

2014-09-18

# Identification and characterization of novel putative virulence factors in *Candida albicans*

Luca Issi

*Worcester Polytechnic Institute*

Follow this and additional works at: <https://digitalcommons.wpi.edu/etd-dissertations>

---

## Repository Citation

Issi, L. (2014). *Identification and characterization of novel putative virulence factors in Candida albicans*. Retrieved from <https://digitalcommons.wpi.edu/etd-dissertations/385>

This dissertation is brought to you for free and open access by [Digital WPI](#). It has been accepted for inclusion in Doctoral Dissertations (All Dissertations, All Years) by an authorized administrator of Digital WPI. For more information, please contact [wpi-etd@wpi.edu](mailto:wpi-etd@wpi.edu).

Identification and characterization of  
novel putative virulence factors in  
*Candida albicans*

by

Luca Issi

A Dissertation

Submitted to the Faculty of

WORCESTER POLYTECHNIC INSTITUTE

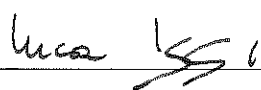
in partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

Biology and Biotechnology

by

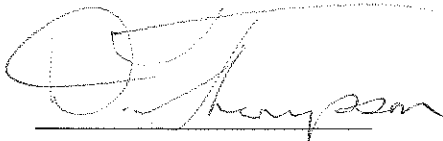


September 2014


Approved by:



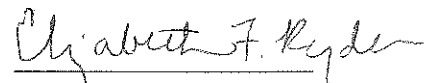
Dr. Reeta Rao



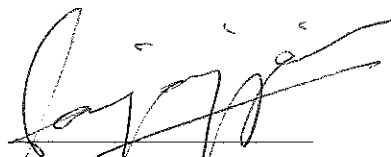
Dr. Dawn Thompson



Dr. Luis Vidali



Dr. Elizabeth Ryder



Dr. Sanjay Jain

## TABLE OF CONTENTS

<b>1</b>	<b>Introduction and background.....</b>	<b>1</b>
1.1	Human fungal infection: a growing threat .....	1
1.2	Types of fungal infections .....	2
1.2.1	Superficial infections .....	2
1.2.2	Mucosal/Invasive infections .....	5
1.3	<i>C. albicans</i> infections: mortality, morbidity and costs .....	9
1.4	<i>C. albicans</i> virulence mechanisms.....	11
1.5	<i>C. albicans</i> therapeutics .....	15
1.5.1	Antifungal drugs .....	15
1.5.2	<i>C. albicans</i> vaccine, promises and challenges .....	21
1.6	Model organisms to study <i>C. albicans</i> infections.....	24
1.6.1	Mammals.....	25
1.6.2	<i>Drosophila melanogaster</i> .....	27
1.6.3	<i>Galleria mellonella</i> .....	29
1.6.4	<i>Danio reio</i> .....	31
1.6.5	<i>Caenorhabditis elegans</i> .....	33
<b>2</b>	<b>A versatile <i>in vivo</i> model to study <i>C. albicans</i> virulence.....</b>	<b>36</b>
2.1	Introduction.....	36
2.2	Materials and methods .....	39
2.2.1	Strains and media.....	39
2.2.2	<i>C. albicans</i> - <i>C. elegans</i> survival assay.....	40
2.3	Results.....	41

## TABLE OF CONTENTS

2.3.1	The <i>C. elegans</i> MAPK pathway is required for resistance to <i>C. albicans</i> .....	41
2.3.2	Genetic and pharmacological modulation of the <i>C. elegans</i> – <i>C. albicans</i> survival assay .....	49
2.4	Discussion .....	51
<b>3</b>	<b>Reverse genetic screen identifies <i>ZCF15</i> a novel virulence factor .....</b>	<b>53</b>
3.1	Introduction.....	53
3.2	Materials and methods .....	55
3.2.1	Strains and media .....	55
3.2.2	<i>C. albicans</i> - <i>C. elegans</i> deformed anal region (Dar) assay.....	59
3.2.3	<i>C. albicans</i> - <i>C. elegans</i> survival assay.....	59
3.2.4	<i>In vitro</i> phenotypic characterization .....	60
3.2.5	<i>Ex vivo</i> macrophage assay .....	61
3.3	Results.....	63
3.3.1	<i>C. albicans</i> reverse genetic in <i>C. elegans</i> identifies four putative virulence factors .....	63
3.3.2	<i>ZCF15</i> is a novel virulence factor with no human homologs .....	66
3.3.3	<i>ZCF15</i> deletion mutant is sensitive to paraquat.....	69
3.3.4	<i>ZCF15</i> is required to withstand <i>C. elegans</i> generated ROS .....	74
3.3.5	<i>ZCF15</i> is required for resistance to murine macrophage killing .....	76
3.3.6	<i>ZCF15</i> deletion has a limited effect in disseminated <i>Candidiasis</i> .....	79
3.4	Discussion .....	81



## TABLE OF CONTENTS

<b>4</b>	<b>ZCF15 and ZCF29 mechanisms of action, some insights from RNASeq and ChIP-Seq.....</b>	<b>83</b>
4.1	Introduction.....	83
4.2	Materials and methods .....	87
4.2.1	Media and growth conditions.....	87
4.2.2	RNA-Seq: RNA extraction and library preparation.....	87
4.2.3	RNA-Seq: sample preparation and sequencing .....	88
4.2.4	ChIP-Seq: Zcf15 and Zcf29 HA tagging .....	89
4.2.5	ChIP-Seq: Immunoprecipitation, library preparation, sequencing and peak calling .....	90
4.3	Results.....	92
4.3.1	<i>In silico</i> analysis of <i>C. albicans</i> (Zn(II) <sub>2</sub> Cys <sub>6</sub> ) transcription factors....	92
4.3.2	ZCF15 and ZCF29 RNA-Seq QC.....	98
4.3.3	ZCF15 and ZCF29 are not regulating canonical genes involved in ROS detoxification .....	107
4.3.4	ZCF29 plays a critical role in N-metabolism.....	110
4.3.4.1	ZCF29 deletion leads to misregulation of N-metabolism .....	110
4.3.4.2	ZCF29 plays a critical role in down regulating ribosome biogenesis upon ROS challenge .....	116
4.3.5	ZCF15 plays an important role in C-metabolism and regulates key thiol peroxidases upon ROS exposure .....	120
4.3.6	Genes involved in iron homeostasis are potentially coregulated by ZCF15 and ZCF29 .....	124

## TABLE OF CONTENTS

4.3.7	<i>ZCF15</i> and <i>ZCF29</i> ChIP-Seq.....	127
4.4	Discussion.....	141
<b>5</b>	<b>Filastatin, a new antimycotic? .....</b>	<b>145</b>
5.1	Introduction and contribution to the paper.....	145
5.2	Chemical screening identifies filastatin, a small molecule inhibitor of <i>Candida albicans</i> adhesion, morphogenesis, and pathogenesis .....	147
5.3	Introduction.....	147
5.4	Results.....	148
5.5	Discussion.....	160
5.6	Methods.....	162
<b>6</b>	<b>Studying evolution of drug resistance in <i>C. albicans</i> clinical isolates .....</b>	<b>167</b>
6.1	Introduction and contribution to the paper.....	167
6.2	The evolution of gradual acquisition of drug resistance in clinical isolates of <i>Candida albicans</i> .....	169
6.3	Introduction.....	170
6.4	Results.....	173
6.5	Discussion.....	191
6.6	Materials and Methods.....	194
6.7	References:.....	203
<b>7</b>	<b>Conclusions and future directions.....</b>	<b>212</b>
<b>8</b>	<b>Appendix.....</b>	<b>214</b>
8.1	Potential vaccine approach against <i>C. albicans</i> Saps .....	214

## TABLE OF CONTENTS

8.2	H <sub>2</sub> O <sub>2</sub> responsive and transcription factor dependent genes: enriched biological functions.....	220
	<b>References.....</b>	<b>224</b>

## LIST OF FIGURES

- Figure 1.1: Examples of fungal superficial infections. ....4
- Figure 1.2: *Candida* species most commonly isolated from fungal infected patients and their characteristics. ....7
- Figure 1.3: Examples of *C. albicans* superficial infection.....10
- Figure 1.4: *C. albicans* common virulence mechanisms .....13
- Figure 1.5: History of fungal infections and antifungal drug development. ....17
- Figure 1.6: amphotericin B and 5-fluorocytosine mechanisms of action. ....18
- Figure 1.7: cellular antifungal targets. ....19
- Figure 1.8: *C. albicans* antigens suggested as potential candidates for epitope-based vaccine approaches.. ....22
- Figure 1.9: *C. albicans* multiplies in mice and rats challenged intravenously. ....25
- Figure 1.10: *D. melanogaster* is susceptible to *C. albicans*.....28
- Figure 1.11: *Galleria mellonella* is susceptible to *C. albicans*.....30
- Figure 1.12: Zebrafish is susceptible to *C. albicans*.....32
- Figure 1.13: *C. elegans* is susceptible to *C. albicans*. ....34
- Figure 2.1: Molecular pathways involved in *C. elegans* immunity and tested in this study. ....42
- Figure 2.2 *C. elegans* mutants' life span assay with heat killed *E. coli*.....46
- Figure 2.3: MAPK pathway is required for *C. elegans* resistance to *C. albicans* .48
- Figure 2.4: *C. albicans* - *C. elegans* model can be pharmacologically or genetically modulated. ....50

## LIST OF FIGURES

- Figure 3.1 Reverse genetic screening identifies four novel putative virulence factors.....65
- Figure 3.2 *zcf15/zcf15* does not differ from wild-type in several in vitro assays...70
- Figure 3.3: *zcf15/zcf15* is hypersusceptible to ROS generator paraquat.....71
- Figure 3.4: *zcf15/zcf15* complementation strategies .....73
- Figure 3.5: *ZCF15* is required to withstanding *C. elegans* generated ROS.....75
- Figure 3.6 *ZCF15* is required to withstand mouse macrophage generated ROS ...78
- Figure 3.7: *ZCF15* loss of function has a limited impact in survival of mice intravenously challenged.....80
- Figure 4.1: canonical structure of Zn(II)<sub>2</sub>Cys<sub>6</sub> DNA binding domain and relative location in *Zcf15* .....84
- Figure 4.2: *C. albicans* *ZCFs* conservation across ascomycetes. ....94
- Figure 4.3: *ZCFs* gene expression in the presence of oxidative stress (Ox), osmotic stress (Os) and heavy metal stress (Hm) .....95
- Figure 4.4: *ZCF29* is required for wildtype ability to resist ROS.....96
- Figure 4.5: *ZCF29* is required for wildtype resistance to purine analogue caffeine and cell wall stressor SDS.....97
- Figure 4.6: RNASeq reads distribution across strains and time points.....100
- Figure 4.7: RNA-Seq reads distribution between exons, introns and intergenic regions across samples. ....101
- Figure 4.8: RNA-Seq coverage distribution metrics across different strains and time points.....102

## LIST OF FIGURES

- Figure 4.9: *ZCF15* and *ZCF29* expression was not detected in the respective knockout strains. ....103
- Figure 4.10: Pearson's correlation coefficients heat map for each of the 27 gene expression profiles obtained in this study. ....105
- Figure 4.11: Transcriptional correlation between biological replicates and within the same replicate before and after the addition of H<sub>2</sub>O<sub>2</sub>. ....106
- Figure 4.12: The Hog1 and Cap1 pathway are the two major known pathways involved in oxidative stress response. ....107
- Figure 4.13: *ZCF15* and *ZCF29* are not regulating canonical ROS resistance pathways. ....109
- Figure 4.14: Volcano plots showing genes differentially expressed between wild type and *zcf29/zcf29*. ....111
- Figure 4.15: Enrichment map of biological functions overrepresented in *zcf29/zcf29* compared to wildtype. ....112
- Figure 4.16: *ZCF29* deletion causes misregulation of threonine, methionine and arginine biosynthesis. ....114
- Figure 4.17: Genes differentially expressed upon H<sub>2</sub>O<sub>2</sub> exposure between wildtype and *zcf29/zcf29*. ....117
- Figure 4.18: *ZCF29* plays an important role in down regulating ribosome biogenesis upon ROS challenge. ....118
- Figure 4.19: Volcano Plots showing the number of genes differentially expressed between wildtype and *zcf15/zcf15* in the absence of H<sub>2</sub>O<sub>2</sub>. ....120

## LIST OF FIGURES

- Figure 4.20: genes differentially expressed between wild type and *zcf15/zcf15* upon H<sub>2</sub>O<sub>2</sub> challenge. ....122
- Figure 4.21: *ZCF15* is required for down regulating genes involved in C-metabolism and simultaneously upregulating genes involved in oxidation reduction processes/ROS detoxification. ....123
- Figure 4.22: genes involved in iron homeostasis are differentially expressed in both *ZCF15* and *ZCF29* deletions.....125
- Figure 4.23: Iron homeostasis is tightly regulated by three major transcription factors, Sef1, Sfu1 and Hap43.....126
- Figure 4.24: An overview of the ChIP-Seq procedure used .....128
- Figure 4.25: *ZCF15* and *ZCF29* HA epitope tagging .....129
- Figure 4.26 *Zcf15* and *Zcf29* change many of their genomic targets upon H<sub>2</sub>O<sub>2</sub> exposure. ....131
- Figure 4.27: various *ZCF29* DNA binding sites flank genes differentially expressed in *ZCF29* deletion.....132
- Figure 4.28: *Zcf29* directly controls orf19.7204 and orf19.7042 expression. ....134
- Figure 4.29: The upregulation of orf19.251 and orf19.1048 upon H<sub>2</sub>O<sub>2</sub> is directly controlled by *Zcf29*. ....136
- Figure 4.30: Few *ZCF15* DNA binding sites flank genes differentially expressed in *ZCF15* deletion. ....137
- Figure 4.31: *Zcf15* may control the expression of orf19.7224 and orf19.871. ....139
- Figure 7.1 *C. albicans* Saps secretion and hypothesized mechanism of action...215

## LIST OF FIGURES

- Figure 7.2 biochemical or genetic inactivation of Saps reduces the pathogen's ability to induce Dar.....217
- Figure 7.3 Sequence alignment showing the five *C. albicans* Saps conserved domains. ....218
- Figure 8.1: Biological functions enriched in genes H<sub>2</sub>O<sub>2</sub> downregulated and *ZCF29* dependent .....220
- Figure 8.2: Biological functions enriched in genes H<sub>2</sub>O<sub>2</sub> upregulated and *ZCF29* dependent .....221
- Figure 8.3: Biological functions enriched in genes H<sub>2</sub>O<sub>2</sub> downregulated and *ZCF15* dependent .....222
- Figure 8.4: Biological functions enriched in genes H<sub>2</sub>O<sub>2</sub> upregulated and *ZCF15* dependent .....223



## LIST OF TABLES

Table 2.1: <i>C. elegans</i> mutant strains used in this study.....	39
Table 3.1: <i>C. albicans</i> transposon mutant library description and relative references.....	55
Table 3.2: <i>C. albicans</i> homozygous knockout mutants and complemented strains used in this study.....	56
Table 3.3: primers used in this study.....	57
Table 4.1: Primers used in this study.....	90
Table 4.2: list of all 77 Zn(II) <sub>2</sub> Cys <sub>6</sub> transcription factors present in <i>C. albicans</i> . ....	93

## LIST OF SYMBOLS AND ABBREVIATIONS

**CD4:** cluster of differentiation 4

**ChIP-Seq:** Chromatin Immunoprecipitation Sequencing

**CFU:** colony forming unit

**DAR:** Deformed Anal Region

**DMEM:** Dulbecco's Modified Eagle Medium

**DMSO:** Dimethyl sulfoxide

**DPI:** Diphenyleneiodonium

**EMS:** Ethyl MethaneSulfonate

**FBS:** Fetal Bovine Serum

**FC:** fold change

**FGSC:** Fungal Genetic Stock Center

**IL-12:** Interleukin 12

**LB:** Luria Broth

**MAPK:** Mitogen Activated Protein Kinase

**NADPH:** Nicotinamide Adenine Dinucleotide Phosphate

**NGM:** Nematode Growth Medium

**PCR:** Polymerase Chain Reaction

**RNAi:** RNA interference

**RNASeq:** RNA sequencing

**ROS:** Reactive Oxygen Species

**SAP:** Secreted Aspartyl Proteases

**SDS:** Sodium Dodecyl Sulfate

**SMAD:** contraction of Sma and Mad. Sma, *C. elegans* small body size, MAD, Drosophila mothers against decapentaplegic

**TMP:** 4,5',8-TriMethylPsoralen

**TGF- $\beta$ :** Transforming Growth Factor beta

**TNF- $\alpha$ :** Tumor Necrosis Factor alpha

## **LIST OF SYMBOLS AND ABBREVIATIONS**

**TOR pathway:** target of rapamycin

**YPD:** Yeast extract Peptone Dextrose

## ABSTRACT

The *C. albicans* community is currently laying the foundation of understanding how this human pathogen causes infection. *C. albicans* infections represent a major medical and economic burden for today's society with an estimated 400,000 blood stream infections worldwide and direct costs exceeding 1\$ billion dollar a year in the U.S. alone. Although finding the biological causes of this disease seemed to be beyond our reach in the past, various aspects of the infection have been recently unveiled including its pathology, immunology, histology, and epidemiology. Here we explored the genetic components of this disease by studying the complex host-pathogen dynamics through a series of *in vivo*, *ex vivo* and *in vitro* experiments. By using a pathogen unbiased reverse genetic approach and a host gene candidate strategy we uncovered some of the genes and pathways that are important for pathogenicity and immunity. In particular we explored the complex host-pathogen dynamics using a *C. albicans* - *C. elegans* model system and identified four novel putative virulence factors. We focused on Zcf15, a *C. albicans* transcription factor that has been poorly characterized in the literature and that plays an important role in the pathogen's ability to resist host generated reactive oxygen species (ROS). By leveraging the power of RNASeq and ChIP-Seq we identified Zcf15 transcriptional targets and DNA binding sites. These studies suggest that Zcf15 plays a critical role in carbon metabolism and that it exerts its ability to protect the pathogen from ROS by controlling the expression of thiol- peroxidases and other detoxifying enzymes.

We also showed here that in *C. elegans*, the host's ability to counteract the infection relies on the MAPK pathway, evidence that mirrors what has been found by others in mammals

and that emphasizes the usefulness of studying *C. albicans* infections in smaller genetically traceable organisms like *C. elegans*.

The nematode model is also shown here to be a powerful tool not only to study the genetic bases that drive infection and immunity but also to identify new compounds that can be used for therapeutic intervention. This model was instrumental in identifying filastatin, a small molecule that was subsequently found by our collaborators to be capable of reducing virulence in mammals. The antifungal properties of filastatin are currently undertaking further preclinical testing.

Overall this thesis shed light on the complex mechanisms of *C. albicans* pathogenicity and host immunity and identified novel virulence determinants that can be used by the larger community for further biological studies or even drug development.

# **1 Introduction and background**

## **1.1 Human fungal infection: a growing threat**

Human fungal infections have alarming morbidity, mortality, and a heavy pharmaco-economical impact on today's society. Fungal infections, also known as mycosis, are caused by various types of fungi capable of colonizing and invading several human tissues. It is estimated that there are approximately 5.1 million different species of fungi on earth and about 300 of those are pathogenic to humans (Blackwell, 2011; Garcia-Solache and Casadevall, 2010). These infections are very common comorbidities of immunocompromised patients and represent a growing threat especially considering that the pool of immunocompromised patients has been growing rapidly in the last few decades (Miceli et al., 2011). In 2008, in the U.S. alone over 23,000 patients were treated with immunosuppressive therapy before organ transplantation, a number two times higher compared to the previous 10 years (Low and Rotstein, 2011). This number coupled with the use of more potent chemotherapeutic agents and the high incidence of HIV infections lead to an estimated 10 million immunocompromised people in the U.S., a remarkable 3.6% of the total population (Kahn, 2008). The increased number of immunocompromised patients in the last three decades has been mirrored by a rise in fungal incidence. In the U.S. from 1980 through 1997, the annual number of deaths caused by fungal infections increased from 1,557 to 6,534, a 320% increase in 17 years (McNeil et al., 2001) and it is estimated that *Candida* infections of the oral cavity affect 2 million people worldwide (Buchacz et al., 2010) and that 70% of all premenopausal women experience at least one episode of vulvovaginal candidiasis in their lives (Fidel et al., 1999).

In the last two decades a large body of research literature has been accumulated and shed some light on some of the common mechanisms causing fungal infections (Casadevall, 2007). However, many of the cellular pathways and signal transduction cascades responsible for the infections remain elusive, and there is a pressing need for more research in order to develop better diagnostics and therapeutics.

## **1.2 Types of fungal infections**

Fungal infections can be divided into two types depending on their sites of infections:

- Superficial: infections occurring on keratinized tissue like skin, nails or hair.
- Mucosal/Invasive: infections taking place on moist surfaces like the oral cavity, gastrointestinal, urogenital and respiratory tract and that have the potential to reach the bloodstream, infecting normally sterile sites like internal organs or the cerebrospinal fluid.

Superficial and mucosal infections, although unpleasant, recurrent and difficult to treat are usually not life-threatening. On the other hand, mucosal/invasive infections have a significantly lower incidence but their prognosis can be very poor.

### **1.2.1 Superficial infections**

Superficial infections affect up to 25% (~1.7 billion) of the general population worldwide (Havlickova et al., 2009) and most people will experience at least one episode of superficial infection during their life time. These infections are primarily caused by dermatophytes like

*Tinea pedis*, *Trichophyton rubrum*, *Microsporum canis* and lead to well-known conditions like athlete's foot, ringworm of the scalp and nail infections (Figure 1.1). Prevalence rates associated with these infections are exceptionally high, with athlete's foot occurring in one every five adults, infection of the nails affecting up to 10% of the general population and ringworm of the scalp estimated to affect more than 200 million individuals worldwide (Havlickova et al., 2009; Thomas et al., 2010). These infections are very difficult to diagnose and are often mistaken for other conditions like eczema or psoriasis which are immune-mediated conditions that have a very different etiology, therapeutic approach and outcome.

Athlete's foot infections are usually long-lasting, recurring and cause symptoms like burning, itching, cracking and scaling and, if left untreated, can lead to severe swelling and even limb amputation in diabetic subjects. The infection is usually caused by one of two types of fungi: *Trichophyton metagrophytes* that cause blisterlike infections that appear suddenly but that are easily treated and *Trichophyton rubrum* that cause long-lasting, recurrent and difficult to treat infections. In contrast to the large majority of other fungal infections, athlete's foot is very contagious and can spread through direct (human-to-human) or indirect contact (through vectors like flies, mites or fleas).

These infections are traditionally treated with a combination of systemic and topical treatment and recently the FDA granted marketing approval for laser therapies that can reach deeper tissues. Although recent gene expression profiling studies have shed some light on the genes involved in pathogenesis (Maranhao et al., 2011), very little is known about the cellular pathways and mechanisms that lead to adhesion, skin layer penetration, tissue invasion and virulence.

Ringworm of the hair is usually caused by dermatophytes in the *Trichophyton* and *Microsporum* genera and can cause severe itching of the scalp and bald patches especially in



areas where the fungus has rooted itself in the skin. *Trichophyton* infections occur predominantly in the United States, Central America and Western Europe while infections from *Microsporum* are more common in South America, East and Central Europe, Africa and the Middle East. Pets are usually the vehicle for the infection, but the infection can also be transmitted between humans. Novel nested PCR based strategies have revolutionized dermatophytes diagnostics (Verrier et al., 2013; Wisselink et al., 2011) but our understanding of how the fungus infects and evades the human immune system remains elusive.



**Figure 1.1: Examples of fungal superficial infections.**

(A) A severe case of athlete's foot caused by the fungus *Trichophyton rubrum*. Adapted from (Heilman, 2011) (B) An example of ringworm of the hair (webMed, 2007). (C) A case of onychomycosis, commonly known as fungal nail infection. Fungal nail infections are difficult to treat because the infection is embedded within the nail and accessing the site of infection is problematic. (Medbullets, 2013)

Nail infections, also known as onychomycosis, are also very common and can be caused by dermatophytes like various species of *Trichophyton*, *Candida* or nondermatophytic molds like *Scytalidium*, *Scopulariopsis* or *Aspergillus*. These infections usually cause the nail to become thickened and discolored first, brittle and fragile next and eventually inflamed and painful if left untreated. These infections are commonly treated with topical antimycotics, however these drugs are often ineffective because they can't reach the site of infection that is embedded within the nail. Recent research shows that laser-driven heat killed strategies might

represent a great therapeutic alternative to topical antimycotics because of their ability to effectively reach the site of infection (Suga et al., 2014).

Superficial infections are very rarely life threatening but they are associated with high morbidity, and their refractory to pharmacological treatments can be the first indication of an underlying immunodeficiency. In addition to the pain and chronic discomfort, superficial infections can cause psychological, social or employment-related difficulties.

### **1.2.2 Mucosal/Invasive infections**

While superficial mycosis are limited to outermost layers of the skin and don't perfuse to deeper tissues, mucosal infections are significantly more serious and can reach moist surfaces like lungs, mouth, intestine or urogenital tract. In severely immunocompromised patients, these infections can reach the blood stream and cause systemic infections. The four most common pathogens that cause mucosal/invasive infections are *Cryptococcus*, *Histoplasma*, *Aspergillus* and *Candida*.

*Cryptococcus* species can cause severe fungal infection called cryptococcosis and are usually caused by *C. gattii* and *C. neoformans*. These two fungi belongs to the phylum of Basidiomycota and are responsible for global deaths of more than 600,000 worldwide (Kwon-Chung et al., 2014). These pathogens live in the soil, and humans can be infected by inhaling airborne, dehydrated yeast cells or spores. These yeasts are weakly encapsulated when found in the wild, however, when they reach the lungs they are capable of forming thick melanin-rich capsule that protect them from the immune system. These encapsulated yeast cells are capable of traveling through the blood stream either on their own or within macrophages and can infect other areas of the body, typically the central nervous system. If they reach the central nervous system they cause meningoencephalitis, a condition that

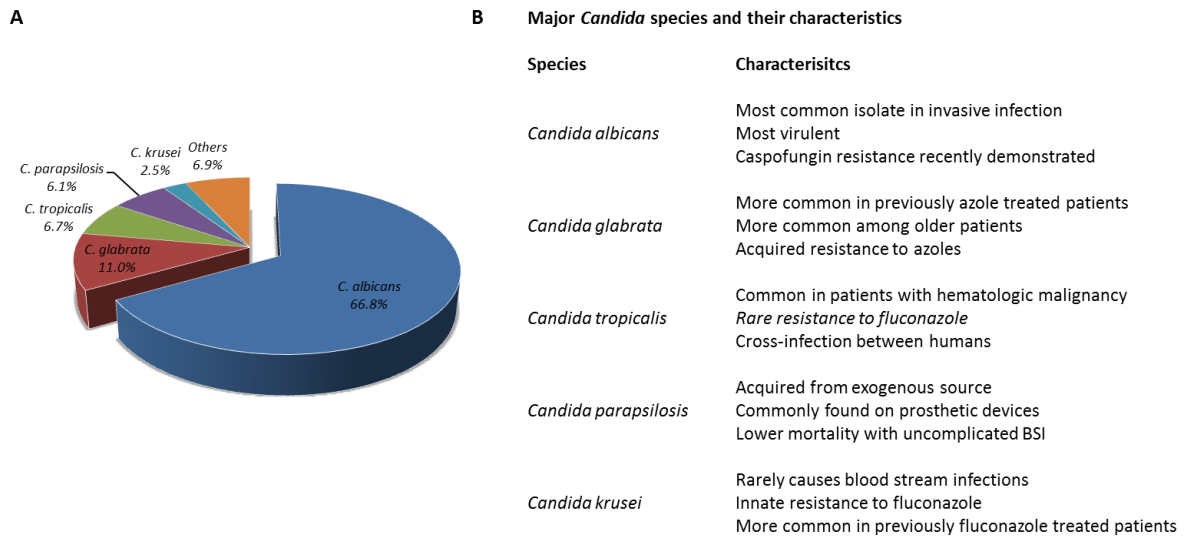
resembles both meningitis (inflammation of the meninges) and encephalitis (inflammation of the brain). Cryptococcosis is usually diagnosed upon detection of cryptococcal antigen either in sputum, urine or cerebrospinal fluid (Antinori, 2013). Although recent gene expression profile studies have elucidated some of the mechanisms that lead to the infections (Fan et al., 2005; Goebels et al., 2013), very little is known about its ability to elude the human immune system and cross the blood brain barrier.

*Histoplasma* is dimorphic fungus usually found in bird and bat droppings that can cause a respiratory disease called histoplasmosis. When inhaled, the fungus can lodge deeply in the lung tissues, enter a latent stage and remain undetected for many years (Rihana et al., 2014). However, if the immune system is compromised (HIV, chemotherapy or organ transplant) the pathogen can colonize the lungs and cause pneumonia-like symptoms like non-productive cough, chest pain and fever. In rare cases of patients severely immunocompromised the infection can disseminate to internal organs like the liver causing severe hepatic histoplasmosis. Histoplasmosis is considered a rare disease; however in HIV-infected patients the mortality rate is estimated to be as high as 10% (Kauffman, 2007).

*Aspergillus* is a member of the deuteromycetes fungi and can cause infections and diseases in patients with tuberculosis or with chronic obstructive pulmonary disease. *Aspergillus* spores can be inhaled by susceptible patients and can colonize the lungs. The warmer environment of the lungs triggers a morphological change that leads to the formation of oval budding yeast cells that can be phagocytized by macrophage and travel to other body parts like lymph nodes or other internal organs. Invasive aspergillosis can be a devastating disease with mortality rates as high as 90% in certain patient populations (Dagenais and Keller, 2009).

Although difficult to diagnose and treat, infections caused by *Cryptococcus*, *Histoplasma* and *Aspergillus* are rare and observed only in severely immunocompromised patients. In contrast,

yeast infections caused by *Candida* are significantly more common. These infections are caused by a large variety of different *Candida* species. Pfaller and Diekema 2007, recently published the results of a global antifungal surveillance program that give us a sense of the most common *Candida* species causing infections.



**Figure 1.2: *Candida* species most commonly isolated from fungal infected patients and their characteristics.**

(A) *Candida* species distribution obtained from the Artemis Disk Global Antifungal Surveillance Study (Pfaller and Diekema, 2007). (B) Among all *Candida* species causing infections, *C. albicans* is considered the most frequent and the most virulent of all. Adapted from (Marr, 2004).

Samples collected from blood, mouth, urine and genitals of patients suffering from fungal infections were collected across 127 medical centers and 39 countries. The study lasted 6.5 years and a total of 134,715 samples were collected. As shown in Figure 1.2, *C. albicans* was by far the most common *Candida* specie isolated (66.8%) followed by *C. glabrata* (11.0%), *C. tropicalis* (6.7%), *C. parapsilosis* (6.1%) and *C. krusei* (2.5%).

This global study highlight that *C. albicans* is and remains the most frequent species responsible for fungal infections. *C. albicans* is considered the fourth most common cause of hospital-acquired infections and the third most common cause of bloodstream infections in

intensive care units (Lewis, 2009). In addition to being the most prevalent, *C. albicans* has been shown to be the most virulent species and the one associated with the most concerning prognosis (Papon et al., 2013). The high prevalence, morbidity and mortality of *C. albicans* infections are some of the reasons why in this dissertation we focused on this organism.

### 1.3 *C. albicans* infections: mortality, morbidity and costs

In December 2013, the Centers for Disease Control and Prevention (CDC) released the 2013 infectious disease threat report, a document that gave the general public a snapshot of the burden and threats posed by drug-resistant pathogens in the United States. Drug-resistant fungal infections caused by *Candida albicans* were classified as the fourth most serious threat, behind bacterial infections caused by *Clostridium difficile*, *Neisseria gonorrhoeae* and *Enterobacteriaceae* (Hampton, 2013). The report estimated 46,000 *Candida* systemic infections a year, 3,400 of which were drug-resistant and that lead to 220 deaths. One major reason that explains these alarming numbers is the remarkable genomic plasticity of *C. albicans*. *C. albicans* generates diversity through a variety of different mechanisms, including parasexual alteration in chromosome number, mating, loss of heterozygosity and recombination between homologous chromosomes (Robinson, 2008; Selmecki et al., 2010a). Its genomic plasticity makes *C. albicans* remarkably capable of coping with various environmental stresses, including host immunity and antifungal drug exposure. A large body of literature has been accumulated over the past two decades (Mayer et al., 2013), however there is still a pressing need for more research in order to facilitate the development of better therapies and diagnostics.

*C. albicans* is a ubiquitous member of the human microbiome that, under certain circumstances, can colonize mucosal and non-mucosal surfaces like the mouth, lining of the gut, nasal passages, airways, urinary tract and genitals. *C. albicans* infections incidence rates are remarkably high as it is estimated that 50 to 75% of women in their childbearing years will suffer from at least one episode of vulvovaginitis caused by *Candida* and up to 8% of them will have at least four episodes annually (Sobel, 2007). The large majority of *C.*

*albicans* infections are either superficial or mucosal (also known as Candidiasis), however, in conditions of severe immunosuppression, *C. albicans* can reach the bloodstream, leading to systemic infection (Candidemia). Once *C. albicans* reaches the blood stream, it can evade many host immunological barrier and colonize several internal organs like liver, spleen and kidney. Interestingly, it has been shown that *C. albicans* has the ability to traverse the blood brain barrier and cause severe and frequently lethal meningitis (Aleixo et al., 2000; Jong et al., 2001; Liu et al., 2011). Estimated mortality rates for Candidemia are as high as 40%, and the associated health care costs are between 1\$ and 2\$ billion per year in the U.S alone (Gudlaugsson et al., 2003a; Rentz et al., 1998). As shown in Figure 1.3 *C. albicans* can cause oropharyngeal thrush even in immunocompetent infants (Figure 1.3 A), skin infection in immunocompromised patients (Figure 1.3 B) or systemic infection that reaches the blood stream usually through the intestinal microvilli (Figure 1.3 C).



**Figure 1.3: Examples of *C. albicans* superficial infection.**

(A) Oral thrush, an infection of the oral cavity very common in low birth weight babies and adult patients affected by HIV. Adapted from (Heilman, 2008) (B) *C. albicans* skin infection. Adapted from (Candidaeffects, 2013) (C) Cartoon depicting an epithelial infection that reaches the blood stream by penetrating the intestinal microvilli. *C. albicans* in yeast form adhere to the intestinal microvilli, penetrate the epithelial tissue (b and c) and reaches the blood stream (d). Once in the blood stream *C. albicans* can infect any internal organs with kidney as one of the preferential targets (e). Adapted from Chauhan et al. 2006

Therapeutic options for patients affected by *C. albicans* infections are limited. The drugs most commonly prescribed were all FDA approved in the 1980-90s and their efficacies are progressively decreasing as the number of drug resistant strains has been exponentially increasing (Cannon et al., 2007; Mukherjee et al., 2003). In order to develop new diagnostics and therapies some of the fundamental biological and molecular questions regarding the pathogen's ability to adhere, infect, penetrate and thrive in the human body need to be answered. In addition, despite a growing list of risks factor that have been recognized, the fundamental mechanisms that allow this pathogen to attack the mucosal tissue and evade the immune system remain elusive. One of the goals of this thesis is to explore new genes and virulence mechanisms that lead to *C. albicans* infections.

#### **1.4 *C. albicans* virulence mechanisms**

The most important etiological factor that controls *C. albicans* infection is the physiological status of the host as *C. albicans* infections are commonly found in immunocompromised patients. However, *C. albicans* is also found in approximately 50% of the population as part of the harmless individual's microflora and it has been shown that slight alterations in the host can turn this normally harmless commensal into a life-threatening pathogen (Naglik et al., 2003). The transition from harmless commensal to pathogen is a fine line in *C. albicans* and it is attributable to its extensive arsenal of virulence factors that are selectively expressed under suitable predisposing conditions. These mechanisms are summarized in Figure 1.4 and include:

- Ability to adhere and penetrate host tissues (Figure 1.4 A and B).
- Ability to form biofilm (Figure 1.4 C).

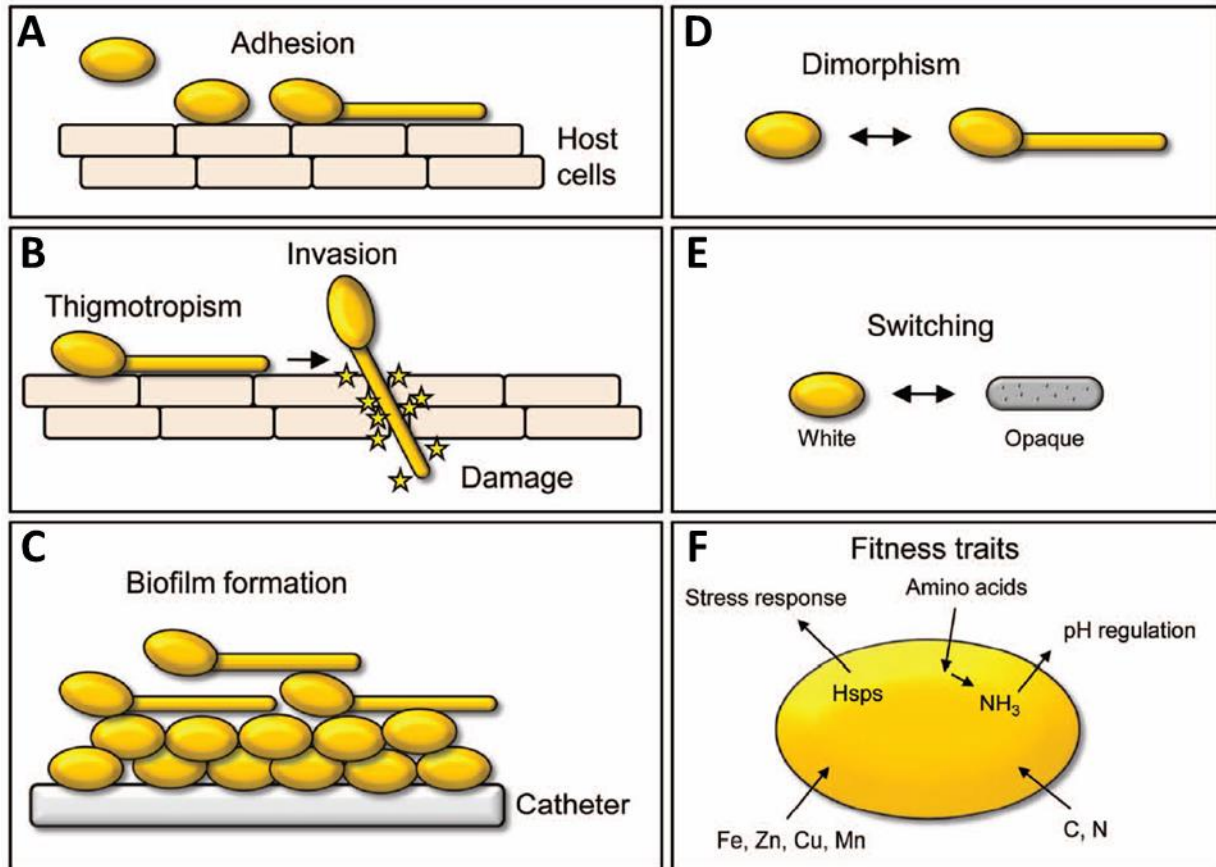


- Ability to switch from yeast-to-hypha (Figure 1.4 D).
- Ability to switch from white to opaque state (Figure 1.4 E).
- Ability to resist stress, uptake amino acids, adapt to pH fluctuations and utilize extracellular carbon, nitrogen and essential trace metals (Figure 1.4 F).

*C. albicans* infections begin with the adhesion of yeast cells to host tissues (Figure 1.4 A). This process is mediated by adhesins proteins like Als (Agglutinin-Like Sequence) (Hoyer, 2001). Als are a family of eight glycosylphosphatidylinositol (GPI)-anchored cell surface proteins that protrude from the *C. albicans* cell wall and that are capable of binding host surface proteins and extracellular matrix.

Once *C. albicans* attaches to host cells, thigmotropism (contact sensing) triggers *C. albicans* filamentation and allow the pathogen to penetrate deeper host tissues (Figure 1.4 B) via the secretion of hydrolytic enzymes like Saps (Secreted Aspartyl Proteases) and phospholipases (Aoki et al., 2011; Schaller et al., 2005).

This remarkable ability to switch from yeast to hyphae has been shown to be critical for virulence. During yeast growth *C. albicans* has the traditional oval shape while during hyphal growth *C. albicans* has an elongated ellipsoid form and can form long filaments that penetrate deeper tissues (Figure 1.4 D). The hyphal form is considered more invasive than the yeast form (Jacobsen et al., 2012) but the ability to switch from one form to the other is critical for virulence. In particular, strains locked in one form or the other have a significantly reduced virulence compared to strains that can freely switch from one form to the other (Lo et al., 1997). The working hypothesis is that hyphal forms are important for tissue invasion in epithelial infections while yeast form, which are smaller in size and can travel more effectively in the blood stream are required for systemic infections (Saville et al., 2003).



**Figure 1.4: *C. albicans* common virulence mechanisms. Adapted from (Mayer et al., 2013)**

*C. albicans* has also a remarkable ability to form biofilms in abiotic surfaces like catheters, dentures and other implants (Figure 1.4 C). Upon adhesion *C. albicans* proliferates, forms hyphal cells in the upper part of the biofilms and finally produce a complex extracellular matrix that renders the biofilm more resistant to antimycotic agents (Finkel and Mitchell, 2010). Recent gene expression profiling and ChIP-ChIP experiments suggest that the cellular circuit controlling biofilm formation is regulated by six transcriptional regulators (*TEC1*, *EFG1*, *BCR1*, *NDT80*, *ROB1* and *BRG1*) that tightly controlled the expression of ~1,000 target genes (Nobile et al., 2012).

*C. albicans* is also capable to switch from the normal yeast morphology (white) to the mating-competent elongated form (opaque) and this influences its ability to colonize the host

(Sasse et al., 2013). In particular white and opaque cells are differentially adapted to specific host niches with opaque cells more virulent in the hypoxic environment of the mammalian intestine (Ramirez-Zavala et al., 2008).

In addition to these direct virulence factors, *C. albicans* has various “fitness traits” that allow it to grow and resist host immunity. These mechanisms include a robust stress response mediated by heat shock proteins (Hsps) as well as an arsenal of detoxifying enzymes and efflux pumps. Moreover, *C. albicans* has a remarkable ability to adapt to the different nutritional environments of the host. Modulating glycolysis, gluconeogenesis, amino acids uptake and starvation responses are all believed to be important for virulence. For example, in the gastrointestinal tract the concentration of nutrients is relatively high but *C. albicans* competes for nutrients with other members of the microbiome. If the gastrointestinal flora is imbalanced (for example in patients undergoing systemic antibiotics therapies), *C. albicans* has access to additional nutrients and can outgrow the other organisms (Brock, 2009). On the other hand, when *C. albicans* is in the blood stream glucose is highly abundant (6-8 mM) but as soon as it is phagocytized the nutritional environment changes completely and *C. albicans* faces nutrient starved conditions. Inside the macrophage *C. albicans* promptly switches from glycolysis to gluconeogenesis and downregulated energetically demanding processes like ribosome biogenesis (Lorenz et al., 2004). Since the environments that *C. albicans* faces during infection are various, a prompt and efficient metabolic plasticity is required for adapting to such different conditions.

## 1.5 *C. albicans* therapeutics

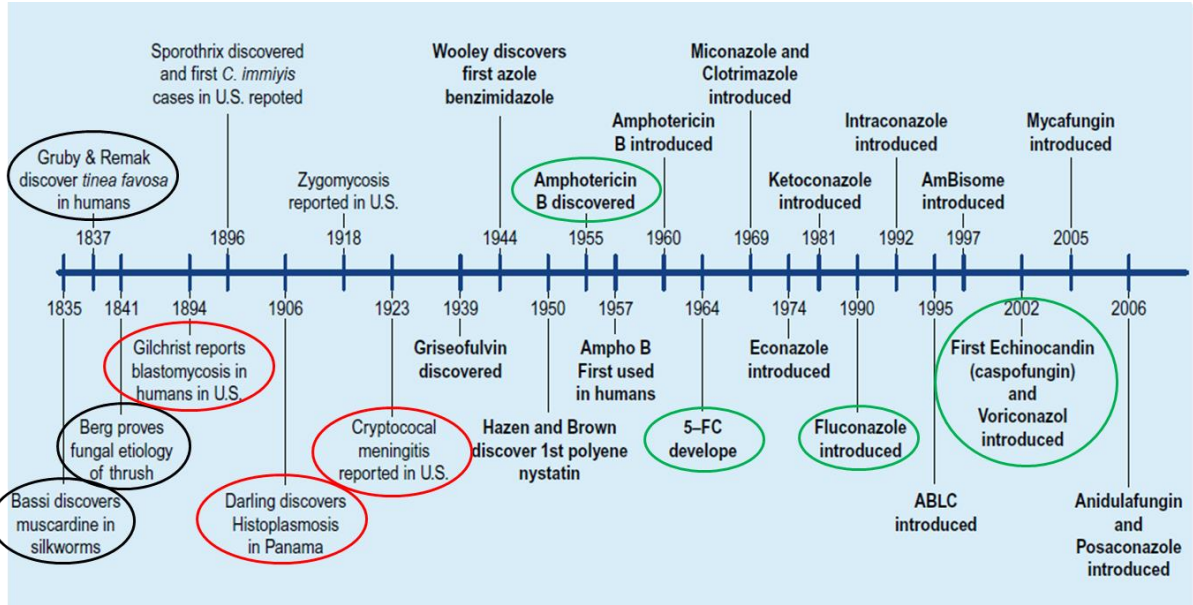
### 1.5.1 Antifungal drugs

As previously discussed in chapter 1.1 the increased use of immunosuppressive therapies, potent chemotherapeutic agents and the surge of AIDS patients caused a dramatic increase in the number of immunocompromised patients, which lead to a dramatic rise in the number of patients suffering from *C. albicans* infections (Wang and Xu, 2008). In addition to the increased number of *C. albicans* infected patients, the clinical outcome of these infections is complicated by the ability of the strain to develop drug resistance and by the limited number of antimycotics currently available (Ferreira et al., 2010). In the last two decades, *C. albicans* developed various clever mechanisms for drug resistance, including genetic alteration of the targeted proteins and overexpression of efflux pumps (White et al., 2002). Patients infected with drug resistant strains don't respond to the pharmacological therapy and, because the number of antimycotic alternatives is limited, the prognosis can be poor. The large majority of antimycotics currently prescribed have been FDA approved in the 1960s and 1990s and there are not very many drugs in company pipelines that have promising safety and efficacy profiles (Ostrosky-Zeichner et al., 2010).

In these paragraphs the antifungal drugs available and their mechanisms of actions will be described in the chronological order by which they were discovered (Figure 1.5). Historically, the antifungal drug discovery process has always faced a big challenge: the eukaryotic nature of the infecting organism. This factor dramatically limits the number of targets available and significantly increases human toxicities (Roemer and Krysan, 2014).

Treatments of fungal infections are further hampered by solubility problems, short half-life and poor pharmacokinetics.

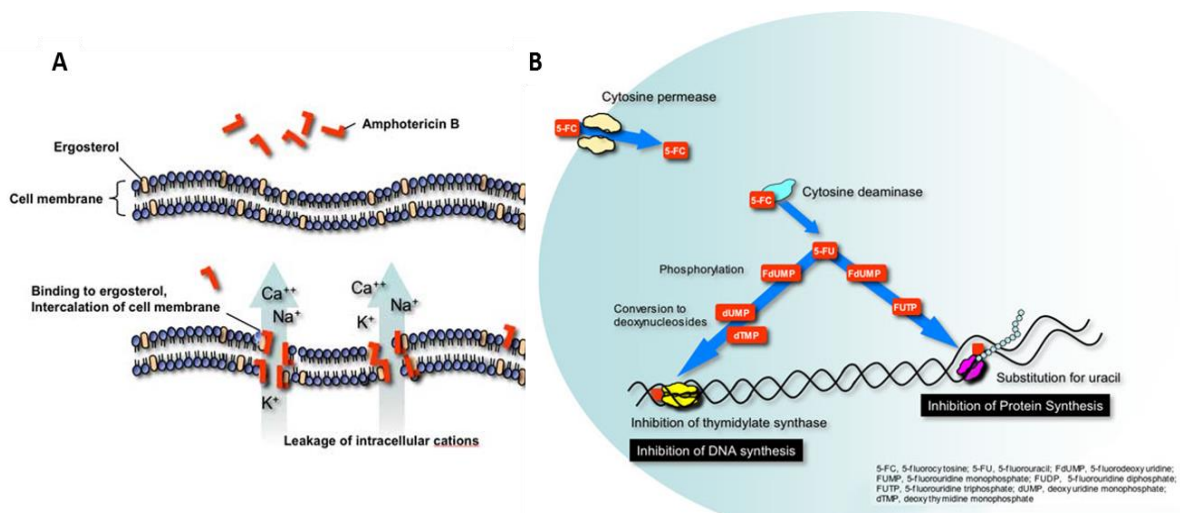
In the pre HIV-era, the incidence of fungal infections was considered low and the pharmaceutical industry invested very little money and time in research and drug discovery. For many years the only antimycotic available was amphotericin B, a polyene delivered intravenously with an atypical mode of action. Instead of inhibiting an enzyme in a metabolic pathway it directly binds ergosterol, the principal sterol in the fungal cell membrane (Baginski and Czub, 2009). Upon binding, amphotericin B interferes with membrane polarity, structure and fluidity causing the leakage of cellular monovalent and divalent cations like  $\text{Ca}^{++}$  and  $\text{K}^+$  (Ellis, 2002), Figure 1.6 A. Amphotericin B is a fungicidal because it causes cell death and it is believed that, on top of the ergosterol binding ability its toxicity may include oxidative damage and host immune stimulation (Hawser and Islam, 1999). Although amphotericin B is still commonly used to treat serious infections, its use is limited by infusion-related side effects like chills, fever, anaphylactoid-like reactions (Eriksson et al., 2001) and dose-limiting nephrotoxicity like renal vasoconstriction or tubular acidosis (Sawaya et al., 1991). After the introduction of amphotericin B in 1957, the discovery of a new safer and more efficacious drug proved to be difficult. In 1964 5-fluorocytosine (5-FC), also known as flucytosine, was discovered, but the development of drug resistance soon limited its use. The drug was originally developed as a potential anticancer but, although ineffective against tumors, it was proven efficacious against fungal infections.



**Figure 1.5: History of fungal infections and antifungal drug development.**

Figure adapted from (Anaissie et al., 2009). In the 1830s and 1840s the first reports documenting fungal etiology for many human and non-human infections were proposed (black circles). Between the end of the 19<sup>th</sup> century and the beginning of the 20<sup>th</sup> century, human diseases like blastomycosis, histoplasmosis and cryptococcosis were discovered (red circles). Four major breakthroughs were made in the history of medical mycology: in 1955, 1964, 1990 and 2002 amphotericin B, 5-flucytosine (5-FC), fluconazole and caspofungin were discovered respectively (green circles). These four drugs belong to four different chemical classes: polyenes, pyrimidine analogs, azoles and echinocandins. Although many chemical analogs have been developed, these 4 drugs are still the most commonly prescribed antimycotics against fungal infections.

As shown in Figure 1.6B, 5-FC is transported by a fungal specific cytosine permease into the cytoplasm and converted to 5-fluoruracil (5-FU). 5-FU is a pyrimidine analog that is incorporated into RNA during transcription and that interferes with mRNA translation, causing the synthesis of miscoded proteins and ultimately translational arrest. 5-FU not only interferes with transcription and translation but also with DNA replication: 5-FU is converted to its deoxynucleoside (a nucleotide without a phosphate group) which interferes with proper DNA synthesis. Although initially efficacious, resistance to 5-FC is quite commonly seen even in treatment-naïve patients and this drug is no longer prescribed as a monotherapy and only used in combination with amphotericin B.



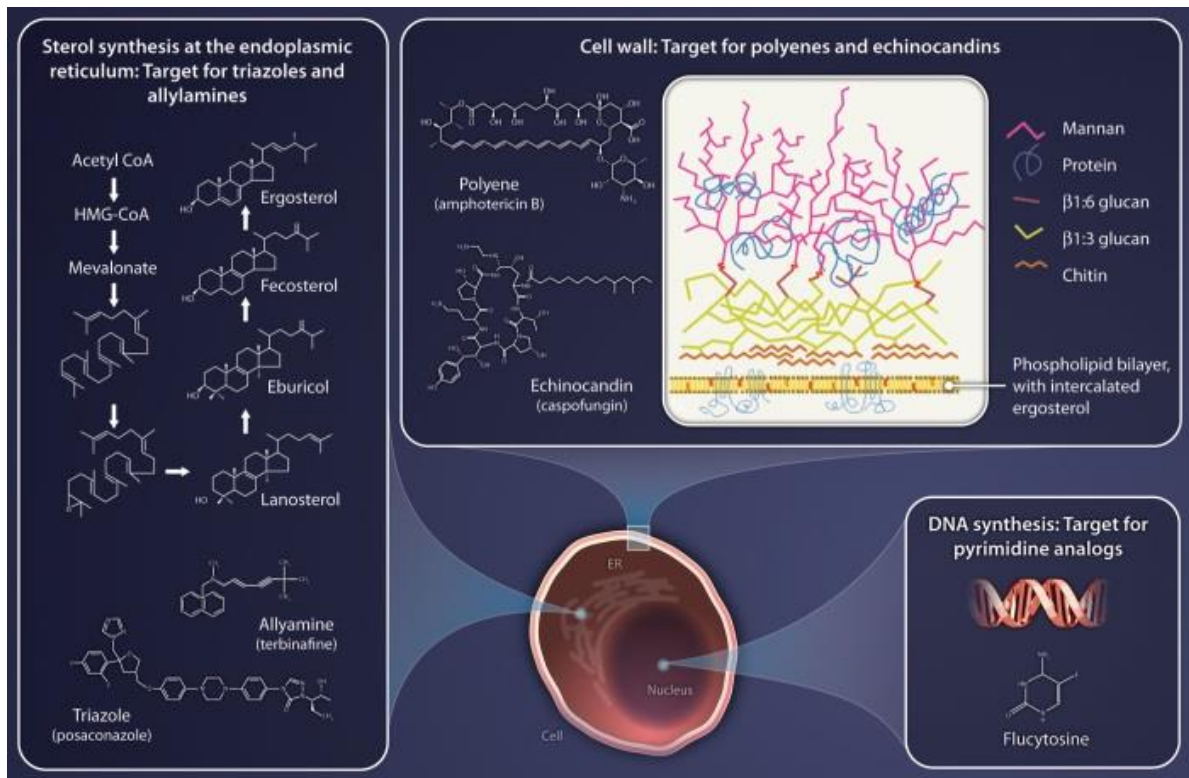
**Figure 1.6: amphotericin B and 5-fluorocytosine mechanisms of action.**

Figure adapted from (Russell, 2010). Amphotericin B binds ergosterol, compromising cell membrane fluidity and polarity. Leakage of important cations like Ca<sup>++</sup>, Na<sup>+</sup> and K<sup>+</sup> are the ultimate cause of the fungicidal effect. 5-fluorocytosine is imported in the cytoplasm by cytosine permease and ultimately converted into 5 fluorouracil, a pyrimidine analog that interferes with gene transcription and mRNA translation.

The 1970s and 1980s represent the nadir in the search for new safer and more effective antimycotics, and for these two decades amphotericin B and 5-FC were essentially the only two therapeutic alternatives. The only drug approved in these two decades was ketoconazole, an imidazole with poor solubility and lack of intravenous formulations.

In 1990, the introduction of fluconazole revolutionized antifungal pharmacological therapy; fluconazole displayed a broader spectrum of antifungal activity, a significantly higher efficacy and a markedly improved safety profile compared to amphotericin B and ketoconazole. In addition, fluconazole was the first antifungal drug to be developed that could be used both orally for minor superficial infections and intravenously for more serious systemic infections. As shown in Figure 1.7, Fluconazole is a triazole that inhibits the synthesis of ergosterol, a fungal specific cell membrane sterol that regulates membrane fluidity and cell signaling. Fluconazole interacts with 14- $\alpha$ -demethylase, a cytochrome P-450

enzyme necessary to convert lanosterol to ergosterol. By inhibiting ergosterol biosynthesis, fluconazole increases cellular permeability and cause leakage of cellular content. In addition to ergosterol biosynthesis inhibition, fluconazole is believed to interfere with endogenous respiration, phospholipid biosynthesis and purine uptake (Pasko et al., 1990). Twenty-five years after its discovery, fluconazole is still the drug of choice to treat fungal infection.



**Figure 1.7: cellular antifungal targets.**

Cell wall, endoplasmic reticulum and nuclear DNA are the three most common antifungal targets. Echinocandins interfere with cell wall architecture by inhibiting the enzyme (1-3)- $\beta$ -D-glucan synthase and preventing  $\beta$ -glucans synthesis. Triazoles and allylamines target the endoplasmic reticulum by inhibiting the synthesis of ergosterol, a sterol essential for cell membrane fluidity. Flucytosine is a pyrimidine analog that inhibits gene transcription and DNA replication as described more in details in Figure 1.6. Figure adapted from (Cowen, 2008)

In 2001 the FDA approved caspofungin, a lipopeptide member of a new family of antifungals called echinocandins. Caspofungin inhibits cell wall synthesis by inhibiting the enzyme 1,3-



beta glucan synthase, a glucosyltransferase involved in the generation of beta-glucans in fungi. Although proven to be efficacious against *C. albicans* infections, caspofungin can only be administered intravenously. Moreover, animal studies have shown that it has embryotoxic effects and it can be administered to pregnant women only if the benefits to the mother clearly outweigh the potential risk for the fetus.

As evident from this chapter, the number of antifungal drugs available is limited and *C. albicans* is rapidly developing drug resistance against them. There is therefore a pressing need to foster our understanding of the mechanisms that lead to *C. albicans* virulence in order to identify new targets and better therapeutics.

### 1.5.2 *C. albicans* vaccine, promises and challenges

On top of the traditional pharmacological therapy, the research community is exploring new therapeutical avenues. Among them, the possibility of generating a *C. albicans* vaccine has recently gained some traction. Vaccines are arguably one of the greatest inventions of modern medicine and the World Health Organization estimated that immunization currently saves two-to three million lives per year worldwide. Many *Candida* antigens have been proposed as vaccine candidates (Raska et al., 2005; Vecchiarelli et al., 2012), including heat shock proteins (Hsp90), proteins involved in energy production and glycolysis like Eno1 and Pfk1 (Li et al., 2011) and enzymes involved in tissue penetration and nutrient utilization like Sap2 (De Bernardis et al., 2012). All of these proteins have two qualities in common; they are required for wild-type virulence and don't have human homologs. In addition to proteins, many fungal polysaccharides and glycoproteins have been prophylactically injected in smaller mammals and proved to induce a strong immunological response and fungal protection (Huang et al., 2010; Xin et al., 2012). Figure 1.8 provides a snapshot of all the fungal antigens that have been studied for their immunological and protective properties against *C. albicans*.

Among all of these approaches, two (NDV-3 and PEV-7) have successfully completed Phase I clinical trials and are planned to enter further clinical investigation in larger patient populations. One of them, NDV-3, developed by NovaDigm Therapeutics, is a formulation containing the N-terminal portion of the *C. albicans* agglutinin-like sequence protein 3 (Als3) and the immunologic adjuvant aluminum hydroxide. This formulation has been shown to have protective efficacy in animal models of oral, vaginal and disseminated candidiasis (Ibrahim et al., 2006; Spellberg et al., 2008). Als3 has sequence and structural homology with

various *S. aureus* cell surface proteins, and preclinical data showed that NDV-3 is safe and efficacious against both candidemia and staphylococemia (Spellberg et al., 2008). In addition, clinical data has shown that NDV-3 is safe and well tolerated in 40 healthy humans (Schmidt et al., 2012). Adverse effects were mild, not statistically significantly compared to placebo, and treatment elicited a robust B- and T-cell immune response. Encouraged by these results, NovaDigm is now recruiting patients for a phase II study, the first ever phase II study for a human fungal vaccine. The double-blind, placebo-controlled interventional 2-year study will recruit a total of 45 women suffering from recurrent vulvovaginal candidiasis, and vaccine efficacy will be evaluated using vulvovaginal recurrence rate as the primary endpoint.

Components	Reference	Protective immunity	Protection
1,3- $\beta$ -glucan	Torosantucci et al. (2009), Cassone et al. (2010), Bromuro et al. (2010)	Abs	Systemic
rHyr1p-N	Luo et al. (2011)	Abs	Systemic
Mannoproteins	De Bernardis et al. (2010) Pietrella et al. (2002)	B cells Th1	Vaginal Systemic
$\beta$ -1,2-mannotriose-Fba	Xin et al. (2012)	Abs	Systemic
Heat shock proteins 90	Raska et al. (2008)	Abs	Systemic and vaginal
<i>C. albicans</i> surface protein, Als3p	Spellberg et al. (2008)	Th1–Th17	Systemic
(Agglutinin-Like Sequence 3) rAls3-N	Baquir et al. (2010), Liu and Filler (2011)	Abs	Vaginal, systemic, and oral
Phosphoglycerate kinase	Calcedo et al. (2012)	Abs	Oral
Secreted aspartic proteases 2	De Bernardis et al. (2012), Cassone and Casadevall (2012)	Abs	Mucosal and vaginal
Yeast derived- $\beta$ glucan particles (GPs)	Huang et al. (2010)	Th1, Th17, Abs	NT
DCs transfected with fungal RNA	Bozza et al. (2004), Perruccio et al. (2004)	T cell response Th1	Systemic
Nanoparticle-mediated target DCs	Roy and Klein (2012)	T cell response	NT
Liposome-mannan (Lmann)	Han et al. (2000)	Abs to $\beta$ -1,2-mannotriose	Systemic and vaginal
Daucosterol	Lim et al. (2007)	Th1	Systemic
Enolase	Li et al. (2011)	Abs Th1	Systemic
Glycopeptide or a peptide synthetic vaccine, ( <i>Candida albicans</i> cell wall-derived)	Xin et al. (2012)	Abs	Systemic
Fba, (peptide derived from fructose biphosphate aldolase which has cytosolic and cell wall distributions in the fungus)	Cutler et al. (2011)	Abs	Systemic
<i>C. albicans</i> dsDNA	Remickova et al. (2009)	T cells (Th1)	Gastrointestinal

NT, not tested.

**Figure 1.8: *C. albicans* antigens suggested as potential candidates for epitope-based vaccine approaches.** Figure adapted from Vecchiarelli et al., 2012.

Another interesting vaccine approach under development is PEV-7, a virosomal formulation containing *C. albicans* Sap2 that has been recently developed by Pevion Biotech. Sap2 is a *C.*

*albicans* secreted aspartyl protease that is able to hydrolyze epithelial proteins that are subsequently used by the pathogen as nutritional nitrogen sources. The recombinant Sap2 is encapsulated in virosomes, influenza-virus envelopes that potentiate immunological response. In mice and rats, PEV-7 delivered intramuscularly triggered a potent immunological response that lead to the production of vaginal anti-Sap2 IgG and IgA that conferred protection against post-immunization *C. albicans* challenge (De Bernardis et al., 2012). The protection was Sap-2 antibody-mediated and it was further confirmed by the fact that murine anti-Sap2 monoclonal antibodies conferred vaginal passive protection in the same animal model. Pevion Biotech recently completed a phase I clinical study to evaluate the safety and tolerability of the vaccine administered either intramuscularly or intravaginally. Although a paper detailing the results hasn't been published yet, the company claims that the study met both primary and secondary end points and that no serious adverse events were reported.

Although expectations for these vaccines are high, they will enter further clinical evaluations only if industrial interest and sufficient financial support will be raised. In addition, development of a vaccine for *C. albicans* is hampered by its remarkable ability to evade the host immune systems. Many researchers argue that a monovalent vaccine like the ones under development won't be effective against recurrent infections and only a vaccine that triggers against multiple, unrelated virulence factors has the potential of being successful (Cassone, 2013).

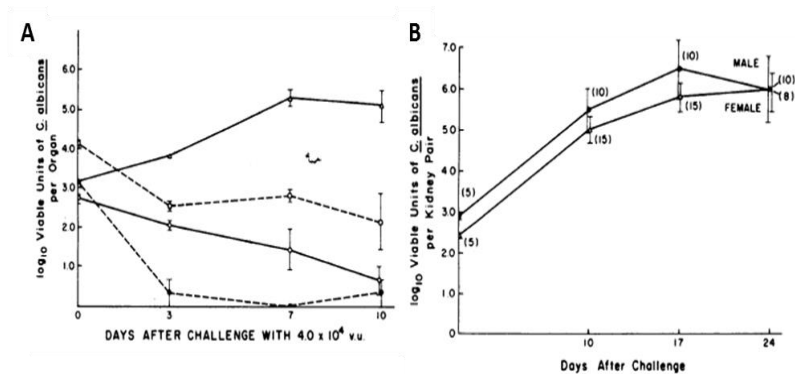
## 1.6 Model organisms to study *C. albicans* infections

Many *in vitro* approaches have been developed to study various aspects of *C. albicans* biology including filamentation, adhesion and phenotype switching. However these *in vitro* approaches can mimic only some aspects of the infection and it is imperative to use *in vivo* whole animal models to fully understand pathogenicity. Several model organisms have been used to study *C. albicans* infections and allowed researcher to explore various virulence aspects in simpler organisms. Because of the common descent of all living organisms and the conservation of metabolic pathways across species, many of the lessons learned in these simpler organisms were proven valid in humans. Model organisms also allow researchers to study both sides of the pathogenesis equation, the pathogen and the host. This approach provides insights on the conditions that the pathogen faces during infection and, at the same time, elucidate the mechanisms the host utilizes to counteract the infection. Model organisms commonly used to study fungal infection include vertebrate mammals like mouse, rats and *Danio rerio* (zebrafish) as well as invertebrate organisms like *Caenorhabditis elegans*, *Drosophila melanogaster*, *Galleria mellonella* (moth). Compared to studying the infection in clinical settings, using simpler organisms has many advantages, including the availability of genetic and molecular resources (knockout mutants, RNAi libraries etc) as well as costs and ethical considerations. In the next few paragraphs we summarized the most commonly used model organisms used to study *C. albicans* and their advantages and disadvantages.

### 1.6.1 Mammals

Mice and rats have been used to model various aspect of fungal infection for at least the last four decades. In 1976 Rogers and Balish were the first to demonstrate that mice challenged intravenously with *C. albicans* accumulate actively dividing yeast cells in the spleen, liver, lungs and kidneys up to 24 days post infection. The same paper also showed that:

- Kidneys are the preferential internal organs targeted by *C. albicans* followed by liver, spleen and lungs (Figure 1.9 A)
- There is no difference in susceptibility between male and female mice (Figure 1.9 B).



**Figure 1.9: *C. albicans* multiplies in mice and rats challenged intravenously.**

(A) Tissue specific time points of mouse challenged intravenously with  $4 \times 10^4$  *C. albicans* cells. Yeast cells are preferentially infecting the kidneys ( $\Delta$ ) followed by liver ( $\square$ ) spleen ( $\circ$ ) and lungs ( $\bullet$ ). (B) Course of infection in male ( $\bullet$ ) and female ( $\circ$ ) mice kidneys. (Figures adapted from (Rogers and Balish, 1976))

Rogers and Balish pioneering work opened up an entire new field of studies, and scientists have developed over the years many innovative models to mimic both systemic and epithelial *C. albicans* infections. Systemic models of infection are usually carried out in immunocompetent mice by lateral tail vein injections (Clancy et al., 2009). Epithelial infections are usually carried out in immunosuppressed mice using a variety of different strategies. These strategies include sublingual inoculation of the oral cavities with saturated

swabs to mimic oropharyngeal candidiasis (Solis and Filler, 2012), intravaginal injections to study *C. albicans* vaginitis (Chen and Kong, 2007) and force feeding (gavage) to study gastrointestinal colonization (Flattery et al., 1996). This repertoire of mammalian models has been instrumental to validate many of the hypotheses that have been formulated clinically. For example, many of the genes that have been shown to be upregulated in clinical samples have also been shown to be required for virulence in mouse models. For example Saps, a family of secreted aspartyl proteases, are upregulated during the infection, and their loss of function leads to a reduced pathogenicity in murine models of epithelial and systemic infection (Naglik et al., 2003a).

In addition to studying the pathogen's strategies used to infect, murine models are powerful tools to study mechanisms exploited by the host to cope with the infection. This is particularly true for mice, which have a genome remarkably similar to humans. The mouse genome was sequenced in 2002 (Waterston et al., 2002) and the study highlighted that:

- the mouse genome is comparable in size to the human one (2.5 Gb vs 2.9 Gb)
- 75% of the mouse ORFs are conserved in humans
- over 90% of the mouse and human genomes can be partitioned into corresponding regions of synteny.

The remarkable conservation of genes and metabolic pathways between these two species allows researchers to validate many clinical observations in this simpler model organism. For example, clinical observations suggested that epithelial infections occur more frequently in patients with defects in cell-mediated immunity, while systemic infections are more frequent in patients with deficiencies in neutrophil number or function. These observations were mirrored in mouse models of infections. The ability of mice to resist epithelial infections is T lymphocyte-dependent, and T lymphocytes play a central role in cell-mediated immunity

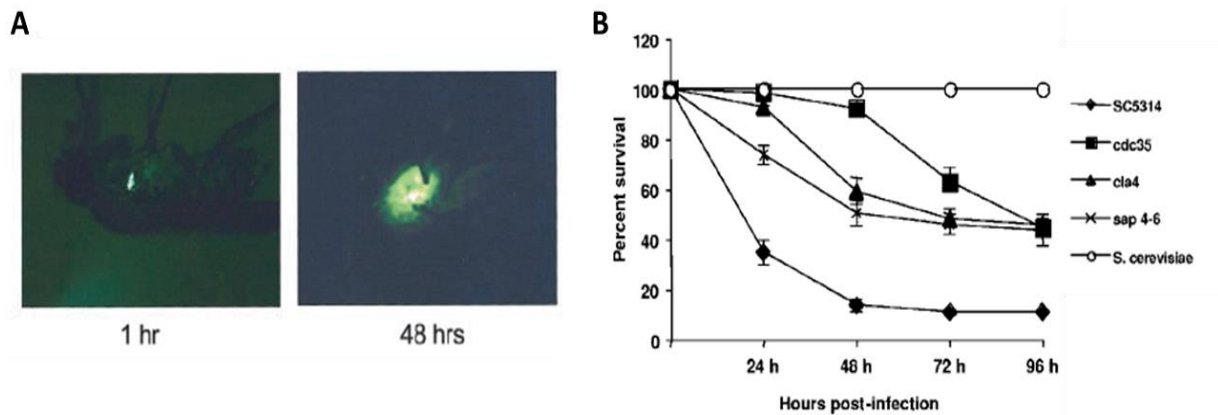
(Ashman et al., 2004). On the other hand, IL-17AR<sup>-/-</sup> mice that have an impaired ability to drive neutrophils to infected internal organs were significantly more susceptible to systemic but not epithelial infections (Huang et al., 2004). Taken together, these studies show the usefulness of using mice as model organisms to study *C. albicans* infection and validate hypotheses formulated either clinically or in other model organisms.

### **1.6.2 *Drosophila melanogaster***

*D. melanogaster* is a species of fly generally known as the common fruit fly. This fly has been widely used as a model organism to study various aspects of microbial pathogenesis. *D. melanogaster* has many advantages over vertebrate models: the fly life cycle is only 8-9 days, they produce a large number of externally laid embryos and they can be genetically modified in numerous ways. Interestingly, not only many individual genes have been conserved between humans and flies (Adams et al., 2000) but also entire cellular signal transduction pathways have been shown to be conserved between the two organisms (Rubin et al., 2000). In addition, *D. melanogaster*, although lacking an adaptive immunity, has an innate immune system with striking similarities with mammalian defense mechanisms (Hoffmann et al., 1999). Immune-deficient *D. melanogaster* has been shown to be susceptible to *C. albicans* infection but not to non-pathogenic baker's yeast *S. cerevisiae* (Alarco et al., 2004). Flies injected in the thorax with *C. albicans* have been shown to accumulate an overwhelming amount of *C. albicans* within hours post injection, suggesting that *C. albicans* can thrive inside the flies' body cavity (Figure 1.10 A). Interestingly, the ability of *C. albicans* to kill *D. melanogaster* requires the presence of *CDC35*, *CLA4*, *SAP4* and *SAP6*, four genes that have also been shown to be important in infections in murine models (Leberer et al., 1997;



Naglik et al., 2003a). This work shows that *C. albicans* genes essential for virulence in mammalian rodents are also critical for virulence in flies, validating the use of this insect model to study *C. albicans* infections.



**Figure 1.10: *D. melanogaster* is susceptible to *C. albicans*.**

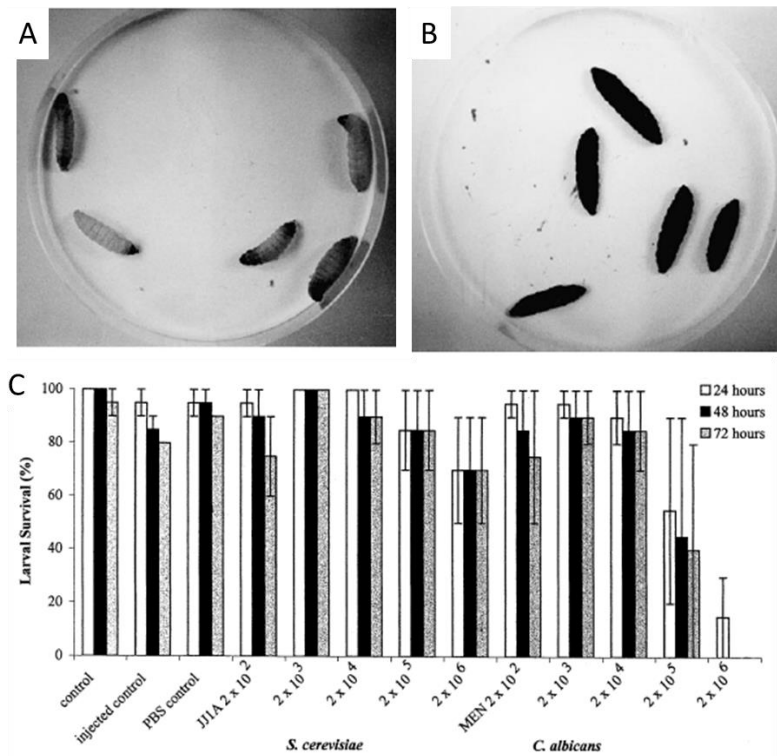
A) Flies injected with GFP-tagged *C. albicans* 1 and 48 hours post injection. B) Toll mutant *D. melanogaster* are sensitive to wildtype *C. albicans* (SC5314), less sensitive to *C. albicans* mutants deleted in key virulence factors (*SAP4-6*, *CLA4* and *CDC35*) and insensitive to *S. cerevisiae*. Figures adapted from (Alarco et al., 2004)

Alarco et al. (2004) pioneering work showed that *D. melanogaster* can be used as a powerful model to dissect host-pathogen interaction in a model of innate immunity that doesn't rely on the complexities of complement and antibody responses.

### 1.6.3 *Galleria mellonella*

In the last decade *Galleria mellonella* has been proposed as a model organism to study fungal infection. Although its genome has not been sequenced and there is not a well-established method to generate mutants, this model organism offers a couple of chief advantages. First, the greater wax moth is capable of supporting interactions with human pathogens at human physiological temperature (37°C). Second, *Galleria mellonella* can mount both cellular and humoral immune responses when the pathogen is injected in the hemocoel, the body cavity that contains hemolymph. Cells in the hemolymph are capable of phagocytosing microbial invaders and their kinetic and mechanism of killing are strikingly similar to human neutrophils (Bergin et al., 2005). For example, human proteins responsible for the production of superoxide via NADPH oxidase are conserved in moth and they play key roles in their immune systems (Bergin et al., 2005). Humoral responses to pathogen relies on the production of lysozyme and small antimicrobial proteins like cecropins (Ibrahim et al., 2006). When larvae were injected with  $2 \times 10^6$  *C. albicans* cells in the hemocoel only 15% and 0% of them survived 24 hours and 48 hours post challenge (Cotter et al., 2000). In contrast, when injected with non-pathogenic *S. cerevisiae*, 75% and 48% of them survived 24 and 72 hours post challenge (Figure 1.11). Interestingly, the authors of the paper also showed that *C. albicans* clinical isolates were more pathogenic than laboratory strains.

The pioneering work of (Cotter et al., 2000) opened up the possibility of using *G. mellonella* as a model organism to study *C. albicans* infection and many other papers substantiate and expanded their initial discoveries. For example (Brennan et al., 2002) showed that *C. albicans* mutants with reduced virulence in mice were also less virulent in *G. mellonella*.



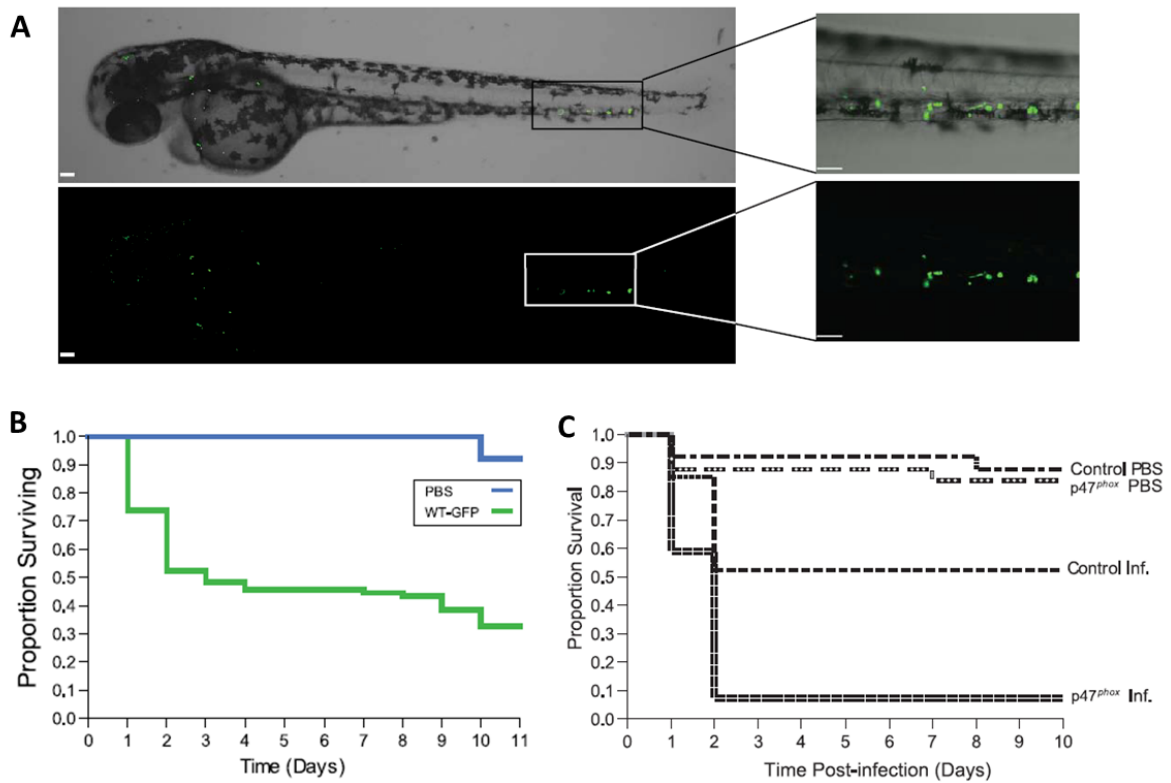
**Figure 1.11: *Galleria mellonella* is susceptible to *C. albicans*.**

A) *G. mellonella* larvae before inoculation with *C. albicans* B) *G. mellonella* larvae 48 hours post injection with  $2 \times 10^6$  *C. albicans* cells in the hemocoel. Typical discoloration of the cadavers due to melanization is visible. Melanization is believed to be a key component of the defense against pathogens and accumulates around microbes in the hemolymph (Kavanagh and Reeves, 2004) C) *G. mellonella* were injected with increasing concentration of wild type *S. cerevisiae* JJ1A or *C. albicans* MEN (clinical isolate from a patient with an eye infection) and survival rates at 24, 48 and 72 hours calculated. Figures adapted from (Cotter et al., 2000)

Taken together, the results demonstrate that *G. mellonella* can be a useful host model to study many aspects of *C. albicans* pathogenesis.

#### **1.6.4 *Danio reiro***

Zebrafish (*D. reiro*), is a small tropical fish originally isolated from Southeast Asia. Zebrafish has a unique combination of genetic and experimental advantages that make it an ideal organism to study many biological mechanisms. Zebrafish has a small generation time, it can lay hundreds of eggs from a single adult, it is optically transparent and its genome has been recently sequenced (Howe et al., 2013). Transient disruption of gene function can be easily achieved by microinjecting mRNA, DNA or short anti-sense oligomers called morpholinos. The availability of these technologies makes transient reverse genetic strategies very rapid, affordable and easy to achieve. A key factor that separates this model organism from others is the presence of both an innate and adaptive immune systems. Zebrafish have cytokines, macrophages, neutrophils, dendritic cells, mast cells, T cells, B cells and eosinophil that have remarkable similarities with humans (Sullivan and Kim, 2008). In recent years zebrafish has been used as a mini-vertebrate host model to study *C. albicans* infections. Brothers et al., 2011 showed that zebrafish embryos can be injected with *C. albicans* through the otic vesicle into the hindbrain. When injected in the hindbrain, *C. albicans* cause lethal disseminated infections. Remarkably, the mortality rates of these infections can be further exacerbated by knocking down zebrafish NADPH oxidase expression (Figure 1.12), suggesting that ROS play a major role in protecting zebrafish from *C. albicans* infections. The macrophage human respiratory burst is one of the first line of defense against *C. albicans*; the fact that the same mechanism is required for zebrafish resistance to *C. albicans* uncovers the promise of using this model to study *C. albicans* infections.



