Time course of global gene expression alterations in *Candida albicans* during infection of HeLa cells

Yi-Bing Lan[1#], Yi-Zhou Huang[1#], Fan Qu[1], Juan-Qing Li[1], Lin-Juan Ma[1], Jie Yan[2,3*], Jian-Hong Zhou[1*]

[1]Department of Gynecology, The Women's Hospital, Zhejiang University School of Medicine, Hangzhou, China, [2]Department of Medical Microbiology and Parasitology, Zhejiang University School of Medicine, Hangzhou, China, [3]Division of Basic Medical Microbiology, State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou, China

ABSTRACT

Candida albicans (*C. albicans*) is an opportunistic fungus that quickly adapts to various microniches. It causes candidiasis, a common fungal infection for which the pathogenic mechanism has not been elucidated yet. To explore the pathogenic mechanism of candidiasis we used several methods, including microscopic observation of morphological changes of HeLa cells and fungus, analysis of differentially expressed genes using gene chips, and a series of biological and bioinformatic analyses to explore genes that are possibly involved in the pathogenesis of *C. albicans*. During the *C. albicans* infection, significant morphological changes of the fungus were observed, and the HeLa cells were gradually destroyed. The gene chip experiments showed upregulated expression of 120 genes and downregulated expression of 178 genes. Further analysis showed that some genes may play an important role in the pathogenesis of *C. albicans*. Overall, morphological variation and adaptive gene expression within a particular microniche may exert important effects during *C. albicans* infections.

KEY WORDS: *Candida albicans*; infection; differential gene expression; HeLa cells
DOI: http://dx.doi.org/10.17305/bjbms.2017.1667 *Bosn J Basic Med Sci. 2017;17(2):120-131.* © *2017 ABMSFBIH*

INTRODUCTION

Candida albicans (*C. albicans*) is an opportunistic fungus that causes infections of different organs in human body, collectively known as candidiasis [1]. More than 70% of women will experience candidal vulvovaginitis once or more in their lives, and *C. albicans* is the most common Candida species associated with this infection [2].

C. albicans is a unicellular eukaryote that exists in two main morphological forms during infection: yeast and hyphal form [3]; in addition, a morphological form that involves a shift to pseudohyphal growth occurs only during germination. In general, morphological changes are considered to play a crucial role in the pathogenesis of *C. albicans* infection [4].

*Corresponding authors: Jian-Hong Zhou, Department of Gynecology, the Women's Hospital, Zhejiang University School of Medicine, 2 Xue-Shi Road, Hangzhou, Zhejiang 310006, China. Phone: +86-571-87061501, Fax: +86-571-87061501. E-mail: zhoujh1117@zju.edu.cn
Jie Yan, Department of Medical Microbiology and Parasitology, Zhejiang University School of Medicine, 866 Yu-Hang-Tang Road, Hangzhou, Zhejiang 310058, China. Phone: +86-571-88208297. Fax: +86-571-88208294. E-mail: med_bp@zju.edu.cn
#These authors equally contributed

Submitted: 15 September 2016/Accepted: 09 December 2016

The genomic sequence of *C. albicans* has been completely annotated [5], which means that it is possible to use genome-wide transcript profiling to identify potential genes that are associated with fungal infections. However, because only a few studies have investigated *C. albicans* vaginal infection using experimental models, the pathogenesis of vaginal candidiasis is still not clear. To investigate the pathogenesis of vaginal candidiasis we simulated *C. albicans* infection in genital tract epithelial cells using HeLa cells. First, morphological changes of *C. albicans* and HeLa cells were observed. Next, changes in gene expression profiles of *C. albicans* were analyzed. Finally, multiple biological and bioinformatic analyses were performed. According to our results, *CHT3*, *SCW11*, *ALG7*, *HYR3*, *LAT1*, *KEL1*, *CRH11*, *FGR41*, *RHD3*, *CHS1*, *MUQ1*, *CPY1*, *FGR41*, *IDP1*, *HXK1*, *SDH41*, *SSK1*, *PDA1*, and *ZRT2* genes were in the center of the co-expression regulatory network and, thus, may play important roles in the pathogenesis of vaginal candidiasis. Our results indicate that the morphological variation and adaptive gene expression, which occur in response to host microniches, markedly influence host-fungus interactions during *C. albicans* infections.

MATERIALS AND METHODS

Strains and growth conditions

C. albicans strain SC5314 was used in our experiments. *C. albicans* cells were grown on YPD agar plates (1% yeast extract, 2% peptone, 2% glucose, and 2% agar) at 30°C and passaged every day. The HeLa cells (provided by the Cell Bank at the Institute of Cytobiology, Chinese Academy of Science, Shanghai, China) were maintained in RPMI-1640 medium (Invitrogen, Carlsbad, CA, USA) supplemented with 10% heat inactivated fetal calf serum (FCS, Gibco/Invitrogen, Carlsbad, CA, USA) at 37°C in an atmosphere of 5% CO_2. The cells were passaged four times before use.

Screening of *C. albicans* morphological changes during infection

The HeLa cells (1×10^5 cells per well) were seeded in 6-well culture plates (Corning, New York, USA) containing a 12×12 mm coverslip in each well, and were incubated for 12 hours at 37°C until they formed cell monolayers. Then, the fungal cells were suspended in RPMI-1640 medium, supplemented with 2% fetal calf serum (FCS), for cell counts. Next, the monolayers of HeLa cells were infected with *C. albicans* strain SC5314 (1×10^7 per well) at a multiplicity of infection (MOI) of 100 and incubated for 1, 3, 6, 12 or 24 hours at 37°C in an atmosphere of 5% CO_2 [6]. At the selected time points, the coverslips were washed twice with phosphate-buffered saline (PBS) to remove unattached fungi, and then were fixed in 4% paraformaldehyde at room temperature for 30 minutes. The morphological changes of *C. albicans* and HeLa cells during infection were observed under a light microscope.

Total RNA extraction from *C. albicans* and complementary RNA synthesis

The suspension of *C. albicans* used for the infection of HeLa cells was prepared as described above. Similarly, to form cell monolayers, HeLa cells (1×10^7) were seeded in tissue culture flasks (150 cm^2, Corning, New York, USA) and incubated for 12 hours at 37°C. The cell monolayers were thoroughly washed with PBS and 20 ml of RPMI-1640 medium, supplemented with 2% FCS, was added to each flask. Then, in the experimental groups, the monolayers were infected with the fungus (1×10^9 cells per flask) at MOI of 100 and incubated for 1, 3, 6, 12 and 24 hours at 37°C in an atmosphere of 5% CO_2. The same protocols (incubation time, culture medium and conditions, and number of fungi [1×10^9]) were applied to control groups, except HeLa cells were not added to these cultures. At various time points, 2 ml of lysis buffer (1% Triton X-100, 150 mmol/L NaCl, 0.02% sodium azide, 100 μg/ml PMSF, 1 μg/ml aprotinin, 50 mmol/L Tris-HCl, pH 8.0), that lyse HeLa but not fungal cells, was added to the flasks in the experimental and control groups, on ice. The mixture was then centrifuged at 10,000×g for 10 minutes (4°C) to precipitate molecules from fungal cells. After 2 washes in PBS and subsequent centrifugation, total RNA from the fungal pellets was extracted by the hot acid phenol method [7]. The total RNA was digested with DNase and purified using a NucleoSpin® RNA Clean-Up Kit (Macherey-Nagel, GmbH & Co. KG, Germany), according to the manufacturer's protocol. The quantity and integrity of total RNA were confirmed by formaldehyde denatured gel electrophoresis.

