Activation of non-zinc cluster transcription factors in *Candida albicans*

Manjari Shrivastava

A Thesis

In the Department

of

Biology

Presented in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy (Biology) at

Concordia University

Montreal, Quebec, Canada

September 9th 2021

© Manjari Shrivastava

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Manjari Shrivastava

Entitled: **Activation of non-zinc cluster transcription factors in *Candida albicans***

and submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy (Biology)

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

______________________________ Chair
Dr. Wayne Brake

______________________________ External Examiner
Dr. Theodore White

______________________________ External to Program
Dr. Brandon Findlay

______________________________ Examiner
Dr. David Kwan

______________________________ Examiner
Dr. David Walsh

______________________________ Thesis Supervisor
Dr. Malcolm Whiteway

______________________________ Thesis Co supervisor
Dr. Steve Shih

Approved by ______________________________
Dr. Robert Weladji (Graduate Program Director)

September 9th 2021 ______________________________
Dr. Pascale Sicotte (Dean of Faculty of Arts and Science)

Abstract

Activation of Non-zinc cluster transcription factors in *Candida albicans*

Manjari Shrivastava, Ph.D.

Concordia University, 2021

The binding of transcription regulators to cis-regulatory sequences is a key step through which cells regulate expression of their genes. Due to gains and losses of cis-regulatory sequences and changes in the transcription regulators themselves, the binding connections between regulators and their target genes can change rapidly over evolutionary time, a process called transcription factor rewiring. The goal of this study is to determine, on a large scale, the rewiring of the transcription factors in *Candida albicans,* an opportunistic human pathogen, compared to *Saccharomyces cerevisiae* the non-pathogenic baker's yeast. Previous work had created a library of activated zinc cluster transcription factors in *C. albicans*. In this work our target is to activate the remaining non-zinc-cluster transcription factors and predict their function through analysis of gene expression and cellular phenotypes. We used VP64, a strong trans-activation domain, for fusion to different DNA-binding domains for the activation of transcription factors, and bioinformatics tools to search for candidate rewired transcription factors in *C. albicans*. Transcription factors play key roles in cellular regulation and are critical in the control of drug resistance in the fungal pathogen *C. albicans*. We found that activation of the transcription factor Orf19.2752 (Adr1) conferred significant resistance against fluconazole. We found that Adr1 is involved in ergosterol biosynthesis in *C. albicans*. The rewiring from ergosterol synthesis to fatty acid metabolism involved all members of the Adr1 regulon except the alcohol dehydrogenase Adh2, which remains under Adr1 control in both circuits. The second part of our studies lead us to investigate a multi-branched sulphur metabolism pathway in detail. In the sulphur assimilation pathway, we have encountered multiple rewiring events whose functional details we are establishing.

Acknowledgments

Ph.D. is a whole different life which you experience and grow from a new, naïve and unshaped potentials clay to well-shaped pot. A professor plays an extremely important role in shaping/ guiding the students not only for the research but also for building our personality. I am blessed that I got you Malcolm, as my guide. In every step I learned many things from you, most importantly the patience which you show towards our stupidity and childishness. Thank you for always being there even after multiple failures specially PCR reactions and frustrations which I shared with you during our weekly meetings. I would also like to thank Dr. Shih, Dr. Kwan and Dr. Walsh, my committee for always being there for me with those excitements which you showed towards my research. Thank you for your valuable comments in completing my thesis project. I would like to extend my special thanks to all the current and former Whiteway lab research group members specially Walters, Jaideep, Amjad and Jinrong. I would like to thank Anna, Saima, Camilla, Raha, Raymond for being there with me during this stressful time.

Ph.D. indeed made me poor, without money but rich in thoughts. But thankyou to Prof Weladji for nominating me for Accelerator award and Split merit scholarship. That saved me a lot. Thank you, Malcolm, and Orly, for selecting me for SynBioApp program I really learned a lot and made some good friends.

Being an international student with no friends around was difficult to mend in but a few people in our department made it easier. I would like to thank my coffee buddies Alaa, and Farhat for always listening to my non-stop complains and teaching me everything about the “Canadian way”. I would also like to thank Subhiksha and Mudabir who are currently

experiencing similar multicultural diversity like me during last 3 years. A family is a wall which protects you from everything but will bend in any way to drive you crazy. Well, I am born crazy, lazy and clumsy. My dad truly believes I will become fully mad after my Ph.D. But you and mom, both always supported me and thank you for always being there and for keeping me on the ground. And when you didn't my sister was there always to fight for me and do all the dirty work. I cannot thank enough to my sister Jhumi di and Brother Shanu for their unconditional support. And off course it was impossible to finish this journey without Thulaj, my fiancé. I don't think I could have completed my PhD voyage without yelling and releasing my daily stress and frustrations on you. Thank you for you support. Last but not least thank you Apu and Srish for your emotional and constant support. Lastly, thank you to all those were a part of this journey.

Contribution of Authors

Chapter 2:

Gaëlle Séréna Kouyoumdjian (490 student): helped for screening the activated transcription factors.

Manjari Shrivastava (PhD project): experiment designing, conceptual understanding, writing manuscript and performing the experiment

Chapter 3:

Eftyhios Kirbizakis (Master student in Hallett's lab): performed in silico experiments and also helped in analysing RNA-seq data.

Gaëlle Séréna Kouyoumdjian (490 student): helped in screening the strains against different drugs.

Daniel Ruiz (Mitacs summer student): Helped in activating Adr1 transcription factor.
Manjari Shrivastava (PhD project): experiment designing, conceptual understanding, writing manuscript and performing the experiment.

Chapter 4:

Dr. Jinrong Feng (exchange scientist) helped in CRISPR cas9 deletion of Met4 and Met32. He also guided and solved few problems during this project.

Mark Coles (490 student): Helped in in activating Met4 transcription factor.

Benjamin Clark (490 student): performed in silico experiments and tried to delete Met28 transcription factor.

Dr. Amjad Islam (senior Ph.D. student): Guided and thought me the basic experiments and lab rules. He also helped me in brainstorming the Met-circuitry.

Manjari Shrivastava (PhD project): experiment designing, conceptual understanding, writing manuscript and performing the experiment

Table of Contents

List of Figures ……………………………………………………………………………..viii

List of Tables.…………………… ……………………………………………..……….ix

List of Abbreviations ………………………………..…………………….…..……...x

Chapter 1: Introduction …………………………………………………..…….……..1

1.1 *Candida* Species and their virulence…….………..……………..………....3

1.2 *Candida* Morphology …………………………….……………………...…4

1.3 Morphology during

Infection………………………………………………………….……...6

1.4 Treatment against Candida infection……………………………………….…7

1.5 Drug resistance in *C. albicans*…………………………………………..………….9

1.6 What is a regulatory network? …..…………………………………..……....13

1.6.A Transcription regulatory network…….……………………….……..……......13

1.6.B Post transcription regulatory network …………..…………………..………..18

1.7 Possible pattern or reason for rewiring events ……...………….…...…………19

1.8 Activation and selection of transcription factor for this study ……….................20

Objectives …………………………………………………….………………....…21

Chapter 2: The screening of activated transcription factor and identification of rewiring

Events 22

2.1 Introduction 22

2.2 Materials and Methods 23

2.3 Results 26

2.4 Discussion 33

Chapter 3: The transcriptional rewiring of Adr1 transcription factor presents novel mode of regulation of the Ergosterol pathway and azole resistance in C. *albicans*…..........................38

3.1 Introduction...38

3.2 Materials and Methods 41

3.3 Results 45

3.4 Discussion 56

Chapter 4: *Modulation of the complex regulatory network for methionine biosynthesis in fungi*

....68

4.1 Introduction 68

4.2 Materials and Methods 69

4.3 Results 72

4.4 Discussion 81

Chapter 5: Conclusions and Future Work...87

References..90

Appendices...105

List of Figures

- Fig. 1.1. Mechanism of drug resistance in *Candida albicans*.
- Fig 1.2. Different types of rewiring in transcription factors between two species. These can be classified as A. Gains and losses of cis-regulatory sequences by simple mutation. B. Trans- regulatory- protein- protein interaction, protein changing its binding. C. Gain of a favorable protein–protein interaction through a small number of point mutations.
- Fig. 2.1.A. *C. albicans* strain with activated Gcn4 growth in 3 amino triazole (3AT) compared with wild type. B. *C. albicans* strain strains with activated Cap1 growth in SC media with fluconazole. C. Microscopic images of *C. albicans* strain activated Tec1. D. Transcriptomic profile of up-regulated genes in Gcn4 activated strain. These genes are unknown to be directly regulated by Gcn4 and involved in amino acid starvation. (Page 27).
- Fig. 2.2.A. Morphological switch from yeast to hyphae was monitored every hour in two different media,Spider and serum. We found that 9 strains showing faster "switch" than the wild type, although only 5 strains showed a faster switch in both media. B. Activation of 4 transcription factors, C2_05640W_a, Nrg2, C4_07150W_a and Gln2 enhanced filamentous morphology predicting they might influence hyphae formation directly or indirectly, C. Activation of C1_13440C_a triggered opaque cells in the presence of opaque specific media with phloxine B in the media, this result indicates that activation of this transcription factor facilitates the faster switching from white to opaque. (Page 29)
- Fig. 2.3.A. Spot assay results showing tolerance to heavy metals like cesium and arsenic through activation of Met4. B. Spot assay results of temperature tolerance in different strains. Hap42, Uga33, Met32 and Adr1 activation generated tolerance to the high temperature whereas C5_04410C_a activation results in low temperature tolerance. C. Spot assay of pH tolerance in various strains containing activated transcription factors. Ash1, Hap41, Met28, Met32 and C3_01220W_a created tolerance to acidic pH whereas C1_04519W_a, Pho4 and Try4 activation allowed alkaline tolerance. (Page 31)

- Fig. 2.4.A. The 4 strains with activated C2_08950W_a, Hap42, Hap41 and Rfx1 respectively, showed growth in the presence of DNA damaging agent hydroxyurea. B. The impact of activated transcription factors in adhesion of *C. albicans* on the surface. (Page 32)
- Fig. 3.1.A. VP64-Gcn4 chimeric transcription factor generates resistance to 3-amino-triazole (3AT). B. VP64-Tec1 TF triggers hyphal elongation in YPD media. C. VP64-Cap1 allows growth in SC media containing fluconazole. D. Schematic representation of workflow involved in activating the transcription factors. (Page 46)
- Fig. 3.2. A. Plate assay of YPD agar with different drugs confirming the VP64 fusion of Orf19.2752p generates fluconazole, amphotericin B and terbinafine resistance. B. Activated Orf19.2752p caused fluconazole, posaconazole, terbinafine and amphotericin B resistance compared to wildtype. C. Sequence alignment shows that Orf19.2752p is similar to the Adr1 protein of S. cerevisiae. (Page 48)
- Fig. 3.3. A. Deletion of *ADR1* (ORF19.2752). B. Sterol estimation of wildtype and the Orf19.2752 activated strain. Presence of the Adr1 motif upstream of ergosterol genes. D. Transcriptomic profile of selected genes of the Orf19.2752-VP64 fusion strain shows upregulation of ergosterol genes. (Page 51)
- Fig. 3.4. A. Adr1 and Upc2 influence response to fluconazole. B. Fold change in MIC from Adr1 and Upc2 activation and deletion. C. Upc2 DNA binding motif in upstream of various regulons including *ADR1*. (Page 54)
- Fig. 3.5. Proposed model of the multiple rewiring of transcription factors between *S. cerevisiae* and *C. albicans* along with coordination of Upc2, Adr1 and Mrr2 during the presence of azoles or depletion of ergosterol leading to drug resistance in *C. albicans*. (Page 61)
- Supplementary fig. 3.1. A. The construct for activation of non-zinc cluster transcription factors. B. The *adr1*Δ/Δ strain (ORF19.2752) showed fluconazole sensitivity. The heterozygous *adr1*Δ/+ shows better growth than the *adr1*Δ/Δ in the presence of fluconazole. C. The *mrr2* Δ/Δ in the Adr1 activated strain caused no changes in drug resistance. (Page 66)

- Fig. 4.1. A. Phylogenetic tree. (Page 73)
- Fig. 4.2.A. Activation and Deletion of Met4 and Met32 in *C. albicans*. (Page 75)
- Fig. 4.3.A. Genes upregulated by Met28 activation. (Page 77)
- Fig. 4.4.A. Search for the Met4 Motif. (Page 79)
- Fig. 4. 5. Phylogenetic analysis of methionine regulatory transcriptional factors across the ascomycetes. (Page 85)
- Supplementary Fig 4.1 The presence within the promoters of methionine genes across the fungal phylogeny of the candidate Met4 motif 5' ATGRYRYCAT 3' is represented by black boxes and the new potential Met4 recognition motif of the CUG clade 5' CAACTCCAAR 3' is represented with grey boxes. Some genes in some species have both motifs; these cases are represented by a box split into a black and a grey section. (Page 86)

List of Tables

Table 2. List of the activated transcription factor with predicted function, Orthologs and possibility for rewiring. (Page 43)

Table 3.1. Screening of activated transcription factors in presences of different drugs. (Page 65)

List of Abbreviations

Transcription factors	Tfs
5-fluoroorotic acid	5-FOA
N-Acetylglucosamine	GlcNAc
Yeast Nitrogen Base	YNB
Phosphate-Buffered Saline	PBS
Polymerase Chain Reaction	PCR
Room Temperature	RT
Tetracycline	TET
Hydroxy Urea	HU
Hydrogen Peroxide	H_2O_2
Ultra Violet Radiation	UV
Wild Type	WT
Yeast extract, Peptone, Dextrose	YPD
Differential Interference Contrast	DIC
Carbon Dioxide	CO_2
Clustered Regularly Interspred Short Palindromic Repeats	CRISPR
CRISPR associated protein 9	Cas9

Chapter 1: Introduction

The fungal kingdom, consisting of molds, mushrooms, lichens, rusts, smuts and yeasts, comprises a group of eukaryotes with remarkably diverse life histories that make essential contributions to the biosphere, to human industry, and to medicine and research. No one knows how many fungal species exist; as many as 100,000 have been described, as many as 1.5 million have been estimated to exist in nature (Stajich et al., 2009) and their phylogenetic tree is proposed to span over more than a billion years of evolution (Hedges et al., 2004). Fungi are known as principal decomposers in ecological systems, and therefore play an important role in various nutrient cycles. The human use of fungi for various purposes including food preparation or preservation is extensive and has a long history. Mushroom farming and mushroom gathering are large industries in many countries. Because of the capacity of this group to produce an enormous range of natural products with antimicrobial or other biological activities, many species have long been used or are being developed for industrial production of antibiotics, vitamins, and anti-cancer and cholesterol-lowering drugs (Copetti, 2019; Hasan et al., 2015). More recently, methods have been developed for genetic engineering of fungi. Yeasts are part of our everyday life, and include important industrial organisms not only involved in the production of numerous fermentation-associated products such as wine, beer, and cider but also purely ethanol-based products such as bioethanol, now being extensively used as a gasoline replacement in countries like Brazil. Yeasts are versatile organisms, and some have even been engineered to produce various natural compounds that are components of human drugs, such as the anti-malaria drug precursor artemisinic acid (Paddon and Keasling, 2014; Ro et al., 2008).

Yeasts frequently serve as model organisms for the study of the molecular genetics of eukaryotes, because some fungi, like the budding yeast *Saccharomyces cerevisiae*, are easily cultured and genetically manipulated. Scientists make use of these advantages to gain insights into diseases such as Parkinson's, Cancer, Alzheimer's, Huntington's and other human diseases by examining homologous genes in *S. cerevisiae.* Thus, while some fungi can cause disease and dramatically threaten human health, overall, their benefits to humans as well as the planetary ecosystem clearly outweigh the negative aspects. Studying fungi is therefore a fascinating area and certainly provides a rich resource to understand processes essential for life (Reece et al., 2011).

Because fungi are increasingly being recognized as ubiquitous members of the human microbiome, having been found on nearly all mucosal surfaces, interest is growing in understanding how these organisms may contribute to health and disease. These microbial communities are under the influence of our immune system and are important for normal immune system development and maintenance of healthy tissues and physiological processes. Individuals who are immune compromised are vulnerable to fungal infections. These includes individuals with AIDS, chronic diseased individuals and patients who have undergone chemotherapy, or are immunosuppressed to facilitate organ transplantation. The severity of fungal diseases varies from superficial infections to life threatening systemic infections. Most of the infections are opportunistic because the infecting agents are commonly found as commensal organisms residing on the human hosts as part of their microbial flora (Limon et al., 2017).

Fungi are members of the normal microflora of the human oral, gastrointestinal and urinogenital tracts. The mycobiome of human body consists of members of the Ascomycota phylum (including *Candida* spp., *Cladosporium* spp., and *Saccharomyces ssp.*) and the Basidiomycota phylum (including *Cryptococcus* spp., *Filobasidium* spp. and *Malassezzia* spp.) that have been shown to dominate most human anatomic sites (Ghannoum et al., 2010; Hoffmann et al., 2013; van Woerden et al., 2013; Zhang et al., 2011). Their natural occurrence makes them candidates for use as probiotics. Several mechanisms of action have been identified in yeasts as a part of our microbiota which includes regulation of intestinal microbial homeostasis, interference with the ability of pathogens to colonize and infect the mucosa, modulation of local and systemic immune responses, stabilization of the gastrointestinal barrier function and induction of enzymatic activity favoring absorption and nutrition (Czerucka et al., 2007; Im and Pothoulakis, 2010; McFarland, 2010) Probiotic yeasts are very commonly used in patients being treated for diseases like inflammatory bowel disease, urinary tract infections, and diarrhea, where they alleviate some of the negative side effects caused by the medications. Compared to bacteria, yeasts like *Saccharomyces cerevisiae* have many advantages as probiotics. Yeast has a well-defined signal transduction pathway (the mating pathway) that can be functionally linked to human G protein-coupled receptors (GPCRs). This helps in enabling controlled expression of proteins in response to stimuli relevant to human physiology (Brown et al., 2000; Shaw et al., 2019). For example, a

probiotic strain of *S. cerevisiae* was developed, where in response to ATP detection through expression of an engineered human P2Y2 GPCR, it secretes the CD39-like ATP-degrading enzyme apyrase. This strain has the potential to be used for the treatment of inflammatory bowel disease.

The most effective and clinically approved probiotic yeast is *Saccharomyces boulardii* (Guslandi et al., 2003; Guslandi et al., 2000). This yeast is used as both a preventive and therapeutic agent for diarrhoea and other GI disorders caused by the action of antimicrobial agents. *S. boulardii* possesses many properties for a potential probiotic agent; it survives transit through the GI tract, its temperature optimum is 37°C, both *in vitro* and *in vivo*, and it inhibits the growth of a number of microbial pathogens. Apart from yeast, filamentous fungi can also act as a probiotics. Saleh and collaborators reported that *Aspergillus awamori*-based probiotic improved growth performance, increased muscle and vitamin E content, modified the skeletal muscle fatty acid profile and decreased lipid peroxidation in the muscles of broiler chickens (Saleh et al., 2011a). In addition to the growth promoting effect, they also demonstrated the capability of *A. awamori* in alleviating the muscle protein breakdown, abdominal fat content and plasma cholesterol of broilers (Saleh et al., 2011b).

1.1 *Candida* species and their virulence

Candida albicans is a very common opportunistic human pathogenic fungi, and can cause serious infections in immunocompromised individuals (Odds, 1987). While there are many Candida species that are known to cause infections, such as *C. albicans*, *Candida tropicalis*, *Candida glabrata*, *Candida dubliniensis*, *Candida parapsilosis*, and now *Candida auris*, the most common and most studied species is *C. albicans*. Candida can cause a range of diseases, from a simple superficial infection to potentially lethal systemic candidiasis. Candidiasis now represents the third-to-fourth most frequent nosocomial infection in hospitals worldwide (Beck-Sague and Jarvis, 1993; Edmond et al., 1999; Wisplinghoff et al., 2014; Wright and Wenzel, 1997). It was the 4th most prevalent pathogen causing sepsis in the United States (Wisplinghoff et al., 2004) and in Europe (Suetens et al., 2013), with a candidemia rate of 0.20-0.38 cases per 1000 hospital admissions (Tortorano et al., 2004). *C. albicans* is the most frequently isolated species from

patients, accounting for almost 50% of all isolates in Italy (Tortorano et al., 2013) and in Europe (Pfaller and Diekema, 2002) and 45.6% in the United States (Horn et al., 2009).

1.2 Candida Morphology

C. albicans is a fundamentally dimorphic fungus. This means that *C. albicans* exists primarily in two different morphological forms, which are an oval shaped yeast form and a branched, elongated hyphal form. Under normal laboratory conditions, *C. albicans* tends to grow as ovoid "yeast" cells but environmental changes in temperature and pH can induce a morphological shift to pseudohyphal growth. When *C. albicans* cells are grown in a medium that mimics the physiological environment of a human host, they grow as "true" hyphae. The hyphal form has an important role in causing disease by invading epithelial cells and causing tissue damage (Sudbery, 2011). Yeast to hyphal switching is the most frequent morphological transition in *C. albicans* (Brand, 2012).

Two key pathways that play central roles in the regulation of filamentous growth are the cyclic-AMP/ protein kinase A (cAMP/PKA) pathway and a mitogen activated protein kinase (MAPK) pathway (Cullen and Sprague, 2012; Si et al., 2013; Sonneborn et al., 1999). The cAMP/PKA pathway acts as a major regulator of filamentation. The adenylyl cyclase Cyr1 is activated by external signals and this activated Cyr1 then converts ATP to cAMP (Hogan and Sundstrom, 2009). cAMP acts as a secondary messenger by binding to Bcy1, the regulatory subunit of PKA, causing conformational changes in PKA thus releasing the two catalytic subunits Tpk1 and Tpk2. These activated catalytic subunits then activate downstream filamentation regulators such as Efg1 via phosphorylation (Bockmuhl and Ernst, 2001; Stoldt et al., 1997). The MAPK pathway, on the other hand, signals through a cascade of kinases (Cst20, Hst7, Cek1) and ends on in the activation of transcription factor Cph1 which then induces the filamentous response genes (Liu et al., 1994).

Another important morphological change possible in *C. albicans* cells is white-opaque switching. *C. albicans* can switch from the mating sterile white form to the mating competent opaque form. White cells are small and round and form white, dome-shaped colonies on solid medium, while opaque cells are slightly larger with an elongated shape, and their colonies give a greyish and

flatter appearance on agar medium. As well, white and opaque forms have other distinguishable characteristics - opaque colonies are stained pink with phloxine B while white form colonies are not; the white form is better in systemic or cutaneous infections, while the opaque forms are good in adhesion to the skin; opaque cells are more resistant to different drugs; the two forms have different lipid contents and show significantly different gene expression profiles (Anderson et al., 1990; Anderson and Soll, 1987; Ghannoum et al., 1990; Kennedy et al., 1988).

Wor1 is the key player in Candida white-opaque switching (Huang et al., 2006). *C. albicans* is a diploid organism and most of the wild type isolates are heterozygous at the Mating Type-Like (*MTL*) locus designated as *MTL a/α*. However, to switch to opaque form, the cells need to be homozygous at the *MTL* locus. The *MTL* locus is located on chromosome 5 and it consists of two heteroallelic loci, *MTLa* and *MTLα*. The *MTLa* locus contains genes for two transcriptional regulators, *a1* and *a2* where the a2 protein is required for activation of *MTLa* specific genes. The *MTLα* locus also contains two transcriptional regulator genes, *α1* and *α2* where α1 is involved in the activation of *MTLα* specific genes. In addition, the a1 and α2 proteins form a heterodimer that can bind to the promoter region of, and repress expression of, Wor1. Therefore, homozygosity at the *MTL* locus allows for Wor1 expression and thus for the white to opaque transition. Homozygosity at the *MTL* locus is generally achieved either through gene conversion or chromosome loss followed by duplication of the retained copy (Soll, 2014). Besides the mating-related regulatory genes, each of the *MTL* loci contains non-sex genes (NSGs), designated as *PIK*, *OBP* and *PAP,* that code for phosphatidylinositol kinase, an oxysterol binding protein-like protein and a poly(A) polymerase respectively (Hull and Johnson, 1999). The *OBP* locus has been implicated in *C. albicans* biofilm formation and in fluconazole resistance (Srikantha et al., 2012).

The MAPK pathway implicated in the hyphal transition is also the main regulatory pathway of Candida mating (Chen et al., 2002). The pathway signaling starts when pheromones bind to the transmembrane receptors; Ste2 (the 'α' pheromone receptor) and Ste3 (the 'a' pheromone receptor) (Dohlman et al., 1998). Upon binding of pheromone to the receptor, the signal is transduced to a heterotrimeric G protein which causes dissociation of Gα from Gβγ (Dignard et al., 2008; Kurjan, 1993; Lu et al., 2014). The Gα and Gβγ subunits transfer the signal through a cascade of kinases that include Cst20, Ste11, Hst7 and Cek1, to activate the transcription factor Cph1 via release of

the inhibitory activity of Dig1 (Bardwell et al., 1998; Regan et al., 2017). Finally, the activated Cph1 induces the pheromone and mating response genes (Monge et al., 2006).

1.3 Morphology during Infection

Both yeast and hyphae/filamentous forms are critical for *C. albicans* virulence (Lo et al., 1997). The yeast form is important for pathogen dissemination via the bloodstream and therefore involved in Candida virulence (Jacobsen et al., 2012). The filamentous form can invade host tissues *in vivo* or agar medium *in vitro*. The invasive filamentous form is induced in response to different environmental signals such as increased temperature, nutrient limitation or the presence of serum (Brown and Gow, 1999). Upon proper signal induction, yeast cells can initiate the invasion of tissues and organs via formation of hyphae or pseudohyphae (Csank et al., 1998; Felk et al., 2002). After host tissue invasion, the pathogen can then enter the bloodstream and spread to other parts of the host thus creating a systemic infection (Felk et al., 2002).

A normal habitat of *C. albicans* cells are mucosal membranes including those of the mouth, gut, vagina, and sometimes the skin of humans and various other mammals. Generally, *C. albicans* does not cause any damage and exists as a commensal organism within the human or animal host. The normal bacterial flora of the gut, mouth, and vaginal mucosa can act as a barrier to the growth of fungal infections. Loss of this normal flora, after for example prolonged antibiotic treatment, is one of the factors may cause an infection by *C. albicans*.

Normally, Candida is predominantly present in the yeast form, but when *C. albicans* becomes pathogenic it typically switches the phenotype to the hyphal form to facilitate invasion of the host cell epithelium. However, interactions between *C. albicans* cells and mammalian host tissues are highly complex, and Candida infections generally arise through a sequence of time scaled steps -

- Adhesion to an epithelial surface is required to initiate the colonization of the actual surface. (Preliminary candidiasis)
- Penetration into the epithelial surfaces is the limit of the infectious process in most cases, leading to the establishment of a superficial candidiasis; normally the fungus is incapable

of further invasion into the immunologically intact host. In immunologically compromised hosts however, this can be followed by tissue invasion.

- When the fungal cells reach the blood stream, they must face the blood-borne cellular host defense system and Candida cells typically must be able to penetrate the endothelial surfaces and invade the tissues in order to cause systemic infection (Prasad, 1993, 2017; Prasad and Ghannoum, 1996).

1.4 Treatment of Candida infection

In general, treating patients suffering from candidiasis is made difficult by the fact that the immune systems of these individuals are already compromised. Treating such patients has always been challenging but was significantly improved in the 1950's when the first antifungal agent, nystatin, was isolated from a soil fungus (El, 1950). Over the next 60 years, antifungal drug development has continued, and today we have several classes of antifungal drugs that are used in the clinic to treat fungal infections. Broadly we can divide the antifungal drugs into the following 4 different categories-

- Polyenes

The class of polyenes includes several hundred different drugs, but the most used ones are nystatin and Amphotericin B deoxycholate (AmB). Polyenes are isolated from soil dwelling bacteria of the genus Streptomyces. While nystatin is usually restricted to treating mucosal Candida infections due to solubility issues, AmB has been used to treat all kinds of fungal infections and has the broadest spectrum of currently available antifungal drugs (Pound et al., 2011). Polyenes have an amphipathic nature, i.e. one part of the molecule is hydrophilic and charged, and another side is hydrophobic and uncharged (Sanglard and White, 2006). Through this property, polyenes target ergosterol, a major lipid in fungal plasma membranes. Through binding to ergosterol, AmB creates a pore that allows ions and other cellular constituents to diffuse across the membrane, ultimately leading to fungal cell death (Brajtburg and Bolard, 1996; Vertut-Croquin et al., 1985).

The main disadvantage of polyenes is host toxicity, as they do not only bind to lipids present in fungi, but also to cholesterol-containing membranes present in human cells. This has considerable disadvantages to the human host, generating conditions including nephrotoxicity, anemia, and

anaphylaxis as well as infusion-related reactions that can occur in approximately 70-90% of patients (Gallis et al., 1990; Goodwin et al., 1995).

- Azoles

The class of azole drugs is one of the major groups of antifungals used in clinics. Ketoconazole was the first azole derivative that was approved in 1981 by the Food and Drug Administration (FDA) for clinical use (Heeres et al., 1979) and was followed by the first generation triazole, fluconazole (FCZ), in 1990 (Chapman et al., 2008). Based on its positive pharmacokinetic properties, FCZ rapidly became the first choice for the treatment of several fungal infections and has been studied extensively in clinical settings (Sheehan et al., 1999). Azoles target the fungal specific ergosterol biosynthesis pathway. The ergosterol pathway can be divided into an early pathway, producing squalene from acetate, and a late pathway that produces ergosterol starting from squalene (Sanglard and Odds, 2002). All azoles block the late pathway by inhibiting the product of the *ERG11* gene, the P450-dependent enzyme lanosterol 14-demethylase. Inhibiting this enzyme results in the accumulation of toxic methylated sterols, which are then incorporated into the plasma membrane, replacing ergosterol. This in turn results in altered membrane fluidity as well as altered activity of membrane bound enzymes, such as chitin synthase (White et al., 1998).

- Echinocandins

The echinocandins are the newest class of antifungals and consist of 3 FDA- approved agents, caspofungin, micafungin and anidulafungin. All three echinocandins are marketed as semisynthetic lipopeptides that have been chemically modified from natural products of fungi. All echinocandins target the cell wall of fungal cells by inhibiting the enzyme glucan synthetase, which is involved in the synthesis of the major cell wall polysaccharide, 1,3-glucan (Akins and Sobel, 2009). *FKS1* encodes the glucan synthetase, which is the direct target of the echinocandins. Given that mammalian cells lack cell walls, echinocandin action is fungal specific, and results in fungicidal interactions against most Candida spp. (Antachopoulos et al., 2007; Nakai et al., 2003).

- Flucytosine

5-flucytosine (5-FC), was originally discovered in 1957 for its antitumor activity (Pound et al., 2011); only four years later was its antifungal property recognized (Chaudhuri et al., 1958). 5-FC is taken up by the same transport system as adenine hypoxanthine and cytosine (Waldorf and Polak, 1983). The uptake is an energy dependent process coupled to proton gradient. The influence of the pH on the kinetic constants points to stoichiometry of the one proton being associated to the transport system together with one substrate molecule. Inside the cells 5-FC is rapidly deaminated to the tumor-inhibitory compound 5-fluorouracil by means of a cytosine amino-hydrolase, an enzyme that appears to occur only in yeast and bacteria. The absence or weak activity of this key enzyme in 5-FC metabolism in the mammalian cells has been the basis for the drug's low toxicity for the mammalian host. It should be noted that 5-FC itself cannot be used as antifungal drug since its uptake by fungi is poor. The activity of 5-FC after uptake and deamination seems to be the consequence of intra-fungal formation of the two metabolites 5-fluorodeoxyuridine monophosphate (FdUMP) and 5-fluorouridine triphosphate (Chandra et al., 2009; Diasio et al., 1978; Waldorf and Polak, 1983).