**Figure 1.12: Zebrafish is susceptible to *C. albicans*.**

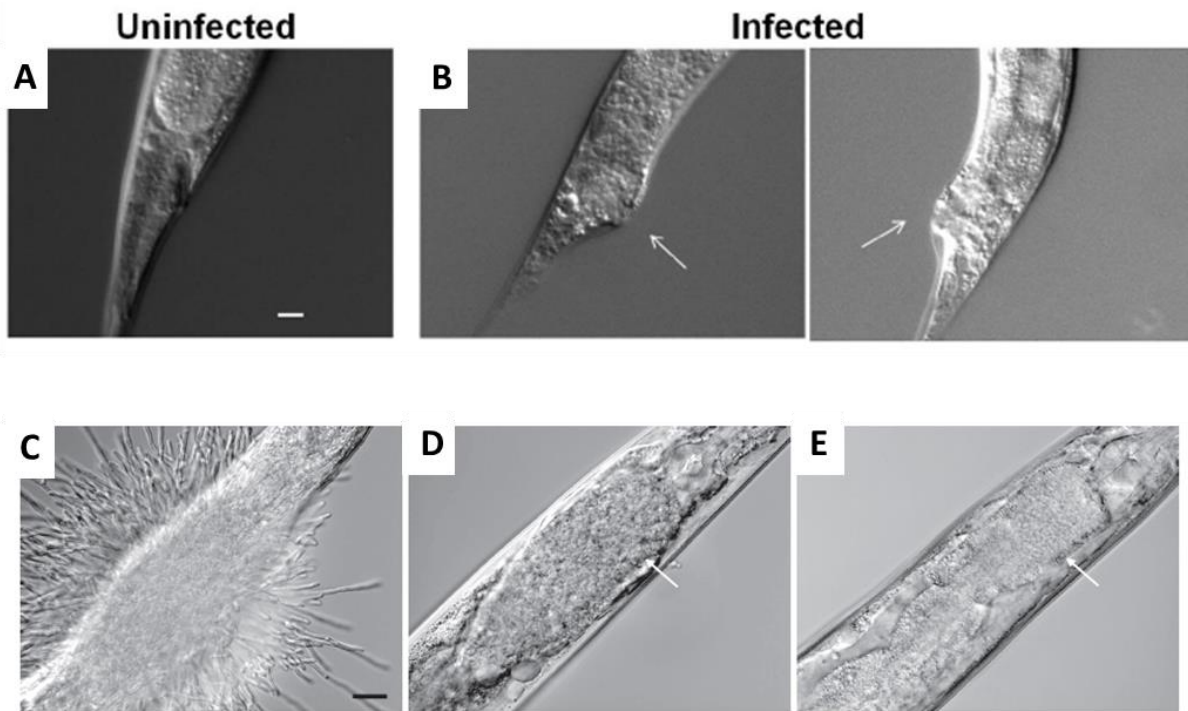
A) Confocal microscopy of zebrafish injected with GFP-expressing wildtype *C. albicans* 24 hours post injection. B) Within 5 days post infection, more than 50% of the zebrafish injected with *C. albicans* succumbed the infection. C) The zebrafish sensitivity was significantly enhanced when the p47<sup>phox</sup> subunit of the NADPH oxidase was knockeddown with p47<sup>phox</sup> morpholino (Control Inf. vs p47<sup>phox</sup> Inf.). Figures adapted from (Brothers et al., 2011)

Every model organisms proposed to study *C. albicans* infection has its own advantages and disadvantages. However *C. elegans* is the only one that doesn't rely on injections but just on the addition of the pathogen to its regular source of food. This makes *C. elegans* the desirable model organism to study *C. albicans* and is briefly described in the next paragraph and its has been extensively used through this thesis.

### ***1.6.5 Caenorhabditis elegans***

*C. elegans* is a transparent nematode that can be easily grown on *E. coli* lawns as the sole source of food. These nematodes are either self-fertilizing hermaphrodites or males, have a generation time of 3-4 days, a life span of 2-3 weeks and can be stored frozen almost indefinitely in liquid nitrogen. In addition, the developmental pattern of its 959 somatic cells has been traced through its transparent cuticle and its genome has been sequenced at high coverage. All of the above qualities, in addition with the availability of functional mutants, fluorescently labeled transgenic strains and RNAi knockdown libraries make *C. elegans* an extraordinarily powerful model organism. *C. elegans* has been used to study a variety of biological processes including neuronal development, cell division, apoptosis and host-pathogen interactions. Historically nematodes have been used to study bacterial infections like the one caused by *Pseudomonas aeruginosa* (Tan et al., 1999), *Staphylococcus aureus* (Sifri et al., 2003) or *Salmonella typhimurium* (Aballay et al., 2000) but it has also been used to study fungal pathogens like *Cryptococcus neoformans* or *Drechmeria coniospora* (Mylonakis et al., 2002). These studies have highlighted that many of the mechanisms used by the pathogen for virulence and by the host for immunity are analogous to the ones observed in human.

Recently our lab and others have shown that *C. elegans* is susceptible to *C. albicans* infections. When *C. albicans* is added to *E. coli* lawns it is ingested by the worms and colonizes the intestine. *C. albicans* infections cause in worms a distinct swelling of the anal region (Figure 1.13 A-B), a significantly shorter survival and, in certain conditions, *C. albicans* is capable of killing worms by piercing their cuticle (Jain et al., 2013b; Jain et al., 2009a; Pukkila-Worley et al., 2011; Pukkila-Worley et al., 2009).



**Figure 1.13: *C. elegans* is susceptible to *C. albicans*.**

A-B) Worms infected with *C. albicans* in solid media show a distinct swelling of the anal region (Jain et al., 2013) C-D-E) Worms infected with *C. albicans* and various other yeasts in liquid media. *C. albicans* is capable of infecting the nematode and piercing their cuticle (C) while other yeasts not capable of hyphal growth like *D. hansenii* (D) and *C. lusitaniae* (E) are not (Pukkila-Worley et al., 2009).

In addition, these studies have shown that many of the genes required for virulence in murine models of infection are also required for virulence in nematodes. For example *RIM101*, a *C. albicans* transcription factor required for alkaline-induced hyphal growth is required for both virulence in murine oropharyngeal candidiasis (Nobile et al., 2008) and virulence in *C. elegans* (Pukkila-Worley et al., 2009). Other genes like *NRG1*, *CAS5*, *ADA2*, *CPH1* and *EFG1* have also shown to be required for virulence in both mice and nematodes. These studies clearly showed that many of the conclusions drawn using this simpler organism remain valid in higher mammals, reiterating the amenability of using *C. elegans* for forward or reverse genetic screening. In addition, this model can be used for small molecule drug screening. (Okoli et al., 2009) developed a high-throughput semi-automated version of the *C.*

*elegans-C. albicans* model system and screened a library of 3,228 compounds. 7 of the 19 compounds capable of extending nematodes survival were known antimycotic validating their approach.

Taken together these studies show that *C. elegans* can be used as a powerful tool to study not only the complex host-pathogen dynamics but also for initial large scale drug screening.



## 2 A versatile *in vivo* model to study *C. albicans* virulence

### 2.1 Introduction

*C. elegans* has been used as a powerful model organism for now more than 50 years. In 1960s South African biologist Sydney Brenner pioneered the use of *C. elegans* to study neuronal development and paved the way for a long lineage of “*C. elegans* scientists” that studied various aspects of cell and animal biology in nematodes. This lineage includes Nobel Price Laureates Craig Mello and Andrew Fire for their RNAi work (Fire et al., 1998), Robert Horvitz and John Sulston for their work on organ development and apoptosis (Ellis and Horvitz, 1986; Hengartner et al., 1992; Sulston and Horvitz, 1977) and Martin Chalfie for his work on green fluorescent protein (Chalfie et al., 1994). Although this model organism has been traditionally used to study molecular and developmental biology, over the past 15 years, researchers have begun to use *C. elegans* to investigate the biology of various human pathogens including *Pseudomonas aeruginosa*, *Staphylococcus aureus*, *Salmonella enterica* and *Serratia marcescens* (Irazoqui et al., 2010; Kaletta and Hengartner, 2006; Kong et al., 2014; Marsh and May, 2012; Sem and Rhen, 2012). These papers revealed that many of the mechanisms involved in the human-pathogen interaction are conserved in nematodes, but also that there are some immunity mechanisms that are unique to this model organism (Bae et al., 2004; Kim et al., 2002). In nature, *C. elegans* encounters a variety of threats from ingested pathogens present in the soil and this has provided a strong selective pressure to evolve and maintain a sophisticated innate immune system in its intestinal lumen. Many of the genes and mechanisms involved in the protection of intestinal lumen are orchestrated by highly-conserved elements that exist also in higher mammals (Couillault et al., 2004; Kim et

al., 2002). *C. elegans* therefore represents a great model to study gastrointestinal pathogens like *Salmonella enterica* (Aballay et al., 2003), *Shigella boydii* (Kesika and Balamurugan, 2012) or *Vibrio cholera* (Cinar et al., 2010).

A pathogenesis assay using *C. albicans* and *C. elegans* has previously been described by our lab (Jain et al., 2013; Jain et al., 2009) and others (Okoli et al., 2009; Pukkila-Worley et al., 2011). Experiments described in these papers show that *C. albicans* cells are quickly ingested by the worms and accumulate in the intestinal lumen causing slower locomotion, distention of the intestine, swelling of the vulva, deformed anal region (Dar) and lethality. *C. albicans* infected worms live 10-12 days compared to 20-22 for uninfected controls.

Here, we expanded the use of the *C. albicans* - *C. elegans* assay and highlight its remarkable versatility. We show here that this assay can be reliably used for studying (a) host effectors important to counteract the infection, (b) pathogen determinants that drive pathogenicity (c) pharmacological compounds that can intervene in pathogenesis.

In order to elucidate the hosts factors that drive *C. elegans* ability to counteract *C. albicans* infections, we used a candidate-gene reverse genetic approach. Our analysis suggests that the MAPK (Mitogen-Activated Protein Kinase) pathway is required for *C. elegans* resistance to *C. albicans*. Functional knockouts in any of the main players in this pathway (*tir-1*, *pmk-1* and *sek-1*) are statistically significantly more susceptible to *C. albicans* infections compared to wild type controls. The MAPK pathway is a highly conserved signaling cascade that has been previously shown to be important in *C. elegans* for heavy metal, starvation resistance and neuronal development (Li et al., 2012; Mizuno et al., 2004). To the best of our knowledge, this is the first time that this pathway has been associated with the nematode's ability to counteract *C. albicans* infections. Interestingly, this pathway is conserved in humans and it has been shown to play a crucial role in protecting the human epithelium

against *C. albicans* (Moyes et al., 2010; Moyes et al., 2014). In humans this pathway plays a critical role in controlling the expression of proinflammatory cytokines like interleukins and tumor necrosis factors (TNF) upon *C. albicans* infection. Our experiments reiterate the usefulness of studying *C. albicans* infection in a simpler genetically tractable organism like *C. elegans*.

In addition to studying host effectors, we show in this chapter proof of principle experiments that demonstrate the amenability of using *C. elegans* to study *C. albicans* virulence. When we infected *C. elegans* with *C. albicans* double knock out mutants in *CPHI* and *EFG1*, two key virulence factors required for infection in mouse, rats and human epithelial models (Chen et al., 2006; Dieterich et al., 2002; Ricicova et al., 2010), nematodes survived significantly longer than controls. These experiments suggest that some of the lessons we learn in this simpler model about *C. albicans* virulence may remain valid in higher mammals and vice versa. These proofs of principle experiments encouraged us to use this model for a large scale *C. albicans* reverse genetic screen that is described in detail in chapter 3.

Finally in this chapter we show that our nematode model system can be pharmacologically modulated. In the presence of fluconazole, the most commonly prescribed antifungal drug, worms challenged with *C. albicans* survived significantly longer than control. This proof of principle experiment suggests that the nematode model can be used for small molecule screening. Indeed, as further discussed in chapter 6, our *C. elegans* model was instrumental in identifying filastatin, a small molecule inhibiting various aspects of fungal virulence (Fazly et al., 2013).

In summary we show here that our *C. elegans* - *C. albicans* model is a versatile and powerful tool that can be used not only to study the genetic bases that drive infection and immunity but also to identify new compounds for therapeutic intervention.

## 2.2 Materials and methods

### 2.2.1 Strains and media

Nematodes were maintained and propagated on *E. coli* OP50 as described here (Brenner, 1974). N2 worms were used as wildtype nematodes and we obtained loss of function mutants strains (summarized in Table 2.1) from the *C. elegans* genetic stock center (University of Minnesota).

Strains	Genotype	Mutagen	Reference
CB502	sma-2(e502) III	EMS	(Brenner, 1974)
NU3	dbl-1(nk3) V	Tc1 transposon	(Morita et al., 1999)
CB1370	daf-2(e1370) III	EMS	Deposited by Riddle D.L.
CF1038	daf-16(mu86) I	TMP + UV	(Lin et al., 1997)
KU25	pmk-1(km25) IV	TMP + UV	Deposited by Kunihiro Matsumoto (3/10/06)
AU1	sek-1(ag1) X	EMS	(Kim et al., 2002)
AU3	nsey-1(ag3) II	EMS	(Kim et al., 2002)
IG685	tir-1(tm3036) III	TMP + UV	Deposited by Shohei Mitani (5/25/08)
IG10	tol-1(nr2033) I	TMP + UV	(Pujol et al., 2001)
CB1482	sma-6(e1482) II	EMS	Deposited by H.R. Horvitz (8/11/87)

**Table 2.1:** *C. elegans* mutant strains used in this study

The *C. albicans* strains used in this study are wild-type SC5314 (Gillum et al., 1984) and CAN34 an *efg1/efg1 cph1/cph1* double mutant strain with the following genotype *ura3Δ::λimm434/ura3Δ::λimm434;cph1Δ::hisG/cph1Δ::hisG;efg1Δ::hisG/efg1Δ::hisG-URA3-hisG* (Lo et al., 1997).

### 2.2.2 *C. albicans* - *C. elegans* survival assay

5 mL *E. coli* and *C. albicans* liquid cultures were grown overnight respectively in LB at 37°C and YPD at 30°C. The *E. coli* culture was centrifuged at max speed for 1 minute in a microcentrifuge tube, washed twice in sterile PBS and resuspended to a final concentration of 200 mg/mL. The OD<sub>600</sub> of the *C. albicans* culture was measured and the culture diluted in PBS to a final OD<sub>600</sub>=3. 10 µ of a streptomycin solution 50 mg/mL was mixed with 2.5 µl of the *E. coli* culture obtained in the previous step to 0.5 µl of the *C. albicans* OD<sub>600</sub>=3 solution and to 7 µl of sterile water. The mixture was plated in the center of a small petri dishes (3.5 cm in diameter) containing NGM agar and incubated overnight at room temperature. The following day, 20 young synchronized adult worms were transferred to the spotted plate (day 0). Synchronous worm's populations were obtained 2-3 days after the addition of *C. elegans* eggs to NGM plates seeded with *E. coli* OP50. In this time frame, the eggs hatched and the larvae reached young adulthood. The worms are maintained at 20°C and scored daily by gentle prodding with a platinum wire. Dead worms were discarded, while live ones were transferred to new seeded plates grown overnight at room temperature. Live worms, dead worms and worms to be censored (worms accidentally killed while transferring or found dead on the edges of the plates) were scored daily. *C. elegans* survivals were examined using Kaplan-Meier method and statistical significance evaluated using the log-rank test (SPSS software). For the drug experiments we added fluconazole to a final concentration of 50 µM and reduced the amount of water added to the mixture.

## 2.3 Results

### 2.3.1 The *C. elegans* MAPK pathway is required for resistance to *C. albicans*

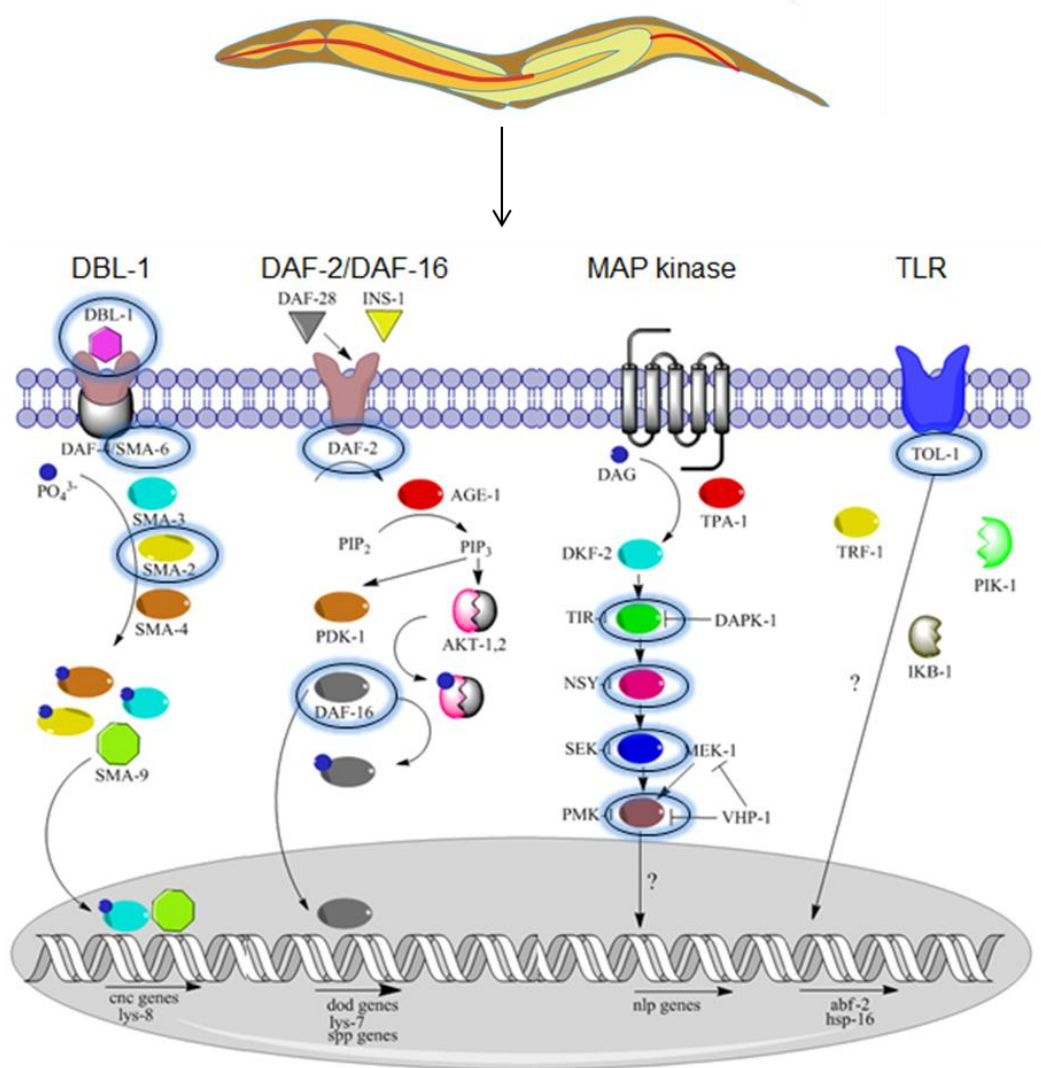
In this study we used a simple reverse genetic approach to pinpoint nematode host effectors that play central roles in mounting an immune response to *C. albicans* infections.

*C. elegans* lacks an adaptive immunity and relies on innate immunity to fight pathogens. Unlike other invertebrates, *C. elegans* not only lacks an adaptive immune systems but also doesn't appear to have specialized immune cells. For example, in contrast with *Drosophila* that has macrophage-like haemocytes that engulf invading pathogens, the only haemocytes present in *C. elegans* are six coelomocytes that don't seem capable of phagocytosis (Ewbank, 2002; Fares and Greenwald, 2001).

In *C. elegans*, four pathways have been previously implicated in immune response to human pathogens (Engelmann and Pujol, 2010):

- DBL-1 pathway, also known as TGF- $\beta$ , regulates the expression of key antimicrobial effectors like lysozymes and lectins (Mallo et al., 2002).
- DAF2/DAF16 pathway, also known as the insulin signaling pathway, is a general stress response pathway that is linked to the expression of multiple antimicrobial proteins like saponins, thaumatin-like proteins and lysozymes (Alper et al., 2007; Murphy et al., 2003).
- MAPK pathway, one of the most ancient signal transduction cascade in immunity shown to be involved in nematode resistance to Gram positive and Gram negative bacterial infections (Aballay et al., 2003; Huffman et al., 2004)
- TLR pathway, a signaling pathway required for the expression of heat-shock proteins and other defense-like molecules (Tenor and Aballay, 2008).

A schematic representation of these pathways and the main players involved in the signaling cascades is summarized in Figure 2.1. In order to test the individual contributions of these pathways in nematode resistance to *C. albicans*, we obtained multiple nematode mutants in these pathways and challenged them with wild-type *C. albicans*.



**Figure 2.1: Molecular pathways involved in *C. elegans* immunity and tested in this study.**

Mutants tested in this study for their roles in nematode resistance to *C. albicans* infections are circled in blue. *dbl-1*, *sma-6* and *sma-2* mutants were chosen to test the DBL-1 pathway; mutants in *daf-2* and *daf-16* to test the DAF-2/DAF-16; mutants in *tir-1*, *nsy-1*, *sek-1* and *pmk-1* to test the MAPK pathway and *tol-1* mutant to test the TLR pathway.

In the DBL-1 pathway, DBL-1 functions as a dose-dependent ligand that binds heterodimeric receptor DAF-4/SMA-6 triggering a phosphorylation cascade that involves SMAD proteins (SMA-2, SMA-3 and SMA-4) and that eventually leads to the activation of transcription factor SMA-9 (Figure 2.1). SMA-9 controls the expression of *lys-1* and *lys-8*, two genes coding for lysozyme, a hydrolytic enzyme that has long been recognized as key innate immune effector (Alper et al., 2007). In order to test the role of this pathway in *C. albicans* resistance, we obtained mutants in the three main components of the pathway, the extracellular ligand (*dbl-1*), the membrane receptor (*sma-6*) and their cytoplasmic target (*sma-2*).

The DAF-2/DAF-16 pathway is the *C. elegans* insulin/IGF-I signaling pathway, an evolutionarily conserved pathway that controls embryonic development, reproduction, fat storage and resistance to high temperature, osmotic and oxidative stress (Lamitina and Strange, 2005; Murphy and Hu, 2013). This pathway plays a critical role in *C. elegans* immunity by regulating the expression of hydrolytic enzymes like Lys-7 (lysozyme), Spp-1 (saposin) and Thn-2 (thaumatin) (Evans et al., 2008). In the absence of microbial threats, the insulin-like peptide DAF-28 binds DAF-2 and triggers a phosphorylation cascade that leads to the cytoplasmic retention of the transcription factor DAF-16. In the presence of a microbial threat, the pathway is repressed and DAF-16 phosphorylation prevented. In these conditions DAF-16 translocates to the nucleus, triggering the expression of antimicrobial peptides. To explore the role of this pathway we obtained mutants in *daf-2*, the receptor tyrosine kinase, and *daf-16*, the forkhead box O (FOXO) homolog that serves as the major transcription factor in the pathway.

The MAPK pathway is probably the most well characterized nematode immune pathway and it has been shown to have various biological functions ranging from coordinated locomotion



and asymmetric neuronal cell fate to resistance to *P. aeruginosa* and *D. conispora* (Kim et al., 2002; Pujol et al., 2008). The details of the first few steps of the cytoplasmic cascade are still under active debate, but they presumably involve the coordinated action of diacylglycerol (DAG) and the protein kinases TPA-1 and DKF-2 (Figure 2.1). Upon activation, DKF-2 activates TIR-1, NSY-1, SEK-1 and PMK-1 in a linear phosphotransfer cascade that leads to the activation of antimicrobial peptides like nlp-29 and nlp-31. The MAPK pathway is a highly conserved pathway and NSY-1, SEK-1 and PMK-1 are the *C. elegans* homologs of human ASK-1, MKK3 and p38 which play major roles in mammalian cellular immune response by controlling the expression of pro-inflammatory cytokines like IL-1 and TNF alpha. In order to evaluate the role of the MAPK pathway in resistance to *C. albicans* we obtained nematode mutants in each of the players involved in the cytoplasmic phosphorylation cascade: *pmk-1*, *nsy-1*, *sek-1* and *pmk-1*.

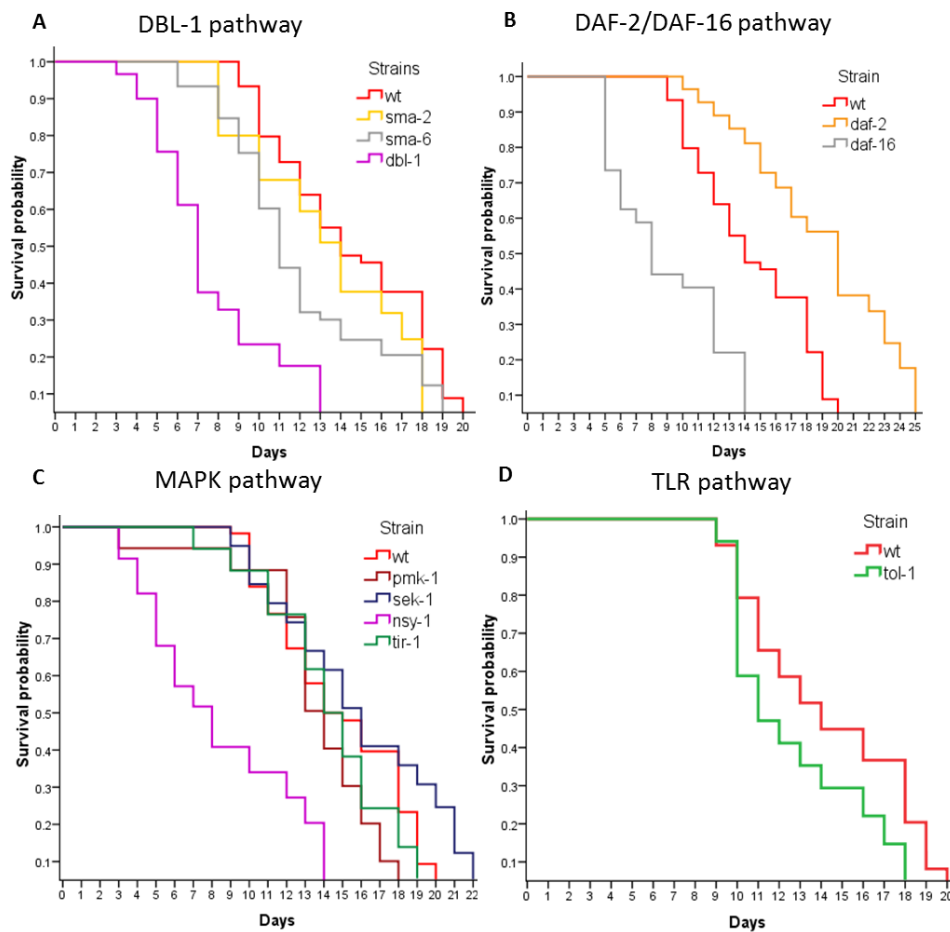
The TLR pathway in *C. elegans*, in contrast with *Drosophila* and mammals, doesn't act downstream of the MAPK pathway and its exact role in immunity is still under active debate. This pathway has a role in a specific avoidance-like behavior in the presence of *Serratia mercenses* and is essential to mount an appropriate immune response against *Salmonella enterica* (Tenor and Aballay, 2008). However, this pathway has also been shown not to be important for resistance to *Pseudomonas aeruginosa*, *Microbacterium nematophilum* and *Drechmeria coniospora* (Pujol et al., 2001). In addition, the TLR pathway hasn't been fully dissected and TOL-1 is the sole *C. elegans* Toll-like receptor (TLR). TRF-1, PIK-1 and IKB-1 are the homologues of the mammalian downstream signal transduction components but they have been shown not to act downstream of the TOL-1 in *C. elegans* (Tenor and Aballay, 2008). In order to evaluate the role of this pathway in resistance to *C. albicans*, we obtained the available mutant in *tol-1*.

Before performing our survival assay in the presence of *C. albicans* we evaluated the life span of each of the mutants. Mutations in key immune system modulators have the potential not only to severely compromise the nematode's immunity but also to shorten its life span. This would add an extra layer of complexity to the interpretation of our *C. albicans* survival results, creating a confounding element difficult to control experimentally. We therefore decided to performed our life span assay first and exclude from our analysis mutants with altered life spans. Life span assays were performed using heat killed *E. coli* as the nematode source of food (Tenor and Aballay, 2008). As shown in Figure 2.2, mutants in *sma-2*, *sma-6*, *pmk-1*, *sek-1*, *tir-1* and *tol-1* (mean time to death  $13.4 \pm 0.5$ ,  $12.1 \pm 0.5$ ,  $13.5 \pm 0.6$ ,  $15.9 \pm 0.7$ ,  $14.3 \pm 0.6$ ,  $12.7 \pm 0.6$  days) didn't show a statistically significant difference in longevity compared to wild-type (mean time to death  $14.6 \pm 0.5$  days). On the other hand, mutations in *dbl-1*, *daf-16* and *nsy-1* caused a significant shortening of nematode life span (mean time to death  $7.7 \pm 0.4$ ,  $9.1 \pm 0.6$ ,  $8.5 \pm 0.6$  days) and mutations in *daf-2* caused life span extension ( $19.1 \pm 0.6$  days).

Loss-of-function mutations in *dbl-1* were known to cause markedly reduced body size mutants (Suzuki et al., 1999) but, to the best of our knowledge, this is the first time that this gene has been implicated in nematode longevity. On the other hand the DAF-2/DAF-16 pathway has been known for its role in controlling life span and the increased and reduced life span of *daf-2* and *daf-16* mutants have been previously reported (Huang and Zhang, 2011; Kenyon et al., 1993).

The working hypothesis is that *daf-2* loss of function leads to the hyperactivation of the transcription factor DAF-16 which controls the expression of antimicrobial peptides *spp-1*, *spp-12* and *lys-7*. The overexpression of these genes leads to an extended life span as well as an increases resistance to pathogens. This hypothesis is corroborated by the fact that *daf-16*

loss of function leads to abrogation of the expression of these genes causing a reduced fitness and shorter life span.

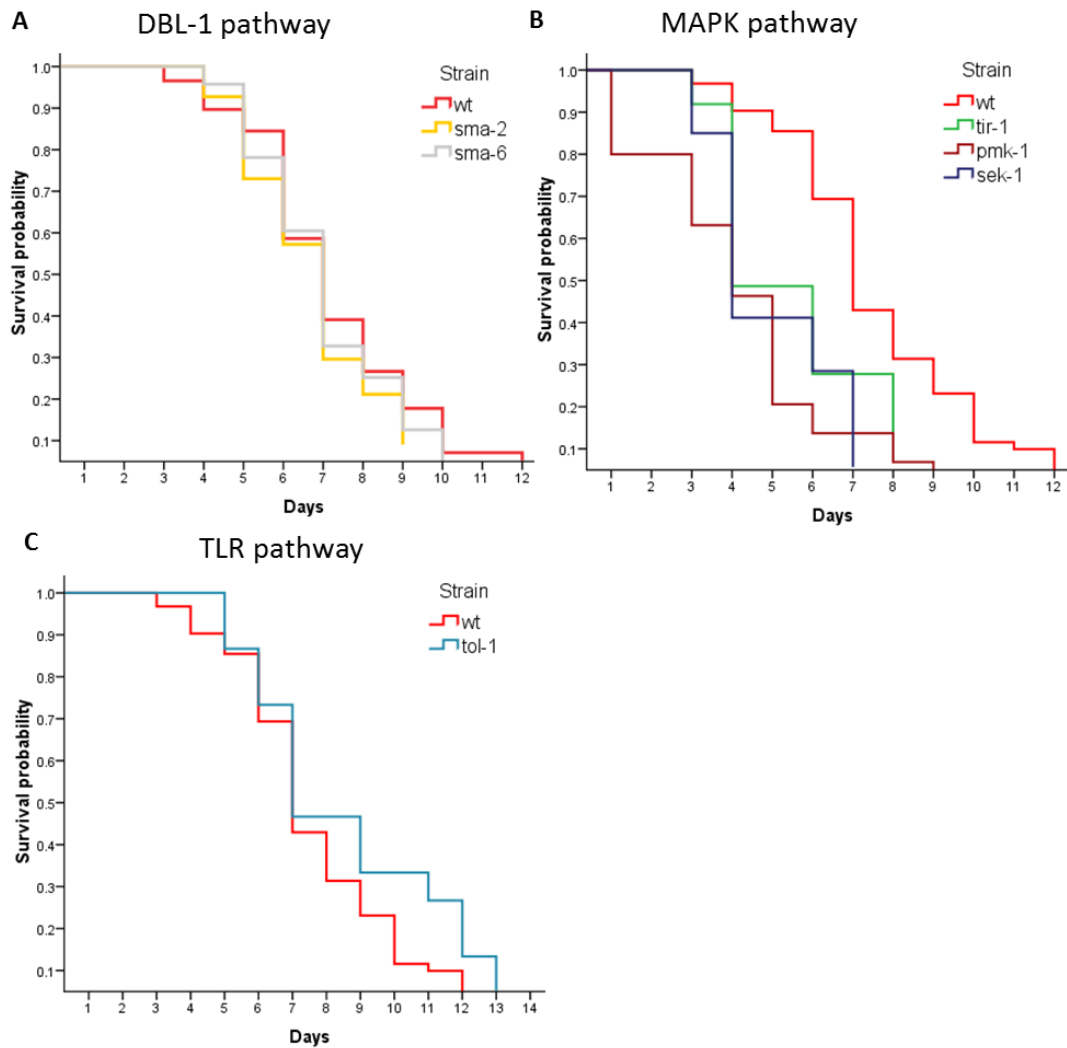


**Figure 2.2 *C. elegans* mutants' life span assay with heat killed *E. coli***

A) DBL-1 pathway: mutants in this pathway live just as long as wild-type except for *dbl-1* which is short-lived B) DAF-2/DAF-16 pathway: mutants in this pathway showed statistically significant different life spans compared to wild-type C) MAPK pathway: only mutants in *nsy-1* show a reduced life span compared to wild-type D) *tol-1* mutants have life spans comparable to wild-type. Survival curves represents average of three biological replicates (n = 60).

Finally, *nsy-1* loss of function leads to a reduced life span, indicating that this gene is not only required for chemotaxis, egg laying, pathogen response and anoxia resistance (Kim et al., 2002; Troemel et al., 1999) but also for longevity. All of the mutants that showed a statistically significant difference life span compared to wild-type were excluded from our subsequent analysis.

Mutants with life spans comparable to wild type were challenged with live *C. albicans*. As shown in Figure 2.3, mutants in either the DBL-1 or TLR pathway lived just as long as controls (mean time to death for *sma-2*, *sma-6*, *tol-1* and wildtype  $6.7 \pm 0.2$ ,  $7.0 \pm 0.3$ ,  $8.6 \pm 0.5$  and  $7.2 \pm 0.3$  days respectively). However, when we tested the MAPK pathway we observed that, remarkably, mutants in any of the major players in this pathway (*tir-1*, *pmk-1* and *sek-1*) were significantly more susceptible to *C. albicans* infection compared to controls ( $p < 0.01$  at both the Log Rank and Breslow statistical test). This evidence suggests that the MAPK pathway plays a crucial role in resistance to *C. albicans* infections. Interestingly, when mutants in the MAPK pathway were infected, we didn't observed a faster or more severe accumulation of *C. albicans* in their intestinal lumen, indicating that these mutations don't impair the nematode's pharynx grinding ability or cause any gross structural changes to the sites of infection. In addition, infected mutants in this pathway didn't show any substantial difference in the tail region swelling (Dar phenotype), suggesting that MAPK mutants' shorter survival depends on immune mechanisms independent from the Dar response.



**Figure 2.3: MAPK pathway is required for *C. elegans* resistance to *C. albicans***

A) DBL-1 pathway: mutants in this pathway don't show increased sensitivity to *C. albicans* infection B) Mutations in any of the major players in the MAP kinase pathway resulted in increased sensitivity to *C. albicans* infections C) *tol-1* mutants don't show increased sensitivity to *C. albicans* infections. Survival curves represents average of three biological replicates (n = 60).

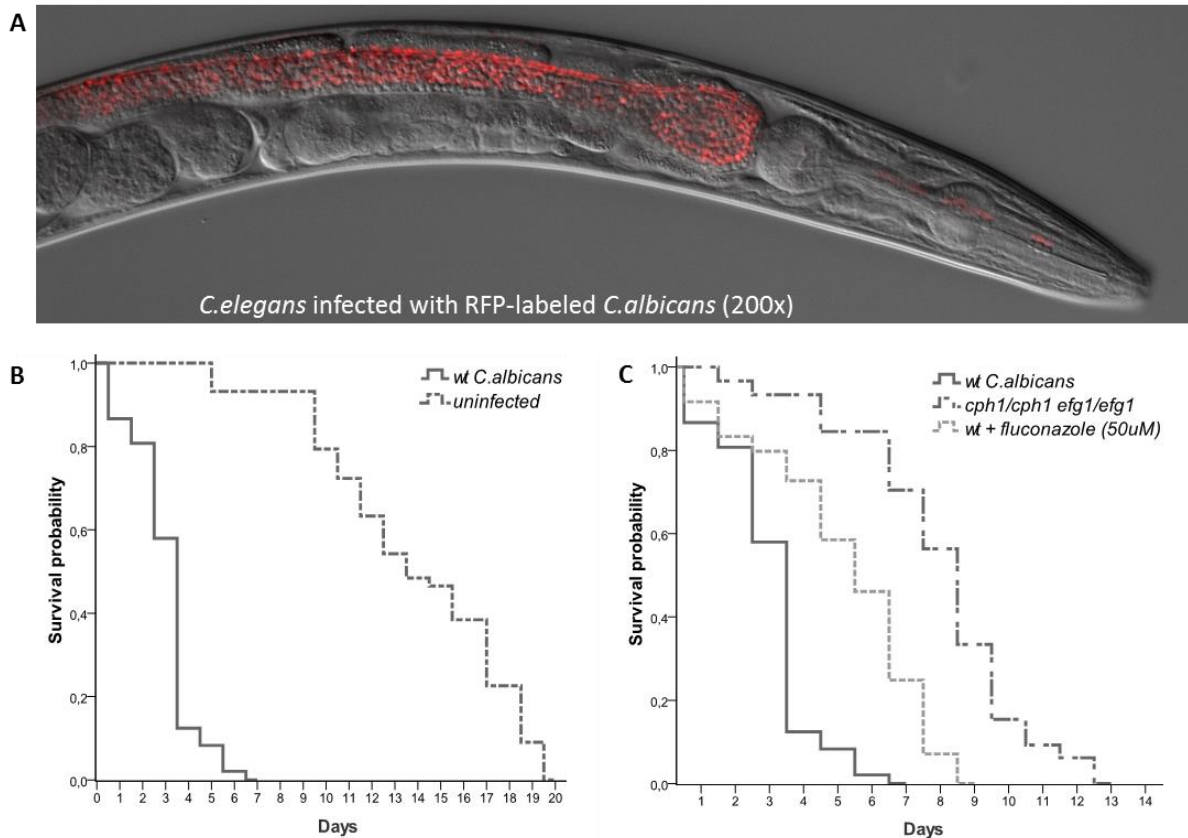
These results suggest that the MAPK pathway has a direct role in nematode defense response to *C. albicans*. These findings were recapitulated by others in mammalian models of infection (Moyes and Naglik, 2011) and reiterate the usefulness of modeling *C. albicans* infections in a simpler genetically tractable model like *C. elegans*.

### 2.3.2 Genetic and pharmacological modulation of the *C. elegans* – *C. albicans* survival assay

In this study we tested the ability of the *C. albicans* - *C. elegans* model system to respond to either genetic or pharmacologic modulation of the yeast.

We first tested the ability of the worms to survive infection caused by *C. albicans efg1/efg1 cph1/cph1* double mutant. Efg1 is a highly-conserved transcription factor and an essential component of the cyclic AMP/protein kinase A (PKA) metabolic pathway. In *C. albicans* this pathway regulates hyphal morphogenesis (McDonough and Rodriguez, 2012), white-opaque switching (Sonneborn et al., 1999) and an arsenal of key virulence factors. These virulence factors include Hwp1 and Hwp2, two yeast-specific cell wall proteins involved in adhesion and biofilm formation, Eap1 a cell-wall adhesin involved in binding to human epithelial cells (Li and Palecek, 2003) and Sap5, an hydrolytic enzyme involved in epithelial tissue invasion (Staib et al., 2002). On the other hand Cph1 is a transcription factor that regulates many metabolic processes including mating, filamentation and biofilm formation and it has been shown to play a critical role in damaging epithelial cells (Korting et al., 2003) and human reconstructed epithelium (Dieterich et al., 2002). Disruption of either of these genes has a significant impact on virulence and simultaneous disruption of both in *cph1/cph1 efg1/efg1* results in dramatic virulence reduction in various animal models including mice (Lo et al., 1997), *Drosophila* (Chamilos et al., 2006), *zebrafish* (Brennan et al., 2002; Brothers et al., 2011) and *moth* (Brennan et al., 2002). *cph1/cph1 efg1/efg1* double mutants is considered by the *C. albicans* community as the a-virulent strain by definition and the gold standard for validation of novel model systems. As shown in Figure 2.4, worms infected with *cph1/cph1 efg1/efg1 C. albicans* double mutant lived statistically significantly longer than controls ( $p <$

0.01 at both Log Rank and Breslow statistical test,  $n = 60$ ) suggesting that these two genes are required for *C. albicans* virulence against *C. elegans*. These proofs of principle experiments suggest that our nematode model can be used to study *C. albicans* virulence and encouraged us to undertake a large scale *C. albicans* reverse genetic screen that is described in chapter 3.



**Figure 2.4: *C. albicans* - *C. elegans* model can be pharmacologically or genetically modulated.**

A) *C. albicans* accumulation in the nematode intestinal lumen day 3 post infection. Yeast cells are quickly ingested by the worms and accumulate in the intestinal lumen completely intact indicating that they are able to survive the mechanical crushing of the pharynx. B) survival curves of nematodes challenged with *C. albicans* versus uninfected controls C) survival curves of nematodes challenged with either wild-type *C. albicans*, *cph1/cph1 efg1/efg1* double mutant or wild-type *C. albicans* + 50  $\mu$ M of fluconazole

In order to explore the possibility of using our *C. elegans* model for potential drug screening, we tested the effect of the addition of fluconazole to the ultimate outcome of the infection.

We used a variety of different concentrations of fluconazole and found that 50  $\mu\text{M}$  gave us the most significant results. As shown in Figure 2.4, worms infected with *C. albicans* and 50  $\mu\text{M}$  of fluconazole lived statistically significantly longer than controls ( $p < 0.01$  at both Log Rank and Breslow statistical test,  $n = 60$ ). Interestingly we also found that lower fluconazole concentrations (below 25  $\mu\text{M}$ ) didn't prolong survivals in a statistically significant fashion and that higher concentrations (above 100  $\mu\text{M}$ ) led to reduced survival suggesting that fluconazole might be toxic for the nematodes over a certain concentration. These proof of principle experiments showed that our nematode model can be used for small molecule antifungal screening. The model was in fact instrumental in the discovery of filastatin, a small molecule inhibiting various aspects of fungal virulence that is currently undergoing further preclinical studies and that is described in chapter 6 (Fazly et al., 2013).

## 2.4 Discussion

We used here a gene candidate reverse genetic approach to identify *C. elegans* immune determinants important for nematode resistance to *C. albicans*. We have shown that mutations in any of the genes involved in the MAPK pathway lead to increased sensitivity to *C. albicans* infections. In mammals, this pathway has been implicated in a variety of different cellular immune responses against *C. albicans* infections. This pathway is critical for *C. albicans* phagocytosis by murine macrophages (Ibata-Ombetta et al., 2001), for the macrophage ability to discriminate between live and heat inactivated *C. albicans* (Martinez-Solano et al., 2009) and for the human epithelia's ability to remain quiescent under low fungal burden while responding strongly upon damage-induced *C. albicans* penetration (Moyes et al., 2010; Moyes et al., 2014). Taken together, these results suggest that the MAPK



pathway plays a critical role in immunity against *C. albicans* in both mammals and nematodes suggesting that some of the lessons we learn in this simpler, genetically-tractable model may hold true in humans. We have also shown here proof of principle experiments that demonstrate that our nematode model can be genetically and pharmacologically modulated and can be used to study pathogen's determinants and for small molecule drug screening.

In conclusion, our model is a very versatile tool that can be used not only to study the contributions of individual host-pathogen genes using forward and reverse genetic but also for exploring new molecules for therapeutic intervention.

### **3 Reverse genetic screen identifies *ZCF15* a novel virulence factor**

#### **3.1 Introduction**

*Candida* infections represent a serious public health threat. With an estimated 46,000 infections per year and direct costs between 1\$ and 2\$ billion, they represent the fourth most common cause of healthcare-associated bloodstream infections in the United States (Rentz et al., 1998; Wilson et al., 2002; Wisplinghoff et al., 2004). These infections can be caused by many *Candida* yeasts species, the most common of which is *Candida albicans*. This polymorphic fungus is part of the human microbiome and it usually lives as a harmless commensal on the skin and mucosa of healthy individuals. However, under certain circumstances, it can cause infections ranging from mild superficial infections to life-threatening systemic infections. Although in the last two decades a large body of literature shed some light on *C. albicans* pathogenicity (Mayer et al., 2013), our understanding of the basic mechanisms that lead this harmless commensal to become pathogenic remain elusive. In addition, because drug resistant has dramatically increased the last two decades (Cowen, 2008), there is an urgent need to identify new targets to develop better diagnostics and therapeutics.

In order to find novel genes involved in virulence, a team of four undergraduate students in our lab screened a large library of *C. albicans* mutants for their abilities to elicit the Dar response in *C. elegans*. This reverse genetic approach identified four novel virulence factors, and in this chapter we will focus on *ZCF15*, a previously uncharacterized Zn(II)<sub>2</sub>Cys<sub>6</sub> DNA binding protein that, unlike the other three, doesn't have a human homolog. This feature is

critical because in order to develop effective antimycotics with high selective toxicity it is imperative to target fungal specific genes or metabolic processes not conserved in humans.

Here we show that *zcf15/zcf15* homozygous knockout not only is significantly less capable of inducing Dar as shown by the team of undergrads, but it is less capable of killing worms, survive mouse macrophages *ex vivo*, and withstanding reactive oxygen species (ROS) *in vitro*.

We used various experimental approaches to demonstrate that the decreased virulence of the *zcf15/zcf15* knockout is likely due to its inability to withstand ROS, a common host defense mechanism. The *C. albicans zcf15/zcf15* knockout is hypo virulent in the nematode model except in worms whose ability to produce ROS has been compromised. Results in the worm model were recapitulated in mouse macrophage cocultures, further supporting our initial findings. In particular, the *C. albicans zcf15/zcf15* knockout is hypersusceptible to macrophage killing except in macrophage in which the ability to produce ROS has been pharmacologically inhibited. Overall, this evidence indicates that *ZCF15* plays a key role in *C. albicans* virulence by contributing to the pathogen's ability to withstand the host's generated ROS. Deletion of *ZCF15* has a limited impact in a disseminated murine bloodstream model of infection, suggesting that additional layers of regulation exist in the context of the murine bloodstream.

Taken together, these results suggest that *ZCF15* is an important virulence determinant and plays a critical role in resistance to host generated ROS. Insights on its precise mechanism of action is elucidated in chapter 4.

## 3.2 Materials and methods

### 3.2.1 Strains and media

*C. albicans* mutants screened in this study and relative references are summarized in Table 3.1. A total of 724 transposon insertion mutants were screened: 84 protein kinase mutants, 98 cell wall mutants, 234 transcription factors mutants and 308 mutants in various other genes (random mutagenesis approach used).

Genes targeted	Fungal genetic stock center name	Number of mutants screened	Reference
Kinases	Kinase set from A. Mitchell	84	(Blankenship et al., 2010)
Cell Wall	CWP plates 1-2	98	(Norice et al., 2007)
Transcription factors	CJNF TF	84	(Nobile and Mitchell, 2005a)
Transcription factors	2 <sup>nd</sup> Set TF	79	(Rauceo et al., 2008)
Transcription factors	Dominique Sanglard TF	71	(Sanglard lab, unpublished)
Random	VB GKO plates 1-5	215	(Davis et al., 2002)
Random	Misc	58	(Sanglard lab, unpublished)
Random	SFY	35	(Sanglard lab, unpublished)

**Table 3.1:** *C. albicans* transposon mutant library description and relative references.

Strains were obtained from the Fungal Genetic Stock Center and routinely grown on YPD (Yeast-Peptone-Dextrose) liquid or agar plates.

To validate the results we obtained from the transposon insertion mutant library screen we obtained a series of homozygous deletions and complemented strains that are described in Table 3.2.

Strain	Phenotype	Genotype	Reference
SN250	Arg-	<i>his1Δ/his1Δ, leu2Δ::C.dubliniensis HIS1 /leu2Δ::C.maltosa LEU2, arg4Δ /arg4Δ, URA3/ura3Δ::imm<sup>434</sup>, IRO1/iro1Δ::imm434</i>	Noble et al., 2010
<i>zcf15/zcf15</i>	Arg-	<i>his1Δ/his1Δ, leu2Δ /leu2Δ, arg4Δ /arg4Δ, URA3/ura3Δ::imm<sup>434</sup>, IRO1/iro1Δ::imm434, orf19.2753Δ::C.dubliniensisHIS1/orf19.2753Δ::C.maltosaLEU2</i>	Noble et al., 2010
<i>dot4/dot4</i>	Arg-	<i>his1Δ/his1Δ, leu2Δ /leu2Δ, arg4Δ /arg4Δ, URA3/ura3Δ::imm<sup>434</sup>, IRO1/iro1Δ::imm434, orf19.3370Δ::C.dubliniensisHIS1/orf19.3370Δ::C.maltosaLEU2</i>	Noble et al., 2011
<i>orf19.1219/orf19.1219</i>	Arg-	<i>his1Δ/his1Δ, leu2Δ /leu2Δ, arg4Δ /arg4Δ, URA3/ura3Δ::imm<sup>434</sup>, IRO1/iro1Δ::imm434, orf19.3370Δ::C.dubliniensisHIS1/orf19.3370Δ::C.maltosaLEU3</i>	Noble et al., 2012
<i>zcf15/ZCF15c (LEU2)</i>	Prototrophic	<i>his1Δ/his1Δ, leu2Δ /leu2Δ::ORF19.2753-C.dubliniensisARG4, arg4Δ /arg4Δ, URA3/ura3Δ::imm<sup>434</sup>, IRO1/iro1Δ::imm434, orf19.2753Δ::C.dubliniensisHIS1/orf19.2753Δ::C.maltosaLEU2</i>	This work
<i>zcf15/ZCF15c (End)</i>	His-	<i>his1Δ/his1Δ, leu2Δ /leu2Δ, arg4Δ /arg4Δ, URA3/ura3Δ::imm<sup>434</sup>, IRO1/iro1Δ::imm434, orf19.2753Δ::C.dubliniensisHIS1::ORF19.2753-C.dubliniensisARG4/orf19.2753Δ::C.maltosaLEU2</i>	This work

**Table 3.2:** *C. albicans* homozygous knockout mutants and complemented strains used in this study.

To complement the *zcf15/zcf15* knockout we reintegrated *ZCF15* either at the endogenous locus (End) or at an ectopic locus (*LEU2*) (Table 3.2). In the former case the expression is controlled by the endogenous regulatory sequences, in the latter by the *LEU2* promoter. For endogenous reintegration, we first PCR amplified a fragment containing the *ZCF15* coding sequence plus ~700 bp upstream using primer 548 and 555 (Table 3.3). As a template for PCR we used genomic DNA extracted from wildtype strain using standard phenol:chloroform extraction (Hoffman and Winston, 1987). The PCR product was BamHI-AflII digested and cloned into a BamHI-AflII digested pSN69 (Noble and Johnson, 2005) in order to obtain plasmid pLI101. We subsequently amplified a ~700 bp fragment downstream of the *ZCF15* open reading frame using primers 554 and 556. The PCR product

was XhoI-XbaI digested and cloned it into a XhoI-XbaI digested pLI101 to obtain plasmid pLI102. Finally we digested pLI102 with BamHI and XbaI to liberate an integrating fragment made of *ZCF15* - *C. dubliniensis*ARG4 flanked by ~700bp upstream and downstream of *ZCF15*. The integrating fragment was transformed into *zcf15/zcf15*- following the LiAc protocol (Wilson et al., 1999) and plated in media lacking arginine. Arg<sup>+</sup> colonies were restreaked on SC-ARG plates to confirm their true Arg<sup>+</sup> nature and then grown overnight in SC-ARG liquid medium. Genomic DNA was extracted from these cultures using standard phenol:chloroform procedures and integration of *ZCF15* was confirmed by PCR using primers designed to check the presence of the new 5' and 3' junctions (Table 3.3, 583-584 for the 5' junction and 585-586 for the 3' junction).

Primer number	Primer sequence
548	TTGCAAGGATCCTCTGGTCCCTCGGATCAACAAGAT
555	GCTCGACTTAAGTATTTATTTAATCAATATATGTAAAAAAGA GTCT
554	GTCGATCTCGAGAGACTCTTTTTTTACATATATTGATTAAATAA ATA
556	GATCATTCTAGAACAAGCACAAATGTTATTCCATGC
583	CCAAAGAGTTAGCTTACTAATTACC
584	GGATTGTGCGTCGATCAATAA
585	ACGAAGGTCACACTGACTTATG
586	ACTGATGATGAGTTGCGATATGA
587	GTAAAATTACAATTGGTATTTTCAACCAATATTACCACACAGG AGGAGGAGGACAAGAAA
589	CTATCTCTCTTTTTTTTGCCATTTGAGTTAGTGCATGACATCA AACATCAGCAGGAGGA
600	GTAACGACGCCAAATCTCAATAC
568	TGGGTTTCCTTCAGCATCAA
601	GTATCGCCTTTGGGATGTCTAT
602	CAATAACTACCTCGGCACCTT

**Table 3.3:** primers used in this study.

To complement *ZCF15* in the *LEU2* locus we leveraged the power of gap repair cloning in *S. cerevisiae* (Bessa et al., 2012). We first amplified *ZCF15* and its ~600 bp upstream and ~600 bp downstream with primers 587 and 589 from genomic SN250. This fragment was digested with BmgBI and co transformed with BmgBI digested pSN105 (Noble et al., 2010) into *S. cerevisiae* BY4741 (Brachmann et al., 1998) using standard LiAc protocol. The primers have been designed so that the PCR product is flanked by 40 bp that are homologous to the extremities of the digested plasmid. The yeast homologous recombination machinery recognized the homologous sequences and ligated them giving the strains the ability to grow in media lacking uracil. Colonies growing in media lacking uracil were restreaked in SC-URA plates and subsequently grown overnight in liquid culture lacking uracil. The “gap repaired” plasmid (pLI103) was extracted using the QIAprep mini kit (Qiagen) and the presence of *ZCF15* was confirmed by both PCR and diagnostic restriction double digestion. pLI103 was electroporated into *E. coli* DH5 $\alpha$  and plated in LB media + 100 $\mu$ g/mL ampicillin. After plasmid mini prep, pLI103 was digested with PmeI to liberate the *ZCF15*-*ARG4* inserting cassette flanked by the ~500 bp fragments upstream and downstream of the *LEU2* open reading frame. The digested plasmid was transformed into *zcf15/zcf15* following the LiAc protocol and plated in media lacking arginine. Arg<sup>+</sup> colonies were restreaked on SC-ARG plates, grown overnight in SC-ARG liquid and genomic DNA was extracted using standard phenol:chloroform procedures. Proper integration of *ZCF15* in the *LEU2* locus was confirmed by checking the presence of the novel 5' and 3' junction using primers 600-568 and 601-602 respectively (Table 3.3).

### **3.2.2 *C. albicans* - *C. elegans* deformed anal region (Dar) assay**

2-3 *C. elegans* N2 young adults were transferred to an NGM agar plate seeded with *E. coli* OP50 and incubated at 20°C for 4-5 days to allow generation of a large number of gravid adult worms. The plate was washed with 10 mL of buffer M9, transferred to a 15 ml conical tube and centrifuged for 2 minutes at 900 g. Supernatant was discarded and pellet resuspended in 15 mL of 0.25 M NaOH plus 5.25% commercial bleach. The mixture was mixed gently for 3 minutes. During this time lapse, worms are lysed and only eggs survive the treatment. Eggs were isolated by centrifugation at 2000 g for 2 minutes. Once the supernatant was discarded, the pellet was washed in sterile water and the eggs finally resuspended in M9 to a final concentration of 5-6 healthy eggs/ $\mu$ L.

*E. coli* OP50 and each of the *C. albicans* strains were grown overnight respectively in LB at 37°C and YPD at 30°C. The *E. coli* and *C. albicans* cultures were centrifuged at full speed in a micro centrifuge and the pellets resuspend in PBS to a final concentration of 200 mg/mL and 10 mg/mL respectively. 10 $\mu$ L of streptomycin solution 50 mg/mL were mixed with 2.5  $\mu$ L of the *E. coli* and 0.5  $\mu$ L of *C. albicans* cultures. The strep-*E. coli*-*C. albicans* mixture was plated in the center of a NGM agar petri dish (6 cm diameter) together with 25-30 healthy *C. elegans* eggs (plated to the side of the Strep-*E. coli*-*C. albicans* mixture). The plates were incubated at 20°C and the percentage of worms showing Dar manually scored under the dissecting microscope day 4 post spotting.

### **3.2.3 *C. albicans* - *C. elegans* survival assay**

*E. coli* OP50 and *C. albicans* were grown overnight respectively in LB at 37°C and YPD at 30°C. The *E. coli* culture was centrifuged at full speed and the pellet resuspended to a final



concentration of 200 mg/mL in sterile water while *C. albicans* strains were diluted in sterile water to  $OD_{600}=3$ . 10 $\mu$ l of streptomycin solution 50 mg/mL were mixed with 2.5  $\mu$ l of the *E. coli* and 0.5  $\mu$ l of *C. albicans* culture. The strep-*E. coli*-*C. albicans* mixture was plated on NGM agar small petri dishes (3.5 cm diameter) and incubated overnight at 20°C. 20 young synchronized *C. elegans* N2 were added to the spotted plate. Plates were incubated at 20°C and worms were scored daily by gentle prodding with a platinum wire. Dead worms were discarded, while live ones were transferred to new plates seeded with the strep-*E. coli*-*C. albicans* mixture and grown overnight at room temperature. Worms accidentally killed while transferring or found dead on the edges of the plates were censored in our analysis.

### **3.2.4 *In vitro* phenotypic characterization**

For growth rate experiments, three independent colonies from each *C. albicans* strain were inoculated in YPD and grown overnight at 30°C. Cultures were resuspended to  $OD_{600} = 0.1$  in fresh YPD and transferred to a microplate reader (Bio-Tek Synergy H4) that measured  $OD_{600}$  every 15 minutes for a total of 48 hours. Growth curves were generated by averaging the individual reads and plotting them over time.

To quantify the strain's ability to form biofilm we essentially followed the protocol detailed here (Reynolds and Fink, 2001). Briefly, liquid cultures were grown in SC + 0.15% glucose at 30° C overnight, resuspended to  $OD_{600} = 0.5$  in fresh SC + 0.15% glucose and transferred to flat bottom 96 well plates (200 $\mu$ l per well). The plates were incubated for 4 hours at 37°C, the content of each well was decanted and the cells attached to the bottom of the wells stained by adding 50  $\mu$ l of 0.5% crystal violet. After 45 minutes at room temperature, the plate was washed ten times by alternate submerging in water and decanting. After having gently tapped

the plates to remove the excess water, 200  $\mu$ l of 75% methanol was added to each well and the plate incubated at room temperature for 30 minutes. Absorbance of each well at 590nm was finally measured using a micro plate reader (Gemini XPS fluorescence microplate reader).

For the *in vitro* spotting assay, strains were grown overnight at 30°C in a roller drum, resuspended to OD = 1 in sterile water, serially diluted 1:5 in sterile water and spotted on agar plates containing SDS, Calcofluor White, NaCl, Sorbitol or paraquat (methyl viologen dichloride) at the various concentrations suggested here (Homann et al., 2009). Plates were incubated overnight at 30°C and photographed using the BioRad Gel Doc XR. For filamentation, strains were spotted on Spider media (1% nutrient broth, 1% D-mannitol, 0.2% K<sub>2</sub>HPO<sub>4</sub>), incubated at 37°C for 7 days before being photographed.

### **3.2.5 *Ex vivo* macrophage assay**

The *ex vivo* macrophage assay was adapted from (Lopes da Rosa et al., 2010). Briefly, two T75 flasks of RAW 264.7 (Raschke et al., 1978) were grown to 80-90% confluency in DMEM + 10% FBS. Cells were scraped off the flasks, resuspended in DMEM + 10% FBS and  $2 \times 10^6$  cells were plated in 6 well plates (total volume 1 mL). Control wells with just DMEM + 10%FBS were also set up. Cells were allowed to adhere for 5 hours at 37°C, 5% CO<sub>2</sub> and then  $1.3 \times 10^5$  *C.albicans* cells resuspended in DMEM + 10% FBS spiked in each well (including the controls lacking macrophages). Plates were incubated at 37°C, 5%CO<sub>2</sub>, overnight. The following day cells were scraped off each well and resuspended in 10 mL 0.02% Triton-X 100 (vol/vol) to osmotically lyse the macrophages. The solution was vortexed at max speed for 1 minute and subsequently centrifuged at 3000 g for 5 minutes.

The pellet was resuspended in 1 mL sterile water and 200  $\mu$ l of the solution transfer to a 96 well plate. A 1:10 serial dilution was set up and 100  $\mu$ l of the  $10^4$  and  $10^5$  dilution plated on YPD plates. Wells containing both macrophages and *C. albicans* were plated to a  $10^4$  dilution while control wells containing just *C. albicans* were plated using the  $10^5$  dilution. After incubating the plates overnight at 30°C, colonies were counted and yeast susceptibility to macrophage killing measured by survival percentage. Survival percentage was measured as a ratio between the number of colonies recovered from wells containing only yeast cells over the number of colonies recovered from wells containing both yeast and macrophage cells.

### 3.3 Results

#### 3.3.1 *C. albicans* reverse genetic in *C. elegans* identifies four putative virulence factors

In order to identify novel *C. albicans* virulence determinants, a team of four undergraduate students (Margaret Chiasson, Benjamin Landry, Kurtis McCannel and Kelly Pastor) screened a collection of 724 *C. albicans* transposon insertion mutants (~12% of the genome) for their abilities to induce the Dar phenotype (Jain et al., 2009) in *C. elegans*.

The mutant library is freely available at the Fungal Genetic Stock Center and it is detailed in Table 3.1. Mutants were obtained using a transposon-based insertional mutagenesis strategy that lead to the disruption of both alleles of the targeted gene. The strategy relied on the bacterial transposon Tn7 (Craig, 1991) for the insertion and the UAU1 cassette for selection (Enloe et al., 2000). As detailed in Table 3.1, the insertional strategy was used in two different ways. Davis et al. 2002 and the Sanglard lab used a random insertional strategy to obtain 308 homozygous insertion mutants in various ORFs (the identities of the mutated genes were determined by sequencing the genomic regions flanking the inserted DNA). On the other hand, Blankenship et al., 2010; Nobile and Mitchell, 2005; Norice et al., 2007; Rauceo et al., 2008 used a targeted approach that yielded 416 insertion mutants in genes coding for transcription factors, cell wall proteins and protein kinases.

Overall, as shown in Figure 3.1 A the library contained 98 mutants in cell wall proteins, 84 mutants in protein kinases, 234 mutants in transcription factors and 308 mutants in various other ORFs.

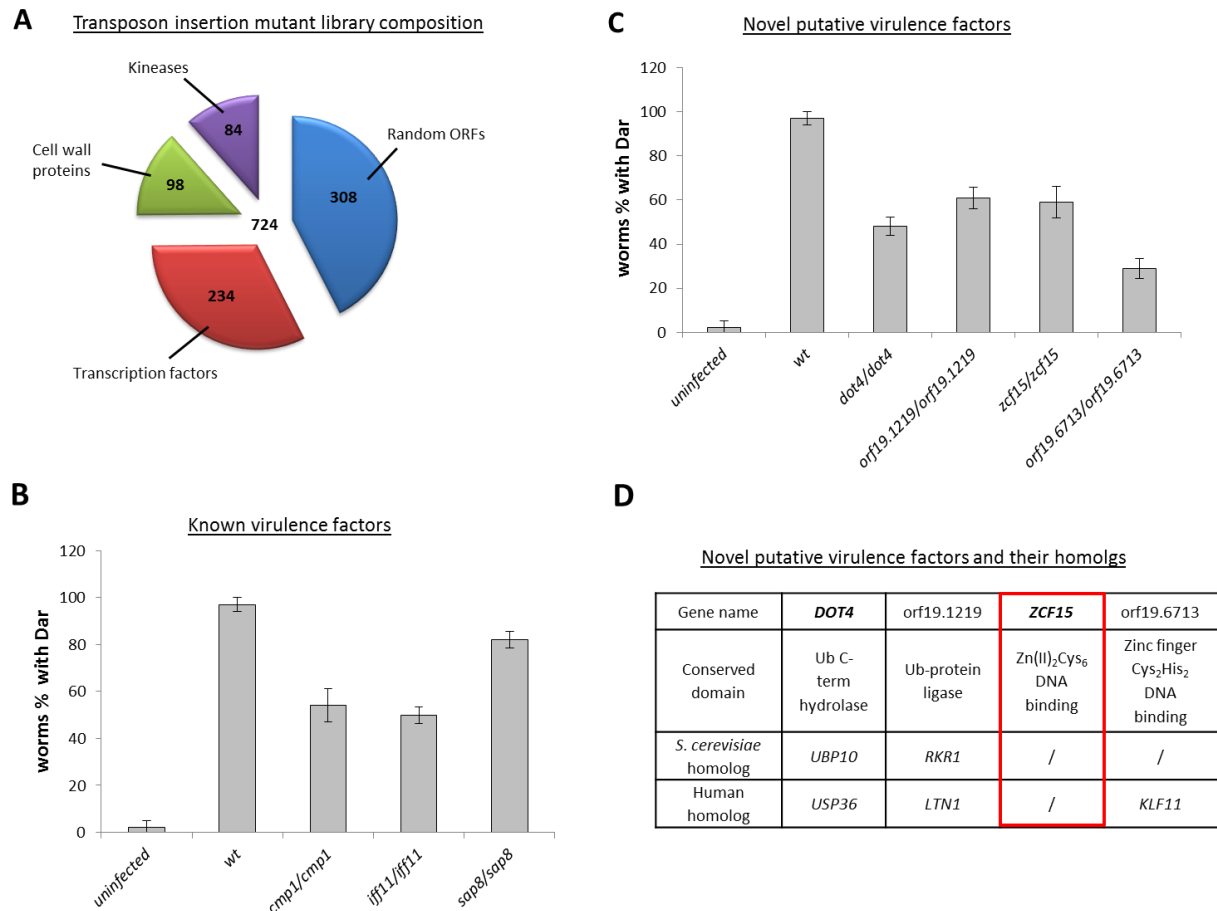
As shown in Figure 3.1 B and C, the team of undergraduate students identified 7 mutants that have a reduced ability to induce the Dar phenotype in *C. elegans*. *C. albicans* insertion mutants in *CMP1*, *IFF11*, *SAP8*, *DOT4*, *ZCF15*, *orf19.1219* and *orf19.6713* induced the Dar

phenotype only in 54%, 50%, 82%, 48%, 29%, 61% and 59% of the nematodes compared to nearly 100% for wildtype *C. albicans*. Three of these seven mutants (*CMPI*, *IFF11* and *SAP8*) are known virulence factors (Bader et al., 2003; Bates et al., 2007; Schaller et al., 2000) and their presence in our screen validated the approach we used.

These preliminary data obtained by the team of undergraduate students were promising but they needed further validation. In particular, it has been recently shown that the selection marker used to obtain these mutants (*URA3*) can interfere with wild type morphogenesis, adhesion and virulence (Brand et al., 2004; Noble et al., 2010a), possibly confounding interpretation of the results. In addition, integration biases, partial loss of function (especially when the insertion occurs near the 3' end of the gene) and genomic rearrangements following transposon insertions have been reported (Green et al., 2012) adding another layer of complexity to the interpretation of the results.

To validate our preliminary results we obtained homozygous knockout mutants in the genes originally shown to reduce incidence of the Dar phenotype. These knockout mutants were obtained by targeted gene replacement through two rounds of transformation to inactivate both alleles of a targeted gene (Noble et al., 2010). This approach doesn't have off-target effects and relies on auxotrophic markers (*ARG4*, *HIS1* and *LEU2*) that don't impact virulence (Noble and Johnson, 2005). We obtained knockout mutants in four of the seven genes originally identified and their reduced ability to cause the Dar phenotype in *C. elegans* was confirmed. Homozygous knockout mutants in *DOT4*, *ORF19.1219*, *ZCF15* and *SAP8* all showed a reduced ability to induce Dar (39%, 66%, 43% and 62% respectively) that was in line with what obtained with the transposon insertion mutants by the team of undergraduate students (48%, 61%, 59% and 82% respectively). The fact that mutants in the same genes but obtained with different strategies showed reduced virulence in *C. elegans* validated their

original findings. As shown in Figure 3.1 B, three (*CMP1*, *IFF11*, *SAP8*) of the seven genes identified in the screen were known virulence factors.



**Figure 3.1 Reverse genetic screening identifies four novel putative virulence factors**

(A) Transposon insertion mutant library composition, a total of 585 mutants were screened, 88 mutants in protein kinases, 98 mutants in cell wall proteins, 162 mutants in transcription factors and 237 mutants in various other ORF (B) Transposon insertion mutant library composition and relative papers describing them. (C) Table summarizing the Dar data obtained using either transposon insertion mutants (TR library) or homozygous knockout mutants (KO library). *ZCF15* has been poorly characterized in the literature, is required to induce Dar in nematodes and doesn't have a human homolog making it a good candidate for potential drug targeting.

*CMP1* codes for the catalytic subunit of calcineurin, a serine/threonine protein phosphatase that regulates physiological processes like cell cycle progression, polarized growth, and resistance to salt and alkaline pH (Bader et al., 2003; Sanglard et al., 2003). *CMP1* loss of

function leads to reduced virulence in a mouse model of disseminated candidiasis and a significantly reduced ability to colonize mouse kidney (Bader et al., 2003).

*IFF11* codes for a secreted protein required for cell wall architecture, and null mutants not only are hypersusceptible to cell-wall damaging agents but they showed attenuated virulence in a mouse model of systemic infection (Bates et al., 2007). *Sap8* is part of a large family of secreted aspartyl protease that plays a crucial role in hydrolyzing host epithelial tissues allowing *C. albicans* to simultaneously penetrate deeper tissues and use the products of the epithelial hydrolysis as N-source for growth (Aoki et al., 2011). *Sap8* has also been found to be upregulated in an *in vitro* model of cutaneous candidiasis (Naglik et al., 2003b) and to have a prominent role in biofilm formation (Ramage et al., 2012). The fact that mutations in genes already known to be important for virulence in mammals were hits in our screen validated our approach and underscored the relevance of studying *C. albicans* infection in a much simpler and genetically tractable host like *C. elegans*.

### **3.3.2 ZCF15 is a novel virulence factor with no human homologs**

As shown in Figure 3.1, in addition to known virulence factors, we identified three genes (*DOT4*, *ORF19.1219* and *ZCF15*) that have been poorly characterized in the literature and, to the best of our knowledge, have never been associated with *C. albicans* virulence. In order to gain some insights on their functions, we identified their *S. cerevisiae* closest homologs and explored their roles in baker's yeast. This process needed to be done cautiously since these two species last shared a common ancestor some 300-900 million years ago (Pesole et al., 1995) and there are many examples in the literature of *S. cerevisiae* genes that in *C. albicans*

have been rewired to have different biological functions (Coste et al., 2008; Murad et al., 2001).

The closest homolog of *C. albicans DOT4* is *S. cerevisiae* is *UBP10*, a ubiquitin-specific hydrolase primarily located in the nucleus and able to deubiquitinate ubiquitin-protein moieties (Kahana and Gottschling, 1999). This class of enzymes usually has a role in promoting protein degradation by clearing the proteasome of ubiquitinated peptides and facilitating ubiquitin recycling. However *UBP10* is involved in preventing, rather than promoting protein degradation (Hochstrasser, 1996). Assuming that the function of *UBP10* is conserved in *C. albicans*, it is reasonable to believe that *DOT4* loss of function leads to an impaired protein turnover, a trait that has been recently hypothesized to be important for virulence (Liu and Xue, 2011). In addition *DOT4* has a close homolog in human *USP36* (Blastp E value =  $1e^{-24}$ ), a poorly characterized gene that contains a well-known ubiquitin carboxyl-terminal hydrolase domain (Chung and Baek, 1999). The purpose of the de-ubiquitination is thought to be two folds: preventing protein degradation on the one hand and recycling of the ubiquitin on the other (Quesada et al., 2004).

Interestingly *C. albicans ORF19.1219* is also potentially involved in protein degradation. *ORF19.1219 S. cerevisiae* closest homolog is *RKR1*, a ubiquitin ligase with a functionally important Ring finger domain located at its carboxyl terminus (Bengtson and Joazeiro, 2010). Rkr1 is predominately cytoplasmic, it colocalize with ribosomes and is an active player of the cells translation quality control machinery (Braun and Johnson, 2000). mRNA lacking stop codons can arise from errors in gene expression and encode aberrant proteins; Rkr1 ubiquitinates nascent nonstop proteins with ubiquitin to signal their proteasomal degradation (Brandman et al., 2012). Interestingly, in the presence of hygromycin B, an antibiotic that affects translation fidelity, cells lacking *RKR1* exhibited greater sensitivity to the drug. *C.*



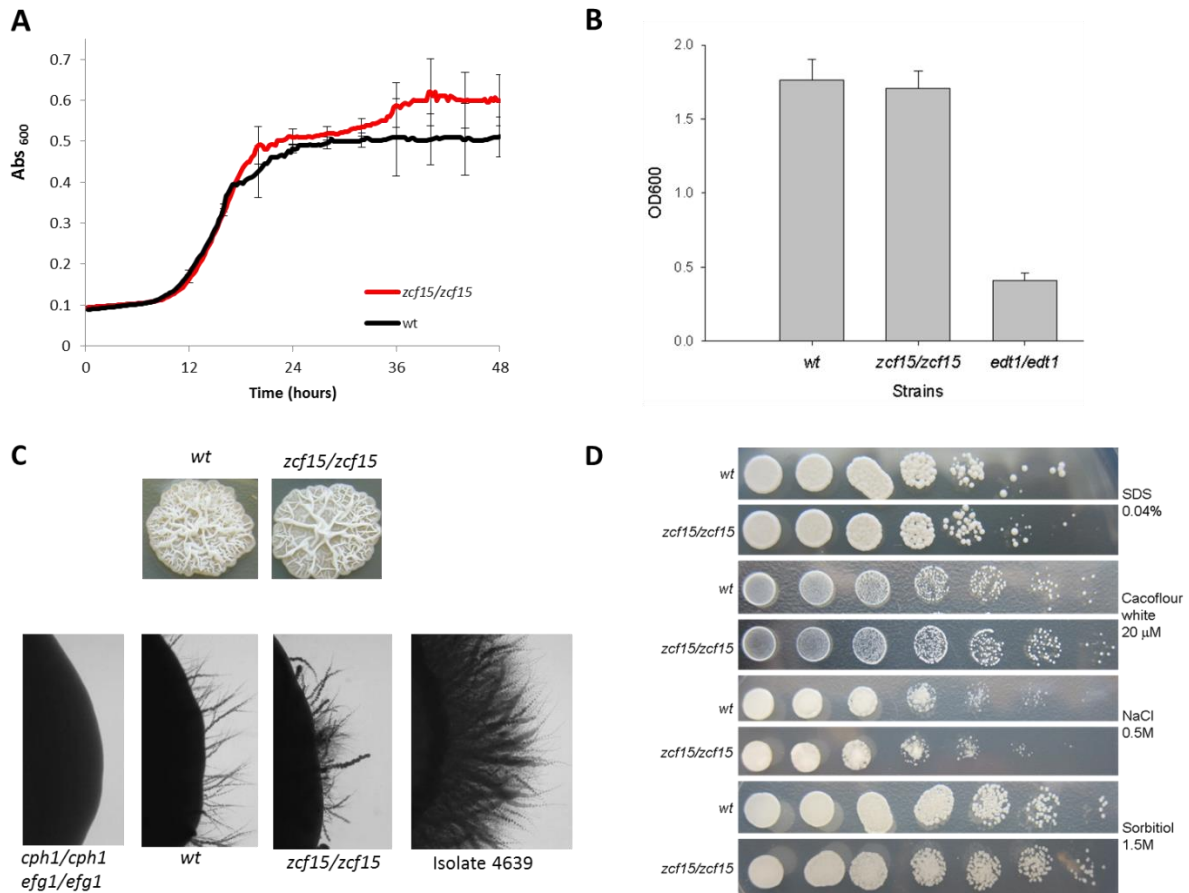
*albicans* *ORF19.1219* also has a close human homolog in Listerin (LTN1) (Blastp E value =  $6e^{-11}$ ) a ubiquitin ligase involved in protein quality control (Bengtson and Joazeiro, 2010). Mice knockout in Listerin, have been shown to experience early-onset and progressive neurological dysfunction (Chu et al., 2009). The working hypothesis is that loss of Listerin cause accumulation of aberrant proteins that ultimately lead to neurodegeneration. Assuming that the function of *RKR1* is conserved in *C. albicans* *ORF19.1219*, it is not surprising its loss of function leads to virulence reduction.

Although interesting, both *DOT4* and *ORF19.1219* are involved in a very fundamental aspect of *C. albicans* cellular biology and, most importantly, have two clear human homologs in *USP36* and *LTN1*. The presence of human homologs makes them significantly less attractive for drug design because potential drugs against these genes will likely have many off target effects and severe toxicities.

In contrast with *orf19.1219* and *Dot4*, *Zcf15* doesn't have an homolog in non-pathogenic *S. cerevisiae* nor in humans making it a potentially drug targetable novel *C. albicans* virulence factor. *ZCF15* belongs to a family of 77 transcription factors that are unique to the fungal kingdom (Maicas et al., 2005). This family of genes is characterized by a zinc binuclear cluster domain called  $Zn(II)_2Cys_6$  because six cysteins in the domain bind two zinc in order to achieve proper protein folding (Maicas et al., 2005). *ZCF15* is an uncharacterized transcription factor and, to the best of our knowledge, it is the first time to be reported to have a role in virulence. In the next paragraphs we'll describe a series of experiments that we run in order to explore its mechanism of action.

### 3.3.3 ZCF15 deletion mutant is sensitive to paraquat

In order to explore *ZCF15* mechanisms of action, we extensively characterized *zcf15/zcf15* *in vitro* looking for phenotypes that could explain its reduced virulence in the nematode model. We tested the knockout's ability to grow on minimal media, to adhere to polystyrene, to filament under N-starvation conditions and to resist various stresses like hyperosmotic, oxidative and cell wall stressors. We tested these conditions because the *C. albicans*' ability to limit/alternate nutrient utilization, form biofilms, switch from a yeast to hyphal forms and resist various stresses are all traits that have been postulated to be important for virulence (Heller and Tudzynski, 2011; Mitchell, 1998; Ramirez and Lorenz, 2007; Wachtler et al., 2011). As shown in Figure 3.2, under many of these conditions *zcf15/zcf15* behaves like wild-type. In particular *zcf15/zcf15* has wild type ability to grow in minimal media (Figure 3.2 A), to form biofilms on commonly used medical devices (Figure 3.2 B), to switch from yeast-to-hyphal form under N-starvation conditions (Figure 3.2 C) and to resist hyperosmotic (NaCl-Sorbitol) and cell wall (SDS-Calcofluor) stresses (Figure 3.2 D). These experiments suggest that the *zcf15/zcf15* hypo virulence in *C. elegans* is not connected to any of these important virulence traits.



