Recent advances in the genomic analysis of *Candida albicans*

Beatrice B. Magee and Paul T. Magee

Department of Genetics, Cell Biology, and Development, University of Minnesota, Minneapolis, USA

Summary

The release of the diploid genomic sequence of *Candida albicans* and its recent community-based annotation have permitted a number of studies which have significantly advanced our understanding of the biology of this important human pathogen. These advances range from analysis of genomic changes to differential gene expression under a variety of conditions. A few general conclusions can be drawn from the data presently in hand; one can expect more and more new insights as the number and kind of experiments grows.

Key words

Gene annotation, Genome rearrangement, Gene discovery, Gene expression, Proteomics, Microarray

Avances recientes en el análisis genómico de *Candida albicans*

Resumen

El conocimiento de la secuencia genómica diploide de *Candida albicans* y su reciente anotación basada en el trabajo de varios grupos ha permitido diversos estudios que han supuesto un avance significativo de nuestro conocimiento de la biología de este importante patógeno humano. Estos avances van desde el análisis de cambios genómicos hasta la expresión genética diferencial bajo condiciones diversas. Pueden obtenerse pocas conclusiones generales de los datos de los que disponemos, pero pueden esperarse descubrimientos más profundos a medida que el número y el tipo de experimentos aumenta.

Palabras clave

Anotación génica, Reagrupamiento genómico, Descubrimiento génico, Expresión génica, Proteómica, Microarray (Matriz)

Candida albicans was the first zoopathogenic fungus to have its genome sequenced. The effort began soon after the release of the sequence of the yeast *Saccharomyces cerevisiae*, the first fungus of any kind to be sequenced. The genome of *C. albicans* poses some difficult problems. First, of course, is that the organism is diploid. More importantly, it is highly heterozygous; the two alleles of many genes are so different that normal sequence assembly puts them into two different contigs. The implications of this for the biology are profound; several pairs of alleles have been shown to be differentially regulated. Despite these intrinsic difficulties, the genome sequence has led to a large number of important advances in the basic biology, the epidemiology, and the pathogenesis of this important human pathogen. New and unexpected genes, such as the *MTL* loci have been identified. Microarray experiments examining gene expression under a variety of conditions have been carried out. Loss of heterozygosity during infection as well as in the laboratory has been demonstrated. Spontaneous triploidy has been show to occur in several widely-used laboratory strains. This review will discuss several of these important findings.

Corresponding address:
Dr. Beatrice B. Magee
Department of Genetics, Cell Biology, and Development
University of Minnesota
321 Church St. SE
Minneapolis, MN 55455, USA
Tel.: +1 612 625 4732
Fax: +1 612 624 0426
E-mail: magee002@umn.edu

1130-1406/01/10.00 Euros

Genome structure

Determining, assembling, and annotating the genomic sequence. The sequencing of the genome began in 1996, funded by the Burroughs Wellcome Fund and the National Institute of Craniofacial and Dental Research. The extent of heterozygosity in *C. albicans* was unknown until the release of the second assembly, assembly 6, when the size of the total number of contigs turned out to be 20 Mb (omitting ribosomal RNA), whereas the diploid DNA content per cell had been measured to be the equivalent of 16 Mb. When the final assembly was released, it became clear that the extent of heterozygosity was great enough that the assembly tool Phrap had assembled many alleles as independent contigs, leading to the excess DNA. Jones et al. published a final sequence, assembly 19, which separated out the heterozygous alleles and presented them as diploid regions while presenting the homozygous regions as haploid, with the understanding that they were present in the cell in two copies [26].

The final assembly has 412 contigs, including 148 pairs which correspond to heterozygous regions of various chromosomes, and the majority of them have been assigned to particular chromosomes via the mapping efforts ongoing at the Biotechnology Research Institute at Montreal and at the University of Minnesota. The final organization of the contigs along the chromosomes is not complete, due in large part to the high amount of repeated DNA in the *Candida* genome. Preparing the map is a collaborative effort of the two laboratories listed above. The Minnesota lab is using an STS technique based on a fosmid library; this technique has yielded a complete map of chromosome 7 [8]. The Montreal group is using hybridization of individual chromosomes, isolated by pulse-field gel electrophoresis, to microarrays containing genes whose location on a specific contig is known. They are also refining the assembly, supported by the emerging sequence of the closely related species *Candida dubliniensis*, to generate new supercontigs and assign them to chromosomes. The high frequency and relatively large size of the repeated DNA in the *Candida* genome has greatly slowed the completion of the *C. albicans* genomic map.