1.5 Drug resistance in *C. albicans.*

The prevalence of resistance to antifungal agents has significantly increased in the past few decades. Such resistance to antifungal agents has important implications for mortality, morbidity and health care in the community. Until recently, fungi were not recognized as increasingly important pathogens because the annual death rate due to *Candidiasis* remained steady from 1950 to 1970. Since 1970, this rate increased significantly due to more widespread use of immunosuppressive therapies, indiscriminate use of broad-spectrum antifungal agents, the common use of indwelling intravenous devices and the epidemic of immunosuppressive viral infections. These developments and the associated increase in fungal infections necessitated the search for new, safer, and more potent agents to combat serious fungal infections. For nearly 30 years, amphotericin B, which has side effects of nephrotoxicity, was the sole drug available to treat serious fungal infections. The imidazoles and the triazoles in late 1980s and early 1990s were major advances in safe and effective treatment of local and systemic fungal infections. Triazoles, in particular fluconazole, have high safety profiles, which has led to widespread use. Fluconazole has been extensively used to treat more than 16 million patients, which includes 300,000 AIDS

patients, in the United States alone since the launch of this drug. Due to selective pressure and the high use of these few antifungal drugs, there have been increasing reports of antifungal resistance. Among them the drug-resistant infection caused by *C. albicans* is most common, and some of the strains of *C. albicans* have even developed multidrug resistance (Prasad and Kapoor, 2005). The mechanisms of resistance to azole antifungal agents have been elucidated in Candida species and can be categorized into 4 main classes -

- Over-expression of the drug target
- Target alteration: Modification in enzymes of the ergosterol biosynthetic pathway affecting azole action.
- Drug degradation
- Increased efflux of drugs by over-expression of drug efflux pump proteins

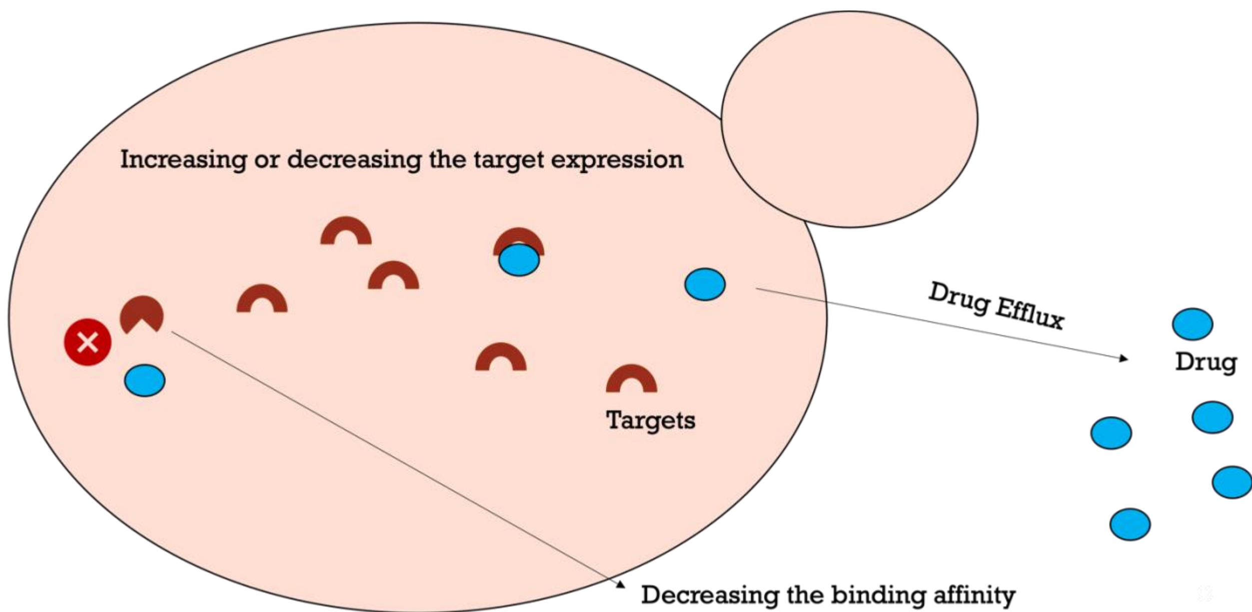


Fig 1.1 Mechanism of drug resistance in Candida albicans

Target alteration and over-expression:

The mechanism by which azoles are imported into the cell remains to be clearly established, although it is suggested that the hydrophobicity of these drugs could facilitate their entry (Ghannoum and Rice, 1999b). There are suggestions that it might proceed via facilitated diffusion through a yet unknown transporter (Zavrel and White, 2015). Once the azole drug enters the cells,

its interaction with lanosterol 14-alpha-demethylase or Erg11 can be modified in two ways - target alteration or over-expression. Several point mutations in the *ERG11* gene that reduce the affinity of the Erg11 enzyme for azoles have been identified; among them the mutation Y132H represents a recent change (Wang et al., 2015). These mutations block normal binding of fluconazole to the Erg11 protein, however there is no effect of the mutations on ergosterol biosynthesis itself. Alterations in other enzymes of the ergosterol biosynthetic pathway can also lead to drug resistance. For example, a defective sterol Δ5, 6 desaturase (Erg3) leads to the development of fluconazole resistance in clinical isolates. This defect results in the accumulation of 14-α-methylfecosterol which compensates for ergosterol depletion (Cheng et al., 2006; Prasad, 2017; Prasad and Kapoor, 2005).

Drug efflux:

Amongst the two super-families of drug transporters, the ABC super-family of transporters is characterized by the presence of nucleotide-binding domains, which, by binding and hydrolyzing ATP, facilitate the efflux of drugs across the membrane. By comparing the protein expression pattern of matched pairs of fluconazole resistant and susceptible clinical isolates, researchers identified several proteins whose expression were upregulated specifically when only the *CaMDR1* gene was overexpressed (Niimi et al., 2006; Wirsching et al., 2000).

In contrast to the ABC drug transporters, the structure-function relationship of major facilitator super-family (MFS) proteins has not been analyzed in detail due to the diversity in their nucleotide and amino acid sequences. There are, however, some reports to suggest that the N-terminal halves of different major facilitator families share greater similarities than their C-terminal halves, which suggests that C-terminal regions are involved in substrate recognition, and N-terminal regions are involved in proton translocation (K. Redhu et al., 2016)

Cdr1p and Cdr2p, ABC transporter family members, are two major drug efflux pumps of *Candida albicans*. These proteins not only efflux azoles and their derivatives but also transport a wide variety of structurally unrelated compounds. Over-expression of the genes *CDR1* and *CDR2* leads to drug resistance in *Candida albicans*. As well, there are two further genes encoding candidate

ABC transporters, *CDR3* and *CDR4*, but they have not been reported to be involved in drug resistance.

Among MFS proteins which act as drug transporters, CaMdr1 is the most associated with drug resistance and its over-expression has been linked to azole resistance in *C. albicans*. The exact mechanism by which the disruption of *CaMDR1* leads to reduced virulence in *C. albicans* is unknown (Chen et al., 2010; Niimi et al., 2004; Prasad et al., 1995a).

Import of drugs:

The hydrophobic nature of many antifungal drugs permits their easy import by passive diffusion. However, the contribution of drug import to the overall scenario of multidrug resistance is not well established since technically it has been difficult to separate efflux of drugs from their import. There are a few studies, particularly with mammalian cells, in which passive diffusion of drugs through a lipid bilayer has been shown to be an important determinant of multidrug resistance. The variations in membrane fluidity are expected to affect passive diffusion of drugs. The enhanced fluidity has been linked to enhanced diffusion of drugs. There are factors, other than membrane fluidity, that can also influence passive diffusion of drugs across the membrane bilayer and thus affect drug susceptibilities in *C. albicans* (Prasad, 2017; Prasad and Kapoor, 2005). Another study suggests that drugs like azoles are imported normally into fungi independent of ATP or pH. As it follows a concentration gradient, and is saturable, uptake apparently proceeds via facilitated diffusion through an unknown transporter (Zavrel and White, 2015).

Apart from these processes, gene or pathway regulators can also influence drug resistance. The changes in the transcriptional and translational regulators of these genes are also one of the strategies behind the drug resistance in C. *albicans*. Upc2 is a global transcriptional factor of the ERG genes (Znaidi et al. 2008; Flowers et al. 2012), and Upc2 is frequently modified in fluconazole-resistant clinical isolates. Constitutive overexpression of *ERG11* via gain-of-function mutations in the transcriptional activator Upc2 is commonly found in resistant isolates of *C. albicans* (Silver et al., 2004). Efg1 is directly involved in the ergosterol biosynthetic process through negative transcriptional regulation of the *ERG3* gene (Lo et al. 2005). A strain containing

a homozygous deletion of the *EFG1* gene shows an enhanced susceptibility to fluconazole that may be caused by overexpression of the *ERG3* gene and consequent accumulation of toxic sterols (Prasad et al. 2010). A third ergosterol biosynthesis-related transcription factor, Ndt80, also plays a role in fluconazole resistance in *C. albicans* (Sellam et al., 2009). A *C. albicans* strain lacking the *NDT80* gene showed increased sensitivity to fluconazole (Chen et al., 2004a; Homann et al., 2009; Sellam et al., 2009).

Apart from ergosterol biosynthesis, the expression of *CDR1* and *CDR2* is regulated by the transcription factor Tac1 in *C. albicans*, with *TAC1* alleles harboring gain-of-function mutations readily identified in resistant isolates (Coste et al., 2006). Similarly, mutations in the transcription factor gene *MRR1* lead to upregulation of Mdr1 in azole-resistant isolates of *C. albicans* (Morschhäuser et al., 2007). Small changes in such regulators can have large impact on *C. albicans* strains and therefore, extensive study of these regulatory networks is needed.

1.6 What is a regulatory network?

A gene (or genetic) regulatory network (GRN) is a collection of molecular regulators that interact with each other and with other substances in the cell to govern the levels of mRNA and proteins relevant to the cellular process being controlled (MacNeil and Walhout, 2011).Transcription factors are one of the key players in such networks. Gene expression is achieved by the action of these regulatory transcription factors that typically bind to cis-regulatory elements that are often located near their target genes. Broadly we can divide regulatory networks into two types:

A. Transcription regulatory networks
B. Post-transcription regulatory networks.

Each of these types of networks can be subdivided into physical and functional networks. Physical networks contain protein-protein, protein-DNA, protein-RNA, and RNA-RNA interactions. Functional networks incorporate the consequences of these physical interactions, e.g., activation or repression of gene expression. Ultimately, transcription and post-transcription regulatory networks need to be combined to obtain a comprehensive picture of all aspects of the regulation

of differential gene expression. Eukaryotic genomes contain thousands of protein and RNA-encoding genes. Some genes are ubiquitously expressed, whereas others are expressed in a tightly controlled manner in only part of the organism, or under specific conditions during development or disease. To understand how differential gene expression is controlled at a genome-wide or systems level, it is important to identify all the *cis*-acting regulatory sequences and *trans*-acting factors involved, and how and when they interact to affect gene expression.

1.6.A Transcription regulatory networks.

Transcription factors (TFs) are typically composed of at least two elements: a DNA binding domain, which serves to interact with its cognate DNA target sequence, and a transcription regulation domain, which serves to activate or repress transcription. Transcription factors play a key role in determining how cells function and respond to different environments, and approximately 4% of *C. albicans* genes code for transcription factors (Homann et al., 2009), which therefore represent the single largest family of proteins in the fungus. TFs are grouped into families based on their predicted DNA binding domains. To date, more than 100 different types of DNA binding domains have been found in eukaryotic organisms (Kummerfeld and Teichmann, 2006). The TFome dynamics of the Saccharomycotina family show the central role of four main TF gene families: Zn cluster, Leucine B zipper (BZIP), C2H2 Zn finger, and homeodomain (HD)-like proteins (Shelest, 2017). TFs in *C. albicans* coordinate critical cellular functions including biofilm formation (Nobile et al., 2009), drug resistance (Cowen et al., 2002), and the transition from a commensal to a pathogenic lifestyle (Liu, 2001).

When transcription regulatory circuits are compared among distantly related eukaryotic organisms, extensive similarities are observed. But there are many incidences where researchers have found that structurally equivalent transcription factors are regulating different genetic circuits in different organisms. This phenomenon has been called “rewiring” of transcription factors. These rewiring events can be seen in various organisms including fungi, plants, and animals. Many times, this rewiring can be seen at a subclass level (Erkenbrack and Davidson, 2015; Tebung et al., 2016), but sometimes rewiring events have been encountered within the same genus (Lind et al., 2015).

An example of a rewiring event in fungi is that of the galactose metabolic genes (needed to convert galactose to glucose) where their transcription activating proteins have shifted from Rgt1-Rtg3 (helix-loop helix proteins) to Gal4 (a Zn finger protein). In *C. albicans* and other ancestral fungi Rgt1 and Rgt3 control the regulation of Gal genes whereas in *S. cerevisiae* it is controlled by Gal4 (Dalal et al., 2016; Martchenko et al., 2007).

In addition, many other rewiring events between these species have been reported (Johnson, 2017a). Another example is when the cis element rewiring between Rap1 and Tbf1 in yeast was also accompanied by a change in the protein domains of co-factors that they interact with. Essentially, a dimer containing Ifh1 is the primary regulator of ribosomal protein genes in both *S. cerevisiae* and *C. albicans*. In *S. cerevisiae*, the dimer is recruited to the ribosomal gene promoter by Rap1 to activate expression, whereas in *C. albicans*, this dimer is directed to its target by Tbf1. Intriguingly, correlated with the transition to the Rap1-regulated circuit in *S. cerevisiae*, the Sc-Ifh1 now contains a Rap1 interaction domain that is not present in the *C. albicans* protein (Mallick and Whiteway, 2013).

Rewiring events have been seen in different organisms. A similar phenomenon can be seen during the development of diverse insect species. The TF Ftz has switched from serving a homeotic role in ancestral insect species, to being involved in segmentation in the Drosophila genus. This switch in Ftz's function is accompanied by the loss of YPWM, a protein sequence motif that is responsible for cofactor interactions with homeotic regulators, and the gain of a LXXLL motif that enables interaction with segmentation-related cofactors and targets (Heffer et al., 2013). Specifically, having acquired the LXXLL motif, Ftz now possesses the capability to dimerize with Ftz-F1 (via the AF-2 domain), a TF controlling segmentation processes across diverse insect species. These TFs cooperatively bind to their target gene engrailed, to activate its expression and initiate this stage of development in Drosophila (Florence, 1997).

Moore and Bornberg-Bauer (Moore and Bornberg-Bauer, 2012) explored the functional implications of protein domain gain, loss and emergence in the proteomes of 20 arthropod species of the pan-crustacean clade. Although they map evolutionary changes in protein structure to function, this study is mainly focused on contrasting the evolution and function of novel protein

domains vs. that of conserved domains across species, and does not venture to explain changes in interaction preferences between TFs on account of sequence changes. Thus it represents primarily a domain-centric analysis, and does not account for the effect of evolutionary changes in short linear motifs (SLiMs), like the LXXLL/YPWM motifs in insects. SLiMs are interaction modules that have been implicated in greatly diversifying functions of protein isoforms (Weatheritt and Gibson, 2012).

The dimorphic lifestyle of *C. albicans* requires specific regulatory networks at the genetic and protein expression level to ensure coordinated expression of genes. A closer look at the evolutionary level can provide multiple examples of changes that make *C. albicans* more adaptable to such lifestyle. One of the examples of the role of transcription factors in adaptation is that of the rewired Rlm1 transcription factor. (Oliveira-Pacheco et al., 2018), studied the impact of lactate, a physiologically relevant carbon source, on the activity of the *C. albicans* Rlm1 transcriptional factor. They found Rlm1 is not only involved in the cell wall integrity pathway but also plays an important role in regulating the flow of carbohydrates into cell wall biosynthesis pathways. Rlm1 mediates cell wall remodelling during carbon adaptation, impacting cellular interaction with immune cells.

Another example of such rewiring is of the *C. albicans* zinc cluster transcription factor Ppr1 that controls the allantoin catabolism regulon (Tebung et al., 2016). Intriguingly, in *S. cerevisiae* Ppr1 serves as a regulator of pyrimidine biosynthesis. It is possible that this transfer of the control of allantoin degradation from Ppr1 to Dal82, together with the repositioning of Ppr1 to the regulation of pyrimidine biosynthesis, may have resulted from a switch to a metabolism that could exploit hypoxic conditions in the lineage leading to *N. castellii* and *S. cerevisiae*.

Such changes in regulatory network appear sometimes very random with the logic difficult to ascertain. For example, in (Lavoie et al., 2010), although transcription of the ribosomal regulon is a critical task for all cells - in *S. cerevisiae* the transcription factors Rap1, Fhl1, Ifh1, and Hmo1 form a multi-subunit complex that controls ribosomal gene expression, while in *C. albicans* this regulation is under the control of Tbf1 and Cbf1. They observe dramatic rewiring of the Cbf1,

Hmo1, Rap1, and Tbf1 regulators, but they did not propose a logical or conclusive reason for this rewiring.

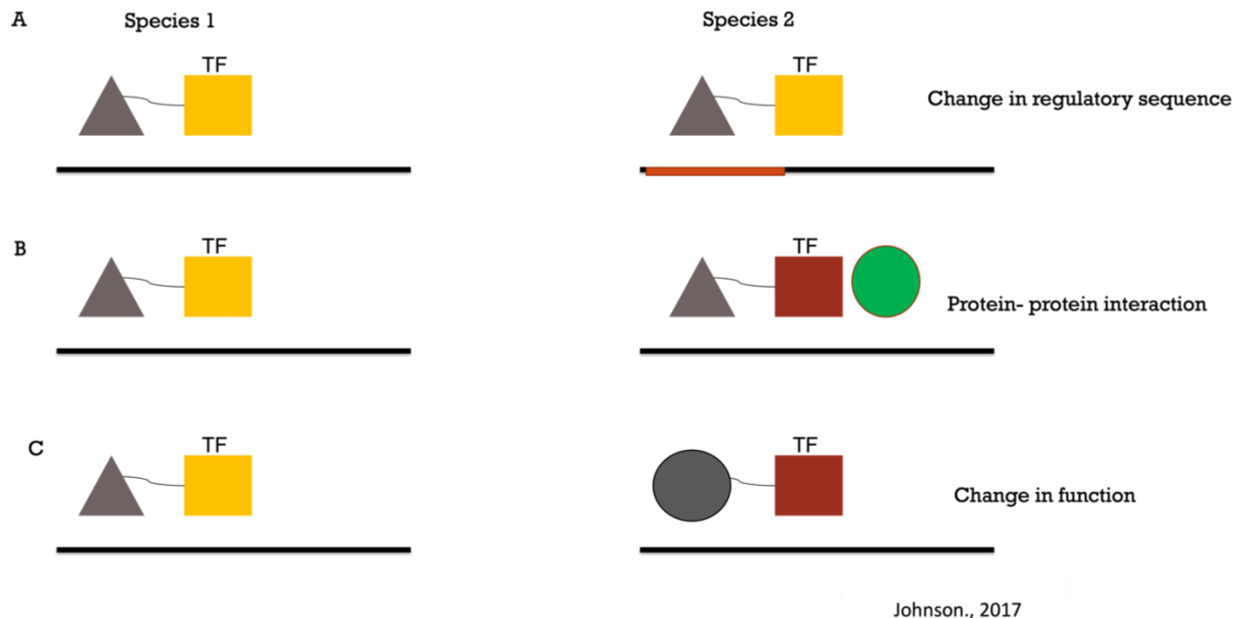


Fig 1.2. Different types of rewiring in transcription factors between two species. These can be classified as A. Gains and losses of cis-regulatory sequences by simple mutation. B. Trans-regulatory- protein- protein interaction, protein changing its binding. C. Gain of a favorable protein–protein interaction through a small number of point mutations.

Understanding how biological networks evolve could eventually help to explain the general mechanism of regulation of cellular systems. (Johnson, 2017b), has described that this rewiring can be classified into three major routes:

(a) Gains and losses of cis-regulatory sequences by simple mutation.
(b) Trans- regulatory- protein- protein interaction, protein changing its binding
(c) Gain of a favorable protein–protein interaction through a small number of point mutations.

a) Gains and losses of cis-regulatory sequences type rewiring

Transcription networks can be rewired through cis-regulatory mutations. Mutations that create or destroy cis-regulatory sequences in the upstream region of a gene may lead to rewiring (Li and Johnson, 2010). One of the examples of such rewiring is the rewiring of the Ppr1 transcription

factor of *C. albicans* and *S. cerevisiae* which we discussed earlier. In *C. albicans* the zinc cluster transcription factor Ppr1 controls the allantoin catabolism regulon whereas in *S. cerevisiae* the Ppr1 ortholog binds the same DNA motif (CGG(N_6)CCG) but regulates pyrimidine biosynthesis (Tebung et al., 2016).

In another study, (Bradley et al., 2010) analyzed the genome-wide binding patterns of six transcription factors involved in initiating segmentation in two closely related fly species. They found that quantitative variation in binding is common and is attributable to the gain and loss of cognate recognition sequences for the factors. These rewirings of transcription factors influence the A-P factor binding during embryogenesis, leading to variations in the development stages. They mentioned that its intriguing how in very similar organisms, with similar timing and structure of embryonic development, there were undoubtedly subtle differences in sampling of developmental stages in the two species.

b) Trans- regulatory and protein- protein interactive rewiring

In protein-protein interactive rewiring, transcription factors work in different combinations to regulate different sets of genes under different conditions. Many combinatorial interactions are due to direct protein–protein contacts between sequence-specific DNA binding proteins. Analysis of the full Mcm1 circuit across species provided clear evidence for transcriptional rewiring via changes in combinatorial interactions. In *S. cerevisiae*, Mcm1 is constitutively expressed and works with different partners to regulate different biological processes, including mating type specification, the cell cycle, and arginine metabolism. To investigate the evolution of regulons defined by Mcm1 and its partners, (Tuch et al., 2008) performed ChIP-chip analysis in three different species; *S. cerevisiae*, *Kluyveromyces lactis* and *C. albicans*. Only about 15% of the direct Mcm1–target gene interactions of *Saccharomyces cerevisiae* were preserved in two other yeast species. Mcm1 binds cooperatively to DNA with a set of cofactors to regulate many genes in each species, and the extensive rewiring observed was traced to high rates of gain and loss of cis-regulatory sequences as well as to the formation of new Mcm1-cofactor combinations and the breaking of old ones.

Another example of this type of this rewiring of Ifh1 and Fhl1 interactions with transcription factors. In *C. albicans* circuit the transcription regulators Ifh1 (interacts with Forkhead 1) and Fhl1 (Forkhead-like 1) is targeted to the ribosomal protein genes by the DNA binding factor Tbf1 whereas in *S. cerevisiae* Rap1 directs the Ifh1-Fhl1 module to the ribosomal protein genes (Mallick and Whiteway, 2013).

1.6.B Post transcription regulatory network

Post-transcriptional regulation is the control of gene expression at the RNA level; between the transcription and the translation of the gene. After being produced, the stability and distribution of the different transcripts is regulated by means of RNA binding proteins (RBPs) that control the various steps and rates controlling events such as alternative splicing, nuclear degradation (exosome), processing, nuclear export (three alternative pathways), sequestration in P-bodies for storage or degradation and ultimately translation. These RNA-binding proteins achieve these events thanks to an RNA recognition motif (RRM) that binds a specific sequence or secondary structure of the transcripts, typically within the UTR region of the transcript.

Post-transcriptional gene regulation plays an important role in determining biological complexity. Post-transcriptional regulons, consisting of RNA-binding proteins (RBPs) and their target genes, are essential for the coordination of gene expression (Keene, 2007; Mata et al., 2005; Woodson and Chory, 2008). Research has demonstrated that post-transcriptional regulatory networks could also be rewired during evolution. Puf3 and Puf4 belong to a group of well-characterized post-transcriptional regulators that contain a PUM-HD-type (PUMilio homology domain, which comprises eight helical repeats) RNA-binding domain to bind mRNAs at the 3'UTR (un-translated region). By binding with their target mRNAs, PUF proteins can recruit factors to transfer the target mRNAs into various subcellular locations, repress translation of mRNAs, and promote mRNA degradation (Jiang et al., 2012). Although Puf4p is highly conserved in fungi, targets of the post-transcriptional regulon are functionally diverse among known fungal species. In the *Saccharomycotina* subdivision, target genes of Puf4p mostly conduct functions in the

nucleolus; however, in the *Pezizomycotina* subdivision, the targets are enriched in the mitochondria. Furthermore, different regulation efficiencies of mitochondrial function by PUF proteins in different fungal clades are observed (Jiang et al., 2012; Jiang et al., 2014).

1.7 Possible patterns or reasons for rewiring events

Deciphering the pathways of network evolution is intriguing. (Carvunis et al., 2015) performed an integrated analysis of transcriptional network evolution by examining mRNA expression, transcription factor binding and cis-regulatory motifs across 25 animal species, including mammals, birds and insects and found that the rewiring occurs at the same rate. Their analysis revealed that the species divergence times are consistent, with similar rates of rewiring events. In rough terms, if the two species diverged 100 million years ago, they should have only 10% similar genetic regulatory networks (Johnson, 2017b). But species like *C. albicans* and *S. cerevisiae* apparently diverged around 300 million years ago yet show more similarity than would be predicted by this rewiring rate. Using computational cis-regulatory sequence detection across many transcriptional regulators, (Habib et al., 2012), estimated that more than 16% of the regulator-target gene connections are preserved between these two species. Separately, (Sarda and Hannenhalli, 2015), also using a computational approach, document extensive rewiring among fungal species over this same time scale. They investigated around 1,700 regulons and 170 TFs and identified around 5,000 significant rewiring events; this leads to less than 10% similarity among these species. It can be argued that focusing on the cis regulatory network rewiring biases these studies. In many cases the transcription factors also show redundancy in their specific binding sequence. For example, (Liang et al., 1996) showed that Gal4 of *S. cerevisiae* has an affinity to bind with more than 15 different sequences.

I suggest a large-scale survey of two species in-silico as well as with experimental determination of rewiring would give a better view of the rewiring patterns. Transcription factors may be expressed only in specific conditions; therefore, deleting the transcription factor may result in mutants with the phenotype of wild type cells unless the specific conditions are identified. Therefore, I created a library of activated non- zinc cluster transcription factors and searched for

activated circuits. These circuits will be compared to the regulatory elements for the circuits in other species to look for rewiring events.

1.5 Activation and selection of transcription factor for this study.

For this study, the search for potential transcription factors starts by BLAST search and comparison of transcription factors in *C. albicans* and *S. cerevisiae.* In *C. albicans* there are around 250 transcription factors whereas in *S. cerevisiae* there are only 216. Based on the type of DNA binding domains we can classify these transcription factors into different categories. These categories are zinc finger, zinc cluster, B-zip, bHLH, C2H2, Mad box, forkhead, homeo domain etc. In *C. albicans* we have 82 zinc cluster transcription factor, 53 C2H2, 41 zinc finger, 13 B-zip, and around 11 bHLH transcription factors. Previously, (Schillig and Morschhäuser, 2013), found that zinc cluster proteins can be artificially turned on by C-terminal fusion to a mutationally activated *S. cerevisiae* Gal4 activation domain. They used this strategy to create a library of *C. albicans* strains expressing all 82 zinc cluster transcription factors of this fungus in a potentially hyperactive form, and this library has been a powerful tool in elucidating the function of a variety of the TFs (Berkow and Lockhart, 2017; Tebung et al., 2016; Tebung et al., 2017).

In this study, we approached the question differently by selecting the transcription factors based on comparison with *S. cerevisiae.* First, we used sequence alignment to compare the DNA-binding domains. Most of them were highly conserved. But as predicted by bioinformatic analysis we were able to find around 10 transcription factors with unique sequences. Next, if the sequences of two DNA binding domains are similar that may predict they will bind to the same DNA sequence. If the DNA binding regions of the proteins are similar, we use the *S. cerevisiae* DNA binding motif (the short DNA sequence to which the transcription factor binds) and search for this motif in the promoter regions of the *C. albicans* genome by using meme-suite software. This approach identified approximately 20 transcription factors with a clear possibility of rewiring. Interestingly, there were a few transcription factors with more than one similar ortholog, for example *C. albicans* Hap4, Hap41 and Hap42 have only one ortholog in *S cerevisiae* (Hap4), even though *S. cerevisiae* has undergone a whole genome duplication, while *C. albicans* has not.

In total we selected a group of 30 non-zinc cluster transcription factors for our investigation. We will focus on activating members of the non-zinc cluster transcriptional factors in *C. albicans* to expand the potential for understanding regulatory circuits.

Objectives of this study

My primary objective was to investigate the function of currently poorly characterized transcription factors in the human pathogen *C. albicans* using transcription factor activation. Transcription factors are responsible for critical cellular functions and behaviors in organisms, and, in *C. albicans*, approximately 4% of genes code for transcription factors. We are interested in studying these transcription factors, most particularly those whose functions are unknown to us and can be rewired for various activities. We approached this by selecting the transcription factors that, based on comparison with *S. cerevisiae,* had no ortholog, or had orthologs that appeared to play different roles in the two yeasts. We selected 30 transcription factors for further investigation and activated them to permit phenotypic assessment of the consequences of this activation.

My secondary objective was to investigate in more depth transcriptional activations that created new insights into *C. albicans* biology. In one case I found that activation of the transcription factor Orf19.2752 (Adr1) conferred significant resistance against fluconazole during screening. Transcription factors play key roles in cellular regulation and are critical in the control of drug resistance in the fungal pathogen *Candida albicans*. However, in *Saccharomyces cerevisiae* Adr1 is carbon source-responsive zinc-finger transcription factor required for transcription of the glucose-repressed gene *ADH2* and of genes required for ethanol, glycerol, and fatty acid utilization. Motif scanning of promoter elements suggests that Adr1 may be rewired in the fungi and involved in the ergosterol synthesis pathway in *C. albicans.*

In a second case, I found during our screening that the regulatory network controlling methionine biosynthesis was distinct between *C. albicans* and *S. cerevisiae*. In *S. cerevisiae* this pathway is regulated by a collection of five transcription factors (Met4, Cbf1, Met28 and Met31/Met32), while in the filamentous fungi the pathway is controlled by a single Met4-like factor. But during our screening of activated Met4, Met28 and Met32 in *C. albicans* we observed they functioned differently from the orthologs in *S. cerevisiae*. We

have investigated this circuitry in the fungal pathogen *Candida albicans*, which is phylogenetically intermediate between the filamentous fungi and *Saccharomyces cerevisiae*. This work provides insight into the evolution of central metabolic circuits in the fungi.

Chapter 2: The screening of activated transcription factors and identification of rewiring.

2.1. Introduction

Candida albicans is one of the major causes of hospital-acquired infections in the USA, and the cost in treatment is nearly 2.6 billion USD annually (Stevens et al., 2014). This opportunistic fungal pathogen has evolved to survive inside its human host and switching among a variety of cellular forms appears important for allowing it to adapt to host cell responses and different environmental stresses (Gasch, 2007; Staib et al., 2001). These different phenotypic forms include the white (Slutsky et al., 1987) and opaque (Pande et al., 2013) yeast forms (Tao et al., 2014), and a filamentous form that appears critical for the fungus to invade host tissues (Calderone and Fonzi, 2001). The white yeast form, with small rounded budding yeast cells, is thought to be important for pathogen dissemination via the blood stream, and to thus have a role in virulence (Jacobsen et al., 2012), while the opaque yeast form, somewhat larger and more elongated, is the mating competent morphotype (Anderson et al., 1990). The invasive or filamentous form is induced in response to environmental cues such as nutrient limitation or the presence of serum. When an appropriate signal is transduced, yeast cells can initiate invasion by forming either hyphae or pseudohyphae (Csank et al., 1998; Felk et al., 2002). These filamentous cells can invade host tissues and ultimately enter the blood stream to cause life threatening systemic infections (Felk et al., 2002). Thus, in addition to the yeast form, the filamentous form is also critical for *C. albicans* virulence (Lo et al., 1997)

Transcription Factors (TFs) in *C. albicans* coordinate critical cellular functions including these morphological switches, biofilm formation (Nobile et al., 2009), drug resistance (Cowen et al., 2002), and the transition from a commensal to a pathogenic lifestyle (Liu, 2001). Approximately 4% of *C. albicans* genes code for transcription factors (Homann et al., 2009), making them the single largest family of proteins. We are trying to decipher the role of different transcription factors in *C. albicans.* We divided the transcription factors that are not yet significantly characterized into two categories. The first category is

transcription factors which do not have any orthologs in *S. cerevisiae*; we defined these TFs as Candida-specific. The second category is transcription factors which do have orthologs but are potentially rewired and thus may perform a distinct function in the pathogen. Keeping these characteristics in mind we selected 30 transcription factors for this study. As mentioned in Chapter 1, we selected these transcription factors based on the in-silico data that they are either rewired compared to *S. cerevisiae,* specific to the Candida clade or predicted to be involved in infection or drug resistance (Table 2.1). We attempted to activate each transcription factor by using the strong activation domain VP64 fused to the DNA binding domain of these 30 transcription factors of interest. Subsequently we investigated the consequences of this potential activation on a range of phenotypes including such things as cell morphology, growth under different pH conditions, growth at different temperatures, ability to metabolise different carbon sources, response to antifungal drugs, and response to metals ions.

2.2. Materials and methods

2.2.1. Bioinformatics analysis

Sequences of genes of all transcription factors were obtained from the Candida Genome Database (CGD- http://www.candidagenome.org/) and the Saccharomyces Genome Database (https://www.yeastgenome.org/). Gene orthogroup assignments for all predicted protein-coding genes across 23 Ascomycete fungal genomes were obtained from the Fungal Orthogroups Repository (Wapinski et al., 2007) maintained by the Broad Institute (broadinstitute.org/regev/orthogroups).

DNA sequence motifs were identified using the Web-based motif-detection algorithm MEME (http://meme.sdsc.edu/meme/intro.html; (Bailey et al., 2015). For more stringent motif identification, we used MAST hits with an E value of <50. An E value of 500 corresponds roughly to a p-value of 0.08 in our analysis, and an E value of 50 roughly corresponds to a p-value of 0.008. We also used AME (http://meme-suite.org/tools/meme), which identifies known motifs throughout the Candida promoter sequences. Protein

domains and linear motifs were detected from each individual TF protein sequence using INTERPROSCAN, PFAM and ELM motif definitions.