Microarray detection

Photolithography-based, monoplex DNA microarray chip was provided by CapitalBio Corp, China. The chip contains 7925 oligo DNA probes (approximately 70-mer per probe) that cover all the predicted 6354 open reading frames (ORFs) and 6 control genes from *C. albicans* genome, based on the data from the Operon *C. albicans* Genome Oligo Set (Version 1.1) (http://www.Operon.com). Briefly, cRNA was prepared from 5 mg of total RNA isolated from *C. albicans* and purified with RNA Clean-Up Kit (Macherey & Nagel, GmbH & Co. KG, Germany). Next, the cRNA was reverse transcribed into cDNA with CbcScript-II reverse transcriptase (Invitrogen, Carlsbad, CA, USA) and labeled with Cy3 or Cy5 dye (GE Healthcare, Little Chalfont, Buckinghamshire, England). The labeled products were purified as described above and then dissolved in 80 μl of hybridization solution (25% formamide, 3×saline-sodium citrate (SSC), 5×Denhardt's Solution, and 0.2% sodium dodecyl sulfate [SDS]); the products were hybridized to a microarray chip slide at 42°C overnight in a rotator, at 180 rpm. After washing and drying, the microarray slides were scanned using a LuxScan 10KA dual-channel laser scanner (CapitalBio, Beijing, China). The image signals were transformed into digital signals using LuxScan 3.0 image analysis software, and data were normalized using the Lowess algorithm.

Microarray and two class differential expression (Dif) analyses

After significance and false discovery rate (FDR) were calculated, we selected differentially expressed genes according to P-value threshold of $P < 0.05$ and FDR threshold of $P < 0.05$ [8]. Genes that showed ratios ≥2.0 were considered to be upregulated, while genes that showed ratios ≤0.5 were considered to be downregulated.

Series tests of cluster (STC) analysis

Based on the above analysis, we identified a set of distinct and representative, temporally expressed genes. The raw

expression values were converted into $\log_2$ ratios, and the permutation test was used to determine the assignment of genes to model profiles through a large number of permutations of the time points [9]. Standard hypothesis testing was used to determine which model profiles have significantly more genes assigned under the true ordering of time points compared with the average number assigned to the model profile in the permutation runs.

Data confirmation by quantitative reverse transcription PCR (RT-qPCR)

To confirm the results of microarray analysis and gene expression trends, 15 genes with typical expression trends were selected for RT-qPCR at different time points; these data were then converted into a heat map and bar chart. The primers used for RT-qPCR are listed in Supplemental Table 1. Briefly, cDNA was synthesized from 2.5 μg of *C. albicans* RNA by reverse transcription (RT) using a cDNA Synthesis Kit (TaKaRa, Otsu, Japan), and prepared as described above. Using cDNA as a template, the mRNA level of each target gene was measured by qPCR using a SYBR® Ex-Taq™ Kit (TaKaRa, Otsu, Japan). The 18S RNA gene of *C. albicans* was used as the internal reference gene [10]. All calculations and statistical analyses were performed on an ABI PRISM 7500 sequence analyzer (Applied Biosystems, Forster City, CA, USA).

STC-Gene Ontology (GO) analysis

Using GO analysis and the annotation of genes based on the Gene Ontology database, we determined which genes participate at different time points. In simpler terms, the GO analysis is applied to genes that show specific tendencies. It is also used to determine the primary function of genes that have the same expression trend according to the Gene Ontology database. At the same time, as the enrichment increases the corresponding function is more specific, which helps us determine the GOs with a more concrete description of function in the experiment.

Pathway analysis and construction of gene co-expression network (GCN)

To determine the significant pathways of the differentially expressed genes, we assigned these genes to different pathways according to the Kyoto Encyclopedia of Genes and Genomes (KEGG), and a significance analysis based on discrete distribution was performed. The Fisher's exact test and χ^2 tests were used to select the significant pathway, and the threshold of significance was defined by P-values <0.05 and FDRs <0.05 [11].

To clarify interactive relationships between genes over time, we selected genes from 12 significant trends according to the STC analysis, and built GCN (Dynamic-GeneNet analysis) to obtain the key genes in the network. For building this network, we performed a dynamic simulation of gene interactions using gene-sample matrix [12]. In the network, the nodes represent genes, and the line segments represent the interactions between the genes. Solid lines represent positively correlated genes and dashed lines represent negatively correlated genes. We calculated the strength of gene relationships according to the position of each gene in the network and its specific network value. Genes that had the highest characteristic value were in the hub position of the network, and may play a critical role in *C. albicans*. k-core, clustering coefficient, and degree were key attributes in the network. Because the network elements represent different genes that can regulate each other, large scale gene networks can be divided into subgraphs, named k-core networks. In each subgraph, all genes are connected to at least k other genes. Consequently, the rank of k-core value indicates the complexity of gene relationships; a greater k-core value indicates higher co-expression of differentially expressed genes. The color of the nodes is the result of the clustering divided by k-core. Same size of the k-core reflects the similarity and correlated function of genes. The identical color nodes are coming from the same subgraph. The complexity of gene relationships increases with k-core value rank. Through disparate k-cores, we expect to define the functions of key genes at different complexity level of the network, at the same time, we can observe snapshots of gene expression regulatory networks among different groups. Edges between two nodes (genes) represent interactions between the genes. These interactions are quantified by degree. The degree indicates how many links one node has to other nodes, and is represented by the node size. Nodes with high degree values are in central positions within the network. The clustering coefficient measures the complexity of interactions between genes that neighbor the core gene, and the core gene is not included. Lower clustering coefficient indicates that the interactions between the neighboring genes are more independent of the core gene. Due to a large number of genes in the first GCN analysis, we could not easily determine the precise location of the genes in the network. To further investigate the significance and interactions between genes, we performed two additional GCN analyses. In the first group, we included all genes from the 12 significant STC-GOs. In the second group, we included all genes that were associated with significant pathways. To construct the two additional GCNs we used genes that were in the intersection and union set of the two groups, respectively. In all three Dynamic-GeneNet analyses, if the k-core of a gene was at maximum, degree was >10, and clustering coefficient was >0.5, we hypothesized an important role for that gene in the GCN.

RESULTS

Attachment and hyphae formation of *C. albicans* during infection

In the early stage of infection (1-3 hours), the *C. albicans* strain SC5314 was attached to the surface of HeLa cells. The fungus also transformed from its yeast form to the hyphal form (Figure 1B and C). Until the end of the 3rd hour, the hyphae adhered extensively to the cell surface; the cell-cell interactions could not be observed clearly, and the cell surface became loose. In the invasion phase (3-12 hours), the hyphae grew abundantly, and the epithelial cells further changed their shapes, some of which demonstrated ruptured cell membranes and the overflow of cell content. However, in some cells, the nuclei were still visible. Finally, in the late phase (12-24 hours), the epithelial cells and the cell nuclei were rarely observed. This was due to the fact that most of the nuclei were completely destroyed by the fungus in the background of the visual fields, and only a few scattered cell fragments remained (Figure 1).

Differentially expressed genes

Compared with the control group, the number of differentially expressed genes in the experimental group increased with time: out of 1076 gene probes 120 differentially expressed genes were upregulated, and out of 1166 gene probes 178 genes were downregulated. These genes were associated with hyphal growth, adherence ability, amino acid metabolism, and drug resistance among other processes.

STC analysis

To further analyze 2242 differentially expressed genes associated with important biochemical processes, we performed the expression trend analysis over five time points and obtained 80 theoretical gene expression tendencies. As shown in Figure 2, 12 gene expression trends were considered significant and marked with red lines; these are shown as profiles 2, 3, 5, 6, 11, 12, 30, 38, 39, 41, 56, and 65, and were used for further analysis.

