

Identification of Infection- and Defense-Related Genes via a Dynamic Host-Pathogen Interaction Network Using a *Candida Albicans*-Zebrafish Infection Model

Zong-Yu Kuo^a Yung-Jen Chuang^b Chun-Cheih Chao^b Fu-Chen Liu^b
Chung-Yu Lan^c Bor-Sen Chen^a

^aLaboratory of Control and Systems Biology, Department of Electrical Engineering, ^bDepartment of Medical Science and Institute of Bioinformatics and Structural Biology, and ^cInstitute of Molecular and Cellular Biology, National Tsing Hua University, Hsinchu, Taiwan, ROC

Key Words

Host-pathogen interaction network · Infection · Hyphal development · Dynamic protein-protein interaction network · Immune response · Host defense

Abstract

Candida albicans infections and candidiasis are difficult to treat and create very serious therapeutic challenges. In this study, based on interactive time profile microarray data of *C. albicans* and zebrafish during infection, the infection-related protein-protein interaction (PPI) networks of the two species and the intercellular PPI network between host and pathogen were simultaneously constructed by a dynamic interaction model, modeled as an integrated network consisting of intercellular invasion and cellular defense processes during infection. The signal transduction pathways in regulating morphogenesis and hyphal growth of *C. albicans* were further investigated based on significant interactions found in the intercellular PPI network. Two cellular networks were also developed corresponding to the different infection stages (adhesion and invasion), and then compared with each other to identify proteins from which we can gain more insight into the pathogenic role of hyphal development in

the *C. albicans* infection process. Important defense-related proteins in zebrafish were predicted using the same approach. The hyphal growth PPI network, zebrafish PPI network and host-pathogen intercellular PPI network were combined to form an integrated infectious PPI network that helps us understand the systematic mechanisms underlying the pathogenicity of *C. albicans* and the immune response of the host, and may help improve medical therapies and facilitate the development of new antifungal drugs.

Copyright © 2013 S. Karger AG, Basel

Introduction

Candida albicans is an opportunistic fungal pathogen responsible for various mucosal infections, such as candidiasis (e.g. oral thrush and vaginitis) and other potentially life-threatening diseases [1]. *C. albicans* is also the species most frequently responsible for hospital-acquired fungal infections. This pathogen can colonize various biomaterials, such as ventricular assist devices and urinary and vascular catheters, forming dense biofilms that are resistant to most antifungal drugs [2]. *C. albicans* infections and candidiasis are difficult to treat and cre-

ate very serious therapeutic challenges. Mortality rates among patients with candidiasis can be as high as 40–60%, especially for those who have bloodstream infections [3]. Therefore, an understanding of the molecular mechanisms underlying the pathogenicity of *C. albicans* and of host defense systems could improve medical therapy and facilitate the development of new antifungal drugs.

Under normal circumstances, *C. albicans* lives in 80% of the human population with no harmful effects, although its overgrowth results in candidiasis, often observed in immunocompromised (e.g. HIV-positive) individuals [4, 5]. *C. albicans* can grow in a variety of morphological forms, ranging from yeast form to pseudohyphae form to true tubular hyphae form, depending on the growth conditions in the host environment [6]. A number of molecular factors have been implicated as associated with the virulence of *C. albicans*, such as host recognition biomolecules, secreted aspartyl proteases and phospholipases, as well as life cycle factors like adhesion and morphogenesis [7]. Among those factors, the transition from yeast to hyphal form is considered to be critical for *C. albicans* pathogenesis. The ability of *C. albicans* to form hyphae has been proposed as a virulence factor, as these structures are often observed in invaded tissue and *C. albicans* strains unable to form hyphae (whether naturally or through introduced mutations) show defective infectivity [6, 8]. Although previous studies have provided some insights, the details of the molecular mechanisms responsible for morphological forms still await elucidation. In this study, we utilized time series microarray data over nine time points to construct two dynamic networks, which represent protein-protein interaction (PPI) in the adhesion stage (i.e. when hyphae are not growing) and the invasion stage (i.e. when they are). By comparing these two dynamic networks, we can investigate the details of molecular mechanisms responsible for changes in *C. albicans* infectivity across morphological forms.

At present, we still lack sufficient high-throughput screening data for *C. albicans* such as PPI and ChIP-chip data, even though the genome for *C. albicans* has been identified, sequenced and released to aid research on this significant pathogen [9]. The *C. albicans* genome for strain SC5314 has already been sequenced, revealing that almost two-thirds of its ~6,000 open reading frames are orthologous to genes of *Saccharomyces cerevisiae*, the most intensively studied eukaryotic model organism to have its entire genome sequenced [10, 11]. The identification also revealed gene orthologs between *C. albicans* and *S. cerevisiae*. Unlike *C. albicans*, *S. cerevisiae* does not

form true hyphae and is generally not considered a human pathogen. *S. cerevisiae* has abundant high-throughput screening data and it is closely related to *C. albicans* (i.e. both fall within the class Hemiascomycetes), the information from *S. cerevisiae* could be usefully adapted for our understanding of *C. albicans* biology and pathogenesis [10].

The zebrafish (*Danio rerio*) has emerged as a powerful new vertebrate model for human disease. Numerous studies have already utilized the zebrafish system to study the pathogenesis of various human infectious diseases, including those caused by bacteria or viruses [12, 13]. The zebrafish immune system displays remarkable similarities to mammalian counterparts. As a demonstration of the zebrafish's utility as a model organism for human disease, in 49 cases of a zebrafish mutant gene being cloned based on a forward genetic screen, the genes were found to have homologs in human disease [14]. Overall, the zebrafish genetic map demonstrates highly conserved synteny with the human genome [15]. Chao et al. [16] have also demonstrated that *C. albicans* can colonize and invade the fish host at multiple anatomical sites and prove fatal in a dose-dependent manner. Therefore, a zebrafish infection model could be used to investigate the details of the *C. albicans* invasive process and infectious mechanisms.

In this study, we construct an infectious *C. albicans* and zebrafish intercellular PPI network by mining and integrating microarray data, PPI information and host-pathogen intercellular interactions in order to investigate how morphology regulates the infectious behavior of *C. albicans* on host tissue. Consequently, we discovered that all major hyphae-related pathways are visible in our hyphal PPI network, confirming the reliability and accuracy of our methods and results. From a systems perspective, we were able to predict the proteins with the largest changes in the number of interactions and the hub proteins for morphological switching processes. We identified several important hyphae growth-related proteins – e.g. Ubi4, Act1, Kex2, Hsl1 and Tsa1 – and some proteins worth further exploration for pathogenicity research such as Hht21, Kre1 and Orf19.5438. Moreover, three noteworthy functions at work in *C. albicans* infection – cellular iron ion homeostasis, glucose transport and cell wall molecular biosynthesis – were named as pathogen invasion mechanisms from analysis of the integrated intercellular protein interaction networks. Furthermore, several functions, such as apoptosis and immune response, were also found to be involved in host defense mechanisms.

Materials and Methods

Ethics Statement

Manipulation of the animal model was approved by the Institutional Animal Care and Use Committee of National Tsing Hua University.

Simultaneous Time-Course Microarray Experiment during *C. Albicans* Infection

SC5314 strain *C. albicans* and adult AB strain zebrafish were used for the experiments. Their maintenance and preparation were performed according to procedures described previously [16]. Zebrafish were anesthetized by immersion in water containing 0.17 g/ml of Tricaine (Sigma) and then intraperitoneally injected with 1×10^8 CFU *C. albicans* cells suspended in 10 μ l sterile phosphate-buffered saline. The infected fish were sacrificed by immersion in ice water at 0.5, 1, 2, 4, 6, 8, 12, 16 and 18 h postinjection and frozen in liquid nitrogen. *C. albicans*-infected zebrafish were treated with Trizol[®] Reagent (Invitrogen, Carlsbad, Calif., USA), pulverized in liquid nitrogen using a small mortar and pestle, and then disrupted using a MagNA Lyser System (Roche) with glass beads (cat. No. G8772-100G, Sigma) by shaking at 5,000 rpm for 15 s. After phase separation by adding chloroform, the total RNA was purified using an RNeasy Mini Kit (Qiagen, Hilden, Germany). Purified RNA was quantified at OD260nm using an ND-1000 spectrophotometer (NanoDrop Technology, Wilmington, Del., USA) and analyzed using a Bioanalyzer 2100 (Agilent Technologies, Santa Clara, Calif., USA) with RNA 6000 Nano LabChip kit (Agilent Technologies). One microgram of the total RNA was amplified using a Quick-Amp Labeling kit (Agilent Technologies) and labeled with Cy3 (CyDye, PerkinElmer, Waltham, Mass., USA) during the in vitro transcription process. 0.625 μ g of Cy3 cRNA for the *C. albicans* array and 1.65 μ g of Cy3 cRNA for the zebrafish array were fragmented to an average size of 50–100 nucleotides by incubation with fragmentation buffer at 60°C for 30 min. The fragmented labeled cRNA was then hybridized to both *C. albicans* and zebrafish oligo microarrays (Agilent Technologies) at 60°C for 17 h. After washing and drying using a nitrogen gun, microarrays were scanned using an Agilent microarray scanner (Agilent Technologies) at 535 nm for Cy3. For each time point, three biological replicates were done for both organisms. The raw data were processed with Loess normalization and the results have been deposited in the Gene Expression Omnibus (accession No. GSE32119; since the data are currently private, reviewers can access the data using the following link: <http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=llqtncceackuwilu&acc=GSE32119>).

Overview of the Process

The global screening method for infection-related proteins was divided into three key steps: (i) data selection and processing for proteins, (ii) constructing dynamic hyphal growth networks for *C. albicans* and dynamic networks for *D. rerio* and (iii) constructing an intercellular PPI network between pathogen and host. The flowchart of the integrated cellular network construction is shown in figure 1. After constructing the overall network – which consists of the hyphal growth PPI network for *C. albicans*, the PPI network for *D. rerio* and the intercellular PPI network between pathogen and host – we searched for potential infection-related proteins and immune response pattern recognition molecules in both *C. albicans* and zebrafish.