2.2.2. Strains and culture conditions:

For general growth and maintenance of the strains, the cells were cultured in fresh YPD medium (1 % w/v yeast extract, 2 % w/v Bacto peptone, 2 % w/v dextrose, 80 mg/L uridine with the addition of 2 % w/v agar for solid medium) at 30 °C. For methionine auxotrophy we used synthetic dextrose (SD) medium (0.67 % w/v yeast nitrogen base, 0.15 % w/v amino acid mix without methionine, 0.1 % w/v uridine, 2 % w/v dextrose and 2 % w/v agar for solid media).

2.2.3. Construction of the library:

For the activation module, the *ACT1* promoter and VP64 were amplified by PCR and homology was created by primer extension such that there is a restriction site for both HindIII and Mlu I between *ACT1* and VP64. This ligated CIPACT-VP64 plasmid was transformed into *E. coli* using the calcium chloride method (Cihlar and Calderone, 2009). After this, the C- terminus of the VP64 cassette was cloned into the multiple cloning site where we inserted different transcription factors DNA binding domains and were ligated using T4 DNA ligase. These ligated CIPACT-VP64-DBDs plasmids were transformed into *E. coli* using the calcium chloride method. Plasmids extracted from colonies that were determined to have the guide sequence successfully cloned in were then used to transform *C. albicans* using a lithium acetate transformation protocol (Vyas et al., 2018). pCIPACT1 was linearized by StuI-HF digest, and 1-2 μg of the linearized plasmid was used in the transformation. *C. albicans* transformants were selected on SD URA-plates.

2.2.4. Screening of the library:

To test sensitivities of strains to different agents and to temperature, overnight cultures grown in YPD at 30°C were serially diluted in 10 fold stages at room temperature to a

dilution concentration of 10^6 and 10^3 cells per ml and 5 µl of each dilution was spotted on YPD agar containing one of the following; hydroxyurea (HU, 10 mM), Congo red (150 µg/ml), NaCl (1 M), arsenic (1M), cesium chloride (1M), and 2% of different sugars including sucrose, galactose, xylose, fructose and lactose. Plates were then incubated at 30°C for 24-48 hours. All strains were tested for temperature sensitivity on YPD agar at 15°C, 20°C, 30°C and 45°C for a period of 48 hours. For pH sensitivity we used YPD agar at pH 4.5, 6.5-7 pH and 8.5-9 pH for a period of 48 hours.

2.2.4. Phenotype Switching:

White and opaque cells were all selected from single colonies on YNB-GlcNAc medium after 1-5 days at room temperature depending on strains. For this study we used a *C. albicans* CaI4 a/a strain. Cells then were suspended in water, the cell concentration was adjusted, and the suspensions plated on agar medium containing 5 µg/ml phloxine B and different carbon sources. Plates were incubated at room temperature. Data were collected and plates were scanned on the 7th day and the frequency of sectored colonies calculated by standard statistical methods.

2.2.5. Microscopy:

For white and opaque cell morphologies, overnight cultures grown in YCB-GlcNAc or other carbon source medium at RT were subjected to phase differential interference contrast (DIC) microscopy. Cells were examined by DIC at 63X and fluorescent microscopy at 100X magnification using a Leica DM 6000 microscope mounted with a TIR camera.

2.2.5. Adhesion:

The WT strain and the mutant strains were cultured in YPD liquid at 30° C overnight with 220 rpm shaking conditions. The cells were washed with PBS and resuspended in Spider medium. The concentration of cells was adjusted at OD_{600} = 0.5. Two-hundred microliters

of the inoculum were transferred to a well of a non-surface-treated 96-well plate for initial incubation at 37° C for 3 h in static conditions to allow cell adherence. Next, the non-adherent cells were washed off with PBS. The metabolic activity of the biofilms was verified by the XTT-colorimetric assay described (Jin et al., 2003). Each well received 158 µL of PBS, 40 µL of 1 mg mL -1 XTT solution (Sigma, St Louis, USA), and 2 µL of fresh prepared 0.4 mM of menadione solution (Sigma, St Louis, USA). The plates were incubated at 37° C for 3 h with shaking in the dark. Following, 100 µL of the supernatant was transferred to a new well and the colorimetric change was read at 492 nm using a microtiter plate reader. The biomass accumulation was examined by the crystal violet assay (Regan et al., 2017). Briefly, the adhered cells were air dried for 45 min. Subsequently, the wells were stained with 0.4% aqueous crystal violet solution for 45 min at room temperature (RT). Next, the biofilms were washed twice with sterile water and then de-stained with 95% ethanol for 45 min at (RT). The de-stained solution was diluted at 1:10 in 95% ethanol and the absorbance measured at OD 595. The procedures were carried out in triplicate on three different occasions.

2.3 Results

We constructed a CIPACT-VPR plasmid with VP64 and a specific multiple cloning site (MCS) downstream of the *ACT1* promoter of the CIPACT plasmid. The MCS site allows the fusion of VP64 to the DNA binding domain of different TFs with linker amino acids like Glycine. To test this strategy, we choose 3 well studied transcriptional factors - the bZIP TF Gcn4 (null mutant gives 3 amino triazole sensitivity), a TEA/ATTS (Homo-domain) TF Tec1 (null mutant blocks hyphal development), and a Leucine bZIP TF Cap1 (involved in fluconazole resistance). We successfully activated Gcn4 to trigger 3AT resistance presumably due to up-regulation of the Gcn4 target *HIS3* (Fig 2.1 A), we activated Cap1 to trigger resistance to fluconazole (Fig 2.1 B) and we activated Tec1 to generate filamentation (Fig 2.1 C).

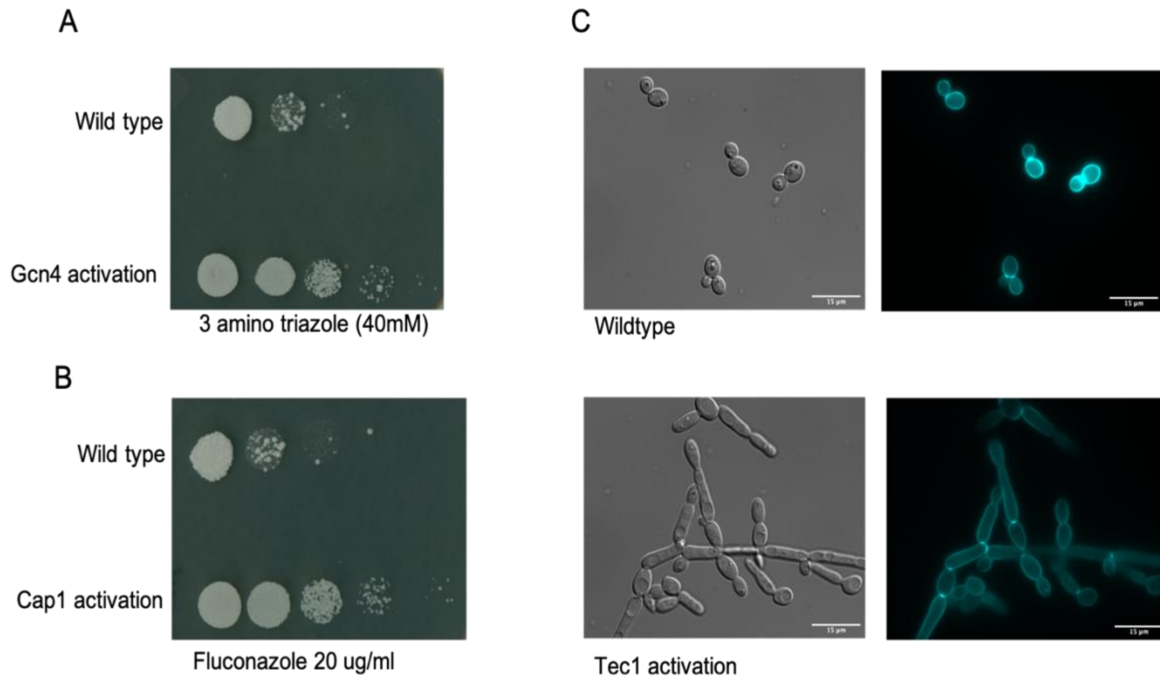


Fig. 2.1.A.C albicans strain with activated Gcn4 growth in 3 amino triazole (3AT) compared with wild type. B. Strain with activated Cap1 growth in SC media with fluconazole. C. Microscopic images of strain with activated Tec1.

Differential growth, metabolism, and morphology of strains upon activating unknown TFs

After activating the unknown transcription factor, we start our screening for carbon utilization. We checked for various sugar utilization including glucose, xylose, fructose, sucrose, and sorbitol as a sole source of carbon. We did not find any variation in the growth of any strain containing a potentially activated transcription factor.

Using the Bioinformatic tool MEME.suite we predicted 6 transcription factors may play a role in the induction of hyphal formation in *C. albicans* (Table 2.1). In microscopic analysis we found that the strains containing activated *C2_05640W_A,* activated *C4_07150W_A,* activated *NRG2,* and activated *GLN3* showed a mixture of both yeast and

filamentous cells (Fig 2.2. B). As predicted by the *in-silico* study Table 2.1, they might be involved in filament formation in *C. albicans* whereas the strains with activated transcription factors *C4_07150W_A* or *CR_05880W_A* do not show any changes. Subsequently, we screened all the strains for their morphology in Spider media and during serum stimulation in both liquid and solid media. During the solid media assay all the strains transitioned to a filamentous form just like the wild type. However, in the liquid media we could see differences in the timing of the filamentous transition. We found that the strains containing activated *C2_05640W_A, C4_07150W_A, NRG2,* and *GLN3* showed a faster transition compared to the wildtype. We also observed other strains containing activated transcription factors like *ADR1*, *MET32*, *UGA33*, *C1_11690W_A* and *PHO4* to switch to the hyphal morphology faster than the wild type (Fig 2.2.A).

Another morphological change in *C. albicans* is the transition from white cells to opaque cells. This can be identified by growing the *C. albicans* in the media containing Phloxine B (Costa et al., 2021). One transcription factor *C1_13440C_A* showed a faster switch from white to opaque cells during the screening (Fig 2.2.C).

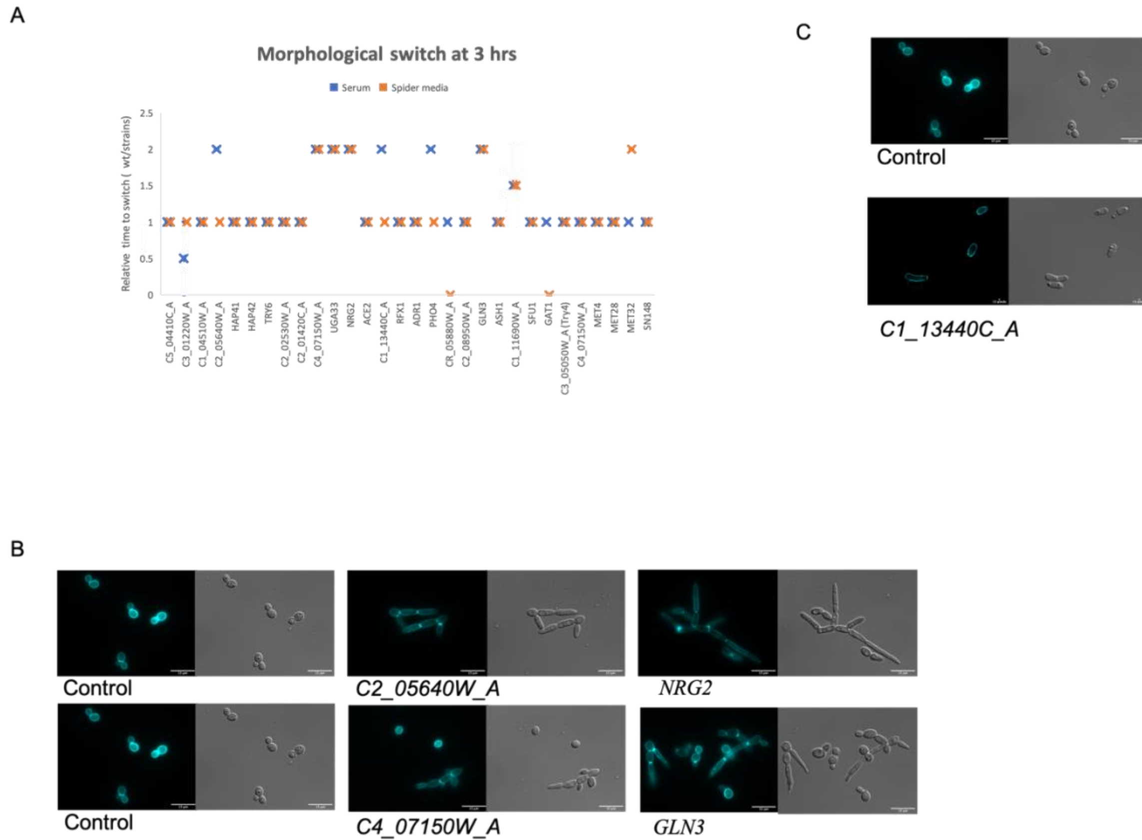


Fig. 2.2.A. Morphological switch from yeast to hyphae was monitored every hour in two different media, Spider and serum. We found that 9 strains showing faster "switch" than the wild type, although only 5 strains showed a faster switch in both media. B. Activation of 4 transcription factors, C2_05640W_a, Nrg2, C4_07150W_a and Gln2 enhanced filamentous morphology predicting they might influence hyphae formation directly or indirectly, C. Activation of C1_13440C_a triggered opaque cells in the presence of opaque specific media with phloxine B in the media, This result indicates that activation of this transcription factor facilitates the faster switching from white to opaque.

2.3.2 Screening of the constructed library for response to cell wall stress and heavy metals.

The cell wall is the first line of defense for most of the unicellular organisms including bacteria and fungi. The cell wall of *C. albicans* contains 10% chitin that is important for the rigid cell wall to protect against harsh external conditions (Ruiz-Herrera et al., 2006).

Congo red and the antifungal drug Caspofungin cause cell wall stress by targeting glucan and chitin synthase (Ghannoum and Rice, 1999a; Roncero and Duran, 1985). Therefore, we checked the growth of our strain with activated transcription factor for growth in the presence of Congo-red. We did not find any changes in the any of the strains. This suggests that none of the transcription factors are involved in cell wall integrity.

We also screened for tolerance to both arsenic and cesium. We found the *MET4* activated strain showed impressive tolerance to both heavy metals. An increased level of methionine biosynthesis has been observed to give tolerance to heavy metals as yeast uptake the sulphate ion from the environment (Barbey et al., 2005b; Yen et al., 2005). Heavy metals along with sulphate ions creates toxic environment for the cells. Excess amount of sulphate ions is reactive and binds to positively charged proteins in the cell non-specifically and therefore has an important role in metal toxicity. The tolerance to the heavy metal might be due to the upregulation of methionine biosynthesis (sulphur utilization). Unlike strain *MET4,* strains *MET28* and *MET32* did not show much resistance to heavy metals (Fig 2.3.A), although Met32 activation strains show slight growth in the presence of arsenic. Met28 and Met32 are known transcription factors which are involved in methionine biosynthesis in *S. cerevisiae* (Petti et al., 2012b; Shrivastava et al., 2021).

2.3.3 Screening of the constructed library for temperature and pH response.

It has been established that a pH around neutrality favours hyphal development of *C. albican*s *in vitro*, while a low pH (pH < 6.5) blocks hyphal formation and stimulates growth of the yeast form (Buffo et al., 1984). Growth of cells in the yeast form is promoted by a growth temperature below 35˚C, a pH of less than 6.5, and glucose as a carbon source whereas pH7 and temperature above 35 ˚C promotes hyphal formation. As well, a temperature of 25˚C is optimal for yeast to opaque switching of the *C. albicans*; higher temperatures maintain the cells in the white state. Therefore, we screened the 30 activated transcription factor containing stains for response to pH and temperature.

We screened the activated transcription factor library for changes in temperature tolerance and found 1 transcription factor which might play role in this process. Transcription factor *C5_04410C_A* shows promising growth at low temperature (Fig 2.3.B), although it did not show any modulation in behaviour at the higher temperatures. Most of the strains behaved similarly to WT at the high temperature, although strains with activated *ADR1, MET32, UGA33*, *C1_11690W_A* and *HAP42* grew slightly better growth in higher temperature compared to wildtype but that might be due to fact that they grow faster than the wildtype and rest of the strains.

We found that activation of transcription factor *ASH1, HAP41, MET28, MET32* significantly, and *C3_01220W_A* to a minor extent gave a tolerance to low pH (Fig 2.2.C), while *C1_04510W_A, PHO4 and TRY4* activation allowed better growth in alkaline pH (Fig 2.3.C).

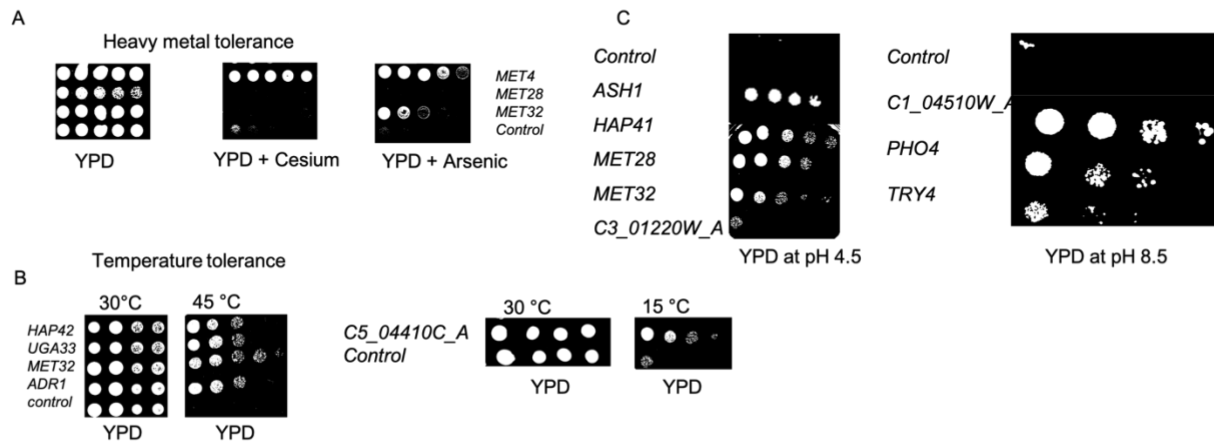


Fig. 2.3.A. Spot assay results showing tolerance to heavy metals like cesium and arsenic through activation of MET4. B. Spot assay results of temperature tolerance in different strains. HAP42, UGA33, MET32 and ADR1 gene activation generated tolerance to the high temperature whereas C5_04410C_A gene activation results in low temperature tolerance.

C. Spot assay of pH tolerance in various strains containing activated transcription factors. ASH1, HAP41, Met28, Met32 and C3_01220W_A gene activation created tolerance to acidic pH whereas C1_04519W_A, PHO4 and TRY4 gene activation allowed alkaline tolerance.

Note: all the above pictures were taken after 12-16 hours, later at 48 hours growth of the control strain was observed.

2.3.5 Screening of the library for hydroxyurea response.

Genotoxic (Genome-Toxic) stress compromises the genome integrity of an organism. There are several mutagens that can induce genotoxic stress, such as the DNA replication stalling chemical hydroxyurea (HU) and the thymine-dimer-causing ultra-violet (UV) radiation. We screened the library of activated transcription factor strains with the DNA-damaging agent hydroxyurea. Among our selected 30 transcription factors, C2_08950W_a transcription factor do not have any ortholog in *S. cerevisiae* but do have a very conserved ortholog in *A. nidulans* which is predicted to be involved in DNA repair (Arnaud et al., 2012; Consortium, 2012). Interestingly, the strain with activation of TF encoded by C2_08950W_a showed resistance to hydroxyurea. This result shows that it might be involved in DNA repair just like it is predicted for it ortholog in *A. nidulans*. As well strains with activation of transcription factors Rfx1, Hap42 and Hap41 also showed resistance to hydroxyurea (Fig 2.4.A).

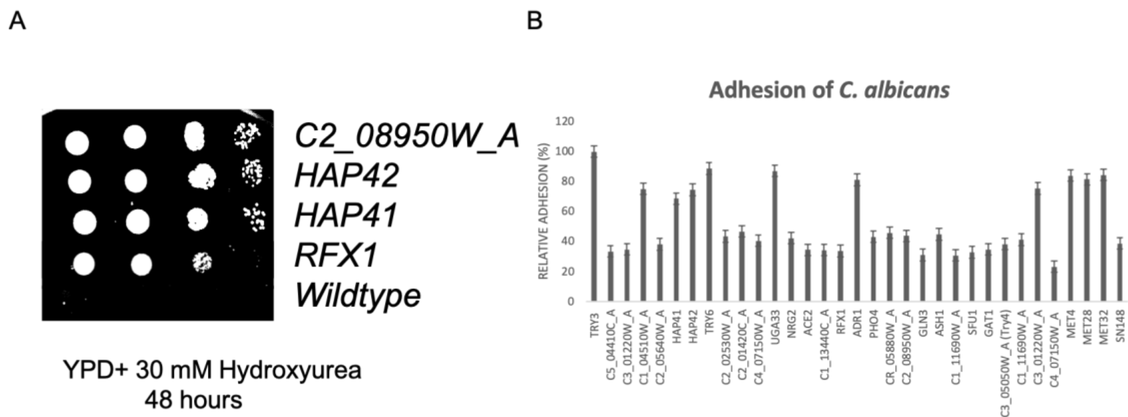


Fig. 2.4.A. The 4 strains with activated C2_08950W_a, Hap42, Hap41 and Rfx1 respectively, showed growth in the presence of DNA damaging agent hydroxyurea. B. The impact of activated transcription factors in adhesion of C. albicans on the surface.

2.3.6 Screening for adhesion.

Adherence of a pathogen to a device surface is a critical early step in formation of clinically problematic biofilms. For device-associated biofilms, definition of the mechanisms that regulate cell-substrate adherence provides insight into how these biofilms form. That understanding may in turn suggest therapeutic or preventive strategies. We found 14 strains with activated transcription factors showing better adhesion than the wildtype. The strains with activation of Hap42, Hap41, Try6, Try3, Adr1, Uga33, C1_04510W_a, C3_01220W_a, Met28, Met32 and Met4 show more than 50% better relative adherence than the wildtype in a standard adhesion assay. The strain with Try3 activation showed the highest level of adherence whereas the strain with Rfx1 activation showed the lowest (Fig 2.3.B). Strains with Try6 or Uga33 activated showed similar amounts of adherence at levels close to that of the strain with Try3 activation. Interestingly, activation of Nrg2 which generated enhanced hyphal formation did not up-regulate adhesion. The activation of Ace2 was also predicted to be involved in adhesion but did not show any enhancement.

2.3. Discussion

Transcription factors play very important roles in all organisms. In *C. albicans* they are involved in process as diverse as infection, morphogenesis, biofilm formation and mating. As discussed in chapter 1, the "rewiring" of transcription factors might play very important roles in host-pathogen interactions, infection or in multi-drug resistance. In this study we have selected 30 transcription factors that are predicted be Candida specific, rewired or to have critical roles in drug resistance.

A critical element of my research was to develop an artificially activated non-zinc cluster TF library. Herpesvirus VP16 (or Vmw65) activates expression of immediate-early genes in virally infected cells (Everett et al., 1998). VP16 has two domains, one for interaction with cellular proteins, which in turn bind to DNA sequences upstream of immediate-early genes, and a second domain that is required for transcriptional activation. VP16 was successfully fused to the N terminal of Gal4 in *Saccharomyces cerevisiae* (Sadowski et al., 1988). This fusion results in the hyper activation of Gal4. We attempted to use the same strategy to activate a set of non-Zinc-finger TFs in C. *albicans.* Our strategy involved the fusion of the VP64 transactivation domain (a tetrameric version of VP16's trans-activating

domain found to be even more active in transcriptional activation (Beerli et al., 1998a)) to the DNA binding domains of different transcription factors. Previously, VP64 has been used to activate transcription factors in plants (Ordiz et al., 2002b) and animals (Maeder et al., 2013b). We used the same strategy and were able to activate known transcription factors Gcn4, Cap1 and Tec1 successfully, so next we attempted to activate our group of selected transcription factors.

Preliminary bioinformatic analysis predicted 12 uncharacterized transcription factors to be rewired (Table 2.1; page 37). Our screening for various phenotypes leads us to believe that at least 8 *C. albicans* transcription factors, specifically Met4, Met28, Met32, Adr1, Try3, Try4, Hap41 and Hap42 may be rewired relative to their orthologs in the model Ascomycete *S. cerevisiae*. Met4, Met28 and Met32 behaved differently in different media although their function in *S. cerevisiae* is methionine biosynthesis (Lee et al., 2010). Met4 activation created heavy metal tolerance with strong adhesion properties, Met28 activation generated tolerance to acidic media and low temperature compared to wildtype. This differential behaviour of these transcription factors intrigued us and therefore we investigated in depth this possible rewiring, which we have discussed in chapter 4.

We screened these transcription factors for several phenotypes like pH, temperature, morphology, carbon source utilization, cell wall stress, adhesion, hydroxyurea, and salt tolerance. We also checked cellular morphology in various media like serum, Phloxine B and Spider media. Serum and Spider media is responsible for the hyphal formation whereas the Phloxine B is a pink color dye which distinguishes between white and opaque cells (Xie et al., 2013).

Transcription factor C1_13440C_*a* activation created pink or opaque colonies during the screening. This transcription factor is known as Ofi1and is known to be overexpressed during white-opaque switching (Du et al., 2015). However, the deletion of this transcription factor did not affect switching to a great extent. This previous research has raised many questions. Wor1 is the known master regulator of the white opaque switching process (Xie et al., 2013); how does the Ofi1 transcription factor function in collaborating with Wor1?

Bioinformatic data reveals that *OFI1* has the Wor1 binding motif in its upstream region thus likely acts downstream of Wor1. Combinational deletions of Wor1, Ofi1 and other transcription factors that might be involved in the mating might help us to understand the exact function of this transcription factor. We also need to find the targets of this transcription factor. This transcription factor activation leads to the faster switch from white to opaque cand can be used as a tool in future for studies.

Hap41 and Hap42 are the transcription factor that are CCAAT-binding factor just like Hap4 which is involved in the iron homeostasis. Although the most closely related of *S. cerevisiae* ScHap4 is CaHap4 which is already known for controlling iron homeostasis. These two-transcription factor Hap41 and Hap42 are similar to Hap4 but have not been characterized yet. Therefore, we tried to look toward the other fungi of this family like *Aspergillus nidulans*, *N. crassa* or *S. pombe* but we could not find any apparent orthologs. Hap42 is essential (Segal et al., 2018).

In the *S. cerevisiae* the CCAAT-binding factor is composed of four subunits, termed Hap2p, Hap3p, Hap4p, and Hap5p, which are involved in the transcriptional activation of numerous genes that encode proteins involved in respiration (Mantovani, 1999), as well as other genes (DeRisi et al., 1997; Forsburg and Guarente, 1989a; Lascaris et al., 2003). In this complex Hap4 is the DNA-binding protein (Forsburg and Guarente, 1989b). Thus, the interaction of Hap4p with Hap2/3/5p modulates the ability of the CCAAT-binding factor to stimulate transcription in yeast. In *C. albicans* Hap2, Hap3, Hap4 and Hap5 are already characterized, and they are highly conserved in sequence as well as function. But intriguingly, *C albicans* also has two extra Hap4-like proteins with very conserved sequences, however preliminary research indicates that they might not be involved in this complex (Johnson et al., 2005).

Thus, these two Candida specific Hap42 and Hap41 are intriguing and indicating that there is more to study. Our screening results shows that their activation results in better adhesion, increases high pH and temperature tolerance as well as generating resistance to the DNA damaging agent hydroxyurea. Although we could not predict the function by

bioinformatics or by screening, hydroxyurea tolerance is directing us toward a possible DNA-repair mechanism role.

Microbial adherence is one of the most important determinants of pathogenesis, yet very few adhesins have been identified from fungal pathogens. Hyphae-associated factors, like the Als family of adhesins promote anchoring to a substratum, cell-cell adhesion, and are important for biofilm establishment (Nobile et al., 2008; Grubb et al., 2009; Finkel et al., 2012). Biofilm-associated hyphal growth is regulated by seven key transcription factors specifically Gal4, Rfx2, Flo8, Bcr1, Efg1, Ndt80, Brg1, Tec1 and Rob1, some of which are also regulators of glycolysis and carbon metabolism and hyphal formation (Bonhomme et al., 2011; Cleary et al., 2012; Fox et al., 2015). In our case, we also found the enhanced adhesion in case of the activated transcription factor that was either predicted to be involved in metabolism or hyphal formation, except Try4 and Try3. These two transcription factors are Candida-specific and already known to be have some role in yeast adhesion in a large-scale study. It would be interesting to know the exact role these two-transcription factors in yeast adhesion.

Another interesting transcription factor was Adr1. Although apart from slightly enhanced growth relative to the wildtype it did not show any other phenotype in our preliminary screening. The predicted function of Adr1, based on transcriptional profiling of the activated TF, was lipid biosynthesis, which can be assessed by resistance of azole drugs which target the lipid biosynthesis pathway. Therefore, as an extension of our screening we tested all our strains for response to various drugs and observed Adr1 activation led to resistance to both azoles tested. We studied Adr1 rewiring extensively in the next chapter.

Table 2.1. List of the activated transcription factor with predicted function, Orthologs and possibility for rewiring.

No.	Transcription factor	Predicted function	Ortholog in *S. cerevisiae*	Rewiring
1	**Try3**	Adhesion	-	-
2	C5_04410c_a	Diauxic shift	Azf1	Y
3	C3_01220w_a	-	-	-

4	C1_04510w_a	-	-	-
5	C2_05640w_a	-	Bas1	Y
6	Hap41	Week acid response	-	-
7	Hap42	-	-	-
8	**Try6**	Yeast form adherence	-	-
9	**C2_02530w_a**	Mating	Gata2	Y
10	**C2_01420c_a**	-	-	-
11	C4_07150w_a	Filaments formation	Atc1?	Y
12	Uga33	Adhesion	-	-
13	Nrg2	Filaments formation	-	-
14	Ace2	Adhesion and Morphogenesis	Ace2	N
15	C1_13440c_a	White opaque Switching	Crz1	Y
16	Rfx1	Paralog involved in DNA damage repair	Rfx1	Y
17	Adr1	Lipid biosynthesis	Adr1	Y
18	Pho4	Resistance to stress	Pho4	Y
19	Cr_05880w_a	Morphogenesis	-	-
20	C2_08950w_a	DNA repair	-	-
21	Gln3	Morphogenesis	Gln3	Y
22	Ash1	Cell division	Ash1	-
23	C1_11690w_a	Oral infection	Gln3	-
24	Sfu1	GATA type of transcription factor But SFU1 is repressor where as GAT1 is the positive regulator of Nitrogen utilization.		
25	Gat1			
26	C3_05050w_a (try4)	ADHESION	-	-
27	C1_11690w_a	infections in epithelial cell		
28	C3_01220w_a	-	-	-
29	C4_07150w_a	Filaments formation -	-	-
30	Met4	Methionine Biosynthesis	Met4	Y
31	Met28	Methionine Biosynthesis	Met4	Y
32	Met32	Methionine Biosynthesis	Met31/met32	Y

Chapter 3: The Adr1 transcription factor directs regulation of the ergosterol pathway and azole resistance in C. *albicans*

3.1 Introduction

Candida albicans is an opportunistic fungal pathogen that is responsible for a variety of fungal infections in humans. In healthy people, this yeast resides as a commensal in the gastrointestinal tract, but it is capable of causing mucosal, cutaneous, and systemic infections in immunocompromised individuals (Calderone and Fonzi, 2001). The prevalence of resistance to antifungal agents has significantly increased in the past few decades, and this resistance has important implications for mortality, morbidity and health care in the community (Arendrup and Patterson, 2017). The development of new antifungal drugs is challenging, as fungi are eukaryotic organisms that share many basic cellular processes with humans, and this evolutionary relatedness makes the identification of specific targets difficult and increases the likelihood of undesired secondary effects. Existing drugs tend to target processes that are divergent between fungi and the human host.

Azoles are generally effective for the management of *C. albicans* infections, but due in part to the fungistatic nature of the drugs, long-term treatment often results in the emergence of azole resistance, ultimately resulting in therapeutic failure (Lopez-Ribot et al., 1998; Morschhauser, 2002; Perea et al., 2001). These azole antifungals bind and inhibit the activity of the enzyme lanosterol 14-alpha-demethylase encoded by *ERG11* (Kelly et al., 1993). Apart from azoles, allylamines (which target Erg1), polyenes (which target ergosterol itself), morpholines (which target Erg2) and statins (which target HMG1/2) also target elements of the sterol pathway (Bhattacharya et al., 2018). Because so many drugs target the *C. albicans* sterol pathway, genetic

changes that perturb the pathway can lead to multidrug resistance (Prasad et al., 1995b; Prasad and Singh, 2013).