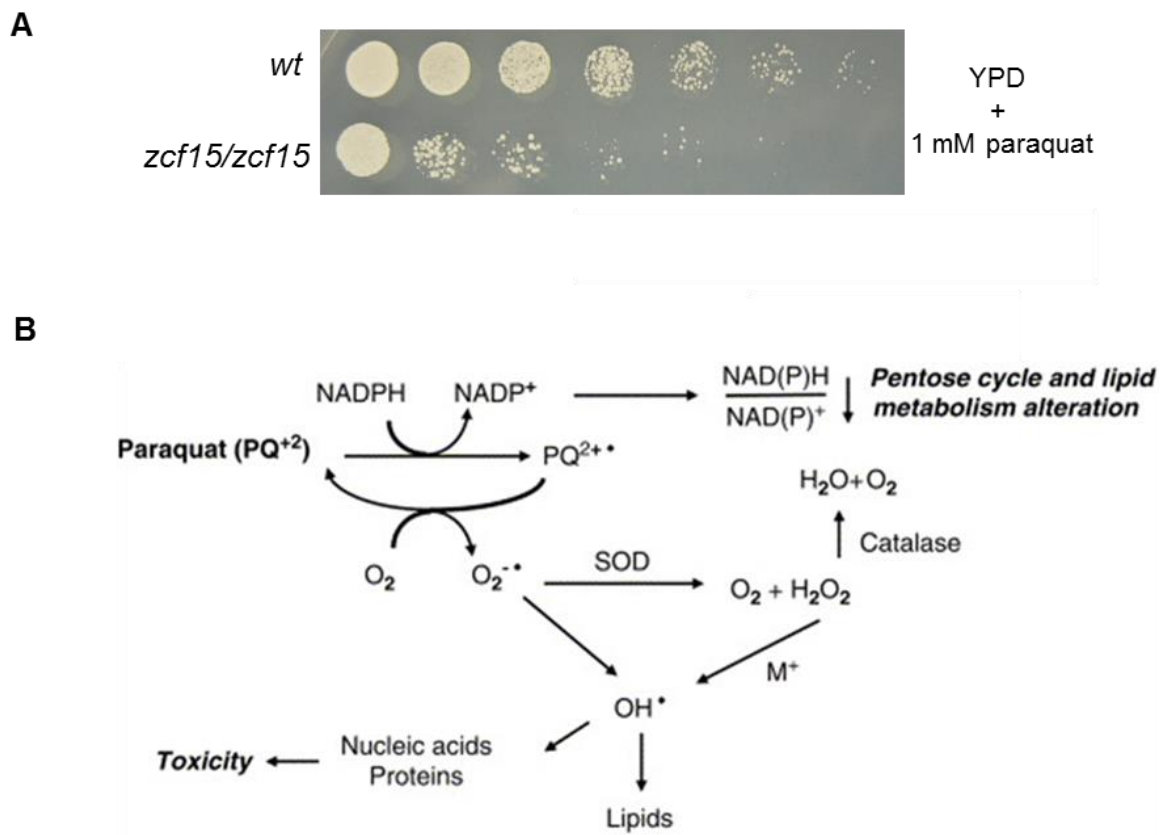
**Figure 3.2 *zcf15/zcf15* does not differ from wild-type in several in vitro assays.**

(A) *zcf15/zcf15* has wildtype ability to grow in minimal media. Overnight cultures were resuspended in SC media at OD<sub>600</sub> = 0.1 and absorbance monitored for 48 hours. Values from 3 independent biological replicates was averaged and plotted over time. Error bars indicate standard deviations. (B) *zcf15/zcf15* has wildtype ability to form biofilm. Overnight cultures were resuspended to OD<sub>600</sub>=0.5 and let adhere in filamentation inducing condition for 4 hours to polystyrene plates. The plates were thoroughly washed and the cells remained attached to the plate (biofilm) stained with crystal violet and quantified by reading absorbances at 590 nm. *edt1/edt1* is a non-adherent control strain (C) *zcf15/zcf15* has wildtype ability to filament under N-starvation conditions. *chp1/cph1 efg1/efg1* double mutant was used as a negative control and isolate 4639 (Perea et al., 2001b) as a positive control. (D) *zcf15/zcf15* has wild type ability to resist cell wall stressors (SDS and Calcoflour white) and hyperosmotic stresses (NaCl and Sorbitol).

However, when we tested the ability of *zcf15/zcf15* to grow in the presence of reactive oxygen species generator paraquat, we found that *ZCF15* is required for wildtype resistance (Figure 3.3 A). This evidence made us hypothesize that *zcf15/zcf15* reduced virulence in *C.*

*elegans* may be connected to its inability to cope with ROS, a commonly used host immune defense mechanism.

Paraquat is a well-established model of intracellular superoxide-mediated cell injury (Castell et al., 2005) and its mechanism of action is depicted in Figure 3.3 B.



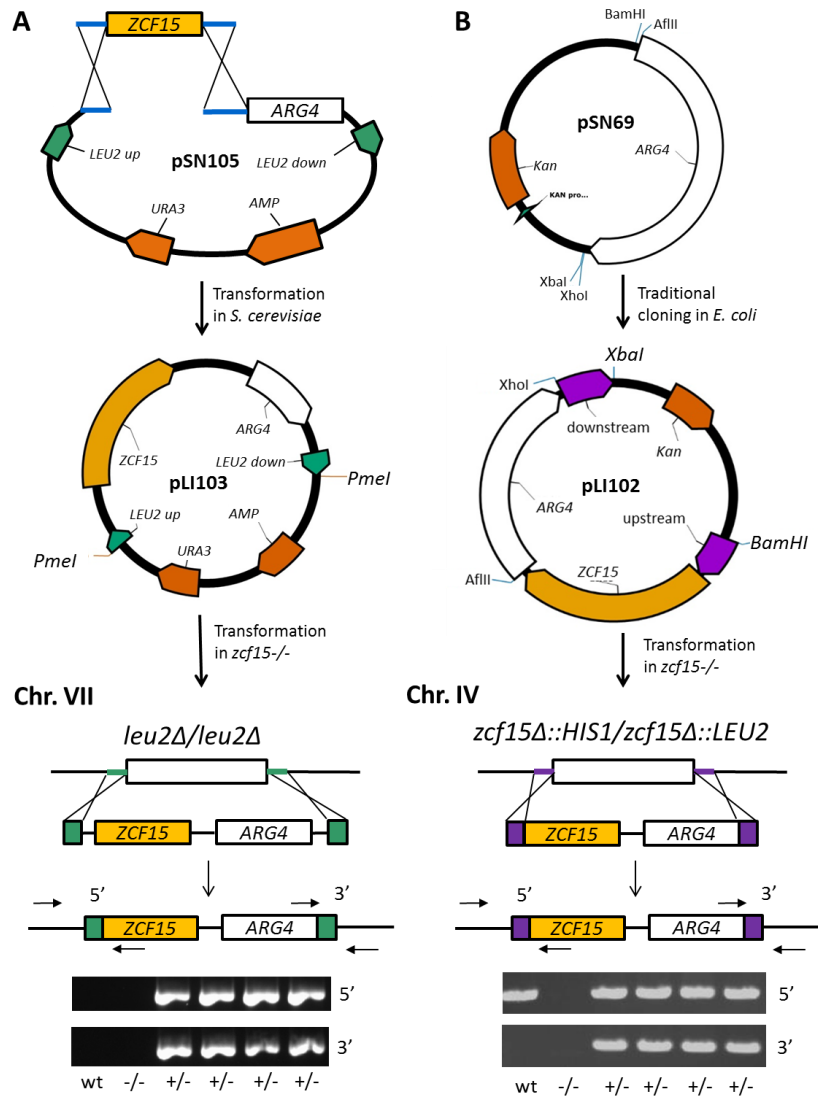
**Figure 3.3: *zcf15/zcf15* is hypersensitive to ROS generator paraquat.**

A) ZCF15 is required for wild-type resistance to reactive oxygen species generator paraquat. Wild-type and *zcf15/zcf15* strains were grown overnight, resuspended to OD = 1, serially diluted 1:5 and plated on YPD + 1mM paraquat.

B) Paraquat is reduced to its free radicals ( $PQ^{2+\bullet}$ ) with concomitant oxidation of NADPH. This reaction reduces the NADPH/NADP<sup>+</sup> ratio altering pentose cycle and lipid metabolism. In aerobic conditions,  $PQ^{2+\bullet}$  is oxidized back to paraquat in a process that leads to the formation of superoxide radicals ( $O_2^{\bullet-}$ ). Superoxide radicals ( $O_2^{\bullet-}$ ) can then lead to the formation of even more toxic reactive oxygen species like hydroxyl radicals ( $OH^\bullet$ ) that cause cellular damages via lipid peroxidation or DNA damage. Figure adapted from Castell et al., 2005.

Paraquat is reduced in the mitochondria to its free radical  $PQ^{2+\bullet}$  with concomitant oxidation of NADPH to  $NADP^+$ . In the presence of atmospheric oxygen, the reduced paraquat is recycled back to paraquat by cytochrome P-450 in a process that leads to the formation of superoxide radical ( $O_2^{\bullet-}$ ). Superoxide radicals can then either be converted to oxygen and hydrogen peroxide by superoxidase dismutase (SOD) or spontaneously form even more toxic radicals like hydroxyl radicals ( $OH^{\bullet}$ ). In the presence of transition metal ( $M^+$ ) hydrogen peroxide can also be converted in hydroxyl radicals ( $OH^{\bullet}$ ). Unlike superoxide, which can be detoxified by superoxide dismutase, hydroxyl radicals cannot be eliminated by enzymatic reaction and cause protein degradation, DNA damage and lipid peroxidation (Bus et al., 1976).

To confirm that *zcf15/zcf15* hypersensitivity to paraquat was a result of the intended deletion, we reintegrated a wild type copy of *ZCF15* into *zcf15/zcf15* knockout to obtain the corresponding complemented strains (*zcf15/ZCF15<sup>c</sup>*). As shown in Figure 3.4 we used two different strategies to obtain the complemented strains. On one side we used a gap repair cloning strategy to reintegrate *ZCF15* into the highly expressed *LEU2* locus (Figure 3.4 A), on the other, we used traditional cloning techniques to reintegrate *ZCF15* to its endogenous locus (Figure 3.4 B). Complementing gene deletion in the highly expressed *LEU2* locus is the most commonly used strategy by the *C. albicans* community however integration of exogenous genes at ectopic location can have position effects on gene expression (Gerami-Nejad et al., 2013) and confound phenotypic analysis. Complementing in the endogenous locus is more laborious but allowed reintegration of *ZCF15* in its endogenous locus and so its gene expression is controlled by its own regulatory sequences.



**Figure 3.4: *zcf15/zcf15* complementation strategies**

- (A) *ZCF15* reintegrated in *zcf15/zcf15* knockout in the ectopic locus *LEU2*. *ZCF15* open reading frame was amplified with primers containing 40 bp homologous sequences to BmgBI digested pSN105 (blue lines). BmgBI digested pSN105 and *ZCF15* were cotransformed into *S. cerevisiae* BY4741 to obtain plasmid pLI103. Plasmid pLI103 was digested with PmeI to liberate the *ZCF15-ARG4* integrating cassette flanked by ~500 bp sequences homologous to the upstream and downstream region of *LEU2* (green boxes). Digested pLI103 was transformed in *C. albicans zcf15/zcf15* to introduce a wild-type copy of *ZCF15* in *LEU2* locus (chromosome VII). *ZCF15* integration was confirmed by PCR checking the presence of the new 5' and 3' junctions at the *LEU2* locus.
- (B) *ZCF15* reintegrated in *zcf15/zcf15* knockout at the endogenous locus. *ZCF15* wild-type sequences (plus ~700 bp upstream of the gene) and ~700 bp sequence downstream of *ZCF15* were cloned into pSN69 to obtain pLI102. This step was achieved using traditional cloning in *E. coli* and restriction enzymes BamHI, AflIII, XbaI and XhoI. pLI102 was subsequently digested with BamHI and XbaI to liberate *ZCF15-ARG4* which integrates at the *ZCF15* endogenous locus on chromosome IV. Proper integration of *ZCF15* was confirmed by PCR checking the presence of the new 5' and 3' junctions.

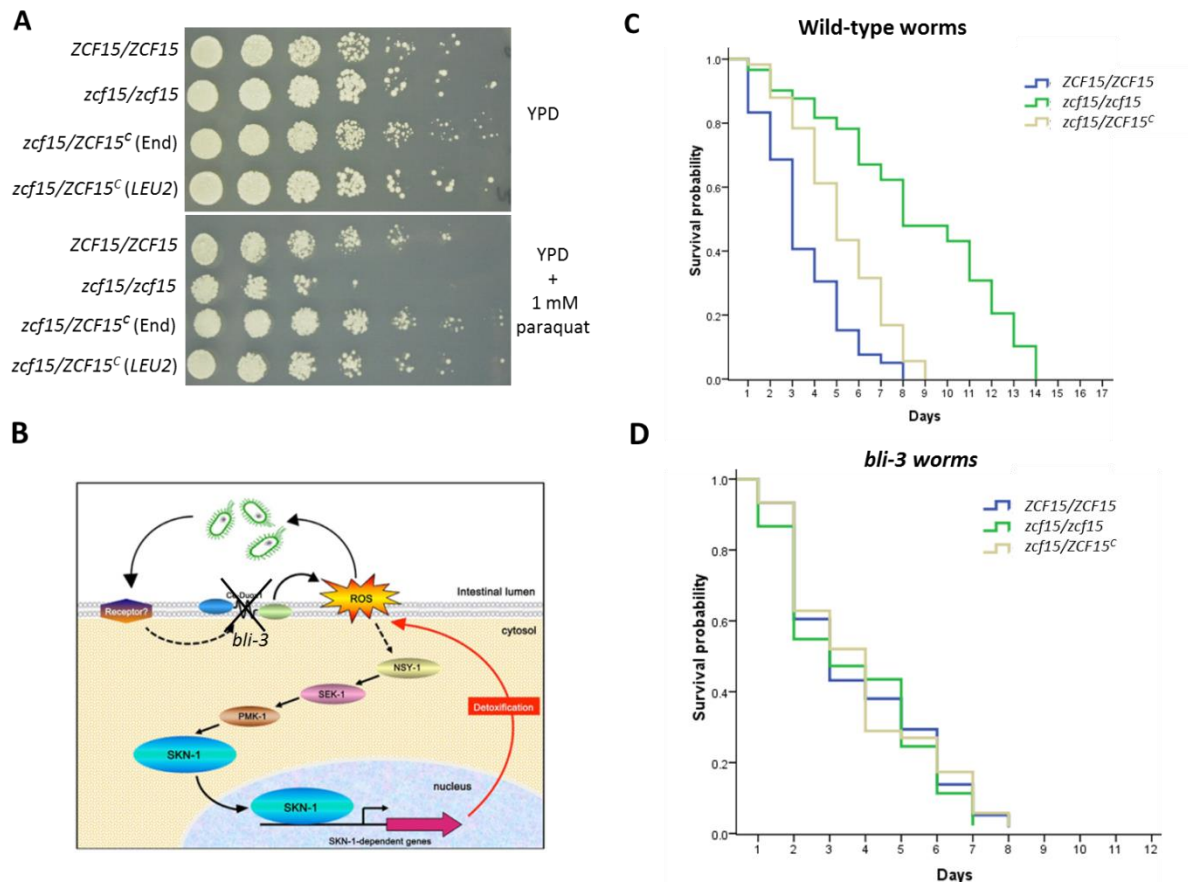
As shown in Figure 3.5 A, reintegration of *ZCF15* in either *LEU2* or in its endogenous locus fully restores wild-type ability to resist paraquat, confirming that the phenotype observed in the deletion is caused by *ZCF15* loss of function and not by other off target effects obtained during the deletion. To the best of our knowledge, this is the first time that *ZCF15* has been associated with the *C. albicans* ability to withstand reactive oxygen species.

### **3.3.4 *ZCF15* is required to withstand *C. elegans* generated ROS**

It is well known that *C. elegans* produces ROS upon both bacterial and fungal infections (Chavez et al., 2007; Jain et al., 2009a; Moy et al., 2004) as part of its defense mechanism. ROS have a biocidal effect on invading organisms and play a major role in innate immunity. *zcf15/zcf15* hypersusceptibility to ROS *in vitro* made us hypothesize that its reduced virulence in *C. elegans* was due to a reduced ability to withstand the host's generated ROS. To test our hypothesis we determined the *zcf15/zcf15* ability to kill either wild-type worms or worms with an impaired ability to produce ROS.

As shown in Figure 3.5 B *C. elegans* responds to pathogen's ingestion by producing extracellular ROS in the intestinal lumen via Ce-Duox1, an NADPH oxidase coded by the gene *bli-3*. During ROS production, the intestinal cells also produce intracellular antioxidants via *DAF-16* to protect its own tissues from the ROS damaging effects (Chavez et al., 2007; Hoeven et al., 2011). Ce-Duox1 is a protein with an N-terminal peroxidase domain, a C-terminal superoxide-generating NADPH-oxidase domain and two central calmodulin-binding sites (Meitzler and Ortiz de Montellano, 2009). Upon microbial infection, this protein uses cytosolic NADPH to generate extracellular toxic ROS in the intestinal lumen to counteract the infection. Throughout a series of biochemical assays, Jain et al., 2009 showed that ROS

are abundantly produced upon yeast infections and that *bli-3* loss of function via *bli-3(e767)* dramatically reduces the nematode ability to produce ROS.



### Figure 3.5: ZCF15 is required to withstand *C. elegans* generated ROS

(A) *ZCF15* is required for wild-type resistance to reactive oxygen species generator paraquat. Wild-type, knockout or complemented strains were grown overnight, resuspended to OD = 1, serially diluted 1:5 and plated on either YPD or YPD + 1mM paraquat (B) *C. elegans* fights pathogens present in the intestinal lumen by producing ROS via *bli-3* and contemporarily producing intracellular antioxidant via *DAF-16* to protect its own tissues. Figure adapted from (Hoeven et al., 2011) (C) Kaplan-Meier survival curves showing that *ZCF15* deletion reduces *C. albicans*' ability to kill wildtype worms. (D) Kaplan-Meier survival curve showing comparable killing kinetics between *zcf15/zcf15*, wild-type and complemented strain when ROS-deficient *C. elegans bli-3(e767)* were challenged.

As shown in Figure 3.5 B *C. elegans* responds to pathogen's ingestion by producing extracellular ROS in the intestinal lumen via Ce-Duox1, an NADPH oxidase coded by the gene *bli-3*. During ROS production, the intestinal cells also produce intracellular antioxidants



via *DAF-16* to protect its own tissues from the ROS damaging effects (Chavez et al., 2007; Hoeven et al., 2011). Ce-Duox1 is a protein with an N-terminal peroxidase domain, a C-terminal superoxide-generating NADPH-oxidase domain and two central calmodulin-binding sites (Meitzler and Ortiz de Montellano, 2009). Upon microbial infection, this protein uses cytosolic NADPH to generate extracellular toxic ROS in the intestinal lumen to counteract the infection. Throughout a series of biochemical assays, Jain et al., 2009 showed that ROS are abundantly produced upon yeast infections and that *bli-3* loss of function via *bli-3(e767)* dramatically reduces the nematode ability to produce ROS.

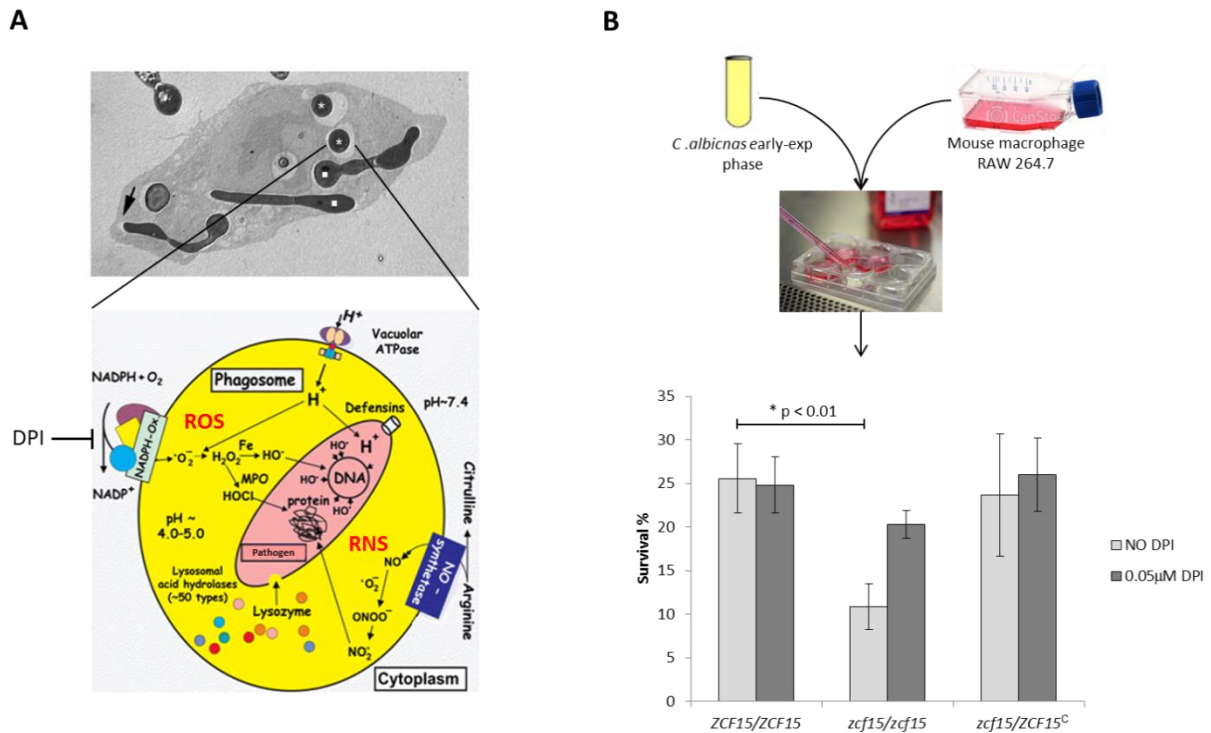
As shown in Figure 3.5C, worms challenged with *zcf15/zcf15* survived significantly longer than wild-type or complemented strains ( $p < 0.01$  by the log-rank test) indicating that *ZCF15* is required for wildtype virulence. However, when we challenged ROS-deficient *C. elegans bli-3(e767)*, we obtained kinetics of killings that were comparable between *zcf15/zcf15*, wild-type and complemented strain (Figure 3.5D). This evidence indicates that *zcf15/zcf15* fail to kill nematodes unless the host ability to produce ROS is compromised. Taken together, these results suggest that *ZCF15* is required for full virulence and that this gene is likely involved in the pathogen's ability to resist host generated ROS. To the best of our knowledge, this is the first time that *ZCF15* has been shown to be involved in pathogenicity and ROS resistance.

### **3.3.5 *ZCF15* is required for resistance to murine macrophage killing**

After having shown that *ZCF15* is required for virulence in nematodes, we wanted to explore its role in a mammalian model of infection. Macrophages, together with other phagocytic cells like neutrophils, monocytes and dendritic cells, play a fundamental role in protecting mammals against *C. albicans* (Ashman et al., 2004; Newman and Holly, 2001). In particular, since there is no evidence that antibodies can mediate the lysis of *C. albicans*, macrophages

and other phagocytic cells are considered the prime effectors in fighting yeast infections (Vazquez-Torres and Balish, 1997). As shown in Figure 3.6 A, when *C. albicans* is coincubated with mouse macrophage cell line RAW264.7 *ex vivo*, *C. albicans* is quickly phagocytized, compartmentalized in the phagosome and killed by a combination of different mechanisms (Lorenz et al., 2004; Marcil et al., 2002). As shown in Figure 3.6 A, these mechanisms include reactive oxygen species (ROS), reactive nitrogen species (RNS), acidic pH and hydrolytic enzymes like lysozyme. We used mouse macrophage RAW264.7 because their ability to produce ROS can be pharmacologically modulated by adding NADPH inhibitor diphenyleneiodium chloride (DPI) (Lopes da Rosa et al., 2010).

We exposed wildtype, *zcf15/zcf15* knockout and the complemented strain to murine macrophages at 1:15 multiplicities of infections (15 macrophage cells per each *C. albicans* cell) and quantified yeast cell survival by end point dilution. As shown in Figure 3.6 B, the *zcf15/zcf15* knockout is significantly more susceptible to macrophage killing compared to either wild-type or complemented strains (Figure 3.6 A). Survival percentages were respectively  $25.6 \pm 4.0$  for *ZCF15/ZCF15*,  $10.8 \pm 2.6$  for *zcf15/zcf15* and  $23.7 \pm 7.0$  for *zcf15/ZCF15*<sup>c</sup> ( $p < 0.01$  for *ZCF15/ZCF15* vs *zcf15/zcf15*). We hypothesized that the *zcf15/zcf15* increased susceptibility to macrophage killing was due to its inability to withstand macrophage generated ROS. To test this hypothesis we repeated the experiments in the presence of NADPH inhibitor diphenyleneiodium chloride (DPI), which it has been previously shown to inhibit macrophage ROS production in the phagosome (Donini et al., 2007). Remarkably, pharmacological inhibition of NADPH oxidase rescued *zcf15/zcf15* hyper susceptibility to macrophage killing suggesting that *ZCF15* is required to neutralize the oxidative environment within the phagosome.



**Figure 3.6 ZCF15 is required to withstand mouse macrophage generated ROS**

- A) Electron micrograph of a macrophage 2 hours post *C. albicans* phagocytosis (top). Upon phagocytosis *C. albicans* is compartmentalized in the phagosome (bottom) where a combination of acidic pH, reactive oxygen species (ROS), reactive nitrogen species (RNS) and hydrolytic enzymes (lysozyme) contribute to *C. albicans* killing. The macrophage ability to ROS can be pharmacologically modulated by adding diphenyleneiodium chloride (DPI). Figures adapted from (Haas, 2007; Heinsbroek et al., 2005).
- B) *C. albicans* growing to early-exponential phase were coincubated with mouse macrophage RAW264.7 at 1:15 multiplicities of infections (1 yeast cell per 15 RAW264.7 cells) and survival percentage determined as the ratio between the number of colonies recovered from wells containing only yeast cells over the number of colonies recovered from wells containing both yeast and macrophage cells. *zcf15/zcf15* was statistically significantly more susceptible to macrophage killing compared to either wild-type or complemented strains unless the macrophage ability to produce ROS was pharmacologically inhibited by adding 0.05  $\mu$ M of DPI.

Taken together, these results suggest that *ZCF15* is required not only for virulence in nematodes but also for virulence in a more relevant mammalian model of infection. In addition, our data suggests that *ZCF15* plays a vital role in the pathogen's ability to withstand

host generated ROS, an evidence that was mirrored in both nematodes and murine macrophages.

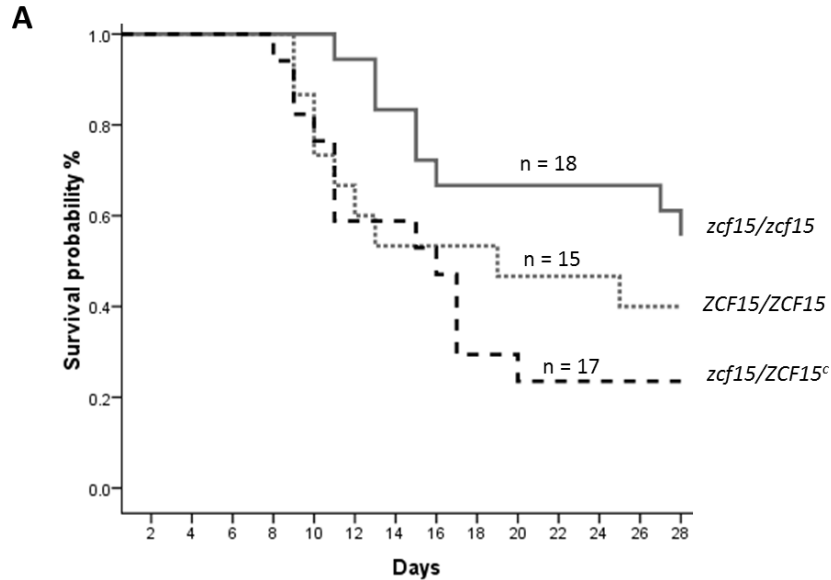
### 3.3.6 *ZCF15* deletion has a limited effect in disseminated *Candidiasis*

*Note: This experiment was performed by Alex Hopke under the supervision of Dr. Robert Wheeler at the University of Maine*

As described in chapter 1.3 *C. albicans* can infect the human gastrointestinal tract, causing epithelial infections but, under certain conditions it can reach the blood stream and cause life threatening systemic infections. In the blood stream a combination of innate and adaptive immune mechanisms are responsible for fighting the infection, including CD4<sup>+</sup> T cells, phagocytic cells, IL-12 and TNF- $\alpha$  (Ashman et al., 2004; Romani, 2000). Both *C. elegans* and murine macrophage systems represent great models to study innate immunity and showed that *ZCF15* plays a critical role in *C. albicans* virulence. Here, we wanted to explore its role in the context of a blood stream infection, where both innate and adaptive immunity contribute to fungal resistance.  $1 \times 10^5$  cells of wild-type, *zcf15/zcf15* or complemented strain were injected into the blood stream of BALB/c mice by tail vein injection and morbidity and mortality were monitored daily. As shown in Figure 3.7, *ZCF15* deletion has only a mild and non-statistically significant impact in virulence.

Mice injected with *zcf15/zcf15* knockout survived longer than those injected with wild-type (mean time to death for *zcf15/zcf15*  $23.2 \pm 1.7$  days versus  $19.1 \pm 2.1$  days for *ZCF15/ZCF15*) however, this difference was not statistically significant (Log Rank p values = 0.193). On the other hand, mice injected with the complemented strain survived statistically significantly longer than those injected with *zcf15/zcf15* knockouts (mean time to death for mice injected with *zcf15/ZCF15c* =  $16.6 \pm 1.7$ , Log Rank p values = 0.030). This

data suggest that *ZCF15* deletion has a modest effect in virulence in a murine model of hematogenously disseminated candidiasis an evidence corroborated by the fact that reintegrating a wild type copy of *ZCF15* slightly increased knockout virulence.



**B** Pairwise Comparisons

	Strains	wt		<i>zcf15</i> <sup>-/-</sup>		<i>zcf15</i> <sup>+/-</sup>	
		Chi-Square	Sig.	Chi-Square	Sig.	Chi-Square	Sig.
Log Rank (Mantel-Cox)	wt			1.691	.193	.718	.397
	<i>zcf15</i> <sup>-/-</sup>	1.691	.193			4.718	.030
	<i>zcf15</i> <sup>+/-</sup>	.718	.397	4.718	.030		

**Figure 3.7: *ZCF15* loss of function has a limited impact in survival of mice intravenously challenged.**

- A) Kaplan-Meier survival curves of mice injected with  $1 \times 10^5$  cells of *C. albicans* wild type (*ZCF15/ZCF15*), *zcf15/zcf15* or *zcf15/ZCF15<sup>c</sup>*.
- B) Pairwise statistical comparison between survival curves using the Log Rank test. While the difference between the knockout and the complemented strain was statistical significant ( $p = 0.030$ ), the difference between wildtype and the knockout was not ( $p = 0.193$ ).

Taken together, our data suggests that *ZCF15* plays an important role in models of infections like *C. elegans* or mouse macrophages that heavily rely on the oxidative burst and innate immunity as their main mechanisms of host response to infections. However, in a much more complex environment like the mouse bloodstream where a wide variety of innate and

adaptive immunity contribute to fight the infection *ZCF15* loss of function has a limited impact on virulence.

### 3.4 Discussion

As discussed in previous chapters, *C. albicans* infections have alarming incidence, morbidity and mortality rates. *C. albicans* infections are common comorbidities in patients affected by various immunodeficiency disorders. Unfortunately, the pharmacological options available to these patients are limited, especially considering that *C. albicans* have found various clever ways to develop drug resistance. In addition, the antimycotics available are all targeting a very limited number of cellular processes and have less than ideal toxicological profiles. For all of these reasons, there is a pressing need for more research in order to identify new targets that can be used for better drugs and diagnostics.

Here we leveraged our nematode *in vivo* model system and the availability of a large collection of *C. albicans* mutants to identify *ZCF15*, a fungal specific transcription factor involved in virulence. Our reverse genetic approach allowed us to screen a total of 724 *C. albicans* mutants (~12% of the genome) and lead to the identification of four novel putative virulence factors: *DOT4*, *ORF19.1219*, *ORF19.6713* and *ZCF15*. Between these four genes we focused on *ZCF15* because, in contrast to the other three genes, it doesn't have a human homolog and can potentially be used for drug targeting. We showed here that *ZCF15* is required for virulence in our nematode model and that it plays a critical role in the pathogen's ability to withstand host generated ROS.

To the best of our knowledge, this is the first time that this gene has been associated with virulence in *C. albicans*. In addition, although its role in a disseminated model of infection

has been shown to be limited, this gene is required for wild type resistance to mammalian macrophage. Macrophages together with dendritic cells and natural killers are a fundamental component of the human innate immunity and are considered the first line of defense against *C. albicans* infections. In particular we have shown here that *ZCF15* is required to withstand macrophage generated reactive oxygen species.

In addition, although the possibility of targeting *ZCF15* for therapeutic purposes goes well beyond the scope of this dissertation, it is worth mentioning that there is growing evidence that transcription factor based therapeutics holds great potential. Transcription factors have been historically considered difficult targets however, significant progress has been made in the last two decades and targeting them is becoming a realistic option (Flood et al., 2011; Yeh et al., 2013). Transcription factor targeting is complicated by the fact that transcription is a nuclear event and many essential components involved in transcription lack enzymatic activity suitable, for chemical intervention (Yan and Higgins, 2013). However in recent years many different strategies that directly inhibit transcription factor expression (RNAi, microRNAs), transcription factor DNA binding (with small molecules competing for binding partners or peptide mimetic that disrupt transcription post translational modifications) have been developed and have recently entered clinical trials (Flood et al., 2011; Ghosh and Papavassiliou, 2005; Kletsas and Papavassiliou, 1999). Transcription factors are the focal points that drive cell signaling and the pathogen's ability to cause disease. Because *ZCF15* is not conserved in humans, it represents a potentially suitable target for therapeutic intervention. As described in the next chapter, we further elucidated *ZCF15* mechanisms of action by identifying its transcriptional targets by RNASeq and its DNA binding partners by ChIP-Seq.

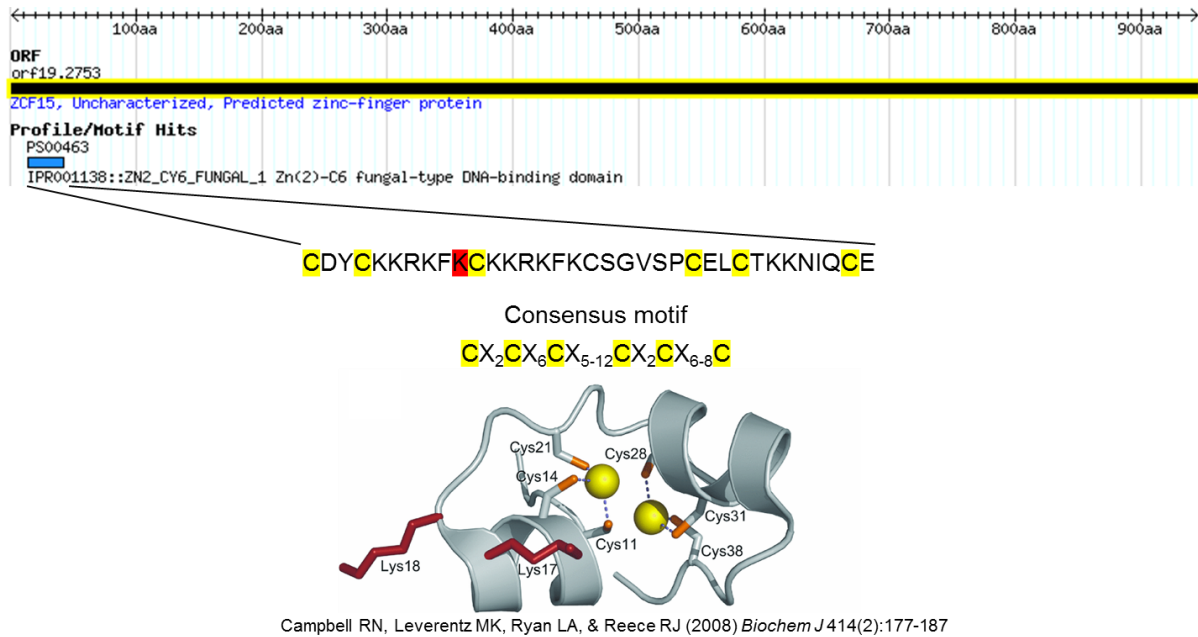
## 4 *ZCF15* and *ZCF29* mechanisms of action, some insights from RNASeq and ChIP-Seq

### 4.1 Introduction

In the previous chapter we described *ZCF15*, a *C. albicans* transcription factor with a clear role in resistance to host generated reactive oxygen species (ROS). This transcription factor belongs to a family of 77 *C. albicans* Zn(II)<sub>2</sub>Cys<sub>6</sub> transcription factors. These proteins contain the well-conserved CysX<sub>2</sub>CysX<sub>6</sub>CysX<sub>5-16</sub>CysX<sub>2</sub>CysX<sub>6-8</sub>Cys domain (Maicas et al., 2005) in which six cysteins (Cys) bind two zinc atoms and coordinate both the proper folding of the protein and its ability to bind DNA (Campbell et al., 2008; Vallee et al., 1991). Figure 4.1 depict the location of the Zn(II)<sub>2</sub>Cys<sub>6</sub> DNA binding domain in the Zcf15 sequence. The domain is located at the C-terminus and it contains a highly conserved lysine (Figure 4.1, red box), which is believed to be a key player in coordinating DNA binding. Indeed, that lysine is conserved in 57 of the 77 *C. albicans* Zn(II)<sub>2</sub>Cys<sub>6</sub> transcription factors, and in 13 of the remaining 20, lysine is replaced by another basic amino acid like histidine or arginine.

*C. albicans* Zn(II)<sub>2</sub>Cys<sub>6</sub> transcription factors regulate a large variety of biological processes ranging from control of meiosis, cell wall architecture, amino acid utilization and multi-drug resistance (Maicas et al., 2005). This family of transcription factors is found in various fungi like *Saccharomyces cerevisiae*, *Aspergillus flavus*, *Fusarium solani*, *Aspergillus niger*, *Kluyveromyces lactis* and *Neurospora crassa* but not in humans, other eukaryotes or prokaryotes. Since they are unique to the fungal kingdom and fundamental to *C. albicans* biology, these transcription factors have the potential to be used for either drug targeting or as epitopes for vaccine design (Vandeputte et al., 2011).





**Figure 4.1: canonical structure of Zn(II)<sub>2</sub>Cys<sub>6</sub> DNA binding domain and relative location in Zcf15**

Zcf15 has the Zn(II)<sub>2</sub>Cys<sub>6</sub> located near its C-terminus end. This domain is made of 35 amino acids in which six cysteins bind to 2 zinc atoms (yellow circles) and drive proper protein folding and DNA binding.

Here, we analyzed this family of transcription factors and looked for genes that (a) have a role in resistance to reactive oxygen species (b) are conserved in pathogenic fungi but not in non-pathogenic ones. This dual approach allowed us to look for genes that, on one side are potentially conferring resistance to host effectors and, on the other, are evolutionarily conserved in fungal pathogenesis. We identified *ZCF29*, a transcription factor that has been previously shown to be upregulated in the presence of H<sub>2</sub>O<sub>2</sub> (Enjalbert et al., 2006), required for resistance to ROS generator menadione (Homann et al., 2009) and that is present only in pathogenic fungi.

In order to gain some insights into *ZCF29* and *ZCF15* biological roles, we leveraged here the power of RNASeq to discovered genes controlled by these two transcription factors. Remarkably, *ZCF29* deletion leads to the misregulation of a staggering number of genes

(1440 genes, ~25% of the genome), indicating that this transcription factor is a master regulator of *C. albicans* metabolism. These genes are mainly involved in amino acid biosynthetic processes and proteolysis, two biological processes highly interconnected with each other. In particular, *ZCF29* deletion leads to downregulation of genes involved in the synthesis of important amino acids like threonine, methionine, tryptophan, tyrosine, lysine and histidine and concomitant upregulation of genes involved in proteolysis. This evidence suggests that upon *ZCF29* deletion *C. albicans* favors amino acids recycling through proteolysis over *de novo* biosynthesis. While both pathways can produce free amino acids as a final product, amino acid biosynthesis is more energy demanding than proteolysis and *C. albicans* wild type favors proteolysis over *de novo* synthesis under N-starving condition (Ramachandra et al., 2014). We believe that *ZCF29* plays a critical role in the *C. albicans* ability to regulate nitrogen metabolism an evidence corroborated by the hyper susceptibility of *zcf29/zcf29* to nitrogen-starvation mimicking caffeine (Homann et al., 2009; Rallis et al., 2013).

On the other hand, *ZCF15* loss of function leads to a more modest transcriptional rewiring with a total of 168 genes differentially expressed compared to wild-type. Many of these genes are involved in carbon metabolism and energy production with genes like *HGT7*, *SHA3*, *TYE7*, *PCK1* and *EHT1* directly involved in glucose transport, acetyl-CoA production or TCA cycle. Many of these genes have been shown to be downregulated in the presence of ROS (Lorenz et al., 2004) and this down regulation is believed to liberate significant energy resources for other cellular processes like DNA repair or cellular detoxification processes. Here, we showed that the *C. albicans* ability to downregulate these genes in the presence of ROS and contemporarily upregulate key enzymes for cellular detoxification processes is compromised in *zcf15/zcf15*. Our working hypothesis is that, in the presence of ROS, *ZCF15*

plays a critical role in rechanneling energy resources from carbon metabolism/energy production to cellular detoxification processes. This hypothesis fits well with the *in vivo*, *in vitro* and *ex vivo* phenotypes observed so far for *zcf15/zcf15* and may be the reason for its hyper susceptibility to ROS.

## **4.2 Materials and methods**

### **4.2.1 Media and growth conditions**

Starting from single colonies, 5 mL cultures were grown overnight in YPD at 30 °C. Cells were collected by centrifugation (2000 g for 5 min), and resuspended to a final OD=0.1 in fresh YPD media. Cultures were allowed to grow for 4-6h to reach mid-exponential phase (OD=0.6-0.8) and then 5mM H<sub>2</sub>O<sub>2</sub> was added to the samples. Cells were cold methanol quenched before and 5 and 15 minutes after the addition of H<sub>2</sub>O<sub>2</sub>. To methanol quench cells conical tubes containing 30 mL 100% methanol were stored at -80 °C overnight. Conical tubes were taken out of the -80 °C and placed in a dry-ice ethanol bath to keep the methanol at around -40 °C. Once the sample (20 mL) was added to the methanol, the conical tubes were centrifuged at 2000 g for 5 minutes and the methanol and media removed. The pellets containing the cells were flash frozen in liquid nitrogen and the tubes stored at -80 °C overnight. The cell pellets were thawed, washed in 5 mL nuclease free water and centrifuged for 5 min at 2000 g at 4 °C. The supernatant was discarded and the pellets re-suspend in 1.5 mL of RNAlater (Ambion) and transferred to smaller tubes for storage. Samples were left at 4 C for 24 hours before being moved to the -80 °C freezer.

### **4.2.2 RNA-Seq: RNA extraction and library preparation**

RNA was extracted using the Qiagen RNeasy Plus Mini Kit. Briefly, 5x10<sup>8</sup> total cells were re-suspended in 600 µl RLT+2-mercaptoethanol (10 µl of 2-mercaptoethanol for 1 mL RLT) in 2 mL bead beating tubes. 500 µl RNAse free Zirconia beads were added and cells lysed in a bead beater at max speed for 3 minutes. The tubes were centrifuged at 1000g for 5 minutes at 4 °C, the supernatant transferred to a gDNA eliminator column, and RNA isolated

following the manufacturer instructions. RNA concentration was determined using a NanoDrop ND-1000 UV-Vis Spectrophotometer (Thermo Scientific) while RNA integrity was determined using an Agilent 2200 Tape Station (RIN score). From total RNA, mRNA was purified using Dynabeads mRNA DIRECT and 300-600 bp fragments obtained by incubating the sample in RNA Storage Solution (Ambion) for 30 min at 98 °C. The fragmented mRNA was reverse transcribed to cDNA using AffinityScript Multiple Temperature Reverse Transcriptase (Agilent) by priming with Oligodt (12,13,14,15,16,17,18) primers. Single stranded cDNA was converted into double stranded cDNA by using NEBNext Second Strand Synthesis Module (New England Biolabs). Double stranded cDNA was purified using Agencourt AMPure XP beads (Beckman Coulter) and strand nonspecific cDNA libraries were prepared for Illumina sequencing (MiSeq) according to the method previously described here (Mardis, 2008a). Once the library was prepared, a size distribution around 300 bp was confirmed by using the Agilent 2200 Tape station and the total DNA concentration (expressed in ng/μl) was measured using a Qubit Fluorometer. DNA concentration in ng/μl was converted in nM using the equation reported below. The equation assumes that the MW of each nucleotide is 650 D and the average length of library is 300bp.

$$Concentration (nM) = \frac{Concentration \left( \frac{\mu g}{\mu l} \right) \times 10^6}{650bp \times 300bp}$$

As an alternative to the Qubit, we used a Kapa library Quant kit that uses a qPCR approach to specifically quantify only PCR-competent DNA molecules (properly ligated to the adaptors).

### 4.2.3 RNA-Seq: sample preparation and sequencing

Each of the samples were resuspended in water to a final concentration of 4 nM and pooled together by adding 5 μl of each sample to a 1.5 mL microcentrifuge. 5 μl of the pooled

sample were denatured by adding 5 µl of NaOH 0.2 N and incubated for 5 minutes at room temperature. 990 µl of HT1 (hybridization buffer) were added to the 10 µl of the denatured libraries to obtain 20 pM denatured library in 1 mM NaOH.

2 µl of 10 nM PhiX library were added to 8 µl 10 mM Tris-Cl (pH 8.5) with 0.1% Tween 20 to obtain a final concentration of 2 nM. 10 µl of this solution was added to 10 µl of 0.2N NaOH, vortexed and incubated at room temperature for 5 minutes to denature the PhiX into single strands and obtain a final PhiX concentration of 1 nM. 8 µl of denatured PhiX is added to 992 µl of HT1 (hybridization buffer) to obtain a final concentration of 8 pM.

300 µl of the 20 pM denatured library was added to 300 µl of 8 pM denatured PhiX library and loaded into the Illumina cartridge for a HiSeq run.

#### **4.2.4 ChIP-Seq: Zcf15 and Zcf29 HA tagging**

Zcf15 and Zcf29 were C-terminus HA tagged using a PCR-mediated homologous recombination strategy described here (Lavoie H, 2008). Figure 4.25 depicts a schematic of the approach used. Briefly, primers pair including ~70 bp of homology flanking *ZCF15* and *ZCF29* stop codon were used to amplify the HA-ARG4 cassettes from pFA-HA-ARG4 plasmid. The plasmid was a generous donation of Dr. Malcom Whiteway (McGill University, Montreal, Canada) and its complete sequence can be found in GenBank, accession number FJ160463.1. Primers pair 551-552 and 561-562 were used to target *ZCF15* and *ZCF29* respectively. The 2.4 Kb PCR products obtained were transformed in *C. albicans* strains SN250 (Noble and Johnson, 2005) using standard LiAc protocol (Hernday et al., 2010). Arg<sup>+</sup> colonies were streaked in media lacking arginine and genomic DNA prepared using standard chloroform:phenol protocols. In order to verify the correct insertion of the HA-ARG4

cassettes we verify the presence of the new 5' and 3' junctions using primers 546-553 and 559-560 for Zcf15-HA tagged strain and 617-553 and 559-616 for Zcf29-HA tagged strain.

Primer number	Primer sequence
551	GACAGATTCATTAGATGTATTTATGTATAATGTTGAATTGAGTGATATTTA TATGATGTGAAACCCAATGGTCGACGGATCCCCGGGT
552	TTGTTAGTAGTTAATTGATTAGTTTATTTATTTTATAAAGTTATTTATTTAAT CAATATATGTAAAAAACGATGAATTCGAGCTCGTT
546	ACGAACCATCTGCCAGAATCCGTA
553	TCATAGGGATAGCCCGCATAGTCA
559	AGAGATGCTCTTGGTGGTACTGCT
560	AGAATGGAGTACTTTACTATGTTGACGA
561	AAATACATTGATTGATATGAATAATTTATTTGAGAATGTTCCGTTTCGATGAG TTGTTCAAAGATTTTTCCGGTCGACGGATCCCCGGGT
562	ATATATAAAAAAAGGGACAGGGGAATAAAAAATGAGCACAGCACATTCAA GACCTCCTCCCAAACAGAATCGATGAATTCGAGCTCGTT
617	AGCACTGGTGTGCAGCATAAA
553	TCATAGGGATAGCCCGCATAGTCA
559	AGAGATGCTCTTGGTGGTACTGCT
616	CTATCTCTCTTTTTTTTGGCCATTTGAGTTAGTGCATGACCACCACCAACAAA CAGAACA

**Table 4.1: Primers used in this study**

We confirmed the expression of the HA tagged proteins by Western Blot using monoclonal anti HA antibody (Thermo Scientific 2-2.2.14) and a Western Blot protocol optimized for *C. albicans* described here (Gerami-Nejad et al., 2009).

#### **4.2.5 ChIP-Seq: Immunoprecipitation, library preparation, sequencing and peak calling**

In order to compare our ChIP-Seq results with the RNA-Seq one we run our experiment in analogous conditions. Zcf15-HA and Zcf29-HA tagged strains were grown to mid-exponential phase in YPD and samples were crosslinked with formaldehyde either before or

5 and 15 minutes after the addition of 5mM H<sub>2</sub>O<sub>2</sub>. Once crosslinked, cells were lysed and chromatin immunoprecipitated as described here (Hernday et al., 2010). Briefly, 37% formaldehyde was added to a final concentration of 1%, and crosslinked for 15 min at room temperature. The crosslink was quenched by adding glycine to a final concentration of 125 mM, cells were washed twice in TBS and pellets frozen in liquid nitrogen and archived at -80C. Pellets were thawed in ice-cold lysis buffer with fresh protease inhibitor and cells lysed by adding 500µm zirconia beads and using 10 one minutes strokes (6m/s) with a FastPrep Instrument (MP Biomedicals). Lysate was subsequently recovered by piercing the bottom of the tube with 26-gauge needle and chromatin sheered by sonication in a Diagenode Bioruptor (15 min, high settings, 30s on 30 s off). Samples were split between “pulldowns” and “inputs” and while “pulldowns” samples were immunoprecipitated using monoclonal anti-HA antibody (Thermo Scientific 2-2.2.14), “inputs” were not and they were used as controls samples. After immunoprecipitation, crosslinking was reversed by incubating samples at 65C for 16 hours and residual proteins hydrolyzed by adding proteinase K. DNA was finally purified using standard phenol:chloroform protocol. DNA libraries were multiplex using standard Illumina protocols, size selected using SPRIselect reagents (Beckman Coulter #B23317) and the presence of a sharp peak at 300 bp confirmed using Agilent 2100 Bioanalyzer. Samples were finally pooled together and sequenced using Illumina HiSeq 2500. Un-demultiplexed FASTQ files were aligned to the *C. albicans* genome to obtained sequence alignment data (BAM files) that were visualized in IGV (Robinson et al., 2011b). BAM files from immunoprecipitated and input controls libraries were used by MACS peak caller (Zhang et al., 2008) in the web-based genomic platform Galaxy (Goecks et al., 2010) to call Zcf15 and Zcf29 bound regions. Peaks were considered statistically significant only if  $p < 1e-3$ .



## 4.3 Results

### 4.3.1 *In silico* analysis of *C. albicans* (Zn(II)<sub>2</sub>Cys<sub>6</sub>) transcription factors

*ZCF15* belongs to a family of 77 Zn(II)<sub>2</sub>Cys<sub>6</sub> transcription factors that are fungal specific and absent in humans. This unique feature makes them attractive targets for potential drug design (Vandeputte et al., 2011). Here, we run an *in silico* analysis to identify other transcription factors in the same family that (a) are conserved in pathogenic fungi (b) have a role in resistance to ROS (c) have been shown to be required for virulence in mammals. This approach allowed us to identify genes that are evolutionarily conserved in virulence and may modulate virulence by conferring resistance to host effectors.

A previous screen (Maicas et al., 2005) identified a total of 77 *C. albicans* Zn(II)<sub>2</sub>Cys<sub>6</sub> transcription factors in the *C. albicans* genome. Forty-two of these transcription factors have clear *S. cerevisiae* homologs and their functions are conserved between *C. albicans* and baker's yeast. As shown in Table 4.2 (highlighted in grey), these transcription factors retain the names of their *S. cerevisiae* homologs. For example, *C. albicans orf19.3012* is called *ARO80* because in both organisms it has a key role in regulating aromatic amino acid catabolism (Ghosh et al., 2008; Nobile et al., 2003). We therefore focused on the remaining 35 “orphan” transcription factors (highlighted in yellow in Table 4.2) because their biological functions is yet to be fully determined. These 35 genes have been assigned the generic name ZCFs (Zinc Cluster transcription Factors) and progressively numbered from 1 to 35. Of these 35 ZCFs we wanted to identify genes that are conserved in pathogenic fungi but absent in non-pathogenic one.

Allele name	Standard name	Allele name	Standard name	Allele name	Standard name	Allele name	Standard name
orf19.7381	AHR1	orf19.4776	LYS143	orf19.6888	ZFU3	orf19.3876	ZCF19
orf19.4766	ARG81	orf19.5380	LYS144	orf19.3187	ZNC1	orf19.4145	ZCF20
orf19.2748	ARG83	orf19.7372	MRR1	orf19.255	ZCF1	orf19.4166	ZCF21
orf19.3012	ARO80	orf19.3986	PPR1	orf19.431	ZCF2	orf19.4251	ZCF22
orf19.166	ASG1	orf19.6203	PUT3	orf19.1168	ZCF3	orf19.4450	ZCF23
orf19.5097	CAT8	orf19.2747	RGT1	orf19.1227	ZCF4	orf19.4524	ZCF24
orf19.7374	CTA4	orf19.4998	ROB1	orf19.1255	ZCF5	orf19.4568	ZCF25
orf19.4288	CTA7	orf19.3753	SEF1	orf19.1497	ZCF6	orf19.4573	ZCF26
orf19.1499	CTF1	orf19.3308	STB5	orf19.1685	ZCF7	orf19.4649	ZCF27
orf19.5849	CWT1	orf19.7319	SUC1	orf19.1718	ZCF8	orf19.4767	ZCF28
orf19.3127	CZF1	orf19.3188	TAC1	orf19.2077	ZCF9	orf19.5133	ZCF29
orf19.3252	DAL81	orf19.6985	TEA1	orf19.2280	ZCF10	orf19.5251	ZCF30
orf19.6817	FCR1	orf19.7570	UGA3	orf19.2423	ZCF11	orf19.5924	ZCF31
orf19.5729	FGR17	orf19.6038	UGA32	orf19.2623	ZCF12	orf19.5940	ZCF32
orf19.6680	FGR27	orf19.7317	UGA33	orf19.2646	ZCF13	orf19.6182	ZCF34
orf19.5338	GAL4	orf19.2745	UME7	orf19.2647	ZCF14	orf19.5992	ZCF33
orf19.3190	HAL9	orf19.391	UPC2	orf19.2753	ZCF15	orf19.7371	ZCF35
orf19.4225	LEU3	orf19.1035	WAR1	orf19.2808	ZCF16		
orf19.5548	LYS14	orf19.5992	WOR2	orf19.3305	ZCF17		
orf19.4778	LYS142	orf19.6781	ZFU2	orf19.3405	ZCF18		

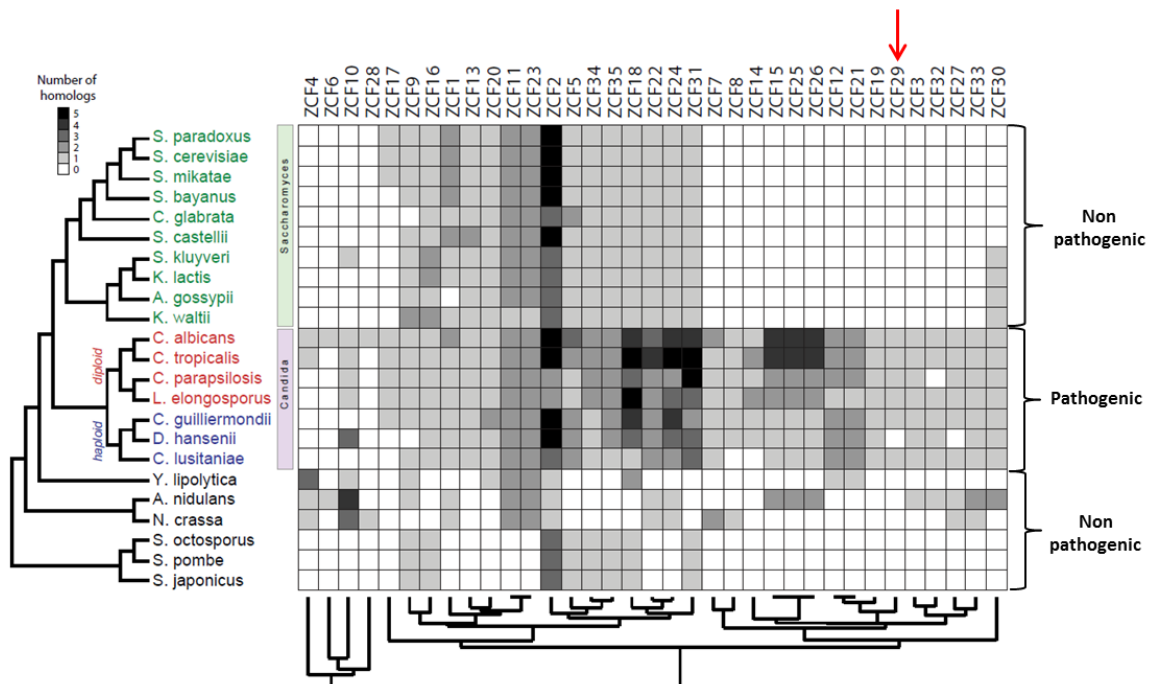
**Table 4.2:** list of all 77 Zn(II)<sub>2</sub>Cys<sub>6</sub> transcription factors present in *C. albicans*.

42 of these genes have been named after their *S. cerevisiae* orthologs because their functions are conserved between the two organisms. 35 of them (yellow) have been called with the generic name *ZCF* (Zinc Cluster transcription Factors) because they either don't have a *S. cerevisiae* orthologs or, if they do, their functions don't seem to be conserved between the two organisms.

We performed a comparative genomic study using a previously developed algorithm (Wapinski et al., 2007) that allowed us to search for *C. albicans* *ZCFs* homologs in 23 other sequenced ascomycetes. These 23 ascomycetes included 7 species (*Candida tropicalis*, *Candida parapsilosis*, *Candida guilliermondii*, *Candida lusitanae*, *Lodderomyces elongisporus*, *Candida albicans* and *Debaryomyces hansenii*) that are considered pathogenic to humans (Chai et al., 2010; Hawkins and Baddour, 2003; Lockhart et al., 2008; Pfaller et al., 2006; Singh and Parija, 2012) and 16 species that are non-pathogenic. These species

include baker's yeast *Saccharomyces cerevisiae*, *Saccharomyces bayanus* (commonly used by the wine making industry), *Saccharomyces kluyveri* (a strict plant pathogen) and *Ashbya gossypii* (commonly found in cotton plants).

For each *ZCF* we determine the number of homologs present in the 23 other ascomycetes and summarized the result in the heat map depicted in Figure 4.2. Overall this study highlights that *ZCFs* are overrepresented in pathogenic fungi compared to non-pathogenic ones as each *ZCF* has an average 1.58 homologs in pathogenic fungi and only 0.55 homologs in non-pathogenic ones. This evidence suggests that this family of transcription factors might have important roles in virulence that are evolutionary conserved.

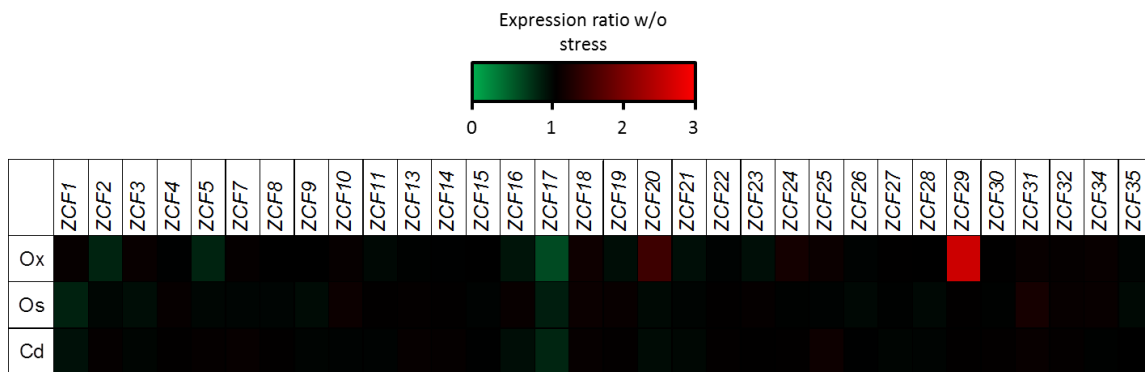


**Figure 4.2: *C. albicans* ZCFs conservation across ascomycetes.**

For each of the 35 *C. albicans* *ZCFs* we determined the number of homologs present in 7 other pathogenic ascomycetes and 16 nonpathogenic one. Number of homologs is represented on a grey scale ranging from 0 to 5. The number of homologs was calculated using SYNERGY as detailed here (Wapinski et al., 2007) and with the help of George Bell, senior Bioinformatic Scientist at the Whitehead Institute.

The fact that *ZCFs* are expanded in pathogenic fungi was encouraging, so we wanted to see if any of them had a role in ROS resistance. ROS plays a fundamental role in the host ability to

counteract infection so a defect in the pathogen's ability to resist ROS may suggest a virulence deficiency. Enjalbert et al., 2006 reported a *C. albicans* gene expression profiling study under oxidative, osmotic and heavy metal stresses. Global transcriptional responses were recorded using microarrays containing ~6000 *C. albicans* probes (~97% of the *C. albicans* genome). When we looked at the *ZCFs* transcriptional changes upon ROS challenge (5 mM H<sub>2</sub>O<sub>2</sub>) we noticed that *ZCF2*, *ZCF5* and *ZCF17* were mildly downregulated (expression ratio w/o H<sub>2</sub>O<sub>2</sub> = 0.81, 0.79 and 0.58 respectively), *ZCF20* mildly upregulated (expression ratio w/o H<sub>2</sub>O<sub>2</sub> = 1.48) and *ZCF29* strongly upregulated (expression ratio w/o H<sub>2</sub>O<sub>2</sub> = 2.6).

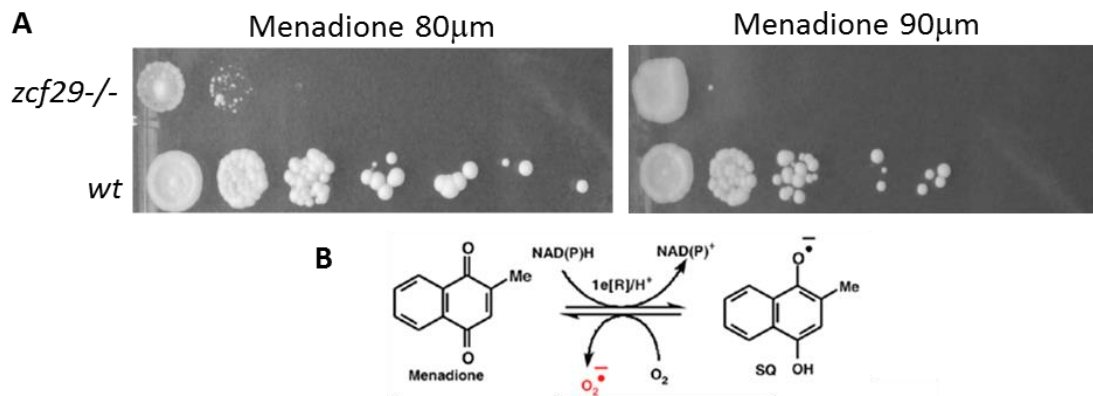


**Figure 4.3: *ZCFs* gene expression in the presence of oxidative stress (Ox), osmotic stress (Os) and heavy metal stress (Hm)**  
 5mM H<sub>2</sub>O<sub>2</sub>, 0.3M NaCl and 0.5 mM CdSO<sub>4</sub> were added to cells growing in mid-exponential phase to trigger Ox, Os and Hm stress respectively. Expression ratios were calculated by comparing stressed cells with their unstressed controls. Data was extracted from (Enjalbert et al., 2006).

Interestingly, *ZCF15* was found not to be upregulated in the presence of H<sub>2</sub>O<sub>2</sub>, suggesting that its role in ROS resistance might take place post transcriptionally or post translationally. The strong *ZCF29* upregulation was observed only in the presence of H<sub>2</sub>O<sub>2</sub> and not in the presence of 0.3M NaCl or 0.5M CdSO<sub>4</sub>, suggesting a specific role of this gene in ROS resistance but not in osmotic or heavy metal stresses. In addition, *ZCF29* has six homologs in

pathogenic fungi and no homologs in non-pathogenic ones (Figure 4.2), potentially indicating a conserved role in virulence. Indeed, Vandeputte et al., 2011 recently demonstrated that *zcf29/zcf29* has a reduced ability to colonize mice kidney when mice were infected by intravenous injection.

In order to gain some insights into *ZCF29* mechanism of action, we obtained the *zcf29/zcf29* knockout from the Fungal Genetic Stock Center and extensively characterized various phenotypes. *zcf29/zcf29* knockout is dramatically susceptible to ROS generator menadione at both 80  $\mu$ M and 90  $\mu$ M (Figure 4.4 A).



**Figure 4.4: *ZCF29* is required for wildtype ability to resist reactive oxygen species**

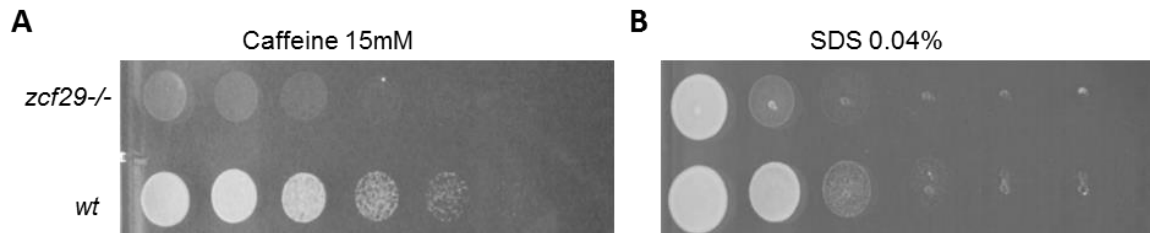
A) *zcf29/zcf29* and wild-type *C. albicans* were serially diluted and spotted in SC media containing 80 and 90  $\mu$ M of menadione and photographed 3 days post spotting.

B) Menadione mechanism of action: menadione is reduced in the mitochondria to its free radical form which is subsequently oxidized back to menadione with concomitant production of highly toxic superoxide radicals (O<sub>2</sub><sup>-•</sup>). Figure adapted from Loor et al., 2010.

In the mitochondria, menadione is reduced to its free radical with concomitant oxidation of NADPH to NADP<sup>+</sup> (Loor et al., 2010). In aerobic condition, menadione radicals are oxidized back to menadione with simultaneous reduction of O<sub>2</sub> to superoxide anion radicals O<sub>2</sub><sup>-•</sup> (Figure 4.4 B). Superoxide anion radicals are an extremely toxic form of reduced oxygen that can lead to protein denaturation, DNA damage and lipid peroxidation (Cooke et al., 2003).

The fact that *ZCF29* is upregulated in the presence of  $H_2O_2$  and that its loss of function leads to increased sensitivity to menadione suggests that it plays an important role in ROS resistance.

In addition, as shown in Figure 4.5, *ZCF29* deletion caused an increased sensitivity to both caffeine and Sodium Dodecyl Sulfate (SDS).



**Figure 4.5: *ZCF29* is required for wildtype resistance to purine analogue caffeine and cell wall stressor SDS.**

A) *zcf29/zcf29* and wild-type *C. albicans* were serially diluted and spotted in SC media containing 15mM caffeine and photographed 3 days post spotting. B) *zcf29/zcf29* and wild-type *C. albicans* were serially diluted and spotted in 0.04% SDS and photographed 7 days post spotting.

Caffeine is a purine analogue and it interferes with the TOR pathway by directly inhibiting the TOR-1 complex (Kuranda et al., 2006). The TOR pathway is a highly conserved protein kinase cascade that, in the presence of abundant nutrients, promotes cell growth and represses genes involved in alternative nutrients utilization. Conversely, when nutrients are limiting, the TOR pathway slows cell growth and redirects cellular resources to scavenge for nutrients (Loewith and Hall, 2011). Caffeine mimics nitrogen starvation conditions and represses the TOR pathway.

SDS is a detergent that interferes with the cell wall polarity and is commonly used to determine cell wall integrity (Delley and Hall, 1999). These two experiments combined suggest that *ZCF29* is required not only for ROS resistance and virulence but also for nutrient regulation and cell wall integrity.

Taken together, our data described in chapter 3 together with data shown here by us and others (Enjalbert et al., 2006; Homann et al., 2009; Vandeputte et al., 2011) indicate that *ZCF15* and *ZCF29* are fungal specific transcription factors important for virulence and ROS resistance. In order to gain some insights into *ZCF15* and *ZCF29* mechanisms of action we leveraged the power of RNASeq to find out which genes and biological pathways are controlled by these two transcription factors.

#### **4.3.2 *ZCF15* and *ZCF29* RNA-Seq QC**

The goal the RNASeq experiment was to compare gene expressions of wild type *C. albicans* with *zcf15/zcf15* and *zcf29/zcf29* deletion both in the presence and absence of ROS. We determined expression profiles of *C. albicans* wildtype, *zcf15/zcf15* and *zcf29/zcf29* either before or 5 and 15 minutes after the addition of 5 mM H<sub>2</sub>O<sub>2</sub>. We chose these time points because the transcriptional response of *C. albicans* to H<sub>2</sub>O<sub>2</sub> has been previously shown to peak around 10-15 minutes post H<sub>2</sub>O<sub>2</sub> addition (Roy et al., 2013) and named these three data points T<sub>0</sub>, T<sub>5</sub> and T<sub>15</sub> respectively.

We decided to use H<sub>2</sub>O<sub>2</sub> to mimic ROS stress *in vitro* because (a) it is the most common ROS used by the immune system to induce respiratory burst (Iles and Forman, 2002) (b) in contrast with other charged ROS, it passes through the cell membrane easily reaching the cytoplasm quickly (Kohchi et al., 2009) (c) various gene expression profiling experiments in wildtype *C. albicans* have been reported and can be used for internal validation of our data (Bruno et al., 2010; Enjalbert et al., 2006).

In order to compare the transcriptome of the different strains in the various condition tested, we performed high-throughput sequencing of cDNA made from poly(A) selected RNAs. The

sequencing libraries were prepared using a strand-specific approach that preserved gene transcription directionality. This approach allowed us to assign every read to a single gene, with no ambiguity. This aspect was especially critical considering that *C. albicans* has a very compact genome with many neighboring genes transcribed in opposite directions. For each strain, three biologically independent samples were analyzed and we estimated gene expression abundance using a pipeline developed at the Broad Institute. The pipeline essentially consisted in 5 steps:

1. RNASeq reads were aligned to the *C. albicans* reference genome using Tophat (Trapnell et al., 2012).
2. Libraries quality was determined using RNA-SeQC, a recently developed tool that provides key measurement of data quality (DeLuca et al., 2012) like number of reads aligned, sequencing depth, continuity of coverage etc.
3. Gene expression for every gene was quantified using Cufflinks (Anders et al., 2012)
4. Reads were normalized by gene length and sample size and expressed as reads per kilobases per million of reads (RPKM) as described here (Li and Dewey, 2011)
5. Gene statistically significantly differentially expressed between conditions were identified using EdgeR (Robinson et al., 2010).

We obtained a total of 153,693,150 reads of which 113,140,942 reads mapped uniquely to the genome. The *mapped uniquely rate* was therefore ~74%, in line with what seen by others (Tuch et al., 2010). Reads were evenly distributed across strains and time points and a snapshot of their distribution is depicted in Figure 4.6.





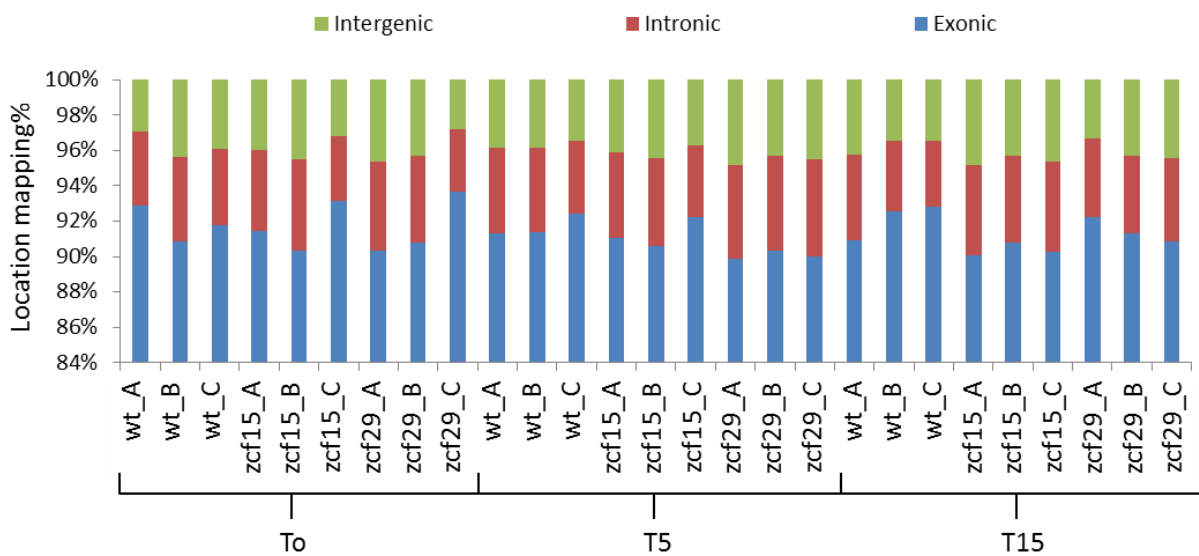
**Figure 4.6: RNASeq reads distribution across strains and time points**

Reads obtained from the RNASeq experiment were equally distributed across strains and time points as shown by the total number of reads per sample (○). We obtained an average of 5.6 million of reads per sample and 70-80% of these reads were uniquely mapped to the genome (●).

For wild-type, *zcf15/zcf15* and *zcf29/zcf29* we obtained 49,251,867, 50,665,642 and 53,775,641 total reads respectively of which 34,408,368, 38,408,587 and 40,323,987 were uniquely aligned to the genome. Total number of reads was also evenly distributed across time points with 44,548,377, 66,575,585 and 42,569,188 for samples extracted at To (before the addition of H<sub>2</sub>O<sub>2</sub>), T<sub>5</sub> (5 minutes post addition of H<sub>2</sub>O<sub>2</sub>) and T<sub>15</sub> (15 minutes after the addition of H<sub>2</sub>O<sub>2</sub>) respectively. Overall we obtained between 2 and 10 million of reads per sample and 70-80% of these reads uniquely aligned to the genome. This number of reads is adequate to draw meaningful conclusions in *C. albicans* as it is estimated that high quality gene expression profiling can be with 20 million reads in humans, 4 million for worms and flies and above 2 million for yeast (Bailey et al., 2013).

Overall we detected expression of 6376 genes of the 6467 genes annotated in the *C. albicans* and the 91 genes for which we couldn't detect expression (for example *ORF19.7340* and *ORF19.3186*) are ORFs annotated as dubious and unlikely to encode for a protein (Inglis et al., 2012).

Figure 4.7 shows read distribution between intronic, exonic and intergenic regions. Spurious reads that aligned to intergenic region of the genome (possibly due to unannotated genes or sequencing error) were minimal (~2-4%) in accordance with what shown by others (Dhamgaye et al., 2012). In addition, very few reads aligned to intronic regions (~3-4%), suggesting that our initial poly-A selection was effective in removing most of the genomic DNA and that only mature mRNAs were sequenced in our samples.

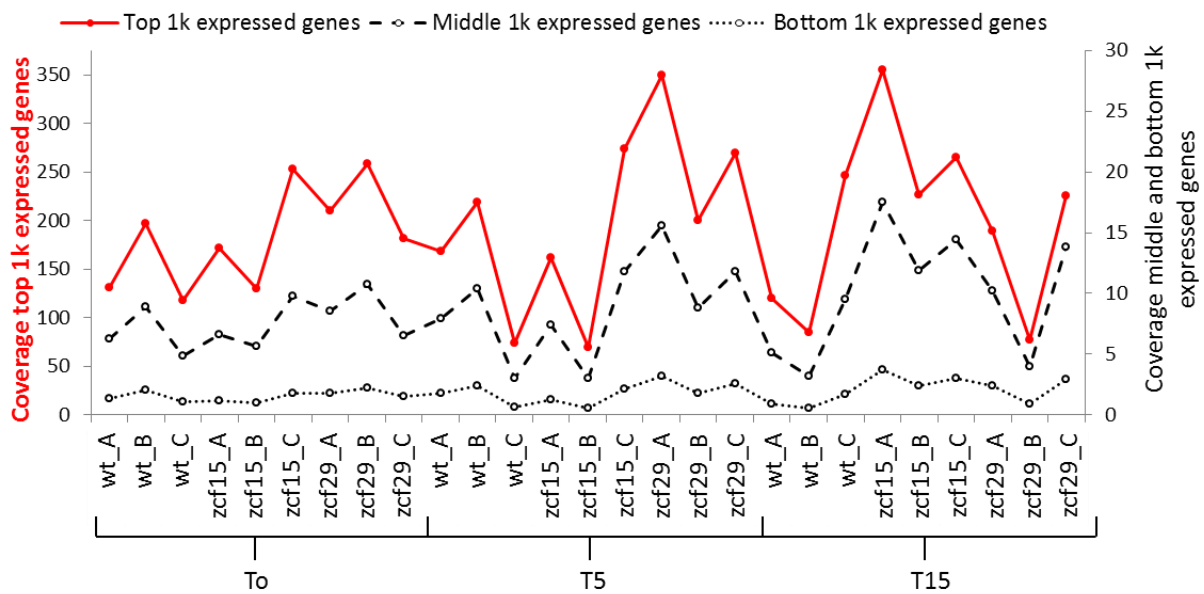


**Figure 4.7: RNA-Seq reads distribution between exons, introns and intergenic regions across samples.**

The large majority of the reads mapped to exonic region (~90-92%) with only few reads mapping to intronic regions (3-4%) and intergenic regions (2-3%). The low levels of intronic reads confirmed that our initial poly-A selection was effective in removing most of the genomic DNA.

We also calculated sequencing depth by determining coverage for the top 1000 most expressed genes (highly expressed genes), middle 1000 expressed genes (medium expressed

genes) and bottom 1000 expressed genes (low expressed genes). Sequencing depth is used as a measure of the number of times a nucleotide is read during sequencing. Highly expressed genes were sequenced at a depth ranging from 354x to 69x, genes expressed at medium levels were sequenced at a depth ranging from 17x to 3x and those expressed at low levels at a depth between 3x and 2x. Figure 4.8 shows sequencing coverage across different samples and time points.



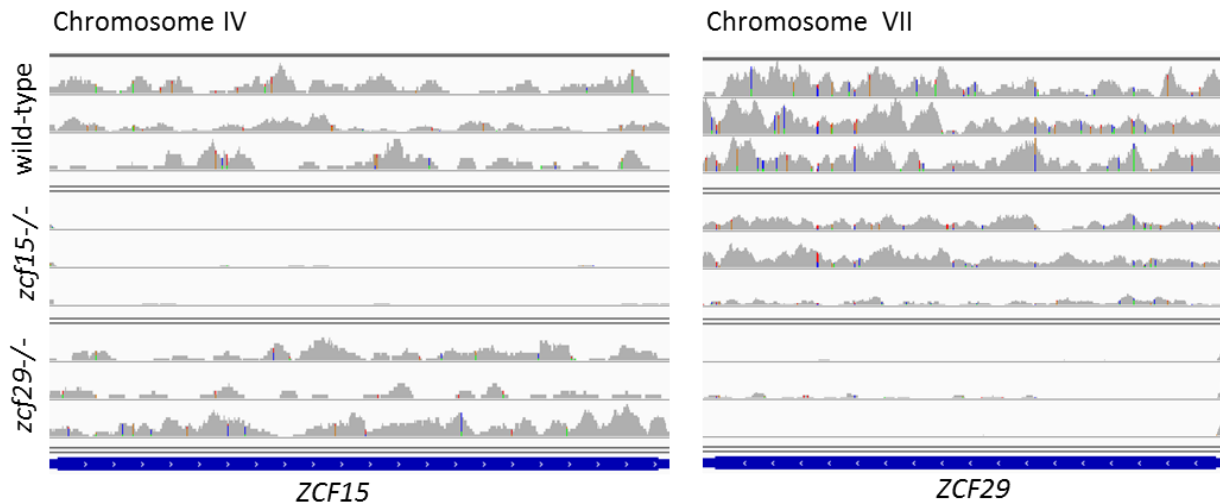
**Figure 4.8: RNA-Seq coverage distribution metrics across different strains and time points.**

Highly expressed genes were sequenced at a depth ranging from 69x to 354x, medium expressed genes at a depth ranging from 3x to 17x and low expressed gene at a depth ranging between 2x and 3x.

