Effect of alterations in non-coding regions in genomic expression reprogramming of two *Candida glabrata* clinical isolates

Tiago Godinho de Ornelas Pedreira

Thesis to obtain the Master of Science Degree in

Microbiology

Supervisors

Prof. Doctor Nuno Gonçalo Pereira Mira

Prof. Doctor Sara Alexandra Cordeiro Madeira

Examination Committee

Chairperson: Prof. Doctor Jorge Humberto Gomes Leitão

Supervisor: Prof. Doctor Nuno Gonçalo Pereira Mira

Member of the committee: Prof. Doctor Miguel Nobre Parreira Cacho Teixeira

December 2015

Acknowledgements

Firstly, I would like to thank my supervisor Professor Nuno Mira, for giving me the opportunity to integrate this project, guiding me through this thesis, for never stop believing in my work, motivating and helping me even when nothing seemed to go well. A deep and sincere thanks to you Nuno. I would like to thank my co-supervisor Professor Sara Madeira for helping me and for always trying to help whenever a problem was encountered. I would also like to acknowledge Professor Isabel Sá-Correia for providing me the conditions to work in Biological Sciences Research Group and for allowing me to develop my thesis at BSRG during this time. I would also like to thank financial support of project PanCandida – Towards the development of a pan-genomic DNA chip for the early detection of invasive candidiasis caused by C. albicans and *C. glabrata*, sponsored by the Gilead Génese Program; of Pfizer research program WI178570; and of FCT (UID/BIO/04565/2013).

I would like to express my most sincere gratitude to Rui Henriques, from INESC-ID, for all the help provided and all the patience showed every time I asked for help. I want to express my gratitude to Professor Miguel Teixeira and Catarina Costa for providing the documented target genes list for CgPdr1, CgAmt1 and CgMsn4 and to Dr. Hiroji Chibana for engineering the KUE100 mutants. A big thank you to Pedro Pais and Sara Salazar for always helping me at the laboratory.

A big thanks to Laura Luzia for all the support, motivation, companionship and a lot of patience during this time. Thank you very much. Thanks to Pedro Moura and João Peça for all the good moments and great friendship. I would like to thank Inês Leonardo, Andreia Lourenço, João Santos and Rita Grenho for all the laughter, quantum-physics debates and great friendship. To my long-time friends, Andreia Pimenta, Pedro Sampaio, Tiago Amaro, Inês Oliveira and Sara Salazar a huge thank you for always being there whenever I needed and whenever I didn't, for always giving me motivation and helping me get through every hurdle in my path. To Tatiana Alves a huge thank you for always making a difference, for always making me laugh and supporting me in this journey.

Finally I would like to thank my family for all the support given during this time, especially my Mother who never gave up believing in me, for the unconditional support, motivation and for letting me follow this path. To my brother Diogo, for all the great moments we shared and keep sharing, for always motivating me, supporting me and try to make me see the light at the end of the tunnel, for becoming my own role model. To my grandparents Ilda and Manuel, for always believing in me and for always supporting me. A huge thank to you all, it wouldn't be possible without you.

I dedicate this thesis to my Mother.

Abstract

Up to now the majority of the genomic analysis undertaken in *C. glabrata* clinical isolates have focused modifications in coding regions ignoring what happens in the non-coding genome. In this work the non-coding genomes of the reference strain CBS138 and of a clinical isolate, named FFUL887, were compared. Among other phenotypic traits, the two strains are differently resistant to antifungals, the FFUL887 isolate being much more tolerant to fluconazole, voriconaozle and caspofungine than CBS138. Massive alterations had been identified in the non-coding genome of the two strains reflecting the great plasticity that characterizes these regions. Emphasis has been put in the examination of how these alterations may impact the governance of the overall regulation of genomic expression in the two strains by modulating the interaction of transcription factors with their binding sites. In particular, the regulatory network controlled by CgPdr1, a key regulator of resistance to antifungals in *C. glabrata*, was focused. 31 new putative CgPdr1 targets whose promoter in the FFUL887 harbored at least one binding sites for CgPdr1, which was absent in the CBS138 counter-partner, were identified. 7 of these putative CgPdr1 targets were confirmed to contribute for maximal *C. glabrata* tolerance to caspofungine and 2 were required for tolerance to fluconazole and voriconazole. The results obtained in this study are expected to contribute for the advance of the understanding of the mechanisms underlying the evolution of regulatory networks controlling the genomic expression of *C. glabrata* cells during colonization of the human host and during acquisition of resistance to antifungals.

Key-words: *C. glabrata*, resistance to antifungals, genomic analysis of non-coding regions, evolution of regulatory networks in *C. glabrata*

Resumo

Até recentemente, a maioria das análises genómicas feitas em isolados clínicos de *C. glabrata* focaram as modificações nas regiões codificantes, ignorando o que acontece no genoma não-codificante. Neste trabalho, os genomas não-codificantes de uma estirpe de referência, CBS138, e de um isolado clínico, chamado FFUL887, foram comparados. Entre outras características fenotípicas, as duas estirpes são diferentemente resistentes a antifúngicos, sendo que o isolado FFUL887 é bastante mais tolerante a fluconazole, voriconazole e caspofungina do que a CBS138. Alterações massivas têm sido identificadas no genoma não-codificante das duas estirpes, reflectindo a elevada plasticidade que caracteriza estas regiões. Tem sido posto ênfase na examinação de como estas alterações podem causar impacto na governação da regulação da expressão genética geral nas duas estirpes por modulação da interação entre factores de transcrição e os seus locais de ligação. Em particular, a rede regulatória controlada pelo CgPdr1, um regulador chave na resistência a antifúngicos em *C. glabrata,* foi estudada em detalhe. 31 Novos alvos putativos do CgPdr1 cujo promotor albergue pelo menos um local de ligação para o CgPdr1 na FFUL887 e que estivesse ausente no promotor da CBS138, foram identificados. 7 Destes alvos putativos do CgPdr1 foram confirmados na atribuição de máxima tolerância à caspofungina e 2 foram identificados como necessários para a tolerância ao fluconazole e voriconazole. Os resultados obtidos neste estudo são esperados que contribuam para o avanço da percepção dos mecanismos subjacentes à evolução das redes regulatórias que controlaram a expressão genómica de células de *C. glabrata* durante a colonização no hospedeiro humano e durante a aquisição de resistência a antifúngicos.

Palavras-chave: *C. glabrata*, resistência a antifúngicos, análise genómica de regiões não-codificantes, evolução da rede regulatória de *C. glabrata*

List of Abbreviations

ABC - ATP-Binding Cassette

bp – Base pairs

***C. albicans* -** *Candida albicans*

***C. glabrata* -** *Candida glabrata*

CIAP – Calf Intestinal Alkaline Phosphatase

DMSO – Dimethyl sulfoxide

DNA - Deoxyribonucleic Acid

dNTP – Deoxynucleotide Triphosphates

ECOFF - Epidemiological Cut-off Value

EPA family - Epithelial Adhesion Family

EUROSCARF – European *Saccharomyces cerevisiae* Archive for Functional Analysis

GOF - Gain-of-function

HOG - High Osmolarity Glycerol

MMB – Minimal Growth Medium

MDR - Multi-drug Resistance

MFS - Major Facilitator (MF) Superfamily

MIC - Minimum Inhibitory Concentration

NCAC - Non-Candida albicans Candida Species

OD_{600} – Optical density at a wavelength of 600 nm

ORF- Open Reading Frame

PCR – Polymerase Chain Reaction

PDR - Pleiotropic Drug Resistance

PDRE - Pdr Response Elements

PEG – Polyethylene glycol

RNA - Ribonucleic Acid

rpm – rotations per minute

RPMI - Roswell Park Memorial Institute Medium

S. cerevisiae - Saccharomyces cerevisiae

SNP - Single-nucleotide Polymorphism

TF – Transcription Factor

YPD – Yeast Peptone Dextrose

List of Figures

List of Tables

Contents

1. Introduction

1.1. Relevance *of Candida glabrata* in the context of fungal infections.

The number of life-threatening infections caused by the members of the fungal genus *Candida*, generally known as invasive candidiasis, has risen significantly in the past decades mainly due to a significant increase in the size of the immunocompromised population as a result of the massification of aggressive therapeutic treatments that negatively affect the activity of the immune system[1–3]. Other recognized risk factors for the development of invasive candidiasis include the use of broad-spectrum antibiotics and the use of indwelling intravenous catheters[1,2,4]. Nowadays, bloodstream infections caused by *Candida* species are more frequent than those attributed to *Streptococci*, *Enterococcus faecalis* or *Escherichia coli*[1,2]. The mortality rate associated to invasive candidiasis can be as high as 40% of the diagnosed infections[5]. Besides invasive infections, Candida spp. are also responsible for a significant proportion of superficial rushes, mainly affecting the mucosas, these being known as superficial candidiasis[2]. Although *C. albicans* is the more common causative agent of invasive and superficial fungal infections, there has been a significant increase in the number of infections caused by non-*Candida albicans* species, generally known as NCAC[2]. These NCAC species include *C. glabrata*, *C. tropicalis*, *C. parapsilosis* and *C. krusei*[2], among others. The relevance of each one of these NCAC species in the context of candidiasis is variable and country-dependent: *C. albicans* is the most common agent in Europe (54.8% of infections) but the least common one in Asia-Pacific (38.2% of infections). Similarly, *C. glabrata* is the most common agent in North America (25.5% of infections) but the least common in Latin America (7.7% of infections)[6]. In general *C. glabrata* ranks as the second most relevant agent of candidiasis, usually followed by *C parapsilosis* and *C. tropicalis*[2,6]. Epidemiological surveys undertaken in Portugal have also reported an increase in the number of infections caused by NCAC[6,7]. However, some discrepancies in the incidence of infections caused by the different NCAC species have also been obtained in these studies: while in healthcare facilities in Portugal *C. glabrata* was found to account for 10% of candidemia occurrences, in epidemiological studies undertaken in Braga *C. glabrata* had the lowest prevalence within patients diagnosed with candidemia, ranking before *C. albicans* (79%) and *C. tropicalis* (5.6%)[8]. The mortality rates attributed to candidemia associated to *C. glabrata* reach about 40% of all diagnosed infections, these being rates comparable to those attributed for *C. albicans*. However, when diagnosis is delayed, infections caused by *C. glabrata* can be more severe and deathly than those caused by *C. albicans*[9–11].

The main factor attributed to the increase in the incidence of infections caused by *C. glabrata* is its reduced susceptibility to the main drug used in the treatment of fungal infections, fluconazole[12]. Besides its extensive use as a prophylactic and treatment agent, which enhances selective pressure[2], the increase in *C. glabrata* resistance to fluconazole (and to other structurally similar azoles) is also though to result from the widespread use of agricultural fungicides structurally similar to clinical azoles[13]. Additionally, *C. glabrata* acquires resistance

to antifungals at a higher rate than other *Candida* spp[14–16]. The emergence of *C. glabrata* strains resistant to newly developed drugs, alternative to fluconazole, such as voriconazole or echinocandins, is increasing, thereby hencing the urgency in understanding which determinants are behind the remarkable antifungal resistance of this yeast species[3,15,17,18]. Some of the knowledge already gathered in the field is reviewed in the following section, with emphasis on the essential role played by the pleiotropic drug resistance transcription factor CgPdr1.

1.2. Overview on the molecular mechanisms of resistance to antifungals in *C. glabrata*: the central role of CgPdr1

Currently there are four different type of drugs that are used against fungal infections: azoles, flucytosine, polyenes and echinocandins. The fact that both fungal and host cells are eukaryotic stands as a limitation while developing new antifungals as the therapeutic targets that can be used during the process are narrowed down. Azoles and flucytosine have a fungistatic effect against Candida spp., whereas echinocandins and polyenes are fungicidal[3,19]. Azoles can be divided in two different groups, based on their chemical structure: imidazoles, which include clotrimazole and ketoconazole; and triazoles, which include fluconazole, voriconazole[20,21] (Figure 1 and 4). The mechanism of action of azoles involve binding to the iron atom located in the heme group of the P450 demethylase enzyme involved in ergosterol biosynthesis. Upon such binding, the activity of the enzyme is inhibited thereby leading to the accumulation of toxic 14α-methylated sterols in the membrane. Altogether with the accumulation of toxic sterols, there is a depletion of ergosterol from the membrane which will compromise the permeability of the plasma membrane[3,22] (Figure 2).

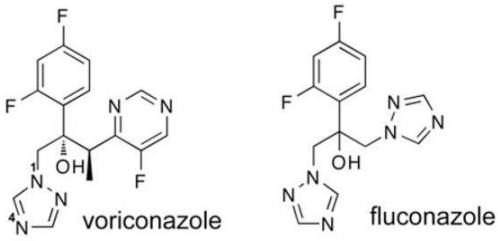


Figure 1 - Chemical structure of two azoles: voriconazole and fluconazole. Adapted from Mast N. et al 2013[22].

Figure 2 - Representation of the accumulation of toxic sterols when *C. glabrata* is in contact with azoles. Azoles will block the production of ergosterol causing the accumulation of toxic sterol intermediate, resulting in cell membrane stress. Adapted from Cowen, LE. 2008[23].

The increased resistance to azoles in *C. glabrata* has been associated to multiple factors including the alteration of the target enzyme Erg11, the overexpression and mutation of other enzymes of the ergosterol biosynthetic pathway and the overexpression of drug-efflux pumps (Figure 3) [3,23]. In *C. glabrata* there are at least four multi-drug efflux pumps of the ABC Superfamily that have been demonstrated to play a role in azole resistance including CgPdh1, CgPdr16, CgSnq2 and CgCdr1[24,25]. Within this set, CgCdr1 is the one showing a more significant protective effect[25–27]. More recently, the involvement of the multridrug resistance transporters of the Major Facilitator Superfamily CgTpo3, CgQdr2, CgAqr1, CgTpo1_1 and CgTpo1_2 in *C. glabrata* tolerance to azoles has also been unveiled[28–30]. However it remains to be established if these transporters play a role in the azole-resistance phenotype exhibited by clinical isolates, something that has been demonstrated at least for CgCdr1[27,31]. The ability of *C. glabrata* to promote the uptake of sterols directly from the growth medium also seems to be an important trait contributing for the high tolerance of this yeast species to azoles[23,32] as these sterols taken from the medium can be further metabolized into ergosterol or even be directly incorporated in the plasma membrane, enabling the compensation the depletion of ergosterol in the membrane during azole stress[32–34]. As the sterol uptake increases the sturdiness of *C. glabrata* during azole stress, the deletion of sterol transporter, CgAus1, was shown to reduce the high tolerance of this yeast species to azoles[34,35].

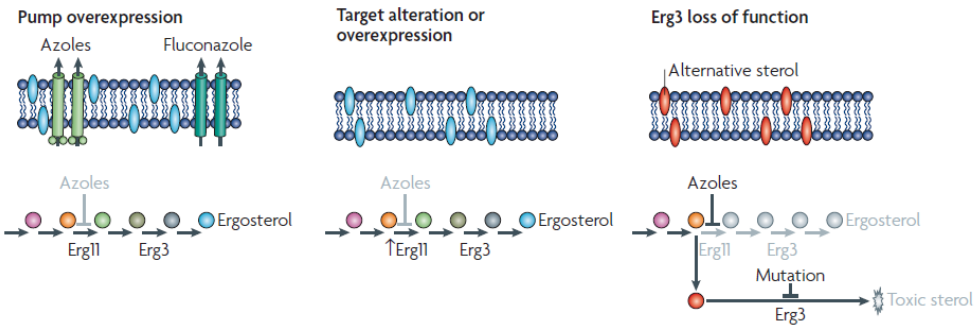


Figure 3 - Representation of azole resistance mechanisms. Resistance to azoles may result from the upregulation of efflux pumps that enable the removal of the drug from the cell; through the mutation or overexpression of some key elements in the ergosterol biosynthetic pathway, such as Erg11, which

minimizes the impact of the drug on the target; by altering the ergosterol biothsynthesis, such as the loss-of-function mutation of Erg3, which blocks the accumulation os toxic sterol intermediate when Erg1 is inhibited by azoles. Adapted from Cowen, LE. 2008[23].

To overcome the emergence of resistance to azoles new antifungal drugs have been/are being developed out of which echinocandins are those more prominent. The first approved echinocandin was caspofungin, in 2002, followed by micafungin in 2005 and anidulafungin in 2006. This new drug class is composed by lipopeptides having a cyclic hexapeptide core N-linked to a varying acyl side chain (Figure 4)[36,37]. The mechanism of action of echinocandins involves the inhibition of 1,3-β-D-glucan synthase which perturbs cell wallsynthesis[3,23,36]. 1,3-β-D-glucan is the major component of the fungal cell wall, ranging from 30 to 60% of the cell wall in *Candida*, maintaining the cell wall's structural integrity. Echinocandins halt cell wall formation, leading to osmotic instability and lastly promoting cell death[37,38](Figure 5).

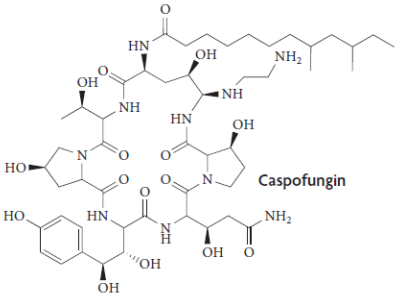


Figure 4 - Chemical structure the echinocandin, caspofungin. Adapted from Cowen, LE. 2008[23].

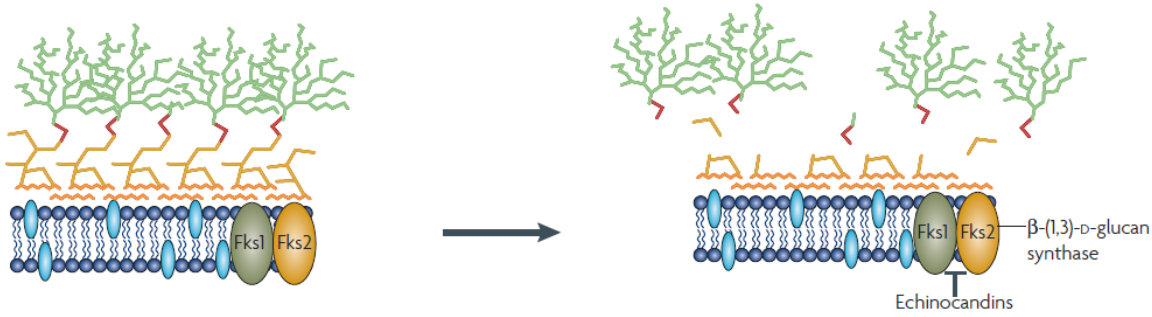


Figure 5 - Representation of echinocandins' mechanism of action. Echinocandins inhibit 1,3-β-D-glucan synthase, the catalytic subunit is encoded by FSK1 and FSK2, and thus disrupting cell wall integrity. Adapted from Cowen, LE. 2008[23].

Up to now the most well described molecular mechanisms for increased resistance of *C. glabrata* isolates to echinocandins involves mutations on the glucan synthase encoding genes CgFKS1 and CgFKS2 (Figure 6). Nevertheless it is not yet clear if the mutations directly contribute to the increased resistance to echinocandins by affecting binding of the drugs to the proteins or by affecting the action of the drug in an indirect manner[23,39–49].

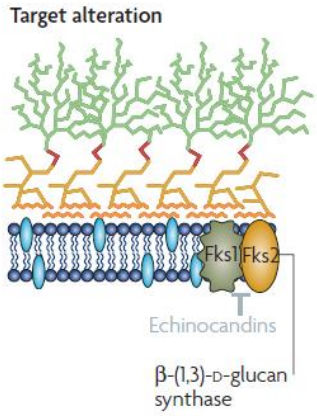


Figure 6 - Representation of echinocandin resistance mechanism. Resistance can result from mutations in Fks1 that minimize the drug impact, maintaining the integrity of cell wall. . Adapted from Cowen, LE. 2008[23].

1.2.1. Role of the transcription factor CgPdr1 in *C. glabrata* response and tolerance to antifungals

The pleiotropic-drug resistance network (PDR) is a well-organized and complex regulatory network under the control of the transcription factor CgPdr1 and that has been found to play a fundamental role in the control of *C. glabrata* response to azoles [31,50]. Although some information regarding the PDR network has been gathered in *C. glabrata*, this network has been much better studied in *S. cerevisiae*. In this latter species, the PDR network is mainly controlled by two orthologous transcription factors: ScPdr1 and ScPdr3[51,52]. The elimination of these transcription factors was found to increase *S. cerevisiae* susceptibility to a wide range of xenobiotics that are structurally unrelated, including antifungals, antibiotics, anticancer drugs and pesticides[53]. In *C. glabrata* there is only one orthologue of Sc*PDR1*, Cg*PDR1*, and its elimination increased the susceptibility to fluconazole, voriconazole, itriconazole and ketoconazole in this species[31,54–56]. Different transcriptomic analysis have unveiled the set of genes regulated by CgPdr1 during *C. glabrata* response to azoles these including several genes encoding ABC- and MFS MDR pumps, such as Cg*SNQ2*, Cg*CDR1*, *CgPDH1*, Cg*YOR1* and Cg*TPO3*, but also genes that are involved in the metabolism of lipids, fatty acids and sterol, such as Cg*AFT2*, Cg*HFD1*, Cg*RTA1* and Cg*ERG4;* stress response (*YIM1*); transcription; and adhesion (such as Cg*PWP6* and Cg*EPA1*)[50,57,58]. Comparison of the results obtained in *C. glabrata* with those gathered in *S. cerevisiae* revealed that the set of genes regulated by CgPdr1 and ScPdr1 is different, on the essential the two regulatory systems appear to control the same biological functions[53]. The mechanism underlying the activation of CgPdr1 by ketoconazole has been recently discovered this being hypothesized as a general mechanism leading to the activation of this transcription factor in response to drugs (Figure 7)[59]. This study showed that this particular drug, ketoconazole, is able to bind directly to a domain of CgPdr1p

that is located at the C-terminal, xenobiotic binding domain (XBD), suggesting that this transcription factor can act as a nuclear receptor. This is possible as upon binding of the drug, the complex Pdr1p-drug binds to the KIX domain of the mediator complex subunit, CgGal11A, which will promote the up-regulation of RNA polymerase II activity, promoting the transcription of Pdr1 target genes, including drug efflux pump encoding genes. Although this mechanism was discovered with ketoconazole, it should be similar to other azole drugs, such as voriconazole and fluconazole [59].

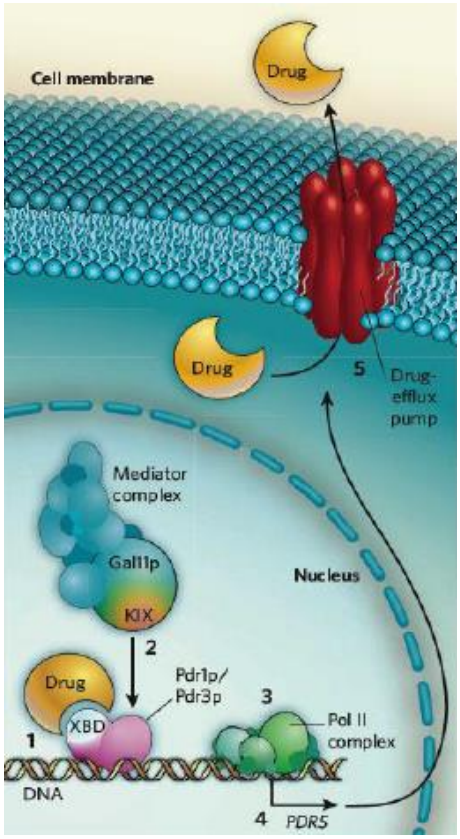


Figure 7 - General mechanism describing the activation of Pdr1/Pdr3 in response to drugs within a fungal cell. Drug binds to the XBD domain of the pleiotropic transcription factors Pdr1/Pdr3, acting as a nuclear receptor. This complex Pdr1/Pdr3-drug can associate to the KIX domain of the Gal11p subunit on the mediator complex and recruit RNA polymerase II to the promoter region of PDR genes, activating the transcription. Image adapted from Goffeau, A., 200860.

Although CgPdr1 is the main player in the control of the PDR network, several other transcription factors have also been found to play an important role in the regulation of PDR genes. In Figure 8 it is depicted a schematic representation of the complexity of the PDR network using results that were gathered in *C. glabrata* and further complemented with those obtained in *S. cerevisiae*. In specific, it has been shown that Cg*PDR1* is activated by Cg*HST1*, Cg*RFM1* and Cg*SUM1*, being repressed by CgStb5. It has also been seen that Cg*PDR1* auto-regulates its own expression and that it represses Cg*RPN4*[50,57,58,60,61–66]. Yrr1 and Yrm1, members of the diverse Zn_2Cys_6 zinc finger transcription factor family[64], were also found to play a critical role in the control of the PDR network in *S. cerevisiae*, however, in *C. glabrata* the

relevance of these transcription factors has not yet been clarified. Sc*YRR1* was found to be involved in the up-regulation of drug-efflux pumps (e.g. *SNQ2*, *PDR15*, *YOR1*, *PDR5 AZR1*, *FLR1* and *SNG1*), transcription factors (*PDR3* and *YRR1* itself) and proteins involved in synthesis and transport of phospholipids (*PDR16*)[64,67]. Rpn4 is a transcription factor recognized as the regulator of proteasome genes in *S. cerevisiae*, but its involvement in regulation of drug-efflux pumps has also been demonstrated [68]. In *C. glabrata* the functional analysis of this transcription factor has not been undertaken at the same extent as it was performed in *S. cerevisiae*, however, the Cg*RPN4* gene is known to be up-regulated when *C. glabrata* cells are in the presence of fluconazole[31,50,58]. CgStb5 is a negative regulator that is known to share its regulon with CgPdr1 these set of co-regulated genes including the drug efflux pumps Cg*CDR1*, Cg*PDH1*, Cg*YOR1* and also Cg*PDR1*[69]. The deletion of CgSTB5 gene was found to slightly increase resistance to fluconazole and voriconazole in *C. glabrata*, while its overexpression leads to increased susceptibility of these cells to these drugs[69]. In *S. cerevisiae* Stb5 was also found to activate the *ERG5*, *ERG11*and *ERG25* genes[70].

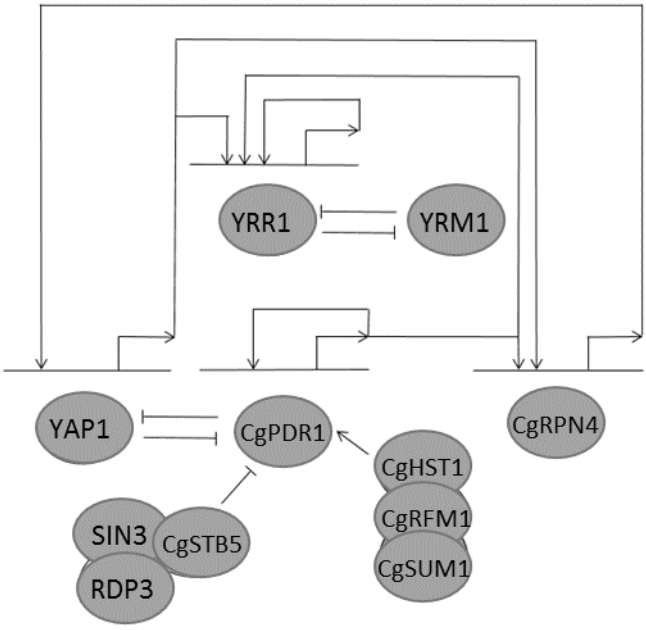


Figure 8 - Interactions of transcription regulators in PDR network that involve MDR response to azoles. Regulation action of Cg*HST1*, Cg*RFM1* and CgSUM1 over Cg*PDR1,* Cg*PDR1* auto-regulation, regulation of Cg*RPN4* and Cg*STB5* repression over Cg*PDR1* were verified in *C. glabrata* in studies from Tsai, H. *et al.*, 2006, Vermitsky, J. *et al.*, 2006, Caudle, K. *et al.*, 2011, Orta-Zavalza, E., *et al.*, 2013, Noble, J. *et al.*, 2013, respectively[31,50,58,60,69]. The remaining interactions were retrieved from *S. cerevisiae* in studies from Cui, Z. *et al.*, 1998, Le Crom, S. *et al.*, 2002, Owsianik, G. *et al.*, 2002, Lucau-Danila, A. *et al.*, 2003, Hahn, J. *et al.*, 2006, Salin, H. *et al.*, 2008[61–66].

The regulation of the genes involved in the PDR network has also been demonstrated to differ in gain-of-function (GOF) CgPdr1 mutants that were recovered from azole resistant clinical isolates. A wide range of GOF point mutations has been described to occur in the coding

sequence of Cg*PDR1* (Figure 9)[57,71,72]. Notably, in CgPdr1 GOF mutants the activity of the transcription factor becomes constitutively high thereby resulting in the constitutive expression of its target genes even in the absence of a xenobiotic stimulus[50,57,58,71]. Recent studies have demonstrated that the distribution of GOF mutations in the CgPDR1 gene are not associated with specific domains (Figure 9) and these can give origin to different transcription profiles by altering the overall structure of the transcriptional regulatory network of the strain[57,58,71].

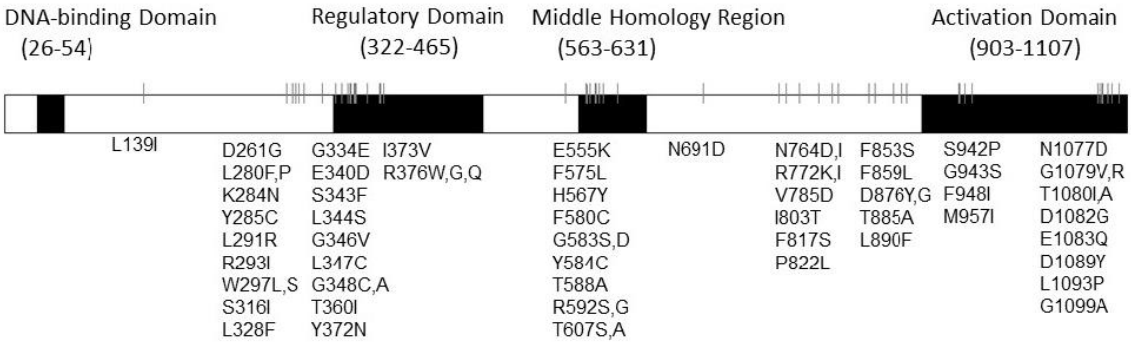


Figure 9 - Described *C. glabrata* PDR1 gain-of-function mutations, according to the information available in Ferrari, S., et al., 2009[71], Tsai, H., et al., 2010[57], Berila, N. and Subik, J.,2010[72]. Image adapted from Salazar, S., 2015[73].

1.3. Introduction to the theme of the thesis

Within the last years a great number of studies has been published studying the molecular mechanisms underlying resistance of *C. glabrata* clinical isolates to azoles and echinocandins, with emphasis played by CgPdr1-regulated genes [55,74]. The majority of these studies have however focused on specific mechanisms and sets of regulated genes and little is known on the alterations occurring at a genome sequence level. Recently a clinical *C. glabrata* isolate recovered from the urinary tract, named FFUL887, was identified as being resistant to fluconazole and voriconazole, also exhibiting increased tolerance to caspofungin, in comparison with the reference strain CBS138 which was found to be susceptible to all these antifungals. One of the traits that was uncovered by the analysis of the genome sequence of the FFUL887 isolate was the fact that it encodes a novel gain-of-function CgPdr1 mutant, this being corroborated not only by the drug-resistance phenotype of the isolate but also by the observation that several genes regulated by CgPdr1 were found to be over-expressed in the FFUL887 strain during cultivation in drug-free growth medium. Different from the approach that was previously performed, the work presented in this thesis is focused on the alterations that have occurred in the non-coding genome of the FFUL887 strain. Until so far, the few genomic analysis that had been undertaken to understand resistance to antifungals in *C. glabrata*, and in fact in most Candida spp. in general, have mostly focused the coding sequences being largely ignored the modifications occurring in non-coding sequences. A better understanding of how these non-coding regions are altered is, however, of paramount importance to fully understand how transcriptional regulatory networks are rewired in resistant isolates. Besides

this, these insights might also help in a better understanding of the molecular mechanisms underlying the control of *C. glabrata* genomic expression during colonization of the mammalian host.