Data Selection and Preprocessing

In this study, several types of data were mined and integrated to construct the integrated cellular network. In *C. albicans*, the required data included some of its microarray gene expression profiles, PPIs from *S. cerevisiae*, data on gene orthologs between *C. albicans* and *S. cerevisiae*, and gene annotations for *C. albicans*. There are 9 time points in the *C. albicans* microarray data spanning from 0.5–18 h postinfection (i.e. 0.5, 1, 2, 4, 6, 8, 12, 16 and 18 h). The gene ortholog data were acquired from the *Candida* Genome Database, and the *C. albicans* gene annotations were retrieved from the Gene Ontology (GO) project. The PPI data for *S. cerevisiae* were extracted from the Biological General Repository for Interaction Datasets (BioGRID).

In zebrafish, the data included microarray gene expression profiles, PPIs from *Homo sapiens*, data on human and zebrafish gene orthologs and functional gene annotations for zebrafish. There were 9 time points in the zebrafish microarray data spanning from 0.5 to 18 h (i.e. 0.5, 1, 2, 4, 6, 8, 12, 16 and 18 h). The gene ortholog data were acquired using the ZFIN, InParanoid, and BLAST databases. The zebrafish gene annotations were retrieved from the GO project. The PPI data for *H. sapiens* were extracted from BioGRID and the Human Protein Reference Database.

Selection of Protein Pool for Rough PPI Networks

Because of the lack of PPI databases between *C. albicans* and zebrafish at present, gene ortholog data between *C. albicans* and *S. cerevisiae* as well as between zebrafish and *H. sapiens* were utilized to set up protein data pools for our candidate *C. albicans* and zebrafish PPI networks, respectively. *C. albicans* PPIs can be inferred by applying ortholog data between *C. albicans* and *S. cerevisiae* to the latter's PPI data; similarly, zebrafish PPIs can be inferred by mapping *H. sapiens* PPI data to ortholog data between humans and zebrafish. Then, we set up the protein pool consisting of differentially expressed proteins. Since large-scale protein activity measurements are unobtainable, mRNA expression profiles were used as a substitute instead. Although the mRNA expression levels cannot be completely representative of the corresponding protein expression levels, they are at least partially and positively correlated [17, 18]. The mRNA expression level for each protein was used to filter differentially expressed proteins using one-way analysis of variance (ANOVA), where the null hypothesis was the average mRNA expression levels at every time point being equivalent. In *C. albicans* and zebrafish alike, the proteins with p values returned by ANOVA that were less than 0.01 were added to the protein pool. In this manner, we selected 4,820 and 9,665 proteins for inclusion in the protein pools of *C. albicans* and zebrafish, respectively. In this step, we found that a set of 4,820 proteins is too large for constructing the PPI network for *C. albicans*. Because the high resultant PPI numbers exceed the size of the microarray data, we narrowed the protein pool of *C. albicans* to avoid over fitting in the parameter identification process for the PPI network construction. So, utilizing the GO database to further select a hyphal growth protein pool within the 4,820 proteins set, we constructed a hyphal PPI network for *C. albicans* consisting of a subset of 403 proteins identified as related to hyphal growth. In addition, we were able to locate the beginning of hyphal growth in the body of the zebrafish at 2–4 h postinfection from microscopy images of the experiment (fig. 2). Therefore, we selected 598 additional proteins of which mRNA levels changed by more than two-fold in 1–6 h to another hyphal growth protein pool. Most of these 598 proteins had not yet

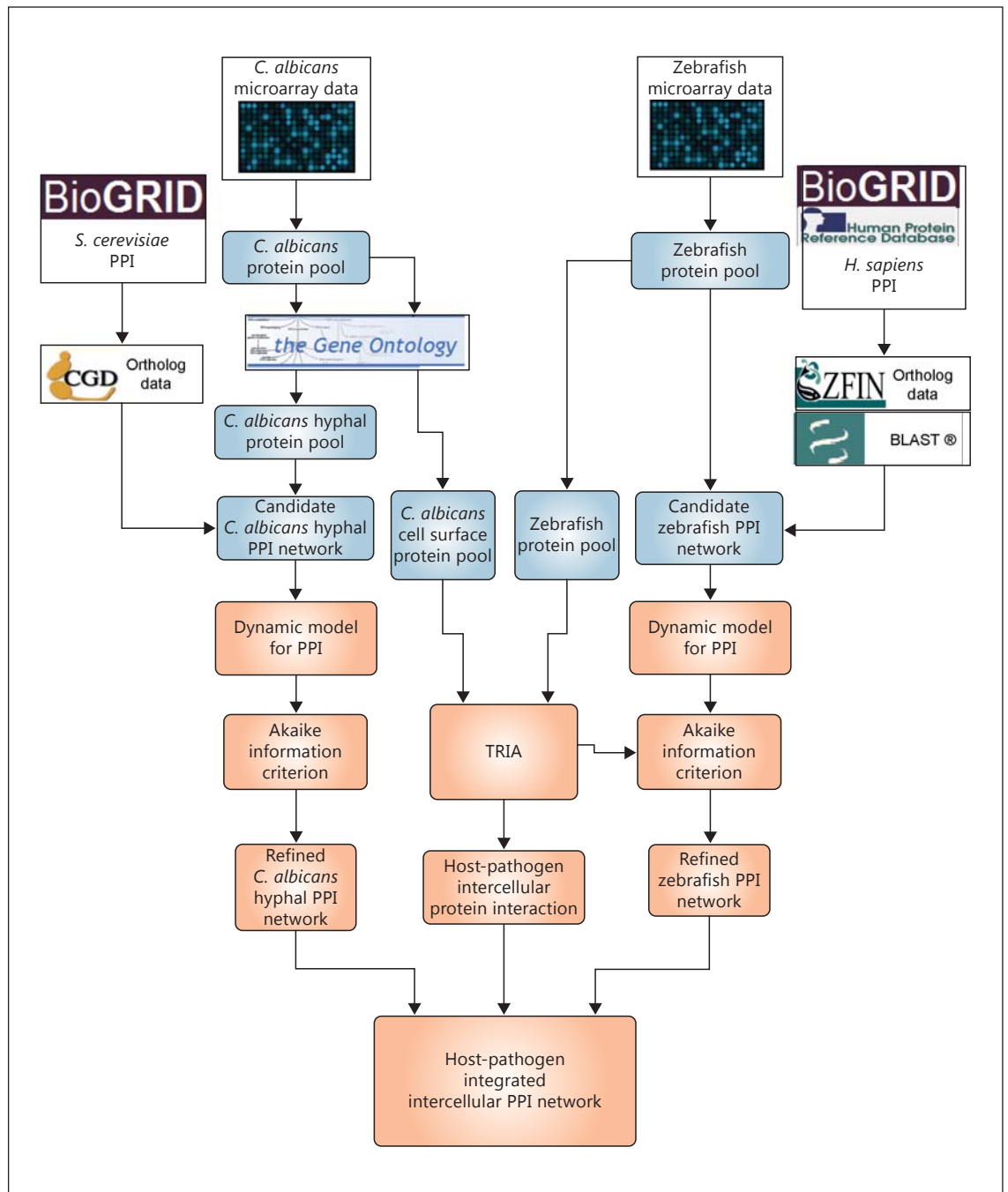


Fig. 1. Flowchart of the construction of the integrated infection intercellular PPI network via database mining and integration. The construction of our integrated intercellular PPI network was performed by database mining and network identification. The network construction combines DNA microarray data with different

types of information from various databases, as shown in the white boxes. The blue boxes show the steps of candidate subnetwork construction. The bottom part (orange boxes) of the flowchart illustrates the steps of network identification and the subsequent construction of the integrated cellular network.

been confirmed as associated with hyphal growth. Combining the 403 hyphae-related proteins with the 598 proteins having a more than two-fold change in expression levels yielded 1,001 proteins for the total hyphal growth protein pool. A candidate PPI network could be constructed based on this protein pool and PPI information. Since candidate PPI networks contain many false positive PPIs, the candidate PPI network was pruned using real-time series microarray data and based on a dynamic interaction model in the following subsection.

Dynamic Model for Organism Protein Interaction Network during Infection

The candidate PPI network can be depicted as a dynamic system in which interactive proteins and mRNA are considered as inputs of the system and protein activities as outputs of the system. All proteins in the candidate PPI network can be considered as target proteins. For a target protein p in the candidate PPI network with N interacting proteins, a dynamic model of the protein's activity can be represented as follows:

$$y_p[t+1] = y_p[t] + \sum_{q=1}^{Q_p} b_{pq} y_q[t] y_p[t] + \alpha_p x_p[t] - \beta_p y_p[t] + \omega_p[t]$$

for $p = 1, 2, \dots, N$ (1)

where $y_p[t]$ represents the activity level of p at time t , b_{pq} denotes the interaction ability of the q -th interactive protein to p , $y_q[t]$ represents the protein activity level of the q -th protein interacting with p , α_p denotes the translation rate from mRNA to protein, $x_p[t]$ represents the mRNA expression level of p , β_p indicates the decay rate of the protein, and $\omega_p[t]$ is stochastic noise. The PPI rate is proportional to the product of two proteins' concentrations [19] (i.e. proportional to the probability of molecular collisions between two proteins), and thus the interaction is modeled as a nonlinear multiplication scheme. The biological interpretation of equation 1 is that the protein activity level of target protein p at time $t + 1$ is a function of the present protein activity level plus regulatory interactions with Q_p interactive proteins, plus additive translation effects from mRNA, minus present protein degradation effects and plus some stochastic noise. Because of the undirected nature of protein interactions, we did not assign direction for a two-protein interaction in the PPI subnetwork in equation 1. After the dynamic interaction model for the p -th protein is constructed as in equation 1, the interaction parameters b_{pq} , translation parameter α_p and decay rate β_p can be estimated from microarray data in the following subsection (see online suppl. methods for more details; for all online suppl. material, see www.karger.com/doi/10.1159/000347104). Since the number of interactions in a candidate PPI network varies in the literature, dependent on the biological situation or condition targeted by a study, there exist many false positives and several interactions may not be relevant for our purposes. Therefore, the estimated interaction parameters b_{pq} should be pruned using the model order selection method Akaike information criterion (AIC), which is detailed next.