However, annotation of the ORFs does not require a genomic map, and there are four websites presenting gene annotation. Several of these were designed to facilitate microarray experiments. The number of ORFs above 100 amino acids was estimated to be 6419 by Jones et al. [26]. A consortium of several laboratories has published a community-based annotation which should be the definitive one for some time at least [5]. They have arrived at a gene number of 6354. There are strong similarities to the relatively closely related model organism *S. cerevisiae* in the number of genes with introns and the number of gene families. In examining the genetic content of the latest annotation, they noted, as had Jones et al., that in contrast to baker's yeast, *Candida* has a large number of genes for expanded metabolic capability, including oligopeptide transporters, amino acid catabolizing enzymes, and lipid degradation enzymes. This difference is suggested to relate to the relatively large number of host niches in which *C. albicans* is found. This annotation has not yet addressed the question of whether some highly heterozygous alleles may have divergent functions. The community-based annotation is the basis for the organization of the *Candida* Genome Database (CGD), a resource located at Stanford University. A list of web sites for sequence data, annotation, proteomics, and mapping is given in Table 1.

Chromosome structure. The diploid genome of *C. albicans* contains eight pairs of homologous chromosomes, ranging in size from 1 Mb to about 4 Mb. They are numbered from chromosome R to chromosome 7 in order of decreasing size. All but chromosome 3 contain at least one copy of a complicated intermediate repeat sequence called the Major Repeat Sequence (MRS) whose function is not known; in fact, the frequency of loss of chromosome 5 on sorbose is correlated with the size of the MRS. There is a significant amount of repeated DNA in the *C. albicans* genome. Much of this is due to the presence of retrotransposon insertions. More than 350 insertions, classified into 34 families, have been identified. The *ALS* gene family is another source of repeated DNA. This family consists of eight or nine members which share a central domain consisting of from four to 36 copies of a 108 bp sequence [22]. The two alleles of *ALS* genes often differ in the number of repeats. The centromeres have been identified on all the chromosomes; there does not seem to be a sequence motif in common [48]. A subtelomeric repeat, *REL1* (or *CARE2)* was found to be present on both ends of chromosome 7 [8] and has so far appeared next to all characterized telomeres (T. Rast, unpublished data).

Genome organization. A great deal of evidence has accumulated suggesting that the genome of *C. albicans* is quite flexible with respect to chromosome number. Barton and Gull were the first to show that aneuploid (2n-1) strains can be isolated [2]. Growth on sorbose often leads to loss of one homologue of chromosome 5 [25]; growth on fluconazole can also lead to chromosome loss [41]. In all cases, the aneuploid (2n-1) strain seems to be at a selective disadvantage with respect to the diploid, and as soon as the selective pressure is removed, diploids which have duplicated the remaining homologue of the monosomic chromosome begin to predominate. An exception to this is the strain WO-2, which is permanently aneuploid since translocations prevent it from achieving euploidy by non-disjunction [35].

Aneuploidy as the result of trisomy has only recently been observed. Chen et al. found that a gene targeted for disruption seemed to have three alleles, as did all

Table 1. Sequence and gene identification.

Assembly 19 supercontigs	http://oldaltoid.stanford.edu:8080/haploid19.html
PUBMED	http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?
Gene annotation	
Agabian	http://agabian.ucsf.edu/
Pasteur	http://genolist.pasteur.fr/CandidaDB/
Minnesota	http://www.cbs.umn.edu/labs/berman/
CGD	http://www.candidagenome.org/
Consortium	http://candida.bri.nrc.ca/candida/index.cfm
Publication	PloS Genetics 1 [5]
Contig mapping	
Minnesota	http://albicansmap.ahc.umn.edu/
Canada	http://cbr-rbc.nrc-cnrc.gc.ca/biovis/candida/
C. dubliniensis sequence	http://www.sanger.ac.uk/Projects/C_dubliniensis
Proteomic maps	
Scotland	http://www.abdn.ac.uk/cogene/
Spain	http://www.babbage.csc.ucm.es/2d/2d.html
Transposon insertion mutants	http://www.tigr.org/tdb/e2k1/caa1/

The sequences of the supercontigs (assembly 19) have been deposited in GenBank. The ORFs can be queried there or at the annotation sites. The Agabian site shows sequence 1000bp before and after the ORF; the Pasteur gives sequence 500bp outside the ORF but gives more information about the gene. CGD provides references and other gene information.

the genes on the same contig. The contig was mapped to chromosome 1, and examination of a number of laboratory strains showed that many were trisomic for chromosome 1 [7]. Trisomic strains were reduced in virulence in a mouse model. Subsequently, Selmecki et al., using a microarray containing probes whose chromosome assignment was known, confirmed the chromosome 1 trisomy and showed that transient trisomy for chromosome 2 is common in laboratory strains [49].

