Using CRISPR-Cas9 to Study the Evolutionary Role of Genome Instability in *Candida albicans*

By Matthew Shaw

2022

Submitted to the University of Kent for the degree of MSc Microbiology

1 Declaration

No part of this thesis has been submitted in support of an application for any degree or other qualification of the University of Kent, or any other University or Institution of learning.

2 Acknowledgements

I would like to thank my supervisor Dr Alessia Buscaino for her insight and guidance throughout this project. Her vast knowledge and enthusiasm have motivated me to achieve my best and I look forward to continuing in her lab.

I am also sincerely grateful to the other members of the Buscaino lab: Dr Elise Iracane, Dr Samuel Vega Estevez and Chloe Uyl. Their instruction, help, and friendship enabled me not only to complete this project, but greatly enjoy doing so. Similarly, I appreciate the other members of the Kent Fungal Group for providing a fun, supportive environment.

I would also like to thank Dr Jordan Price at NIAB EMR for the genome sequencing he contributed to this project.

My final thanks go to my family, friends, and partner Jasmine, all of whose support has kept me going throughout.

Contents

3 List of Figures

4 List of Tables

5 Abbreviations

- Cas9 — CRISPR-associated protein 9
- CCL — Concerted chromosome loss
- CHEF — Clamped-homogenous electric field
- CRISPR/Cas9 — Clustered regularly interspaced short palindromic repeats
- DMSO — Dimethyl sulfoxide
- DSB — Double Strand break
- EDTA — Ethylenediaminetetraacetic acid
- HIV — Human immunodeficiency virus
- HR — Homologous recombination
- IC — Invasive candidiasis
- LOH — Loss of heterozygosity
- MIC — Minimum inhibitory concentration
- MRS — Major repeat sequence
- NHEJ — Non-homologous end joining
- NAHR — Non-allelic homologous recombination
- ORF — Open reading frame
- SNP — Single nucleotide polymorphism
- TF — Transcription factor
- WT — Wild-Type

6 Abstract

Candida albicans is a typical member of the human microbiome which can also cause a range of infections from superficial to systemic. The ability of *C. albicans* to survive the broad range of conditions it encounters in the human host is partially attributed to its genomic instability. This instability generates diversity and allows selection of fitter genotypes which can thrive in the hostile host microenvironments. Consequently, *C. albicans* clinical isolates have diverse karyotypes. *C. albicans* chromosomal rearrangements often occur around repetitive elements including the Major Repeat Sequence (MRS), a conserved repeat array of unknown function occurring on seven of the eight *C. albicans* chromosomes. We hypothesise that MRS elements serve as instability hotspots to facilitate genomic rearrangements and rapid evolution.

This study aims to establish a cause-and-effect relationship between MRS-driven chromosome rearrangements and generation of novel, fitter genotypes. To this end, we have used a CRISPR-Cas9 approach to generate double strand breaks within the MRS, inducing chromosome rearrangements in unstressed standard laboratory growth conditions. These unstable strains have then been evolved in clinically relevant stresses, including antifungal drugs.

Long read sequencing shows numerous translocations have been successfully generated around the MRS in these unstable strains. CRISPR-Cas9 can therefore be used to generate chromosomal rearrangements in *C. albicans.* Strains bearing translocations had morphological and fitness changes, indicating that rearrangements at the MRS can generate diversity. Evolution experiments showed that MRS-destabilized strains had higher genome instability than WT *C. albicans*, undergoing significantly more rearrangements throughout. In one instance, this increased plasticity resulted in faster adaptation to stress.

This study demonstrates the ability to generate karyotypic alterations in *C. albicans* using CRISPR-Cas9. Furthermore, it supports the evolutionary role of the MRS as a facilitator of genome plasticity, helping to generate karyotypic diversity. There is some evidence that this enables *Candida* to adapt to environmental stress, but this requires further investigation.

7 Introduction

7.1 *Candida* Taxonomy and Phylogeny

Candida albicans is a pleomorphic yeast of the order Saccharomycetales, a division of the fungal phylum Ascomycota. Within the genus *Candida,* there are over 200 species, but these do not all share a single evolutionary origin. Most *Candida* species, including three of the five most associated with disease, are part of the CTG clade (fig. 1) (Butler et al., 2009; Maguire et al., 2013). The defining feature of this clade is that the CTG codon which usually encodes leucine, instead encodes serine (Santos et al., 1993). This is true for around 97% of instances where this codon is translated in *C. albicans* (Suzuki et al., 1997). Addition of leucine is still tolerated in *C. albicans,* the rate of which increases under stress (Gomes et al., 2007; Turner and Butler, 2014). The CTG clade can be further divided by ploidy, with some species including *C. albicans* usually or always diploid, while others are haploid.

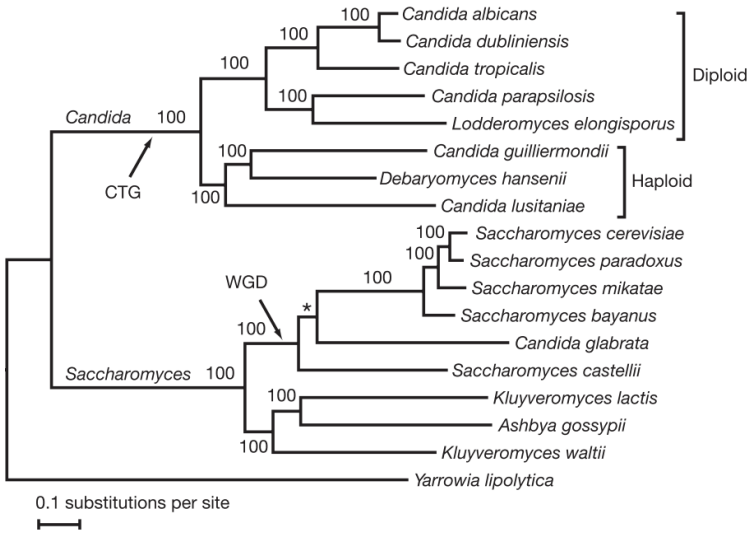

Figure 1. Candida and Saccharomyces phylogenetic tree. *From Butler et al., 2009.*

Tree based on whole genome sequences. Constructed with MrBayes, a program for Baysian inference, utilizing the Markov chain Monte Carlo method. Branch lengths indicate sequence divergence and posterior probabilities are written on each branch. C. albicans is most closely related to C. dubliniensis followed by C. tropicalis. C. glabrata is more closely related to Saccharomyces species than to other Candida species.

7.2 Epidemiology of *Candida* Infection

In over 60% of healthy humans, *C. albicans* forms part of the microbiome, colonising the oropharynx, oesophagus, gastrointestinal tract, genitourinary tract, and skin (Bertolini and Dongari-Bagtzoglou, 2019; Pappas et al., 2018). The fungal component of the microbiome, the mycobiome, often contains multiple *Candida* species, which have been identified as the most frequently fungal genus of the oral cavity (Ghannoum et al., 2010).

C. albicans is also an opportunistic pathogen and alterations to the host environment, host health or bacterial microbiome can initiate this pathogenicity. This change in pathogenicity state is typically associated with a morphological switch from yeast to filamentous growth (Jacobsen et al., 2014).

Candidiases, the infections caused by *Candida* species vary significantly in severity. These include superficial skin or mucosal infections including oral and vulvovaginal candidiasis as well as invasive candidiases (IC) which affect the blood (candidemia), brain, eyes, bones, liver, and any other tissue. At the extreme, IC can lead to fulminant sepsis, which has a mortality rate of >70% (Pappas et al., 2018).

IC is particularly prevalent in immunocompromised individuals including neonates (Greenberg and Benjamin, 2014), HIV patients (Torssander et al., 2009) and those undergoing chemotherapy (Teoh and Pavelka, 2016). Furthermore, prolonged ICU stays, organ transplants, surgery (especially abdominal), central vascular catheter use (Cleveland et al., 2015), dialysis and antibiotic use are also significant factors which are associated with elevated rates of IC (Kullberg and Arendrup, 2015).

This association with immunocompromised individuals and medical treatment, in addition to the ubiquity of *Candida* species means *Candida* account for a significant proportion of hospital acquired infections, including around 20% of all healthcare-associated blood stream infections (Pappas et al., 2018; Simonetti et al., 2013; Vincent et al., 2009; Wisplinghoff et al., 2004). Such interaction with other diseases makes it difficult to determine the mortality rate of systemic *Candida* infection. Unadjusted

estimates range from 29-75% (Pfaller and Diekema, 2007; Strollo et al., 2017), but actual rates are likely closer to 10-20% (Pappas et al., 2018). This does however vary geographically and temporally.

At least 15 *Candida* species can cause Candidiasis, however *C. albicans* is the most common, generally responsible for half of all Candidiases (Cleveland et al., 2015; Guinea, 2014; Pfaller et al., 2011; Wisplinghoff et al., 2004). Along with *C. glabrata*, *C. tropicalis*, *C. parapsilosis* and *C. krusei, C. albicans* is also responsible for most IC cases (Cleveland et al., 2015; Pappas et al., 2018; Wisplinghoff et al., 2004). While *C. albicans* is almost always the most prevalent, the incidence of the non-*albicans Candida* species varies geographically. Of the non-*albicans* species, the most common agent of IC in Northwestern Europe, North America and Australia is *C. glabrata,* which is not a member of the CTG clade and is morphologically and phylogenetically closer to *S. cerevisiae* (fig. 1) (Astvad et al., 2018; Chapman et al., 2017; Lortholary et al., 2017; Pappas et al., 2018; Pfaller et al., 2011; St-Germain et al., 2008). In South America and Asia, *C. parapsilosis* dominates (Doi et al., 2016; Morii et al., 2014; Pappas et al., 2018).

Globally, IC is estimated to affect at least 250,000 people every year, and accounts for >50,000 deaths (Kullberg and Arendrup, 2015). Large population studies in the US have estimated occurrence of IC as being between 3 and 5 in 100,000 individuals (Pappas et al., 2018; Strollo et al., 2017). The cost per case in the US is estimated around $45,000 (Strollo et al., 2017), where in 2017, *Candida* infections were estimated to have cost over $3 billion, including >$1.2 billion for IC (Benedict et al., 2019).

7.3 Drug Treatment of *Candida* Infection

7.3.1 Classes and Modes of Action of Current Antifungal Drugs

Despite the global impact of IC and other fungal diseases, only three classes of antifungal drug are available: polyenes, echinocandins and azoles. Polyenes such as Amphotericin B (Am B) were the first developed antifungals. This drug readily aggregates outside of cells and fungicidally extracts ergosterol from the fungal cell membrane (Anderson et al., 2014; Lee et al., 2021) (fig. 2). The risk of toxicity towards the human host due to similarity between ergosterol and cholesterol makes this class of drug unfavourable (Lee et al., 2021; Ostrosky-Zeichner et al., 2010). Resistance to Am B is low in *C. albicans* due to fitness costs associated with acquired resistance, which leave the yeast very susceptible to stress and severely impair its pathogenicity (Vincent et al., 2013).

Echinocandins, including Caspofungin competitively bind (1,3)-β-D-glucan synthase, impairing (1,3)-β-D-glucan biosynthesis, resulting in cell wall destabilisation and osmotic stress with fungicidal effects (Lee et al., 2021; Letscher-Bru and Herbrecht, 2003) (fig. 2). The lack of cell wall in human cells significantly reduces the risk of host toxicity making this class of drugs favourable to polyenes, however both have poor oral absorption and are predominantly used intravenously (Lee et al., 2021).

Azoles are the most widely used class of antifungal and act by inhibiting the cytochrome P450 (*ERG11*) binding to the haem group in its active site. This impairs ergosterol biosynthesis, destabilising the membrane with toxic effects (fig. 2). Stalling of this synthesis also results in the accumulation of sterol intermediates which are cytotoxic (Lee et al., 2021; Ostrosky-Zeichner et al., 2010; Robbins et al., 2016). One such intermediate is 14-α-methyl-3,6-diol, which is produced by an Δ-5,6-desaturase, encoded by *ERG3* (Lee et al., 2021). Azoles are generally fungistatic rather than fungicidal which can allow the evolution of drug resistance (Lee et al., 2021; Robbins et al., 2016; Roemer and Krysan, 2014).

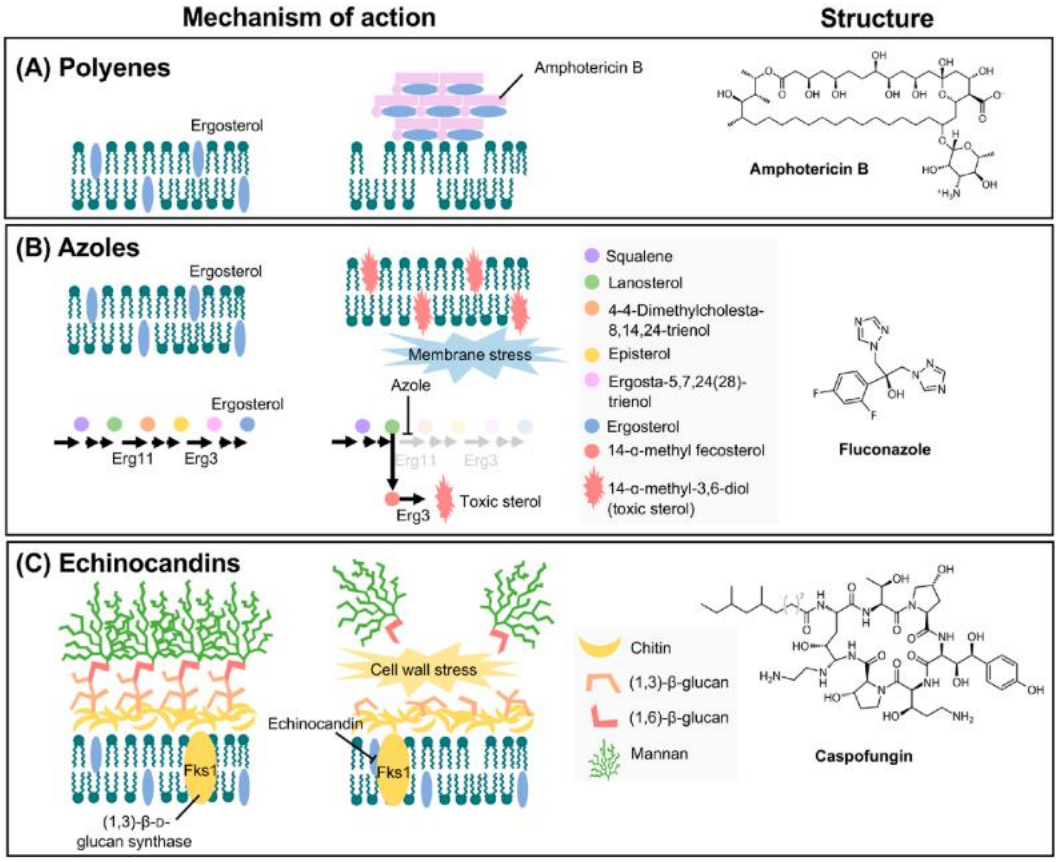

Figure 2. Modes of action of antifungal drug classes*. From Lee et al., 2021*

Cartoons display the modes of action of the key classes of antifungal drugs. Skeletal structure of the most common drug of each class is shown adjacent. Polyenes (A) accumulate into a 'sterol sponge' which extracts ergosterol from fungal cell membranes. Azoles (B) prevent ergosterol biosynthesis by Erg11, destabilizing the membrane and generating toxic intermediates. Echinocandins (C) target fungal cell walls by inhibiting (1,3)-β-D-glucan biosynthesis.

7.3.2 Development of antifungal drugs

The very limited range of classes of antifungal drugs is partially responsible for the high mortality rates of systemic *Candida* infections. Development of these drugs is however challenging; fungi are eukaryotic organisms, which are significantly closer related to the human host than bacteria or viruses. Consequently, molecules with toxic effects on fungi are more likely to have a similar effect on the host (Roemer and Krysan, 2014). Clinical trials looking at invasive mycoses are also challenged by the epidemiology of the diseases, which often occur alongside other conditions or diseases (Tollemar, 2001).

This slow progress in developing new antifungals, coupled with the global increase in antifungal drug resistance (Lee et al., 2021), is concerning and illustrates the importance of studying fungal pathogens.

7.3.3 Occurrence and Mechanisms of Azole Drug Resistance in *Candida* Species

Resistance to Fluconazole is more common than other antifungals due to both the fungistatic nature of the drug, and it being among the most prescribed antifungals (Whaley et al., 2017). *Candida* isolates vary in their susceptibility to fluconazole as demonstrated by the variation in Minimal Inhibitory Concentration (MIC) (tab. 1). This depends on both intrinsic resistance, which differs between species, and acquired resistance. Relatively, *C. albicans* has high intrinsic Fluconazole susceptibility while species including *C. glabrata* and *C. krusei* are intrinsically more resistant (Lee et al., 2021) (tab. 1). In *C. glabrata*, this is attributed to elevated efflux of Fluconazole by ABC transporters, as regulated by the TF *Pdr1* (Ferrari et al., 2011; Lee et al., 2021). The mechanism of resistance is poorly understood for *C. krusei* (Lee et al., 2021). Perhaps most concerning is the extremely high (>90%) frequency of Fluconazole resistance among isolates of the emerging pathogen *C. auris*, in which multi-drug resistance is common (Lee et al., 2021; Lockhart et al., 2017; Montoya et al., 2019; Rhodes and Fisher, 2019).

Table 1. Fluconazole resistance among common Candida species. *From Espinel-Ingroff et al., 2014*

Antifungal agent	Species	No. of isolates/no. of labs	MIC (range) (µg/ml)	MIC Mode (µg/ml)
Fluconazole	*C. albicans*	5,265/9	0.06-≥128	0.12
	C. dubliniensis	162/7	0.06-64	0.25
	C. glabrata	7,538/14	0.12-≥128	4
	C. guilliermondii	373/11	0.12-64	2
	C. krusei	1,073/11	0.25-≥128	16
	C. lusitaniae	574/10	0.12-64	0.5
	C. parapsilosis	6,023/15	0.06- ≥128	0.5
	C. tropicalis	3,748/14	0.06- ≥128	0.25

In addition to intrinsic resistance, *Candida* species readily acquire Fluconazole resistance via various mechanisms (fig. 3). While different mechanisms have been observed in different *Candida* species (fig. 3), these generally involve increased drug efflux and alteration or elevated expression of drug targets (Lee et al., 2021). These genetic mechanisms are summarised in the figure but will be discussed in section 7.7.2.

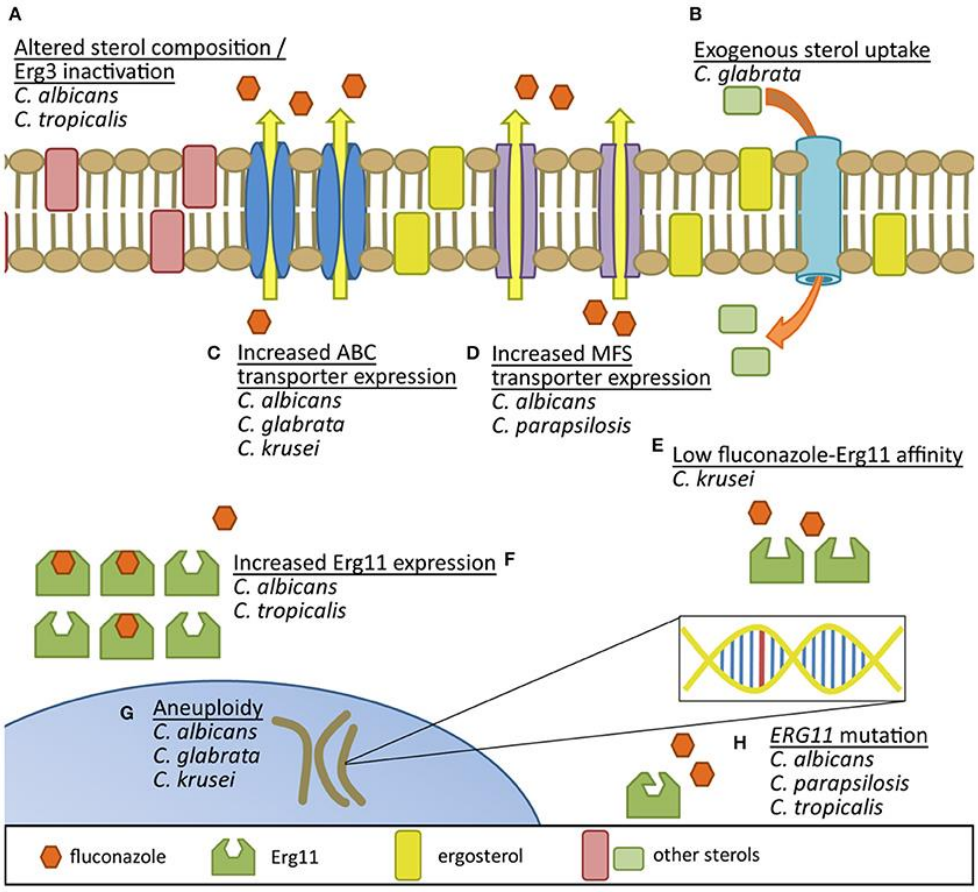

Figure 3. Mechanisms of Fluconazole resistance among common Candida species. *From Whaley et al., 2017.*

Figure illustrates the eight mechanisms by which Candida species can acquire azole resistance and lists beneath each mechanism the Candida species in which it has been confirmed to occurred. In the cartoon, the top of the figure, above the cell membrane is the external environment, bellow the membrane is the intracellular space.

7.4 Biofilms

While many studies looking at *C. albicans* and its drug resistance study the species in suspension, this is not the state in which the yeast is usually found in the environment. Instead, *C. albicans* typically grows as biofilm on abiotic surfaces which can also form on mucosal surfaces (Andes et al., 2004; Chandra et al., 2001; Gulati and Nobile, 2016; Tsui et al., 2016; Williams et al., 2013). Biofilms are structures formed by a community of cells adhering to a surface, and producing an extracellular matrix (fig. 4) (Costeton et al., 1995; Mathé and Van Dijck, 2013). Importantly, cells within a biofilm have attributes which distinguish them from planktonic cells, including greater resistance to physical and chemical stress, including antifungal drugs (Gulati and Nobile, 2016; Lee et al., 2021; Nett et al., 2007; Ramage et al., 2002). Biofilm also protect *C. albicans* from host immune defences including phagocytes (Chandra et al., 2007; Katragkou et al., 2010; Mathé and Van Dijck, 2013). This protection is largely attributed to the extracellular matrix of polysaccharides and other macromolecules which encase the biofilm, as well as the upregulation of drug efflux pumps (Gulati and Nobile, 2016; Ramage et al., 2002; Zarnowski et al., 2014).

The pleomorphic nature of *C. albicans,* whereby it switches between yeast and filamentous growth is key to its ability to form biofilm as this structure consists of different cell types (fig. 4) (Tsui et al., 2016). Briefly, a biofilm forms when yeast cells aggregate together, adhere to a surface, and proliferate to form a basal layer. From this layer, hyphae then grow and act as a scaffold as the biofilm continues to be populated by yeast, hyphae, and pseudo-hyphae. During this maturation phase, the extracellular matrix is formed. From this mature biofilm, yeast cells can disperse and colonise other surfaces (fig. 4) (Gulati and Nobile, 2016).

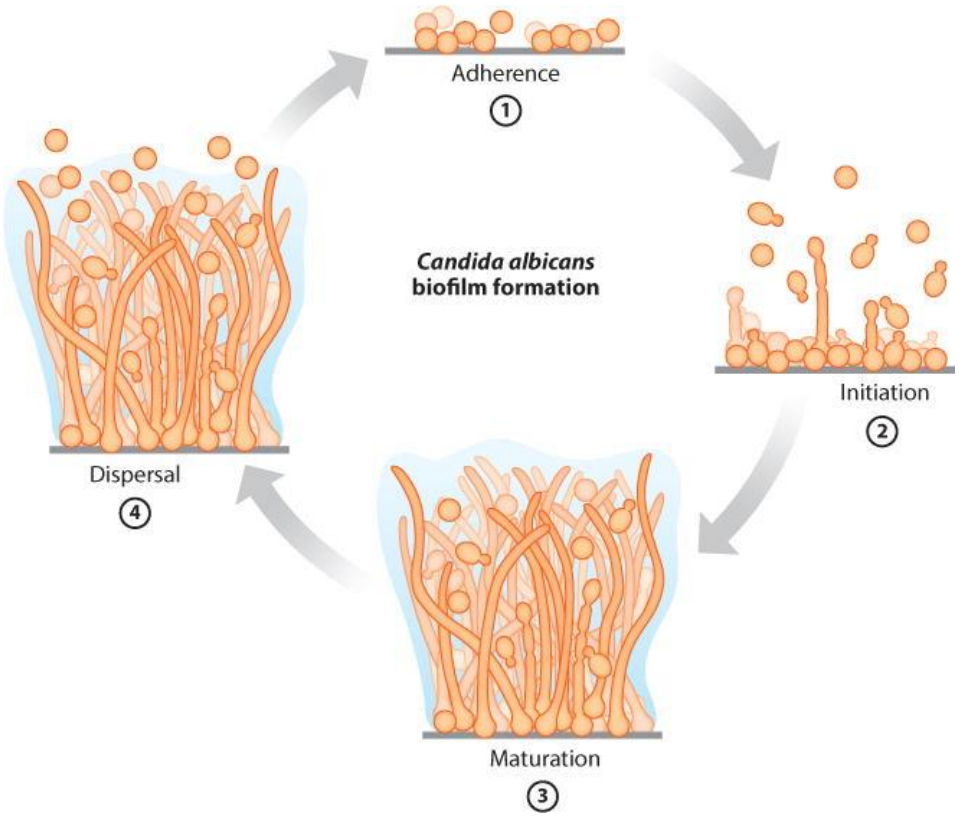

Figure 4. Four stages of biofilm formation in <u>Candida albicans</u>. *From Nobile and Johnson, 2015*

(1) Planktonic yeast cells aggregate and adhere to a surface. (2) Hyphae then start to grow from this basal layer. (3) Yeast, pseudo-hyphae, and hyphae form a mature, dense biofilm surrounded by extracellular matrix. (4) Yeast cells are released and colonize new sites (1).

Biofilms are key to *Candida* pathogenicity and a common cause of candidemia is implantation of intravenous catheters or medical devices which are colonised by *Candida* biofilms (Gulati and Nobile, 2016). Indeed, >50% of central venous catheters implanted in the USA result in biofilm infection (Fox and Nobile, 2012). The biofilms on these abiotic surfaces act as reservoirs from which *Candida* can disseminate, leading to recurrent candidemia. Further, the drug-resistant nature of biofilms results in mortality rates of ~50% (Douglas, 2003; Gulati and Nobile, 2016; Li et al., 2018; Ponde et al., 2021; Tumbarello et al., 2012) and often necessitates removal of the device (Andes et al., 2012; Kojic and Darouiche, 2004).

C. albicans biofilms do not necessarily only contain this species. Polymicrobial biofilms are clinically significant and in addition to *C. albicans,* can contain other *Candida* (Coco et al., 2008; Pathak et al., 2012), or bacterial species (Ellepola et al., 2017; Ponde et al., 2021; Valentine et al., 2019). Such interactions can be synergistic or antagonistic and can lead to increased virulence and further complications in treatment (Ponde et al., 2021). Most famous is the interaction between *C. albicans* and *Streptococcus mutans* in oral plaque*,* in which the bacterium binds directly to *Candida* yeast and hyphal cells, forming a dual-species biofilm (Ellepola et al., 2017; Hwang et al., 2015). This biofilm increases the ability of *S. mutans* to colonize the mouth, and increases oral pathology (Khoury et al., 2020; Ponde et al., 2021). This demonstrates the importance of inter-species interactions within the host microbiota and alludes to the plethora of ways in which *C. albicans* can affect human health.

7.5 *C. albicans* Reference Genome

Candida albicans is a diploid organism with eight chromosomes (2n=16) in its 15.47 Mb haploid genome (Skrzypek et al., 2017). These chromosomes are numbered according to size, with chromosome 1 being the largest at 3.19 Mb, and chromosome 7 the smallest at 950 kb (fig. 5). The exception to this is chromosome R, which carries the ribosomal DNA (rDNA) locus and is second largest.

Genome data are based on Assembly 22 of the common/reference lab strain SC5314, which was used in this study. Indeed, SC5314 is the most common *C. albicans* WT lab strain. It is a clinical isolate obtained from a patient with a generalized *Candida* infection and is virulent in mouse models. This strain is in the most common *C. albicans* clade, which includes ~40% of all *C. albicans* strains (A. Tavanti, A.D. Davidson, N.A.R.G., M.C.J. Maiden and F.C.O., *unpublished observations*; Odds, Brown and Gow, 2004).

Assembly 22 covers the entire *C. albicans* genome, including >6000 ORFs, of which, less than a third have been verified (Skrzypek et al., 2017). The assembly was constructed predominantly using Illumina sequence data. This short-read sequencing is very accurate at the nucleotide level but is poor at sequencing repetitive elements such as the rDNA locus and telomeres. Sanger sequences from previous assemblies were used to resolve these loci, but it is recognised that long read sequencing technologies such as Oxford nanopore is necessary to completely resolve these (Skrzypek et al., 2017).

Diversity is relatively high between *C. albicans* genomes, with a SNP approximately every 330–390 bp-significantly more than in *C. dubliniensis,* closest relative to *C. albicans* (Butler et al., 2009; Jackson et al., 2009). This variation is also evident when considering individual *C. albicans* genomes, which have a high degree of heterozygosity. The extent of this heterozygosity has led to the hypothesis that *C. albicans* originated from a hybridisation of two distinct species (Mixão and Gabaldón, 2020).

C. albicans chromosomes can be separated by CHEF electrophoresis to karyotype strains. This technique separates whole chromosomes by alternating the direction of current while the DNA migrates through

the gel over several days. The banding patter of WT SC5314 is below (fig. 5) along with the reference lengths of the chromosomes corresponding to each band. Note two to four bands are often visible at the bottom of the gel and are caused by chromosome 6 and 7 homologues. The chromosome length polymorphisms of these homologues are visible as CHEF has a higher resolution for shorter chromosomes. Chromosome R typically appears as a smear due to the length polymorphism generated by the rDNA encoded thereon.

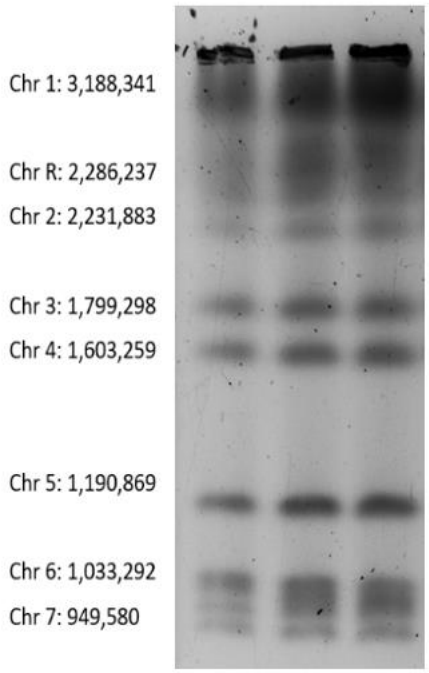

Figure 5. CHEF banding pattern of WT SC5314.

Gel shows three replicates of strain SC5314 imaged as part of this study with chromosome numbers and their length as of Assembly 22 adjacent the corresponding band. Note poor resolution of chromosomes 1 and R due to the frequency of length polymorphism at the rDNA locus.

7.6 *C. albicans* Parasexual Cycle

C. albicans usually divides mitotically and lacks traditional meiosis. Classically, it was believed that *C. albicans* was obligately diploid and asexual, but a mating type locus (*MTL*) (Hull et al., 2000) in addition to haploid (Hickman et al., 2013) and tetraploid (Legrand et al., 2004) forms have been identified. This led to the identification of a parasexual cycle in *C. albicans* and other related species.

During parasexual reproduction, two diploid *C. albicans* cells join to form a tetraploid (Chibana et al., 2000; Hull et al., 2000). Tetraploids are unstable and undergo concerted chromosome loss (CCL) during division to generate diploid or aneuploid offspring. During CCL, chromosomes are reorganised and there is evidence of significant recombination between homologous chromosomes. Despite the lack of meiosis, Spo11 is expressed. This protein is responsible for inducing DSBs during meiosis across the Eukaryotes. Deletion of the *Spo11* gene results in a significant reducing of recombination, indicating this meiotic function was conserved and that this cycle evolved from classical meiosis (Forche et al., 2008; Keeney, 2008; Wang et al., 2018).

In addition to the diploid and tetraploid forms, *C. albicans* has been observed in a haploid state (Hickman et al., 2013). This is thought to arise from further CCL and is less fit that the heterozygous diploid. Haploid cells of opposite mating types can join to make a heterozygous diploid, or haploid cells can undergo auto-diploidization to produce a homozygous diploid, likely by mitotic defects (fig. 6). The haploid state is typically unstable and transient, with significantly impaired fitness relative to the diploid (Hickman et al., 2013; Mba et al., 2022). Indeed, there is a ploidy drive, whereby lab evolved haploids or tetraploids eventually become stable diploids (Gerstein et al., 2017).