Determination of Significant Interaction Pairings

When the regulatory interaction parameters b_{pq} have been identified, AIC [20] is then employed for both model order selection and determination of significant interactions in the infection PPI networks: i.e. to determine the number of interactions Q_p in equation 1. The AIC, which attempts to include both the estimated

residual error and model complexity in one statistical measure, decreases as the residual variance decreases and increases as the number of interactions (i.e. complexity) increases [21].

$$AIC(Q_p) = 2\log \varepsilon_p + \frac{2Q_p}{L}, \quad \text{where } \varepsilon_p = Y_p - \Phi_p \hat{\theta}_p \quad (2)$$

As the expected residual error decreases with increasing interactions for inadequate model complexities, there should be a minimum located near the correct interaction number [20, 21]. Therefore, AIC can be used to select model order (i.e. the number of interactions) based on the protein interaction abilities b_{pq} identified above. In other words, we use the AIC model order selection method to reduce the likelihood of false positive PPIs from the candidate network using time profile microarray data to achieve a more realistic PPI network. After constructing the PPI networks for host and pathogen, we constructed a network for the protein interactions between pathogen and host to gain more insight into the offensive and defensive schemes of pathogen and host during the infection process.

Construction of an Intercellular PPI Network between Pathogen and Host

To identify the intercellular PPIs between pathogen and host during infection of zebrafish with *C. albicans*, we utilized the Temporal Relationship Identification Algorithm (TRIA), which uses gene expression data to identify a given transcription factor's regulatory targets from its binding targets as inferred from ChIP-chip data [22]. The first step was to build a pool of *C. albicans* cell surface proteins. We used the GO database to select 195 cell surface proteins from the 4,031-protein pool to build the resultant protein pool for *C. albicans*. Because host resistance against *C. albicans* infections is mediated predominantly by phagocytes, namely neutrophils and macrophages [23, 24], we assumed that cell surface proteins of *C. albicans* may interact with any protein of zebrafish in the infectious process. So, we let $\bar{x} = (x_1, \dots, x_N)$ denote the gene expression time profile of *C. albicans* cell surface protein x and $\bar{y} = (y_1, \dots, y_N)$ denote the gene expression time profile of zebrafish protein y . We constructed the protein interactions between *C. albicans* and zebrafish via cross-correlation calculations of their time series microarray data.

We compute the cross-correlation between \bar{x} and \bar{y} with a lag of k time points as follows:

$$c(k) = \left(\sum_{i=1}^{N-k} (y_{i+k} - \bar{y})(x_i - \bar{x}) \right) / \left(\sqrt{\sum_{i=1}^{N-k} (y_{i+k} - \bar{y})^2} \sqrt{\sum_{i=1}^{N-k} (x_i - \bar{x})^2} \right),$$

$k = 0, 1, \dots, T$ (3)

where

$$\bar{y} \triangleq \left(\sum_{i=1}^{N-k} y_{i+k} \right) / (N-k),$$

$$\bar{x} \triangleq \left(\sum_{i=1}^{N-k} x_i \right) / (N-k) \quad (4)$$

and T is the maximal time lag after the *C. albicans* infection. We interpolated the 9 time points available for both *C. albicans* and zebrafish into 36 time points. The interval between each time point was 0.5 h. In this study, we set $T = 8$, meaning that we computed the cross-correlation between a *C. albicans* cell surface protein and a zebrafish protein for all possible time lags less than 4 h. Although

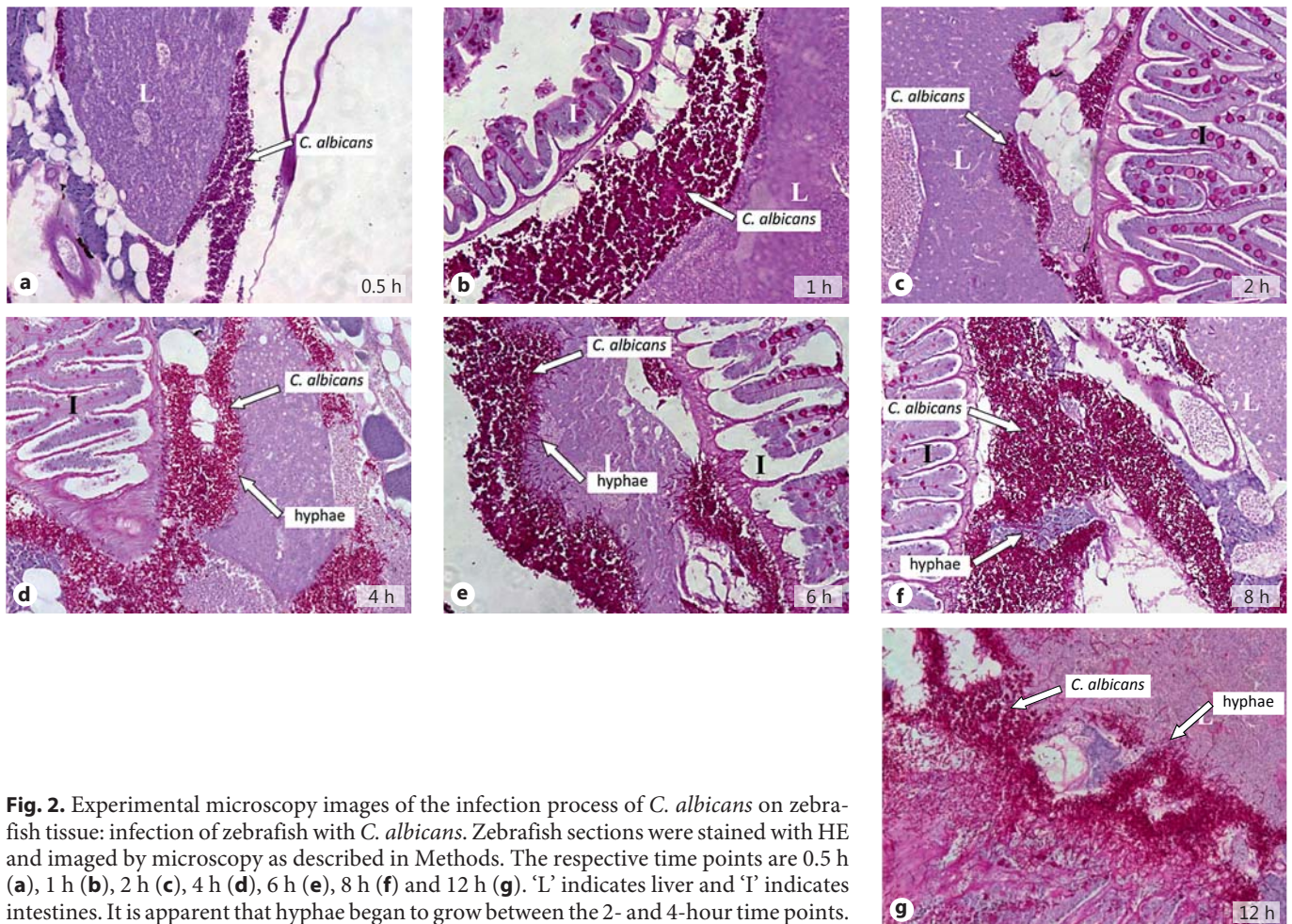


Fig. 2. Experimental microscopy images of the infection process of *C. albicans* on zebrafish tissue: infection of zebrafish with *C. albicans*. Zebrafish sections were stained with HE and imaged by microscopy as described in Methods. The respective time points are 0.5 h (a), 1 h (b), 2 h (c), 4 h (d), 6 h (e), 8 h (f) and 12 h (g). 'L' indicates liver and 'I' indicates intestines. It is apparent that hyphae began to grow between the 2- and 4-hour time points.

the beginning of hyphal growth in the body of the zebrafish occurs at 2–4 h postinfection, we assumed the hyphae-related proteins of *C. albicans* might influence zebrafish proteins ahead of 4 h postinfection. Then, we tested the null hypothesis (H_0): $c(k) = 0$ (i.e. the cell surface proteins of *C. albicans* and zebrafish proteins are uncorrelated) and the alternative hypothesis (H_a): $c(k) \neq 0$ by the bootstrap method [25] to obtain a p value $p(k)$ (see online suppl. methods for details). After all cross-correlations were calculated, we set the constraint that cross-correlation levels must be higher than 0.95. The PPIs satisfying this constraint were considered as potential intercellular PPIs between *C. albicans* and zebrafish.

Results

Construction of the Integrated Intercellular PPI Network during Infection

This study aimed to construct the integrated intercellular interaction network between the hyphal proteins of *C. albicans* and zebrafish proteins during the infection

process. The flowchart detailing its construction is shown in figure 1 and has three main routes, among which two separately construct the hyphal PPI network of *C. albicans* and the PPI network of zebrafish. The third constructs the host-pathogen intercellular PPI network. Based on the microarray data, we selected 4,820 and 9,665 proteins for inclusion in the source protein pools of *C. albicans* and zebrafish, respectively. In addition, we selected 1,002 proteins for inclusion in the hyphal growth protein pool from the *C. albicans* protein pool due to the need to investigate what factors are behind the transition from yeast form to hyphal form in the infection process. In the candidate *C. albicans* hyphal PPI network, there were 3,604 PPIs; in the candidate zebrafish PPI network, there were 1,129.