A promising approach for drug development is to identify synergistic targets that can enhance the antifungal effect of currently available drugs (Lu et al., 2021). Transcription factors (TFs) play a key role in determining how cells function and respond to different environments, and approximately 4% of *C. albicans* genes code for transcription factors (Homann et al., 2009), making them the single largest family of proteins in the pathogen. TFs in *C. albicans* coordinate critical cellular functions including biofilm formation (Nobile et al., 2009), drug resistance (Cowen et al., 2002), and the transition from a commensal to a pathogenic lifestyle (Liu, 2001).

Most transcription factors are conserved, in that they fall into a limited number of groups of structurally similar proteins, such as the zinc finger, the basic helix loop helix, and the leucine zipper classes. However, evolutionary changes in transcription networks are an important source of diversity across species, driven primarily not by major changes in the structures of the factors themselves, but in the connections between the transcription factors and their regulated genes. There are many incidences where researchers have identified structurally equivalent transcription factors regulating different genetic circuits in different organisms (Erkenbrack and Davidson, 2015; Lavoie et al., 2010; Oliveira-Pacheco et al., 2018; Tebung et al., 2016). This phenomenon has been called "rewiring" of transcription factors. Studies suggest that this rewiring happens at a relatively constant rate, and for two species that are diverged by 100 million years, only a fraction of transcription factor/target gene combinations will likely have remained conserved (Carvunis et al., 2015; Johnson, 2017b). *C. albicans* belongs to the same family as *Saccharomyces cerevisiae,*

but the two fungi are suggested to have diverged as long ago as 300 million years, allowing for considerable rewiring. While transcription factor studies have tended to focus on similarities between these two species, it has been estimated that only 16% of the regulator-target gene connections are preserved between the *C. albicans* and *S. cerevisiae* (Habib et al., 2012).

In our study we have activated a group of transcription factors of *C. albicans* for which there was limited information and which had a potential of being rewired. Among our tested set we found that Orf19.2752 gives resistance to several cell-membrane targeting drugs. This resistance arises because Orf19.2752 is a central regulator of the ergosterol pathway in *C. albicans*. Further analysis shows that this transcription factor is the ortholog of *S. cerevisiae* Adr1, and that the two proteins play distinct cellular roles in the two species.

3.2 Material and Methods

3.2.1 Bioinformatics analysis

Sequences of genes Ca*ADR1*and *ScADR1* were obtained from the Candida Genome Database (CGD- http://www.candidagenome.org/) and the Saccharomyces Genome Database (https://www.yeastgenome.org/). Gene orthogroup assignments for all predicted protein-coding genes across 23 ascomycete fungal genomes were obtained from the Fungal Orthogroups Repository (Wapinski et al., 2007) maintained by the Broad Institute (broadinstitute.org/regev/orthogroups).

DNA sequence motifs were identified using the Web-based motif-detection algorithm MEME (http://meme.sdsc.edu/meme/intro.html; (Bailey et al., 2015). For more stringent motif identification, we used MAST hits with an E value of <50. An E value of 500 corresponds roughly to a p-value of 0.08 in our analysis, and an E value of 50 roughly corresponds to a p-value of 0.008. We also used AME (http://meme-suite.org/tools/meme), which identifies known motifs throughout the Candida upstream sequences.
Protein domains and linear motifs were detected from each individual TF protein sequence using INTERPROSCAN, PFAM and ELM motif definitions.

3.2.2 Strains and culture conditions:

For general growth and maintenance of the strains, the cells were cultured in fresh YPD medium (1 % w/v yeast extract, 2 % w/v Bacto peptone, 2 % w/v dextrose, 80 mg/L uridine with the addition of 2 % w/v agar for solid medium) at 30 °C. For drug assays we used synthetic dextrose (SD) medium (0.67 % w/v yeast nitrogen base, 0.15 % w/v amino acid

mix, 0.1 % w/v uridine, 2 % w/v dextrose and 2 % w/v agar for solid media) along with the various concentrations of drugs in liquid and solid media.

3.2.3 Gene knockouts using CRISPR

All *C. albicans* mutants were constructed in the wild type strain CAI4. The protocol used for the CRISPR-mediated knockout of *ADR1*, *CTF1* and *UPC2* was adapted from (Vyas et al., 2015); we used *URA3* replacements in our study. CRIPSR-mediated knockouts used the lithium acetate method of transformation with the modification of growing transformants overnight in liquid YPD at room temperature after removing the lithium acetate-PEG. *C. albicans* transformants were selected on SD-URA plates.

Activation of transcription factors

For the activation module, the *ACT1* promoter and VP64 were amplified by PCR and homology was created by primer extension such that there is an Mlu I site between *ACT1* and VP64. After this, the C- terminus of the VP64 cassette was cloned into the multiple cloning site. pCIPACT and the ligated Act1-VP64-MCS were digested with Hind III and were ligated using T4 DNA ligase. This ligated CIPACT-VP64 plasmid was transformed into *E. coli* using the calcium chloride method. High-throughput equipment at the Concordia Genome Foundry, the Biomek FXP Automated Workstation otherwise known as Biomek FXP liquid handler was used to create and screen this library.

Plasmids extracted from colonies that were determined to have the guide sequence successfully cloned in were then used to transform *C. albicans* using a lithium acetate

transformation protocol. pCIPACT1 was linearized by StuI-HF digest, and 1-2 μg of the linearized plasmid was used in the transformation. *C. albicans* transformants were selected on SD URA-plates.

3.2.4 RNA seq analysis

The CaAdr1 and SC5143 strains were grown in YPD overnight at 30 °C, diluted to OD600 of 0.1 in YPD at 30 °C, and then grown to an OD_{600} of approximately 1.0 on a 220-rpm shaker. Total RNA was extracted using the Qiagen RNeasy minikit protocol, and RNA quality and quantity were determined using an Agilent bioanalyzer. Paired-end sequencing (150bp) of extracted RNA samples was carried out at the Quebec Genome Innovation Center located at McGill University using an Illumina miSEQ sequencing platform. Raw reads were pre-processed with the sequence-grooming tool cutadapt version 0.4.1 (Martin, 2011) with the following quality trimming and filtering parameters (`--phred33 --length 36 -q 5 --stringency 1 -e 0.1`). Each set of paired-end reads was mapped against the *C. albicans* SC5314 haplotype A, version A22 downloaded from the Candida Genome Database (CGD) (http://www.candidagenome.org/) using HISAT2 version 2.0.4. SAM tools was then used to sort and convert SAM files. The read alignments and *C. albicans* SC5314 genome annotation were provided as input into StringTie v1.3.3 (Pertea et al., 2015), which returned gene abundances for each sample. Raw and processed data have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002).

3.2.5 Sterol estimation

The cells were grown overnight (16 hrs) at 30°C and harvested by centrifugation. Non-treated cells were maintained separately and considered as control. The cell pellets were washed with sterile distilled water twice. We followed the same method which has been described in (Breivik and Owades, 1957) with slight modifications. Cell pellets were resuspended in 2.5 ml methanol, 1.5 ml potassium hydroxide (60% [wt/vol]), and 1 ml methanol-dissolved pyrogallol (0.5% [wt/vol]) and heated at 90°C for 2 h. The cell extracts were cooled and then sterols were extracted with two rounds of treatment with 5ml of hexane. The extracted sterols indicated a four-peak spectral absorption pattern produced by ergosterol and 24(28)-dehydroergosterol [24 (28)-DHE] spectrophotometrically (DU530 life science UV spectrophotometer). Both ergosterol and 24(28)-DHE absorb at 281.5 nm, whereas only 24(28)-DHE absorbs at 230 nm. The ergosterol content is determined by subtracting the amount of 24(28)-DHE (calculated from the A230) from the total ergosterol plus- 24(28)-DHE content (calculated from the A281.5). Ergosterol content was calculated as a percentage of the wet weight of the cells with the following equations: % ergosterol + % 24(28)-DHE [(A281.5/290) * F]/pellet weight, % 24(28)-DHE _ [(A230/518) _ F]/pellet weight, and % ergosterol = [% ergosterol + % 24(28)-DHE] – [% 24(28)-DHE], where F is the factor for dilution in petroleum ether and 290 and 518 are the E values (in percent per centimeter) determined for crystalline ergosterol and 24(28)-DHE, respectively.

3.3 Results

3.3.1 Fusion of different transcription factors to the strong activation domain VP64.

In *Saccharomyces cerevisiae* the fusion of VP16 to the N terminus of Gal4 resulted in the hyper-activation of Gal4 (Daniel et al., 1998). We have used a similar strategy in *C. albicans,* fusing a tetrameric version of the VP16 trans-activating domain (VP64), found to be even more potent in transcriptional activation (Beerli et al., 1998b), to the DNA binding domains of different *C. albicans* transcription factors. Previously, VP64 has been used to successfully activate transcription factors in both plants (Ordiz et al., 2002a) and animals (Maeder et al., 2013a). We constructed plasmid CIPACT-VPR containing the VP64 module and a multiple cloning site (MCS) downstream of the *ACT1* promoter of the CIPACT plasmid (Fig S1A).

To test this strategy, we choose 3 well-studied transcriptional factors - the bZIP TF Gcn4 (null mutant gives 3 amino-triazole sensitivity), a TEA/ATTS (Homeo-domain) TF Tec1 (null mutant blocks hyphal development), and a leucine bZIP TF Cap1 (involved in fluconazole resistance). The Gcn4 construct generated resistance to 3 amino-triazole consistent with up-regulation of the Gcn4 target *HIS3* (Fig 3.1.A). The Tec1 construct triggered filamentation under yeast morphology growth conditions (Fig 3.1.B), and the Cap1 activation construct enhanced resistance to fluconazole (Fig 3.1.C).

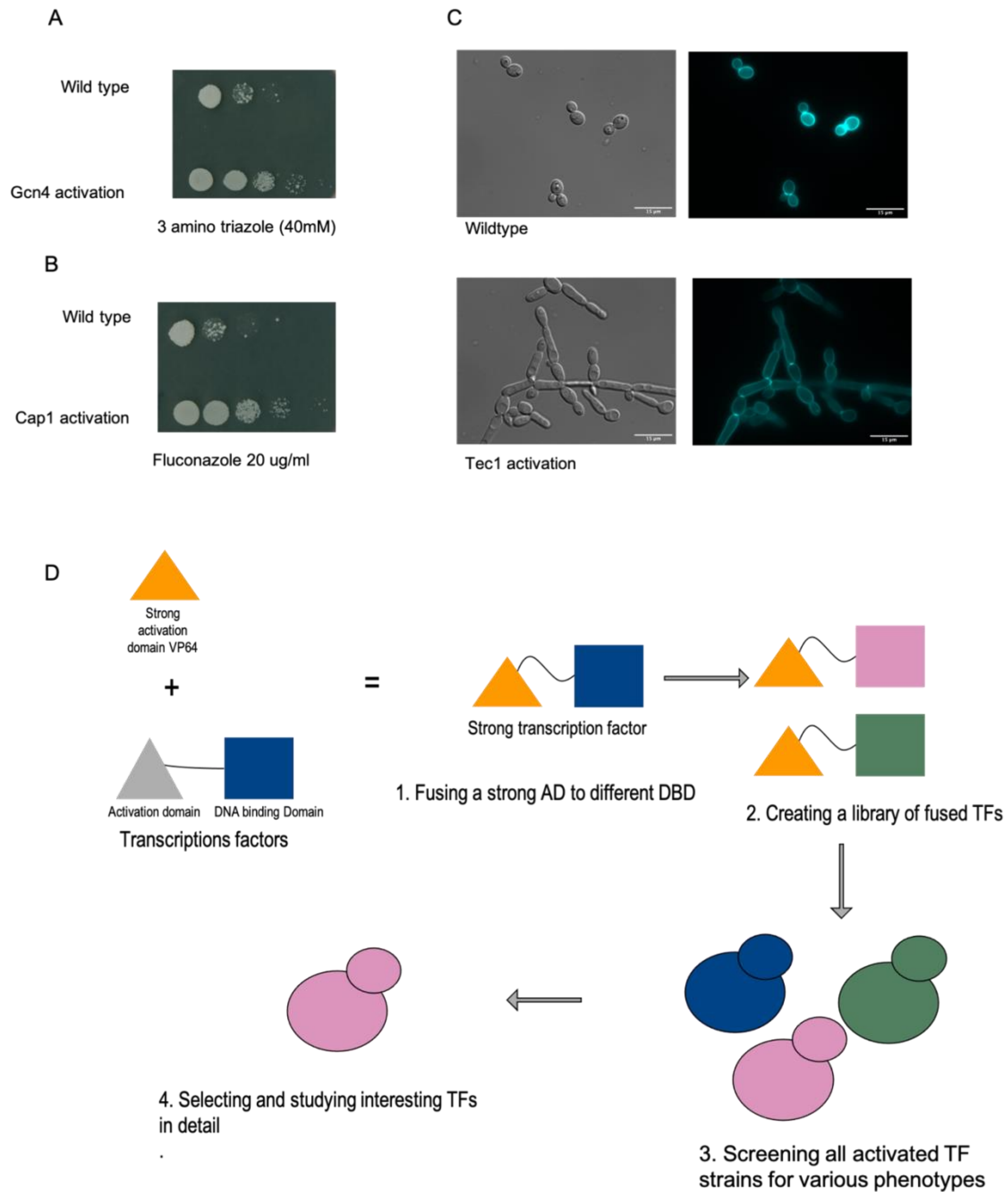


Fig. 3.1.A. VP64-Gcn4 chimeric transcription factor generates resistance to 3-amino-triazole (3AT). *To construct the Gcn4- Vp64 fusion, we PCR amplified the Gcn4 DNA-binding domain and ligated it in the CIP-ACT-CYC plasmid in frame with VP64 at the N terminus. After transforming this plasmid into C. albicans we observed resistance to 3AT*

consistent with the VP64 module activating the transcription factor. ***B. VP64-Tec1 TF triggers hyphal elongation in YPD media.*** *Tec1 is a transcription factor implicated in the morphological switch from the C. albicans yeast form to the hyphal form. A construct containing the fusion of the N terminus of Tec1 to the VP64 module triggers elongated cellular growth.* ***C. VP64-Cap1 allows growth in SC media containing fluconazole.*** *Cap1 is a poorly characterized transcription factor in C. albicans that gives resistance to azoles through Mdr1; activation of Cap1 upregulates MDR1 expression. Fusion of the VP64 module to Cap1 increased cellular tolerance to the azole fluconazole.* ***D. Schematic representation of workflow involved in activating the transcription factors.*** *Based on the success of the control constructs, we selected a set of transcription factors for fusion constructs and characterized the consequences of these fusions through phenotypic analysis.*

We next selected 30 different transcription factors based on their phylogenetic uniqueness, their possible involvement in drug resistance, and their potential of functioning in rewired circuits. After generating these 30 transcription factor fusions we initially investigated their involvement in antifungal drug resistance. We selected 3 drugs for our preliminary screening - fluconazole, posoconazole, and amphotericin B. All three drugs target the cell membrane; fluconazole and posoconazole target lanosterol 14-alpha-demethylase (Erg11), an enzyme of the ergosterol pathway, whereas amphotericin B targets the end product ergosterol. We identified two transcription factor fusions, encoded by *ORF19.2752* and *ORF19.6339*, which gave resistance to all three drugs. As the Orf19.2752 fusion created a higher level of resistance than the Orf19.6339 fusion, and we prioritized Orf19.2752 for further study (Table 3.1; page 62).

3.3.2 Activation of Orf19.2752 confers multi-drug resistance.

We tested whether the VP64 fusion to Orf19.2752 (C4_02500C_A) could trigger resistance to a variety of drugs - fluconazole, posaconazole, terbinafine, nystatin, caspofungin, anidulafungin and amphotericin B (Fig. 3.2.A-B). The fusion of Orf19.2752 to VP64 increases the minimal inhibitory concentration as well as the minimal fungicidal concentration of fluconazole, amphotericin B and terbinafine (Fig 3.2.A), and also increased the MIC by more than 3-fold for these drugs (Fig 3.2.B) as well as posaconazole and nystatin (not shown). However, for the drugs caspofungin and anidulafungin that target the cell wall, there was no change in the MIC or growth rate for the activated strain relative to the control. Thus, activation of Orf19.2752 did not cause a general drug resistance, but did seem effective in generating resistance to cell membrane targeting drugs.

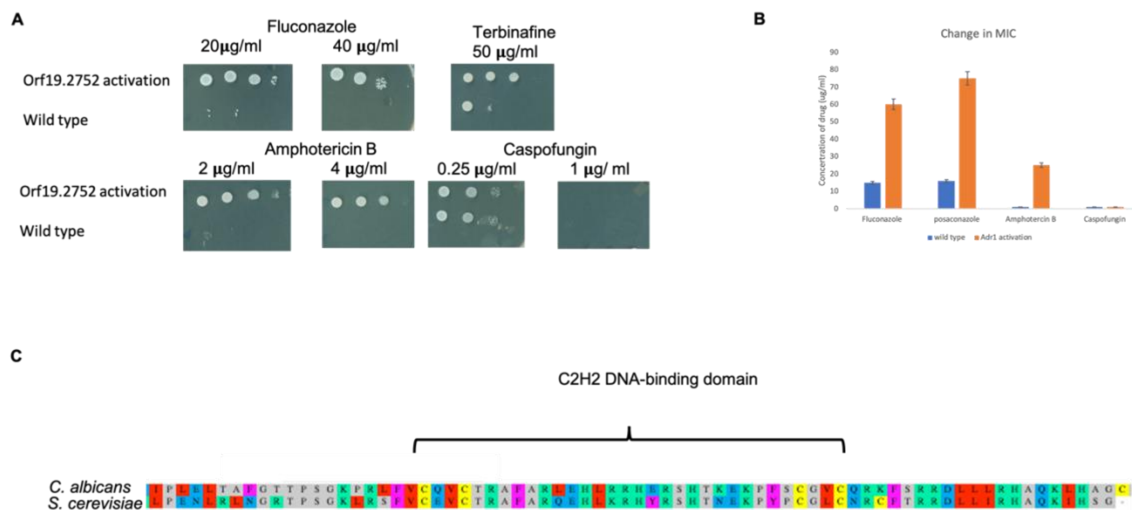


Fig. 3.2. A. Plate assay of YPD agar with different drugs confirming the VP64 fusion of Orf19.2752 generates fluconazole, amphotericin B and terbinafine resistance. *We used minimal inhibitory concentration (MIC) and 2MIC concentrations to check if the fusion*

construct creates resistance to the drugs. ***B. Activated Orf19.2752 caused fluconazole, posaconazole, terbinafine and amphotericin B resistance compared to wildtype.*** *Graphical representation of the change in minimal inhibitory concentrations of the Orf19.2752 activated strain in liquid YPD media containing the various drugs. After 24 hrs cells were checked for residual growth activity on YPD agar.* ***C. Sequence alignment shows that Orf19.2752 is similar to the Adr1 protein of S. cerevisiae****. Blast alignment of the zinc cluster DNA binding domain of the Adr1 transcription factor in S. cerevisiae with that of Orf19.2752 in C. albicans.*

3.3.3 Orf19.2752 is an ortholog of *S. cerevisiae* Adr1

Because transcription factors are frequently conserved across species, we looked for orthologs of the *C. albicans ORF19.2752 (C4_02500C_A)* gene. We found it to be highly similar to the *S. cerevisiae ADR1* gene; the two proteins have about 50% sequence identity and the N-terminal DNA binding domain is highly conserved between them (Fig 3.2.C).

In *S. cerevisiae*, Adr1, acting through a conserved binding motif 5' RCCCCM 3', is required for transcriptional regulation of ethanol, glycerol and fatty acid utilization (Denis, 1987; Young et al., 2002a). Due to the highly conserved DNA binding domains of the two orthologs, we searched for this ScAdr1 binding motif upstream of *C. albicans* ORFs. We found 221 genes with this motif in their predicted promoter regions using Meme-suite software as described in the methods. Of the genes with this promoter motif, a significant fraction (one tenth, or 20 genes) were implicated in ergosterol biosynthesis, while a further

one quarter (52 genes) were categorized as of unknown function. However, in contrast to the situation in *S. cerevisiae*, this motif is not enriched in the ethanol, glycerol and fatty acid metabolism genes in *C. albicans*. Because of the large number of motif-containing genes in the pathway of sterol biosynthesis, it appeared CaAdr1 might instead be linked to sterol production.

3.3.4 CaAdr1 influences sterol metabolism

Azole drugs target Erg11 of the ergosterol pathway in *C. albicans* (Flowers et al., 2015), and upregulation of Erg11 is one of the known mechanisms for drug resistance against fluconazole (Henry et al., 2000; Xiang et al., 2013). To test if the *ADR1* gene of *C. albicans* was involved in sterol metabolism, we deleted the gene and checked the consequences of loss of function; consistent with a role in sterol biosynthesis, *ADR1* deletion causes sensitivity to cell membrane targeting drugs like fluconazole (Fig S3.1). However, unlike the situation with *ScADR1*, the deletion of *CaADR1* did not block growth on fatty acid substrates. By contrast, deletion of the gene for the transcription factor Ctf1, identified as a regulator of fatty acid metabolism genes in the pathogen (Ramírez and Lorenz, 2009), completely blocked *C. albicans* growth on linoleic acid and severely compromised growth on olive oil (Fig 3.3.A). This suggests that in *C. albicans* Ctf1 is controlling fatty acid utilization, while Adr1 is not involved. In *Saccharomyces cerevisiae*, ScAdr1 was found to be haplo-insufficient for ethanol, glycerol and fatty acid metabolism. Similarly, in *C. albicans*, CaAdr1 is haplo-insufficient for fluconazole resistance, as the heterozygote showed sensitivity to fluconazole relative to the WT, but was clearly more resistant than the homozygous null (Fig S3.1.C).

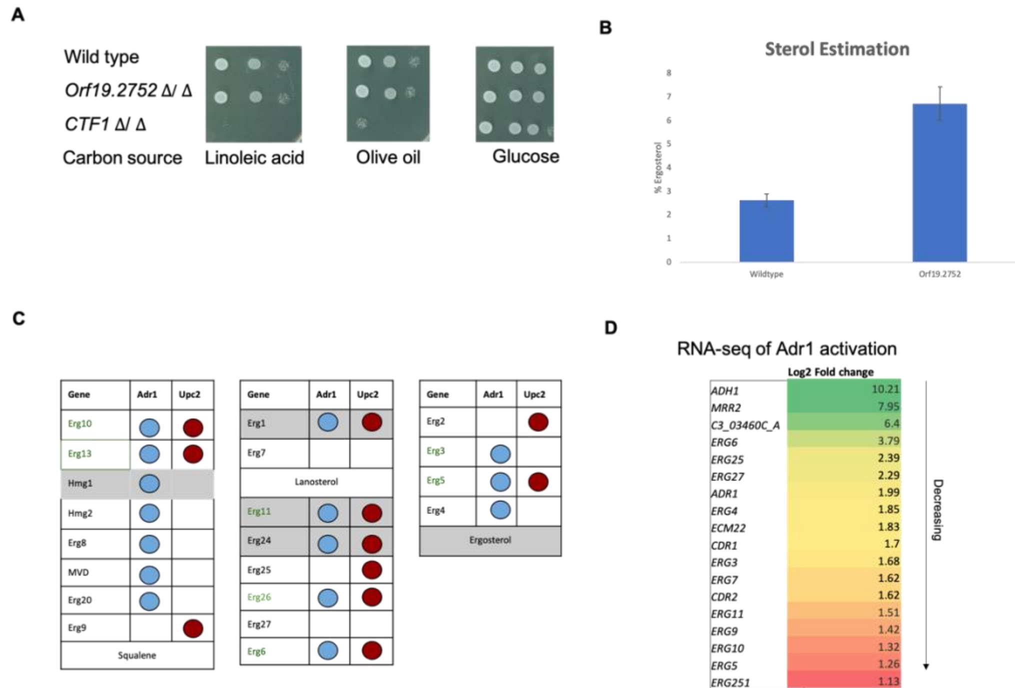


Fig. 3.3. A. Deletion of ADR1 (ORF19.2752). *We deleted the ADR1 gene and checked the resulting strain for fatty acid, glycerol and alcohol utilization. We found that it does not have any effect on growth on a fatty acid substrate, while a CTF1 deletion strain shows impaired growth on either linoleic acid or olive oil as a substrate.* ***B. Sterol estimation of wildtype and the Orf19.2752 activated strain.*** *We extracted the sterols from overnight grown C. albicans by the pargolol-hexane extraction method. The extracted sterols showed a four-peak spectral absorption pattern characteristic of ergosterol and 24(28)-dehydroergosterol [24 (28)-DHE]. The activated strain showed an approximately 2-fold increase in measured sterols.* ***C. Presence of the Adr1 motif upstream of ergosterol genes.*** *The presence of the candidate Adr1 motif is represented by a blue circle. We found most of the ergosterol genes have the candidate Adr1 motif 5'NRCCCCM 3' in their promoter regions.* ***D. Transcriptomic profile of selected genes of the Orf19.2752-VP64 fusion strain shows upregulation of ergosterol genes.*** *After activation of the Adr1 transcription*

factor we did RNAseq comparison of the activated strain and the wild type. We identified various ergosterol pathway genes being upregulated and observed high expression of the genes for the transcription factor Mrr2 and the alcohol dehydrogenase Adh1.

We directly checked the sterol content of the Adr1 activated strain and wildtype by extracting sterols with the organic solvents pargolol and hexane followed by spectrophotometric assessment. Activated CaAdr1 enhances the production of ergosterol (Fig 3.3.B). We also directly assessed the transcriptional consequences of Adr1 activation through RNAseq analysis. *ADH1*, which encodes the alcohol dehydrogenase that oxidizes ethanol to acetaldehyde (Bertram et al., 1996) was the most highly upregulated gene in our profile, and intriguingly, the orthologous gene in *S. cerevisiae* is a direct target of ScAdr1. We also found that the transcription factor Mrr2 is dramatically up-regulated; this could help explain the observed multi-drug resistance phenotype, as Mrr1 acts to up-regulate the *CDR1* and *CDR2* genes encoding transporter proteins, and our RNA-seq analysis shows both *CDR1* and *CDR2* are more highly expressed in the Adr1-VP64 fusion strain than in the control (Fig 3.3.D). To establish if the fluconazole resistance is a direct result of this upregulation of Mrr2, we deleted *MRR2* in the Adr1 activated strain. However, deletion of *MRR2* had essentially no effect on fluconazole resistance driven by activation of Adr1, suggesting the up-regulation of the Mrr2 transcription factor was in fact not critical in creating the resistance phenotype (Fig S3.1.D).

3.3.5 Adr1 and Upc2 roles in azole response

Consistent with the presence of the candidate Adr1 binding motif in their promoters, we found the expression level of most of the ergosterol pathway genes, including Erg11, was upregulated by the activated Adr1 construct. This increase in ergosterol pathway gene expression was however not associated with up-regulation of the classic ergosterol biosynthesis pathway transcription factor Upc2, which functions as a key regulator of the pathway in both *S. cerevisiae* and *C. albicans* (Flowers et al., 2012; Lohberger et al., 2014; Vasicek et al., 2014). Therefore, it appears that Adr1 activation of the *C. albicans* ergosterol pathway genes is likely direct and thus the fluconazole resistance generated by the Adr1VP64 fusion protein may be due to the generalized increase in the expression of these ergosterol biosynthetic pathway genes. We then asked whether the fluconazole resistance generated by Adr1 activation was fully independent of Upc2. First, we compared fluconazole resistance levels in strains with Adr1 activated and with Upc2 activated, as well as in strains with Upc2 deleted and with Adr1 deleted. As is shown in Fig. 4A-B, both Upc2 activation and Adr1 activation gave similar high levels of resistance to fluconazole, while both Adr1 deletion and Upc2 deletion conferred sensitivity to fluconazole, with the Upc2 deletion strain being somewhat more sensitive. Secondly, we assessed the resistance to fluconazole of the doubly activated strain; in this case the strain grew poorly in the absence of drug but was resistant to fluconazole at similar levels to that of the single activated mutants (Fig 3.4.A-B). Finally, we investigated the behaviour of cells with activated Adr1 that lacked functional Upc2, and cells with an activated Upc2 that lacked Adr1. Here we observed that loss of Upc2 had essentially no effect on the fluconazole resistance caused by activation of Adr1, suggesting the effect of Adr1 on drug resistance

is independent of Upc2, while loss of Adr1 significantly compromised fluconazole resistance caused by activation of Upc2 (Fig 3.4.A-B). This is consistent with part of the effect of Upc2 activation on azole resistance working through Adr1. We checked the upstream sequences of the *UPC2* and *ADR1* genes for regulatory motifs and found that the promoter sequence of the *ADR1* gene contains a potential *UPC2* motif (Fig 3.4.C), consistent with Adr1 being part of the Upc2 regulon, while the *UPC2* promoter lacks any potential Adr1 binding motif.

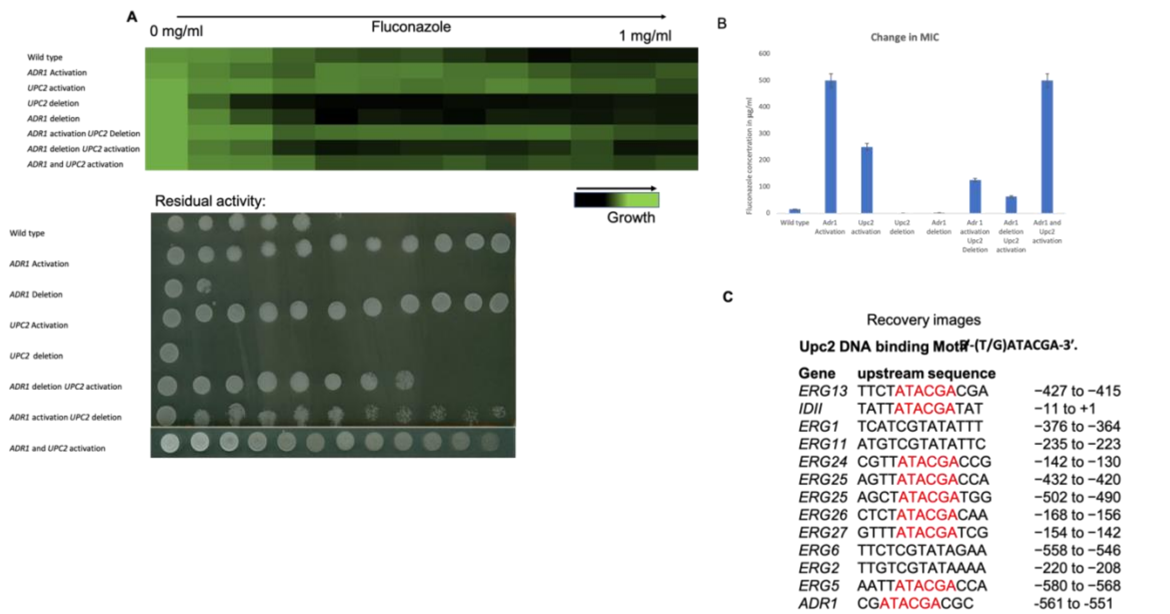


Fig. 3.4. A. Adr1 and Upc2 influence response to fluconazole. *Fluconazole MIC assay of ADR1 deletion, ADR1 activation ADR1 activation in UPC2 deletion, ADR1 deletion in UPC2 activation and double activation, followed by the residual growth on solid medium*. ***B. Fold change in MIC from Adr1 and Upc2 activation and deletion***. *Graphical representation of the change in minimal inhibitory concentrations of various drugs in the Orf19.2752 activated strain.* **C. Upc2 DNA binding motif in upstream of various regulons including ADR1.** *Upstream regions of genes including the transcription factors Adr1, Upc2, Mrr2 and Mrr1 as well as the structural genes for the ergosterol biosynthesis*

pathway were examined for candidate TF binding motifs. We identified a potential Upc2 binding motif upstream of the ADR1 gene.

3.3.6 Phylogenetic analysis

We examined a phylogenetic profile of the Ascomycota and characterized the upstream sequences of all ergosterol biosynthesis genes for the regulatory motifs associated with both the Adr1 and the Upc2 transcription factors. In *C. albicans* our data suggest that Adr1 works together with Upc2 to control the ergosterol biosynthesis pathway, with Adr1 also controlling the expression of genes such *ADH1,* encoding the alcohol dehydrogenase which oxidizes ethanol to acetaldehyde, and *MRR2,* encoding a stress-responsive transcription factor. The more basal filamentous fungi also have the candidate binding motifs for both Adr1 and Upc2 the promoters of the ergosterol biosynthesis genes. However, the presence of the Adr1 motif in the promoter of *MRR2* is very specific to the CTG clade species (Supplementary data file 3.1).