Analyzing the 12 significant different gene expression tendencies, we found that *PDR16*, *CDR1*, *CDR4*, *ARE2*, *RTA3*, *ERG4*, *IPL1*, *ADR1*, *GPX1*, *IDP2*, *ADK1*, *GPX1*, *IDP2*, *IFE2*, and *TPO4* genes, among others, were significantly upregulated, while *SAM4*, *ALS2*, *HEX1*, *MUQ1*, *GAL7*, *CHS7*, *AYR2*, *CPY1*, *IHD1*, *DAC1*, *GPD1*, *HGT14*, *ALD6*, *TPO3*, *DUR3*, *ARR3*, *PMI1*, *HEM14*, and *HGT14* genes, among others, were markedly downregulated.

Validation of mRNA expression in *C. albicans*

Fifteen selected differentially expressed genes were in accordance with the results of gene chip and differential gene expression tendency analyses (Figure 3 and Supplemental Figure 1).

STC-GO analysis outcomes

Based on the analysis of significant gene expression tendencies, we performed GO annotation using the Gene Ontology database. Taking the profile 3 as an example (Figure 4), the

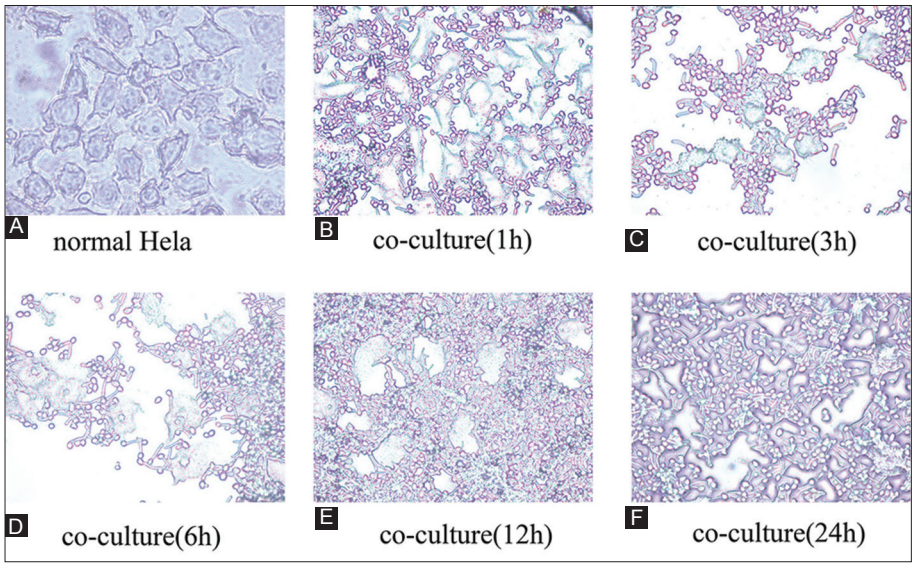


FIGURE 1. Attachment and hyphal formation of *C. albicans* during infection of HeLa cells. (A) In uninfected (normal) HeLa cells, we could clearly see the cell membranes and nuclei. (B) After 1 hour, the fungus attached to the surface of HeLa cells and transformed from its yeast to hyphal form. (C) After 3 hours, the hyphae adhered extensively to the cell surface so that the cell-cell interactions could not be clearly observed, and the cell surface became loose. (D-E) In the invasion phase (3-12 hours), the hyphae were abundant, the shape of the epithelial cells changed, the membranes were ruptured and the cell content overflowed. (F) In the late phase (12–24 hours), the epithelial cells were completely destroyed and only a large number of *C. albicans* hyphal forms was observed.

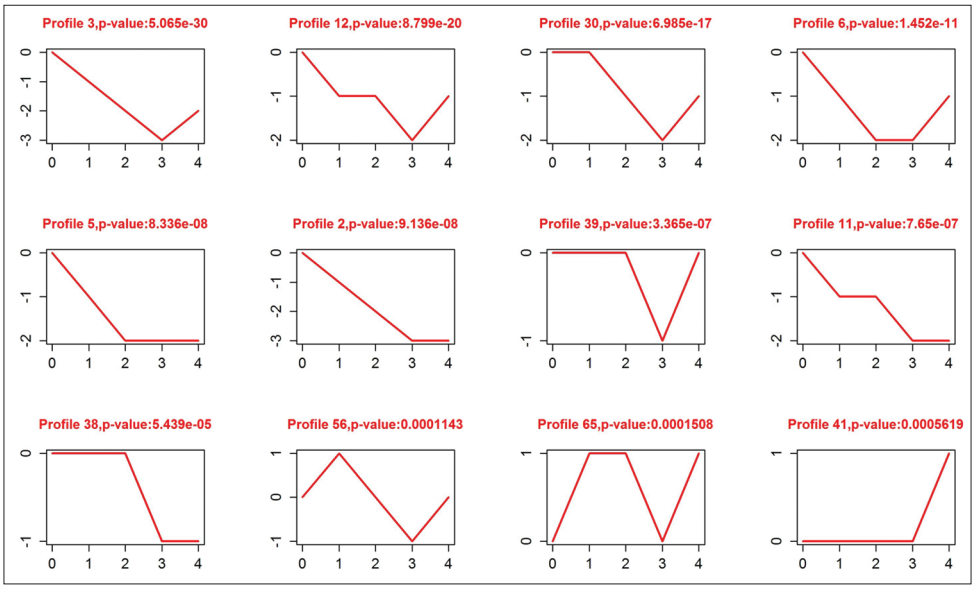


FIGURE 2. Twelve significant differential expression tendencies (profiles 2, 3, 5, 6, 11, 12, 30, 38, 39, 41, 56, and 65). Eighty theoretical expression tendencies were observed, but only 12 expression trends were considered significant. The most significant expression trend was observed in the profile 3, where gene expression decreased initially but after that increased slightly at the last time point.

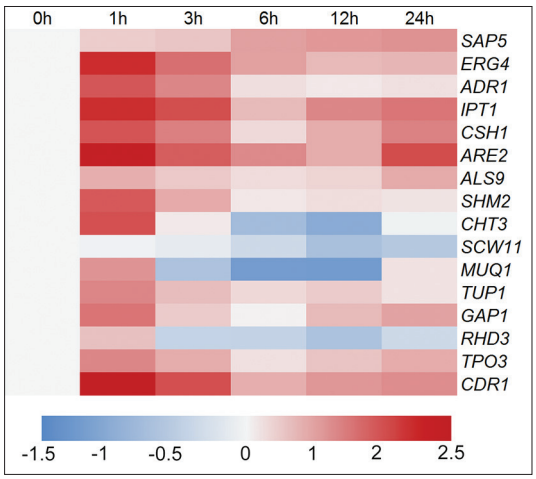


FIGURE 3. Heat map of genes detected by quantitative reverse transcription PCR (RT-qPCR). The heat map illustrates that the 15 selected differentially expressed genes were in accordance with the results of gene chip and differential gene expression tendency analyses. The rows represent genes and the columns represent replicates of each time points. Blue indicates down-regulation, red up-regulation, and white indicates the gene expression of *Candida albicans* before infection. The average and standard deviation (STD) can be seen in Supplemental Figure 1.