Overall, our RNA-SeQC quality control confirmed that we sequenced our samples at very high depth of sequencing and obtained reads that are equally distributed across time points and samples. Assessing sequencing performance and library quality was critical before proceeding to further computational our analysis.

In addition to the RNA-SeQC quality control, we manually confirmed that *ZCF15* and *ZCF29* were not expressed from the respective knockouts. We used Interactive Genomics

Viewer (IGV), an algorithm developed for large genomic dataset visualization (Robinson et al., 2011b) to visualize expression levels in the knockout loci. As shown in Figure 4.9 we couldn't detect *ZCF15* and *ZCF29* expression from *zcf15/zcf15* and *zcf29/zcf29* strains suggesting that the genes were properly deleted.



**Figure 4.9: *ZCF15* and *ZCF29* expression was not detected in the respective knockout strains.**

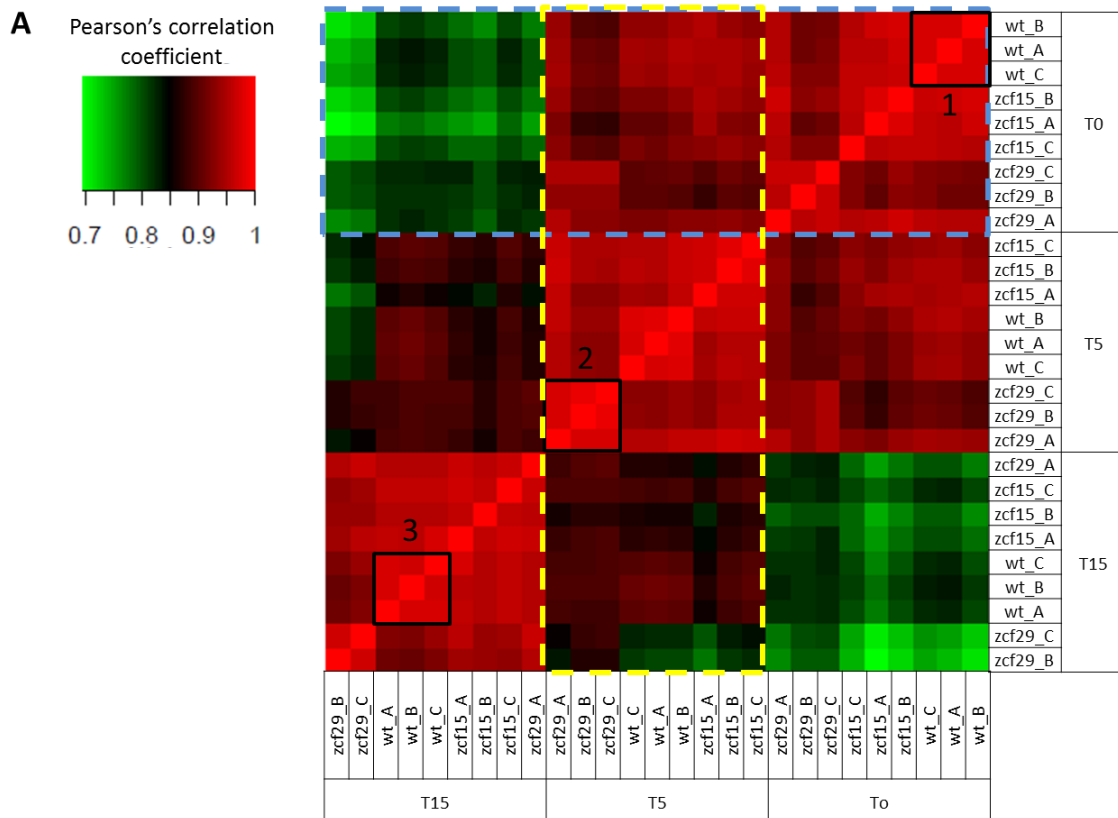
The density of the reads piling up in the *ZCF15* (chromosome IV) and *ZCF29* (chromosome VII) locus were visualized using Integrative Genome Viewer (IGV). *ZCF15* expression was not detected in any of the three *zcf15/zcf15* biological replicates. *ZCF29* expression was not detected in any of the three *zcf29/zcf29*- biological replicates.

Once we confirmed that the libraries obtained in our experiment were high quality, we estimated gene expression abundance using the algorithm Cufflinks. Cufflinks determined the absolute number of reads coming from each gene. However, the absolute number is not informative because it is a function of sequencing depth and gene lengths. Longer transcripts are more likely to have sequences mapping to their regions. Concurrently, deeper sequencing results in higher absolute number of reads. In order to control for these biases, we determined gene expression levels using reads per kilobases per million of reads mapped (RPKM) as described here (Li and Dewey, 2011). RPKMs are standard measurements of gene expression

that can be used for comparative studies since their absolute values don't depend on gene size or sequencing depth.

We first used RPKM to evaluate gene expression reproducibility across biological replicates. We used the  $\log_2(\text{RPKM})$  as our variable and made pairwise comparison across all 27 samples. Sample correlation was measured using the Pearson Correlation Coefficient (PCC) which measures the strengths of a linear relationship between variables. We obtained a total of 729 ( $27 \times 27$ ) correlation coefficients and summarized them as a heat map in Figure 4.11.

Biological replicates were highly correlated between each other (Figure 4.10, small black boxes for example) reassuring us on the reproducibility of our replicates. This high degree of transcriptional correlation between biological replicates was observed both before (Figure 4.10, black box 1) and after the addition of  $\text{H}_2\text{O}_2$  (Figure 4.10, black box 2 and 3). In contrast, transcriptional profiles obtained from the same strains before and after the addition of  $\text{H}_2\text{O}_2$  were dramatically different as suggested by their poor degree of correlation (Figure 4.10 dashed blue rectangle). This evidence indicates clearly that  $\text{H}_2\text{O}_2$  triggered a large transcriptional rewiring especially if we compared samples at  $T_0$  and  $T_{15}$ . Interestingly, samples at  $T_5$  were highly correlated within each other but poorly correlated with samples at both  $T_0$  and  $T_{15}$  (Figure 4.10, yellow dashed rectangle), suggesting that the  $T_5$  samples effectively captured an intermediate step in the  $\text{H}_2\text{O}_2$  driven transcriptional rewiring obtained between  $T_0$  and  $T_{15}$ .

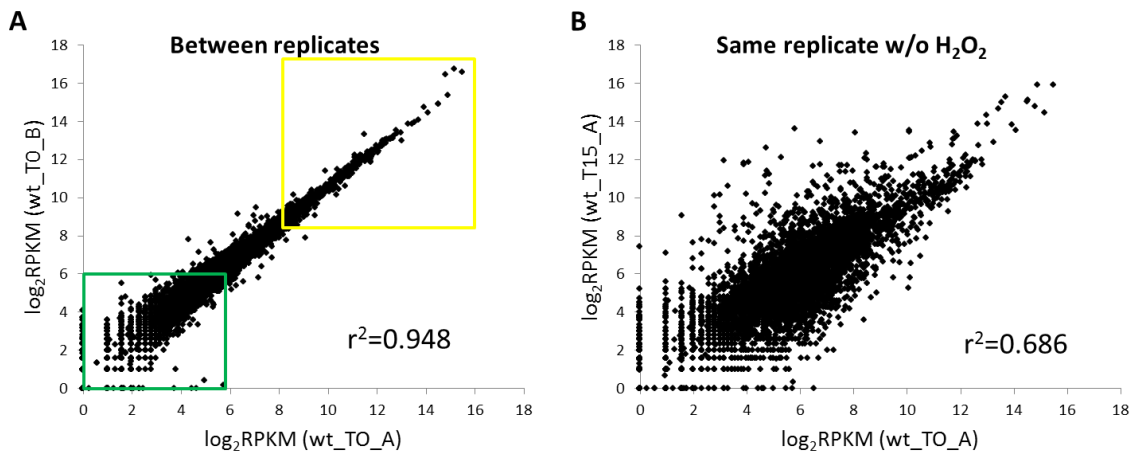


**Figure 4.10: Pearson's correlation coefficients heat map for each of the 27 gene expression profiles obtained in this study.**

Black boxes indicate high degree of correlation between biological replicates. Blue dashed rectangle indicates a progressively decreasing degree of correlation between samples at T<sub>0</sub>, T<sub>5</sub> and T<sub>15</sub> suggesting that the addition of H<sub>2</sub>O<sub>2</sub> triggered a large transcriptional change across all strains. Yellow dashed rectangle shows low degree of correlation between T<sub>5</sub> samples and both T<sub>0</sub> and T<sub>15</sub> samples. This evidence highlights that T<sub>5</sub> samples effectively captured an intermediate step in the H<sub>2</sub>O<sub>2</sub> driven transcriptional change observed between T<sub>0</sub> and T<sub>15</sub>. As rule of thumb, PCC is typically from 0.3-0.4 for unrelated samples to > 0.9 for highly correlated replicate samples (Bailey et al., 2013).

Biological replicates were highly correlated within each other but we wanted to check if this high degree of correlation was observed for both high and low expressed genes. We therefore plotted genes expressions levels of one replicate against another. As an example (Figure 4.11 A) when we plotted gene expression of wild type (replicate A) at T<sub>0</sub> against gene expression of wild type (replicate B) at T<sub>0</sub> we observed a high degree of correlation ( $r^2 = 0.948$ ). The high degree of correlation was observed across genes expressed at both low levels (Figure

4.11 A green box) and high levels (Figure 4.11 A, yellow box), demonstrating the robustness of our RNA-Seq experiments and the absence of biases towards genes expressed at different levels. In contrast, when we compared expression profiles of the same replicate before and after the addition of H<sub>2</sub>O<sub>2</sub>, (Figure 4.11 B) we observed a poor degree of correlation ( $r^2 =$  suggesting that H<sub>2</sub>O<sub>2</sub> triggered a large transcriptional change).



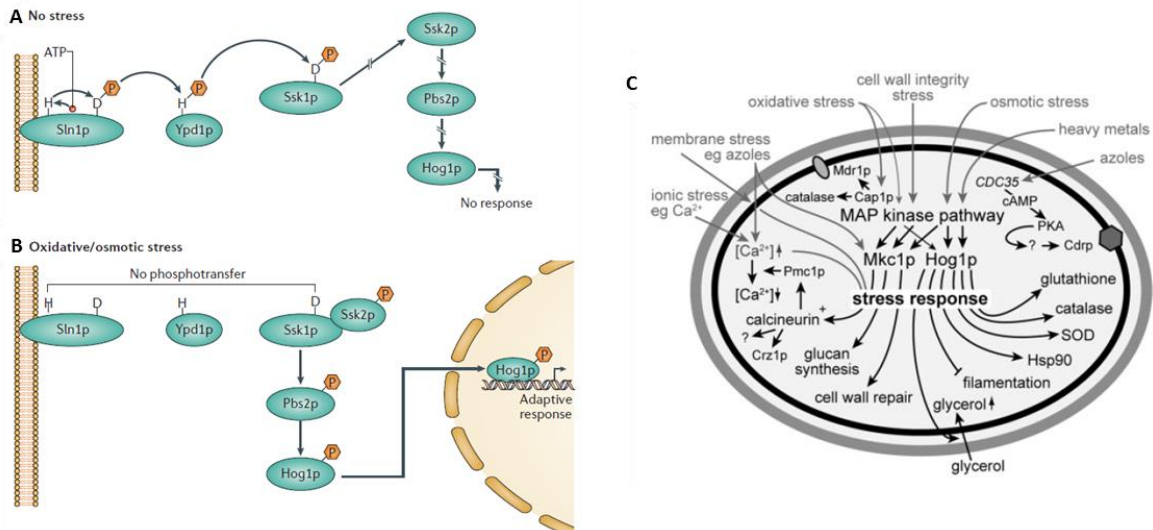
**Figure 4.11: Transcriptional correlation between biological replicates and within the same replicate before and after the addition of H<sub>2</sub>O<sub>2</sub>.**

- A) At T<sub>0</sub>, transcriptional profiles of *C. albicans* wildtype replicate A is highly correlated with replicate B. The high degree of correlation was observed for both highly expressed genes (yellow box) and lowly expressed genes (green box).  $r^2 = 0.86$  for lowly expressed genes,  $r^2 = 0.97$  for highly expressed genes and  $r^2 = 0.948$  across genes expressed at all levels.
- B) Linear correlation between transcriptional profiles of the same replicate before (X-axis) and after (Y-axis) the addition of H<sub>2</sub>O<sub>2</sub>. The poor correlation between the two samples suggests that H<sub>2</sub>O<sub>2</sub> triggered a large transcriptional change.

Once we normalized gene expression for all samples, we used the algorithm EdgeR to identify genes statistically significantly differentially expressed between conditions. EdgeR has been coded so reads coming from biological replicates are combined and pairwise comparisons between conditions can be assessed using an overdispersed Poisson model (Robinson et al., 2010).

### 4.3.3 ZCF15 and ZCF29 are not regulating canonical genes involved in ROS detoxification

We first explored the possibility that *ZCF15* and *ZCF29* are regulating the canonical genes involved in ROS detoxification processes. In *C. albicans* two major pathways are known to be involved in ROS resistance: the Hog1 and the Cap1 pathway (Figure 4.12).



**Figure 4.12: The Hog1 and Cap1 pathway are the two major known pathways involved in oxidative stress response.**

- In the absence of oxidative stress the histidine kinase Sln1 is autophosphorylated on its histidine (H) and aspartate (D) residues. In these conditions, Sln1 phosphorylates Ypd1, which subsequently transfer its phosphate group to Ssk1. Ssk1 phosphorylation inhibits the activation of the MAPK pathway and cause Hog1 cytoplasmic retention. Figure adapted from (Chauhan et al., 2006).
- In the presence of oxidative or osmotic stress, the Sln1-Ypd1-Ssk1 phosphotransfer cascade is inhibited and Ssk1 activates the MAPK pathway. Activation of the MAPK pathway leads to the phosphorylation of Hog1 that translocate to the nucleus and activates oxidative/osmotic stress genes.
- Cartoon that summarizes the pathways involved in *C. albicans* resistance to various stresses. In addition to the MAPK pathway, oxidative stress activates the transcription factor Cap1 that controls the expression of catalase and efflux pumps like Multi Drug Resistance genes (Mdr1). Figure adapted from Enjalbert et al., 2006

The Hog1 pathway (high osmolality glycerol) is a general stress response pathway involved in oxidative, osmotic and heavy metal stress (Smith et al., 2004). This pathway is believed to be activated only at the high concentration of peroxide and relies on a series of signal

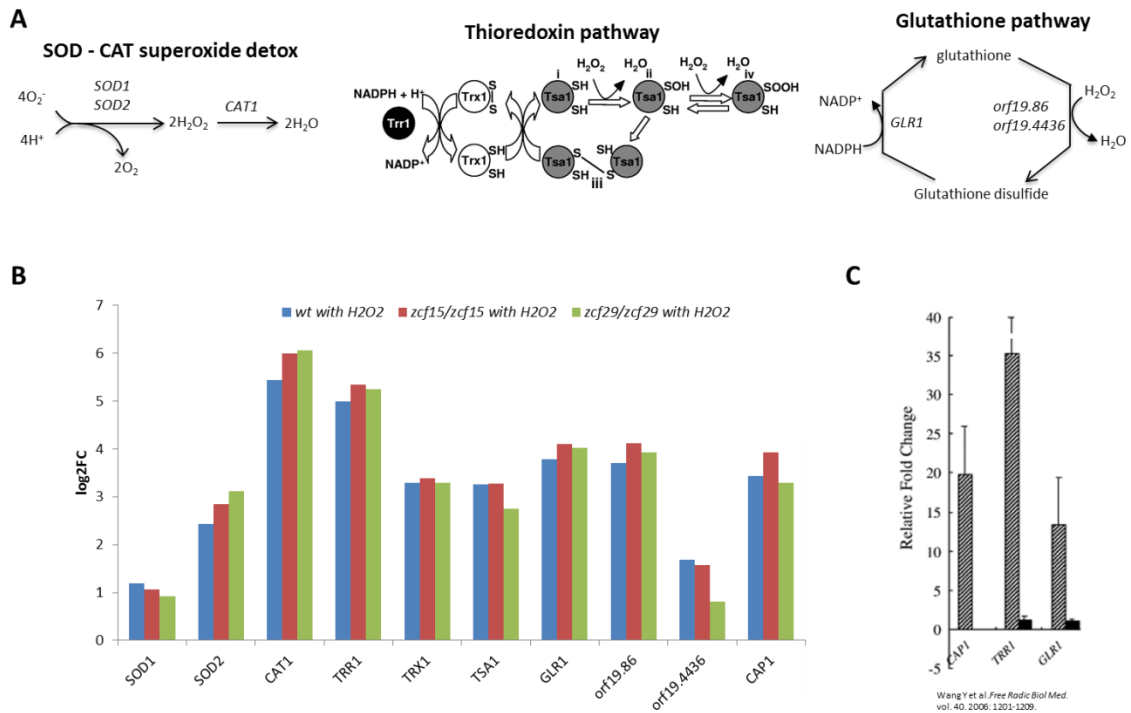


transducer proteins that phosphorylate the transcription factor Hog1 in the presence of oxidative stress (Figure 4.12 A and B). The Hog1 pathway plays a central role in regulation of stress response genes and various mutants in this pathway have been shown to be hypersensitive to ROS (Nagahashi et al., 1998).

In addition to the Hog1 pathway, *C. albicans* relies on the Cap1 pathway to withstand ROS stress. Cap1 is a bZIP transcription factor that is activated at both high and low ROS concentrations and controls the expressions of detoxifying enzymes like superoxide dismutase (*SOD1* and *SOD2*), catalase (*CAT1*) and Multi Drug Resistance efflux pumps like *MDR1* (Figure 4.12 C).

Together, Hog1 and Cap1 control the expression of genes involved in ROS defense mechanisms like SOD-CAT superoxide detoxification, thioredoxin antioxidant mechanism and glutathione radical scavenging process (Figure 4.13 A). In order to evaluate if any of these canonical ROS resistance processes are controlled by *Zcf15* or *Zcf29* we measured their H<sub>2</sub>O<sub>2</sub> dependent upregulation in wildtype and deletion strains. As shown in Figure 4.13, genes involved in these pathways were upregulated in *zcf15/zcf15* and *zcf29/zcf29* to the same extent as wildtype. This evidence suggests that *ZCF15* and *ZCF29* don't control these detoxifying processes and are likely involved in novel pathways that likely don't crosstalk with the very well characterized Hog1 and Cap1 pathways.

Our analysis allowed us to compare the degree of upregulation we observed in our dataset with literature results. Wang et al., 2006 reported a 20, 35 and 13 folds upregulation for *CAP1*, *TRR1* and *GLR1* respectively upon exposure of 1 mM H<sub>2</sub>O<sub>2</sub> (Figure 4.13 C). We obtained similar results with *CAP1*, *TRR1* and *GLR1* upregulated on average 12, 37 and 16 folds (Figure 4.13 B), suggesting that our results recapitulate what has been shown by others and confirming experimental reproducibility between microarray and RNA-Seq data.



**Figure 4.13: ZCF15 and ZCF29 are not regulating canonical ROS resistance pathways.**

- A) Upon ROS exposure *C. albicans* upregulates superoxide dismutase *SOD1* and *SOD2* that convert superoxide ions to  $H_2O_2$  which is subsequently converted into  $H_2O$  by *CAT1*. The thioredoxins pathway acts as an electron donor to various peroxidases and other reductases and helps in scavenging toxic ROS. The glutathione pathway functions as a cellular redox buffer scavenging toxic ROS through the action of *GLR1*, *orf19.86* and *orf19.4436*.
- B) Genes involved in SOD-CAT superoxide detoxification and thioredoxin/glutathione ROS scavenging are upregulated in *zcf15/zcf15* and *zcf29/zcf29* to the same extent as wildtype. This evidence suggests that *ZCF15* and *ZCF29* don't control these detoxifying processes and are likely involved in novel pathways.
- C) The upregulation observed in our dataset is comparable to what has been shown by Wang et al., 2006 confirming experimental reproducibility between microarray and RNA-Seq experiments.

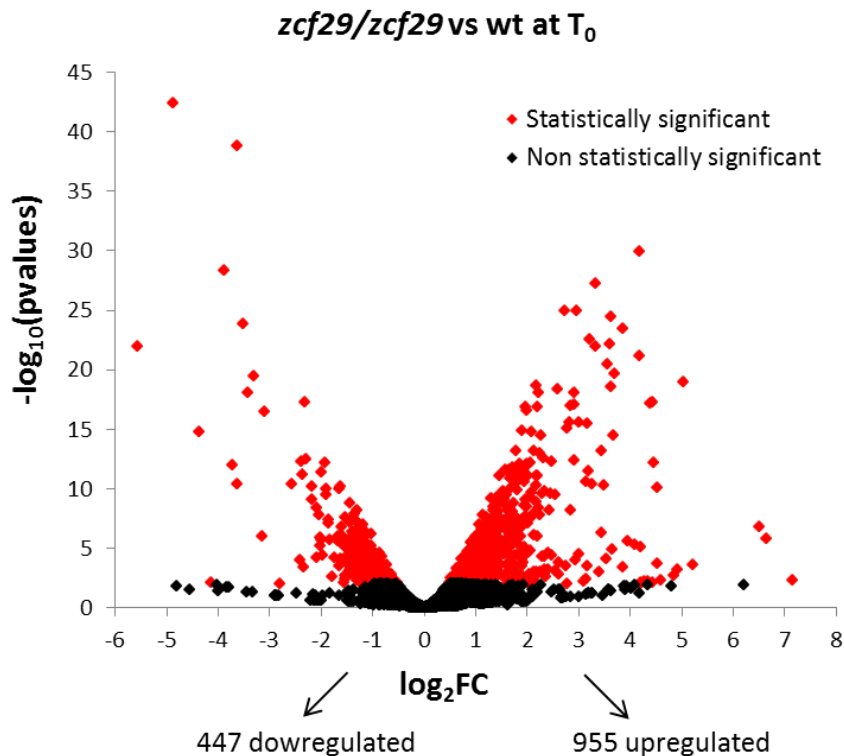
Taken together, our results showed that a) the addition of  $H_2O_2$  activated the canonical genes involved in ROS detoxification processes in both wild type and deletions strains b) the upregulation in our dataset is comparable to what have been observed by others (Wang et al., 2006) c) *ZCF15* and *ZCF29* are likely not controlling the canonical ROS resistance pathways. This evidence suggests that *ZCF15* and *ZCF29* are likely controlling novel uncharacterized ROS resistance pathways.

#### 4.3.4 *ZCF29* plays a critical role in N-metabolism

##### 4.3.4.1 *ZCF29* deletion leads to misregulation of N-metabolism

We first identified genes differentially expressed between wildtype and *zcf29/zcf29* strains growing in mid- exponential phase in the absence of H<sub>2</sub>O<sub>2</sub> (T<sub>0</sub>). Genes differentially expressed between these two conditions are represented using Volcano Plots (Figure 4.14), a type of scatter-plot in which statistical significance (pvalues) is plotted against fold-change (FC) on the y- and x-axes respectively. Obviously, pvalues and FC are expressed in a logarithmic scale since both variables are heavily skewed in a linear scale (Li, 2012). As shown in Figure 4.14, genes statistically significantly differentially expressed ( $p < 0.01$ ) are reported as red dots, while genes whose differential expression is not statistically significant are reported as black dots. Genes upregulated in *zcf29/zcf29* compared to wildtype are reported on the right side of the graph, while genes downregulated in *zcf29/zcf29* compared to wild-type are reported to the left side of the graph.

As shown in Figure 4.14, deletion of *ZCF29* causes a major transcriptional change with a total of 1402 genes statistically significantly differentially expressed. This large genes set correspond to a remarkable 22% of the genome suggesting that this transcription factor may regulate a variety of biological functions. The fact that *ZCF29* deletion causes a pronounced transcriptional change in *C. albicans* recapitulates the variety of different phenotypes observed for the deletion. These phenotypes include hypersuseptability to cell wall stressor SDS, ROS and caffeine (Figure 4.4 and Figure 4.5).

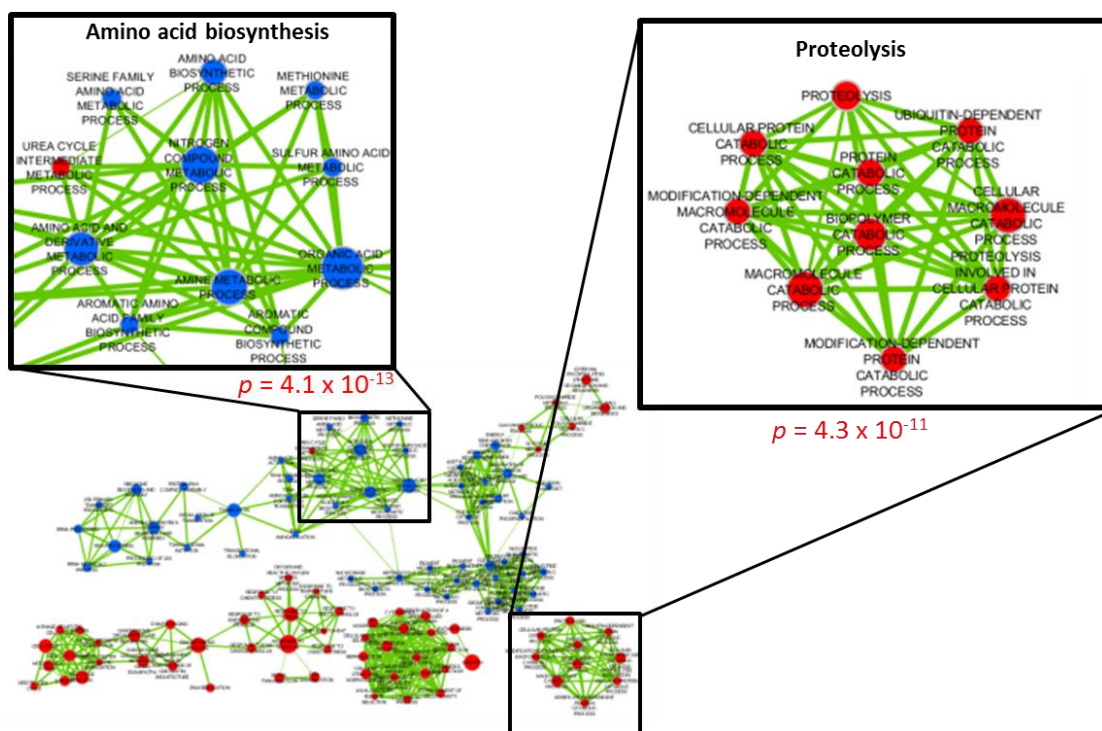


**Figure 4.14: Volcano plots showing genes differentially expressed between wild type and *zcf29/zcf29*.**

*ZCF29* deletion leads to the statistically significant differential expression of 1402 genes: 955 upregulated and 447 downregulated. Differential gene expression was reported as log<sub>2</sub>FC on the X-axis and statistical significance (–log<sub>10</sub> (pvalue)) on the Y-axis. Genes statistically significantly upregulated in *zcf29/zcf29* compared to wildtype are reported as red dots on the right side of the graph while genes statistically significantly downregulated as red dots to the left of it. Genes not statistically significantly differentially expressed are reported as black dots.

In order to extract biological insights from the group of genes differentially expressed in *ZCF29* deletion, we used a powerful analytical algorithm called GSEA (Gene Set Enrichment Analysis). GSEA identifies statistically over-represented functions given the differentially expressed genes and an *a priori* set of genes and biological functions that are known in *C. albicans* (Marotta et al., 2013; Subramanian et al., 2005). Once we obtained the biological functions enriched in our dataset we used Cytoscape (Shannon et al., 2003; Smoot et al., 2011) and the Enrichment map plugin (Merico et al., 2010) to clearly visualize enriched biological functions. Figure 4.15 shows the enrichment map for *zcf29/zcf29* strains compared

to wildtype. Enrichment maps are network representations made of nodes (circles) connected within each other by edges (lines). Blue nodes represent downregulated biological functions, red upregulated one. Node size represents the number of genes differentially expressed present in that function while edges indicate the relatedness between nodes (the larger the number of genes in common between biological functions the thicker the line). When *ZCF29* is deleted the two most statistically significant enriched biological functions are amino acid biosynthesis ( $p = 4.1 \times 10^{-13}$ ) and proteolysis ( $p = 6.9 \times 10^{-11}$ ).



**Figure 4.15: Enrichment map of biological functions overrepresented in *zcf29/zcf29* compared to wildtype.**

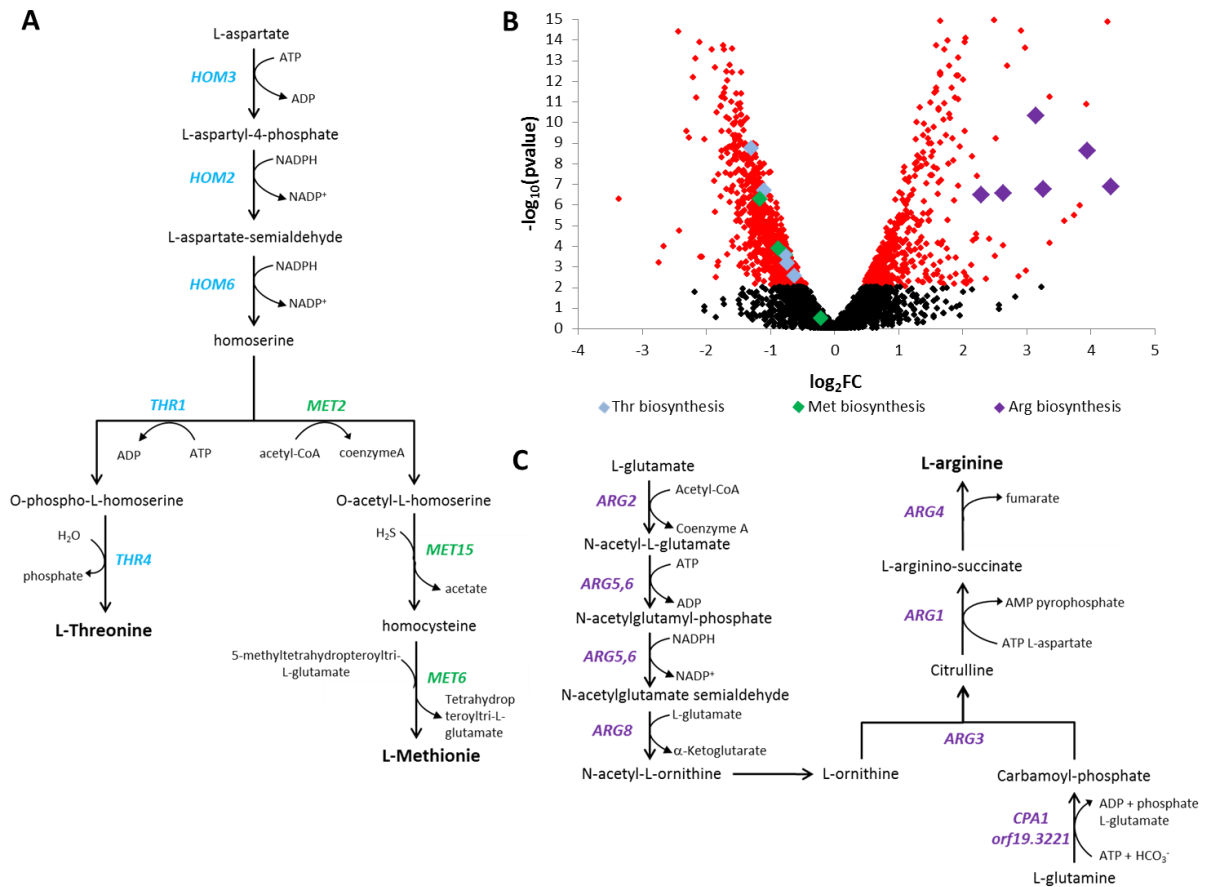
Circles (nodes) represent biological functions enriched in *zcf29/zcf29* deletion compared to wildtype. Blue nodes represent downregulated biological functions, while red one upregulated one. The two most statistically significant enriched biological functions are amino acids biosynthesis and proteolysis. Lines (edges) connecting nodes indicate the relatedness between nodes, the larger the number of genes in common between biological functions the thicker is the line.

Amino acid biosynthesis and proteolysis are two highly interconnected biological functions as *C. albicans* is constantly striking a balance between amino acids *de novo* synthesis and

amino acid recycling through proteolysis. Interestingly, wild type *C. albicans* favors the latter over the former under N-starvation condition (Ramachandra et al., 2014). While both pathways can produce free amino acids as a final product, amino acid biosynthesis is more energy demanding than proteolysis and this is believed to be the reason why *C. albicans* favors the latter under N-starving condition. *ZCF29* deletion causes a transcriptional change that resembles the transcriptional change observed in wild type *C. albicans* under N-limiting conditions (Ramachandra et al., 2014).

In particular, when we looked more in depth at the specific metabolic pathways differentially expressed in *zcf29/zcf29* we found that essentially every gene involved in methionine and threonine biosynthesis were downregulated. The biological pathways that lead to the synthesis of these two amino acids are reported in Figure 4.16 A. These two pathways are highly interconnected within each other as the first three steps (from aspartate to homoserine) are in common between them. Each of the five genes involved in the synthesis of threonine are statistically significantly down regulated (blue dots Figure 4.16 B) in *zcf29/zcf29* and the same is true for 5 of the 6 genes involved in the synthesis of methionine (green dots Figure 4.16 B). The 5 genes involved in threonine biosynthesis (*HOM3*, *HOM2*, *HOM6*, *THR1* and *THR4*) are statistically significantly downregulated between 2.5 to 1.5 folds while genes involved in the methionine biosynthesis (*HOM3*, *HOM2*, *HOM6*, *MET2* and *MET6*) between 2.1 and 1.7 folds. In addition to methionine and threonine, many other genes involved in amino acid biosynthesis were statistically significantly downregulated. These genes include *ARO3*, *ARO4*, *ARO2*, *TRP2* and *TRP4* required for the biosynthesis of aromatic amino acids and *LYS4*, *LYS12*, *HIS1*, *HIS4*, *HIS5*, *SER1*, *SER2*, *LEU1*, *LEU2*, *LEU4* required for the biosynthesis of lysine, histidine, serine and leucine. This data suggest that overall amino acid

biosynthesis is highly downregulated in *zcf29/zcf29* a condition commonly observed in *C. albicans* under N-starvation conditions.



**Figure 4.16: ZCF29 deletion causes misregulation of threonine, methionine and arginine biosynthesis.**

- A) *C. albicans* threonine and methionine biosynthetic pathway. The first 3 steps of the pathways are in common between the two amino acids. Once homoserine is synthesized, it can either be acetylated leading to the subsequent formation of methionine or phosphorylated leading to the subsequent formation of threonine.
- B) Volcano plot showing that *ZCF29* deletion leads to the statistically significant downregulation of genes involved in threonine and methionine biosynthesis (blue and green dots) and concurrent upregulation of genes involved in arginine biosynthesis (purple dots).
- C) *C. albicans* arginine biosynthetic pathway. There are 7 genes involved in this pathway, 6 of the 7 genes are upregulated in *ZCF29* deletion. Interestingly, the only gene in the pathway not upregulated in our data is *ARG2*, a gene that is largely regulated post-transcriptionally in *S. cerevisiae* (Wipe and Leisinger, 1979)

Interestingly, in this context of extensive downregulation of amino acid biosynthesis, we found that the arginine biosynthetic pathway was substantially upregulated in *zcf29/zcf29*. This data was somewhat surprising as 6 of the 7 genes involved in this pathway (*ARG1*, *ARG3*, *CPA1*, *ARG8*, *CPA2*, *ARG5,6*) were upregulated between 23 and 6 folds and the only gene in the pathway not upregulated in our dataset is *ARG2*, a gene that in *S. cerevisiae* has been shown to be largely regulated posttranslationally (Wipe and Leisinger, 1979). This result was surprising especially considering that many other genes involved in amino acid biosynthesis were downregulated in *zcf29/zcf29*. However, arginine biosynthetic pathway has been recently linked to ROS resistance. Jimenez-Lopez et al., 2013 recently showed that *C. albicans* upregulates the arginine biosynthetic pathway upon exposure to macrophage generated ROS. Interestingly, similarly to what we have seen, the only gene in the pathway not upregulated upon exposure to macrophage generated ROS was also *ARG2*. Although the biological reasons behind this upregulation are still under active investigation, our data reinforce a yet poorly understood connection between arginine biosynthesis upregulation and reactive oxygen species resistance.

As shown in Figure 4.15, the second most enriched biological function upon *ZCF29* deletion is proteolysis. Proteolysis is a complex biological process that in *C. albicans* includes 4 major steps: proteasome assembly, protein target ubiquitination, protein degradation and amino acids recycling. Interestingly genes involved in each of these four major steps were upregulated in *zcf29/zcf29*. Genes directly involved in protein degradation like *APE3*, *CPY1*, *LAP3*, *LAP41*, *PRC2*, *PRC3* coding for various amino- and carboxy-peptidases were upregulated between 2 and 18 folds in *ZCF29* knockout. In addition genes involved in 26s proteasome assembly, ubiquitination and amino acid recycling like *DOA1*, *DOA4*, *HRT1*, *PR26*, *RPN1*, *RPN2*, *RPT1*, *RPT2*, *RPT4*, *RPT5*, *RPT6* and *UBP6* were all statistically

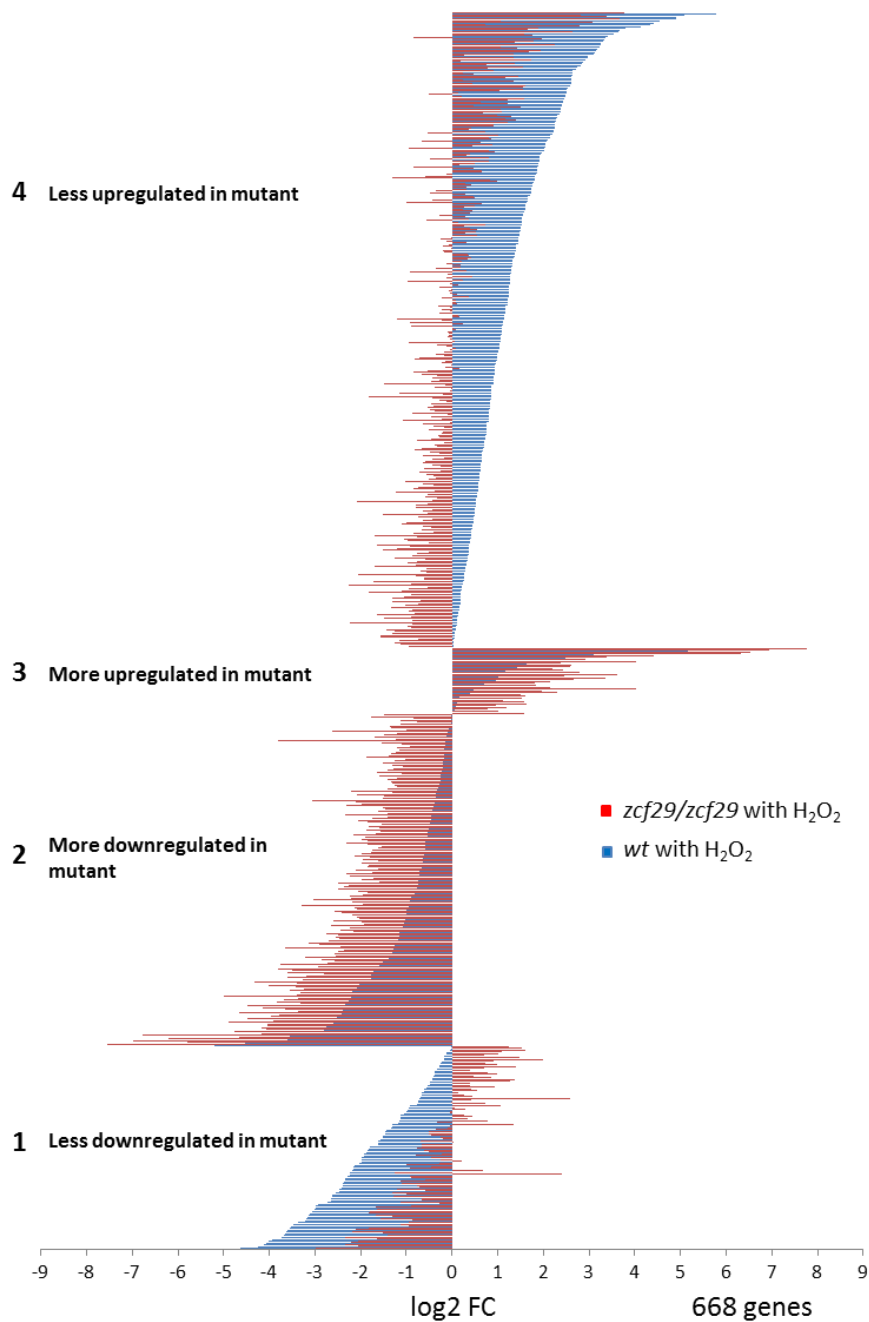


significantly upregulated. In wild type cells many of these genes have been shown to be upregulated after switching cells from rich media to media lacking nitrogen (Ramachandra et al., 2014) reinforcing the idea that *ZCF29* deletion leads to a transcriptional response that recapitulates the one observed in wild type cells under N-starvation.

Our RNAseq experiments allowed us to identify genes differentially expressed upon *ZCF29* deletion. Although we can't discriminate if these genes are directly or indirectly controlled by *Zcf29* and further biological experiments are required to pinpoint its regulatory network, we believe that *ZCF29* plays a critical role in the *C. albicans* ability to regulate amino acids biosynthesis, proteolysis and N-metabolism. This evidence is corroborated by the fact that *ZCF29* deletion leads to an increased sensitivity to caffeine, which mimics N-starvation (Figure 4.5 A).

#### **4.3.4.2 *ZCF29* plays a critical role in down regulating ribosome biogenesis upon ROS challenge**

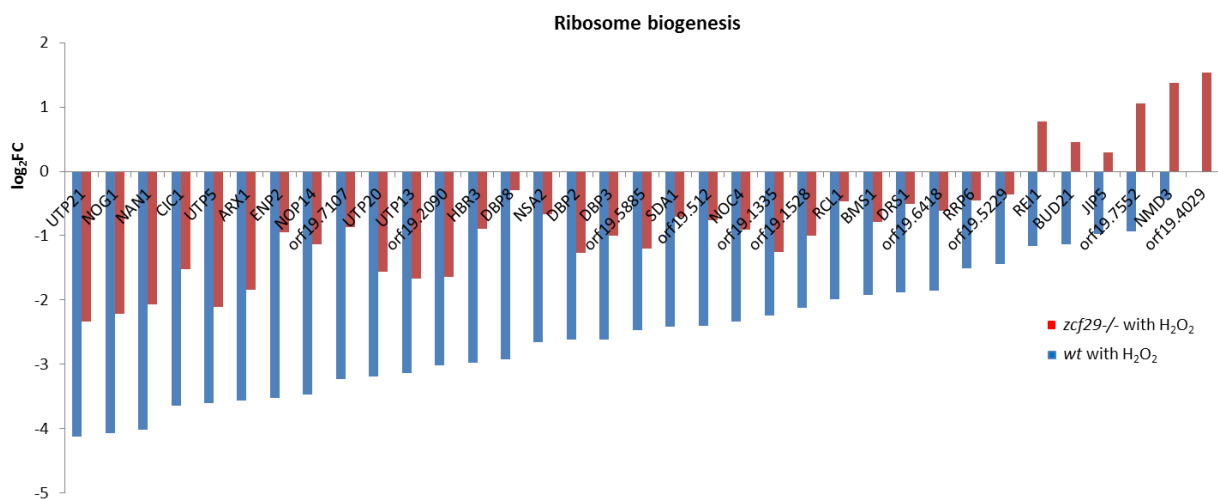
In order to identify genes and pathways regulated by *ZCF29* upon ROS challenge we compared the transcriptional changes caused by H<sub>2</sub>O<sub>2</sub> in wildtype with the once observed in *zcf29/zcf29* 15 minutes after the addition of H<sub>2</sub>O<sub>2</sub>. We found a total of 668 genes statistically significantly differentially expressed (Figure 4.17) and clustered them in four groups. Group 1 and 2 represents genes downregulated by H<sub>2</sub>O<sub>2</sub> in wild type (blue bars) but to a statistically significant lower (group 1) or higher (group 2) degree in *zcf29/zcf29* (red bars). Group 3 and 4 represents genes that are upregulated by H<sub>2</sub>O<sub>2</sub> in wild type but to a statistically significant higher (group 3) or lower (group 4) degree in *zcf29/zcf29*.



**Figure 4.17: Genes differentially expressed upon  $H_2O_2$  exposure between wildtype and *zcf29/zcf29*.**

Group 1 and 2 represents genes that are downregulated by  $H_2O_2$  but to a statistically significant lower (group 1) or higher (group 2) degree in *zcf29/zcf29* compared to wild type. Full list of enriched biological functions in group 1 is reported in Figure 8.1 Group 3 and 4 represents genes that are upregulated by  $H_2O_2$  but to a statistically significant higher (group 3) or lower (group 4) degree in *zcf29/zcf29*. Full list of enriched biological functions in group 4 is reported in Figure 8.2

To identify enriched biological functions in our dataset we ran a GO term analysis (Inglis et al., 2012) on each of the four groups reported in Figure 4.17. This analysis highlighted ribosome biogenesis as one of the most overrepresented biological functions of group 1 (corrected p value  $1.4 \times 10^{-18}$ ) suggesting that wild type downregulation of genes involved in ribosome biogenesis is hampered in *ZCF29* deletion. Figure 4.18 shows 35 genes involved in ribosome biogenesis that are downregulated to a statistically lower degree in *ZCF29* knockout compared to wildtype.



**Figure 4.18: *ZCF29* plays an important role in down regulating ribosome biogenesis upon ROS challenge.**

35 genes downregulated upon  $H_2O_2$  exposure in wildtype (blue) are downregulated to a statistically significant lower extend in *ZCF29* deletion (red).

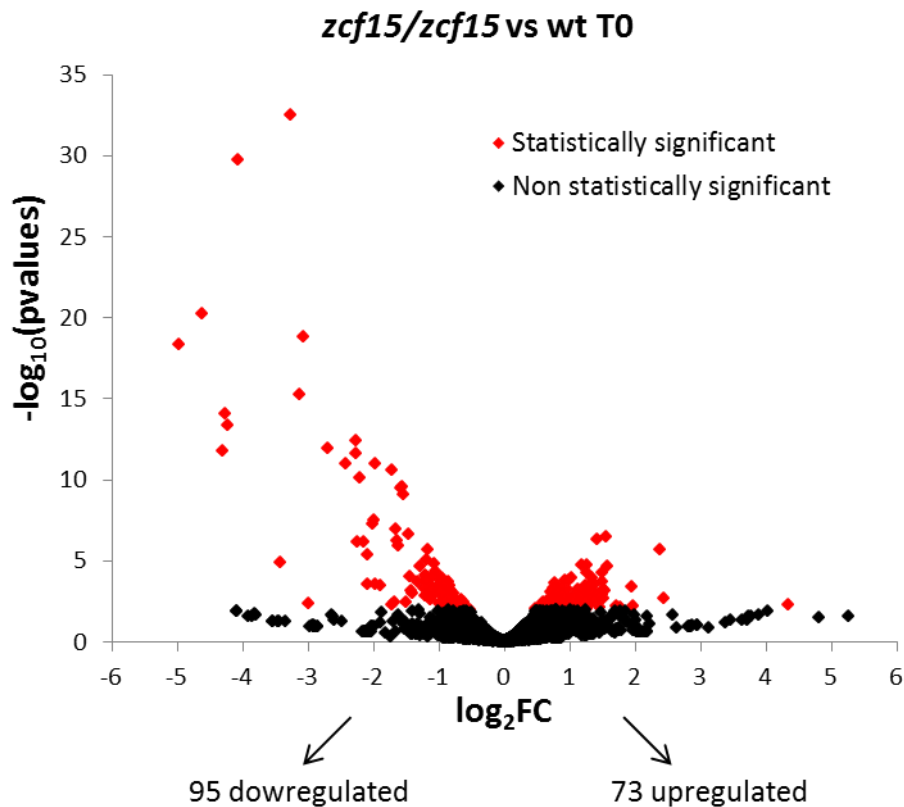
Ribosome assembly in *C. albicans* requires a dynamic series of steps in which protein-protein, RNA-protein, and RNA-RNA interactions needs to be carefully orchestrated. Biogenesis begins in the nucleolus where rRNA is transcribed and subsequently goes through a series of maturation steps including RNA folding, spacer sequence removal, and ribosomal protein binding (Woolford and Baserga, 2013). This maturation continues as preribosomal particles like the 90s Preribosome bind to rRNAs-protein complexes and translocate from the

nucleolus to the nucleoplasm to the cytoplasm ultimately forming the two mature subunits (60s subunit and 40s subunits). Overall the ribosome is an enormous machine composed of four rRNAs and more than 80 ribosomal proteins and yeast cells in log-exponential phase are believed to produce 40 ribosomes per second (Warner, 1999). Ribosome biogenesis is very energetically demanding and under ROS stress, *C. albicans* coordinately downregulate ribosome biogenesis in order to liberate significantly energy resources for other cellular processes (Loar et al., 2004). Our data suggests that the ability to downregulate ribosome biogenesis upon ROS challenge is compromised in *zcf29/zcf29* as many of the genes involved in the various steps of ribosome biogenesis are downregulated to a significant lower degree in *zcf29/zcf29*. For example *UTP21*, *NAN1*, *UTP5*, *UTP20*, *UTP13* five U3 snoRNA proteins involved in pre-rRNA processing, *DBP2* and *DBP3* two DEAD-box RNA helicases involved in rRNA maturation, *RCL1* an endonuclease involved in 90S preribosome processing and *REI1*, a cytoplasmic pre-60S subunit protein are all downregulated to a statistically significantly lower degree in *zcf29/zcf29* compared to wildtype. Many of these genes have been shown to be downregulated upon macrophage generated ROS (Lorenz et al., 2004), and suggesting that these genes are downregulated not only in the presence of ROS *in vitro* but also in the context of a mammalian infection.

Taken together our data suggests that the canonical downregulation of ribosome biogenesis under ROS stress is compromised in *zcf29/zcf29*. The *C. albicans* ability to downregulate this very energy demanding process upon ROS exposure is critical and we believe that the *zcf29/zcf29* inability to do so may be the underlying cause of its hypersusceptibility to ROS.

### 4.3.5 *ZCF15* plays an important role in C-metabolism and regulates key thiol peroxidases upon ROS exposure

We identified genes and pathways regulated by *ZCF15* using a strategy similar to the one described in paragraph 4.3.1.1. We first looked at genes differentially expressed between wild type and *zcf15/zcf15* strains growing in mid exponential phase in the absence of H<sub>2</sub>O<sub>2</sub>. In these conditions, the deletion of *ZCF15* leads to the statistically significant differential expressions of 168 genes, 95 downregulated and 73 upregulated (Figure 4.19).

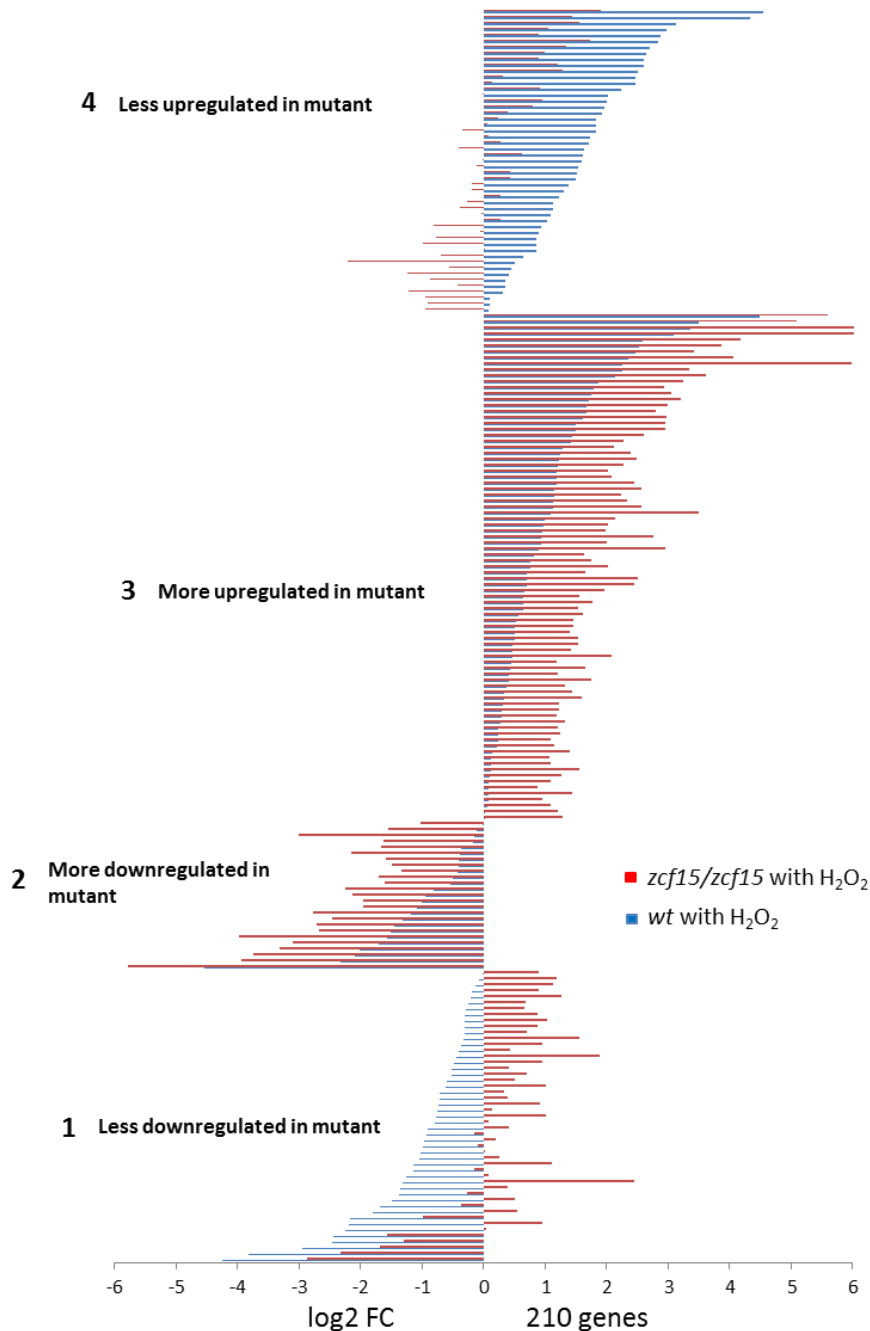


**Figure 4.19: Volcano Plots showing the number of genes differentially expressed between wildtype and *zcf15/zcf15* in the absence of H<sub>2</sub>O<sub>2</sub>.**

When we run a GSEA analysis we found that carbon metabolism, TCA cycle and energy production are all downregulated in *zcf15/zcf15* compared to wild type (p values between 0.00123 and 0.00410). In particular glucose transporters like *SHA3*, *HGT7* and *HGT6* are

downregulated between 16 and 3 folds, genes involved in glycolysis/TCA cycle like glycolysis regulator *TYE7*, phosphoenolpyruvate carboxykinase (*PCK1*) and pyruvate carboxylase (*PYC2*) between 8 and 2 folds and genes involved in cellular respiration like NADH dehydrogenase (*NDE1*) and malate dehydrogenase (*MAE1*) between 4 and 2 folds. This data suggests that *ZCF15* plays an important role in regulating C-metabolism and energy production an hypothesis that was confirmed when we compared the transcriptional changes triggered by H<sub>2</sub>O<sub>2</sub> in wildtype versus *zcf15/zcf15*.

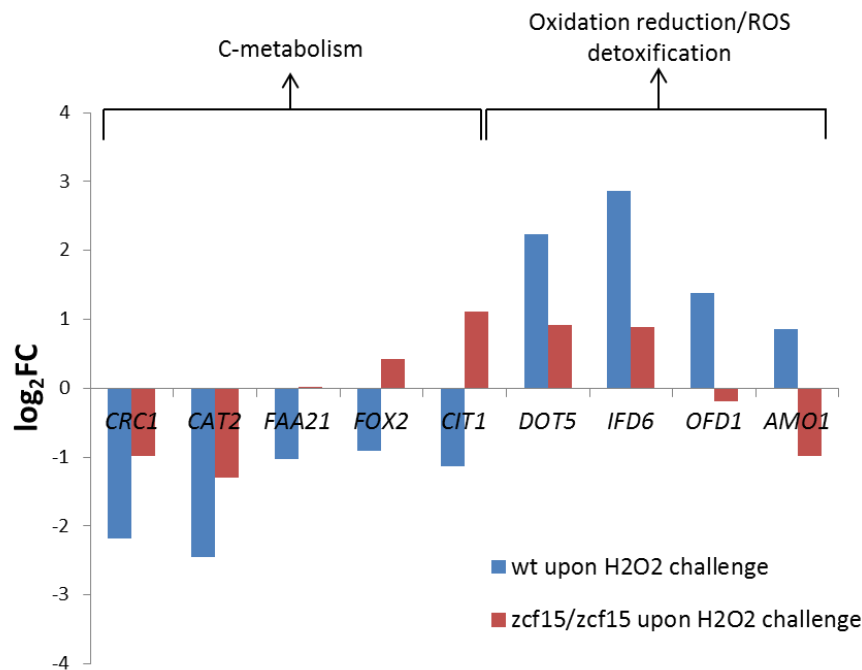
As shown in Figure 4.20:, when we run the EdgeR generalized linear model as described in 4.3.3.2 we found 210 genes statistically significantly differentially expressed. We clustered these genes in four groups. Group 1 and 2 represents H<sub>2</sub>O<sub>2</sub> repressed genes downregulated to a statistically significant lower (group 1) or higher (group 2) degree in *zcf15/zcf15* compared to wild type. Group 3 and 4 represents H<sub>2</sub>O<sub>2</sub> induced genes upregulated to a statistically significant higher or lower degree in the *zcf15/zcf15*. When we looked for biological functions enriched in group 1 and 4 using GO term analysis (Inglis et al., 2012) we found that carbon utilization and oxidation-reduction processes were the two most highly statistically significant enriched processes (carbon utilization  $p = 2.5 \times 10^{-5}$ , oxidation reduction process  $p = 1 \times 10^{-4}$ ). These two biological processes are highly interconnected as it has been shown that in the presence of ROS *C. albicans* goes in a “starvation like” mode by downregulating genes involved in carbon metabolism and simultaneously upregulating genes involved in oxidation reduction processes and ROS detoxification (Lorenz et al., 2004).



**Figure 4.20: genes differentially expressed between wild type and *zcf15/zcf15* upon H<sub>2</sub>O<sub>2</sub> challenge.**

Group 1 represents genes downregulated by H<sub>2</sub>O<sub>2</sub> in wildtype but to a statistically significant lower degree in *zcf15/zcf15*. Full list of enriched biological functions in group 1 is reported in Figure 8.3. Group 2 represents genes downregulated in wildtype but to a statistically significant higher degree in *zcf15/zcf15*. Using the same logic group 3 and 4 represents H<sub>2</sub>O<sub>2</sub> induced genes but to a significant higher (group 3) or lower degree (group 4) in *zcf15/zcf15*. Full list of enriched biological functions in group 4 is reported in Figure 8.4.

As shown in Figure 4.21 many of the genes involved in carbon metabolism like *CRC1*, *CAT2*, *FAA21*, *FOX2* and *CIT1* are downregulated to a statistically significant lower degree in the mutant compared to wild type. Concurrently genes involved in oxidation reduction processes/ROS detoxification like *DOT5* (a thiol peroxidase), *IFD6* (a aldo-keto reductase), *OFD1* (prolyl hydroxylase) and *AMO1* (a peroxisomal oxidase) are upregulated in *zcf15/zcf15* to a statistically significantly lower extent.



**Figure 4.21: ZCF15 is required for down regulating genes involved in C-metabolism and simultaneously upregulating genes involved in oxidation reduction processes/ROS detoxification.**

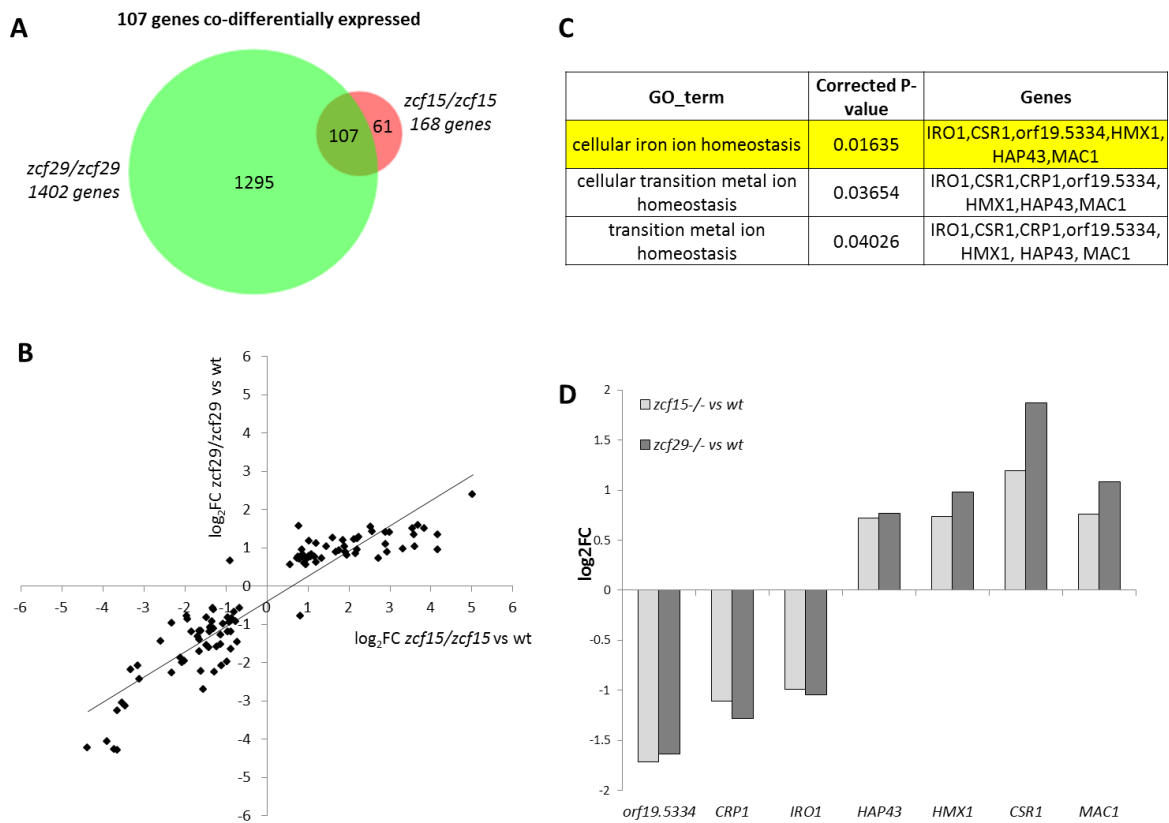
This evidence suggests that the canonical wildtype ability to simultaneously downregulate genes involved in carbon metabolism and upregulate genes involved in ROS detoxification upon H<sub>2</sub>O<sub>2</sub> exposure is compromised in *zcf15/zcf15*. We believe that *zcf15/zcf15* inability to downregulate carbon metabolism and rechannel energy resources towards oxidation reduction processes/ROS detoxification is the underlying cause of its increased susceptibility to ROS.



#### **4.3.6 Genes involved in iron homeostasis are potentially coregulated by *ZCF15* and *ZCF29***

As discussed in the previous chapters, in the absence of H<sub>2</sub>O<sub>2</sub>, *ZCF29* deletion leads to the differential expression of 1402 genes while the deletion of *ZCF15* leads to a more modest transcriptional rewiring of 168 genes. In order to find biological functions potentially coregulated by both transcription factors we identified genes differentially expressed in both mutants. As shown in Figure 4.22 A, we found 1295 genes differentially expressed only in *zcf29/zcf29*, 61 genes differentially expressed only in *zcf15/zcf15* and 107 genes differentially expressed in both mutants. Remarkably the expression directionality of these genes is highly conserved in both deletions with only 2 out of the 107 genes found differentially expressed in both mutants but in opposite directions (Figure 4.22 B). This evidence suggests that the remaining 105 genes are not only co-occurring in the lists of differentially expressed genes but that they are likely regulating common biological functions.

When we looked for enriched biological functions in these 105 genes we found that cellular iron homeostasis was the most statistically significant enriched functions (Figure 4.22 C). In particular we found that genes involved in iron utilization like *IROI*, *CRP1* and *orf19.5334* (Garcia et al., 2001; Weissman et al., 2000) are downregulated in both mutants and contemporarily genes involved in iron uptake like *HAP43*, *HMX1*, *CSRI* and *MAC1* (Kim et al., 2008; Lan et al., 2004) are upregulated. This transcriptional response resembles the transcriptional response observed by Lan et al., 2004 in wildtype *C. albicans* grown under low iron conditions. When iron is limiting *C. albicans* upregulates genes involved in iron uptake and downregulates genes involved in iron utilization.

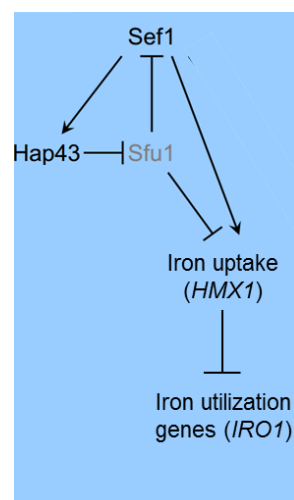


**Figure 4.22: genes involved in iron homeostasis are differentially expressed in both *ZCF15* and *ZCF29* deletions.**

- A) Deletion of *ZCF15* and *ZCF29* leads to the differential expression of a common set of 107 genes.
- B) The directionality of the potentially coregulated genes is respected in both mutants.  $\log_2FC$  of genes differentially expressed upon *ZCF15* deletion is plotted against  $\log_2FC$  of the same genes upon *ZCF29* deletion. 105 of the 107 genes are co-differentially expressed in the same direction and only 2 genes are differentially expressed in both deletion but in opposite directions.
- C) Cellular iron homeostasis is the most statistically significantly biological function in the 107 genes differentially expressed in both *zcf15/zcf15* and *zcf29/zcf29*.
- D) Deletion of *ZCF15* and *ZCF29* leads to the upregulation of *HAP43*, *HMX1*, *CSR1*, *MAC1* and simultaneous downregulation of *orf19.5334*, *CRP1* and *IRO1*. Upregulation of *HAP43*, *HMX1* and simultaneous downregulation of *IRO1* is commonly observed in wildtype *C. albicans* under nitrogen limiting conditions.

In *C. albicans* iron uptake is tightly regulated. Too little iron can cause core metabolic processes such as TCA cycle to be shut down and too much iron can cause the formation of toxic oxygen radicals (Blankenship and Mitchell, 2011). In addition, iron homeostasis plays a

major role in virulence since *C. albicans* needs to quickly adapt and thrive in both the iron replete environment of the mammalian gut as well as in the iron limited environment of the blood stream. As shown in Figure 4.23, genes involved in iron homeostasis are tightly regulated by three transcription factors: Sfu1, Sef1 and Hap43 (Chen et al., 2011). Sfu1 is major repressor of iron uptake and its expression is repressed by regulated by Hap43. In other words, Hap43 activates iron uptake by repressing the repressor. (Blankenship and Mitchell, 2011).



**Figure 4.23: Iron homeostasis is tightly regulated by three major transcription factors, Sef1, Sfu1 and Hap43.** Figure adapted from Blankenship and Mitchell, 2011

In our dataset the expression of Hap43 is statistically significantly upregulated in both *zcf15/zcf15* and *zcf29/zcf29*. Our working hypothesis is that upon deletion of *ZCF15* and *ZCF29*, *HAP43* upregulation represses Sfu1 leading to upregulation of iron uptake genes like *HMX1* and downregulation of genes involved in iron utilization like *IRO1* (Figure 4.23).

The fact that *ZCF15* and *ZCF29* deletion resembles the transcriptional response observed in wildtype under low iron conditions suggests that these two transcription factors may play a role in the regulating cellular iron uptake and utilization.

#### 4.3.7 ZCF15 and ZCF29 ChIP-Seq

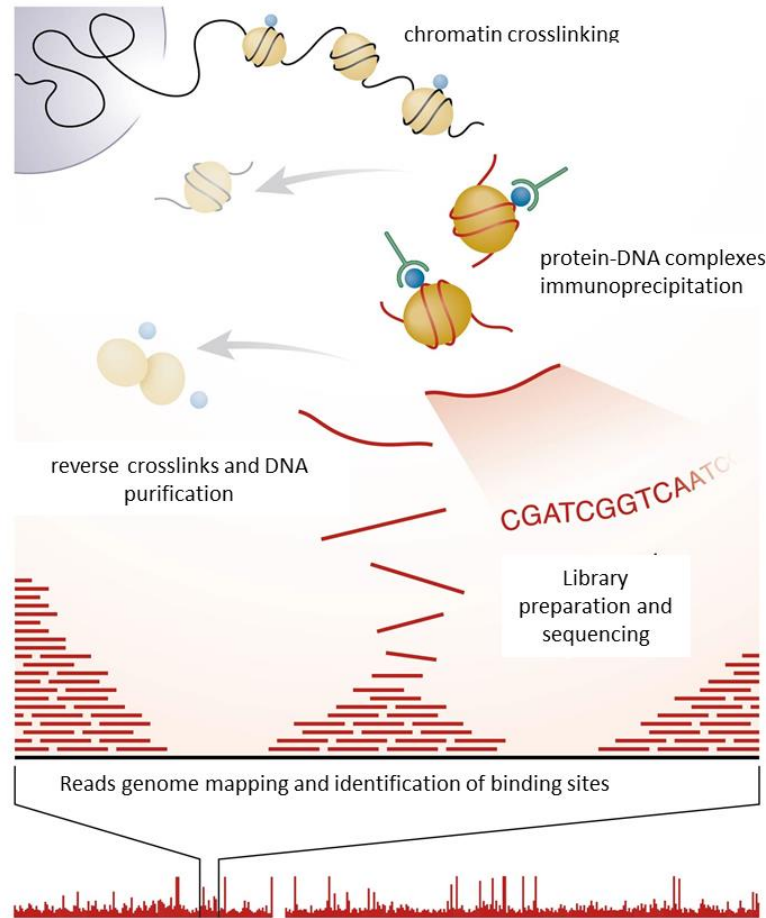
The RNA-Seq experiment described in the previous paragraphs gave us some insights on the genes regulated by *ZCF15* and *ZCF29*. However, genes differentially expressed in transcription factor (TF) knockouts can be either directly or indirectly regulated by the TF. In the former case, the TF binds to specific region of the genome directly regulating the expression of genes flanking that region. In the latter case, the regulation is indirect because, for example, the TF can either be a distant-acting regulator or can control the expression of other regulators (silencer/enhancer) which control the expression of the differentially expressed gene.

In order to identify *Zcf15* and *Zcf29* transcriptional targets *in vivo*, we used a **chromatin immunoprecipitation** approach coupled with DNA **Sequencing** (ChIP-Seq). This method allowed us to explore protein-DNA interaction of *Zcf15* and *Zcf29* with their respective transcriptional binding sites. This approach allowed us to retroactively look back at our transcriptomic experiments and identify genes directly controlled by *Zcf15* and *Zcf29*. Combining RNA-Seq and ChIP-Seq data have been shown to be a powerful way to explore novel transcriptional mechanisms (Hull et al., 2013).

The ChIP-Seq experimental approach we used relies on five steps (Figure 4.24):

- Crosslink DNA binding proteins and their transcriptional targets with formaldehyde.
- Selectively immunoprecipitate protein-DNA complexes.
- Reverse formaldehyde cross-links, hydrolyze residual proteins and purify DNA.
- Prepare DNA multiplexed libraries and deep sequence them.
- Map reads to the genome and identifies transcriptional binding sites.

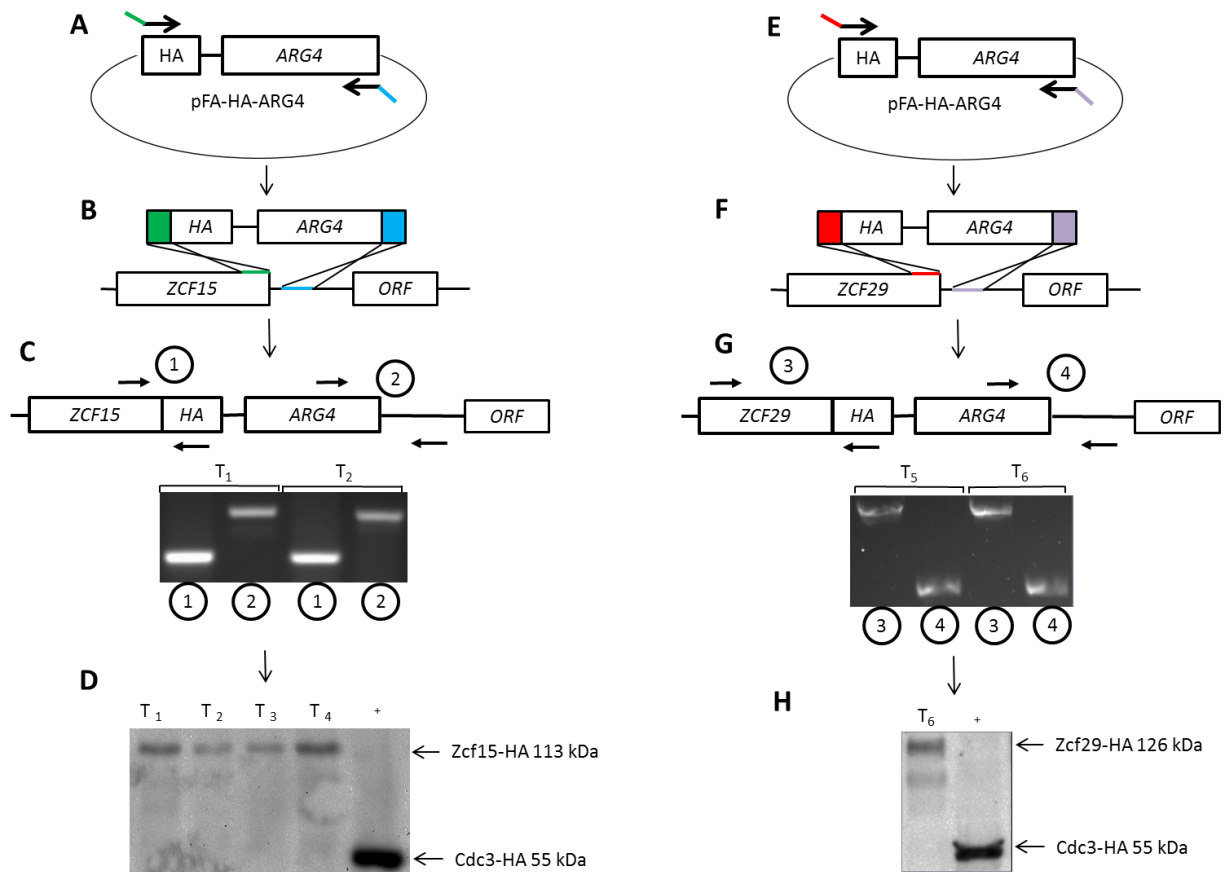
These five steps are summarized in Figure 4.24 and are thoroughly reviewed in Botcheva et al., 2011.



**Figure 4.24: An overview of the ChIP-Seq procedure used**

In our ChIP-Seq experiment we first cross-linked DNA binding proteins to their transcriptional targets. We subsequently immunoprecipitate our protein-DNA complexes of interest, reverse the crosslink and purified DNA. The purified DNA was used for library preparation and sequencing. Reads obtained from the sequencing were mapped to the *C. albicans* genome and binding sites identified using the algorithm MACS. Figure adapted from (Botcheva et al., 2011)

In order to selectively immunoprecipitate Zcf15 and Zcf29 protein-DNA complexes we first epitope tagged Zcf15 and Zcf29 at their C-termini with an HA-containing epitope. A cassette containing 3-tandem copies of the HA-epitope was integrated at the 3' end of each gene using a PCR-mediated homologous recombination strategy (Figure 4.25).



**Figure 4.25: ZCF15 and ZCF29 HA epitope tagging**

- A) *HA-ARG4* insertion cassette is amplified from pFA-*HA-ARG4* plasmid using primers containing ~70bp sequence homology for the region upstream and downstream of *ZCF15* stop codon.
- B) Wildtype *C. albicans* is transformed with the *HA-ARG4* insertion cassette that integrates by homologous recombination at the 3' end of *ZCF15* and consequently HA tagging Zcf15 at its C-terminus.
- C) Proper insertion of the *HA-ARG4* insertion cassette was confirmed by diagnostic PCR targeting the new 5' and 3' junctions. Two independent transformants (T<sub>1</sub> and T<sub>2</sub>) were confirmed.
- D) Proper expression of the Zcf15-HA tagged protein was confirmed by Western Blot. Four independent transformants (T<sub>1</sub>, T<sub>2</sub>, T<sub>3</sub> and T<sub>4</sub>) were tested and Cdc3-HA described here (Gerami-Nejad et al., 2009) was used as a positive control.
- E-F-G-H) Zcf29p was tagged using different primers but essentially the same logic described in A-B-C-D for *ZCF15*.

In order to epitope tag Zcf15 and Zcf29 we first amplified the integrating cassette with primers containing overhanging sequences of homology for the regions flanking their respective stop codons (Figure 4.25 green-blue for *ZCF15* and red-purple for *ZCF29*). Upon

transformation, proper genomic integration was confirmed by diagnostic PCR (Figure 4.25 C and G) and proper expression of the tagged protein by Western Blot (Figure 4.25 D and H). We decided to epitope tag our transcription factors at their 3' and not 5' end so that their expression remained under the control of their own regulatory sequences.

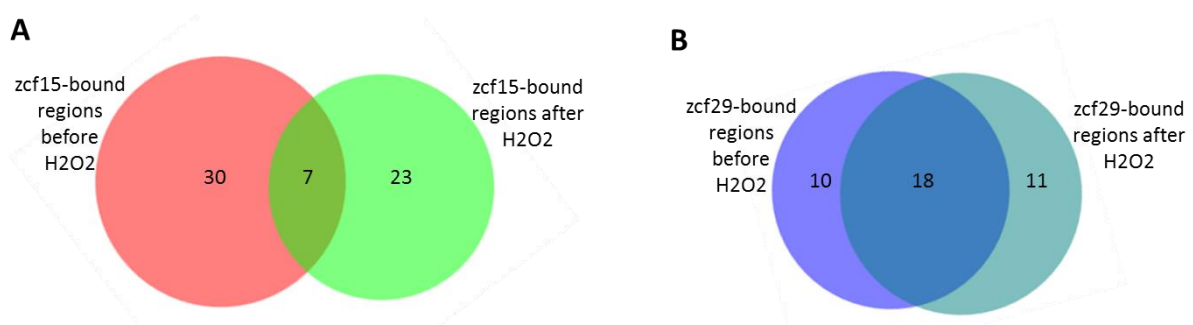
In order to compare the ChIP-Seq results with the transcriptomics studies described in 4.3.3, we ran our experiments in analogous conditions. We identified Zcf15 and Zcf29 DNA binding partners 5 and 15 minutes after the addition of 5mM H<sub>2</sub>O<sub>2</sub> and compared them to unexposed samples. Upon H<sub>2</sub>O<sub>2</sub> exposure samples were quickly cross-linked with formaldehyde, TF-bound DNA purified and DNA libraries prepared and sequenced.

Our sequencing runs yielded a total of 148 million 25-base sequence reads with an average of 4.2 million reads per sample. Reads were successfully aligned to the *C. albicans* genome and we obtained a mean coverage between 50x and 75x. Zcf15 and Zcf29 bound regions were obtained using MACS peak caller (Zhang et al., 2008) in the web-based genomic platform Galaxy (Goecks et al., 2010). Peaks were called by comparing reads between anti-HA pulldown samples and input controls.

In the absence of H<sub>2</sub>O<sub>2</sub>, we identified 37 and 28 regions bound by Zcf15 and Zcf29 respectively. In the presence of H<sub>2</sub>O<sub>2</sub>, we identified 30 and 29 regions bound by Zcf15 and Zcf29 respectively. These genomic regions were identified using MACS and only enriched regions with p values < 1e-3 were considered. The average length of these TF bound regions is 3.9 kb and these regions are evenly distributed across the genome in all samples suggesting no biases toward specific chromosomes or genomic locations.