In the evolutionary trajectory leading to *S. cerevisiae* it appears that the Adr1 transcription factor was repurposed to control alcohol, fatty acid and glycerol utilization, taking over from Ctf1 orthologs that do the job in the filamentous fungi and the CTG clade species. Based on the search of promoter motifs we found that Upc2 motif is found in the promoters of ergosterol pathway genes throughout the fungi along with the Adr1 motif. But after *C. guilliermondii* in the phylogeny the Adr1 motif signal weakens, leaving only Upc2 with a strong signal associated with the ergosterol pathway. The Adr1-binding motif signal connected the ergosterol pathway genes in the CTG clade appears to be gradually

transferred to genes involved in control of alcohol and fatty acid utilization in *S. cerevisiae* and its relatives (Supplementary data file 3.1).

3.4 Discussion

One of the common approaches to investigating the regulatory networks controlling drug resistance in fungal pathogens is through comparison with the *S. cerevisiae* circuits; this approach has led to the discovery of many transcription factors responsible for drug resistance in both *S. cerevisiae* and *C. albicans* including Fcr1, Ndt80, Mrr1 and Upc2 (Chen et al., 2004a; Chen et al., 2004b; Chen et al., 2009; Flowers et al., 2012; Morschhauser et al., 2007). However, these two fungi diverged as long as 300 million years ago, and for species diverged by such an evolutionary distance, the majority of the DNA-binding patterns of a given regulator in one species are unlikely preserved in the other species (Johnson 2017). Overall, genome-wide correlations converge on about 10% overlap for species with this level of divergence (Johnson, 2017b), and therefore, there is a significant probability that many of the transcription factors responsible for drug resistance will be different between *C. albicans* and the budding yeast.

To identify candidate transcription factors with Candida-specific roles in drug resistance we selected a set of TF-encoding genes that were either not found in *S. cerevisiae* or that were predicted to have potentially changed function between the two species. We identified 30 such transcription factors and activated them in order to identify potential roles in drug resistance (as well as other cellular processes) (Table 3.1; page 62). Among these transcription factors, Orf19.2752 activation resulted in clear resistance to a set of drugs

generally targeting the cell membrane; activation of this transcription factor generated resistance to azoles, allylamines and polyenes (Fig 3.2). Sequence comparisons established that *ORF19.2752* was the *C. albicans* ortholog of the *S. cerevisiae* gene *ADR1*, a gene not linked to drug resistance in the budding yeast. These two transcription factors share a highly conserved DNA binding domain (Fig 3.1.C).

In *S. cerevisiae,* Adr1 is involved in the transcriptional regulation of genes involved in the catabolism of ethanol, glycerol and fatty acids (Denis, 1987; Young et al., 2002a), and is proposed to act through a candidate DNA-binding motif, 5'NRCCCCM3' in the promoter regions of these genes (Tachibana et al., 2005). Interestingly, in *C. albicans* this same DNA-binding motif is enriched in the upstream regions of the ergosterol-pathway genes, whereas it was generally absent from the promoters of the ethanol, glycerol and fatty acid utilization genes of the pathogen. This suggests that Adr1 may have been rewired from the ergosterol biosynthesis pathway in other fungi to the metabolic utilization of ethanol, glycerol and fatty acids in *S. cerevisiae*. Further investigation established that activation of CaAdr1 generated transcriptional upregulation of most of the ergosterol pathway genes in the pathogen (Fig 3.3.D). However, Upc2, the well-established regulator of the Erg-pathway genes in both *C. albicans* and *S. cerevisiae*, was not transcriptionally upregulated, suggesting that in *C. albicans* Adr1 activation of the ergosterol pathway genes was not going through Upc2 (Fig 3.3.A).

Thus it appears that in *C. albicans,* Adr1 acts in addition to Upc2 in regulating the ergosterol biosynthetic pathway. As well, in our RNA seq analysis we found the *ADH2*

gene highly upregulated in response to Adr1 activation, similar to the situation in *S. cerevisiae*. The promoter of *ADH2* in *C. albicans* contains the Adr1 motif, suggesting that Adh2 is under direct control of Adr1. However, the bulk of the circuit of ethanol, fatty acid and glycerol metabolism controlled by Adr1 in *S. cerevisiae* is under the control of Ctf1 in *C. albicans* (Ramírez and Lorenz, 2009). We confirmed this by deleting both Adr1 and Ctf1 in *C. albicans* and testing growth on fatty acid substrates. We found that Adr1 deletion did not affect growth on substrates like olive oil and linoleic acid, whereas Ctf1 deletion gives a clear auxotrophy (Fig 3.3.B).

These data underline the multiple circuit restructurings modifying the control of these pathways in different fungi. In *S. cerevisiae* Cat8 and Adr1 both appear to have rewired to connect to the module controlling alcohol, acetic acid and fatty acid utilization; Adr1 from the ergosterol circuit, Cat8 from gluconeogenesis (ScSip4). Another event is the disappearance of the Ctf1 transcription factor from the *S. cerevisiae* genome, as there is no apparent ortholog of Ctf1 in S. *cerevisiae.* This loss could be facilitated by transfer of the Ctf1 regulon to Adr1 control in *S. cerevisiae*. In *S. cerevisiae* Upc2 gains a paralog (Emc22) and apparently unique control over the ergosterol pathway (Yang et al., 2015).

In *C. albicans*, Adr1 activation causes upregulation of many ergosterol pathway members, Erg11 (target of azoles), Erg1 (target of allylamines), Erg2 (target of morpholines), HMG1/2 (target of statins), and causes increases in ergosterol itself (target of polyenes), which has the potential of generating multidrug resistance. The activation of Adr1 dramatically enhances the MIC of fluconazole, amphotericin B, terbinafine and statins (Fig

3.2). We created a series of strains in order to determine the how Upc2 and Adr1 are influencing the ergosterol pathway. Activation of either Upc2 or Adr1 enhanced azole resistance, while deletion of either gene created azole sensitivity. Activation of both transcription factors at the same time caused poor growth, perhaps due to the disturbed circuits or due to the activation of an Erg3-driven side branch of the pathway generating the toxic 14a methyllergosta 8-24(28) dienol. However, the relative resistance to azoles remained similar to the individually activated strains, suggesting the actions of the two transcription factors are not additive or synergistic.

The fluconazole resistance caused by activation of Upc2 is significantly dependent on the presence of Adr1, but loss of Upc2 had very little effect the resistance profile of strains with an activated Adr1 (Fig 3.4.A). In addition, there is a potential Upc2 DNA binding site in the promoter region of *ADR1* (Fig 3.4.C), while *UPC2* has no candidate site for Adr1. These results are consistent with Upc2 serving as a master regulator of ergosterol biosynthesis, controlling ERG gene expression both directly and through regulation of the Adr1 transcription factor that also serves as an activator of ERG gene expression. In the absence of Upc2, Adr1 is sufficient to ensure ERG expression, although response to azole drugs is compromised by loss of either TF.

Among the highly up-regulated genes resulting from Adr1 activation is the gene encoding Mrr2 (Fig 3.3.D), itself a transcription factor involved in the expression of the multi-drug-resistance regulating membrane transporter Cdr1. However, even though *CDR1* expression was somewhat up-regulated in the Adr1 activated strain, the Mrr2 up-regulation

did not seem critical for the observed fluconazole resistance, because deletion of *MRR2* in the Adr1 activated strain did not compromise this resistance (Fig S3.1.D).

While it appears that in the filamentous fungi and the CTG clade species the transcription factor Adr1 is linked to ergosterol biosynthesis, in the Saccharomycotina it has been shifted to controlling the pathway for ethanol, glycerol and fatty acid degradation (Young et al., 2003; Young et al., 2002b) replacing Ctf1 that controls the process in the non-Saccharomycotina species. This transfer appears so complete that the Ctf1 factor has been lost in the Saccharomycotina. Intriguingly, throughout this transition, Adr1 has maintained a role in control of the expression of the alcohol dehydrogenase catalyzing the oxidizes ethanol to acetaldehyde (*ADH1* in *C. albicans*, *ADH2* in *S. cerevisiae*). To deal with ethanol toxicity, in *S. cerevisiae* ethanol is modified into unsaturated lipids and ergosterol (Ma and Liu, 2010). The rewiring of Adr1 from the ergosterol pathway to the ethanol utilization process (Tachibana et al., 2005; Thukral et al., 1991) may have been driven by the shift to the Crabtree positive life-style of *S. cerevisiae*, requiring the ability to both tolerate and process high concentrations of ethanol.

In the fungal pathogen *C. albicans* the transcription factor Adr1 functions, in conjunction with the zinc cluster TF Upc2, acts as a regulator of the ergosterol pathway. This contrasts to the model yeast *S. cerevisiae,* where Adr1 functions in the regulation of alcohol and fatty acid metabolism while ergosterol biosynthesis remains under the specific control of Upc2. Adr1 appears to act as a component of the Upc2 regulon; activation of Upc2 is less effective at generating azole resistance in the absence of Adr1, while the effects of activation of both elements are epistatic. Loss of both elements may be lethal, consistent

with a requirement of at least one regulator for proper production of the cell membrane. We propose a model which explains the coordination of Upc2, Mrr1, Mrr2 and Adr1 for ergosterol biosynthesis and for their role in drug resistance. Upc2, when bound to ergosterol, remains inactive in the cytoplasm. Once there is depletion of ergosterol, whether by environmental factors or due to the presence of the drugs targeting the ergosterol pathway, it becomes activated. This activated Upc2 transcription factor goes to the nucleus and activates key players of ergosterol biosynthetic pathway together with the transcription factors Mrr1 and Adr1. Adr1 acts to activate the ergosterol pathway genes as well as the TF Mrr2. This activation of Upc2 and Adr1 leads to the upregulation of the ergosterol biosynthesis pathway, as well as the activation of phospholipid transporters Cdr1 and Cdr2 that can also function to export antifungal drugs (Fig 3.5).

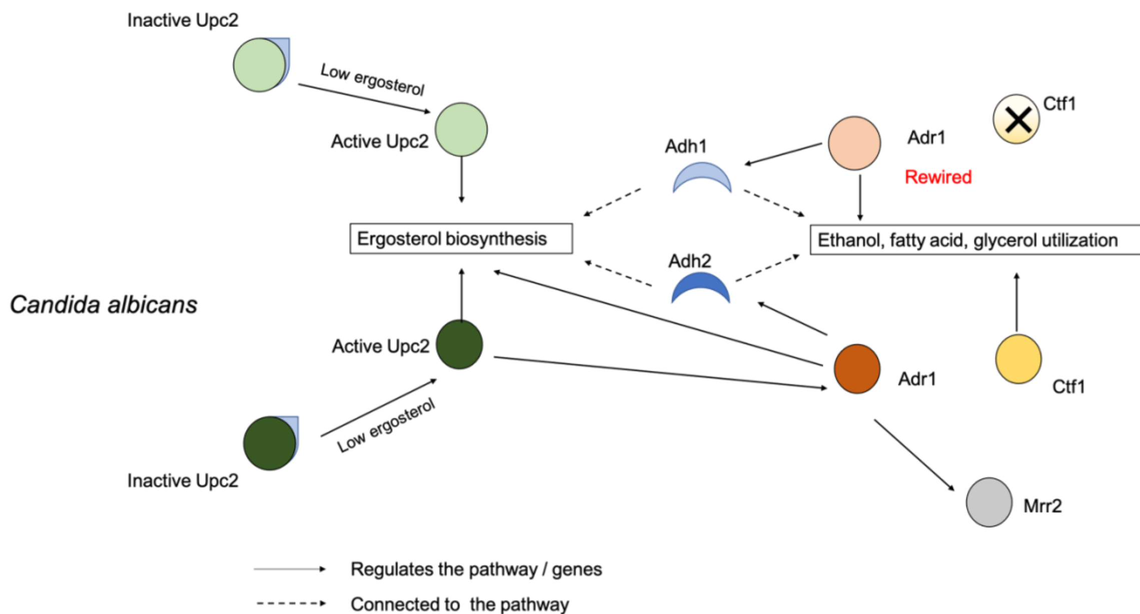


Fig. 3.5. Proposed model of the multiple rewiring of transcription factors between S. cerevisiae and C. albicans along with coordination of Upc2, Adr1 and Mrr2 during the presence of azoles or depletion of ergosterol leading to drug resistance in C. albicans.

In both S. cerevisiae and C. albicans, the Upc2 transcription factor binds to ergosterol and remains inactive in cytosol. Depletion of ergosterol changes Upc2 to an active form, which enters the nucleus and initiates the transcription of the genes for ergosterol biosynthesis. In C. albicans activated Upc2 also triggers expression of the Adr1 transcription factor which further serves to direct expression of ergosterol biosynthesis genes, as well as the alcohol metabolism gene Adh2 and the Mrr2 transcription factor. In S. cerevisiae, Adr1 has been rewired to control other parts of the alcohol utilization circuit in addition to alcohol dehydrogenase, as well as to control both fatty acid and glycerol utilization circuits.

Table 3.1. Screening of activated transcription factors in presences of different drugs:

Transcription factor	Fluconazole	Posaconazole	Amphotericin B	Caspofungin
Try3 **Orf19.1971**	-	-	+	-
C5_04410C_A **Orf19.3928**	-	-	-	-
C3_01220W_A	+	++	-	-

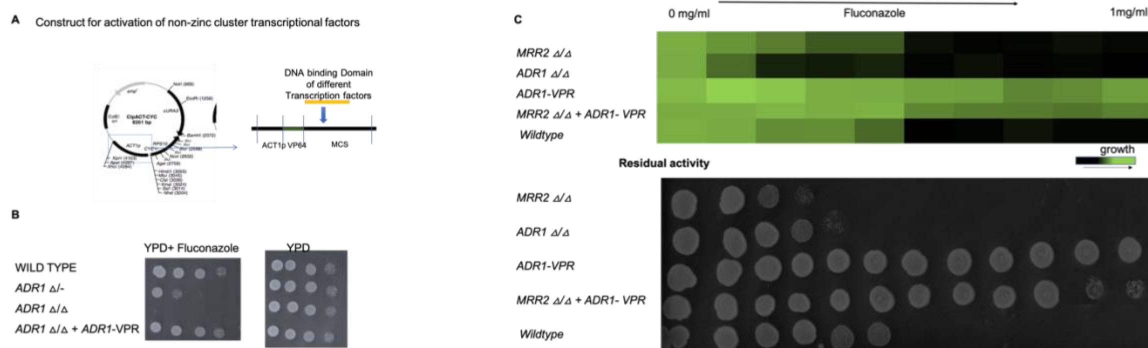
Orf19.1729				
C1_04510W_A **Orf19.6845**	++	++	+	+
C2_05640W_A **Orf19.6874**	-	-	-	-
Hap41 **Orf19.740**	-	-	+	-
Hap42 **Orf19.1481**	++	++	+	+
Try6 **Orf19.6824**	-	-	-	+++
C2_02530W_A **Orf19.1577**	-	-	-	-
C2_01420C_A **Orf19.1447**	+	+	++	-
C4_07150W_A **Orf19.3088**	+	+	+	+

Uga33 **Orf19.7317**	-	-	-	-
Nrg2 **Orf19.6339**	++	++	++	++
Ace2 **Orf19.6124**	-	-	-	-
C1_13440C_A **Orf19.4972**	-	-	-	-
Rfx1 **Orf19.3865**	-	-	+	++
Grf10 **Orf19.4000**	+	-	+	+++
C4_05680W_A **Orf19.1253**	+	+	-	++
CR_05880W_A **Orf19.6626**	++	++	+	-
C2_08950W_A **Orf19.211**	+	+	-	+

C5_04280C_A Orf19.3912	-	-	-	-
C4_02500C_A Orf19.2752 (Adr1)	+++	+++	+++	+
C2_10660W_A Orf19.5343	+	+	+	+
C1_11690W_A Orf19.1150	+	+	-	-
C1_10020W_A Orf19.4869	-	-	+	++
C4_05880W_A Orf19.1275	-	-	-	-
Try4 Orf19.5975	-	-	-	+++
C4_05880W_A Orf19.1275	-	-	-	-
Met4 Orf19.5312	-	-	-	+

Gcn4	+	+	+	-
Cap1 (Control)	+++	+++	++	+
Tec1 (control)	++	++	+	+

The optical densities are represented by plus signs (+) where + indicates growth equivalent to that of the positive control, ++ indicates a higher OD growth than + and +++ higher than ++ (-<+<++<+++ OD at 600nm after 24 hours). *Results focused on in this paper are in red.*



Supplementary fig. 3.1. *A. The construct for activation of non-zinc cluster transcription factors.* B. *The adr1Δ/Δ strain (ORF19.2752) showed fluconazole sensitivity. The heterozygous adr1Δ/+ shows better growth than the adr1Δ/Δ in the presence of fluconazole.* C. *The mrr2 Δ/Δ in the Adr1 activated strain caused no changes in drug resistance.*

Supplementary Data file-

https://drive.google.com/file/d/1qy9C0oM_GR5Px14letdReJMLlFrFdDcX/view?usp=sharing

Chapter 4: Modulation of the complex regulatory network for methionine biosynthesis in fungi

4.1 Introduction

Methionine occupies a central role in metabolism and growth control in fungi. It is synthesized by the methionine (Met) biosynthesis pathway, also known as the sulphur assimilation pathway (Hebert et al., 2011)This pathway is regulated by a number of signals, including sulphur-containing compounds and amino acids, cadmium, arsenite, zinc, and potentially the diauxic shift (Barbey et al., 2005a; Kuras et al., 2002; Lee et al., 2010).
The transcriptional regulation of the genes encoding this methionine metabolic circuitry has been investigated in a number of fungi. The simplest circuits, found in the filamentous fungi, appear to be controlled by a single transcription factor. These regulators are members of the basic leucine zipper class of transcription factors, contain highly similar DNA binding domains, and are orthologs of the Met4 protein of *S. cerevisiae*. These include the Cys3 protein of *N. crassa* and the MetR protein of *A. nidulans*. In *A. nidulans* and its close relatives (*A. fumigatus, A. niger, A. terreus A. oryzae and P. rubens*), a gene duplication event has given rise to two paralogs, MetR and MetZ, that are both part of sulphur assimilation control (Pilsyk et al., 2015). Between these two paralogs, MetR can be considered most likely the direct ortholog of Met4 based on synteny and the overall level of sequence identity. These two paralogs are predicted to form dimers of MetR-MetR, MetR-MetZ and MetZ-MetZ and bind to the palindromic sequence 5′-ATGRYRYCAT-3′ (Pilsyk et al., 2015).

A highly sophisticated circuitry regulating the methionine biosynthesis has been identified in *S. cerevisiae* and its close relatives (Baker et al., 1989; Baker and Masison, 1990; Blaiseau and Thomas, 1998; Bram and Kornberg, 1987; Kaiser et al., 2006; Thomas and Surdin-Kerjan, 1997), In *S. cerevisiae*, the Met4 protein associates with at least four other transcription factors, the basic helix-loop-helix protein Cbf1, the basic leucine zipper protein Met28, and two paralogous zinc finger transcription factors, Met31 and Met32. Met4 generates the central transactivating activity in these complexes but depends on the

DNA binding activity of either Cbf1 or Met31/32 for promoter recruitment, potentially due to an insertion event which disrupted the structure of the DNA-binding leucine zipper module of the Met4 protein.

In *S. cerevisiae* the basic leucine zipper protein Met28 also does not appear to independently bind DNA but has been shown to enhance promoter binding of the Cbf1-Met4 complex (Kuras et al., 1997). Cbf1 recognizes the sequence TCACGTG whereas Met31 and Met32 recognize AAACTGTG. Interestingly, the co-activators Met4 and Met28 also show a binding affinity to the so-called recruitment motif 'RYAAT' (Siggers et al., 2011), and thus the association of Cbf1 and Met32 to their respective binding sites appears to be facilitated by Met4 and Met28.

The entire sulphur assimilation circuitry has not been as extensively studied in the *Candida* (CTG) clade. Here we have investigated transcriptional regulation of sulphur metabolism in *Candida albicans* to establish the structure of the regulatory circuitry. This circuit represents a transition between the simple Met4 superfamily member core regulator found in the filamentous fungi, and the highly sophisticated complex of Cbf1, Met4, Met28 and Met32/31 found in *S. cerevisiae*.

4.2 Materials and Methods

4.2.1 Bioinformatics analysis

Sequences of genes *MET4*, *MET28*, *MET32* and *CBF1* were obtained from the Candida Genome Database (CGD- http://www.candidagenome.org/) and the Saccharomyces Genome Database (https://www.yeastgenome.org/). Gene orthogroup assignments for all predicted protein-coding genes across 23 Ascomycete fungal genomes were obtained from the Fungal Orthogroups Repository (Wapinski et al., 2007) maintained by the Broad Institute (broadinstitute.org/regev/orthogroups).
DNA sequence motifs were identified using the Web-based motif-detection algorithm MEME (http://meme.sdsc.edu/meme/intro.html; (Bailey et al., 2015). For more stringent

motif identification, we used MAST hits with an E value of <50. An E value of 500 corresponds roughly to a p-value of 0.08 in our analysis, and an E value of 50 roughly corresponds to a p-value of 0.008. We also used AME (http://meme-suite.org/tools/meme), which identifies known motifs throughout the Candida upstream sequences.
Protein domains and linear motifs were detected from each individual TF protein sequence using INTERPROSCAN, PFAM and ELM motif definitions.

4.2.2 Strains and culture conditions:

For general growth and maintenance of the strains, the cells were cultured in fresh YPD medium (1 % w/v yeast extract, 2 % w/v Bacto peptone, 2 % w/v dextrose, 80 mg/L uridine with the addition of 2 % w/v agar for solid medium) at 30 °C. For methionine auxotrophy we used synthetic dextrose (SD) medium (0.67 % w/v yeast nitrogen base, 0.15 % w/v amino acid mix without methionine, 0.1 % w/v uridine, 2 % w/v dextrose and 2 % w/v agar for solid media).

4.2.3 Gene knockout using CRISPR

All *C. albicans* mutants were constructed in the wild type strain CaI4. The protocol used for the CRISPR-mediated knockout of *MET4*, *MET28* and *MET32* was adapted from (Vyas et al., 2015); we used *URA3* replacements in our study. CRIPSR-mediated knockouts used the lithium acetate method of transformation with the modification of growing transformants overnight in liquid YPD at room temperature after removing the lithium acetate-PEG. *C. albicans* transformants were selected for on SD URA-plates.

4.2.4 Activation of Met4, Met28 and Met32

As described in chapter 2, for the activation module, the *ACT1* promoter and VP64 were amplified by PCR and homology was created by primer extension such that there is a restriction site of restriction enzyme Mlu I in between *ACT1* and VP64. After this, the C-terminus of the VP64 cassette was cloned into the multiple cloning site. The pCIPACT and

the ligated Act1-VP64-MCS were digested with Hind III and were ligated using T4 DNA ligase. This ligated CIPACT-VP64 plasmid was transformed into *E. coli* using the calcium chloride method.

Plasmids extracted from colonies that were determined to have the guide sequence successfully cloned in were then used to transform *C. albicans* using a lithium acetate transformation protocol. pCIPACT1 was linearized by StuI-HF digest, and 1-2 µg of the linearized plasmid was used in the transformation. *C. albicans* transformants were selected on SD URA-plates.

4.2.5 RNA seq analysis

As described in chapter 3, the CaMet4, CaMet28, CaMet32 and SC5143 strain cultures were grown in YPD overnight at 30 °C, diluted to OD600 of 0.1 in YPD at 30 °C, and then grown to an OD600 of 0.8 to 1.2 on a 220-rpm shaker. Total RNA was extracted using the Qiagen RNeasy minikit protocol, and RNA quality and quantity were determined using an Agilent bioanalyzer. Paired-end sequencing (150bp) of extracted RNA samples was carried out at the Quebec Genome Innovation Center located at McGill University using an Illumina miSEQ sequencing platform. Raw reads were pre-processed with the sequence-grooming tool cutadapt version 0.4.1 (Martin, 2011) with the following quality trimming and filtering parameters (`--phred33 --length 36 -q 5 --stringency 1 -e 0.1`). Each set of paired-end reads was mapped against the *C. albicans* SC5314 haplotype A, version A22 downloaded from the Candida Genome Database (CGD) (http://www.candidagenome.org/) using HISAT2 version 2.0.4. SAM tools was then used to sort and convert SAM files. The read alignments and *C. albicans* SC5314 genome annotation were provided as input into StringTie v1.3.3 (Pertea et al., 2015), which returned gene abundances for each sample. Raw and processed data have been deposited in NCBI's Gene Expression Omnibus (Edgar et al., 2002).

4.3 Results

4.3.1 Orthologs of methionine biosynthesis regulators in *C. albicans.*

In the ascomycetes, key methionine regulating TFs have different names, such as Met4 in *S. cerevisiae*, MetR in *A. nidulans* and Cys3 in *N. crassa*, but they are structurally and functionally orthologous. A binding target for this class of transcription factors, 5′-ATGRYRYCAT-3′, has been identified that is conserved among the filamentous fungi (Li and Marzluf, 1996). In *S. pombe* the equivalent TF, Zip1, appears rewired and is involved in a cadmium-responsive checkpoint pathway (Harrison et al., 2005) rather than in methionine biosynthesis. We found 2 proteins in *C. albicans* that showed strong structural similarities to this Met4 class of TFs; Orf19.5312 (designated CaMet4) and Orf19.7046 (designated Met28). Thus, it appears, similar to the situation in *A. nidulans*, that the CUG clade has had a duplication of the basal *MET4* gene leading to two paralogs, *MET4* and *MET28*.

The Met4 protein of *C. albicans* is 385 amino acids long, with a leucine bZIP domain (aa303 to aa354) that includes a specific DNA binding region (aa308 – aa322) (Fig 4.1). Among the orthologs, CaMet4 is intermediate in size between the *A. nidulans* MetR of 294 amino acids, and the *S. cerevisiae* Met4 of 672 amino acids. The well-studied ScMet4 contains multiple domains in addition to the DNA binding region; these include an activation domain (AD), a ubiquitin-interacting motif (UIM), an inhibitory region required for repression of Met4 activity by methionine (IR), an auxiliary domain required to fully relieve IR mediated repression (AUX), and a protein-protein interaction domain that binds Met31 and Met32 (INT). We searched for these domains in AnMetR, ScMet28, CaMet28, CaMet4 and NcCys3. CaMet4 shows considerable similarity to ScMet4, and even though the overall protein is almost half the size we were able to identify a candidate activation domain, a UIM region, and the B-zip region (Fig 4.1).

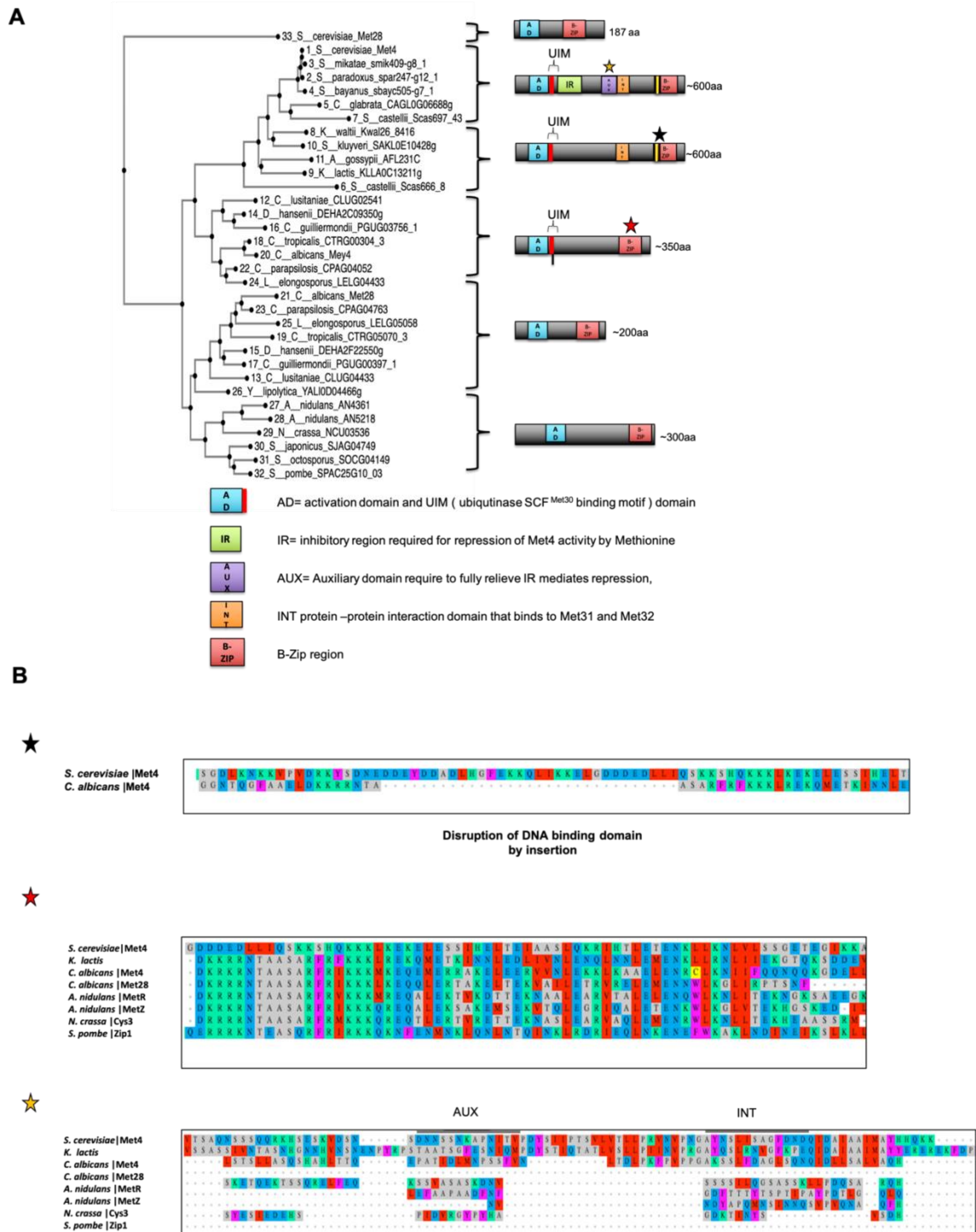


Fig. 4.1. A. Phylogenetic tree. *The tree is constructed based on the sequence similarities of Met4 and Met28 across the fungal kingdom. A block diagram showing all the domains of the transcription factors. B. The sequence alignment of subdomain that are highlighted by star with respective color shows that the disrupted B-zip domain, AUX and INT domain*

are present only in S. cerevisiae and K. lactis clade whereas in the rest of fungi complete B-Zip domain is conserved.

CaMet28 shows limited similarity with ScMet28 or ScMet4; although it is essentially the same length as ScMet28 and has the characteristic BZip domain, overall CaMet28 show the most similarity with MetR of *A. nidulans*. In particular, the DNA-binding domains of CaMet28 and MetR were highly conserved (Fig 4.1), suggesting the possibility that CaMet28 could recognize a 5′-ATGRYRYCAT-3′ motif in promoter regions of Candida genes. Using Meme-suit online tool, we screened for this motif in the 1000 bp upstream regions of all the open reading frames of the *C. albicans* genome, and found only 12 genes that have this exact sequence in their promoters; these genes are not involved in any obvious common cellular function, and none are implicated in methionine biosynthesis. This result suggests that Met28 may not be involved in control of the Met regulon of *C. albicans*.

4.3.2 Deletion of the *MET4* and *MET28* genes in *C. albicans* suggests Met4 is regulating methionine biosynthesis.

The two alleles of *MET4* were deleted from the prototrophic strain SC5314 using the CRISPR-Cas9 system (Vyas et al., 2015). The *MET4* deleted strain was somewhat slow-growing compared to the wild type in rich YPD medium. When grown on media lacking methionine, the *MET4* deleted strain was unable to grow (Fig 4.2.A), suggesting that *MET4* deletion inhibits methionine biosynthesis in *C. albicans* and makes cells dependent on supplemented methionine. We were unable to construct a *MET28* homozygous deleted strain using the same approach, although heterozygotes were easy to obtain. Previously, it was reported that *MET28* disruptants were not obtained by the UAU1 method (Segal et al., 2018), which suggests that *MET28* might be an essential gene in *C. albicans*.