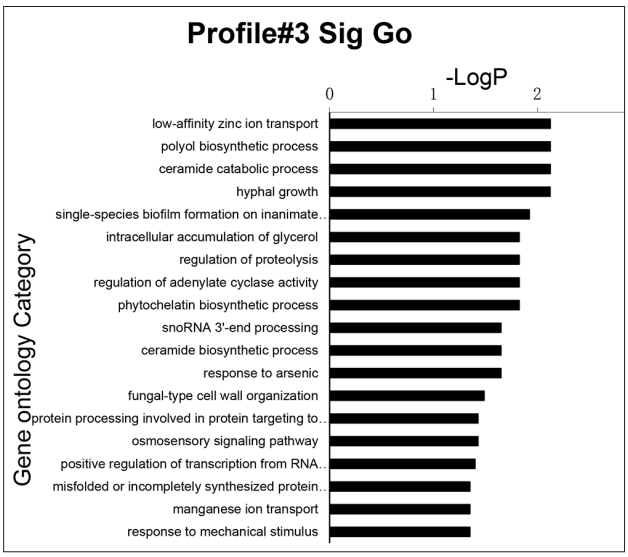


FIGURE 4. Series tests of cluster (STC) - Gene Ontology (GO) analysis of the third significant expression tendency (profile 3). The STC-GO analysis showed that the genes from profile 3 were associated with a variety of biological functions, including hyphal growth, pseudohyphal growth, budding cell bud growth, cell-matrix adhesion, biofilm formation, mRNA polyadenylation, and response to osmotic stress among others. When the P-value was <0.05, distribution map of significant functions was performed with the negative logarithm of P value; the abscissa was -LogP, and the ordinate was significant GO. In this figure, the closer it was to the top, the more significant the GO function was.

STC-GO analysis showed that the genes involved in this trend are associated with a variety of biological functions, including: hyphal growth, pseudohyphal growth, budding cell bud growth, cell-matrix adhesion, biofilm formation, mRNA polyadenylation, and response to osmotic stress among others. Other significant gene expression trends were also related to different biological processes. The 12 STC-GO analysis outcomes are provided in Supplemental Figure 2. The genes with significant expression tendencies may have important roles in the reproduction and pathogenesis of *C. albicans*.

Signaling pathway analysis and GCN analysis

We found 17 significant signaling pathways (Figure 5). These pathways were involved in glycerophospholipid metabolism, and amino sugar and nucleotide sugar metabolism among others. Based on the similarity of gene expression, we constructed GCN to analyze the interaction between

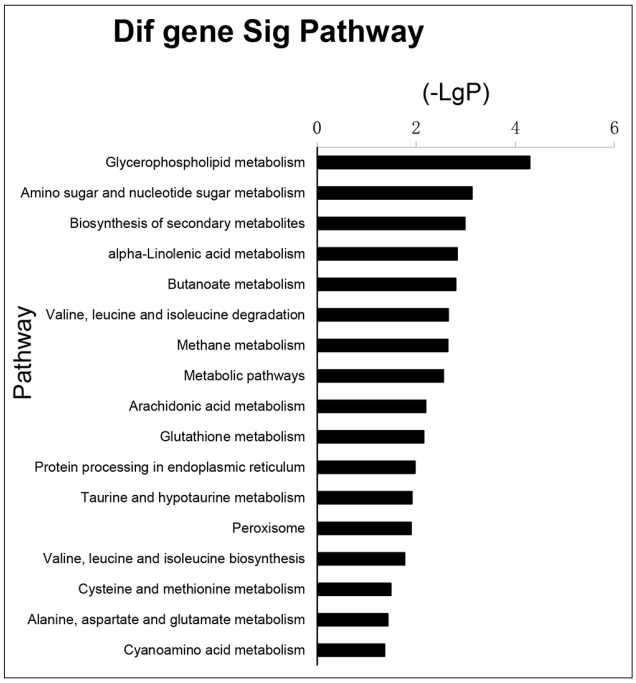


FIGURE 5. Analysis of signaling pathways. Overall, 17 significant signaling pathways with differentially expressed genes were observed, and these included glycerophospholipid metabolism, amino sugar and nucleotide sugar metabolism, biosynthesis of secondary metabolites, and alpha-linolenic acid metabolism.

these genes. The GCN contained 557 gene probes (Figure 6). According to the degree, clustering coefficient and k-core, 33 genes were in the center of the network; among these genes, we speculated that *SCW11*, *ALG7*, *HYR3*, *KEL1*, *LAT1*, *CRH11*, *FGR41*, and *CHT3* genes have very important functions in *C. albicans* (Table 1), because they acted as bridges to other genes. Moreover, *AYR2*, *IPF168*, *CPY1*, *RAC1*, *CSH1*, *FGR24*, *ALG7*, *HYR3*, *PHO15*, *IPF2137*, *KEL1*, *SHM2*, *CRH11*, *TRP99*, *PRC2*, *UBP15*, *LYS7*, *RRD1*, *APE3*, *SPT6*, *IDP1*, and *IFF6* genes may also have important functions, even though they were not in the center of the network. These genes showed a higher complexity in terms of gene-gene associations and interactions with other genes.

Thirty-four genes from the intersection set were included in the GCN. After eliminating unknown gene fragments, we found that *CHT3*, *SHM2*, *PDA1*, *SDH41*, *ADE13*, *LAT1*, *IDP1*, *SER33*, *VMA2*, *HEM12*, *MUQ1*, and *HXK1* genes from the intersection set may have an important role in *C. albicans* (Figure 7A). Initially, 191 genes from the union set were included in the GCN. After eliminating genes for which names were not assigned, the GCN of the union set contained 134 genes. Among 134 genes, *PDA1*, *ZRT2*, *CHS8*, *KEL1*, *ALG7*, *IPF168*, *HOG1*, *LAT1*, *DCW1*, and *HXK1* were in the center of the GCN and possibly have important roles in *C. albicans* (Figure 7 and 7B). Furthermore, after we excluded genes for which names were not assigned, the GCN analysis showed that *GLC7*, *APR1*, *IPF168*, *TRP99*, *TUP1*, *CPY1*, *APE3*, and *CHS1* genes may also have an important role in *C. albicans*.

DISCUSSION

C. albicans primarily exists in three forms: yeast (blastospore), pseudohyphae, and hypha. The fungus switches between these forms depending on the environmental cues. The hyphal form is always pathogenic, while the yeast form is parasitic and not infectious [13]. *C. albicans* can rapidly adapt to changing environment by altering the expression of genes. The products of these genes are involved in budding, proliferation, adhesion, colonization, persistence, and changes in virulence among other processes [14,15].

Based on the previous research [16,17], three different substages are assumed in the pathogenesis of *C. albicans*: an early or colonization phase (0-3 hours), which is characterized by the proliferation and adhesion of fungus; an invasion phase (6-12 hours), which is associated with hyphal formation and penetration; and a late phase (12-24 hours), associated with substantial tissue destruction. These stages were demonstrated in our infection model, by infecting cervical cancer-derived epithelial HeLa cells with a typical strain of *C. albicans* (SC5314) (Figure 1). To identify *C. albicans* genes that are associated with the pathogenicity of this fungus, we examined gene expression profiles of *C. albicans* (SC5314) cells at 1, 3, 6, 12, and 24 hours following the infection of the Hella cells, using gene chip analysis. Then, through a series of biological and bioinformatic analyses, we investigated potential roles of differentially expressed genes in the pathogenesis of candidiasis, as well as interactions between these genes.

The gene expression trend analysis showed 80 theoretical gene expression tendencies. Out of the 80 expression tendencies, 12 gene expression tendencies were considered significant, and similar results were obtained with the STC-GO analysis. The genes with significant expression trends were then analyzed and identified using the KEGG and GCN. Genes that were in the core of the network had a top k-core level, degree ≥10, and the clustering coefficient ≥0.05. The products of these core genes are associated with hyphal growth, biofilm formation, cell-matrix adhesion, drug resistance, and response to osmotic and oxidative stress, among other processes.