1.4. Evolution of regulatory networks in *Candida glabrata*

Recent studies suggest that a significant portion of the phenotypic variation, within and between species, is due to a differential gene regulation[75,76]. Among the phenotypic traits that were found to be determined by a differential gene expression are the ability to reversibly switch between distinct cell and colony morphologies[76–78]. In *C. glabrata* colony phenotype switching has been described, these being characterized by changes in colony morphology and colour, as the colonies can shift from a normal white colour to a very-dark brown colour [76,78]. The process of *C. glabrata* evolution inside the human host has been found to include an increased rate of gene loss, when compared to *S. cerevisiae*[79]. Among the lost genes are those required for assimilation and use of galactose and sucrose, phosphate, nitrogen and sulfur and those required for nicotinic acid biosynthesis, these responses reflecting the adaptation of the microbe as a human commensal and opportunistic pathogen[76]. Another relevant trait of Candida spp in general, and of *C. glabrata* in general, is the significant expansion of genes encoding adhesins and proteins related with adhesion, namely of genes of the EPA and YPS family[79,80]. The expansion of *EPA* and *YPS* gene families has been previously linked to virulence, reinforces the idea that pathogens modulate their genes according to their evolutionary needs[76].

Other aspects of adaptation of *C. glabrata* to the human host involves the capacity to evade and/or defeat the host immune cells. *C. albicans* uses two strategies to undergo this process, one concealing the production of immunostimulatory molecules on the cell-wall surface, such as β-1,3-glucan, and by destroying the phagocytic cells of the host's innate immunity by hyphal outgrowth[76]. Differently *C. glabrata* has not been found to induce such a strong response over macrophages being able to survive and replicate in the phagosome of these cells, exhibiting a high resistance to starvation, which might also support survival during engulfment in phagocytic immune cells, although this mechanism is still unclear[76,81–83]. Recently, results of a study described microevolution of *C. glabrata* within macrophages[84]. This study, analysed the effect of a single point mutation in *CHS2* gene allowing the growth form to change from typical spherical yeast to pseudohyphae-like structures. This filamentous-like morphology permitted faster escape from macrophages and increased damage of macrophages. Additionally, this strain showed increased virulence in a systemic mouse infection model, suggesting that the microevolutionary processes can modify the adaptations of *C. glabrata* to distinct host niches and finally lead to hypervirulent strains[84].

Comparative genomic approaches between species, together with comparative gene expression profiling have been remarkably useful to understand regulatory network principles, as well as to develop insights into the possible evolution of these networks between different

species[75,76,85–87]. The principles of evolution of the networks include the conservation of binding sites of a certain transcription factor across species in a regulon, the target genes of this conserved regulon often having related functions and exhibiting a coherent expression under a number of different conditions, such as the presence or absence of drug in the medium. This has been demonstrated to be the case of the ScYap1/CgAp1 transcription factors which have been found to mediate response and tolerance of *S. cerevisiae* and *C. glabrata* to oxidative stress through the regulation of similar cohorts of target genes[66]. These changes in regulatory networks play an important role in driving specific adaptations in fungal species and, in particular, pathogenicity in *Candida* species[76,85,86]. An example of these changes is the rearrange of oxidative stress response network in *C. glabrata* when comparing to *C. cerevisiae. C. glabrata* oxidative stress resistance is regulated similarly to that of *S. cerevisiae* by the transcription factors CgSkn7, CgYap1 and CgMsn2/4. However, the high intrinsic oxidative stress resistance of *C. glabrata* is explained by the transcriptional differences between the two species, whereas only CgYap1 is induced and CgSkn7 is repressed in *S. cerevisiae*[76,88].

The advances in genomic tools and computational methods have been playing a critical role in the identification of regulatory network components (such as transcription factors and their target genes) and to study their patterns of conservation and divergence across multiple species. Changes in regulatory mechanisms can be learned from patterns of evolutionary divergence in regulatory properties at multiple levels: gene expression (mRNA levels), characterization of *cis*-regulatory elements in orthologous promoter sequences, protein-DNA interactions measured across organisms (transcription factor binding), and duplication and divergence of regulators or comparing chromatin organization. Regulatory variations can be difficult to detect as they can occur at multiple levels, such as in the transcription, post-transcriptional, translation, post-translational and epigenetic levels[86]. As an example, the comparison of organismal phenotypes, including those already mentioned, during development in a single species and through evolutionary time in different species allowed the linking of specific phenotypes to change expression of distinct genes as well as more global transcriptional rewiring. Indeed, gene expression show variations within and between species, suggesting that it is influenced by the evolution of gene regulation[86]. The emergence of drug resistance may be a result of this transcriptional rewiring that is resultant of the change of expression of several genes, such as Cg*PDR1* and the genes regulated by CgPdr1[31,50,58,76]. However, it is difficult to tell whether gene expression levels vary according to natural selection or neutral drift. In *C. glabrata* the fact that this species exhibits higher tolerance to drugs may have resulted from the pressured selection exerted during extensive treatments using azoles, such as fluconazole[3,15,86].

Despite the likely importance of variation in gene expression, little is known about the evolution of gene-expression regulation or how this evolution contributes to microbial diversification. A gene's expression pattern is dictated by flanking noncoding sequences that contain, among other components, binding sites that are recognized by sequence-specific

nucleotide-binding proteins that module transcript abundance, transcription factors. These regions, *cis*-regulatory elements, govern the expression of co-regulated genes[89–91]. Functional regulatory sequences can be identified in a set of co-regulated genes, meaning that the enriched part of these same genes contain the same sequences within their flanking regions. This might suggested that co-regulation can be preserved in related species, such as the cis-regulatory elements, while the surrounding DNA will be exposed to higher rates of mutation[90,91]. Recent studies examined the evolution of *cis*-regulatory networks across species by analysing the orthologs of genes co-regulated in *S. cerevisiae*[91]. The results strongly suggest that many of the known *cis*-regulatory systems from *S. cerevisiae* have been conserved over the past hundred millions of years of evolution among orthologous regulatory regions from related *Saccharomyces* species. Although it is hard to identify if every *cis*-regulatory element is conserved in more distantly related species, the characterization of the pattern of evolution within transcription factor binding sites allow to explore the nature of functional constraints of these sequences. Comparing to some well-known protein sequences, it is to be expected that the pattern of evolution in transcription factors binding sites reflect the constraint under which they function, meaning that important regions or residues should be constrained, while unimportant positions may show fixed changes. Indeed, functionally important positions are expected to be under stronger purifying selection and therefore show stronger conservation[90]. Recent studies showed that the patterns of *cis*-sequence enrichment strongly suggest that many of the genes co-regulated in *S. cerevisiae* are also co-regulated in other fungal species, meaning that the expression of those genes is likely to be regulated by the same *cis*-regulatory systems[91]. It is to be expected, however, that species closer to *S. cerevisiae* have more identified *cis*-regulatory elements than those more distantly related, suggesting that the number of regulatory systems conserved across species correlates with their divergence times. Nonetheless, there are some *cis*-regulatory systems that are conserved across all fungal genomes, such as the group of G1-phase cell cycles genes[91]. The ability to identify novel enriched sequences of co-regulated *S. cerevisiae* genes implies that, even though the genes are still co-regulated in those species, the systems governing their expression have changed, indicating that the regulatory regions of those genes co-evolved to contain the same *cis*-sequences. This means that is possible that the regulation of the genes' expression evolved even though their expression patterns did not[91]. Sequences that match *cis*-regulatory elements can readily appear in noncoding DNA through drift. In the same way that this occurrence can promote binding site turnover within a given regulatory region, it can also create *de novo* elements in the regulatory regions of random genes, giving rise to new targets of that regulatory system[91,92]. An accurate understanding of the evolution of functional regulatory sequences is critical to the optimal use of comparative sequence data when analysing transcriptional regulation. This way, comparative functional genomics approaches are increasingly collecting and measuring multiple regulatory genomic data across species. Comparison of such datasets across species and strains has been fruitful in gaining an initial understanding of how changes in regulatory network drive the evolution of new phenotypes[76,91].

In a recent study, a novel softclustering algorithm was used to compare the global transcriptional response to fluconazole in *S. cerevisiae*, *C. glabrata* and *K. lactis (REF). The results obtained* illustrated how divergence of regulatory networks may dictate the intrinsic resistance of *C. glabrata*, or the increased sensitivity of *K. lactis*, to fluconazole[76,93]. The two most induced co-expressed genes in all three species were related with ergosterol synthesis, presumably to respond to the fluconazole-induced depletion of ergosterol. While most ergosterol genes were coordinately induced in all three species, the expression levels of the first reactions of the ergosterol biosynthetic pathway, in particular ithe genes involved in isoprenoid biosynthesis, were found to be remarkably divergent across all three species[76,93]. The divergent expression of ABC and MFS transporters also shed lights into the different susceptibility of the three species to fluconazole. In specific, it was observed an up-regulation of the orthologs *PDR5*/*10*/*15* in *S. cerevisiae* and in *C. glabrata* but not in *K. lactis*. Furthermore the *CgSNQ2* gene was solely induced in *C. glabrata*[76,93]. Additionally, ABC transporter genes, *AUS1* and *PDR11*, which regulate the sterol uptake under anaerobic conditions, were found to be induced in *S. cerevisiae*, but the expression of CgAus1 do not change[94]. Thus, the striking divergence in expression of fluconazole export and sterol import pathways suggests differing strategies in the azole response after exposure to fluconazole of the different species. In particular, the *S. cerevisiae* response seems to depend on the influx of sterol, by the up-regulation of *PDR11* and *AUS1*; while the response of *C. glabrata* relies on the efflux of fluconazole through the strong induction of *CgSNQ2* and *PDR5*/*10*/*15* orthologues. However, none of these responses evolved in *K. lactis*, and is the likely cause of the severe drug sensitivity in this species[76]. The application of this set of functional genomics tools to generate both molecular and phenotypic measurements, altogether with computational methods to construct phenotype-predictive regulatory networks and their comparison between species, helps to clarify the genes that contribute to virulence and drug resistance. These type of research also gives critical insight into the evolutionary forces that have acted to rewire the regulatory networks during adaptation of species[76].

2. Materials and Methods

2.1. Strains, plasmids and growth media

In this work, it was used a cohort of 13 *C. glabrata* strains which are listed in Table 1. The yeast *Saccharomyces cerevisiae* BY4741, acquired from the Euroscarf collection, was also used.

Table 1 - *C. glabrata* strains and *S. cerevisiae* strain used in this work.

Strain	*Description*
FFUL887	FFUL clinical isolate, in study strain
CBS138	CBS clinical isolate and reference strain
KUE100	Wild-type strain
BY4741	MATa, his3Δ1, leu2Δ0, met15Δ0, ura3Δ0
*KUE100 Δ*CgA	Δ*CgA* strain, the promoter region of the *CgA* gene was replaced by a tetracycline-repressible promoter
KUE100 ΔB	Δ*B* strain, the promoter region of the *B* gene was replaced by a tetracycline-repressible promoter
KUE100 ΔC	Δ*C* strain, *C* gene was deleted
KUE100 ΔD	Δ*D* strain, *D* gene was deleted
KUE100 ΔE	Δ*E* strain, *E* gene was deleted
KUE100 ΔF	Δ*F* strain, *F* gene was deleted
KUE100 ΔG	Δ*G* strain, *G* gene was deleted
KUE100 ΔH	Δ*H* strain, *H* gene was deleted
KUE100 ΔI	Δ*I* strain, the promoter region of *CgI* was replaced by a tetracycline-repressible promoter
KUE100 ΔJ	ΔJ strain, J promoter region was deleted

C. glabrata cells and *S. cerevisiae* BY4741 were batch-cultured at 30ºC, with orbital agitation (250rpm) in Yeast Peptone Dextrose (YPD) or in Minimal Medium (MM). YPD growth medium contains, per liter: 20 g glucose (Merck), 20 g yeast extract (Difco) and 10 g bactopeptone (Difco). MM growth medium contains, per liter, 20 g glucose (Merck), 1.7 g Yeast Nitrogen Base without amino acids and ammonium sulphate (Difco) and 2.65 g ammonium sulphate (Merck). To grow *S. cerevisiae* cells MM medium was further supplemented with 20 mg/L histidine, 60 mg/L leucine and 20 mg/L methionine to complement the auxotrophies of the BY4741 strain. Solid YPD or MM media were obtained by supplementing the corresponding liquid growth medium with 2% agar (Iberagar). YPD and MMB media were sterilized by autoclaving them for 15 minutes at 121ºC and 1 atm. The plasmid used, pYEP354W, in this study is described in Table 2.

Table 2 – Plasmid, pYEP354W used in this study (Annex V).

Plasmid	*Description*
pYEP354W	Yeast episomal vector with *URA3* selection marker and *lacZ* gene. Used for construction of promoter regions pYEP354W-*MET10* and pYEP354W-*PDH1* vectors.

2.2. Genomic DNA extraction

Gnomic DNA was extracted from FFUL887 and CBS138 cells cultivated in solid YPD. After growth, three loops of biomass were inserted in a 1 mL-Eppendorf tube with approximately 100µl of glass beads (0.5mm) and 200µl of lysis buffer (Tris 50 mM, EDTA 50 mM, NaCl 250 mM, SDS 0.3%). The tubes were vortexed for 2 minutes at maximum speed and then incubated at 65ºC for 1h, after which a second round of vortexing was performed. The obtained disrupted cell suspension was then centrifuged at 13000 rpm for 15 minutes at 4ºC and the supernatant transferred to a clean eppendorf. 20µl NaAC 3M (pH 4,8) and 400µl cold Ethanol were added to the suspensions to induce DNA precipitation. The samples were left at -20ºC for, at least, 30 minutes after which these were centrifuged at 13000 rpm, during 20min at 4ºC. The pellet obtained was washed with 500µl ethanol 70%, dried in the speed vacuum and finally resuspended in 50µl deionized water.

2.3. In vitro DNA manipulations

The PCR reactions were performed using Taqmed polymerase (Citomed) and a set of primers that were designed to specifically hybridize in the selected promoter regions of the genes *CDR1*, *PDH1*, and *MET10*. Beside the region that allows PCR amplification, the selected primers also contain the necessary regions of homology with the plasmid (pYEP354W), essential for homologous recombination. The nucleotide sequences of the primers used are indicated in Table 3, the reaction mixture used and the experimental conditions used in the PCR amplification are described in Tables 4 and 5, respectively. Amplification of the fragments was confirmed by running the PCR products in a 0.8% agarose gel. The obtained fragments, having a size of approximately 1000 bp, were excised and purified from the agarose gel using the JETQUICK Gel Extraction Spin kit (Genomed). The purified DNA was then concentrated in a speed vacuum and finally resuspended in 10 µl ddH_2O. The purified DNA was stored at -20ºC until further use.

The pYEP354W plasmid, 1 µg, was digested with 15 U of HindIII (Takara) at 37ºC overnight. After this period, the digested product was then incubated with 1 U of Calf Intestinal Alkaline Phosphatase (CIAP) (Invitrogen) at 37ºC for 45 minutes to prevent re-circularization.

Table 3– Primers used in the amplification of the promoter fragments. Sequence includes specific promoter regions of each gene (red) and the homology region with the plasmid pYEP354W (black).

***CDR1*_Fw:** TCATTCCCGGAATTCCCGGGGATCCGTCGACCTGCAGCCA TAGCGCATGGAAATCCTTGG
***CDR1*_Rev:** TCCCAGTCACGACGTTGTAAAACGACGGCGGGAGCAAGCTT CTTGTCACTTGCAAGAGACA
***PDH1*_Fw:** TCATTCCCGGAATTCCCGGGGATCCGTCGACCTGCAGCCA TTGAACTATCCCAACACCAG
***PDH1*_Rev:** TCCCAGTCACGACGTTGTAAAACGACGGCGGGAGCAAGCT TAGAGTCATCGGGTGTGTTCA

***MET10*_Fw:** TCATTCCCGGAATTCCCGGGGATCCGTCGACCTGCAGCCA GTAAGTATGTCCCAGTATAC
***MET10*_Rev:** TCCCAGTCACGACGTTGTAAAACGACGGCGGGAGCAAGCT TCACAGACATTGCTTGGCTTT

Table 4 – Reaction mixture used for the amplification of Cg*CDR1*, Cg*PDH1* and Cg*MET10* promoters in a total volume of 50 µl.

Component	Volume per reaction
dNTP's	1 µl
Primer forward	1 µl (50 pMoles)
Primer reverse	1 µl (50 pMoles)
Template DNA	1 µl (~200ng)
$MgCl^{2+}$	2.5 µl (50mM)
Taqmed buffer	5 µl
Taqmed polymerase	0.5 µl (5U/µl)
Water	38 µl
Total	50 µl

Table 5– Conditions used in the thermocycler to allow amplification of the different promoters.

Time	Temperature	
1'	94ºC	
1'	94ºC	x30
40''	50ºC	
1'	72ºC	
7'	72ºC	
∞	4ºC	

2.4. Yeast transformation

The amplified promoter regions and the digested pYEP354W plasmid were transformed in *S. cerevisiae* BY4741 cells using the Alkali-Cation TM Yeast Transformation kit (MP Biomedicals). For this, yeast cells were cultivated until mid-exponential phase (OD600 nm=0.4) in 50 mL of liquid YPD and then centrifuged for 5 minutes at 6000 rpm and 4ºC. The obtained pellet was resuspended in 4.5 mL TE (pH 7.5) and then centrifuged again using the same conditions described previously. The pellet obtained was resuspended in 2.5 mL of Lithium/Cesium Acetate Solution and the obtained suspension was incubated at 30ºC for 30 minutes with orbital agitation (100 rpm). After this time, the cells were centrifuged again using the same conditions described above and the pellet obtained was resuspended in 500 µL of TE (pH 7.5). After this set of steps, cells are now considered to be competent and ready for transformation.

Each transformation mixture used 100 µL of competent cells, 5 µL of carrier DNA, 5 µL of histamine solution, 2 µL of digested pYEP354W and 10 µL of the purified promoter fragments.

The mixture was incubated at room temperature for 15 minutes with gentle mixing. Afterwards, 0.2 mL of TE/Cation Mix and 0.8 mL were added to each reaction and were incubated for 10 minutes at 30ºC. Cells were heat shocked for 10 minutes at 42ºC followed by cooling to 30ºC for 5 minutes. The reaction mixtures were centrifuged at 13000 rpm, for 30 seconds, and the pellet obtained was resuspended in 100 µL of YPD and finally plated on MMB plates.

Confirmation of the integration of the different promoters in the pYEP354W plasmid in the clones that grew in MM plates was performed by colony PCR. For this, a loop of biomass obtained from each candidate was resuspended in 10 µL of 20mM NaOH and incubated at 100ºC for 10 minutes. The suspension was afterwards centrifuged for 1 minute in a table-top centrifuge at 13000 rpm. 5 µL of the supernatant was used for the subsequent PCR reaction, which was performed under the same conditions described in Table 3 and Table 4. Colonies in which the amplification of a ~1000 bp product was observed were considered to be positive candidates.

2.5. Fluconazole, voriconazole and caspofungin resistance assays using *C. glabrata* FFUL887 derived mutants and *C. glabrata* KUE100 strain in microplate

Susceptibility of KUE100 and of the derived deletion mutants to fluconazole, voriconazole and caspofungin was tested using an experimental setup based on 96-microwell plates as schematically represented in Figure 10. Because some of the mutants were unable to grow in MM, these were tested in the rich medium YPD. Three concentrations of each drug were tested, the concentration that is known to correspond to the clinical resistance breakpoint (as defined by EUCAST, www.eucast.org); one below and one above that resistance breakpoint (Table 6). The stock solutions of the drugs (all having 5 mg/mL) were prepared in DMSO (Sigma). The cells used to inoculate the microplates were cultivated in 50 mL of YPD or MM at 30ºC and 250 rpm orbital agitation, for approximately 7h. Cell suspensions of *C. glabrata* KUE100 and each derived mutant were prepared in a stock solution containing 10 mL of 1X MM or YPD for an initial DO_{600} of approximately 0.05. Each column of the microplate was inoculated with 50 µL of these cell suspensions for a total volume of 200 µL in each well, as represented in Figure10 and Table 6. The microplates were cultivated for 24h for fluconazole (FLC) and 48h for caspofungin (CASPO) and voriconazole (VC).

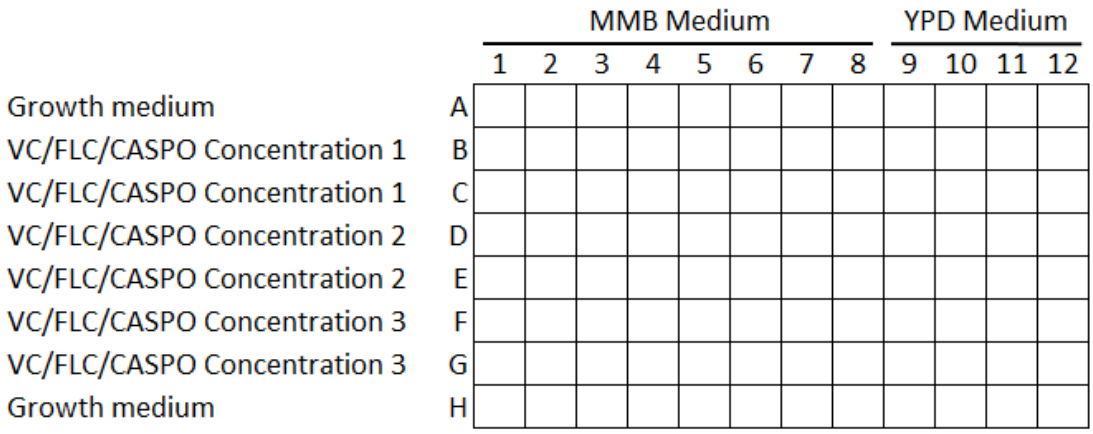


Figure 10 - Schematic representation of the 96-well microplate used to test the susceptibility of KUE100 strain and of the derived deletion mutants to voriconazole, fluconazole and caspofungin. Strains distribution: 1 - KUE100; 2 - KUE100 Δ*B*; 3 - KUE100 Δ*C*; 4 - KUE100 Δ*D*; 5 - KUE100 Δ*G*; 6 - KUE100 Δ*H*; 7 - KUE100 Δ*I*; 8 - KUE100 ΔJ; 9 - KUE100 Δ*CgA*; 10 - KUE100 Δ*E*; 11 - KUE100 Δ*F*; 12 – Medium only.

Table 6 - Concentrations used for the antifungal susceptibility tests performed. Concentration 2 is equivalent to the defined clinical resistance breakpoint, as defined by EUCAST (www.eucast.org). As in the case of caspfungin there is no defined clinical resistance breakpoint, the recommended ECOFF value, as defined by EUCAST, was used.

	Concentration 1 (µg/L)	*Concentration 2 (µg/L)*	*Concentration 3 (µg/L)*
Voriconazole	0.5	1	1.5
Fluconazole	16	32	64
Caspofungin	0.0625	0.125	0.25

2.6. Testing of the "white/dark brown" phenotype

For the white/dark brown phenotype assay, solid YPD was supplemented with 1mM sterile $CuSO_4$ after medium sterilization. FFUL887, CBS138 and KUE100 *C. glabrata* cells were cultivated in liquid YPD growth medium until mid-exponential phase (DO_{600} of 0.4). After reaching the desired DO_{600} of 0.4, 1mL of each culture batch was used to performed dilution series to a final dilution of 10^{-6}. From these dilutions, only 10^{-4}, 10^{-5} and 10^{-6} cell suspensions were plated into solid YPD supplemented with 1mM of CuSO4. Cells were then incubated at 30ºC for 5 days and daily observations were performed to evaluate the phenotype switching progression between strains.

2.7. Spot assays

For the spot assay, solid MMB medium was prepared without glucose as a carbon source, instead three different concentrations of acetate, butyrate and propionate were used as carbon sources: 0.2%; 0.5% and 2%. FFUL887 and CBS138 strains were cultivated in liquid MMB until-mid-exponential phase (DO_{600} of 0.4). After reaching DO_{600} of 0.4, cell suspensions were prepared for an initial DO_{600} of 0.05 from the original suspension. Two more dilutions were performed for 1:5 and 1:25 from the initial DO_{600} of 0.05 cell suspension. Afterwards, 4µL of each cell dilution was used to plate in solid MMB medium with the different acetate, butyrate and propionate concentration. Cells were then incubated for 3 days at 30ºC.

3. Results and discussion

3.1. Overview

To identify and subsequently characterize the alterations occurring in the non-coding genome of the FFUL887 isolate, the assembled reads obtained after sequencing of this isolate's genome[73] were compared with the publicly available genome of the CBS138 reference strain. The results of this comparative analysis focused on the non-coding genomes of the two strains are schematically represented in Figures 11 and 12 and summarized in Table 7. A total of 35438 SNPs were identified in the non-coding genome of the two strains strains, the vast majority of these being found in gene promoters (which were considered to the composed by the 1000 bp located upstream of each ATG start codon) (Table 7). One previously identified ncRNAs[95] CaglfMr14, described as a component of mitochondrial 15S rRNA, was also found to differ in two nucleotides (A19324G, C19449T) in the FFUL887 isolate and in the CBS138 strain.

Table 7 – Overview of the number of SNPs identified in the coding and non-coding genome of the FFUL887 isolate, when compared with the genome sequence of the reference strain CBS138[73].

	Number of SNPs
Coding regions[73]	79079
Non-coding regions	35438
Non-coding RNAs	2

Using a bioinformatics script and the promoter sequences of the CBS138 strain that are available at the RSA tools website (http://rsat-tagc.univ-mrs.fr/rsat/) it was possible to reconstruct all the promoter regions of FFUL887 genes. Closer comparison of these promoters with those of the CBS138 strain showed that around 73% of all genes predicted to be encoded by the genome of the FFUL887 strain harboured at least one SNP in their promoter, when compared with their CBS138 counter-partners (Figure 11 and Figure 12). Most of the changed promoters differed by only 1 nucleotide (Figure 11 and Figure 12) but 26 genes were found to harbour more than 30 mutations in their promoters, the more evident case being the Cg*PWP4* gene whose promoter in the FFUL887 strain exhibited 136 different nucleotides comparing to the corresponding promoter found in the CBS138 genome (Figure 12). In Table 8 those promoters which changed the most between FFUL887 and CBS138 strains are further detailed. It was evident in this dataset the presence of promoters for genes encoding adhesins or proteins related with adhesion (Cg*EPA11*, CAGL0C00968g and CAGL0C00253g) (Table 8). The observation that the promoter region of genes involved in adhesion are apparently under a strong selective pressure is interesting considering that this was also one of the biological functions that were found to be under more pressure when only the coding sequences of CBS138 and FFUL887 were compared[73]. In fact, the modification of proteins involved in adhesion was found to be an hallmark of *C. glabrata* adaptation to the human host, this presumably reflecting the need of the microbe to adjust its surface characteristics to the specificities of the colonization niche[96]. This high rate of mutation in adhesion properties is in fact considered a primary mechanism of pathogenicity evolution in *C. glabrata*[97]*.* Adhesion proteins have an important role in crucial host-fungus interactions in facilitating the establishment of human mycoses[79,80]. In particular, the *EPA* genes family have been linked to the adhesion to human tissues and to other infection-related processes, such as invasion, biofilm formation and iron acquisition[96]. Up to now it has not been examined how the expression of adhesins is affected by the alterations in their promoter regions, although the results gathered in this study may suggest that may occur to adjust the expression of some adhesins in detriment of others. The modifications in the promoter region of these genes might contribute for their different transcriptional regulation which could thereby favour adaptation to the host, although this hypothesis has to be further confirmed in particular by comparing the expression of these adhesion-related genes in FFUL887 and CBS138 in different conditions. Another promoter which was also found to be under a strong selective pressure was the one of the *ICL1* gene, encoding a 2-methylisocitrate lyase. The Icl1 gene is expressed mainly in conditions of glucose limitation[98] providing to *C. glabrata* cells the enhanced ability to consume acetate and other intermediates of the glyoxylate cycle[98]. The promoter region of *CgSTE12* gene was also found to harbour a high number of SNPs. This gene encodes a transcription factor which was found to regulate genes involved in filamentous growth in response to nitrogen starvation and that was also described to play a role in *C. glabrata* virulence[99]. Furthermore, C*gSTE12* was also linked to epithelial adhesion during host adaptation[99]. The promoter region of the gene CAGL0C03916g, which encodes a protein with predicted role in protein glycosylation, was also

identified as having high number of mutations as well. However, little is known regarding the function of this protein in *C. glabrata*.

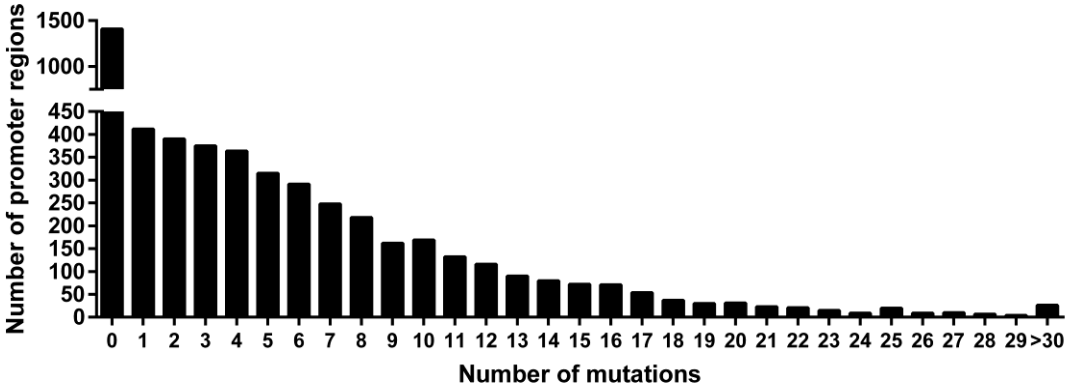


Figure 11 – Representation of the number of SNPs found in the promoter region of genes encoded by the FFUL887 when compared to CBS138 strain.