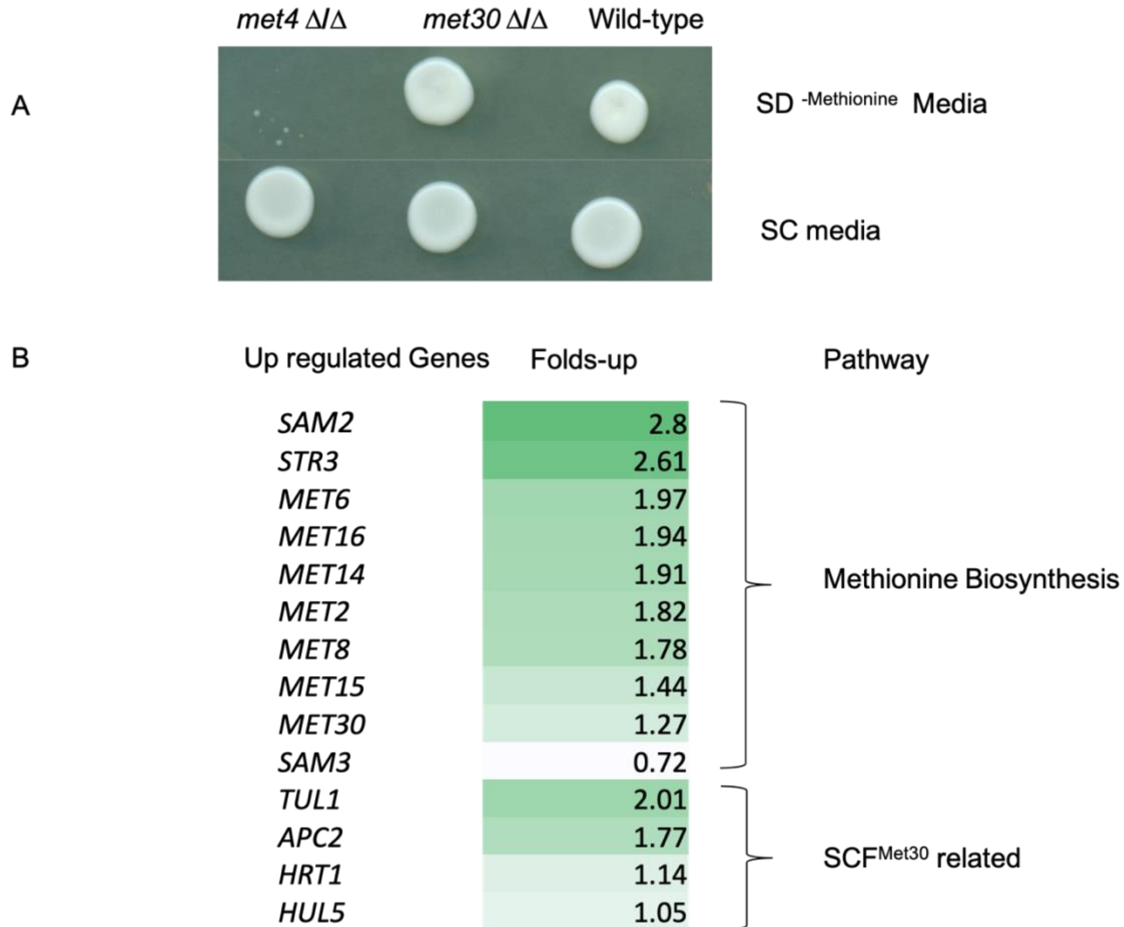


Fig. 4.2.A. Activation and Deletion of Met4 and Met32 in C. albicans. *Methionine starvation of strain SC5314, Met4 deleted strain and Met32 deleted strain. The Met4 deleted strain shows no growth during methionine starvation. B. Transcriptomic profile of top 50 up-regulated genes in the Met4 activated strain show up regulation of methionine biosynthesis-related genes along with some important SCF^{Met30} related genes. There is no obvious functional enrichment within the Met4 up-regulated genes that are not linked to methionine regulation.*

4.3.3Activation of Met4 and Met28 suggests that only Met4 is involved in methionine biosynthesis

Since we were unable to create null mutants of *MET28*, we could not test its involvement in methionine biosynthesis through loss of function, so we investigated both Met4 and

Met28 function in methionine biosynthesis by adding the strong VP64 activation domain to each transcription factor (Ndisdang et al., 1998). The transcriptomic profile of cells containing Met4-VP64 showed up-regulation of SCFMet30, several methionine biosynthesis genes and some important Met30 related genes (Fig 4.2.B). When we divide methionine biosynthesis into 3 different modules; module 1- the sulphur assimilation pathway; module 2- methionine biosynthesis (homoserine to methionine and cysteine) and module 3- the S-adenosylmethionine cycle; then the up-regulated genes are in modules 1 and 3, with *CYS3*, *CYS4* and *MET2* of the methionine and cysteine biosynthesis and glutathione producing circuit not up-regulated. By contrast, Met28 activation shows a different set of up-regulated genes, with none of them related to methionine biosynthesis. Among these genes 23/50 were annotated as biological function unknown, 7/50 are candidate secreted or cell wall-associated proteins, while 11/50 do not have convincing orthologs in other fungi. Furthermore, 10 out of 50 genes have the candidate 5' ATGRYRYCAT 3' MetR DNA binding site in their promoter regions. Among these 10 genes 2 are essential, and out of the top 40 up-regulated genes 9 genes are essential, which could explain the Met28 essentiality, but the up-regulated genes do not allow a clear prediction of Met28 function (Fig 4.3).

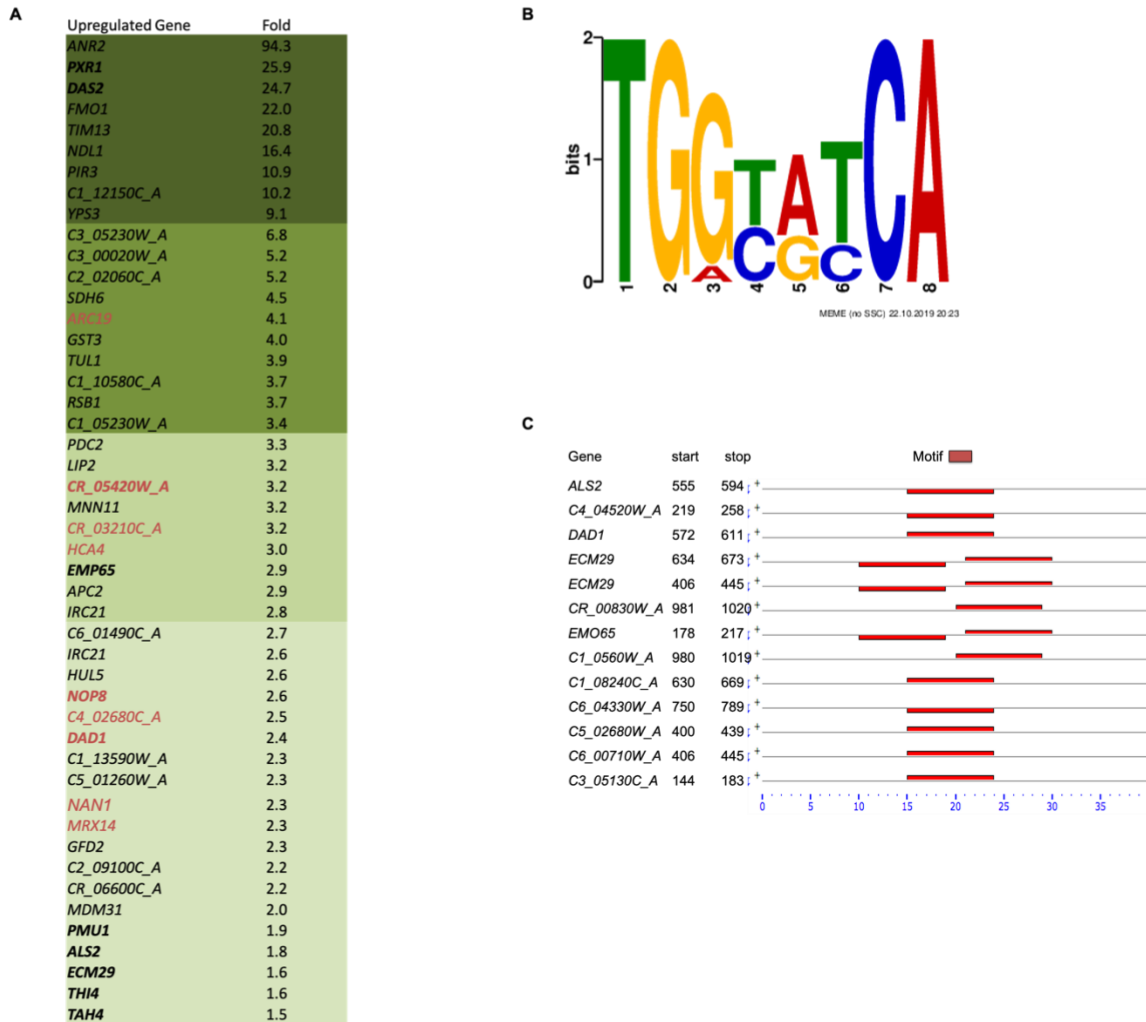


Fig. 4.3.A. Genes upregulated by Met28 activation. *Top 50 upregulated genes generated by activation of the CaMet28 transcription factor. There are 9 essential genes (in red color), 10 genes that have the AnMet28 DNA binding motif 5'TGRYRYCA 3', 23 genes that are annotated as biological function unknown, and 7 genes that encode candidate secreted or cell wall associated proteins (in blue colour). B. AnMet28 DNA binding motif 5'TGRYRYCA 3'. C. The location of the AnMet28 DNA binding motif in the upstream regions of specific upregulated genes.*

4.3.4 Potential Met4 binding sites in *C. albicans*

Alignment of the BZip domains shows the DNA binding region of CaMet4 is somewhat different from that of the CaMet28 and AnMetR/MetZ proteins, suggesting that it may bind

a motif distinct from 5’ ATGRYRYCAT 3’ (Fig 4.1). We used the MEME-suit software and checked the presence of potential motifs in the upstream region of the methionine genes in *Candida*. The two sequences identified with best E values were 5’ GTWGTRGTGG 3’ and the Cbf1 binding site TCACGTG. (Fig 4.4.A). Because ChIP-chip analysis supports Cbf1’s involvement in methionine biosynthesis in *C. albicans* (Lavoie et al., 2010) it is possible that 5’ GTWGTRGTGG 3’represents the binding target of CaMet4. Interestingly, across the different phylogenies studied, while the 5’ ATGRYRYCAT 3’ motif is strongly associated with the Met regulon genes in the filamentous fungi, it becomes much less prevalent in the promoters of the Met regulon genes of *L. elongosporus*, and in species closely related to *C. albicans* or *S. cerevisiae*. In these species where the 5’ ATGRYRYCAT 3’ motif is not strongly associated with the MET regulon, many of the MET regulon promoters have the 5’ GTWGTRGTGG 3’ motif identified in *C. albicans* as a candidate binding motif for Met4. This pattern is consistent with a transfer of the 5’ ATGRYRYCAT 3’/Met28 module away from the Met regulon, and the persistence of the Met4 connection with the Met regulon in *C. albicans* working through the 5’ GTWGTRGTGG 3’ motif (Fig S4.1).

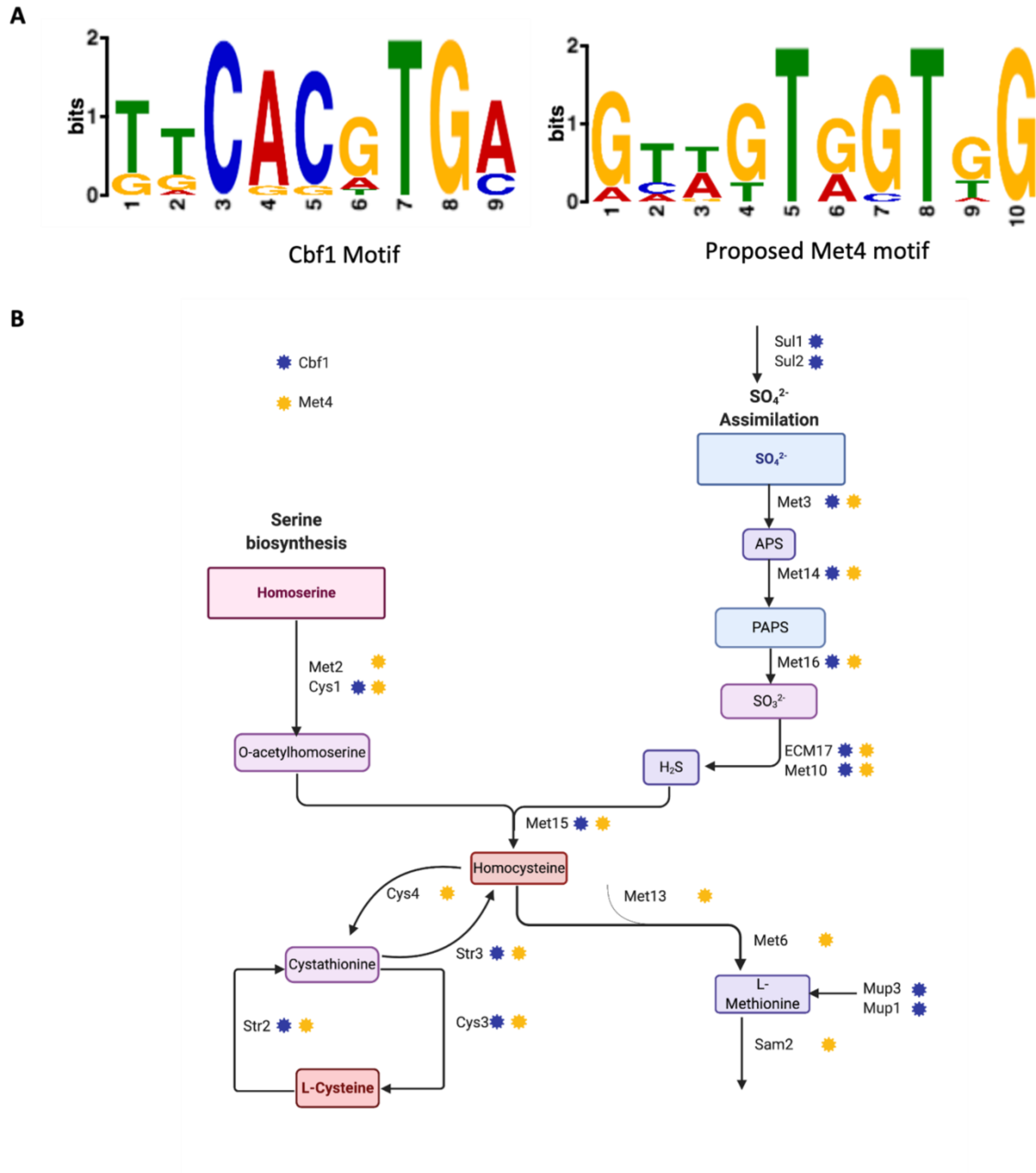


Fig. 4.4.A. Search for the Met4 Motif- *We used the Meme-suit online tool to look for potential regulatory binding sites. This approach identifies the Cbf1 motif in the upstream region of the up-regulated methionine genes. The second-best motif we found is 5'CAACTCCAAR 3' that may represent a Met4 binding motif. B. A diagram highlighting the potential binding motifs of Cbf1 and Met4 throughout the methionine biosynthesis circuit.*

4.3.5 Other transcription factors potentially involved in *C. albicans* methionine biosynthesis

While the regulation of methionine metabolism in Ascomycetes seems primarily controlled by the Met4 class of transcription factors, the regulation of the process in *S. cerevisiae* was found to be more complex. In particular, two transcription factor complexes have been identified to play a role, one centered on the DNA binding protein Cbf1, the other on the DNA binding paralogs Met31 and Met32. Each of these DNA binding elements associates with a pair of key cofactors, the bZIP proteins Met4 and Met28. In *S. cerevisiae* the binding of Met4 and Met28 with Cbf1 involves the interaction of HLH domain of Cbf1 with the bZIP domains of Met28 and Met4 (Kuras et al., 1997).

Previous work had already linked the Cbf1 ortholog in C. *albicans* to methionine regulation (Lavoie et al., 2010) so we investigated whether orthologs of the Met31 and Met32 paralogs of *S. cerevisiae* are present in the genome of *C. albicans*. A single Met32 ortholog is conserved between *S. cerevisiae* and *C. albicans* but versions are not found in the filamentous fungi. Since the DNA binding domain is completely conserved between the two fungi, it could be anticipated that CaMet32 would have same DNA binding motif (5' YYACTGTG 3') as seen for ScMet32. We used the MEME software to search for this 5' YYACTGTG 3' motif within the promoter sequences of the Candida genome. Approximately 400 genes had this motif upstream of their translation start sites; none of these, with the exception of Met4, were methionine biosynthesis genes. This observation suggests that the Met32 transcription factor may be rewired between C. *albicans* and *S. cerevisiae*. The two alleles of *MET32* were deleted from the prototrophic strain SC5314. This *MET32* deletion and did not cause any methionine auxotrophy (Fig 4.2.A), but the deletion strain shows a moderately enhanced growth rate compared to the wild type in enriched media. This might be due to the fact that Met32 binding site was found in many genes of central metabolism like Ade2, Arg1, Dal2, Lys5, and Faa2.

4.4 Discussion

The metabolic pathways for the assimilation of sulfur and the biosynthesis of methionine are extensive, and play a key roles in the central metabolism of most cells (Hebert et al., 2011; Marzluf, 1997; Mendoza-Cozatl et al., 2005; Petti et al., 2012a; Thomas et al., 1995; Thomas and Surdin-Kerjan, 1997). Within the ascomycete fungi, the regulation of the circuitry ranges from relatively simple in filamentous fungi, to quite complex in the baker's yeast. In filamentous fungi it involves a single (or sometimes duplicated) bZIP transcription factor (Marzluf and Metzenberg, 1968). However, in the model yeast *S. cerevisiae*, regulation involves two protein complexes made up of distinct DNA binding components and a set of associating factors (Baker et al., 1989; Baker and Masison, 1990; Bram and Kornberg, 1987; Kaiser et al., 2006; Rouillon et al., 2000). Here we have investigated the regulatory circuitry in the opportunistic fungal pathogen C. *albicans*. Consistent with its phylogenetic position between the filamentous fungi and the bakers' yeast, regulation of methionine biosynthesis in *C. albicans* appears intermediate; more complex than that in the filamentous fungi, but not as complex as that in *S. cerevisiae*.

C. albicans contains two key transcription factors controlling methionine biogenesis, the bHLH factor Cbf1 and the bZIP factor Met4. We had shown that methionine genes including Met4 have a Cbf1 DNA binding motif in their promoters, and confirmed binding to these and other promoters by ChIP-CHIP analysis (Lavoie et al., 2010). Some key genes like *MET2*, *SAH1*, *CYS4 MET13, MET6* and *SAM2* lack Cbf1 binding sites, while the transporter proteins for sulfate (Sul1/2) and methionine (Mup1/3) are under the sole control of Cbf1. Mup1 has an essential role in methionine-induced morphogenesis, biofilm formation and survival inside the macrophages and virulence (Schrevens et al., 2018). We established that deletion of Met4 causes methionine auxotrophy, which suggests that, as found with other fungi, *C. albicans* methionine biosynthesis is dependent on a Met4 ortholog. In addition to their general requirement in normal metabolism, methionine genes are up-regulated in biofilms (Garcia-Sanchez et al., 2004; Uppuluri et al., 2018; Zhu et al., 2013), as well as in planktonic cells derived from biofilms (Uppuluri et al., 2018). Neither

Cbf1 nor Met4 are transcriptionally upregulated during biofilm formation, although the majority of the biofilm up-regulated methionine pathway genes are regulated by either Met4 alone or both Met4 and Cbf1.

CaMet4 shows convincing molecular evidence as a regulator of methionine circuitry, as activation of the protein through the creation of a Met4-VP64 fusion up-regulated many methionine biosynthesis genes, although cysteine biosynthesis (*CYS3* and *CYS4*), and glutathione-associated genes were not affected. We also observed enhanced expression of serine and glycine related genes that directly or indirectly are involved in methionine and glutathione production (Wen et al., 2004). We looked for potential DNA binding motifs in the promoters of the methionine regulon and found 5' GTWGTRGTGG 3' and 5' TCACGTG 3'. Because 5' CACGTG 3' represents the Cbf1 binding target (Lavoie et al., 2010) it is possible that this 5' GTWGTRGTGG 3'motif represents the Met4 target. This motif is quite different from the 5' ATGRYRYCAT 3' motif that represents the binding motif of the Met4 orthologs in the filamentous fungi; it appears that the 5' ATGRYRYCAT 3' motif is disconnected from the *C. albicans* MET regulon, while Met4 remains associated with the *C. albicans* MET regulon but uses the 5' GTWGTRGTGG 3' motif as a binding site.

A paralog of CaMet4, Orf19.7046, has been annotated in *C. albicans* as Met28; this protein shows highest sequence similarity to AnMetR and NcZip1 (Fig 4.1), while its similarity to Met28 of *S. cerevisiae* is limited to its bZIP domain and its very similar size. The *MET28* gene appears essential in *Candida*, and transcription profiling of cells with the activating Met28-VP64 fusion suggests that this TF is not involved in methionine biosynthesis (Fig 4.3). Among the upregulated genes were several essential genes and searching the promoters of the upregulated set identified a candidate binding sequence similar to the motif 5' AGRYRYCAT 3' recognized by AnMetR. Because the AnMet4 and CaMet28 proteins potentially share this DNA binding motif, the genes upregulated by the Met28-VP64 fusion and containing this promoter motif may represent a Met28 controlled regulon. This suggests that the 5' AGRYRYCAT 3' motif remains connected to a Met4-type protein, in this case the paralog Met28, but the motif has been disconnected from the MET regulon. No clear cellular process is evident among these upregulated genes, but because a

large fraction of them have currently undefined functions in *C. albicans*, it is possible further work may uncover some functional relationships. While the activation of the Met28 does not create any obvious morphological or growth phenotype, it is interesting to note that expression of *MET28* is highly upregulated during host infection (Xu et al., 2015). A repositioning of the Met4 ortholog of *S. pombe*, SpZip1, away from control of the methionine regulon has also been observed (Harrison et al., 2005), showing that Met4 orthologs are not uniquely associated with methionine pathway regulation.

We were interested in the potential trajectory of the evolution of the methionine regulatory circuitry from the simple pattern in the filamentous fungi to the more complex arrangement in *S. cerev*isiae. One key step that occurred during this transition is the recruitment of the Cbf1 factor into methionine regulation. The orthologs of Cbf1 in the filamentous fungi like Eurotials, Onygenales and Taphrinomycotina clades do not appear involved in methionine biosynthesis, as connection of the Cbf1 binding motif is found in the promoters of the methionine regulon only in the Candida, Protoploid (*K. lactis*) and Saccharomyces clades (Fig 4.4). In the Candida clade there are therefore two transcription factors (Cbf1 and Met4) co-ordinating the expression of the methionine regulon, but there is no evidence for a direct physical interaction between these two factors. However, they may frequently be found in close proximity on the DNA molecule, as the potential binding sites for the Cbf1 and Met4 proteins in the promoters are often closely adjacent.

The transition from the *C. albicans* framework to that seen in *S. cerevisiae* involves other steps. One is the recruitment of the Met31/32 protein to methionine regulation. In *S. cerevisiae* Met31 and its paralog Met32 are both involved in methionine biosynthesis, although not uniquely, as 455 genes are primarily dependent on Met32 (Petti et al., 2012a). In *Candida* deletion of Met32 does not confer methionine auxotrophy, and although we found almost 400 genes (at a 5% false discovery rate) with the candidate Met32 binding sequence (5’ YYACTGTG 3’) in their promoter, only *MET4* of the sulphur or methionine related genes had this motif. *A. nidulans* also has this sequence in the promoters of many genes, but not methionine genes, suggesting that throughout the ascomycetes Met32 functions as a general TF. In *K. lactis* a few methionine genes like *MET2* and *MET6* have

the Met32 motif in their promoters, while after the whole genome duplication the 5' YYACTGTG 3' sequence becomes extensively linked to methionine biosynthesis.
An additional step in the formation of the *S. cerevisiae* regulatory circuitry is the transition of Met4 from a DNA-binding transcription factor to an activation-domain-containing cofactor that uses Cbf1 and Met31/32 as targeting components. This transition may have been precipitated by the appearance of new sequence in the protein that both disrupted the bZIP DNA binding region of ScMet4 (Hebert et al., 2011) (Fig 4.1), and established protein-protein interactions with Cbf1 and Met31/32. Such a transition could be facilitated by the proximity of Cbf1 and Met4 binding sites on DNA; a weakening of Met4 DNA binding could be compensated by a protein-protein association, thus keeping the functional relationship intact (Met4 and Cbf1 associating with the methionine regulon promoters) but modifying the details of the association (independent DNA binding switching to protein association and use of only one element for DNA binding). bZIP factors are known for their extensive and drastic changes in rewiring and interactions. Previous research suggested that protein-protein interactions (dimer formation) are often associated with DNA binding activity, so modifications in one process frequently affects the other (Reinke et al., 2013). ScMet4 lost dimer-formation interactions resulting in partial DNA-protein interactions. The combination of BHLH and BZip factors is very common in higher eukaryotes; for example, in mammals the transactivation factors Myc (BHLH) and Max (bZIP) show interactions just like the Met4-CBF complex (Vancraenenbroeck and Hofmann, 2018).

There thus appears to be a gradual transition from single protein regulation to a complex circuit in methionine biosynthesis in the ascomycete fungi (Fig 4.5). In filamentous ascomycetes there are no clear orthologs of Cbf1, Met32-like proteins are functioning as general transcription factors, and the methionine circuitry is uniquely controlled by Met4 orthologs. In the Candida clade Met32 is still acting as a general factor, while Met4 and Cbf1 orthologs have been recruited to methionine regulation. In *K. lactis*, more Met related genes have been incorporated into the Met32 control circuit, and this connection is made stronger after the WGD. Finally, in the *S. cerevisiae* lineage, Met4 and another B-zip factor Met28 apparently transition from autonomously acting factors to become transactivating

cofactors through a reduction in DNA binding, and an enhancement in protein-protein associations with the DNA binding elements Cbf1 and Met31/32. It is interesting that throughout these changes the control modules tend to be dual (MetR/MetZ in *A. nidulans*; Met4/Cbf1 in *C. albicans*; Met31/32 and Cbf1 complexes in *S. cerevisiae*), suggesting the complexity of methionine regulation may be best served by more than a simple single transcription regulator.

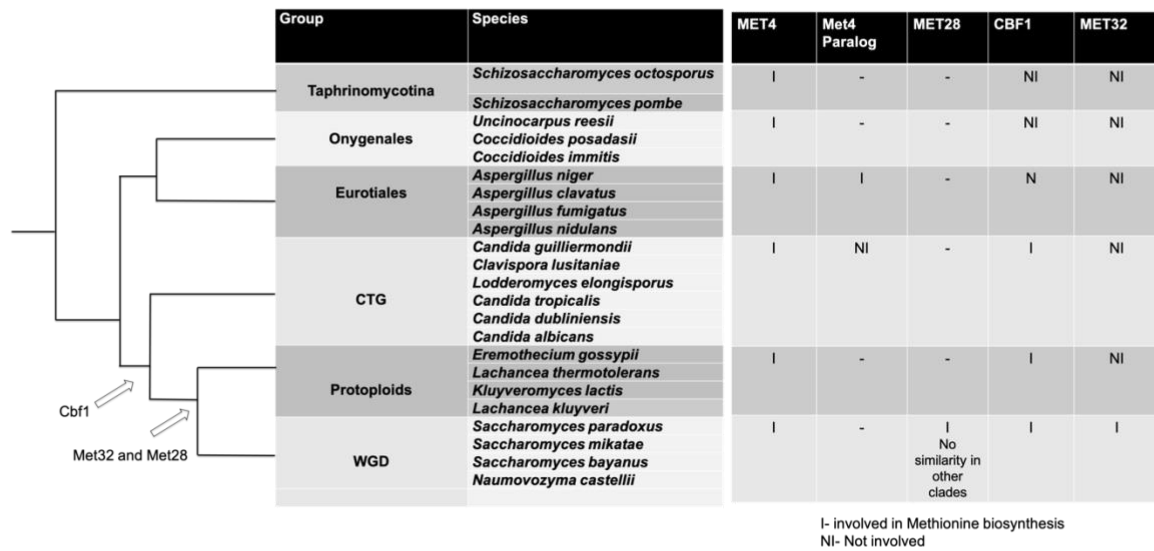


Fig. 4. 5. Phylogenetic analysis of methionine regulatory transcriptional factors across the ascomycetes.

In the phylogenetic tree a Met4 ortholog is involved with the methionine biosynthesis in most of the fungi. Cbf1 was rewired and connects to methionine regulation in the CTG clade, whereas Met32 and Met28 are connected to the pathway later in the protoploids. The designation "I" for involved in methionine biosynthesis, "NI" for not involved, and " - " for not present in the genome, represent the situation for all species in the group identified.

	Cys1	Sam2	Met8	Met16	Str2	Mer3	Met14	Cys4	Met10	Met6	Cys3	Str3	Met2
S.pombe													
S. octosporus													
S. japonicus													
N. crassa													
A. nidulans													
Y. lipolytica													
L. elongosporu													
C. parapsilosis													
C. albicans													
C. tropicalis													
C. guilliermondii													
D. hansenii													
C. lusitaniae													
A. gossypii													
S. kluyveri													
K. lactis													
K. waltii													
S. castellii													
C. glabrata													
S. bayanus													
S. mikatae													
S. paradoxus													
S. cerevisiae													

Supplementary Fig 4.1 *The presence within the promoters of methionine genes across the fungal phylogeny of the candidate Met4 motif 5' ATGRYRYCAT 3' is represented by black boxes and the new potential Met4 recognition motif of the CUG clade 5' CAACTCCAAR 3' is represented with grey boxes. Some genes in some species have both motifs; these cases are represented by a box split into a black and a grey section.*

Chapter 5: Conclusions and Future Work

This thesis has used artificial activation to investigate a group of transcription factors in the human pathogen *Candida albicans*. These factors are either not found in the related non-pathogenic fungus *Saccharomyces cerevisiae*, or likely play different roles in the two fungi. By studying the functions of transcription factors at a large scale, we can identify important information about distinct behaviors and evolved traits. For example, the rewiring of transcription factors may affect critical functions in *C. albicans* or direct the evolution of new functions compared to its highly studied relative the bakers' or brewers' yeast *S. cerevisiae.* Most of the large-scale studies searching for rewiring involve specific *in silico* approaches, while our focus toward the discovery of distinct functioning of transcription factors between these two species was based on a combination of both *in silico* analyses and direct *in vivo* studies. We performed large-scale screening of 30 different activated-transcription-factor strains of *C. albicans* against 9 phenotypic behaviors including morphological changes, carbon source utilization, osmotic and cell wall stress, heavy metal tolerance, environmental factors like temperature and pH, genotoxicity, and drug resistance. Interestingly, we found 5 transcription factors affecting morphology, 8 affecting pH tolerance and 5 temperature tolerance, 4 affecting genotoxicity, 11 affecting adhesions, and 1 resulting in multidrug resistance. The large-scale screening results not only reveals multiple rewired transcription factors compared to the well-known species, *S. cerevisiae,* but also predicted functions for previously uncharacterized transcription factors in *C. albicans*.

We encountered an intriguing rewiring involving the transcription factor Adr1. In *C. albicans* the Adr1 transcription factor functions, in conjunction with the zinc cluster TF Upc2, as a regulator of the ergosterol biosynthetic pathway. We examined the relationship of Adr1 and Upc2 in sterol biosynthesis in *C. albicans*. Phenotypic profiles of either *ADR1* and *UPC2* modulation in *C. albicans* showed differential growth in the presence of fluconazole; either *adr1* and *upc2* homozygous deletion results in sensitivity to the drug while their activation generates a fluconazole resistant strain. In *S. cerevisiae*, Adr1 has

been shown to regulate alcohol and fatty acid metabolism. This rewiring from ergosterol synthesis to fatty acid metabolism involved all members of the Adr1 regulons except the alcohol dehydrogenase Adh2, which remains under Adr1 control in both circuits. This contrasts to the model yeast *S. cerevisiae*, where Adr1 functions in the regulation of alcohol and fatty acid metabolism, while ergosterol biosynthesis remains under the specific control of Upc2. This study shows that the rewiring is directional and not a random event. Comparing the lifestyle of these two fungi, *S. cerevisiae* uses fermentation in the presence of oxygen and at high glucose concentrations (Crabtree-positive) whereas *C. albicans* does not (Crabtree-negative). Therefore, the Crabtree-positive lifestyle of S. *cerevisiae* may need enhanced control over the utilization of alcohols, fatty acids and glycerol and could be a reason of the rewiring of this transcription factor.

Another clear rewiring which we found was that of methionine biosynthesis. This rewiring looks to be a step-wise progression toward complexity, from low complexity in filamentous fungi, intermediate complexity in Candida, and greatest complexity in Saccharomyces, yet we could not establish a clear reason behind this rewiring and currently attribute it to being be a random process. *C. albicans* regulates its methionine biosynthetic pathways using the transcription factors Met4 and Cbf1 which is less complex than that of *S. cerevisiae*, which requires two transcription complexes, involving Cbf1 and Met31/32 as DNA-binding modules, and Met4 and Met28 as co-activators, to regulate Met biosynthesis. In this case Met31/32 and Met28 are rewired to this circuit, as in *C. albicans* they do not appear to be linked to regulation of the methionine metabolism. Thus, the phylogenetic progression from filamentous fungi to *S. cerevisiae* is characterized by increasing complexity of the regulators of this circuit, although the structural elements of the pathways are fundamentally the same. This study showed us that evolution of such networks could be very unpredictable and difficult to figure out the logical aspects of rewiring. Therefore, the answer to the question whether the rewiring is random or not remains an answered. But we were able to contribute a clearer picture of this journey toward the rewiring events in evolution.