C. albicans can reversibly transform into yeast and hyphal forms. The hyphal phase is considered to be the primary *C. albicans* form of infection [18,19]. In our experiment, most of the genes that are related to yeast-hyphae switching [20] (i.e., *CHS1, MUQ1, CHT3, CPY1, FGR41*, and *RHD3* among others) were downregulated, which may be important during the yeast-hyphae switching phase.

In the natural environment, the cell wall of *C. albicans* encloses the fungal cells and acts as a skeletal support system that allows for interactions with the surrounding environment. In our experiment, several structure-related genes found in the key locations of the network, may be important in the

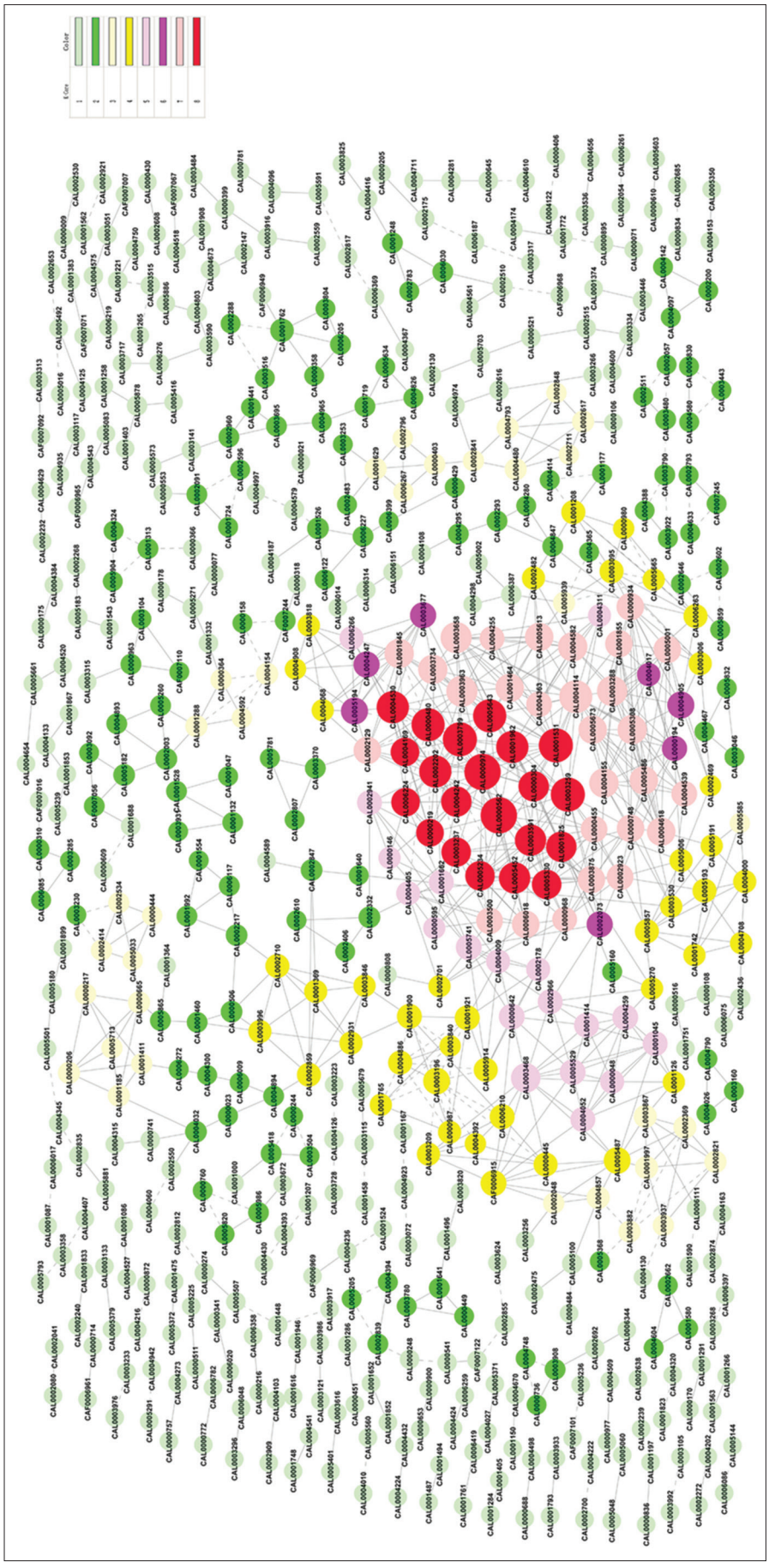


FIGURE 6. Gene Co-expression Network (GCN) analysis of genes from 12 significant differential expression trends. The GCN contained 557 genes and the results varied in relation to the infection time points. The nodes represent genes and the line segments represent the interactions between genes. The solid lines represent positively correlated genes, which means that the products of one or more genes can promote the expression of other genes. The dashed lines represent negatively correlated genes, meaning mutual inhibition between genes and their products. The colors represent the value of k-core, and k-cores of the same value reflect a similarity and indicate a correlation in gene functions. The complexity of gene relationships increases with k-core value rank. The identically colored nodes are from the same subgraph.

TABLE 1. Genes that may have important functions according the Gene Co-expression Network analysis

Gene	Gene name	Description	Clustering coefficient	Degree	k-core
CAL0002202	*SCW11*	Cell wall protein; transcription is decreased in mutant lacking *ACE2*; caspofungin repressed; transcriptionally regulated by iron; expression greater in high-iron medium	0.55	16	8
CAL0004530		Predicted open reading frame (ORF) from Assembly 19	0.51666667	16	8
CAL0003591		Predicted ORF from Assembly 19; hyphal induced	0.56190476	15	8
CAL0005452	*ALG7*	Protein involved in cell wall mannan biosynthesis; transcription is elevated in chk1, nik1, and sln1 homozygous null mutants	0.55238095	15	8
CAL0000304	*HYR3*	Putative GPI-anchored protein of unknown function; similar to Hyr1p; transcriptionally regulated by iron; expression greater in high-iron medium	0.57142857	14	8
CAL0004242		Predicted ORF from Assembly 19	0.56043956	14	8
CAL0005234	*KEL1*	Protein containing a Kelch-repeat domain, similar to *Saccharomyces cerevisiae (S. cerevisiae)* Kel1p; not required for adherence to buccal epithelial cells or virulence in mouse systemic infection; null mutant colonies exhibit slightly decreased filamentation ratio	0.62820513	13	8
CAL0005330		Predicted ORF from Assembly 19	0.62820513	13	8
CAL0003237	*LAT1*	Predicted ORF from Assembly 19	0.53846154	13	8
CAL0004169	*CRH11*	Putative ortholog of *S. cerevisiae* Crh1p; predicted glycosyl hydrolase domain; similar to Csf4p, to *S. cerevisiae* Utr2p, and to antigenic *Aspergillus fumigatus Aspf9*; predicted Ke×2p substrate; putative GPI-anchor; localizes to cell wall	0.66666667	12	8
CAL0006224	*FGR41*	Putative GPI-anchored protein; transposon mutation affects filamentous growth	0.60606061	12	8
CAL0000219	*CHT3*	Chitinase, major; functional homolog of S. cerevisiae Cts1p; 4 N-glycosylation motifs; possible O-mannosylated region; putative signal peptide; hyphal-repressed; farnesol upregulated in biofilm; regulated by Efg1p, Cyr1p, Ras1p	0.72727273	11	8