The *MTL* of *C. albicans* is found on chromosome 5 and is approximately 8.8 kb. As in *S. cerevisiae,* two alleles exist, *a* and *α*. Rates of heterozygosity at this locus are >90% for diploid *C. albicans* isolates (Odds et al., 2007; Ropars et al., 2018) which are generally fitter and more virulent that *MTL* homozygotes (Wu

et al., 2007). To become mating competent, *C. albicans* must first undergo a phenotypic change, from a white to an opaque state. White cells are round, and form domed, white colonies on solid media, while opaque cells are elongated, forming flatter, darker appearing colonies (Miller and Johnson, 2002; Slutsky et al., 1987; Soll, 1997).

White-opaque switching is initiated through environmental stress. This is regulated by homeodomain proteins at the *MTL* and results in alterations to the chromatin structure of the *MTL* (Lockhart et al., 2002; Miller and Johnson, 2002). Heterozygosity at the *MTL* represses the switch to the competent opaque phenotype and so only homozygotes can sexually reproduce (Lockhart et al., 2002; Miller and Johnson, 2002; Soll, 2007). Homozygotes are produced either by loss of heterozygosity (LOH) at the *MTL,* or auto-diploidization (Mba et al., 2022).

Opaque diploid cells signal one another using pheromones (Alby et al., 2009; Bennett et al., 2003), which initiates the formation of a conjugation tube to facilitate nuclear and cell fusion (Usher, 2019). Mating can be homothallic or heterothallic, meaning that while the two diploid cells must be *MTL* homozygous to mate, they can have either the same, or different *MTL alleles* (Alby et al., 2009; Usher, 2019) (fig. 6).

The ability to reproduce both asexually and sexually is advantageous. Sexual reproduction is favourable in heterogenous environments (Becks and Agrawal, 2010), as it generates diversity, thus facilitating adaptation (McDonald et al., 2016). It is however energetically costly and in addition to generating fitter genotypes, generates deleterious or less fit genotypes. Asexual reproduction is favourable in homogenous environments, where selection maintains beneficial allele combinations which are not disrupted by sex (Becks and Agrawal, 2010; Wang et al., 2018). *C. albicans* is believed to be primarily asexual (Nébavi et al., 2006).

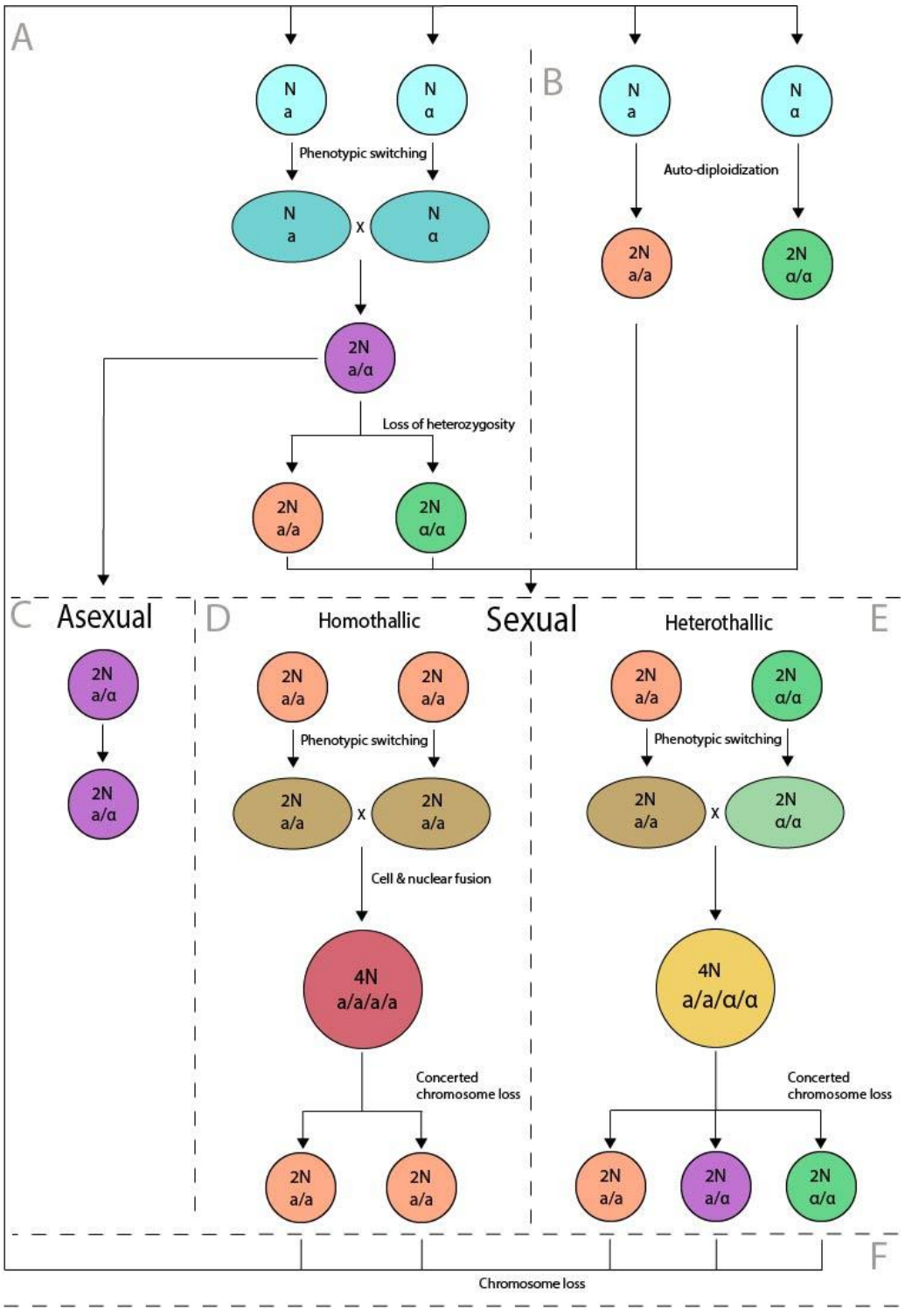

Figure 6. <u>Candida albicans</u> parasexual cycle. *Made in adobe illustrator.*

Cells are colour coded based on ploidy, white/opaque switching and MTL allele. Ploidy is indicated as multiples of the haploid genome (N), 'X' indicates a mating event. Boxes A and B indicate diploidization of haploid <u>C. albicans</u> cells, generating heterozygous cells (A) or homozygous cells by auto-diploidization (B). C shows typical asexual replication, in contrast to both types of sexual replication, homothallic (D) and heterothallic (E). Diploid cells can undergo further LOH or errors during mitosis to generate haploid cells and complete the cycle (F).

7.7 *C. albicans* Intraspecific Variation

Despite its predominantly clonal lifecycle, *C. albicans* isolates display a high degree of diversity, and the population is divided into five large and thirteen small clades based on molecular data (Odds et al., 2007; Shin et al., 2011; Sitterlé et al., 2019). This diversity allows *C. albicans* to colonise a vast range of niches, including within human (da Silva Dantas et al., 2016) or animal (Barnett, 2008) hosts or less commonly on plants (Bensasson et al., 2019) or transiently on abiotic surfaces (Gulati and Nobile, 2016). It also facilitates adaptation to stresses including drug treatment. Diversity can be considered at both the phenotypic level, and the genetic level which will both be briefly discussed.

7.7.1 Phenotypic Variation

C. albicans isolates have a range of phenotypic attributes, including divergent metabolic abilities, growth rates, morphologies, and virulence. For example, clinical isolates of *C. albicans* exhibits signification variation in virulence during murine and *Galleria melonella* infection models (Allen and Beck, 1983; Frenkel et al., 2016; Hirakawa et al., 2015; MacCallum et al., 2009; Schönherr et al., 2017). Virulence is a complex trait and differences between strains can result from various sources, including variation in the expression levels or coding sequences of virulence factors (Braunsdorf and Leibundgut-Landmann, 2018; Moyes et al., 2016; Schönherr et al., 2017), and differences in the ability to form biofilms (Li et al., 2003). *C. albicans* strains also display a high degree of variability in antifungal drug resistance as previously discussed (Espinel-Ingroff et al., 2014).

Furthermore, *C. albicans* strains can reversibly switch between a range of morphological states, including yeast, hyphae, pseudo-hyphae, and chlamydospore (fig. 7). Cells of these states differ morphologically and transcriptionally and switching occurs in response to environmental cues (Noble et al., 2017). Hyphae are generally more associated with disease, typically superficial infection due to their

ability to invade tissue (fig. 7) (Dalle et al., 2010; Moyes et al., 2010; Noble et al., 2017). Yeast, pseudo-hyphae, and hyphae are all involved in disseminating disease (Chin et al., 2014; Noble et al., 2017).

In addition to the white, 'true' yeast state, there are yeast-like states: opaque cells, grey cells, and GUT cells which are all ellipsoid in shape and have distinct properties (Noble et al., 2017). Grey cells are a third state which some strains can form and are the smallest cell type (Tao et al., 2014). GUT (gastrointestinally induced transition) cells are another distinct state which is induced in the host gastrointestinal tract (Noble et al., 2017; Pande et al., 2013).

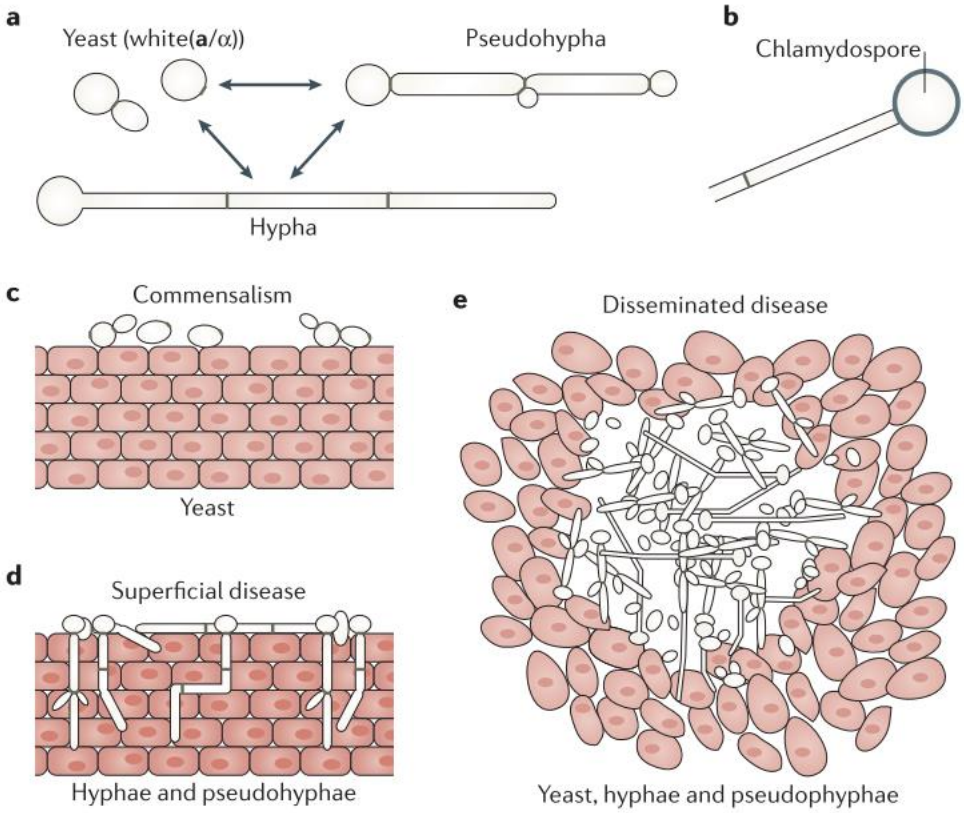

Figure 7. Morphological states of <u>C. albicans</u> and their association with disease. *From Noble, Gianetti and Witchley, 2017.*

(A) Yeast cells are round or ovoid and upon cell division, separate completely into two independent cells. Hyphae and pseudo-hyphae form multicellular structures, where daughter cells do not separate from parents. Hyphae are cylindrical and form a continuous tube whereas pseudo-hyphae are ellipsoid with indentations at cell septa. Strains can reversibly switch between these three states. (B) Chlamydospores are a rarer state and have only been seen in vitro under stringent stress conditions. They are large cells which are produced from the end of hyphae and have thick cell walls. (C) Yeast cells are the typical commensal state, growing on epithelial surfaces. (D) Hyphae and pseudo-hyphae can penetrate epithelial cells and are associated with superficial infections. (E) These three cell types are seen during disseminating disease.

7.7.2 Genomic Variation

The high degree of phenotypic diversity of *C. albicans,* and indeed its success both as a commensal organism and as a pathogen are due to its genetic diversity (Braunsdorf and Leibundgut-Landmann, 2018). As already discussed, the *C. albicans* genome has as high degree of heterozygosity (Butler et al., 2009) and genomes of different strains differ significantly from one another on both small and large scales. Small scale changes include SNPs and indels, while large scale changes include LOH, segmental deletions/duplications, translocations, inversions, and aneuploidies (Mba et al., 2022).

While *C. albicans* is primarily asexual, its parasexual cycle can generate novel genotypes, with altered virulence and drug resistance and facilitates gene flow between genetic clusters (Hirakawa et al., 2017; Ropars et al., 2018). This rare event is however insufficient to account for the extensive genomic diversity of *C. albicans.* Large scale genomic changes are due to the highly plastic nature of the *C. albicans* genome which is very tolerant of chromosomal rearrangements (Fischer et al., 2006). These events can have strong phenotypic effects as they potentially influence the gene expression levels of many genes, often by altering their copy number and are key to *Candida* evolution. Large-scale genomic rearrangements arise primarily due to repair of DSBs as well as mitotic non-disjunction and mitotic non-allelic homologous recombination (NAHR) but also occur during parasexual recombination events (Mba et al., 2022; Ropars et al., 2018; Wang et al., 2018). The DSB repair pathways in *C. albicans* will now be briefly discussed in addition to both large, and small-scale genomic changes in the context of acquiring azole resistance.

7.7.2.1 DSB repair pathways in C. albicans

DSBs are typically either repaired by non-homologous end joining (NHEJ), which typically incurs small deletions, or by homology directed repair pathways, including homologous recombination (HR). There are numerous homology-based pathways, but they all require detection of the broken ends which are

then resected, to generate a single strand which searches for a homologous sequence to act as a repair template (Marton et al., 2021). Following homology searching, the single stranded DNA invades the double stranded homolog, and DNA replication occurs. The two homologous chromosomes or chromatids then need to be separated by resolving the holiday junctions which join them. The way in which these are resolved can lead to cross-over products- where the chromosome arms originate from different chromosomes- or non-crossover products. Even in non-crossover products, there is a region where heterozygosity is lost at the break point due to the phenomenon of gene conversion (Legrand et al., 2019).

Homology based pathways are generally favourable as they prevent loss of genetic information, however at repetitive elements, sequence similarity to other repeats can result in non-allelic homologous recombination (NAHR). This can occur between repeats on the same chromosome (in *cis*) or another chromosome (in *tans*). Either pathway can result in chromosomal rearrangements (Janssen et al., 2018) (fig. 8). Failure to repair DBS results in chromosome truncations, thus generating CNVs (Legrand et al., 2019).

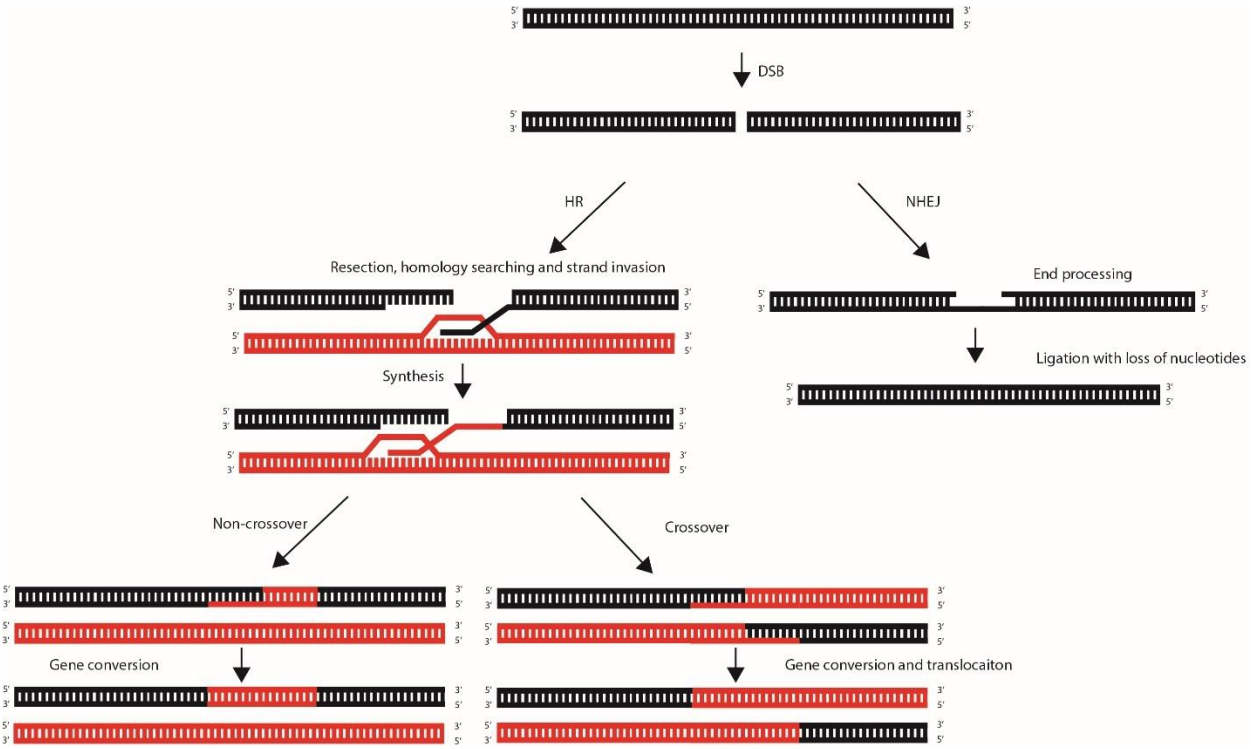

Figure 8. DSB repair pathways. *Made in Adobe Illustrator.*

DSBs are repaired by Homologous recombination mechanisms or NHEJ. NHEJ involves resection of the blunt ends which are then annealed, resulting in loss of genetic material. HR methods rely upon a repair template; the broken ends are resected and then search for this homologous sequence. Upon finding it, they invade the strand and DNA replication occurs to replace the lost nucleotides. This junction between the two chromosomes is then resolved. This can occur in various ways depending on where the strands are cut and can produce either non-crossover, or crossover products. In both cases, gene conversion occurs, but in crossover, there are also translocations of the downstream regions.

While there is evidence of both NHEJ and HR in *C. albicans*, the latter is the favoured pathway with NHEJ being reported as inefficient and rare in the organism (Legrand et al., 2019; Ng and Dean, 2017; Vyas et al., 2018). While this is consistent with *S. cerevisiae*, NHEJ is the far more frequent in other species including *C. glabrata* (Vyas et al., 2018).

7.7.2.2 Karyotype changes and aneuploidies

Aneuploidy is where the copy number of whole chromosome, or segments thereof are altered. *C. albicans* is diploid and thus has two copies of each chromosome. Mitotic non-disjunction, and concerted chromosome loss following parasexual mating, can result in changes in this copy number (Selmecki et al., 2010; Yang et al., 2021). Further karyotype changes can involve segmental duplications or deletions

which again, can generate CNVs and exert strong phenotypic effects. Reciprocal events also occur including translocations and inversions which do not results in any gain or loss of genetic material, but simply shuffle the organisation of the chromosomes. These are all typically generated by DSB repair, or non-allelic homologous recombination (NAHR). Reciprocal chromosome changes are believed to generally have less extreme phenotypes but may influence the expression of genes or affect gene function if the breakpoint is within a coding or regulatory sequence. Such changes are frequently observed in drug resistant clinical isolates (Legrand et al., 2007). The true fitness costs of reciprocal rearrangements are unclear as they have traditionally been difficult to generate experimentally without also generating SNPs (Fleiss et al., 2019). Reciprocal translocations in *S. cerevisiae* can increase the frequency of aneuploidy by generating multivalents during meiosis, leading to non-disjunction (Fischer et al., 2000). It is unclear whether similar mechanisms affect *C. albicans* during parasexual mating or mitosis.

Aneuploidies are commonly observed among drug resistant *C. albicans* strains (Whaley et al., 2017). This resistance is typically due to CNVs of genes encoding drug targets. Most famously, fluconazole treatment of *C. albicans* often results in the generation and fixation of isochromosome 5L, which contains genes *ERG11* and *TAC1* (Selmecki et al., 2008, 2006, 2009). It is however uncertain how common these events are outside of clinical settings. While studies looking at clinical *C. albicans* isolates identified aneuploidy as a common genomic feature (Hirakawa et al., 2015), recent population genetic studies suggest these events are only common following antifungal drug treatment (Ropars et al., 2018). Another recent study generated *C. albicans* strains with trisomy of each chromosome and found that each is survivable but incur fitness costs under normal conditions. However, in the presence of stress, some provided a selective advantage over the euploid WT (Yang et al., 2021). This supports the hypothesis that while aneuploidy is generated quite frequently, it is typically only maintained in populations exposed to stresses including antifungal drugs.

7.7.2.3 SNPs and Indels

SNPs and indels are small-scale changes and affect individual nucleotides or regions <1 kb, respectively. They are very common in the *C. albicans* genome, and the rate of de novo SNPs and indels increases during host infection (Ene et al., 2018). They are important as if they fall within coding sequences, they can alter the functioning of proteins. Similarly, if they fall within regulatory elements, then can influence expression of genes. These can therefore influence azole resistance in *C. albicans*, often by changing the coding sequence of the drug target *ERG11*. This includes direct mutation of *ERG11*, resulting in lower binding affinity of fluconazole to the target (Lee et al., 2021; Morio et al., 2010) (fig. 3). Alternatively, gain-of-function mutations in *UPC2,* a zinc-cluster TF regulating *ERG11,* can result in higher *ERG11* expression and azole resistance (Flowers et al., 2012; Hoot et al., 2011).

Another key mechanism in *C. albicans* is upregulation of drug efflux. This can be achieved by overexpression of ATP-binding cassette (ABC) proteins, or major facilitator (MF) superfamily efflux pumps (Lee et al., 2021; Whaley et al., 2017) (fig. 3). Hyperactivating amino acid substitutions in the active site of Tac1, a regulator of the ABC transporters *CDR1* and *CDR2* increases expression of the two efflux pumps, leading to elevated azole resistance (Coste et al., 2006; Lee et al., 2021; Revie et al., 2018; Tsao et al., 2009). Similarly, SNPs in *MRR1*, TF of the MF transporter Mdr1 can result in overexpression of the pump and increased resistance (Dunkel et al., 2008; Lee et al., 2021; Morschhäuser et al., 2007; White, 1997; Wirsching et al., 2000).

There are additional, less frequently observed mechanisms, including loss-of-function mutations in *ERG3*. Such mutations have been identified as alleviating cellular stress by preventing the synthesis of 14-α-methyl-3,6-diol, instead leading to an accumulation of non-toxic 14-α-methyl fecosterol (Kelly et al., 1997; Lee et al., 2021; Morio et al., 2012). In these strains, ergosterol deficiency is supplemented with 14α-methyl-3,6-diol (Kelly et al., 1997; Lee et al., 2021). Further mechanisms involve stress

signalling, including heat shock protein 90 (Hsp90) and protein kinase C (PKC) pathways (Lafayette et al., 2010; Lee et al., 2021).

7.7.2.4 Loss of Heterozygosity

Loss of heterozygosity (LOH) is a common phenomenon which shapes the *C. albicans* genome by distributing SNPs and is largely responsible for the high degree of sequence divergence among *C. albicans* strains (Wang et al., 2018). A study looking at two common lab strains identified that all chromosomes had tracts of homozygosity, and that total homozygosity varied significantly between the strains, 30% in WO-1 compared to 15% in SC5314 (Butler et al., 2009).

LOH occurs on different scales: short-tract, long-tract, and whole chromosome (Legrand et al., 2019). This is dependent on the cause of the LOH event. Short tracts are often generated during gene conversion at the site of DSB repair as already discussed (Legrand et al., 2019). Long tracts are typically the result of mitotic recombination or break induced replication- where one arm following a DSB is lost and repair synthesises everything downstream of the break from the template (Marton et al., 2021; Ropars et al., 2018). Whole chromosome LOH is a rarer event (Ropars et al., 2018) but can result from the parasexual cycle, or mitotic non-disjunction (Legrand et al., 2019).

LOH is functionally important as it can alter the phenotype of a diploid organism by revealing the effect of recessive alleles which are usually masked in heterozygotes (Forche et al., 2011). There are therefore some strains which are known to harbour recessive lethal alleles because LOH is never seen at particular loci (Feri et al., 2016). While LOH can be deleterious, it can also provide a selective advantage in some conditions and is important for the evolution of *Candida.* For example, co-dominance of the WT and hyperactive *TAC1* alleles means that the full azole resistance phenotype is only seen in homozygotes, which can be generated by LOH (Coste et al., 2006; Whaley et al., 2017). Similar mechanisms likely

account for the increased rate of LOH in response to stress, including antifungal drug treatment (Avramovska and Hickman, 2019; Forche et al., 2011).

7.8 Repetitive Elements

7.8.1 Regulators of Genome Stability

Repetitive elements are responsible for modulating genome stability throughout the natural world. In addition to promoting NAHR, they can destabilise genomes by generating chromosome length polymorphisms (CLPs), facilitating chromosomal rearrangements, or generating indels and copy number variants (CNVs) (Dunn and Anderson, 2019). These events are often caused by repetitive elements forming non B-form DNA, which can result in stalling of replication forks (Dunn and Anderson, 2019; Madireddy and Gerhardt, 2017). Stalling can lead to fork slippage which results in expansion/reduction of repeats or can lead to the formation of DSBs. As already discussed, repair of DSBs can generate chromosomal rearrangements. Given the lack of traditional meiosis in *C. albicans,* repetitive elements, in addition to mitotic recombination, are likely key to generating and maintaining genome diversity (Buscaino, 2019).

7.8.2 Epigenetic Regulation of Repeat Locus Heterochromatin in *C. albicans*

To offset the genomic instability which repeat loci incur, eukaryotic repetitive elements are usually condensed into heterochromatin, as regulated by histone deacetylation, DNA methylation and RNAi pathways (Buscaino, 2019; Janssen et al., 2018; Peng and Karpen, 2006). This makes DNA less accessible and reduces the risk of DNA damage and recombination. Epigenetic control of the chromatin structure at repetitive elements is therefore a way of regulating genome plasticity. While highly unstable genomes are generally deleterious, some instability facilitates rapid evolution to changing environments (Janssen et al., 2018). This regulation is therefore a dynamic process; levels of genome stability change depending on the environment of the organism (Buscaino, 2019).

A key regulator of chromatin structure in *C. albicans* is Sir2. This protein deacetylates H3K16 (lysine 16 of histone 3) within nucleosomes, allowing them to become more compact (Freire-Benéitez et al., 2016a) (fig. 9). These is also some evidence of 5-Methylcytosine (5mC) in some *C. albicans* repetitive elements which is typical of heterochromatin, however these findings are controversial. The DNA methyl transferases necessary to methylate these nucleotides have not yet been identified in *C. albicans* (Buscaino, 2019; Erlendson et al., 2017). The *C. albicans* genome encodes the necessary components of an RNA interference (RNAi) pathway, which is implicated in regulation of heterochromatin in other yeasts (Drinnenberg et al., 2009; Volpe et al., 2002). It is however unclear whether this pathway is active in *C. albicans* (Buscaino, 2019; Staab et al., 2011).

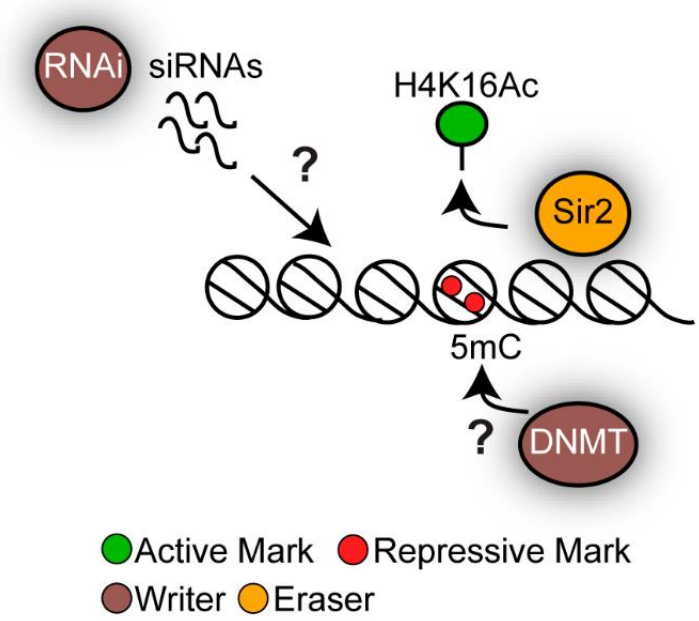

Figure 9. Epigenetic regulation of heterochromatin in C. albicans. *From Buscaino, 2019.*

Heterochromatin is maintained at repeat loci by Sir2-dependant deacetylation of lysine 16 of histone 3. This may be aided by 5-Methylcytosine formation and siRNAs as in other fungal species, although it is uncertain whether either of these mechanisms is active in C. albicans.

7.8.3 Repetitive Elements in *C. albicans*

The role of repetitive elements in regulating genome stability is understood to be a fundamental evolutionary mechanism in pathogenic fungi, including *C. albicans* (Buscaino, 2019; Galhardo et al., 2008; Selmecki et al., 2010). Indeed, serine-129 phosphorylation of histone H2A (γH2A) (a DNA damage indicator) is enriched at *C. albicans* subtelomeric repeats, supporting the understanding that these are unstable sites (Price et al., 2019). Furthermore, instability at repeat loci increases in response to host-specific stresses including elevated temperatures and antifungal drugs (Buscaino, 2019; Forche et al., 2011; Freire-Benéitez et al., 2016b). This instability generates karyotypic diversity, producing novel phenotypes which may be selectively advantageous (Buscaino, 2019; Selmecki et al., 2010). This mechanism of generating diversity is particularly important for non-meiotic organisms such as *C. albicans.*

7.8.3.1 Classes of Repetitive Element in C. albicans

C. albicans has four main types of repetitive element: telomeres, subtelomeres, rDNA and the Major Repeat Sequence (MRS) (fig. 10) (Dunn and Anderson, 2019). The MRS is the focus of this study and will be discussed in detail after the other three repeat regions are briefly introduced.

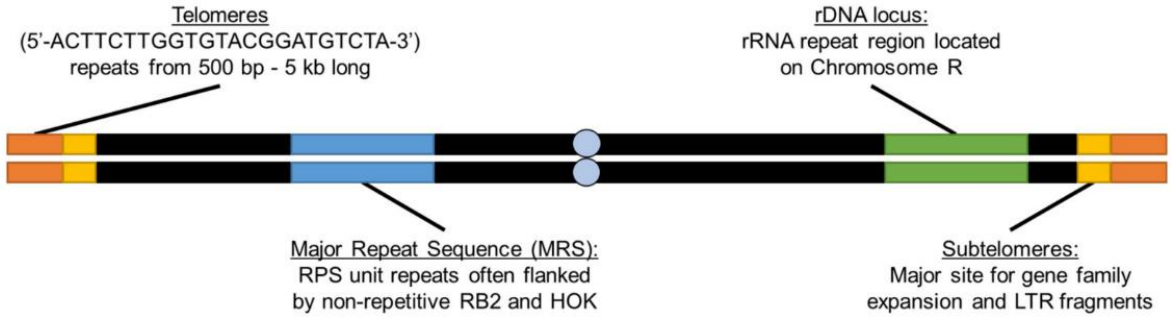

Figure 10. Major repetitive elements of the C. albicans genome. *From Dunn and Anderson., 2019.*

Figure shows the four major classes of repetitive element in the C. albicans genome. Telomeres cap the ends of chromosomes and protect them from decay. They are formed from 23 bp repeats. Adjacent these are the subtelomeres which contain expanded gene families, LTRs and transposons. The rDNA locus on ChrR is a long repetitive array encoding the ribosomal RNAs. The MRS is an element unique to Candida and occurs eight times in the haploid genome. It consists of a central repeat array flanked by non-repetitive elements.

7.8.3.1.1 Telomeres

Telomeres protect the ends of chromosomes from fusion or degradation and are typically G-rich tandem repeats of short sequences. The telomeres of *C. albicans* are 0.5-5 kb in length, made up of a 23 bp repeat: ACTTCTTGGTGTACGGATGTCTA. This repeated sequence is unusually long and is highly divergent from that of closely related *Candida* species (Dunn and Anderson, 2019; Mceachernt and Hicks, 1993). Telomeres are lengthened by the telomerase complex following replication to overcome the chromosome end replication problem, and by recombination (Ciudad et al., 2004; Dunn and Anderson, 2019; Lundblad and Blackburn, 1993).