We utilized the 9 time point *C. albicans* time series microarray data to construct two dynamic networks for different infection stages. Since hyphae appear to begin to

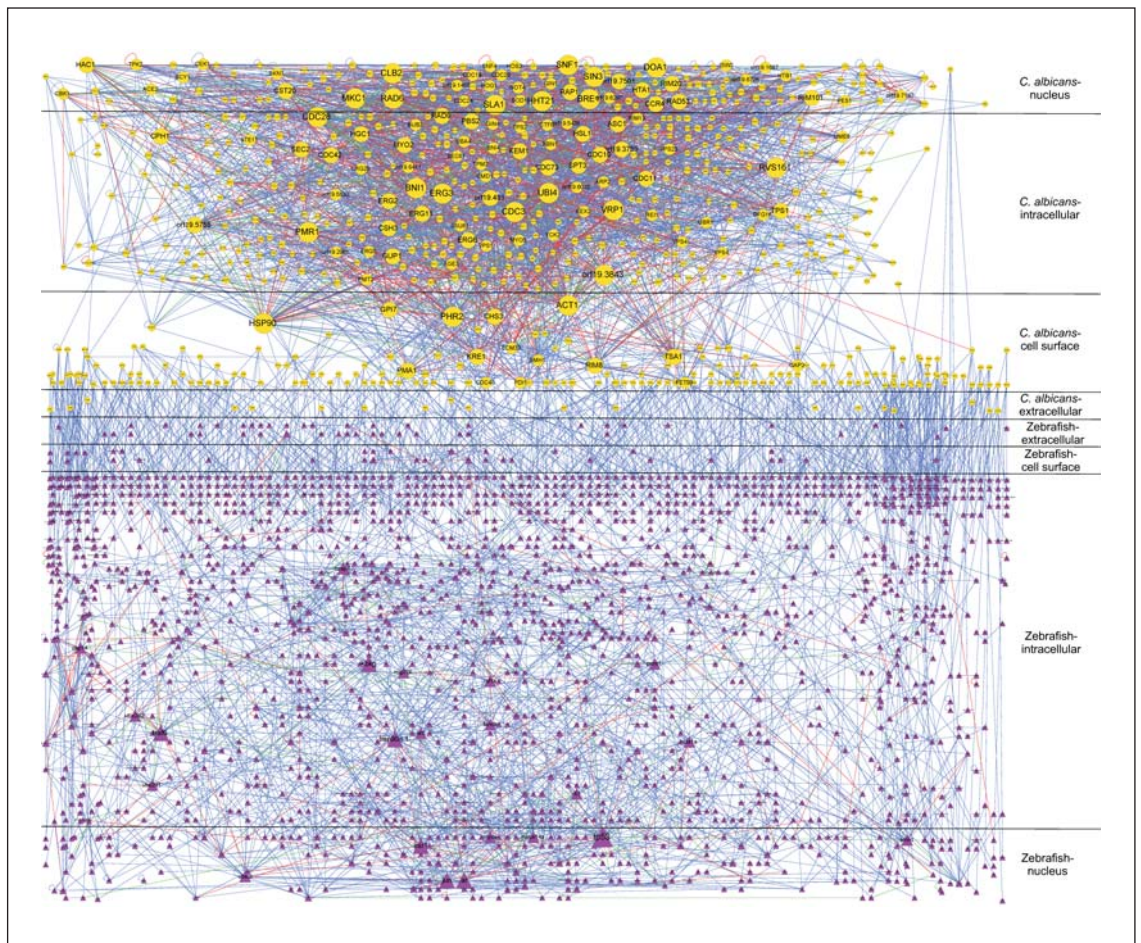


Fig. 3. *C. albicans* and zebrafish integrated intercellular dynamic PPI network during *C. albicans* infection of zebrafish: the infectious intercellular network is composed of three subnetworks. The upper subnetwork is the dynamic hyphal PPI network of *C. albicans*. The middle subnetwork shows the host-pathogen intercellular interaction network. For simplicity, only the top five correlated interactions of the *C. albicans* cell surface proteins are listed. The bottom subnetwork is the dynamic defensive protein interac-

tion network of zebrafish. This infectious intercellular PPI network contains lines and nodes of three different colors. The red lines denote PPIs that did not appear in the stage 1 network but did in the stage 2 network. The green lines denote PPIs that appeared in the stage 1 network but did not in the stage 2 network. The blue lines denote PPIs that appeared in both the stage 1 and 2 networks. The node size denotes connectivity degree. The drawing of the PPI network was created using Cytoscape.

grow in the zebrafish body from 2–4 h postinfection in the experimental microscopy images (fig. 2), we collected two groups of data at different stages of infection to construct two separate networks. With the *C. albicans* microarray data spanning 0.5–4 h, we constructed a network called the ‘adhesive stage network’, which represents *C. albicans* cells in the adhesion stage. Since cubic spline interpolation requires at least four data points to solve a cubic polynomial [26], we included the 4-hour data point to construct this network. With the *C. albicans* microarray data spanning 2–12 h, we constructed another net-

work called the ‘hyphal stage network’, which represents *C. albicans* cells transitioning to the hyphal form. Similarly, we collected two groups of data at different stages of infection to construct two separate PPI networks for zebrafish as well: one for microarray data from 0.5 to 4 h, and another for data from 2 to 12 h, named the zebrafish stage 1 network and zebrafish stage 2 network, respectively. By estimating the system parameters using the time series microarray data and selecting model order using the AIC measurement [20, 21], the likelihood of false positive interactions in the potential PPI network for the in-

fection process was reduced. Network refinement yielded 550 proteins with 2,725 PPIs in the adhesive stage network and 555 proteins with 3,171 PPIs in the hyphal stage network: these two networks could then be combined into the *C. albicans* dynamic hyphal PPI network for the infection process (online suppl. fig. S1). Similar refinements in the zebrafish data returned 1,248 proteins with 2,344 PPIs in the zebrafish stage 1 network and 1,265 proteins with 2,379 PPIs in the zebrafish stage 2 network, and these two networks could then be combined into the zebrafish dynamic PPI network for the defensive process (online suppl. fig. S2). The *C. albicans* dynamic hyphal PPI network, the zebrafish dynamic PPI network and the host-pathogen intercellular PPI network could be merged into an integrated infection intercellular PPI network.

The global system view of the *C. albicans*- and zebrafish-integrated infection intercellular PPI network is illustrated in figure 3. The entire integrated infection intercellular network can be divided into eight levels according to the location of protein action (i.e. nucleus, intracellular, cell surface or extracellular) and species (i.e. *C. albicans* or zebrafish) and is composed of three subnetworks. The upper subnetwork is the dynamic hyphal PPI network of *C. albicans*. The middle subnetwork shows the host-pathogen intercellular interaction network. For simplicity, only the top five correlated interactions of the *C. albicans* cell surface proteins are listed. The bottom subnetwork is the dynamic defensive protein interaction network of zebrafish.

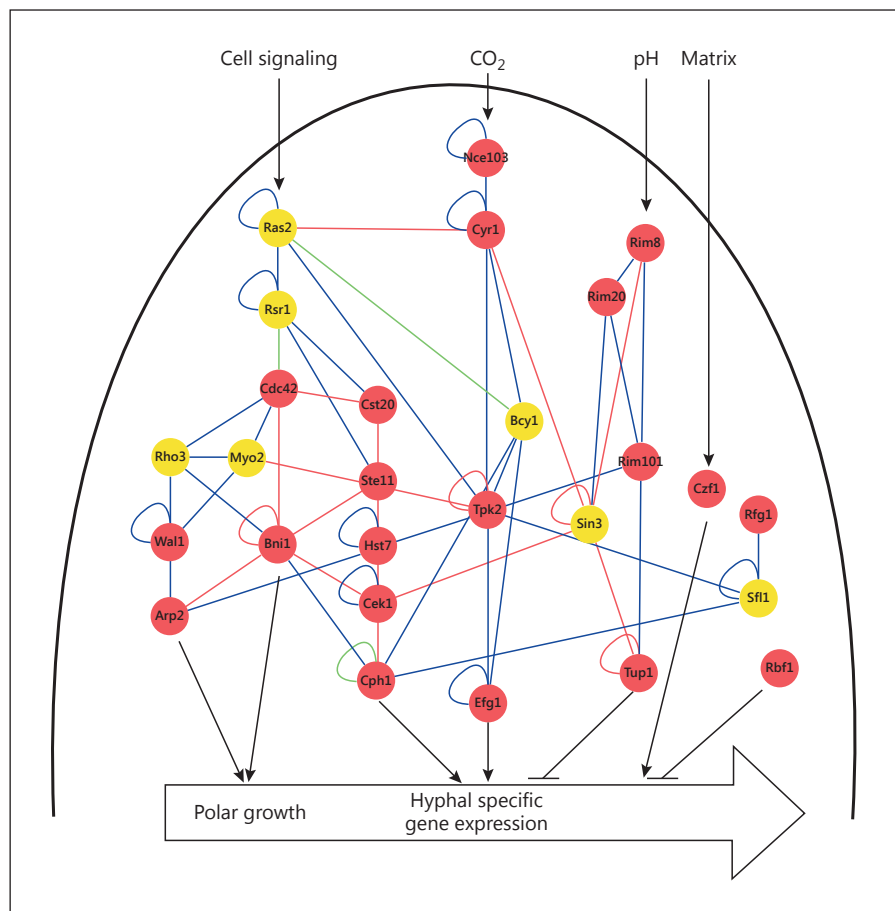
Inspection of the Dynamic Hyphal Growth PPI Network of C. albicans