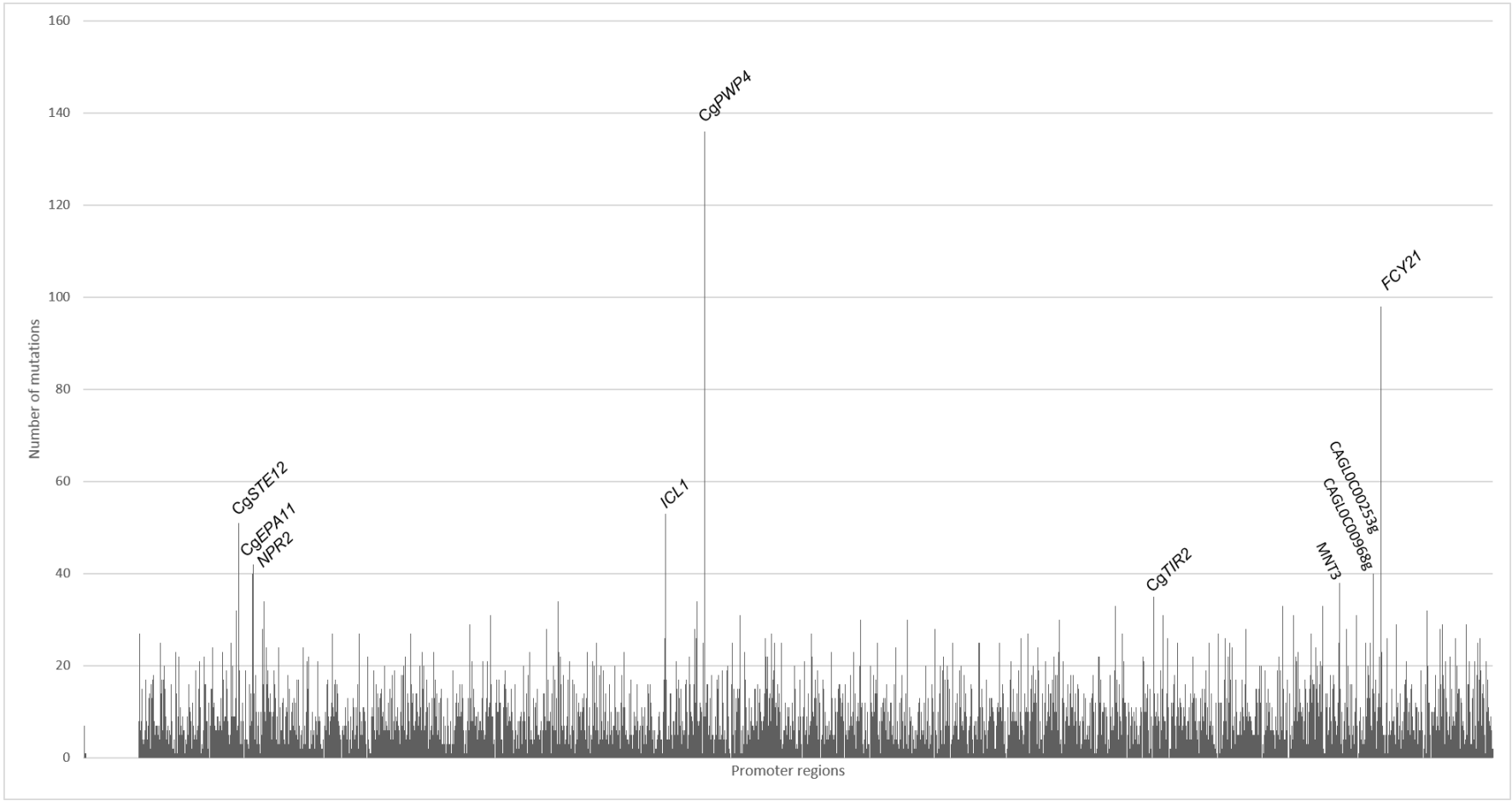


Figure 12 - Number of SNPS found across all gene promoters of the FFUL887 strain, when compared with their CBS138 counter-partners. Promotor regions were defined as being the 1000 bp nucleotides upstream of the ATG start codon. The names of the genes whose promoters changed the most in the FFUL887 strain (Table 2) are highlighted.

Table 8 - List of the genes whose promoters changed more significantly in the FFUL887, in comparison with their CBS138 counter-partners. The function of each gene was based on the information available at the *Candida* genome database.

ORF	Name	Function	Number of SNPs found
CAGL0I10362g	*CgPWP4*	Lectin-like cell wall protein (flocculin) involved in flocculation	136
CAGL0C00231g	*CgFCY21*	Putative purine-cytosine permease	98
CAGL0J03058g	*CgICL1*	Isocitrate lyase	53
CAGL0M01254g	*CgSTE12*	Putative transcription factor, required for filamentous growth induced by nitrogen starvation and for virulence	51
CAGL0C00253g		Putative cell wall adhesin	49
CAGL0L13288g	*CgNPR2*	Subunit of the Iml1p/SEACIT complex	42
CAGL0C00968g		Adhesin-like protein with a predicted role in cell adhesion	40
CAGL0L13299g	*CgEPA11*	Putative cell wall adhesin	40
CAGL0C03916g	*CgMNT3*	Role in protein glycosylation	38
CAGL0F01485g	*CgTIR2*	Putative GPI-linked cell wall mannoprotein of the Srp1p/Tip1p family	35

Previous analysis of the evolution of promoters in species of the *Saccharomycetaceae* family has showed that the different segments of a promoter evolve at a different rate. In specific, a lower degree of mutation in regions of the promoter that serve as binding sites for transcriptional regulators was observed[91]. In this context, to take a hint into how the SNPs found were distributed throughout FFUL887 gene promoters' these were divided in 50bp windows and the number of mutations in each of these windows was computed. The results of this analysis are shown in Figure 13 and represented in a more detailed manner in the heat maps shown in Figure 14. As it is not possible to show in this figure all the promoters, only those which changed the most in the FFUL887 and CBS138 strains are shown. The results obtained showed that the mutations in the promoter regions are distributed all over the genes' promoters, however, a slight increase in the frequency of mutation was observed to occur closer to the ATG start codon (Figure 13).

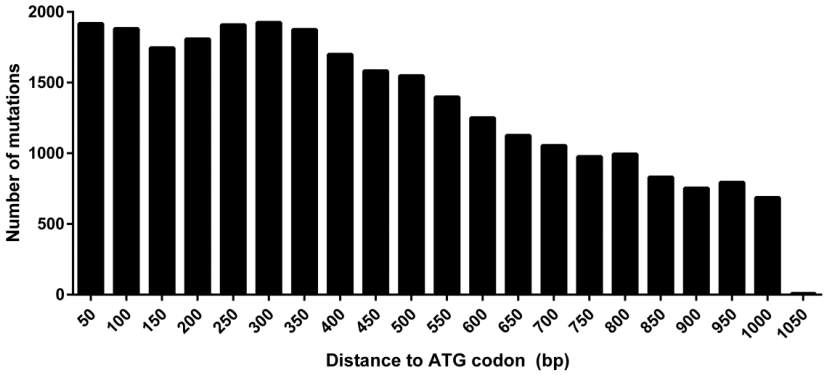


Figure 13 – Distribution of the SNPs throughout FFUL887 gene promoters. Each promoter region was divided in 50 bp segments and the number of SNPs (assessed upon comparing each promoter with its corresponding CBS138 counter-partner) was computed.

In fact, 1925 mutations identified in the promoters of FFUL887 strain are included in the 300 bp window, this number increasing to 13060 if we take into consideration the first 350 bp (Figure 13). Interestingly, a previous similar analysis undertaken within the Saccharomycetea family as shown that this initial 350 bp window of these yeast's gene promoters is particularly enriched in binding sites for transcription factors[91]. If such is the case for *C. glabrata*, then such a high rate of mutation in this region might reflect the capacity of the microbe to rewire its genomic expression program by extensively affecting the transcription factor-DNA interactions. On the overall the results render clear that gene promoters being regulatory regions appear to be highly plastic, although it is important to stress that the expression pattern of a certain gene and the regulatory system that governs its expression can remain intact even though mutations might occur in this region[75,91].

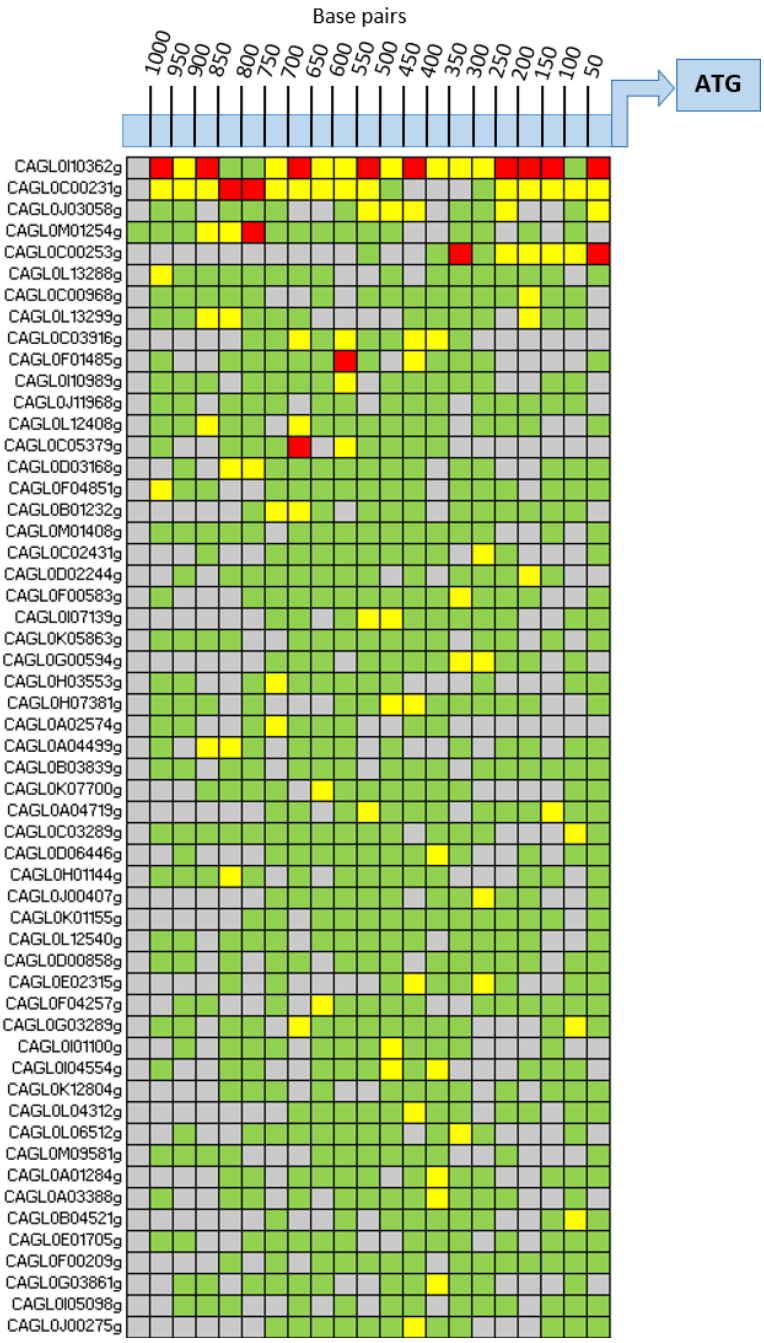


Figure 14 - Heatmap representation of the distribution of mutations found in the 55 promoters that changed the most between FFUL887 and CBS138 strains. Legend: Grey – number of mutations in the base pair window equal to 0; Green – number of mutations in the base pair window greater than 0 and lesser than 5; Yellow - number of mutations in the base pair window greater than 5 but lesser than 7; Red – number of mutations in the base pair window greater than 7.

3.2. Correlation between gene expression and the SNPs found in promoters of FFUL887 genes

To assess the relevance of the identified mutations in promoters of FFUL887 genes the results of a previous study that had compared the transcriptome of these cells with the one of CBS138 cells during cultivation in RPMI growth medium were used[73]. It is important to stress that this is only a snapshot of the overall transcriptomic landscape of the FFUL887 strain and many other differences in genomic expression could be obtained, however, it would be impossible to perform such assessment for a great variety of experimental conditions. In case the mutations in gene promoters of the FFUL887 strain are relevant for gene transcription then one could expect that the corresponding genes to be differently expressed in FFUL887 and in CBS138. In Figure 15 it is shown the correlation between the number of mutations found in each promoter with the difference in expression registered in the FFUL887 and in the CBS138 strains, according to the microarray analysis performed. No significant correlation between the number of mutations found in one promoter and the corresponding level of gene expression in the FFUL887 and in the CBS138 strains was observed, which was somehow expected since one single mutation could have a strong impact in expression. On the overall, we could identify 82 genes over-expressed and 173 repressed (bellow a 0.7-fold) in the FFUL887 strain that do not harbour mutations in their promoter regions (e.g. Cg*INO1 (*over-expressed 7.4-fold), *CgPIL1* (3.6-fold) and *CgYPS11* (3.4-fold). The higher expression of these genes in the FFUL887 strain represents precisely those cases in which a higher gene expression is achieved independently of the alteration of the promoter region, probably resulting from a higher activity of the transcription factors controlling the expression of these genes. Some of the genes whose promoters were heavily mutated were found to have little differences in expression in FFUL887 and CBS138 cells, at least during the experimental conditions that were used to assess the transcriptome of these two strains. This is the case for example of Cg*PWP4*, which harboured 136 mutations in its promoter and whose expression in FFUL887 was only reduced by 20%, comparing with the levels attained in CBS138. The expression of Cg*FCY21* gene, whose promoter harboured 98 SNPs, was also found to be 30% lower in the FFUL887 strain, comparing with the levels attained in the CBS138 strain. It will be interesting to compare the expression of these genes whose promoters were heavily mutated in FFUL887 strain using different experimental conditions.

Despite these observations, the analysis undertaken allowed the identification of a set of genes that were differently expressed (above a threshold of 3-fold) in FFUL887 and CBS138 cells and whose promoters also differed in these two strains, these being shown in a more detailed manner in Figure 16.

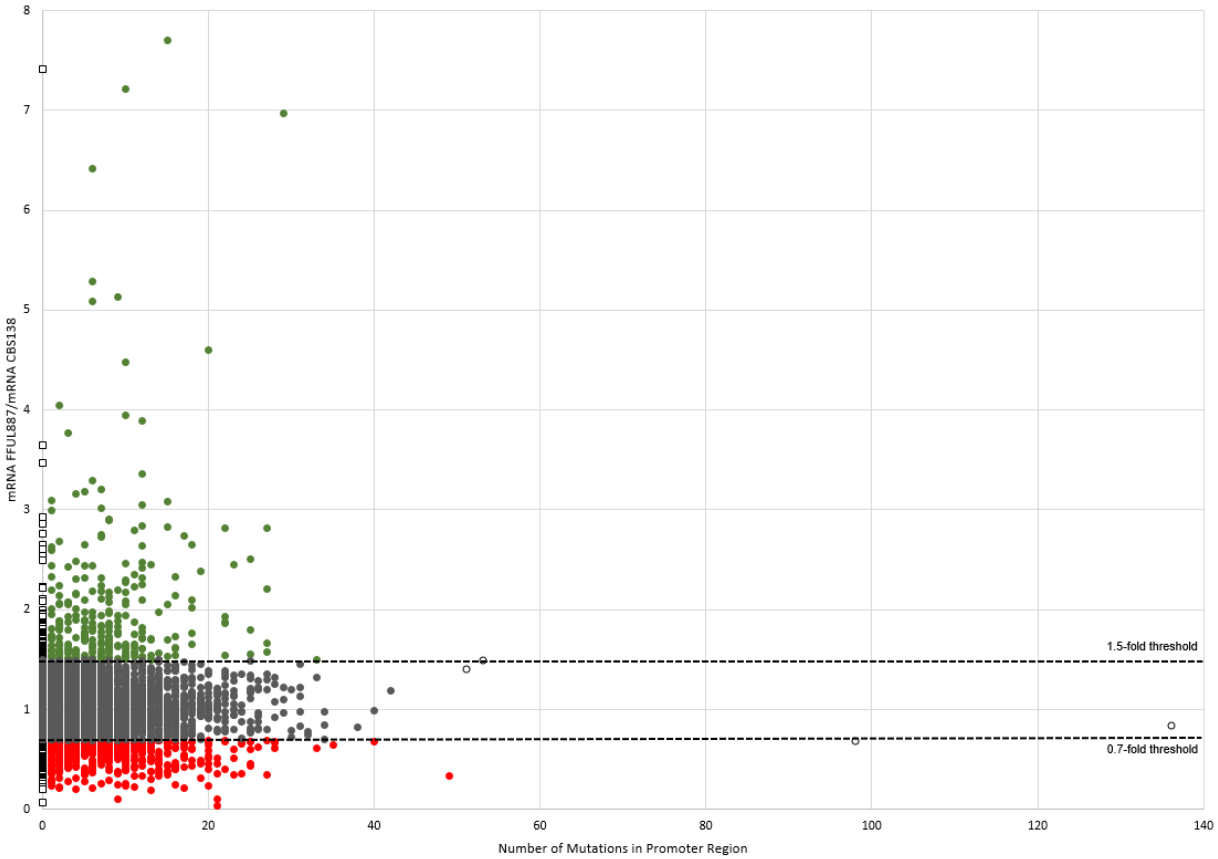


Figure 15 – Correlation between the numbers of mutations found harbouring promoter regions in FFUL887 when compared to CBS138 genome and the expression of the corresponding genes in the two strains during cultivation in RPMI growth medium. Genes that have no mutations in their promoter regions, but are up- or down-regulated in the FFUL887 strain are highlighted as open squares. The genes whose promoters were found to be more strongly mutated (above 50 mutations) are highlighted as open circles. Genes that are above the 1.5-fold threshold are considered up-regulated and represented in green circles. Genes that are below the 0.7 threshold are considered down-regulated and are represented in red circles. Genes that were found to have no alterations in their expression are represented in grey circles.

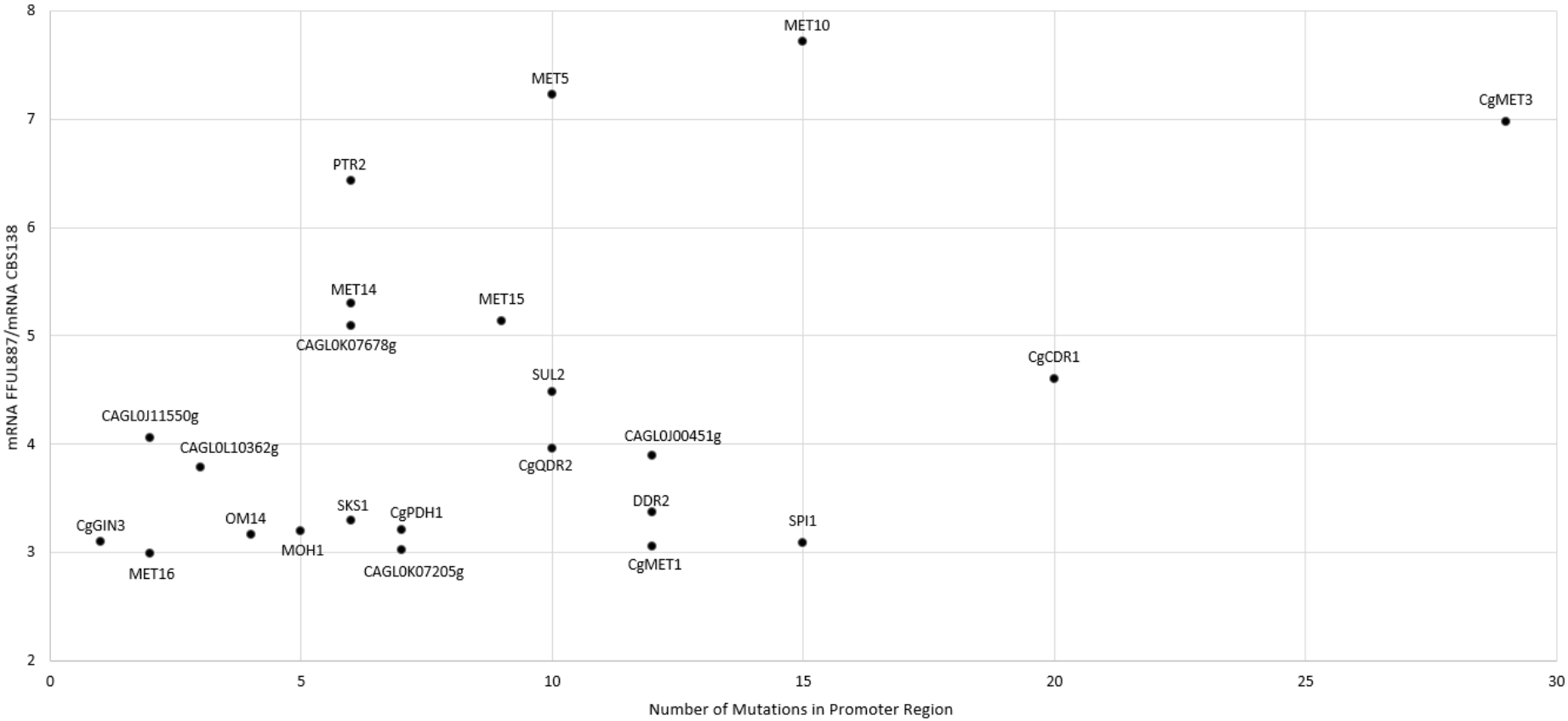


Figure 16 – Representation of the identified genes harbouring mutations in their promoter regions that were differently expressed (above a threshold of 3-fold) in FFUL887 and CBS138 cells.

Functional clustering of the genes up-regulated and whose promoters changed the most revealed a significant over-representation of genes involved in sulphate metabolism and, in particular, in biosynthesis of methionine: Cg*SUL1*, Cg*MET1*, Cg*MET3*, Cg*MET5*, *CgMET10*, Cg*MET14*, Cg*MET15*, Cg*MET16* (Figure 16). Three multidrug-resistance transporters, CgQDR2, CgPHD1 and CgCDR2 were also found in this dataset as well as Cg*PTR2*, involved in transport of peptides; Cg*OM14*, whose *S. cerevisiae* orthologue encodes a mitochondrial outer membrane receptor for cytosolic ribosomes; Cg*SPI1*, which *S. cerevisiae* orthologue encodes a GPI-anchored cell wall protein involved in weak acid resistance; Cg*SKS1*, which *S. cerevisiae* orthologue is involved in the adaptation to low concentrations of glucose and *DDR2*, whose *S. cerevisiae* is involved in response to xenobiotic agents and environmental or physiological stresses. The over-expression of the genes related with biosynthesis of sulphur amino acids in the FFUL887 strain was particularly interesting as the higher expression of these genes was previously associated with the transition of the white to brown" phenotypic switch in *C. glabrata*[77]. These studies implicated the up-regulation of Cg*CBF1* and Cg*AMT1* in the up-regulation of these genes[77]. Cg*CBF1* exhibited 20% less expression in the FFUL887 strain but Cg*AMT1* was 80% more expressed in this strain, comparing with the levels attained in the CBS138 strain. Altogether these observations lead to the hypothesis that the white-brown phenotypic switching could be altered in the FFUL887 strain. To test this hypothesis FFUL887 and CBS138 cells were cultivated in YPD rich growth medium supplemented with 1 mM $CuSO_4$, a compound that is widely used to induce white to brown phenotype in *C. glabrata* (Figure 17)[77,78]. As a positive control, the KUE100 strain was used since this strain was previously observed to exhibit a fast transition to the brown phenotype (unpublished observations from our laboratory). In the figure below are shown some pictures of the colonies that were obtained in the Cu-supplemented plates, being evident a stronger brown colour observed in FFUL887 colonies, comparing with the one observed in colonies of the CBS138 strain. Interestingly, all colonies from the same strain displayed the same phenotype during the observed time-lapse.

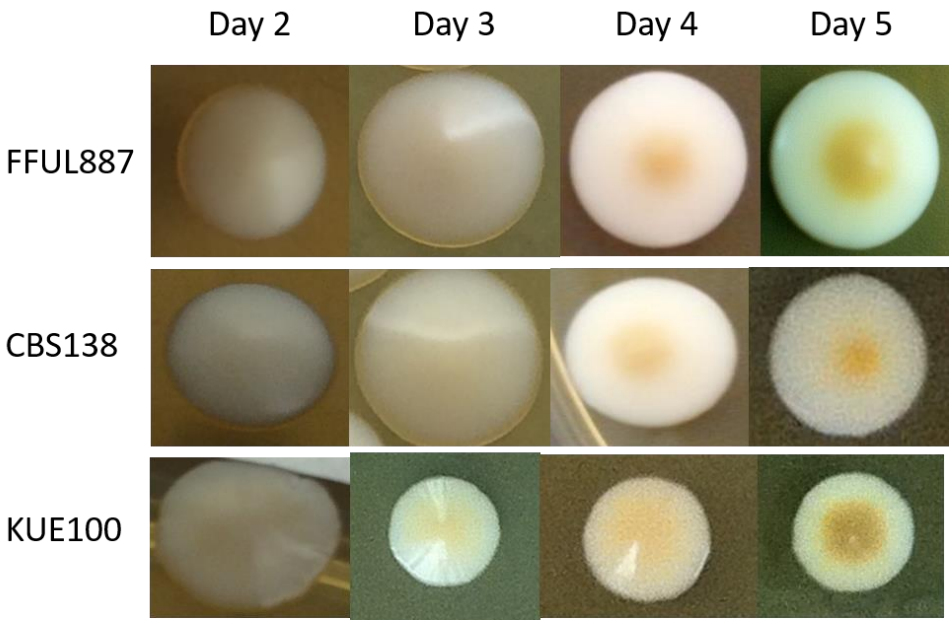


Figure 17 - Dark Brown phenotype assay with FFUL887, CBS138 and KUE100 strains. Cells were grown in YPD solid medium supplemented with 1mM $CuSO_4$ and incubated for 5 days at 30ºC.

The higher expression in FFUL887 cells of the genes shown in Figure 16 could result from the alteration(s) that were registered in their promoter or from the activity of the different factors involved in their transcriptional regulation, or even from the combination of these two things. To better understand what is the effect in gene expression of the mutations observed to occur in the FFUL887 strain, an experimental strategy based on the construction of a different set of lacZ fusions was used. In specific, in this strategy it was attempted to clone the promoter region of Cg*MET10*, Cg*PHD1* and Cg*CDR1* genes obtained from the genome of the FFUL887 or of the CBS138 strains upstream of the *lacZ* gene present in the pYEP354w plasmid (map in Annex V). The experimental strategy that was used is schematically represented in Figure 18. The promoter of *CgMET10* was selected because this gene was among those having the highest number of mutations in their promoter and Cg*CDR1* and Cg*PDH1* were chosen due to their important role in azole resistance, a phenotypic trait exhibited by the FFUL887 isolate[27,31,58,73].

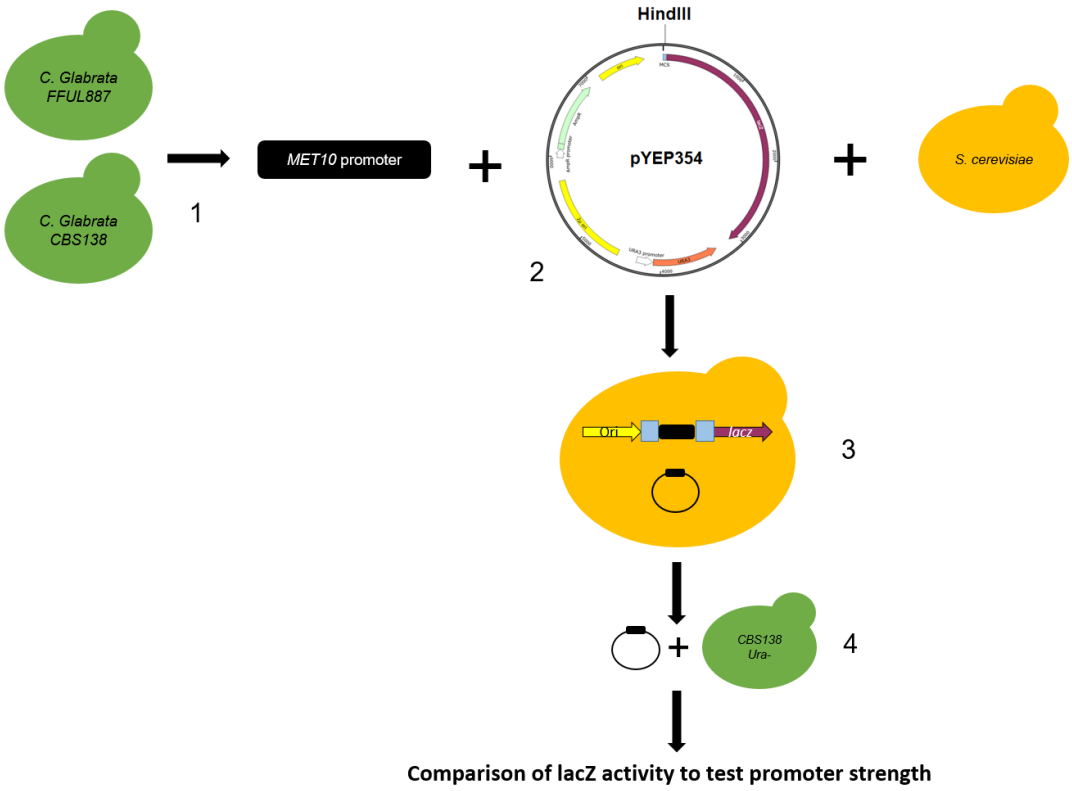


Figure 18 – Genetic engineering strategy used to clone the promoter region of *MET10* gene in the pYEP354w vector by homologous recombination in *S. cerevisiae*. 1- Amplification of promoter region by PCR; 2 – Digestion of the plasmid; 3 – Homologous recombination in *S. cerevisiae* (representation of linearized plasmid after insertion of the promoter region: Yellow arrow refers to *ori* region; blue blocks refer to multiple cloning site region, open in HindIII cutting region; black rectangle refers to MET10 promoter region; purple arrow refers to l*acZ* region); 4 – Expression test in *C. glabrata* CBS138 Ura$^-$ strain using the cloned promoters after the selection of positive candidates in colony PCR. The same strategy was used to clone the promoters of Cg*CDR1* and Cg*PDH1* genes.

The promoters of interest were at first amplified, by PCR, using as a template the genomic DNA of the FFUL887 strain. The result obtained is shown in the agarose gel shown in Figure 19. After recovery from the gel, and subsequent purification, this PCR product was transformed into *S. cerevisiae* cells together with the HindIII-digested pYEP354w plasmid. The experimental conditions used for the PCR reaction only allowed the amplification of the promoters of Cg*PDH1* (lane 6 and 7) and Cg*MET10* (lane 2 and 3) and therefore it was decided to move forward only using these two promoters. After two days of incubation at 30ºC, the plates obtained from the transformations performed with Cg*MET10* and Cg*PDH1* promoters had about 60 colonies each. The correct insertion of each promoter in the pYEP354w plasmid in these candidates was tested by colony PCR (Figure 20). The results obtained allowed the identification of three positive candidates for Cg*MET10* (Figure 20A) and three also for Cg*PDH1* (Figure 20B), suggesting that only these colonies have the promoter region fused with pYEP354w plasmid. At the time this thesis was being written the correct insertion of the two promoters in the plasmid was being confirmed by Sanger sequencing. Due to the lack of time it was not possible to

further continue this analysis, however, if the two constructs herein obtained are confirmed to be correct these will be used to transform the CBS138 strain to compare the strength of the FFUL887/CBS138 promoters in driving expression of the lacZ gene, as indicated in Figure 18.

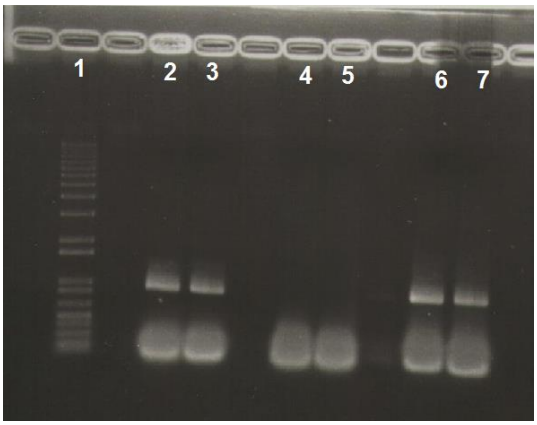


Figure 19 - Agarose gel showing the result of the amplification of Cg*MET10*, Cg*CDR1* and Cg*PDH1* promoters using as template the genomic DNA of the FFUL887 strain. 1 – 1kb DNA plus ladder; 2 and 3 – Cg*MET10* amplification result, fragment size expected to be around 1000 bp; 4 and 5 – Cg*CDR1* amplification result, fragment size expected to be around 1000 bp; 6 and 7 – Cg*PDH1* amplification result, fragment size expected to be around 1000 bp.

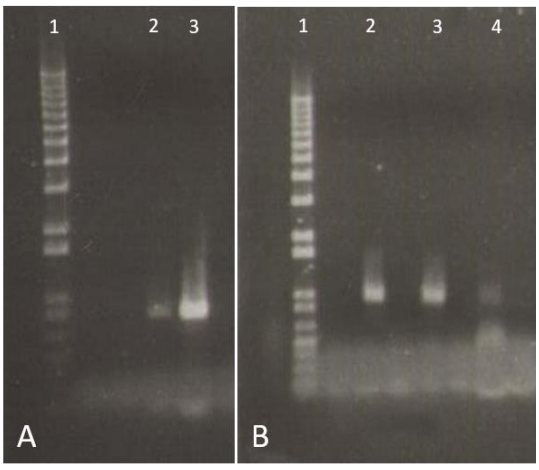


Figure 20 – Agarose gel showing the result of the Colony PCR amplification for *MET10* (A) and Cg*PDH1* (B). A: 1 – 1kb DNA plus ladder, 2 and 3 – amplification of CgMET10 from the CgMET10-pYEP354w candidates; B: 1- 1kb DNA plus ladder, 2, 3 and 4 – amplification of CgPDH1 from the CgPDH1-pYEP354w candidates.