Misregulation of telomere length, structure and recombination can significantly affect the length of chromosomes, generate excess T-circles, and increases the risk of chromosome fusion (Tomaska et al., 2000; Yu et al., 2010). T-circles are extra-chromosomal loops of telomeric DNA produced by recombination at the telomeres (Tomaska et al., 2000).

Telomeres are typically condensed into heterochromatin, as regulated by Sir2 deacetylation (Freire-Benéitez et al., 2016a). This chromatin structure is not fixed, and changes in response to environmental stress. For example, at higher temperatures (39°C), there is an upregulation of heterochromatin at the telomeres, incurring more gene silencing in the neighboring sub telomere. This may be an adaptive mechanism to regulate gene expression in subtelomeres, or to regulate genome plasticity (Freire-Benéitez et al., 2016b).

7.8.3.1.2 Subtelomeres

Subtelomeres are regions immediately adjacent the telomeres and are relatively gene poor. These loci are rich in repetitive elements, including transposons and are common sites of gene family expansion and remnants thereof (Dunn and Anderson, 2019). The subtelomeric sequences remain poorly characterized due to their repetitive nature (Dunn and Anderson, 2019; Freire-Benéitez et al., 2016a;

Van het Hoog et al., 2007). The 55 known subtelomeric genes are involved primarily in biofilm formation, filamentous growth, and metabolism. Many of these genes are associated with pathogenicity of *C. albicans* (Dunn and Anderson, 2019).

C. albicans subtelomeres are usually organized into Sir2 dependent heterochromatin, which extends from the telomeres. Regions of euchromatin are often found at sub telomeric genes, including the *TLO* genes (Freire-Benéitez et al., 2016b, 2016a; Froyd et al., 2013). The *TLO* genes are highly related paralogues which have recently expanded, existing in 14 copies in the *C. albicans* genome compared to twice in the genome of *C. dubliniensis.* These have sequence similarity the *Med2* subunit of the mediator complex and are believed to regulate virulence factors in *C. albicans* (Dunn and Anderson, 2019; Freire-Benéitez et al., 2016a; Haran et al., 2014; Liu et al., 2016; Zhang et al., 2012).

Subtelomeres frequently recombine, which is often associated with LOH. Ectopic recombination events at the sub telomeres can change the copy number of pathogenicity associated genes, including *TLO*s, and recombination events between *TLOs* can generate new sequences. This leads to novel phenotypes with altered pathogenicity and may be important for the rapid evolution of *C. albicans* during systemic infection (Anderson et al., 2015; Dunn and Anderson, 2019). Furthermore, the subtelomeres have a unique *TRE* element which promotes recombination of *TLO* genes. This is usually repressed by Sir2, but this repression is masked under stress conditions (Freire-Benéitez et al., 2016a, 2016b).

7.8.3.1.3 rDNA locus

Ribosomal DNA (rDNA) is localized to the *RDN1* locus of chromosome R in *C. albicans*. This locus is a repeat array of 50-200 repeats, each approximately 12 kb (fig. 11) (Dunn and Anderson, 2019; Freire-Benéitez et al., 2016a; Rustchenko et al., 1993). rDNA encodes a large rRNA precursor, 35S, which is processed into 18S, 5.8S, and 25S rRNAs (Hamperl et al., 2013). In the precursor, these three regions are separated by ITS (Internal transcribed spacer) 1 and 2 which are often sequenced to confirm species

identity. This region is transcribed by RNA pol I. Another rDNA region, encoding 5S rRNA, is transcribed by RNA pol III. 35S and 5S regions, as well as individual repeats are separated by Non-Transcribed Spacers (NTS) (fig. 11).

NTS regions are condensed into transcriptionally silenced Sir2-dependant heterochromatin, marked by histone deacetylation and DNA methylation (Freire-Benéitez et al., 2016a). Recombination events at the rDNA locus are very frequent, and in addition to replication fork slippage, account for the high degree of ChrR length polymorphism (Dunn and Anderson, 2019). As seen by the inability to resolve ChrR during CHEF electrophoresis (fig. 5), it is the most variable in length of any *C. albicans* chromosome, and 92% of length polymorphisms on this chromosome are due to the rDNA locus (Rustchenko et al., 1993). Recombination at the rDNA locus can also generate extrachromosomal rDNA repeats which have been identified in *C. albicans* (Dunn and Anderson, 2019; Rustchenko-bulgac et al., 1991).

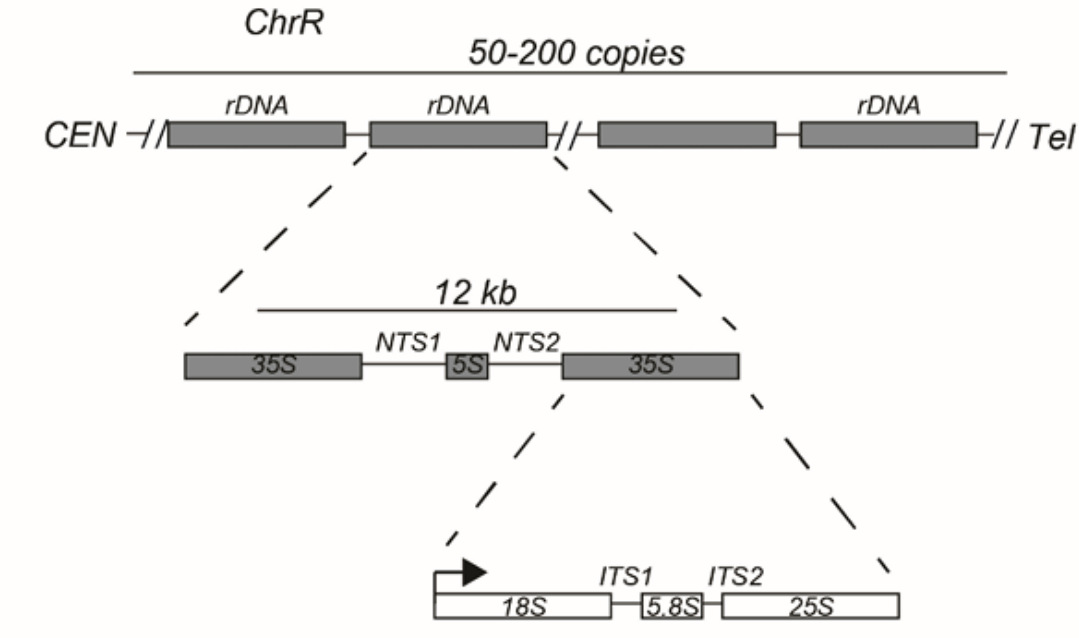

Figure 11. The rDNA locus. *From Freire-Benéitez, Price, et al., 2016.*

Figure shows the structure of the rDNA array on ChrR. The array consists of 50-200 tandemly repeated subunits separated by NTS regions. Each subunit consists of 35S and 5S regions, also separated by NTS regions. The 35S regions is a precursor to 18S, 5.8S and 25S subunits which are separated by ITS sequences.

7.9 MRS

The Major Repeat Sequence (MRS) is a repetitive element unique to the genomes of *Candida albicans,* and its closest relative, *C. dubliniensis* (Jackson et al., 2009). The role of the MRS is unknown, although its conservation throughout the *C. albicans* genome suggests some selective advantage. Most likely, the MRS has a role in generating karyotypic diversity and influencing gene expression.

7.9.1 Structure of the MRS

The MRS consists of three subunits: a central repeat array, the RPS which is flanked by two non-tandemly repetitive elements, the RB2 and Hok (Chibana and Magee, 2009; Chindamporn et al., 1998; Iwaguchi et al., 1992) (fig. 12). The MRS is transcriptionally accessible, with the heterochromatic marker: H3K4 hypomethylation in addition to euchromatic markers: H3K9 and H4K16 acetylation (Buscaino, 2019; Freire-Benéitez et al., 2016b).

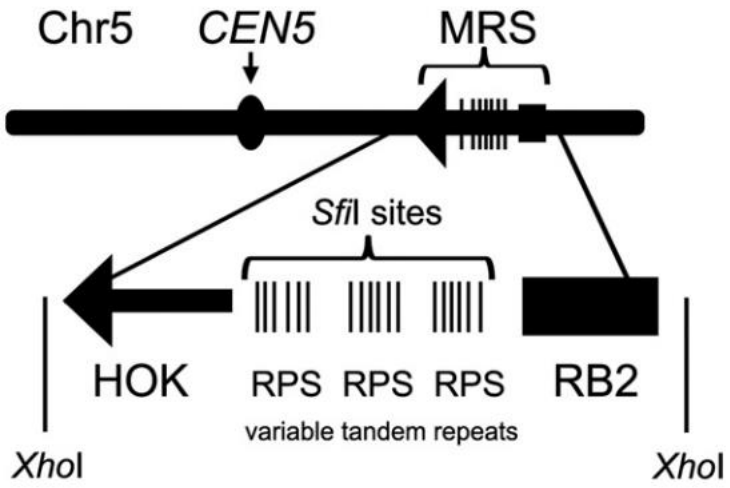

Figure 12. MRS structure on Chr5*. From Selmecki, Forche and Berman, 2010*

*The MRS consists of a central array of RPS subunits which are also intrinsically repetitive. Flanking this array are the non-repetitive Hok and RB2 subunits. Xho*I *restriction sites flank the MRS and have classically been used in conjunction with the Sfi*I *restriction sites in the RPS repeats to study MRS structure.*

7.9.1.1 RPS

The RPS is an approximately 2 kb sequence tandemly repeated. Each repeat is in turn made of 172 bp sequences known as 'alt' sequences which are tandemly repeated (Chibana et al., 1994). These alt sequences each contain 29 bp elements known as COM29. COM29 contains an *Sfil* endonuclease recognition sequence which have been utilised in MRS research (fig. 12) (Chu et al., 1993; Selmecki et al., 2010).

The repetitive nature of this array results in frequent expansion or reduction, meaning that RPS elements often vary in length between homologues (Chibana and Magee, 2009; Iwaguchi et al., 1992). While the assembly 22 genome only shows a single 2 kb RPS repeat at each MRS, the arrays are typically much longer. Long read sequencing shows the RPS array of Chr1 in SC5314 is >37 kb (Price, *unpublished data*), and studies indicate the overall length of the MRS can range from 10 kb to 100 kb in different *C. albicans* strains (Chibana et al., 1994; Freire-Benéitez et al., 2016b; Iwaguchi et al., 1992; Lephart et al., 2005). A study which used T7 insertion and haploinsufficiency to identify genes responsible for filamentation noted that of the 45 strains with insertion within the RPS, 73% had hyperfilamentous morphology (Uhl et al., 2003). This indicates a role of the MRS in regulating filamentation, a phenotype considered instrumental to the pathogenicity of *C. albicans.*

7.9.1.2 Hok

Hok is a non-repetitive sequence. It is widely reported as being 8 kb (Chibana and Magee, 2009; Chindamporn et al., 1998), but as of Assembly 22, ranges from 6.2-6.8 kb, with the exception of *Hok-5* which is significantly truncated, at 2.9 kb.

7.9.1.3 RB2

RB2 is another non-repetitive sequence. It is ~6.1 kb at full length, but several chromosomes carry truncated copies. RB2 elements are the most numerous of the three MRS elements, frequently appearing independently of the other MRS subunits (Chibana and Magee, 2009).

The RB2 subunit contains an ORF, named *FGR6* (filamentous growth regulator 6) (Uhl et al., 2003). This 1539 bp ORF has no homologues in *S. cerevisiae.* The gene is believed to be involved in regulation of filamentation; T7 insertion into a single copy of *FGR6* induces hyper or hypo-filamentation (Uhl et al., 2003).

7.9.2 Distribution

The MRS accounts for approximately 1-3% of the *C. albicans* genome on average (Chibana and Magee, 2009; Paul R. Lephart et al., 2005) and is found on every chromosome, except Chr3 which only has RB2 subunits (Chindamporn et al., 1998) (fig. 13). Chr4 and 7 each have two complete MRS elements. Chr2 and 4 have partial Hok subunits, independent of the other subunits, and all chromosomes except Chr6 have at least one partial RB2 subunit, outside of an MRS (fig. 13). Complete MRS elements are always proximal to centromeres- the furthest being 0.4 Mb away on ChrR- and are orientated such that the Hok subunit is towards the centromere. The significance of this distribution, and the relationship between the MRS and centromeres is unclear (Chibana and Magee, 2009).

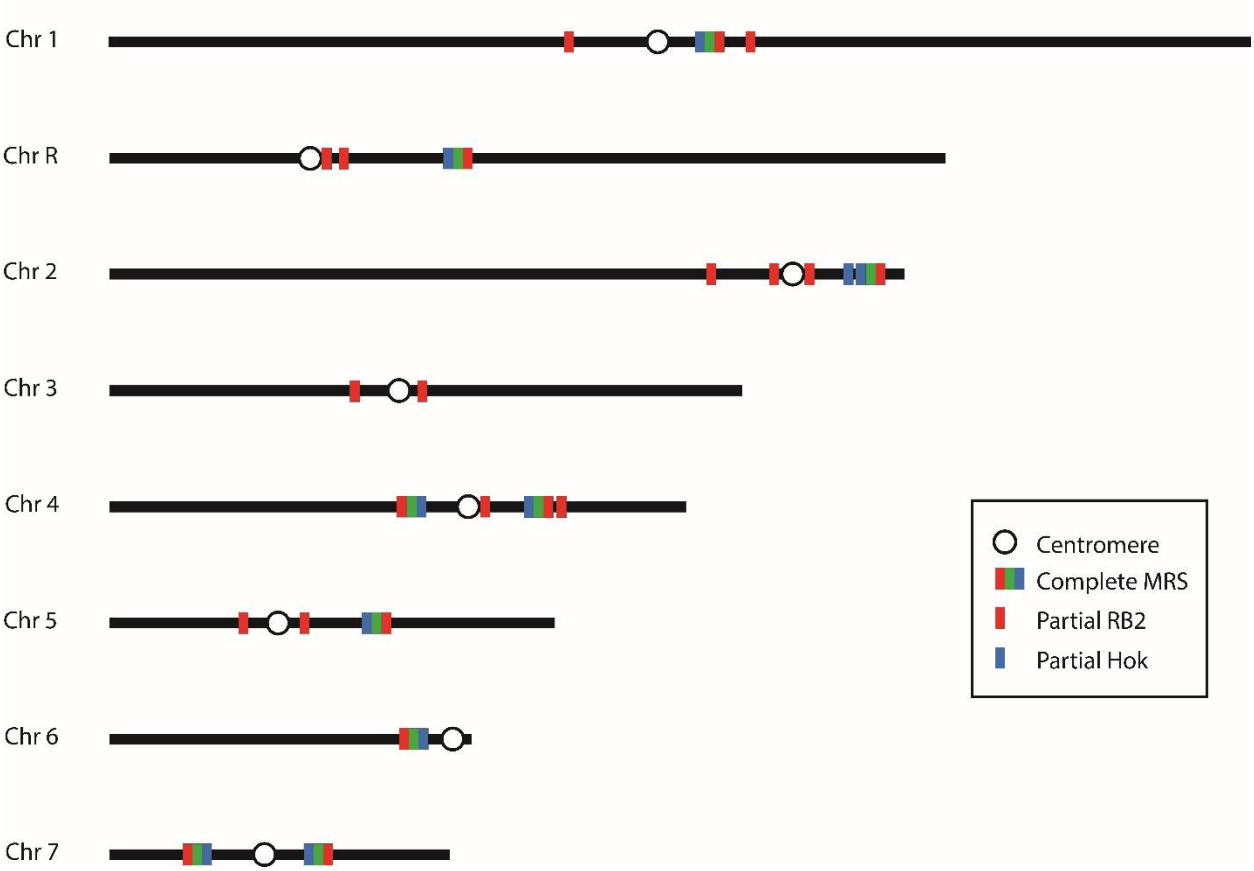

Figure 13. Distribution of MRS elements in C. albicans.

Made in Adobe illustrator based on long read MinION data from Price (unpublished data). This data differs slightly from Assembly 22 data discussed above. Centromeres, complete MRS elements and two MRS subunits: RB2 and Hok are plotted on the eight chromosomes. Note while chromosomes are to scale, these elements are not as they have been enlarged to be more visible. All MRS elements are found proximal to centromeres, and where there is a complete MRS the Hok element is closest to the centromere. A full MRS occurs on every chromosome except Chr3, which has two partial RB2 subunits. Chr4 and Chr7 have two complete MRS elements.

7.9.3 Evolutionary Significance of the MRS

The role that the MRS performs is unknown. However, numerous studies have identified that translocations are highly enriched at the MRS in both clinical isolates and lab strains of *C. albicans* (Butler et al., 2009; Chibana et al., 2000; Chu et al., 1993). Additionally, the *C. dubliniensis* genome, despite high sequence similarity to that of *C. albicans*, appears to have undergone multiple translocations at MRS elements since the divergence of the two species (fig. 14) (Magee et al., 2008). These observations suggest that the MRS has a significant evolutionary role in promoting karyotypic diversity, and that this can contribute to the generation of novel strains and species with divergent

phenotypes. Insertion of a *Ura* cassette into the RB2 subunit, but not RPS or Hok has been shown to generate translocations, further supporting the role of the MRS is generating karyotypic diversity (Iwaguchi et al., 2004).

The repetitive nature of the MRS makes it a likely target for non-allelic homologous recombination, which can generate chromosomal rearrangements, inducing inversions, deletions, and translocations. It was further hypothesised that the MRS also acts as a recombination hotspot, which would also help explain the high frequency of translocations at these loci (Pujol et al., 1999). Analysis of the recombination rate of MRS-5, compared to an adjacent genomic region has indicated no significant increase in recombination in the MRS (Lephart and Magee, 2006). It is however unknown whether this observation is applicable to other MRS elements throughout the genome (Chibana and Magee, 2009), and further investigation would be useful.

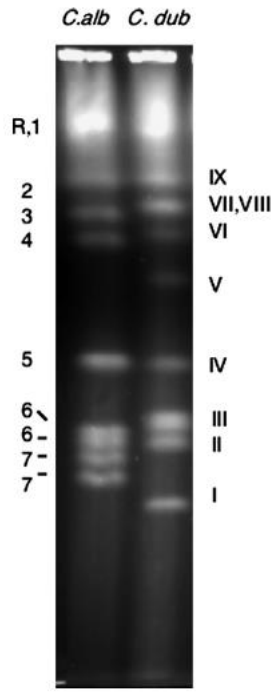

Figure 14. CHEF of <u>C. albicans</u> and <u>C. dubliniensis</u>. *From Magee et al., 2008.*

Whole chromosomes are separated by length. Despite sequence similarity, the karyotype of <u>C. dubliniensis</u> differs significantly from that of <u>C. albicans</u> resulting from multiple structural rearrangements.

Chromosome length polymorphism (CLP) is a common source of diversity in *C. albicans* (Magee and Magee, 1987; Rustchenko-bulgac et al., 1991). While the MRS is implicated in the generation of CLP by facilitating translocations, it also frequently generates them by repeat array expansion/reduction. This is evident in clinical isolates (Chibana et al., 2000; Chibana and Magee, 2009) and occurs when replication fork slippage at the RPS results in the nascent DNA strand having either more or less copies of the repeats.

This generation of MRS elements of different lengths is an important factor in generating karyotypic diversity as it can contribute to mitotic non-disjunction, leading to aneuploidies. The repetitiveness of the MRS means that it separates late in anaphase, and chromosomes with longer RPS subunits have a greater risk of non-disjunction (Lephart, Chibana and Magee, 2005). This results in chromosomes with large MRS elements being preferentially lost from the population. Indeed, deletion of a single MRS-5 copy resulted in preferred loss of the full length Chr5 (Lephart, Chibana and Magee, 2005).

While the MRS has a significant impact on shaping the *C. albicans* genome, it is unclear whether the element is essential to the survival of the yeast. It is generally reported as being non-essential, based on the viability of strains generated with a homozygous MRS-5 deletion (Lephart, Chibana and Magee, 2005), however to date, no study has reported the effect of deleting all MRS elements.

7.9.4 MRS in Non-*albicans Candida* Species

The MRS is almost unique to *C. albicans*. Whole MRS elements have only been identified in the closest relative, *C. dubliniensis* (fig. 1) which has 9 complete MRS elements (Joly et al., 2002, 1999). Despite very high genome-wide sequence similarity between the two species, the *C. dubliniensis* genome has numerous chromosomal rearrangements relative to *C albicans*. Some such arrangements appear to have occurred around the MRS, implicating it in the speciation event ~10 MYA (Chibana and Magee, 2009; Joly et al., 2002; Magee et al., 2008; Singh-Babak et al., 2021).

There is some sequence similarity between the MRS and the *C. tropicalis* genome (fig. 1) (Butler et al., 2009; Zhang et al., 1997). Again, *C. tropicalis* has a divergent karyotype from that of *C. albicans* and this may have been influenced by the MRS (Butler et al., 2009).

7.10 Experimental Generation of Chromosome Rearrangements

Chromosomal rearrangements have been identified many times in *C. albicans* isolates and lab strains, including around the MRS (Butler et al., 2009; Chibana et al., 2000; Chu et al., 1993; Legrand et al., 2004). Studies have also identified rearrangements following passage of *C. albicans* though mice (Forche et al., 2009) and during murine models of oral infection (Forche et al., 2018). To understand the effect of these genomic events however, they must be generated experimentally in a controlled setting.

Experimental generation of chromosomal rearrangements has been challenging with traditional genetic strategies which have limited our understanding (Fleiss et al., 2019). Random rearrangements have been generated with low efficiency in *S. cerevisiae* by inducing DSBs with ionizing radiation (Argueso et al., 2008), and targeted rearrangements have been generated via recombination strategies (Fasullot and Davis, 1988), but such approaches are limited in utility. More understanding has come from recent studies using Cre-loxP recombination systems in *Schizosaccharomyces pombe* (Teresa Avelar et al., 2013) and synthetic yeast chromosomes based on *S. cerevisiae* (Annaluru et al., 2014; Hochrein et al., 2018). However, these approaches typically have low efficacy, and leave markers in the genome (Fleiss et al., 2019).

CRISPR-Cas9 technology has revolutionized genetics and made genome editing far more efficient and precise. Using this system, any genomic region adjacent to a PAM (protospacer adjacent motif, typically NGG), can be targeted. A single guide RNA (sgRNA) is designed with homology to the desired cut site, which directs the Cas9 endonuclease to the locus, where it then generates a DSB. Typically, exogenous repair templates are provided to induce genome editing by homology directed repair pathways (Fleiss et

al., 2019). In addition to the specificity and efficiency with which DSBs can be generated, the system is easy to customize as only the sgRNA needs to be altered.

A recent study in *S. cerevisiae* has successfully used a CRIPSR-Cas9 system to induce chromosomal rearrangements. This study targeted long tandem repeats (LTRs) in the genome and did not provide a repair template. This meant that uncut copies of the LTS elsewhere in the genome were used as repair templates, resulting in the generation of chromosomal rearrangements and distinct phenotypes (Fleiss et al., 2019).

The efficiency of this approach makes it an attractive option for studying the evolutionary role of chromosomal rearrangements in *C. albicans.* To the author's knowledge, chromosomal rearrangements have not been generated in *C. albicans* by genome editing.

7.10.1 CRISPR-Cas9 Generation of Chromosomal Rearrangements in *C. albicans*

This study designed a sgRNA to target a conserved region of the RB2 subunit of the MRS which occurs on all eight chromosomes. This guide directs the CRISPR-Cas9 system to make DSBs within the RB2, potentially cutting multiple sites in the genome simultaneously. By not providing a repair template, we aimed to promote the use of other, uncut MRS elements as repair templates, leading to chromosomal rearrangements as has been achieved in *S. cerevisiae* (Fleiss et al., 2019).

The way in which DSBs are repaired can result in numerous translocations (fig. 15). Repair in cis- where the two ends being repaired originate from the same chromosome- can result in reconstruction of the WT chromosome. However, resolution of the Holliday junctions formed during repair can generate a mix of crossover and non-crossover products. Crossover products will have translocations between the cut chromosome and the repair template chromosome. Repair in trans- where the two halves being repaired are from different chromosomes- will always generate translocations. The non-crossover

products will leave the template chromosome intact but join the two halves from different chromosomes together. The crossover products will generate multiple translocations, affecting the cut chromosome arms in addition to the repair template chromosome (fig. 15).

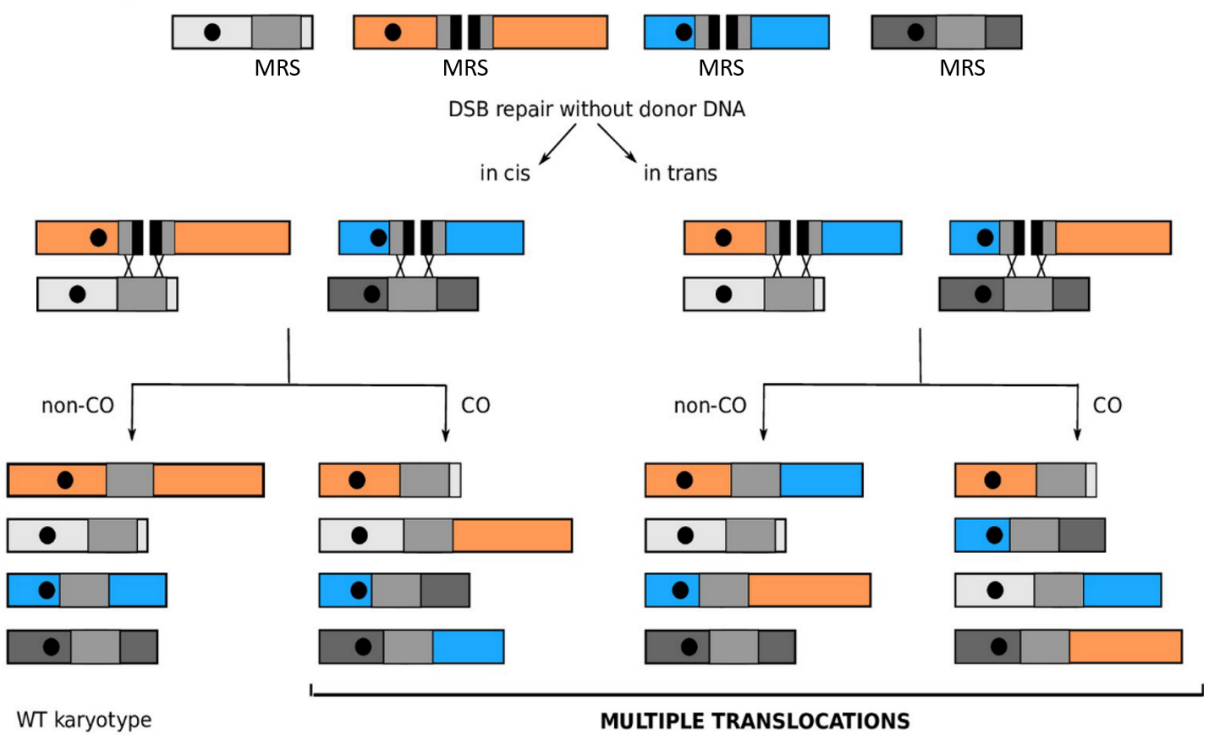

Figure 15. Generation of translocations using CRISPR-Cas9. *Adapted from Fleiss* et al.*, 2019.*

DSBs are induced at multiple MRS elements by Cas9. Repair of these breaks utilized uncut MRS elements as repair templates. Repair of chromosome arms from the same chromosome (in cis) or different chromosomes (in trans) can generate non-crossover or cross-over products. This repair can result in multiple translocations.

7.11 Hypothesis and Aims

Genome plasticity is an important evolutionary mechanism for *C. albicans* as it facilitates chromosomal rearrangements, which can lead to distinct, novel genotypes. Generating this diversity provides the basis for evolution as these changes can be selected for, or against. This enables *C. albicans* to rapidly adapt to new niches, including during systemic infection and drug treatment. It is well documented that repeat loci in the genome can facilitate chromosomal rearrangements via various mechanisms and this appears to be common in *C. albicans,* as many clinical isolates have rearrangements around such loci. The MRS is an intriguing repetitive genomic element and its conservation in *C. albicans* suggests it has some selective advantage. Many strains have been identified with translocations around the MRS and so it is likely it acts as a hotspot for chromosomal rearrangements. Such mechanisms are key to generating diversity in this largely asexual species which lacks meiosis.

This study hypothesises that the MRS acts as a hotspot for chromosomal rearrangements and that this mechanism allows *C. albicans* to rapidly adapt to stress. This study therefore aims to establish a cause-and-effect relationship between MRS-driven chromosome rearrangements and generation of novel, fitter genotypes. To this end, we used a CRISPR-Cas9 system to induce DSBs at the MRS elements. By not providing a repair template, homology directed repair will use uncut MRS elements as repair templates. This should generate chromosomal rearrangements representative of those which naturally occur at these loci. Strains with such rearrangements have then been evolved in clinically relevant stresses. In doing so, this study aims to show that:

I. A CRIPSR-Cas9 system can be used to induce chromosomal rearrangements in *C. albicans.*
II. Constantly targeting the MRS elements with Cas9 acts as a model of increased genome instability.

III. Chromosomal rearrangements around the MRS can generate distinct phenotypes which can be selectively advantageous and facilitate rapid adaptation to clinically relevant stresses.

8 Materials and Methods

All strains used in this study are derivatives of SC5314 which have been generated in this study (tab. 2). Oligonucleotides and primers were ordered from Sigma Aldrich unless otherwise specified (tab. 3). At times in this study YPD (tab. 4) was supplemented with Caffeine (Sigma-Aldrich) to a concentration of 12 mM and Nourseothricin (NAT) (Melford) to a concentration 1 µl/ml. SC media (tab. 5) was supplemented with Fluconazole (Sigma-Aldrich) to various concentrations.

Table 2. Strains used in this study.

Strain	Parental	Description	Reference
SC5314		WT *C. albicans* lab strain. Usually grows as white, round colonies.	Gillum et al., 1984
1095	SC5314	RB2 targeted with CRIPSR. Cassette still in. Pseudohyphal growth. Extra band on CHEF gel.	This study
1096	SC5314	RB2 targeted with CRIPSR. Cassette still in. WT-like growth after initial observation of hyphae on NAT plate. Two additional band on CHEF gel.	This study
986	SC5314	RB2 targeted with CRIPSR. Cassette still in. WT-like colony morphology. Extra band on CHEF gel.	This study
1097	1096	Strain 1096 with cassette removed.	This study
1098	986	Strain 1097 with cassette removed.	This study

Table 3. Primers used in this study.

ID	Name	Sequence	Source
Ab_796	ITS1	TCCGTAGGTGAACCTGCGG	Villa-Carvajal et al., 2006
Ab_797	ITS4	TCCTCCGCTTATTGATATGC	Villa-Carvajal et al., 2006
Ab_ 1179	Ca_RB2_His/FLP	CGTAAACTATTTTTAATTTGGATCGGATAAAC CGTCCCAGGTTTTAGAGCTAGAAATAGC	This paper, based on: Hernday et al., 2017
Ab_1137	AHO1096	GACGGCACGGCCACGCGTTTAAACCGCC	Hernday et al., 2017
Ab_1138	AHO1098	CAAATTAAAAATAGTTTACGCAAG	Hernday et al., 2017
Ab_1139	AHO1097	CCCGCCAGGCGCTGGGGTTTAAACACCG	Hernday et al., 2017
Ab_1140	AHO1237	AGGTGATGCTGAAGCTATTGAAG	Hernday et al., 2017
Ab_1141	AHO1236	TAAAGCTGCCACAAGAGGTATTTC	Hernday et al., 2017

Table 4. Plasmids used in this study.

ID	Name	Description	Source
pAB_264	pADH99	*C.albicans* HIS/FLP CAS9 expression plasmid	Hernday et al., 2017
pAB_266	pADH147	*C.albicans* HIS/FLP plasmid template for Fragment B for cloning-free stitching of gRNA expression cassette.	Hernday et al., 2017
pAB_240	pADH110	Half gRNA plasmid for fusion PCR gRNA cassette	Hernday et al., 2017

Table 5. Recipes for media used in this study.

Media	Volume	Recipe
SC	1 L	2.002 g Supplement mix (Formedium DSCK1000), 6.70 g Yeast nitrogen base without amino acids (Difco 291940), 20 g Agar* (Melford), 20 g Glucose (Fisher Chemical), 10 mL 32.8 mM Uridine (Sigma), 990 ml H2O.
YPD	1 L	10 g Yeast extract (Melford y1333), 20 g Bacto-peptone (BD 211677), 20 g Agar* (Melford), 20 g Glucose (Fisher Chemical), 20 ml 37 mM Adenine (Sigma), 10 ml 32.8 mM Uridine (Sigma), 970 mL H2O.