In order to verify the accuracy of our dynamic hyphal growth protein interaction network, we investigated whether this network contains previously identified pathways related to hyphal growth, which are illustrated in online supplementary figure S3 [27]. This figure displays signal transduction pathways involved in regulating morphogenesis in *C. albicans*. An inspection of the integrated infectious intercellular network seen in figure 3 confirmed that our *C. albicans* dynamic hyphal PPI network includes the MAP kinase cascade, cyclic AMP/PKA pathway and other hyphae-associated pathways. We isolated these pathways from figure 3, and then constructed a new hyphae-related subnetwork as figure 4. It is apparent that this subnetwork is very similar to online supplementary figure S3. Our new hyphae-related subnetwork contains almost all of the proteins and interactions of the already known hyphae-related pathways. The GTP binding protein Ras1 was not contained in our dynamic hyphal

growth PPI network because its p value is greater than 0.01 in the original protein pool selection step for *C. albicans*; however, Ras2, which is in the same family as Ras1, appeared in the new subnetwork. Ras2 is similar to *S. cerevisiae* Ras2p, which can activate adenylate cyclase and is involved in *S. cerevisiae* pseudohyphal growth, and Ras2 mutants could have altered filamentous growth patterns [28]. Similar to Ras1 in online supplementary figure S3, Ras2 also stimulates Cyr1 (Cdc35), which in turn acts as an intracellular second messenger during morphological switching. Ras2 also stimulates Cdc42 through Ras-related protein (Rsr1), which is involved in budding, cell morphogenesis and hyphal development processes [29]. In addition, an interaction between Cdc42 and Wall1 is represented with a dotted line in online supplementary figure S3, meaning this link is not completely known. However, we can see that Cdc42 links to Wall1 via Myo2 and Rho3 in our new hyphae-related subnetwork. Myo2 is required for polarized cell growth and dimorphic switching in *C. albicans* and is also involved in hyphal development [30]. Rho3 is required for polarized cell growth and cell separation and also involved in hyphal development [31]. Fortunately, the previously uncertain pathway between Cdc42 and Wall1 was more completely elucidated in our dynamic hyphal growth PPI network.

Aside from these three well-known hyphae-related signaling pathways – i.e. the MAP kinase and cyclic AMP signaling pathways and the polarized cell growth pathway – the pH-dependent Rim101 pathway was also identified from the dynamic integrated infection intercellular network. In this pathway, Nrg1 was not identified for inclusion in figure 4 due to its p value being greater than 0.01. However, Tup1, which has the same function as Nrg1, fits into the pathways shown in figure 4 and hence the pH-dependent pathway would seemingly be uninterrupted. In conclusion, 20 out of 22 proteins from already known pathways (i.e. those in online suppl. fig. S3) are included in our *C. albicans* dynamic hyphal PPI network, in which the major hyphae-related pathways are all visible. These results verify the high accuracy of our infection intercellular PPI network. Moreover, our hyphal growth protein interaction subnetwork provides a dynamic and more complete hyphal network in comparison with online supplementary figure S3. Specifically, the yellow nodes in figure 4 represent the proteins that are not contained in online supplementary figure S3. These proteins are all related to hyphal growth or filamentous growth, and the figure reflects the true pathways of these proteins.

Fig. 4. Signaling cascades involved in the dynamic hyphal growth protein interaction subnetwork of *C. albicans* at different stages of infection: the MAP kinase signaling pathway, cyclic AMP signaling pathway, polarized cell growth pathway and Rim101 signaling pathway were identified in our dynamic hyphal growth PPI network of *C. albicans* as occurring during *C. albicans* infection of zebrafish. These signal transduction pathways are involved in the dynamic regulation of *C. albicans* morphological transitions. The red lines indicate PPIs that did not appear in the adhesive stage network but did in the hyphal stage network. The green lines indicate PPIs that appeared in the adhesive stage network but did not appear in the hyphal stage network. The blue lines indicate PPIs that appeared in both the adhesive and hyphal stage networks. The red nodes indicate proteins that are also included in online supplementary figure S3; the yellow nodes indicate proteins that are not included in online supplementary figure S3.



Utilization of Dynamic Intracellular PPI Networks to Identify Proteins with Important Roles in Hyphal Development

We utilized the dynamic intercellular PPI networks to investigate which proteins play important roles in hyphal development. Based on the dynamic hyphal PPI network of *C. albicans* during infection (online suppl. fig. S1), we explored proteins of which interactions with other proteins displaying large variations between values in the adhesive and hyphal stage networks. In other words, the number of increased interactions and the number of reduced interactions for each protein would be summed to find out the largest interaction difference between the two stages. Table 1 lists the top 15 proteins by magnitude of their PPI changes. In addition, these 15 proteins almost completely overlap with the hubs in the hyphal stage network (online suppl. table S1). In general, the number of possible interactions for any given protein is paralleled by changes in protein interaction values. Therefore, the variations of the hub protein interactions should be larger

than the conventional nodes in the infectious PPI network.

From online supplementary figure S4 it can be seen that our hyphal growth PPI network is scale free. In a scale-free network, the probability that a node is highly connected is significantly higher than in a random network, and the network's properties are often determined by a relatively small number of highly connected nodes known as hubs [32]. The scale-free networks are particularly resistant to random node removal but extremely sensitive to the targeted removal of hubs [33]. Hence, the hubs are believed to be essential to the robustness of information transmission in the morphological transition from yeast to hyphal form. In our hyphal growth PPI network, Ubi4 is the protein for which the number of interactions had the largest changes in table 1 and it is also the biggest hub in the hyphal stage network (online suppl. table S1). Ubi4 is involved in the negative control of morphological switching in *C. albicans*, as well as in maintaining yeast cell morphology [34]. From the time profile data

Table 1. Proteins with the largest changes in PPI number between the adhesive and hyphal stage networks during infection

<i>C. albicans</i> protein	Increased inter-actions	Reduced inter-actions	Total changed inter-actions	GO functional annotation	
				hyphal growth	filamentous growth
Ubi4	37	23	60		+
Act1	31	12	43	+	
Hsp90	20	15	35		+
Sla1	22	8	30	+	
Bni1	22	7	29	+	+
Sin3	22	6	28		+
Hht21	19	7	26		
Mkc1	16	10	26	+	+
Phr2	17	8	25		+
Pmr1	17	7	24	+	
Rvs161	14	9	23	+	
Rad6	16	6	22		+
Hgc1	15	6	21	+	
Vrp1	16	4	20	+	
Clb2	12	7	19		+

The top 15 *C. albicans* proteins in the hyphal growth PPI networks are ranked by change in the number of protein interactions. The number of increased PPIs of these proteins between the adhesive and hyphal stages is given, along with the number of reduced ones. Total changed interactions displays the sum of PPI changes. GO functional annotation provides the GO terms of proteins, which we have filtered to only list GO terms more concerned with hyphal growth and filamentous growth. The functional annotations are from the GO database (<http://www.geneontology.org/>).

of Ubi4, the expression of Ubi4 was reduced from the 0.5- to 4-hour time points (online suppl. fig. S5). It is clear that Ubi4 expression drops when hyphae began to grow. Act1 is the second-ranked protein in table 1. Act1 is an actin that influences both cAMP synthesis and hyphal morphogenesis [35]. Consequently, Ubi4 and Act1 may play the most important roles in hyphal growth development, because of their high PPI variation between the adhesive and hyphal stage networks. Similarly, other highly ranked proteins in table 1 – Hsp90, Sla1, Bni1, etc. – have also been identified as potentially important in hyphal development [36–38]. In fact, Hht21 is the only protein in table 1 that has not been verified as related to hyphal growth at present. We predict that Hht21 also plays an important role in hyphal development and was a worthwhile protein to identify in the present study.

Subsequently, we wanted to investigate which proteins had many increased interactions but fewer reduced interactions in the adhesive stage network than in the hyphal

stage network. We supposed that these proteins would have a lot of influence on hyphal growth development. In order to discard the many proteins with only minor interaction variations between the networks, the minimum number of changed interactions for a protein was raised to 10. Column 6 of table 2 indicates the ratio of increased interactions to total changed interactions. From table 2, we can see that 13 of the top 15 proteins from the GO database are related to hyphal development – i.e. having GO terms of hyphal growth or filamentous growth. The first-ranked protein is Kex2, which influences *C. albicans* proteinase secretion and hyphae formation. The disruption of Kex2 function in *C. albicans* has pleiotropic effects that may impinge on the ability of the organism to colonize and invade tissues [39]. The second-ranked protein, Hsl1, is a probable protein kinase involved in morphological determination during the cell cycle of both yeast-form and hyphal cells via regulation of Swe1 and Cdc28 [40]. The third-ranked protein is Tsa1, indispensable for the yeast-to-hyphal transition when *C. albicans* is cultured under oxidative stress [41]. The fourth- to seventh-ranked proteins are Cek1, Chs3, Top1 and Orf19.3843, respectively, and are also all related to hyphal development [42–44].

The eighth- and fifteenth-ranked proteins are Kre1 and Orf19.5438, respectively. From the current literature, it is unknown whether Kre1 and Orf19.5438 are related to hyphal growth. However, we predict that Kre1 and Orf19.5438 may have some relationship to hyphal development, because the remaining 13 of the top 15 proteins in table 2 have been confirmed to relate to this process.