3.3. Effect of the SNPs found in gene promoters in the overall regulatory network controlling the genomic expression of the FFUL887 and CBS138 strains

The observed modifications in the promoter region of FFUL887 genes, when compared to the gene promoters of the CBS138 strain, could contribute for the remodelling of the networks governing the control of gene expression in these strains for example by creating/inactive DNA motifs that may serve as binding sites for transcription factors. To see how the alterations observed to occur in FFUL887 gene promoters could alter the potential regulatory associations between target genes and transcription factors in this strain, the promoters of this strain were searched for DNA motifs that may serve as binding sites for predicted *C. glabrata* transcription factors. The identification of these transcription factors was performed using the information present in the Candida Genome Database and using as a reference the set of transcription factors described in *S. cerevisiae*, closely related to *C. glabrata*. Since the binding sites of most *C. glabrata* transcription factors are not yet characterized, this analysis was performed using the DNA sequences that are known to be recognized by the corresponding *S. cerevisiae* orthologues. It is important to stress that in the majority of the cases the *C. glabrata*/*S. cerevisiae* orthologous transcription factors pairs that were examined exhibited very strong homologies at the level of the DNA binding domain, thereby turning reasonable to hypothesize that these may recognize similar DNA sequences[75,90]. The exception to this were the transcription factors Cg*PDR1*, involved in the control of drug resistance in *C. glabrata*, and Cg*MSN4*, involved in the control of response to environmental stress, whose binding sites were already characterized being HYCCRKGGRN and CCYYCCYYM, respectively[100,101]. The association between *S. cerevisiae* transcription factors and their recognition sequences was based on the information available at the YEASTRACT database (http://www.yeastract.com/consensuslist.php). In the end, this genome-wide search for transcription factors' binding sites establishes the basis of what can be the potential regulatory network of the FFUL887 strain.

The results of the search for TF binding sites in the promoter region of FFUL887 genes is summarized in Table 9 and in Figure 21, the full data being available in supplementary table shown in Annex I. Around 8042 new putative TF binding sites were found in FFUL887 gene promoters, these corresponding to motifs that were absent in the corresponding CBS138 promoters. On the other side, 3033 putative TF binding sites were lost in gene promoters of the FFUL887 strain. This difference in the number of putative TF binding sites corresponds to 2511 (48%) promoters being identical in the two strains, 914 (17%) promoters that gained binding sites in the FFUL887 isolate, 1383 (26%) that lost binding sites in the FFUL887 isolate and 491 (9%) that suffered both types of modifications (Figure 21).

Table 9 – Total number of identified genes with new and lost consensus binding sites in the respective mutated promoter region in FFUL887 when comparing with CBS138.

	Number of promoter region
Promoter regions with new binding sites	914
Promoter regions with binding sites lost	1383

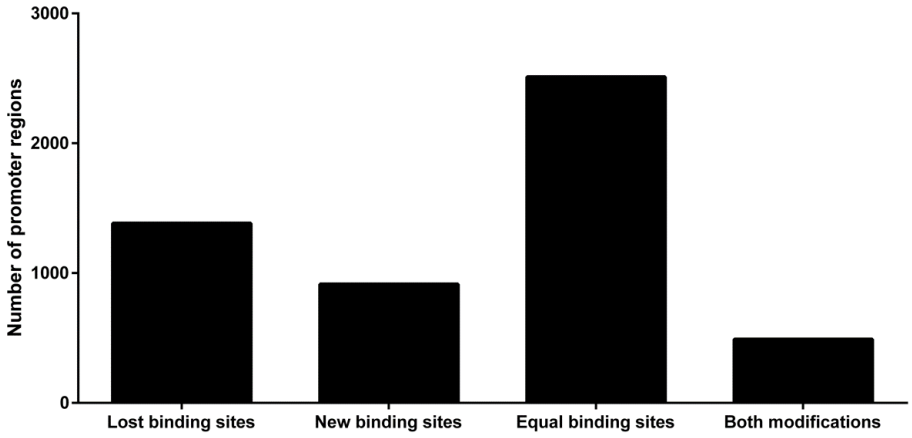


Figure 21 – Representation of the number of FFUL887 gene promoters that lost, gained and equal binding sites for transcription factors. The promoter regions of FFUL887 and CBS138 genes were searched for DNA motifs presumed to serve as binding sites for *C. glabrata* transcription factors

In Figure 22 it is shown a representation of those binding sites that appeared/disappeared with higher frequencies in the genome of the FFUL887 isolate, in comparison with the genome of the reference strain. In general the same DNA motifs that emerged within the dataset of those that were “gained” by the FFUL887 strain were also highly prevalent in the dataset of those that were lost in this same strain (Figure 22). The relative high degeneracy of some of these DNA motifs (such as “RATCA” (Ash1) and “RYMAAYA” (Fkh1)) may explain the high frequency of their existence within the promoter regions in general.

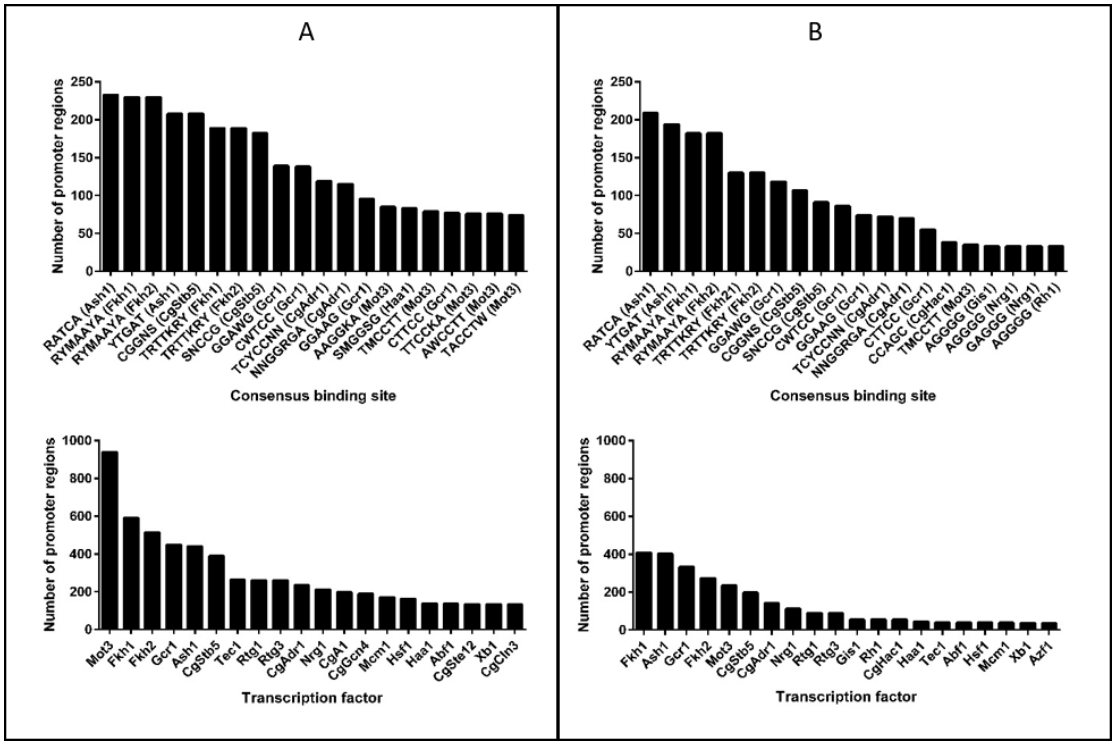


Figure 22 – Representation of gained binding sites (A) and lost binding sites (B). Top graphic in box A refers to the consensus binding sites with more frequency of appearance in FFUL887 promoter regions, while bottom graphic in box A refers to transcription factors whose recognized consensus binding sites were more gained in FFUL887 promoter regions. Top graphic in box B refers to the consensus binding sites with more frequency of disappearance in FFUL887 promoter regions, while bottom graphic in box B refers to transcription factors whose recognized consensus binding sites were more lost in FFUL887 promoter regions.

It is interesting to observe that the transcription factors whose number of binding sites increased/decreased the most in the genome of the FFUL887 strain have biological functions that are relevant for adaptation of *C. glabrata* isolates to the human host including pseudohyphal differentiation (Ash1, Fkh1, Fkh2, Tec1, Nrg1, CgSte12), metabolism of alternative carbon sources (Adr1), stress response (Hsf1, Haa1, CgAp1). The CgStb5 transcription factor, involved in resistance of *C. glabrata* to azoles[69], was also found to be over-represented in the dataset of TFs which gained more binding sites in the FFUL887 strain. The "new possible target genes" of these transcription factors were clustered in MIPS Functional Catalogue (http://mips.helmholtz-muenchen.de/funcatDB/) and the more relevant results are shown in Table 10. None of the new genes identified as having a CgStb5 binding site in the FFUL887 strain has a biological function associated to drug resistance in *C. glabrata*. Remarkably, in the FFUL887 the gene encoding the Ste12 has a now a DNA motif that serves as a binding site for this regulator, suggesting that auto-regulation may occur (Table 10).

Table 10 – Genes with new consensus binding site in FFUL887 of the more represented transcription factors in pseudohyphal growth and metabolism of alternative carbon sources.

	Regulator	*ORF*	*Name*
Pseudohyphal growth	Ash1	CAGL0G03597g	*CgSHO1*
		CAGL0L00693g	*CgDFG10*
	Fkh1	CAGL0K02673g	Cg*STE20*
	Fkh2	CAGL0H08129g	*CgDFG16*
	CgSte12	CAGL0M01254g	Cg*STE12*
Metabolism of alternative carbon sources	Adr1	CAGL0H09460g	*CgFAA2*

Interestingly, Cg*SHO1* gene harbours an Ash1 binding site (Table 10) in its promoter region, suggesting that a link between pseydohyphal growth and the HOG pathway, which controls osmotic regulation, in FFUL887 isolate. The promoter region of Cg*DFG10* and Cg*DFG16* genes, already described to be involved in invasive and pseudohyphal growth, were also found to harbour binding sites for Ash1 and Fkh2 in the FFUL887 strain (Table 10). This computational analysis also lead to the identification of a binding site for Adr1 in the promoter region of Cg*FAA2* gene, which encodes a medium chain fatty acyl-CoA synthetase and is involved in fatty acid metabolism.

It is clear that the above reported modifications create a very strong alteration in what can be the potential transcriptional regulatory network of the FFUL887 isolate. To get insights into how these modifications may have affected gene expression in the two strains under scrutiny, the dataset of genes that were differently expressed in the CBS138 or in the FFUL887 isolate was re-examined with the objective of identifying eventual gains or losses of binding sites that could contribute to explain the different levels of gene expression registered in the two strains. This way, up- and down-regulated genes were clustered according to their gain or loss of binding sites harbouring their promoter region (Figure 23) and the ten promoter regions of up- and down-regulated genes that suffered the most gains of binding sites are represented in Table 11 and highlighted in Figure 23.

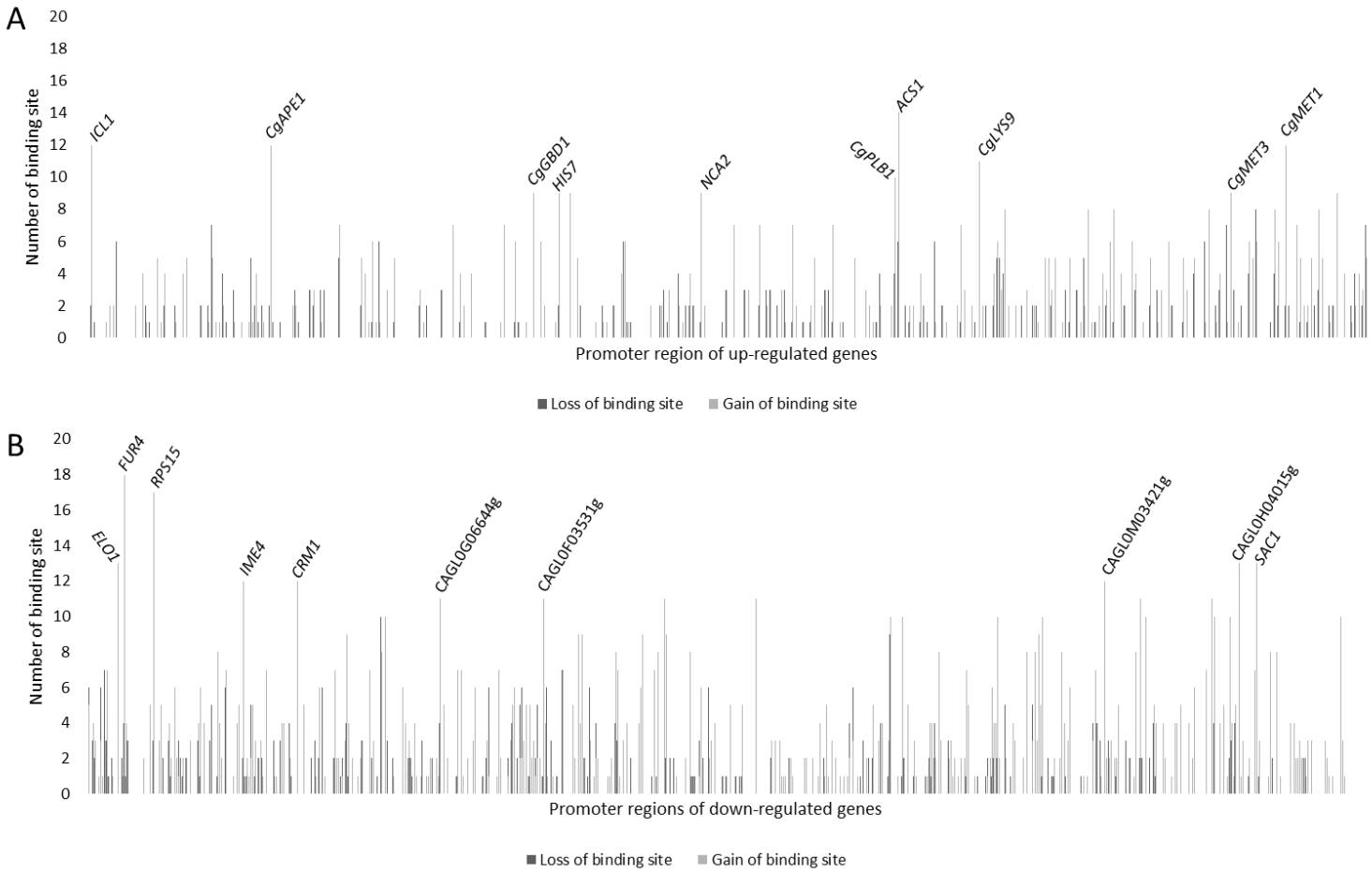


Figure 23 – Promoter regions of up- (A) and down-regulated (B) genes correlation with their respective number of gained and lost transcription factor's binding sites. Up- and down-regulated genes with increased number of gained binding sites in their promoter region are highlighted in the figure.

Table 11 – Promoter regions of up- (A) and down-regulated (B) genes that suffered the most gains of binding sites. Promoter region of up-regulated genes were searched using binding sites of transcription factors that act as activators, while promoter region of down-regulated genes were searched using binding sites of transcription factors that act as repressors.

A ORF	Name	mRNAFFUL887/ mRNACBS138	Binding Sites
CAGL0L00649g	*CgACS1*	2	GAGCCC (CgCrz1) WCAGCNDWAW (CgAmt1) ACGCGT (Mbp1) CYNNWWWRGGH (Mcm1) CYNNWWWRGG (Mcm1)
CAGL0G01903g	*CgMET1*	3	AGGGG (Gis1) CACGTG (CgGcn4) CAYGTGAY (CgCbf1) SMGGSG (Haa1) CACGTGA (CgMet4)
CAGL0J03058g	*CgICL1*	1.5	AGGGG (Gis1) ATCCCWTA (Gis1) SMGGSG (Haa1) NTTCNNGAAN (Hsf1)
CAGL0K08536g	*CgAPE1*	1.61	TCYCCNN (Adr1) CCCCT (Gis1) CSCCKS (Haa1) CCAGC (CgHac1) AAGGK (Mot3)
CAGL0C03443g	*CgLYS9*	2.11	CGGN{10}CCG (Cha4) GAGCCC (CgCrz1) RYMAAYA (Fkh1) CGGN{10}CCG (Put3) CGGN{10}CCG (Tea1) GGAWG (Gcr1)
CAGL0J11770g	*CgPLB1*	1.88	TCN{7}ACG (Abf1) NNGGRGA (Adr1) TCYCCNN (Adr1) TATACGA (CgUpc2B/2A) RYMAAYA (Fkh1) AAGGKA (Mot3) AAGGWT (Mot3)
CAGL0B03839g	*CgMET3*	6.97	AKAGTCA (CgGcn4) TRTTKRY (Fkh2) RYMAAYA (Fkh1) TTTGTMA (CgAp1)

B ORF	Name	mRNAFFUL887/ mRNACBS138	Binding Sites
CAGL0B04433g	*CgFUR4*	0.69	TRTTKRY (Fkh2) AGGGG (Nrg1, Rph1) GAGGG (Nrg1)
CAGL0F01045g	*CgRPS15*	0.39	TRTTKRY (Fkh2) TTCCKA (Mot3) CCCCT (Nrg1, Rph1) AGGGG (Nrg1, Rph1) CCCTC (Nrg1) GACGG (Nrg1) SNCCG (CgStb5)
CAGL0K00583g	*CgELO1*	0.46	CYNNWWWRGGH (Mcm1) ATGGGAT (Mot3) CCCCT (Nrg1, Rph1) AGGGG (Nrg1, Rph1) SNCCG (CgStb5)
CAGL0H04015g	-	0.6	RATCA (Ash1) RTCACRTG (CgCbf1) AAGAGG (Mot3) CCCTC (Nrg1) GAGGG (Nrg1)
CAGL0K00297g	*CgSAC1*	0.37	CCGN{3}CGG (Leu3) AGGGG (Nrg1, Rph1) GAGGG (Nrg1) CGGNS (CgStb5)
CAGL0M03421g	-	0.66	RYMAAYA (Fkh1) AWCCTT (Mot3) CGGNS (CgStb5)
CAGL0A03300g	*IME4*	0.51	AACGTGA (CgCbf1) CCYWWWNNRG (Mcm1) CAGGYA (Mot3) AGGGG (Nrg1, Rph1) GAGGG (Nrg1)

							CGGNS (CgStb5)
CAGL0C01595g	*CgHIS7*	1.71	GCTGC (CgHac1) TRTTKRY (Fkh2) RYMAAYA (Fkh1) AAGWT (Mot3) GGTAC (Rtg1/3) GGCYGGC (CgSkn7)	CAGL0B02189g	*CgCRM1*	0.63	RYMAAYA (Fkh1) GATAAG (Gzf3) GAGG (Nrg1) CGGNS (CgStb5)
CAGL0F04851g	*CgNCA2*	1.51	TCN{7}ACG (Abf1) TCYCNN (Adr1) CTACGTGC (CgGcn4) AGGGG (Gis1) CTTATC (CgCln3)	CAGL0G06644g	-	0.51	YTGAT (Ash1) RATCA (Ash1) RYMAAYA (Fkh1) TRTTKRY (Fkh2) GATAAG (Gzf3)
CAGL0G09977g	*CgGDB1*	1.7	CGTN{4}VRYGAY (Abf1) CGTN{7}GA (Abf1) YNCGTN{6}YGAY (Abf1) NCCDTYNVCCGN (Cat8) NCCDTYNVCCGN (Sip4)	CAGL0F03531g	-	0.52	RYMAAYA (Fkh1)

Importantly, in this analysis the binding sites of transcription factors that act as activators were crossed with the promoter regions of up-regulated genes and the down-regulated promoter regions were searched for binding sites of down-regulated transcription factors that act as repressors. This analysis lead to the identification of 139 up-regulated genes which do not lose or gain binding sites in their promoter regions in FFUL887 isolate, among them being Cg*CDR1* and Cg*PDR1*. 30 genes up-regulated in the FFUL887 strain were registered only losses of binding sites in their promoter region in this strain, such as CgQDR2 which lost the binding site "RYMAAYA" recognized by the transcription factor Fkh1; 86 up-regulated registered only gains of binding sites, such as Cg*MSN4*, harbouring the binding site of Cln3 and Tec1, both transcription factors involved in pseudohyphal growth, which could suggest that in FFUL887 there is a link between the pseudohyphal growth and the response to oxidative stress; 92 genes were identified as lost and gained binding sites in their promoter region, is the case of genes of *MET* family, such as Cg*MET15, CgMET14, CgMET5, CgMET10* and Cg*MET28* gaining and losing binding sites of Fkh1, Fkh2 and Ash1 transcription factors. On the other hand, within the down-regulated genes 291 of these genes maintained their promoter regions unaltered in FFUL887 when comparing to CBS137; 43 genes only lost binding sites in their promoter regions in strain, 151 down-regulated genes were identified to only gain transcription factor binding sites while 155 genes both lost and gained binding sites. Remarkably, it is noteworthy that some genes may be down- or up-regulated induced through the interaction of activated transcription factors that have the function to repress or activate the expression of genes, respectively. Interestingly, there are up- and down-regulated genes that, despite not having new binding sites in their promoter regions, have altered expression, suggesting that in those cases there are other contributions for their expression levels, as is the case of Cg*PDR1* which, as case of gain-of-function, is enough to render the increased expression.

A closer look at Table 11 tell us that similarly to what was observed before, the motifs with higher representation are those with higher degeneracy, such as “RYMAAYA” and “TRTTKRY”, recognized by Fkh1 and Fkh2 respectively. Motifs with shorter sequences and repeated nucleotides are also heavily represented in both cases, such as Mot3 motifs. In the down-regulated genes, there is also a clear representation of the motifs “CGGNS”, recognized from CgStb5. As expected, these results are in agreement with those observed in Figure 22.

3.4. Examination of the CgPdr1-regulatory network active in the FFUL887 isolate

As said before the FFUL887 isolate is resistant to fluconazole, voriconazole and exhibits increased tolerance to caspofungin[73]. This resistance phenotype was attributed, at least in part, to the fact that this strain encodes a gain-of-function mutant of the transcription factor CgPdr1[73]. 81 of the genes found to be up-regulated in the FFUL887 strain were previously documented to be regulated by CgPdr1, this corresponding to approximately 22% of the overall set of genes over-expressed in this strain (Annex II, Figure 24). On the other direction, there are 39 documented CgPdr1 targets within the set of genes down-regulated in the FFUL887 isolate (Figure 24, Annex II). Closer analysis of the promoter region of these up-regulated documented CgPdr1 targets differently expressed in the FFUL887 and in CBS138 strains shows that only 16 of them (Cg*PDR1*, Cg*GIN3*, Cg*MET15*, Cg*YPS9*, *CgREV1*, *CgPDH1*, Cg*YOR1*, *CgCIR2*, Cg*QDR2*, *CgHXT2*, Cg*HSP12*, *CgPOR1*, *CgPLB3*, CAGL0K12958g, Cg*CDR1* and Cg*YIM1*) harbour the DNA motif HYCCRKGGRN, the binding site recognized by CgPdr1[100]. Among these up-regulated genes that harbour the PDRE motif are the drug efflux pumps Cg*PDH1*, Cg*YOR1*, Cg*QDR2* and Cg*CDR1* (Figure 24). Since CgPdr1 is an activator its effect in the regulation of genes up-regulated in the FFUL887 strain that have the PDRE motif in its promoter region is likely to be direct. On the other side, the effect of CgPdr1 in the genes that are down-regulated in the FFUL887 isolate should be indirect, eventually through its role in regulating the expression of other transcription factors. The expression of other CgPdr1 documented targets that are up-regulated in the FFUL887 strain but that do not harbour a PDRE motif in their promoter should also be indirect. Indeed, the Amt1 transcription factor, up-regulated in the FFUL887 isolate by 1.83-fold is a known documented target of CgPdr1[50,58,100] and 10 of the genes up-regulated in FFUL887 isolate are documented targets of CgAmt1 (Cg*AMT1*, Cg*LYS9*, Cg*UGA2*, CAGL0H02519g, Cg*YCP4*, CAGL0J07084g, CAGL0K07205g, Cg*HIS1*, Cg*FOX2* and *DCS2* in Figure 24). These observations suggest that CgPdr1 might be activating these genes through the activation of CgAmt1. In Figure 24 it was compiled a putative transcriptional regulatory network controlled by the CgPdr1 transcription factor encoded by the FFUL887 isolate. At least 31 genes up-regulated in the FFUL887 isolate were found to harbour at least one PDRE motif in their promoter region (Cg*MET16*, Cg*MSA1*, CAGL0M11902g, Cg*INO2*, Cg*TCB1*, Cg*CPR1*, CgILV2, CAGL0G05566g, Cg*HBT1*, Cg*CDC25*, Cg*SEF1*, Cg*PBI1*, Cg*ADR1*, Cg*PFK26*, CAGL0I01650g, Cg*BOP2*, Cg*BOR1*, Cg*FLC1*, Cg*SKO1*, Cg*TRX3*,

Cg*CYS3*, Cg*MET10*, Cg*GRX8*, Cg*MET3*, Cg*INO4*, Cg*LYS12*, Cg*PTC6*, *CgARO3*, Cg*UPB11*, Cg*HAP1*, CAGL0I01276g), although, in this case there is no documented association between these target genes and CgPdr1. It is interesting to observe within this dataset the inclusion of several genes involved in biosynthesis of methionine (Cg*MET3*, Cg*MET10* and Cg*MET16*) and of the Cg*ADR1* gene, an essential regulator of *C. glabrata* metabolization of alternative carbon sources.

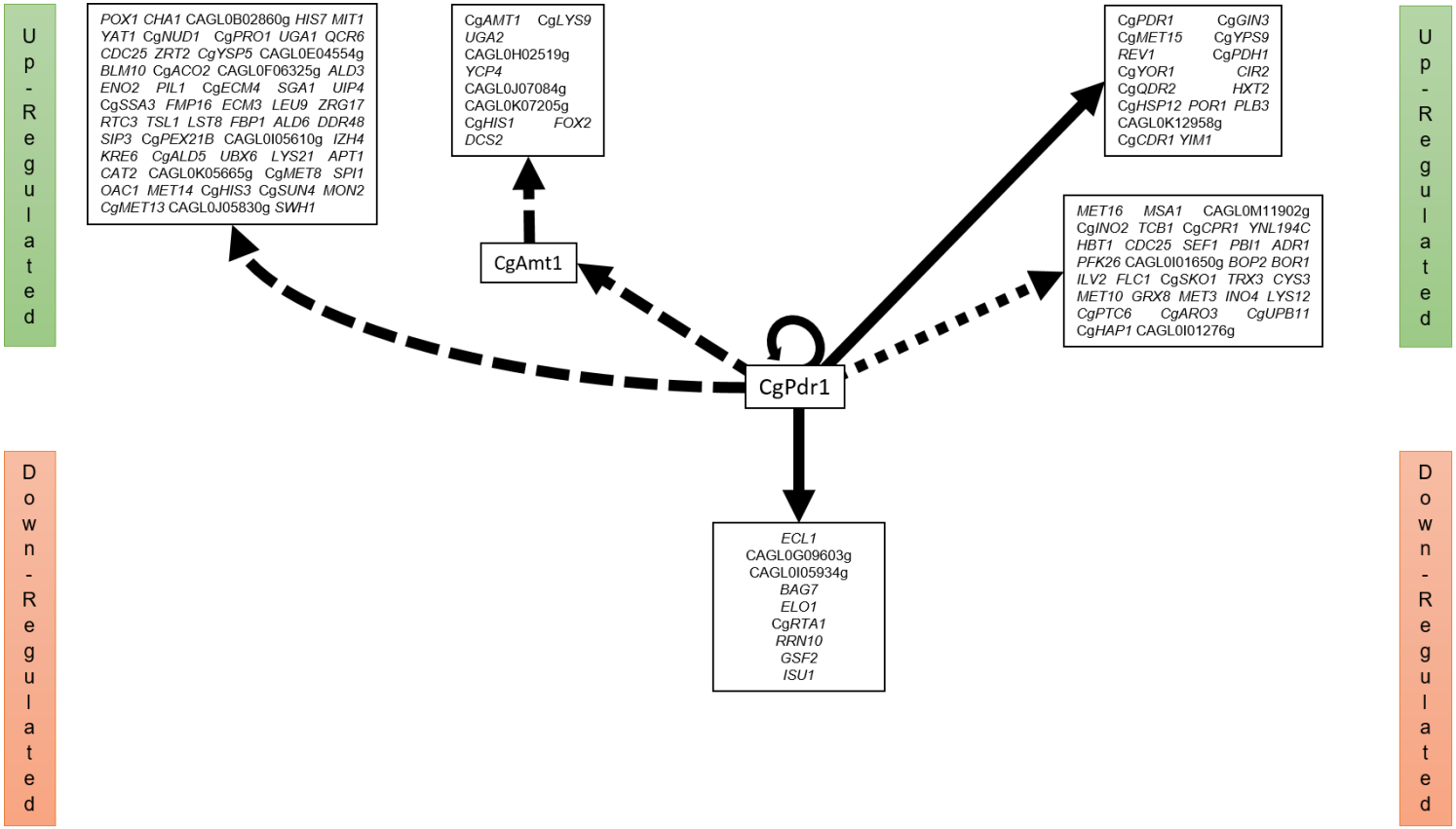


Figure 24 – Representation of CgPdr1 putative transcriptional regulatory network. Solid arrows represent documented target genes that harbour CgPdr1 consensus binding site; dashed-line arrows represent documented target genes that do not harbour CgPdr1 consensus binding site; dotted-line arrow represents undocumented targets that harbour CgPdr1 consensus binding site.

The regulatory network controlled by the CgPdr1 transcription factor mutant encoded by FFUL887 genome could also be altered as the result of modifications in the promoter region of target genes which can gain or lose PDRE motifs. To examine this hypothesis all the promoter regions of FFUL887 genes were searched for the PDRE motif. The results obtained led to the identification of 19 genes that are candidates to be under CgPdr1 regulation by harbouring a PDRE motif in their promoter region, this being absent in the promoter region of the corresponding gene encoded by the genome of the CBS138 strain. These genes are listed in the Table 12. Differently, 4 FFUL887 genes (Cg*ECL1*, CAGL0G01122g, Cg*SKO1* and Cg*RTA1*) lost the PDRE motif from their promoter (Figure 25A). Notably, the Cg*E*, J, Cg*C*, Cg*G*, Cg*SCM4*, Cg*PRM15*, Cg*VNX1* and Cg*TMA10* genes were found to be more expressed in the FFUL887 strain than in the CBS138 strain, within the range of 10-60% (Table 12). In Figure 25B it is possible to see that the overall number of CgPdr1 binding sites harbouring in the FFUL887 strain reaches 420 while in the CBS138 only 405 promoter harbour this motif. Within these, the Cg*CDR1*, Cg*PDH1*, Cg*QDR2* and Cg*YOR1* harbour the PDRE motif in both FFUL887 isolate and CBS138 strain. In the future, it may be interesting to compare the level of expression of these genes in the two strains in the presence of the drugs to which FFUL887 is resistant to.