The existence of a diploid assembly allows investigation of the haplotypes found in *C. albicans.* Forche et al. constructed a SNP (Single Nucleotide Polymorphism) map of the *C. albicans* genome, based on the sequence contigs and the emerging genome map [13]. Because of the availability of strains homozygous for one or the other homologue of chromosome 5, the SNP haplotype for this chromosome in strain SC5314 is available. A selection for strains that have undergone loss of heterozygosity at the *GAL1* locus on chromosome 1 was used to isolate strains from infected mice [14]. These isolates were analyzed using a SNP microarray for markers on chromosomes 5 and 7, and loss of heterozygosity due to mitotic recombination was observed. Several of the affected strains had altered colony phenotypes or karyotypic changes [15]. Taken together, the studies on genome organization strongly support the idea that karyotypic variability, including reversible changes in chromosome number and loss of heterozygosity, seems to be quite common and reversible in *C. albicans* and can affect the phenotype.

Gene discovery

As the sequence was emerging, several laboratories mined the expanding information base for information about the genetic capacity of the fungus. Perhaps the most striking example of gene discovery enabled by the sequence led to the demonstration of sexuality in what had been thought of as an asexual fungus. After observing that *C. albicans* contains a gene with very high similarity to the *MAT***a**1 gene of *S. cerevisiae*, Hull and Johnson went on to show that the strain being sequenced, SC5314, has orthologues to both the *MAT***a** and *MAT*α genes. They called the corresponding loci *MTL* for Mating Type Like [23]. *MTL***a** and *MTL*α were shown to be located on separate homologues of chromosome 5 [34]. From this discovery came a series of papers demonstrating mating and characterizing the process in strains engineered to be either homozygous or hemizygous for one mating type and in strains that were homozygous as isolated from patients [24,30,31,34]. Three groups independently mined the genomic sequence and found a potential α-pheromone sequence which turned out to be the biologically active one [3,32,40].

One of the most interesting findings to arise from the discovery of the *MTL* loci was the demonstration that the white-opaque colony phenotype transition, discovered by Slutsky et al. in 1987 [50], was found only in strains which were hemizygous or homozygous for one or the other *MTL* allele. The opaque cell-type was found to be the sexually active form and to mate at a rate some 10^6-fold higher than white cells [36].

Tzung et al. examined the presence or absence of genes required for meiosis in *S. cerevisiae* and found that much, but not all, of the capacity was present in *Candida.* They concluded that the missing genes might be either dispensable or undetectable by homology searches and that *C. albicans* is likely to be capable of a complete sexual cycle [53].

Three genomic approaches have been used to examine gene function in *C. albicans:* a library containing 18,000 transposon-interrupted genes was constructed and screened for mutants in the yeast-to-hyphal transition. One hundred forty-six haploinsufficient genes were identified [54]. Specific targeting of the genes identified in Assembly 6 of the *Candida* genome was carried out using a two-step process which disrupted one allele and placed the other under control of the tetracycline promoter. Five hundred sixty-seven essential genes were identified [46]. Interestingly, only 61% of the genes found to be essential in *S. cerevisiae* were essential in *C. albicans.* A resource for screening the genome using the UAU cassette, which allows disruption of both copies of a gene with a single transformation [12], is available. This resource was constructed by the Mitchell laboratory and TIGR. The website is listed in table 1.

In a clever use of microarray technology, Moran et al. compared whole-genome hybridization of *C. albicans* and the closely related species *C. dubliniensis* and found that about 4% of the *C. dubliniensis* genome failed to hybridize to the *C. albicans* microarray, and that many genes previously implicated in virulence were either absent or highly divergent in sequence in *C. dubliniensis.* This provides insight into the diminished virulence of *C. dubliniensis* compared to *C. albicans* [37].

Gene expression

The availability of the genomic sequence has led to a large variety of microarray experiments to study gene expression. At this time more than 20 papers detailing microarray experiments have been published. The conditions investigated have ranged from growth on blood after transfer from minimal medium to changes in gene expression associated with the acquisition of drug resistance. Table 2 lists many of the microarray experiments which have been published recently. These microarray experiments have provided a large amount of information about gene expression in *C. albicans.* Among other results, they have identified the genes regulated by pH [4], shown that *Candida* lacks a stress response found in *S. cerevisiae* [11], looked at the patterns of gene expression in cells lacking such important transcription factors as Nrg1p, Mig1p, and Tup1p [38].