In order to evaluate if the addition of H<sub>2</sub>O<sub>2</sub> drove our TF to different genomic loci, we identified TF-bound regions that remain bound both before and after the addition of H<sub>2</sub>O<sub>2</sub>. As

shown in Figure 4.26, only 7 and 18 of the Zcf15 and Zcf29 bound regions remained bound both before and after the addition of H<sub>2</sub>O<sub>2</sub>. This evidence suggest that, upon H<sub>2</sub>O<sub>2</sub> exposure, Zcf15 and Zcf29 switch many of their genomic targets and relocate themselves to other genomic loci likely controlling different genes sets. This evidence confirms and validates our RNA-Seq experiments that showed that Zcf15 and Zcf29 control different biological functions in the presence and absence of H<sub>2</sub>O<sub>2</sub> (paragraph 4.3).



**Figure 4.26 Zcf15 and Zcf29 change many of their genomic targets upon H<sub>2</sub>O<sub>2</sub> exposure.** Zcf15 has 30 and 23 unique genomic binding sites in the absence and presence of H<sub>2</sub>O<sub>2</sub> respectively and only 7 that are present in both conditions. Zcf29 has 10 and 11 unique genomic binding sites and 18 that are present in both conditions. Together these evidences suggest that Zcf15 and Zcf29 relocate to different genomic loci upon H<sub>2</sub>O<sub>2</sub> exposure confirming our transcriptome analysis that showed regulations of different sets of genes in the presence or absence of H<sub>2</sub>O<sub>2</sub>. Genomic regions between conditions were considered in common if the location of their summits were within 2 kb of each other.

In order to validate these ChIP-Seq results and define Zcf15 and Zcf29 regulons we compared the ChIP-Seq data with our transcriptome analysis. Overall, we found various examples in which the genes differentially expressed in our RNA-Seq experiment were flanked by TF-bound regions. For example, in the absence of H<sub>2</sub>O<sub>2</sub> 14 of the 28 *ZCF29* DNA binding sites are flanking genes differentially expressed upon *ZCF29* deletion (Figure 4.27 A genes highlighted in yellow). In addition, 7 of the 29 DNA binding sites that *ZCF29* recognizes in the presence of H<sub>2</sub>O<sub>2</sub> are flanking genes that we considered H<sub>2</sub>O<sub>2</sub> responsive and *ZCF29* dependent (Figure 4.27 B, genes highlighted in yellow).



**A**

No H <sub>2</sub> O <sub>2</sub>			
ChIPSeq data		RNASeq data	
Peak	Gene flanknig	log2FC	pvalue
1	<b>PDR16</b>	0.572321994	0.032777
2	<b>orf19.4476</b>	3.502381126	4.43E-11
3	<b>orf19.1048</b>	3.563049102	2.98E-21
4	<b>FGR6-3</b>	1.070965553	0.000289
5	<b>ZWF1</b>	1.138607373	8.32E-07
6	DAK2	0.149342523	0.57079
7	<b>orf19.4780</b>	0.995209691	5.91E-06
8	<b>HSP70</b>	2.08912282	1.93E-15
9	<b>RTA3</b>	0.530528766	0.073655
10	<b>orf19.344</b>	#N/A	#N/A
11	<b>CDR1</b>	0.778123933	0.001163
12	<b>AHR1</b>	0.668113614	0.009231
13	<b>orf19.4653</b>	#N/A	#N/A
14	<b>FGR6-10</b>	1.027286073	0.0002
15	<b>orf19.1240</b>	0.633055229	0.005431
16	SFL2	-0.0674016	0.837488
17	ADH1	0.470792008	0.039418
18	<b>MDR1</b>	0.515868401	0.274047
19	<b>orf19.5775</b>	#N/A	#N/A
20	<b>orf19.7042</b>	-2.40294088	9.9E-05
21	<b>orf19.6586</b>	-0.89305706	0.028421
22	<b>orf19.6585</b>	#N/A	#N/A
23	PFK2	-0.57415706	0.016528
24	<b>FGR6-1</b>	0.903703155	0.003431
25	<b>orf19.5131</b>	0.667876512	0.004188
26	<b>orf19.7204</b>	1.164179687	0.000797
27	CRZ2	0.384495024	0.184462
28	<b>orf19.7307</b>	-0.04266887	0.85761

**B**

H <sub>2</sub> O <sub>2</sub>			
ChIPSeq data		RNASeq data	
Peak	Gene flanknig	log2FC	pvalue
1	<b>PDR16</b>	0.284805863	0.43427281
2	<b>orf19.1048</b>	2.682734278	3.3258E-08
3	<b>FGR6-3</b>	1.60247454	0.0016576
4	<b>orf19.4780</b>	0.867537437	0.00767006
5	GND1	-0.15370106	0.57969487
6	orf19.1450	#N/A	#N/A
7	orf19.850	0.630040101	0.19863784
8	<b>RTA3</b>	-0.03586489	0.93016632
9	orf19.1613	0.328441464	0.50273796
10	<b>orf19.251</b>	1.944483355	8.0526E-07
11	<b>orf19.344</b>	#N/A	#N/A
12	<b>CDR1</b>	-0.02516934	0.94094124
13	<b>orf19.4653</b>	#N/A	#N/A
14	HIT1	#N/A	#N/A
15	<b>FGR6-10</b>	1.248816941	0.00542261
16	orf19.2892	0.315750494	0.34647666
17	orf19.979	-0.04087522	0.89473581
18	PFK1	0.771848299	0.03222483
19	<b>SFL2</b>	0.836916402	0.0779138
20	<b>MDR1</b>	-1.79149416	0.0017081
21	<b>orf19.5775</b>	#N/A	#N/A
22	<b>orf19.7042</b>	-2.16038807	0.00154588
23	<b>orf19.6586</b>	-1.18052063	0.01525928
24	<b>orf19.6585</b>	#N/A	#N/A
25	orf19.6522	#N/A	#N/A
26	<b>FGR6-1</b>	0.756180119	0.14120716
27	<b>orf19.7204</b>	0.589380187	0.20336137
28	orf19.727	0.577207718	0.24552403
29	<b>orf19.7307</b>	#N/A	#N/A

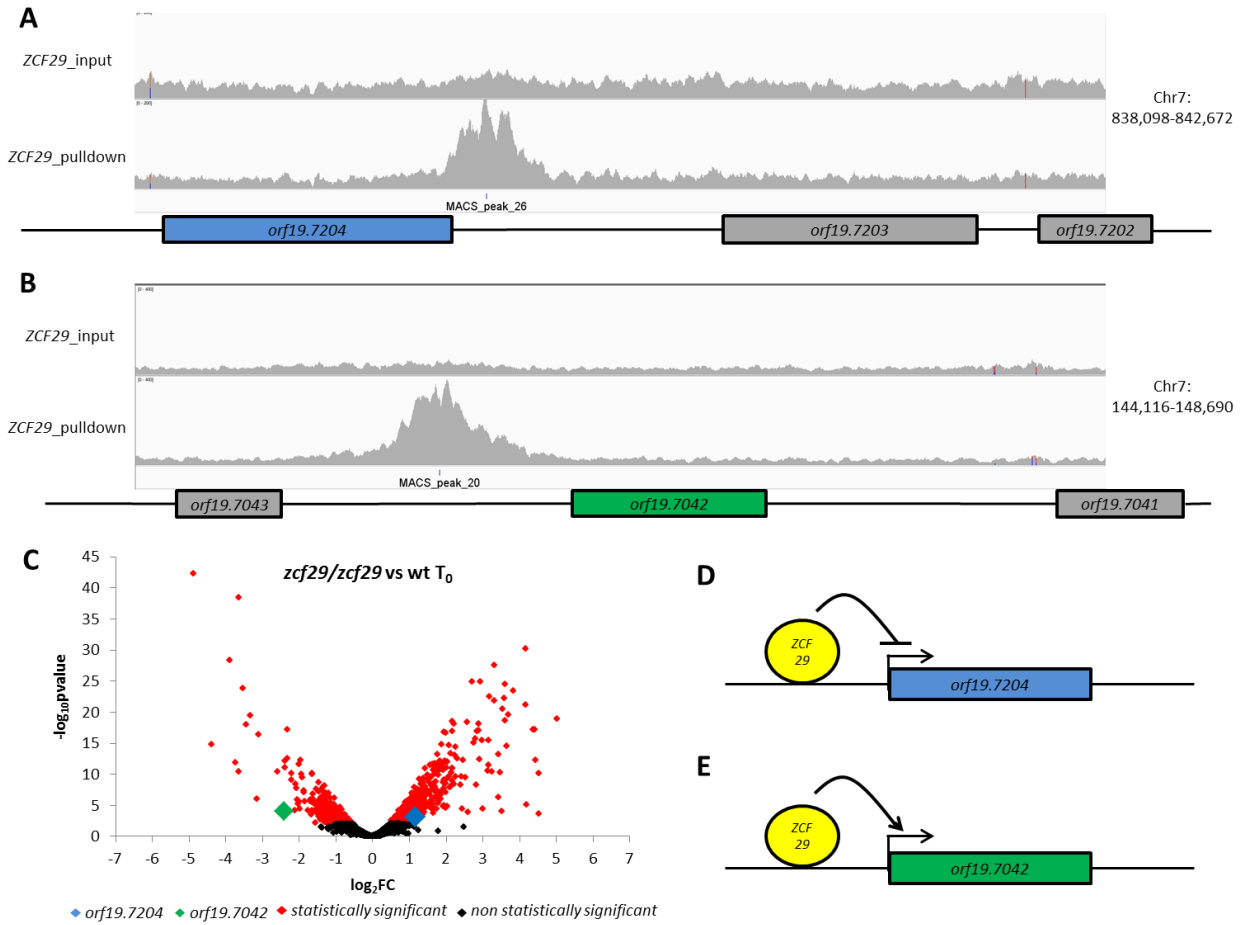
**Figure 4.27: various *ZCF29* DNA binding sites flank genes differentially expressed in *ZCF29* deletion.**

- A) In the absence of H<sub>2</sub>O<sub>2</sub>, 14 of the 28 *ZCF29* genomic binding sites (ChIP-Seq data) flank genes that are differentially expressed upon *ZCF29* deletion (RNA-Seq data). Genes reported in bold indicate *ZCF29* genomic binding partners bound both in the presence and absence of H<sub>2</sub>O<sub>2</sub> (18 total, shown also in Figure 4.26 B)
- B) In the presence of H<sub>2</sub>O<sub>2</sub>, 7 of the 29 *ZCF29* DNA binding sites (ChIP-Seq data) flank genes that are shown to be H<sub>2</sub>O<sub>2</sub> responsive and *ZCF29* dependent.

Figure 4.28 depicts two genes that are likely directly regulated by *Zcf29* in the absence of H<sub>2</sub>O<sub>2</sub>. The ChIP-Seq experiment showed a peak (3.5 folds enriched,  $p < 0.01$ ) right upstream of *orf19.7204* (Figure 4.28 A) and the same gene was shown to be statistically significantly

upregulated in *zcf29/zcf29* (Figure 4.28 C blue dot). Together these results suggest that *orf19.7204* is a direct Zcf29 target and that Zcf29 functions as a repressor of this gene (Figure 4.28 D). On the other hand Zcf29 is likely a direct activator of *orf19.7042* since we found a Zcf29 bound region just upstream of this gene in wild type *C. albicans* (Figure 4.28 B) and contemporarily showed that this genes is statistically significantly downregulated in *zcf29/zcf29* (Figure 4.28 C green dot).

*orf19.7204* is an uncharacterized gene with orthologs that have glyoxysome localization. Glyoxysome are specialized peroxisomes in which fatty acids are hydrolyzed to acetyl-Coa leading to the production of ROS (Michels et al., 2005). Various antioxidant proteins are expressed in this organelle to detoxify ROS and it is possible that *orf19.7204* is one of them. *orf19.7042* is also an uncharacterized gene but it has been shown to be induced in *MDR1* overexpressing strains (Karababa et al., 2004). *MDR1* is a multidrug efflux pump that is commonly overexpressed in drug resistant clinical isolates (White, 1997a) and is directly involved in cellular detoxification processes (Morschhauser et al., 2007). We believe that *orf19.7042* is activated by Zcf29 and that it cooperates in cellular detoxification process, an hypothesis that is corroborated by the fact that this gene was found to be upregulated upon nitric oxide exposure (Hromatka et al., 2005).



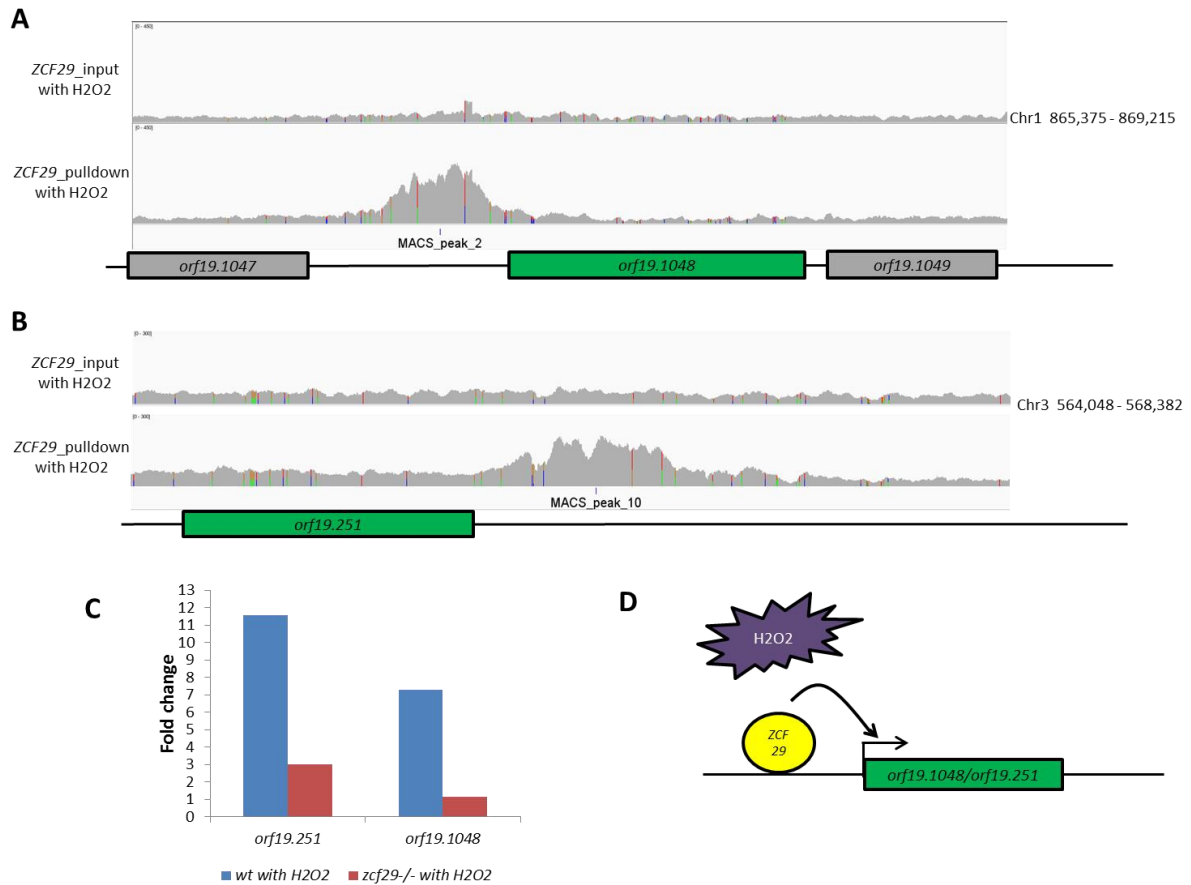
**Figure 4.28: Zcf29 directly controls the expressions of orf19.7204 and orf19.7042.**

- A) A 4.5 kb span of *C. albicans* genome displayed in IGV (chromosome 7 location coordinates 838,098-842,672). The two tracks labeled Zcf29\_pulldown and Zcf29\_input represent reads from pulldowns samples and control respectively. A large statistically significant peak is present right upstream of orf19.7204
- B) A 4.5 kb span of *C. albicans* genome displayed in IGV (chromosome 7 location coordinates 144,116-148,690) showing a large Zcf29 binding region right upstream of orf19.7042
- C) RNASeq:Volcano Plots showing that orf19.7042 and orf19.7204 are statistically significantly up and downregulated upon ZCF29 deletion.
- D) orf19.7204 is a direct Zcf29 targets and its expression is likely repressed by Zcf29.
- E) orf19.7042 is a direct Zcf29 targets and its expression is likely activated by Zcf29.

In addition to genes directly regulated by Zcf29 in the absence of H<sub>2</sub>O<sub>2</sub>, the ChIP-Seq experiment allowed us to determine H<sub>2</sub>O<sub>2</sub> responsive genes whose expression is directly controlled by Zcf29.

For example, as shown in Figure 4.29A and B upon H<sub>2</sub>O<sub>2</sub> exposure Zcf29 binds to the regions directly upstream of orf19.1048 and orf19.251 (Figure 4.29 A and B). The presence of these two Zcf29 bound regions upon H<sub>2</sub>O<sub>2</sub> exposure validates our RNA-Seq experiments which showed that Zcf29 is required for the H<sub>2</sub>O<sub>2</sub> dependent upregulation of these two genes. While in wild type *C. albicans* these two genes are upregulated 11 and 7 folds upon H<sub>2</sub>O<sub>2</sub> exposure, the same genes were only 3 and one fold upregulated in *zcf29/zcf29* (Figure 4.29D). Together these results suggest that upon H<sub>2</sub>O<sub>2</sub> exposure Zcf29 binds the regions right upstream of these two genes directly upregulating their expressions.

orf19.251 is a glutathione-independent glyoxalase that has a critical role in stress response and in glycoalyx detoxification (Hasim et al., 2014), while orf19.1048 is an alpha-keto reductase that is induced in strains overexpressing the cellular detoxification efflux pump MDR1 (Rogers and Barker, 2003). To the best of our knowledge this is the first time that the expression of these two H<sub>2</sub>O<sub>2</sub> responsive genes have been shown to be directly controlled by Zcf29.



**Figure 4.29: The upregulation of orf19.251 and orf19.1048 upon H<sub>2</sub>O<sub>2</sub> is directly controlled by Zcf29.**

- A) A 3.8 kb span of *C. albicans* genome displayed in IGV (chromosome 1 location coordinates 865,375 - 869,215). The two tracks labeled Zcf29\_pulldown and Zcf29\_input represent reads from pulldowns samples and controls obtained after the addition of H<sub>2</sub>O<sub>2</sub>. A large statistically significant peak is present right upstream of orf19.1048 (3.5 fold enriched p < 0.01).
- B) A 4.3 kb span of *C. albicans* genome displayed in IGV (chromosome 3 location coordinates 564,048-568,382) showing a large peak (2.7 fold enriched p < 0.01) right upstream of orf19.251.
- C) RNASeq experiment showing that the upregulation upon H<sub>2</sub>O<sub>2</sub> is hampered in *ZCF29* deletion.
- D) Zcf29 is likely a direct activator of orf19.1048 and orf19.251 upon H<sub>2</sub>O<sub>2</sub> exposure.

Overall, we were able to find many examples of Zcf29 bound regions flanked by differentially expressed genes. On the other hand, we found a significantly smaller correlation when we repeated the analysis for Zcf15.

**A**

No H <sub>2</sub> O <sub>2</sub>			
ChIPSeq data		RNASeq data	
Peak	Gene flanknig	log2FC	pvalue
1	orf19.3671	#N/A	#N/A
2	RPL6	0.11619983	0.658055
3	ACS2	-0.2693238	0.274418
4	SSA2	-0.2233416	0.532388
5	orf19.2478.2	-0.031692	0.907999
6	orf19.6220.5	0.2461191	0.405734
<b>7</b>	<b>ENO1</b>	<b>0.34969299</b>	<b>0.128081</b>
<b>8</b>	<b>ZWF1</b>	<b>0.0190911</b>	<b>0.934003</b>
9	RPL2	0.09990207	0.746237
<b>10</b>	<b>TYE7</b>	<b>-3.0537689</b>	<b>1.62E-19</b>
11	EFT2	-0.0959304	0.664726
<b>12</b>	<b>orf19.871</b>	<b>-0.1851423</b>	<b>0.458167</b>
13	CDC19	0.37092821	0.130538
14	orf19.6882.2	0.20843186	0.456201
<b>15</b>	<b>FRP3</b>	<b>-0.5358984</b>	<b>0.046123</b>
16	TEF1	0.02995589	0.90425
17	YWP1	-0.2431512	0.318956
18	GCN4	-0.0003293	0.998731
19	RPP2B	-0.0739404	0.822604
20	RPL40B	-0.0226974	0.949103
21	orf19.4628	-0.3125137	0.313813
22	orf19.2724	0.57456844	0.088873
23	SPO75	-0.034984	0.895948
24	orf19.2892	-0.0270237	0.910712
<b>25</b>	<b>PDC11</b>	<b>0.1467881</b>	<b>0.577585</b>
26	RPT5	0.14906478	0.488462
27	CEF3	-0.0947932	0.710738
28	SAH1	0.01509147	0.938865
<b>29</b>	<b>ADH1</b>	<b>0.24192421</b>	<b>0.289152</b>
30	ANB1	0.05152644	0.818749
31	RPP0	0.03387182	0.918
32	LAT1	-0.1350025	0.514925
33	orf19.6522	#N/A	#N/A
<b>34</b>	<b>orf19.2529.2</b>	<b>-2.4189076</b>	<b>1.16E-11</b>
35	MET6	0.03086155	0.882218
<b>36</b>	<b>orf19.3874</b>	<b>0.70241846</b>	<b>0.00379</b>
37	MIS11	-0.1419725	0.518651

**B**

H <sub>2</sub> O <sub>2</sub>			
ChIPSeq data		RNASeq data	
Peak	Gene flanknig	log2FC	pvalue
1	RPL10	-0.11038	0.738314
2	PRN3	-0.30217	0.599754
3	CAT1	-0.559	0.116387
4	GCS1	0.238942	0.511522
5	orf19.5066	#N/A	#N/A
<b>6</b>	<b>ENO1</b>	<b>0.491536</b>	<b>0.130186</b>
7	TPO3	-0.64324	0.112658
<b>8</b>	<b>ZWF1</b>	<b>-0.29278</b>	<b>0.363926</b>
9	orf19.4758	-0.92208	0.029086
10	orf19.4908	#N/A	#N/A
11	orf19.4922	-0.47707	0.223794
12	GND1	-0.16027	0.56382
<b>13</b>	<b>orf19.871</b>	<b>-0.46292</b>	<b>0.165536</b>
<b>14</b>	<b>FRP3</b>	<b>-1.49178</b>	<b>0.002524</b>
15	orf19.2165	-0.23016	0.566803
<b>16</b>	<b>UCF1</b>	<b>-3.44628</b>	<b>6.59E-07</b>
17	CCP1	-0.37899	0.251816
<b>18</b>	<b>NDE1</b>	<b>-1.05969</b>	<b>0.000482</b>
19	FBA1	0.564545	0.103612
<b>20</b>	<b>PDC11</b>	<b>0.651891</b>	<b>0.08057</b>
21	GLR1	-0.31952	0.318941
22	TRR1	-0.34476	0.247071
<b>23</b>	<b>ADH1</b>	<b>0.376283</b>	<b>0.243473</b>
24	EBP1	-0.28773	0.441275
25	OYE2	-0.94438	0.035044
26	OYE23	-0.02644	0.952878
27	orf19.5517	-0.1072	0.761578
28	orf19.2825	0.164612	0.696466
<b>29</b>	<b>orf19.3874</b>	<b>0.014266</b>	<b>0.967379</b>
30	TRX1	-0.10484	0.767027

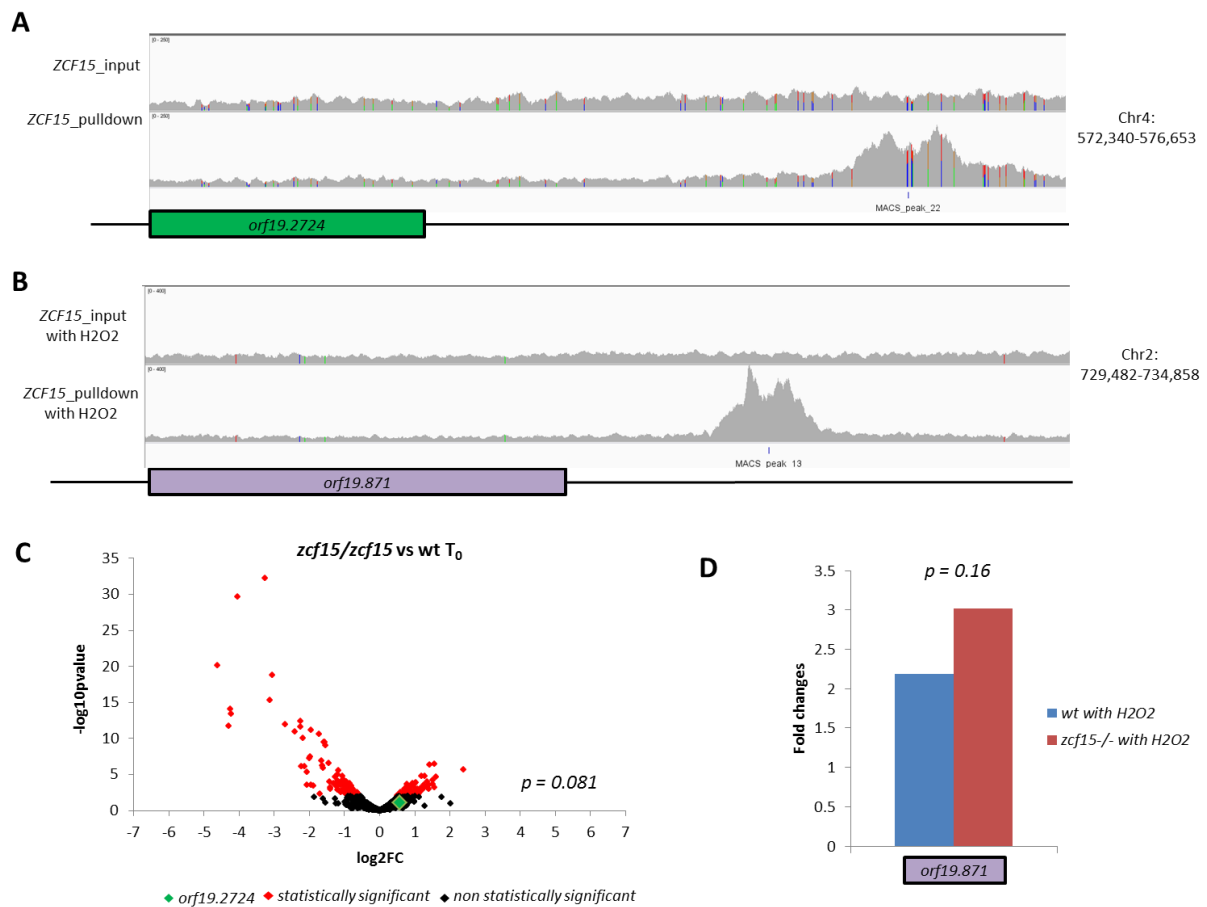
**Figure 4.30: few *ZCF15* DNA binding sites flank genes differentially expressed in *ZCF15* deletion.**

- A) In the absence of H<sub>2</sub>O<sub>2</sub>, 3 of the 37 *ZCF15* genomic binding sites (ChIP-Seq data) flank genes that are differentially expressed upon *ZCF15* deletion (RNA-Seq data). Genes reported in bold indicate *ZCF15* genomic binding partners bound both in the presence and absence of H<sub>2</sub>O<sub>2</sub> (7 total, shown also in Figure 4.26 A)
- B) In the presence of H<sub>2</sub>O<sub>2</sub>, 3 of the 30 *ZCF29* DNA binding sites (ChIP-Seq data) flank genes that are shown to be H<sub>2</sub>O<sub>2</sub> responsive and *ZCF29* dependent.

For example in the absence of H<sub>2</sub>O<sub>2</sub>, we found that only 3 of the 37 *ZCF15* DNA binding sites are flanking genes differentially expressed upon *ZCF15* deletion (Figure 4.30 A, genes highlighted in yellow). Moreover, 3 of the 30 DNA binding sites that *ZCF15* recognizes in the presence of H<sub>2</sub>O<sub>2</sub> are flanking genes that are H<sub>2</sub>O<sub>2</sub> responsive and *ZCF15* dependent (Figure 4.30 B highlighted in yellow). This evidence wasn't surprising because (a) compared to *Zcf29* we showed that the deletion of *Zcf15* causes a significantly more modest transcriptional rewiring (168 differentially expressed genes for *ZCF15* deletion versus 1402 differentially expressed genes for *ZCF29*, Figure 4.19 and Figure 4.14) (b) it is not uncommon for TF to binding regions not to flank differentially expressed genes. For example it is possible that many of the *Zcf15* bound regions either regulate more distal transcripts or that the binding is structural (help maintain chromosome structure). The presence of distant-acting silencers or enhancers is well documented in *C. albicans* (Tuch et al., 2010) and it is possible that many of the genes controlled by *Zcf15* are located to genomic loci not flanking *Zcf15* bound regions.

In addition, we can't rule out the hypothesis that our stringent requirements for differential expression identification have partially limited our ability to correlate *Zcf15* ChIP-Seq and RNASeq data. For example, as shown in Figure 4.31, we found a large *Zcf15* bound regions upstream of *orf19.2724* (Figure 4.31 A) in the absence of H<sub>2</sub>O<sub>2</sub> and a large *Zcf15* bound region upstream of *orf19.871* in the presence of H<sub>2</sub>O<sub>2</sub> (Figure 4.31 B). These two genes were found to be differentially expressed in analogous conditions in our RNASeq but not to statistically significant levels. In particular, in the absence of H<sub>2</sub>O<sub>2</sub>, deletion of *ZCF15* leads to the upregulation of *orf19.2724* (1.5 folds) but not to statistically significant levels ( $p = 0.081$ ), Figure 4.31 C. Similarly, upon H<sub>2</sub>O<sub>2</sub> exposure, *orf19.871* is upregulated 2 fold and 3

fold in wild type and *zcf15/zcf15* respectively but their differential expression was not statistically significant ( $p = 0.16$ ) Figure 4.31 D.



**Figure 4.31: Zcf15 may control the expression of orf19.7224 and orf19.871.**

- A) A 4.3 kb span of *C. albicans* genome displayed in IGV (chromosome 4 location coordinates 572,340 – 576,653). The two tracks labeled *ZCF15\_pulldown* and *ZCF15\_input* represent reads from pulldowns samples and controls obtained before the addition of  $H_2O_2$ . A large statistically significant peak is present right upstream of *orf19.2724* (3.1 fold enriched  $p < 0.01$ ).
- B) A 5.3 kb span of *C. albicans* genome displayed in IGV (chromosome 2 location coordinates 729,482 – 734,858). The two tracks labeled represent pulldown and controls samples obtained after the addition of 5mM  $H_2O_2$ . A large statistically significant peak is present right upstream of *orf19.871* (3.6 fold enriched  $p < 0.01$ ).
- C) Volcano plot showing that *Zcf15* deletion leads to a non-statistically significant ( $p = 0.081$ ) upregulation of *orf19.2724*.
- D) *orf19.871* is upregulated 2 folds upon  $H_2O_2$  exposure in wildtype and 3 folds in *Zcf15* deletion but their difference is non-statistically significant ( $pvalue = 0.16$ ).

Overall the ChIP-Seq data validated many of the genes found to be differentially expressed in our RNA-Seq experiment. We were able to find many examples in which genes differentially



expressed were bound by Zcf29 in their respective upstream regulatory sequences. We highlighted here four emblematic examples in which the expression of orf19.7042, orf19.7204, orf19.1048 and orf19.251 is directly regulated by Zcf29. Zcf29 functions as a direct activator and repressor of orf19.7042 and orf19.7204 respectively in the absence of H<sub>2</sub>O<sub>2</sub>. On the other hand, upon H<sub>2</sub>O<sub>2</sub> exposure Zcf29 binds to the orf19.1048 and orf19.251 upstream regulatory sequences directly up regulating their expressions.

Overall our data showed that Zcf15 and Zcf29 relocate themselves to different genomic loci upon H<sub>2</sub>O<sub>2</sub> exposure and sheds some light on the genes that are directly regulated by these two transcription factors. This study also highlights that many of the TF bound regions are not flanking differentially expressed genes, in agreement with evidence previously shown by others (Tuch et al., 2010) and that reiterates that distant acting TF regulation is common in *C. albicans*.

#### 4.4 Discussion

In the previous chapter we described *ZCF15*, a *C. albicans* transcription factor poorly characterized in the literature that has a key role in resistance to host generated reactive oxygen species (ROS). Encouraged by the fact that this gene belongs to a family of transcription factor that is not conserved in humans and therefore can be potentially used for either epitope-based vaccine approaches or drug targeting, we expanded our study to other genes within the same family. We looked for genes that (a) have a role in resistance to reactive oxygen species (b) are conserved in pathogenic fungi but not in non-pathogenic one. We identified *ZCF29*, a transcription factor that has been previously shown to be upregulated in the presence of H<sub>2</sub>O<sub>2</sub> (Enjalbert et al., 2006), required for resistance to ROS generator menadione (Homan et al., 2009) and present only in pathogenic fungi.

In order to explore the biological roles of *Zcf15* and *Zcf29*, we leveraged the power of RNA-Seq and ChIP-Seq. Our analysis showed that *zcf29* deletion leads to the misregulation of a staggering number of genes (1440 genes, ~25% of the genome). These genes are mainly enriched for biological functions involved in N-metabolism. *ZCF29* deletion leads to the downregulation of genes involved in biosynthesis of threonine, methionine, tryptophan, tyrosine, lysine and histidine and simultaneous upregulation of genes involved in proteolysis. This evidence suggests that upon *ZCF29* deletion *C. albicans* favors amino acids recycling through proteolysis over *de novo* biosynthesis. Overall, amino acid *de novo* biosynthesis is repressed in *zcf29/zcf29* but we found one noticeable exception in our dataset: arginine biosynthesis. Six of the seven genes of the arginine biosynthetic pathway are upregulated between 23 and 6 folds in *zcf29/zcf29* and the only genes not upregulated in this pathway is largely regulated post transcriptionally in *S. cerevisiae*. In murine macrophages, *C. albicans* is killed by a variety of different mechanisms that are dependent on arginine availability. In

particular, as shown in Figure 3.6, reactive nitrogen species (RNS) are produced by nitric oxide synthase an enzyme that converts arginine to citrulline with simultaneous production of nitric oxide (Marletta et al., 1988). Nitric oxide is subsequently converted to more toxic RNS like nitrogen dioxide radicals ( $\text{NO}_2^\bullet$ ) and peroxynitrite ( $\text{ONOO}^-$ ) that have the fungicidal effect on *C. albicans*. We hypothesized that the upregulation of arginine in *zcf29/zcf29* can be the underlying cause of the reduced virulence reported by Vandeputte et al. 2011. We believe that the availability of more arginine in *zcf29/zcf29* leads to a more efficient production of macrophage generated RNS and this leads to the *zcf29/zcf29* hypo-virulence. Although this hypothesis requires further biological validations, our data reinforce the yet poorly understood connection between arginine and virulence that have been recently reported by others (Jimenez-Lopez et al., 2013). In addition, our RNASeq experiment revealed for the first time that *Zcf29* plays a critical role in the well-documented ability of the pathogen to rechannel energy resources from ribosome biogenesis to ROS detoxification processes upon ROS exposure.

In contrast with what has been shown for *ZCF29* deletion, the deletion of *ZCF15* caused a significantly more modest transcriptional rewiring with only 168 differentially expressed genes at  $T_0$ . This result can potentially be explained by gene redundancy which is defined as the existence of several genes in the genome that perform the same or similar functions. In particular, while *ZCF29* doesn't have any homologs in *C. albicans*, *ZCF15* has three other homologs that can potentially compensate for the lack of this gene in the knockout. This hypothesis is further supported by the fact that the deletion of *ZCF15* causes hypersusceptibility to ROS only while deletion of *ZCF29* caused hypersusceptibility not only to ROS but also to cell wall stressor and nitrogen starvation mimicking caffeine.

Overall, while Zcf29 plays a critical role in N-metabolism, we found that Zcf15 plays a critical role for C-metabolism, TCA cycle and energy production. In particular our RNASeq experiment showed for the first time that Zcf15 plays a critical role in the well-documented ability of the pathogen to downregulate genes involved in C-metabolism and contemporarily up regulate key enzymes involved in ROS detoxification processes like thiol peroxidases.

Taken together, our data suggests that *ZCF15* and *ZCF29* play an important role in regulating two of the most important cellular nutritional requirement: C- and N- metabolism. As discussed in paragraph 1.4, during the infection the main nutrient sources for *C. albicans* are likely to be host-derived glucose, proteins, amino acids and lipids whose concentration varies dramatically in the different anatomical niches. The ability of *C. albicans* to rapidly and dynamically respond to host changes in micro-environmental nutrient availability contributes to its success as a pathogen. Our data suggests that genes involved in these processing are miss regulated in *zcf15/zcf15* and *zcf29/zcf29* and may be the cause of their reduced virulence.

While our RNASeq experiment highlighted genes that are either directly or indirectly controlled by Zcf15 and Zcf29, the ChIP-Seq analysis allowed us to find direct DNA binding targets. Our analysis revealed that Zcf15 and Zcf29 relocate themselves to different genomic loci upon H<sub>2</sub>O<sub>2</sub> exposure. This evidence, on one side reiterates the fact these two TFs play key roles in ROS resistance and, on the other, validate our RNASeq data that suggested different roles of these two TFs in the presence or absence of H<sub>2</sub>O<sub>2</sub>. The ChIP-Seq data showed that Zcf29 plays a critical role in the H<sub>2</sub>O<sub>2</sub> dependent upregulation of a glutathione-independent glyoxalase and of an aldo-keto reductase. Both of these genes have been shown to be important for cellular detoxification processes and, to the best of our knowledge, this is

the first time that the expression of these two critical H<sub>2</sub>O<sub>2</sub> responsive genes have been shown to be directly controlled by Zcf29.

Overall, our data shed some lights on the genes and biological functions controlled by Zcf15 and Zcf29, two new virulence determinants that plays a critical role in resistance to reactive oxygen species in *C. albicans*.

## 5 Filastatin, a new antimycotic?

### 5.1 Introduction and contribution to the paper

In this chapter we described how we used our *C. albicans* - *C. elegans* virulence model for small molecule screening. Data described in chapter 2 showed that this model can be pharmacologically modulated (Figure 2.4 C), hence potentially useful for drug discovery. In particular, we showed that the addition of 50  $\mu$ M of fluconazole significantly extended the survival of *C. albicans* infected worms (Figure 2.4). Encouraged by this proof of principle experiment, we expanded our study and tested both fluconazole and amphotericin B at various concentrations for their abilities to alter the course of infection as measured by extend worm's life spans and reduce Dar. These two drugs are the most commonly prescribed drugs against *C. albicans* infections, but they have very different mechanisms of action. Fluconazole reduces cell growth by inhibiting the synthesis of ergosterol (a sterol that maintains membrane fluidity) and it is considered fungistatic because it can inhibit growth but not cause cell death. On the other side, amphotericin B is fungicidal, because it directly binds ergosterol depolarizing the cell membrane causing leakage of intracellular cations and ultimately cell death. Fluconazole is more commonly prescribed (can be delivered orally) but poses the threat of generating drug resistance; amphotericin B is more potent but it can only be administrated intravenously and is associated with high nephrotocity. Here, we show that both drugs added to our assays at the optimal concentration can reduce the Dar phenotype and increase nematode survival.

Once we optimized our assay and showed that it responds to commonly prescribed antimycotics, we used it to test a series of 26 compounds recently identified in collaboration with the Kaufman lab at the University of Massachusetts Medical School (UMMS) for their

abilities to inhibit *C. albicans* adhesion to polystyrene. These 26 compounds were identified from a library of 30,000 using a stratified screening method combined with a chemical structure analysis. We treated infected *C. elegans* with each of these 26 compounds and used the Dar assay to test for efficacy and toxicity. We found that one of these compounds, named filastatin, was very potent in reducing Dar. We rigorously tested filastatin alongside with fluconazole and amphotericin B and found that it also alleviate disease and improves nematode survival to levels comparable to fluconazole. Our data in nematodes was subsequently validated by our collaborators at UMMS and at Louisiana State University, who showed that filastatin is also capable of inhibiting binding to human epithelial cells and mouse mucosal infection assays. Overall our collaboration with UMMS led to the publication of the following paper in the Proceedings of the National Academy of Sciences (PNAS) (Fazly et al., 2013). My contribution to the paper is described in the narrative of the paper (paragraph titled “Filastatin Exhibits Antifungal Activity in a Nematode Model of Infection”) and summarized in Figure 5.

## 5.2 Chemical screening identifies filastatin, a small molecule inhibitor of *Candida albicans* adhesion, morphogenesis, and pathogenesis

Ahmed Fazly<sup>a</sup>, Charu Jain<sup>b</sup>, Amie C. Dehner<sup>a</sup>, Luca Issi<sup>b</sup>, Elizabeth A. Lilly<sup>c</sup>, Akbar Ali<sup>d</sup>, Hong Cao<sup>d</sup>, Paul L. Fidel, Jr.<sup>c</sup>, Reeta P. Rao<sup>b,1</sup>, and Paul D. Kaufman<sup>a,1</sup>

<sup>a</sup>Program in Gene Function and Expression and <sup>d</sup>Small Molecule Screening Facility, University of Massachusetts Medical School, Worcester, MA 01605; <sup>b</sup>Department of Biology and Biotechnology, Worcester Polytechnic Institute, Worcester, MA 01609; and <sup>c</sup>Department of Oral and Craniofacial Biology, Louisiana State University Health Sciences Center School of Dentistry, New Orleans, LA 70119

Edited by Jasper Rine, University of California, Berkeley, CA, and approved July 9, 2013 (received for review April 1, 2013)

**Infection by pathogenic fungi, such as *Candida albicans*, begins with adhesion to host cells or implanted medical devices followed by biofilm formation. By high-throughput phenotypic screening of small molecules, we identified compounds that inhibit adhesion of *C. albicans* to polystyrene. Our lead candidate compound also inhibits binding of *C. albicans* to cultured human epithelial cells, the yeast-to-hyphal morphological transition, induction of the hyphal-specific HWP1 promoter, biofilm formation on silicone elastomers, and pathogenesis in a nematode infection model as well as alters fungal morphology in a mouse mucosal infection assay. We term this compound filastatin based on its strong inhibition of filamentation, and we use chemical genetic experiments to show that it acts downstream of multiple signaling pathways. These studies show that high-throughput functional assays targeting fungal adhesion can provide chemical probes for study of multiple aspects of fungal pathogenesis.**

## 5.3 Introduction

In humans, the most widespread fungal pathogen is *Candida albicans*, which is also one of the most frequent causes of hospital-acquired infections (1, 2). As an opportunistic pathogen, *C. albicans* is responsible for common clinical conditions, including oral thrush and vaginitis, but it can also lead to life-threatening systemic infections (candidemia) in immunocompromised patients (3), resulting in ~30–50% mortality rates (2). Further-more, the estimated annual cost of treating systemic *Candida* infections exceeds \$1 billion per year (1, 2). A contributing factor to these statistics is the ability of *C. albicans* to develop resistance to antifungal drugs (4). For these reasons, new strategies for combating fungal infections without toxicity to humans are a high medical priority.

Adhesion to surfaces is the first step in establishing a fungal infection. *Candida* cells with a planktonic yeast morphology initiate adhesion, and a subsequent transition from yeast to



hyphal morphology leads to tissue invasion and biofilm formation (5–8). Biofilm formation is an important step in pathogenesis, because after biofilms form on intravascular devices and catheters, they normally do not respond to conventional treatment and subsequently, can give rise to life-threatening systemic infections (9). Therefore, biofilm-associated medical devices often have to be surgically replaced (10).

Here, we report a chemical screen for compounds that prevent adhesion of *C. albicans* to polystyrene surfaces. One of the identified compounds also inhibited adhesion to human cells. Further-more, this compound prevented adhesion of other *Candida* species to polystyrene, diminished the *C. albicans* yeast-to-hyphal transition, impaired biofilm formation on silicone elastomers, reduced fungal pathogenesis in a nematode infection model, and altered biofilm morphology in a mouse mucosal infection model. Based on its strong inhibition of filamentation, we term this compound filastatin, and we show that it blocks filamentation induced by some but not all external signals. In sum, our chemical screen identified an inhibitor of multiple pathogenesis-related functions that will also be a useful probe of morphological mechanisms as well as a lead for the development of unique antifungal therapeutics.

## 5.4 Results

### Identifying Small Molecule Inhibitors of *C. albicans* Adhesion

To detect small molecules that inhibit *C. albicans* adhesion, we modified a previous protocol for measuring adhesion of *Saccharomyces cerevisiae* to polystyrene (11). We observed that *C. albicans* cells bind strongly to 96-well polystyrene plates that have been optimized for protein binding (Methods), and we used this robustly bound substrate to set a high threshold for inhibitors of adhesion. Bound cells were stained with crystal violet followed by washing to remove unbound dye and cells. Cells that remained bound after washing were then quantified by measuring dye absorbance. Wells that displayed low  $A_{590}$  values were confirmed visually (Fig. 1A).

We screened a library of 30,000 small molecules for compounds that inhibit adhesion of WT clinical isolate *C. albicans* strain SC5314 (12). All compounds were dissolved in DMSO, and therefore, we used SC5314 cells exposed only to DMSO as a negative control (Fig. 1A, column 12). As a positive control for poor adhesion, we used a *C. albicans* *edt1*<sup>-/-</sup> strain (13)

that lacks a cell wall adhesion protein (Fig. 1A, column 1). We ranked the normalized scores for each compound (Methods). Of 30,000 compounds, 40 (1.3%) compounds inhibited adhesion by >75%, and many of these compounds fell into two scaffold families [1-benzoyl-4-phenylpiperazines (termed Scaffold 1) and N-phenylbenzamides (termed Scaffold 2)] (Table S1). Omitting compounds within this list that were very similar, we reordered 26 of the candidate compounds (termed 1–26) (Table S1) for subsequent characterization.

We prioritized the candidate compounds by testing the dose dependence of their effects. Because of the laborious washing required to remove background crystal violet staining, we used an alternative detection reagent, the vital dye alamarBlue. We found that, after two rounds of washing, alamarBlue robustly detected the adhesion differences between the WT SC5314 and *edt1*<sup>-/-</sup> mutant cells (Fig. 1C and Fig. S1 A). The majority of the reordered compounds were effective at reducing *Candida* adhesion at 50  $\mu$ M, which is the concentration originally tested during the screen (Methods). However, on reducing the concentration to 25  $\mu$ M, some of these compounds were substantially less effective (Fig. 1B). However, many candidates were able to inhibit adhesion by >50%, even at concentrations as low as 7.5  $\mu$ M (Fig. 1C).

Because alamarBlue measures the metabolic activity of live cells, we confirmed these data with an assay that directly detected cells, regardless of viability or metabolism, by measuring adhesion using a *C. albicans* strain encoding GFP (13). Both fluorescence measurements and microscopy (Fig. 1 D and E) of the GFP-expressing strain confirmed that all of the compounds effective in the alamarBlue based assay impair *C. albicans* adhesion to polystyrene. We conclude that our chemical screen detected compounds that block adhesion and yielded consistent results, regardless of the detection method.

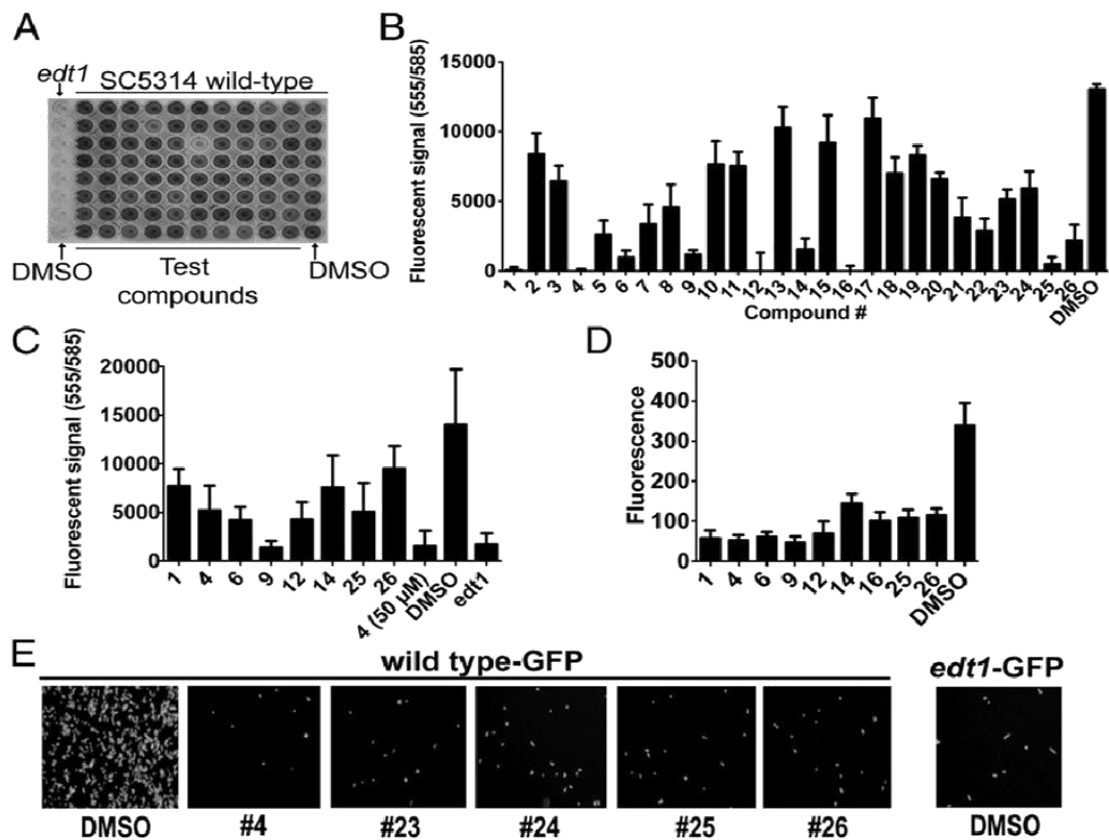


Fig. 1. Chemical inhibition of *C. albicans* adhesion to polystyrene. (A) Crystal violet-stained plate from small molecule screen. WT *C. albicans* strain SC5314 or adhesion-defective *edt1*<sup>-/-</sup> cells were plated in the presence of 1% DMSO vehicle or small molecules at 50 μM followed by extensive washing to remove unbound cells. Cells that remained were then stained with crystal violet and quantified by A<sub>590</sub>. (B) alamarBlue-based polystyrene adhesion assay. Using the vital dye alamarBlue as the detection reagent, 26 primary candidates from the screen were retested at 25 μM alongside a DMSO control. The mean and SD of data from eight wells are shown. (C) The same as in B but with compounds at 7.5 μM. The mean and SD of data from eight wells are shown. (D) GFP-based adhesion assay. *C. albicans* strains expressing GFP (SC5314-GFP and *edt1*<sup>-/-</sup>-GFP) or untagged SC5314 were mixed with compounds at 25 μM or DMSO and bound to polystyrene plates. Unbound cells were removed by washing. GFP fluorescence was measured, and background fluorescence of wells containing un-tagged SC5314 was subtracted from signals. Mean and SD for eight wells are shown. (E) Wells from the experiment in D were photographed using a 20× objective and FITC filters.

In addition to experiments with inert surfaces, we also tested how the candidate compounds would affect *C. albicans* adherence to human cells by measuring binding of GFP-encoding *C. albicans* cells to monolayers of human lung epithelial A549 cells, which are an effective substrate for *Candida* adhesion (14). We tested each of the reordered compounds and observed that compound 4 was, by far, the strongest inhibitor of adhesion to human epithelial

cells, which was shown by both fluorescence quantitation (Fig. 2A) and microscopy (Fig. 2B). We, therefore, prioritized study of this compound (Fig. 2F), which we termed filastatin based on its inhibition of *C. albicans* filamentation as described below. We showed that filastatin does not affect the viability of human A549 cells, even at concentrations much larger than those concentrations used in the adhesion assay (e.g., 250 $\mu$ M) (Fig.2C). We conclude that filastatin is not toxic to this human cell line under our assay concentrations, but it is unique among our candidates in that it can impair fungal adhesion to both inert surfaces and cultured human epithelial cells.

We next tested whether filastatin could affect adhesion after cells had already bound polystyrene, because this activity could prove useful in a clinical setting. We compared cells continuously incubated with filastatin for 8 h with cells adhered for 4 h (as in our previous assays) (Fig. 1B) followed by 4 h in the presence of filastatin. We observed that filastatin did, indeed, reduce the amount of bound cells when added after adhesion, although not as efficiently as when the compound and cells are coincubated from the start of the experiment (Fig. 2D). We conclude that filastatin does display some activity against adhered cells, consistent with potential use against pathogen biofilms.

To test the applicability of filastatin to other fungal pathogens, we examined adhesion by additional pathogenic *Candida* species (15). Specifically, we observed that filastatin inhibited polystyrene adhesion by *C. dubliniensis* and *C. tropicalis* and to a lesser extent, *C. parapsilosis* (Fig. 2E). Filastatin, therefore, inhibits adhesion by multiple pathogenic *Candida* species. Additionally, filastatin (molecular weight = 360 g/mol) has an  $IC_{50} \sim 3 \mu$ M in the GFP-based adhesion assay (Fig. S1 B), equivalent to a concentration of  $\sim 1.1 \mu$ g/ml. This value is similar to or less than the effective concentrations of existing antifungal compounds (16).

Additionally, data in Figs. S2, S3, and S4 compare the activities of filastatin with other compounds from this study and prior studies (17). Of all these studies, only filastatin inhibited *C. albicans* adhesion to human cell monolayers, and therefore, we focused additional studies on this molecule.

### **Filastatin Inhibits *C. albicans* Biofilm Formation**

Fungal biofilm formation on implanted medical devices is a serious medical problem, because these biofilms can lead to life-threatening systemic infections (18). An established *in vitro* assay to study device colonization measures biofilm formation on surgical silicone elastomers (19). In the presence of DMSO vehicle, we confirmed that WT *C. albicans* cells efficiently formed biofilms on silicone elastomers; in contrast, *edt1*<sup>-/-</sup> mutant cells did not, resulting in the majority of the cells dispersing throughout the media rather than adhering to the elastomer surface (Fig. 3A). Visual inspection showed that filastatin, but not compounds 6 or 9, effectively kept the cells dispersed in the media rather than on the elastomer surface (Fig. 3A). Measurements of the media turbidity and biofilm dry weight (Fig. S5) confirmed these assessments. Additionally, we observed that compounds 27, Q1, and Q2 also inhibited biofilm formation in this assay (Fig. S5). Therefore, filastatin has multiple activities, with inhibition of human cell adhesion unique among our candidate compounds (Tables S1 and S2).

### **Filastatin Inhibits Hyphal Morphogenesis and Colony Morphology**

Visual inspection of cells remaining at the end of adhesion assays suggested that some of the candidate compounds inhibit generation of hyphae. Because the ability to interconvert between yeast and hyphal morphologies is usually correlated with pathogenicity (5, 7, 8), we explored this finding in more detail, examining induction of hyphae on carbon starvation (Spider media) (20). We used a strain that contains the Red Fluorescent Protein ORF driven by the hyphal-specific HWP1 promoter (21), which provides a molecular reporter for hyphae formation in addition to cell morphology. We observed that filastatin as well as a subset of the other compounds (Fig. S6 A and Table S1) blocked formation of hyphae. Together, our data indicate that the related compound pairs (filastatin/27 and Q1/Q2) were effective inhibitors of both hyphal morphogenesis and biofilm formation on silicone elastomers (Fig. 3A and Figs. S5 and S6). Therefore, the shared backbone of these compounds correlates with inhibition of multiple activities (summarized in Table S2). On titration, we observed that filastatin is effective at inhibiting hyphae formation at concentrations > 2.5  $\mu$ M (Fig. S6 B), similar to the results observed in the polystyrene adhesion assay (Fig. S1). Using this protein reporter assay, the effect of filastatin could be detected within 2 h of Spider media addition (Fig. S6 C).

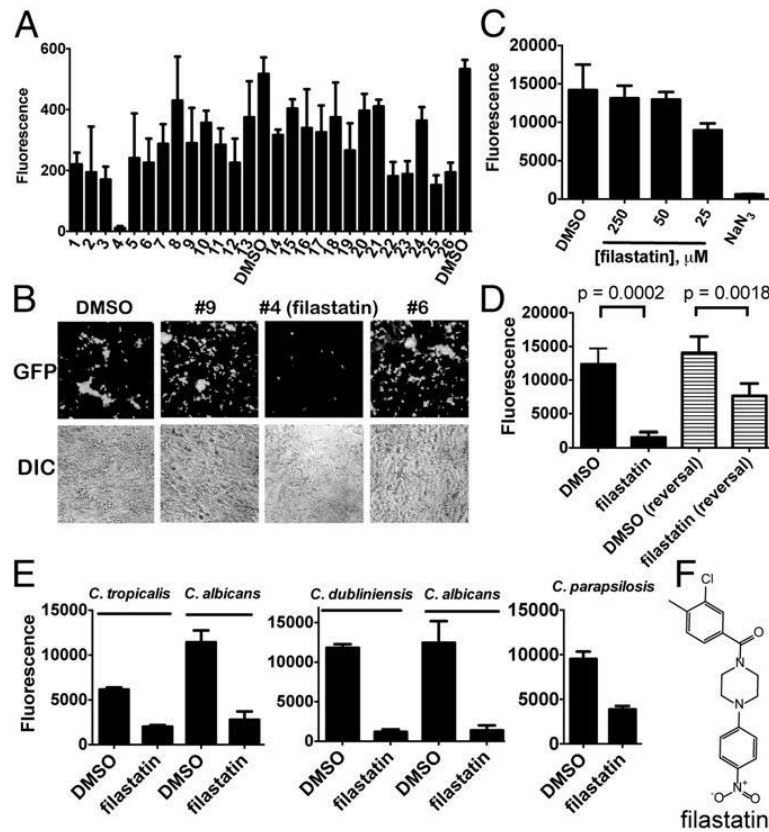


Fig. 2. Filastatin is a nontoxic inhibitor of adhesion to human cells, and it affects multiple *Candida* species. (A) Adhesion to A549 cells. Human A549 cells were grown to confluence on 48-well plates. SC5314-GFP, *edt1*<sup>-/-</sup>-GFP, or untagged SC5314 were added to triplicate wells and incubated at 37 °C for 90 min with 1% DMSO or 25 μM indicated compounds in 1% DMSO. Wells were washed extensively to remove unbound fungi, and GFP fluorescence was measured on a plate reader. Mean and SD for three wells are shown. (B) Representative differential interference contrast (DIC) and cor-responding GFP fluorescence images from the experiment in A. (C) Human cell toxicity. A549 cells were coincubated with 1% DMSO, 250, 100, or 50 μM filastatin, or 1% NaN<sub>3</sub> for 24 h, and cell viability was then measured using alamarBlue. Mean and SD for three wells are shown. (D) Reversal of adhesion. WT *C. albicans* SC5314 cells were seeded onto 96-well plates. One-half the samples were incubated continuously in the presence of 1% DMSO or 50 μM filastatin as indicated for 8 h at 37 °C (left black bar). One-half of the samples were allowed to adhere for 4 h in the absence of compounds, and then, 1% DMSO or 50 μM filastatin was added followed by additional incubation for 4 h at 37 °C (reversal; left striped bar). After incubations, plates were washed two times, and bound cells were detected with alamarBlue. Means and SDs (n = 5 independent experiments, 24 wells measured for each) with P value comparisons are shown (unpaired two-tailed t test). (E) Other *Candida* species. Adhesion assays are the same as in Fig. 1B, with 25 μM filastatin tested for *C. dubliniensis* and *C. tropicalis* and 50 μM filastatin tested for *C. parapsilosis*. Mean and SD for 16 wells are shown. (F) Structure of filastatin.

Also, because filastatin blocked filamentation, even after 16 h of incubation (Fig. S6 A), it

seemed unlikely that this inhibitory function relied on a reduced growth rate. Although filastatin did modestly reduce growth rates of yeast-form cells in rich media (Fig. S7 A), this effect was less pronounced during hyphal induction, where we assessed growth by both OD and metabolic activity (Fig. S7 B and C). Furthermore, filastatin maintained the yeast morphology of cells throughout an 8-h time course ( Fig. S7 D). We conclude that filastatin is a long lasting inhibitor of *C. albicans* filamentation.

Hyphal morphogenesis is induced by multiple stimuli that act through multiple signal transduction pathways (5, 22, 23) (Fig. 4E) and mediated by a complex transcriptional network (8, 24). In addition to nutrient-poor Spider media, mammalian serum also induces hyphae (25, 26), and we observed that filastatin inhibited the response to both these stimuli (Fig. 3B). We detected morphological difference between control and filastatin-treated cultures as early as 60 min in 10% (vol/vol) serum and 90 min in Spider media. Additionally, a wrinkled colony morphology is displayed by *C. albicans* when grown on solid Spider media, reflecting transitions between hyphal and yeast forms (5, 27). Notably, this wrinkled phenotype was also abolished by filastatin (Fig. 3B). Together, our data indicate that filastatin inhibits hyphal morphogenesis and colony morphology phenotypes that depend on the hyphal transition.

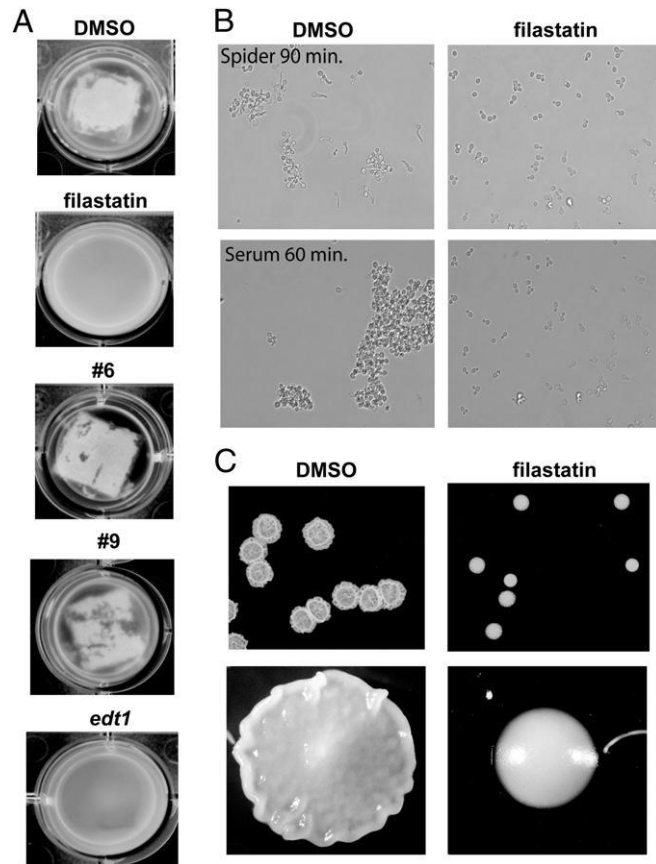


Fig. 3. Filastatin inhibits *C. albicans* biofilm formation and hyphal morphogenesis. (A) Biofilm formation on silicone elastomers. Biofilms formed by SC5314-GFP (or *edt1*<sup>-/-</sup>-GFP where indicated) cells were photographed after 60 h. Indicated compounds were added at 50  $\mu$ M. Turbid media indicate planktonic cells unattached to silicone elastomers when biofilm formation is inhibited. Clear media is observed when robust biofilm formed. Representatives of triplicate wells are shown. Additional biofilm data are in Fig. S5. (B) Hyphal morphogenesis. SC5314 cells were diluted into Spider media or Supplemented YNB + 10% bovine serum with DMSO or 50  $\mu$ M filastatin as indicated, grown at 37  $^{\circ}$ C for the indicated times, and photographed. Additional hyphal development experiments are in Fig. S6. (C) Colony morphology. SC5314 cells were spotted onto solid Spider media agar plates overlaid with either 1% DMSO or 50  $\mu$ M filastatin and incubated at 37  $^{\circ}$ C for 2 d before photography.



## **Filastatin Inhibits Induction of Hyphal Morphogenesis by Multiple but Not All Signaling Pathways**

To explore the mechanism of filastatin action, we tested whether it would alter hyphal induction in hyperfilamentous mutants. For example, hyphal induction by Spider media requires the cAMP-PKA pathway (23, 26) (Fig. 4E). Cells constitutively overexpressing the G protein-coupled receptor Gpr1 become hyperfilamentous on solid Spider media by PKA stimulation (26, 28), and we observed that filastatin blocked hyphal morphogenesis in these cells (Fig. 4A). Stimulation of the PKA pathway also drives phosphorylation of transcription factor Efg1, activating Efg1 to increase expression of genes required for hyphal morphogenesis. We confirmed that cells with a hyperactive Ras1 signaling protein (ras1-G13V) (29), an overexpressed, constitutively active G- $\alpha$  subunit that acts up-stream of Efg1 (gpa2-Q355L) (28), or a constitutively expressed Efg1 transcription factor with a phosphomimetic mutation that simulates constitutive PKA signaling (PCKpr-efg1-T206E) (30), indeed, were all hyperfilamentous compared with WT cells in Spider media (Fig. 4B). However, in the presence of filastatin, cells from all these mutant strains retained a planktonic, budded morphology (Fig. 4A). Together, these data suggest that filastatin acts downstream of transcription factor Efg1 (Fig. 4E).

However, other experiments suggested that filastatin affects more than one signaling pathway. For example, the modified sugar GlcNac also stimulates hyphal morphogenesis but does so independently of the cAMP-PKA pathway; instead, it activates the transcription factor Cph1 (26). On testing morphogenesis driven by GlcNac-containing media or constitutive overexpression of Cph1, we found that filastatin also inhibits formation of hyphae in these cases (Fig. 4C). These data indicated that filastatin may affect multiple signaling pathways or could act by destroying the ability of the cell to form elongated structures, regardless of the inducing signal.

In addition to nutrient-mediated signals, genotoxic stress promotes hyphal morphogenesis in *C. albicans*, and the DNA damage signaling kinase Rad53 is required for the hyphal transition (31). In contrast to the above studies, cells treated with the DNA replication inhibitor hydroxyurea (HU) form hyphae, regardless of the presence of filastatin (Fig. 4D). Together, these data suggest that filastatin is not incompatible with hyphal morphogenesis per se but blocks multiple signaling mechanisms that promote it (Fig. 4E).

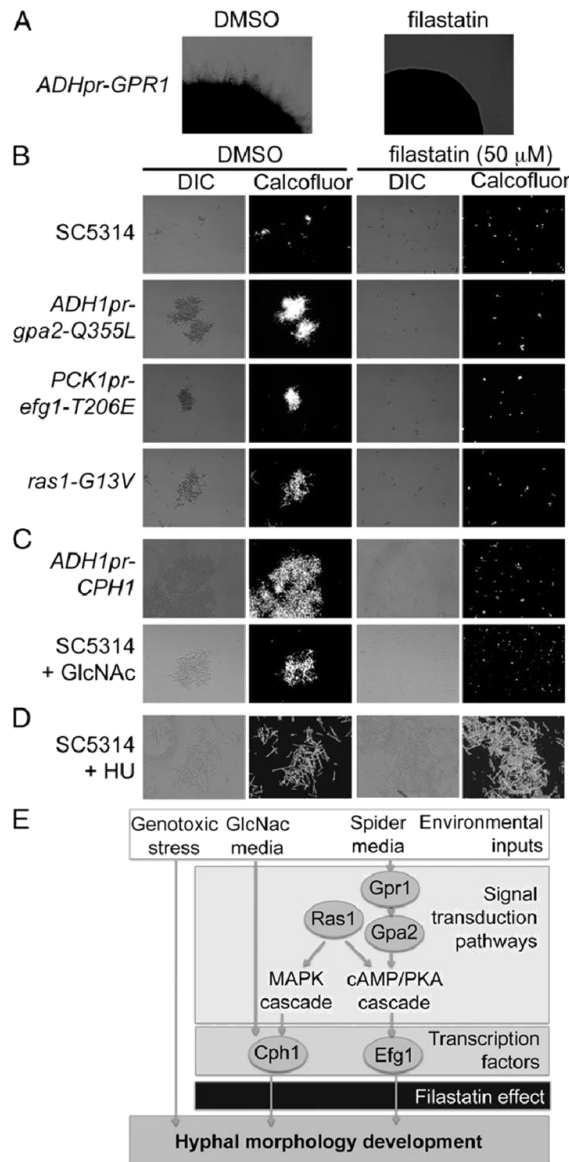


Fig. 4. Filastatin inhibits hyphal development induced by some but not all external signals. (A) Filastatin inhibits the hyperfilamentous growth of a strain that constitutively overexpresses Gpr1. Colonies were grown on solid Spider media agar overlaid with either 1% DMSO or 50  $\mu$ M filastatin and incubated at 37  $^{\circ}$ C for 3 d. (B) Effects of filastatin on strains hyperactivating the cAMP-PKA pathway. Cells of the indicated genotypes were grown in the presence of 1% DMSO or 50  $\mu$ M filastatin in Spider medium at 37  $^{\circ}$ C for 4 h, stained with Calcofluor, and photographed using DIC and fluorescence microscopy. (C) Filastatin blocks filamentation induced by GlcNAc or constitutive overexpression of the GlcNAc-activated transcription factor Cph1. (D) Filastatin does not block genotoxic stress-induced filamentation. SC5314 cells were grown in yeast extract-peptone-dextrose (YPD)-rich media + 50  $\mu$ M HU at 30  $^{\circ}$ C. (E) Schematic diagram of some of the signaling pathways that govern hyphal morphogenesis (22, 26). For simplicity, only those genes investigated are shown. Filastatin is shown blocking hyphal development induced by Spider media and GlcNAc but not genotoxic stress.

### **Filastatin Exhibits Antifungal Activity in a Nematode Model of Infection**

In addition to our in vitro assays, we tested whether filastatin affects fungal infections in vivo. We used the previously described deformed anal region (Dar) tail-swelling phenotype (16, 32) as a biomarker to monitor fungal infections in nematode hosts. A small quantity of *C. albicans* was introduced with the nematode's standard diet of *Escherichia coli* OP50 cells, and *E. coli* growth was attenuated with the antibiotic streptomycin to avoid interactions with *C. albicans* (33). Here, we used a dose of *C. albicans* that caused all untreated worms to display Dar dis-ease (Fig. 5A). In contrast, ~30% of worms treated with either 12.5  $\mu$ M filastatin or fluconazole during this experiment maintained a healthy phenotype and did not display Dar disease. Worms treated with the same concentration of amphotericin B displayed an even larger decrease in Dar disease, perhaps due to amphotericin B being a fungicidal drug, while fluconazole is fungistatic (34). In any case, these data indicate that filastatin can ameliorate a fungal pathogenesis phenotype in vivo.

We next tested whether filastatin protects *Caenorhabditis elegans* against killing by *C. albicans*. *C. elegans* was exposed to *C. albicans* in the presence of DMSO, fluconazole, or filastatin. The *C. albicans* infections greatly shortened the lifespan of the *C. elegans* hosts compared with uninfected controls (Fig. 5B). We observed that both fluconazole and filastatin increased the lifespan of the infected worms (Fig. 5 B and C), indicating that both these compounds are protective during infection of a live animal. Increased lifespan in the presence of filastatin was statistically significant as confirmed by Kaplan–Meier statistical analysis (P value < 0.012 or less for all three experiments). Therefore, consistent with its in vitro activities, filastatin functions as an antifungal agent in this model infection system.

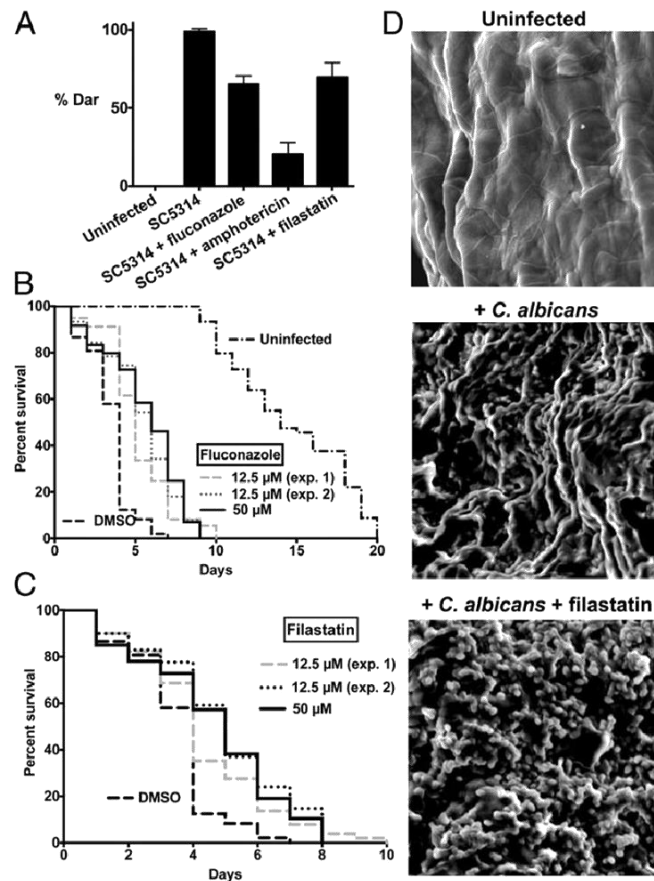


Fig. 5. Effects on metazoan hosts. (A) Nematodes were either cultured without *C. albicans* (Uninfected) or with WT *C. albicans* SC5314 cells with or without 12.5  $\mu\text{M}$  indicated compounds. Dar disease rates were compared for worms treated with filastatin and untreated worms ( $P = 2.9 \times 10^{-3}$ ). Infected worms were treated with antifungal agents fluconazole ( $P = 2.6 \times 10^{-4}$ ) and amphotericin B ( $P = 2.9 \times 10^{-3}$ ) as positive controls. Experiments were done in triplicate, with 14–28 animals in each biological replicate sample, with 63–81 total animals for each experimental condition. Student t test (one tail, equal variance) was used to calculate P values. (B) Lifespan of *C. elegans* exposed to *C. albicans* treated with DMSO, 12.5  $\mu\text{M}$  fluconazole (two experiments, plotted separately), or 50  $\mu\text{M}$  fluconazole;  $n = 60$  individuals for each experiment. All fluconazole-treated populations were significantly different from the DMSO control, with P values for each  $< 0.0001$  by the Mantel–Cox test. Lifespan of uninfected *C. elegans* is plotted on the same graph. (C) As in B, except worms were treated with filastatin at 12.5 (two experiments, plotted separately) or 50  $\mu\text{M}$ . All filastatin-treated populations were significantly different from the DMSO control, with  $P < 0.012$  for 12.5  $\mu\text{M}$  experiment 1,  $P < 0.0001$  for 12.5  $\mu\text{M}$  experiment 2, and  $P < 0.0002$  for the 50  $\mu\text{M}$  experiment. (D) Microscopic analysis of ex vivo vaginal mucosal biofilm formation. Tissues were processed for scanning EM. (Magnification: 1,000 $\times$ .) These representative images show areas of biofilm growth from two independent repeats ( $n = 6$  vaginal explants per experiment).

## **Filastatin Alters Biofilm Morphology in a Mouse Model of Vulvovaginal Candidiasis**

As an initial test of filastatin in a mammalian infection model, we examined a mouse model of vulvovaginal candidiasis (VVC). This system is ideal for evaluating our compounds, because VVC infections depend on fungal morphogenesis and biofilm formation (35). This assay is medically important, because VVC affects 75% of all women at least once in their lifetimes, and because *C. albicans* is responsible for 85–95% of these infections (36).

In this assay, excised murine vaginal mucosal tissue is tested as a substrate for infection by *Candida* cells. This protocol provides a rapid and simple method for optimizing treatment conditions before in vivo infection assays, although it does not measure adhesion to the tissue, because the tissue is not washed before processing for microscopy. One readout in this assay is a measure of the number of viable fungal cells expressed as cfu. We did not observe significant differences between the cfu re-covered from untreated and filastatin-treated samples, although because unadhered cells were not washed away, direct comparison with the in vitro adhesion assays cannot be made. Also, in the presence of filastatin, the *Candida* cells still formed a carbohydrate-rich ECM, which was detected by Con A staining and confocal microscopy (Fig. S8). However, filastatin modulated the morphology of the *C. albicans* biofilm during ex vivo VVC, which was revealed by scanning EM (Fig. 5D), with a predominance of individual yeast-form cells appearing. Therefore, filastatin restrained formation of hyphae in this infection model, consistent with its hyphal inhibition activity in vitro (Figs. 3 and 4). Future experiments will be required to determine how filastatin affects ultrastructural aspects of biofilm formation and how it may affect adhesion during biofilm formation on mammalian cells.

## **5.5 Discussion**

In recent years, the rate at which new drugs are brought to market has slowed (37). Traditional drug discovery methods have often relied on rational drug design, where agents that inhibit a particular protein are selected. However, this approach is time-consuming and can miss important but unanticipated targets. Here, we describe an alternate function-based approach that relies on phenotypic screening without bias to a particular target. To our knowledge, there has not, to date, been a published screen that directly targets fungal adhesion. By targeting adhesion, our screen required removal of unbound cells. Although

laborious in a high-throughput setting, this extra effort allowed us to detect a class of compounds that has unique properties. Importantly, because our assays involve intact cells, we avoided the complication of compounds unable to cross the cell wall and/or membranes. Here, we report discovery of filastatin, which displays multiple activities related to fungal pathogenesis. Filastatin inhibits fungal adhesion to polystyrene and human cells (Figs. 1 and 2), hyphal morphogenesis and colony morphology (Figs. 3 and 4), biofilm formation (Fig. 3), pathogenesis in a worm infection model (Fig. 5C), and biofilm morphology on mouse tissue explants (Fig. 5D). Notably, inhibition of hyphal morphogenesis occurs in liquid culture in the absence of adhesion, suggesting that filastatin affects multiple cellular processes. The effects on morphology were not selected in our primary screen, showing the need for multiple assays to characterize our candidate compounds fully.

We find that small changes in the compounds can lead to different combinations of activities. This observation was most pronounced in the comparison of filastatin with compound 27, which differ by the substitution of a methyl group for a chlorine substitution on one aryl ring (Fig. S3 and Table S1). Filastatin, but not compound 27, effectively inhibited human cell binding, but these two compounds shared the ability to inhibit polystyrene binding, hyphae formation, and biofilm formation (Table S1). As another example, the quinoline ring-containing compounds Q1 and Q2 do not inhibit polystyrene binding or human cell binding, but they do inhibit hyphae formation and biofilm formation. Conversely, filastatin is not synergistically toxic to fluconazole-resistant cells like Q1 and Q2 (Fig. S4). More systematic structure–activity relationship studies are underway to better understand the complex pharmacophore of filastatin and explore its molecular target(s). We note that the multiple activities of filastatin leave open the possibility that multiple targets exist.

Several previous experiments suggest a strong link between hyphal morphogenesis and fungal pathogenesis, including tissue-associated biofilm formation. For example, experiments with a doxycycline-inducible transcription factor (NRG1) that governs hyphae formation have shown that the ability to form hyphae is continually required for the lethality associated with systemic candidiasis (6). Likewise, many mutants defective in hyphae formation are nonpathogenic (7) and poor at biofilm formation, although this correlation is not absolute (38).

Our data show that filastatin blocks the transcriptional induction of the HWP1 promoter (Fig. S6), which is an early and essential event in the process of hyphal development (23).

Notably, filastatin blocks hyphal formation induced by serum, Spider media, and GlcNac but not the genotoxic stress agent HU (Fig. 4). These data indicate that the structural machinery required for cell elongation and hyphae formation is able to function in the presence of the compound. Therefore, our screen has provided a powerful chemical probe of the signaling mechanisms that govern hyphal formation. Biochemical and molecular genetics experiments are underway to better understand how filastatin affects hyphal development and fungal adhesion.

## 5.6 Methods

During the screen for small molecule adhesion inhibitors, one 5-mL culture of SC5314 cells and two 5-mL cultures of *edt1*<sup>-/-</sup> cells were inoculated with single colonies in yeast nitrogen broth without amino acids supplemented with 30 µg/mL L-leucine, 20 µg/mL L-histidine, 20 µg/mL L-tryptophan, and 25 µg/mL uridine (Supplemented YNB) + 0.15% glucose media and grown overnight with shaking at 200 rpm at 30 °C. The next day, these starter cultures were used to inoculate 100-mL cultures for each strain grown under the same conditions. One day later, OD at 600 nm was measured for both cultures. Cells were pelleted at 3,000 × g and resuspended in fresh Supplemented YNB + 0.15% glucose media to a final concentration of 0.5 OD/mL; 200 µL/well 0.5 OD/mL *edt1*<sup>-/-</sup> cells were added to the first column of each 96-well Immulon 2HB flat-bottom microtiter plate (Part No. 3455; Thermo Scientific), and 200 µL/well 0.5 OD/mL SC5314 cells were added to the remaining columns. Cell plating was followed by robotic addition of 2 µL DMSO vehicle to columns 1 and 12; compounds from the University of Massachusetts Medical School Small Molecule Facility DIVERset Library (Chembridge) at a stock concentration of 5 mM were added to columns 2–11 (Fig. 1A). The wells were mixed by robotic pipetting three times, yielding a final compound concentration of 50 µM. The plates were covered with foil and incubated at 37 °C for 4 h. The contents of the wells were then decanted, and 50 µL 0.5% crystal violet (Sigma) in water were added to each well. The plates were covered again and incubated at room temperature for 45 min. The dye was removed by decanting, and the plates were gently rinsed by 10 rounds of submersion in an ice bucket filled with distilled water followed by decanting the water. The water in the bucket was changed after the fifth wash. The plates were then gently inverted onto a paper towel to remove excess water; 200 µL 75% methanol were then added to each well. The

plates were incubated for 30 min at room temperature, and then, absorbance at 590 nm was measured.