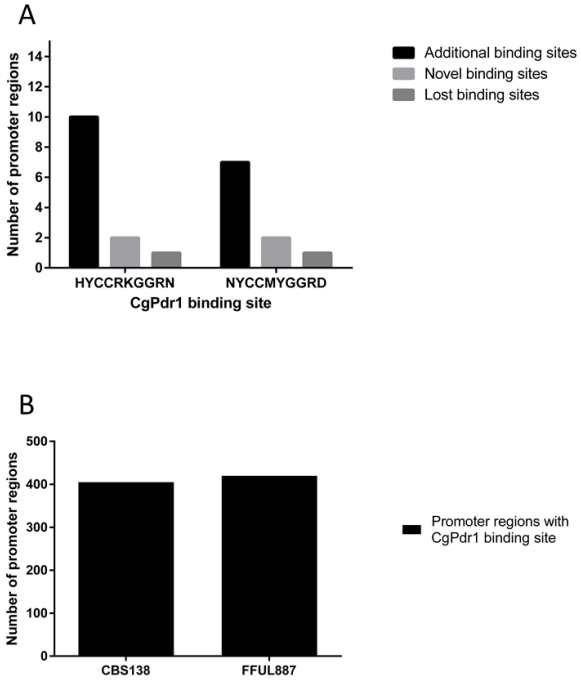


Figure 25 – Representation of alterations in the promoter regions harbouring CgPdr1 binding site. A – Number of gained and lost binding sites. Additional binding sites – addition of CgPdr1 binding site in a promoter already harbouring its binding site; novel binding site – promoter regions harbouring a CgPdr1 binding site in FFUL887 strain only; lost binding sites – promoter regions harbouring a CgPdr1 binding

site in CBS138 strain only. B – Overall number of promoters in CBS138 and FFUL887 harbouring CgPdr1 binding site.

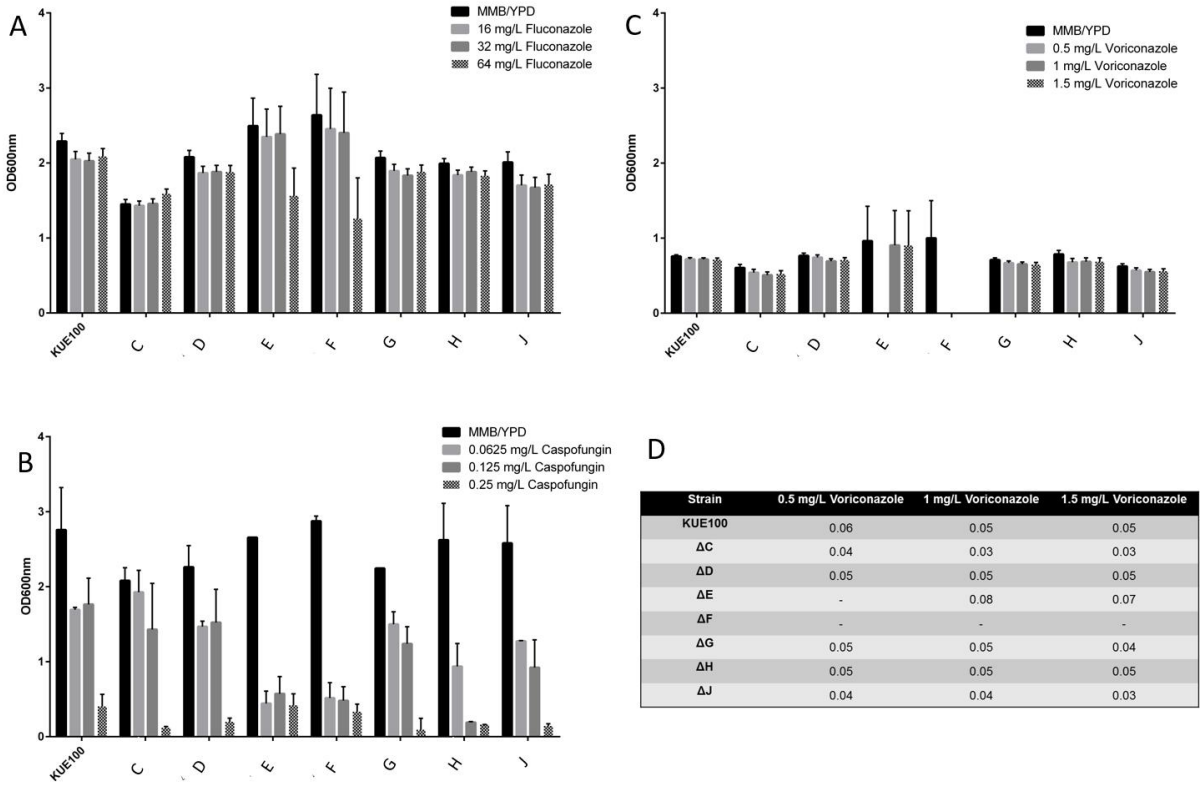
Table 12 –FFUL887 887 genes whose promoter was found to harbour one (or more) PDRE motifs and that were absent in the promoter region of the corresponding gene encoded by the genome of the CBS138 strain.

	ORF	Name	mRNA FFUL/mRNA CBS18
Promoters that "gained" a PDRE motif	CAGL0A01366g	Cg*EPA9*	0.99
	-	*CgE*	1.26
	J	-	1.38
	-	*CgH*	0.84
	-	*CgC*	1.14
	-	*CgD*	0.62
	-	*CgG*	1.12
	-	*CgF*	0.91
	-	Cg*A*	0.39
	-	*CgI*	0.96
	CAGL0A03080g	*CgSDH6*	0.69
	CAGL0D05060g	*CgPPT1*	0.92
	CAGL0J11220g	*CgRPS3*	0.64
	CAGL0K04279g	*CgSCM4*	1.18
	CAGL0M02981g	*CgPRM15*	1.59
	CAGL0M03487g	*RPC34*	0.67
	CAGL0M04147g	*VNX1*	1.35
Promoters that already harboured a PDRE motif and that gained other(s)	-	*B*	0.92
	CAGL0F08745g	*CgTMA10*	1.10

To examine whether these genes that gained a PDRE motif in their promoter play a role in *C. glabrata* tolerance to azoles or echinocandins, the susceptibility of strains devoid of a selected set of these genes (Cg*A*, Cg*B*, CgC, CgD, Cg*E*, Cg*F*, Cg*G*, Cg*H*, Cg*I*, J) was compared with the one of the parental strain KUE100. These assays were undertaken during cultivation of the strains in liquid MMB growth medium supplemented with inhibitory concentrations of the different drugs. In specific, three concentrations of voriconazole, fluconazole or caspofungin were tested: one corresponding to the considered clinical resistance breakpoint, one below that value and another one above it. Growth of the strains in the presence of the different concentrations of fluconazole and voriconazole was accompanied along 48h based on the increase of the OD600nm of the cultures. In the case of caspofungin an end-point assay was used in which the final OD600nm of the different cultures was compared after 48h of incubation in the presence or absence of the drugs. The mutants KUE100 ΔCg*A*, KUE100 Δ*E* and KUE100

Δ*F* were found to be unable to grow in the minimal medium used in the susceptibility assays and for that reason the susceptibility of these mutants was tested in rich YPD growth medium.

The results shown in Figure 26 show that the deletion of genes Cg*C*, Cg*D*, Cg*E*, Cg*F*, Cg*G*, Cg*H* and J did not significantly affected growth of *C. glabrata* cells under non-stressful conditions since all strains grew equally in the absence of the drugs. The elimination of *CgC*, *CgD*, *CgE*, *CgF*, *CgG*, *CgH* and J led to an increase in susceptibility to the three concentrations of caspofungin tested, the more prominent effect being registered for the Δ*E* and Δ*F* mutants which did not exhibited any significant growth in the lower two concentrations tested (Figure 26B). The expression of these genes did not exerted any significant protective effect against the inhibitory concentrations of fluconazole or voriconazole tested, at least under the experimental conditions used in this study (Figure 26A, 26C, 26D), although in the case of fluconazole both Δ*E* and Δ*F* achieved higher growth values (Figure 26A). Up to now it has not been seen the involvement of these genes, Cg*E*, which encodes a subunit of the vacuolar-ATPase V0 domain, and Cg*F*, which encodes an ornithine transporter, in tolerance to caspofungin. Despite this, it is evident that the susceptibility of strains devoid of these genes is higher when compared to their parental and control strain KUE100, suggesting that there might be new underlying resistance mechanisms not yet explored. Further studies are required to uncover the potential role of these genes in *C. glabrata* tolerance to caspofungin. In particular, it will be necessary to test if the expression of these genes is indeed under the regulation of CgPdr1.



Strain	0.5 mg/L Voriconazole	1 mg/L Voriconazole	1.5 mg/L Voriconazole
KUE100	0.06	0.05	0.05
ΔC	0.04	0.03	0.03
ΔD	0.05	0.05	0.05
ΔE	-	0.08	0.07
ΔF	-	-	-
ΔG	0.05	0.05	0.04
ΔH	0.05	0.05	0.05
ΔJ	0.04	0.04	0.03

Figure 26 – Growth of the KUE100 strain and of the derived deletion mutants in MM growth medium supplemented with three concentrations of fluconazole (A), voriconazole (B) or caspofungin (C). In the case of voriconazole and fluconazole growth was compared after 24h of growth while for caspofungin this was extended to 48h due to the more inhibitory effect of the drug. Voriconazole growth values (C) and ratios (D) are based on the growth curved shown in Annex VI. Derived mutants KUE100 Δ*E* and KUE100 Δ*F* were grown in rich YPD medium supplemented with the mentioned drug concentrations; derived mutants KUE100 Δ*C*, KUE100 Δ*D*, KUE100 Δ*G*, KUE100 Δ*H* and KUE100 ΔJ were grown in MMB medium supplemented with the indicated drug concentrations.

Because the deletion of the genes Cg*A*, *CgB* and *CgI* was found to significantly reduce viability of *C. glabrata* cells it was not possible to use a true deletion strain and therefore in these cases the suppression of these genes was achieved by changing their promoter for a tetracycline-repressible promotor. The experimental setup used in these cases was similar to the one described above with the exception that in this case the growth medium was supplemented with tetracycline to promote gene repression. In this case it was only compared the susceptibility of the wild-type cells with the mutant strains in the presence of caspofungine or fluconazole. Similarly to the other tested genes, the results shown in Figure 27 show that the deletion of genes Cg*A*, Cg*B* and Cg*I* did not significantly affected growth of *C. glabrata* cells under non-stressful conditions since all strains grew equally in the absence of the drugs (Figure 27) The elimination of Cg*A* and *CgI* genes lead to an increase in susceptibility to the three drug concentrations of fluconazole (Figure 27A) and caspofungin (Figure 27B) tested. The elimination of *CgB* appears to lead to an increase in susceptibility only to the three concentrations of caspofungin tested (Figure 27B).

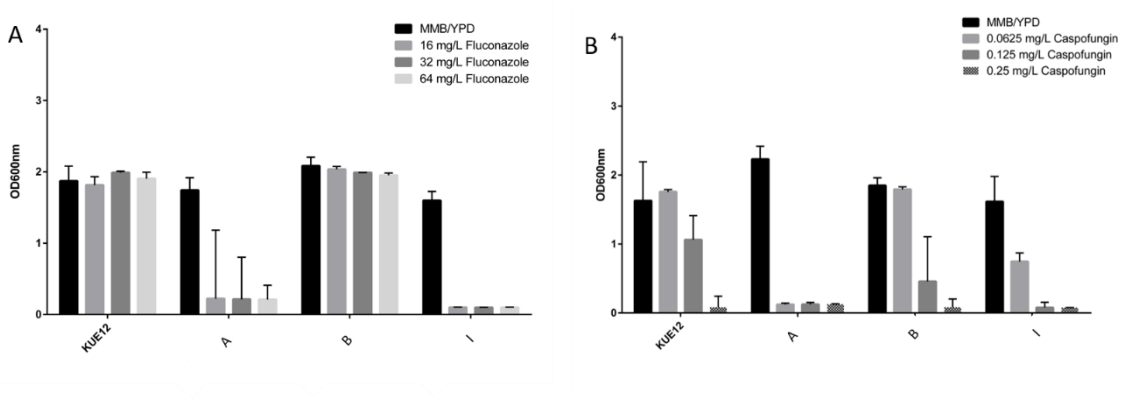


Figure 27 – Growth values for KUE100 and derived tetracycline-repressible promotor mutants under three different concentrations of fluconazole (A) after 24h of growth and caspofungin (B) after 48h of growth. Derived mutants CgB and CgI were grown in MMB medium supplemented with the indicated drug concentrations and tetracycline; derived mutant CgA was grown in rich YPD medium supplemented with the indicated drug concentrations and tetracycline.

Cg*A*, encodes a protein component of the small 40S ribosomal subunit; *CgB*, encodes a cell wall mannoprotein, and *CgI*, encodes a subunit of the Set1C histone H3K4 methiltransferase. Up to now these genes had not been previously implicated in *C. glabrata* tolerance to caspofungin and fluconazole resistance, however, it is clear that their deletion, in particular of Cg*A* and *CgI* genes, strongly alter the susceptibility to these drugs, while the deletion of *CgB* causes an increase of susceptibility only to caspofungin (Figure 27). This suggests that although not documented, there may be unstudied mechanisms that could be underlying the resistance to these drugs. Similarly to the genes studied before in this work, it will be interesting

to further examine the role of these genes in potential mechanisms of tolerance of *C. glabrata* to caspofungin and fluconazole. It will also be required to determine whether they are indeed under the regulation of CgPdr1, particularly in response to antifungal stress.

3.5. Analysis of other CgPdr1 GOF's regulatory networks

Transcriptomic analysis of *C. glabrata* isolates expressing different gain-of-function Pdr1 mutants revealed a considerable dispersion in the different set of over-expressed genes thereby leading to the hypothesis that the different mutations in the transcription factor impact the regulatory network in a different manner[58]. Having this in mind, the set of genes up-regulated in the FFUL887 isolate was compared with the set of genes described to be regulated by two other Pdr1 gain-of-function mutants: SM3, harbouring the T2837C (L946A) mutation, and 6955 mutant, harbouring the G822T (K274N) mutation[58]. Although other CgPdr1 GOF mutants are described in the literature[50,57], SM3 and 6955 are the only ones having accessible datasets on what concerns to the regulated genes. In Figure 28 it is shown the results of this comparative analysis.

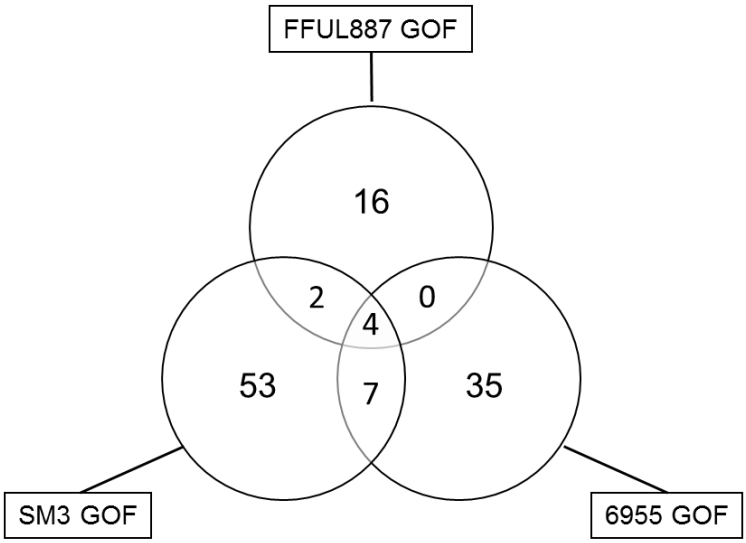


Figure 28 – Representation of the comparison of the set of genes up-regulated and regulated by CgPdr1 in different clinical isolates harbouring different GOF CgPdr1 mutations[58].

Four genes were found to be in common in all the three datasets examined, out of which it stands out the Cg*PDR1* gene. This observation suggests that up-regulation of this gene is a necessary feature for the different GOF mutants to exhibit increased tolerance to antifungals.

The other 3 genes in the common dataset, Cg*YPS9*, *REV1* and *HXT2* were previously documented to be regulated by CgPdr1, however, their functional activities have not been directly implicated in the drug resistance. Cg*YPS9*, however, has been linked to the increased virulence in *C. glabrata*[102]. Interestingly, and despite all GOFs suggest the auto-regulation of CgPdr1, only FFUL887 CgPdr1 GOF appears to be regulating the efflux pumps involved in MDR, such as Cg*YOR1*, Cg*PDH1*, Cg*QDR2* and Cg*CDR1*. On the other hand, SM3 GOF suggests that Cg*MSN4* may be under the regulation of CgPdr1, while 6955 GOF indicates that *STE12* might be regulated by CgPdr1. Both transcription factors are indeed up-regulated in FFUL887 CgPdr1 GOF, however, the deletion of Cg*PDR1* in both SM3 and 6955 GOF indicate that these transcriptions factors might establish interactions with CgPdr1. With these results, it should be interesting to delete FFUL887 Cg*PDR1* and re-assess this data to see whether the predicted CgPdr1 target genes in FFUL887 are indeed being regulated by this transcription factor, such as the MDR efflux pumps.

3.6. Examination of the role of Cg*Stb5*, Cg*Rpn4* and *Yrr1* in the FFUL887 CgPdr1-regulated network

The control of the expression of the genes of the PDR network involves, besides CgPdr1, at least three other transcription factors, CgRpn4, CgYrr1 and CgStb5. It is the combined activity of these transcription factors that leads to the final expression of genes required for resistance to antifungals in *C. glabrata*[60,64,69]. In view of this, the FFUL887 gene promoters' were searched for binding sites for CgStb5, Yrr1 and CgRpn4, the results obtained being shown in Figure 29. The number of CgRpn4 binding sites was the same in the two strains, thus not being represented, while 390 new genes were found to harbouring new CgStb5-binding sites and 5 were found to harbour new CgYrr1-binding sites (Figure 29A and 29B, respectively). The frequency of losses of binding sites for these two transcription factors was much smaller (Figure 29A and 29B). It was interesting to observe that despite these alterations the total number of CgStb5- and CgYrr1- binding sites identified in the genomes of both CBS138 strain and FFUL887 isolate was similar (Figure 29C). From the 390 genes that harboured a new CgStb5 motif, 28 were found to up-regulated in the FFUL887 isolate (Cg*PBI1*, Cg*HXT3*, Cg*MET3*, Cg*HIS7*, Cg*LYS9*, Cg*YAT1*, Cg*RAD59*, CAGL0D01782g, Cg*AAT2*, Cg*PRO1*, Cg*MET15*, Cg*SOD2*, Cg*ADR1*, Cg*APE2*, CAGL0F02101g, Cg*IRC24*, Cg*YOR1*, Cg*SRC1*, Cg*GRX7*, Cg*PIR2*, Cg*TDH3*, Cg*ICL1*, Cg*BOR1*, Cg*SFA1*, Cg*FOX2*, Cg*SUL2*, Cg*CDR1* and Cg*PIR3*) while 54 were down-regulated (Cg*RPL15B*, CAGL0B00572g, Cg*PDI1*, CAGL0C00968g, Cg*SOD1*, Cg*DRS1*, Cg*RPS15*, CAGL0F01815g, Cg*RPL12A*, Cg*HEK2*, Cg*LSC2*, Cg*RPS4A*, Cg*ZUO1*, Cg*REC102*, Cg*RPL40B*, Cg*SUR4*, Cg*D,* Cg*TIF35*, Cg*RPL43A*, CAGL0H09834g, Cg*POP3*, Cg*PBP4*, Cg*PUP3*, Cg*ADH1*, *CgPWP3*, *RPS25B*, CAGL0J00715g, Cg*RPL2B*, Cg*RPS24B*, Cg*MEP2*, Cg*RPP1B*, Cg*RPS3*, Cg*RML2*, Cg*EAP1*, Cg*TWF1*, CAGL0K01155g, Cg*GLN1*, Cg*RPS17A*, Cg*GUK1*, Cg*RPL20B*, Cg*KRE33*, Cg*PTR1*, Cg*A*, Cg*EBP2*, Cg*SMI1*,

CAGL0L09911g, Cg*RPL6A*, Cg*RPL33A*, CAGL0M02541g, CAGL0M03179g, CAGL0M03421g, Cg*PCH2*, Cg*RPL21B* and Cg*CUE5*). It is interesting to observe that the dataset of down-regulated genes in the FFUL887 strain that acquired one or more CgStb5 binding sites is highly enriched for genes involved in ribosome biogenesis. Up to now it has not been described any involvement of CgStb5 in regulation of ribosomal genes. Interestingly, CgStb5 has been described to share most of its regulated genes with CgPdr1 and indeed 10 of these genes that had gained new CgStb5 binding sites were found to be documented targets of CgPdr1, 1 of them being up-regulated in the FFUL887 strain, *HIS7*, encoding an imidazole glycerol phosphate synthase. CgStb5 is a repressor of the PDR genes and in this sense the acquisition of new binding sites by these genes that are co-regulated with CgPdr1 does not alone explain their up-regulation in the FFUL887 strain, but rather the loss of CgStb5 binding site in the promoter region of up-regulated genes. This way, 18 genes were identified as having lost CgStb5 binding site in their promoter region, 6 of them being up-regulated in the FFUL887 strain: *ZRT2*, encoding a zinc transporter, *ACO2*, encoding a putative aconitase isozyme, *CAT2*, encoding a carnitine O-acetyltransferase, Cg*PDH1*, Cg*QDR2*, and Cg*CDR1*, all these three genes encoding drug-efflux pumps. It is possible that the loss of binding sites for CgStb5 in the promoter regions of Cg*PDH1*, Cg*QDR2* and Cg*CDR1* could contribute for the up-regulation of these genes. In the case of CgYrr1, 50 new genes were found to harbour a binding site for this transcription factor that was absent in the CBS138 strain (Figure 29B). Only 5 of these genes (Cg*YAT1*, Cg*EFM1,* Cg*POM33*, Cg*QDR2* and Cg*FAA2*) were up-regulated in FFUL887 isolate, among them being the drug efflux pump Cg*QDR2*, demonstrated to play an important role in *C. glabrata* resistance to imidazoles[28].

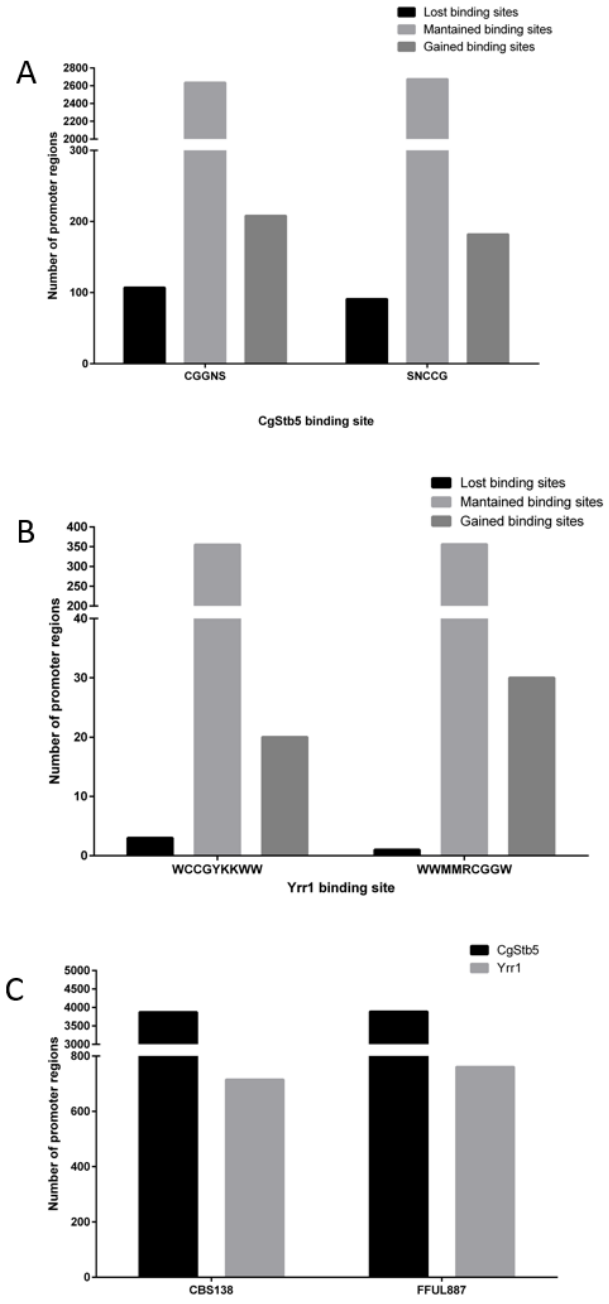


Figure 29 – Global regulatory network analysis of the transcription factors CgStb5 and Yrr1. A – Number of promoter regions which gained, lost or maintained CgStb5 binding site in FFUL887 isolate when compared to CBS138 strain. B - Number of promoter regions which gained, lost or maintained Yrr1 binding site in FFUL887 isolate when compared to CBS138 strain. C – Number of genes that harbour CgStb5 or Yrr1 binding sites in their promoter regions in both CBS138 strain and FFUL887 isolate.

Interestingly, CgStb5 potentially regulates 3870 genes in CBS138 and 4498 genes in FFUL887, while Yrr1 potentially regulates 715 genes in CBS138 and 761 genes in FFUL887 (Figure 29C). Differently, CgPdr1 can potentially regulate 348 genes in CBS138 strain and 346 genes in FFUL887, as these were the genes found to harbour PDRE motif (Figure 30).

As it is possible to observe in Figure 30A, 87 potential target genes are co-regulated by CgPdr1, CgStb5 and CgYrr1 in CBS138, while 93 potential target genes are co-regulated by these transcription factors (Figure 30B).

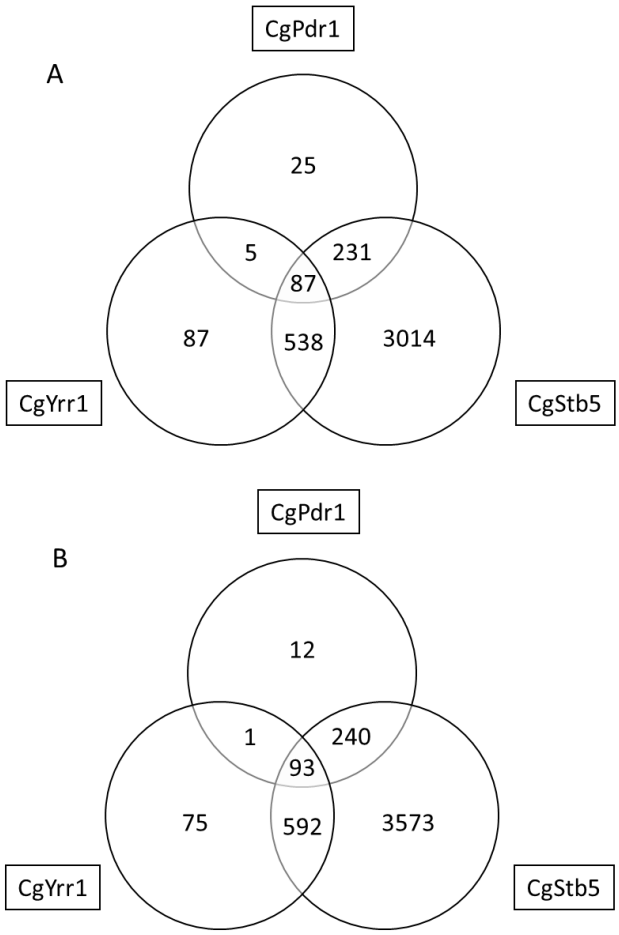


Figure 30 – Potential regulatory network between CgPdr1, CgYrr1 and CgStb5. A – Interactions predicted in CBS138 strain. B – Interactions predicted in FFUL887 isolate.

With this results it was possible to identify a set of 83 genes that are co-regulated by these transcription factors in both CBS138 strain and FFUL887 isolate, whereas only 4 genes are predicted to be solely regulated by this network in CBS138 (CAGL0G09603G, encoding a putative protein with unknown function, Cg*YGP1*, encoding a cell wall-related secretory glycoprotein, Cg*PST2*, a protein with similarity to a family of flavodoxin-like proteins and Cg*RAD14*, a protein that recognizes and binds damaged DNA during nucleotide excision repair) and 10 genes predicted to be regulated by this network in FFUL887 isolate (Cg*SDH6*, involved in in mitochondrial respiratory chain complex II, Cg*TMA10*, protein of unknown function that associates with ribosomes, Cg*C*, encoding a syntaxin-like vacuolar t-SNARE, Cg*QDR2*,

encoding a drug:H+ antiporter, Cg*NCE102*, encoding a protein of unknown function, Cg*F*, encoding an ornithine transporter of the mitochondrial inner membrane, Cg*GSF2*, encoding a putative protein of the ER membrane involved in hexose transporter secretion reported to be up-regulated in azole resistant strains[50], Cg*RSE1*, protein involved in pre-mRNA splicing, Cg*SEC21*, encoding a gamma subunit of coatomer and Cg*LAC1*, encoding a ceramide synthase component). Interestingly, FFUL887 CgPdr1 predicted target genes include those already identified above, such as Cg*C*, Cg*F* and Cg*E*, which is also predicted as a CgPdr1 in CBS138, strengthening the idea that these genes are involved in the resistance to drugs and, as shown in this work, contribute to the higher tolerance levels to caspofungin (Figure 26). Crossing the information with CgPdr1 documented target lead to the identification of 33 genes already documented as CgPdr1 targets in CBS138 strain and 34 genes identified as documented CgPdr1 target genes in FFUL887 isolate (Figure 31). Remarkably, in both CBS138 and FFUL887 strains, Cg*PDR1* was identified as being co-regulated by this cohort of transcription factors, whereas only in FFUL887 isolate Cg*QDR2* appears to be regulated by this cohort (Figure 31). These results suggest that the majority of the network is conserved between the two strains, however, FFUL887 appears to gain a set of targets that may have a role in drug resistance, such as Cg*QDR2*, Cg*C* and Cg*F*, possibly contributing to FFUL887's higher tolerance to drugs.

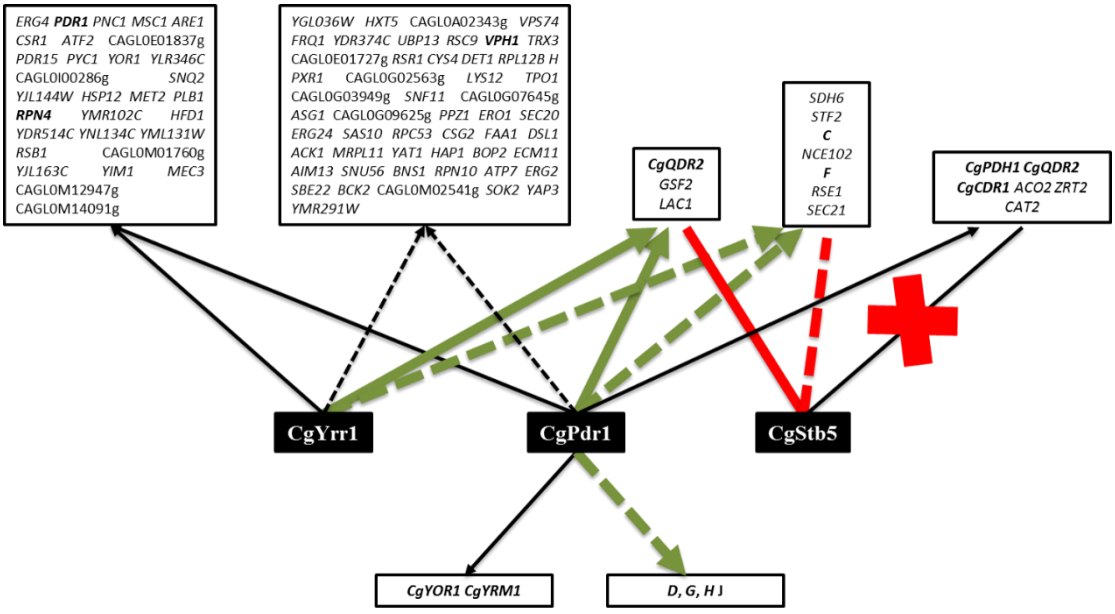


Figure 31 – Potential target genes of the CgPdr1, CgStb5 and CgYrr1 regulatory network in FFUL887. Target prediction was accomplished through the search of CgPdr1, CgStb5 and CgYrr1 binding sites in all CBS138 and FFUL887 promoter regions. Solid thicker arrows represent an interaction predicted by the CgPdr1 documented target genes; dashed thicker arrow represents an interaction not predicted by the CgPdr1 documented target genes; green arrows indicates an interaction that activates the expression of the target gene(s); red arrows indicates an interaction which represses the expression of the target gene(s); red arrow with a red cross (X) represents the loss of CgStb5 binding site in the target gene(s); black solid arrows indicate an interaction observed in both CBS138 and FFUL887 strains that is predicted by the documented target genes; black dashed arrows indicate an interaction observed in both CBS1338

and FFUL887 strain that is not predicted by the documented target genes, in these two cases the repressive action of CgStb5 is maintained; genes in bold represent genes which are described to be involved in drug resistance.