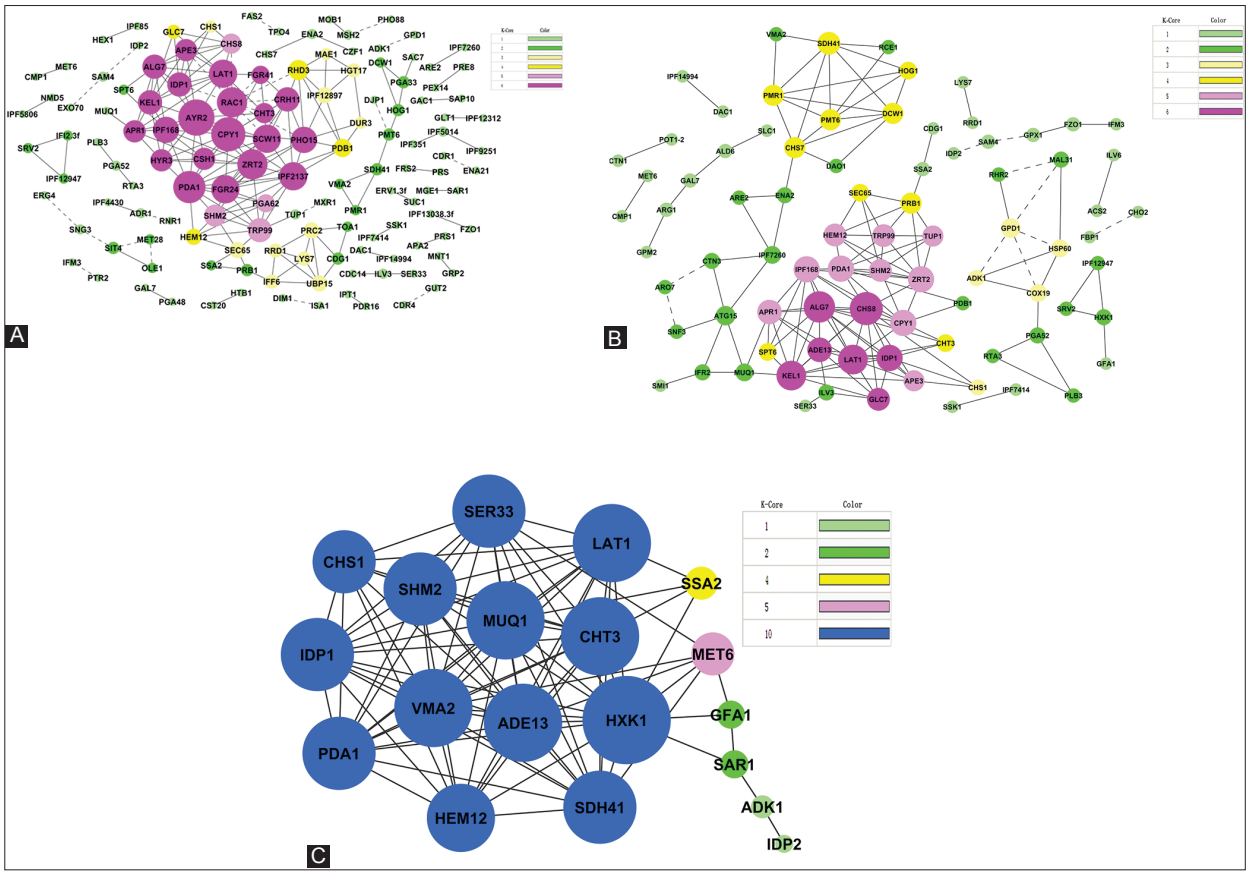


FIGURE 7. (A) The co-expression network with excluded genes for which names were not assigned. Some genes corresponded to multiple probes, some probes did not have a gene name; for a more intuitive analysis of a correlation between genes, we excluded the genes for which names were not assigned and reconstructed the co-expression network. This network contained 134 genes, where *PDA1* and *ZRT2* genes appeared to have important roles in *C. albicans*. (B) Gene Co-expression Network (GCN) analysis of genes from the intersection set. According to the *GCN, CHT3, SHM2, PDA1, SDH41, ADE13, LAT1, IDP1, SER33, VMA2, HEM12, MUQ1,* and *HXK1* genes may have an important role in *C. albicans*. (C) Gene Co-expression Network (GCN) analysis of genes from the union set. *PDA1, ZRT2, CHS8, KEL1, ALG7, IPF168, HOG1, LAT1, DCW1,* and *HXK1* genes were in the center of the GCN, indicating their possible important roles in *C. albicans.*

pathogenesis of *C. albicans*. Among these genes, *ALG7*, *SSK1* and *SCW11* code for structural proteins of the cell wall [21,22]; the products of *CHT2*, *FGR41*, *RHD3*, *HYR3*, and *DCW1* genes belong to glycosylphosphatidylinositol (GPI)-anchored proteins (GAPs), which can attach covalently to the cell wall or to the plasma membrane.

Chitin is synthesized by a family of enzymes encoded by chitin synthase (CHS) genes and is regulated by an array of gene products that regulate CHS activation and localization [23]. Dunkler et al. [24] found that the product of *CHT3* gene may be attributed to the main chitinase activity in the growth medium [24]. In our study, the *CHT3* gene was down-regulated, which may indicate its role in the pathogenesis of *C. albicans*, but additional studies are required to confirm this result.

Cell surface hydrophobicity (CSH) contributes to the virulence of *C. albicans* and can be conveniently regulated in planktonic cultures by altering the growth temperature [25]. The product of *CSH1* gene was the first protein to be implicated in the CSH in *C. albicans*; a 75% loss of CSH is observed when the *CSH1* is knocked out [26]. In our study, the expression of *CSH1* was significantly upregulated, indicating that this gene may be involved in the pathogenicity of *C. albicans*.

Three nicotinamide adenine dinucleotide phosphate (NADP+)-specific isocitrate dehydrogenases (IDPs) have been reported, which are, in *Saccharomyces cerevisiae*, located in different cellular compartments [27]. IDP1 is localized in the mitochondria, IDP2 in the cytoplasm, and IDP3 is required to supply fatty acid β-oxidation with NADPH in the cytoplasm. The *IDP1* and *IDP2* were both occupying vital positions in the GCN of our intersection set, indicating that the *IDP* genes are involved in the pathogenesis of *C. albicans*, by affecting the metabolic activity of the fungus.

Other genes that were significant according to the GCN analysis are associated with metabolic enzymes, and these included *ADE13*, *HXK1*, *MET6*, *PDA1*, *SHM2*, *SER33*, *MUQ1*, and *CPY1*. In addition, the *ALS2* gene, which is related to biofilm formation, adhesion ability, and destruction of reconstituted human epithelium (RHE) [28,29] was also significant according to our GCN.

CONCLUSION

Our screening of differentially expressed genes using STC and STC-GO methods revealed important genes that are possibly involved in *C. albicans* infections. Some of our results are in agreement with previous studies, but we also obtained results that are different from those reported previously. The genes that change the expression at a single time point were outside the scope of this paper. Nevertheless, some important genes were not even identified as differentially expressed in our experiments (e.g., genes related to hyphae formation, such as *ADH1* and *YWP1*); these time-dependent genes may not correlate with the time points examined in our study, or these genes may have a very short transient expression, and thus we could not detect their expression at the selected time points. Additionally, the magnitude of expression of these genes at the RNA level may be below the detection level of our microarray analysis, or important changes may occur at the posttranscriptional level. Finally, this study was conducted *in vitro*, meaning that the expression of *C. albicans* genes in an infected host may be different, as a result of the host immune response to the infection. In the present study, some genes that may be important in the virulence of *C. albicans* have not been analyzed thoroughly, hence, additional studies are required in this field.

ACKNOWLEDGEMENTS

This work was supported by two grants, 2009C33037 grant from the Zhejiang Province Science and Technology Project of P.R. China and 2012BAI32B08 grant from The Key Projects in the National Science & Technology Pillar Program during the Twelfth Five-year Plan Period of P.R. China.