*Agar added for solid media only.

Table 6. Composition of uncommon Buffers and reagents in this study.

Buffer	Volume	Recipe
LET	1 L	200 ml 0.5 M EDTA (Fisher Chemical), 10 ml, 1 M Tris-HCL (pH 7.5), 790 ml H_2O
NDS	1 L	200 ml 0.5 M EDTA (Fisher Chemical), 50 ml 1M Tris-HCL (pH 7.5), 50 ml 10% SDS (Sigma), 700 ml H_2O
PLATE	1 L	875 ml 50% PEG 3350, 100 ml 10X TE (100 mM Tris pH 7.4, 10 mM EDTA pH8), 25 ml 1 M LiOAc, pH 7

8.1 General Lab Methods

8.1.1 Yeast Maintenance and Culturing

C. albicans stocks were made by suspending colonies in 50% glycerol in 1.8 ml cryotubes (Thermo Scientific) and stored at -80°C in an Ultra-low temperature Freezer (New Brunswick Scientific). Unless otherwise specified, when needed strains were taken from the -80°C, streaked onto YPD (tab. 5) plates and grown at 30°C.

Overnight (o/n) cultures were generally made by inoculating 5 ml YPD liquid media with cells of a single colony, collected with a pipette tip which remained in the test tube. These cultures were grown at 30°C and shaken at 180 rpm to prevent sedimentation.

The optical density at 600 nM (OD_{600}) of cells in the cultures was measured with a Kinetic BioSpectrometer (Eppendorf), following dilution in Milli-Q (Merck) water sufficient to obtain a raw OD_{600} reading <1. This OD_{600} reading was then used to approximate the concentration of *Candida* cells based on the typical estimate of an OD_{600} reading of 1 being equivalent to $3x10^7$ cells per ml.

When more accurate quantification was necessary, or where OD_{600} readings became inaccurate due to sedimentation/flocculation, haemocytometry was used to estimate cell concentration. A glass coverslip was placed over a haemocytometer, and 10 µl of o/n media, diluted 1:100 was added to the edge of the coverslip, passing underneath it and into the counting chamber. A light microscope was used at 10x magnification to view the cells in the 5x5 grid. Cells in five of these squares were counted and this was used to estimate the concentration of cells in the o/n culture.

8.1.2 PCR

PCR cycles were run in 0.2 ml Eppendorf tubes in a C1000 Thermo Cycler (BioRad). Enzymes used, buffers and cycling conditions varied depending on the reaction and are specified in the relevant section. If not specified, Taq DNA polymerase (PCRBIO) was used as per the manufacturer's instructions.

8.1.3 Gel Electrophoresis

Typical electrophoresis gels were made by dissolving agarose in 0.5x TBE to a final concentration of 1%. Ethidium bromide (Sigma) was added to a final concentration of 0.005%. Set gels were added to electrophoresis chamber filled with TBE. 5 µl sample was combined with 1 µl 6x loading dye and loaded into wells alongside 5 µl PCRBIO Ladder II: 250 bp-10 kb. Electrophoresis was run at 100 v for typically 20-25 minutes depending on DNA length, using a Fisher Scientific Power300. Gels were then imaged with a Syngene GBox Chemi XX6 Gel imaging system with 302 nm UV transilluminator which excites the DNA-bound EtBr.

8.1.4 Spotting Assays

Spotting assays are a qualitative test to compare the relative fitness of strains on solid media. Known volumes and concentrations of strains are spotted onto a plate, each strain serially diluted across the plate.

Overnight cultures were diluted to an OD_{600} of 1 and this dilution was added to the first column of a 96-well plate (Greiner). These were then serially diluted 10-fold across the five adjacent wells with a multichannel pipette. A sterilized 8x6 replica plater (Sigma-Aldrich) was inserted into the wells for 30 seconds and then pressed gently onto the solid medium. Each assay was performed in technical triplicate. Where drugs were used, a drug-free control plate was also included. Additionally in drug

assays, overnight cultures were centrifuged and washed with sterile water to prevent changes to the drug concentrations.

Plates were typically incubated for 48-72 hours depending on the media and/or conditions, after which time they were imaged with an Epson Perfection V33 scanner. The ability to form colonies at lower cell concentrations was taken as an indication of higher relative fitness.

8.1.5 Growth Curves

1 ml of overnight culture was taken for each strain, centrifuged at 3000 rpm for 3 minutes and washed twice with sterile water. OD_{600} readings were taken, and media was inoculated to an OD of 0.1. 200 µl of media inoculated with each strain was added to three wells of a flat bottom 96-well plate (Griener) to act as technical replicates. Samples were loaded onto the plate, preferentially using central wells; outermost wells were not used. Blank wells were filled with media alone and empty wells filled with sterile water to prevent evaporation of sample media. For growth assays at 37 °C, three sides of the plate were wrapped in parafilm to prevent evaporation. Plates were then put into a BMG Labtech SPECTROstar Nano and put on a cycle of two programs each with double orbital shaking at 400 rpm, recording the OD_{600} every 120 seconds. These runs typically lasted for 48 or 72 hours.

8.1.6 MIC Calculations

Wells of a Griener flat bottom 96-well plates were filed with 100 µl SC media containing serially diluted concentrations of Fluconazole (fig. 16) (tab. 5). Blank wells around the edge of the plate were filled with 200 µL SC, except two which acted as DMSO controls. 1 ml of overnight cultures were washed twice with sterile water before their concentration was estimated by haemocytometry. SC media was then inoculated to a concentration of 2000 cells/ml and 100 µl of this inoculated media was added to the remaining wells. Plates were then sealed with parafilm along three edges.

Plates were incubated at 37°C, shaken at 750 rpm for up to 72 hours in Grant-bio Thermo-shaker PHMP. At hour 0 of incubation, an OD_{600} reading was taken of each well using a BMG Labtech FLUOstar Omega. Subsequent readings were then taken every 24 hours. Data was analyzed with Growthcurver in R Studio (http://www.r-project.org/).

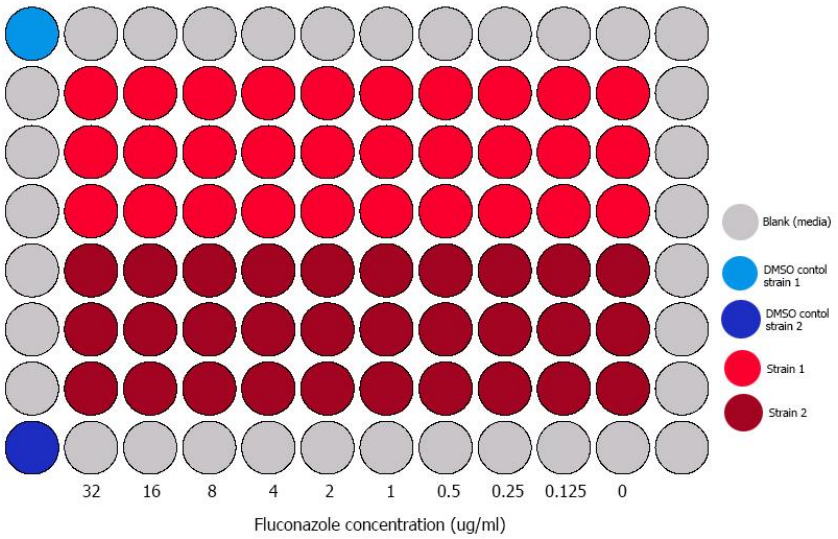

Figure 16. MIC plate layout. *Made in Adobe Photoshop.*

All outermost wells except A1 and H1 (top and bottom left) are filed with blank SC medium. A1 and H1 are filled with SC + DMSO. The volume of DMSO is equivalent to that in the highest fluconazole concentration wells. Central wells are filled with SC + Fluconazole. The concentration of fluconazole is 32ug/ml in column 2 and halves until column 10, column 11 has no fluconazole. Medium inoculated with 2 different strains were added to the wells as seen. The three rows for each strain act as technical replicates.

8.1.7 Microscopy and Quantification of Filamentous Morphology

Strains S5314, 1095, 1096 and 986 were grown overnight in YPD at 30 and 37°C. These cultures were then diluted 20x and 3 µl of diluted cultures were added to microscope slides and covered with a cover slip. Images were taken using an Olympus IX2-ILL100 microscope with an 60/1.35 oil objective and Andor Xyla 4.1 cMOS camera with CoolLED illumination. Images were captured using Metamorph. At least 8 images were taken for each strain/condition with a total of at least 100 cells between the images. Images were also taken with an 100/1.40 oil objective for better visualization of morphology, but these were not quantified. The frequency of yeast and hyphae cells was then quantified in ImageJ for each

image, excluding daughter cells which were not approximately the same size as the parent. Averages of these percentages were taken for each strain/stress in R, where Chi-squared tests were then performed (http://www.r-project.org/). Where cell images are presented in this thesis, they have been cropped to highlight specific morphologies and as such not to scale. Size comparisons therefore cannot be made between images. Original data is available upon request.

For visualization of colony morphology, the same overnight cultures in YPD were diluted and 100 cells of each strain were plated onto YPD. After Three days these colonies were imaged using a Leica S9I scope with integrated 10 MP CMOS-camera. Colony morphology can also be seen in whole-plate images taken at various stages throughout the study with an Epson Perfection V33 scanner.

8.2 CRISPR Protocol

This study adapted a CRIPSR-Cas9 protocol from Nguyen, Quail and Hernday, 2017. This protocol is designed for genome editing in *C. albicans*, is highly efficient, does not use molecular cloning and integrates into the *HIS1* locus. Adaptation of this system to target different sequences is cloning-free. A 60 bp oligonucleotide with two 20 bp non-variable ends complementary to the pSRN52 promoter and a central 20 bp target-specific sequence is designed and ordered. The non-variable 20 bp ends are then used in two sequential PCR reactions to assemble half of the cassette (fig. 17). In the first reaction, this custom sgRNA is used as a primer to amplify a fragment of the cassette containing a flip recombinase target, and a *HIS1* homology region from pADH147. This amplicon is then stitched to another amplicon which is amplified from pADH110 which contains a promoter and half of a Nourseothricin N-acetyl transferase (NAT) resistance gene (fig. 17).

This gRNA cassette is then co-transformation with a Cas9 cassette which both integrate into the *HIS1* locus. Each cassette contains half of a *NAT* gene, which come together to form a complete, functional form of the gene upon correct co-integration (fig. 17). Integration into the genome ensures continued expression of the Cas9 and sgRNA, so the system will continue to act until all available targets have been destroyed, or the cassette is removed by the maltose induced FLP recombinase (fig. 17-18). Following removal of the cassette, the genome is left with only one FLP recombinase target (FRT) integrated in the *HIS1* locus.

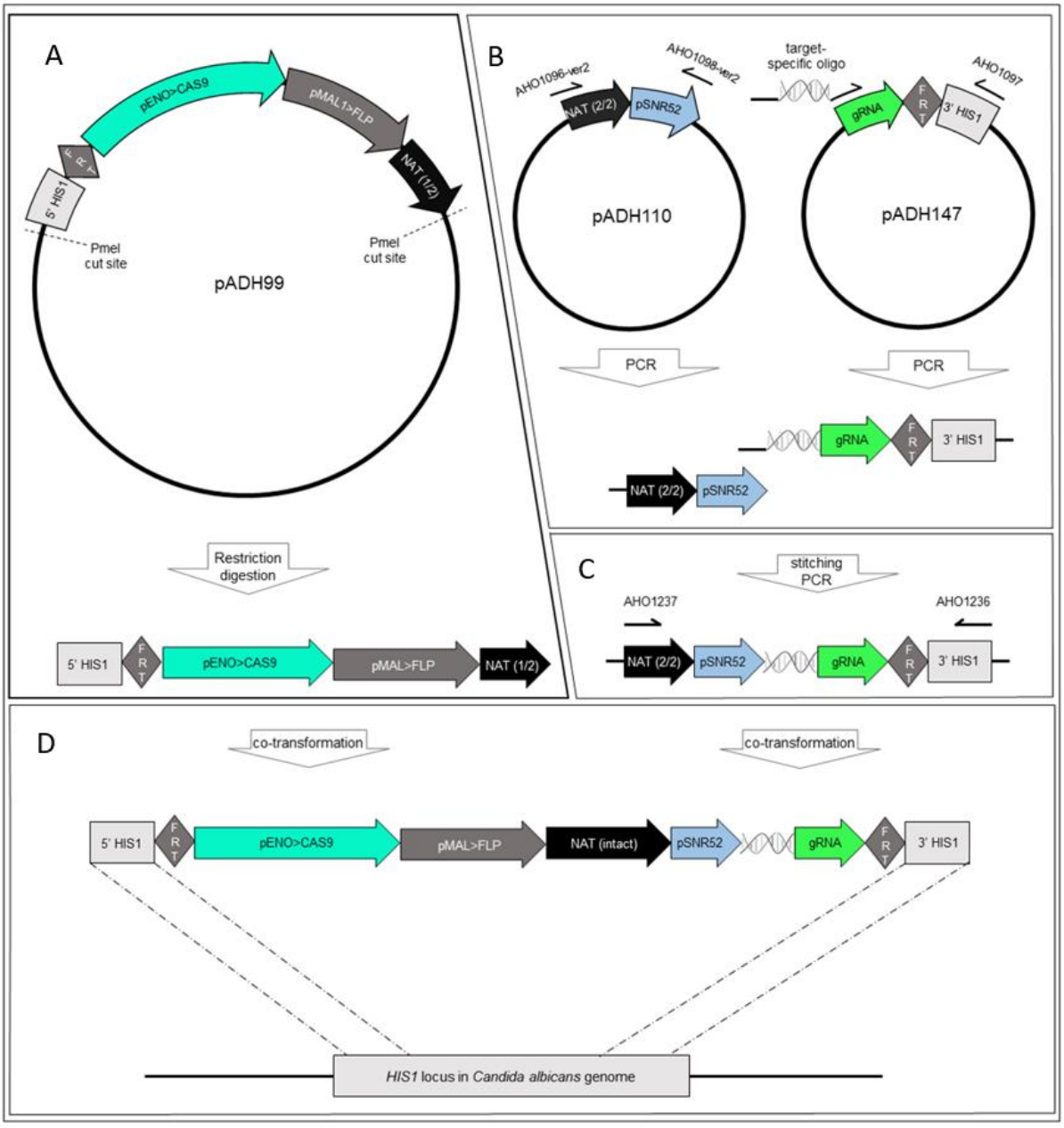

Figure 17. CRSPR cassette assembly and insertion. *From Nguyen, Quail and Hernday, 2017.*

Plasmid pADH99 is digested to release the Cas9 cassette (A). The promoter cassette is PCR amplified out of plasmid pADH110 to make fragment A and the gRNA cassette is PCR amplified out of plasmid pADH147 with a primer containing the variable 20 bp gRNA sequence to make fragment B (B). Fragments A and B are PCR stitched together to make fragment C (C). The Cas9 and gRNA cassettes are co-transformed into the C albicans HIS1 locus (D). Upon successful cointegration, the two halves of the NAT gene come together to make a complete, functional gene allowing selection with Nourseothricin.

8.2.1 PCR

8.2.1.1 Fragment A

Fragment A, containing half of the NAT gene and the pSNR52 promoter was first PCR amplified out of plasmid pADH110. The reaction mixture contained 20 µl 5x Phusion HiFi buffer (PCR BioSystems), 1 µl 3 ng/µl pADH110, 5ul 10uM AHO1096, 5 µl 10 µM AHO1098, 0.5 µl Phusion polymerase (PCR BioSystems) and 2 µl of a dNTP mix, with each dNTP at 10 mM. This mix was made up to 100 µl with Milli-Q water (Merck). The PCR cycle then began with a 30 second initial denaturation at 98°C, before 30 cycles of 20 second 98°C denaturation, 20 second 58°C annealing and 30 second 72°C extension in a C1000 Thermo Cycler (BioRad).

8.2.1.2 Fragment B

Fragment B was then amplified out of plasmid pADH147 with one generic primer as well as the target-specific oligo. The reaction mixture comprised 4 µl 5x Phusion GC buffer (PCR BioSystems), 1 µl 10 µM AHO1097, 1 µl 10 µM specific gRNA, 0.6 µl 1ng/µl pADH147, 0.2 µl Phusion polymerase (PCR BioSystems) and 0.4 µl of the above dNTP mix. This was made up to 20 µl with water. The PCR cycle then began with a 30 second 98°C initial denaturation, followed by 10 cycles of 20 second 98°C denaturation, 20 second 65°C annealing (reducing by 1°C each cycle) and 30 second 72°C extension. After the 10th cycle, when the annealing temperature had dropped to 55°C, the conditions remained constant for another 15 cycles.

8.2.1.3 Fragment C

Gel electrophoresis was then preformed to check the A and B amplicons, expected to be 1066 and 704 bp, respectively. Thereafter, the fragments were PCR stitched together to make fragment C. 10 µl 5x PCRBIO reaction buffer (PCR BioSystems), 1 µl of each fragments A and B, 0.5 µl PCRBIO HIFI polymerase (PCR BioSystems) were added to 35.5 µl water. Stitching PCR was then done with 1-minute initial

denaturation at 95°C followed by 5 cycles of 15 second denaturation at 95°C, 15 second annealing at 58°C and 1 minute extending at 72°C. Thereafter, 2 µl of the each 10 µM AHO1236 and 10 µM AHO1237 was added to the reaction mixture which cycled with an initial denaturation of 1 minute at 95°C, followed by 30 cycles of 15 second 95°C denaturation, 15 second 66°C annealing, 1 minute 72°C extension, with a final extension time of 2 minutes. Fragment C was then checked by gel electrophoresis, with an expected size of 1730 bp.

8.2.2 Plasmid Digest

pADH99 was digested in the following mixture, 2 µg pADH99 (tab. 4), 1 µl 10X Fast Dig buffer (Thermo Scientific), 0.5 µl MssI (Thermo Scientific) and 10ul water. This was heated to 37°C for at least 30 minutes.

8.2.3 *C. albicans* Transformation

WT SC5314 was gown overnight in standard conditions. Overnights were then diluted in fresh YPD to achieve an OD_{600} of around 0.2. Cultures were grown in conical flasks at 30°C with agitation for 4-6 hours until an OD_{600} of 0.5-0.8 was achieved. Once the desired OD_{600} was achieved, cultures were transferred to 50 ml falcons (Sarstedt) and centrifuged at 4000 rpm for 5 minutes at 4°C in a 5810 R Centrifuge (Eppendorf). The pellet was washed twice with 1ml sterile water before being resuspended in water $1/100^{th}$ of the volume of the culture.

50 µl of this cell suspension was then added to 2 µg of the digested pADH99, 50 µl of fragment C, 10 µl 10 mg/ml denatured salmon sperm DNA and 1 ml PLATE (tab. 6). This mixture was then incubated statically at 30°C overnight.

The following day the mixture was heat shocked at 44°C for 15 minutes, and the cells pelleted for 2 minutes at 5000 rpm and washed twice with 1 ml YPD. 1 ml YPD was added to a 15 ml falcon and

inoculated with this 1 ml cell suspension. This culture was then incubated at 30°C, agitated for 5 hours. The culture was then centrifuged for 2 minutes at 5000 rpm, the supernatant poured away, and the pellet resuspended in the residual supernatant. This suspension was then plated onto YPD plates supplemented with 1 µl /ml nourseothricin (YPD+NAT). Plates were incubated at 30°C for 3 days.

8.2.4 Screening

The transformation was performed on two separate occasions, each time with four replicates. On the first occasion, few colonies grew and so all those large enough to re-streak were screened. On the second occasion there were many more colonies and seven were selected from each plate, representing a range of colony sizes and morphologies. Colonies were then taken and re-streaked onto YPD and YPD+NAT (tab. 5) plates to confirm nourseothricin resistance. Screening was then performed by CHEF electrophoresis. Colonies screened by CHEF were first stocked in glycerol.

8.2.5 Cassette Removal

Overnight cultures were made in 5 ml YP maltose. The following day cultures were diluted and approximately 1000 cells were plated onto YP maltose and incubated at 30°C for 3 days. Colonies were again re-streaked on YPD and YPD+NAT to confirm removal of the cassette, with successful colonies having lost nourseothricin resistance. These colonies were then stocked in glycerol.

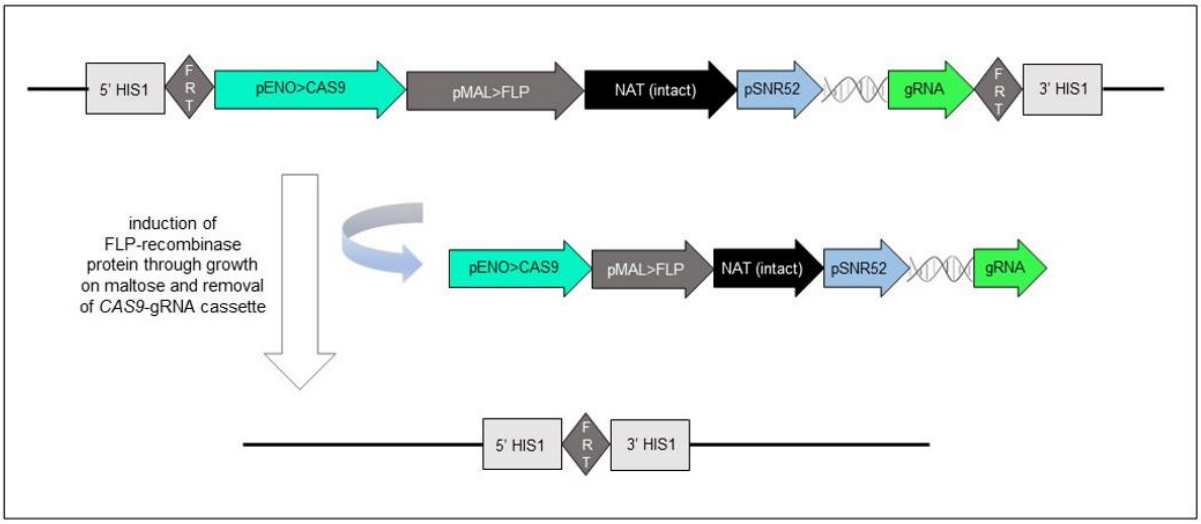

Figure 18. Removal of CRIPSR cassette. *From Nguyen, Quail and Hernday, 2017.*

Strains with the CRISPR cassette integrated can be plated onto YP maltose to induce removal of the cassette. This medium activates maltose induced FLP recombinase which acts upon the FLP recombinase target (FRT) sequences flanking the cassette. This removes the cassette, leaving only one FRT remaining in the HIS1 locus. Removal of the cassette is confirmed by the loss of nourseothricin resistance.

8.3 CHEF Electrophoresis

Clamped homogenous electrical field (CHEF) electrophoresis is a type of pulsed-field gel electrophoresis where electric fields generated around the gel change direction to allow the separation of DNA strands between 10 kb and 9 Mb in length (fig. 19). This technique allows the accurate separation of *C. albicans* chromosomes which are too large to be separated by normal gel electrophoresis (Magee *et al.*, 1988). The voltage, run time, switch time and in some cases, angle can be adjusted for the separation of different sized DNA molecules. Switch times are the length of time the electric field is pulsed in either direction and are often ramped, gradually increasing from an initial to a final time over the run.

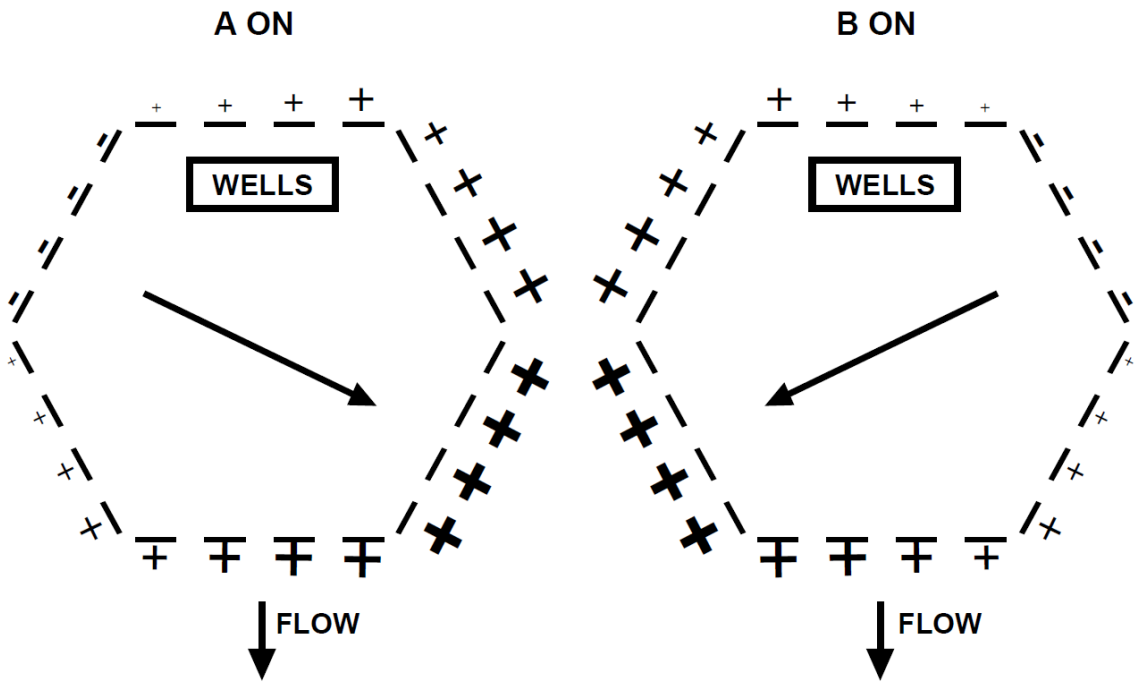

Figure 19. CHEF electrophoresis electrode potential and sample migration. *From Bio Rad, 1992.*

Figure shows the potential of electrodes when Net vector angle of +60° (A) and -60° (B) is applied. The 24 electrodes are arranged in a hexagon in the electrophoresis chamber. The size of the +/- signs indicate the relative potentials of electrode pairs. The highest potential switches between the bottom right, and bottom left segment and decreases along the neighboring segments. These two net vectors produce an overall field angle of 120° which drives the migration of the DNA samples from the wells, down the gel.

Strains were grown overnight in 5 ml YPD at 30°C, shaken at 180 rpm. Low melting agarose (LMA) was added to 100 mM EDTA to a concentration of 1%, heated until dissolved then kept at 50°C. OD_{600} of the

overnight cultures was measured and the volume required to make a 1 ml dilution of OD_{600} 7 was added to 1.5 ml Eppendorf tubes. Cultures were then centrifuged at 3000 rpm for 3 minutes to pellet the cells which were subsequently washed with 50 mM EDTA. 20 µl 10 mg/ml zymolyase (Amsbio) was then added to degrade the *Candida* cell wall and 300 µl LMA was quickly added, before pipetting the mix into the plug mold.

Once set, plugs were transferred to 5 ml tubes containing 3 ml LET (tab. 6). 25 µl β-mercaptoethanol (Sigmg-Aldrich) was then added to each of the tubes which were then incubated overnight statically at 37°C.

LET was then removed from the tubes and the plugs were washed twice with 50 mM EDTA (Fisher Chemical) for 30 minutes. Proteinase K (Melford) was then added to NDS (tab. 6) to a final concentration of 0.2 mg/ml. 3 ml of this solution was added to each tube. Tubes were then incubated at 50°C for two nights after which, the Proteinase K solution as removed. The plugs were then washed with, and resuspended in EDTA and stored at 4°C until needed.

The gel was made from 1.5 g megabase agarose (BioRad) dissolved in 150 ml 0.5x TBE and allowed to set before plugs were cut and inserted into the wells. Wells were then sealed with remaining megabase agarose solution. The electrophoresis chamber of the CHEF-DR® II Pulsed Field Electrophoresis System (BioRad) was filled with 2 L 0.5x TBE and cooled to 14°C. The gel was inserted, and the machine programmed as below.

- Block 1: 60-120 (initial-final switch time in seconds), 6 V/cm, 12 hours.
- Block 2: 120-300 (initial-final switch time in seconds), 5 V/cm, 26 hours.

Upon completion, the gel was removed and stained with 20 µl ethidium bromide (EtBr) in 400 ml 0.5x TBE for 45 minutes and de-stained with 400 ml water for 20 minutes. It was then imaged with a Syngene GBox Chemi XX6 Gel imaging system with 302 nm UV transilluminator.

8.4 ITS Sequencing

Colony PCR was performed to amplify ITS regions for species confirmation. 40 µl 0.02 M NaOH was added to a 1.5 ml Eppendorf. Using a sterile tip, a small amount of one colony was collected, then suspended in this solution. The tube was then boiled for 10 minutes at 105°C, and then iced for another 10 minutes. The tube was then centrifuged at 5000 rpm in room temperature for 2 minutes.

2 µl of supernatant was added to a PCR tube with 10 µl 5x PCRBIO reaction buffer, 0.5 µl PCRBIO HiFi polymerase, 2 µl each 10 mM ITS1 and ITS4 (ab_796 and ab_797) primers (tab. 3) and 33.5 µl water. The tube was added to a C1000 Thermo Cycler (BioRad) and underwent an initial denaturing at 95°C for 1 minutes, followed by 30 cycles of 95°C 15 second denaturing, 55°C 15 second annealing and 72°C 30 second extension.

Gel electrophoresis was done to identify the presence of the expected band at 535 bp. The concentration of each amplicon was estimated and diluted to 5 ng/µl. 15 µl of this dilution was added to 2 µl of ITS1 primer and added to a Eurofins Mix2Seq kit. Samples were Sanger sequenced by Eurofins and sequences were then run in NCBI BLAST to confirm species identity.

8.5 MinION Sequencing

Genomic DNA extractions were performed using Qiagen Genomic DNA Buffer set and Qiagen Genomic-tip 100/g columns. The purity and yield were then calculated using a Nanodrop spectrophotometer (Thermo Fisher Scientific). Adequate samples had a yield of >300 ng/µl, a 260:280 of approximately 1.8 (+/-0.1) and a 260/230 around 2.0 (+/-0.1). These samples were then kept at 4°C before being sent to NIAB EMR for MinION long read sequencing.

8.6 Sequence Analysis

Sequence data (FASTA) was initially viewed by aligning assembled contigs to the reference SC5314 using D-genies. This allowed crude identification of translocations or karyotype abnormalities for the haploid genome. To assess read depth and identify CNVs, the reads were plotted against the reference genome using Circos. Reads from SC5314 genome long read sequencing (Price, *unpublished data*) were also mapped to the reference to compare coverage.

To characterize rearrangement, the raw read data was sorted with Samtools and mapped to the reference genome using Minimap2. BAM alignment files were then exported and viewed in IGV, where changes in read depth could be identified. Breakpoint loci apparent from the D-genies alignment were identified in IGV, where chimeric reads spanning translocation breakpoints were identified to confirm translocations. These reads identified in IGV were chosen if they were long (>30 kb), spanned the breakpoint and mapped to multiple chromosomes. Such reads were copied from IGV and then analyzed in Snapgene, and the regions either side of the breakpoint were run through NCBI BLAST to confirm they could only have originated from different chromosomes. Reads were only regarded as significant if they extended outside of the MRS region as this is highly homologous between chromosomes. Searches for telomeric repeats were also performed on reads around breakpoints which were extracted with Samtools to try to characterize the rearrangements which had occurred. In all cases, the proximity of the predicted breakpoints relative to the sgRNA target was mapped to ensure all observed rearrangements were at the MRS.

8.7 Evolution Experiment

Once MRS targeted strains with chromosomal rearrangements were identified, an evolution experiment was conducted to assess (1) the genome stability of these strains in different conditions compared to the WT and (2) whether the rate of adaption to stress differed between MRS targeted strains and the WT. During the evolution experiment, karyotypes and relative fitness of these strains and the evolved WT were assessed at regular intervals. It was expected that genome instability would be greater in the MRS targeted strains than the evolved WT and that these strains would be able to adapt to stress more quickly. MRS targeted strains with the CRIPSR cassette still integrated were used in the evolution experiment to simulate increased genome instability at the MRS.

8.7.1 Solid Media

Most of the evolution experiments were done on solid media to allow visualization of colony morphology. This allowed changes in colony morphology to be more easily identified throughout the experiment.

Strains were grown overnight in liquid YPD and then plated onto solid media with the number of CFUs plated depending on the culture conditions (tab. 7). The conditions used were all on YPD media (tab. 5), and were: 30°C, 37°C, 37°C + 5% CO_2, or 30°C with 12 mM caffeine. 30°C acted as a non-selective control. 37°C and 37°C + 5% CO2 are clinically relevant stresses as they are conditions which *Candida* would encounter within a human host. Caffeine (1,3,7- trimethylxanthine) was used as it has previously been used in our lab and is a fungal cell wall stress which inhibits adhesion and biofilm formation (Fuchs and Mylonakis, 2009; Raut et al., 2013). SDS was also used as a stress in preliminary experiments, and a concentration of 0.1% was chosen as an optimal stress, eliminating ~90% of WT SC5314 colonies. This concentration proved too high for the MRS mutant strains which did not grow, and this condition was abandoned.

