both feasible and cost effective.

Even though the number of fungal genome sequences available for comparison is small, and most of these are incomplete, certain trends can be discerned. For example, there are major differences in G+C (guanine plus cytosine) content, which ranges from about 40% in yeast to 56% in *N. crassa*, with values for other fungi scattered in between (Figure 1b). The functional significance of these differences is unknown, but may reflect fungal lifestyle or ecological niche. At the very least, G+C content can serve as a signature to assist gene-finding efforts and to signal the presence of horizontally transferred genes.

Another recent observation is that synteny may exist among fungi. Synteny has been documented between the hemiascomycetes *S. cerevisiae* and *Ashbya gossypii*, between the pyrenomycetes *N. crassa* and *Magnaporthe grisea*, and between the pyrenomycete *M. grisea* and the plectomycete *A. nidulans*. Synteny has not been observed between pyrenomycetes and hemiascomycetes (on the basis of reports by R Dean, F Dietrich, L Hamer, T Mitchell, XXI Asilomar Fungal Genetics Conference, 2001). Comparison of the genomes of the yeasts *Candida albicans* and *S. cerevisiae* has shown that less than 10% of gene adjacencies have been conserved; this lack of microsynteny appears to be caused by a combination of small inversions and large chromosomal rearrangements [27•]. These results suggest that synteny needs to be assessed at both the macro and the micro levels; that is, the positions of gene sets may be conserved but the order of genes within any given set may be scrambled. If synteny, either local or long-range, is widespread among fungi, disruptions of synteny may indicate genomic rearrangements that cause key differences between saprophytes and pathogens, or among pathogens of different hosts or different tissues.

A possible emerging trend is that pathogens may carry more genes dedicated to secondary metabolism than do saprophytes. Preliminary documentation for this can be found in Table 2. A subset of protein families, chosen because one or more members of each family is a proven fungal virulence factor (see footnotes to Table 2) is listed. In certain protein families, major differences exist between fungi that cause serious plant diseases and those that do not. Note that three pathogenic fungi are rich in nonribosomal peptide synthetases (NRPSs) and polyketide synthases (PKSs), whereas the genomes of saprophytes encode few or none of these proteins. In other protein families, no differences exist among these fungi (Table 2).
Comparison remains to be seen. If the frequency of orphans is high in fungi generally, the initial assignment of gene function in fungi will require broad, rather than focused, phenotypic screening. Yeast and N. crassa share a very small set of genes that have no homologs in nonfungal organisms, that is, there appear to be fungus-specific genes. Conversely, there is a second small set of N. crassa genes that have homologs in nonfungal databases but not in yeast; these genes may reflect gene loss from yeast or horizontal transfer into N. crassa, although there is little evidence for the latter.

Genome comparisons can be used effectively to address specific questions about fungal lifestyle, such as choice of reproductive strategy. C. albicans, for example, has historically been thought to lack a sexual cycle. Recently, C. albicans was shown to have homologs of the S. cerevisiae mating-type genes and to be capable of initiating, but not completing, sexual development if strains of opposite mating type were constructed and paired [31]. A 10.4X shotgun sequence coverage of the C. albicans genome provided a database that could be searched for the 500-odd genes known to be required for S. cerevisiae mating [32•]. A 10.4X shotgun sequence coverage of the C. albicans genome provided a database that could be searched for the 500-odd genes known to be required for S. cerevisiae mating [32•]. All of the genes involved in the pheromone response pathway were found but some of the genes required for meiosis were missing, which may explain the natural sterility of C. albicans.

**Fungal pathogenicity gene set**

Global genome queries allow issues regarding host–fungal interactions to be addressed with a realistic expectation of resolution. The fundamental question is, ‘Why are certain fungi pathogenic whereas most of them are obligate saprophytes?’ Secondly, there are other issues: why are virtually all pathogenic fungi host-specific (no fungus is pathogenic to all plants or all animals)? On a given host, why do most fungal pathogens display tissue specificity? Underlying these questions is the history of pathogenicity gene flow: what are the genetic mechanisms responsible for fungal acquisition of the capacity for aggressive behavior? Are genes for pathogenicity transmitted by vertical inheritance or by horizontal processes, or both? Horizontal mechanisms have been documented for human pathogenic bacteria [33–37], and there is evidence for the same phenomenon in plant pathogenic fungi [38–41]. Do pathogenic fungi possess genes with functions dedicated to infectious activities or are genes for normal metabolic processes co-opted for these purposes? Evidence obtained to date indicates that both modes of evolution operate in fungi. Although a variety of unique pathogenesis factors have been proven [1–3,5,42], the so-called host-specific toxins are among the best examples. Each of these molecules is produced by only one genotype of one fungal species, and is required for pathogenesis by that genotype [43–45]. The alternative situation, in which a common metabolite is redirected to pathogenic functions, is illustrated by the roles of glycerol and the pigment melanin, both of which are required for host penetration by the rice blast fungus M. grisea [46–49].