There have been several studies examining genome-wide gene expression in cells either drug-resistant or growing in the presence of drugs. The first of these was carried out by DeBacker et al. who looked at genes over- and under-expressed in the presence of itraconazole. They found that more than 150 showed increased expression while about 100 showed down-regulation [10]. Cowen et al. carried out a long-term study of the acquisition of drug resistance by cells subjected to increasing concentrations of fluconazole over hundreds of generations. Microarray analysis of several independent resistant lines showed that there were three different resistant phenotypes, one characterized by overexpression of the *CDR2* ABC-transporter efflux pump, and two by overexpression of the *MDR1* multidrug major facilitator transporter gene. One of the *MDR1*-related phenotypes occurred part-way through the experiment in two of the lines but was, in the end, replaced by the other *MDR1* phenotype [9]. An analysis of benomyl-induced *MDR1* and fluphenazine-induced *CDR1*-related phenotypes showed that while some of the genes were affected similarly to resistant clinical strains, there were significant differences. However, in both kinds of strains, drug-resistance and stress-related genes are

Table 2. Some microarray studies in *C. albicans.*

1st Author [ref.]	Date	Subject	Conclusions
Genome Analysis			
Moran [37]	2004	Comparison of genes in *C. albicans* and *C. dubliniensis*	Absence and divergence of 4.4% (247) genes in *C. dubliniensis* including virulence genes
Selmeki [49]	2005	Comparative genome hybridization	Detect aneuploidy and gene deletions
Forche [15]	2005	Single nucleotide polymorphism array	Changes after infection
Drug resistance and response			
De Backer [10]	2001	Response to itraconazole	116 genes down, 180 up including ergosterol biosynthetic genes and genes of unknown function
Cowen [9]	2002	Population genomics of drug resistance	301 genes; three patterns of response to long-term growth on fluconazole
Rogers [47]	2002	Fluconazole-resistant vs. fluconazole-sensitive strains	New genes involved beside transporters
Barker [1]	2004	Induced amphotericin and fluconazole-resistant strains	134 genes including upregulation of sterol biosynthesis genes
Karababa [27]	2004	Compares drug-resistant laboratory strains with clinical strains	Clinical strains upregulate both drug-resistant and cell damage genes
Unique biological properties			
Lan [28]	2002	Metabolic response to white-opaque switching	373 genes; 221 up in opaque (especially oxidative metabolism), 152 in white (fermentative)
Bennett [3]	2003	Mating pheromone response	Induction of 62 genes including cell surface “virulence” genes
Tsong [52]	2003	Mating type and factors	Regulation important for survival in mammalian host
Nantel [39]	2002	Yeast to hyphal transition	Known virulence genes regulated by transition
Sohn [51]	2003	Cell wall dynamics (yeast to hypha)	The yeast to hyphal transition upregulates cell-wall biosynthesis genes
Enjalbert [11]	2003	Stress induced gene expression	Absence of general stress response
Lan [29]	2004	Regulation by iron concentration	526 up in low iron, 626 up in high iron
Bensen [4]	2004	Response to pH and Rim101p	514 pH response genes; alkaline pH affects response to iron
Harcus [20]	2004	Cyclic AMP signaling	cAMP affects genes involved in many processes, including growth. Not all cAMP regulation involves *RAS1*
Fungus-host interactions			
Fradin [17]	2003	Stage-specific response to blood	Gene responses in cells incubated in blood
Fradin [16]	2005	Response to blood fractions	Granulocytes govern the *Candida* response
Lorenz [33]	2004	Internalization by macrophages	Uptake by macrophages induces a starvation response
Garcia-Sanchez [19]	2004	Biofilms vs planktonic cells	~320 genes affected, some independent of hypha formation; sulphur amino acid biosynthesis important
Cao [6]	2005	*C. albicans* biofilm exposed to farnesol	104 genes upregulated by farnesol; 170 downregulated
Proteomics			
Yin [55]	2004	Proteomic response to aminoacids starvation	Comparison with *S. cerevisiae;* differences reflect niches
Pitarch [42]	2004	Identification of serum proteins after *Candida* infection	*C. albicans* housekeeping genes are antigenic

highly expressed [27]. Other studies using matched clinical fluconazole-sensitive and -resistant strains have shown that a variety of new genes are upregulated in the resistant strains, including catabolic and stress-related pathways [47]. Examination of the gene expression changes in an experimentally induced amphotericin B-resistant strain (which had simultaneously become fluconazole-resistant) found expression changes in 134 genes, with sterol biosynthesis and cell stress genes being upregulated and some genes whose products are very important to the cell (histones, protein synthesis components, energy generating components) down-regulated [1]. The common thread through these studies is that drugs induce both specific resistance genes and genes related to cell damage. Interestingly, no one has reported up-regulation of DNA repair genes in these studies, perhaps because none of drugs used target DNA.