3.7. Unveiling other transcriptional regulators that could be involved in the control of genomic expression of the FFUL887 isolate: emphasis on the CgAdr1 transcription factor

To better understand the overall regulatory network governing the genomic expression of the FFUL887 isolate the documented regulatory network associations between transcription factors found to be up- or down-regulated in this strain, in comparison with the expression levels registered in the CBS138 strain, were searched. For this an initial list of regulatory associations between transcription factors and target genes in *C. glabrata*, kindly provided by Dr Catarina Costa and Prof Miguel Teixeira, was searched (Annex VII). Since this list was not updated a further survey of the regulatory associations was performed focused on more recent publications and only focusing the TFs whose expression was altered in the FFUL887 strain. Although the change in expression may not necessarily mean that the activity of a given TF is higher in the FFUL887 strain is higher, in comparison with the one registered in the CBS138 strain, it is an important indicative and in fact the only one that it would be possible to use since the list of regulatory associations between TFs and their target genes in *C. glabrata* is very scarce. The list of transcription factors that were used for this is shown in Table 7 and Table 8.

Besides CgPdr1, which regulated 23% of the genes up-regulated in the FFUL887 isolate, the transcription factor CgMsn4 was documented to regulate 2% of the up-regulated genes in this strain (Table 13). Those transcription factors whose expression decreased in the FFUL887 strain had no documented targets within the set of genes down-regulated in this strain (Table 14).

Since most transcription factors found to be up- or down-regulated in the FFUL887 strain do not have documented targets in the literature, the promoter region of genes found to be up- or down- regulated in this strain, in comparison with the expression values registered in the CBS138 strain, was searched for their binding sites. These results are shown in Table 15 and Table 16 and they indicate the potential regulatory associations that may occur between transcription factors and their documented targets. To note that in this analysis the identified up-regulated transcription factor CgSef1 and the down-regulated transcription factors CgCom2 and CgUsv1 were not included as there is scarce information regarding these transcription factor's binding site.

Table 13 – Transcription factors found to be up-regulated (>1.5-fold). Number of documented target genes described in the literature and the respective number of up-regulated genes are also represented. (-) – Indicates that no documented targets were found for this transcription factor.

ORF	*Transcription factor*	*mRNA FFUL887/mRNA CBS138*	*Number of documented target genes*	*Number of genes up-regulated In the FFUL887 isolate that are documented to be regulated by the TF*
CAGL0A04455g	CgSef1	1.52	-	-
CAGL0K05841g	Cg*Hap1*	1.53	-	-
CAGL0J06182g	Cg*Sko1*	1.59	-	-
CAGL0L03916g	CgAzf1	1.73	-	-
CAGL0K08756g	CgYap5	1.64	-	-
CAGL0M13189g	CgMsn*4*	1.99	21	7 (33%)
CAGL0I04180g	CgAmt1	1.83	363	81 (22%)
CAGL0A00451g	CgPdr1	1.96	7	0
CAGL0E04884g	Adr1	2.1	-	-

Table 14 - Transcription factors found to be down-regulated (<0.7-fold). Number of documented target genes described in the literature and the respective number of down-regulated genes are also represented. (-) – Indicates that no documented targets were found for this transcription factor.

ORF	*Transcription factor*	*mRNA FFUL887/mRNA CBS138*	*Number of documented target genes*	*Number of documented target genes down-regulated*
CAGL0C05335g	CgRtg1	0.51	-	-
CAGL0L06072g	CgCom2	0.65	-	-
CAGL0E06116g	CgUsv1	0.64	-	-
CAGL0K08624g	CgHap4	0.62	-	-
CAGL0F06259g	CgArg80	0.68	-	-

Table 15 – Possible interactions between the identified up-regulated transcription factors which have no documented target genes described and FFUL887 isolate genome. Consensus binding site acquired at the online platform YEASTRACT (http://www.yeastract.com/consensuslist.php).

Transcription Factor	*Consensus binding site*	*Possible number of target genes*	*Possible up regulated target genes*
CgHap1	CCGN{3}CCG	238	27
CgYap5	TTASTAA	751	56
CgSko1	TKACGTMA	171	18
CgAzf1	AARARAAA	1829	144
CgAdr1	NNGGAGA	3560	294

Table 16 - Possible interactions between the identified down-regulated transcription factors which have no documented target genes described and FFUL887 isolate genome. Consensus binding site acquired at the online platform YEASTRACT (http://www.yeastract.com/consensuslist.php).

Transcription Factor	*Consensus binding site*	*Possible number of target genes*	*Possible down-regulated target genes*
CgRTG1	GKNAC	3479	455
CgHAP4	YHRTTKMT	334	0
CgARG80	CGGTNTCGG	5	1

The transcription factor Adr1 (Table 15) emerges as a potential important transcription factor as a large number of the up-regulated genes was found to harbour a binding site for this transcription factor. It is important to stress that in *S. cerevisiae* Adr1 is involved in the control of the regulation of genes required for metabolization of alternative carbon sources, which as said before it is an hallmark of *C. glabrata* adaptation to the mammalian host[76]. Among the genes up-regulated in the FFUL887 isolate that could be potentially regulated by CgAdr1 are: Cg*CIT1*, involved in Krebs cycle; *ALD3*, encoding a putative aldehyde dehydrogenase; *SFC1*, involved succinate:fumarate antirporter activity; Cg*ACO2*, involved in Krebs cycle, ICL1, *MLS1*, and *GOR1*, involved in glyoxylate cycle, *POX1*, *FOX2, TES1* and *FAA2* involved in fatty acids metabolism (Figure 32 and 33).

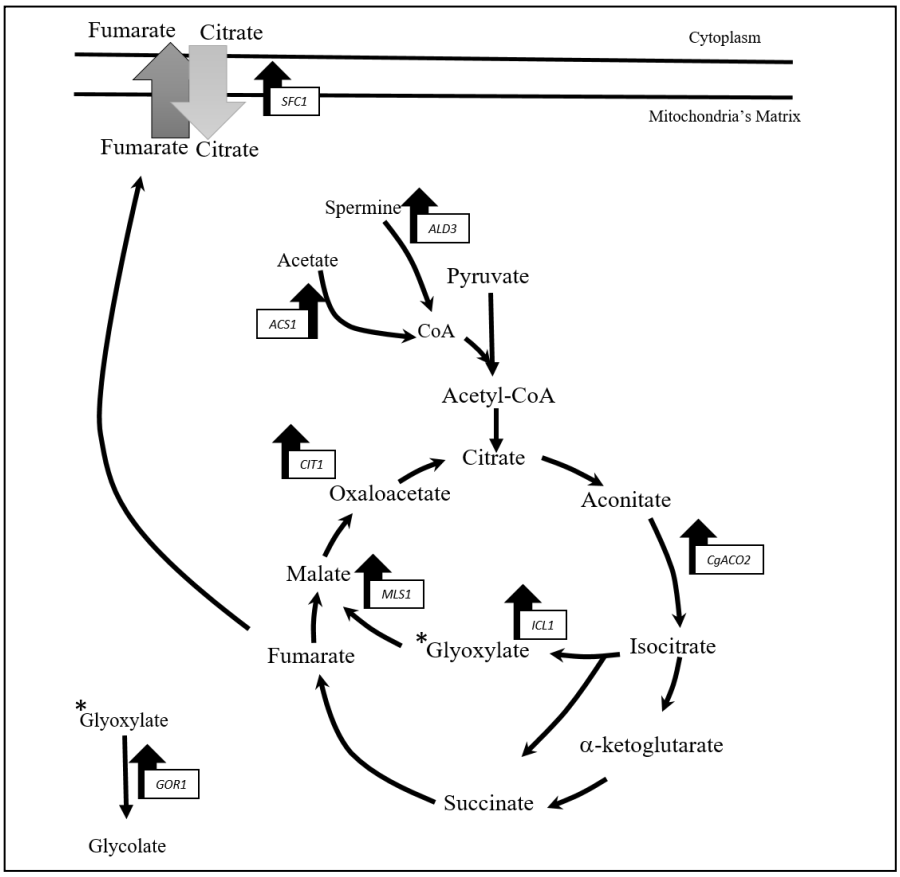


Figure 32 – Representation of the up-regulated tricarboxylic acid cycle and glyoxylate reactions potentially targeted by the transcription factor CgAdr1.

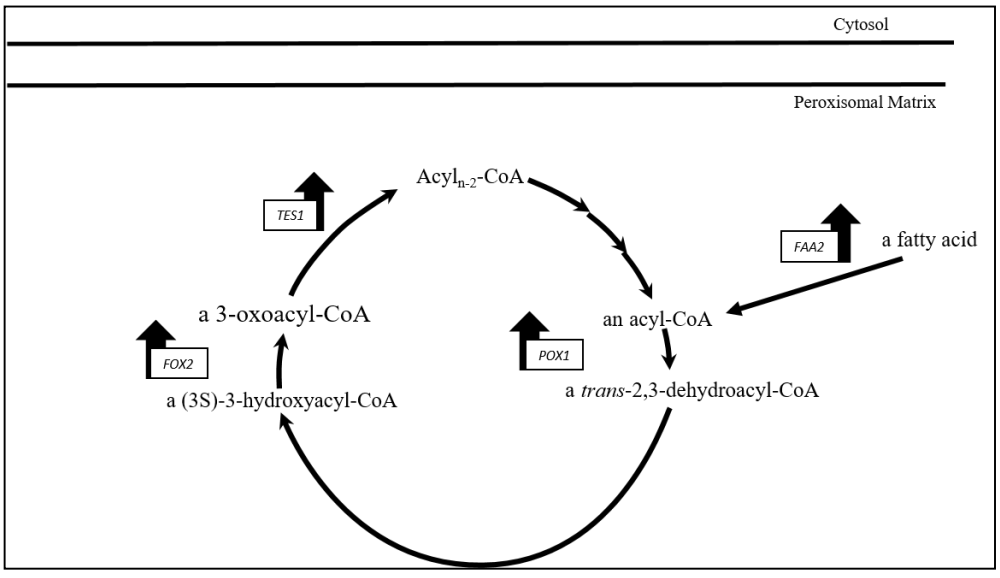


Figure 33 - Representation of the up-regulated fatty-acid oxidation reactions potentially targeted by the transcription factor CgAdr1.

The over-expression of these genes in the FFUL887 strain even during cultivation in the glucose-supplemented RPMI growth medium is interesting and suggests that the transcription factors involved in their regulation could be constitutively active. The CgAdr1 transcription factor encoded by the FFUL887 strain has two point mutations (Ala650Gly and Gly627Ser) relative to its CBS138 counter-partners, however, it is not know the role of residues in the control of the activity of the transcription factor. CgGcr1 and CgCat8, two transcription factors that in *S. cerevisiae* are also essential for the control of the expression of genes required for metabolization of alternative carbon sources, were also found to harbour point mutations in the FFUL887 isolate: 8 in the case of Gcr1 (Ala981Ser, Ser970Pro, His919Gln, Ser744Pro, Ser479Asn, Thr471Ser, Ser360Gly, and Thr245Ser) and 4 in the case of CgCat8 (Leu1128Val, Ala1127Ser, Val910Gly and Ile404Leu). These different insights prompted the examination of whether the FFUL887 and CBS138 strains had a different capacity to grow in the presence of different carbon sources, the results obtained being shown in Figure 34. Consistently, the FFUL887 strain exhibited a much faster growth than the CBS138 strain when acetate, butyrate and propionate were used as the sources of carbon in the growth medium.

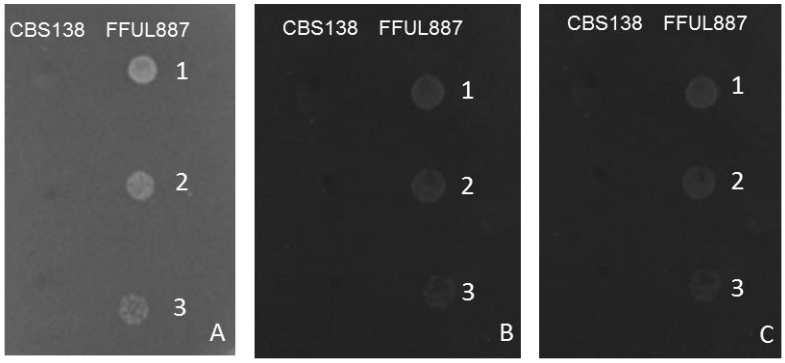


Figure 34 –Spot assays using MMB medium with 0.5% concentration of acetate (A), butyrate (B) and propionate (C) as the only carbon source for CBS138 and FFUL887 strains. Spots were prepared by plating a cellular suspension with initial OD600nm of 0.05 (1), followed by the dilution of 1:5 (2) and 1:25 (3) from the initial suspension (1).

4. Concluding Remarks

One of the clear observations that comes out of this study was the massive alteration in the non-coding genome of the FFUL887 strain, in comparison with the one of the reference strain CBS138. Although some of these modifications are likely to be the mere result of evolution and may not represent meaningful events affecting the physiology of *C. glabrata* cells, it is clear that many of these modifications will contribute to shape the evolution of the regulatory network of the strains thereby allowing a modification in the transcriptional landscape to fit better the needs of the microbe along colonization of the host. To better understand the true relevance of the herein reported alterations extensive transcription profiling of the FFUL887 and CBS138 strains under different experimental

conditions would be required, along with a much better knowledge of the binding sites recognized by each *C. glabrata* transcription factor. In the case of CgPdr1 it was possible to identify a set of new putative target genes whose promoters gained at least one PDRE motif which was absent from the genome of the CBS138 strain. It remains to be confirmed if the expression of these genes is indeed regulated by CgPdr1, especially in the presence of antifungals to which CgPdr1 responds to. In case this hypothesis is confirmed, then it can be hypothesized that the expansion of regulatory networks through the acquisition of novel TF-DNA interactions could represent an important mechanism by which *C. glabrata* cells acquire resistance to antifungals. It will also be of interest to further confirm if the loss of binding sites for CgStb5 may contribute to enhance the expression of the drug-efflux pumps CgQDR2, CgPHD1 and CgCDR1 in the FFUL887 isolate. If this is confirmed, then it will be an important demonstration of the evolution of the PDR network occurring by alteration in the promoter region of target genes.

5. References

1. Kamran, M. *et al.* Inactivation of transcription factor gene ACE2 in the fungal pathogen Candida glabrata results in hypervirulence. *Eukaryot. Cell* **3,** 546–52 (2004).
2. Fidel, P. L., *et al*. Candida glabrata: review of epidemiology, pathogenesis, and clinical disease with comparison to C. albicans. *Clin. Microbiol. Rev.* **12,** 80–96 (1999).
3. Sanglard, D. *et al*. Resistance of Candida species to antifungal agents: molecular mechanisms and clinical consequences. *Lancet Infect. Dis.* **2,** 73–85 (2002).
4. Rodloff, C., *et al*. Epidemiology and antifungal resistance in invasive candidiasis. *Eur. J. Med. Res.* **16,** 187–195 (2011).
5. Wisplinghoff, H. *et al.* Nosocomial bloodstream infections in US hospitals: analysis of 24,179 cases from a prospective nationwide surveillance study. *Clin. Infect. Dis.* **39,** 309–317 (2004).
6. Castanheira, M., *et al*. Activity of echinocandins and triazoles against a contemporary (2012) worldwide collection of yeast and moulds collected from invasive infections. *Int. J. Antimicrob. Agents* **44,** 320–326 (2014).
7. Paulo, C. *et al.* Retrospective analysis of clinical yeast isolates in a hospital in the centre of Portugal: spectrum and revision of the identification procedures. *Med. Mycol.* **47,** 836–44 (2009).
8. Correia, A., Sampaio, P. *et al*. Study of Molecular Epidemiology of Candidiasis in Portugal by PCR Fingerprinting of Candida Clinical Isolates Study of Molecular Epidemiology of Candidiasis in Portugal by PCR Fingerprinting of Candida Clinical Isolates. *J. Clin. Microbiol.* **42,** 5889–5903 (2004).
9. Klevay, M. J. *et al*. Initial treatment and outcome of Candida glabrata versus Candida albicans bloodstream infection. *Diagn. Microbiol. Infect. Dis.* **64,** 152–157 (2009).
10. Moran, C. Comparison of Costs, Length of Stay, and Mortality Associated with Candida glabrata and Candida albicans Bloodstream Infections. **38,** 78–80 (2012).
11. Paulo, C. *et al.* Retrospective analysis of clinical yeast isolates in a hospital in the centre of Portugal: spectrum and revision of the identification procedures. *Med. Mycol.* **47,** 836–44 (2009).
12. Blot, S. *et al*. Risk factors for Candida non-albicans candidemia. *Diagn. Microbiol. Infect. Dis.* **61,** 362–3; author reply 364 (2008).
13. Faria-Ramos, I. *et al.* Environmental azole fungicide, prochloraz, can induce cross-resistance to medical triazoles in Candida glabrata. *FEMS Yeast Res.* **14,** 1119–23 (2014).
14. Pfaller, M. *et al*. Candida bloodstream infections: Comparison of species distributions and antifungal resistance patterns in community-onset and nosocomial isolates in the SENTRY Antimicrobial Surveillance Program, 2008-2009. *Antimicrob. Agents Chemother.* **55,** 561–566 (2011).
15. Pfaller, M. *et al*. Echinocandin and triazole antifungal susceptibility profiles for clinical opportunistic yeast and mold isolates collected from 2010 to 2011: Application of new CLSI clinical breakpoints and epidemiological cutoff values for characterization of geographic . *J. Clin. Microbiol.* **51,** 2571–2581 (2013).
16. Borst, A. *et al*. Rapid acquisition of stable azole resistance by Candida glabrata isolates obtained before the clinical introduction of fluconazole. *Antimicrob. Agents Chemother.* **49,** 783–7 (2005).
17. Arendrup, M. C. Candida and Candidaemia: Susceptibility and Epidemiology. *Dan. Med. J.* **60,** 1–32 (2013).
18. Pfaller, M. *et al.* Use of epidemiological cutoff values to examine 9-year trends in susceptibility of Candida species to anidulafungin, caspofungin, and micafungin. *J. Clin. Microbiol.* **49,** 624–629 (2011).
19. Cowen, L. E. *et al*. Stress, drugs, and evolution: the role of cellular signaling in fungal drug resistance. *Eukaryot. Cell* **7,** 747–64 (2008).
20. Sheehan, D. J. *et al*. Current and emerging azole antifungal agents. *Clin. Microbiol. Rev.* **12,** 40–79 (1999).
21. Odds, F. C. *et al*. Antifungal agents: mechanisms of action. *Trends Microbiol.* **11,** 272–279 (2003).

22. Mast, N. *et al*. Antifungal Azoles: Structural Insights into Undesired Tight Binding to Cholesterol-Metabolizing CYP46A1. *Mol. Pharmacol.* **84,** 86–94 (2013).
23. Cowen, L. E. The evolution of fungal drug resistance: modulating the trajectory from genotype to phenotype. *Nat. Rev. Microbiol.* **6,** 187–198 (2008).
24. Torelli, R. *et al.* The ATP-binding cassette transporter-encoding gene CgSNQ2 is contributing to the CgPDR1-dependent azole resistance of Candida glabrata. *Mol. Microbiol.* **68,** 186–201 (2008).
25. Sanglard, D. *et al*. Role of ATP-binding-cassette transporter genes in high-frequency acquisition of resistance to azole antifungals in Candida glabrata. *Antimicrob. Agents Chemother.* **45,** 1174–1183 (2001).
26. Sanguinetti, M. *et al.* Mechanisms of Azole Resistance in Clinical Isolates of Candida glabrata Collected during a Hospital Survey of Antifungal Resistance. **49,** 668–679 (2005).
27. Vermitsky, J.-P. *et al*. Azole resistance in Candida glabrata: coordinate upregulation of multidrug transporters and evidence for a Pdr1-like transcription factor. *Antimicrob. Agents Chemother.* **48,** 3773–81 (2004).
28. Costa, C. *et al.* Candida glabrata drug: H+ antiporter CgQdr2 confers imidazole drug resistance, being activated by transcription factor CgPdr1. *Antimicrob. Agents Chemother.* **57,** 3159–3167 (2013).
29. Costa, C. *et al.* The dual role of candida glabrata drug:H+ antiporter CgAqr1 (ORF CAGL0J09944g) in antifungal drug and acetic acid resistance. *Front. Microbiol.* **4,** 170 (2013).
30. Costa, C. *et al.* Candida glabrata drug:H+ antiporter CgTpo3 (ORF CAGL0I10384G): Role in azole drug resistance and polyamine homeostasis. *J. Antimicrob. Chemother.* **69,** 1767–1776 (2014).
31. Tsai, H., *et al*. Candida glabrata PDR1, a transcriptional regulator of a pleiotropic drug resistance network, mediates azole resistance in clinical isolates and petite mutants. *Antimicrob. Agents Chemother.* **50,** 1384–92 (2006).
32. Hull, C. M. *et al.* Facultative Sterol Uptake in an Ergosterol-Deficient Clinical Isolate of Candida glabrata Harboring a Missense Mutation in ERG11 and Exhibiting Cross-Resistance to Azoles and Amphotericin B. *Antimicrob. Agents Chemother.* **56,** 4223–4232 (2012).
33. Nakayama, H. *et al*. Depletion of the squalene synthase (ERG9) gene does not impair growth of Candida glabrata in mice. *Antimicrob. Agents Chemother.* **44,** 2411–2418 (2000).
34. Nakayama, H. *et al.* The Candida glabrata putative sterol transporter gene CgAUS1 protects cells against azoles in the presence of serum. *J. Antimicrob. Chemother.* **60,** 1264–1272 (2007).
35. Bard, M. *et al.* Sterol uptake in Candida glabrata: rescue of sterol auxotrophic strains. *Diagn. Microbiol. Infect. Dis.* **52,** 285–93 (2005).
36. Emri, T. *et al*. Echinocandins: Production and applications. *Appl. Microbiol. Biotechnol.* **97,** 3267–3284 (2013).
37. Deresinski, S. C. *et al*. Caspofungin. *Oxford Journals* **2699,** 1445–1457 (2003).
38. Perlin, D. S. Resistance to echinocandin-class antifungal drugs. *Public Health* **10,** 121–130 (2009).
39. Pfeiffer, C. D. *et al.* Breakthrough Invasive Candidiasis in Patients on Micafungin. *J. Clin. Microbiol.* **48,** 2373–2380 (2010).
40. Healey, K. R.*et al*. Candida glabrata mutants demonstrating paradoxical reduced caspofungin susceptibility but increased micafungin susceptibility. *Antimicrob. Agents Chemother.* **55,** 3947–3949 (2011).
41. Arendrup, M. C. *et al*. Evaluation of CLSI M44-A2 disk diffusion and associated breakpoint testing of caspofungin and micafungin using a well-characterized panel of wild-type and fks hot spot mutant Candida isolates. *Antimicrob. Agents Chemother.* **55,** 1891–5 (2011).
42. Garcia-Effron, G. *et al*. Effect of Candida glabrata FKS1 and FKS2 Mutations on Echinocandin Sensitivity and Kinetics of 1,3- -D-Glucan Synthase: Implication for the Existing Susceptibility Breakpoint. *Antimicrob. Agents Chemother.* **53,** 3690–3699 (2009).
43. Thompson, G. R. *et al.* Development of caspofungin resistance following prolonged therapy for invasive candidiasis secondary to Candida glabrata infection. *Antimicrob. Agents Chemother.* **52,** 3783–3785 (2008).
44. Costa-de-Oliveira, S. *et al.* FKS2 Mutations Associated with Decreased Echinocandin

Susceptibility of Candida glabrata following Anidulafungin Therapy. *Antimicrob. Agents Chemother.* **55,** 1312–1314 (2011).

45. Zimbeck, A. J. *et al.* FKS Mutations and Elevated Echinocandin MIC Values among Candida glabrata Isolates from U.S. Population-Based Surveillance. *Antimicrob. Agents Chemother.* **54,** 5042–5047 (2010).
46. Cleary, J. D. *et al*. Reduced Candida glabrata Susceptibility Secondary to an FKS1 Mutation Developed during Candidemia Treatment. *Antimicrob. Agents Chemother.* **52,** 2263–2265 (2008).
47. Katiyar, S., *et al*. Candida albicans and Candida glabrata Clinical Isolates Exhibiting Reduced Echinocandin Susceptibility. *Antimicrob. Agents Chemother.* **50,** 2892–2894 (2006).
48. Fardeau, V. *et al.* The central role of PDR1 in the foundation of yeast drug resistance. *J. Biol. Chem.* **282,** 5063–74 (2007).
49. Garcia-Effron, G. *et al.* Novel FKS Mutations Associated with Echinocandin Resistance in Candida Species. *Antimicrob. Agents Chemother.* **54,** 2225–2227 (2010).
50. Vermitsky, J.-P. *et al.* Pdr1 regulates multidrug resistance in Candida glabrata: gene disruption and genome-wide expression studies. *Mol. Microbiol.* **61,** 704–722 (2006).
51. Balzi, E. *et al*. The multidrug resistance gene PDR1 from Saccharomyces cerevisiae. *J. Biol. Chem.* **262,** 16871–9 (1987).
52. Delaveau, T., Delahodde, a, Carvajal, E., Subik, J. & Jacq, C. PDR3, a new yeast regulatory gene, is homologous to PDR1 and controls the multidrug resistance phenomenon. *Mol. Gen. Genet.* **244,** 501–11 (1994).
53. Rogers, B. *et al.* The pleitropic drug ABC transporters from Saccharomyces cerevisiae. *J. Mol. Microbiol. Biotechnol.* **3,** 207–214 (2001).
54. Klimova, N. *et al*. Phenotypic analysis of a family of transcriptional regulators, the zinc cluster proteins, in the human fungal pathogen Candida glabrata. *G3 (Bethesda).* **4,** 931–40 (2014).
55. Kaur, R. *et al*. Functional genomic analysis of fluconazole susceptibility in the pathogenic yeast Candida glabrata: roles of calcium signaling and mitochondria. *Antimicrob. Agents Chemother.* **48,** 1600–13 (2004).
56. Schwarzmüller, T. *et al.* Systematic phenotyping of a large-scale Candida glabrata deletion collection reveals novel antifungal tolerance genes. *PLoS Pathog.* **10,** e1004211 (2014).
57. Tsai, H. F. *et al.* Microarray and molecular analyses of the azole resistance mechanism in Candida glabrata oropharyngeal isolates. *Antimicrob. Agents Chemother.* **54,** 3308–3317 (2010).
58. Caudle, K. E. *et al.* Genomewide Expression Profile Analysis of the Candida glabrata Pdr1 Regulon. *Eukaryot. Cell* **10,** 373–383 (2011).
59. Thakur, J. K. *et al.* A nuclear receptor-like pathway regulating multidrug resistance in fungi. *Nature* **452,** 604–609 (2008).
60. Orta-Zavalza, E. *et al.* Local silencing controls the oxidative stress response and the multidrug resistance in Candida glabrata. *Mol. Microbiol.* **88,** 1135–1148 (2013).
61. Hahn, J.-S. *et al*. A stress regulatory network for co-ordinated activation of proteasome expression mediated by yeast heat shock transcription factor. *Mol. Microbiol.* **60,** 240–251 (2006).
62. Lucau-Danila, A. *et al*. Competitive promoter occupancy by two yeast paralogous transcription factors controlling the multidrug resistance phenomenon. *J. Biol. Chem.* **278,** 52641–50 (2003).
63. Owsianik, G. *et al*. Control of 26S proteasome expression by transcription factors regulating multidrug resistance in Saccharomyces cerevisiae. *Mol. Microbiol.* **43,** 1295–1308 (2002).
64. Le Crom, S. *et al.* New insights into the pleiotropic drug resistance network from genome-wide characterization of the YRR1 transcription factor regulation system. *Mol Cell Biol* **22,** 2642–2649 (2002).
65. Cui, Z. *et al*. Yeast gene YRR1, which is required for resistance to 4-nitroquinoline N-oxide, mediates transcriptional activation of the multidrug resistance transporter gene SNQ2. *Mol. Microbiol.* **29,** 1307–15 (1998).
66. Salin, H. *et al.* Structure and properties of transcriptional networks driving selenite stress response in yeasts. *BMC Genomics* **9,** 333 (2008).
67. Gallagher, J. E. G. *et al.* Divergence in a master variator generates distinct phenotypes and

transcriptional responses. *Genes Dev.* **28,** 409–421 (2014).

68. Jelinsky, S. *et al*. Regulatory networks revealed by transcriptional profiling of damaged Saccharomyces cerevisiae cells: Rpn4 links base excision repair with proteasomes. *Mol. Cell. Biol.* **20,** 8157–8167 (2000).
69. Noble, J. a. *et al.* STB5 is a negative regulator of azole resistance in candida glabrata. *Antimicrob. Agents Chemother.* **57,** 959–967 (2013).
70. Larochelle, M. *et al*. Oxidative stress-activated zinc cluster protein Stb5 has dual activator/repressor functions required for pentose phosphate pathway regulation and NADPH production. *Mol. Cell. Biol.* **26,** 6690–6701 (2006).
71. Ferrari, S. *et al.* Gain of Function Mutations in CgPDR1 of Candida glabrata Not Only Mediate Antifungal Resistance but Also Enhance Virulence. *PLoS Pathog.* **5,** e1000268 (2009).
72. Berila, N. *et al*. Molecular analysis of Candida glabrata clinical isolates. *Mycopathologia* **170,** 99–105 (2010).
73. Salazar, S. Master thesis: Genomic analysis of a Candida glabrata clinical isolate resistant to antifungals unveils novel features of drug resistance in this pathogenic yeast. (2015).
74. Singh-Babak, S. D. *et al.* Global analysis of the evolution and mechanism of echinocandin resistance in Candida glabrata. *PLoS Pathog.* **8,** e1002718 (2012).
75. Thompson, D. A. *et al*. Fungal regulatory evolution: cis and trans in the balance. *FEBS Lett.* **583,** 3959–3965 (2009).
76. Roy, S. *et al*. Evolution of regulatory networks in *Candida glabrata* : learning to live with the human host. *FEMS Yeast Res.* **15,** fov087 (2015).
77. Srikantha, T. *et al*. Phenotypic switching in Candida glabrata accompanied by changes in expression of genes with deduced functions in copper detoxification and stress. *Eukaryot. Cell* **4,** 1434–1445 (2005).
78. Lachke, S. *et al*. Phenotypic switching and filamentation in Candida glabrata. *Microbiology* **148,** 2661–2674 (2002).
79. Dujon, B. *et al.* Genome evolution in yeasts. *Nature* **430,** 35–44 (2004).
80. Castaño, I. *et al.* Telomere length control and transcriptional regulation of subtelomeric adhesins in Candida glabrata. *Mol. Microbiol.* **55,** 1246–1258 (2005).
81. Seider, K. *et al.* The Facultative Intracellular Pathogen Candida glabrata Subverts Macrophage Cytokine Production and Phagolysosome Maturation. *J. Immunol.* **187,** 3072–3086 (2011).
82. Seider, K. *et al.* Immune Evasion, Stress Resistance, and Efficient Nutrient Acquisition Are Crucial for Intracellular Survival of Candida glabrata within Macrophages. *Eukaryot. Cell* **13,** 170–183 (2014).
83. Seider, K. *et al*. Interaction of pathogenic yeasts with phagocytes: survival, persistence and escape. *Curr. Opin. Microbiol.* **13,** 392–400 (2010).
84. Brunke, S. *et al.* One small step for a yeast--microevolution within macrophages renders Candida glabrata hypervirulent due to a single point mutation. *PLoS Pathog.* **10,** e1004478 (2014).
85. Li, H. *et al*. Evolution of Transcription Networks - Lessons from Yeasts. *Curr Biol* **18,** 1199–1216 (2013).
86. Zheng, W. *et al*. Regulatory variation within and between species. *Annu. Rev. Genomics Hum. Genet.* **12,** 327–346 (2011).
87. Necsulea, A. *et al*. Evolutionary dynamics of coding and non-coding transcriptomes. *Nat. Rev. Genet.* **15,** 734–748 (2014).
88. Cuéllar-Cruz, M. *et al.* High resistance to oxidative stress in the fungal pathogen Candida glabrata is mediated by a single catalase, Cta1p, and is controlled by the transcription factors Yap1p, Skn7p, Msn2p, and Msn4p. *Eukaryot. Cell* **7,** 814–825 (2008).
89. McGuire, A. M. *et al*. Conservation of DNA regulatory motifs and discovery of new motifs in microbial genomes. *Genome Res.* **10,** 744–757 (2000).
90. Moses, A. M. *et al*. Position specific variation in the rate of evolution in transcription factor binding sites. *BMC Evol. Biol.* **3,** 19 (2003).
91. Gasch, A. P. *et al.* Conservation and evolution of cis-regulatory systems in ascomycete fungi. *PLoS Biol.* **2,** e398 (2004).
92. Stone, J. R. *et al*. Rapid evolution of cis-regulatory sequences via local point mutations. *Mol. Biol. Evol.* **18,** 1764–70 (2001).