Additional methods are in *SI Methods*.

## References

1. Miller LG, Hajjeh RA, Edwards JE, Jr. (2001) Estimating the cost of nosocomial candidemia in the united states. *Clin Infect Dis* 32(7):1110.
2. Pappas PG, et al. (2003) A prospective observational study of candidemia: Epidemiology, therapy, and influences on mortality in hospitalized adult and pediatric patients. *Clin Infect Dis* 37(5):634–643.
3. Fidel PL, Jr., Sobel JD (1996) Immunopathogenesis of recurrent vulvovaginal candidiasis. *Clin Microbiol Rev* 9(3):335–348.
4. Cowen LE, Anderson JB, Kohn LM (2002) Evolution of drug resistance in *Candida albicans*. *Annu Rev Microbiol* 56:139–165.
5. Sudbery P, Gow N, Berman J (2004) The distinct morphogenic states of *Candida albicans*. *Trends Microbiol* 12(7):317–324.
6. Saville SP, Lazzell AL, Monteagudo C, Lopez-Ribot JL (2003) Engineered control of cell morphology in vivo reveals distinct roles for yeast and filamentous forms of *Candida albicans* during infection. *Eukaryot Cell* 2(5):1053–1060.
7. Lo HJ, et al. (1997) Nonfilamentous *C. albicans* mutants are avirulent. *Cell* 90(5):939–949.
8. Finkel JS, Mitchell AP (2011) Genetic control of *Candida albicans* biofilm development. *Nat Rev Microbiol* 9(2):109–118.
9. Nobile CJ, Mitchell AP (2006) Genetics and genomics of *Candida albicans* biofilm formation. *Cell Microbiol* 8(9):1382–1391.
10. Bauters TG, Moerman M, Vermeersch H, Nelis HJ (2002) Colonization of voice prostheses by *albicans* and non-*albicans* *Candida* species. *Laryngoscope* 112(4):708–712.
11. Reynolds TB, Fink GR (2001) Bakers' yeast, a model for fungal biofilm formation. *Science* 291(5505):878–881.
12. Gillum AM, Tsay EY, Kirsch DR (1984) Isolation of the *Candida albicans* gene for or-otidine-5'-phosphate decarboxylase by complementation of *S. cerevisiae* *ura3* and *E.*



coli pyrF mutations. Mol Gen Genet 198(1):179–182.

13. Wheeler RT, Kombe D, Agarwala SD, Fink GR (2008) Dynamic, morphotype-specific *Candida albicans*  $\beta$ -glucan exposure during infection and drug treatment. PLoS Pathog 4(12):e1000227.
14. Kitamura A, et al. (2009) Effect of beta-1,6-glucan inhibitors on the invasion process of *Candida albicans*: Potential mechanism of their in vivo efficacy. Antimicrob Agents Chemother 53(9):3963–3971.
15. Junqueira JC, et al. (2011) Oral *Candida albicans* isolates from HIV-positive individuals have similar in vitro biofilm-forming ability and pathogenicity as invasive *Candida* isolates. BMC Microbiol 11:247.
16. Okoli I, et al. (2009) Identification of antifungal compounds active against *Candida albicans* using an improved high-throughput *Caenorhabditis elegans* assay. PLoS One 4(9):e7025.
17. Youngsaye W, et al. (2011) Piperazinyl quinolines as chemosensitizers to increase fluconazole susceptibility of *Candida albicans* clinical isolates. Bioorg Med Chem Lett 21(18):5502–5505.
18. Ganguly S, Mitchell AP (2011) Mucosal biofilms of *Candida albicans*. Curr Opin Microbiol 14(4):380–385.
19. Nobile CJ, et al. (2006) Critical role of Bcr1-dependent adhesins in *C. albicans* biofilm formation in vitro and in vivo. PLoS Pathog 2(7):e63.
20. Chauhan N, Kruppa MD (2009) Standard growth media and common techniques for use with *Candida albicans*. Methods Mol Biol 499:197–201.
21. Ganguly S, et al. (2011) Zap1 control of cell-cell signaling in *Candida albicans* biofilms. Eukaryot Cell 10(11):1448–1454.
22. Shareck J, Belhumeur P (2011) Modulation of morphogenesis in *Candida albicans* by various small molecules. Eukaryot Cell 10(8):1004–1012.
23. Lu Y, Su C, Wang A, Liu H (2011) Hyphal development in *Candida albicans* requires two temporally linked changes in promoter chromatin for initiation and maintenance. PLoS Biol 9(7):e1001105.
24. Nobile CJ, et al. (2012) A recently evolved transcriptional network controls biofilm development in *Candida albicans*. Cell 148(1-2):126–138.

25. Kadosh D, Johnson AD (2005) Induction of the *Candida albicans* filamentous growth program by relief of transcriptional repression: A genome-wide analysis. *Mol Biol Cell* 16(6):2903–2912.
26. Midkiff J, Borochoff-Porte N, White D, Johnson DI (2011) Small molecule inhibitors of the *Candida albicans* budded-to-hyphal transition act through multiple signaling pathways. *PLoS One* 6(9):e25395.
27. Homann OR, Dea J, Noble SM, Johnson AD (2009) A phenotypic profile of the *Candida albicans* regulatory network. *PLoS Genet* 5(12):e1000783.
28. Miwa T, et al. (2004) Gpr1, a putative G-protein-coupled receptor, regulates morphogenesis and hypha formation in the pathogenic fungus *Candida albicans*. *Eukaryot Cell* 3(4):919–931.
29. Feng Q, Summers E, Guo B, Fink G (1999) Ras signaling is required for serum-induced hyphal differentiation in *Candida albicans*. *J Bacteriol* 181(20):6339–6346.
30. Bockmühl DP, Ernst JF (2001) A potential phosphorylation site for an A-type kinase in the Efg1 regulator protein contributes to hyphal morphogenesis of *Candida albicans*. *Genetics* 157(4):1523–1530.
31. Shi QM, Wang YM, Zheng XD, Lee RT, Wang Y (2007) Critical role of DNA checkpoints in mediating genotoxic-stress-induced filamentous growth in *Candida albicans*. *Mol Biol Cell* 18(3):815–826.
32. Jain C, Pastor K, Gonzalez AY, Lorenz MC, Rao RP (2013) The role of *Candida albicans* AP-1 protein against host derived ROS in in vivo models of infection. *Virulence* 4(1):67–76.
33. Jain C, Yun M, Politz SM, Rao RP (2009) A pathogenesis assay using *Saccharomyces cerevisiae* and *Caenorhabditis elegans* reveals novel roles for yeast AP-1, Yap1, and host dual oxidase BLI-3 in fungal pathogenesis. *Eukaryot Cell* 8(8):1218–1227.
34. Hawser S, Islam K (1999) Comparisons of the effects of fungicidal and fungistatic antifungal agents on the morphogenetic transformation of *Candida albicans*. *J Antimicrob Chemother* 43(3):411–413.
35. Harriott MM, Lilly EA, Rodriguez TE, Fidel PL, Jr., Noverr MC (2010) *Candida albicans* forms biofilms on the vaginal mucosa. *Microbiology* 156(Pt 12):3635–3644.
36. Sobel JD, et al. (1998) Vulvovaginal candidiasis: Epidemiologic, diagnostic, and therapeutic considerations. *Am J Obstet Gynecol* 178(2):203–211.

37. Swinney DC, Anthony J (2011) How were new medicines discovered? *Nat Rev Drug Discov* 10(7):507–519.
38. Noble SM, French S, Kohn LA, Chen V, Johnson AD (2010) Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nat Genet* 42(7):590–598.

## 6 Studying evolution of drug resistance in *C. albicans* clinical isolates

### 6.1 Introduction and contribution to the paper

In this chapter we described how we used our nematode virulence model together with other phenotypic analyses as part of a large-scale collaborative study to explore the evolution of *C. albicans* drug resistance. 43 *C. albicans* clinical isolates from 11 HIV patients were obtained and extensively characterized both genotypically (Broad Institute and University of Minnesota) and phenotypically (WPI and Broad Institute). These isolates were obtained before and after fluconazole was prescribed as a therapy and some patients were sampled multiple times post fluconazole prescription. These samples represent powerful time series to study the progressive evolution of *C. albicans* drug resistance in humans.

The overall goal of the study was to identify genomic changes that alter *C. albicans* drug resistance and other phenotypes associated with virulence. We determined the ability of each of these strains to form biofilms by testing adhesion to polystyrene, to transition from yeast-to-hyphae and to infect and kill nematodes. These three phenotype are highly linked to virulence as the ability to form biofilm on medical devices, to transition from yeast-to-hyphae and to kill nematodes have been all connected to mammalian virulence (Chandra et al., 2001; Jain et al., 2013b; Lu et al., 2011). On the other hand, genomic analysis included genome deep sequencing to study the occurrence of single-nucleotide polymorphisms (SNPs), copy number variations (CNV) and loss of heterozygosity (LOH) events.

The study shows that LOH events were recurrent and tracked with increased in drug resistance while aneuploidies were transient and didn't correlate with fluconazole resistance.

The study also identified 240 genes that accumulate persistent SNPs that recur between patients, suggesting that these genes might play a primary role in the complex process of host adaptation and fluconazole exposure. Although we found substantial variations in many of the phenotypes observed between isolates, we found that (a) *C. albicans in vitro* fitness is anti-correlated with phenotypic markers of virulence (b) a progressive increase in drug resistance *in vitro* is mirrored by a progressively increased virulence. Overall our combined data supports the notion that the process of *C. albicans* adaptation to the host during antimycotic therapy is complex and multifactorial and that its ability to adapt to the human environment relies on the *C. albicans* remarkable genotypic and phenotypic plasticity.

Our contribution to the paper is described in the narrative of the paper (paragraph titled “Changes in virulence phenotypes in evolved drug-resistant isolates”) and summarized in Figure 6. The paper has been submitted to the journal *eLIFE* and is in the process of resubmission after many of the reviewers comments has been addressed. The paper reported below represented the most recent version (August 8<sup>th</sup> 2014).

## 6.2 The evolution of gradual acquisition of drug resistance in clinical isolates of *Candida albicans*

Christopher B. Ford<sup>#1</sup>, Jason M. Funt<sup>#1,2</sup>, Darren Abbey<sup>§3</sup>, Luca Issi<sup>§4</sup>, Candace Guiducci, Diego A. Martinez, Toni Delorey, Theodore C. White<sup>6</sup>, Christina Cuomo, Reeta P. Rao<sup>4</sup>, Judith Berman<sup>3,7</sup>, Dawn A. Thompson<sup>1,\*</sup>, and Aviv Regev<sup>1,2,\*</sup>

<sup>1</sup>Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA 02142

<sup>2</sup>Howard Hughes Medical Institute, Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Ave, Cambridge, MA 02140

<sup>3</sup>University of Minnesota, Minneapolis MN 55455 USA

<sup>4</sup> Worcester Polytechnic Institute, Department of Biology and Biotechnology, 100 Institute Road, Worcester MA 01609

<sup>5</sup>School of Biological Sciences, University of Missouri at Kansas City, MS

<sup>6</sup>Tel Aviv University, Ramat Aviv, 69978 Israel

<sup>#</sup>These authors contributed equally to this work

<sup>§</sup> These authors contributed equally to this work

\* To whom correspondence should be addressed. Email: [dawnt@broadinstitute.org](mailto:dawnt@broadinstitute.org) (DAT), [aregev@broad.mit.edu](mailto:aregev@broad.mit.edu) (AR)

### 6.3 Introduction

Virtually all humans are colonized with *Candida albicans*, but in some individuals this benign commensal organism becomes a serious, life-threatening pathogen. *C. albicans* possesses an arsenal of traits that promote its pathogenicity, including phenotypic switching (Alby and Bennett, 2009), yeast-hyphae transition (Kumamoto and Vences, 2005) and the secretion of molecules that promote adhesion to abiotic surfaces (Chandra et al., 2001). As a commensal, an intricate balance is maintained between the ability of *C. albicans* to invade host tissues and the host defense mechanisms (Kim and Sudbery, 2011; Kumamoto and Pierce, 2011). Alteration of this delicate host-fungus balance can result in high levels of patient mortality (Charles et al., 2003; Pittet et al., 1994): systemic *C. albicans* infections are fatal in 42% of cases (Wisplinghoff et al., 2003), despite the use of antifungal therapies, and *C. albicans* is the fourth most common infection in hospitals (Gudlaugsson et al., 2003b; Pappas et al., 2003). While compromised immune function contributes to pathogenesis (Berman and Sudbery, 2002), it is less clear how *C. albicans* evolves to better exploit the host environment during the course of infection.

Two classes of antifungals in clinical use target ergosterol, a major component of the fungal cell membrane: polyenes and azoles. Polyenes (e.g., Amphotericin B) can cause high levels of morbidity when used at fungicidal levels and thus are used sparingly (Rex et al., 1994), whereas azoles (e.g., fluconazole) are used widely because they can be administered orally, and have few side effects (Rex et al., 2003). However, resistance to the azoles arises within the commensal population of the treated individual, primarily because azoles are fungistatic (inhibit growth but do not kill) (Cowen et al., 2002). Indeed, epidemiological data suggest that the intensity of fluconazole use is driving the appearance of resistant isolates (Pfaller et al., 1998). Studies of clinical isolates *C. albicans* suggest drug-resistance increases during an infection through the acquisition of aneuploidies (Selmecki et al., 2009) suggesting genomic plasticity and rapid evolutionary selection during infection.

Previous studies have identified two main classes of azole resistance in *C. albicans*. First, increased activity or level of the enzymes of the ergosterol pathway (e.g., *ERG11*) (Asai et al., 1999; Oliver et al., 2007) reduces the direct effect on the drug target. Second, increased efflux of the drug from cells by ABC transporters (encoded by *CDR1* and *CDR2*) (Coste et al., 2006) or by the major facilitator superfamily efflux pump (encoded by *MDR1*) (Dunkel et

al., 2008) reduces the effective intracellular drug concentration. In both cases, such alterations can result from point mutations in genes encoding the proteins (Marichal et al., 1999), in transcription factors that regulate them (Coste et al., 2006; Dunkel et al., 2008; MacPherson et al., 2005), or from increased copy number of the relevant genes, via genome rearrangements such as whole chromosome and segmental aneuploidies (Selmecki et al., 2006; Selmecki et al., 2008; Selmecki et al., 2009). Indeed, the genomes of drug resistant strains isolated following clinical treatment often exhibit large-scale changes, such as loss of heterozygosity (LOH)(Coste et al., 2006; Dunkel and Morschhauser, 2011), copy number variation (CNV), including short segmental CNV, and whole chromosome aneuploidy (Selmecki et al., 2010b) accompanied by point mutations.

While we understand some aspects of the molecular basis of resistance, we understand less about the dynamic mechanisms that drive the evolution of drug resistance and overall pathogenicity in *C. albicans*. The diploid genome and lack of a complete sexual cycle have complicated efforts to use forward genetic approaches. An alternative approach is to use isolates sampled consecutively from the same patient to study the *in vivo* evolution of drug resistance and pathogenicity. Studies in evolved isolates have implicated multiple mechanisms in drug resistance, but have focused on large-scale aberrations (Selmecki et al., 2008; Selmecki et al., 2009) or candidate genes (Perea et al., 2001a; White et al., 2002), and do not comprehensively chart the genetic basis of adaptation.

Here, we used genome sequencing of isolates sampled consecutively from patients that were clinically treated with fluconazole to systematically analyze the genetic dynamics that accompany the appearance of drug resistance during oral candidiasis in human HIV patients. Most isolates from each individual patient were highly related, suggesting a clonal population structure and facilitating the identification of variation. As each clinical sample was colony purified, we cannot assess the population structure at any single time point; therefore, we focus on the occurrence of single-nucleotide polymorphisms (SNPs), CNV and LOH events in each isolate. Consistent with previous studies, we found that LOH events were recurrent and associated with increased drug resistance. To identify SNPs with likely functional impact in the context of substantial genetic diversity, we focused on those that were both persistent within a patient and recurrent across multiple patients. We found 240 genes that recurrently contain persistent SNPs, many of which may be related to the complex process of adaptation to the host and antifungal exposure. In contrast, aneuploidies were prevalent in the isolates,



yet they were more likely to be transient, and aneuploidy, *per se*, did not correlate with changes in drug resistance. Taken together, our study suggests new molecular mechanisms of emergent drug resistance and virulence in *C. albicans*, and provides a general model for such studies in other eukaryotic pathogens.

## 6.4 Results

### Whole genome sequencing of 43 serial clinical isolates from 11 patients

To study the *in vivo* evolution of azole resistance in *C. albicans*, we analyzed 43 longitudinal isolates from 11 HIV-infected patients with oropharyngeal candidiasis (Perea et al., 2001a; White, 1997b) (Table 1).

Publication Name	Patient	Strain	Entry Date	Drug Treatment	Dose (mg/day)	E-test MIC (ug/mL)	Depth of Coverage	Reads	Percent Aligned
White, T.C.	1	1	9/10/1990	Fluconazole	100	0.25	111.9610	9,896,468	87.17%
		2	12/14/1990	Fluconazole	100	1	69.2023	12,797,328	87.43%
		3	12/21/1990	Fluconazole	100	4	92.0415	16,987,814	86.87%
		4	12/31/1990	Fluconazole	100	3	80.6876	14,858,710	87.81%
		5	2/8/1991	Fluconazole	100	4	110.7990	20,484,584	86.75%
		6	2/22/1991	Fluconazole	100	4	101.9360	18,837,954	86.63%
		7	3/25/1991	Fluconazole	100	4	81.6485	15,123,020	86.66%
		8	4/8/1991	Fluconazole	100	4	112.5280	20,778,562	86.64%
		9	6/4/1991	Fluconazole	100	4	113.1840	22,223,228	83.20%
		11	7/15/1991	Fluconazole	100	4	53.2780	9,896,468	87.17%
		12	11/26/1991	Fluconazole	200	4	96.1041	18,282,472	85.54%
		13	12/13/1991	Fluconazole	400	32	123.6710	22,070,518	89.13%
		14	1/28/1992	Fluconazole	400	24	98.6643	18,114,916	87.41%
		15	2/21/1992	Clotrimazole	50	24	120.9000	22,401,374	86.57%
		16	4/1/1992	Fluconazole	400	96	87.4437	16,061,560	87.17%
		17	8/25/1992	Fluconazole	800	96	97.8316	18,317,118	85.91%
		Perea, S. et al.	7	412	2/15/1995	Fluconazole	0	0.25	93.1525
2307	11/22/1995			Fluconazole	400	0.75	95.7885	18,014,242	85.25%
Perea, S. et al.	9	1002	4/20/1995	Fluconazole	100	0.125	188.4890	34,834,970	86.74%
		2823	4/6/1996	Fluconazole	800		282.6230	52,839,288	86.30%
		3795	2/26/1997	Fluconazole	800	128	77.6274	13,901,062	88.78%
Perea, S. et al.	14	580	3/13/1995	Fluconazole	0	1.5	77.0755	14,711,804	85.00%
		2440	1/3/1996	Fluconazole	800	1.5	82.9294	15,446,882	85.69%
		2501*	1/4/1996	Fluconazole	800	96	88.5917	17,480,274	81.98%
Perea, S. et al.	15	945	4/14/1995	Fluconazole	300	4	108.5880	20,591,044	85.19%
		1619	7/11/1995	Fluconazole	500	64	93.1423	17,565,080	84.69%
Perea, S. et al.	16	3107	6/5/1996	Fluconazole	800	4	97.0099	18,361,266	84.84%
		3119	6/5/1996	Fluconazole	800	96	87.9159	16,615,462	84.67%
		3120	6/5/1996	Fluconazole	800	96	105.9450	19,442,016	86.79%
		3184	7/1/1996	Fluconazole	800		101.8920	18,487,462	87.50%
		3281	7/16/1996	Fluconazole	800		76.4442	14,327,376	85.69%
Perea, S. et al.	30	5106	1/7/1998	Fluconazole	800	0.5	87.2055	16,466,524	84.67%
		5108	1/7/1998	Fluconazole	800	0.75	82.3160	17,480,274	81.98%
Perea, S. et al.	42	1691	8/3/1995	Fluconazole	100		122.5980	22,072,562	88.38%
		3731	12/27/1996	Fluconazole	400	256	119.8970	21,436,034	88.72%
		3733	12/27/1996	Fluconazole	400	256	95.5084	17,295,888	88.00%
Perea, S. et al.	43	1649	7/19/1995	Fluconazole	0	0.125	102.0970	19,545,530	84.08%
		3034	5/15/1996	Fluconazole	400	0.75	92.9746	17,300,040	85.64%
Perea, S. et al.	59	3917	2/19/1997	Fluconazole	800	2	113.2650	21,549,704	83.86%
		4617	8/28/1997	Fluconazole	400	64	75.3678	15,242,904	81.42%
		4639	9/2/1997	Fluconazole	400	128	115.3160	25,468,190	75.69%
Perea, S. et al.	64	4018	4/2/1997	Fluconazole	200		110.1550	20,118,736	86.78%
		4380	7/14/1997	Fluconazole	200		18.0328	20,970,946	9.26%

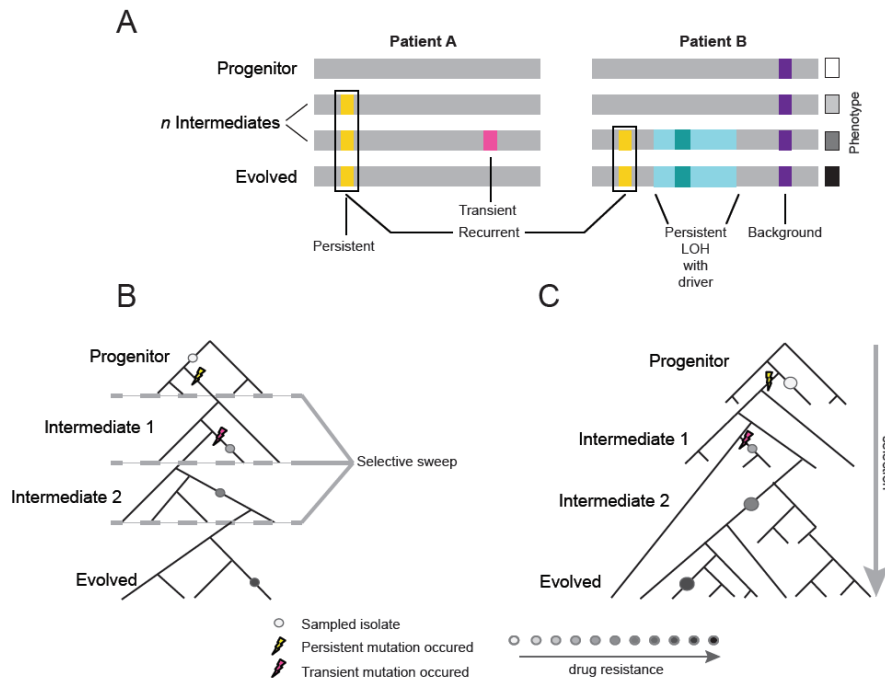
Not clonally derived from progenitor

\* isolated on same day from same patient as previously published strain, 2500

**Table 1** *C. albicans* clinical isolates used in this study

The isolates were collected during incidences of infection and form a time series from each patient (2-16 isolates per series; **Figure 1, Figure 2A**). Each isolate was derived from a single colony and thus represents a single diploid genotype sampled from the within-host *C. albicans* population at the respective time point. In each series, the first isolate ('progenitor') was collected prior to any treatment with azole antifungals and the remaining isolates were

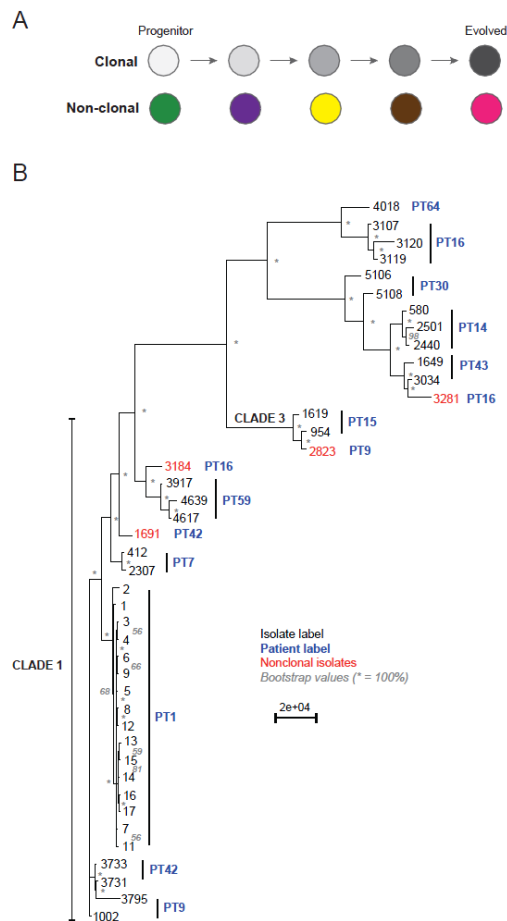
collected at later, typically consecutive, time points, culminating in the final ‘endpoint’ isolate (**Table 1**).



### Figure 1 Overview of study design

(A) Background, persistent, transient, recurrent, and driver mutations in patient time courses. Shown is a schematic illustration of the genomes of isolates (gray bars) from two patient time courses (Patient A and B, left and right panels, respectively), ordered from the first isolate (progenitor, top) to the last (evolved, bottom). Background mutations (purple) exist in the progenitor; persistent mutations (yellow) are not in the progenitor, but found in all subsequent isolates; transient mutations (pink) are not in the progenitor and only in some later isolates; recurrent mutations are persistent mutations that occur in the same gene in more than one patient. LOH events were also evaluated for persistence (light teal bar). Driver mutations, where a new persistent homozygous allele appears (*e.g.* G/T > A/A) are annotated in association with persistent LOH events (dark teal) and independent of these events (not shown). Each of these can be associated with a change in phenotype, such as drug resistance (boxes, right). (B) Sampling in the context of *de novo* mutation and selection bottlenecks. Each strain is a single clone (circle) isolated from an evolving population (represented by a phylogenetic tree). The population evolves and undergoes selective sweeps (dashed lines), with phenotypic changes occurring during the course of infection and treatment (*i.e.* drug resistance, black: high, white: low; gray scale at bottom). Persistent mutations (yellow lightning bolt) have likely swept through the population, whereas transient mutations (pink lightning bolt) have not. (C) Sampling in the context of selection on existing variation. Selection acts to vary the frequency of different pre-existing genotypes in the population. Persistent mutations (yellow lightning bolt) have risen in the population to a frequency that they are repeatedly sampled (large circles) whereas transient mutations (pink lightning bolt) have not (small circle).

The progenitor isolates were more sensitive to fluconazole than subsequent isolates, as defined by the minimum inhibitory concentration (MIC) (Table 1, Methods). Previous studies with some of these patient isolates identified several mechanisms that may contribute to azole resistance, including segmental aneuploidy (Selmecki et al., 2006), increased expression of drug efflux genes (Coste et al., 2006), loss of heterozygosity (LOH) across large chromosomal segments (Coste et al., 2006; Dunkel et al., 2008), mutations in ergosterol biosynthetic genes (Asai et al., 1999; Oliver et al., 2007), and facilitation by the chaperone heat shock protein 90 (Hsp90) (Cowen and Lindquist, 2005).



**Figure 2: Most isolates from the same patient are clonal.** (A) Two possible models of infection may underlie serial isolates. In the ‘clonal model’ (top) each subsequent sample (circle) is related to the other isolates. In the non-clonal model (bottom) isolates in a series are un-related. (B) Determining clonal infections from SNP calls. The phylogenetic relationship of the isolates (black) from 11 patients (blue) was inferred based 201,793 informative SNP positions using maximum parsimony in PAUP\*. Isolates from the same patient separated by a branch distance greater than 20,000 were considered non-clonal (3281, 2823, 3184, 1691, red). Most nodes were supported by 100% of 1,000 bootstrap replicates (indicated by \*), except as indicated (in grey). Clade identifiers were included as appropriate.

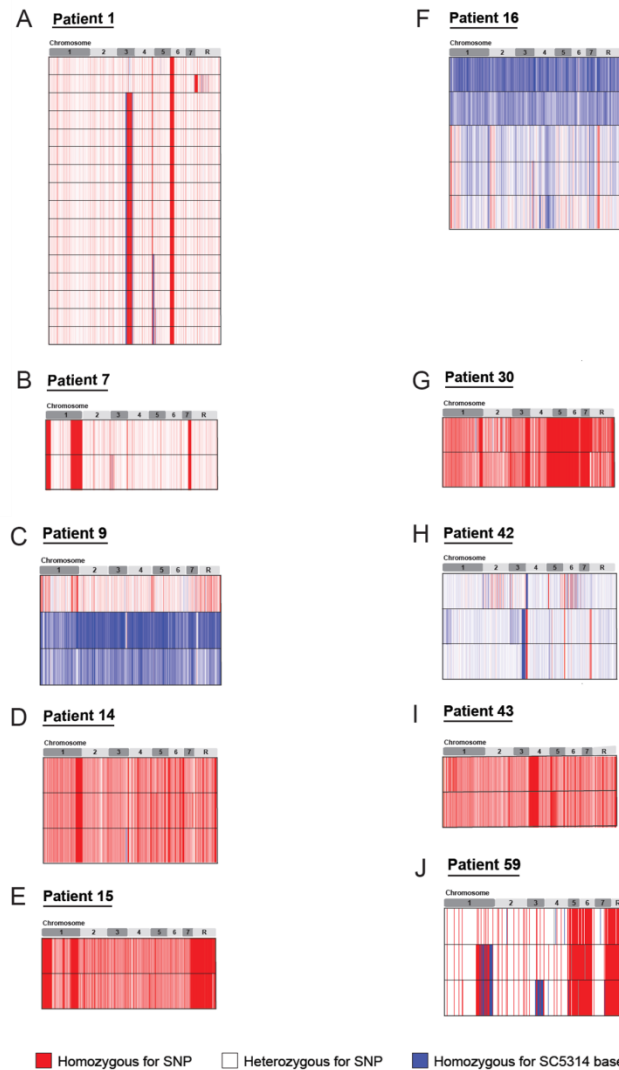
We sequenced the genomic DNA of each of the isolates as well as the *C. albicans* lab strain, SC5314, using Illumina sequencing (average coverage, 100X, **Methods, Table 1**) and identified in each series point mutations, LOH events and aneuploidies that were not present in the progenitor of that series. All mutations were defined relative to SC5314, the *C. albicans* genome reference strain. We validated our pipeline for detection of point mutations using the Sequenom iPlex genotyping technology (**Methods**, (Storm et al., 2003)). We interrogated 2499 randomly chosen SNPs in 27 isolates from 9 clinical series, and found that the iPlex base calls matched 2270 (90.8%, Table 2) of the calls from the computational pipeline used to analyze the sequencing data.

### **Most series are clonal, but there is significant genetic diversity between isolates**

We designated polymorphisms common to all isolates in a series, including the progenitor (Figure 1, purple) as background polymorphisms. Based on background SNPs, we first determined that most isolates within most series were clonal, suggesting a single (primary) infection source (Figure 1B, Figure 2, Figure 3, Methods). To distinguish between a single primary infection (Figure 2A, top) and repeated, independent infections (Figure 2A, bottom) we determined the distance between every two isolates based on their SNP profile, and used as a heuristic a neighbor-joining algorithm to construct a phylogenetic tree from this distance metric, with the reference strain SC5314 as an outgroup for rooting (Methods, Figure 2B). All but one (10/11) of the series were predominantly clonal (Figure 2B, Figure 2-figure supplement 1). Nevertheless, we detected at least one non-clonal isolate in 4/11 patient series, indicating that at least ~36% of patients carried more than one unrelated strain of *C. albicans*. We removed non-clonal samples from further consideration, and all subsequent analyses focused on samples from the 10 patients with at least two clonal isolates.

Despite these clonal relationships, the distance between isolates indicated significant genetic diversity *within* each patient series (**Figure 2B**), typically with each isolate differing by several thousand SNPs from its ‘progenitor’ isolate (**Figure 2 - source data 1 and 2**). These data could reflect two evolutionary scenarios: accumulation of *de novo* mutations followed by selection (**Figure 1B**), or largely pre-existing variation with selection acting to vary the frequency of different genotypes in the population (**Figure 1C**). The large number of SNPs detected suggests that isolates from later time points in a series are not simply direct

descendants of the earlier isolate; however, since mutation and mitotic recombination rates can be elevated under stressful conditions (e.g. drug treatment (Forche et al., 2011; Galhardo et al., 2007)), we cannot rule out the possibility that some of the variation may be due to *de novo* events. Formally distinguishing between these two models is not possible with the data at hand. However, the observation that different isolates collected on the same day from patient 14 (2440 and 2501) and patient 16 (3107 and 3119) differed by 9,668 and 18,291 SNPs, respectively (**Figure 2 - source data 1 and 2**), and had very different fluconazole MIC levels (**Table 1**) and different fitness phenotypes (**Figure 6, 7, and 8**) supporting the role of standing diversity. Thus, we conclude that a population of related but divergent genotypes of the same lineage exists within a given patient. We next sought to identify potentially adaptive genetic changes by focusing on large-scale events (LOH and aneuploidies) as well as single nucleotide polymorphisms.



**Figure 2 - figure supplement: SNP heterozygosity profiles for each strain.** The heterozygosity profiles shows, in chromosomal order (top), each variant locus that exists in at least one strain in the series (white is a heterozygous SNP, blue is homozygous for the SC5314 allele, red is a homozygous SNP relative to SC5314). (**A**) Patient 1; (**B**) Patient 7; (**C**) Patient 9; (**D**) Patient 14; (**E**) Patient 15; (**F**) Patient 16; (**G**) Patient 30; (**H**) Patient 42; (**I**) Patient 43; (**J**) Patient 59. Only Patient 9 (C), Patient 16 (F), Patient 42 (H), and Patient 64 (*not shown*\*\*\*) contain un-related isolates. \*\* Patient 64 contained an isolate (4380) whose genome aligned poorly to the *C. albicans* reference, but aligned well to *C. dubliniensis*.

### **Genetic alterations absent from the progenitor, persistent within a patient, and recurrent across patients are likely adaptive**

Given the high number of SNPs, LOH events and aneuploidies, we next devised a strategy to identify those changes that are likely to play an adaptive role in drug resistance and host adaptation. Focusing on genetic alterations absent from the (drug sensitive) progenitor but present in later isolates, we defined alterations as *persistent* if present within the same patient at all subsequent time points after the isolate in which they are first identified. We reasoned that such persistent changes will include those variants that were driven to sufficiently high frequency by selection to ensure repeated sampling (**Figure 1B** and **1C**, yellow lightning bolt), whereas non-persistent (transient) ones do not (**Figure 1B** and **1C**, pink lightning bolt). We consider the special case of a genetic change detected only in the endpoint strain as ‘persistent’ as well, since several of the time courses consist of only two or three isolates. We apply the persistence filter to better identify potentially adaptive aneuploidies, LOH events, and SNPs.

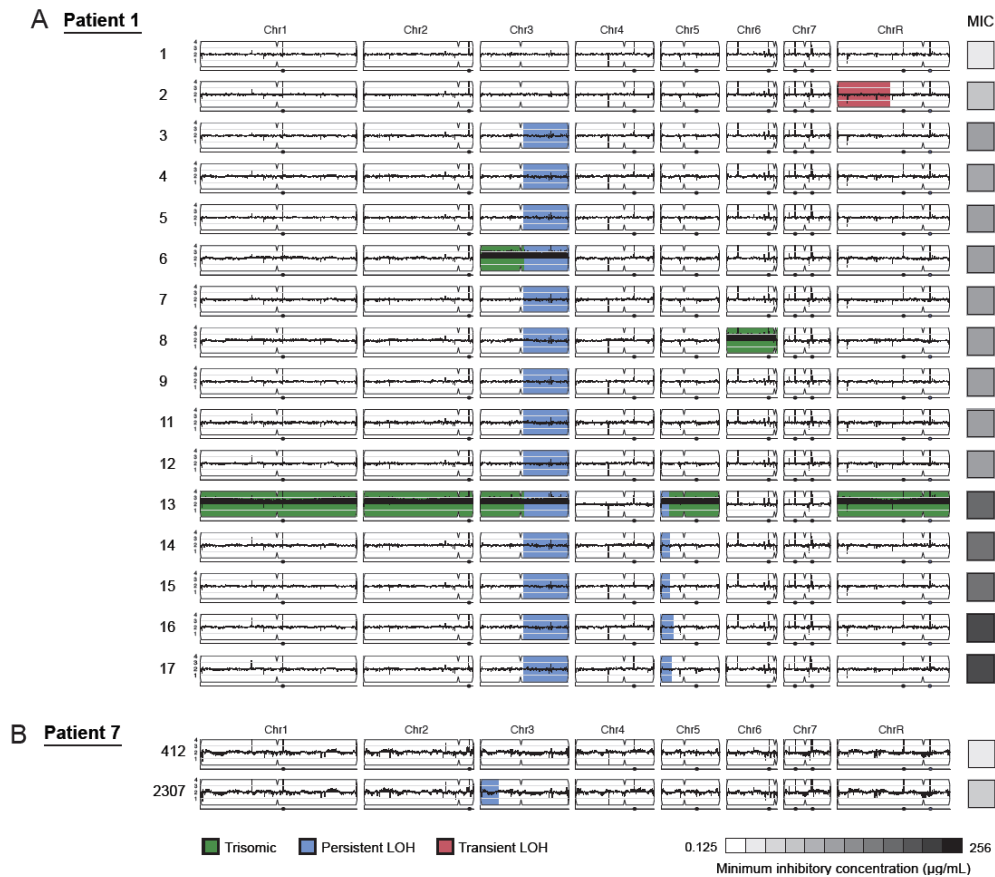
The large number of SNPs detected required additional filtering; therefore, to further restrict our analysis to potentially adaptive SNPs, we first excluded all *background* mutations, defined as any SNP relative to the reference present in all isolates from a series. Next, we focused on nonsynonymous polymorphisms in coding regions and employed two different strategies to identify potentially adaptive changes. First, to identify potential *drivers* of adaptation, we focused on non-synonymous SNPs that were homozygous for a genotype not found in the progenitor strain that persisted in the subsequent isolates (e.g. G/T > A/A) consistent with positive selection. In the second strategy, polymorphic genes were considered *recurrent* across patients when persistent, nonsynonymous polymorphisms appeared within the same open reading frame (ORF) in different patient series (**Figure 1**, **Figure 5 - source data 1**). We considered only those that were not included in LOH regions, as these regions artificially inflate the estimates of persistence and recurrence. Recurrence allows us to better handle polymorphisms from the endpoint isolate in a series for which ‘persistence’ does not provide a meaningful filter. Thus, polymorphisms only in the terminal isolate in one patient will only be further considered if polymorphisms also recurred in the same ORF in a series from two other patients. For example, filtering for both persistence and recurrence across at least three series brought the number of polymorphisms for patient 1



from 13,562 polymorphisms to 23 recurrent genes (**Figure 2 - source data 1 and 2, Figure 5 - source data 1**).

### **LOH events are commonly associated with increased resistance**

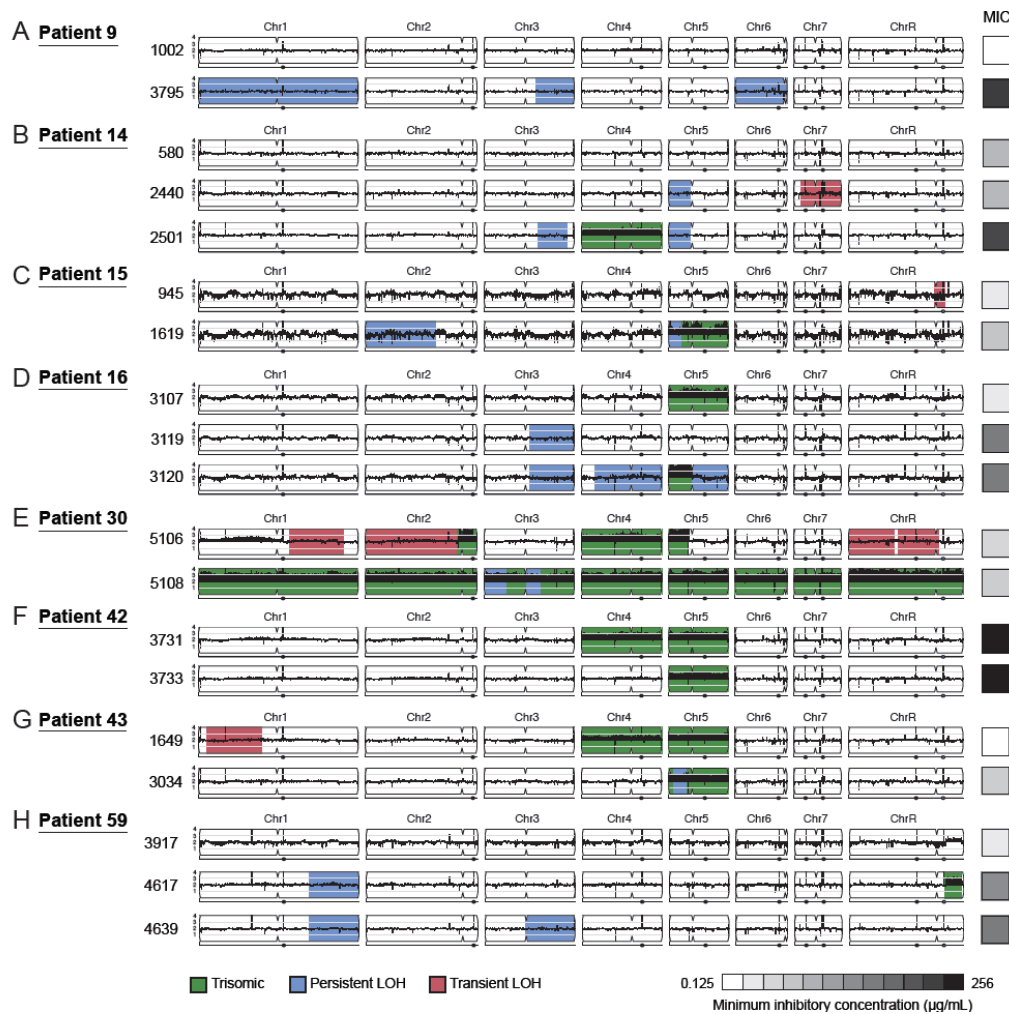
LOH events were detected in the vast majority (9/10) of the series, and were often persistent, recurrent and associated with increased drug resistance (**Figure 3 and 4, Figure 3 and 4 - source data 1**). For example, two of three LOH events in Patient 1 were persistent and associated with an increase in MIC and both of these events were recurrent, such that LOH events in these genomic regions coincided with increases in MIC in other patients. Highly recurrent LOH events occurred on the right arm of chromosome 3 (in Patients 1, 9, 14, 16, 30, and 59; **Figure 3A,B, Figure 4B,C, Figure 3 and 4- source data 1**) and on the left arm of chromosome 5 (in Patients 1, 14, 15, and 43; **Figure 3A, Figure 4C,H, Figure 3 and 4 - source data 1**). These regions include key genes implicated in drug resistance: on Chromosome 3, genes encoding the Cdr1 and Cdr2 efflux pumps and the Mrr1 transcription factor that regulates the Mdr1 major facilitator superfamily efflux pump (Schubert et al., 2011), and on Chromosome 5 genes encoding the drug target Erg11, and Tac1, a transcription factor that positively regulates expression of *CDR1* and *CDR2* (Coste et al., 2006). The extent of persistence and recurrence of these two LOH events is statistically significant under a naïve binary model ( $p < 0.009$  for the Chr3R LOH;  $p < 0.009$  for the Chr5L LOH). The recurrence of LOH events that coincide with changes in MIC suggests that they have been positively selected to rise in frequency relative to the progenitor strain. Notably, some of the recurrent LOH events may have been difficult to detect previously on SNP arrays (Forche et al., 2008; Forche et al., 2004; Forche et al., 2005), due to the relative paucity of SNPs in those regions in the reference strain, SC5413, itself a clinical isolate.



### Figure 3 LOH events were often persistent while aneuploidies were often transient

For each time series shown are the genomes of all isolates (rows) from a patient, ordered from the first isolate (progenitor, top) to the last (evolved, bottom). Boxes on right indicate the MIC of the respective strain (black: high, white: low; gray scale at bottom). Persistent LOHs: blue, transient LOHs: pink; trisomies (all transient): green. The sequence coverage along each chromosome is indicated by black tickmarks. **(A)** Patient 1 has four LOH events, each coinciding with an increase in MIC (gray scale boxes, right). One LOH is transient (isolate 2, chromosome R, pink), and three are persistent (isolate 3, chromosome 3; isolate 13, chromosome 5; and isolate 16, chromosome 5, blue). The ploidy changes (isolates 6, 8, 13) are all transient. **(B)** Patient 7 has one LOH event (isolate 2307, chromosome 3, blue) which coincides with an increase in MIC.

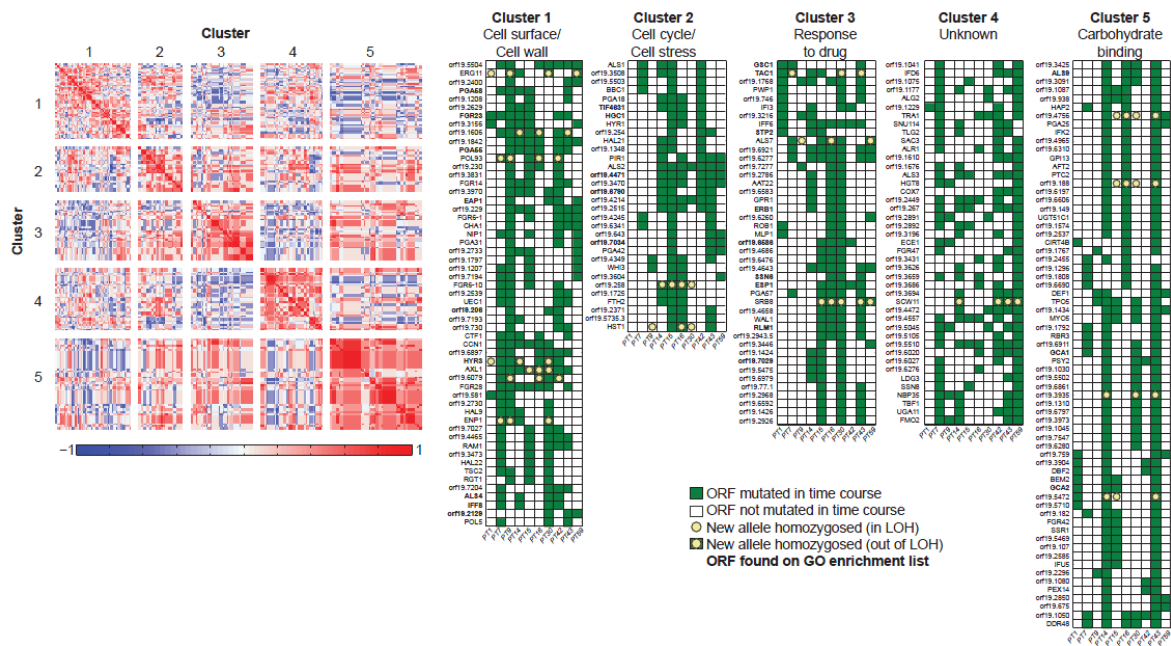
Driver mutations (e.g.  $G/T > A/A$ ) in these regions are suggestive of a point mutation followed by an LOH of the mutant allele that confers an advantage. There were 131 such mutations in 86 ORFs from 18 LOH regions from the 10 clonal patient series (**Table 1 and Figure 5 Figure 5 - source data 3**).



**Figure 4 Persistent and transient LOH and aneuploidies.** For each time series shown are the genomes of all isolates (rows), ordered from the first isolate (progenitor, top) to the last (evolved, bottom). Boxes on right indicate the MIC of the respective strain (black: high, white: low, grey scale at bottom). Persistent LOHs: light blue, transient LOHs: pink; trisomies (all transient): green. The coverage along each chromosome is indicated by black tickmarks. (A) Patient 9; (B) Patient 14; (C) Patient 15; (D) Patient 16; (E) Patient 30; (F) Patient 42; (G) Patient 43; (H) Patient 59. Several LOHs are recurrent (right arm of chromosome 3, left arm of chromosome 5, and chromosome 1). We also observe i5L in patients 16, isolate 3120 (D) and patient 30, isolate 5106 (E), but neither coincides with an increase in MIC.

Some of the SNPs were in genes that encode key players in drug resistance and were associated with large LOH events. For example, a nonsynonymous homozygous change in the fluconazole drug target *ERG11* was associated with the formation of the persistent LOH on the left arm of chromosome 5 in Patient 1 (Figure 3A), consistent with previous reports (White, 1997c) as was *TAC1* in Patient 42. In another example, the persistent and recurrent

LOH on the right arm of chromosome 3 in Patients 9 and 16 was associated with the appearance of a homozygous mutation in *MRR1* (Schubert et al., 2011), a regulator of *MDR1* expression. Other mutations were in genes not previously related to fluconazole resistance, including cell adhesion (*ALS3,5* and *7*, and *HYR3*, (Hoyer et al., 1998) (Sheppard et al., 2004) (Hoyer et al., 2008) , filamentous growth (*FGR14, FGR28* and *EFH1*, (Connolly et al., 2013; Uhl et al., 2003) and biofilm formation (*BCR1* and *YAK1*, (Goyard et al., 2008; Nobile et al., 2012; Nobile and Mitchell, 2005b; Noble et al., 2010b).



**Figure 5 Co-occurrence of nonsynonymous substitutions across isolates reveals functional clusters.** For each of the recurrently mutated 240 genes (genes in which nonsynonymous persistent SNPs appear in more than 3 patients and are not within an LOH region), we constructed a patient-by-gene binary vector (shown as cluster membership matrices, right, green: gene is persistently mutated in patient per definition above). We clustered the resulting patient by gene matrix using NMF clustering to reveal 5 coherent clusters (correlation matrix of the clusters left; red: positive correlation; blue: negative correlation; white: no correlation). Functional enrichment of clusters was revealed using gene ontology, and genes matching the enriched cluster function are bolded. We have overlaid recurrent driver mutations (e.g. G/T > A/A) (n = 17) occurring outside of LOH regions (yellow circle, green box) and inside LOH regions (yellow circle, white box).

### **Aneuploidies are not predictive of MIC, but may facilitate the appearance of drug resistance**

Aneuploidies, either whole chromosome or segmental, were evident in at least one isolate from 80% (8/10) of the clonal patient series, with the most prevalent aneuploidies involving Chromosome 5 (6 of 8 patients with at least one aneuploid isolate). In contrast to the persistent nature of LOH events, persistent aneuploidies were more rare, and were not consistently associated with adaptive increases in MIC levels (**Figure 3 and 4**).

While we cannot definitively infer an ordering of events from our singly sampled isolates, we hypothesize that aneuploidy contributes to the evolution of LOH by increasing the likelihood of its occurrence. For example, in four of the six patients with a Chromosome 5 LOH, the LOH event is preceded by or is concomitant with Chromosome 5 trisomy. Thus, the additional copy may increase the likelihood of an LOH event on that chromosome. In three of these cases, *ERG11*, located on the region of Chromosome 5 with LOH, was mutated. Additionally, isolates in two of the seven patients with a Chromosome 3 LOH were trisomic for this chromosome.

### **Persistent SNPs in 240 recurrently polymorphic genes identify targets likely associated with and drug resistance and host adaptation.**

We identified persistent, nonsynonymous coding SNPs within 1,471 genes that fell outside LOH tracts. Among these we identified 338 driver polymorphisms in 168 genes (**Figure 5-source data 2**). These again include *ERG11* in patient 9, 14, 30, and 59 and *TAC1* in patients 1, 7, 14, 15, 30 and 43 (**Figure 5**). Applying the recurrence filter (*i.e.*, persistent nonsynonymous SNPs that appeared in the same ORFs in 3 or more patient series to the list of 1,471 genes), we identified 240 polymorphic genes that are more likely to have contributed to adaptation (**Figure 5 -source data 1**). This number of genes is more likely than expected by chance ( $p < 0.0001$ , by simulation). Though the coding sequence for these 240 recurrent genes is longer than average, and thus a larger target for mutation, ( $2.21 \pm 1.53\text{kb}$  vs.  $1.83 \pm 1.29\text{kb}$  for non-recurrent persistent genes,  $p < 3.68 \times 10^{-5}$ , *t*-test), our simulation accounts for gene length, suggesting we see significantly more recurrence than expected by chance alone. Notably, we identified 17 persistent recurrently polymorphic genes that were also drivers (8 of which were also homozygosed in an LOH tract) (**Figure 5, Figure 5 – source data 1, 2,**

**and 3**). Finally, polymorphisms in 166 of the 240 genes appeared together with and increase in MIC and are thus stronger candidates for making a significant functional contribution to resistance (**Figure 5- figure supplement 1, Figure 5 – source data 4**,  $p < 0.00001$ , by simulation).

The set of 240 recurrently mutated genes was enriched for fungal-type cell wall and cell surface genes, including several members in each of three cell wall gene families important for biofilm formation and virulence: the Hyr/Iff proteins (HYR1,3 and IFF8,6); the Als adhesins (ALS1-4,7,9); the PGA-30-like proteins (7 genes) (**Figure 5- source data 1** (Hoyer et al., 2008)). All three gene families have previously shown to be specifically enriched in pathogenic *Candida* species (Butler et al., 2009). There are also seven members of the FGR gene family, involved in filamentous growth that is specifically expanded in *C. albicans* (**Figure 5- source data 1**, (Butler et al., 2009)).

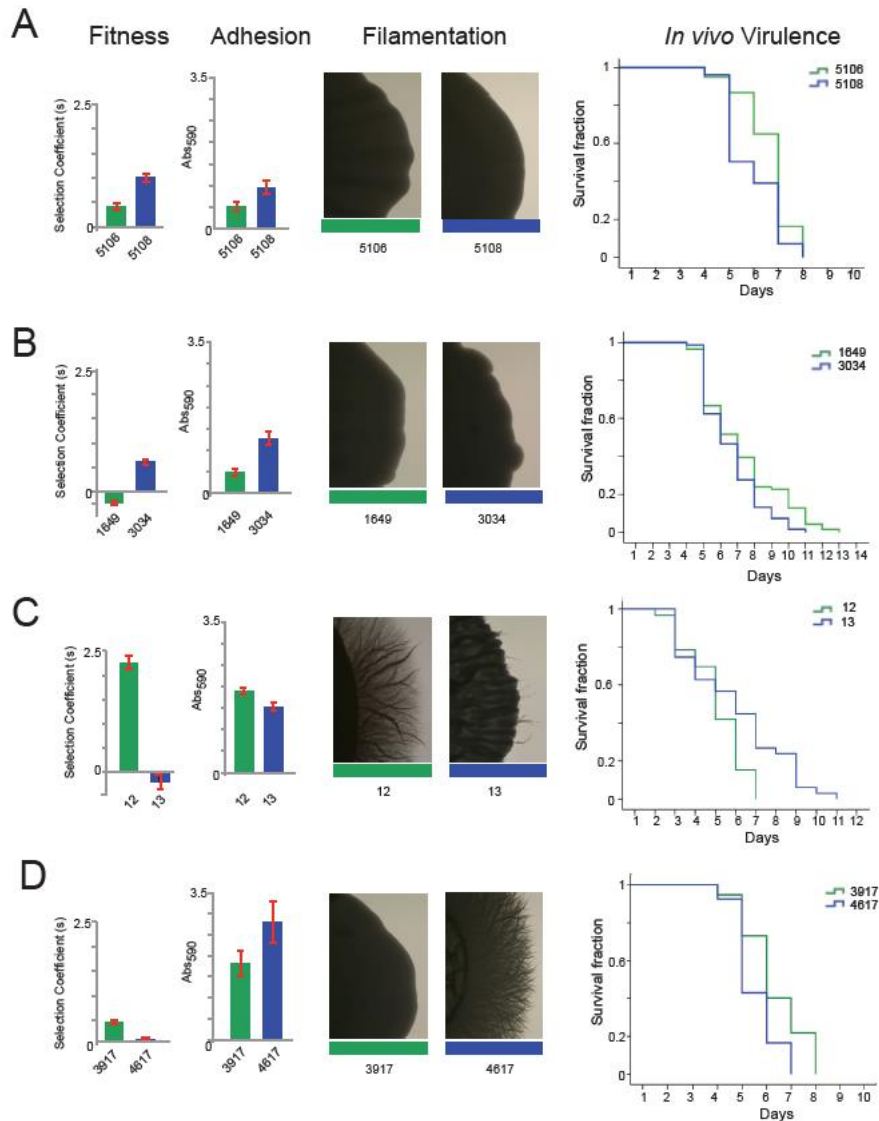
The most recurrently mutated gene outside of an LOH region was *Axl1* (persistently mutated in eight series), followed by ten genes mutated in seven series. Notably, persistent SNPs in *ERG11*, which encodes the drug target, appeared in Patients 9, 14, 30 and 59) in addition to mutations in the LOH events in Patients 1, 15, and 43 (**Figure 5 – source data 1**). Thus, *ERG11* was affected in 70% (7/10) of the patient series. More generally, 171 of the 240 genes were also mutated in an LOH tract in at least one additional patient (15/171 in 3 or more additional patients and 34/171 in 2 additional patients). Likewise Hyr3 was persistently mutated in nine of the patients, three of which occurred in LOH tracts, including one in which a new allele was homozygosed (**Figure 5, Figure 5 – source data 1, 2, and 3**).

Next, we partitioned the 240 recurrently mutated genes into 5 ‘co-occurrence clusters’ based on the correlation in their mutation occurrence patterns (**Figure 5, Figure 5 – source data 1**). These correlations are significantly higher than expected in a null model (KS test,  $P < 2.2 \times 10^{-126}$ , Wilcoxon Rank Sum test,  $P < 0.0196$ , permutation test). The characterized genes in most of the clusters have coherent functions. Cluster 1 is enriched for cell wall and cell surface genes, Cluster 2 for cell cycle and stress genes, Cluster 3 for genes involved in drug response and Cluster 5 for carbohydrate binding (**Figure 5, Figure 5 – source data 1**). Most of the genes in these clusters are not well characterized and represent new candidates involved drug resistance and adaptation to the host environment.

### **Changes in virulence phenotypes in evolved drug-resistant isolates**

To explore the possibility that some of the mutations reflect adaptation to other factors besides drug, we next measured four additional phenotypes (**Methods**). Filamentation, adhesion and competitive fitness were measured in competition assays in standard tissue culture medium with and without drug *in vitro*; and virulence was measured in a nematode model of infection (Jain et al., 2013a).

We found substantial variation in many of these phenotypes between isolates in the same series (**Figure 6, 7 and 8**), supporting the notion that the isolates are samples from a broad range of genetic variants within clonal (single infection) populations. In general, increased fitness *in vitro* (in the absence of drug) correlated with an increase in traits associated with virulence (adhesion, filamentation and virulence in nematode). For example, the later isolates in the series from patients 30 and 43 had increased fitness and higher virulence by all three measures (**Figure 6A, B**). Similarly, a decrease in fitness in isolate 13 of patient 1 was accompanied by a decrease in virulence (**Figure 6C**). A notable exception was patient 59, where fitness *in vitro* decreased while virulence phenotypes increased (**Figure 6D**).

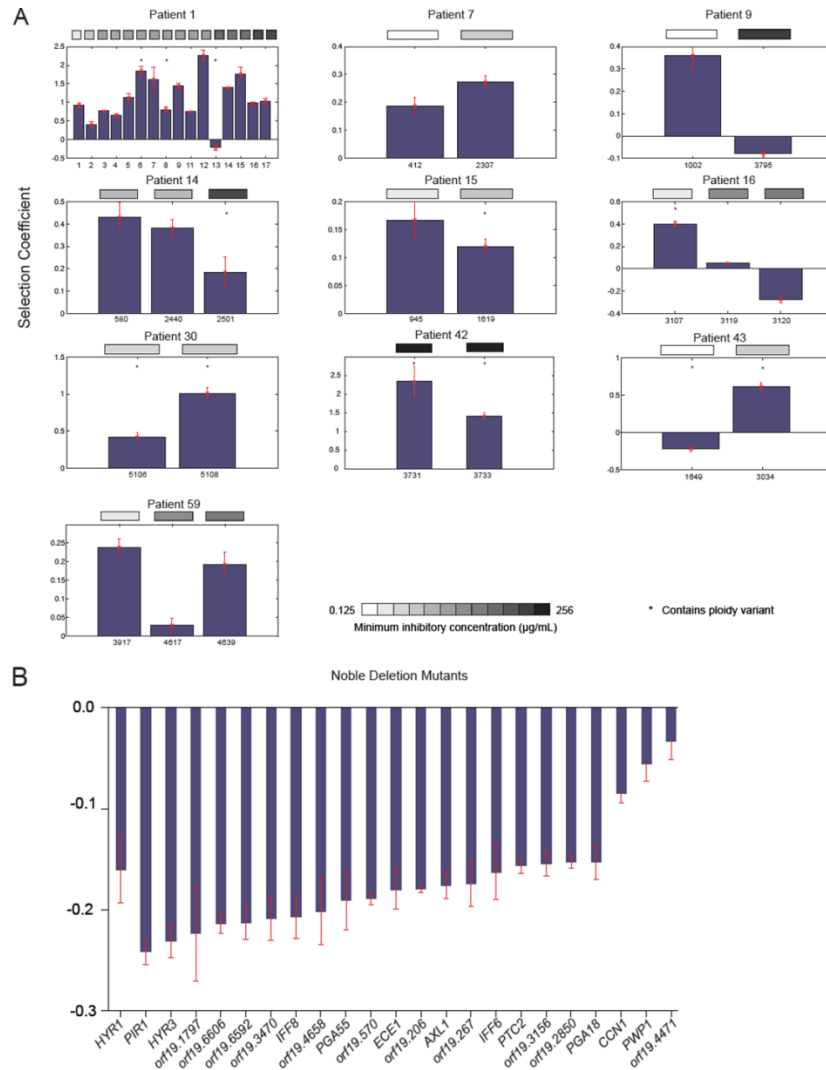


**Figure 6 Filamentation, adhesion and virulence increase concurrently with fitness.** For each pair of consecutive isolates (green preceding blue), shown are the fitness (far left, Y axis, **Methods**), adhesion (left, Y axis, **Methods**), filamentation (middle, **Methods**) and virulence in a worm model of infection (right). For virulence, shown are Kaplan-Meier plots of survival rates for each isolate, where significant changes in virulence were observed between the two isolates ( $P < 0.001$ , log-rank test). (A) Patient 30 isolates 5106 and 5108; (B) patient 43 isolates 1649 and 3034; (C) Patient 1 isolates 12 and 13; (D) patient 59 isolates 3917 and 4617.

Initially in a series, drug resistance (MIC) and *in vitro* fitness (in the absence of drug) were anti-correlated, suggesting that these are competing selective pressures. When MIC increases first appeared, they were usually accompanied by a *decrease* in fitness in the absence of drug (Patient 1, isolates 2-4, 13, 15 and 16, Patients 9, 14, 15 16, and 59, **Figure 7**). Of course, consistent with the elevated MIC, these isolates exhibited *increased* relative fitness in the

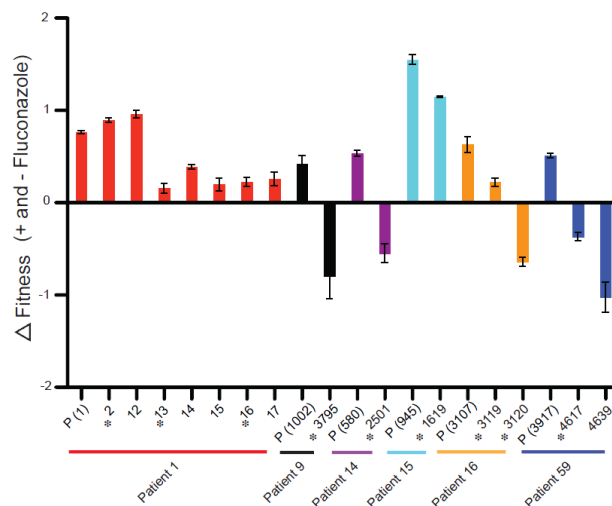


presence of the drug (**Figure 8**). This is also consistent with a recent study in *C. albicans* showing that resistance conferred by gain-of-function mutations in the transcription factors Mrr1, Tac1 and Upc2 are associated with reduced fitness under non-selective conditions *in vitro* as well as *in vivo* during colonization of a mammalian host (Sasse et al., 2012).



**Figure 7. Emergence of increased drug resistance often coincides with reduction in fitness in absence of drug. (A)** For each patient (panel) shown is the fitness (**Methods**) of each strain (Y axis, mean  $\pm$  STDV), ordered from the progenitor to evolved isolates (left to right, X axis). The MIC of each strain is shown in the grey boxes on top (white: low; black: high, color bar at bottom). Isolates with aneuploidies are marked by asterisks. **(B)** Shown is the fitness (**Methods**) for each double deletion mutant strain and corresponding wild type strain (Y axis, mean  $\pm$  STDV). The wild type strain (SN50) is on the far left and locus names are given for the mutant isolates (X axis). Double asterisks denote the MIC of the one mutant strain that differs from that of SN50.

Furthermore, consistent with selection of strains with compensatory variations, isolates from later time points often were more fit than those from earlier time points (measured *in vitro*, in the absence of drug) without further changes in MIC (e.g., patient 1, isolates 5-7, isolate 14, **Figure 7**). This is consistent with previous studies in bacteria (Bjorkman and Andersson, 2000) and in a single documented case in *C. glabrata* (Singh-Babak et al., 2012), suggesting that compensatory mutations arise to offset the major fitness cost of mutations conferring drug resistance.



**Figure 8 The difference in fitness in the presence or absence of Fluconazole.** Shown is the mean difference in fitness in the absence and presence of drug (Y axis, error bars are +/- STDV;  $n \geq 3$ ) for isolates (X axis) that showed a decrease in fitness (**Figure 6A**) in the absence of drug concomitant with an increase in MIC (asterisks), and flanking isolates in patient 1 and 59 (ordered from the progenitor (P) to evolved isolates (left to right, X axis). The difference in fitness is calculated as the difference in selection coefficient ( $s$ , Y axis) between matching competition experiments in RPMI and those in RPMI with one half the MIC for Fluconazole (**Table 1**) for each isolate tested (X axis). Negative values indicate that the strain had higher fitness in the presence of fluconazole versus assays without fluconazole.

In this context it appears that aneuploidies (**Figure 7**, asterisks), while largely transient, may be an important intermediate giving rise to more stable adaptive genotypes in some cases (Yona et al., 2012). For example, in Patient 1 isolate 13, an increase in MIC and a trisomy of 5 of 8 chromosomes accompanies a large decrease in fitness (in the absence of drug) relative to the preceding isolate 12 (**Figure 7**) but increased fitness in the presence of drug (**Figure 8**). Isolate 14 has a similar MIC phenotype to isolate 13, but is euploid (**Figure 3**) and is much more fit (**Figure 7**). Consistent with the generally negative effect of aneuploidy on

fitness (Tang and Amon, 2013), absence of the extra chromosomes resulted in improved overall fitness.

### **Candidate mutated genes associated with drug resistance or virulence**

Overall, the analysis of clinical isolates identified a range of new candidate genes that may affect drug resistance, fitness and/or virulence. To test the contribution of some of the recurrently identified genes to specific *C. albicans* phenotypes, we analyzed 23 of these persistent and recurrently mutated genes using available homozygous deletion mutants from a deletion strain collection (Noble et al., 2010b). We tested all those loci recurrently mutated in our data for which a deletion mutant was available.

We measured the MIC of fluconazole and *in vitro* fitness of each of 23 deletion isolates. One gene (*orf19.4658*) caused a 2-fold decrease in MIC, whereas the other 22 mutants tested did not have a significant effect on MIC. Deletion mutants are loss of function mutations, whereas the previously identified mechanisms of fluconazole resistance are “gain of function”, resulting in increases in the amount or activity level of Erg11 (Asai et al., 1999; Oliver et al., 2007) or the efflux of drug transporters (Coste et al., 2006; Dunkel et al., 2008). Therefore, it is possible that the recurrent nonsynonymous coding SNPs in the other loci we identified in the clinical isolates confer resistance. Alternatively, these loci may not be involved in fluconazole resistance *per se* but have a more general role in adaptation to the complex host environment. Consistent with a role in host adaptation, the majority of the deletion mutants affected *in vitro* fitness in a culture medium purported to approximate *in vivo* conditions (**Figure 7**, Methods). Three were more fit (**Figure 7**), including *CCN1*, that encodes a G1 cyclin required for hyphal growth maintenance (Loeb et al., 1999) and *orf19.4471*, an ortholog of *S. cerevisiae VPS64*, which is required for cytoplasm to vacuole targeting of proteins (Bonangelino et al., 2002), is involved in recycling pheromone receptors (Kemp and Sprague, 2003), and is identified as an ‘aneuploidy-tolerating mutant’ (Torres et al., 2010). Among the least fit were cell wall protein genes (*HYR1*, *HYR3* and *PIR1* (de Groot et al., 2004)).

## 6.5 Discussion

Here, we sequenced and analyzed the genomes of *C. albicans* serial clinical isolates by comparing ‘consecutive’ isolates from one patient to reach novel insights into the evolution of drug resistance within the human host. This approach allowed us to distinguish (and remove from further analysis) isolates that were non-clonal and to estimate that at least ~45% of the patients carried at least one unrelated strain of *C. albicans*. We used the clonal isolates to identify persistent SNPs, and the different series to identify those persistent SNPs that recurred within the same ORFs, thereby focusing the analysis on a small set of loci where the identified variants are more likely to be adaptive, and excluding the substantial background of likely neutral variation that hitchhike along with the others that are selective beneficial mutations.

Our study identified a surprisingly large amount of genetic diversity in each series, especially in light of the report of only 26 SNPs detected in a single clinical series of *Candida glabrata* isolates that spanned a 10 month period (Singh-Babak et al., 2012). Several reasons may account for this difference. **First**, fluconazole, the antifungal drug used to treat the patients in our series, is fungistatic, such that many cells exposed to the drug arrest their growth but do not die. Thus, the range of diversity in the initial population is not entirely lost. In contrast, the *C. glabrata* study involved exposure to caspofungin, an echinocandin fungicidal drug. Therefore, most cells likely died upon drug exposure and only the rare survivors went on to seed the remaining population. Accordingly, the *C. glabrata* isolates may have been subjected to selection that would have removed much of the initial diversity in the population, whereas in the *C. albicans* series diversity persisted and selection acted mostly to change the relative proportions of different genotypes. **Second**, *C. albicans* is a heterozygous diploid, while *C. glabrata* is haploid, and thus by its very nature, homozygous. Mutations can be more readily assimilated in a diploid than in a haploid organism, since deleterious mutations are potentially buffered by a functional version (Thompson et al., 2006). Furthermore, because *C. albicans* genomes are initially much more diverse (with tens of thousands of heterozygous SNPs in a given isolate), LOH is a high frequency mechanism available to reveal mutations more readily. **Third**, *C. albicans* lab isolates undergo a stress-induced elevation of mutation and mitotic recombination rates (Forche et al., 2011; Ponder et al., 2005; Rosenberg, 2011) and exposure to a mammalian host clearly results in elevated

frequencies of LOH and aneuploidy (Forche et al., 2009). Thus, we hypothesize that *C. albicans* isolates within the human host also undergo elevated levels of LOH, and possibly of mutation to generate a wider range of diversity. Additionally, they may not only tolerate genome and proteome diversity (Santos et al., 2004; Selmecki et al., 2010b), but also generate large-scale genetic variation as a means of adaptation (Chang et al., 2013; Gresham et al., 2008; Pavelka et al., 2010; Yona et al., 2012).

LOH of several genes important for fluconazole resistance have been reported previously (*ERG11*) (Oliver et al., 2007), *TAC1* (Coste et al., 2006) and *MRR1* (Schubert et al., 2011), however, the degree to which LOH is important in clinical infections was not known. In the 10 patients studied here, LOH was commonly observed (9/10) and was associated with changes in MIC. As we detected LOH of mutations in *ERG11* in 3 patients, it would be of interest to know if the LOH in these known genes was sufficient to increase MIC, or if other genes within the homozygous region make important contributions.

Aneuploidies appeared frequently within the drug resistant isolates, which is consistent with previous reports (Selmecki et al., 2006). Unlike LOH events, aneuploidies were often transient, and thus not uniquely correlated to increases in drug resistance. Perhaps these aneuploidies provide a mechanism akin to genetic assimilation ('phenotype precedes genotype'), in which cells are provided with a phenotypic mechanism that facilitates survival until a more stable and/or less costly mechanism is attained. In this case, the mechanism would be genetic but unstable—the acquisition of one or more extra chromosomes. Nevertheless, aneuploidies may cause increased frequencies of LOH events through whole chromosome loss, as well as by increasing the likelihood of recombination events. A transient role for aneuploidy is consistent with recent findings from *in vitro* evolution studies in *S. cerevisiae* (Yona et al., 2012) in which a transient aneuploidy was responsible for fitness at elevated temperature, but was eventually replaced by a more stable mutation.

In addition, a substantial number of persistent and recurrent SNPs, and clusters of co-occurring SNPs, implicate a broad range of pathways and functions that likely provide some growth advantage in the presence of the complex selective pressures found in the host. In particular there was strong enrichment for cell wall gene families purported to be critical determinants of the transition from commensalism to pathogenesis (Gow and Hube, 2012). The genes in several of these families (e.g. Als1, 7 and Hyr/Iff genes) frequently contain

intragenic tandem repeats. Variation in intragenic repeat number modulates phenotypic diversity in adhesion and biofilm formation (Verstrepen et al., 2005). This functional diversity of cell surface antigens has been proposed, in fungi and other pathogens, to allow rapid adaptation to the environment and evasion of the host immune system (Gemayel et al., 2010). Indeed, many of the isolates evolved additional phenotypes, including *in vitro* fitness, filamentation, adhesion and *in vivo* virulence, and the data presented here points to candidate genes that underlie some of these evolved traits. For example, the evolved isolate (4617) in patient 59 had a dramatic increase in filamentation relative to the progenitor that was concomitant with the appearance of persistent SNPs in genes associated with filamentous growth: *CHO1*, *MNN2* and 5 different *FGR* (filamentous growth regulator) genes (Uhl et al., 2003)..

The evolution of drug resistance in *C. albicans* has many parallels with somatic evolution of cancer cells undergoing chemotherapy or treated with specific inhibitors. These include variation on a background of clonal descent, lack of sexual recombination, acquisition of drug resistance, lack of diploid fidelity in successive mitotic divisions, and increased mutation and mitotic recombination rates (under stress). Indeed, several recent studies have shown a similar spectrum of genetic alterations during the somatic evolution of cancers in patients undergoing chemotherapy (Landau et al., 2013; Podlaha et al., 2012) or treated with specific inhibitors (Ding et al., 2012) to those observed here.

Finally, our data and analyses provide a rich and novel resource for *Candida* researchers and a host of candidate genes for further functional studies. While our analysis focused on recurrent SNPs in ORFs, we nonetheless cataloged the many genetic alterations found in intergenic regions (**Figure 2 - source data 1**), some of which could affect gene regulation. It will be especially interesting to analyze the similarities and differences in additional *C. albicans* genome sequences that are likely to become available in the near future. Our results suggest there may be complex population dynamics during the transition from commensal to pathogen, and across the course of treatment. As sequencing capacity continues to grow, it will be especially interesting to more fully sample this population level diversity during longitudinal collection to better understand these dynamics. In particular, it will be interesting to determine the degree to which specific mutations recur in different isolates prior to and after the acquisition of drug resistance.

## 6.6 Materials and Methods

### Isolates

Isolates were isolated from HIV-infected patients with oropharyngeal candidiasis, as previously described (Perea et al., 2001a; White, 1997b). The patients were not on azole antifungal treatment at time of enrollment; subsequent samples were collected during recurrence of infection. The isolates and are detailed in **Table 1**.

### Drug susceptibility

Minimal inhibitory concentrations (MIC) were determined for each strain using fluconazole E-test strips (0.016–256 µg/ml, bioMérieux, Durham, NC) on RPMI 1640-agar plates (Remel). Overnight YPD cultures were diluted in sterile 0.85% NaCl to an OD600 of 0.01 and 250 uL was plated using beads. After a 30-minute pre-incubation, 2-3 E-test strips were applied and plates were incubated at 35°C for 48 hours. The susceptibility endpoint was read at the first growth-inhibition ellipse, and the median value is reported here.

### Illumina sequencing

Genomic DNA was prepared from different clinical time courses via a Qiagen Maxiprep kit and sequenced using 101base paired-end Illumina sequencing (Mardis, 2008b). Library preparation included an 8 base barcode (Grabherr et al., 2011) ;43 samples from 11 patients were sequenced. All reads were mapped to the SC5314 reference genome (Candida Genome Database, gff downloaded on January 4, 2010) using the BWA alignment tool (version 0.5.9) (Li and Durbin, 2009). To minimize false positive SNP calls near insertion/deletion (indel) events, poorly aligning regions were identified and realigned using the GATK RealignerTargetCreator and IndelRealigner (GATK version 1.4-14, (version 1.4-14) (McKenna et al., 2010). Coverage for each strain is reported in **Table 1**. Coverage was defined as the total number of bases with BWA mapping quality greater than 10 divided by the total number of sites in the nuclear genome. Isolate 4380 aligned poorly to the SC5314 genome; however, this sequence aligned at high identity to the *C. dubliniensis*, and was therefore removed from further analysis. Reads are deposited for access to the NCBI SRA

under project accession number SRP028114.

### **SNP calling**

SNP were identified using Unified Genotyper (GATK version 1.4.14) (McKenna et al., 2010), using read alignments to the SC5314 reference sequence. Unreliable SNPs were identified using the GATK Variant Filtration module, with the version 3 best practice recommended annotation filters (QD < 2.0, MQ < 40.0, FS > 60.0, HaplotypeScore > 13.0, MQRankSum < -12.5, ReadPosRankSum < -8.0) except that the HaplotypeScore was also filtered if greater than two standard deviations above the mean of all HaplotypeScore values. The combined list of SNP positions across all strains was used to evaluate those matching the reference allele; by emitting all sites using Unified Genotyper, high quality reference matches were identified as positions with quality of 30 or greater, with positions with extremes of read depth (top or bottom 0.5% quantile) eliminated. A matrix of all strains by all positions was created from the SNP calls, with reference calls added where identified. Non-clonal isolates (see below) were removed from further analysis.

### **Sequenom iPLEX genotyping assay**

Loci chosen for validation were selected because they were: (1) persistent within their time courses; and (2) at least 150 bp away another SNP in either direction. 578 targets from 9 time courses were interrogated. This allowed us to compare our sequencing and genotyping predictions for 2499 cases (genetic locus X strain combination); 2270 predictions were confirmed as correct by Sequenom genotyping and 229 were incorrect (Table 2). Based on the 2499 cases at which comparison was possible, the SNP calling accuracy was determined to be 90.8%.

### **Determination of relatedness to determine clonality**

We investigated the phylogenetic relationship of all strains using SNP calls to determine relatedness between strains; positions with missing data in 10% or more of strains were eliminated, resulting in a total of 201,793 parsimony informative positions. A distance based tree was estimated using maximum parsimony with PAUP\* (4.0) (Swofford, 2002); a step matrix was used to estimate the distance between homozygous and heterozygous positions,



where each of the homozygotes are two steps apart from each other and one step from the heterozygote. SNP positions were resampled using 1,000 bootstrap replicates, and the phylogeny re-estimated to test the branch support. We define isolates with a branch distance of greater than 20,000 as non-clonal.

### **Copy-number determination**

For each strain, we calculated a per-locus depth-of-coverage using GATK(McKenna et al., 2010), with a minimal mapping quality of 10. The number of reads aligning to each 5 kb window across the nuclear genome was calculated and then normalized to the genome median. Each bin was then multiplied to the ploidy for the majority of the genome as determined by a FACS assay (below). We then applied a sliding window across each bin, defining a potential CNV if 70% of 10 consecutive bins had a normalized count > 2.5x. Regional amplifications are identified if > 15% of the chromosome is identified as having a CNV. Boundaries were confirmed by visual inspection in the Integrative Genome Viewer (Robinson et al., 2011a).