A number of groups have taken advantage of microarray technology to examine gene expression under conditions of particular importance in *Candida* biology. The yeast-to-hyphal transition has been of particular interest. Nantel et al. showed that known virulence genes are regulated by the transition from yeast to hyphae [39]. The transition causes derepression of several cell-wall genes, and this derepression is mediated through the transcription factor *EFG1* [51]. Temperature, osmotic, and oxidative stress seem not to induce the kind of stress response in *C. albicans* that one finds in *S. cerevisiae* [11]. However, iron deprivation and pH change both seem to cause widespread gene expression changes. Growth in low iron medium causes the upregulation of more than 500 genes, while growth in excess iron leads to high expression of 626. A number of these genes seem to be subject to regulation by a newly discovered gene called *SFU1* [29]. The effect of pH on gene regulation seems equally global; more than 514 genes change their regulation in response to pH changes. Alkaline pH upregulates several iron acquisition genes, and, in contrast to the case in *S. cerevisiae*, the major pH response gene *RIM101* does not work in *C. albicans* by repressing the *NRG1* gene. Instead, these two transcriptional repressors seem to work in parallel [4]. Harcus et al. used transcriptional profiling to look at the role of cAMP in gene regulation in *Candida.* They used a knockout of *CDC35,* the adenyl cyclase, and found that although cAMP affected genes involved in cellular processes ranging from growth to stress response, there was a significant degree of differentiation among the genes affected [20]. For example, a knockout of *RAS1* only affected a subset of the cAMP-responsive genes. These studies demonstrate clearly that the regulatory circuits in *C. albicans* resemble those in *S. cerevisiae* in some ways but have significant and important differences.

Lan et al. compared the gene expression patterns of white and opaque cells [28], while Tsong looked at the control exerted by the *MTL***a**1, *MTL***a**2, *MTL*α1, and *MTL*α2 genes [52]. These two studies taken together suggest that much of the gene regulation specific to mating is exerted by the white-opaque switch. Bennett et al. looked at the genes induced by the α-pheromone and found that in addition to genes that seem to be important in mating, a number of genes implicated in virulence are also induced [3]. Lockhart et al. using RT-PCR technology, found similar results independently [32].

Fradin et al. looked at the changes in gene expression upon incubation in blood or plasma and were able to identify several sets of genes whose expression went up (oxidative stress-related genes, alkaline-pH-induced genes, both glycolytic and glyoxalate-cycle genes, and hypha-related genes) [18]. Almost as many genes were down-regulated (240) as were up-regulated (283), prominent among them heat-shock genes [18]. Much of this response was shown to be due to neutrophils in the blood; experiments in which the blood was depleted of granulocytes, monocytes, or all cells showed that the granulocytes seemed to induce the largest number of changes in gene expression [16].

Since uptake by macrophages is one of the early steps in the infection process, Lorenz et al. looked at transcriptional changes occurring after engulfment. They found that the cells shifted to a starvation response immediately, followed by recovery characterized by hyphal growth and a return to normal metabolism. A large number of the differentially regulated ORFs are uncharacterized [33].

Two laboratories have looked at gene expression in biofilms. García-Sánchez et al. using an apparatus in which nutrient composition or flow can be varied, found that the pattern of gene expression was more stable than in cells undergoing planktonic growth [19]. More than 300 genes were shown to be differentially expressed, among them a number of amino acid biosynthetic genes. Since hypha formation seems likely to be intricately associated with biofilm formation, a biofilm formed by a *cph/cph efg/efg* double mutant was compared with one formed by a wild-type strain. Three hundred and seventeen genes were expressed independently of these transcriptional regulators, while 86 depended on them. Farnesol has been shown to be a quorum-sensing molecule in *C. albicans* [21] and to inhibit biofilm formation [45]. Treatment of a biofilm with farnesol leads to upregulation of 104 genes and downregulation of 170 [6].

As the sequence became available, the *C. albicans* proteome became accessible. Pitarch and coworkers began early to develop proteomic techniques [43,44], and the application of those techniques to the examination of the antigens important in systemic candidiasis has been published. These workers found that housekeeping gene products are antigenic, and that responses to specific antigens could be reflect a differentiation of the immune response. Furthermore, they suggest that high enolase-antibody concentrations may be associated with recovery [42].