From the dynamic PPI network of zebrafish during infection detailed in online supplementary figure S2, we also investigated which proteins have the largest interaction difference between zebrafish stage 1 and stage 2 networks. These protein interaction changes are related to molecular defensive mechanisms activated in response to *C. albicans* invasion. Table 3 lists the top 10 proteins with the largest changes in number of protein interactions. Moreover, we also listed the hub proteins of the zebrafish stage 2 network in online supplementary table S2. Tp53 is the protein with the largest variation in interaction between zebrafish stage 1 and stage 2 networks. Tp53 is already well known as an apoptosis protein [45]. Esr1 is the second-ranked protein in table 3, which is identified as an apoptosis-related protein in the GO database. Traf6 is the third-ranked protein; a recent study using a zebrafish embryo model has analyzed the in vivo function of Traf6 in innate immune response without interference of adaptive immunity [46]. Traf6 can activate the NF- κ B signal transduction pathway in zebrafish [47]. We can see that of the

Table 2. Proportion of increased PPIs to total changed PPIs between the adhesive and hyphal stage PPI network during infection

Ranking	<i>C. albicans</i> protein	Increased interactions	Reduced interactions	Total changed interactions	Ratio of increased interactions	GO functional annotation	
						hyphal growth	filamentous growth
1	Kex2	10	0	10	1	+	
2	Hsl1	15	1	16	0.9375		+
3	Tsa1	10	0	10	0.9167	+	
4	Cek1	9	1	10	0.9	+	
5	Chs3	9	1	10	0.9	+	
6	Top1	9	1	10	0.9	+	
7	Orf19.3843	15	2	17	0.8826		+
8	Kre1	10	2	12	0.8333		
9	Sec2	10	2	10	0.8333	+	
10	Cdc42	14	3	17	0.8235	+	
11	Cst20	14	3	17	0.8235	+	
12	Kem1	14	3	17	0.8235	+	
13	Erg5	9	2	11	0.8182		+
14	Gpi7	9	2	11	0.8182	+	
15	Orf19.5438	9	2	11	0.8182		

The top 15 *C. albicans* proteins in the hyphal growth PPI networks are ranked by the proportion of increased PPIs to reduced PPIs. The number of increased PPIs of these proteins between the adhesive and hyphal stages is given, along with the number of reduced PPIs. Total changed interactions displays the sum of PPI changes. The ratio of increased interactions shows the proportion of increased PPIs to total changed PPIs. GO functional annotation provides the GO terms of proteins, which we have filtered to only list the GO terms more concerned with hyphal growth and filamentous growth. The functional annotations are from the GO database (<http://www.geneontology.org/>).

top 10 proteins in table 3, many are related to innate immune response and apoptosis. It makes sense that the immunization-related proteins of a host would have a greater response to pathogen invasion. Additionally, previous studies have indicated that many bacteria are able to trigger apoptosis in the host cell [48]. Induction of apoptosis in epithelial or endothelial cells might break the epithelia/endothelial cell barrier and permit the bacteria to reach the submucosa. On the other hand, apoptosis might be beneficial for the infected organ, since apoptotic cell death of the infected target cell permits other cells to phagocytose the apoptotic bodies containing bacteria, possibly resulting in the rapid digestion of the pathogen. It is therefore also reasonable that apoptosis-associated protein interactions would have a greater defense response to pathogen invasion.

Intercellular PPI Network between C. Albicans and Zebrafish during Infection

After investigating the dynamic PPI networks of *C. albicans* hyphal growth and zebrafish defense, respectively,

we investigated the intercellular PPI network between *C. albicans* and zebrafish shown in figure 3. The cross-correlation of each protein interaction between *C. albicans* and zebrafish was calculated using TRIA, and the interactions higher than 0.95 were chosen as potential PPIs. In order to find out which *C. albicans* proteins have the greatest impact on zebrafish, we analyzed which cell surface proteins of *C. albicans* (taken from the 4,031-protein set) had the most potential interactions with zebrafish. The top 20 *C. albicans* hub proteins ranked by number of interactions are listed in table 4, with higher-ranked proteins deemed more important in the infection process. These hub proteins could be classified into seven major functional groups according to the GO database, including hyphal growth, cell adhesion, biofilm formation, cellular response to neutral pH, cellular iron ion homeostasis, glucose transport and cell wall molecular biosynthesis. The first four are well known functions that occur during *C. albicans* infection. When *C. albicans* infects a host, it adheres to the host at first, then grows hyphae to invade host and further forms biofilms to parasitize in host. The ability to respond to am-

Table 3. Proteins with the most changes in PPI number between the zebrafish stage 1 and stage 2 networks during infection

<i>D. rerio</i> protein	Increased inter-actions	Reduced inter-actions	Total changed interactions	<i>D. rerio</i> GO annotation	<i>H. sapiens</i> GO annotation
Tp53	10	10	20	apoptosis	apoptosis
Esr1	8	5	13	metal ion binding	regulation of apoptosis
Traf6	6	5	11	innate immune response regulation of apoptosis response to bacterium	apoptosis innate immune response
Jun	3	6	9	canonical Wnt receptor signaling pathway	innate immune response
Ar	5	4	9	metal ion binding	cell death positive regulation of NF-kappaB transcription factor activit
Usp14	5	2	7	ubiquitin-dependent protein catabolic process	ubiquitin thiolesterase activity
Hsp90a.1	5	2	7	myofibril assembly	
Rb1	2	4	6	myoblast differentiation ubiquitin protein ligase binding	negative regulation of cell growth
Mdm2	4	1	5	negative regulation of apoptosis p53 binding	fibroblast growth factor receptor signaling pathway
Casp8	1	3	4	apoptosis proteolysis	activation of innate immune response apoptosis

The names of the top 10 zebrafish proteins in the PPI networks are ranked by change in the number of protein interactions. The number of increased PPIs of these proteins between stages 1 and 2 are given, along with the number of reduced PPIs. The final two columns provide the GO terms of proteins. *D. rerio* proteins were mapped to *H. sapiens* proteins by using ortholog data.

bient pH is also critical to the growth and virulence of *C. albicans*. It is well established that a near-neutral pH (~6.5) favors hyphal development of *C. albicans* in vitro, while a low pH (<6.5) blocks hyphal formation and stimulates growth of the yeast form [49]. It is interesting that iron and glucose and cell wall molecular biosynthesis appear in this important table. Notably, in the last several years iron ion uptake has been thought to be an important factor in pathogenesis. The ability to acquire iron from host tissues is a major contributing factor to the virulence of pathogenic microorganisms. *C. albicans*, like many pathogenic bacteria, is able to utilize heme and hemoglobin as iron sources [50]. The availability of iron can have a significant impact on both pathogen virulence and host antimicrobial defenses. Some studies have shown that pretreatment of endothelial cells with an iron chelator can reduce the damage inflicted by *C. albicans* [51, 52]. Glucose has previously been reported to induce germ-tube formation in *C. albicans* [53, 54], and Paranjape and Datta [55] recognized it as critical for its pH-regulated pathways. Chitin, β -glucan

and mannose are essential molecules constituting most fungal cell walls. It is intriguing that the cell wall proteins of *C. albicans* are predicted by our modeling to have so many interactions with zebrafish proteins. Recent studies have shown that cell wall β -glucan is a key fungal signature molecule targeted by the innate immune system to clear fungal infection and that *C. albicans* masks β -glucan from immune recognition by using a mannoprotein coat [56, 57]. The immune system may also be able to counteract this fungal defense by unmasking the signature components of the fungus during the course of infection. Wheeler et al. [57] showed that β -glucan is initially masked but subsequently exposed on the surface of *C. albicans* in the normal course of infection. Although it is unknown how this unmasking occurs, it is possible that with time immune cells accumulate in sufficient numbers to directly damage the cell wall and expose the β -glucan of *C. albicans*. The chitin of the cell wall may be similarly destroyed such that *C. albicans* needs to synthesize more essential cell wall molecules to protect itself.

Table 4. Numbers of PPIs of *C. albicans* cell surface proteins in the host-pathogen intercellular PPI network

Interactions, n	<i>C. albicans</i> protein	GO functional annotation						
		cell adhesion	hyphal growth	biofilm formation	cellular response to neutral pH	iron ion homeostasis	glucose transport	cell wall molecular biosynthesis
1,447	Csh1	+		+				
1,089	Rbt5			+		+		
1,078	Pga10			+		+		
1,027	Pga7							
964	Csa1			+		+		
747	Mp65	+	+	+	+		+	
728	Hgt2						+	
701	Chs3		+					+
643	Cht2		+					+
532	Chs2							+
532	Hgt9						+	
532	Mnt1	+	+					+
491	Hgt12		+		+		+	+
451	Phr1	+	+		+			+
435	Pga49							
419	Cdc48							
408	Hgt7						+	
367	Als1	+	+	+	+			
357	Tos1							
352	Mid1		+					

The top 20 *C. albicans* hub proteins with the most potential interactions with zebrafish are listed; their cross-correlation values between protein interactions were larger than 0.95. The number of PPIs of *C. albicans* cell surface proteins are sorted in descending order, along with the corresponding protein name.

Csh1 (cell surface hydrophobicity 1) is the largest hub in the intercellular *C. albicans*-zebrafish PPI network. Knockout of the Csh1 gene has been undertaken to address the potential contribution of its antigen in mediating fungal cell adhesion to host tissue [58]. Another study has demonstrated that Csh1 contributes to virulence of *C. albicans* in mice [59]. The fourth hub in table 4 is Pga7, and its function and those of two other proteins (Pga49 and Tos1) are unknown at present. As the top 20 hub-proteins in table 4 have the most potential interactions with zebrafish, any one of these proteins should have a significant impact on zebrafish defense. It would be worthwhile to investigate the functions of these three unknown proteins.

In addition, we have also listed the top 25 zebrafish hub proteins that have the most intercellular interactions with *C. albicans* in online supplementary table S3. There are several proteins related to cell proliferation, blood coagulation and proteolysis. It makes sense that more cells

would need to proliferate to defend and maintain the normal operation of the host body while *C. albicans* invades, and zebrafish may also secrete proteolytic compounds to damage *C. albicans*. Proteolytic cascades also play a crucial role in innate immune response because they can be triggered more quickly than the adaptive immune response, which requires that gene expression be altered [60]. To summarize our findings, table 5 provides the biological functions of *C. albicans* and zebrafish that are observed in the intercellular infection PPI network.