### **High-resolution ploidy analysis by flow cytometry**

*C. albicans* cultures were grown to log phase. 200µl of culture was centrifuged in a round bottom microtiter plate and pellets were resuspended in 20µl of 50mM Tris pH8 / 50mM EDTA (50/50 TE). 180µl of 95% ethanol was added and suspensions were stored overnight at 4°C. Cells were centrifuged and pellets washed twice with 200µl of 50/50 TE, then resuspended in 50µl of RNase A at 1mg/ml in 50/50TE and incubated 1 hour at 37°C. Cells were centrifuged and pellets resuspended in 50µl of Proteinase K at 5mg/ml in 50/50 TE for 30 minutes at 37°C. Cells were washed in 50/50 TE and pellets resuspended in 50µl of a 1:85 dilution SYBR Green I (Invitrogen, Carlsbad, CA) in 50/50 TE and incubated overnight in the dark at 4°C. Cells were centrifuged and pellets were resuspended in 700µl 50/50 TE and read on a FACScaliber flow cytometer (BD Biosciences, San Jose, CA). Flow data was fitted with a multi-Gaussian cell cycle model to produce estimates for whole genome ploidy.

## LOH determination

For each time course, we assembled the high quality SNPs (post-filtering, above) from multi-sample calling into the columns of a matrix, ordered by genome position, with the isolates in rows, ordered temporally. The genetic state of each locus in each sample was coded to distinguish loci homozygous for the haploid reference (-1), heterozygous SNPs (0), and homozygous SNPs for the non-reference state (1). We then applied a sliding window method across each chromosome, only looking at sites in which a SNP call was made in at least one isolate. An LOH event was defined as occurring if **(1)** at least one isolate had a heterozygosity content  $> 40\%$ , and **(2)** at least one other isolate had a heterozygosity content  $< 5\%$ . Window sizes were of length 500. Boundaries were trimmed such that if a window terminated in a heterozygous site in the isolate for which the LOH occurred, it was trimmed back until it was homozygous. If two 500+ windows were within 7 KB of each, the region was assessed to determine if the event was actually one event and merged if the heterozygous sites in the inter-window space had homozygosed. If two isolates had LOHs that overlapped but did not have precisely identical boundaries, the LOH regions were combined such that the LOH interval for both isolates was the same. All LOH regions were confirmed by visual inspection and are listed in **Figures 3 and 4-source data 1**.

## Classification of SNPs

For each time course, each SNP was classified for its position in the genome (**Figure 2-source data 1**). If the SNP fell within an ORF, the reference and altered amino acids were reported. If the SNP fell outside of an ORF, the distance to the closest flanking ORFs was reported, as well as the SNP's orientation with respect to these ORFs. SNP genotypes that are common to all isolates (including the 'progenitor') were classified as background mutations. Genotypes not present in the progenitor or evolved strain, but that occur in one or more intermediate strain, are classified as transient. Finally, genotypes that occur after the progenitor, and persist through the terminally evolved time point, are classified as persistent.

To determine if the number of persistent non-synonymous SNPs (nsSNPs) occurring in conjunction with changes in MIC was greater than expected, we developed a simple model to simulate the occurrence of nsSNPs outside of LOH regions at each time point. For each time point (i), a random variable  $X_i \sim \text{Pois}(\lambda_i)$  was assigned, where  $\lambda_i$  represents the Poisson

parameter for each time point:

$$\lambda_i = m/T * t_i$$

$m = 1471$ , the number of ORFs with persistent nsSNP

$T = 23$ , the number of time points

$t_i$  = the length of time (days) for time point (i) divided

by the mean length of time

The number of persistent nsSNP containing ORFs for each of the 14 time points associated with a change in MIC was summed, and this was repeated 100,000 times to build a probability distribution, where  $p(\text{observing } x \text{ mutated ORFs})$  was determined by dividing the number of successes for each bin by the number of trials.

To determine the probability of observing  $x$  recurrent, persistently mutated ORFs outside of LOH regions, we developed an additional stochastic simulation model. For each patient series (i), at each non-LOH ORF (j), a random variable,  $X_{ij} \sim B(n, p_{ij})$  was assigned, where  $n$  represents the number of trials, 1, and  $p$  represents the probability of a SNP occurring in that ORF:

$$p = m_i/M_i * h_j$$

$m_i$  = number of ORFs with persistent nsSNPs found

outside of LOH events for patient series (i)

$M_i$  = number of ORFs outside of LOH events for patient series (i)

$h_j$  = log normalized ORF length divided by mean log

normalized ORF length for ORF (j).

This was repeated 10,000 times to build the a distribution, and  $p(\text{observing } x \text{ recurrent nsSNPs})$  was determined by dividing the number of successes in each bin the cumulative number of trials.

### **Analysis of co-occurring mutations**

For co-occurrence analysis we focused only SNPs that (1) had persistent nonsynonymous

coding SNPs that did not occur in LOH regions and (2) recurred in three or more time courses. We generated for each such gene a binary patient vector, and we used NMF (Brunet et al., 2004) clustering to identify the optimal number of clusters, based on local maximas. This was accomplished using the “NMFCConsensus” module (version 5) in GenePattern. To determine the most appropriate number of clusters,  $k$  was selected such that it was the smallest value for which the cophenetic correlation begins decreasing. We then tested each of the co-occurrence gene clusters for functional enrichment (below). To determine if the degree of co-occurrence would have arisen by chance, we ran 1,000 iterations of one million edge-pair swaps from the original binary matrix, calculating a Pearson correlation matrix for each of the 1,000 iterations. We compared the distribution of Pearson correlations on the real and permuted vectors using a two-sample Kolmogorov-Smirnov (KS) test and Wilcoxon Rank Sum test.

### **Functional enrichment**

We calculated the overlap of each co-occurring cluster with Gene Ontology gene sets using the Gene Ontology toolset from the Candida Genome Database. Bonferroni adjusted P-values as well as the False Discovery Rate are reported (**Figures 5 – source data 1**).

### **Competition assay to assess fitness**

We measured the relative fitness of the progenitor and evolved lines in RPMI medium, competing them against a reference strain, expressing *ENO1::YFP*. Isolates stored at  $-80^{\circ}\text{C}$  were revived on rich media petri plates and then grown overnight in 3 ml cultures of minimal media in a roller drum at  $35^{\circ}\text{C}$ . An aliquot of cells in each culture was removed, sonicated in a Branson 450 sonifier, and the concentration of cells was determined using a Cellometer M10 (Nexcelom). The reference strain and experimental competitors were added to fresh RPMI medium in a 1:1 ratio and a final cell concentration of  $1 \times 10^7$  cells/ml. The cultures were grown for 24 hours in a roller drum at  $35^{\circ}\text{C}$ . Cells were then counted as above, and  $3 \times 10^6$  cells were transferred to fresh RPMI medium grown for 24 hours in a roller drum at  $35^{\circ}\text{C}$  (transfer cycle 1). This procedure was repeated (transfer cycle 2). This protocol represents 5–10 generations of growth, depending on the strain genotype. The ratio of the two competitors was quantified at the initial and final time points by flow cytometry (Accuri). Three to six

independent replicates for each fitness measurement were performed. The selective advantage,  $s$ , or disadvantage of the evolved population was calculated as previously described (Thompson et al., 2006), where  $E$  and  $R$  are the numbers of evolved and reference cells in the final ( $f$ ) and initial ( $i$ ) populations, and  $T$  is the number of generations that reference cells have proliferated during the competition.

### **Competition assay to assess fitness with and without fluconazole**

Fitness assays were performed as described above except that the reference strain used was a derivative of the drug resistant isolate 4639 from patient 59 (**Table 1**) expressing *ENO1::YFP* and competition experiments were performed in RPMI (cell culture medium, Gibco) and in RPMI with one half the MIC for fluconazole (**Table 1**) initiated from replicate 1:1 mixtures of the same population of cells for each isolate tested.

### ***C. elegans* survival assay**

A *C. elegans* survival assay was performed as previously described (Jain et al., 2009b). Briefly, *E. coli* OP50 and the different *C. albicans* clinical isolates were grown overnight respectively in LB at 37°C and YPD at 30°C. *E. coli* was then centrifuged and resuspended to a final concentration of 200 mg/ml, while *C. albicans* isolates were diluted with sterile water to OD<sub>600</sub>=3. Small petri dishes (3.5 cm) containing NGM agar were spotted with a mixture of 10 ml streptomycin (stock solution 50 mg/ml), 2.5 ml of *E. coli*, 0.5 ml of *C. albicans* and 7 ml of sterile water. The plates were incubated overnight at 25°C and 20 young synchronized N2 *C. elegans* adults were transferred on the spotted plates. Synchronous populations of adult worms were obtained by plating eggs on NGM plates seeded with *E. coli* OP50 at 20°C for 2-3 days. In this time frame, the eggs hatch and the larvae reach young adulthood. The survival assay was carried at 20°C and worms were scored daily by gentle prodding with a platinum wire; dead worms were discarded while live ones were transferred to seeded plates grown overnight at 25°C. Worms accidentally killed while transferring or found dead on the edges of the plates were censored. Statistical analysis was performed using SPSS software; survival curves were obtained using the Kaplan-Meier method and p-values by using the log-rank test.



### **Filamentation assay**

Overnight cultures grown in YPD at 30°C were normalized to OD<sub>600</sub>=1 with sterile water and spotted on Spider agar media (1% mannitol, 1% Difco nutrient broth, 0.2% K<sub>2</sub>HPO<sub>4</sub>). Plates were incubated at 37 °C and colonies photographed 3, 7 and 10 days post spotting. As a negative control for filamentation *cph1/cph1 efg1/efg1*(Lo et al., 1997) double mutant strain was used.

### ***In vitro* adhesion assay**

The *In vitro* adhesion assay was performed as previously described for *S. cerevisiae* (Reynolds and Fink, 2001). Briefly, cultures were grown in Synthetic Complete (SC) media + 0.15% glucose at 30°C overnight. Cells were then centrifuged at maximal speed and resuspended to OD<sub>600</sub>=0.5 in fresh media. 200 ml of each culture were dispensed into 8 wells of a flat bottom 96-well plate and incubated at 37°C for 4 hours. The content of the plate was then decanted and 50 ml of crystal violet added to each well. After 45 minutes of incubation at room temperature, the content of the plate was decanted and the plate was rinsed ten times in DI water by alternate submerging and decanting. 200 ml of 75% methanol was added to each well and absorbance was measured after 30 minutes at OD<sub>590</sub>. An *edt1/edt1* knockout mutant (Zakikhany et al., 2007) was used as a negative control for adhesion.

## 6.7 References:

- Alby, K., & Bennett, R. J. (2009). Stress-induced phenotypic switching in *Candida albicans*. *Molecular biology of the cell*, 20(14), 3178-3191. doi: 10.1091/mbc.E09-01-0040
- Andrulis, E. D., Zappulla, D. C., Ansari, A., Perrod, S., Laiosa, C. V., Gartenberg, M. R., & Sternglanz, R. (2002). Esc1, a nuclear periphery protein required for Sir4-based plasmid anchoring and partitioning. *Molecular and cellular biology*, 22(23), 8292-8301.
- Asai, K., Tsuchimori, N., Okonogi, K., Perfect, J. R., Gotoh, O., & Yoshida, Y. (1999). Formation of azole-resistant *Candida albicans* by mutation of sterol 14-demethylase P450. *Antimicrobial agents and chemotherapy*, 43(5), 1163-1169.
- Berman, J., & Sudbery, P. E. (2002). *Candida Albicans*: a molecular revolution built on lessons from budding yeast. *Nature reviews. Genetics*, 3(12), 918-930. doi: 10.1038/nrg948
- Bharucha, N., Chabrier-Rosello, Y., Xu, T., Johnson, C., Sobczynski, S., Song, Q., . . . Krysan, D. J. (2011). A large-scale complex haploinsufficiency-based genetic interaction screen in *Candida albicans*: analysis of the RAM network during morphogenesis. *PLoS genetics*, 7(4), e1002058.
- Bjorkman, J., & Andersson, D. I. (2000). The cost of antibiotic resistance from a bacterial perspective. *Drug resistance updates : reviews and commentaries in antimicrobial and anticancer chemotherapy*, 3(4), 237-245. doi: 10.1054/drup.2000.0147
- Blankenship, J. R., Fanning, S., Hamaker, J. J., & Mitchell, A. P. (2010). An extensive circuitry for cell wall regulation in *Candida albicans*. *PLoS pathogens*, 6(2), e1000752. doi: 10.1371/journal.ppat.1000752
- Bonangelino, C. J., Chavez, E. M., & Bonifacino, J. S. (2002). Genomic screen for vacuolar protein sorting genes in *Saccharomyces cerevisiae*. *Molecular biology of the cell*, 13(7), 2486-2501. doi: 10.1091/mbc.02-01-0005
- Bonhomme, J., Chauvel, M., Goyard, S., Roux, P., Rossignol, T., & d'Enfert, C. (2011). Contribution of the glycolytic flux and hypoxia adaptation to efficient biofilm formation by *Candida albicans*. *Molecular microbiology*, 80(4), 995-1013. doi: 10.1111/j.1365-2958.2011.07626.x
- Brunet, J. P., Tamayo, P., Golub, T. R., & Mesirov, J. P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proceedings of the National Academy of Sciences of the United States of America*, 101(12), 4164-4169. doi: 10.1073/pnas.0308531101
- Butler, G., Rasmussen, M. D., Lin, M. F., Santos, M. A., Sakthikumar, S., Munro, C. A., . . . Cuomo, C. A. (2009). Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature*, 459(7247), 657-662. doi: 10.1038/nature08064
- Chandra, J., Kuhn, D. M., Mukherjee, P. K., Hoyer, L. L., McCormick, T., & Ghannoum, M. A. (2001). Biofilm formation by the fungal pathogen *Candida albicans*: development, architecture, and drug resistance. *Journal of bacteriology*, 183(18), 5385-5394.



- Chang, S. L., Lai, H. Y., Tung, S. Y., & Leu, J. Y. (2013). Dynamic large-scale chromosomal rearrangements fuel rapid adaptation in yeast populations. *PLoS genetics*, 9(1), e1003232. doi: 10.1371/journal.pgen.1003232
- Charles, P. E., Doise, J. M., Quenot, J. P., Aube, H., Dalle, F., Chavanet, P., . . . Blettery, B. (2003). Candidemia in critically ill patients: difference of outcome between medical and surgical patients. *Intensive care medicine*, 29(12), 2162-2169. doi: 10.1007/s00134-003-2002-x
- Chico, L., Ciudad, T., Hsu, M., Lue, N. F., & Larriba, G. (2011). The *Candida albicans* Ku70 modulates telomere length and structure by regulating both telomerase and recombination. *PloS one*, 6(8), e23732. doi: 10.1371/journal.pone.0023732
- Connolly, L. A., Riccombeni, A., Grozer, Z., Holland, L. M., Lynch, D. B., Andes, D. R., . . . Butler, G. (2013). The APSES transcription factor Efg1 is a global regulator that controls morphogenesis and biofilm formation in *Candida parapsilosis*. *Molecular microbiology*, 90(1), 36-53. doi: 10.1111/mmi.12345
- Coste, A., Turner, V., Ischer, F., Morschhauser, J., Forche, A., Selmecki, A., . . . Sanglard, D. (2006). A mutation in Tac1p, a transcription factor regulating CDR1 and CDR2, is coupled with loss of heterozygosity at chromosome 5 to mediate antifungal resistance in *Candida albicans*. *Genetics*, 172(4), 2139-2156. doi: 10.1534/genetics.105.054767
- Cowen, L. E., Anderson, J. B., & Kohn, L. M. (2002). Evolution of drug resistance in *Candida albicans*. *Annual review of microbiology*, 56, 139-165. doi: 10.1146/annurev.micro.56.012302.160907
- Cowen, L. E., & Lindquist, S. (2005). Hsp90 potentiates the rapid evolution of new traits: drug resistance in diverse fungi. *Science*, 309(5744), 2185-2189. doi: 10.1126/science.1118370
- de Groot, P. W., de Boer, A. D., Cunningham, J., Dekker, H. L., de Jong, L., Hellingwerf, K. J., . . . Klis, F. M. (2004). Proteomic analysis of *Candida albicans* cell walls reveals covalently bound carbohydrate-active enzymes and adhesins. *Eukaryotic cell*, 3(4), 955-965. doi: 10.1128/EC.3.4.955-965.2004
- Ding, L., Ley, T. J., Larson, D. E., Miller, C. A., Koboldt, D. C., Welch, J. S., . . . DiPersio, J. F. (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature*, 481(7382), 506-510. doi: 10.1038/nature10738
- Dunkel, N., Blass, J., Rogers, P. D., & Morschhauser, J. (2008). Mutations in the multi-drug resistance regulator MRR1, followed by loss of heterozygosity, are the main cause of MDR1 overexpression in fluconazole-resistant *Candida albicans* strains. *Molecular microbiology*, 69(4), 827-840. doi: 10.1111/j.1365-2958.2008.06309.x
- Dunkel, N., & Morschhauser, J. (2011). Loss of heterozygosity at an unlinked genomic locus is responsible for the phenotype of a *Candida albicans* sap4Delta sap5Delta sap6Delta mutant. *Eukaryotic cell*, 10(1), 54-62. doi: 10.1128/EC.00281-10
- Felsenstein, J. (2005). PHYLIP (Phylogeny Inference Package) (Version 3.6). Department of Genome Sciences, University of Washington, Seattle: by author.
- Forche, A., Abbey, D., Pisithkul, T., Weinzierl, M. A., Ringstrom, T., Bruck, D., . . . Berman, J. (2011). Stress alters rates and types of loss of heterozygosity in *Candida albicans*. *mBio*, 2(4). doi: 10.1128/mBio.00129-11

- Forche, A., Alby, K., Schaefer, D., Johnson, A. D., Berman, J., & Bennett, R. J. (2008). The parasexual cycle in *Candida albicans* provides an alternative pathway to meiosis for the formation of recombinant strains. *PLoS biology*, 6(5), e110. doi: 10.1371/journal.pbio.0060110
- Forche, A., Magee, P. T., Magee, B. B., & May, G. (2004). Genome-wide single-nucleotide polymorphism map for *Candida albicans*. *Eukaryotic cell*, 3(3), 705-714. doi: 10.1128/EC.3.3.705-714.2004
- Forche, A., Magee, P. T., Selmecki, A., Berman, J., & May, G. (2009). Evolution in *Candida albicans* populations during a single passage through a mouse host. *Genetics*, 182(3), 799-811. doi: 10.1534/genetics.109.103325
- Forche, A., May, G., & Magee, P. T. (2005). Demonstration of loss of heterozygosity by single-nucleotide polymorphism microarray analysis and alterations in strain morphology in *Candida albicans* strains during infection. *Eukaryotic cell*, 4(1), 156-165. doi: 10.1128/EC.4.1.156-165.2005
- Galhardo, R. S., Hastings, P. J., & Rosenberg, S. M. (2007). Mutation as a stress response and the regulation of evolvability. *Critical reviews in biochemistry and molecular biology*, 42(5), 399-435. doi: 10.1080/10409230701648502
- Gemayel, R., Vinces, M. D., Legendre, M., & Verstrepen, K. J. (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual review of genetics*, 44, 445-477. doi: 10.1146/annurev-genet-072610-155046
- Gow, N. A., & Hube, B. (2012). Importance of the *Candida albicans* cell wall during commensalism and infection. *Current opinion in microbiology*, 15(4), 406-412. doi: 10.1016/j.mib.2012.04.005
- Goyard, S., Knechtle, P., Chauvel, M., Mallet, A., Prevost, M. C., Proux, C., . . . d'Enfert, C. (2008). The Yak1 kinase is involved in the initiation and maintenance of hyphal growth in *Candida albicans*. *Molecular biology of the cell*, 19(5), 2251-2266. doi: 10.1091/mbc.E07-09-0960
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., . . . Regev, A. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology*, 29(7), 644-652. doi: 10.1038/nbt.1883
- Gresham, D., Desai, M. M., Tucker, C. M., Jenq, H. T., Pai, D. A., Ward, A., . . . Dunham, M. J. (2008). The Repertoire and Dynamics of Evolutionary Adaptations to Controlled Nutrient-Limited Environments in Yeast. *PLoS genetics*, 4(12). doi: Artn E1000303
- Doi 10.1371/Journal.Pgen.1000303
- Gudlaugsson, O., Gillespie, S., Lee, K., Vande Berg, J., Hu, J., Messer, S., . . . Diekema, D. (2003). Attributable mortality of nosocomial candidemia, revisited. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 37(9), 1172-1177. doi: 10.1086/378745
- Hayek, P., Dib, L., Yazbeck, P., Beyrouthy, B., & Khalaf, R. A. (2010). Characterization of Hwp2, a *Candida albicans* putative GPI-anchored cell wall protein necessary for invasive growth. *Microbiological research*, 165(3), 250-258. doi: 10.1016/j.micres.2009.03.006

- Hoyer, L. L., Green, C. B., Oh, S. H., & Zhao, X. (2008). Discovering the secrets of the *Candida albicans* agglutinin-like sequence (ALS) gene family--a sticky pursuit. *Medical mycology : official publication of the International Society for Human and Animal Mycology*, *46*(1), 1-15. doi: 10.1080/13693780701435317
- Hoyer, L. L., Payne, T. L., Bell, M., Myers, A. M., & Scherer, S. (1998). *Candida albicans* ALS3 and insights into the nature of the ALS gene family. *Curr Genet*, *33*(6), 451-459.
- Jain, C., Pastor, K., Gonzalez, A., Lorenz, M. C., & Rao, R. P. (2013). The role of *Candida albicans* AP-1 protein against host derived ROS in in vivo models of infection. *Virulence*, *4*(1).
- Jain, C., Yun, M., Politz, S. M., & Rao, R. P. (2009). A pathogenesis assay using *Saccharomyces cerevisiae* and *Caenorhabditis elegans* reveals novel roles for yeast AP-1, Yap1, and host dual oxidase BLI-3 in fungal pathogenesis. *Eukaryotic cell*, *8*(8), 1218-1227. doi: 10.1128/EC.00367-08
- Karababa, M., Coste, A. T., Rognon, B., Bille, J., & Sanglard, D. (2004). Comparison of gene expression profiles of *Candida albicans* azole-resistant clinical isolates and laboratory strains exposed to drugs inducing multidrug transporters. *Antimicrobial agents and chemotherapy*, *48*(8), 3064-3079. doi: 10.1128/AAC.48.8.3064-3079.2004
- Kemp, H. A., & Sprague, G. F., Jr. (2003). Far3 and five interacting proteins prevent premature recovery from pheromone arrest in the budding yeast *Saccharomyces cerevisiae*. *Molecular and cellular biology*, *23*(5), 1750-1763.
- Kempf, M., Cottin, J., Licznar, P., Lefrancois, C., Robert, R., & Apaire-Marchais, V. (2009). Disruption of the GPI protein-encoding gene IFF4 of *Candida albicans* results in decreased adherence and virulence. *Mycopathologia*, *168*(2), 73-77. doi: 10.1007/s11046-009-9201-0
- Kim, J., & Sudbery, P. (2011). *Candida albicans*, a major human fungal pathogen. *Journal of microbiology*, *49*(2), 171-177. doi: 10.1007/s12275-011-1064-7
- Kinneberg, K. M., Bendel, C. M., Jechorek, R. P., Cebelinski, E. A., Gale, C. A., Berman, J. G., . . . Wells, C. L. (1999). Effect of INT1 gene on *Candida albicans* murine intestinal colonization. *The Journal of surgical research*, *87*(2), 245-251. doi: 10.1006/jsre.1999.5755
- Kumamoto, C. A., & Pierce, J. V. (2011). Immunosensing during colonization by *Candida albicans*: does it take a village to colonize the intestine? *Trends in microbiology*, *19*(6), 263-267. doi: 10.1016/j.tim.2011.01.009
- Kumamoto, C. A., & Vences, M. D. (2005). Alternative *Candida albicans* lifestyles: growth on surfaces. *Annual review of microbiology*, *59*, 113-133. doi: 10.1146/annurev.micro.59.030804.121034
- Kunze, D., Melzer, I., Bennett, D., Sanglard, D., MacCallum, D., Norskau, J., . . . Hube, B. (2005). Functional analysis of the phospholipase C gene CaPLC1 and two unusual phospholipase C genes, CaPLC2 and CaPLC3, of *Candida albicans*. *Microbiology*, *151*(Pt 10), 3381-3394. doi: 10.1099/mic.0.28353-0
- Landau, D. A., Carter, S. L., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M. S., . . . Wu, C. J. (2013). Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell*, *152*(4), 714-726. doi: 10.1016/j.cell.2013.01.019

- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25(14), 1754-1760. doi: 10.1093/bioinformatics/btp324
- Lo, H. J., Kohler, J. R., DiDomenico, B., Loebenberg, D., Cacciapuoti, A., & Fink, G. R. (1997). Nonfilamentous *C. albicans* mutants are avirulent. *Cell*, 90(5), 939-949.
- Loeb, J. D., Sepulveda-Becerra, M., Hazan, I., & Liu, H. (1999). A G1 cyclin is necessary for maintenance of filamentous growth in *Candida albicans*. *Molecular and cellular biology*, 19(6), 4019-4027.
- MacPherson, S., Akache, B., Weber, S., De Deken, X., Raymond, M., & Turcotte, B. (2005). *Candida albicans* zinc cluster protein Upc2p confers resistance to antifungal drugs and is an activator of ergosterol biosynthetic genes. *Antimicrobial agents and chemotherapy*, 49(5), 1745-1752. doi: 10.1128/AAC.49.5.1745-1752.2005
- Mardis, E. R. (2008). The impact of next-generation sequencing technology on genetics. *Trends in genetics : TIG*, 24(3), 133-141. doi: 10.1016/j.tig.2007.12.007
- Marichal, P., Koymans, L., Willemsens, S., Bellens, D., Verhasselt, P., Luyten, W., . . . Bossche, H. V. (1999). Contribution of mutations in the cytochrome P450 14 $\alpha$ -demethylase (Erg11p, Cyp51p) to azole resistance in *Candida albicans*. *Microbiology*, 145 ( Pt 10), 2701-2713.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., . . . DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9), 1297-1303. doi: 10.1101/gr.107524.110
- Munro, C. A., Winter, K., Buchan, A., Henry, K., Becker, J. M., Brown, A. J., . . . Gow, N. A. (2001). Chs1 of *Candida albicans* is an essential chitin synthase required for synthesis of the septum and for cell integrity. *Molecular microbiology*, 39(5), 1414-1426.
- Murillo, L. A., Newport, G., Lan, C. Y., Habelitz, S., Dungan, J., & Agabian, N. M. (2005). Genome-wide transcription profiling of the early phase of biofilm formation by *Candida albicans*. *Eukaryotic cell*, 4(9), 1562-1573. doi: 10.1128/EC.4.9.1562-1573.2005
- Nobile, C. J., Fox, E. P., Nett, J. E., Sorrells, T. R., Mitrovich, Q. M., Hernday, A. D., . . . Johnson, A. D. (2012). A recently evolved transcriptional network controls biofilm development in *Candida albicans*. *Cell*, 148(1-2), 126-138. doi: 10.1016/j.cell.2011.10.048
- Nobile, C. J., & Mitchell, A. P. (2005). Regulation of cell-surface genes and biofilm formation by the *C. albicans* transcription factor Bcr1p. *Current biology : CB*, 15(12), 1150-1155. doi: 10.1016/j.cub.2005.05.047
- Noble, S. M., French, S., Kohn, L. A., Chen, V., & Johnson, A. D. (2010). Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nature genetics*, 42(7), 590-598. doi: 10.1038/ng.605
- Oliver, B. G., Song, J. L., Choiniere, J. H., & White, T. C. (2007). cis-Acting elements within the *Candida albicans* ERG11 promoter mediate the azole response through transcription factor Upc2p. *Eukaryotic cell*, 6(12), 2231-2239. doi: 10.1128/EC.00331-06

- Pappas, P. G., Rex, J. H., Lee, J., Hamill, R. J., Larsen, R. A., Powderly, W., . . . Dismukes, W. E. (2003). A prospective observational study of candidemia: epidemiology, therapy, and influences on mortality in hospitalized adult and pediatric patients. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 37(5), 634-643.
- Pavelka, N., Rancati, G., Zhu, J., Bradford, W. D., Saraf, A., Florens, L., . . . Li, R. (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature*, 468(7321), 321-325. doi: 10.1038/nature09529
- Perea, S., Lopez-Ribot, J. L., Kirkpatrick, W. R., McAtee, R. K., Santillan, R. A., Martinez, M., . . . Patterson, T. F. (2001). Prevalence of molecular mechanisms of resistance to azole antifungal agents in *Candida albicans* strains displaying high-level fluconazole resistance isolated from human immunodeficiency virus-infected patients. *Antimicrobial agents and chemotherapy*, 45(10), 2676-2684. doi: 10.1128/AAC.45.10.2676-2684.2001
- Pfaller, M. A., Lockhart, S. R., Pujol, C., Swails-Wenger, J. A., Messer, S. A., Edmond, M. B., . . . Soll, D. R. (1998). Hospital specificity, region specificity, and fluconazole resistance of *Candida albicans* bloodstream isolates. *Journal of clinical microbiology*, 36(6), 1518-1529.
- Pittet, D., Monod, M., Suter, P. M., Frenk, E., & Auckenthaler, R. (1994). *Candida* colonization and subsequent infections in critically ill surgical patients. *Annals of surgery*, 220(6), 751-758.
- Podlaha, O., Riester, M., De, S., & Michor, F. (2012). Evolution of the cancer genome. *Trends in genetics : TIG*, 28(4), 155-163. doi: 10.1016/j.tig.2012.01.003
- Ponder, R. G., Fonville, N. C., & Rosenberg, S. M. (2005). A switch from high-fidelity to error-prone DNA double-strand break repair underlies stress-induced mutation. *Molecular cell*, 19(6), 791-804. doi: 10.1016/j.molcel.2005.07.025
- Rex, J. H., Bennett, J. E., Sugar, A. M., Pappas, P. G., van der Horst, C. M., Edwards, J. E., . . . et al. (1994). A randomized trial comparing fluconazole with amphotericin B for the treatment of candidemia in patients without neutropenia. Candidemia Study Group and the National Institute. *The New England journal of medicine*, 331(20), 1325-1330. doi: 10.1056/NEJM199411173312001
- Rex, J. H., Pappas, P. G., Karchmer, A. W., Sobel, J., Edwards, J. E., Hadley, S., . . . Booth, J. (2003). A randomized and blinded multicenter trial of high-dose fluconazole plus placebo versus fluconazole plus amphotericin B as therapy for candidemia and its consequences in nonneutropenic subjects. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America*, 36(10), 1221-1228. doi: 10.1086/374850
- Reynolds, T. B., & Fink, G. R. (2001). Bakers' yeast, a model for fungal biofilm formation. *Science*, 291(5505), 878-881. doi: 10.1126/science.291.5505.878
- Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., & Mesirov, J. P. (2011). Integrative genomics viewer. *Nature biotechnology*, 29(1), 24-26. doi: 10.1038/nbt.1754
- Rosenberg, S. M. (2011). Stress-induced loss of heterozygosity in *Candida*: a possible missing link in the ability to evolve. *mBio*, 2(5). doi: 10.1128/mBio.00200-11

- Ruiz-Herrera, J., Elorza, M. V., Valentin, E., & Sentandreu, R. (2006). Molecular organization of the cell wall of *Candida albicans* and its relation to pathogenicity. *FEMS yeast research*, *6*(1), 14-29. doi: 10.1111/j.1567-1364.2005.00017.x
- Sandini, S., Stringaro, A., Arancia, S., Colone, M., Mondello, F., Murtas, S., . . . De Bernardis, F. (2011). The MP65 gene is required for cell wall integrity, adherence to epithelial cells and biofilm formation in *Candida albicans*. *BMC microbiology*, *11*, 106. doi: 10.1186/1471-2180-11-106
- Santos, M. A., Moura, G., Massey, S. E., & Tuite, M. F. (2004). Driving change: the evolution of alternative genetic codes. *Trends in genetics : TIG*, *20*(2), 95-102. doi: 10.1016/j.tig.2003.12.009
- Sasse, C., Dunkel, N., Schafer, T., Schneider, S., Dierolf, F., Ohlsen, K., & Morschhauser, J. (2012). The stepwise acquisition of fluconazole resistance mutations causes a gradual loss of fitness in *Candida albicans*. *Molecular microbiology*, *86*(3), 539-556. doi: 10.1111/j.1365-2958.2012.08210.x
- Schubert, S., Popp, C., Rogers, P. D., & Morschhauser, J. (2011). Functional dissection of a *Candida albicans* zinc cluster transcription factor, the multidrug resistance regulator Mrr1. *Eukaryotic cell*, *10*(8), 1110-1121. doi: 10.1128/EC.05100-11
- Selmecki, A., Forche, A., & Berman, J. (2006). Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science*, *313*(5785), 367-370. doi: 10.1126/science.1128242
- Selmecki, A., Forche, A., & Berman, J. (2010). Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryotic cell*, *9*(7), 991-1008. doi: 10.1128/EC.00060-10
- Selmecki, A., Gerami-Nejad, M., Paulson, C., Forche, A., & Berman, J. (2008). An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. *Molecular microbiology*, *68*(3), 624-641. doi: 10.1111/j.1365-2958.2008.06176.x
- Selmecki, A. M., Dulmage, K., Cowen, L. E., Anderson, J. B., & Berman, J. (2009). Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. *PLoS genetics*, *5*(10), e1000705. doi: 10.1371/journal.pgen.1000705
- Sheppard, D. C., Yeaman, M. R., Welch, W. H., Phan, Q. T., Fu, Y., Ibrahim, A. S., . . . Edwards, J. E., Jr. (2004). Functional and structural diversity in the Als protein family of *Candida albicans*. *Journal of Biological Chemistry*, *279*(29), 30480-30489. doi: 10.1074/jbc.M401929200
- Singh-Babak, S. D., Babak, T., Diezmann, S., Hill, J. A., Xie, J. L., Chen, Y. L., . . . Cowen, L. E. (2012). Global analysis of the evolution and mechanism of echinocandin resistance in *Candida glabrata*. *PLoS pathogens*, *8*(5), e1002718. doi: 10.1371/journal.ppat.1002718
- Storm, N., Darnhofer-Patel, B., van den Boom, D., & Rodi, C. P. (2003). MALDI-TOF mass spectrometry-based SNP genotyping. *Methods in molecular biology*, *212*, 241-262.
- Swofford, D. L. (2002). PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods). *Sinauer Associates, Sunderland, Massachusetts*.

- Tang, Y. C., & Amon, A. (2013). Gene copy-number alterations: a cost-benefit analysis. *Cell*, 152(3), 394-405. doi: 10.1016/j.cell.2012.11.043
- Thompson, D. A., Desai, M. M., & Murray, A. W. (2006). Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. *Current biology : CB*, 16(16), 1581-1590. doi: 10.1016/j.cub.2006.06.070
- Torres, E. M., Dephoure, N., Panneerselvam, A., Tucker, C. M., Whittaker, C. A., Gygi, S. P., . . . Amon, A. (2010). Identification of aneuploidy-tolerating mutations. *Cell*, 143(1), 71-83. doi: 10.1016/j.cell.2010.08.038
- Uhl, M. A., Biery, M., Craig, N., & Johnson, A. D. (2003). Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen *C.albicans*. *The EMBO journal*, 22(11), 2668-2678. doi: 10.1093/emboj/cdg256
- Vandeputte, P., Ischer, F., Sanglard, D., & Coste, A. T. (2011). In vivo systematic analysis of *Candida albicans* Zn2-Cys6 transcription factors mutants for mice organ colonization. *PloS one*, 6(10), e26962. doi: 10.1371/journal.pone.0026962
- Verstrepen, K. J., Jansen, A., Lewitter, F., & Fink, G. R. (2005). Intragenic tandem repeats generate functional variability. *Nature genetics*, 37(9), 986-990. doi: 10.1038/ng1618
- Wachtler, B., Wilson, D., Haedicke, K., Dalle, F., & Hube, B. (2011). From attachment to damage: defined genes of *Candida albicans* mediate adhesion, invasion and damage during interaction with oral epithelial cells. *PloS one*, 6(2), e17046. doi: 10.1371/journal.pone.0017046
- Wapinski, I., Pfeffer, A., Friedman, N., & Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. *Nature*, 449(7158), 54-61. doi: 10.1038/nature06107
- White, T. C. (1997a). Increased mRNA levels of ERG16, CDR, and MDR1 correlate with increases in azole resistance in *Candida albicans* isolates from a patient infected with human immunodeficiency virus. *Antimicrobial agents and chemotherapy*, 41(7), 1482-1487.
- White, T. C. (1997b). The presence of an R467K amino acid substitution and loss of allelic variation correlate with an azole-resistant lanosterol 14alpha demethylase in *Candida albicans*. *Antimicrobial agents and chemotherapy*, 41(7), 1488-1494.
- White, T. C., Holleman, S., Dy, F., Mirels, L. F., & Stevens, D. A. (2002). Resistance mechanisms in clinical isolates of *Candida albicans*. *Antimicrobial agents and chemotherapy*, 46(6), 1704-1713.
- Wisplinghoff, H., Seifert, H., Tallent, S. M., Bischoff, T., Wenzel, R. P., & Edmond, M. B. (2003). Nosocomial bloodstream infections in pediatric patients in United States hospitals: epidemiology, clinical features and susceptibilities. *Pediatr Infect Dis J*, 22(8), 686-691.
- Xu, D., Jiang, B., Ketela, T., Lemieux, S., Veillette, K., Martel, N., . . . Roemer, T. (2007). Genome-wide fitness test and mechanism-of-action studies of inhibitory compounds in *Candida albicans*. *PLoS pathogens*, 3(6), e92. doi: 10.1371/journal.ppat.0030092
- Yona, A. H., Manor, Y. S., Herbst, R. H., Romano, G. H., Mitchell, A., Kupiec, M., . . . Dahan, O. (2012). Chromosomal duplication is a transient evolutionary solution to

stress. *Proceedings of the National Academy of Sciences of the United States of America*. doi: 10.1073/pnas.1211150109

Zakikhany, K., Naglik, J. R., Schmidt-Westhausen, A., Holland, G., Schaller, M., & Hube, B. (2007). In vivo transcript profiling of *Candida albicans* identifies a gene essential for interepithelial dissemination. *Cellular microbiology*, 9(12), 2938-2954. doi: 10.1111/j.1462-5822.2007.01009.x

Zhang, T., Li, W., Li, D., Wang, Y., & Sang, J. (2008). Role of CaECM25 in cell morphogenesis, cell growth and virulence in *Candida albicans*. *Science in China. Series C, Life sciences / Chinese Academy of Sciences*, 51(4), 362-372. doi: 10.1007/s11427-008-0051-7



## 7 Conclusions and future directions

Overall this thesis shed light on the complex mechanisms that lead *C. albicans* to its virulence. By using an *in vivo* model system previously developed in the lab, we identified four virulence factors that have been poorly characterized in the literature and that are important for inducing pathogenicity in *C. elegans*. Between these four novel virulence factors we focused on *ZCF15*, a transcription factor never found in prokaryotes or other eukaryotes and that can therefore be used as a target for potential therapeutic intervention.

We showed here that *ZCF15* plays a major role in the pathogen's ability to resist toxic reactive oxygen species generated by the host. *ZCF15* deletion leads to a reduced virulence in nematodes and murine macrophages unless we crippled the ability of the host to produce reactive oxygen species. Moreover, *ZCF15* deletion leads to a reduced virulence even in the significant more complex environment of the mouse blood stream reiterating its important role in virulence.

In addition to *ZCF15* we functionally characterized in this thesis *ZCF29*, a gene in the same family of transcription factors that has been shown by others to have a role in virulence and ROS resistance. By using genome scale approaches like RNA-Seq and Chip-Seq we showed that *ZCF15* plays a major role in controlling carbon metabolism and in inducing ROS detoxification processes when *C. albicans* is exposed to ROS. On the other hand, *ZCF29* plays an important role in regulating N-metabolism and is required for down regulating ribosome biogenesis upon ROS exposure.

Although this thesis pinpointed the genes and biological processes that *ZCF15* and *ZCF29* are regulating, more experiments are required to fully dissect the metabolic pathways that are controlled by *ZCF15* and *ZCF29*. For example, in order to discriminate if any of the genes

found differentially expressed in *zcf15/zcf15* and *zcf29/zcf29* are directly or indirectly controlled by *ZCF15* and *ZCF29* we can obtain deletion strains in these genes and phenotypically characterized them. Knockouts that have phenotypes that recapitulate the once observed in *zcf15/zcf15* and *zcf29/zcf29* are likely working in the same pathway and further epistasis experiments can indicate their order in the pathway. Moreover, because TF are known to cooperate with other TF in the process of binding DNA, a yeast two hybrid screen can be used to discover potential binding partners that can subsequently be validated by co-immunoprecipitation experiments.

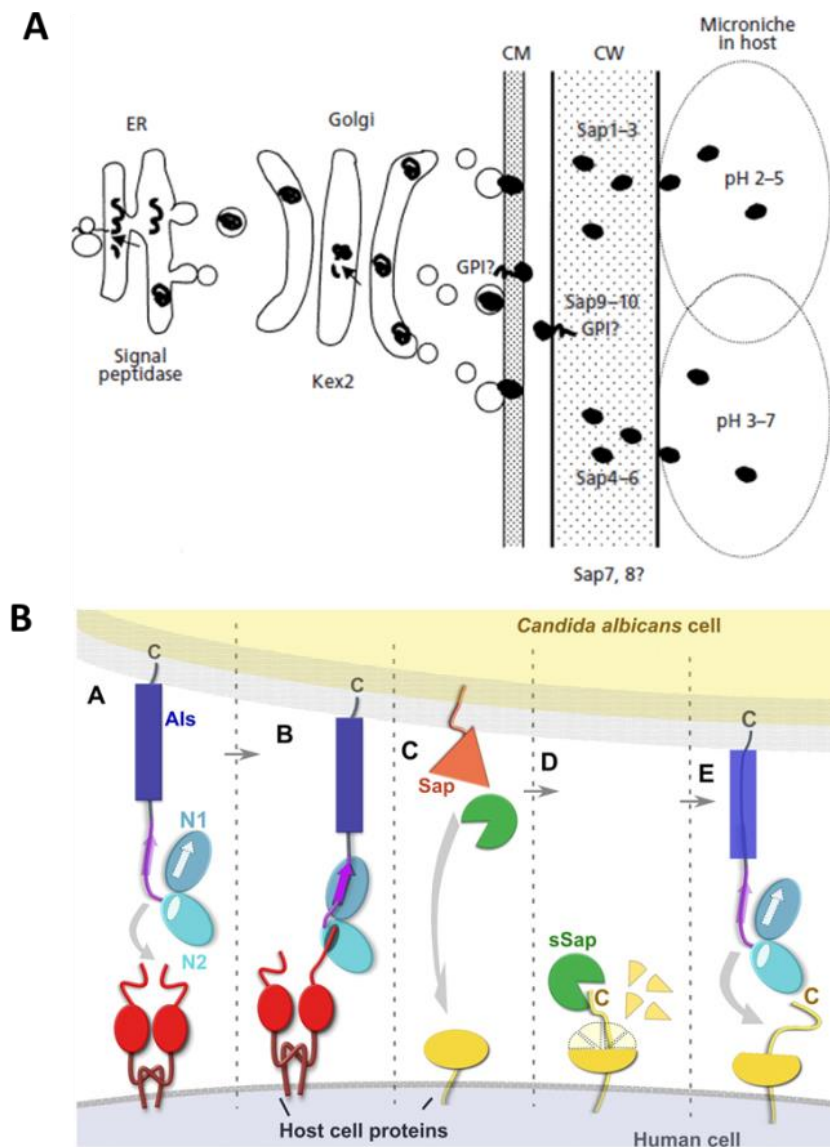
Because *ZCF15* and *ZCF29* are so fundamental for *C. albicans* biology and they are not conserved in humans, they can also be considered for potential drug targeting. Once their DNA binding partners are fully characterized it is possible to design peptide mimetics that recognized these sites and prevent their binding. Overall, this thesis identified novel *C. albicans* virulence determinants that can be used by the larger community for further biological studies or even drug targeting.

## 8 Appendix

### 8.1 Potential vaccine approach against *C. albicans* Saps

Saps are a family of 10 secreted aspartyl proteases that have been long recognized as an important *C. albicans* virulence factors (Fallon et al., 1997). These hydrolytic enzymes are believed to be secreted during the infection in order to degrade host tissues and allow *C. albicans* penetration to deeper epithelial layers (Naglik et al., 2003a). Saps have a highly conserved aspartate residue that plays a major role in the catalysis and Figure 0.1 depict more in detail their mechanisms of action. Saps are translated in the endoplasmic reticulum as preproteins. As shown in Figure 0.1 A, Kex2, a subtilisin-like protease cleaves the preproteins in the Golgi generating mature Saps that are subsequently secreted to the extracellular space (Newport and Agabian, 1997). In the extracellular space Saps have different optimum pHs of activities with Sap1-3 at pH 3-7 and Sap 4-6 at pH 3-7 (Figure 0.1 A) indicating that their individual roles in infection might varies depending on the local tissue acidities (Parnanen et al., 2010).

Upon secretion, Saps cooperate with *C. albicans* Als (agglutinin-like sequences) glycoproteins and facilitate pathogen's adhesion and tissue invasion (Figure 0.1 B). In particular it has been recently theorized that the apical domain of Als protrude from the *C. albicans* cells wall and bind the C-terminus domain of host cell surface proteins (Salgado et al., 2011). Saps cooperates to the Als binding by partially hydrolyzing host cell proteins and consequently exposing more free C-terminus domains that are available for Als binding (Figure 0.1 B).



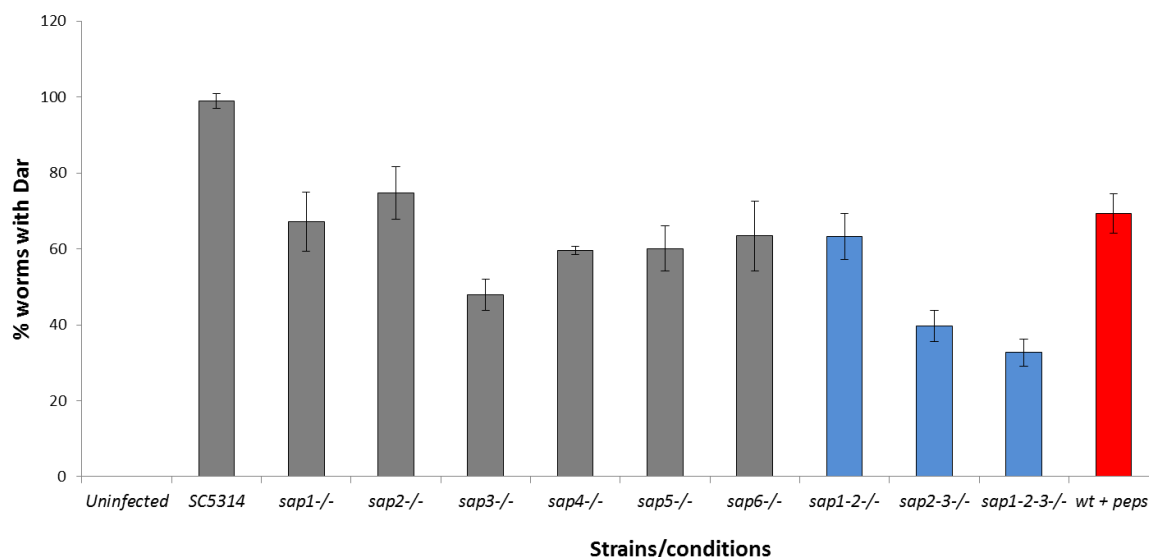
**Figure 0.1 *C. albicans* Saps secretion and hypothesized mechanism of action.**

- A) Saps are translated in the endoplasmic reticulum as preprotein, cleaved by Kex2 in the Golgi and subsequently secreted in the extracellular space by a complex system of vesicles. Figure adapted from Naglik et al., 2003
- B) Saps and Als proteins facilitate pathogen's adhesion and penetration to deeper human tissues. Als N-termini protrude out of the *C. albicans* cell wall (N1-N2) and bind free C-terminus ends of host cell proteins (red). The subsequent secretion of Saps leads to hydrolysis of host cell proteins and facilitate the binding of other Als proteins to human tissue. Figure adapted from Salgado et al., 2011

During our large scale reverse genetic study described in chapter 3 we showed that homozygous deletion of *C. albicans* sap8 was significantly less capable of inducing Dar in

nematodes. In this chapter we expanded our study from just *SAP8* to all other Saps and evaluated the contribution of individual genes in this family to the *C. albicans*' ability to induce Dar in nematodes. We showed here that either the pharmacological or genetic inhibition of Saps reduces the *C. albicans* ability to induce Dar and that the deletion of multiple Saps has an additive effect on the phenotype. Encouraged by these results we propose a possible vaccine alternative that, in contrast with what is currently under development (Sandini et al., 2011), doesn't target just one Sap but potentially the entire family.

In order to expand our study on the role of Saps in inducing the Dar phenotype we obtained homozygous knockout mutants in each of the Saps from *sap1*<sup>-/-</sup> to *saps*<sup>-/-</sup> as well as two double mutants (*sap1*<sup>-/-</sup> *sap2*<sup>-/-</sup> and *sap2*<sup>-/-</sup> *sap3*<sup>-/-</sup>) and one triple mutants (*sap1*<sup>-/-</sup> *sap2*<sup>-/-</sup> *sap3*<sup>-/-</sup>). All of these strains were a kind donation of Joachim Morschhauser from the University of Wuzburg and they are described in (Lermann and Morschhauser, 2008). As shown in Figure 0.2 the deletion of each of the Saps tested reduces the *C. albicans* ability to induce the Dar phenotype and the deletion of multiple saps makes the effect even more pronounced. In addition when we added to our assay 200mM of pepstatin A, a proteinase inhibitor previously shown to be very potent against Saps (Schaller et al., 2003), we observed a statistically significant reduction in the *C. albicans* ability to induce Dar. Taken together, our data mirrors what has been shown by others in mammalian system of infections and (a) reiterates the importance of Saps for *C. albicans* virulence (b) validates our approach of studying infection in nematodes.



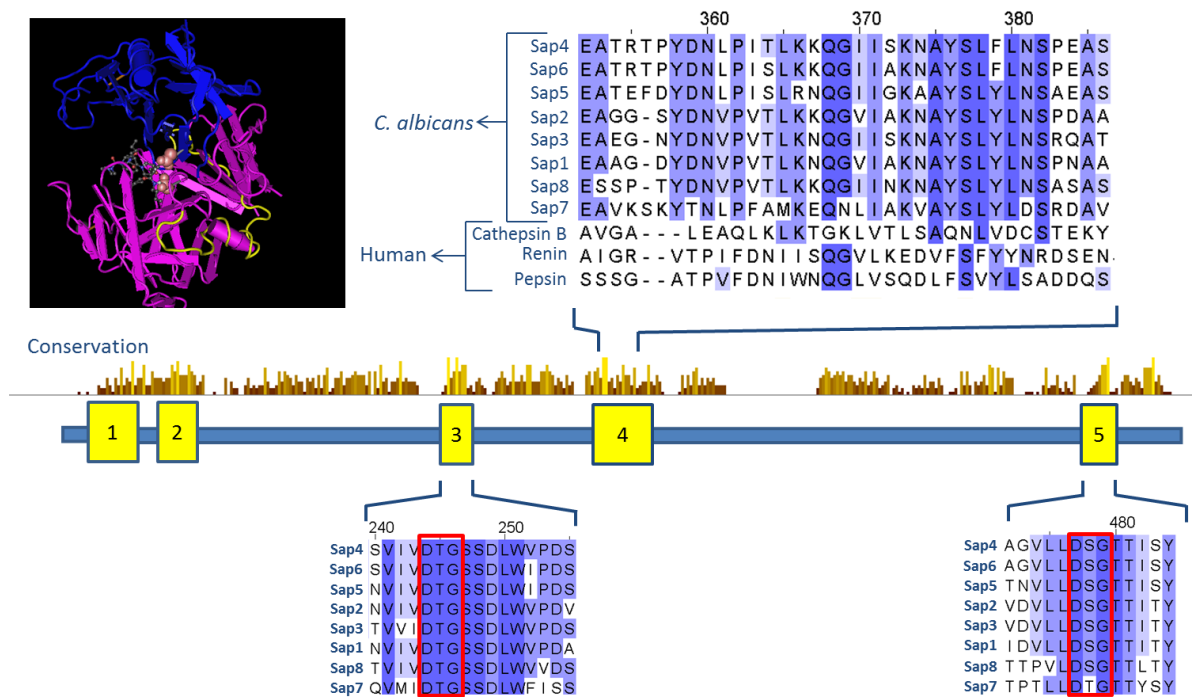
**Figure 0.2 biochemical or genetic inactivation of Saps reduces the pathogen’s ability to induce Dar.**

Single, double and triple *C. albicans* homozygous knockout mutants have a reduced ability in inducing Dar in nematodes (grey and blue bars). The addition of pepstatin A (a specific Saps inhibitor) reduces *C. albicans* wildtype ability to induce the Dar phenotype in nematodes (red bars).

Saps are small, secreted enzymes that have limited homology with human proteins. The combination of these three features makes them ideal candidates for potential drug targeting or for epitope-based vaccine approaches (Cassone, 2013). Indeed, a recombinant truncated version of Sap2 has been expressed *in E. coli* and shown to be effective in protecting mice from *C. albicans* infection when injected either prophylactically or post infection (De Bernardis et al., 2012). A vaccine approach using a recombinant Sap2 delivered through influenza-like virosome have recently entered phase I clinical trial (NCT010671310).

The results from these studies are encouraging however injecting a truncated version of Sap2 is expected to trigger an immunological response that raises antibody specifically for Sap2. Here, we propose an alternative approach in which multiples Saps can potentially be targeted simultaneously. We aligned the amino acid sequences of all secreted Saps with the three human closest homologs (cathepsin B, pepsin A and renin) using the algorithm clustalW

(Larkin et al., 2007). We identified 5 regions that are conserved within all *C. albicans* Saps and noticed that domain 4 (Figure 0.3 top) is highly conserved within *C. albicans* Saps and simultaneously highly divergent to human homologs. This region represents a stretch of 35 amino acids with the following consensus sequence EATGTPYDNPVTLKKQGIIAKNAYSLEYLNSPEAS. When we looked for the location of this sequence on the available Sap2p crystal structure (Cutfield et al., 1995) we found that it lies on a highly exposed region of the protein that is ideal for potential antibody binding (Figure 0.3 top left corner, yellow loop).



**Figure 0.3 Sequence alignment showing the five *C. albicans* Saps conserved domains.**

The five conserved domains are reported in yellow. Domains 3 and 5 contain the two catalytic subunits (highlighted in red) while domain 4 is the domain proposed for potential epitope based vaccine development. Top left corner represents the 3-D structure of Sap2 in which domain 4 has been highlighted in yellow.

We believe that the 35 amino acids long consensus sequence can be recombinantly expressed and potentially used for some preliminary pre-clinical testing in mice. The first tests would

involve toxicological and immunological studies to determine if the sequence can be safely administrated and if it does trigger an immunological response.

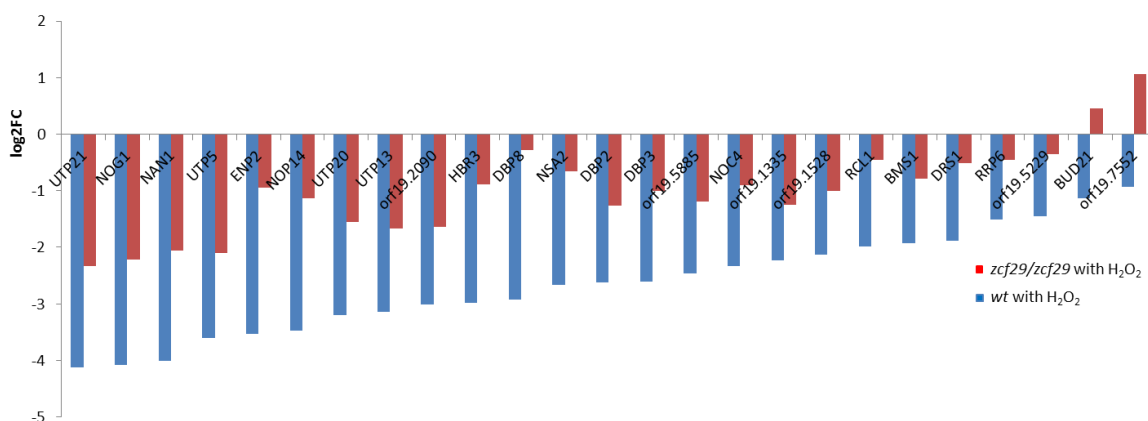
Overall our data shows that Saps are important for triggering the Dar response in nematodes and that the deletion of multiple Saps has an additive effect. We therefore proposed an potential alternative vaccine approach in which multiple Saps can be targeted simultaneously. In order to follow up on this idea further immunological and toxicological studies are required. For example it would be critical to evaluate the immunogenicity of this epitope-based approach and its ability to prevent infection in mouse models of infections.



## 8.2 H<sub>2</sub>O<sub>2</sub> responsive and transcription factor dependent genes: enriched biological functions

In chapter 4.3.4.2 we described that ribosome biogenesis was the most statistically significant enriched biological function between the genes found to be H<sub>2</sub>O<sub>2</sub> responsive and *ZCF29* dependent. Here we report the other enriched biological functions for completion. In Figure 4.17 we defined a cluster of genes called “1, less downregulated in *zcf29/zcf29*”. In addition to ribosome biogenesis this cluster has enriched biological function highly related to ribosome biogenesis like rRNA metabolic process, ribosome small subunit biogenesis, cleavage involved in rRNA processing as well as others like nitrogen compound metabolic process (Figure 8.1)

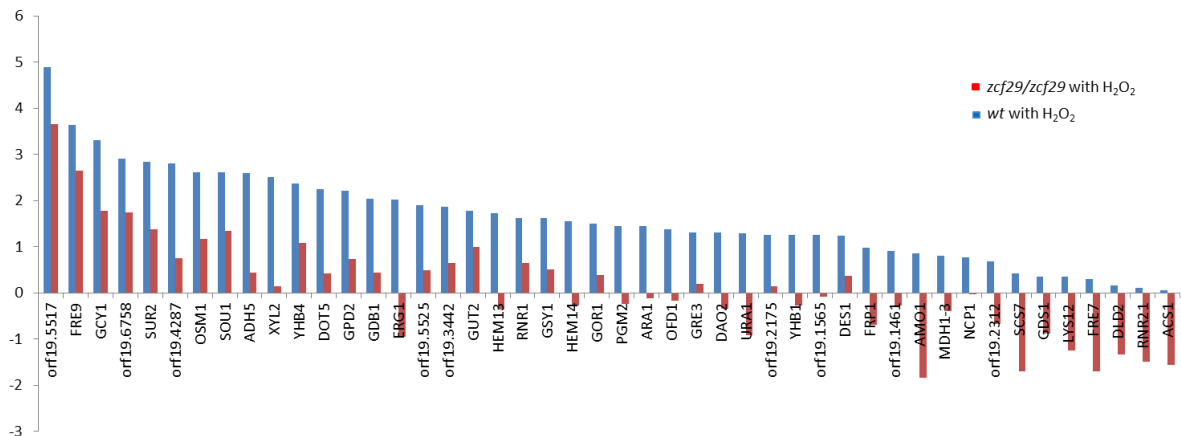
Biological function	Corrected P-value	Gene(s) annotated to the term
ribosome biogenesis	1.38E-18	CIC1 orf19.1335 orf19.1528 UTP21 DBP2 RCL1 NOC4 orf19.2090 BMS1 NAN1 ARX1 UTP20 orf19.4029 UTP13 JIP5 DBP3 orf19.512 orf19.5229 BUD21 RRP6 orf19.5885 REI1 NOP14 orf19.6418 SDA1 DBP8 ENP2 HBR3 NMD3 orf19.7107 NOG1 NSA2 orf19.7552 UTP5 DRS1
rRNA metabolic process	1.51E-12	orf19.1335 orf19.1528 UTP21 DBP2 RCL1 NOC4 orf19.2090 BMS1 NAN1 UTP20 UTP13 DBP3 orf19.5229 BUD21 RRP6 orf19.5885 NOP14 DBP8 ENP2 HBR3 NOG1 NSA2 orf19.7552 UTP5 DRS1
ribosomal small subunit biogenesis	1.37E-07	RCL1 NOC4 orf19.2090 NAN1 UTP20 UTP13 orf19.512 BUD21 orf19.5885 NOP14 DBP8 ENP2 HBR3 orf19.7552 UTP5
cleavage involved in rRNA processing	3.28E-07	orf19.1335 RCL1 NOC4 UTP20 UTP13 DBP3 orf19.5229 BUD21 RRP6 NOP14 DBP8 orf19.7552
RNA phosphodiester bond hydrolysis	5.69E-07	orf19.1335 RCL1 NOC4 UTP20 UTP13 DBP3 orf19.5229 BUD21 RRP6 NOP14 DBP8 orf19.7552
nitrogen compound metabolic process	1.18E-06	orf19.1335 orf19.1528 UTP21 AFG3 DBP2 RPC19 IMH3 RPA190 RCL1 NOC4 RRN3 orf19.2017 orf19.2090 URA2 BMS1 NAN1 RPC53 MET16 UTP20 YAH1 BAS1 orf19.3840 URA7 ECM17 UTP13 MSS116 GUA1 DBP3 TYE7 orf19.500 MET3 orf19.5229 orf19.5334 BUD21 MET15 RRP6 orf19.5847 orf19.5885 NOP14 RPO41 orf19.6297 ADE6 DBP8 ENP2 HBR3 GLN4 orf19.7152 NOG1 ELP3 NSA2 orf19.7552 UTP5 DRS1 MET14 orf19.969
maturation of SSU-rRNA	1.66E-06	RCL1 NOC4 orf19.2090 NAN1 UTP20 UTP13 BUD21 orf19.5885 NOP14 DBP8 HBR3 orf19.7552 UTP5



**Figure 8.1: Biological functions enriched in genes H<sub>2</sub>O<sub>2</sub> downregulated and *ZCF29* dependent**

In Figure 4.17 we defined a cluster of genes called “4, less upregulated in *zcf29/zcf29*”. Biological functions enriched in this group of genes are reported in Figure 8.2. Enriched biological functions include oxidation-reduction process but to a statistically lower degree compared to the biological functions we found in Figure 8.1. In Figure 8.1 we found “ribosome biogenesis” to be the most statistically significant biological function with *pvalue* =  $1.38 \times 10^{-18}$  compared to “oxidation-reduction process” that was enriched with *pvalue* =  $3.1 \times 10^{-4}$  in Figure 8.2.

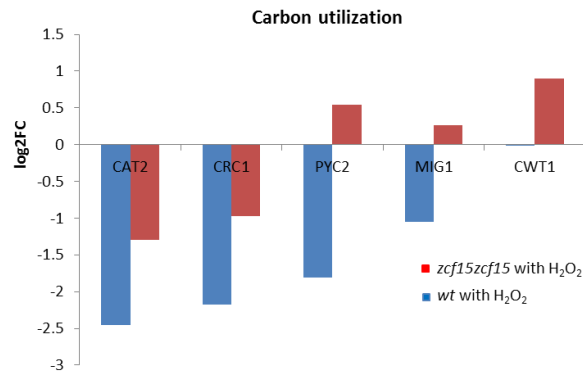
Biological functions	Corrected P-value	Gene(s) annotated to the term
single-organism metabolic process	0.00024	MNN21 HBR2 MRF1 GAD1 LEU42 orf19.1434 orf19.1461 BMT5 orf19.1565 PMS1 ACS1 OFD1 APN2 GDS1 CDC45 DFG10 ARA1 orf19.2175 AAH1 ZCF10 orf19.2312 BIG1 POL32 orf19.2494 GLX3 LYS12 SLD1 ADH5 orf19.2670 NCP1 orf19.276 AMS1 HEM13 PGM2 SOU1 RAD16 GOR1 COO4 orf19.3043 TRP4 GUT2 CTF18 GSY1 orf19.3281 VTC4 DAO2 orf19.338 ADK1 orf19.3442 FRE9 ERG9 orf19.3648 PGK1 YHB1 DCR1 SCS7 HIS1 MUM2 DES1 ERG1 RIB4 orf19.4107 HIS5 RAD52 MNN13 MNN1 orf19.4287 GRE3 orf19.4523 orf19.4581 NAT4 orf19.4699 orf19.4736 HEM14 DAK2 MLS1 URA1 orf19.4864 DAP1 MNN12 orf19.4982 SMI1 CAT8 RAD1 MDH1-3 LAP3 RHD1 DOT5 ILV1 orf19.5517 orf19.5525 BMT6 FRP1 PSF2 PAD1 RNR1 AMO1 SMP3 RNR21 SUR2 orf19.6016 FRE7 ATC1 TPS1 HMO1 orf19.666 PNC1 orf19.6698 DLD2 GCY1 orf19.6758 OSM1 PLB1 GPD2 orf19.6912 orf19.7310 orf19.7404 MMS21 HNT2 GDB1 NTH1 RTT109 YHB4 PRO1 XYL2 BAT21 INO4 THR1
oxidation-reduction process	0.00031	MRF1 orf19.1461 orf19.1565 ACS1 OFD1 GDS1 ARA1 orf19.2175 orf19.2312 LYS12 ADH5 NCP1 HEM13 PGM2 SOU1 GOR1 GUT2 GSY1 DAO2 orf19.3442 FRE9 YHB1 SCS7 DES1 ERG1 orf19.4287 GRE3 HEM14 URA1 MDH1-3 DOT5 orf19.5517 orf19.5525 FRP1 RNR1 AMO1 RNR21 SUR2 FRE7 DLD2 GCY1 orf19.6758 OSM1 GPD2 GDB1 YHB4 XYL2
carbohydrate metabolic process	0.00243	MNN21 CHT4 ARA1 BIG1 AMS1 PGM2 SOU1 GUT2 GSY1 orf19.338 PGK1 MNN13 MNN1 GRE3 DAK2 MLS1 MNN12 orf19.4984 SMI1 MDH1-3 RHD1 ATC1 TPS1 HMO1 HEX1 orf19.6739 DLD2 GCY1 GPD2 orf19.7214 GDB1 NTH1 INO4
glycosphingolipid metabolic process	0.06967	BMT5 SLD1 SCS7 BMT6 SUR2



**Figure 8.2: Biological functions enriched in genes  $H_2O_2$  upregulated and *ZCF29* dependent**

In Figure 4.20 we defined a cluster of genes called “1, less downregulated in *zcf15/zcf15*” and biological functions enriched in this group of genes is reported in Figure 8.3. As we articulated in paragraph 4.3.5 carbon metabolism was the most statistically enriched biological function.

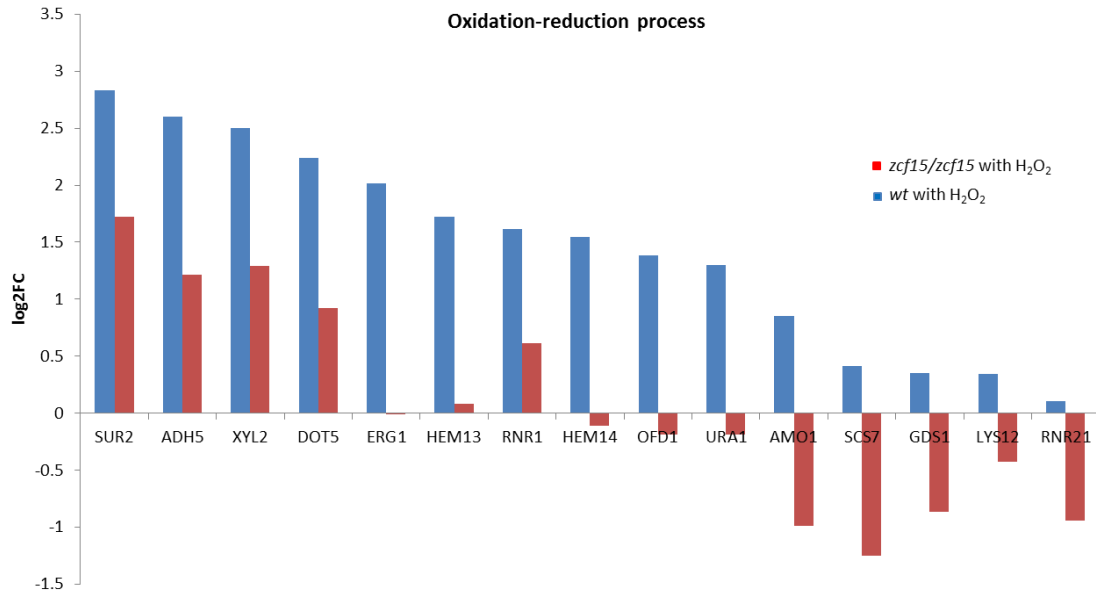
GO_term	Corrected P-value	Gene(s) annotated to the term
carbon utilization	2.48E-05	<i>CRC1 MIG1 CAT2 CWT1 PYC2</i>
single-organism process	0.00011	<i>CPH2 FRP3 FOX2 REG1 CRC1 FAA21 RGT1 DIP5 LYS2 orf19.3406 MAE1 SHA3 orf19.3689 MET10 ECM17 FET3 MIG1 CIT1 CAT2 RGA2 PTC8 TYE7 MET3 QDR1 SDS24 orf19.5334 GAL4 PLC2 ALD5 CWT1 CDR1 KGD2 orf19.6423 orf19.6435 CUP9 NRG1 MDH1 orf19.7670 PYC2 HMS1 orf19.93</i>
carboxylic acid metabolic process	0.00022	<i>FOX2 CRC1 FAA21 LYS2 MAE1 ECM17 FET3 CIT1 CAT2 MET3 ALD5 KGD2 MDH1</i>
carbon catabolite regulation of transcription	0.00143	<i>MIG1 TYE7 GAL4 CWT1 NRG1</i>



**Figure 8.3: Biological functions enriched in genes H<sub>2</sub>O<sub>2</sub> downregulated and *ZCF15* dependent**

In Figure 4.20 we defined a cluster of genes called “4, less upregulated in *zcf15/zcf15*” and biological functions enriched in this cluster is reported in Figure 8.4.

Biological function	Corrected P-value	Gene(s) annotated to the term
oxidation-reduction process	0.0001	OFD1 GDS1 LYS12 ADH5 HEM13 SCS7 ERG1 HEM14 URA1 DOT5 RNR1 AMO1 RNR21 SUR2 XYL2
single-organism metabolic process	0.00168	orf19.1434 OFD1 GDS1 AAH1 LYS12 SLD1 ADH5 HEM13 SCS7 HIS1 ERG1 MNN1 orf19.4581 NAT4 HEM14 URA1 DOT5 PSF2 RNR1 AMO1 RNR21 SUR2 ATC1 PLB1 GPI1 NTH1 XYL2 GPM1
single-organism biosynthetic process	0.00343	orf19.1434 AAH1 LYS12 SLD1 HEM13 SCS7 HIS1 ERG1 orf19.4581 HEM14 URA1 RNR1 RNR21 SUR2 GPI1 GPM1
glycolipid metabolic process	0.00357	SLD1 SCS7 orf19.4581 SUR2 GPI1
organophosphate metabolic process	0.00931	orf19.1889 AAH1 UBA4 SCS7 orf19.4581 URA1 RNR1 RNR21 SUR2 PLB1 GPI1



**Figure 8.4: Biological functions enriched in genes H<sub>2</sub>O<sub>2</sub> upregulated and *ZCF15* dependent**

## REFERENCES

- Aballay, A., Drenkard, E., Hilbun, L.R., and Ausubel, F.M. (2003). *Caenorhabditis elegans* innate immune response triggered by *Salmonella enterica* requires intact LPS and is mediated by a MAPK signaling pathway. In *Curr Biol (England)*, pp. 47-52.
- Aballay, A., Yorgey, P., and Ausubel, F.M. (2000). *Salmonella typhimurium* proliferates and establishes a persistent infection in the intestine of *Caenorhabditis elegans*. In *Curr Biol (England)*, pp. 1539-1542.
- Adams, M.D., Celniker, S.E., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Scherer, S.E., Li, P.W., Hoskins, R.A., Galle, R.F., *et al.* (2000). The genome sequence of *Drosophila melanogaster*. In *Science (United States)*, pp. 2185-2195.
- Alarco, A.M., Marcil, A., Chen, J., Suter, B., Thomas, D., and Whiteway, M. (2004). Immune-deficient *Drosophila melanogaster*: a model for the innate immune response to human fungal pathogens. *J Immunol* 172, 5622-5628.
- Alby, K., and Bennett, R.J. (2009). Stress-induced phenotypic switching in *Candida albicans*. *Mol Biol Cell* 20, 3178-3191.
- Aleixo, M.J., Caldeira, L., and Ferreira, M.L. (2000). *Candida albicans* meningitis: clinical case. *J Infect* 40, 191-192.
- Alper, S., McBride, S.J., Lackford, B., Freedman, J.H., and Schwartz, D.A. (2007). Specificity and complexity of the *Caenorhabditis elegans* innate immune response. *Mol Cell Biol* 27, 5544-5553.
- Anaissie, E.J., McGinnis, M.R., and Pfaller, M.A. (2009). *Clinical mycology*, 2nd edn (Edinburgh: Churchill Livingstone).
- Anders, S., Reyes, A., and Huber, W. (2012). Detecting differential usage of exons from RNA-seq data. *Genome Res* 22, 2008-2017.
- Antinori, S. (2013). New Insights into HIV/AIDS-Associated Cryptococcosis. *ISRN AIDS* 2013, 471363.
- Aoki, W., Kitahara, N., Miura, N., Morisaka, H., Yamamoto, Y., Kuroda, K., and Ueda, M. (2011). Comprehensive characterization of secreted aspartic proteases encoded by a virulence gene family in *Candida albicans*. *J Biochem* 150, 431-438.
- Asai, K., Tsuchimori, N., Okonogi, K., Perfect, J.R., Gotoh, O., and Yoshida, Y. (1999). Formation of azole-resistant *Candida albicans* by mutation of sterol 14-demethylase P450. *Antimicrob Agents Chemother* 43, 1163-1169.
- Ashman, R.B., Farah, C.S., Wanasaengsakul, S., Hu, Y., Pang, G., and Clancy, R.L. (2004). Innate versus adaptive immunity in *Candida albicans* infection. In *Immunol Cell Biol (Australia)*, pp. 196-204.
- Bader, T., Bodendorfer, B., Schroppel, K., and Morschhauser, J. (2003). Calcineurin is essential for virulence in *Candida albicans*. *Infect Immun* 71, 5344-5354.
- Bae, T., Banger, A.K., Wallace, A., Glass, E.M., Aslund, F., Schneewind, O., and Missiakas, D.M. (2004). *Staphylococcus aureus* virulence genes identified by bursa aurealis mutagenesis and nematode killing. *Proc Natl Acad Sci U S A* 101, 12312-12317.
- Baginski, M., and Czub, J. (2009). Amphotericin B and its new derivatives - mode of action. *Curr Drug Metab* 10, 459-469.
- Bailey, T., Krajewski, P., Ladunga, I., Lefebvre, C., Li, Q., Liu, T., Madrigal, P., Taslim, C., and Zhang, J. (2013). Practical guidelines for the comprehensive analysis of ChIP-seq data. *PLoS Comput Biol* 9, e1003326.

## REFERENCES

- Bates, S., de la Rosa, J.M., MacCallum, D.M., Brown, A.J., Gow, N.A., and Odds, F.C. (2007). *Candida albicans* Iff11, a secreted protein required for cell wall structure and virulence. *Infect Immun* 75, 2922-2928.
- Bengtson, M.H., and Joazeiro, C.A. (2010). Role of a ribosome-associated E3 ubiquitin ligase in protein quality control. *Nature* 467, 470-473.
- Bergin, D., Reeves, E.P., Renwick, J., Wientjes, F.B., and Kavanagh, K. (2005). Superoxide production in *Galleria mellonella* hemocytes: identification of proteins homologous to the NADPH oxidase complex of human neutrophils. *Infect Immun* 73, 4161-4170.
- Berman, J., and Sudbery, P.E. (2002). *Candida Albicans*: a molecular revolution built on lessons from budding yeast. *Nature reviews Genetics* 3, 918-930.
- Bessa, D., Pereira, F., Moreira, R., Johansson, B., and Queiros, O. (2012). Improved gap repair cloning in yeast: treatment of the gapped vector with Taq DNA polymerase avoids vector self-ligation. *Yeast* 29, 419-423.
- Bjorkman, J., and Andersson, D.I. (2000). The cost of antibiotic resistance from a bacterial perspective. *Drug Resist Updat* 3, 237-245.
- Blackwell, M. (2011). The fungi: 1, 2, 3 ... 5.1 million species? In *Am J Bot* (United States), pp. 426-438.
- Blankenship, J.R., Fanning, S., Hamaker, J.J., and Mitchell, A.P. (2010). An extensive circuitry for cell wall regulation in *Candida albicans*. *PLoS Pathog* 6, e1000752.
- Blankenship, J.R., and Mitchell, A.P. (2011). *Candida albicans* adds more weight to iron regulation. In *Cell Host Microbe* (United States: 2011 Elsevier Inc), pp. 93-94.
- Bonangelino, C.J., Chavez, E.M., and Bonifacino, J.S. (2002). Genomic screen for vacuolar protein sorting genes in *Saccharomyces cerevisiae*. *Mol Biol Cell* 13, 2486-2501.
- Botcheva, K., McCorkle, S.R., McCombie, W.R., Dunn, J.J., and Anderson, C.W. (2011). Distinct p53 genomic binding patterns in normal and cancer-derived human cells. *Cell Cycle* 10, 4237-4249.
- Brachmann, C.B., Davies, A., Cost, G.J., Caputo, E., Li, J., Hieter, P., and Boeke, J.D. (1998). Designer deletion strains derived from *Saccharomyces cerevisiae* S288C: a useful set of strains and plasmids for PCR-mediated gene disruption and other applications. In *Yeast* (England), pp. 115-132.
- Brand, A., MacCallum, D.M., Brown, A.J., Gow, N.A., and Odds, F.C. (2004). Ectopic expression of *URA3* can influence the virulence phenotypes and proteome of *Candida albicans* but can be overcome by targeted reintegration of *URA3* at the *RPS10* locus. *Eukaryot Cell* 3, 900-909.
- Brandman, O., Stewart-Ornstein, J., Wong, D., Larson, A., Williams, C.C., Li, G.W., Zhou, S., King, D., Shen, P.S., Weibezahn, J., *et al.* (2012). A ribosome-bound quality control complex triggers degradation of nascent peptides and signals translation stress. *Cell* 151, 1042-1054.
- Braun, B.R., and Johnson, A.D. (2000). TUP1, CPH1 and EFG1 make independent contributions to filamentation in *Candida albicans*. *Genetics* 155, 57-67.
- Brennan, M., Thomas, D.Y., Whiteway, M., and Kavanagh, K. (2002). Correlation between virulence of *Candida albicans* mutants in mice and *Galleria mellonella* larvae. In *FEMS Immunol Med Microbiol* (Netherlands), pp. 153-157.
- Brenner, S. (1974). The genetics of *Caenorhabditis elegans*. *Genetics* 77, 71-94.

## REFERENCES

- Brock, M. (2009). Fungal metabolism in host niches. In *Curr Opin Microbiol* (England), pp. 371-376.
- Brothers, K.M., Newman, Z.R., and Wheeler, R.T. (2011). Live imaging of disseminated candidiasis in zebrafish reveals role of phagocyte oxidase in limiting filamentous growth. *Eukaryot Cell* *10*, 932-944.
- Brunet, J.P., Tamayo, P., Golub, T.R., and Mesirov, J.P. (2004). Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A* *101*, 4164-4169.
- Bruno, V.M., Wang, Z., Marjani, S.L., Euskirchen, G.M., Martin, J., Sherlock, G., and Snyder, M. (2010). Comprehensive annotation of the transcriptome of the human fungal pathogen *Candida albicans* using RNA-seq. *Genome Res* *20*, 1451-1458.
- Buchacz, K., Baker, R.K., Palella, F.J., Jr., Chmiel, J.S., Lichtenstein, K.A., Novak, R.M., Wood, K.C., and Brooks, J.T. (2010). AIDS-defining opportunistic illnesses in US patients, 1994-2007: a cohort study. *AIDS* *24*, 1549-1559.
- Bus, J.S., Aust, S.D., and Gibson, J.E. (1976). Paraquat toxicity: proposed mechanism of action involving lipid peroxidation. *Environ Health Perspect* *16*, 139-146.
- Butler, G., Rasmussen, M.D., Lin, M.F., Santos, M.A., Sakthikumar, S., Munro, C.A., Rheinbay, E., Grabherr, M., Forche, A., Reedy, J.L., *et al.* (2009). Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. *Nature* *459*, 657-662.
- Campbell, R.N., Leverenz, M.K., Ryan, L.A., and Reece, R.J. (2008). Metabolic control of transcription: paradigms and lessons from *Saccharomyces cerevisiae*. *Biochem J* *414*, 177-187.
- Candidaeffects (2013). Symptoms and treatment for *Candida* yeast skin infection (<http://candidaeffects.com/symptoms-and-treatment-for-candida-yeast-skin-infection/>; last accessed 7/21/2014).
- Cannon, R.D., Lamping, E., Holmes, A.R., Niimi, K., Tanabe, K., Niimi, M., and Monk, B.C. (2007). *Candida albicans* drug resistance another way to cope with stress. *Microbiology* *153*, 3211-3217.
- Casadevall, A. (2007). Determinants of virulence in the pathogenic fungi. *Fungal Biol Rev* *21*, 130-132.
- Cassone, A. (2013). Development of vaccines for *Candida albicans*: fighting a skilled transformer. In *Nat Rev Microbiol* (England), pp. 884-891.
- Castell, J.V., Donato, M.T., and Gomez-Lechon, M.J. (2005). Metabolism and bioactivation of toxicants in the lung. The in vitro cellular approach. *Exp Toxicol Pathol* *57 Suppl 1*, 189-204.
- Chai, L.Y., Denning, D.W., and Warn, P. (2010). *Candida tropicalis* in human disease. *Crit Rev Microbiol* *36*, 282-298.
- Chalfie, M., Tu, Y., Euskirchen, G., Ward, W.W., and Prasher, D.C. (1994). Green fluorescent protein as a marker for gene expression. *Science* *263*, 802-805.
- Chamilos, G., Lionakis, M.S., Lewis, R.E., Lopez-Ribot, J.L., Saville, S.P., Albert, N.D., Halder, G., and Kontoyiannis, D.P. (2006). *Drosophila melanogaster* as a facile model for large-scale studies of virulence mechanisms and antifungal drug efficacy in *Candida* species. In *J Infect Dis* (United States), pp. 1014-1022.
- Chandra, J., Kuhn, D.M., Mukherjee, P.K., Hoyer, L.L., McCormick, T., and Ghannoum, M.A. (2001). Biofilm formation by the fungal pathogen *Candida albicans*: development, architecture, and drug resistance. *J Bacteriol* *183*, 5385-5394.