DECLARATION OF INTERESTS

The authors declare no conflict of interests.

REFERENCES

[1] Sobel JD. Vulvovaginal candidosis. Lancet 2007;369(9577):1961-71. https://doi.org/10.1016/S0140-6736(07)60917-9.

[2] Sobel JD. Management of patients with recurrent vulvovaginal candidiasis. Drugs 2003;63(11):1059-66. https://doi.org/10.2165/00003495-200363110-00002.

[3] Sudbery P, Gow N, Berman J. The distinct morphogenic states of Candida albicans. Trends Microbiol 2004;12(7):317-24. https://doi.org/10.1016/j.tim.2004.05.008.

[4] Kabir MA, Hussain MA. Human fungal pathogen Candida albicans in the postgenomic era: An overview. Expert Rev Anti Infect Ther 2009;7(1):121-34. https://doi.org/10.1586/14787210.7.1.121.

[5] Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, et al. The diploid genome sequence of Candida albicans. Proc Natl Acad Sci U S A 2004;101(19):7329-34 https://doi.org/10.1073/pnas.0401648101.

[6] Cheng G, Wozniak K, Wallig MA, Fidel PJ, Trupin SR, Hoyer LL. Comparison between Candida albicans agglutinin-like sequence gene expression patterns in human clinical specimens and models of vaginal candidiasis. Infect Immun 2005;73(3):1656-63 https://doi.org/10.1128/IAI.73.3.1656-1663.2005.

[7] Fradin C, De Groot P, MacCallum D, Schaller M, Klis F, Odds FC, et al. Granulocytes govern the transcriptional response, morphology and proliferation of Candida albicans in human blood. Mol Microbiol 2005;56(2):397-415. https://doi.org/10.1111/j.1365-2958.2005.04557.x.

[8] Clarke R, Ressom HW, Wang A, Xuan J, Liu MC, Gehan EA, et al. The properties of high-dimensional data spaces: Implications for exploring gene and protein expression data. Nat Rev Cancer

2008;8(1):37-49.
https://doi.org/10.1038/nrc2294.

[9] Ramoni MF, Sebastiani P, Kohane IS. Cluster analysis of gene expression dynamics. Proc Natl Acad Sci U S A 2002;99(14):9121-6. https://doi.org/10.1073/pnas.132656399.

[10] Lamanna AC, Karbstein K. An RNA conformational switch regulates pre-18S rRNA cleavage. J Mol Biol 2011;405(1):3-17. https://doi.org/10.1016/j.jmb.2010.09.064.

[11] Yi M, Horton JD, Cohen JC, Hobbs HH, Stephens RM. WholePathwayScope: A comprehensive pathway-based analysis tool for high-throughput data. Bmc Bioinformatics 2006;7:30. https://doi.org/10.1186/1471-2105-7-30.

[12] Nikiforova VJ, Willmitzer L. Network visualization and network analysis. Experientia Suppl 2007;97:245-75. https://doi.org/10.1007/978-3-7643-7439-6_11.

[13] Kumamoto CA, Vinces MD. Contributions of hyphae and hypha-co-regulated genes to Candida albicans virulence. Cell Microbiol 2005;7(11):1546-54. https://doi.org/10.1111/j.1462-5822.2005.00616.x.

[14] Calderone RA, Fonzi WA. Virulence factors of Candida albicans. Trends Microbiol 2001;9(7):327-35. https://doi.org/10.1016/S0966-842X(01)02094-7.

[15] Naglik J, Albrecht A, Bader O, Hube B. Candida albicans proteinases and host/pathogen interactions. Cell Microbiol 2004;6(10):915-26. https://doi.org/10.1111/j.1462-5822.2004.00439.x.

[16] Wilson D, Thewes S, Zakikhany K, Fradin C, Albrecht A, Almeida R, et al. Identifying infection-associated genes of Candida albicans in the postgenomic era. Fems Yeast Res 2009;9(5):688-700. https://doi.org/10.1111/j.1567-1364.2009.00524.x.

[17] Zakikhany K, Naglik JR, Schmidt-Westhausen A, Holland G, Schaller M, Hube B. In vivo transcript profiling of Candida albicans identifies a gene essential for interepithelial dissemination. Cell Microbiol 2007;9(12):2938-54. https://doi.org/10.1111/j.1462-5822.2007.01009.x.

[18] Ramage G, VandeWalle K, Lopez-Ribot JL, Wickes BL. The filamentation pathway controlled by the Efg1 regulator protein is required for normal biofilm formation and development in Candida albicans. Fems Microbiol Lett 2002;214(1):95-100. https://doi.org/10.1111/j.1574-6968.2002.tb11330.x.

[19] Whiteway M, Oberholzer U. Candida morphogenesis and host-pathogen interactions. Curr Opin Microbiol 2004;7(4):350-7. https://doi.org/10.1016/j.mib.2004.06.005.

[20] de Boer AD, de Groot PW, Weindl G, Schaller M, Riedel D, Diez-Orejas R, et al. The Candida albicans cell wall protein Rhd3/Pga29 is abundant in the yeast form and contributes to virulence. Yeast 2010;27(8):611-24. https://doi.org/10.1002/yea.1790.

[21] Cappellaro C, Mrsa V, Tanner W. New potential cell wall glucanases of Saccharomyces cerevisiae and their involvement in mating. J Bacteriol 1998;180(19):5030-7.

[22] Bensen ES, Filler SG, Berman J. A forkhead transcription factor is important for true hyphal as well as yeast morphogenesis in Candida albicans. Eukaryot Cell 2002;1(5):787-98. https://doi.org/10.1128/EC.1.5.787-798.2002.

[23] Henar VM, Duran A, Roncero C. Chitin synthases in yeast and fungi. Experientia Suppl 1999;87:55-69.

[24] Dunkler A, Walther A, Specht CA, Wendland J. Candida albicans CHT3 encodes the functional homolog of the Cts1 chitinase of Saccharomyces cerevisiae. Fungal Genet Biol 2005;42(11):935-947. https://doi.org/10.1016/j.fgb.2005.08.001.

[25] Hazen BW, Hazen KC. Isolation of hydrophobic and hydrophilic variants of Candida albicans. Fems Microbiol Lett 1989;48(2):167-71. https://doi.org/10.1111/j.1574-6968.1989.tb03293.x.

[26] Singleton DR, Fidel PJ, Wozniak KL, Hazen KC. Contribution of cell surface hydrophobicity protein 1 (Csh1p) to virulence of hydrophobic Candida albicans serotype a cells. Fems Microbiol Lett 2005;244(2):373-7. https://doi.org/10.1016/j.femsle.2005.02.010.

[27] Contreras-Shannon V, McAlister-Henn L. Influence of compartmental localization on the function of yeast NADP+ - specific isocitrate dehydrogenases. Arch Biochem Biophys 2004;423(2):235-46. https://doi.org/10.1016/j.abb.2003.12.038.

[28] Zhao X, Oh SH, Yeater KM, Hoyer LL. Analysis of the Candida albicans Als2p and Als4p adhesins suggests the potential for compensatory function within the Als family. Microbiology 2005;151(Pt 5):1619-30. https://doi.org/10.1099/mic.0.27763-0.