In a metabolic analysis, Yin et al. looked at the response of the proteome to 3-aminotriazole, a histidine analogue which triggers the general amino acid control response regulated by the *GCN4* gene, in *Candida* and in baker's yeast, where it has been well characterized [55]. Many of the proteins induced were the same; the exceptions seem to relate to the different niches of the two fungi. In general the proteomic studies support and extend the transcription analyses, but as more appear, one can expect evidence for post-transcriptional regulation to surface.

Conclusion

The availability of the genome sequence of *C. albicans* has made possible a variety of experiments which have greatly expanded our knowledge of the biology of this important pathogen. Advances have been made in genome structure and dynamics, in the identification of genes which may play a role in the special niches (commensalism and pathogenicity) of the organism, and in the patterns of gene expression under conditions important to the biology. The complete annotation will stimulate gene and protein expression analysis under an increasingly varied number of conditions; these data will in turn cast light on the function of the large number of genes of un-known function.

References

1. Barker KS, Crisp S, Wiederhold N, Lewis RE, Bareither B, Eckstein J, Barbuch R, Bard M, D. Rogers P. Genome-wide expression profiling reveals genes associated with amphotericin B and fluconazole resistance in experimentally induced antifungal resistant isolates of *Candida albicans.* J Antimicrob Chemother 2004; 54: 376-385.
2. Barton RC, Gull K. Isolation, characterization, and genetic analysis of monosomic, aneuploid mutants of *Candida albicans.* Mol Microbiol 1992; 6: 171-177.
3. Bennett RJ, Uhl MA, Miller MG, Johnson AD. Identification and characterization of a *Candida albicans* mating pheromone. Mol Cell Biol 2003; 23: 8189-8201.
4. Bensen ES, Martin SJ, Li M, Berman J, Davis DA. Transcriptional profiling in *Candida albicans* reveals new adaptive responses to extracellular pH and functions for Rim101p. Mol Microbiol 2004; 54: 1335-1351.
5. Braun BR, van Het Hoog M, d'Enfert C, Martchenko M, Dungan J, Kuo A, Inglis DO, Uhl MA, Hogues H, Berriman M, Lorenz M, Levitin A, Oberholzer U, Bachewich C, Harcus D, Marcil A, Dignard D, Iouk T, Zito R, Frangeul L, Tekaia F, Rutherford K, Wang E, Munro CA, Bates S, Gow NA, Hoyer LL, Kohler G, Morschhauser J, Newport G, Znaidi S, Raymond M, Turcotte B, Sherlock G, Costanzo M, Ihmels J, Berman J, Sanglard D, Agabian N, Mitchell AP, Johnson AD, Whiteway M, Nantel A. A human-curated annotation of the *Candida albicans* genome. PLOS Genet 2005; 1: 36-57.
6. Cao YY, Cao YB, Xu Z, Ying K, Li Y, Xie Y, Zhu ZY, Chen WS, Jiang YY. cDNA microarray analysis of differential gene expression in *Candida albicans* biofilm exposed to farnesol. Antimicrob Agents Chemother 2005; 49: 584-589.
7. Chen X, Magee BB, Dawson D, Magee PT, Kumamoto CA. Chromosome 1 trisomy compromises the virulence of *Candida albicans.* Mol Microbiol 2004; 51: 551-565.
8. Chibana H, Magee BB, Grindle S, Ran Y, Scherer S, Magee PT. A physical map of Chromosome 7 of *Candida albicans.* Genetics 1998; 149: 1739-1752.
9. Cowen LE, Nantel A, Whiteway MS, Thomas DY, Tessier DC, Kohn LM, Anderson JB. Population genomics of drug resistance in *Candida albicans.* Proc Natl Acad Sci USA 2002; 99: 9284-9289.
10. 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.
11. Enjalbert B, Nantel A, Whiteway M. Stress-induced gene expression in *Candida albicans:* absence of a general stress response. Mol Biol Cell 2003; 14:1460-1467.
12. Enloe B, Diamond A, Mitchell AP. A single-transformation gene function test in diploid *Candida albicans.* J Bacteriol 2000; 182: 5730-5736.
13. Forche A, Magee PT, Magee BB, May G. Genome-wide single-nucleotide polymorphism map for *Candida albicans.* Eukaryot Cell 2004; 3:705-714.
14. Forche A, May G, Beckerman J, Kauffman S, Becker J, Magee PT. A system for studying genetic changes in *Candida albicans* ns during infection. Fungal Genet Biol 2003; 39: 38-50.
15. Forche A, May G, Magee PT. Demonstration of LOH by SNP microarray analysis and alterations in strain morphology in strains of *Candida albicans* during infection. Eukaryot Cell 2005; 4: 156-165.
16. Fradin C, De Groot P, MacCallum D, Schaller M, Klis F, Odds FC, Hube B. Granulocytes govern the transcriptional response, morphology, and proliferation of *Candida albicans* in human blood. Mol Microbiol 2005; 56: 379-415.
17. Fradin C, Hube B. Tissue infection and site-specific gene expression in *Candida albicans.* Adv Appl Microbiol 2003; 53: 271-290.
18. Fradin C, Kretschmar M, Nichterlein T, Gaillardin C, d'Enfert C, Hube B. Stage-specific gene expression of *Candida albicans* in human blood. Mol Microbiol 2003; 47: 1523-1543.
19. Garcia-Sanchez S, Aubert S, Iraqui I, Janbon G, Ghigo JM, d'Enfert C. *Candida albicans* biofilms: a developmental state associated with specific and stable gene expression patterns. Eukaryot Cell 2004; 3: 536-545.