93. Kuo, D. *et al.* Evolutionary divergence in the fungal response to fluconazole revealed by soft clustering. *Genome Biol.* **11,** R77 (2010).
94. Wilcox, L. J. Transcriptional Profiling Identifies Two Members of the ATP-binding Cassette Transporter Superfamily Required for Sterol Uptake in Yeast. *J. Biol. Chem.* **277,** 32466–32472 (2002).
95. Linde, J. *et al.* Defining the transcriptomic landscape of Candida glabrata by RNA-Seq. *Nucleic Acids Res.* **43,** 1392–1406 (2015).
96. De Groot, P. W. J. *et al.* The cell wall of the human pathogen Candida glabrata: Differential incorporation of novel adhesin-like wall proteins. *Eukaryot. Cell* **7,** 1951–1964 (2008).
97. Moran, G. P. *et al*. Comparative genomics and the evolution of pathogenicity in human pathogenic fungi. *Eukaryot. Cell* **10,** 34–42 (2011).
98. Luttik, M. a H. *et al.* The Saccharomyces cerevisiae ICL2 gene encodes a mitochondrial 2-methylisocitrate lyase involved in propionyl-coenzyme a metabolism. *J. Bacteriol.* **182,** 7007–7013 (2000).
99. Calcagno, A. M. *et al.* Candida glabrata STE12 is required for wild-type levels of virulence and nitrogen starvation induced filamentation. *Mol Microbiol* **50,** 1309–1318 (2003).
100. Paul, S. *et al*. Identification of Genomic Binding Sites for Candida glabrata Pdr1 Transcription Factor in Wild-Type and 0 Cells. *Antimicrob. Agents Chemother.* **58,** 6904–6912 (2014).
101. Roetzer, A. *et al. Candida glabrata* environmental stress response involves *Saccharomyces cerevisiae* Msn2/4 orthologous transcription factors. *Mol. Microbiol.* **69,** 603–620 (2008).
102. Kaur, R. *et al*. A family of glycosylphosphatidylinositol-linked aspartyl proteases is required for virulence of Candida glabrata. *Proc. Natl. Acad. Sci. U. S. A.* **104,** 7628–33 (2007).

Annex

Annex I

Table 17 – Computational results of motif binding search in FFUL887 isolate when comparing with CBS138 strain for all known binding sites. Binding sites acquired from online platform YEASTRACT (http://www.yeastract.com/consensuslist.php).

ORF	*Name*	*Loss of Binding Sites*	*Novel Binding Sites*	*Additional Binding Sites*	*Maintained Binding Sites*
CAGL0J01177g	Abf1p	39	109	28	4268
CAGL0I05170g	CgCST6	0	7	0	101
CAGL0M04323g	CgAce2	9	35	5	932
CAGL0E04884g	CgAdr1	142	86	148	4638
CAGL0H03487g	Aft1p	1	26	0	618
CAGL0F06259g	Arg80p	0	1	0	4
CAGL0H06875g	Arg81p	3	50	5	787
CAGL0D00462g	Ash1p	403	76	365	7396
CAGL0I02838g	Azf1p	35	97	17	2606
CAGL0L02585g	Bas1p	1	32	3	1018
CAGL0F03069g	CgCAD1	7	36	6	748
CAGL0M03025g	Cat8p	1	7	0	168
CAGL0H01815g	CgCBF1	11	24	5	1169
CAGL0E05434g	Cha4p	0	31	1	496
CAGL0M08800g	CgYAP6	7	36	6	748
CAGL0M06831g	CgCRZ1	1	47	3	907
CAGL0I05170g	CgCST6	0	14	0	202
CAGL0I04180g	CgAMT1	1	44	3	863
CAGL0L03157g	Dal80p	0	0	0	11
CAGL0F06743g	Dal81p	0	2	0	35
CAGL0M04015g	Dal82p	0	2	0	19
CAGL0F07865g	CgUPC2B	3	21	0	290
CAGL0G08866g	Fkh1p	408	182	410	10230
CAGL0G08866g	Fkh2p	274	222	292	9566
CAGL0H07557g	Fzf1p	0	0	0	1
CAGL0K07634g	Gat1p	10	24	3	906
CAGL0L02475g	CgGCN4	12	176	13	4941
CAGL0H05379g	Gcr1p	333	161	288	9440
CAGL0G08107g	Gis1p	55	59	44	2525
CAGL0C02277g	CgCLN3	25	112	20	2988
CAGL0L03674g	Gsm1p	0	21	2	445
CAGL0L03157g	Gzf3p	14	50	9	1715
CAGL0L09339g	Haa1p	44	80	58	2335
CAGL0K12540g	CgHAC1	54	71	36	2997
CAGL0B03421g	Hap1p	1	14	1	235
CAGL0H07843g	Hap2p	6	30	3	620
CAGL0J04400g	Hap3p	6	30	3	620

CAGL0K08624g	Hap4p	6	30	3	620
CAGL0K09900g	CgHAP5	6	30	3	620
CAGL0J03608g	Hcm1p	0	9	0	161
CAGL0H03443g	Hsf1p	38	125	38	4495
CAGL0M09042g	Ime1p	0	22	0	323
CAGL0B01947g	CgINO2	1	11	0	189
CAGL0I07359g	CgINO4	2	13	0	218
CAGL0K03047g	Ixr1p	0	0	0	2
CAGL0B00462g	Kar4p	0	1	0	13
CAGL0H00396g	Leu3p	0	15	0	227
CAGL0K11902g	Lys14p	0	10	0	146
CAGL0M07590g	Mac1p	0	25	2	410
CAGL0D01012g	Mbp1p	2	16	2	472
CAGL0I10769g	Mcm1p	38	140	31	3951
CAGL0B02651g	Met31p	4	10	0	360
CAGL0B02651g	Met32p	0	32	2	446
CAGL0G06688g	CgMET4	1	7	3	412
CAGL0A01628g	CgMIG1	0	0	0	12
CAGL0K09372g	Mig2p	0	0	0	12
CAGL0C02519g	Mig3p	0	0	0	12
CAGL0K03003g	Mot3p	236	692	248	18670
CAGL0F05995g	CgMSN2	18	23	11	966
CAGL0M13189g	CgMSN4	18	23	11	966
CAGL0F07667g	CgMSS11	0	2	0	4
CAGL0L13090g	Ndt80p	2	17	0	341
CAGL0L07480g	Nrg1p	112	114	97	5132
CAGL0J07150g	Oaf1p	1	8	2	200
CAGL0A00451g	CgPDR1	4	17	2	401
CAGL0A00451g	Pdr3p	0	12	0	171
CAGL0L04576g	Pdr8p	0	5	0	113
CAGL0D05170g	CgPHO4	6	46	10	1235
CAGL0J07150g	Pip2p	1	8	2	200
CAGL0D02904g	Ppr1p	2	2	0	122
CAGL0L09691g	Put3p	0	10	0	174
CAGL0K04917g	CgRAP1	1	2	0	91
CAGL0H09988g	Reb1p	1	1	0	253
CAGL0B04895g	Rfx1p	0	0	0	0
CAGL0L01903g	Rgt1p	31	78	31	2151
CAGL0E03762g	Rim101p	1	19	4	360
CAGL0H05621g	CgRLM1	6	57	5	1201
CAGL0K04257g	Rme1p	0	1	0	6
CAGL0G08107g	Rph1p	55	51	44	2351
CAGL0K01727g	CgRPN4	0	0	0	25
CAGL0C05335g	Rtg1p	89	172	88	5405
CAGL0M09405g	Rtg3p	89	172	88	5405
CAGL0I07183g	Sfl1p	0	8	0	141

CAGL0L03377g	Sip4p	1	7	0	175
CAGL0F09097g	CgSKN7	2	22	1	497
CAGL0J06182g	CgSKO1	0	22	1	327
CAGL0M06325g	Smp1p	0	0	0	2
CAGL0I02552g	CgSTB5	198	116	274	5308
CAGL0M01254g	CgSTE12	7	124	10	2411
CAGL0E04312g	Stp1p	4	14	2	284
CAGL0E04312g	Stp2p	3	34	0	828
CAGL0J10956g	CgSUM1	2	17	0	341
CAGL0A04565g	CgSWI4	0	26	0	459
CAGL0E01331g	CgSWI5	9	35	5	932
CAGL0E05434g	Tea1p	0	5	0	87
CAGL0M01716g	Tec1p	40	216	48	5117
CAGL0F05357g	Ume6p	0	1	0	36
CAGL0C01199g	CgUPC2A	3	29	0	408
CAGL0H04367g	CgWAR1	0	2	0	76
CAGL0G02739g	Xbp1p	37	88	45	3149
CAGL0H04631g	CgAP1	32	170	28	4017
CAGL0K02585g	CgYAP3	7	47	6	1054
CAGL0K08756g	CgAP5	7	36	6	748
CAGL0L04400g	Yrr1p	4	47	3	711
CAGL0J05060g	Zap1p	0	0	0	17

Annex II

Table 18 – Result from the computational search of the PDRE HYCCRKGGRN in FFUL887 isolate and CBS138 strain.

	Binding Site HYCCRKGGRN	Name	Function	Fold
Novel Binding Sites	CAGL0A01366g	CgEPA9	Putative Adhesin	0.99
	-	E	Subunit a of vacuolar-ATPase V0 domain	1.26
	J	-	Unknown Function	1.38
	-	H	Monoglyceride lipase (MGL)	0.84
	-	C	Syntaxin-like vacuolar t-SNARE	1.14
	-	D	Component of the ESCRT-I complex	0.63
	-	G	Putative protein kinase	1.12
	-	F	Ornithine transporter of the mitochondrial inner membrane	0.92
	-	CgA	Protein component of the small (40S) ribosomal subunit	0.39
	-	I	Subunit of the COMPASS (Set1C) histone H3K4 methyltransferase complex	0.96
Loss of Binding Sites	CAGL0A01650g	ECL1	Unknown Function	0.69
	CAGL0G01122g	YLR346C	Unknown Function	1.21
Addition of Binding Sites	-	B	Cell wall mannoprotein	0.92
	Binding Site NYCCMYGGRD	**Name**	**Function**	**Fold**
Novel Binding Sites	CAGL0A03080g	SDH6	Mitochondrial protein involved in assembly of succinate dehydrogenase	0.69
	CAGL0D05060g	PPT1	Protein serine/threonine phosphatase	0.93
	CAGL0J11220g	RPS3	Protein component of the small (40S) ribosomal subunit	0.64
	CAGL0K04279g	CgSCM4	Mitochondrial mitophagy-specific protein	1.18

	CAGL0M02981g	PRM15	Phosphoribomutase	1.60
	CAGL0M03487g	RPC34	RNA polymerase III subunit C34	0.67
	CAGL0M04147g	VNX1	Calcium/H+ antiporter localized to the endoplasmic reticulum membrane;	1.36
Loss of Binding Sites	CAGL0J06182g	CgSKO1	bZIP domain-containing protein	1.59
	CAGL0K00715g	CgRTA1	Putative protein involved in 7-aminocholesterol resistance	0.55
Addition of Binding Sites	CAGL0F08745g	TMA10	Unknown Function	1.10

Annex III

Table 19 – New possible target genes that harbour Adr1 binding site in FFUL887 isolate when compared to CBS138 strain.

	Binding Site NNGGRGA	Name	Function	Fold-value	Binding Site TCYCCNN	Name	Function	Fold-value
Novel Binding Sites	CAGL0A04741g	RET2	Delta subunit of the coatomer complex (COPI)	1.19	CAGL0A01193g	ACM1	Pseudosubstrate inhibitor of the APC/C	1.36
	CAGL0B03663g	CIT1	Citrate synthase	1.52	CAGL0A04191g	SBH2	Ssh1p-Sss1p-Sbh2p complex component	1.01
	CAGL0B03927g	TMA22	Unknown Function	0.94	CAGL0A04719g	PRE4	Beta 7 subunit of the 20S proteasome	1.26
	CAGL0C01705g	CgGPX2	Putative glutathione peroxidase	1.35	CAGL0B00418g	PRD1	Zinc metalloendopeptidase	1.19
	CAGL0C02431g	FIR1	Protein involved in 3' mRNA processing	1.14	CAGL0C02189g	CgSAH1	adenosylhomocysteinase activity	1.02
	CAGL0C03828g	EST3	Component of the telomerase holoenzyme	1.02	CAGL0C04026g	MNT3	Alpha-1,3-mannosyltransferase	0.57
	CAGL0D00858g	RPS29A	Protein component of the small (40S) ribosomal subunit	0.36	CAGL0C04917g	CPA2	Large subunit of carbamoyl phosphate synthetase	1.36
	CAGL0D01166g	MUD1	U1 snRNP A protein	1.10	CAGL0D03168g	CYR1	Adenylate cyclase	1.32
	CAGL0D05478g	TFG1	TFIIF (Transcription Factor II) largest subunit	1.13	CAGL0E00605g	AHC2	Component of the ADA histone acetyltransferase complex	0.90
	CAGL0E00275g	CgEPA20	Putative adhesin	0.78	CAGL0E01639g	VPS68	Unknown Function	1.38
	CAGL0F01683g	RPL22A	Ribosomal 60S subunit protein L22A	0.11	CAGL0F01639g	PER33	Protein that localizes to the endoplasmic reticulum	0.93
	CAGL0F02475g	PRP21	Subunit of the SF3a splicing factor complex	0.89	CAGL0F06743g	DAL81	Positive regulator of genes in multiple nitrogen degradation pathways	1.34
	CAGL0F03333g	RRN9	Protein involved in promoting high level transcription of rDNA	0.83	CAGL0G00726g	CgRFA1	DNA replication factor A, 69 KD subunit	1.18
	CAGL0F04477g	PRE7	Beta 6 subunit of the 20S proteasome	1.07	CAGL0G01144g	KAP95	Karyopherin beta	0.93
	CAGL0F07711g	MME1	Transporter of the mitochondrial inner membrane that exports magnesium	0.92	CAGL0G10021g	PZF1	Transcription factor IIIA (TFIIIA)	1.32

CAGL0F07887g	TCP1	Alpha subunit of chaperonin-containing T-complex	0.87	CAGL0H03289g	YPL191C	Unknown Function	0.77
CAGL0G01056g	CgGAS3	Putative glycoside hydrolase of the Gas/Phr family	0.62	CAGL0H07249g	CgDYS1	Putative deoxyhypusine synthase	0.94
CAGL0G09042g	-	Unknown Function	1.20	CAGL0H07667g	PDE1	Low-affinity cyclic AMP phosphodiesterase	1.27
CAGL0H04191g	ATP18	Subunit of the mitochondrial F1F0 ATP synthase	1.20	CAGL0H09460g	FAA2	Medium chain fatty acyl-CoA synthetase	2.06
CAGL0H07029g	MRPL35	Mitochondrial ribosomal protein of the large subunit	0.90	CAGL0H10202g	TSC3	Protein that stimulates the activity of serine palmitoyltransferase	1.42
CAGL0H07601g	HFM1	Meiosis specific DNA helicase	0.91	CAGL0I01694g	RAD26	Protein involved in transcription-coupled nucleotide excision repair	1.00
CAGL0I02882g	YOR111W	Unknown Function	0.91	CAGL0I04576g	APC11	Catalytic core subunit	0.84
CAGL0I02904g	-	Unknown Function	1.34	CAGL0I07051g	CLG1	Cyclin-like protein that interacts with Pho85p	1.58
CAGL0J03124g	ARG5,6	Acetylglutamate kinase and N-acetyl-gamma-glutamyl-phosphate reductase	1.05	CAGL0I10200g	CgPWP3	putative adhesin-like protein	0.36
CAGL0J04488g	VPS36	Component of the ESCRT-II complex	0.99	CAGL0J00869g	RPC25	RNA polymerase III subunit C25	0.69
CAGL0J04796g	MAG2	Cytoplasmic protein of unknown function	0.97	CAGL0J10780g	OSH3	Member of an oxysterol-binding protein family	0.81
CAGL0J08162g	LYP1	Lysine permease	1.11	CAGL0K00363g	STE6	Plasma membrane ATP-binding cassette (ABC) transporter	0.59
CAGL0J08800g	PNP1	Purine nucleoside phosphorylase	1.19	CAGL0K02739g	LAG1	Ceramide synthase component	1.02
CAGL0J10450g	SEC14	Phosphatidylinositol/phosphatidylcholine transfer protein	1.12	CAGL0K07111g	DAD3	Essential subunit of the Dam1 complex (aka DASH complex)	1.13
CAGL0J11770g	CgPBL1	Putative phospholipase B	1.88	CAGL0K09878g	APP1	Phosphatidate phosphatase	0.84
CAGL0K00341g	-	Unknown Function	0.58	CAGL0K10010g	RRP8	Nucleolar S-adenosylmethionine-dependent rRNA methyltransferase	1.21
CAGL0K08316g	CgRHO4	Non-essential small GTPase	1.12	CAGL0L00407g	SGS1	RecQ family nucleolar DNA helicase	1.34

	CAGL0K10208g	PPH3	Catalytic subunit of protein phosphatase PP4 complex	0.80	CAGL0L00781g	CgMET30	F-box protein containing five copies of the WD40 motif	1.07
	CAGL0L00935g	CgAPQ12	Protein required for nuclear envelope morphology	1.09	CAGL0L01067g	PAR32	Unknown Function	0.84
	CAGL0L01551g	SUR7	Plasma membrane protein	0.75	CAGL0L04026g	NCS2	Protein required for uridine thiolation of Lys(UUU) and Glu(UUC) tRNAs	1.00
	CAGL0L04202g	NRK1	Nicotinamide riboside kinase	1.07	CAGL0L06006g	ATG1	Protein serine/threonine kinase	1.49
	CAGL0L05324g	-	Unknown Function	0.54	CAGL0L08954g	ULA1	Protein that activates Rub1p (NEDD8) before neddylation	1.05
	CAGL0L09449g	MSL1	U2B component of U2 snRNP	0.73	CAGL0M01254g	CgSTE12	Putative transcription factor, required for filamentous growth	1.41
	CAGL0L10318g	LPL1	Phospholipase	1.45	CAGL0M03751g	PAP2	Non-canonical poly(A) polymerase	0.79
	CAGL0L11616g	RTC5	Unknown Function	1.03	CAGL0M04345g	USB1	Putative phosphodiesterase specific for U6 snRNA 3' end modification	1.19
	CAGL0L12518g	-	Unknown Function	0.44	CAGL0M07700g	SEC59	Dolichol kinase	1.17
	CAGL0I05720g	TOM40	Component of the TOM (translocase of outer membrane) complex	0.68	CAGL0M09339g	MIX23	Unknown Function	0.92
	CAGL0K06193g	ADA2	Transcription coactivator	0.89				
	CAGL0M07447g	YMR027W	Unknown Function	1.24				
Addition of Binding Sites	CAGL0A01199g	CgDIP5	Putative dicarboxylic amino acid permease	1.33	CAGL0A00363g	LEU1	Isopropylmalate isomerase	1.76
	CAGL0A01243g	CgGIT1	predicted transmembrane transporter activity	0.80	CAGL0A02321g	HXT3	Low affinity glucose transporter of the major facilitator superfamily	1.11
	CAGL0A01738g	OCH1	Mannosyltransferase of the cis-Golgi apparatus	0.98	CAGL0A03190g	NKP1	Central kinetochore protein and subunit of the Ctf19 complex	0.82
	CAGL0A04411g	SHC1	Sporulation-specific activator of Chs3p (chitin synthase III)	0.93	CAGL0A03388g	RPL15B	Ribosomal 60S subunit protein L15B	0.63
	CAGL0B01727g	YDR109C	Putative kinase	1.05	CAGL0A04477g	UBP13	Ubiquitin-specific protease that cleaves Ub-protein fusions	0.99

CAGL0B02860g	-	Unknown Function	1.80	CAGL0B00286g	CHA1	Catabolic L-serine (L-threonine) deaminase	2.24
CAGL0C00231g	FCY21	Putative purine-cytosine permease	0.70	CAGL0B00616g	SPS2	Protein expressed during sporulation	0.82
CAGL0C03267g	CgFPS1	Glycerol transporter	1.28	CAGL0B01683g	EMP70	Protein with a role in cellular adhesion and filamentous growth	1.34
CAGL0C03982g	MNT3	Alpha-1,3-mannosyltransferase	0.89	CAGL0B02541g	RSC9	Component of the RSC chromatin remodeling complex	1.30
CAGL0C04169g	DSF2	Deletion suppressor of mpt5 mutation	1.06	CAGL0B02838g	CgMUP1	High affinity methionine permease	1.39
CAGL0C04499g	PTC1	Type 2C protein phosphatase (PP2C)	0.97	CAGL0B02926g	CgBMT3	Beta mannosyltransferase	0.67
CAGL0C05379g	CgSSB1	Heat shock protein	0.62	CAGL0C00253g	-	Putative cell wall adhesin	0.34
CAGL0D01496g	ISA2	Protein required for maturation of mitochondrial [4Fe-4S] proteins	1.17	CAGL0C00473g	MEC1	Genome integrity checkpoint protein and PI kinase superfamily member	1.36
CAGL0D05456g	-	Unknown Function	0.49	CAGL0C03850g	DOT5	Nuclear thiol peroxidase	1.11
CAGL0E01287g	KGD2	Dihydrolipoyl transsuccinylase	1.29	CAGL0D00286g	CgBMT1	Beta mannosyltransferase	1.72
CAGL0E03113g	RSR1	GTP-binding protein of the Ras superfamily	1.22	CAGL0D00946g	CSR1	Phosphatidylinositol transfer protein	1.13
CAGL0E04312g	STP2	Transcription factor	0.97	CAGL0D01892g	AAT2	Cytosolic aspartate aminotransferase involved in nitrogen metabolism	1.64
CAGL0E04378g	TDA3	Putative oxidoreductase involved in late endosome to Golgi transport	1.94	CAGL0D03894g	CgPRO1	Putative gamma-glutamyl phosphate reductase	2.66
CAGL0E05170g	CgGRE2	Putative methylglyoxal reductase (NADPH-dependent)	2.33	CAGL0D05940g	CgERG1	Squalene epoxidase with a role in ergosterol synthesis	1.37
CAGL0E05676g	TYW1	Iron-sulfer protein required for synthesis of Wybutosine modified tRNA	1.45	CAGL0D06050g	-	Unknown Function	0.79
CAGL0F00583g	GRH1	Acetylated cis-Golgi protein	0.85	CAGL0E02101g	PAP2	Non-canonical poly(A) polymerase	0.95
CAGL0F00759g	PMT2	Protein O-mannosyltransferase of the ER membrane	0.73	CAGL0E02211g	-	Unknown Function	1.17
CAGL0F01815g	-	Unknown Function	0.67	CAGL0E04312g	STP2	activated by proteolytic processing in response to signals from the SPS sensor system for external amino acids	0.97

CAGL0F03003g	HKR1	Mucin family member that functions as an osmosensor in the HOG pathway	1.20	CAGL0E05170g	CgGRE2(A)	Putative methylglyoxal reductase (NADPH-dependent)	2.33
CAGL0F07535g	H	Monoglyceride lipase (MGL)	0.84	CAGL0E05368g	MRS2	Mitochondrial inner membrane Mg(2+) channel	1.13
CAGL0F08745g	TMA10	Unknown Function	1.10	CAGL0E05588g	REV1	Deoxycytidyl transferase	2.43
CAGL0G01034g	FKS1	Putative 1,3-beta-glucan synthase component	1.05	CAGL0E06116g	USV1	Putative transcription factor containing a C2H2 zinc finger;	0.64
CAGL0G03773g	POM33	Transmembrane nucleoporin	1.91	CAGL0E06600g	-	Putative adhesin-like protein	1.20
CAGL0G03905g	ISA1	Protein required for maturation of mitochondrial [4Fe-4S] proteins	0.72	CAGL0F00913g	PSK2	PAS-domain containing serine/threonine protein kinase	1.01
CAGL0G04609g	CgPKH2	Serine/threonine protein kinase	1.06	CAGL0F01111g	OPI10	Protein with a possible role in phospholipid biosynthesis	0.90
CAGL0G04807g	C	Syntaxin-like vacuolar t-SNARE	1.14	CAGL0F01199g	PRE6	Alpha 4 subunit of the 20S proteasome	0.76
CAGL0G05489g	SHS1	Component of the septin ring that is required for cytokinesis	1.00	CAGL0F01485g	CgTIR2	Putative GPI-linked cell wall mannoprotein of the Srp1p/Tip1p family	0.65
CAGL0G06182g	-	Unknown Function	1.03	CAGL0F04213g	AAC3	Mitochondrial inner membrane ADP/ATP translocator	0.83
CAGL0G09273g	YJR015W	Unknown Function	0.81	CAGL0F04763g	ESC8	Protein involved in telomeric and mating-type locus silencing;	0.98
CAGL0G09493g	ULS1	Swi2/Snf2-related translocase	1.00	CAGL0F04851g	NCA2	Protein that regulates expression of Fo-F1 ATP synthase subunits	1.51
CAGL0H01177g	DPP1	Diacylglycerol pyrophosphate (DGPP) phosphatase	1.03	CAGL0F05269g	CDC42	Small rho-like GTPase	0.85
CAGL0H01837g	PTK2	Serine/threonine protein kinase;	1.35	CAGL0F08371g	CgTNA1	High-affinity nicotinic acid transporter; strongly induced under niacin-limiting conditions	0.81
CAGL0H02563g	DDR2	Multi-stress response protein	3.37	CAGL0G04059g	SSL1	Subunit of the core form of RNA polymerase transcription factor TFIIH	0.94

CAGL0H02959g	CgTOS8	RNA polymerase II core promoter proximal region sequence-specific DNA binding	0.73	CAGL0G05632g	YDL218W	Unknown Function	0.44
CAGL0H03575g	SPO1	Meiosis-specific prospore protein	0.77	CAGL0G05720g	NPR1	Protein kinase	1.90
CAGL0H03993g	CIT1	Citrate synthase	1.06	CAGL0H00462g	RPS5	Protein component of the small (40S) ribosomal subunit	0.26
CAGL0H04081g	CgERG13	3-hydroxy-3-methylglutaryl coenzyme A synthase	0.73	CAGL0H04499g	HLJ1	Co-chaperone for Hsp40p	1.36
CAGL0H04675g	TIF35	eIF3g subunit of the eukaryotic translation initiation factor 3 (eIF3)	0.50	CAGL0H06215g	CgGAL11A	Component of the transcriptional Mediator complex that provides interfaces between RNA polymerase II and upstream activator proteins	0.72
CAGL0H05379g	GCR1	Transcriptional activator of genes involved in glycolysis	0.49	CAGL0H07513g	CgIFA38	Predicted ketoreductase involved in production of very long chain fatty acid for sphingolipid biosynthesis	0.93
CAGL0H05577g	SEC16	COPII vesicle coat protein required for ER transport vesicle budding	1.14	CAGL0H08041g	RTT106	Histone chaperone	0.94
CAGL0H06633g	PCK1	Phosphoenolpyruvate carboxykinase	1.08	CAGL0H08541g	NHP6A	High-mobility group (HMG) protein	0.68
CAGL0H08283g	HSP10	chaperone binding	0.96	CAGL0I02904g	-	Unknown Function	1.34
CAGL0I02090g	-	Unknown Function	1.09	CAGL0I03256g	POL32	Third subunit of DNA polymerase delta	0.71
CAGL0I04972g	IES5	Non-essential INO80 chromatin remodeling complex subunit	0.94	CAGL0I05610g	YNR014W	Putative protein of unknown function	2.30
CAGL0I08591g	YER158C	Unknown Function	1.42	CAGL0I06160g	CgPIR4	Pir protein family member, putative cell wall component	1.07
CAGL0J04950g	IRC8	Unknown Function	0.92	CAGL0I07139g	LSC1	Alpha subunit of succinyl-CoA ligase	1.46
CAGL0J05918g	NSG2	Protein involved in regulation of sterol biosynthesis	1.25	CAGL0I07425g	NDJ1	Protein that regulates meiotic SPB cohesion and telomere clustering	0.75
CAGL0J09944g	CgAQR1	Plasma membrane drug:H+ antiporter involved in resistance to drugs and acetic acid	0.75	CAGL0I07513g	PKH2	Serine/threonine protein kinase	1.32

CAGL0J11110g	MRX12	Protein that associates with mitochondrial ribosome	1.48	CAGL0I08745g	PSD2	Phosphatidylserine decarboxylase of the Golgi and vacuolar membranes	0.99
CAGL0K02079g	CgGLC7	Putative serine-threonine phosphoprotein phosphatase 1	1.07	CAGL0J02948g	CgFCY2	Purine-cytosine transporter	0.85
CAGL0K06215g	UTP6	Nucleolar protein	0.85	CAGL0J03146g	RNR3	Minor isoform of large subunit of ribonucleotide-diphosphate reductase	1.09
CAGL0K07414g	RPL20B	Ribosomal 60S subunit protein L20B	0.39	CAGL0J04224g	MDJ1	Co-chaperone that stimulates HSP70 protein Ssc1p ATPase activity	0.92
CAGL0K08646g	KTI12	Protein that plays a role in modification of tRNA wobble nucleosides	0.86	CAGL0J09064g	ARF2	ADP-ribosylation factor	0.73
CAGL0K10956g	COX14	Mitochondrial cytochrome c oxidase (complex IV) assembly factor	0.81	CAGL0J10076g	YNL058C	Unknown Function	0.84
CAGL0K11440g	HTA1	Histone H2A	0.61	CAGL0J11220g	RPS3	Protein component of the small (40S) ribosomal subunit	0.64
CAGL0K11748g	CgA	Protein component of the small (40S) ribosomal subunit	0.39	CAGL0J11242g	RHO5	Non-essential small GTPase of the Rho/Rac family of Ras-like proteins	1.23
CAGL0L03608g	BCH2	Member of the ChAPs (Chs5p-Arf1p-binding proteins) family	1.16	CAGL0K01661g	G	Putative protein kinase, potentially phosphorylated by Cdc28p	1.12
CAGL0L04950g	ERB1	Constituent of 66S pre-ribosomal particles	0.67	CAGL0K02651g	YHL008C	Unknown Function	0.85
CAGL0L07590g	-	Unknown Function	0.68	CAGL0K03707g	EPO1	Protein involved in septin-ER tethering	1.00
CAGL0L07722g	CgPGK1	Putative 3-phosphoglycerate kinase	1.05	CAGL0K05203g	YOP1	Reticulon-interacting protein	1.15
CAGL0L07788g	RSV161	Amphiphysin-like lipid raft protein	0.96	CAGL0K08294g	DYN1	Cytoplasmic heavy chain dynein	1.15
CAGL0L09691g	PUT3	Transcriptional activator	1.13	CAGL0K08536g	CgAP1	Vacuolar aminopeptidase I	1.62
CAGL0L12298g	PDS1	Securin	1.36	CAGL0K11462g	HTB1	Histone H2B	0.41
CAGL0M04147g	VNX1	Calcium/H+ antiporter localized to the endoplasmic reticulum membrane	1.36	CAGL0L01111g	SFA1	Bifunctional alcohol dehydrogenase and formaldehyde dehydrogenase	1.65
CAGL0M08426g	YJL163C	Unknown Function	1.41	CAGL0L04246g	CgRPB2	RNA polymerase II second largest subunit	1.10

CAGL0M08492g	CgPIR3	Pir protein family member	2.36	CAGL0L06138g	TPN1	Plasma membrane pyridoxine (vitamin B6) transporter	1.55
				CAGL0L09361g	-	Unknown Function	0.99
				CAGL0L09691g	PUT3	Transcriptional activator	1.13
				CAGL0L10092g	-	Putative adhesin-like cell wall protein	1.13
				CAGL0L10428g	-	Unknown Function	1.21
				CAGL0L12056g	CgBMH1	14-3-3 protein	0.46
				CAGL0M06523g	RPL21B	Ribosomal 60S subunit protein L21B	0.35
				CAGL0M07656g	CgERG5	Putative C22 sterol desaturase	1.01

Annex IV

Table 20 – Gcr1 identified putative new target genes. The identified genes harbour Gcr1 binding site in FFUL887 isolate when compared to CBS138 strain.