Discussion

To discover *C. albicans* proteins or pattern recognition molecules that play a critical role in the infection of zebrafish, we constructed an intercellular infection network consisting of *C. albicans* dynamic hyphal, zebrafish dynamic and host-pathogen intercellular PPI networks. To

Table 5. Important biological processes of the *C. albicans* invasion mechanism and zebrafish defense mechanism during infection

	<i>C. albicans</i> (invasion mechanism)	Zebrafish (defense mechanism)
Biological process	hyphal growth cell adhesion biofilm formation cellular response to neutral pH cellular iron ion homeostasis glucose transport cell wall molecular biosynthesis	apoptosis innate immune response cell proliferation blood coagulation proteolysis

verify the reliability and accuracy of our methods for constructing the molecular networks and our results, we compared the *C. albicans* dynamic hyphal PPI network with already known pathways implicated in hyphal growth. Twenty out of 22 proteins in already known pathways (online suppl. fig. S3) were included in the *C. albicans* dynamic hyphal PPI network, and all major hyphae-related pathways were visible in our hyphal PPI network. Our results also provided a comprehensive and dynamic PPI network for hyphal development during the infection process, and we additionally found that Ras2, Myo2, Rho3, Rsr1, etc. may play significant roles in hyphal growth processes. Based on the time course information of *C. albicans* microarray data, we were also able to construct two PPI networks that represented prehyphal growth (i.e. adhesion) mechanisms and hyphal growth (i.e. invasion) mechanisms for their respective stages. Using a similar approach for zebrafish, we were able to construct two stages of defense networks corresponding to these two invasion networks of *C. albicans*. The proteins with the largest PPI variation between the two networks were elucidated for both *C. albicans* and zebrafish. In the *C. albicans* hyphal growth network, Ubi4 was the protein with the most interaction changes between the adhesive stage and hyphal stage networks, and was also the biggest hub in the latter. Ubi4 in *C. albicans* is involved in the negative control of morphological switching, as well as in maintaining yeast cell morphology. We additionally listed other proteins with large variations – Act1, Hsp90, Sla1, Bni1, etc. – in table 1, which also appear to be network hubs in online supplementary table S1. We predict that deletion of these proteins in *C. albicans* could strongly impact the robustness of information transmission during the morphological transition from yeast to hyphal form. Hht21 is a noteworthy exception in table 1 and online supplementary table S1 that has not been verified as related to hyphal growth.

In addition, we investigated some proteins with many increased and fewer reduced interactions in the hyphal stage network in comparison with the adhesive stage network. We found that 13 of our top 15 proteins were identified as related to hyphal growth development by the GO database. The top-ranked proteins – such as Kex2, Hsl1, Tsa1, Cek1, Chs3 and Top1 – are all related to hyphal growth, and we predict they are very important in hyphal development. Moreover, whether the eighth- and fifteenth-ranked proteins – Kre1 and Orf19.5438, respectively – are related to hyphal growth is still unknown in the present literature; however, we might predict that Kre1 and Orf19.5438 are indeed related to hyphal development, because the other 13 of the top 15 proteins in table 3 have all been confirmed as such. Interestingly, a recent study of deletion mutants in *S. cerevisiae* revealed that the Kre1-deleted strain significantly suppresses the hyperpseudohyphal phenotype [61]. This result implies that Kre1 may also be related to filamentous growth in *C. albicans*. Due to the limited relevant research on zebrafish PPIs, PPI mapping between zebrafish and *H. sapiens* still remains incompletely characterized, with the variations in protein interaction levels in zebrafish dynamic PPI networks appearing very small in comparison with the *C. albicans* hyphal PPI network. Nevertheless, we were able to list the top 10 proteins having the largest changes in the number of protein interactions, and found them to be mainly related to apoptosis and innate immunity.

The intercellular PPI network between *C. albicans* and zebrafish is also considered in our analysis. The top 20 *C. albicans* hub proteins having the most intercellular interactions with zebrafish proteins are listed in table 4. These hub proteins can be classified into seven major functions according to their annotations in the GO database: hyphal growth, cell adhesion, biofilm formation, cellular response to neutral pH, cellular iron ion homeostasis, glu-

glucose transport and cell wall molecular biosynthesis. The first four are the well-known functions that would occur when *C. albicans* infects its host. Ionic iron is necessary for hyphal growth and its availability has been thought as a factor connected with pathogenesis in the last several years. The ability to acquire iron from host tissues is a major factor affecting the pathogenicity of microorganisms. However, a detailed mechanism for the participation of iron in *C. albicans* infection is still unclear. Glucose has previously been reported to induce germ-tube formation in *C. albicans*. *C. albicans* needs energy and nutrients while it infects its host. To obtain them it may acquire glucose from host tissues and so glucose transport functionality in *C. albicans* might have many intercellular interactions with zebrafish. Moreover, it would be worthwhile to investigate why cell wall molecular proteins have so many interactions with host proteins. Recent studies have shown that the immune system may be able to counteract this fungal defense by unmasking the signature molecular components of the fungus during the course of infection. Although it is not known how this unmasking occurs, it is possible that with time, immune cells accumulate in sufficient numbers to directly damage the cell wall and expose β -glucan of *C. albicans*. The chitin of the cell wall may also be destroyed during this accumulation as well, so *C. albicans* would need to synthesize more cell wall molecules to protect itself.

In this study, our aim was to construct a host-pathogen intercellular integrated PPI network by using the microarray data of *C. albicans* and zebrafish, and then utilize it to predict which proteins play critical roles in the infection process of pathogens and the defense process of

hosts. In conclusion, we identified several important proteins related to *C. albicans* infection such as Ubi4, Act1, Kex2, Hsl1 and Tsa1, and some proteins whose contribution to pathogenicity warrants further investigation, such as Hht21, Kre1 and Orf19.5438. These proteins may exert a tremendous influence on morphological transition of *C. albicans* and hence they may provide useful drug targets for broad-spectrum treatments of *C. albicans* infection. Moreover, three noteworthy functions in *C. albicans* host infection – cellular iron ion homeostasis, glucose transport and cell wall molecular biosynthesis – were also discovered from the perspective of the intercellular PPI network. Several significant proteins related to innate immune and apoptotic function such as Tp53, Esr1 and Traf6 were found in the molecular defensive mechanisms of zebrafish responsive to *C. albicans* invasion. Furthermore, biological processes like apoptosis, innate immune response, cell proliferation, blood coagulation and proteolysis were also found in systematic defensive mechanisms of zebrafish during *C. albicans* infection. We hope that our proposed intercellular protein interaction method implemented through our dynamic host-pathogen interaction scheme may ultimately help provide useful medical therapies and facilitate the development of new antifungal drugs.

Acknowledgements

The work was supported by the National Science Council of Taiwan under grants NSC 100-2745-E-007-001-ASP, NSC 101-2745-E-007-001-ASP (to B.S.C.), NSC 100-2627-B-007-002 (to C.Y.L.) and NSC 100-2627-B-007-003 (to Y.J.C.).

References

- 1 Leroy O, Gangneux JP, Montravers P, Mira JP, Gouin F, Sollet JP, Carlet J, Reynes J, Rosenheim M, Regnier B, et al: Epidemiology, management, and risk factors for death of invasive *Candida* infections in critical care: a multicenter, prospective, observational study in France (2005–2006). *Crit Care Med* 2009; 37:1612–1618.
- 2 Kojic EM, Darouiche RO: *Candida* infections of medical devices. *Clin Microbiol Rev* 2004; 17:255–267.
- 3 Seneviratne CJ, Jin L, Samaranyake LP: Biofilm lifestyle of *Candida*: a mini review. *Oral Dis* 2008;14:582–590.
- 4 Pfaller MA, Diekema DJ: Epidemiology of invasive candidiasis: a persistent public health problem. *Clin Microbiol Rev* 2007;20:133–163.
- 5 Olorode OA, Okpokwasili GC: The efficacy of disinfectants on abattoirs' *Candida albicans* isolates in Niger Delta region. *Mycoses* 2012; 55:323–323.
- 6 Lo HJ, Kohler JR, DiDomenico B, Loebenberg D, Cacciapuoti A, Fink GR: Nonfilamentous *C. albicans* mutants are avirulent. *Cell* 1997; 90:939–949.
- 7 Calderone RA, Fonzi WA: Virulence factors of *Candida albicans*. *Trends Microbiol* 2001; 9:327–335.
- 8 Leberer E, Ziegelbauer K, Schmidt A, Harcus D, Dignard D, Ash J, Johnson L, Thomas DY: Virulence and hyphal formation of *Candida albicans* require the Ste20p-like protein kinase CaClp4p. *Curr Biol* 1997;7:539–546.
- 9 Jones T, Federspiel NA, Chibana H, Dungan J, Kalman S, Magee BB, Newport G, Thorstenson YR, Agabian N, Magee PT, et al: The diploid genome sequence of *Candida albicans*. *P Natl Acad Sci USA* 2004;101:7329–7334.
- 10 Ihmels J, Bergmann S, Berman J, Barkai N: Comparative gene expression analysis by differential clustering approach: application to the *Candida albicans* transcription program. *PLoS Genet* 2005;1:e39.
- 11 Goffeau A, Barrell BG, Bussey H, Davis RW, Dujon B, Feldmann H, Galibert F, Hoheisel JD, Jacq C, Johnston M, et al: Life with 6,000 genes. *Science* 1996;274:546–567.
- 12 Meeker ND, Trede NS: Immunology and zebrafish: spawning new models of human disease. *Dev Comp Immunol* 2008;32:745–757.
- 13 Sullivan C, Kim CH: Zebrafish as a model for infectious disease and immune function. *Fish Shellfish Immunol* 2008;25:341–350.