20. Harcus D, Nantel A, Marcil A, Rigby T, Whiteway M. Transcription profiling of cyclic AMP signaling in *Candida albicans.* Mol Biol Cell 2004; 15: 4490-4499.
21. Hornby JM, Jensen EC, Lisec AD, Tasto JJ, Jahnke B, Shoemaker R, Dussault P, Nickerson KW. Quorum sensing in the dimorphic fungus *Candida albicans* is mediated by farnesol. Appl Environ Microbiol 2001; 67: 2982-2992.
22. Hoyer LL. The ALS gene family of *Candida albicans.* Trends Microbiol 2001; 9: 176-180.
23. Hull CM, Johnson AD. Identification of a mating type-like locus in the asexual pathogenic yeast *Candida albicans.* Science 1999; 285: 1271-1275.
24. Hull CM, Raisner RM, Johnson AD. Evidence for mating of the "Asexual" yeast *Candida albicans* in a mammalian host. Science 2000; 289: 307-310.
25. Janbon G, Sherman F, Rustchenko E. Monosomy of a specific chromosome determines L-sorbose utilization: a novel regulatory mechanism in *Candida albicans.* Proc Natl Acad Sci USA 1998; 95: 5150-5155.
26. Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, Newport G, Thorstenson YR, Agabian N, Magee PT, Davis RW, Scherer S. The diploid genome sequence of *Candida albicans.* Proc Natl Acad Sci USA 2004; 101: 7329-7334.
27. Karababa M, Coste AT, Rognon B, Bille J, Sanglard D. Comparison of gene expression profiles of *Candida albicans* azole-resistant clinical isolates and laboratory strains exposed to drugs inducing multidrug transporters. Antimicrob Agents Chemother 2004; 48: 3064-3079.
28. Lan CY, Newport G, Murillo LA, Jones T, Scherer S, Davis RW, Agabian N. Metabolic specialization associated with phenotypic switching in *Candida albicans.* Proc Natl Acad Sci USA 2002; 99: 14907-14912.
29. Lan CY, Rodarte G, Murillo LA, Jones T, Davis RW, Dungan J, Newport G, Agabian N. Regulatory networks affected by iron availability in *Candida albicans.* Mol Microbiol 2004; 53: 1451-1469.
30. Legrand M, Lephart P, Forche A, Mueller FM, Walsh T, Magee PT, Magee BB. Homozygosity at the MTL locus in clinical strains of *Candida albicans:* karyotypic rearrangements and tetraploid formation. Mol Microbiol 2004; 52: 1451-1462.
31. Lockhart SR, Daniels KJ, Zhao R, Wessels D, Soll DR. Cell Biology of Mating in *Candida albicans.* Eukaryot Cell 2003; 2: 49-61.
32. Lockhart SR, Zhao R, Daniels KJ, Soll DR. Alpha-pheromone-induced "shmooing" and gene regulation require white-opaque switching during *Candida albicans* mating. Eukaryot Cell 2003; 2: 847-855.
33. Lorenz MC, Bender JA, Fink GR. Transcriptional response of *Candida albicans* upon internalization by macrophages. Eukaryot Cell 2004; 3: 1076-1087.
34. Magee BB, Magee PT. Induction of mating in *Candida albicans* by construction of *MTL* a and *MTL* alpha strains. Science 2000; 289: 310-313.
35. Magee BB, Magee PT. WO-2, a stable aneuploid derivative of *Candida albicans* strain WO-1, can switch from white to opaque and form hyphae. Microbiology 1997; 143: 289-295.
36. Miller MG, Johnson AD. White-opaque switching in *Candida albicans* is controlled by mating-type locus homeodomain proteins and allows efficient mating. Cell 2002; 110: 293-302.
37. Moran G., Stokes C, Thewes S, Hube B, Coleman DC, Sullivan D. Comparative genomics using *Candida albicans* DNA microarrays reveals absence and divergence of virulence-associated genes in *Candida dubliniensis.* Microbiology 2004; 150: 3363-3382.
38. Murad AM, d'Enfert C, Gaillardin C, Tournu H, Tekaia 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.
39. Nantel A, Dignard D, Bachewich C, Harcus D, Marcil A, Bouin AP, Sensen CW, Hogues H, van het Hoog M, Gordon P, Rigby T, Benoit F, Tessier DC, Thomas DY, Whiteway M. Transcription profiling of *Candida albicans* cells undergoing the yeast-to-hyphal transition. Mol Biol Cell 2002; 13: 3452-3465.
40. Panwar SL, Legrand M, Dignard D, Whiteway M, Magee PT. MFalpha1, the gene encoding the alpha mating pheromone of *Candida albicans.* Eukaryot Cell 2003; 2: 1350-60.
41. Perepnikhatka V, Fischer FJ, Niimi M, Baker RA, Cannon RD, Wang YK, Sherman F, Rustchenko E. Specific chromosome alterations in fluconazole-resistant mutants of *Candida albicans.* J Bacteriol 1999; 181: 4041-4049.
42. Pitarch A, Abian J, Carrascal M, Sanchez M, Nombela C, Gil C. Proteomics-based identification of novel *Candida albicans* antigens for diagnosis of systemic candidiasis in patients with underlying hematological malignancies. Proteomics 2004; 4: 3084-3106.
43. Pitarch A, Sanchez M, Nombela C, Gil C. Analysis of the *Candida albicans* proteome. I. Strategies and applications. J Chromatogr B Analyt Technol Biomed Life Sci 2003; 787: 101-128.
44. Pitarch A, Sanchez M, Nombela C, Gil C. Analysis of the *Candida albicans* proteome. II. Protein information technology on the Net (update 2002). J Chromatogr B Analyt Technol Biomed Life Sci 2003; 787: 129-148.
45. Ramage G, Saville SP, Wickes BL, Lopez-Ribot JL. Inhibition of *Candida albicans* biofilm formation by farnesol, a quorum-sensing molecule. Appl Environ Microbiol 2002; 68: 5459-5463.