## REFERENCES

- Chang, S.L., Lai, H.Y., Tung, S.Y., and Leu, J.Y. (2013). Dynamic large-scale chromosomal rearrangements fuel rapid adaptation in yeast populations. *PLoS Genet* 9, e1003232.
- Charles, P.E., Doise, J.M., Quenot, J.P., Aube, H., Dalle, F., Chavanet, P., Milesi, N., Aho, L.S., Portier, H., and Blettery, B. (2003). Candidemia in critically ill patients: difference of outcome between medical and surgical patients. *Intensive Care Med* 29, 2162-2169.
- Chauhan, N., Latge, J.P., and Calderone, R. (2006). Signalling and oxidant adaptation in *Candida albicans* and *Aspergillus fumigatus*. *Nat Rev Microbiol* 4, 435-444.
- Chavez, V., Mohri-Shiomi, A., Maadani, A., Vega, L.A., and Garsin, D.A. (2007). Oxidative stress enzymes are required for DAF-16-mediated immunity due to generation of reactive oxygen species by *Caenorhabditis elegans*. *Genetics* 176, 1567-1577.
- Chen, C., Pande, K., French, S.D., Tuch, B.B., and Noble, S.M. (2011). An iron homeostasis regulatory circuit with reciprocal roles in *Candida albicans* commensalism and pathogenesis. *Cell Host Microbe* 10, 118-135.
- Chen, C.G., Yang, Y.L., Cheng, H.H., Su, C.L., Huang, S.F., Chen, C.T., Liu, Y.T., Su, I.J., and Lo, H.J. (2006). Non-lethal *Candida albicans* cph1/cph1 efg1/efg1 transcription factor mutant establishing restricted zone of infection in a mouse model of systemic infection. In *Int J Immunopathol Pharmacol (Italy)*, pp. 561-565.
- Chen, Z., and Kong, X. (2007). Study of *Candida albicans* vaginitis model in Kunming mice. *J Huazhong Univ Sci Technol Med Sci* 27, 307-310.
- Chu, J., Hong, N.A., Masuda, C.A., Jenkins, B.V., Nelms, K.A., Goodnow, C.C., Glynne, R.J., Wu, H., Masliah, E., Joazeiro, C.A., *et al.* (2009). A mouse forward genetics screen identifies LISTERIN as an E3 ubiquitin ligase involved in neurodegeneration. *Proc Natl Acad Sci U S A* 106, 2097-2103.
- Chung, C.H., and Baek, S.H. (1999). Deubiquitinating enzymes: their diversity and emerging roles. In *Biochem Biophys Res Commun (United States: 1999 Academic Press.)*, pp. 633-640.
- Cinar, H.N., Kothary, M., Datta, A.R., Tall, B.D., Sprando, R., Bilecen, K., Yildiz, F., and McCardell, B. (2010). *Vibrio cholerae* hemolysin is required for lethality, developmental delay, and intestinal vacuolation in *Caenorhabditis elegans*. *PLoS One* 5, e11558.
- Clancy, C.J., Cheng, S., and Nguyen, M.H. (2009). Animal models of candidiasis. *Methods Mol Biol* 499, 65-76.
- Connolly, L.A., Riccombeni, A., Grozer, Z., Holland, L.M., Lynch, D.B., Andes, D.R., Gacser, A., and Butler, G. (2013). The APSES transcription factor Efg1 is a global regulator that controls morphogenesis and biofilm formation in *Candida parapsilosis*. *Mol Microbiol* 90, 36-53.
- Cooke, M.S., Evans, M.D., Dizdaroglu, M., and Lunec, J. (2003). Oxidative DNA damage: mechanisms, mutation, and disease. In *FASEB J (United States)*, pp. 1195-1214.
- Coste, A., Turner, V., Ischer, F., Morschhauser, J., Forche, A., Selmecki, A., Berman, J., Bille, J., and Sanglard, D. (2006). A mutation in Tac1p, a transcription factor regulating CDR1 and CDR2, is coupled with loss of heterozygosity at chromosome 5 to mediate antifungal resistance in *Candida albicans*. *Genetics* 172, 2139-2156.
- Coste, A.T., Ramsdale, M., Ischer, F., and Sanglard, D. (2008). Divergent functions of three *Candida albicans* zinc-cluster transcription factors (CTA4, ASG1 and CTF1)



## REFERENCES

- complementing pleiotropic drug resistance in *Saccharomyces cerevisiae*. *Microbiology* 154, 1491-1501.
- Cotter, G., Doyle, S., and Kavanagh, K. (2000). Development of an insect model for the in vivo pathogenicity testing of yeasts. In *FEMS Immunol Med Microbiol* (Netherlands), pp. 163-169.
  - Couillault, C., Pujol, N., Reboul, J., Sabatier, L., Guichou, J.F., Kohara, Y., and Ewbank, J.J. (2004). TLR-independent control of innate immunity in *Caenorhabditis elegans* by the TIR domain adaptor protein TIR-1, an ortholog of human SARM. In *Nat Immunol* (United States), pp. 488-494.
  - Cowen, L.E. (2008). The evolution of fungal drug resistance: modulating the trajectory from genotype to phenotype. In *Nat Rev Microbiol* (England), pp. 187-198.
  - Cowen, L.E., Anderson, J.B., and Kohn, L.M. (2002). Evolution of drug resistance in *Candida albicans*. *Annu Rev Microbiol* 56, 139-165.
  - Cowen, L.E., and Lindquist, S. (2005). Hsp90 potentiates the rapid evolution of new traits: drug resistance in diverse fungi. *Science* 309, 2185-2189.
  - Craig, N.L. (1991). Tn7: a target site-specific transposon. *Mol Microbiol* 5, 2569-2573.
  - Cutfield, S.M., Dodson, E.J., Anderson, B.F., Moody, P.C., Marshall, C.J., Sullivan, P.A., and Cutfield, J.F. (1995). The crystal structure of a major secreted aspartic proteinase from *Candida albicans* in complexes with two inhibitors. *Structure* 3, 1261-1271.
  - Dagenais, T.R., and Keller, N.P. (2009). Pathogenesis of *Aspergillus fumigatus* in Invasive Aspergillosis. *Clin Microbiol Rev* 22, 447-465.
  - Davis, D.A., Bruno, V.M., Loza, L., Filler, S.G., and Mitchell, A.P. (2002). *Candida albicans* Mds3p, a conserved regulator of pH responses and virulence identified through insertional mutagenesis. *Genetics* 162, 1573-1581.
  - De Bernardis, F., Amacker, M., Arancia, S., Sandini, S., Gremion, C., Zurbriggen, R., Moser, C., and Cassone, A. (2012). A virosomal vaccine against candidal vaginitis: immunogenicity, efficacy and safety profile in animal models. In *Vaccine* (Netherlands: 2012 Elsevier Ltd), pp. 4490-4498.
  - de Groot, P.W., de Boer, A.D., Cunningham, J., Dekker, H.L., de Jong, L., Hellingwerf, K.J., de Koster, C., and Klis, F.M. (2004). Proteomic analysis of *Candida albicans* cell walls reveals covalently bound carbohydrate-active enzymes and adhesins. *Eukaryot Cell* 3, 955-965.
  - Delley, P.A., and Hall, M.N. (1999). Cell wall stress depolarizes cell growth via hyperactivation of RHO1. *J Cell Biol* 147, 163-174.
  - DeLuca, D.S., Levin, J.Z., Sivachenko, A., Fennell, T., Nazaire, M.D., Williams, C., Reich, M., Winckler, W., and Getz, G. (2012). RNA-SeQC: RNA-seq metrics for quality control and process optimization. *Bioinformatics* 28, 1530-1532.
  - Dhamgaye, S., Bernard, M., Lelandais, G., Sismeiro, O., Lemoine, S., Coppee, J.Y., Le Crom, S., Prasad, R., and Devaux, F. (2012). RNA sequencing revealed novel actors of the acquisition of drug resistance in *Candida albicans*. *BMC Genomics* 13, 396.
  - Dieterich, C., Schandar, M., Noll, M., Johannes, F.J., Brunner, H., Graeve, T., and Rupp, S. (2002). In vitro reconstructed human epithelia reveal contributions of *Candida albicans* EFG1 and CPH1 to adhesion and invasion. *Microbiology* 148, 497-506.
  - Ding, L., Ley, T.J., Larson, D.E., Miller, C.A., Koboldt, D.C., Welch, J.S., Ritchey, J.K., Young, M.A., Lamprecht, T., McLellan, M.D., *et al.* (2012). Clonal evolution in relapsed acute myeloid leukaemia revealed by whole-genome sequencing. *Nature* 481, 506-510.

## REFERENCES

- Donini, M., Zenaro, E., Tamassia, N., and Dusi, S. (2007). NADPH oxidase of human dendritic cells: role in *Candida albicans* killing and regulation by interferons, dectin-1 and CD206. *Eur J Immunol* 37, 1194-1203.
- Dunkel, N., Blass, J., Rogers, P.D., and Morschhauser, J. (2008). Mutations in the multi-drug resistance regulator MRR1, followed by loss of heterozygosity, are the main cause of MDR1 overexpression in fluconazole-resistant *Candida albicans* strains. *Mol Microbiol* 69, 827-840.
- Dunkel, N., and Morschhauser, J. (2011). Loss of heterozygosity at an unlinked genomic locus is responsible for the phenotype of a *Candida albicans* sap4Delta sap5Delta sap6Delta mutant. *Eukaryot Cell* 10, 54-62.
- Ellis, D. (2002). Amphotericin B: spectrum and resistance. *J Antimicrob Chemother* 49 Suppl 1, 7-10.
- Ellis, H.M., and Horvitz, H.R. (1986). Genetic control of programmed cell death in the nematode *C. elegans*. In *Cell* (United States), pp. 817-829.
- Ene, I.V., Adya, A.K., Wehmeier, S., Brand, A.C., MacCallum, D.M., Gow, N.A., and Brown, A.J. (2012). Host carbon sources modulate cell wall architecture, drug resistance and virulence in a fungal pathogen. *Cell Microbiol* 14, 1319-1335.
- Engelmann, I., and Pujol, N. (2010). Innate immunity in *C. elegans*. *Adv Exp Med Biol* 708, 105-121.
- Enjalbert, B., Smith, D.A., Cornell, M.J., Alam, I., Nicholls, S., Brown, A.J., and Quinn, J. (2006). Role of the Hog1 stress-activated protein kinase in the global transcriptional response to stress in the fungal pathogen *Candida albicans*. *Mol Biol Cell* 17, 1018-1032.
- Enloe, B., Diamond, A., and Mitchell, A.P. (2000). A single-transformation gene function test in diploid *Candida albicans*. *J Bacteriol* 182, 5730-5736.
- Eriksson, U., Seifert, B., and Schaffner, A. (2001). Comparison of effects of amphotericin B deoxycholate infused over 4 or 24 hours: randomised controlled trial. *BMJ* 322, 579-582.
- Evans, E.A., Kawli, T., and Tan, M.W. (2008). *Pseudomonas aeruginosa* suppresses host immunity by activating the DAF-2 insulin-like signaling pathway in *Caenorhabditis elegans*. *PLoS Pathog* 4, e1000175.
- Ewbank, J.J. (2002). Tackling both sides of the host-pathogen equation with *Caenorhabditis elegans*. In *Microbes Infect* (France), pp. 247-256.
- Fallon, K., Bausch, K., Noonan, J., Huguenel, E., and Tamburini, P. (1997). Role of aspartic proteases in disseminated *Candida albicans* infection in mice. *Infect Immun* 65, 551-556.
- Fan, W., Kraus, P.R., Boily, M.J., and Heitman, J. (2005). *Cryptococcus neoformans* gene expression during murine macrophage infection. *Eukaryot Cell* 4, 1420-1433.
- Fares, H., and Greenwald, I. (2001). Genetic analysis of endocytosis in *Caenorhabditis elegans*: coelomocyte uptake defective mutants. *Genetics* 159, 133-145.
- Fazly, A., Jain, C., Dehner, A.C., Issi, L., Lilly, E.A., Ali, A., Cao, H., Fidel, P.L., Jr., R, P.R., and Kaufman, P.D. (2013). Chemical screening identifies filastatin, a small molecule inhibitor of *Candida albicans* adhesion, morphogenesis, and pathogenesis. In *Proc Natl Acad Sci U S A*.
- Ferreira, C., Silva, S., Faria-Oliveira, F., Pinho, E., Henriques, M., and Lucas, C. (2010). *Candida albicans* virulence and drug-resistance requires the O-acyltransferase Gup1p. *BMC Microbiol* 10, 238.

## REFERENCES

- Fidel, P.L., Jr., Vazquez, J.A., and Sobel, J.D. (1999). *Candida glabrata*: review of epidemiology, pathogenesis, and clinical disease with comparison to *C. albicans*. *Clin Microbiol Rev* 12, 80-96.
- Finkel, J.S., and Mitchell, A.P. (2010). Genetic control of *Candida albicans* biofilm development. *Nat Rev Microbiol* 9, 109-118.
- Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., and Mello, C.C. (1998). Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* 391, 806-811.
- Flattery, A.M., Abruzzo, G.K., Gill, C.J., Smith, J.G., and Bartizal, K. (1996). New model of oropharyngeal and gastrointestinal colonization by *Candida albicans* in CD4+ T-cell-deficient mice for evaluation of antifungal agents. *Antimicrob Agents Chemother* 40, 1604-1609.
- Flood, P.M., Qian, L., Peterson, L.J., Zhang, F., Shi, J.S., Gao, H.M., and Hong, J.S. (2011). Transcriptional Factor NF-kappaB as a Target for Therapy in Parkinson's Disease. *Parkinsons Dis* 2011, 216298.
- Forche, A., Abbey, D., Pisithkul, T., Weinzierl, M.A., Ringstrom, T., Bruck, D., Petersen, K., and Berman, J. (2011). Stress alters rates and types of loss of heterozygosity in *Candida albicans*. *MBio* 2.
- Forche, A., Alby, K., Schaefer, D., Johnson, A.D., Berman, J., and Bennett, R.J. (2008). The parasexual cycle in *Candida albicans* provides an alternative pathway to meiosis for the formation of recombinant strains. *PLoS Biol* 6, e110.
- Forche, A., Magee, P.T., Magee, B.B., and May, G. (2004). Genome-wide single-nucleotide polymorphism map for *Candida albicans*. *Eukaryot Cell* 3, 705-714.
- Forche, A., Magee, P.T., Selmecki, A., Berman, J., and May, G. (2009). Evolution in *Candida albicans* populations during a single passage through a mouse host. *Genetics* 182, 799-811.
- Forche, A., May, G., and Magee, P.T. (2005). Demonstration of loss of heterozygosity by single-nucleotide polymorphism microarray analysis and alterations in strain morphology in *Candida albicans* strains during infection. *Eukaryot Cell* 4, 156-165.
- Galhardo, R.S., Hastings, P.J., and Rosenberg, S.M. (2007). Mutation as a stress response and the regulation of evolvability. *Crit Rev Biochem Mol Biol* 42, 399-435.
- Garcia, M.G., O'Connor, J.E., Garcia, L.L., Martinez, S.I., Herrero, E., and del Castillo Agudo, L. (2001). Isolation of a *Candida albicans* gene, tightly linked to URA3, coding for a putative transcription factor that suppresses a *Saccharomyces cerevisiae* aft1 mutation. In *Yeast* (England: 2000 John Wiley & Sons, Ltd.), pp. 301-311.
- Garcia-Solache, M.A., and Casadevall, A. (2010). Global warming will bring new fungal diseases for mammals. *MBio* 1.
- Gemayel, R., Vincens, M.D., Legendre, M., and Verstrepen, K.J. (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu Rev Genet* 44, 445-477.
- Gerami-Nejad, M., Dulmage, K., and Berman, J. (2009). Additional cassettes for epitope and fluorescent fusion proteins in *Candida albicans*. *Yeast* 26, 399-406.
- Gerami-Nejad, M., Zacchi, L.F., McClellan, M., Matter, K., and Berman, J. (2013). Shuttle vectors for facile gap repair cloning and integration into a neutral locus in *Candida albicans*. *Microbiology* 159, 565-579.
- Ghosh, D., and Papavassiliou, A.G. (2005). Transcription factor therapeutics: long-shot or lodestone. *Curr Med Chem* 12, 691-701.

## REFERENCES

- Ghosh, S., Kebaara, B.W., Atkin, A.L., and Nickerson, K.W. (2008). Regulation of aromatic alcohol production in *Candida albicans*. *Appl Environ Microbiol* 74, 7211-7218.
- Gillum, A.M., Tsay, E.Y., and Kirsch, D.R. (1984). Isolation of the *Candida albicans* gene for orotidine-5'-phosphate decarboxylase by complementation of *S. cerevisiae* *ura3* and *E. coli* *pyrF* mutations. *Mol Gen Genet* 198, 179-182.
- Goebels, C., Thonn, A., Gonzalez-Hilarion, S., Rolland, O., Moyrand, F., Beilharz, T.H., and Janbon, G. (2013). Introns regulate gene expression in *Cryptococcus neoformans* in a *Pab2p* dependent pathway. *PLoS Genet* 9, e1003686.
- Goecks, J., Nekrutenko, A., and Taylor, J. (2010). Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biol* 11, R86.
- Gow, N.A., and Hube, B. (2012). Importance of the *Candida albicans* cell wall during commensalism and infection. *Curr Opin Microbiol* 15, 406-412.
- Goyard, S., Knechtle, P., Chauvel, M., Mallet, A., Prevost, M.C., Proux, C., Coppee, J.Y., Schwarz, P., Dromer, F., Park, H., *et al.* (2008). The *Yak1* kinase is involved in the initiation and maintenance of hyphal growth in *Candida albicans*. *Mol Biol Cell* 19, 2251-2266.
- Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., *et al.* (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29, 644-652.
- Green, B., Bouchier, C., Fairhead, C., Craig, N.L., and Cormack, B.P. (2012). Insertion site preference of *Mu*, *Tn5*, and *Tn7* transposons. *Mob DNA* 3, 3.
- Gresham, D., Desai, M.M., Tucker, C.M., Jenq, H.T., Pai, D.A., Ward, A., DeSevo, C.G., Botstein, D., and Dunham, M.J. (2008). The Repertoire and Dynamics of Evolutionary Adaptations to Controlled Nutrient-Limited Environments in Yeast. *PLoS Genet* 4.
- Gudlaugsson, O., Gillespie, S., Lee, K., Vande Berg, J., Hu, J., Messer, S., Herwaldt, L., Pfaller, M., and Diekema, D. (2003a). Attributable mortality of nosocomial candidemia, revisited. *Clin Infect Dis* 37, 1172-1177.
- Gudlaugsson, O., Gillespie, S., Lee, K., Vande Berg, J., Hu, J., Messer, S., Herwaldt, L., Pfaller, M., and Diekema, D. (2003b). Attributable mortality of nosocomial candidemia, revisited. *Clin Infect Dis* 37, 1172-1177.
- Haas, A. (2007). The phagosome: compartment with a license to kill. *Traffic* 8, 311-330.
- Hampton, T. (2013). Report reveals scope of US antibiotic resistance threat. In *JAMA (United States)*, pp. 1661-1663.
- Hasim, S., Hussin, N.A., Alomar, F., Bidasee, K.R., Nickerson, K.W., and Wilson, M.A. (2014). A glutathione-independent glyoxalase of the DJ-1 superfamily plays an important role in managing metabolically generated methylglyoxal in *Candida albicans*. *J Biol Chem* 289, 1662-1674.
- Havlickova, B., Czaika, V.A., and Friedrich, M. (2009). Epidemiological trends in skin mycoses worldwide (vol 51, Suppl 4, pg 2, 2008). *Mycoses* 52.
- Hawkins, J.L., and Baddour, L.M. (2003). *Candida lusitanae* infections in the era of fluconazole availability. In *Clin Infect Dis (United States)*, pp. e14-18.
- Hawser, S., and Islam, K. (1999). Comparisons of the effects of fungicidal and fungistatic antifungal agents on the morphogenetic transformation of *Candida albicans*. *J Antimicrob Chemother* 43, 411-413.

## REFERENCES

- Heilman, J. (2008). Human tongue with oral candidiasis.
- Heilman, J. (2011). FeetFungal ([http://en.wikipedia.org/wiki/Athlete's\\_foot#mediaviewer/File:FeetFungal.JPG](http://en.wikipedia.org/wiki/Athlete's_foot#mediaviewer/File:FeetFungal.JPG)): last accessed 7/21/2014).
- Heinsbroek, S.E., Brown, G.D., and Gordon, S. (2005). Dectin-1 escape by fungal dimorphism. *Trends Immunol* 26, 352-354.
- Heller, J., and Tudzynski, P. (2011). Reactive oxygen species in phytopathogenic fungi: signaling, development, and disease. *Annu Rev Phytopathol* 49, 369-390.
- Hengartner, M.O., Ellis, R.E., and Horvitz, H.R. (1992). *Caenorhabditis elegans* gene *ced-9* protects cells from programmed cell death. *Nature* 356, 494-499.
- Hernday, A.D., Noble, S.M., Mitrovich, Q.M., and Johnson, A.D. (2010). GENETICS AND MOLECULAR BIOLOGY IN CANDIDA ALBICANS. *Methods in Enzymology, Vol 470: Guide to Yeast Genetics:: Functional Genomics, Proteomics, and Other Systems Analysis, 2nd Edition* 470, 737-758.
- Hochstrasser, M. (1996). Ubiquitin-dependent protein degradation. *Annu Rev Genet* 30, 405-439.
- Hoeven, R., McCallum, K.C., Cruz, M.R., and Garsin, D.A. (2011). Ce-Duox1/BLI-3 generated reactive oxygen species trigger protective SKN-1 activity via p38 MAPK signaling during infection in *C. elegans*. *PLoS Pathog* 7, e1002453.
- Hoffman, C.S., and Winston, F. (1987). A ten-minute DNA preparation from yeast efficiently releases autonomous plasmids for transformation of *Escherichia coli*. *Gene* 57, 267-272.
- Hoffmann, J.A., Kafatos, F.C., Janeway, C.A., and Ezekowitz, R.A. (1999). Phylogenetic perspectives in innate immunity. *Science* 284, 1313-1318.
- Homann, O.R., Dea, J., Noble, S.M., and Johnson, A.D. (2009). A phenotypic profile of the *Candida albicans* regulatory network. *PLoS Genet* 5, e1000783.
- Howe, K., Clark, M.D., Torroja, C.F., Torrance, J., Berthelot, C., Muffato, M., Collins, J.E., Humphray, S., McLaren, K., Matthews, L., *et al.* (2013). The zebrafish reference genome sequence and its relationship to the human genome. *Nature* 496, 498-503.
- Hoyer, L.L. (2001). The ALS gene family of *Candida albicans*. In *Trends Microbiol (England)*, pp. 176-180.
- Hoyer, L.L., Green, C.B., Oh, S.H., and Zhao, X. (2008). Discovering the secrets of the *Candida albicans* agglutinin-like sequence (ALS) gene family--a sticky pursuit. *Med Mycol* 46, 1-15.
- Hoyer, L.L., Payne, T.L., Bell, M., Myers, A.M., and Scherer, S. (1998). *Candida albicans* ALS3 and insights into the nature of the ALS gene family. *Current genetics* 33, 451-459.
- Hromatka, B.S., Noble, S.M., and Johnson, A.D. (2005). Transcriptional response of *Candida albicans* to nitric oxide and the role of the YHB1 gene in nitrosative stress and virulence. *Mol Biol Cell* 16, 4814-4826.
- Huang, H., Ostroff, G.R., Lee, C.K., Specht, C.A., and Levitz, S.M. (2010). Robust stimulation of humoral and cellular immune responses following vaccination with antigen-loaded beta-glucan particles. *MBio* 1.
- Huang, W., Na, L., Fidel, P.L., and Schwarzenberger, P. (2004). Requirement of interleukin-17A for systemic anti-*Candida albicans* host defense in mice. In *J Infect Dis (United States)*, pp. 624-631.

## REFERENCES

- Huang, X., and Zhang, H. (2011). The zinc-finger protein SEA-2 regulates larval developmental timing and adult lifespan in *C. elegans*. In *Development (England)*, pp. 2059-2068.
- Huffman, D.L., Abrami, L., Sasik, R., Corbeil, J., van der Goot, F.G., and Aroian, R.V. (2004). Mitogen-activated protein kinase pathways defend against bacterial pore-forming toxins. *Proc Natl Acad Sci U S A* *101*, 10995-11000.
- Hull, R.P., Srivastava, P.K., D'Souza, Z., Atanur, S.S., Mechta-Grigoriou, F., Game, L., Petretto, E., Cook, H.T., Aitman, T.J., and Behmoaras, J. (2013). Combined ChIP-Seq and transcriptome analysis identifies AP-1/JunD as a primary regulator of oxidative stress and IL-1beta synthesis in macrophages. *BMC Genomics* *14*, 92.
- Ibata-Ombetta, S., Jouault, T., Trinel, P.A., and Poulain, D. (2001). Role of extracellular signal-regulated protein kinase cascade in macrophage killing of *Candida albicans*. *J Leukoc Biol* *70*, 149-154.
- Ibrahim, A.S., Spellberg, B.J., Avanesian, V., Fu, Y., and Edwards, J.E., Jr. (2006). The anti-*Candida* vaccine based on the recombinant N-terminal domain of Als1p is broadly active against disseminated candidiasis. *Infect Immun* *74*, 3039-3041.
- Iles, K.E., and Forman, H.J. (2002). Macrophage signaling and respiratory burst. In *Immunol Res (United States)*, pp. 95-105.
- Inglis, D.O., Arnaud, M.B., Binkley, J., Shah, P., Skrzypek, M.S., Wymore, F., Binkley, G., Miyasato, S.R., Simison, M., and Sherlock, G. (2012). The *Candida* genome database incorporates multiple *Candida* species: multispecies search and analysis tools with curated gene and protein information for *Candida albicans* and *Candida glabrata*. *Nucleic Acids Res* *40*, D667-674.
- Irazoqui, J.E., Troemel, E.R., Feinbaum, R.L., Luhachack, L.G., Cezairliyan, B.O., and Ausubel, F.M. (2010). Distinct pathogenesis and host responses during infection of *C. elegans* by *P. aeruginosa* and *S. aureus*. *PLoS Pathog* *6*, e1000982.
- Jacobsen, I.D., Wilson, D., Wachtler, B., Brunke, S., Naglik, J.R., and Hube, B. (2012). *Candida albicans* dimorphism as a therapeutic target. *Expert Rev Anti Infect Ther* *10*, 85-93.
- Jain, C., Pastor, K., Gonzalez, A., Lorenz, M.C., and Rao, R.P. (2013a). The role of *Candida albicans* AP-1 protein against host derived ROS in in vivo models of infection. *Virulence* *4*.
- Jain, C., Pastor, K., Gonzalez, A.Y., Lorenz, M.C., and Rao, R.P. (2013b). The role of *Candida albicans* AP-1 protein against host derived ROS in in vivo models of infection. *Virulence* *4*, 67-76.
- Jain, C., Yun, M., Politz, S.M., and Rao, R.P. (2009a). A pathogenesis assay using *Saccharomyces cerevisiae* and *Caenorhabditis elegans* reveals novel roles for yeast AP-1, Yap1, and host dual oxidase BLI-3 in fungal pathogenesis. *Eukaryot Cell* *8*, 1218-1227.
- Jain, C., Yun, M., Politz, S.M., and Rao, R.P. (2009b). A pathogenesis assay using *Saccharomyces cerevisiae* and *Caenorhabditis elegans* reveals novel roles for yeast AP-1, Yap1, and host dual oxidase BLI-3 in fungal pathogenesis. *Eukaryot Cell* *8*, 1218-1227.
- Jimenez-Lopez, C., Collette, J.R., Brothers, K.M., Shepardson, K.M., Cramer, R.A., Wheeler, R.T., and Lorenz, M.C. (2013). *Candida albicans* induces arginine biosynthetic genes in response to host-derived reactive oxygen species. *Eukaryot Cell* *12*, 91-100.
- Jong, A.Y., Stins, M.F., Huang, S.H., Chen, S.H., and Kim, K.S. (2001). Traversal of *Candida albicans* across human blood-brain barrier in vitro. *Infect Immun* *69*, 4536-4544.

## REFERENCES

- Kahana, A., and Gottschling, D.E. (1999). DOT4 links silencing and cell growth in *Saccharomyces cerevisiae*. *Mol Cell Biol* 19, 6608-6620.
- Kahn, L. (2008). The growing number of immunocompromised (Bull At. Sci.).
- Kaletta, T., and Hengartner, M.O. (2006). Finding function in novel targets: *C. elegans* as a model organism. In *Nat Rev Drug Discov* (England), pp. 387-398.
- Karababa, M., Coste, A.T., Rognon, B., Bille, J., and Sanglard, D. (2004). Comparison of gene expression profiles of *Candida albicans* azole-resistant clinical isolates and laboratory strains exposed to drugs inducing multidrug transporters. *Antimicrob Agents Chemother* 48, 3064-3079.
- Kauffman, C.A. (2007). Histoplasmosis: a clinical and laboratory update. *Clin Microbiol Rev* 20, 115-132.
- Kavanagh, K., and Reeves, E.P. (2004). Exploiting the potential of insects for in vivo pathogenicity testing of microbial pathogens. In *FEMS Microbiol Rev* (Netherlands), pp. 101-112.
- Kemp, H.A., and Sprague, G.F., Jr. (2003). Far3 and five interacting proteins prevent premature recovery from pheromone arrest in the budding yeast *Saccharomyces cerevisiae*. *Mol Cell Biol* 23, 1750-1763.
- Kenyon, C., Chang, J., Gensch, E., Rudner, A., and Tabtiang, R. (1993). A *C. elegans* mutant that lives twice as long as wild type. *Nature* 366, 461-464.
- Kesika, P., and Balamurugan, K. (2012). Studies on *Shigella boydii* infection in *Caenorhabditis elegans* and bioinformatics analysis of immune regulatory protein interactions. In *Biochim Biophys Acta* (Netherlands: 2012 Elsevier B.V), pp. 1449-1456.
- Kim, D.H., Feinbaum, R., Alloing, G., Emerson, F.E., Garsin, D.A., Inoue, H., Tanaka-Hino, M., Hisamoto, N., Matsumoto, K., Tan, M.W., *et al.* (2002). A conserved p38 MAP kinase pathway in *Caenorhabditis elegans* innate immunity. In *Science* (United States), pp. 623-626.
- Kim, J., and Sudbery, P. (2011). *Candida albicans*, a major human fungal pathogen. *J Microbiol* 49, 171-177.
- Kim, M.J., Kil, M., Jung, J.H., and Kim, J. (2008). Roles of Zinc-responsive transcription factor Csr1 in filamentous growth of the pathogenic Yeast *Candida albicans*. In *J Microbiol Biotechnol* (Korea South), pp. 242-247.
- Kletsas, D., and Papavassiliou, A.G. (1999). The therapeutic potential of targeting drugs at transcription factors. *Expert Opin Investig Drugs* 8, 737-746.
- Kohchi, C., Inagawa, H., Nishizawa, T., and Soma, G. (2009). ROS and innate immunity. In *Anticancer Res* (Greece), pp. 817-821.
- Kong, C., Yehye, W.A., Abd Rahman, N., Tan, M.W., and Nathan, S. (2014). Discovery of potential anti-infectives against *Staphylococcus aureus* using a *Caenorhabditis elegans* infection model. *BMC Complement Altern Med* 14, 4.
- Korting, H.C., Hube, B., Oberbauer, S., Januschke, E., Hamm, G., Albrecht, A., Borelli, C., and Schaller, M. (2003). Reduced expression of the hyphal-independent *Candida albicans* proteinase genes SAP1 and SAP3 in the *efg1* mutant is associated with attenuated virulence during infection of oral epithelium. *J Med Microbiol* 52, 623-632.
- Kumamoto, C.A., and Pierce, J.V. (2011). Immunosensing during colonization by *Candida albicans*: does it take a village to colonize the intestine? *Trends Microbiol* 19, 263-267.

## REFERENCES

- Kumamoto, C.A., and Vines, M.D. (2005). Alternative *Candida albicans* lifestyles: growth on surfaces. *Annu Rev Microbiol* 59, 113-133.
- Kuranda, K., Leberre, V., Sokol, S., Palamarczyk, G., and Francois, J. (2006). Investigating the caffeine effects in the yeast *Saccharomyces cerevisiae* brings new insights into the connection between TOR, PKC and Ras/cAMP signalling pathways. In *Mol Microbiol* (England), pp. 1147-1166.
- Kwon-Chung, K.J., Fraser, J.A., Doering, T.L., Wang, Z., Janbon, G., Idnurm, A., and Bahn, Y.S. (2014). *Cryptococcus neoformans* and *Cryptococcus gattii*, the Etiologic Agents of Cryptococcosis. In *Cold Spring Harb Perspect Med* (2014 Cold Spring Harbor Laboratory Press; all rights reserved.).
- Lamitina, S.T., and Strange, K. (2005). Transcriptional targets of DAF-16 insulin signaling pathway protect *C. elegans* from extreme hypertonic stress. In *Am J Physiol Cell Physiol* (United States), pp. C467-474.
- Lan, C.Y., Rodarte, G., Murillo, L.A., Jones, T., Davis, R.W., Dungan, J., Newport, G., and Agabian, N. (2004). Regulatory networks affected by iron availability in *Candida albicans*. In *Mol Microbiol* (England: 2004 Blackwell Publishing Ltd), pp. 1451-1469.
- Landau, D.A., Carter, S.L., Stojanov, P., McKenna, A., Stevenson, K., Lawrence, M.S., Sougnez, C., Stewart, C., Sivachenko, A., Wang, L., *et al.* (2013). Evolution and impact of subclonal mutations in chronic lymphocytic leukemia. *Cell* 152, 714-726.
- Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X version 2.0. In *Bioinformatics* (England), pp. 2947-2948.
- Lavoie H, S.A., Askew C, Nantel A, Whiteway M (2008). A toolbox for epitope-tagging and genome-wide l... [BMC Genomics. 2008] - PubMed - NCBI.
- Leberer, E., Ziegelbauer, K., Schmidt, A., Harcus, D., Dignard, D., Ash, J., Johnson, L., and Thomas, D.Y. (1997). Virulence and hyphal formation of *Candida albicans* require the Ste20p-like protein kinase CaCla4p. In *Curr Biol* (England), pp. 539-546.
- Lermann, U., and Morschhauser, J. (2008). Secreted aspartic proteases are not required for invasion of reconstituted human epithelia by *Candida albicans*. In *Microbiology* (England), pp. 3281-3295.
- Lewis, R.E. (2009). Overview of the changing epidemiology of candidemia. *Curr Med Res Opin* 25, 1732-1740.
- Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323.
- Li, C., Hisamoto, N., Nix, P., Kanao, S., Mizuno, T., Bastiani, M., and Matsumoto, K. (2012). The growth factor SVH-1 regulates axon regeneration in *C. elegans* via the JNK MAPK cascade. In *Nat Neurosci* (United States), pp. 551-557.
- Li, F., and Palecek, S.P. (2003). EAP1, a *Candida albicans* gene involved in binding human epithelial cells. *Eukaryot Cell* 2, 1266-1273.
- Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25, 1754-1760.
- Li, W. (2012). Volcano plots in analyzing differential expressions with mRNA microarrays. *J Bioinform Comput Biol* 10, 1231003.
- Li, W., Hu, X., Zhang, X., Ge, Y., Zhao, S., Hu, Y., and Ashman, R.B. (2011). Immunisation with the glycolytic enzyme enolase confers effective protection against *Candida albicans* infection in mice. In *Vaccine* (Netherlands: 2011 Elsevier Ltd), pp. 5526-5533.



## REFERENCES

- Lin, K., Dorman, J.B., Rodan, A., and Kenyon, C. (1997). *daf-16*: An HNF-3/forkhead family member that can function to double the life-span of *Caenorhabditis elegans*. *Science* 278, 1319-1322.
- Liu, T.B., and Xue, C. (2011). The Ubiquitin-Proteasome System and F-box Proteins in Pathogenic Fungi. *Mycobiology* 39, 243-248.
- Liu, Y., Mittal, R., Solis, N.V., Prasadarao, N.V., and Filler, S.G. (2011). Mechanisms of *Candida albicans* trafficking to the brain. *PLoS Pathog* 7, e1002305.
- Lo, H.J., Kohler, J.R., DiDomenico, B., Loebenberg, D., Cacciapuoti, A., and Fink, G.R. (1997). Nonfilamentous *C. albicans* mutants are avirulent. *Cell* 90, 939-949.
- Loar, J.W., Seiser, R.M., Sundberg, A.E., Sagerson, H.J., Ilias, N., Zobel-Thropp, P., Craig, E.A., and Lycan, D.E. (2004). Genetic and biochemical interactions among *Yar1*, *Ltv1* and *Rps3* define novel links between environmental stress and ribosome biogenesis in *Saccharomyces cerevisiae*. *Genetics* 168, 1877-1889.
- Lockhart, S.R., Messer, S.A., Pfaller, M.A., and Diekema, D.J. (2008). *Lodderomyces elongisporus* masquerading as *Candida parapsilosis* as a cause of bloodstream infections. *J Clin Microbiol* 46, 374-376.
- Loeb, J.D., Sepulveda-Becerra, M., Hazan, I., and Liu, H. (1999). A G1 cyclin is necessary for maintenance of filamentous growth in *Candida albicans*. *Mol Cell Biol* 19, 4019-4027.
- Loewith, R., and Hall, M.N. (2011). Target of rapamycin (TOR) in nutrient signaling and growth control. *Genetics* 189, 1177-1201.
- Loor, G., Kondapalli, J., Schriewer, J.M., Chandel, N.S., Vanden Hoek, T.L., and Schumacker, P.T. (2010). Menadione triggers cell death through ROS-dependent mechanisms involving PARP activation without requiring apoptosis. *Free Radic Biol Med* 49, 1925-1936.
- Lopes da Rosa, J., Boyartchuk, V.L., Zhu, L.J., and Kaufman, P.D. (2010). Histone acetyltransferase *Rtt109* is required for *Candida albicans* pathogenesis. *Proc Natl Acad Sci U S A* 107, 1594-1599.
- Lorenz, M.C., Bender, J.A., and Fink, G.R. (2004). Transcriptional response of *Candida albicans* upon internalization by macrophages. *Eukaryot Cell* 3, 1076-1087.
- Low, C.Y., and Rotstein, C. (2011). Emerging fungal infections in immunocompromised patients. *F1000 Med Rep* 3, 14.
- Lu, Y., Su, C., Wang, A., and Liu, H. (2011). Hyphal development in *Candida albicans* requires two temporally linked changes in promoter chromatin for initiation and maintenance. *PLoS Biol* 9, e1001105.
- MacPherson, S., Akache, B., Weber, S., De Deken, X., Raymond, M., and Turcotte, B. (2005). *Candida albicans* zinc cluster protein *Upc2p* confers resistance to antifungal drugs and is an activator of ergosterol biosynthetic genes. *Antimicrob Agents Chemother* 49, 1745-1752.
- Maicas, S., Moreno, I., Nieto, A., Gomez, M., Sentandreu, R., and Valentin, E. (2005). In silico analysis for transcription factors with Zn(II)(2)C(6) binuclear cluster DNA-binding domains in *Candida albicans*. *Comp Funct Genomics* 6, 345-356.
- Mallo, G.V., Kurz, C.L., Couillault, C., Pujol, N., Granjeaud, S., Kohara, Y., and Ewbank, J.J. (2002). Inducible antibacterial defense system in *C. elegans*. In *Curr Biol* (England), pp. 1209-1214.

## REFERENCES

- Maranhao, F.C.A., Silveira, H.C.S., Rossi, A., and Martinez-Rossi, N.M. (2011). Isolation of transcripts overexpressed in the human pathogen *Trichophyton rubrum* grown in lipid as carbon source. *Canadian Journal of Microbiology* 57, 333-338.
- Marcil, A., Harcus, D., Thomas, D.Y., and Whiteway, M. (2002). *Candida albicans* killing by RAW 264.7 mouse macrophage cells: effects of *Candida* genotype, infection ratios, and gamma interferon treatment. *Infect Immun* 70, 6319-6329.
- Mardis, E.R. (2008a). Next-generation DNA sequencing methods. In *Annual Review of Genomics and Human Genetics* (Palo Alto: Annual Reviews), pp. 387-402.
- Mardis, E.R. (2008b). The impact of next-generation sequencing technology on genetics. *Trends Genet* 24, 133-141.
- Marichal, P., Koymans, L., Willemsens, S., Bellens, D., Verhasselt, P., Luyten, W., Borgers, M., Ramaekers, F.C., Odds, F.C., and Bossche, H.V. (1999). Contribution of mutations in the cytochrome P450 14alpha-demethylase (Erg11p, Cyp51p) to azole resistance in *Candida albicans*. *Microbiology* 145 ( Pt 10), 2701-2713.
- Marletta, M.A., Yoon, P.S., Iyengar, R., Leaf, C.D., and Wishnok, J.S. (1988). Macrophage oxidation of L-arginine to nitrite and nitrate: nitric oxide is an intermediate. *Biochemistry* 27, 8706-8711.
- Marotta, D.H., Nantel, A., Sukala, L., Teubl, J.R., and Rauceo, J.M. (2013). Genome-wide transcriptional profiling and enrichment mapping reveal divergent and conserved roles of Sko1 in the *Candida albicans* osmotic stress response. *Genomics* 102, 363-371.
- Marr, K.A. (2004). Invasive *Candida* infections: the changing epidemiology. *Oncology* (Williston Park) 18, 9-14.
- Marsh, E.K., and May, R.C. (2012). *Caenorhabditis elegans*, a model organism for investigating immunity. *Appl Environ Microbiol* 78, 2075-2081.
- Martinez-Solano, L., Reales-Calderon, J.A., Nombela, C., Molero, G., and Gil, C. (2009). Proteomics of RAW 264.7 macrophages upon interaction with heat-inactivated *Candida albicans* cells unravel an anti-inflammatory response. *Proteomics* 9, 2995-3010.
- Mayer, F.L., Wilson, D., and Hube, B. (2013). *Candida albicans* pathogenicity mechanisms. *Virulence* 4, 119-128.
- McDonough, K.A., and Rodriguez, A. (2012). The myriad roles of cyclic AMP in microbial pathogens: from signal to sword. *Nat Rev Microbiol* 10, 27-38.
- McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., *et al.* (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20, 1297-1303.
- McNeil, M.M., Nash, S.L., Hajjeh, R.A., Phelan, M.A., Conn, L.A., Plikaytis, B.D., and Warnock, D.W. (2001). Trends in mortality due to invasive mycotic diseases in the United States, 1980-1997. In *Clin Infect Dis* (United States), pp. 641-647.
- Medbullets (2013). Onychomycosis (<http://www.medbullets.com/step2-3-dermatology/20060/onychomycosis>: last accessed 7/21/2014).
- Meitzler, J.L., and Ortiz de Montellano, P.R. (2009). *Caenorhabditis elegans* and human dual oxidase 1 (DUOX1) "peroxidase" domains: insights into heme binding and catalytic activity. *J Biol Chem* 284, 18634-18643.
- Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G.D. (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5, e13984.

## REFERENCES

- Miceli, M.H., Diaz, J.A., and Lee, S.A. (2011). Emerging opportunistic yeast infections. In *Lancet Infect Dis* (United States: 2011 Elsevier Ltd), pp. 142-151.
- Michels, P.A., Moyersoen, J., Krazy, H., Galland, N., Herman, M., and Hannaert, V. (2005). Peroxisomes, glyoxysomes and glycosomes (review). *Mol Membr Biol* 22, 133-145.
- Mitchell, A.P. (1998). Dimorphism and virulence in *Candida albicans*. In *Curr Opin Microbiol* (England), pp. 687-692.
- Mizuno, T., Hisamoto, N., Terada, T., Kondo, T., Adachi, M., Nishida, E., Kim, D.H., Ausubel, F.M., and Matsumoto, K. (2004). The *Caenorhabditis elegans* MAPK phosphatase VHP-1 mediates a novel JNK-like signaling pathway in stress response. *EMBO J* 23, 2226-2234.
- Morita, K., Chow, K.L., and Ueno, N. (1999). Regulation of body length and male tail ray pattern formation of *Caenorhabditis elegans* by a member of TGF-beta family. *Development* 126, 1337-1347.
- Morschhauser, J., Barker, K.S., Liu, T.T., Bla, B.W.J., Homayouni, R., and Rogers, P.D. (2007). The transcription factor Mrr1p controls expression of the MDR1 efflux pump and mediates multidrug resistance in *Candida albicans*. *PLoS Pathog* 3, e164.
- Moy, T.I., Mylonakis, E., Calderwood, S.B., and Ausubel, F.M. (2004). Cytotoxicity of hydrogen peroxide produced by *Enterococcus faecium*. *Infect Immun* 72, 4512-4520.
- Moyes, D.L., and Naglik, J.R. (2011). Mucosal immunity and *Candida albicans* infection. *Clin Dev Immunol* 2011, 346307.
- Moyes, D.L., Runglall, M., Murciano, C., Shen, C., Nayar, D., Thavaraj, S., Kohli, A., Islam, A., Mora-Montes, H., Challacombe, S.J., *et al.* (2010). A biphasic innate immune MAPK response discriminates between the yeast and hyphal forms of *Candida albicans* in epithelial cells. *Cell Host Microbe* 8, 225-235.
- Moyes, D.L., Shen, C., Murciano, C., Runglall, M., Richardson, J.P., Arno, M., Aldecoa-Otalora, E., and Naglik, J.R. (2014). Protection Against Epithelial Damage During *Candida albicans* Infection Is Mediated by PI3K/Akt and Mammalian Target of Rapamycin Signaling. In *J Infect Dis*.
- Mukherjee, P.K., Chandra, J., Kuhn, D.M., and Ghannoum, M.A. (2003). Mechanism of fluconazole resistance in *Candida albicans* biofilms: phase-specific role of efflux pumps and membrane sterols. *Infect Immun* 71, 4333-4340.
- Murad, A.M., d'Enfert, C., Gaillardin, C., Tournu, H., Tekaia, F., Talibi, D., Marechal, D., Marchais, V., Cottin, J., and Brown, A.J. (2001). Transcript profiling in *Candida albicans* reveals new cellular functions for the transcriptional repressors CaTup1, CaMig1 and CaNrg1. *Mol Microbiol* 42, 981-993.
- Murphy, C.T., and Hu, P.J. (2013). Insulin/insulin-like growth factor signaling in *C. elegans*. *WormBook*, 1-43.
- Murphy, C.T., McCarroll, S.A., Bargmann, C.I., Fraser, A., Kamath, R.S., Ahringer, J., Li, H., and Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. In *Nature* (England), pp. 277-283.
- Mylonakis, E., Ausubel, F.M., Perfect, J.R., Heitman, J., and Calderwood, S.B. (2002). Killing of *Caenorhabditis elegans* by *Cryptococcus neoformans* as a model of yeast pathogenesis. *Proc Natl Acad Sci U S A* 99, 15675-15680.
- Nagahashi, S., Mio, T., Ono, N., Yamada-Okabe, T., Arisawa, M., Bussey, H., and Yamada-Okabe, H. (1998). Isolation of CaSLN1 and CaNIK1, the genes for osmosensing

## REFERENCES

- histidine kinase homologues, from the pathogenic fungus *Candida albicans*. *Microbiology* *144* ( Pt 2), 425-432.
- Naglik, J.R., Challacombe, S.J., and Hube, B. (2003a). *Candida albicans* secreted aspartyl proteinases in virulence and pathogenesis. *Microbiol Mol Biol Rev* *67*, 400-428, table of contents.
  - Naglik, J.R., Rodgers, C.A., Shirlaw, P.J., Dobbie, J.L., Fernandes-Naglik, L.L., Greenspan, D., Agabian, N., and Challacombe, S.J. (2003b). Differential expression of *Candida albicans* secreted aspartyl proteinase and phospholipase B genes in humans correlates with active oral and vaginal infections. In *J Infect Dis* (United States), pp. 469-479.
  - Newman, S.L., and Holly, A. (2001). *Candida albicans* is phagocytosed, killed, and processed for antigen presentation by human dendritic cells. *Infect Immun* *69*, 6813-6822.
  - Newport, G., and Agabian, N. (1997). KEX2 influences *Candida albicans* proteinase secretion and hyphal formation. *J Biol Chem* *272*, 28954-28961.
  - Nobile, C.J., Bruno, V.M., Richard, M.L., Davis, D.A., and Mitchell, A.P. (2003). Genetic control of chlamydospore formation in *Candida albicans*. *Microbiology* *149*, 3629-3637.
  - Nobile, C.J., Fox, E.P., Nett, J.E., Sorrells, T.R., Mitrovich, Q.M., Hernday, A.D., Tuch, B.B., Andes, D.R., and Johnson, A.D. (2012). A recently evolved transcriptional network controls biofilm development in *Candida albicans*. *Cell* *148*, 126-138.
  - Nobile, C.J., and Mitchell, A.P. (2005a). Regulation of cell-surface genes and biofilm formation by the *C. albicans* transcription factor Bcr1p. *Current Biology* *15*, 1150-1155.
  - Nobile, C.J., and Mitchell, A.P. (2005b). Regulation of cell-surface genes and biofilm formation by the *C. albicans* transcription factor Bcr1p. *Curr Biol* *15*, 1150-1155.
  - Nobile, C.J., Solis, N., Myers, C.L., Fay, A.J., Deneault, J.S., Nantel, A., Mitchell, A.P., and Filler, S.G. (2008). *Candida albicans* transcription factor Rim101 mediates pathogenic interactions through cell wall functions. *Cell Microbiol* *10*, 2180-2196.
  - Noble, S.M., French, S., Kohn, L.A., Chen, V., and Johnson, A.D. (2010a). Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nat Genet* *42*, 590-598.
  - Noble, S.M., French, S., Kohn, L.A., Chen, V., and Johnson, A.D. (2010b). Systematic screens of a *Candida albicans* homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nat Genet* *42*, 590-598.
  - Noble, S.M., and Johnson, A.D. (2005). Strains and strategies for large-scale gene deletion studies of the diploid human fungal pathogen *Candida albicans*. *Eukaryot Cell* *4*, 298-309.
  - Norice, C.T., Smith, F.J., Solis, N., Filler, S.G., and Mitchell, A.P. (2007). Requirement for *Candida albicans* sun41 in biofilm formation and virulence. *Eukaryotic Cell* *6*, 2046-2055.
  - Okoli, I., Coleman, J.J., Tampakakis, E., An, W.F., Holson, E., Wagner, F., Conery, A.L., Larkins-Ford, J., Wu, G., Stern, A., *et al.* (2009). Identification of antifungal compounds active against *Candida albicans* using an improved high-throughput *Caenorhabditis elegans* assay. *PLoS One* *4*, e7025.
  - Oliver, B.G., Song, J.L., Choiniere, J.H., and White, T.C. (2007). cis-Acting elements within the *Candida albicans* ERG11 promoter mediate the azole response through transcription factor Upc2p. *Eukaryot Cell* *6*, 2231-2239.

## REFERENCES

- Ostrosky-Zeichner, L., Casadevall, A., Galgiani, J.N., Odds, F.C., and Rex, J.H. (2010). An insight into the antifungal pipeline: selected new molecules and beyond. In *Nat Rev Drug Discov* (England), pp. 719-727.
- Papon, N., Courdavault, V., Clastre, M., and Bennett, R.J. (2013). Emerging and emerged pathogenic *Candida* species: beyond the *Candida albicans* paradigm. *PLoS Pathog* 9, e1003550.
- Pappas, P.G., Rex, J.H., Lee, J., Hamill, R.J., Larsen, R.A., Powderly, W., Kauffman, C.A., Hyslop, N., Mangino, J.E., Chapman, S., *et al.* (2003). A prospective observational study of candidemia: epidemiology, therapy, and influences on mortality in hospitalized adult and pediatric patients. *Clin Infect Dis* 37, 634-643.
- Parnanen, P., Meurman, J.H., Samaranayake, L., and Virtanen, I. (2010). Human oral keratinocyte E-cadherin degradation by *Candida albicans* and *Candida glabrata*. In *J Oral Pathol Med* (Denmark), pp. 275-278.
- Pasko, M.T., Piscitelli, S.C., and Van Slooten, A.D. (1990). Fluconazole: a new triazole antifungal agent. *DICP* 24, 860-867.
- Pavelka, N., Rancati, G., Zhu, J., Bradford, W.D., Saraf, A., Florens, L., Sanderson, B.W., Hattem, G.L., and Li, R. (2010). Aneuploidy confers quantitative proteome changes and phenotypic variation in budding yeast. *Nature* 468, 321-325.
- Perea, S., Lopez-Ribot, J.L., Kirkpatrick, W.R., McAtee, R.K., Santillan, R.A., Martinez, M., Calabrese, D., Sanglard, D., and Patterson, T.F. (2001a). Prevalence of molecular mechanisms of resistance to azole antifungal agents in *Candida albicans* strains displaying high-level fluconazole resistance isolated from human immunodeficiency virus-infected patients. *Antimicrob Agents Chemother* 45, 2676-2684.
- Perea, S., Lopez-Ribot, J.L., Kirkpatrick, W.R., McAtee, R.K., Santillan, R.A., Martinez, M., Calabrese, D., Sanglard, D., and Patterson, T.F. (2001b). Prevalence of molecular mechanisms of resistance to azole antifungal agents in *Candida albicans* strains displaying high-level fluconazole resistance isolated from human immunodeficiency virus-infected patients. *Antimicrob Agents Chemother* 45, 2676-2684.
- Pesole, G., Lotti, M., Alberghina, L., and Saccone, C. (1995). Evolutionary origin of nonuniversal CUGSer codon in some *Candida* species as inferred from a molecular phylogeny. *Genetics* 141, 903-907.
- Pfaller, M.A., and Diekema, D.J. (2007). Epidemiology of invasive candidiasis: a persistent public health problem. *Clin Microbiol Rev* 20, 133-163.
- Pfaller, M.A., Diekema, D.J., Mendez, M., Kibbler, C., Erzsebet, P., Chang, S.C., Gibbs, D.L., and Newell, V.A. (2006). *Candida guilliermondii*, an opportunistic fungal pathogen with decreased susceptibility to fluconazole: geographic and temporal trends from the ARTEMIS DISK antifungal surveillance program. *J Clin Microbiol* 44, 3551-3556.
- Pfaller, M.A., Lockhart, S.R., Pujol, C., Swails-Wenger, J.A., Messer, S.A., Edmond, M.B., Jones, R.N., Wenzel, R.P., and Soll, D.R. (1998). Hospital specificity, region specificity, and fluconazole resistance of *Candida albicans* bloodstream isolates. *J Clin Microbiol* 36, 1518-1529.
- Pittet, D., Monod, M., Suter, P.M., Frenk, E., and Auckenthaler, R. (1994). *Candida* colonization and subsequent infections in critically ill surgical patients. *Ann Surg* 220, 751-758.
- Podlaha, O., Riester, M., De, S., and Michor, F. (2012). Evolution of the cancer genome. *Trends Genet* 28, 155-163.

## REFERENCES

- Ponder, R.G., Fonville, N.C., and Rosenberg, S.M. (2005). A switch from high-fidelity to error-prone DNA double-strand break repair underlies stress-induced mutation. *Mol Cell* 19, 791-804.
- Pujol, N., Cypowyj, S., Ziegler, K., Millet, A., Astrain, A., Goncharov, A., Jin, Y., Chisholm, A.D., and Ewbank, J.J. (2008). Distinct innate immune responses to infection and wounding in the *C. elegans* epidermis. *Curr Biol* 18, 481-489.
- Pujol, N., Link, E.M., Liu, L.X., Kurz, C.L., Alloing, G., Tan, M.W., Ray, K.P., Solari, R., Johnson, C.D., and Ewbank, J.J. (2001). A reverse genetic analysis of components of the Toll signaling pathway in *Caenorhabditis elegans*. In *Curr Biol (England)*, pp. 809-821.
- Pukkila-Worley, R., Ausubel, F.M., and Mylonakis, E. (2011). *Candida albicans* infection of *Caenorhabditis elegans* induces antifungal immune defenses. *PLoS Pathog* 7, e1002074.
- Pukkila-Worley, R., Peleg, A.Y., Tampakakis, E., and Mylonakis, E. (2009). *Candida albicans* hyphal formation and virulence assessed using a *Caenorhabditis elegans* infection model. *Eukaryot Cell* 8, 1750-1758.
- Quesada, V., Diaz-Perales, A., Gutierrez-Fernandez, A., Garabaya, C., Cal, S., and Lopez-Otin, C. (2004). Cloning and enzymatic analysis of 22 novel human ubiquitin-specific proteases. In *Biochem Biophys Res Commun (United States)*, pp. 54-62.
- Rallis, C., Codlin, S., and Bahler, J. (2013). TORC1 signaling inhibition by rapamycin and caffeine affect lifespan, global gene expression, and cell proliferation of fission yeast. *Aging Cell* 12, 563-573.
- Ramachandra, S., Linde, J., Brock, M., Guthke, R., Hube, B., and Brunke, S. (2014). Regulatory networks controlling nitrogen sensing and uptake in *Candida albicans*. *PLoS One* 9, e92734.
- Ramage, G., Coco, B., Sherry, L., Bagg, J., and Lappin, D.F. (2012). In vitro *Candida albicans* biofilm induced proteinase activity and SAP8 expression correlates with in vivo denture stomatitis severity. *Mycopathologia* 174, 11-19.
- Ramirez, M.A., and Lorenz, M.C. (2007). Mutations in alternative carbon utilization pathways in *Candida albicans* attenuate virulence and confer pleiotropic phenotypes. *Eukaryot Cell* 6, 280-290.
- Ramirez-Zavala, B., Reuss, O., Park, Y.N., Ohlsen, K., and Morschhauser, J. (2008). Environmental induction of white-opaque switching in *Candida albicans*. *PLoS Pathog* 4, e1000089.
- Raschke, W.C., Baird, S., Ralph, P., and Nakoinz, I. (1978). Functional macrophage cell lines transformed by Abelson leukemia virus. In *Cell (England)*, pp. 261-267.
- Raska, M., Belakova, J., Wudattu, N.K., Kafkova, L., Ruzickova, K., Sebestova, M., Kolar, Z., and Weigl, E. (2005). Comparison of protective effect of protein and DNA vaccines hsp90 in murine model of systemic candidiasis. *Folia Microbiol (Praha)* 50, 77-82.
- Rauceo, J.M., Blankenship, J.R., Fanning, S., Hamaker, J.J., Deneault, J.S., Smith, F.J., Nantel, A., and Mitchell, A.P. (2008). Regulation of the *Candida albicans* cell wall damage response by transcription factor Sko1 and PAS kinase Psk1. *Molecular Biology of the Cell* 19, 2741-2751.
- Rentz, A.M., Halpern, M.T., and Bowden, R. (1998). The impact of candidemia on length of hospital stay, outcome, and overall cost of illness. *Clin Infect Dis* 27, 781-788.

## REFERENCES

- Rex, J.H., Bennett, J.E., Sugar, A.M., Pappas, P.G., van der Horst, C.M., Edwards, J.E., Washburn, R.G., Scheld, W.M., Karchmer, A.W., Dine, A.P., *et al.* (1994). A randomized trial comparing fluconazole with amphotericin B for the treatment of candidemia in patients without neutropenia. Candidemia Study Group and the National Institute. *N Engl J Med* *331*, 1325-1330.
- Rex, J.H., Pappas, P.G., Karchmer, A.W., Sobel, J., Edwards, J.E., Hadley, S., Brass, C., Vazquez, J.A., Chapman, S.W., Horowitz, H.W., *et al.* (2003). A randomized and blinded multicenter trial of high-dose fluconazole plus placebo versus fluconazole plus amphotericin B as therapy for candidemia and its consequences in nonneutropenic subjects. *Clin Infect Dis* *36*, 1221-1228.
- Reynolds, T.B., and Fink, G.R. (2001). Bakers' yeast, a model for fungal biofilm formation. *Science* *291*, 878-881.
- Ricicova, M., Kucharikova, S., Tournu, H., Hendrix, J., Bujdakova, H., Van Eldere, J., Lagrou, K., and Van Dijck, P. (2010). *Candida albicans* biofilm formation in a new in vivo rat model. In *Microbiology (England)*, pp. 909-919.
- Rihana, N.A., Kandula, M., Velez, A., Dahal, K., and O'Neill, E.B. (2014). Histoplasmosis presenting as granulomatous hepatitis: case report and review of the literature. *Case Rep Med* *2014*, 879535.
- Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011a). Integrative genomics viewer. *Nat Biotechnol* *29*, 24-26.
- Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011b). Integrative genomics viewer. *Nat Biotechnol* *29*, 24-26.
- Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* *26*, 139-140.
- Robinson, R. (2008). Birds do it, bees do it, but *Candida albicans* does it differently. *PLoS Biol* *6*, e121.
- Roemer, T., and Krysan, D.J. (2014). Antifungal drug development: challenges, unmet clinical needs, and new approaches. In *Cold Spring Harb Perspect Med (United States)*.
- Rogers, P.D., and Barker, K.S. (2003). Genome-wide expression profile analysis reveals coordinately regulated genes associated with stepwise acquisition of azole resistance in *Candida albicans* clinical isolates. *Antimicrob Agents Chemother* *47*, 1220-1227.
- Rogers, T., and Balish, E. (1976). Experimental *Candida albicans* infection in conventional mice and germfree rats. *Infect Immun* *14*, 33-38.
- Romani, L. (2000). Innate and adaptive immunity in *Candida albicans* infections and saprophytism. *J Leukoc Biol* *68*, 175-179.
- Rosenberg, S.M. (2011). Stress-induced loss of heterozygosity in *Candida*: a possible missing link in the ability to evolve. *MBio* *2*.
- Roy, S., Wapinski, I., Pfiffner, J., French, C., Socha, A., Konieczka, J., Habib, N., Kellis, M., Thompson, D., and Regev, A. (2013). Arboretum: reconstruction and analysis of the evolutionary history of condition-specific transcriptional modules. *Genome Res* *23*, 1039-1050.
- Rubin, G.M., Yandell, M.D., Wortman, J.R., Gabor Miklos, G.L., Nelson, C.R., Hariharan, I.K., Fortini, M.E., Li, P.W., Apweiler, R., Fleischmann, W., *et al.* (2000). Comparative genomics of the eukaryotes. *Science* *287*, 2204-2215.
- Russell, L. (2010). Antifungal pharmacology.

## REFERENCES

- Salgado, P.S., Yan, R., Taylor, J.D., Burchell, L., Jones, R., Hoyer, L.L., Matthews, S.J., Simpson, P.J., and Cota, E. (2011). Structural basis for the broad specificity to host-cell ligands by the pathogenic fungus *Candida albicans*. *Proc Natl Acad Sci U S A* 108, 15775-15779.
- Sandini, S., La Valle, R., Deaglio, S., Malavasi, F., Cassone, A., and De Bernardis, F. (2011). A highly immunogenic recombinant and truncated protein of the secreted aspartic proteases family (rSap2t) of *Candida albicans* as a mucosal anticandidal vaccine. *FEMS Immunol Med Microbiol* 62, 215-224.
- Sanglard, D., Ischer, F., Marchetti, O., Entenza, J., and Bille, J. (2003). Calcineurin A of *Candida albicans*: involvement in antifungal tolerance, cell morphogenesis and virulence. In *Mol Microbiol (England)*, pp. 959-976.
- Santos, M.A., Moura, G., Massey, S.E., and Tuite, M.F. (2004). Driving change: the evolution of alternative genetic codes. *Trends Genet* 20, 95-102.
- Sasse, C., Dunkel, N., Schafer, T., Schneider, S., Dierolf, F., Ohlsen, K., and Morschhauser, J. (2012). The stepwise acquisition of fluconazole resistance mutations causes a gradual loss of fitness in *Candida albicans*. *Mol Microbiol* 86, 539-556.
- Sasse, C., Hasenberg, M., Weyler, M., Gunzer, M., and Morschhauser, J. (2013). White-opaque switching of *Candida albicans* allows immune evasion in an environment-dependent fashion. *Eukaryot Cell* 12, 50-58.
- Saville, S.P., Lazzell, A.L., Monteagudo, C., and Lopez-Ribot, J.L. (2003). Engineered control of cell morphology in vivo reveals distinct roles for yeast and filamentous forms of *Candida albicans* during infection. *Eukaryot Cell* 2, 1053-1060.
- Sawaya, B.P., Weihprecht, H., Campbell, W.R., Lorenz, J.N., Webb, R.C., Briggs, J.P., and Schnermann, J. (1991). Direct vasoconstriction as a possible cause for amphotericin B-induced nephrotoxicity in rats. *J Clin Invest* 87, 2097-2107.
- Schaller, M., Borelli, C., Korting, H.C., and Hube, B. (2005). Hydrolytic enzymes as virulence factors of *Candida albicans*. In *Mycoses (Germany)*, pp. 365-377.
- Schaller, M., Krnjaic, N., Niewerth, M., Hamm, G., Hube, B., and Korting, H.C. (2003). Effect of antimycotic agents on the activity of aspartyl proteinases secreted by *Candida albicans*. *J Med Microbiol* 52, 247-249.
- Schaller, M., Schackert, C., Korting, H.C., Januschke, E., and Hube, B. (2000). Invasion of *Candida albicans* correlates with expression of secreted aspartic proteinases during experimental infection of human epidermis. In *J Invest Dermatol (United States)*, pp. 712-717.
- Schmidt, C.S., White, C.J., Ibrahim, A.S., Filler, S.G., Fu, Y., Yeaman, M.R., Edwards, J.E., Jr., and Hennessey, J.P., Jr. (2012). NDV-3, a recombinant alum-adjuvanted vaccine for *Candida* and *Staphylococcus aureus*, is safe and immunogenic in healthy adults. *Vaccine* 30, 7594-7600.
- Schubert, S., Popp, C., Rogers, P.D., and Morschhauser, J. (2011). Functional dissection of a *Candida albicans* zinc cluster transcription factor, the multidrug resistance regulator Mrr1. *Eukaryot Cell* 10, 1110-1121.
- Selmecki, A., Forche, A., and Berman, J. (2006). Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* 313, 367-370.
- Selmecki, A., Forche, A., and Berman, J. (2010a). Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryot Cell* 9, 991-1008.
- Selmecki, A., Forche, A., and Berman, J. (2010b). Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryot Cell* 9, 991-1008.



## REFERENCES

- Selmecki, A., Gerami-Nejad, M., Paulson, C., Forche, A., and Berman, J. (2008). An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. *Mol Microbiol* 68, 624-641.
- Selmecki, A.M., Dulmage, K., Cowen, L.E., Anderson, J.B., and Berman, J. (2009). Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. *PLoS Genet* 5, e1000705.
- Sem, X., and Rhen, M. (2012). Pathogenicity of *Salmonella enterica* in *Caenorhabditis elegans* relies on disseminated oxidative stress in the infected host. *PLoS One* 7, e45417.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 13, 2498-2504.
- Sheppard, D.C., Yeaman, M.R., Welch, W.H., Phan, Q.T., Fu, Y., Ibrahim, A.S., Filler, S.G., Zhang, M., Waring, A.J., and Edwards, J.E., Jr. (2004). Functional and structural diversity in the Als protein family of *Candida albicans*. *J Biol Chem* 279, 30480-30489.
- Sifri, C.D., Begun, J., Ausubel, F.M., and Calderwood, S.B. (2003). *Caenorhabditis elegans* as a model host for *Staphylococcus aureus* pathogenesis. *Infect Immun* 71, 2208-2217.
- Singh, R., and Parija, S.C. (2012). *Candida parapsilosis* : an emerging fungal pathogen. *Indian J Med Res* 136, 671-673.
- Singh-Babak, S.D., Babak, T., Diezmann, S., Hill, J.A., Xie, J.L., Chen, Y.L., Poutanen, S.M., Rennie, R.P., Heitman, J., and Cowen, L.E. (2012). Global analysis of the evolution and mechanism of echinocandin resistance in *Candida glabrata*. *PLoS Pathog* 8, e1002718.
- Smith, D.A., Nicholls, S., Morgan, B.A., Brown, A.J., and Quinn, J. (2004). A conserved stress-activated protein kinase regulates a core stress response in the human pathogen *Candida albicans*. *Mol Biol Cell* 15, 4179-4190.
- Smoot, M.E., Ono, K., Ruscheinski, J., Wang, P.L., and Ideker, T. (2011). Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 27, 431-432.
- Sobel, J.D. (2007). Vulvovaginal candidosis. *Lancet* 369, 1961-1971.
- Solis, N.V., and Filler, S.G. (2012). Mouse model of oropharyngeal candidiasis. *Nat Protoc* 7, 637-642.
- Sonneborn, A., Tebarth, B., and Ernst, J.F. (1999). Control of white-opaque phenotypic switching in *Candida albicans* by the Efg1p morphogenetic regulator. *Infect Immun* 67, 4655-4660.
- Spellberg, B., Ibrahim, A.S., Yeaman, M.R., Lin, L., Fu, Y., Avanesian, V., Bayer, A.S., Filler, S.G., Lipke, P., Otoo, H., *et al.* (2008). The antifungal vaccine derived from the recombinant N terminus of Als3p protects mice against the bacterium *Staphylococcus aureus*. *Infect Immun* 76, 4574-4580.
- Staib, P., Kretschmar, M., Nichterlein, T., Hof, H., and Morschhauser, J. (2002). Transcriptional regulators Cph1p and Efg1p mediate activation of the *Candida albicans* virulence gene SAP5 during infection. *Infect Immun* 70, 921-927.
- Storm, N., Darnhofer-Patel, B., van den Boom, D., and Rodi, C.P. (2003). MALDI-TOF mass spectrometry-based SNP genotyping. *Methods Mol Biol* 212, 241-262.
- Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S., *et al.* (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 102, 15545-15550.

## REFERENCES

- Suga, Y., Kimura, U., and Hiruma, M. (2014). Can persistent toenail fungus be successfully treated with a laser? In *Med Mycol J (Japan)*, pp. J65-71.
- Sullivan, C., and Kim, C.H. (2008). Zebrafish as a model for infectious disease and immune function. In *Fish Shellfish Immunol (England)*, pp. 341-350.
- Sulston, J.E., and Horvitz, H.R. (1977). Post-embryonic cell lineages of the nematode, *Caenorhabditis elegans*. In *Dev Biol (United States)*, pp. 110-156.
- Suzuki, Y., Yandell, M.D., Roy, P.J., Krishna, S., Savage-Dunn, C., Ross, R.M., Padgett, R.W., and Wood, W.B. (1999). A BMP homolog acts as a dose-dependent regulator of body size and male tail patterning in *Caenorhabditis elegans*. *Development* 126, 241-250.
- Swofford, D.L. (2002). PAUP\*. Phylogenetic Analysis Using Parsimony (\*and Other Methods). Sinauer Associates, Sunderland, Massachusetts.
- Tan, M.W., Mahajan-Miklos, S., and Ausubel, F.M. (1999). Killing of *Caenorhabditis elegans* by *Pseudomonas aeruginosa* used to model mammalian bacterial pathogenesis. *Proc Natl Acad Sci U S A* 96, 715-720.
- Tang, Y.C., and Amon, A. (2013). Gene copy-number alterations: a cost-benefit analysis. *Cell* 152, 394-405.
- Tenor, J.L., and Aballay, A. (2008). A conserved Toll-like receptor is required for *Caenorhabditis elegans* innate immunity. *EMBO Rep* 9, 103-109.
- Thomas, J., Jacobson, G.A., Narkowicz, C.K., Peterson, G.M., Burnet, H., and Sharpe, C. (2010). Toenail onychomycosis: an important global disease burden. *Journal of Clinical Pharmacy and Therapeutics* 35, 497-519.
- Thompson, D.A., Desai, M.M., and Murray, A.W. (2006). Ploidy controls the success of mutators and nature of mutations during budding yeast evolution. *Curr Biol* 16, 1581-1590.
- Torres, E.M., Dephoure, N., Panneerselvam, A., Tucker, C.M., Whittaker, C.A., Gygi, S.P., Dunham, M.J., and Amon, A. (2010). Identification of aneuploidy-tolerating mutations. *Cell* 143, 71-83.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat Protoc* 7, 562-578.
- Troemel, E.R., Sagasti, A., and Bargmann, C.I. (1999). Lateral signaling mediated by axon contact and calcium entry regulates asymmetric odorant receptor expression in *C. elegans*. In *Cell (United States)*, pp. 387-398.
- Tuch, B.B., Mitrovich, Q.M., Homann, O.R., Hernday, A.D., Monighetti, C.K., De La Vega, F.M., and Johnson, A.D. (2010). The transcriptomes of two heritable cell types illuminate the circuit governing their differentiation. *PLoS Genet* 6, e1001070.
- Uhl, M.A., Biery, M., Craig, N., and Johnson, A.D. (2003). Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen *C.albicans*. *The EMBO journal* 22, 2668-2678.
- Vallee, B.L., Coleman, J.E., and Auld, D.S. (1991). Zinc fingers, zinc clusters, and zinc twists in DNA-binding protein domains. *Proc Natl Acad Sci U S A* 88, 999-1003.
- Vandeputte, P., Ischer, F., Sanglard, D., and Coste, A.T. (2011). In Vivo Systematic Analysis of *Candida albicans* Zn2-Cys6 Transcription Factors Mutants for Mice Organ Colonization. *Plos One* 6.

## REFERENCES

- Vazquez-Torres, A., and Balish, E. (1997). Macrophages in resistance to candidiasis. *Microbiol Mol Biol Rev* 61, 170-192.
- Vecchiarelli, A., Pericolini, E., Gabrielli, E., and Pietrella, D. (2012). New approaches in the development of a vaccine for mucosal candidiasis: progress and challenges. *Front Microbiol* 3, 294.
- Verrier, J., Krahenbuhl, L., Bontems, O., Fratti, M., Salamin, K., and Monod, M. (2013). Dermatophyte identification in skin and hair samples using a simple and reliable nested polymerase chain reaction assay. *British Journal of Dermatology* 168, 295-301.
- Verstrepen, K.J., Jansen, A., Lewitter, F., and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. *Nat Genet* 37, 986-990.
- Vylkova, S., Carman, A.J., Danhof, H.A., Collette, J.R., Zhou, H., and Lorenz, M.C. (2011). The fungal pathogen *Candida albicans* autoinduces hyphal morphogenesis by raising extracellular pH. *MBio* 2, e00055-00011.
- Wachtler, B., Wilson, D., Haedicke, K., Dalle, F., and Hube, B. (2011). From attachment to damage: defined genes of *Candida albicans* mediate adhesion, invasion and damage during interaction with oral epithelial cells. *PLoS One* 6, e17046.
- Wang, Y., Cao, Y.Y., Jia, X.M., Cao, Y.B., Gao, P.H., Fu, X.P., Ying, K., Chen, W.S., and Jiang, Y.Y. (2006). Cap1p is involved in multiple pathways of oxidative stress response in *Candida albicans*. In *Free Radic Biol Med (United States)*, pp. 1201-1209.
- Wang, Y., and Xu, X.L. (2008). Bacterial peptidoglycan-derived molecules activate *Candida albicans* hyphal growth. *Commun Integr Biol* 1, 137-139.
- Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Natural history and evolutionary principles of gene duplication in fungi. In *Nature (England)*, pp. 54-61.
- Warner, J.R. (1999). The economics of ribosome biosynthesis in yeast. In *Trends Biochem Sci (England)*, pp. 437-440.
- Waterston, R.H., Lindblad-Toh, K., Birney, E., Rogers, J., Abril, J.F., Agarwal, P., Agarwala, R., Ainscough, R., Alexandersson, M., An, P., *et al.* (2002). Initial sequencing and comparative analysis of the mouse genome. In *Nature (England)*, pp. 520-562.
- webMed (2007). Ringworm of the Scalp (*Tinea Capitis*) (<http://www.webmd.com/skin-problems-and-treatments/picture-of-ringworm-of-the-scalp-tinea-capitis>: last accessed 7/21/2014).
- Weissman, Z., Berdicevsky, I., Cavari, B.Z., and Kornitzer, D. (2000). The high copper tolerance of *Candida albicans* is mediated by a P-type ATPase. *Proc Natl Acad Sci U S A* 97, 3520-3525.
- White, T.C. (1997a). Increased mRNA levels of ERG16, CDR, and MDR1 correlate with increases in azole resistance in *Candida albicans* isolates from a patient infected with human immunodeficiency virus. *Antimicrob Agents Chemother* 41, 1482-1487.
- White, T.C. (1997b). Increased mRNA levels of ERG16, CDR, and MDR1 correlate with increases in azole resistance in *Candida albicans* isolates from a patient infected with human immunodeficiency virus. *Antimicrob Agents Chemother* 41, 1482-1487.
- White, T.C. (1997c). The presence of an R467K amino acid substitution and loss of allelic variation correlate with an azole-resistant lanosterol 14 $\alpha$  demethylase in *Candida albicans*. *Antimicrob Agents Chemother* 41, 1488-1494.
- White, T.C., Holleman, S., Dy, F., Mirels, L.F., and Stevens, D.A. (2002). Resistance mechanisms in clinical isolates of *Candida albicans*. *Antimicrob Agents Chemother* 46, 1704-1713.

## REFERENCES

- Wilson, L.S., Reyes, C.M., Stolpman, M., Speckman, J., Allen, K., and Beney, J. (2002). The direct cost and incidence of systemic fungal infections. *Value Health* 5, 26-34.
- Wilson, R.B., Davis, D., and Mitchell, A.P. (1999). Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. *J Bacteriol* 181, 1868-1874.
- Wipe, B., and Leisinger, T. (1979). Regulation of activity and synthesis of N-acetylglutamate synthase from *Saccharomyces cerevisiae*. *J Bacteriol* 140, 874-880.
- Wisplinghoff, H., Bischoff, T., Tallent, S.M., Seifert, H., Wenzel, R.P., and Edmond, M.B. (2004). Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. *Clin Infect Dis* 39, 309-317.
- Wisplinghoff, H., Seifert, H., Tallent, S.M., Bischoff, T., Wenzel, R.P., and Edmond, M.B. (2003). Nosocomial bloodstream infections in pediatric patients in United States hospitals: epidemiology, clinical features and susceptibilities. *Pediatr Infect Dis J* 22, 686-691.
- Wisselink, G.J., van Zanten, E., and Kooistra-Smid, A.M.D. (2011). Trapped in keratin; a comparison of dermatophyte detection in nail, skin and hair samples directly from clinical samples using culture and real-time PCR. *Journal of Microbiological Methods* 85, 62-66.
- Woolford, J.L., Jr., and Baserga, S.J. (2013). Ribosome biogenesis in the yeast *Saccharomyces cerevisiae*. *Genetics* 195, 643-681.
- Xin, H., Cartmell, J., Bailey, J.J., Dziadek, S., Bundle, D.R., and Cutler, J.E. (2012). Self-adjuvanting glycopeptide conjugate vaccine against disseminated candidiasis. *PLoS One* 7, e35106.
- Yan, C., and Higgins, P.J. (2013). Drugging the undruggable: transcription therapy for cancer. *Biochim Biophys Acta* 1835, 76-85.
- Yeh, J.E., Toniolo, P.A., and Frank, D.A. (2013). Targeting transcription factors: promising new strategies for cancer therapy. *Curr Opin Oncol* 25, 652-658.
- Yona, A.H., Manor, Y.S., Herbst, R.H., Romano, G.H., Mitchell, A., Kupiec, M., Pilpel, Y., and Dahan, O. (2012). Chromosomal duplication is a transient evolutionary solution to stress. *Proc Natl Acad Sci U S A*.
- Zakikhany, K., Naglik, J.R., Schmidt-Westhausen, A., Holland, G., Schaller, M., and Hube, B. (2007). In vivo transcript profiling of *Candida albicans* identifies a gene essential for interepithelial dissemination. *Cell Microbiol* 9, 2938-2954.
- Zhang, Y., Liu, T., Meyer, C.A., Eeckhoute, J., Johnson, D.S., Bernstein, B.E., Nusbaum, C., Myers, R.M., Brown, M., Li, W., *et al.* (2008). Model-based analysis of ChIP-Seq (MACS). *Genome Biol* 9, R137.