- 14 Amsterdam A, Hopkins N: Mutagenesis strategies in zebrafish for identifying genes involved in development and disease. *Trends Genet* 2006;22:473–478.
- 15 Postlethwait J, Amores A, Force A, Yan YL: The zebrafish genome. *Methods Cell Biol* 1999;60:149–163.
- 16 Chao CC, Hsu PC, Jen CF, Chen IH, Wang CH, Chan HC, Tsai PW, Tung KC, Wang CH, Lan CY, et al: Zebrafish as a Model Host for *Candida albicans* Infection. *Infect Immun* 2010;78:2512–2521.
- 17 Orntoft TF, Thykjaer T, Waldman FM, Wolf H, Celis JE: Genome-wide study of gene copy numbers, transcripts, and protein levels in pairs of non-invasive and invasive human transitional cell carcinomas. *Mol Cell Proteomics* 2002;1:37–45.
- 18 Newman JRS, Ghaemmaghami S, Ihmels J, Breslow DK, Noble M, DeRisi JL, Weissman JS: Single-cell proteomic analysis of *S-cerevisiae* reveals the architecture of biological noise. *Nature* 2006;441:840–846.
- 19 Alon U: An Introduction to Systems Biology: Design Principles of Biological Circuits. Boca Raton, Chapman & Hall/CRC, 2007.
- 20 Akaike H: New look at statistical-model identification. *IEEE Trans Automat Contr* 1974;19:716–723.
- 21 Johansson R: System Modeling and Identification. Englewood Cliffs, Prentice Hall, 1993.
- 22 Wu WS, Li WH, Chen BS: Identifying regulatory targets of cell cycle transcription factors using gene expression and ChIP-chip data. *BMC Bioinformatics* 2007;8:188.
- 23 Edwards JE Jr, Rotrosen D, Fontaine JW, Haudenschild CC, Diamond RD: Neutrophil-mediated protection of cultured human vascular endothelial cells from damage by growing *Candida albicans* hyphae. *Blood* 1987;69:1450–1457.
- 24 Hummert S, Hummert C, Schroter A, Hube B, Schuster S: Game theoretical modelling of survival strategies of *Candida albicans* inside macrophages. *J Theor Biol* 2010;264:312–318.
- 25 Efron B, Tibshirani R: An introduction to the bootstrap. New York: Chapman & Hall, 1993.
- 26 Dyer SA, Dyer JS: Cubic-spline interpolation: Part 1. *IEEE Instrum Meas Mag* 2001;4:44–46.
- 27 Esser K: The Mycota: A Comprehensive Treatise on Fungi as Experimental Systems for Basic and Applied Research, ed 2. Berlin, Springer, 2008.
- 28 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.
- 29 Yaar L, Mevarech M, Koltin Y: A *Candida albicans* RAS-related gene (CaRSR1) is involved in budding, cell morphogenesis and hypha development. *Microbiology* 1997;143:3033–3044.
- 30 Woo M, Lee K, Song K: MYO2 is not essential for viability, but is required for polarized growth and dimorphic switches in *Candida albicans*. *FEMS Microbiol Lett* 2003;218:195–202.
- 31 Dunkler A, Wendland J: *Candida albicans* RHO-type GTPase-encoding genes required for polarized cell growth and cell separation. *Eukaryotic Cell* 2007;6:844–854.
- 32 Barabasi AL, Oltvai ZN: Network biology: understanding the cell's functional organization. *Nat Rev Genet* 2004;5:101–113.
- 33 Jeong H, Mason SP, Barabasi AL, Oltvai ZN: Lethality and centrality in protein networks. *Nature* 2001;411:41–42.
- 34 Roig P, Gozalbo D: Depletion of polyubiquitin encoded by the UBI4 gene confers pleiotropic phenotype to *Candida albicans* cells. *Fungal Genet Biol* 2003;39:70–81.
- 35 Zou H, Fang HM, Zhu Y, Wang Y: *Candida albicans* Cyr1, Cap1 and G-actin form a sensor/effector apparatus for activating cAMP synthesis in hyphal growth. *Mol Microbiol* 2010;75:579–591.
- 36 Shapiro RS, Uppuluri P, Zaas AK, Collins C, Senn H, Perfect JR, Heitman J, Cowen LE: Hsp90 Orchestrates Temperature-Dependent *Candida albicans* Morphogenesis via Ras1-PKA Signaling. *Curr Biol* 2009;19:621–629.
- 37 Reijnt P, Jorde S, Wendland J: *Candida albicans* SH3-domain proteins involved in hyphal growth, cytokinesis, and vacuolar morphology. *Curr Genet* 2010;56:309–319.
- 38 Rida PCG, Nishikawa A, Won GY, Dean N: Yeast-to-hyphal transition triggers formin-dependent golgi localization to the growing tip in *Candida albicans*. *Mol Biol Cell* 2006;17:4364–4378.
- 39 Newport G, Agabian N: KEX2 influences *Candida albicans* proteinase secretion and hyphal formation. *J Biol Chem* 1997;272:28954–28961.
- 40 Uehara Y, Umeyama T, Kaneko A, Nagai Y, Hanaoka N, Tanabe K, Takano Y, Niimi M: *Candida albicans* protein kinase CaHsl1p regulates cell elongation and virulence. *Mol Microbiol* 2005;55:381–395.
- 41 Shin DH, Jung S, Park SJ, Kim YJ, Ahn JM, Kim W, Choi W: Characterization of thiol-specific antioxidant 1 (TSA1) of *Candida albicans*. *Yeast* 2005;22:907–918.
- 42 Whiteway M, Csank C, Schroppel K, Leberer E, Harcus D, Mohamed O, Meloche S, Thomas DY: Roles of the *Candida albicans* mitogen-activated protein kinase homolog, Cek1p, in hyphal development and systemic candidiasis. *Infect Immun* 1998;66:2713–2721.
- 43 Gow NAR, Brand A, Lee K, Veses V: Calcium homeostasis is required for contact-dependent helical and sinusoidal tip growth in *Candida albicans* hyphae. *Mol Microbiol* 2009;71:1155–1164.
- 44 Jiang WD, Gerhold D, Kmiec EB, Hauser M, Becker JM, Koltin Y: The topoisomerase I gene from *Candida albicans*. *Microbiology* 1997;143:377–386.
- 45 Ghiselli G: SMC3 knockdown triggers genomic instability and p53-dependent apoptosis in human and zebrafish cells. *Mol Cancer* 2006;5:52.
- 46 Stockhammer OW, Rauwerda H, Wittink FR, Breit TM, Meijer AH, Spaink HP: Transcriptome analysis of Tráf6 function in the innate immune response of zebrafish embryos. *Mol Immunol* 2010;48:179–190.
- 47 Phelan PE, Mellon MT, Kim CH: Functional characterization of full-length TLR3, IRAK-4, and TRAF6 in zebrafish (*Danio rerio*). *Mol Immunol* 2005;42:1057–1071.
- 48 Grassme H, Jendrosseck V, Gulbins E: Molecular mechanisms of bacteria induced apoptosis. *Apoptosis* 2001;6:441–445.
- 49 Buffo J, Herman MA, Soll DR: A characterization of Ph-regulated dimorphism in *Candida albicans*. *Mycopathologia* 1984;85:21–30.
- 50 Weissman Z, Kornitzer D: A family of *Candida* cell surface haem-binding proteins involved in haemin and haemoglobin-iron utilization. *Mol Microbiol* 2004;53:1209–1220.
- 51 Fratti RA, Belanger PH, Ghannoum MA, Edwards JE Jr, Filler SG: Endothelial cell injury caused by *Candida albicans* is dependent on iron. *Infect Immun* 1998;66:191–196.
- 52 Almeida RS, Brunke S, Albrecht A, Thewes S, Laue M, Edwards JE, Filler SG, Hube B: The hyphal-associated adhesin and invasin Als3 of *Candida albicans* mediates iron acquisition from host ferritin. *PLoS Pathog* 2008;4:e1000217.
- 53 Vidotto V, Accattatis G, Zhang Q, Campanini G, Aoki S: Glucose influence on germ tube production in *Candida albicans*. *Mycopathologia* 1996;133:143–147.
- 54 Hudson DA, Sascia OL, Sanders RJ, Norris GE, Edwards PJB, Sullivan PA, Farley PC: Identification of the dialysable serum inducer of germ-tube formation in *Candida albicans*. *Microbiology* 2004;150:3041–3049.
- 55 Paranjape V, Datta A: Role of nutritional status of the cell in pH-regulated dimorphism of *Candida albicans*. *FEMS Microbiol Lett* 1991;64:333–336.
- 56 Wheeler RT, Fink GR: A drug-sensitive genetic network masks fungi from the immune system. *Plos Pathogens* 2006;2:328–339.
- 57 Wheeler RT, Kombe D, Agarwala SD, Fink GR: Dynamic, morphotype-specific *Candida albicans* beta-glucan exposure during infection and drug treatment. *PLoS Pathog* 2008;4:e1000227.
- 58 Singleton DR, Masuoka J, Hazen KC: Cloning and analysis of a *Candida albicans* gene that affects cell surface hydrophobicity. *J Bacteriol* 2001;183:3582–3588.
- 59 Singleton DR, Fidel PL Jr, 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;244:373–377.
- 60 Cerenius L, Kawabata SI, Lee BL, Nonaka M, Soderhall K: Proteolytic cascades and their involvement in invertebrate immunity. *Trends Biochem Sci* 2010;35:575–583.
- 61 Ryan O, Shapiro RS, Kurat CF, Mayhew D, Baryshnikova A, Chin B, Lin ZY, Cox MJ, Vizeacoumar F, Cheung D, et al: Global gene deletion analysis exploring yeast filamentous growth. *Science* 2012;337:1353–1356.