46. Roemer T, Jiang B, Davison J, Ketela T, Veillette K, Breton A, Tandia F, Linteau A, Sillaots S, Marta C, Martel N, Veronneau S, Lemieux S, Kauffman S, Becker J, Storms R, Boone C, Bussey H. Large-scale essential gene identification in *Candida albicans* and applications to antifungal drug discovery. Mol Microbiol 2003; 50: 167-181.
47. Rogers PD, Barker KS. Evaluation of differential gene expression in fluconazole-susceptible and -resistant isolates of *Candida albicans* by cDNA microarray analysis. Antimicrob Agents Chemother 2002; 46: 3412-3417.
48. Sanyal K, Baum M, Carbon J. Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. Proc Natl Acad Sci USA 2004; 101: 11374-11379.
49. Selmecki A, Bergmann S, Berman J. Comparative genome hybridization reveals widespread aneuploidy in *Candida albicans* laboratory strains. Mol Microbiol 2005; 55: 1553-1565.
50. Slutsky B, Staebell M, Anderson J, Risen L, Pfaller M, Soll DR. "White-opaque transition": a second high-frequency switching system in *Candida albicans.* J Bacteriol 1987; 169: 189-197.
51. Sohn K, Urban C, Brunner H, Rupp S. EFG1 is a major regulator of cell wall dynamics in *Candida albicans* as revealed by DNA microarrays. Mol Microbiol 2003; 47: 89-102.
52. Tsong AE, Miller MG, Raisner RM, Johnson AD. Evolution of a combinatorial transcriptional circuit: a case study in yeasts. Cell 2003; 115: 389-399.
53. Tzung KW, Williams RM, Scherer S, Federspiel N, Jones T, Hansen N, Bivolarevic V, Huizar L, Komp C, Surzycki R, Tamse R, Davis RW, Agabian N. Genomic evidence for a complete sexual cycle in *Candida albicans.* Proc Natl Acad Sci USA 2001; 98: 3249-3253.
54. Uhl MA, Biery M, Craig N, Johnson AD. Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen *C. albicans.* Embo J 2003; 22: 2668-2678.
55. Yin Z, Stead D, Selway L, Walker J, Riba-Garcia I, McLnerney T, Gaskell S, Oliver SG, Cash P, Brown AJ. Proteomic response to amino acid starvation in *Candida albicans* and *Saccharomyces cerevisiae.* Proteomics 2004; 4: 2425-2436.