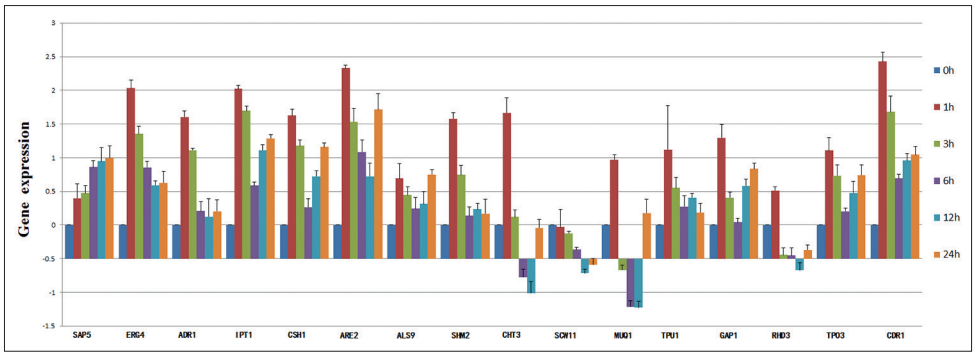
[29] Green CB, Cheng G, Chandra J, Mukherjee P, Ghannoum MA, Hoyer LL. RT-PCR detection of Candida albicans ALS gene expression in the reconstituted human epithelium (RHE) model of oral candidiasis and in model biofilms. Microbiology 2004;150(Pt 2):267-75. https://doi.org/10.1099/mic.0.26699-0.

SUPPLEMENTAL TABLE

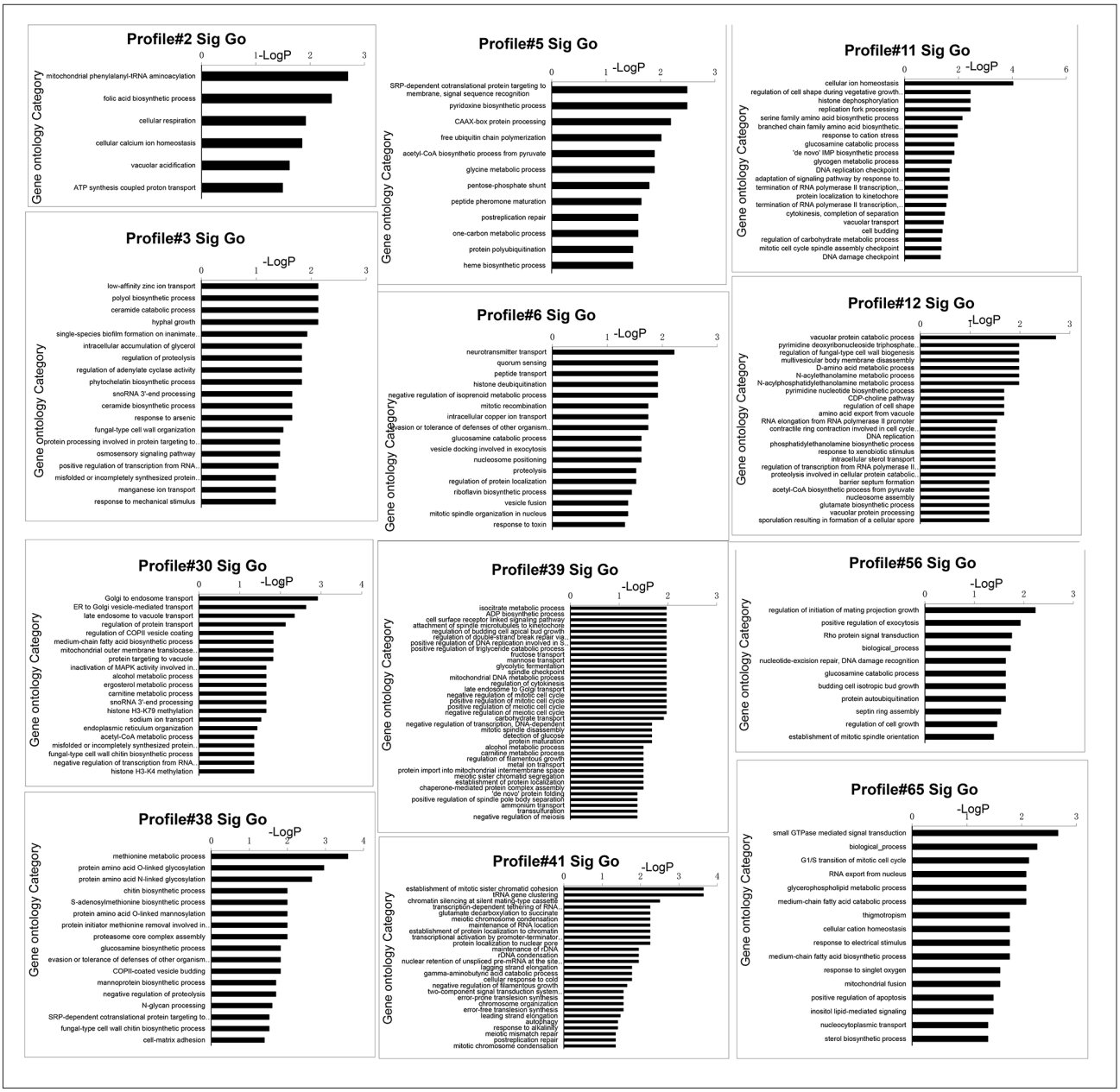
SUPPLEMENTAL TABLE 1. Primers used for quantitative reverse transcription PCR (RT-qPCR)

Serial number	Gene name	Upstream primers (5' to 3')	Downstream primers (5' to 3')
1	18s rRNA	TGACGCAATCGGCACCTT	GCACCACCACCCACAAAATC
2	*SAP5*	CCAGCATCTTCCCGCACTT	GGTAGCTTCGCCGCTTTGA
3	*CSH1*	CAAGTTTCATTGGCATGGTGT	TCTTTGGCGTGATATGGCTC
4	*ERG4*	ATCGTCAAAAGAACAGTTTCAGAA	CGTACCACCCATCAGTCAACA
5	*ARE2*	GGGGAGAACAAATACACTGGAC	TGCTTCCTCAATCGCAACTC
6	*ALS9*	ATCATCGGGAAACTCGCATAA	CGTGTAAGATACCCTTGGAACGT
7	*ADR1*	TGCCAGGAGGTTACTCGTTT	GTGAAGCACCACTAGGATGTGA
8	*IPT1*	AGGTCATATCACTGTGCCCATTA	CAGCGGCATAACCTGGAGTA
9	*SCW11*	TGCTGCTGGTGTTGATTCTATT	GATAAGTGGCAGGTGGTTCAG
10	*CHT3*	GATGATGCCACTGCCAAACA	TCGGGATAAGGACATTGAGGA
11	*MUQ1*	CAAAGACCATCACTATCCAGCAA	ATCAAAGCCACCGTCAATGTAG
12	*TUP1*	CAAAGAGGCATACGAGGAAGA	GTCGATAACCGATAATGATGTGA
13	*GAP1*	GGGTGTAAGATGGTATGGTGAAG	CACTTGTTTAATGGCTTTGGGTA
14	*RHD3*	ATCCATACCGTTACTCTGAAAGC	ACATGACTAATCCAGCAACAACA
15	*TPO3*	CCTGTGCCTTGTCTCCAAAT	TGCAATAAACCCATTGACGA
16	*CDR1*	GCCATGACTCCTGCTACCG	ACCTGGACCACTTGGAACAT

SUPPLEMENTAL FIGURES



SUPPLEMENTAL FIGURE 1. Validation of mRNA expression. As shown in the bar chart, the mRNA expression was in accordance with the gene chip results and differential expression tendencies.



SUPPLEMENTAL FIGURE 2. Series tests of cluster (STC) - Gene Ontology (GO) analysis of 12 significant gene expression tendencies. The STC-GO analysis showed associations between different genes and biological functions, apart from the biological functions presented in the profile 3.