Table 7. Culture conditions for evolution experiment on solid media.

Media	Conditions	Number of cells plated
YPD	30°C	100
YPD	37°C	100
YPD	37°C 5% CO_2	100
YPD + 12 mM caffeine	30°C	1000

Plated strains were then grown for one week before being imaged and three colonies from each plate were stocked and re-streaked onto the same media and condition (fig. 20). Thereafter, a blob of cells was taken from each strain and re-streaked in the same way every two days for three weeks. At approximately weekly intervals, each plate was imaged, and the strains stocked. This continued until the end of the experiment, when final stocks were made of the evolved strains four weeks from when they were originally plated.

Upon completion of the evolution, spotting assays were performed to identify the relative fitness of the evolved strains relative to their parental strain in addition to both the evolved, and parental SC5314 WT in the condition (fig. 20). These were done in the conditions to which the strains had been evolved.

Once particularly evolved strains were identified, CHEF electrophoresis was performed again, to identify any chromosomal rearrangements relative to both the WT and the pre-evolved parental strain.

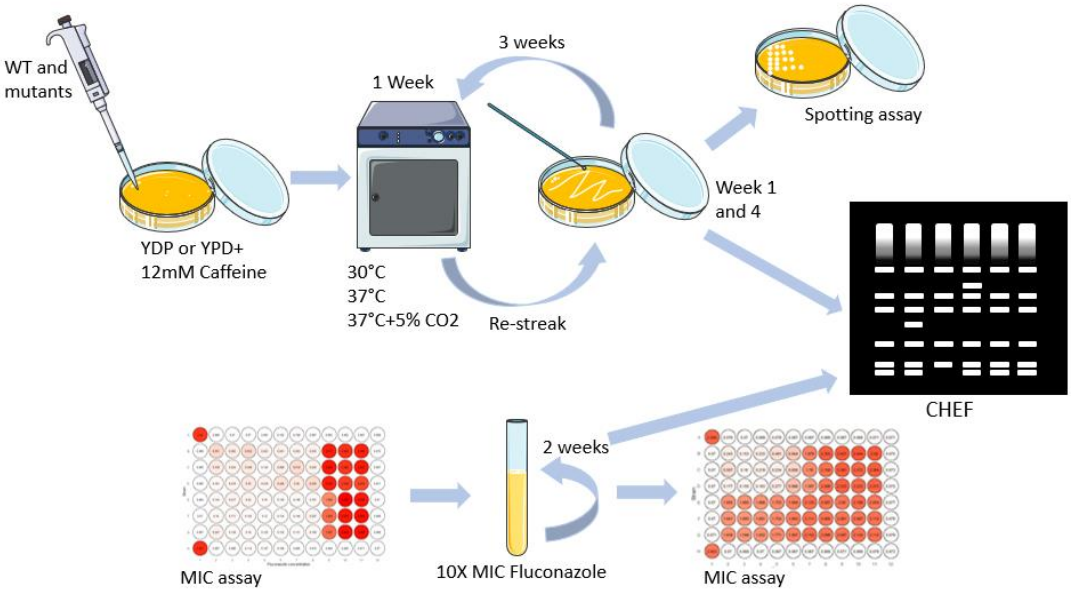

Figure 20. Flowchart of evolution experiment.

Strains are grown in stresses on solid media, being regularly re-streaked. They are then karyotyped and changes in fitness are identified by spotting assay. For fluconazole evolution, a base MIC is calculated, then strains are evolved in liquid media supplemented with Fluconazole. Changes in MIC and karyotype are then calculated.

8.7.2 Evolution Experiment in Liquid Media

As it is the drug most frequently used to treat *C. albicans* infection, this study was particularly interested in understanding how *Candida* become resistant to Fluconazole. To assess this, strain 1096 and WT SC5314 were evolved in liquid SC (tab. 5) containing Fluconazole.

The MICs of the strains were ascertained. Overnight cultures of the strains in SC media were then washed with sterile water. Tubes containing SC media and 10x of each strain's MIC of Fluconazole, or the equivalent volume of DMSO were inoculated with each strain to an OD of 0.1. These tubes were incubated at 37°C, shaken at 180 rpm for 72 hours. A glycerol stock was then made of each, before removing 1 ml of the culture and adding it to a tube containing 9 ml of fresh SC containing Fluconazole/DMSO. This was then repeated every 72 hours for 15 days. After the 15 days, the MIC of the evolved strains was then calculated, with the Fluconazole and DMSO evolved derivatives of each parental strain on the same plate. These strains were also karyotyped again by CHEF electrophoresis.

9 Genomic Results and Identification of MRS Mutants

9.1 CRISPR-Cas9 System

This study aimed to generate chromosomal rearrangements around the MRS to understand the evolutionary significance of this repetitive element. We induce DSBs at the MRS, and by not providing a repair template, intend on promoting homology-based repair using uncut MRS elements as templates, resulting in rearrangements.

To achieve this, we adapted a CRISPR protocol which had been developed for gene editing in *C. albicans* (Nguyen, Quail and Hernday, 2017). This system overcomes limitations of previous systems as changing the target sequence does not require cloning into large plasmids and does not disrupt the *ENO1* antigen gene (Vyas et al., 2015). The integration of the Cas9 cassette into the genome also facilitates lasting expression, better suited to this study than the transient expression achieved by non-integrating systems (Min et al., 2016).

9.2 Design of a sgRNA to Target the RB2 Subunit

The CRIPSR sgRNA was designed based on maximizing the number of MRS-specific targets within the genome. However, targets which were tandemly repeated within the same element, or across the MRS were excluded. This was to promote translocation rather than segmental deletions or repeat array expansion/reduction.

All RB2, Hok and RPS sequences in assembly 22 were downloaded from the *Candida* genome database (http://www.candidagenome.org/). RB2 sequences were available for chromosomes 1, 2, 3, 4, 4a, 5, 5a, 6, 7a, 7b and R, RPS sequences were available for chromosomes 1, 2, 4, 6, 7a, 7b, R and Hok sequences were available for chromosomes 1, 2, 4, 5, 6, 7a, 7b, R. The homologues were aligned in Jalview, using Clustal with defaults (fig. 21).

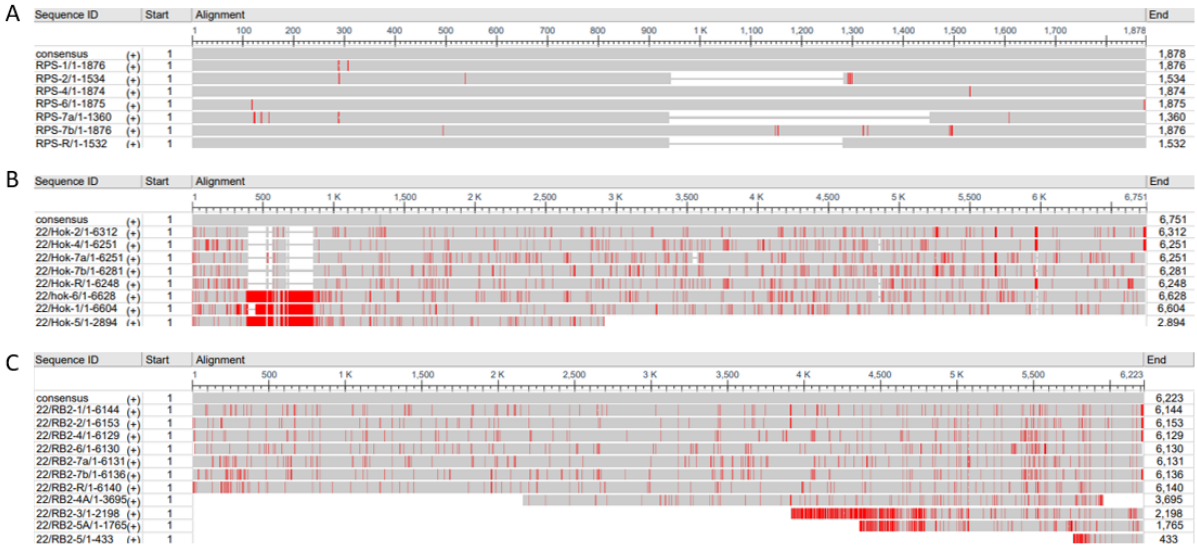

Figure 21. Alignment of RPS (A), Hok (B) and RB2 (C) homologues.

Constructed on NCBI Multiple Sequence Alignment Viewer version 1.22.0 following alignment in Jalview. Consensus for each element is at the top, followed by the homologues. Red bars indicate mismatches between that homologue and the consensus, grey areas match the consensus and white areas are present in the consensus but not in the homologue.

Of the three elements, the RPS was the most conserved, with very high sequence similarity between the chromosomes (fig. 21). The RB2 and Hok elements had a similar degree of homology, with significantly more mismatches than the RPS elements. Furthermore, chromosome 5 had a truncated Hok region,

around half the length of its homologues and chromosomes 3, 4A, 5 and 5A had truncated RB2 elements (fig. 21).

Non tandemly repetitive regions of at least 20 bp which were well conserved between chromosomes were then identified. Regions of greatest homology were then loaded into Benchling (https://www.benchling.com/) to identify potential CRIPSR targets based on a 20 bp sgRNA and nGG PAM sight. Targets were chosen based on high on-target scores. Off-target scores were ignored due to the intention of having multiple targets.

Prospective target sequences were then put into NCBI BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify any homology to non-MRS elements and identify the frequency of their occurrence in the genome. Targets with homology to other genomic regions, particularly those within or proximal to coding regions were rejected.

This led to three potential targets, with one in each of the MRS elements. The RB2 target was chosen for this study due to it having the highest on-target score (75.3) and because the RB2 is the most commonly occurring MRS element. This target had a low off target score (4), indicating multiple sites in the genome would be targeted by the same gRNA as intended.

The RB2 target is around 2.4 kb from the *FGR6* coding region of the RB2 and is near the distal end of the element. The target sequence is completely conserved between all the RB2 units for which data was available expect for chromosome 4A, which was truncated at this region.

9.2.1 Identification of all sgRNA Targets in *C. albicans* Genome

RB2 sgRNA targets were identified in Snapgene. A search was done on each chromosome to identify sequences within 2 mismatches of the 20 bp sequence (fig. 22). The presence of a PAM site was also checked for each.

There are 11 exact matches to the RB2 sgRNA sequence in the haploid *C. albicans* genome (Assembly 22). Each is also adjacent a PAM site and are therefore viable cut sites. Each chromosome has at least one of these sites and chromosomes R, 5 and 7 have two. An additional two sites were identified which were highly similar, one on chromosome 1 bearing a single mismatch, and another on chromosome 3, with a mismatch and an insertion. These are also adjacent PAM sites and are potential, but less likely cut sites.

Clearly, this exceeds the number of MRS elements in the genome. There are however 14 partial RB2 subunits in the *C. albicans* genome (Assembly 22). BLAST analysis was conducted to identify all RB2 elements to the reference genome chromosomes which were mapped on Snapgene. This was done to ensure all sgRNA targets were in RB2 elements. All 13 potential sgRNA targets are within RB2 elements.

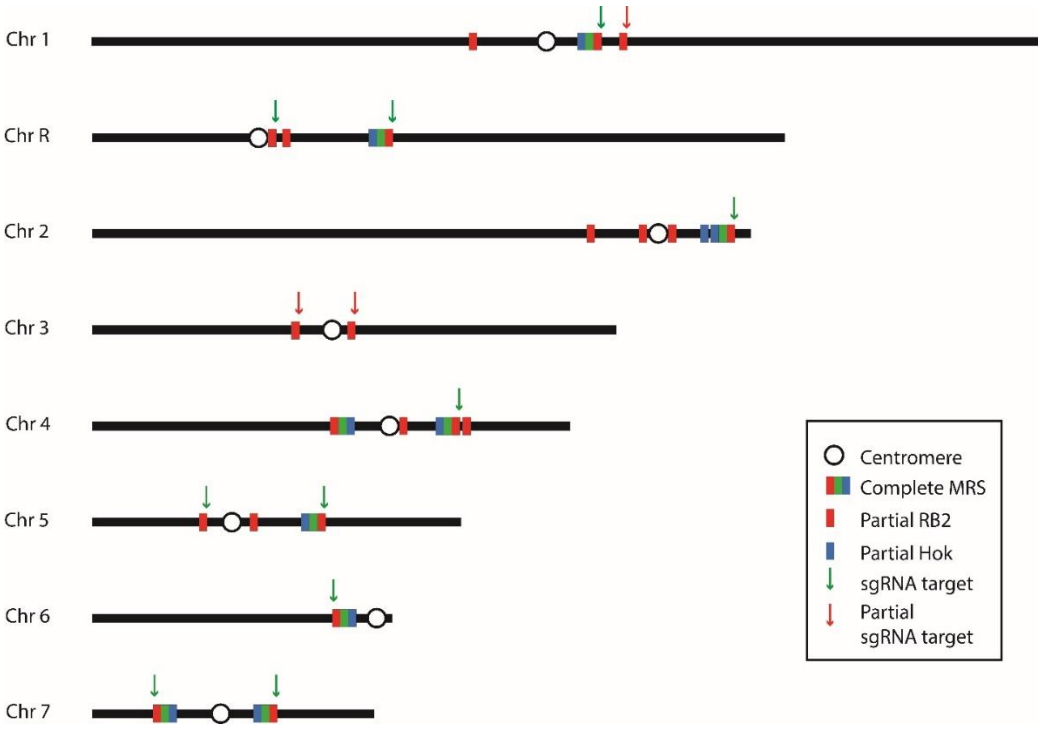

Figure 22. sgRNA targets in the Candida albicans genome.

Figure as in section 7.9.2, constructed from Long read data. Data differs slightly from Assembly 22 discussed above. Onto the figure, the sgRNA targets have been added to show the distribution of potential cut sites. Green arrows mark perfect matches with 20/20 matching bases, while red arrows are partial matches, allowing for up to two mismatches, gaps, or insertions. Note is at least one potential target on each chromosome. While it is possible for sequences with more than two mismatches to still be targeted, it is far less likely.

9.3 Screening Transformed Colonies by CHEF Identifies Multiple Rearrangements

The RB2 CRISPR transformation was performed on two occasions, each time with four replicates. Successful integration of the complete CRISPR cassette allowed colonies to grow in Nat. These colonies varied in size and some displayed filamentation as may be expected by targeting the MRS (fig. 23).

A total of 62 colonies were screened by CHEF electrophoresis. CHEF separates whole chromosomes by size, and so allowed the identification of large chromosomal rearrangements. Colonies were selected to represent the different morphologies. A random selection of colonies was selected matching the approximate proportion of large, medium, and small colonies on the plates and all those with increased filamentation were also screened. Two colonies were particularly unusual: strain 1095 and 1096 (fig. 23). 1095 formed dense pseudo hyphae and its high degree of filamentation made it difficult to re-streak or suspend in liquid. Strain 1096 grew at the edge of the plate and had hyphae which subsequently radiated out across the media (fig. 24).

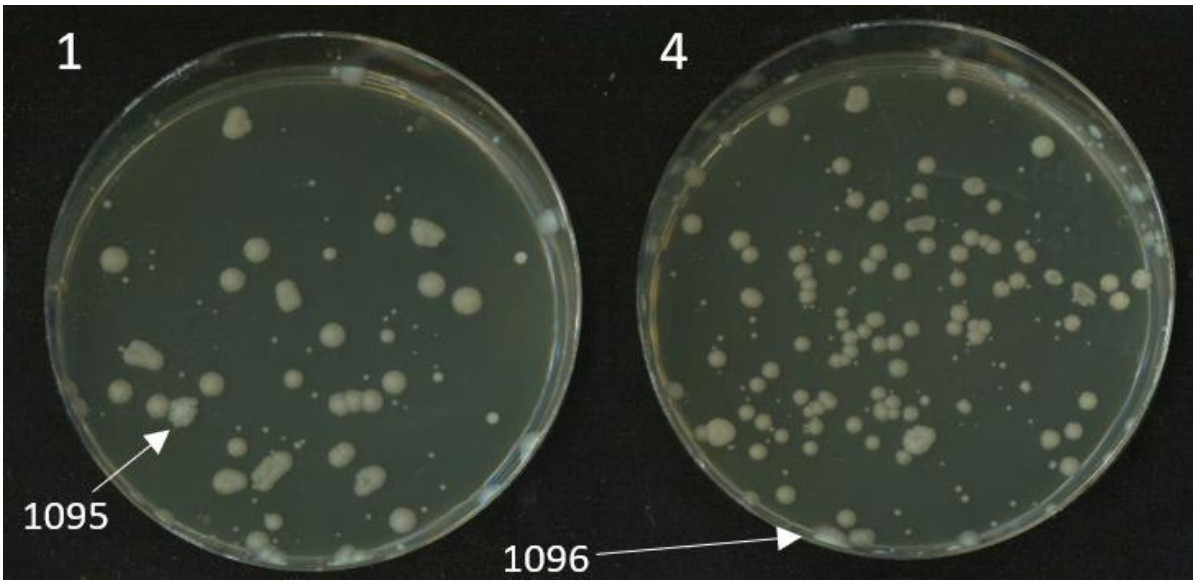

Figure 23. YPD + NAT plates 1 and 4 flowing RB2 transformation.

Ten days following transformation, a range of colony morphologies were present. Most colonies were smooth and round and differentiated as small, medium, or large. Arrows indicate pseudohyphal morphology of strain 1095 identified on plate 1, and the unusual hyphal morphology of strain 1096 on plate 4.

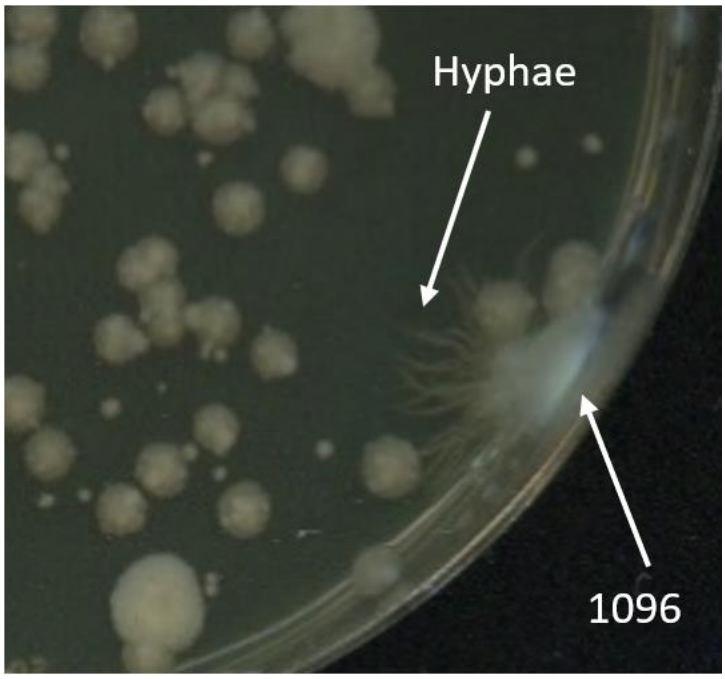

Figure 24. Hyphal morphology of strain 1096.

Strain 1096 can be seen on the bottom right edge of the YPD + NAT plate. This is the same plate as figure 23, following another 7 days of growth. After 17 days, it had grown long hyphae and was morphologically distinct from neighboring colonies.

CHEF electrophoresis of the 62 colonies following transformation identified three strains with distinct karyotypes (fig. 25). The first, strain 986 was chosen at random and did not have a distinct morphology, growing as a circular, entire, raised colony. It has an additional band corresponding to a chromosome of around 1.9 Mb (fig. 25). Both stains which were identified as having particularly irregular morphologies: 1095 and 1096 were subsequently found to also have atypical karyotypes. It therefore seems that filamentation is a good indicator of successful generation of rearrangements around the MRS.

Strain 1095 has an additional band corresponding to a chromosome of around 1.1 Mb. Strain 1096 has two additional bands corresponding to around 2.1 Mb and 1.5 Mb and appears to have lost the band corresponding to chromosome 7 (fig. 25).

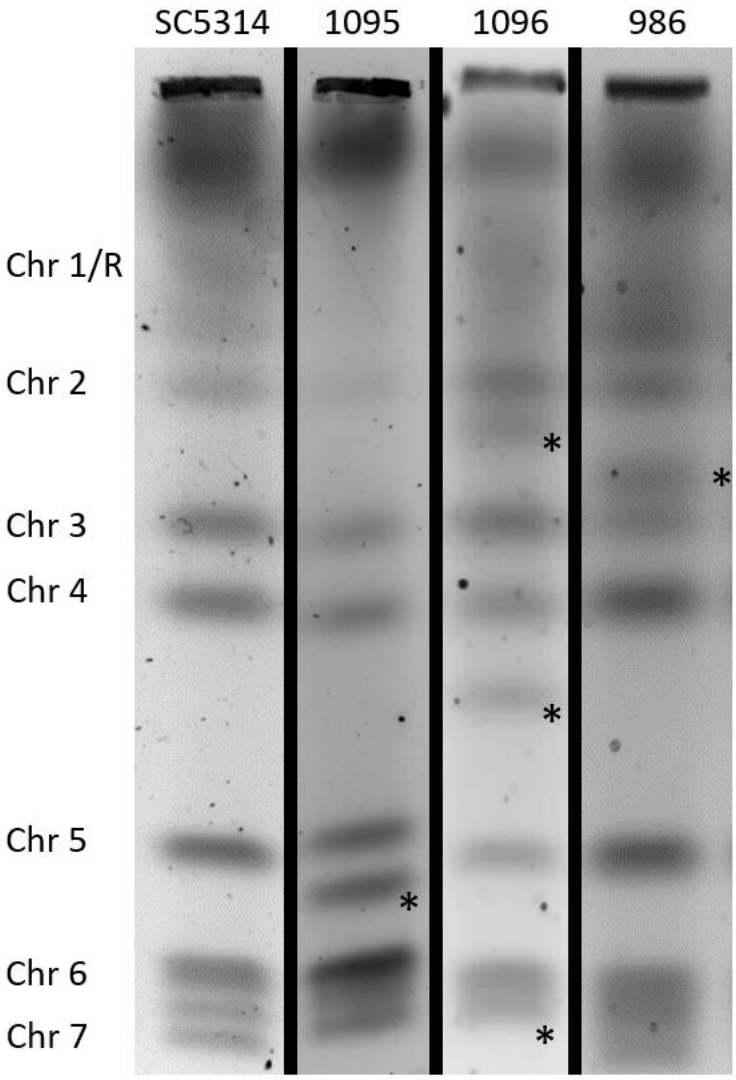

Figure 25. CHEF electrophoresis of MRS mutants.

Figure shows the CHEF banding pattern of WT SC5314, alongside three strains with mutant karyotypes identified following transformation of SC5314 with a CRISPR cassette containing sgRNA targeting the RB2 subunit of the MRS. The figure is a composite from four independent CHEF runs and non-WT bands or the loss of WT bands are marked with asterisks (*). The chromosome corresponding to each band is labelled against the WT SC5314 lane. Each original gel was run with a WT control.

9.4 Species Confirmation by ITS Sequencing

The ITS (internally transcribed spacer) region of the three strains: 986, 1095 and 1096 was PCR amplified and the amplicons sanger sequenced (Eurofins Genomics) to confirm that these strains were indeed *C. albicans* and not contaminants. The sequences were then analyzed on NCBI BLAST. The ITS amplicons of strains 1095 (E=0), 1096 (E=0) and 986 (E=0) only aligned to the *C. albicans* genome with high sequence identity (fig. 26). This confirms that the strains, despite karyotypic divergence and morphological abnormalities, are *C. albicans.*

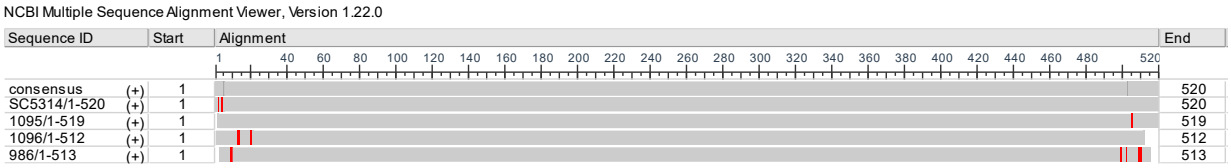

Figure 26. Sequence alignment of ITS amplicons.

ITS amplicons from strains 1095, 1096 and 986 were sequenced and aligned to the ITS region of the SC5314 reference genome. Red lines indicate mismatches. Note there are few mismatches at the ends of the amplicons, but that sequence identity is 100% for 480 bp of the 520 bp. Generated in NCBI Multiple Sequence Alignment Viewer version 1.22.0 following alignment in Jalview.

9.5 Removal of CRISPR Cassette

The CRISPR cassette contains a gene encoding FLP recombinase and has a FLP recombinase target (FRT) at each end. This FLP recombinase is maltose induced, and so growth on maltose media results in excision of the cassette from the genome. Removal of the cassette results in loss of NAT resistance. Cassettes were removed where possible prior to sequencing to maintain genome integrity and prevent further rearrangements.

1000 cells of strains 986, 1095 and 1096 were plated onto YP maltose to induce the FLP recombinase. Colonies grew successfully for each strain (fig. 27) but were significantly larger for strain 1096.

Figure 27. Plating transformed strains on YP maltose.

Strains 1095, 1096 and 986 were plated on YP maltose to induce removal of CRISPR cassette by FLP recombinase.

Five colonies of each strain grown on YP maltose were re-streaked onto YPD + Nat and YPD as a positive control. Successful removal of the cassette would include removal of the *NAT* gene, and so strains should grow on the YPD plate only.

While all five colonies of 1096 and 986 grew on YPD but not in the presence of Nat, all five colonies of 1095 grew on both plates (fig. 28). This indicates that the cassette was not successfully removed from strain 1095. 1095 was plated again onto YP maltose and then 10 colonies were screened from this plate.

All colonies retained Nat resistance. It appears that the cassette cannot be removed from strain 1095, likely due to mutations to the FLP recombinase gene, its promoter, or either FLP recombinase target.

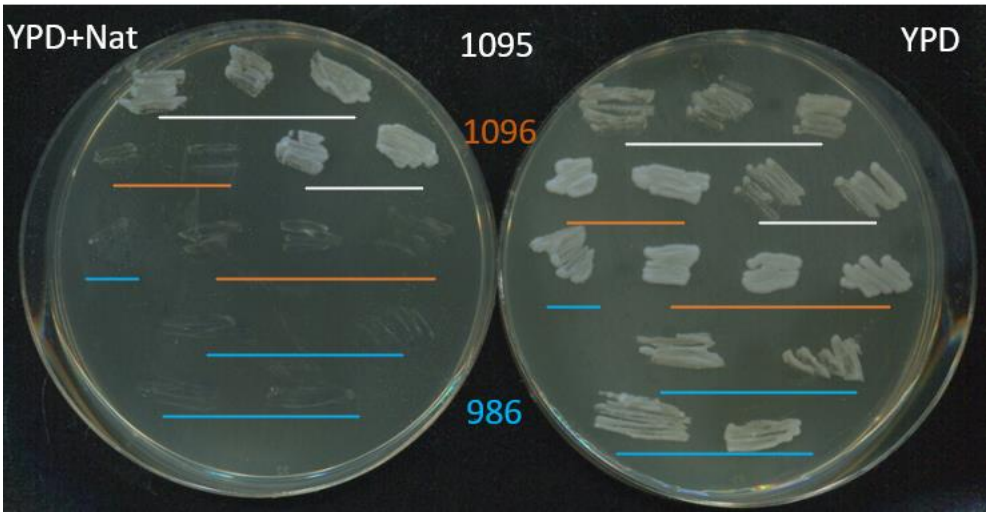

Figure 28. YPD+NAT and YPD plates, screening for cassette removal.

From right to left, five colonies of each strain 1095, 1096 and 986 were streaked onto both plates. All colonies grow on YPD as expected and all colonies of 1096 and 986 do not grow in the presence of Nat, indicating removal of cassette. All 1095 colonies grow in presence of Nat, indicating retention of cassette.

CHEF electrophoresis was performed on the strains from which the cassette had been removed to ensure doing so had not altered their karyotype. Two of the five screened colonies above were run for each strain 1096 and 986. These were run along their corresponding strain with the cassette still integrated, the WT, and strain 1095 (fig. 29). This identified that one of the cassette-removed derivatives of strain 1096 had the same karyotype as 1096, while the second had a karyotype which appeared to match that of the WT. Both strains from 986 appeared to maintain the same karyotype (fig. 29).

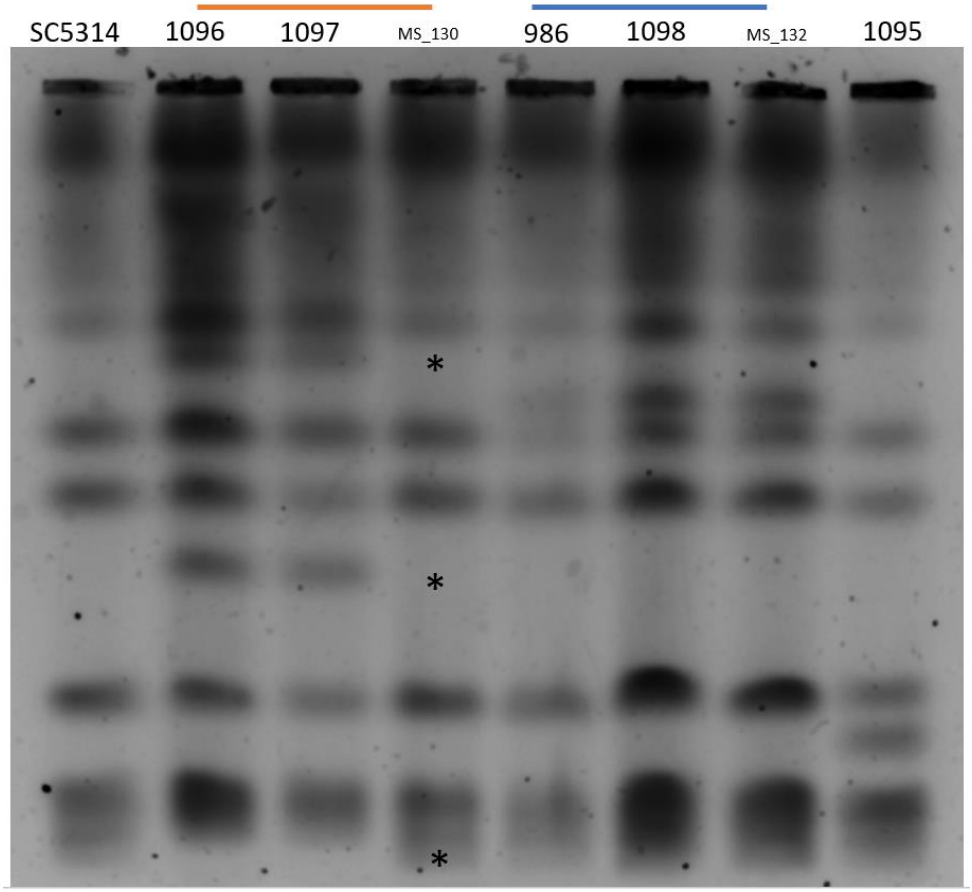

Figure 29. CHEF of MRS mutants following removal of CRISPR cassette.

CHEF was performed on WT SC5314, MRS mutants 1096, 986 and 1095 as well as two strains derived from CRISPR cassette removal from each strain 1096 and 986. The cassette could not be removed from 1095. Once cassette-removed strain from 1096 (1097) maintained its parental karyotype and went on to be sequences, while the other (MS_130) appears to have a WT-like karyotype- the differences in banding pattern between this strain and 1096 are marked with asterisks (*). Both cassette-removed strains from 986 maintained their parental karyotype and 1098 went on to be sequenced.

9.6 Long Read Sequencing of MRS Mutants

To better understand the chromosomal rearrangements in strains 1095, 1097 (1096 without cassette) and 1098 (986 without cassette), MinION long read sequencing was performed. Long read sequencing allows identification of chromosomal rearrangements as reads are long enough to span the breakpoints and extend into unique flanking sequences. Conversely, short read sequencing technologies, including Illumina, would be unlikely to identify such changes. Similarly, short read sequencing is poor at resolving repetitive elements whereas long reads can extend straight across them and accurately represent their length, and distribution in the genome. As this study aimed to generate chromosomal rearrangements at a repetitive element, long read sequencing was clearly most suitable.

9.6.1 Coverage of Genome Suggests CNVs in Strains 1095 and 1098

Following sequencing of the 1095, 1097 and 1098 genomes, reads were mapped to the SC5314 reference genome Assembly 22 to view the coverage of each region of the genome and identify CNVs. This was compared to reads from SC5314 genome MinION long read sequencing (Price, *Unpublished data)*. The overall read depth varies between the three mutant strains; 1095 being the lowest, followed by 1097 and then 1098.