A primary objective of this work was to find the functions of currently poorly characterized transcription factors in the human pathogen either based on ortholog comparison with factors in other fungi, particularly *S. cerevisiae*, or by direct assessment of factors unique to *C. albicans*. Among our 30 selected transcription factors we found 4 transcription factors (Met4, Met32, Met28 and Adr1) having clearly different behaviour relative to their orthologs in the budding yeast. We found that in *C. albicans* the Met4 ortholog is also a core regulator of methionine biosynthesis, where it functions together with Cbf1. While *C. albicans* encodes this Met4 protein, a Met4 paralog designated Met28 (orf19.7046), and a Met31 protein, deletion and activation constructs suggest that of these proteins only Met4 is actually involved in regulation of methionine biosynthesis. Both Met28 and Met31 are linked to other functions; Met28 appears essential, and Met32 appears implicated in regulation of genes of central metabolism. Therefore, while *S. cerevisiae* and *C. albicans* share Cbf1 and Met4 as central elements of the methionine biosynthesis control, the other proteins that make up the circuit in *S. cerevisiae* are not members of the *C. albicans* control network, and so the *S. cerevisiae* circuit likely represents a recently evolved arrangement.

As a continuation of this work analysing activated TFs, the transcription factor Ofi1 could be very intriguing. Our data showed that activation of Ofi1 facilitates the morphological switch from white cells to opaque cells. Efg1, Wor1, Wor2 are already known transcription factors which are involved in this type of switch. It will be interesting to decipher the role of Ofi1 in regulation of morphological changes and in the mating process itself. Ofi1 do not have any ortholog in *S cerevisiae* or other close related fungi. It is intriguing how this unique transcription factor plays and important role in morphogenesis. The activation of other transcription factors also has revealed new results for adhesion regulation which can be further investigated for impacts on biofilm formation and in infections. Another interesting study would be Hap42 and Hap43 as they both are similar proteins structurally, yet their activation generated different phenotypes. Also, the Candida-specific transcription factors showed a lot of potential and need further investigation. One experiment which is required for this further study would be Chip-Seq where we can find the unique DNA-binding motifs of these transcription factors.

Overall then the strategy of transcription factor activation has proven an extremely powerful tool for the investigation of regulatory circuits in *C. albicans*. Expanding the phenotypic analyses to include TF binding targets and an expanded set of environmental conditions and investigating in depth those factors connected to new circuitry, will provide a framework for extensive continuing research.

References

- Akins, R.A., and Sobel, J.D. (2009). Antifungal targets, mechanisms of action, and resistance in Candida albicans. In Antimicrobial drug resistance (Springer), pp. 347-407.
- Anderson, J., Mihalik, R., and Soll, D.R. (1990). Ultrastructure and antigenicity of the unique cell wall pimple of the Candida opaque phenotype. J Bacteriol *172*, 224-235.
- Anderson, J.M., and Soll, D.R. (1987). Unique phenotype of opaque cells in the white-opaque transition of Candida albicans. J Bacteriol *169*, 5579-5588.
- Antachopoulos, C., Meletiadis, J., Sein, T., Roilides, E., and Walsh, T.J. (2007). Concentration-dependent effects of caspofungin on the metabolic activity of Aspergillus species. Antimicrobial agents and chemotherapy *51*, 881-887.
- Arendrup, M.C., and Patterson, T.F. (2017). Multidrug-resistant Candida: epidemiology, molecular mechanisms, and treatment. The Journal of infectious diseases *216*, S445-S451.
- Arnaud, M.B., Cerqueira, G.C., Inglis, D.O., Skrzypek, M.S., Binkley, J., Chibucos, M.C., Crabtree, J., Howarth, C., Orvis, J., and Shah, P. (2012). The Aspergillus Genome Database (AspGD): recent developments in comprehensive multispecies curation, comparative genomics and community resources. Nucleic acids research *40*, D653-D659.
- Bailey, T.L., Johnson, J., Grant, C.E., and Noble, W.S. (2015). The MEME suite. Nucleic acids research *43*, W39-W49.
- Baker, R.E., Fitzgerald-Hayes, M., and O'Brien, T.C. (1989). Purification of the yeast centromere binding protein CP1 and a mutational analysis of its binding site. J Biol Chem *264*, 10843-10850.
- Baker, R.E., and Masison, D.C. (1990). Isolation of the gene encoding the Saccharomyces cerevisiae centromere-binding protein CP1. Mol Cell Biol *10*, 2458-2467.
- Barbey, R., Baudouin-Cornu, P., Lee, T.A., Rouillon, A., Zarzov, P., Tyers, M., and Thomas, D. (2005a). Inducible dissociation of SCF(Met30) ubiquitin ligase mediates a rapid transcriptional response to cadmium. EMBO J *24*, 521-532.

- Barbey, R., Baudouin-Cornu, P., Lee, T.A., Rouillon, A., Zarzov, P., Tyers, M., and Thomas, D. (2005b). Inducible dissociation of SCFMet30 ubiquitin ligase mediates a rapid transcriptional response to cadmium. The EMBO journal *24*, 521-532.
- Bardwell, L., Cook, J.G., Zhu-Shimoni, J.X., Voora, D., and Thorner, J. (1998). Differential regulation of transcription: repression by unactivated mitogen-activated protein kinase Kss1 requires the Dig1 and Dig2 proteins. Proc Natl Acad Sci U S A *95*, 15400-15405.
- Beck-Sague, C., and Jarvis, W.R. (1993). Secular trends in the epidemiology of nosocomial fungal infections in the United States, 1980-1990. National Nosocomial Infections Surveillance System. J Infect Dis *167*, 1247-1251.
- Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F. (1998a). Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proceedings of the National Academy of Sciences *95*, 14628-14633.
- Beerli, R.R., Segal, D.J., Dreier, B., and Barbas, C.F., 3rd (1998b). Toward controlling gene expression at will: specific regulation of the erbB-2/HER-2 promoter by using polydactyl zinc finger proteins constructed from modular building blocks. Proc Natl Acad Sci U S A *95*, 14628-14633.
- Berkow, E.L., and Lockhart, S.R. (2017). Fluconazole resistance in Candida species: a current perspective. Infection and drug resistance *10*, 237.
- Bertram, G., Swoboda, R.K., Gooday, G.W., Gow, N.A., and Brown, A.J. (1996). Structure and regulation of the Candida albicans ADH1 gene encoding an immunogenic alcohol dehydrogenase. Yeast *12*, 115-127.
- Bhattacharya, S., Esquivel, B.D., and White, T.C. (2018). Overexpression or deletion of ergosterol biosynthesis genes alters doubling time, response to stress agents, and drug susceptibility in Saccharomyces cerevisiae. MBio *9*.
- Blaiseau, P.L., and Thomas, D. (1998). Multiple transcriptional activation complexes tether the yeast activator Met4 to DNA. EMBO J *17*, 6327-6336.

- Bockmuhl, D.P., and Ernst, J.F. (2001). A potential phosphorylation site for an A-type kinase in the Efg1 regulator protein contributes to hyphal morphogenesis of Candida albicans. Genetics *157*, 1523-1530.
- Bradley, R.K., Li, X.-Y., Trapnell, C., Davidson, S., Pachter, L., Chu, H.C., Tonkin, L.A., Biggin, M.D., and Eisen, M.B. (2010). Binding site turnover produces pervasive quantitative changes in transcription factor binding between closely related Drosophila species. PLoS Biol *8*, e1000343.
- Brajtburg, J., and Bolard, J. (1996). Carrier effects on biological activity of amphotericin B. Clinical microbiology reviews *9*, 512-531.
- Bram, R.J., and Kornberg, R.D. (1987). Isolation of a Saccharomyces cerevisiae centromere DNA-binding protein, its human homolog, and its possible role as a transcription factor. Mol Cell Biol *7*, 403-409.
- Brand, A. (2012). Hyphal growth in human fungal pathogens and its role in virulence. Int J Microbiol *2012*, 517529.
- Breivik, O., and Owades, J. (1957). Yeast analysis, spectrophotometric semimicrodetermination of ergosterol in yeast. Journal of Agricultural and Food Chemistry *5*, 360-363.
- Brown, A.J., Dyos, S.L., Whiteway, M.S., White, J.H., Watson, M.A.E., Marzioch, M., Clare, J.J., Cousens, D.J., Paddon, C., and Plumpton, C. (2000). Functional coupling of mammalian receptors to the yeast mating pathway using novel yeast/mammalian G protein α-subunit chimeras. Yeast *16*, 11-22.
- Brown, A.J., and Gow, N.A. (1999). Regulatory networks controlling Candida albicans morphogenesis. Trends Microbiol *7*, 333-338.
- Buffo, J., Herman, M.A., and Soll, D.R. (1984). A characterization of pH-regulated dimorphism in Candida albicans. Mycopathologia *85*, 21-30.
- Calderone, R.A., and Fonzi, W.A. (2001). Virulence factors of Candida albicans. Trends Microbiol *9*, 327-335.
- Carvunis, A.-R., Wang, T., Skola, D., Yu, A., Chen, J., Kreisberg, J.F., and Ideker, T. (2015). Evidence for a common evolutionary rate in metazoan transcriptional networks. Elife *4*, e11615.

- Chandra, J., Mohammad, S., and Ghannoum, M.A. (2009). Flucytosine: Site of action, mechanism of resistance and use in combination therapy. In Antimicrobial Drug Resistance (Springer), pp. 313-326.
- Chapman, S.W., Dismukes, W.E., Proia, L.A., Bradsher, R.W., Pappas, P.G., Threlkeld, M.G., and Kauffman, C.A. (2008). Clinical practice guidelines for the management of blastomycosis: 2008 update by the Infectious Diseases Society of America. Clinical Infectious Diseases *46*, 1801-1812.
- Chaudhuri, N., Montag, B.J., and Heidelberger, C. (1958). Studies on fluorinated pyrimidines: III. The metabolism of 5-fluorouracil-2-C14 and 5-fluoroorotic-2-C14 acid in vivo. Cancer Research *18*, 318-328.
- Chen, C.-G., Yang, Y.-L., Shih, H.-I., Su, C.-L., and Lo, H.-J. (2004a). CaNdt80 is involved in drug resistance in Candida albicans by regulating CDR1. Antimicrobial Agents and Chemotherapy *48*, 4505-4512.
- Chen, C.G., Yang, Y.L., Shih, H.I., Su, C.L., and Lo, H.J. (2004b). CaNdt80 is involved in drug resistance in Candida albicans by regulating CDR1. Antimicrob Agents Chemother *48*, 4505-4512.
- Chen, C.G., Yang, Y.L., Tseng, K.Y., Shih, H.I., Liou, C.H., Lin, C.C., and Lo, H.J. (2009). Rep1p negatively regulating MDR1 efflux pump involved in drug resistance in Candida albicans. Fungal Genet Biol *46*, 714-720.
- Chen, J., Chen, J., Lane, S., and Liu, H. (2002). A conserved mitogen-activated protein kinase pathway is required for mating in Candida albicans. Mol Microbiol *46*, 1335-1344.
- Chen, L., Xu, Y., Zhou, C., Zhao, J., Li, C., and Wang, R. (2010). Overexpression of CDR1 and CDR2 genes plays an important role in fluconazole resistance in Candida albicans with G487T and T916C mutations. Journal of International Medical Research *38*, 536-545.
- Cheng, G., Yeater, K.M., and Hoyer, L.L. (2006). Cellular and molecular biology of Candida albicans estrogen response. Eukaryotic cell *5*, 180-191.
- Cihlar, R.L., and Calderone, R.A. (2009). Candida albicans: methods and protocols (Springer).

- Consortium, G.O. (2012). The gene ontology: enhancements for 2011. Nucleic acids research *40*, D559-D564.
- Copetti, M.V. (2019). Fungi as industrial producers of food ingredients. Current opinion in food science *25*, 52-56.
- Costa, A.C.B.P., Omran, R.P., Law, C., Dumeaux, V., and Whiteway, M. (2021). Signal-mediated localization of Candida albicans pheromone response pathway components. G3 *11*, jkaa033.
- Coste, A., Turner, V., Ischer, F., Morschhäuser, 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.
- Cowen, L.E., Nantel, A., Whiteway, M.S., Thomas, D.Y., Tessier, D.C., Kohn, L.M., and Anderson, J.B. (2002). Population genomics of drug resistance in Candida albicans. Proc Natl Acad Sci U S A *99*, 9284-9289.
- Csank, C., Schroppel, K., Leberer, E., Harcus, D., Mohamed, O., Meloche, S., Thomas, D.Y., and Whiteway, M. (1998). Roles of the Candida albicans mitogen-activated protein kinase homolog, Cek1p, in hyphal development and systemic candidiasis. Infect Immun *66*, 2713-2721.
- Cullen, P.J., and Sprague, G.F., Jr. (2012). The regulation of filamentous growth in yeast. Genetics *190*, 23-49.
- Czerucka, D., Piche, T., and Rampal, P. Review article: yeast as probiotics–Saccharomyces boulardii. Aliment Pharmacol. 2007; 26 (6): 767-78.
- Czerucka, D., Piche, T., and Rampal, P. (2007). yeast as probiotics–Saccharomyces boulardii. Alimentary pharmacology & therapeutics *26*, 767-778.
- Dalal, C.K., Zuleta, I.A., Mitchell, K.F., Andes, D.R., El-Samad, H., and Johnson, A.D. (2016). Transcriptional rewiring over evolutionary timescales changes quantitative and qualitative properties of gene expression. Elife *5*, e18981.
- Denis, C.L. (1987). The effects of ADR1 and CCR1 gene dosage on the regulation of the glucose-repressible alcohol dehydrogenase from Saccharomyces cerevisiae. Mol Gen Genet *208*, 101-106.

- DeRisi, J.L., Iyer, V.R., and Brown, P.O. (1997). Exploring the metabolic and genetic control of gene expression on a genomic scale. Science *278*, 680-686.
- Diasio, R.B., Bennett, J.E., and Myers, C.E. (1978). Mode of action of 5-fluorocytosine. Biochemical pharmacology *27*, 703-707.
- Dignard, D., Andre, D., and Whiteway, M. (2008). Heterotrimeric G-protein subunit function in Candida albicans: both the alpha and beta subunits of the pheromone response G protein are required for mating. Eukaryot Cell *7*, 1591-1599.
- Dohlman, H.G., Song, J., Apanovitch, D.M., DiBello, P.R., and Gillen, K.M. (1998). Regulation of G protein signalling in yeast. Semin Cell Dev Biol *9*, 135-141.
- Du, H., Li, X., Huang, G., Kang, Y., and Zhu, L. (2015). The zinc-finger transcription factor, Ofi1, regulates white–opaque switching and filamentation in the yeast Candida albicans. Acta biochimica et biophysica Sinica *47*, 335-341.
- Edgar, R., Domrachev, M., and Lash, A.E. (2002). Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res *30*, 207-210.
- Edmond, M.B., Wallace, S.E., McClish, D.K., Pfaller, M.A., Jones, R.N., and Wenzel, R.P. (1999). Nosocomial bloodstream infections in United States hospitals: a three-year analysis. Clin Infect Dis *29*, 239-244.
- El, H. (1950). Two antifungal agents produced by a soil actinomycete. Science (New York, NY) *112*, 423-423.
- Erkenbrack, E.M., and Davidson, E.H. (2015). Evolutionary rewiring of gene regulatory network linkages at divergence of the echinoid subclasses. Proceedings of the National Academy of Sciences *112*, E4075-E4084.
- Everett, R.D., Orr, A., and Preston, C.M. (1998). A viral activator of gene expression functions via the ubiquitin–proteasome pathway. The EMBO Journal *17*, 7161-7169.
- Felk, A., Kretschmar, M., Albrecht, A., Schaller, M., Beinhauer, S., Nichterlein, T., Sanglard, D., Korting, H.C., Schafer, W., and Hube, B. (2002). Candida albicans

hyphal formation and the expression of the Efg1-regulated proteinases Sap4 to Sap6 are required for the invasion of parenchymal organs. Infect Immun *70*, 3689-3700.

- Florence, A.T. (1997). The oral absorption of micro-and nanoparticulates: neither exceptional nor unusual. Pharmaceutical research *14*, 259-266.
- Flowers, S.A., Barker, K.S., Berkow, E.L., Toner, G., Chadwick, S.G., Gygax, S.E., Morschhäuser, J., and Rogers, P.D. (2012). Gain-of-function mutations in UPC2 are a frequent cause of ERG11 upregulation in azole-resistant clinical isolates of Candida albicans. Eukaryotic cell *11*, 1289-1299.
- Flowers, S.A., Colón, B., Whaley, S.G., Schuler, M.A., and Rogers, P.D. (2015). Contribution of clinically derived mutations in ERG11 to azole resistance in Candida albicans. Antimicrobial agents and chemotherapy *59*, 450-460.
- Forsburg, S.L., and Guarente, L. (1989a). Communication between mitochondria and the nucleus in regulation of cytochrome genes in the yeast Saccharomyces cerevisiae. Annual review of cell biology *5*, 153-180.
- Forsburg, S.L., and Guarente, L. (1989b). Identification and characterization of HAP4: a third component of the CCAAT-bound HAP2/HAP3 heteromer. Genes Dev *3*, 1166-1178.
- Gallis, H.A., Drew, R.H., and Pickard, W.W. (1990). Amphotericin B: 30 years of clinical experience. Reviews of infectious diseases *12*, 308-329.
- Garcia-Sanchez, S., Aubert, S., Iraqui, I., Janbon, G., Ghigo, J.M., and d'Enfert, C. (2004). Candida albicans biofilms: a developmental state associated with specific and stable gene expression patterns. Eukaryotic Cell *3*, 536-545.
- Gasch, A.P. (2007). Comparative genomics of the environmental stress response in ascomycete fungi. Yeast *24*, 961-976.
- Ghannoum, M.A., Jurevic, R.J., Mukherjee, P.K., Cui, F., Sikaroodi, M., Naqvi, A., and Gillevet, P.M. (2010). Characterization of the oral fungal microbiome (mycobiome) in healthy individuals. PLoS Pathog *6*, e1000713.
- Ghannoum, M.A., and Rice, L.B. (1999a). Antifungal agents: mode of action, mechanisms of resistance, and correlation of these mechanisms with bacterial resistance. Clin Microbiol Rev *12*, 501-517.

- Ghannoum, M.A., and Rice, L.B. (1999b). Antifungal agents: mode of action, mechanisms of resistance, and correlation of these mechanisms with bacterial resistance. Clinical microbiology reviews *12*, 501-517.
- Ghannoum, M.A., Swairjo, I., and Soll, D.R. (1990). Variation in lipid and sterol contents in Candida albicans white and opaque phenotypes. J Med Vet Mycol *28*, 103-115.
- Goodwin, S.D., Cleary, J.D., Walawander, C.A., Taylor, J.W., and Grasela Jr, T.H. (1995). Pretreatment regimens for adverse events related to infusion of amphotericin B. Clinical infectious diseases *20*, 755-761.
- Guslandi, M., Giollo, P., and Testoni, P.A. (2003). A pilot trial of Saccharomyces boulardii in ulcerative colitis. Eur J Gastroenterol Hepatol *15*, 697-698.
- Guslandi, M., Mezzi, G., Sorghi, M., and Testoni, P.A. (2000). Saccharomyces boulardii in maintenance treatment of Crohn's disease. Dig Dis Sci *45*, 1462-1464.
- Habib, N., Wapinski, I., Margalit, H., Regev, A., and Friedman, N. (2012). A functional selection model explains evolutionary robustness despite plasticity in regulatory networks. Molecular systems biology *8*, 619.
- Harrison, C., Katayama, S., Dhut, S., Chen, D., Jones, N., Bahler, J., and Toda, T. (2005). SCF(Pof1)-ubiquitin and its target Zip1 transcription factor mediate cadmium response in fission yeast. EMBO J *24*, 599-610.
- Hasan, S., Ansari, M.I., Ahmad, A., and Mishra, M. (2015). Major bioactive metabolites from marine fungi: A Review. Bioinformation *11*, 176.
- Hebert, A., Casaregola, S., and Beckerich, J.M. (2011). Biodiversity in sulfur metabolism in hemiascomycetous yeasts. FEMS Yeast Res *11*, 366-378.
- Hedges, S.B., Blair, J.E., Venturi, M.L., and Shoe, J.L. (2004). A molecular timescale of eukaryote evolution and the rise of complex multicellular life. BMC Evol Biol *4*, 2.
- Heeres, J., Backx, L., Mostmans, J., and Van Cutsem, J. (1979). Antimycotic imidazoles. Part 4. Synthesis and antifungal activity of ketoconazole, a new potent orally active broad-spectrum antifungal agent. Journal of medicinal chemistry *22*, 1003-1005.

- Heffer, A., Grubbs, N., Mahaffey, J., and Pick, L. (2013). The evolving role of the orphan nuclear receptor ftz-f1, a pair-rule segmentation gene. Evolution & development *15*, 406-417.
- Henry, K.W., Nickels, J.T., and Edlind, T.D. (2000). Upregulation of ERG Genes inCandida Species by Azoles and Other Sterol Biosynthesis Inhibitors. Antimicrobial agents and chemotherapy *44*, 2693-2700.
- Hoffmann, C., Dollive, S., Grunberg, S., Chen, J., Li, H., Wu, G.D., Lewis, J.D., and Bushman, F.D. (2013). Archaea and fungi of the human gut microbiome: correlations with diet and bacterial residents. PLoS One *8*, e66019.
- Hogan, D.A., and Sundstrom, P. (2009). The Ras/cAMP/PKA signaling pathway and virulence in Candida albicans. Future Microbiol *4*, 1263-1270.
- 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.
- Horn, D.L., Neofytos, D., Anaissie, E.J., Fishman, J.A., Steinbach, W.J., Olyaei, A.J., Marr, K.A., Pfaller, M.A., Chang, C.-H., and Webster, K.M. (2009). Epidemiology and outcomes of candidemia in 2019 patients: data from the prospective antifungal therapy alliance registry. Clinical infectious diseases *48*, 1695-1703.
- Huang, G., Wang, H., Chou, S., Nie, X., Chen, J., and Liu, H. (2006). Bistable expression of WOR1, a master regulator of white-opaque switching in Candida albicans. Proc Natl Acad Sci U S A *103*, 12813-12818.
- Hull, C.M., and Johnson, A.D. (1999). Identification of a mating type-like locus in the asexual pathogenic yeast Candida albicans. Science *285*, 1271-1275.
- Im, E., and Pothoulakis, C. (2010). Recent advances in Saccharomyces boulardii research. Gastroenterologie clinique et biologique *34*, S62-S70.
- 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.
- Jiang, H., Guo, X., Xu, L., and Gu, Z. (2012). Rewiring of posttranscriptional RNA regulons: Puf4p in fungi as an example. Molecular biology and evolution *29*, 2169-2176.

- Jiang, H., Xu, L., Wang, Z., Keene, J., and Gu, Z. (2014). Coordinating expression of RNA binding proteins with their mRNA targets. Scientific reports *4*, 1-5.
- Johnson, A.D. (2017a). The rewiring of transcription circuits in evolution. Current opinion in genetics & development *47*, 121-127.
- Johnson, A.D. (2017b). The rewiring of transcription circuits in evolution. Curr Opin Genet Dev *47*, 121-127.
- Johnson, D.C., Cano, K.E., Kroger, E.C., and McNabb, D.S. (2005). Novel regulatory function for the CCAAT-binding factor in Candida albicans. Eukaryot Cell *4*, 1662-1676.
- K. Redhu, A., Shah, A.H., and Prasad, R. (2016). MFS transporters of Candida species and their role in clinical drug resistance. FEMS yeast research *16*, fow043.
- Kaiser, P., Su, N.Y., Yen, J.L., Ouni, I., and Flick, K. (2006). The yeast ubiquitin ligase SCFMet30: connecting environmental and intracellular conditions to cell division. Cell Div *1*, 16.
- Keene, J.D. (2007). RNA regulons: coordination of post-transcriptional events. Nature Reviews Genetics *8*, 533-543.
- Kelly, S.L., Arnoldi, A., and Kelly, D.E. (1993). Molecular genetic analysis of azole antifungal mode of action. Biochem Soc Trans *21*, 1034-1038.
- Kennedy, M.J., Rogers, A.L., Hanselmen, L.R., Soll, D.R., and Yancey, R.J., Jr. (1988). Variation in adhesion and cell surface hydrophobicity in Candida albicans white and opaque phenotypes. Mycopathologia *102*, 149-156.
- Kummerfeld, S.K., and Teichmann, S.A. (2006). DBD: a transcription factor prediction database. Nucleic Acids Res *34*, D74-81.
- Kuras, L., Barbey, R., and Thomas, D. (1997). Assembly of a bZIP-bHLH transcription activation complex: formation of the yeast Cbf1-Met4-Met28 complex is regulated through Met28 stimulation of Cbf1 DNA binding. EMBO J *16*, 2441-2451.
- Kuras, L., Rouillon, A., Lee, T., Barbey, R., Tyers, M., and Thomas, D. (2002). Dual regulation of the met4 transcription factor by ubiquitin-dependent degradation and inhibition of promoter recruitment. Mol Cell *10*, 69-80.

- Kurjan, J. (1993). The pheromone response pathway in Saccharomyces cerevisiae. Annu Rev Genet *27*, 147-179.
- Lascaris, R., Bussemaker, H.J., Boorsma, A., Piper, M., van der Spek, H., Grivell, L., and Blom, J. (2003). Hap4p overexpression in glucose-grown Saccharomyces cerevisiae induces cells to enter a novel metabolic state. Genome Biol *4*, R3.
- Lavoie, H., Hogues, H., Mallick, J., Sellam, A., Nantel, A., and Whiteway, M. (2010). Evolutionary tinkering with conserved components of a transcriptional regulatory network. PLoS Biol *8*, e1000329.
- Lee, T.A., Jorgensen, P., Bognar, A.L., Peyraud, C., Thomas, D., and Tyers, M. (2010). Dissection of combinatorial control by the Met4 transcriptional complex. Mol Biol Cell *21*, 456-469.
- Li, H., and Johnson, A.D. (2010). Evolution of transcription networks—lessons from yeasts. Current Biology *20*, R746-R753.
- Li, Q., and Marzluf, G.A. (1996). Determination of the Neurospora crassa CYS 3 sulfur regulatory protein consensus DNA-binding site: amino-acid substitutions in the CYS3 bZIP domain that alter DNA-binding specificity. Curr Genet *30*, 298-304.
- Liang, S.D., Marmorstein, R., Harrison, S.C., and Ptashne, M. (1996). DNA sequence preferences of GAL4 and PPR1: how a subset of Zn2 Cys6 binuclear cluster proteins recognizes DNA. Molecular and cellular biology *16*, 3773-3780.
- Limon, J.J., Skalski, J.H., and Underhill, D.M. (2017). Commensal Fungi in Health and Disease. Cell Host Microbe *22*, 156-165.
- Lind, A.L., Wisecaver, J.H., Smith, T.D., Feng, X., Calvo, A.M., and Rokas, A. (2015). Examining the evolution of the regulatory circuit controlling secondary metabolism and development in the fungal genus Aspergillus. PLoS Genet *11*, e1005096.
- Liu, H. (2001). Transcriptional control of dimorphism in Candida albicans. Curr Opin Microbiol *4*, 728-735.
- Liu, H., Kohler, J., and Fink, G.R. (1994). Suppression of hyphal formation in Candida albicans by mutation of a STE12 homolog. Science *266*, 1723-1726.

- 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.
- Lohberger, A., Coste, A.T., and Sanglard, D. (2014). Distinct roles of Candida albicans drug resistance transcription factors TAC1, MRR1, and UPC2 in virulence. Eukaryotic Cell *13*, 127-142.
- Lopez-Ribot, J.L., McAtee, R.K., Lee, L.N., Kirkpatrick, W.R., White, T.C., Sanglard, D., and Patterson, T.F. (1998). Distinct patterns of gene expression associated with development of fluconazole resistance in serial candida albicans isolates from human immunodeficiency virus-infected patients with oropharyngeal candidiasis. Antimicrob Agents Chemother *42*, 2932-2937.
- Lu, H., Shrivastava, M., Whiteway, M., and Jiang, Y. (2021). Candida albicans targets that potentially synergize with fluconazole. Critical Reviews in Microbiology, 1-15.
- Lu, H., Sun, Y., Jiang, Y.Y., and Whiteway, M. (2014). Ste18p is a positive control element in the mating process of Candida albicans. Eukaryot Cell *13*, 461-469.
- Ma, M., and Liu, Z.L. (2010). Mechanisms of ethanol tolerance in Saccharomyces cerevisiae. Applied microbiology and biotechnology *87*, 829-845.
- MacNeil, L.T., and Walhout, A.J. (2011). Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. Genome research *21*, 645-657.
- Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013a). CRISPR RNA-guided activation of endogenous human genes. Nat Methods *10*, 977-979.
- Maeder, M.L., Linder, S.J., Cascio, V.M., Fu, Y., Ho, Q.H., and Joung, J.K. (2013b). CRISPR RNA–guided activation of endogenous human genes. Nature methods *10*, 977-979.
- Mallick, J., and Whiteway, M. (2013). The evolutionary rewiring of the ribosomal protein transcription pathway modifies the interaction of transcription factor heteromer Ifh1-Fhl1 (interacts with forkhead 1-forkhead-like 1) with the DNA-binding specificity element. Journal of Biological Chemistry *288*, 17508-17519.

- Mantovani, R. (1999). The molecular biology of the CCAAT-binding factor NF-Y. Gene *239*, 15-27.
- Martchenko, M., Levitin, A., Hogues, H., Nantel, A., and Whiteway, M. (2007). Transcriptional rewiring of fungal galactose-metabolism circuitry. Current Biology *17*, 1007-1013.
- Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal *17*, 10-12.
- Marzluf, G.A. (1997). Molecular genetics of sulfur assimilation in filamentous fungi and yeast. Annu Rev Microbiol *51*, 73-96.
- Marzluf, G.A., and Metzenberg, R.L. (1968). Positive control by the cys-3 locus in regulation of sulfur metabolism in Neurospora. J Mol Biol *33*, 423-437.
- Mata, J., Marguerat, S., and Bähler, J. (2005). Post-transcriptional control of gene expression: a genome-wide perspective. Trends in biochemical sciences *30*, 506-514.
- McFarland, L.V. (2010). Systematic review and meta-analysis of Saccharomyces boulardii in adult patients. World journal of gastroenterology: WJG *16*, 2202.
- Mendoza-Cozatl, D., Loza-Tavera, H., Hernandez-Navarro, A., and Moreno-Sanchez, R. (2005). Sulfur assimilation and glutathione metabolism under cadmium stress in yeast, protists and plants. FEMS Microbiol Rev *29*, 653-671.
- Monge, R.A., Roman, E., Nombela, C., and Pla, J. (2006). The MAP kinase signal transduction network in Candida albicans. Microbiology *152*, 905-912.
- Moore, A.D., and Bornberg-Bauer, E. (2012). The dynamics and evolutionary potential of domain loss and emergence. Molecular biology and evolution *29*, 787-796.
- Morschhauser, J. (2002). The genetic basis of fluconazole resistance development in Candida albicans. Biochim Biophys Acta *1587*, 240-248.
- 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.
- Morschhäuser, J., Barker, K.S., Liu, T.T., Blaß-Warmuth, 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.

- Nakai, T., Uno, J., Ikeda, F., Tawara, S., Nishimura, K., and Miyaji, M. (2003). In vitro antifungal activity of micafungin (FK463) against dimorphic fungi: comparison of yeast-like and mycelial forms. Antimicrobial Agents and Chemotherapy *47*, 1376-1381.
- Ndisdang, D., Morris, P.J., Chapman, C., Ho, L., Singer, A., and Latchman, D.S. (1998). The HPV-activating cellular transcription factor Brn-3a is overexpressed in CIN3 cervical lesions. J Clin Invest *101*, 1687-1692.
- Niimi, K., Maki, K., Ikeda, F., Holmes, A., Lamping, E., Niimi, M., Monk, B., and Cannon, R. (2006). Overexpression of Candida albicans CDR1, CDR2, or MDR1 does not produce significant changes in echinocandin susceptibility. Antimicrobial agents and chemotherapy *50*, 1148-1155.
- Niimi, M., Niimi, K., Takano, Y., Holmes, A.R., Fischer, F.J., Uehara, Y., and Cannon, R.D. (2004). Regulated overexpression of CDR1 in Candida albicans confers multidrug resistance. Journal of Antimicrobial Chemotherapy *54*, 999-1006.
- Nobile, C.J., Nett, J.E., Hernday, A.D., Homann, O.R., Deneault, J.S., Nantel, A., Andes, D.R., Johnson, A.D., and Mitchell, A.P. (2009). Biofilm matrix regulation by Candida albicans Zap1. PLoS Biol *7*, e1000133.
- Odds, F.C. (1987). Candida infections: an overview. Crit Rev Microbiol *15*, 1-5.
- Oliveira-Pacheco, J., Alves, R., Costa-Barbosa, A., Cerqueira-Rodrigues, B., Pereira-Silva, P., Paiva, S., Silva, S., Henriques, M., Pais, C., and Sampaio, P. (2018). The Role of Candida albicans Transcription Factor RLM1 in Response to Carbon Adaptation. Front Microbiol *9*, 1127.
- Ordiz, M.I., Barbas, C.F., 3rd, and Beachy, R.N. (2002a). Regulation of transgene expression in plants with polydactyl zinc finger transcription factors. Proc Natl Acad Sci U S A *99*, 13290-13295.
- Ordiz, M.I., Barbas, C.F., and Beachy, R.N. (2002b). Regulation of transgene expression in plants with polydactyl zinc finger transcription factors. Proceedings of the National Academy of Sciences *99*, 13290-13295.