The foregoing questions can be addressed on the broadest level by whole-genome comparisons, as illustrated in Figure 2. If the genome sequence of a pathogen (e.g., Cochliobolus) were to be compared to that of an obligate saprophyte (e.g., Neurospora), we might expect that a portion of the total sequences would be common to both fungi, with the remainder divided into two classes, one unique to each fungus (Figure 2). Among the sequences unique to the pathogen, we would expect to find genes for pathogenicity. Similarly, if a comparison were made between the genome sequences of different pathogens, each specific to a particular host, many would be common, but among those unique to each fungus would be genes that determine host specificity. A similar strategy could be used to identify genes as candidates for roles in determining fungal specificity to host tissues (Figure 2). Comparison of phylogenetically distant fungi is desirable in the short term to survey the spectrum of genes that occur in fungi. Such a comparison will, however, identify a pool of ‘common’ genes that is predicted to be large, which would make it difficult to recognize the subset of genes involved in pathogenicity. One way to approach this problem is to sequence the genomes of ever more closely related pathogens to narrow the choice of candidates that might determine host or tissue specificity. An alternative approach might be to probe a DNA oligo array of one fungus with genomic DNA fragments of the same fungus and

### Table 2

<table>
<thead>
<tr>
<th>Protein Family</th>
<th>Cochliobolus</th>
<th>Fusarium</th>
<th>Botrytis</th>
<th>Neurospora</th>
<th>Ashbya</th>
<th>Saccharomyces</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptide synthetases</td>
<td>30</td>
<td>37</td>
<td>29</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Polyketide synthases</td>
<td>40</td>
<td>35</td>
<td>42</td>
<td>7</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>ABC transporters</td>
<td>51</td>
<td>54</td>
<td>46</td>
<td>39</td>
<td>17</td>
<td>29</td>
</tr>
<tr>
<td>Cytochrome P450s</td>
<td>63</td>
<td>40</td>
<td>33</td>
<td>44</td>
<td>ND</td>
<td>4</td>
</tr>
<tr>
<td>Protein kinases</td>
<td>112</td>
<td>94</td>
<td>70</td>
<td>120</td>
<td>ND</td>
<td>117</td>
</tr>
</tbody>
</table>

These protein families, compared here for six fungi, were chosen because each has at least one member known to be involved in fungal pathogenesis (the numbers for the filamentous fungi are estimates; they are used to illustrate whether zero, a few, or many members of a particular protein family are encoded by a genome). Examples of virulence factors: nonribosomal peptide synthetase, HC-toxin [39]; polyketide synthase, T-toxin [38]; ABC transporter, ABC1 [53]; cytochrome P450, pisatin demethylase, [54]; protein kinase, PMK1 [55]. ND, not determined.
of a second fungus differing in some pathological function. Genes present in the first fungus but missing in the second would fail to hybridize [50**].

A case study: genome-wide analysis of NRPSs
Among enzymes, NRPSs are appealing as candidates for roles in fungal pathogenesis because the products of several NRPSs have been proven already as virulence factors. For example, enniatin is required for the virulence of *Fusarium oxysporum* on potatoes [51], HC-toxin for *Cochliobolus carbonum* race 1 on corn [39], and AM-toxin for *Alternaria alternata* on apple [52]. The availability of fungal genome sequences has allowed the quantification of NRPSs in fungi. Surprisingly, the plant-pathogenic fungi *Cochliobolus*, *Fusarium*, and *Botrytis* have an abundance of NRPSs genes, whereas *S. cerevisiae* and the related fungus *A. gossypii* have none (Table 2). *N. crassa* has only a few NRPSs (Table 2), the functions of which are unknown. The *N. crassa* NRPSs are not expected to have pathological functions, however, as this fungus is known only as a saprophyte. Therefore, the *N. crassa* NRPSs have predictive value, that is, their orthologs in pathogenic fungi would have low priority as candidates for virulence functions. Systematic deletion of each NRPS-encoding gene in pathogenic fungi will determine the extent to which this family of enzymes is dedicated to aggressive activities.

Conclusions
Rapid advances in the comprehensive analyses of entire gene sets from a broad array of filamentous fungal genomes are imminent. Sequencing of a few fungal genomes is already complete and sequencing of a number of others is in progress. Genome-wide comparisons coupled with expression profiling will unearth a treasure trove of genes and pathways involved in complex developmental behaviors such as pathogenicity, conidiation, and mating, and will identify pathway overlaps. Equally attractive are the prospects of finding genes unique to fungi that determine the fungal lifestyle generally and genes that distinguish between filamentous and single-cell growth specifically. It should be emphasized that although genomics technologies are powerful, their value is substantially reduced without a genetic system that allows gene validation. Gene validation can be carried out in those fungi that are amenable to procedures used to disrupt, delete, or silence candidate genes. Methods for gene validation are crucial to conclusively establishing the biological roles of gene products.

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

Fungal genomics and pathogenicity Yoder and Turgeon 321


The authors address the issue of how small and large genome rearrangements affect micro- and macrosynteny: the genomes of two yeasts, S. cerevisiae and C. albicans, are compared.


This paper presents the first genome-wide comparison of a filamentous fungus with other eukaryotes, including S. cerevisiae and various nonfungal organisms. Lacking access to a full genome sequence at the time, the authors made use of several EST libraries.


The genome sequence of C. albicans, a human pathogen long considered to be asexual, was searched for homologs of S. cerevisiae and various nonfungal organisms. Lacking access to a full genome sequence at the time, the authors made use of several EST libraries.


This is a comprehensive review of microarrays and how they can be used in whole-genome analysis. It begins with details of array and probe preparation, then proceeds to an assessment of the many ways in which this technology can be used to extract information from biological systems, and ends with considerations of how the large amounts of data generated by chip experiments can be handled.