Coverage of strain 1097 is highly consistent with the SC5314 genome sequencing, with no fluctuations in read depth along the chromosomes and peaks in coverage conforming with those in the SC5314 read alignment (fig. 30). This indicates no CNVs, including segmental or chromosomal aneuploidies have occurred which suggests that all chromosomal rearrangements which took place were reciprocal (fig. 30).

The strain 1095 genome sequence has very low coverage overall (fig. 30), which could be indicative of haploidy or may be simply due to less efficient genomic DNA extraction. Indeed, the latter is highly likely due to the pseudo-hyphal colony morphology of the strain, which did not readily resuspend. The read

depth of chromosomes 5 and 6 appears significantly higher than the other chromosomes in strain 1095, which may indicate aneuploidy of these two chromosomes. This observation however requires further analysis. There are also numerous peaks seen in the reference strain and strains 1097 and 1098, which appear to have been lost from strain 1095 (fig. 30). This observation also needs further investigation as does the reason for the inability to excise the CRISPR cassette from this strain.

The genome sequencing of 1098 had very high coverage (fig. 30), however this was consistent between chromosomes and is not believed to be significant. Nearly all peaks in coverage conform with strain SC5314, however, there is increased coverage of Chr3 between 633-786 kb (fig. 30). At the center of this region, is a partial RB2 sgRNA target sequence. This locus was also where the contigs corresponding to Chr3 each end. Contig 11, 11.6 kb, maps within this region but has low homology. Analysis of raw reads identified a long (>100 kb) read spanning this region. Pairwise BLAST analysis of this read against itself identified it as an inverted repeat. The same alignment was then done between the read and the Chr3 reference. This showed that this strain has an inverted duplication at this locus, which is not present in the reference genome. As this occurs at the site of a partial sgRNA target sequence, it seems highly likely this is due to Cas9 action, despite the absence of a complete MRS on Chr3. This would account for the increased coverage identified in this region (fig. 30). This duplication is ~153 kb and contains 17 verified genes and many additional unconfirmed ORFs. This lengthening of Chr3 is very likely to account for the extra chromosome band identified by CHEF which is appears slightly longer than Chr3 (fig. 25).

Overall, only this one CNV, on Chr3 of 1098 was identified with certainty. This CNV can explain the CHEF banding pattern of the strain. No CNVs were identified in the 1097 genome, indicating that the non-WT CHEF banding pattern of this strain is due to reciprocal chromosomal rearrangements which will be identified in the next section. The coverage of the 1095 genome sequencing appears to show some changes in copy number, but these require further analysis, and the genome of this strain will not be further analyzed in the present study.

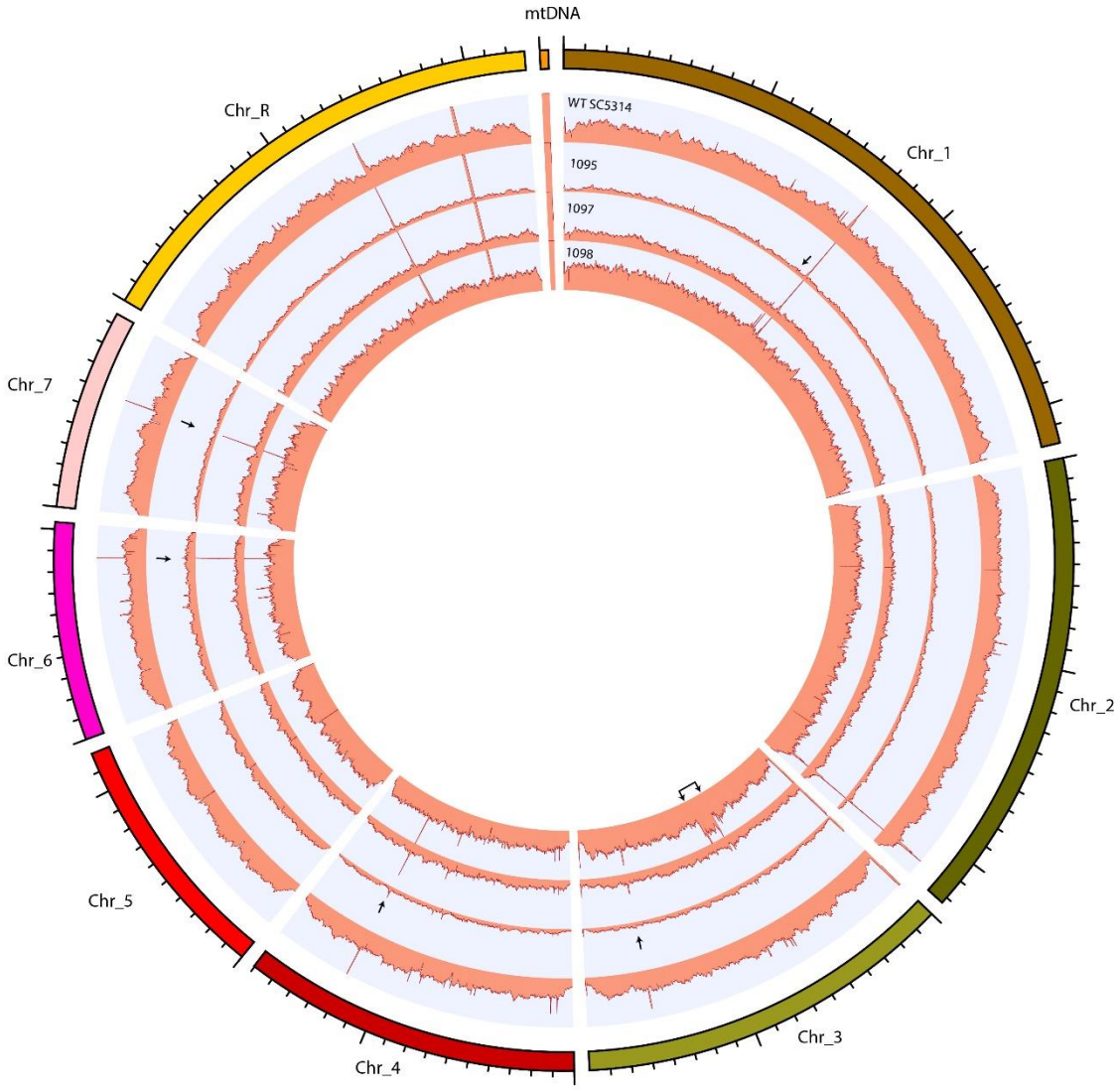

Figure 30. Circos plot of WT SC5314 and strains 1095, 1097, 1098.

Figure shows the read depth of the sequenced genomes mapped against the reference genome Assembly 22. Coverage of Chr5 and 6 in strain 1095 are significantly higher than the other chromosomes and several peaks present in all other strains are absent (marked with arrows). Coverage of the strain 1097 genome appears to conform with the WT. The 1098 genome has a segmental duplication on Chr3 (marked by arrows) which is evident from the high coverage of the corresponding region.

9.6.2 Sequence Alignment and Characterization of Translocations

Reads from the sequencing of the 1097 genome were assembled into contigs which were then aligned to the reference genome using D-genies to identity chromosomal rearrangements. Initially, these were

identified when a contig mapped to more than one chromosome. Most small inversions were ignored as they likely represent repeat loci, which are omitted or poorly represented in the reference genome. Chromosomes were characterized as WT-like if they mapped to only one reference chromosome, matched its approximate length (allowing for discrepancy due to repeats) and had telomeric repeats at both ends. Often, the whole chromosome appears as one contig, but in some instances these could not be assembled. In such cases, chromosomes were deemed WT-like if the sections on different contigs overlapped, covered the whole reference chromosome between them and between them, had two telomeres.

Upon identifying translocations, the position of the breakpoint on each chromosome was identified relative to the sgRNA target. Read spanning these breakpoints were then identified in IGV and characterized using NCBI BLAST and Snapgene. Importantly, the alignments only represent the haploid genome. As *C. albicans* is a diploid organism, additional karyotype changes may not be represented. Indeed, while analyzing raw reads spanning breakpoints, evidence of additional translocations was found which were not apparent from the alignment.

9.6.2.1 The 1097 Genome has Multiple Translocations around MRS Elements

Alignment of the 1097 contigs to the reference genome showed that at least four contigs align to multiple chromosomes, indicating translocations have occurred (fig. 31-32). These involve chromosomes 1, 4 and 7.

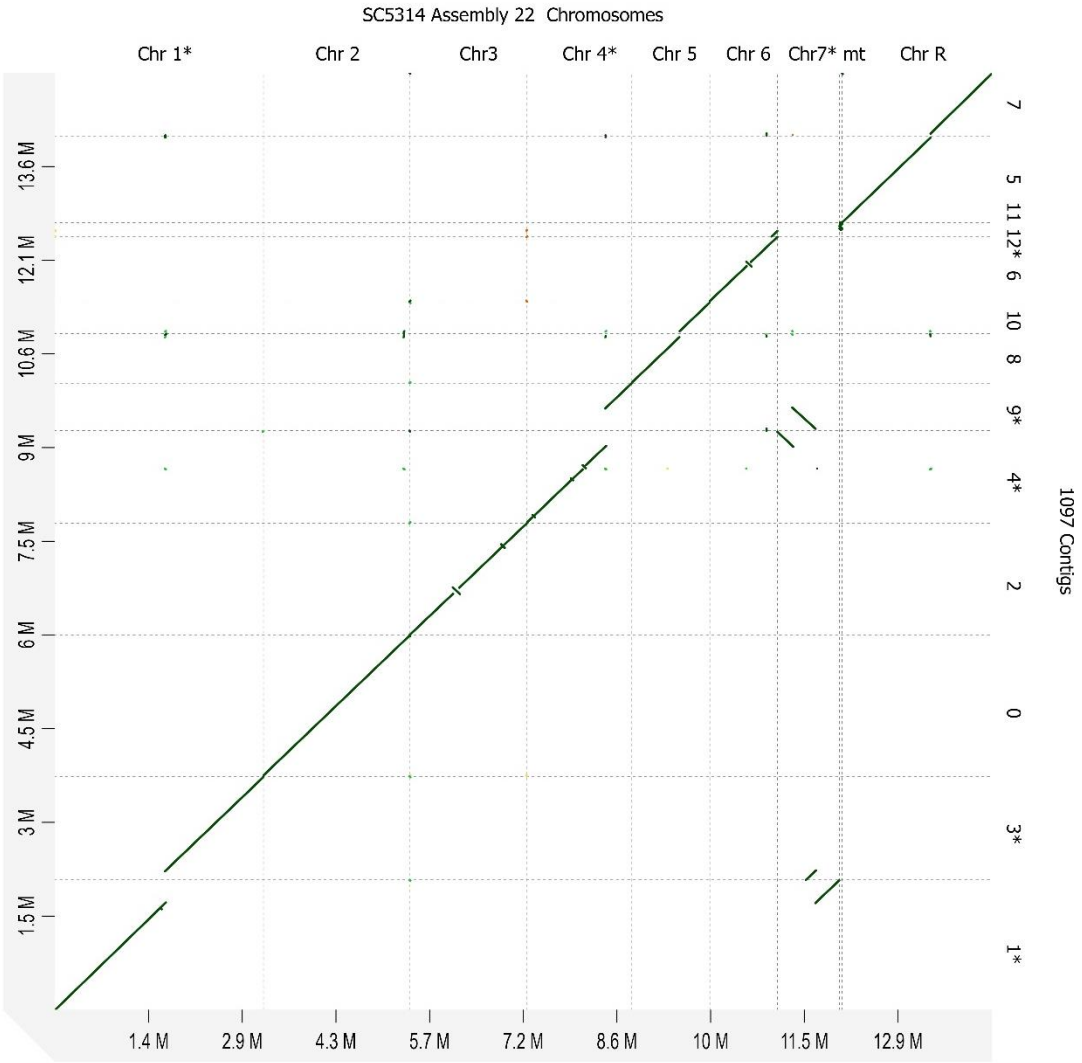

Figure 31. D-genies alignment of strain 1097 contigs to Assembly 22.

Figure shows the alignment of the 1097 contigs against the reference genome. Contigs and chromosomes of interest are marked by asterisks. Contigs which align continuously to a complete chromosome can be considered WT, whereas contigs 1, 3, 4 and 9 each align to two reference chromosomes, indicating multiple translocations have occurred. Contig 12 is very short and overlaps with contig 6- this will require further investigation but may be as artifact of assembly. It appears Chr 1 and 4 have been cut once and chr 7 has been cut twice.

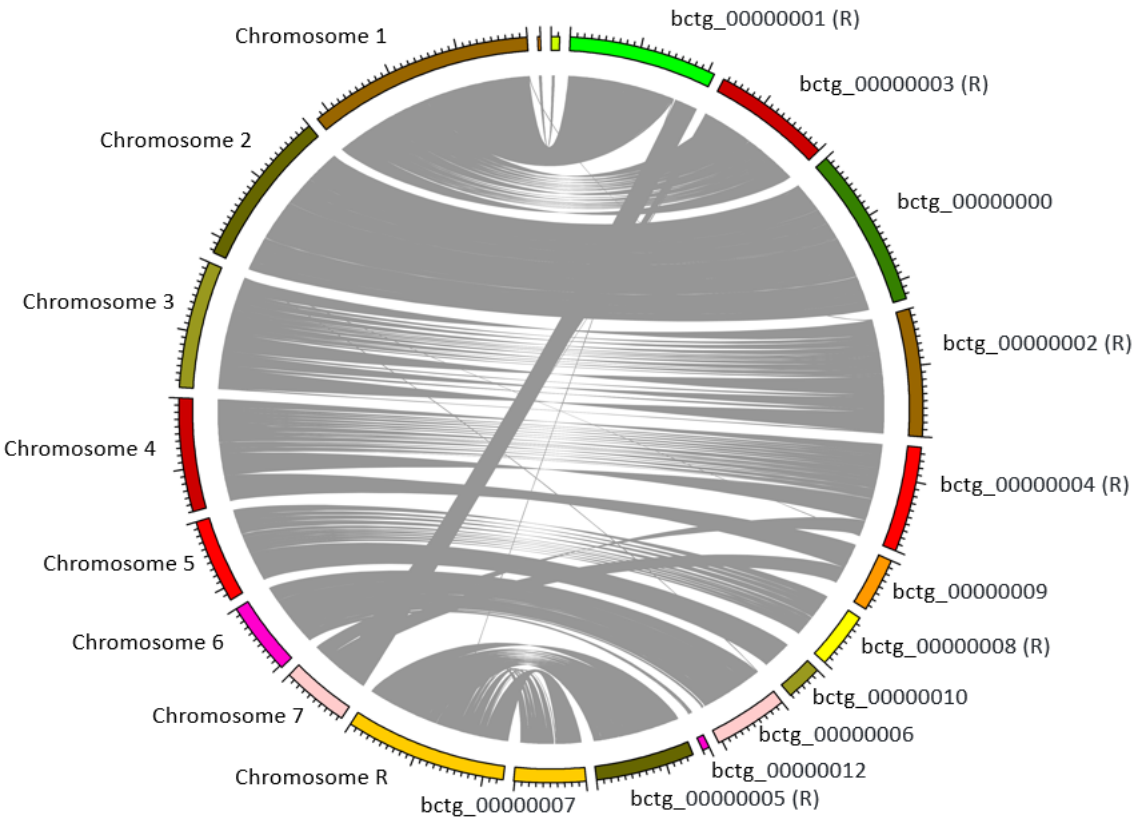

Figure 32. Circos plot of 1097 mapped to Assembly 22.

This plot represents the same data as the above alignment. Again, translocations involving Chr 1, 4 and 7 can be seen as several contigs each map to two of these chromosomes. The unlabeled, small alignment at the top is the mtDNA and can be ignored.

Contig 1 indicates a translocation between chromosomes 1 and 7 forming a 2.081 Mb chimeric chromosome (fig. 31). This contig has telomeric repeats at both ends, and the breakpoints on Chr1 and Chr7 are at the MRS. The other half of Chr1 is on contig 3. No overlap exists between these two halves of Chr1, indicating that both originate from the same copy of Chr1, cut at the MRS and not homologous chromosomes. No read spanning this breakpoint has been identified yet but analyses are ongoing.

This second half of Chr 1 on contig 3 appears to be joined to a central region of Chr7 (fig. 31). This contig only has telomeres at the end of the Chr1 arm. Further analysis is needed to identify whether reads at the Chr7 end have telomeric repeats. These have not yet been identified. This end of Chr7 is proximal to the Chr7 centromere, which may have affected the assembly. Interestingly, this central region of Chr 7

on contig 3 overall overlaps with a region of contig 9, which maps to a larger central region of Chr3 (fig. 31). A ~41 kb read was identified which spanned this breakpoint, extending significantly outside of the MRS in both directions (fig. 33). BLAST analysis of this read confirmed that the regions flanking the MRS on either side can only be from Chr 7 and Chr 1, respectively, and that it is therefore a chimeric chromosome formed by a translocation at the MRS.

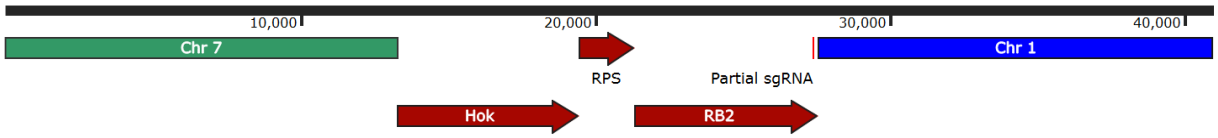

Figure 33. Chr7-Chr1 translocation breakpoint.

Generated in Snapgene. A read spanning the Contig 3 breakpoint was extracted from IGV. The MRS elements were mapped, and BLAST analysis confirmed that the flanking regions originate from Chr 7 and Chr 1.

Contig 9, 765.3 kb, contains the whole central region of Chr7, between the two MRS elements, with each breakpoint around the MRS (fig. 31). This indicates that both sgRNA targets on Chr7 were cut on a single copy of the chromosome. The rest of contig 9 maps to a distal region of chromosome 4 (fig. 31). The breakpoint on Chr 4 is also proximal to a sgRNA target. This contig only has telomeres at the end of the Chr4 section. Further analysis needs to be done to try to identify telomeric repeats in reads covering the Chr7 breakpoint to confirm whether this contig represents a whole chimeric chromosome as well as to identify reads spanning the breakpoint where Chr 7 joins Chr 4.

The rest of Chr4 appears on contig 4, attached to the remaining end of Chr7 (fig. 31). Both breakpoints have already been discussed and are proximal to the sgRNA target. This contig has telomeric repeats at both ends indicating it represents an entire chimeric chromosome. A 64 kb read has been identified which spans this region (fig. 34). This read extends outside of the MRS on both sides, with flanking

regions aligning to chromosomes 7 and 4 only. Another two reads of 49 kb and 30 kb were subsequently identified also spanning this region (not shown).

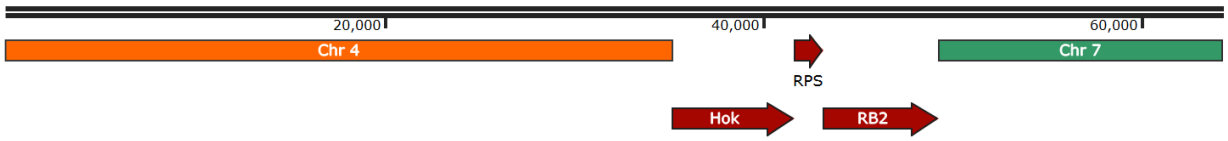

Figure 34. Chr4-Chr7 translocation breakpoint.

Generated in Snapgene. This read spans the MRS, and flanking regions map to Chr 4 and Chr 7.

Contig 12 is very short, at 101 kb, has telomeric repeats only at one end, and maps only to Chr6. This region of Chr6 also appears on contig 6, which spans the entire chromosome and has telomeres at both ends (fig. 31). No reads have been identified which map to Chr 6 and any other chromosome. Because of these reasons, it is assumed that this contig an artifact of assembly and does not signify chromosomal rearrangements involving Chr6.

The two chimeric chromosomes which are completely resolved, appearing as contigs 1 and 4, could account for the extra bands seen on the CHEF gel (fig. 25, 31). These contigs are 2.081 Mb and 1.491 Mb, respectively which match the estimated sizes of the additional bands (fig. 25).

The assembly shows Chr: 2, 3, 5, 6 and R as being WT-like (fig. 31). Although Chr5 and R are split over two contigs, these contigs overlap and each only have a telomere at one end. It therefore appears that this is an assembly issue and that these WT-like chromosomes are present in the strain. However, analysis of individual reads implicated some of these chromosomes in structural rearrangements which are not apparent from the alignment.

A 42 kb read was identified which spans the MRS with flanking regions mapping exclusively to chromosomes 5 and 7 (fig. 35).

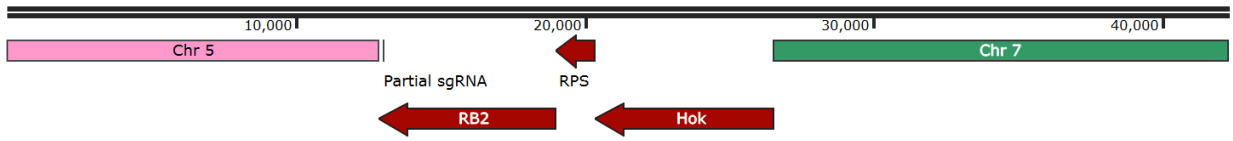

Figure 35. Chr5-Chr7 translocation breakpoint.

Generated in Snapgene. This read spans the MRS and flanking regions map to Chr 5 and Chr 7.

A ~57 kb read was identified which contains the MRS and flanking regions with homology to Chr 1 and Chr 4 only (fig. 36).

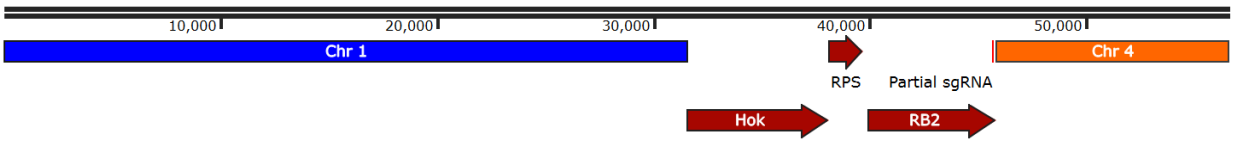

Figure 36. Chr1-Chr4 translocation breakpoint.

Generated in Snapgene. This read spans the MRS, and flanking regions map to Chr 1 and Chr 4.

Another 69 kb read also maps to chr1 and a 15 kb region of chr4. This second break does not appear to be at MRS. A 42 kb read was also identified which mapped to Chr 7 and a 3 kb region of Chr R. This again, does not appear to be proximal to an MRS element and both require further investigation. If these reads are valid, then it seems both copies of Chr 1, 4 and 7 may have been cut. Due to their absence from the alignment, it is difficult to characterize these events presently.

9.7 Assessing the Phenotypic Effect of MRS-Driven Chromosomal Rearrangements

The phenotypes of the MRS mutants were also assessed to determine the impact of the observed chromosomal rearrangements on morphology and fitness. We aim to show that chromosomal rearrangements around the MRS can generate phenotypic diversity which would support the predicted evolutionary role of the MRS: facilitating chromosomal rearrangements to generate phenotypic diversity.

Growth rates at 30°C (non-stress) and 37°C (host-specific stress) were assessed to indicate fitness. Thereafter, morphology was assessed by colony and cell microscopy, with the specific goal of identifying and quantifying variations in filamentation. Filamentation is both easy to identify and is also strongly associated with virulence in *C. albicans*. Note that phenotyping was generally done with strains with the CRISPR cassette integrated: 1095, 1096 (1097 parental) and 986 (1098 parental).

9.7.1 MRS-Driven Rearrangements Alter Growth Rate and Temperature Sensitivity

Growth rate in YPD at 30°C was significantly reduced for strain 1095, while both 986 and 1096 were approximately consistent with the WT (fig. 37a). This reduction in growth of 1095 was exacerbated at 37°C, and 986 also exhibited a moderately reduced growth rate compared to the WT at this elevated temperature but not to the extent of 1095 (fig. 37b). 1096 had a growth rate equivalent to that of the WT at both temperatures (fig. 37). Growth rates were also assessed for strains 1097 and 1098 to ensure that cassette removal had not affected their phenotype (not shown).

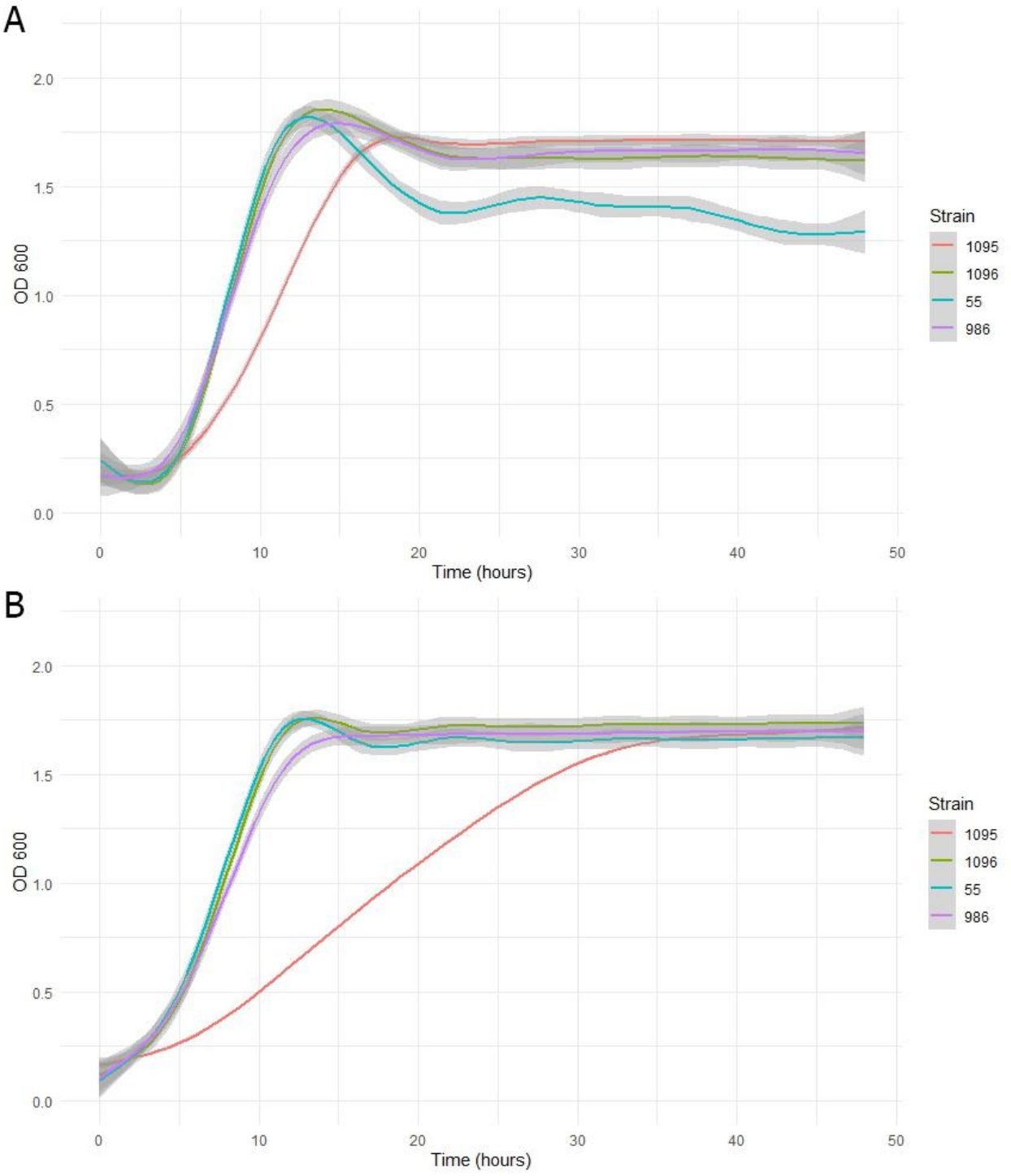

Figure 37. Growth curves of MRS mutant strains.

Figure shows the growth data of WT SC5314, and MRS mutants 1095, 1096 and 986 at 30 (A) and 37 (B) degrees Celsius in YPD media. Data is from 12 technical replicates. Lines have been fitted using a generalized additive model (GAM) and shaded areas indicate 95% confidence intervals.

9.7.2 MRS-Driven Rearrangements Produce Pseudohyphal Colony Morphology

As previously discussed, strains 1095 and 1096 were initially screened due to hyper-filamentous phenotypes. 1095 grew as a darker, wrinkled, pseudohyphal colony which was difficult to streak or suspend, tending to remain clumped (fig. 23). 1097 was a very white, large colony which grew on the YPD + Nat media and up the side of the plate. Further, 1096 grew long hyphae which extended across the plate (fig. 24). 986 was screened randomly, and nothing about its colony morphology was exceptional, forming WT-like white, circular, raised colonies.

While the pseudohyphal colony morphology of 1095 has remained highly consistent, the hyphal phenotype of 1097 has not been replicated since this initial observation (fig. 38).

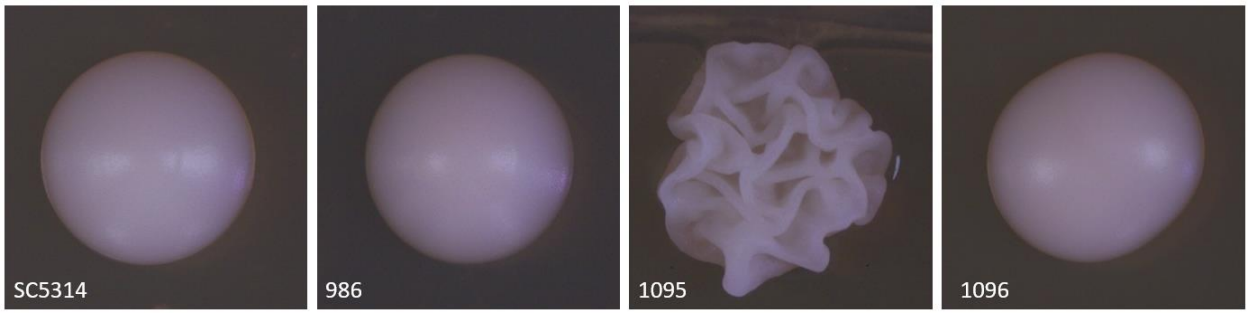

Figure 38. Colony morphology of SC5314 and MRS mutants.

Colony morphology of 986 and 1096 are consistent with that of SC5314 on YPD at 30°C. This is despite the original observation of 1096 as forming hyphae. These colonies are all roughly circular, entire, raised colonies of approximately equal size. They are all white with a shiny appearance. 1095 is constantly distinct from the other strains in that it forms pseudohyphal colonies which appear wrinkled and irregular in shape.

9.7.3 MRS-Driven Rearrangements Generate Hyperfilamentous Cellular Morphology

The proportion of filamentous to yeast cells was measured by light microscopy for WT SC5314 and the three MRS-targeted strains (1095, 1096 and 986). This was measured at both 30 and 37°C. At least 100 cells were counted in total for each strain. Quantification was done from images taken at 60x magnification. A minimum of eight images were taken for each strain, chosen at random co-ordinates.

The percentage of yeast and filamentous cells was calculated for each image and the mean of these percentages was calculated and plotted (fig. 9).

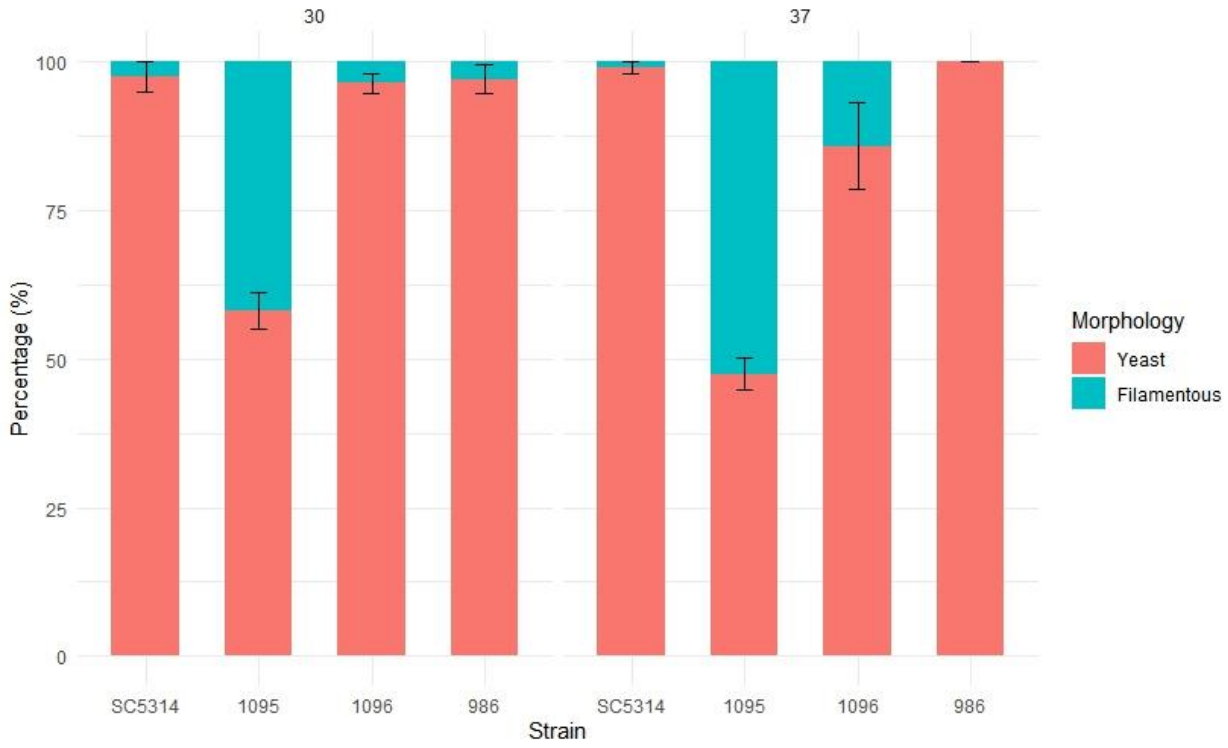

Figure 39. Frequency of filamentous morphologies at 30 and 37°C.