- Paddon, C.J., and Keasling, J.D. (2014). Semi-synthetic artemisinin: a model for the use of synthetic biology in pharmaceutical development. Nature Reviews Microbiology *12*, 355-367.
- Pande, K., Chen, C., and Noble, S.M. (2013). Passage through the mammalian gut triggers a phenotypic switch that promotes Candida albicans commensalism. Nat Genet *45*, 1088-1091.
- 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. (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. Antimicrob Agents Chemother *45*, 2676-2684.
- Pertea, M., Pertea, G.M., Antonescu, C.M., Chang, T.C., Mendell, J.T., and Salzberg, S.L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. Nat Biotechnol *33*, 290-295.
- Petti, A.A., McIsaac, R.S., Ho-Shing, O., Bussemaker, H.J., and Botstein, D. (2012a). Combinatorial control of diverse metabolic and physiological functions by transcriptional regulators of the yeast sulfur assimilation pathway. Mol Biol Cell *23*, 3008-3024.
- Petti, A.A., McIsaac, R.S., Ho-Shing, O., Bussemaker, H.J., and Botstein, D. (2012b). Combinatorial control of diverse metabolic and physiological functions by transcriptional regulators of the yeast sulfur assimilation pathway. Molecular biology of the cell *23*, 3008-3024.
- Pfaller, M.A., and Diekema, D. (2002). Role of sentinel surveillance of candidemia: trends in species distribution and antifungal susceptibility. Journal of Clinical microbiology *40*, 3551-3557.
- Pilsyk, S., Natorff, R., Sienko, M., Skoneczny, M., Paszewski, A., and Brzywczy, J. (2015). The Aspergillus nidulans metZ gene encodes a transcription factor involved in regulation of sulfur metabolism in this fungus and other Eurotiales. Curr Genet *61*, 115-125.

- Pound, M.W., Townsend, M.L., Dimondi, V., Wilson, D., and Drew, R.H. (2011). Overview of treatment options for invasive fungal infections. Medical mycology *49*, 561-580.
- Prasad, R. (1993). Candida albicans: cellular and molecular biology. Revista do Instituto de Medicina Tropical de São Paulo *35*, 372-372.
- Prasad, R. (2017). Candida albicans (Springer).
- Prasad, R., De Wergifosse, P., Goffeau, A., and Balzi, E. (1995a). Molecular cloning and characterization of a novel gene of Candida albicans, CDR1, conferring multiple resistance to drugs and antifungals. Current genetics *27*, 320-329.
- Prasad, R., and Ghannoum, M.A. (1996). Lipids of pathogenic fungi (CRC press).
- Prasad, R., and Kapoor, K. (2005). Multidrug resistance in yeast Candida. International review of cytology *242*, 215-248.
- Prasad, R., Murthy, S.K., Gupta, V., and Prasad, R. (1995b). Multiple drug resistance in Candida albicans. Acta Biochim Pol *42*, 497-504.
- Prasad, R., and Singh, A. (2013). Lipids of Candida albicans and their role in multidrug resistance. Curr Genet *59*, 243-250.
- Ramírez, M.A., and Lorenz, M.C. (2009). The transcription factor homolog CTF1 regulates β-oxidation in Candida albicans. Eukaryotic cell *8*, 1604-1614.
- Regan, H., Scaduto, C.M., Hirakawa, M.P., Gunsalus, K., Correia-Mesquita, T.O., Sun, Y., Chen, Y., Kumamoto, C.A., Bennett, R.J., and Whiteway, M. (2017a). Negative regulation of filamentous growth in Candida albicans by Dig1p. Mol Microbiol *105*, 810-824.
- Reinke, A.W., Baek, J., Ashenberg, O., and Keating, A.E. (2013). Networks of bZIP protein-protein interactions diversified over a billion years of evolution. Science *340*, 730-734.
- Ro, D.K., Ouellet, M., Paradise, E.M., Burd, H., Eng, D., Paddon, C.J., Newman, J.D., and Keasling, J.D. (2008). Induction of multiple pleiotropic drug resistance genes in yeast engineered to produce an increased level of anti-malarial drug precursor, artemisinic acid. BMC Biotechnol *8*, 83.

- Roncero, C., and Duran, A. (1985). Effect of Calcofluor white and Congo red on fungal cell wall morphogenesis: in vivo activation of chitin polymerization. J Bacteriol *163*, 1180-1185.
- Rouillon, A., Barbey, R., Patton, E.E., Tyers, M., and Thomas, D. (2000). Feedback-regulated degradation of the transcriptional activator Met4 is triggered by the SCF(Met30)complex. EMBO J *19*, 282-294.
- Ruiz-Herrera, J., Elorza, M.V., Valentin, E., and Sentandreu, R. (2006). Molecular organization of the cell wall of Candida albicans and its relation to pathogenicity. FEMS Yeast Res *6*, 14-29.
- Sadowski, I., Ma, J., Triezenberg, S., and Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. Nature *335*, 563-564.
- Saleh, A.A., Eid, Y.Z., Ebeid, T.A., Amber, K., Kamizono, T., Ohtsuka, A., and Hayashi, K. (2011a). Aspergillus awamori as probiotic in broiler chickens. In Proceedings of the 9th Asia Pacific Poultry Conference, Taipei, Taiwan.
- Saleh, A.A., Eid, Y.Z., Ebeid, T.A., Kamizono, T., Ohtsuka, A., and Hayashi, K. (2011b). Effects of feeding Aspergillus awamori and Aspergillus niger on growth performance and meat quality in broiler chickens. The journal of poultry science *48*, 201-206.
- Sanglard, D., and Odds, F.C. (2002). Resistance of Candida species to antifungal agents: molecular mechanisms and clinical consequences. The Lancet infectious diseases *2*, 73-85.
- Sanglard, D., and White, T.C. (2006). Molecular principles of antifungal drug resistance. Molecular principles of fungal pathogenesis, 197-212.
- Sarda, S., and Hannenhalli, S. (2015). High-throughput identification of cis-regulatory rewiring events in yeast. Molecular biology and evolution *32*, 3047-3063.
- Schillig, R., and Morschhäuser, J. (2013). Analysis of a fungus-specific transcription factor family, the C andida albicans zinc cluster proteins, by artificial activation. Molecular microbiology *89*, 1003-1017.
- Schrevens, S., Van Zeebroeck, G., Riedelberger, M., Tournu, H., Kuchler, K., and Van Dijck, P. (2018). Methionine is required for cAMP-PKA mediated

morphogenesis and virulence of Candida albicans (vol 108, pg 258, 2018). Molecular Microbiology *109*, 415-416.

- Segal, E.S., Gritsenko, V., Levitan, A., Yadav, B., Dror, N., Steenwyk, J.L., Silberberg, Y., Mielich, K., Rokas, A., Gow, N.A.R., *et al.* (2018). Gene Essentiality Analyzed by In Vivo Transposon Mutagenesis and Machine Learning in a Stable Haploid Isolate of Candida albicans. mBio *9*.
- Sellam, A., Tebbji, F., and Nantel, A. (2009). Role of Ndt80p in sterol metabolism regulation and azole resistance in Candida albicans. Eukaryotic cell *8*, 1174-1183.
- Shaw, W.M., Yamauchi, H., Mead, J., Gowers, G.-O.F., Bell, D.J., Öling, D., Larsson, N., Wigglesworth, M., Ladds, G., and Ellis, T. (2019). Engineering a model cell for rational tuning of GPCR signaling. Cell *177*, 782-796. e727.
- Sheehan, D.J., Hitchcock, C.A., and Sibley, C.M. (1999). Current and emerging azole antifungal agents. Clinical microbiology reviews *12*, 40-79.
- Shelest, E. (2017). Transcription Factors in Fungi: TFome Dynamics, Three Major Families, and Dual-Specificity TFs. Front Genet *8*, 53.
- Shrivastava, M., Feng, J., Coles, M., Clark, B., Islam, A., Dumeaux, V., and Whiteway, M. (2021). Modulation of the complex regulatory network for methionine biosynthesis in fungi. Genetics.
- Si, H., Hernday, A.D., Hirakawa, M.P., Johnson, A.D., and Bennett, R.J. (2013). Candida albicans white and opaque cells undergo distinct programs of filamentous growth. PLoS Pathog *9*, e1003210.
- Siggers, T., Duyzend, M.H., Reddy, J., Khan, S., and Bulyk, M.L. (2011). Non-DNA-binding cofactors enhance DNA-binding specificity of a transcriptional regulatory complex. Mol Syst Biol *7*, 555.
- Silver, P.M., Oliver, B.G., and White, T.C. (2004). Role of Candida albicans transcription factor Upc2p in drug resistance and sterol metabolism. Eukaryotic Cell *3*, 1391-1397.
- Slutsky, B., Staebell, M., Anderson, J., Risen, L., Pfaller, M., and Soll, D.R. (1987). "White-opaque transition": a second high-frequency switching system in Candida albicans. J Bacteriol *169*, 189-197.

- Soll, D.R. (2014). The role of phenotypic switching in the basic biology and pathogenesis of Candida albicans. J Oral Microbiol *6*.
- Sonneborn, A., Bockmuhl, D.P., and Ernst, J.F. (1999). Chlamydospore formation in Candida albicans requires the Efg1p morphogenetic regulator. Infect Immun *67*, 5514-5517.
- Srikantha, T., Daniels, K.J., Pujol, C., Sahni, N., Yi, S., and Soll, D.R. (2012). Nonsex genes in the mating type locus of Candida albicans play roles in a/alpha biofilm formation, including impermeability and fluconazole resistance. PLoS Pathog *8*, e1002476.
- Staib, P., Wirsching, S., Strauss, A., and Morschhauser, J. (2001). Gene regulation and host adaptation mechanisms in Candida albicans. Int J Med Microbiol *291*, 183-188.
- Stajich, J.E., Berbee, M.L., Blackwell, M., Hibbett, D.S., James, T.Y., Spatafora, J.W., and Taylor, J.W. (2009). The fungi. Curr Biol *19*, R840-845.
- Stevens, V., Geiger, K., Concannon, C., Nelson, R.E., Brown, J., and Dumyati, G. (2014). Inpatient costs, mortality and 30-day re-admission in patients with central-line-associated bloodstream infections. Clin Microbiol Infect *20*, O318-324.
- Stoldt, V.R., Sonneborn, A., Leuker, C.E., and Ernst, J.F. (1997). Efg1p, an essential regulator of morphogenesis of the human pathogen Candida albicans, is a member of a conserved class of bHLH proteins regulating morphogenetic processes in fungi. EMBO J *16*, 1982-1991.
- Sudbery, P.E. (2011). Growth of Candida albicans hyphae. Nat Rev Microbiol *9*, 737-748.
- Suetens, C., Hopkins, S., Kolman, J., and Hogberg, L. (2013). European Centre for Disease Prevention and Control: point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals. Stockholm: European Centre for Disease Prevention and Control.
- Tachibana, C., Yoo, J.Y., Tagne, J.-B., Kacherovsky, N., Lee, T.I., and Young, E.T. (2005). Combined global localization analysis and transcriptome data identify genes that are directly coregulated by Adr1 and Cat8. Molecular and cellular biology *25*, 2138-2146.

- Tao, L., Du, H., Guan, G., Dai, Y., Nobile, C.J., Liang, W., Cao, C., Zhang, Q., Zhong, J., and Huang, G. (2014). Discovery of a "white-gray-opaque" tristable phenotypic switching system in candida albicans: roles of non-genetic diversity in host adaptation. PLoS Biol *12*, e1001830.
- Tebung, W.A., Choudhury, B.I., Tebbji, F., Morschhauser, J., and Whiteway, M. (2016). Rewiring of the Ppr1 Zinc Cluster Transcription Factor from Purine Catabolism to Pyrimidine Biogenesis in the Saccharomycetaceae. Curr Biol *26*, 1677-1687.
- Tebung, W.A., Omran, R.P., Fulton, D.L., Morschhauser, J., and Whiteway, M. (2017). Put3 Positively Regulates Proline Utilization in Candida albicans. mSphere *2*.
- Thomas, D., Kuras, L., Barbey, R., Cherest, H., Blaiseau, P.L., and Surdin-Kerjan, Y. (1995). Met30p, a yeast transcriptional inhibitor that responds to S-adenosylmethionine, is an essential protein with WD40 repeats. Mol Cell Biol *15*, 6526-6534.
- Thomas, D., and Surdin-Kerjan, Y. (1997). Metabolism of sulfur amino acids in Saccharomyces cerevisiae. Microbiol Mol Biol Rev *61*, 503-532.
- Thukral, S.K., Eisen, A., and Young, E.T. (1991). Two monomers of yeast transcription factor ADR1 bind a palindromic sequence symmetrically to activate ADH2 expression. Molecular and cellular biology *11*, 1566-1577.
- Tortorano, A., Peman, J., Bernhardt, H., Klingspor, L., Kibbler, C., Faure, O., Biraghi, E., Canton, E., Zimmermann, K., and Seaton, S. (2004). Epidemiology of candidaemia in Europe: results of 28-month European Confederation of Medical Mycology (ECMM) hospital-based surveillance study. European Journal of Clinical Microbiology and Infectious Diseases *23*, 317-322.
- Tortorano, A., Prigitano, A., Lazzarini, C., Passera, M., Deiana, M., Cavinato, S., De Luca, C., Grancini, A., Cascio, G.L., and Ossi, C. (2013). A 1-year prospective survey of candidemia in Italy and changing epidemiology over one decade. Infection *41*, 655-662.
- Tuch, B.B., Galgoczy, D.J., Hernday, A.D., Li, H., and Johnson, A.D. (2008). The evolution of combinatorial gene regulation in fungi. PLoS Biol *6*, e38.

- Uppuluri, P., Zaldivar, M.A., Anderson, M.Z., Dunn, M.J., Berman, J., Ribot, J.L.L., and Kohler, J.R. (2018). Candida albicans Dispersed Cells Are Developmentally Distinct from Biofilm and Planktonic Cells. Mbio *9*.
- van Woerden, H.C., Gregory, C., Brown, R., Marchesi, J.R., Hoogendoorn, B., and Matthews, I.P. (2013). Differences in fungi present in induced sputum samples from asthma patients and non-atopic controls: a community based case control study. BMC Infect Dis *13*, 69.
- Vancraenenbroeck, R., and Hofmann, H. (2018). Occupancies in the DNA-binding pathways of intrinsically disordered helix-loop-helix leucine-zipper proteins. The Journal of Physical Chemistry B *122*, 11460-11467.
- Vasicek, E.M., Berkow, E.L., Flowers, S.A., Barker, K.S., and Rogers, P.D. (2014). UPC2 is universally essential for azole antifungal resistance in Candida albicans. Eukaryotic cell *13*, 933-946.
- Vertut-Croquin, A., Bolard, J., and Gary-Bobo, C. (1985). Transfer of amphotericin B from gel state vesicles to mycoplasma cells: biphasic action on potassium transport and permeability. Antimicrobial agents and chemotherapy *28*, 167-171.
- Vyas, V.K., Barrasa, M.I., and Fink, G.R. (2015). A Candida albicans CRISPR system permits genetic engineering of essential genes and gene families. Sci Adv *1*, e1500248.
- Vyas, V.K., Bushkin, G.G., Bernstein, D.A., Getz, M.A., Sewastianik, M., Barrasa, M.I., Bartel, D.P., and Fink, G.R. (2018). New CRISPR mutagenesis strategies reveal variation in repair mechanisms among fungi. MSphere *3*, e00154-00118.
- Waldorf, A.R., and Polak, A. (1983). Mechanisms of action of 5-fluorocytosine. Antimicrobial agents and chemotherapy *23*, 79-85.
- Wang, B., Huang, L.-H., Zhao, J.-X., Wei, M., Fang, H., Wang, D.-Y., Wang, H.-F., Yin, J.-G., and Xiang, M. (2015). ERG11 mutations associated with azole resistance in Candida albicans isolates from vulvovaginal candidosis patients. Asian Pacific Journal of Tropical Biomedicine *5*, 909-914.
- Wapinski, I., Pfeffer, A., Friedman, N., and Regev, A. (2007). Automatic genome-wide reconstruction of phylogenetic gene trees. Bioinformatics *23*, i549-i558.

- Weatheritt, R.J., and Gibson, T.J. (2012). Linear motifs: lost in (pre) translation. Trends in biochemical sciences *37*, 333-341.
- Wen, S., Zhang, T., and Tan, T. (2004). Utilization of amino acids to enhance glutathione production in Saccharomyces cerevisiae. Enzyme and Microbial Technology *35*, 501-507.
- White, T.C., Marr, K.A., and Bowden, R.A. (1998). Clinical, cellular, and molecular factors that contribute to antifungal drug resistance. Clinical microbiology reviews *11*, 382-402.
- Wirsching, S., Michel, S., and Morschhäuser, J. (2000). Targeted gene disruption in Candida albicans wild-type strains: the role of the MDR1 gene in fluconazole resistance of clinical Candida albicans isolates. Molecular microbiology *36*, 856-865.
- 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. Clinical infectious diseases *39*, 309-317.
- Wisplinghoff, H., Ebbers, J., Geurtz, L., Stefanik, D., Major, Y., Edmond, M.B., Wenzel, R.P., and Seifert, H. (2014). Nosocomial bloodstream infections due to Candida spp. in the USA: species distribution, clinical features and antifungal susceptibilities. International journal of antimicrobial agents *43*, 78-81.
- Woodson, J.D., and Chory, J. (2008). Coordination of gene expression between organellar and nuclear genomes. Nature Reviews Genetics *9*, 383-395.
- Wright, W.L., and Wenzel, R.P. (1997). Nosocomial Candida. Epidemiology, transmission, and prevention. Infect Dis Clin North Am *11*, 411-425.
- Xiang, M.-J., Liu, J.-Y., Ni, P.-H., Wang, S., Shi, C., Wei, B., Ni, Y.-X., and Ge, H.-L. (2013). Erg11 mutations associated with azole resistance in clinical isolates of Candida albicans. FEMS yeast research *13*, 386-393.
- Xie, J., Tao, L., Nobile, C.J., Tong, Y., Guan, G., Sun, Y., Cao, C., Hernday, A.D., Johnson, A.D., Zhang, L., *et al.* (2013). White-opaque switching in natural MTLa/alpha isolates of Candida albicans: evolutionary implications for roles in host adaptation, pathogenesis, and sex. PLoS Biol *11*, e1001525.

- Xu, W., Solis, N.V., Ehrlich, R.L., Woolford, C.A., Filler, S.G., and Mitchell, A.P. (2015). Activation and alliance of regulatory pathways in C. albicans during mammalian infection. PLoS Biol *13*, e1002076.
- Yang, H., Tong, J., Lee, C.W., Ha, S., Eom, S.H., and Im, Y.J. (2015). Structural mechanism of ergosterol regulation by fungal sterol transcription factor Upc2. Nat Commun *6*, 6129.
- Yen, J.L., Su, N.-Y., and Kaiser, P. (2005). The yeast ubiquitin ligase SCFMet30 regulates heavy metal response. Molecular biology of the cell *16*, 1872-1882.
- Young, E.T., Dombek, K.M., Tachibana, C., and Ideker, T. (2003). Multiple pathways are co-regulated by the protein kinase Snf1 and the transcription factors Adr1 and Cat8. Journal of Biological Chemistry *278*, 26146-26158.
- Young, E.T., Kacherovsky, N., and Van Riper, K. (2002a). Snf1 protein kinase regulates Adr1 binding to chromatin but not transcription activation. J Biol Chem *277*, 38095-38103.
- Young, E.T., Kacherovsky, N., and Van Riper, K. (2002b). Snf1 protein kinase regulates Adr1 binding to chromatin but not transcription activation. Journal of Biological Chemistry *277*, 38095-38103.
- Zavrel, M., and White, T.C. (2015). Medically important fungi respond to azole drugs: an update. Future microbiology *10*, 1355-1373.
- Zhang, E., Tanaka, T., Tajima, M., Tsuboi, R., Nishikawa, A., and Sugita, T. (2011). Characterization of the skin fungal microbiota in patients with atopic dermatitis and in healthy subjects. Microbiol Immunol *55*, 625-632.
- Zhu, Z.Y., Wang, H., Shang, Q.H., Jiang, Y.Y., Cao, Y.Y., and Chai, Y.F. (2013). Time Course Analysis of Candida albicans Metabolites during Biofilm Development. Journal of Proteome Research *12*, 2375-2385.

Appendices

Appendix 1 *Candida albicans* strains used in this study.

Strain	Parent	Mating type	Description	Source
CAI4	CAI4	*a/α*	*his3::hisG/his3::hisG leu2::tetRGAL4AD-URA3/LEU2*	Roemer *et al*., (2003) Molecular Microbiology, **50**, 167-181
SN148	SN76	*a/α*	*arg4Δ/arg4Δ leu2Δ/leu2Δ his1Δ/his1Δ ura3Δ::imm434/ura3Δ::imm434 iro1Δ::imm434/iro1Δ::imm434*	Noble *et al*., (2005) Eukaryotic Cell, **4**, 298-309
SN148	SN76	*aa*	*arg4Δ/arg4Δ leu2Δ/leu2Δ his1Δ/his1Δ ura3Δ::imm434/ura3Δ::imm434 iro1Δ::imm434/iro1Δ::imm434*	Lab Stock

Appendix 2 Plasmid used for this study

Plasmid Name	Description	Reference
pFA-URA3	Disruption cassette	Gola *et al.,* (2003) Yeast **20**, 1339-1347
pFA-ARG4	Disruption cassette	Gola *et al.,* (2003) Yeast **20**, 1339-1347
pFA-HIS1	Disruption cassette	Gola *et al.,* (2003) Yeast **20**, 1339-1347
pFA-TAP ARG4	Disruption cassette and/or epitope tagging	Lavoie *et al.*, (2008) BMC Genomics **9**, 578
pFA-GFP ARG4	Disruption cassette and/or GFP tagging	Gola *et al.,* (2003) Yeast **20**, 1339-1347
CIp10	Complementation plasmid	Murad *et al.*, (2000) Yeast **16**, 325-327

Appendix 3 Strains constructed for this study

Strains	**Genotype**	**Source**
SN148	*arg4/arg4 leu2/leu2 his1/his1 ura3::imm434/ura3::imm434 iro1::imm434/iro1::imm43* Lab stock	
TRY3	SN148 *::CIP10-VPR- TRY3*	Feng, 2013
C5_04410C_A	SN148 *::CIP10-VPR C5_04410C_A*	this study
C3_01220W_A	SN148 *::CIP10-VPR C3_01220W_A*	this study
C1_04510W_A	SN148 *::CIP10-VPR C1_04510W_A*	this study
C2_05640W_A	SN148 *::CIP10-VPR C2_05640W_*	this study
HAP41	SN148 *::CIP10-VPR HAP41*	this study
HAP42	SN148 *::CIP10-VPR HAP42*	this study
TRY6	SN148 *::CIP10-VPR TRY6*	this study
C2_02530W_A	SN148 *::CIP10-VPR C2_02530W_A*	this study
C2_01420C_A	SN148 *::CIP10-VPR C2_01420C_A*	this study
C4_07150W_A	SN148 *::CIP10-VPR C4_07150W_A*	this study
UGA33	SN148 *::CIP10-VPR UGA33*	this study
NRG2	SN148 *::CIP10-VPR NRG2*	this study
ACE2	SN148 *::CIP10-VPR ACE2*	
C1_13440C_A	SN148 *::CIP10-VPR C1_13440C_A*	this study
RFX1	SN148 *::CIP10-VPR RFX1*	this study
ADR1	SN148 *::CIP10-VPR ADR1*	this study
PHO4	SN148 *::CIP10-VPR PHO4*	this study
CR_05880W_A	SN148 *::CIP10-VPR CR_05880W_A*	this study
C2_08950W_A	SN148 *::CIP10-VPR C2_08950W_A*	this study
GLN3	SN148 *::CIP10-VPR GLN3*	this study
ASH1	SN148 *::CIP10-VPR ASH1*	this study
C1_11690W_A	SN148 *::CIP10-VPR C1_11690W_A*	this study
SFU1	SN148 *::CIP10-VPR SFU1*	this study
GAT1	SN148 *::CIP10-VPR GAT1*	this study
C3_05050W_A (Try4)	SN148 *::CIP10-VPR C3_05050W_A*	this study
C1_11690W_A	SN148 *::CIP10-VPR C1_11690W_A*	this study
C3_01220W_A	SN148 *::CIP10-VPR C3_01220W_A*	this study
C4_07150W_A	SN148 *::CIP10-VPR C4_07150W_A*	this study
MET4	SN148 *::CIP10-VPR MET4*	this study
MET28	SN148 *::CIP10-VPR MET28*	this study
MET32	SN148 *::CIP10-VPR MET32*	this study
MET4Δ/Δ	SN148 *met4::ARG4/met4::ARG4*	this study

MET28Δ/-	SN148 *met28::ARG4/-*	this study
MET32Δ/Δ	SN148 *met32::ARG4/met32::ARG4*	this study
MET4Δ/-	SN148 *met4::ARG4/-*	this study
MRR2Δ/-	SN148 *mrr2::ARG4/mrr2::ARG4*	this study
MET32Δ/-	SN148 *met32::ARG4/-*	this study
ADR1Δ/Δ	SN148 *adr14::ARG4/adr1::ARG4*	this study
MRR2Δ/-	SN148 *mrr2::ARG4/-*	this study
ADR1Δ/-	SN148 *adr1::ARG4/adr1::ARG4*	this study
ADR1 Rev	SN148 *mrr2::ARG4/mrr2::ARG4 ::CIP10-VPR ADR1*	this study
Met4 Rev	SN148 *met4::ARG4/met4::ARG4 ::CIP10-VPR MET4*	this study

Appendix 4 Primers used in this study

Name	SEQUENCE
Met28 DND Forward	CGACGCGTTTACTGACCTTAAATCTTGATTTCTCC
Met28 DND REVERSE	TCCCCCCGGGACCCCCAACAACGGCCTATAAA
MET4 DND FORWARD	CGACGCGTGACAACCAGGAAATTGGCAGT
MET4 DND REVERSE	TCCCCCCGGGGGTTTCCTTATATGACAGTTCTAAC
GCN4 F	CGACGCGTTGGTGTCCCCTTTACTACTGCT
GCN4 R	TCCCCCCGGGGTGCCAACGTGACTTTAGGGC
TEC1 F	CGACGCGTTCGCAAGCTACTCCTAGTGCT
TEC1 R	CCCCCCGGGAGCCGCTAAACTAATGTATCCAACA
Mark.ADR1-Forward primer	CGACGCGTCCAACTCACCAGAGTCAATAC
Mark. ADR1 reverse primer	TCCCCCCGGGCGTCAACATAGTAAAGTACTCC
Mark.Dal80 Forward primer	CGACGCGTTCTATGAGTGATATCCAACAACGTC
Mark.Dal80 Forward primer	TCCCCCCGGGTTGTGAAGTCCCTTGGCGGTAT
Manjari CRISPR P1	AAGAAAGAAAGAAAACCAGGAGTGAA
Manjari CRISPR P4	ACAAATATTTAAACTCGGGACCTGG
Manjari Cas9 F-	ATCTCATTAGATTTGGAACTTGTGGGTT
Manjari Cas9 R-	TTCGAGCGTCCCAAAACCTTCT
VP64F PRIMER	CCCAAGCTTCGGTTCCGGACGGGCTATGATG
VP64R PRIMER	CGACGCGTAGATCCGGAACTTCTAGA
GCN4 F PRIMER GOF	CGACGCGTCCTGCTACTACTCCTATTAT
GCN4 R PRIMER GOF	CCCCCCGGGATCTCTCTCTCTCCCTGCTGCT
CAP1 F PRIMER GOF	CGACGCGTGATATTGCCTCACCAGCAAATCTA
CAP1 R PRIMER GOF	CCCCCCGGGATTAAACCCACCACTAACTTGA
CRISPR-sgRNA F Met4	GCAACTGAGCAATACCACAAGTTTTAGAGCTAGAAATAG CAAGTTAAA
CRISPR-sgRNA R Met4	TTGTGGTATTGCTCAGTTGCCAAATTAAAAATAGTTTACG CAAGTC
Repair F met4	ATGAGTGACGAACCAACATCAGCAGCATTATTGGAGCAG TTAGTGTATATTGATAATTATTCCAATACCTCGCCAGAAC

Repair R Met4	TTATTGATTTAATCTAGCCTTTTCACGTAATAATTTAACTT CATCATCAGATTTTTGTGTCTCACTATAGGGAGACCG
CRISPR-sgRNA F Met32	AGATGATGATGATGAAGAGGGTTTTAGAGCTAGAAATAG CAAGTTAAA
CRISPR-sgRNA F Met32	CCTCTTCATCATCATCATCTCAAATTAAAAATAGTTTACGC AAGTC
GAT 1 F	CGACGCGTCACAACAGCAACCTTCTGCCAA
GAT1 R	TCCCCCCGGGAGAAACCTGCAGACTTTTCACT
GRF10 F	CGACGCGTACACCTACTCCGCCAACATCAT 3'
GRF 10 R	TCCCCCCGGGTTGGCAGCAAAAGCACAACTCA
PHO4 F	CGACGCGTTCATTCTCCCCTCAACCTGCTG
PHO4 R	TCCCCCCGGGCGATCAACCCCTCCTCTTTGGT
GST Tag Met4 F	CCGCTCGAGACAGCTACAACCACAACCACACC
GST tag Met R	CGGGATCCGACTCGCTCCAAAACCCCTTTG
GST TAG MET28 F	CCGCTCGAGTTACTGACCTTAAATCTTGATTTCTCC
GST TAG MET28 R	CGGGATCCACCCCCAACAACGGCCTATAAA
VP64 nls f	CCCAAGCTTCAGCAGGGCTGACATGCCCAAGAA
vp64 nls R	CGACGCGTGGTGAAGGGGTAAGGCTGGGG
GSt tag Adr1 F	CCGCTCGAGATAGATGCGCCTACTCCCGCCT
Gst tag Adr1 R	CGGGATCCAGTTGGCACCCGATTGAGCAGA
Ura3 guideF	ACCACCAACCAAGAGCCAAGGTTTTAGAGC TAGAAATAGCAAGTTAAA
Ura3 guide R	CTTGGCTCTTGGTTGGTGGTCAAATTAAAAATAGTTTAC GCAAGTC
CTF1 Guide F p2	ATCAAATATAAACGTCCTAGGTTTTAGAGC TAGAAATAGCAAGTTAAA
Ctf1 Guide R p3	CTAGGACGTTTATATTTGATCAAATTAAAAATAGTTTAC GCAAGTC
Reoain Ctf1 F	TAAAAAAGATAAACCAAAGGCATTTACTATCAAATATA AACGTCCTCCAATACCTCGCCAGAAC
reoair R	TGATCGACAAACTGTACAAGCACGAGAACCTCTCTCACT ATAGGGAGACCG
Met32-F-cripsr del-	5'ATTTGCAATATGCTAGTGCTCCTCA 3'
Met32-R-cripsr del	-5'AAAACTGAGGAGCACTAGCATATTG 3'
Met28-F-Cripsr del	-5'ATTTGCAACACTGTTCCATCAACAG 3'
Met28-R-Crispr del	-5'AAAACCTGTTGATGGAACAGTGTTG 3'
Met4-F-Crispr del	5' ATTTGAGCCACGCCTAATCTTGACA 3'
Met4-R-Crispr del	5'AAAACTGTCAAGATTAGGCGTGGCT 3'
GCN4-F	5' CGACGCGTCGTGTTCTGAATAGAGTAAAAAGGCCT 3'

GCN4-R	5’ CCATCGATGG ATCTCTCTCTCTCCCTGCTGCT 3'
Vp64- F	GTCGACGCTGCAGAAAGGTAACGAGCTG
Vp64-R	CTAGCTAGCTAGGAGCAGAGGCCTGAGAGATCTG
Orf19.3328 F	CGACGCGTCGTATACCGCTCAACTCCGTTCCC
Orf19.3328 R:	CCATCGATGGCATTGGCTGGTTGATCTGGTCG
NAT- F	ATTTGCGGCCGCTTTACCTCTGACACATGCAGCTCCCGG
NAT –R	CGGGATCCCG CCTCGCTCACTGACTCGCT