ORF	Name	Function
CAGL0A01738g	OCH1	Mannosyltransferase of the cis-Golgi apparatus, initiates the polymannose outer chain elongation of N-linked oligosaccharides of glycoproteins
CAGL0A03051g	LSM6	Lsm (Like Sm) protein; part of heteroheptameric complexes (Lsm2p-7p and either Lsm1p or 8p): cytoplasmic Lsm1p complex involved in mRNA decay; nuclear Lsm8p complex part of U6 snRNP and possibly involved in processing tRNA, snoRNA, and rRNA
CAGL0B00440g	CAGL0B00440g	NA
CAGL0B03899g	YJR012C	Essential protein of unknown function; proposed involvement in transport based on mass spectrometry analysis of copurifying proteins; partially overlaps neighboring ORF, GPI14/YJR013W
CAGL0C00693g	CCC1	Putative vacuolar Fe2+/Mn2+ transporter; suppresses respiratory deficit of yfh1 mutants, which lack the ortholog of mammalian frataxin, by preventing mitochondrial iron accumulation
CAGL0D03454g	RBG2	Protein with similarity to mammalian developmentally regulated GTP-binding protein
CAGL0E04548g	YOR020W-A	Putative protein of unknown function, conserved in A. gossypii; the authentic, non-tagged protein is detected in highly purified mitochondria in high-throughput studies
CAGL0F02101g	BLM10	Proteasome activator subunit; found in association with core particles, with and without the 19S regulatory particle; required for resistance to bleomycin, may be involved in protecting against oxidative damage; similar to mammalian PA200
CAGL0F07469g	RMD9	Mitochondrial protein required for respiratory growth; mutant phenotype and genetic interactions suggest a role in delivering mt mRNAs to ribosomes; located on matrix face of the inner membrane and loosely associated with mitoribosomes
CAGL0F08657g	PEX31	Peroxisomal integral membrane protein, involved in negative regulation of peroxisome size; partially functionally redundant with Pex30p and Pex32p; probably acts at a step downstream of steps mediated by Pex28p and Pex29p
CAGL0G01232g	KEX1	Protease involved in the processing of killer toxin and alpha factor precursor; cleaves Lys and Arg residues from the C-terminus of peptides and proteins
CAGL0G02255g	HBS1	GTP binding protein with sequence similarity to the elongation factor class of G proteins, EF-1alpha and Sup35p; associates with Dom34p, and shares a similar genetic relationship with genes that encode ribosomal protein components

CAGL0G03729g	NDD1	Transcriptional activator essential for nuclear division; localized to the nucleus; essential component of the mechanism that activates the expression of a set of late-S-phase-specific genes
CAGL0G03861g	TAM41	Mitochondrial protein involved in protein import into the mitochondrial matrix; maintains the functional integrity of the TIM23 protein translocator complex; viability of null mutant is strain-dependent; mRNA is targeted to the bud
CAGL0G04345g	MDV1	Peripheral protein of the cytosolic face of the mitochondrial outer membrane, required for mitochondrial fission; interacts with Fis1p and with the dynamin-related GTPase Dnm1p; contains WD repeats
CAGL0G04653g	AIM8	Protein of unknown function; overexpression causes a cell cycle delay or arrest; null mutant displays severe respiratory growth defect and elevated frequency of mitochondrial genome loss; detected in mitochondria in high-throughput studies
CAGL0G09361g	PDX1	Dihydrolipoamide dehydrogenase (E3)-binding protein (E3BP) of the mitochondrial pyruvate dehydrogenase (PDH) complex, plays a structural role in the complex by binding and positioning E3 to the dihydrolipoamide acetyltransferase (E2) core
CAGL0H00737g	GYP5	GTPase-activating protein (GAP) for yeast Rab family members, involved in ER to Golgi trafficking; exhibits GAP activity toward Ypt1p that is stimulated by Gyl1p, also acts on Sec4p; interacts with Gyl1p, Rvs161p and Rvs167p
CAGL0H05269g	BTS1	Geranylgeranyl diphosphate synthase, increases the intracellular pool of geranylgeranyl diphosphate, suppressor of bet2 mutation that causes defective geranylgeranylation of small GTP-binding proteins that mediate vesicular traffic
CAGL0H07513g	IFA38	Microsomal beta-keto-reductase; contains oleate response element (ORE) sequence in the promoter region; mutants exhibit reduced VLCFA synthesis, accumulate high levels of dihydrosphingosine, phytosphingosine and medium-chain ceramides
CAGL0H08800g	YPL225W	Protein of unknown function that may interact with ribosomes, based on co-purification experiments; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm
CAGL0I02552g	STB5	Transcription factor, involved in regulating multidrug resistance and oxidative stress response; forms a heterodimer with Pdr1p; contains a Zn(II)2Cys6 zinc finger domain that interacts with a pleiotropic drug resistance element in vitro
CAGL0I08107g	NBA1	Protein of unknown function, localizes to the bud neck and cytoplasm; interacts with Nap1p; may interact with ribosomes, based on co-purification experiments; potential Cdc28p substrate
CAGL0I08217g	STE24	Highly conserved zinc metalloprotease that functions in two steps of a-factor maturation, C-terminal CAAX proteolysis and the first step of N-terminal proteolytic processing; contains multiple transmembrane spans
CAGL0J01848g	RPA135	RNA polymerase I subunit A135
CAGL0J04290g	FUS3	Mitogen-activated serine/threonine protein kinase involved in mating; phosphoactivated by Ste7p; substrates include Ste12p, Far1p, Bni1p, Sst2p; inhibits invasive growth during mating by phosphorylating Tec1p, promoting its degradation

CAGL0J08635g	VPS21	GTPase required for transport during endocytosis and for correct sorting of vacuolar hydrolases; localized in endocytic intermediates; detected in mitochondria; geranylgeranylation required for membrane association; mammalian Rab5 homolog
CAGL0J11198g	MRPL22	Mitochondrial ribosomal protein of the large subunit
CAGL0K06325g	HEH2	Inner nuclear membrane (INM) protein; contains helix-extension-helix (HEH) motif, nuclear localization signal sequence; targeting to the INM requires the Srp1p-Kap95p karyopherins and the Ran cycle
CAGL0K08140g	BET3	Hydrophilic protein that acts in conjunction with SNARE proteins in targeting and fusion of ER to Golgi transport vesicles; component of the TRAPP (transport protein particle) complex
CAGL0K12166g	EXO84	Essential protein with dual roles in spliceosome assembly and exocytosis; the exocyst complex (Sec3p, Sec5p, Sec6p, Sec8p, Sec10p, Sec15p, Exo70p, and Exo84p) mediates polarized targeting of secretory vesicles to active sites of exocytosis
CAGL0K12716g	YFL040W	Putative transporter, member of the sugar porter family; YFL040W is not an essential gene
CAGL0L03938g	YNL115C	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to mitochondria; YNL115C is not an essential gene
CAGL0L04928g	NUP100	Subunit of the nuclear pore complex (NPC) that is localized to both sides of the pore; contains a repetitive GLFG motif that interacts with mRNA export factor Mex67p and with karyopherin Kap95p; homologous to Nup116p
CAGL0L05632g	PBS2	MAP kinase kinase that plays a pivotal role in the osmosensing signal-transduction pathway, activated under severe osmotic stress; plays a role in regulating Ty1 transposition
CAGL0L10428g	MSA1	Activator of G1-specific transcription factors, MBF and SBF; involved in regulation of the timing of G1-specific gene transcription and cell cycle initiation; localization is cell-cycle dependent and regulated by Cdc28p phosphorylation
CAGL0M00616g	VPS70	Protein of unknown function involved in vacuolar protein sorting
CAGL0M01166g	THI4	Thiazole synthase, catalyzes formation of a thiazole intermediate during thiamine biosynthesis; required for mitochondrial genome stability in response to DNA damaging agents
CAGL0M01826g	ECM33	GPI-anchored protein of unknown function, has a possible role in apical bud growth; GPI-anchoring on the plasma membrane crucial to function; phosphorylated in mitochondria; similar to Sps2p and Pst1p
CAGL0M02167g	PRM4	Pheromone-regulated protein proposed to be involved in mating; predicted to have 1 transmembrane segment; transcriptionally regulated by Ste12p during mating and by Cat8p during the diauxic shift
CAGL0A03080g	YDR379C-A	Putative protein of unknown function; identified by homology with other fungi; GFP-fusion protein localized to the mitochondrion

CAGL0A04103g	UPS1	Mitochondrial intermembrane space protein that regulates mitochondrial cardiolipin levels, null has defects in Mgm1p processing, integrity of mitochondrial inner membrane complexes, and mitochondrial morphology; ortholog of human PRELI
CAGL0A04257g	TOD6	PAC motif binding protein involved in rRNA and ribosome biogenesis; subunit of the RPD3L histone deacetylase complex; Myb-like HTH transcription factor, similar to Dot6p; hypophosphorylated by rapamycin treatment in a Sch9p-dependent manne
CAGL0B00990g	FRM2	Protein of unknown function, involved in the integration of lipid signaling pathways with cellular homeostasis; expression induced in cells treated with the mycotoxin patulin; has similarity to bacterial nitroreductases
CAGL0B01595g	CAGL0B01595g	NA
CAGL0B01793g	YDR115W	Putative mitochondrial ribosomal protein of the large subunit, has similarity to E. coli L34 ribosomal protein; required for respiratory growth, as are most mitochondrial ribosomal proteins
CAGL0B02013g	MTC5	Protein of unknown function; mtc5 is synthetically sick with cdc13-1
CAGL0C03399g	POP2	RNase of the DEDD superfamily, subunit of the Ccr4-Not complex that mediates 3' to 5' mRNA deadenylation
CAGL0C03784g	URM1	Ubiquitin-like protein with weak sequence similarity to ubiquitin; depends on the E1-like activating enzyme Uba4p; molecular function of the Urm1p pathway is unknown, but it is required for normal growth, particularly at high temperature
CAGL0D01496g	ISA2	Protein required for maturation of mitochondrial and cytosolic Fe/S proteins, localizes to the mitochondrial intermembrane space, overexpression of ISA2 suppresses grx5 mutations
CAGL0E00671g	UBC1	Ubiquitin-conjugating enzyme that mediates selective degradation of short-lived and abnormal proteins; plays a role in vesicle biogenesis and ER-associated protein degradation (ERAD); component of the cellular stress response
CAGL0E01705g	MDH2	Cytoplasmic malate dehydrogenase, one of three isozymes that catalyze interconversion of malate and oxaloacetate; involved in the glyoxylate cycle and gluconeogenesis during growth on two-carbon compounds; interacts with Pck1p and Fbp1
CAGL0F00253g	TRP3	Bifunctional enzyme exhibiting both indole-3-glycerol-phosphate synthase and anthranilate synthase activities, forms multifunctional hetero-oligomeric anthranilate synthase:indole-3-glycerol phosphate synthase enzyme complex with Trp2p
CAGL0F00671g	MYO2	One of two type V myosin motors (along with MYO4) involved in actin-based transport of cargos; required for the polarized delivery of secretory vesicles, the vacuole, late Golgi elements, peroxisomes, and the mitotic spindle
CAGL0F05599g	NUP42	Subunit of the nuclear pore complex (NPC) that localizes exclusively to the cytoplasmic side; involved in RNA export, most likely at a terminal step; interacts with Gle1p
CAGL0F07557g	YJU2	Essential protein required for pre-mRNA splicing; associates transiently with the spliceosomal NTC ("nineteen complex") and acts after Prp2p to promote the first catalytic reaction of splicing

CAGL0G00308g	SCW4	Cell wall protein with similarity to glucanases; scw4 scw10 double mutants exhibit defects in mating
CAGL0G05896g	DSE2	Daughter cell-specific secreted protein with similarity to glucanases, degrades cell wall from the daughter side causing daughter to separate from mother; expression is repressed by cAMP
CAGL0G06248g	MAK16	Essential nuclear protein, constituent of 66S pre-ribosomal particles; required for maturation of 25S and 5.8S rRNAs; required for maintenance of M1 satellite double-stranded RNA of the L-A virus
CAGL0G06644g	CAGL0G06644g	NA
CAGL0G09273g	SNG1	Protein involved in nitrosoguanidine (MNNG) resistance; expression is regulated by transcription factors involved in multidrug resistance
CAGL0H01705g	FPR2	Membrane-bound peptidyl-prolyl cis-trans isomerase (PPIase), binds to the drugs FK506 and rapamycin; expression pattern suggests possible involvement in ER protein trafficking
CAGL0H02607g	SPE2	S-adenosylmethionine decarboxylase, required for the biosynthesis of spermidine and spermine; cells lacking Spe2p require spermine or spermidine for growth in the presence of oxygen but not when grown anaerobically
CAGL0H06545g	ATG32	Mitochondrial-anchored transmembrane receptor that interacts with the autophagy adaptor protein, Atg11p, and is essential for mitophagy, the selective vacuolar degradation of mitochondria in response to starvation
CAGL0I00572g	ATP14	Subunit h of the F0 sector of mitochondrial F1F0 ATP synthase, which is a large, evolutionarily conserved enzyme complex required for ATP synthesis
CAGL0I03432g	YDL160C-A	Putative protein of unknown function
CAGL0I09064g	CDC14	Protein phosphatase required for mitotic exit; located in the nucleolus until liberated by the FEAR and Mitotic Exit Network in anaphase, enabling it to act on key substrates to effect a decrease in CDK/B-cyclin activity and mitotic exit
CAGL0J06374g	RPP1B	Ribosomal protein P1 beta, component of the ribosomal stalk, which is involved in interaction of translational elongation factors with ribosome; accumulation is regulated by phosphorylation and interaction with the P2 stalk component
CAGL0J08657g	PTC5	Mitochondrial type 2C protein phosphatase involved in regulation of pyruvate dehydrogenase activity by dephosphorylating the ser-133 residue of the Pda1p subunit; acts in concert with kinases Pkp1p and Pkp2p and phosphatase Ptc6p
CAGL0J09064g	ARF1	ADP-ribosylation factor, GTPase of the Ras superfamily involved in regulation of coated vesicle formation in intracellular trafficking within the Golgi; functionally interchangeable with Arf2p
CAGL0J09636g	GLE1	Cytoplasmic nucleoporin required for polyadenylated RNA export but not for protein import; component of Nup82p nuclear pore subcomplex; contains a nuclear export signal

CAGL0J09834g	PTC4	Cytoplasmic type 2C protein phosphatase; identified as a high-copy number suppressor of the synthetic lethality of a cnb1 mpk1 double deletion; overexpression decreases high-osmolarity induced Hog1p phosphorylation and kinase activity
CAGL0K00539g	APM1	Mu1-like medium subunit of the clathrin-associated protein complex (AP-1); binds clathrin; involved in clathrin-dependent Golgi protein sorting
CAGL0K01375g	MRP10	Mitochondrial ribosomal protein of the small subunit; contains twin cysteine-x9-cysteine motifs
CAGL0K01661g	YDL025C	Putative protein kinase, potentially phosphorylated by Cdc28p; YDL025C is not an essential gene
CAGL0K04279g	SCM4	Potential regulatory effector of CDC4 function, suppresses a temperature-sensitive allele of CDC4, tripartite protein structure in which a charged region separates two uncharged domains, not essential for mitosis or meiosis
CAGL0K06501g	TLG1	Essential t-SNARE that forms a complex with Tlg2p and Vti1p and mediates fusion of endosome-derived vesicles with the late Golgi; binds the docking complex VFT (Vps fifty-three) through interaction with Vps51p
CAGL0K09614g	KRE33	Essential protein of unknown function; heterozygous mutant shows haploinsufficiency in K1 killer toxin resistance
CAGL0K10428g	YFR017C	Putative protein of unknown function; green fluorescent protein (GFP)-fusion protein localizes to the cytoplasm and is induced in response to the DNA-damaging agent MMS; YFR017C is not an essential gene
CAGL0L01045g	CAGL0L01045g	NA
CAGL0L03960g	DMA2	Protein involved in ubiquitination; plays a role in regulating spindle position and orientation; functionally redundant with Dma1p; orthologous to human RNF8 protein, also has sequence similarity to human Chfr
CAGL0L04686g	CLD1	Mitochondrial cardiolipin-specific phospholipase; functions upstream of Taz1p to generate monolyso-cardiolipin; transcription increases upon genotoxic stress; involved in restricting Ty1 transposition; has homology to mammalian CGI-58
CAGL0L04752g	DAM1	Essential subunit of the Dam1 complex (aka DASH complex), couples kinetochores to the force produced by MT depolymerization thereby aiding in chromosome segregation; Ipl1p target for regulating kinetochore-MT attachments
CAGL0L10846g	BUD23	Methyltransferase, methylates residue G1575 of 18S rRNA; required for rRNA processing and nuclear export of 40S ribosomal subunits independently of methylation activity; diploid mutant displays random budding pattern
CAGL0L11528g	BIG1	Integral membrane protein of the endoplasmic reticulum, required for normal content of cell wall beta-1,6-glucan
CAGL0L12804g	MNN9	Subunit of Golgi mannosyltransferase complex also containing Anp1p, Mnn10p, Mnn11p, and Hoc1p that mediates elongation of the polysaccharide mannan backbone; forms a separate complex with Van1p that is also involved in backbone elongation
CAGL0L13112g	RGR1	Subunit of the RNA polymerase II mediator complex; associates with core polymerase subunits to form the RNA polymerase II holoenzyme; required for glucose repression, HO repression, RME1 repression and sporulation

CAGL0M06457g	GDT1	Putative protein of unknown function; expression is reduced in a gcr1 null mutant; GFP-fusion protein localizes to the vacuole; expression pattern and physical interactions suggest a possible role in ribosome biogenesis
CAGL0M06699g	SDS3	

Annex V

Figure 35 – Restriction map of the vector used in this work, pYEP354w. Map acquired at Snapgene (http://www.snapgene.com/resources/plasmid_files/yeast_plasmids/YEp354/)

Annex VI

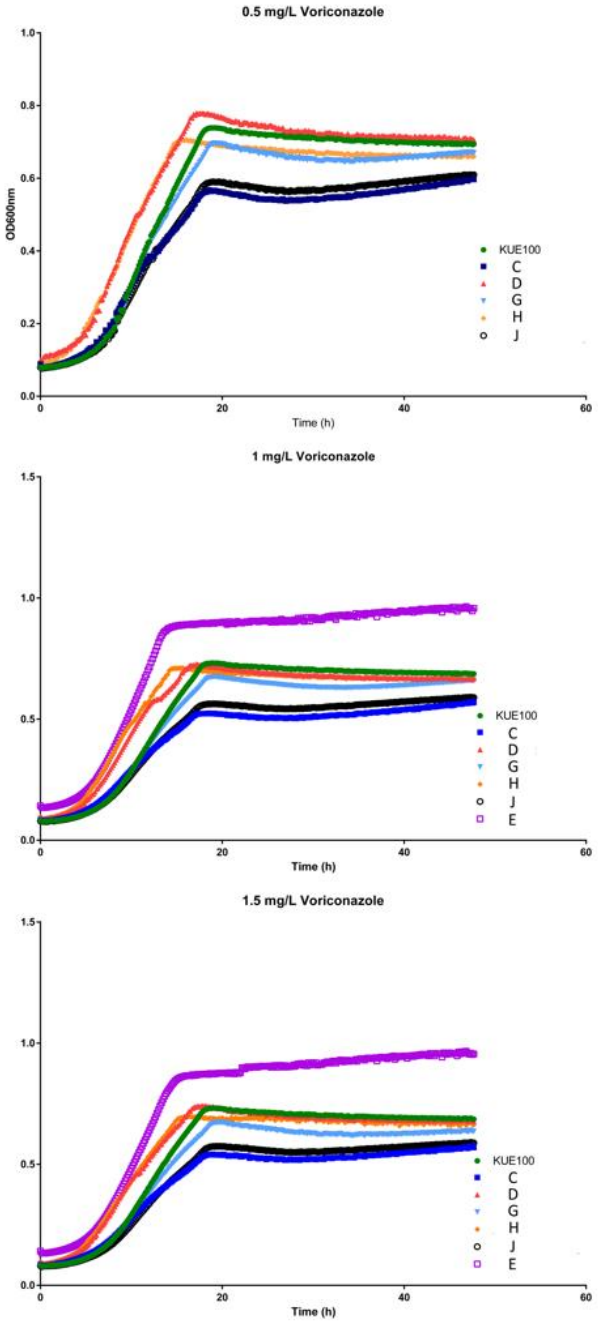


Figure 36 – KUE100 and derived mutants growth curves with 0.5 mg/L , 1 mg/L and 1.5 mg/L voriconazole supplemented in growth medium. Mutants were grown for 48h. Voriconazole concentrations used are according to EUCAST MIC (1 mg/L), ECOFF (www.eucast.org)

Annex VII

Table 21 – List of CgPdr1, CgMsn4 and CgAmt1 documented target genes provided by Professor Miguel Teixeira and Catarina Costa.

TF	*RG ORF*	*RG*	*Reference*
CgPdr1		CgCDR2	http://www.ncbi.nlm.nih.gov/pubmed/18385733/
CgPdr1		CgCDR1	http://www.ncbi.nlm.nih.gov/pubmed/21131438?dopt=Abstract
CgPdr1		CgCDR2	http://www.ncbi.nlm.nih.gov/pubmed/21131438?dopt=Abstract
CgPdr1		CgCDR1	http://www.ncbi.nlm.nih.gov/pubmed/21131438?dopt=Abstract
CgPdr1		CgPDR1	http://www.ncbi.nlm.nih.gov/pubmed/21131438?dopt=Abstract
CgPdr1	CAGL0B02101g		http://www.ncbi.nlm.nih.gov/pubmed/25199772
CgPdr1	CAGL0C02937g		http://www.ncbi.nlm.nih.gov/pubmed/25199773
CgPdr1	CAGL0G05313g	IPT1	http://www.ncbi.nlm.nih.gov/pubmed/25199774
CgPdr1	CAGL0I02574g	OYE2	http://www.ncbi.nlm.nih.gov/pubmed/25199775
CgPdr1	CAGL0J05852g		http://www.ncbi.nlm.nih.gov/pubmed/25199776
CgPdr1	CAGL0J07436g	PDR16	http://www.ncbi.nlm.nih.gov/pubmed/25199777
CgPdr1	CAGL0K01749g	RPN4	http://www.ncbi.nlm.nih.gov/pubmed/25199778
CgPdr1	CAGL0K05819g		http://www.ncbi.nlm.nih.gov/pubmed/25199779
Msn4	CAGL0E00803g	HSP42	http://www.ncbi.nlm.nih.gov/pubmed/18547442
Msn4	CAGL0C04323g	NTH1	http://www.ncbi.nlm.nih.gov/pubmed/18547443
Msn4	CAGL0I07865g	PHM7	http://www.ncbi.nlm.nih.gov/pubmed/18547444
Msn4	CAGL0H06699g	GUT2	http://www.ncbi.nlm.nih.gov/pubmed/18547446
Msn4	CAGL0J04202g	HSP12	http://www.ncbi.nlm.nih.gov/pubmed/18547447
Msn4	CAGL0F04895g	GPH1	http://www.ncbi.nlm.nih.gov/pubmed/18547448
Msn4	CAGL0H08844g	DDR48	http://www.ncbi.nlm.nih.gov/pubmed/18547449
Msn4	CAGL0J09812g	TPS1	http://www.ncbi.nlm.nih.gov/pubmed/18547450
Msn4	CAGL0J06050g	YGP1	http://www.ncbi.nlm.nih.gov/pubmed/18547451
Msn4	CAGL0G09977g	GDB1	http://www.ncbi.nlm.nih.gov/pubmed/18547452
Msn4	CAGL0M13651g	PRC1	http://www.ncbi.nlm.nih.gov/pubmed/18547453
Msn4	CAGL0D00198g	BDH1	http://www.ncbi.nlm.nih.gov/pubmed/18547454
Msn4	CAGL0M03377g	GLC3	http://www.ncbi.nlm.nih.gov/pubmed/18547455
Msn4	CAGL0G05335g	TPS2	http://www.ncbi.nlm.nih.gov/pubmed/18547456
Msn4	CAGL0K03421g	PGM2	http://www.ncbi.nlm.nih.gov/pubmed/18547457
Msn4	CAGL0M07205g	NVJ1	http://www.ncbi.nlm.nih.gov/pubmed/18547458
Msn4	CAGL0K07480g	PGM2	http://www.ncbi.nlm.nih.gov/pubmed/18547459
Msn4	CAGL0J10846g	PCL5	http://www.ncbi.nlm.nih.gov/pubmed/18547460
Msn4	CAGL0E00803g	HSP42	http://www.ncbi.nlm.nih.gov/pubmed/18547442
Msn4	CAGL0C04323g	NTH1	http://www.ncbi.nlm.nih.gov/pubmed/18547443
Msn4	CAGL0H06699g	GUT2	http://www.ncbi.nlm.nih.gov/pubmed/18547446
Msn4	CAGL0J04202g	HSP12	http://www.ncbi.nlm.nih.gov/pubmed/18547447
Msn4	CAGL0F04895g	GPH1	http://www.ncbi.nlm.nih.gov/pubmed/18547448
Msn4	CAGL0H08844g	DDR48	http://www.ncbi.nlm.nih.gov/pubmed/18547449
Msn4	CAGL0J09812g	TPS1	http://www.ncbi.nlm.nih.gov/pubmed/18547450
Msn4	CAGL0J06050g	YGP1	http://www.ncbi.nlm.nih.gov/pubmed/18547451
Msn4	CAGL0G05335g	TPS2	http://www.ncbi.nlm.nih.gov/pubmed/18547456
Msn4	CAGL0K07480g	PGM2	http://www.ncbi.nlm.nih.gov/pubmed/18547459
Msn4	CAGL0J10846g	PCL5	http://www.ncbi.nlm.nih.gov/pubmed/18547460
Msn4	CAGL0L10582g	YMR196W	http://www.ncbi.nlm.nih.gov/pubmed/18547461
Msn4	CAGL0L11132g	SML1	http://www.ncbi.nlm.nih.gov/pubmed/18547462
Msn4	CAGL0E00803g	HSP42	http://www.ncbi.nlm.nih.gov/pubmed/18547442
Msn4	CAGL0C04323g	NTH1	http://www.ncbi.nlm.nih.gov/pubmed/18547443
Msn4	CAGL0I07865g	PHM7	http://www.ncbi.nlm.nih.gov/pubmed/18547444

Msn4	CAGL0B04323g	ETR1	http://www.ncbi.nlm.nih.gov/pubmed/18547445
Msn4	CAGL0H06699g	GUT2	http://www.ncbi.nlm.nih.gov/pubmed/18547446
Msn4	CAGL0J04202g	HSP12	http://www.ncbi.nlm.nih.gov/pubmed/18547447
Msn4	CAGL0F04895g	GPH1	http://www.ncbi.nlm.nih.gov/pubmed/18547448
Msn4	CAGL0H08844g	DDR48	http://www.ncbi.nlm.nih.gov/pubmed/18547449
Msn4	CAGL0J09812g	TPS1	http://www.ncbi.nlm.nih.gov/pubmed/18547450
Msn4	CAGL0J06050g	YGP1	http://www.ncbi.nlm.nih.gov/pubmed/18547451
Msn4	CAGL0G09977g	GDB1	http://www.ncbi.nlm.nih.gov/pubmed/18547452
Msn4	CAGL0M13651g	PRC1	http://www.ncbi.nlm.nih.gov/pubmed/18547453
Msn4	CAGL0D00198g	BDH1	http://www.ncbi.nlm.nih.gov/pubmed/18547454
Msn4	CAGL0M03377g	GLC3	http://www.ncbi.nlm.nih.gov/pubmed/18547455
Msn4	CAGL0G05335g	TPS2	http://www.ncbi.nlm.nih.gov/pubmed/18547456
Msn4	CAGL0K03421g	PGM2	http://www.ncbi.nlm.nih.gov/pubmed/18547457
Msn4	CAGL0M07205g	NVJ1	http://www.ncbi.nlm.nih.gov/pubmed/18547458
Msn4	CAGL0K07480g	PGM2	http://www.ncbi.nlm.nih.gov/pubmed/18547459
Msn4	CAGL0J10846g	PCL5	http://www.ncbi.nlm.nih.gov/pubmed/18547460
Msn4	CAGL0L10582g	YMR196W	http://www.ncbi.nlm.nih.gov/pubmed/18547461
Msn4	CAGL0L11132g	SML1	http://www.ncbi.nlm.nih.gov/pubmed/18547462
Msn4	CAGL0E00803g	HSP42	http://www.ncbi.nlm.nih.gov/pubmed/18547442
Msn4	CAGL0H08844g	DDR48	http://www.ncbi.nlm.nih.gov/pubmed/18547449
Msn4	CAGL0J09812g	TPS1	http://www.ncbi.nlm.nih.gov/pubmed/18547450
Msn4	CAGL0J06050g	YGP1	http://www.ncbi.nlm.nih.gov/pubmed/18547451
Msn4	CAGL0G09977g	GDB1	http://www.ncbi.nlm.nih.gov/pubmed/18547452
Msn4	CAGL0M13651g	PRC1	http://www.ncbi.nlm.nih.gov/pubmed/18547453
Msn4	CAGL0G05335g	TPS2	http://www.ncbi.nlm.nih.gov/pubmed/18547456
Msn4	CAGL0M07205g	NVJ1	http://www.ncbi.nlm.nih.gov/pubmed/18547458
Msn4	CAGL0J10846g	PCL5	http://www.ncbi.nlm.nih.gov/pubmed/18547460
Amt1	CAGL0H04257g		http://www.ncbi.nlm.nih.gov/pubmed/1508182
Amt1	CAGL0H04279g		http://www.ncbi.nlm.nih.gov/pubmed/1508182
Amt1	CAGL0L01831g		http://www.ncbi.nlm.nih.gov/pubmed/1508182
Amt1	CAGL0H04257g		http://www.ncbi.nlm.nih.gov/pubmed/1508182
Amt1	CAGL0H04279g		http://www.ncbi.nlm.nih.gov/pubmed/1508182
Amt1	CAGL0L01831g		http://www.ncbi.nlm.nih.gov/pubmed/1508182
Amt1	CAGL0H04257g		http://www.ncbi.nlm.nih.gov/pubmed/8509391
Amt1	CAGL0H04279g		http://www.ncbi.nlm.nih.gov/pubmed/8509391
Amt1	CAGL0L01831g		http://www.ncbi.nlm.nih.gov/pubmed/8509391