The stacked percentage plot shows the proportion of yeast and filamentous cell morphologies of WT SC5315 and the three MRS targeted strains. This was measured at both 30 and 37°C. Strain 1095 has significantly more filamentation than the other three strains at both temperatures. At 37°C, strain 1096 also has significantly more filamentation than WT and strain 986. Filamentation significantly increases with temperatures for strains 1095 and 1096. The opposite trend is seen with strains SC5314 and 986, this is not however significant. Error bars show the standard error of the mean.

Filamentous morphology was significantly associated with strain at both 30°C (X^2= 122.91, df = 3, p < 2.2e-16) and 37°C (X^2 = 64.336, df = 3, p= 6.958e-14). Strain 1095 has consistently more filamentation at both temperatures, significantly increasing between the two temperatures (X^2 = 5.0837, df = 1, p-value = 0.0242). Observed differences between the other strains at 30°C are non-significant. At 37°C, 1096 also has more filamentation than the WT. These results are consistent with colony morphology observations, with 1095 being consistently filamentous, 1096 occasionally showing filamentation and neither SC5134 nor 986 showing any filamentation under normal lab conditions. In addition to hyphae and

pseudohyphae, a range of yeast forms were visible under light microscopy, including large, ovoid yeast cells, particularly at 37°C (fig. 40).

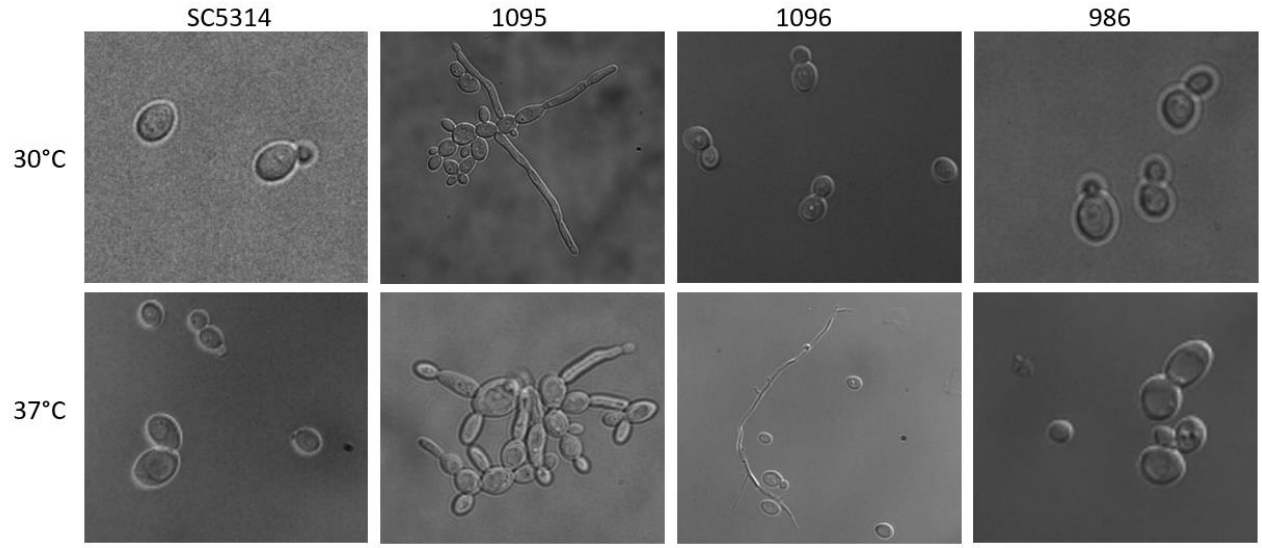

Figure 40. Cellular morphologies of SC5314 and MRS mutants at 30 and 37°C.

Images were captured at 100X and have been cropped and magnified to highlight morphological differences between strains and at different temperatures. They are not to scale. Images were selected to represent the typical morphology of the strain. Only 1095 had significant filamentation at both temperatures. 1096 had significant filamentation at 37°C only. A greater range of yeast cell sizes is seen at 37°C for strains SC5314, 1096 and 986. Cell sizes cannot be compared between images.

10 Evolution Experiment: Results

Evolution experiments were performed to assess the genome stability of the MRS-targeted strains relative to WT SC5314 as well their ability to adapt to stress. These experiments included strains 1095, 1096 and 986, which still have the CRIPSR-Cas9 cassette integrated into their genomes, in addition to the WT. Each strain was evolved with three independent replicates. The expectation was that Cas9 would continue to cut available target sequences throughout the experiment to artificially increase the rate of MRS-driven rearrangements and genome instability. This increased rate can be compared to the evolved WT. This is intended to show that:

1. Continued expression of Cas9 can simulate increased genome instability.
2. MRS-mediated genome instability can facilitate rapid evolution and adaptation to stress.

These strains were evolved at 30°C (non-stress), 37°C, 37°C + 5% CO_2 and 12 mM caffeine, each in triplicate. At the end of week 1, and at the end of week 4, evolved strains were karyotyped by CHEF electrophoresis and their stress-specific fitness was assessed. Karyotype data for CO_2 and Caffeine evolved strains at week 4 are unavailable.

Table 8. Strains generated during evolution experiment.

ID	Strain name	Parental	Condition	Timepoint	Replicate
MS_8	55/30/1/1	SC5315	30	Week 1	1
MS_9	55/30/1/2	SC5315	30	Week 1	2
MS_10	55/30/1/3	SC5315	30	Week 1	3
MS_11	1/30/1/1	1095	30	Week 1	1
MS_12	1/30/1/2	1095	30	Week 1	2
MS_13	1/30/1/3	1095	30	Week 1	3
MS_14	986/30/1/1	986	30	Week 1	1
MS_15	986/30/1/2	986	30	Week 1	2
MS_16	986/30/1/3	986	30	Week 1	3
MS_17	55/37/1/1	SC5315	37	Week 1	1
MS_18	55/37/1/2	SC5315	37	Week 1	2
MS_19	55/37/1/3	SC5315	37	Week 1	3

MS_20	1/37/1/1	1095	37	Week 1	1
MS_21	1/37/1/2	1095	37	Week 1	2
MS_22	1/37/1/3	1095	37	Week 1	3
MS_23	986/37/1/1	986	37	Week 1	1
MS_24	986/37/1/2	986	37	Week 1	2
MS_25	986/37/1/3	986	37	Week 1	3
MS_26	55/C02/1/1	SC5315	CO2	Week 1	1
MS_27	55/C02/1/2	SC5315	CO2	Week 1	2
MS_28	55/C02/1/3	SC5315	CO2	Week 1	3
MS_29	1/CO2/1/1	1095	CO2	Week 1	1
MS_30	1/CO2/1/2	1095	CO2	Week 1	2
MS_31	1/CO2/1/3	1095	CO2	Week 1	3
MS_32	986/CO2/1/1	986	CO2	Week 1	1
MS_33	986/CO2/1/2	986	CO2	Week 1	2
MS_34	986/CO2/1/3	986	CO2	Week 1	3
MS_35	55/Caff/1/1	SC5315	12mM Caffeine	Week 1	1
MS_36	55/Caff/1/2	SC5315	12mM Caffeine	Week 1	2
MS_37	55/Caff/1/3	SC5315	12mM Caffeine	Week 1	3
MS_38	1/Caff/1/1	1095	12mM Caffeine	Week 1	1
MS_39	1/Caff/1/2	1095	12mM Caffeine	Week 1	2
MS_41	986/Caff/1/1	986	12mM Caffeine	Week 1	1
MS_42	986/Caff/1/2	986	12mM Caffeine	Week 1	2
MS_43	986/Caff/1/3	986	12mM Caffeine	Week 1	3
MS_44	1/Caff/1/3	1095	12mM Caffeine	Week 1	3
MS_45	55/30/2/1	SC5315	30	Week 2	1
MS_46	55/30/2/2	SC5315	30	Week 2	2
MS_47	55/30/2/3	SC5315	30	Week 2	3
MS_48	1/30/2/1	1095	30	Week 2	1
MS_49	1/30/2/2	1095	30	Week 2	2
MS_50	1/30/2/3	1095	30	Week 2	3
MS_51	986/30/2/1	986	30	Week 2	1
MS_52	986/30/2/2	986	30	Week 2	2
MS_53	986/30/2/3	986	30	Week 2	3
MS_54	55/37/2/1	SC5315	37	Week 2	1
MS_55	55/37/2/2	SC5315	37	Week 2	2
MS_56	55/37/2/3	SC5315	37	Week 2	3
MS_57	1/37/2/1	1095	37	Week 2	1
MS_58	1/37/2/2	1095	37	Week 2	2
MS_59	1/37/2/3	1095	37	Week 2	3
MS_60	986/37/2/1	986	37	Week 2	1
MS_61	986/37/2/2	986	37	Week 2	2
MS_62	986/37/2/3	986	37	Week 2	3
MS_63	55/C02/2/1	SC5315	CO2	Week 2	1

MS_64	55/C02/2/2	SC5315	CO2	Week 2	2
MS_65	55/C02/2/3	SC5315	CO2	Week 2	3
MS_66	1/CO2/2/1	1095	CO2	Week 2	1
MS_67	1/CO2/2/2	1095	CO2	Week 2	2
MS_68	1/CO2/2/3	1095	CO2	Week 2	3
MS_69	986/CO2/2/1	986	CO2	Week 2	1
MS_70	986/CO2/2/2	986	CO2	Week 2	2
MS_71	986/CO2/2/3	986	CO2	Week 2	3
MS_72	55/Caff/2/1	SC5315	12mM Caffeine	Week 2	1
MS_73	55/Caff/2/2	SC5315	12mM Caffeine	Week 2	2
MS_74	55/Caff/2/3	SC5315	12mM Caffeine	Week 2	3
MS_75	1/Caff/2/1	1095	12mM Caffeine	Week 2	1
MS_76	1/Caff/2/2	1095	12mM Caffeine	Week 2	2
MS_77	1/Caff/2/3	1095	12mM Caffeine	Week 2	3
MS_78	986/Caff/2/1	986	12mM Caffeine	Week 2	1
MS_79	986/Caff/2/2	986	12mM Caffeine	Week 2	2
MS_80	986/Caff/2/3	986	12mM Caffeine	Week 2	3
MS_81	55/30/3/1	SC5315	30	Week 3	1
MS_82	55/30/3/2	SC5315	30	Week 3	2
MS_83	55/30/3/3	SC5315	30	Week 3	3
MS_84	1/30/3/1	1095	30	Week 3	1
MS_85	1/30/3/2	1095	30	Week 3	2
MS_86	1/30/3/3	1095	30	Week 3	3
MS_87	986/30/3/1	986	30	Week 3	1
MS_88	986/30/3/2	986	30	Week 3	2
MS_89	986/30/3/3	986	30	Week 3	3
MS_90	55/37/3/1	SC5315	37	Week 3	1
MS_91	55/37/3/2	SC5315	37	Week 3	2
MS_92	55/37/3/3	SC5315	37	Week 3	3
MS_93	1/37/3/1	1095	37	Week 3	1
MS_94	1/37/3/2	1095	37	Week 3	2
MS_95	1/37/3/3	1095	37	Week 3	3
MS_96	986/37/3/1	986	37	Week 3	1
MS_97	986/37/3/2	986	37	Week 3	2
MS_98	986/37/3/3	986	37	Week 3	3
MS_99	55/C02/3/1	SC5315	CO2	Week 3	1
MS_100	55/C02/3/2	SC5315	CO2	Week 3	2
MS_101	55/C02/3/3	SC5315	CO2	Week 3	3
MS_102	1/CO2/3/1	1095	CO2	Week 3	1
MS_103	1/CO2/3/2	1095	CO2	Week 3	2
MS_104	1/CO2/3/3	1095	CO2	Week 3	3
MS_105	986/CO2/3/1	986	CO2	Week 3	1
MS_106	986/CO2/3/2	986	CO2	Week 3	2

MS_107	986/CO2/3/3	986	CO2	Week 3	3
MS_108	55/Caff/3/1	SC5315	12mM Caffeine	Week 3	1
MS_109	55/Caff/3/2	SC5315	12mM Caffeine	Week 3	2
MS_110	55/Caff/3/3	SC5315	12mM Caffeine	Week 3	3
MS_111	1/Caff/3/1	1095	12mM Caffeine	Week 3	1
MS_112	1/Caff/3/2	1095	12mM Caffeine	Week 3	2
MS_113	1/Caff/3/3	1095	12mM Caffeine	Week 3	3
MS_114	986/Caff/3/1	986	12mM Caffeine	Week 3	1
MS_115	986/Caff/3/2	986	12mM Caffeine	Week 3	2
MS_116	986/Caff/3/3	986	12mM Caffeine	Week 3	3
MS_117	2/30/1/1	1096	30	Week 1	1
MS_118	2/30/1/2	1096	30	Week 1	2
MS_119	2/30/1/3	1096	30	Week 1	3
MS_120	2/37/1/1	1096	37	Week 1	1
MS_121	2/37/1/2	1096	37	Week 1	2
MS_122	2/37/1/3	1096	37	Week 1	3
MS_123	2/Co2/1/1	1096	CO2	Week 1	1
MS_124	2/Co2/1/2	1096	CO2	Week 1	2
MS_125	2/Co2/1/3	1096	CO2	Week 1	3
MS_126	2/Caff/1/1	1096	12mM Caffeine	Week 1	1
MS_127	2/Caff/1/2	1096	12mM Caffeine	Week 1	2
MS_128	2/Caff/1/3	1096	12mM Caffeine	Week 1	3
MS_143	55/30/4/1	SC5315	30	Week 4	1
MS_144	55/30/4/2	SC5315	30	Week 4	2
MS_145	55/30/4/3	SC5315	30	Week 4	3
MS_146	1/30/4/1	1095	30	Week 4	1
MS_147	1/30/4/2	1095	30	Week 4	2
MS_148	1/30/4/3	1095	30	Week 4	3
MS_149	986/30/4/1	986	30	Week 4	1
MS_150	986/30/4/2	986	30	Week 4	2
MS_151	986/30/4/3	986	30	Week 4	3
MS_152	55/37/4/1	SC5315	37	Week 4	1
MS_153	55/37/4/2	SC5315	37	Week 4	2
MS_154	55/37/4/3	SC5315	37	Week 4	3
MS_155	1/37/4/1	1095	37	Week 4	1
MS_156	1/37/4/2	1095	37	Week 4	2
MS_157	1/37/4/3	1095	37	Week 4	3
MS_158	986/37/4/1	986	37	Week 4	1
MS_159	986/37/4/2	986	37	Week 4	2
MS_160	986/37/4/3	986	37	Week 4	3
MS_161	55/C02/4/1	SC5315	CO2	Week 4	1
MS_162	55/C02/4/2	SC5315	CO2	Week 4	2
MS_163	55/C02/4/3	SC5315	CO2	Week 4	3

MS_164	1/CO2/4/1	1095	CO2	Week 4	1
MS_165	1/CO2/4/2	1095	CO2	Week 4	2
MS_166	1/CO2/4/3	1095	CO2	Week 4	3
MS_167	986/CO2/4/1	986	CO2	Week 4	1
MS_168	986/CO2/4/2	986	CO2	Week 4	2
MS_169	986/CO2/4/3	986	CO2	Week 4	3
MS_170	55/Caff/4/1	SC5315	12mM Caffeine	Week 4	1
MS_171	55/Caff/4/2	SC5315	12mM Caffeine	Week 4	2
MS_172	55/Caff/4/3	SC5315	12mM Caffeine	Week 4	3
MS_173	1/Caff/4/1	1095	12mM Caffeine	Week 4	1
MS_174	1/Caff/4/2	1095	12mM Caffeine	Week 4	2
MS_175	1/Caff/4/3	1095	12mM Caffeine	Week 4	3
MS_176	986/Caff/4/1	986	12mM Caffeine	Week 4	1
MS_177	986/Caff/4/2	986	12mM Caffeine	Week 4	2
MS_178	986/Caff/4/3	986	12mM Caffeine	Week 4	3
MS_179	2/30/2/1	1096	30	Week 2	1
MS_180	2/30/2/2	1096	30	Week 2	2
MS_181	2/30/2/3	1096	30	Week 2	3
MS_182	2/37/2/1	1096	37	Week 2	1
MS_183	2/37/2/2	1096	37	Week 2	2
MS_184	2/37/2/3	1096	37	Week 2	3
MS_185	2/Co2/2/1	1096	CO2	Week 2	1
MS_186	2/Co2/2/2	1096	CO2	Week 2	2
MS_187	2/Co2/2/3	1096	CO2	Week 2	3
MS_188	2/Caff/2/1	1096	12mM Caffeine	Week 2	1
MS_189	2/Caff/2/2	1096	12mM Caffeine	Week 2	2
MS_190	2/Caff/2/3	1096	12mM Caffeine	Week 2	3
MS_191	2/30/3/1	1096	30	Week 3	1
MS_192	2/30/3/2	1096	30	Week 3	2
MS_193	2/30/3/3	1096	30	Week 3	3
MS_194	2/37/3/1	1096	37	Week 3	1
MS_195	2/37/3/2	1096	37	Week 3	2
MS_196	2/37/3/3	1096	37	Week 3	3
MS_197	2/Co2/3/1	1096	CO2	Week 3	1
MS_198	2/Co2/3/2	1096	CO2	Week 3	2
MS_199	2/Co2/3/3	1096	CO2	Week 3	3
MS_200	2/Caff/3/1	1096	12mM Caffeine	Week 3	1
MS_201	2/Caff/3/2	1096	12mM Caffeine	Week 3	2
MS_202	2/Caff/3/3	1096	12mM Caffeine	Week 3	3
MS_203	2/30/4/1	1096	30	Week 4	1
MS_204	2/30/4/2	1096	30	Week 4	2
MS_205	2/30/4/3	1096	30	Week 4	3
MS_206	2/37/4/1	1096	37	Week 4	1

MS_207	2/37/4/2	1096	37	Week 4	2
MS_208	2/37/4/3	1096	37	Week 4	3
MS_209	2/Co2/4/1	1096	CO2	Week 4	1
MS_210	2/Co2/4/2	1096	CO2	Week 4	2
MS_211	2/Co2/4/3	1096	CO2	Week 4	3
MS_212	2/Caff/4/1	1096	12mM Caffeine	Week 4	1
MS_213	2/Caff/4/2	1096	12mM Caffeine	Week 4	2
MS_214	2/Caff/4/3	1096	12mM Caffeine	Week 4	3
MS_239	55/DMSO	SC5314	DMSO	15 days	1
MS_241	2/DMSO	1096	DMSO	15 days	1
MS_242	55/Flc	SC5314	5ug/ml Flc	15 days	1
MS_244	2/Flc	1096	5ug/ml Flc	15 days	1

10.1.1 Continued Expression of Cas9 Simulates Increases Genome Instability

10.1.1.1 All MRS-Targeted Strains have Higher Genome Instability than WT

No karyotype changes were observed in any evolved WT strain throughout the experiment (fig. 41) (tab. 9). Overall, 33% of the evolved MRS-targeted strains had further karyotype changes after 1 week. This rose to 61% after 4 weeks (tab. 9). The number of strains which had undergone structural rearrangements was recorded, as opposed to the frequency of individual rearrangements due to the poor resolution of CHEF. It is important to consider that some evolved strain underwent multiple rearrangements.

Of note, only one strain 1096 evolved derivative appeared to have a karyotype change at week 4 (Wk4) but a strain which had karyotype changes at Wk1, had reverted to a WT-like karyotype. This Wk4 strain is still regarded as having undergone karyotype change.

Table 9. Frequency of karyotype changes at week 1 and week 4 of solid media evolution experiment per strain.

Strain	Karyotype changes Wk1	Karyotype changes Wk4
55	0/12	0/6
1095	7/12	6/6
1096	4/12	2/6
986	1/12	3/6

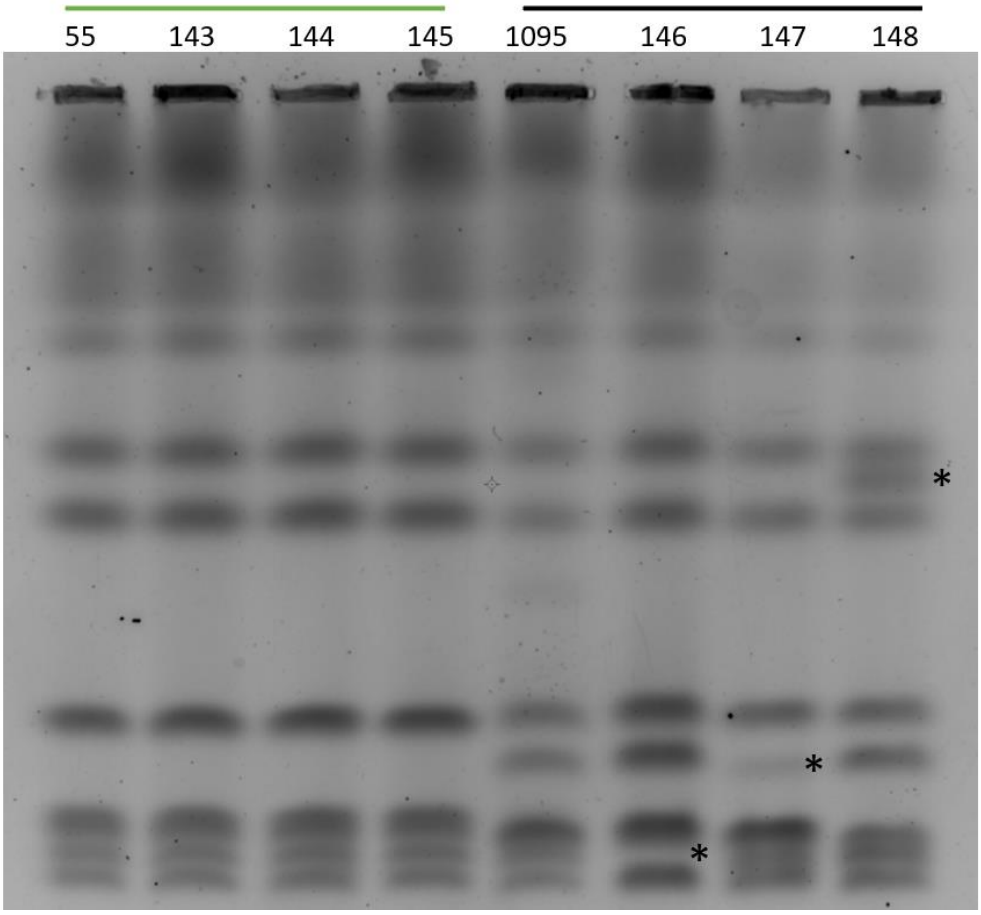

Figure 41. CHEF karyotype of 30°C-evolved SC5314 and 1095 after 4 weeks.

Evolved strains are shown adjacent their parental strain. Karyotype changes are marked with asterisks. All three of the evolved 1095 strains have undergone structural rearrangements. All three evolved SC5314 strain have maintained WT-like karyotypes.

10.1.1.2 Karyotype Changes are More Frequent in Stress Conditions

The highest frequency of karyotype changes was seen in strains evolved at 37°C (fig. 42), followed by Caffeine and CO_2 (tab. 10). 30°C acts as a non-stress control and gives a baseline for genome stability. Importantly, the karyotypes of CO_2 and Caffeine evolved strains are not known for Wk4. As these conditions generate more karyotype changes in Wk1 than the non-stress condition, it seems likely that the actual proportion of karyotype changes at Wk4 is higher than estimated.

Table 10. Frequency of karyotype changes at week 1 and week 4 of solid media evolution experiment per condition.

Stress/condition	Karyotype changes Wk1	Karyotype changes Wk2
30°C	1/12	4/12
37°C	5/12	6/12
CO_2	2/12	N/a
Caffeine	4/12	N/a

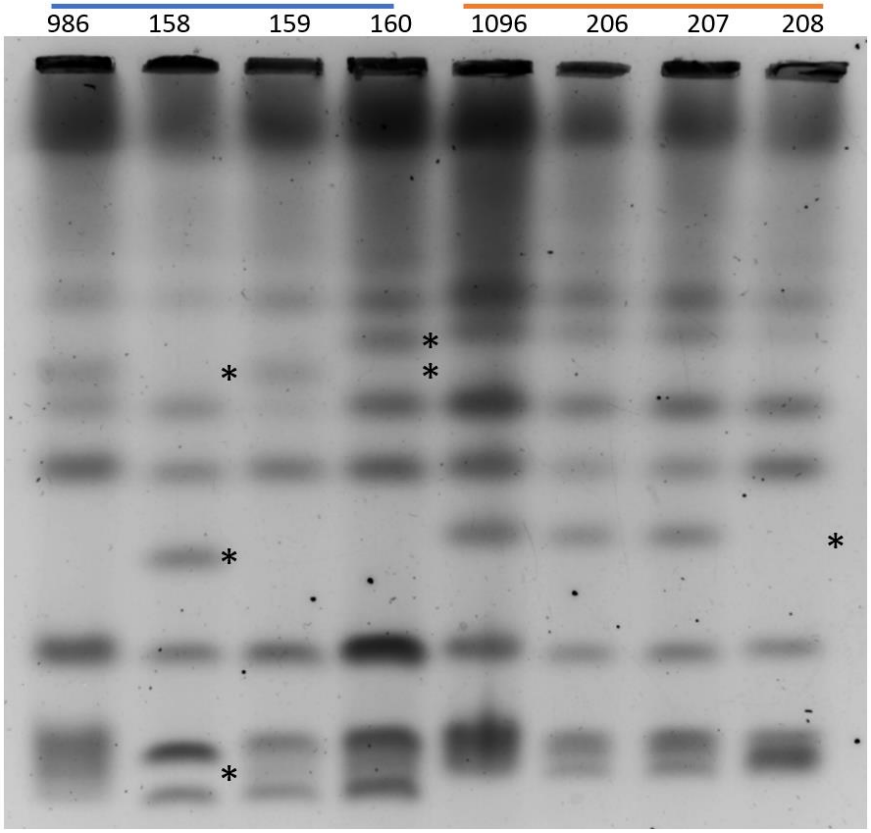

Figure 42. CHEF karyotype of 37°C-evolved 986 and 1096 after 4 weeks.

Evolved strains are shown adjacent their parental strain. Karyotype changes are marked with asterisks. Three of the six evolved strains have undergone structural rearrangements. Strain 158 appears to have experienced multiple rearrangements.

10.1.2 Frequency of Fitness Changes Varies Between Strains and Stress Conditions

Changes in fitness were identified by spotting assay and were measured by comparing the fitness of the evolved strain to its parental strain, not to the WT. This is because strains 1095, 1096 and 986 parentals have significantly different stress-specific fitness to one another and the WT. Changes in fitness are

reported simply as increase or decrease, those which are not reported do not significantly differ in stress-specific fitness from their parental strain.

10.1.2.1 Changes in Fitness are More Frequent in MRS Targeted Strains

Fitness changes were observed in all evolved MRS targeted strain and evolved SC5314 (tab. 11). At Wk1, 986 increased in stress-specific fitness significantly more frequently than any other strain. 1095 was the only strain to experience decreases in fitness at week one. At Wk4, the frequency of fitness increase was approximately equal between strains and 1095 and 986 both experiences fitness decreases. 1095 had 6/11 evolved strains decrease in fitness at Wk1 (in 37°C and CO_2), but only 3/11 at Wk4, three of these strains (three CO_2 replicates) underwent fitness increases in this time. This increase did not exceed the fitness of parental 1095. Note that 1095 is out of 11 as one replicate of 1095 in Caffeine never grew following re-streaking of the colony. Evolved strains 986 appear to have decreased in fitness between Wk1 and Wk4, following an initial increase in fitness. Both SC5314 and 1096 only changed in fitness during caffeine evolution.

Table 11. Frequency of fitness changes at week 1 and week 4 of solid media evolution experiment per strain.

Strain	Fitness increase Wk1	Fitness decrease Wk1	Fitness increase Wk4	Fitness Decrease Wk4
SC5314	3/12	0/12	3/12	0/12
1095	5/11	6/11	2/11	3/11
1096	3/12	0/12	3/12	0/12
986	9/12	0/12	3/12	3/12

10.1.2.2 Caffeine Selects for the Most Fitness Increases

At Wk1, 30°C, 37°C and CO_2 selected for the same frequency of fitness improvements while fitness declines were only experienced in 37°C and CO_2. Caffeine is the strongest selective force, with all strains evolved in this condition increasing in fitness at both timepoints (tab. 12). No fitness improvements

were seen in the other three conditions at Wk4, while half of the strains evolved in 37°C decreased in stress-specific fitness (tab. 12).

Table 12. Frequency of fitness changes at week 1 and week 4 of solid media evolution experiment per stress.

Stress/condition	Fitness increase Wk 1	Fitness decrease Wk 1	Fitness increase Wk 4	Fitness Decrease Wk 4
30°C	3/12	0/12	0/12	0/12
37°C	3/12	3/12	0/12	6/12
CO_2	3/12	3/12	0/12	0/12
Caffeine	11/11	0/11	11/11	0/11

Overall, there appears to be no clear relationship between karyotypic rearrangements and improved fitness. There are however some cases where MRS-targeted strains increased in stress-specific fitness faster than the evolved WT (fig. 43).

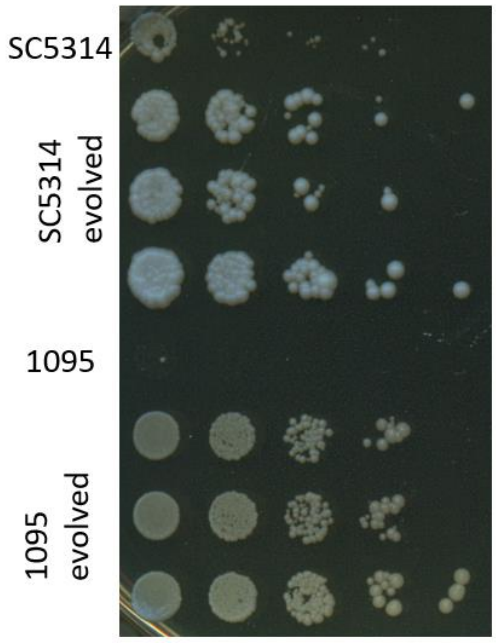

Figure 43. Spotting assays of Caffeine evolved strain 1095 and WT.

Figure shows spotting assay of 1095 evolved in caffeine for 1 week. It also has the parental strain as well as the evolved WT from the same timepoint and the parental WT. The third replicate of 1095 is a mixed stock while other are from single colonies. Greater relative increases in caffeine-specific fitness are achieved by strain 1095 than SC5314. Fitness is compared to the parental strain. Cell suspensions are spotted at an OD of 1 in the first column, diluted 10-fold sequentially thereafter. Plates are YPD + 12 mM caffeine. Fitness is inferred from the ability to grow from higher dilutions.

10.1.3 Fluconazole Evolution Experiment

Another evolution experiment was conducted with strain 1096 and WT SC5314 in the presence of Fluconazole (Flc). Fluconazole is the most frequently prescribed drug to treat *Candida* infections and so understanding how *C. albicans* can acquire resistance is of the utmost clinical importance.

Strain 1096 was chosen as it has the most chromosome rearrangements, and its genome is the most studied of the three mutants. 1096 and the WT were evolved in SC liquid media supplemented with 5 µg/ml Flc or the equivalent volume of DMSO (control) (tab. 5) for 15 days, with media replaced every three days. This was done in liquid culture because a pilot study in liquid media yielded more definitive results than the alternative on solid media, and because the MIC testing protocol is in liquid media. A concentration of 5 µg/ml Flc was chosen as it is approximately 10x the MIC_{50} of the two strains. Again, at the end of the experiment, the evolved strains were karyotyped by CHEF electrophoresis and their stress-specific fitness was measured.

10.1.3.1 Fluconazole Selects for Chromosome Rearrangements

No karyotype changes were observed in either Fluconazole or DMSO-evolved WT SC5314 (fig. 44). The DMSO-evolved strain 1096 retained its parental karyotype, but the Flc-evolved 1096 appears to have lost a chromosome compared to its parent (fig. 44). This indicates that Flc was a strong enough stress to select for karyotype changes, while the DMSO control was not. It also provides more evidence that the MRS-targeted strains do indeed have higher levels of genome plasticity than the WT.

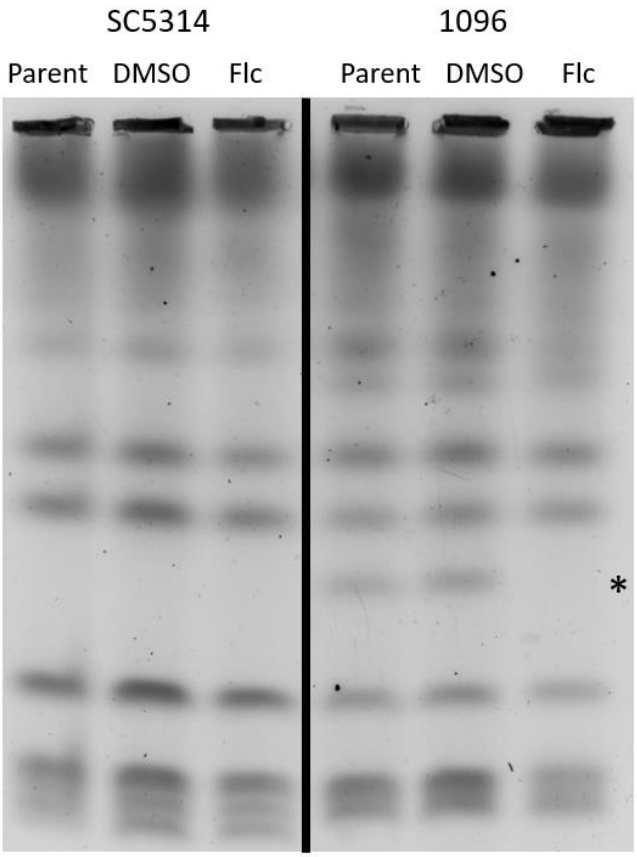

Figure 44. Karyotype changes following Fluconazole evolution.

No structural genomic rearrangements are observed in either the Flc or DMSO evolved WT (SC5314), or DMSO evolved 1096. One rearrangement is apparent in Flc-evolved 1096- marked with an asterisk.

10.1.3.2 Relative Increase in Flc Resistance is Greater in WT than MRS Mutant

Both strains 1096 and SC5314 began with the same MIC_{50}, of 0.5 µg/ml. Increases in Fluconazole resistance were observed in both the evolved WT and evolved 1096 (fig. 45-46). In all cases, fluconazole-evolved strains had higher fluconazole resistance than DMSO-evolved controls (fig. 45-46). The greatest increase in resistance was achieved by the evolved SC5314.

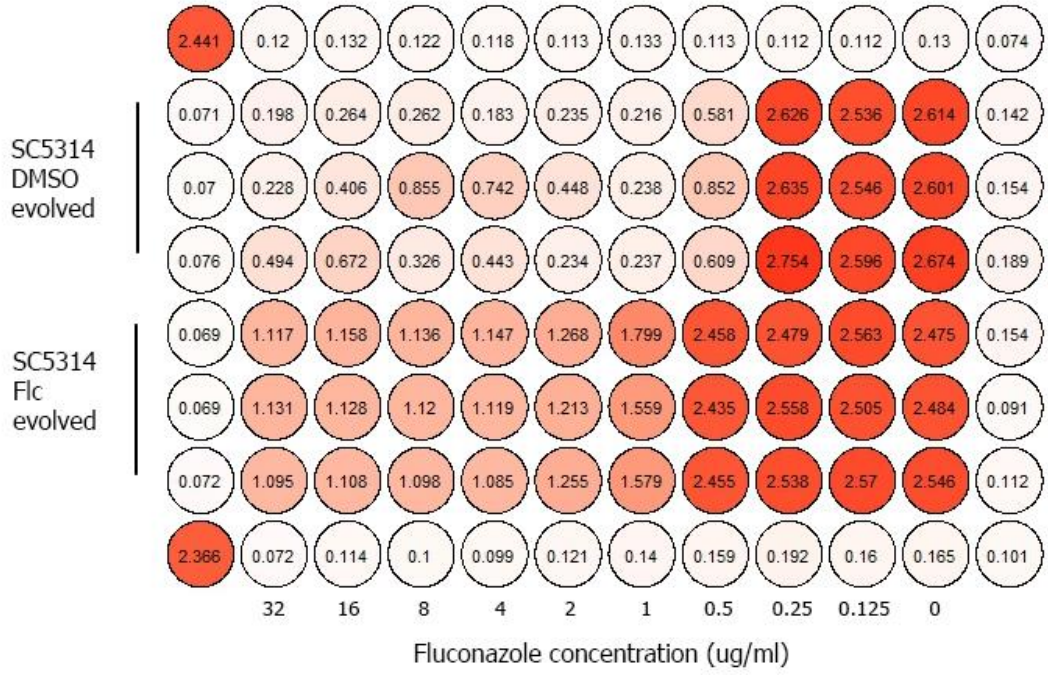

Figure 45. MIC of DMSO/Flc evolved WT SC5314.

DMSO evolved SC5314 has an MIC_{50} of 0.5 µg/ml but grows slightly up to 32 µg/ml. All three replicates of the Flc evolved SC5314 grow at 32 µg/ml and the MIC_{50} is approximately 2 µg/ml.

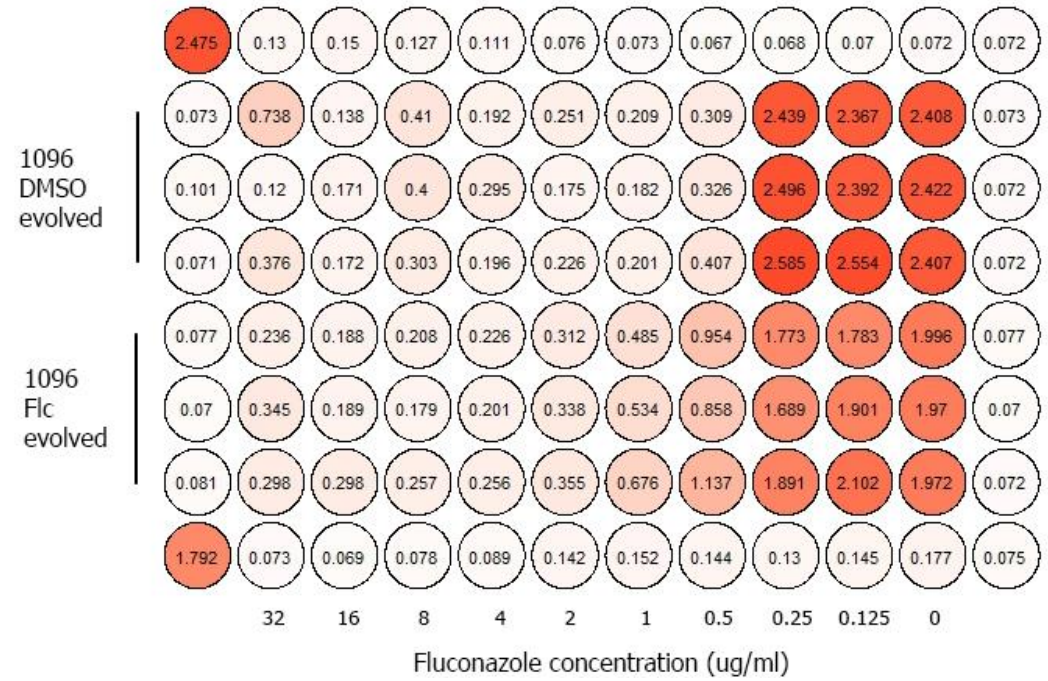

Figure 46. MIC of DMSO/Flc evolved 1096.

DMSO evolved 1096 has an MIC_{50} of 0.5 µg/ml and shows very little growth at higher concentrations. Flc evolved 1096 can grow slightly more at higher Flc concentrations but still has an MIC_{50} of 0.5 µg/ml.

11 Discussion

11.1 CRISPR-Cas9 Driven Chromosomal Rearrangements around the MRS Produce Distinct Phenotypes

This study, for the first time used genome editing techniques to generate chromosomal rearrangements in *C. albicans.* Directing Cas9 to form DSBs at the MRS in the absence of an exogenous repair template successfully promoted the use of other copies of this repetitive element as a template for homology directed repair. As intended, this resulted in the generation of chromosomal rearrangements which were identified by CHEF electrophoresis (fig. 25). Due to the relatively low resolution of CHEF, including the inability to resolved chromosomes 1 and R, it is highly likely that generation of such structural changes was more efficient than results initially indicated. Indeed, analyses of long read genome sequencing data from strain 1097 indicate multiple interchromosomal translocations which were not evident during initial screening (fig. 31).

Alignment of the assembled contigs from this 1097 genome sequencing to the reference genome identified multiple chromosomal rearrangements, resulting in at least four chimeric chromosomes. The breakpoints of which all occur at the MRS. Additional events are apparent from analysis of individual reads. These cannot be seen in the alignment due to the diploid nature of the *C. albicans* genome. The full extent of the chromosomal rearrangements which have been generated is therefore uncertain. Overall, these data suggest that this CRISPR-Cas9 technique can be used to rapidly generate diverse karyotypes, which facilitates the study of their effect on *Candida* host adaptation and acquired drug resistance.

As there appears to be no CNVs in strain 1097 (fig. 30), all genomic rearrangements are assumed to be reciprocal. This system therefore provides a means of studying the phenotypic effect of structural rearrangements without generating CNVs which have thus far been very difficult. Further, as repair has

occurred at the MRS, it seems likely that coding sequences will not have been altered; the only gene in the MRS is 2.4 kb from the cut site. Therefore, observed phenotypes are due solely to rearrangement of the genetic material. As this strain had a distinct phenotype from the WT, this study provides evidence that rearranging the genome can generate phenotypic diversity, independent of CNVs and/or alterations to coding sequences. Such an effect would account for the high frequency of karyotype changes observed in clinical isolates and following passage of *C. albicans* though murine systems (Forche et al., 2018, 2009).

Conversely, analysis of the strain 1098 genome indicated that the action of Cas9 resulted in an inverted duplication on Chr3. This demonstrates that the CRISPR-Cas9 technique used in this study can generate a diverse range of structural variants. This strain did not have a significantly divergent phenotype. As the effect of CNVs depends largely upon the genes they impact, further work is needed to characterize the genes which have been duplicated in this strain. Generally, *Candida* is highly amenable to CNVs, which are well studied in the organism and were not the primary focus of the present study.

This study also supports the role of the MRS in facilitating genome plasticity as these elements were sufficiently homologous to serve as endogenous repair templates, leading to gross chromosomal rearrangement. Furthermore, rearrangements specifically at these loci were able to generate phenotypic diversity. This not only suggests that reciprocal translocations are an important evolutionary mechanism for *Candida* but supports the hypothesized role of the MRS is facilitating these events.

Initial phenotyping of the strains in this study was limited to colony/cell morphology, growth rate and temperature sensitivity. This however identified that some strains with atypical karyotypes had reduced growth rates. Strain 1095 had severely reduced growth under physiological conditions (37°C), which warrants further investigation for clinical application. Hitherto, the karyotype changes in this strain have however not been characterized. Further, two of the strains identified as having aberrant karyotypes

following CRIPSR transformation were initially hyperfilamentous (fig. 23). While filamentation is a key pathogenicity factor this observation is perhaps unsurprising given the existing association between genotoxic stress and polarised growth as well as between insertion into the RPS and hyperfilamentation (Da et al., 2010; Legrand et al., 2019; Uhl et al., 2003).

The CRISPR protocol used in this study (Nguyen et al., 2017) incorporated a FLP recombinase system which facilitated excision of the cassette from the genome with minimal alteration to the *HIS1* locus and no permanent markers. Intriguingly, this was not the case for strain 1095, from which the cassette could not be excised; there was no loss of the *NAT* marker, and the cassette was subsequently identified in the genome sequencing (data not shown). Further analysis will be conducted to understand this phenomenon, but it is assumed that either the FLP recombinase or its targets were mutated.

11.2 Targeting the MRS with Cas9 Models Increased Genome Instability and can Affect Adaptation to Stress

The evolution experiment in this study served to estimate the stability of the MRS targeted strains, relative to the WT, as well as their relative ability to adapt to stress. Throughout the experiment, karyotype changes were only observed in the MRS-targeted strains which continually expressed Cas9 and not in any evolved WT (tab. 9). This confirms that the observed rearrangements were due to the action of Cas9, thereby demonstrating that this strategy provides a model of increased genome instability. Importantly, we would eventually expect to see rearrangements in the WT as we are modelling a natural phenomenon, simply at a faster rate. Indeed, further work may be done to estimate the rate of karyotypic change in the WT under the same stresses, by continuing this evolution experiment over a longer time.

A second observation was that karyotype changes were observed more frequently in stronger stress conditions (those which inhibited the growth of SC5314 most) (tab. 10). While more stringent stresses may promote rearrangements, it seems more likely that rearrangement only provide a selective advantage under stress and are otherwise lost from the population by drift. Similar observations have been made of aneuploidies in *C. albicans*, whereby they are generally deleterious under standard conditions but can be beneficial in particular stresses (Yang et al., 2021). Future experiments may therefore use more stringent stresses.

Throughout the experiment, there were few cases where the MRS targeted strains adapted to stress at a faster rate than the WT (tab. 11). This was true even when accounting for the generally reduced fitness of these strain prior to the evolution experiment. There was however an example where strain 1095 had a greater relative increase in fitness compared to the evolved WT following one week of evolution in 12 mM caffeine. This therefore suggests that increased genome instability at the MRS can increase the rate

of evolution. However, further work is needed to support this, which again, would likely benefit from the use of more stringent stresses.

11.3 Limitations

This study was the first time where chromosomal rearrangements have been induced in *C. albicans* using CRISPR, and despite the success of this strategy, there were some limitations to the study which should be considered in subsequent experiments.

Designing the sgRNA for this study, the RB2 subunit of the MRS was chosen as a target. While this sequence is itself non-repetitive, it is adjacent the RPS repeat array. Though chromosomal aberrations may have been generated at higher frequency by targeting the RPS itself, this would have resulted in up to hundreds of target sequences per chromosome. It was decided that this would fragment the genome too much, and generate array expansion/reduction, which is not the aim of the study, and would not be identifiable by CHEF electrophoresis and would therefore confound results.

While screening these strains following CRIPSR transformation, it was not immediately known whether translocations evident from the CHEF were due to the directed action of Cas9 and if they involved the MRS. Subsequent genome sequence alignment of strains 1095 and 1097 to the WT confirmed that all events had involved the MRS and were therefore almost certainly due to Cas9 action. Analysis of individual reads however then indicated the presence of some additional events which may not involve the MRS. This highlights the difficult of working with a diploid strain. Further work is needed to confirm and characterise these rearrangements. Another unlikely possibility is that the strains identified in this study as having aberrant karyotypes, may have less stable genomes by chance, and not due to the action of Cas9. This would account of their abnormal CHEF banding pattern initially, and their increased rate of change during the evolution experiment. This eventuality is almost certainly untrue, but to address these concerns, another evolution experiment will be conducted with the MRS targeted strains from which the CRISPR cassette has been removed (1097 and 1098), although clearly this cannot be done for strain 1095. It would be expected that strains without the cassette have higher genome stability, and

that karyotype changes may occur in these strains at a rate between that of the strains with the cassette integrated and that of the WT.

While we evolved the WT SC5314 in triplicate alongside the MRS-targeted strains to act as a control, this WT had not undergone any transformation. While unlikely, it is possible that integration of the cassette itself influences genome stability and fitness. *HIS1* is however a common locus for cassette integration and should not have such effects (Nguyen et al., 2017). Future evolution experiments may transform WT SC5314 with a version of the cassette lacking the sgRNA sequence to act as a better control.

The nature of this study- generating karyotype changes around the MRS almost at random- also excludes the possibility of generating true biological replicates for the evolution experiment. While this is important to consider, it is an inevitability of this type of experiment and technical replicates are sufficient.

Furthermore, despite the continued expression of Cas9 providing a model of elevated genome instability, this system is not sustainable for long durations. It is likely that upon repair of DSBs, the target sequence or PAM site may be altered or lost. It is unclear how quickly these sites would be exhausted; gene conversion will conserve them in some instances, but there are numerous repair pathways, not all of which will maintain the sequence.

As already mentioned, during the evolution experiment, greater changes were seen in both fitness and karyotype under more stringent stresses: elevate temperatures and 12 mM caffeine. Therefore, future studies will use stronger stresses with the intention of selecting for karyotype changes which may not be beneficial under the current conditions. This will include higher concentrations of fluconazole. Stronger stresses will be identified as those which significantly reduce the growth rate of WT SC5314 or prevent it from growing from higher dilutions of inoculum during spotting assay.

While the elevated genome instability of the MRS targeted strains during the evolution experiment was as expected, the changes in relative fitness were very surprising. Not only did the WT evolved at a greater rate under most conditions, but some MRS targeted strains decreased in fitness throughout. While again, this may be due to insufficiently stringent conditions, it may also signify inconsistency in fitness measurements. Spotting assays were used as a measure of fitness due to their simplicity. However, these are not quantifiable and have relatively low reliability. Particularly the filamentous nature of strain 1095 may have distorted initial OD readings and may confound the spotting assay results. In future, haemocytometry may be used for all strains prior to spotting assay to quantify cell concentrations more reliably.

11.4 Next steps

Despite the issues with the evolution experiment which have already been addressed, this study was successful in using CRIPSR to generate chromosomal rearrangements and increase genome instability in *C. albicans.* Further work will focus initially on fully characterizing the genome sequences of the three MRS-targeted strains; 1095, 1097 and 1098, to completely understand the events which have taken place. Once this is done, we will identify genes which have been affected by these changes to see if they could account for the observed phenotype changes. We will then also repeat/continue the evolution experiment under harsher conditions, including higher concentrations of fluconazole, and addition stresses with a more accurate fitness calculation procedure. Subsequently we plan to delete all MRS copies from the *C. albicans* genome and evolve this *ΔMRS* under stress conditions to better understand its evolutionary significance.

We also intend to generate individual, deliberate translocations, as in Fleiss *et al.*, (2019), by providing a chimeric exogenous repair template. Establishing such a protocol to generate translocations in *Candida* would facilitate the study of their phenotypic effects in a highly controlled setting.

Thereafter we will shift our attention to other repetitive elements in the *C. albicans* genome. While the MRS is particularly interesting due to its uniqueness and size, chromosomal rearrangements are also enriched at other repetitive elements. It would therefore be interesting to compare the effects of chromosomal rearrangements around different repetitive elements. Such data may then translate more to other species which do not have MRS elements.

In this study, we have demonstrated the utility and efficacy of a CRIPSR-Cas9 system to generate chromosomal rearrangements in *C. albicans*. Furthermore, we have alluded to the significance of repeat-facilitated genome instability in fungal pathogen evolution. With this in mind, we may utilise our established CRIPSR system to target the MRS-like elements of *C. dubliniensis* or *C. tropicalis* as well as

other repetitive elements in these, and further fungal pathogens to establish a broader understanding of this evolutionary mechanism.

12 Conclusion

This study highlights the role of repeat-associated chromosomal rearrangements and demonstrates how they can rapidly generate diverse phenotypes and influence evolution. While our focused has been *C. albicans*, this mechanism is ubiquitous throughout eukaryotic life. Chromosomal rearrangements are implicated in adaptation and speciation in animals (Fuller et al., 2019), plants (Wang et al., 2020), and fungi (Shi-Kunne et al., 2018). They are also associated with various human genetic diseases and cancers (Holland and Cleveland, 2012; Jiang et al., 2016; Pellestor, 2019). Equivalent mechanisms of genome reordering are also present in archaea (Ausiannikava et al., 2018), bacteria (Darling et al., 2008) and viruses (Karamitros et al., 2020).

CRISPR-Cas9 technology has revolutionized our ability to study these events and their impact. Methods similar to those used for this study have demonstrated the evolutionary role of chromosomal rearrangements in *Saccharomyces* (Fleiss et al., 2019), *Arabidopsis* (Beying et al., 2020) and *Cryptococcus* (Yadav et al., 2020). The utilization of CRIPSR-Cas9 approaches in mouse embryonic stem cells (Jiang et al., 2016) and human cells (Torres et al., 2014) has also improved our understanding of the role of chromosomal rearrangements in cancer.

The present study further supports the evolutionary role of chromosomal rearrangements in generating novel phenotypes, as well as the utility of CRISPR in their study. This is the first study to apply this approach to a diploid human fungal pathogen and to the authors best knowledge, it the first instance where Cas9 has been continually expressed to generate rearrangements throughout the course of an evolution experiment. This is a powerful tool which could help us to understand the effects of increased genome instability throughout the natural world.

13 References

Alby, K., Schaefer, D., Bennett, R.J., 2009. Homothallic and heterothallic mating in the opportunistic pathogen *Candida albicans*. Nat. 2009 4607257 460, 890–893.

Allen, C.M., Beck, F.M., 1983. Strain-Related Differences in Pathogenicity of *Candida albicans* for Oral Mucosa. J. Infect. Dis. 147, 1036–1040.

Anderson, M.Z., Wigen, L.J., Burrack, L.S., Berman, J., 2015. Real-time evolution of a subtelomeric gene family in *Candida albicans*. Genetics 200, 907–919.

Anderson, T.M., Clay, M.C., Cioffi, A.G., Diaz, K.A., Hisao, G.S., Tuttle, M.D., Nieuwkoop, A.J., Comellas, G., Maryum, N., Wang, S., Uno, B.E., Wildeman, E.L., Gonen, T., Rienstra, C.M., Burke, M.D., 2014. Amphotericin forms an extramembranous and fungicidal sterol sponge. Nat. Chem. Biol. 2014 105 10, 400–406.

Andes, D., Nett, J., Oschel, P., Albrecht, R., Marchillo, K., Pitula, A., 2004. Development and characterization of an in vivo central venous catheter *Candida albicans* biofilm model. Infect. Immun. 72, 6023–6031.

Andes, D.R., Group, for the M.S., Safdar, N., Group, for the M.S., Baddley, J.W., Group, for the M.S., Playford, G., Group, for the M.S., Reboli, A.C., Group, for the M.S., Rex, J.H., Group, for the M.S., Sobel, J.D., Group, for the M.S., Pappas, P.G., Group, for the M.S., Kullberg, B.J., Group, for the M.S., 2012. Impact of Treatment Strategy on Outcomes in Patients with Candidemia and Other Forms of Invasive Candidiasis: A Patient-Level Quantitative Review of Randomized Trials. Clin. Infect. Dis. 54, 1110–1122.

Annaluru, N., Muller, H., Mitchell, L.A., Ramalingam, S., Stracquadanio, G., Richardson, S.M., Dymond,

J.S., Kuang, Z., Scheifele, L.Z., Cooper, E.M., Cai, Y., Zeller, K., Agmon, N., Han, J.S., Hadjithomas, M., Tullman, J., Caravelli, K., Cirelli, K., Guo, Z., London, V., Yeluru, A., Murugan, S., Kandavelou, K., Agier, N., Fischer, G., Yang, K., Martin, J.A., Bilgel, M., Bohutski, P., Boulier, K.M., Capaldo, B.J., Chang, J., Charoen, K., Choi, W.J., Deng, P., DiCarlo, J.E., Doong, J., Dunn, J., Feinberg, J.I., Fernandez, C., Floria, C.E., Gladowski, D., Hadidi, P., Ishizuka, I., Jabbari, J., Lau, C.Y.L., Lee, P.A., Li, S., Lin, D., Linder, M.E., Ling, J., Liu, Jaime, Liu, Jonathan, London, M., Henry, M., Mao, J., McDade, J.E., McMillan, A., Moore, A.M., Oh, W.C., Ouyang, Y., Patel, R., Paul, M., Paulsen, L.C., Qiu, J., Rhee, A., Rubashkin, M.G., Soh, I.Y., Sotuyo, N.E., Srinivas, V., Suarez, A., Wong, A., Wong, R., Xie, W.R., Xu, Y., Yu, A.T., Koszul, R., Bader, J.S., Boeke, J.D., Chandrasegaran, S., 2014. Total synthesis of a functional designer eukaryotic chromosome. Science (80-.). 344, 55–58.

Argueso, J.L., Westmoreland, J., Mieczkowski, P.A., Gawel, M., Petes, T.D., Resnick, M.A., 2008. Double-strand breaks associated with repetitive DNA can reshape the genome. Proc. Natl. Acad. Sci. 105, 11845–11850.

Astvad, K.M.T., Johansen, H.K., Røder, B.L., Rosenvinge, F.S., Knudsen, J.D., Lemming, L., Schønheyder, H.C., Hare, R.K., Kristensen, L., Nielsen, L., Gertsen, J.B., Dzajic, E., Pedersen, M., Østergård, C., Olesen, B., Søndergaard, T.S., Arendrup, M.C., 2018. Update from a 12-year nationwide fungemia surveillance: Increasing intrinsic and acquired resistance causes concern. J. Clin. Microbiol. 56.

Ausiannikava, D., Mitchell, L., Marriott, H., Smith, V., Hawkins, M., Makarova, K.S., Koonin, E. V., Nieduszynski, C.A., Allers, T., 2018. Evolution of Genome Architecture in Archaea: Spontaneous Generation of a New Chromosome in *Haloferax volcanii*. Mol. Biol. Evol. 35, 1855–1868.

Avramovska, O., Hickman, M.A., 2019. The Magnitude of *Candida albicans* Stress-Induced Genome Instability Results from an Interaction Between Ploidy and Antifungal Drugs. G3 Genes|Genomes|Genetics 9, 4019–4027.

Barnett, J.A., 2008. A history of research on yeasts 12: medical yeasts part 1, *Candida albicans*. Yeast 25, 385–417.

Becks, L., Agrawal, A.F., 2010. Higher rates of sex evolve in spatially heterogeneous environments. Nat. 2010 4687320 468, 89–92.

Benedict, K., Jackson, B.R., Chiller, T., Beer, K.D., 2019. Estimation of Direct Healthcare Costs of Fungal Diseases in the United States. Clin. Infect. Dis. 68, 1791–1797.

Bennett, R.J., Uhl, M.A., Miller, M.G., Johnson, A.D., 2003. Identification and Characterization of a *Candida albicans* Mating Pheromone . Mol. Cell. Biol. 23, 8189–8201.

Bensasson, D., Dicks, J., Ludwig, J.M., Bond, C.J., Elliston, A., Roberts, I.N., James, S.A., 2019. Diverse Lineages of *Candida albicans* Live on Old Oaks. Genetics 211, 277–288.

Bertolini, M., Dongari-Bagtzoglou, A., 2019. The Relationship of *Candida albicans* with the Oral Bacterial Microbiome in Health and Disease. Adv. Exp. Med. Biol. 1197, 69–78.

Beying, N., Schmidt, C., Pacher, M., Houben, A., Puchta, H., 2020. CRISPR–Cas9-mediated induction of heritable chromosomal translocations in *Arabidopsis*. Nat. Plants 2020 66 6, 638–645.

Bio Rad, 1992. Pulsed Field Electrophoresis Systems Instruction Manual and Applications Guide Catalog Numbers, Electrophoresis.

Braunsdorf, C., Leibundgut-Landmann, S., 2018. Modulation of the Fungal-Host Interaction by the Intra-Species Diversity of *C. albicans*. Pathog. 2018, Vol. 7, Page 11 7, 11.

Buscaino, A., 2019. Chromatin-mediated regulation of genome plasticity in human fungal pathogens. Genes. 10, 855.

Butler, G., Rasmussen, M.D., Lin, M.F., Santos, M.A.S., Sakthikumar, S., Munro, C.A., Rheinbay, E.,

Grabherr, M., Forche, A., Reedy, J.L., Agrafioti, I., Arnaud, M.B., Bates, S., Brown, A.J.P., Brunke, S., Costanzo, M.C., Fitzpatrick, D.A., De Groot, P.W.J., Harris, D., Hoyer, L.L., Hube, B., Klis, F.M., Kodira, C., Lennard, N., Logue, M.E., Martin, R., Neiman, A.M., Nikolaou, E., Quail, M.A., Quinn, J., Santos, M.C., Schmitzberger, F.F., Sherlock, G., Shah, P., Silverstein, K.A.T., Skrzypek, M.S., Soll, D., Staggs, R., Stansfield, I., Stumpf, M.P.H., Sudbery, P.E., Srikantha, T., Zeng, Q., Berman, J., Berriman, M., Heitman, J., Gow, N.A.R., Lorenz, M.C., Birren, B.W., Kellis, M., Cuomo, C.A., 2009. Evolution of pathogenicity and sexual reproduction in eight *Candida* genomes. Nat. 2009 4597247 459, 657–662.

Chandra, J., Kuhn, D.M., Mukherjee, P.K., Hoyer, L.L., McCormick, T., Ghannoum, M.A., 2001. Biofilm formation by the fungal pathogen *Candida albicans*: Development, architecture, and drug resistance. J. Bacteriol. 183, 5385–5394.

Chandra, J., McCormick, T.S., Imamura, Y., Mukherjee, P.K., Ghannoum, M.A., 2007. Interaction of *Candida albicans* with adherent human peripheral blood mononuclear cells increases *C. albicans* biofilm formation and results in differential expression of pro- and anti-inflammatory cytokines. Infect. Immun. 75, 2612–2620.

Chapman, B., Slavin, M., Marriott, D., Halliday, C., Kidd, S.E., Arthur, I., Bak, N., Heath, C., Kennedy, K., Morrissey, C.O., Sorrell, T., van Hal, S., Keighley, C., Goeman, E., Underwood, N., Hajkowicz, K., Hofmeyr, A., Leung, M., Macesic, N., Botes, J., Blyth, C., Cooley, L., George, C.R., Kalukottege, R.P., Kesson, A., McMullan, B., Baird, R., Robson, J., Korman, T.M., Pendle, S., Weeks, K., Liu, E., Cheong, E., Chen, S., Daveson, K., Wilson, H., Thomas, M., Peachey, G., Gottlieb, T., Lui, E., Jozwiak, F., Bell, S.M., Kotsiou, G., Chen, S.C.A., Marshall, C., Ferguson, C., Bryant, M., Reed, C., Darvall, J., Maddigan, V., Boan, P., Menon, V., 2017. Changing epidemiology of *Candida*emia in Australia. J. Antimicrob. Chemother. 72, 1103–1108.

Chibana, H., Beckerman, J.L., Magee, P.T., 2000. Fine-Resolution Physical Mapping of Genomic Diversity in *Candida albicans*. Genome Res. 10, 1865–1877.

Chibana, H., Iwaguchi, S.I., Homma, M., Chindamporn, A., Nakagawa, Y., Tanaka, K., 1994. Diversity of tandemly repetitive sequences due to short periodic repetitions in the chromosomes of *Candida albicans*. J. Bacteriol. 176, 3851–3858.

Chibana, H., Magee, P.T., 2009. The enigma of the major repeat sequence of *Candida albicans* 171–179.

Chin, V.K., Foong, K.J., Maha, A., Rusliza, B., Norhafizah, M., Chong, P.P., 2014. Multi-Step Pathogenesis and Induction of Local Immune Response by Systemic *Candida albicans* Infection in an Intravenous Challenge Mouse Model. Int. J. Mol. Sci. 2014, Vol. 15, Pages 14848-14867 15, 14848–14867.

Chindamporn, A., Nakagawa, Y., Mizuguchi, I., Chibana, H., Doi, M., Tanaka, K., 1998. Repetitive sequences (RPSs) in the chromosomes of *Candida albicans* are sandwiched between two novel stretches, HOK and RB2, common to each chromosome. Microbiology 144, 849–857.

Chu, W.-S., Magee, B.B., Magee, P.T., Sadhu, C., Mceachern, M.J., Rustchenko-Bulgac, E.P., Schmid, J., Soll, D.R., Hicks, J.B., Bacteriol, J., 1993. Construction of an SfiI macrorestriction map of the *Candida albicans* genome. J. Bacteriol. 175, 6637–6651.

Ciudad, T., Andaluz, E., Steinberg-Neifach, O., Lue, N.F., Gow, N.A.R., Calderone, R.A., Larriba, G., 2004. Homologous recombination in *Candida albicans*: role of CaRad52p in DNA repair, integration of linear DNA fragments and telomere length. Mol. Microbiol. 53, 1177–1194.

Cleveland, A.A., Harrison, L.H., Farley, M.M., Hollick, R., Stein, B., Chiller, T.M., Lockhart, S.R., Park, B.J., 2015. Declining Incidence of Candidemia and the Shifting Epidemiology of *Candida* Resistance in Two US Metropolitan Areas, 2008–2013: Results from Population-Based Surveillance. PLoS One 10, e0120452.

Coco, B.J., Bagg, J., Cross, L.J., Jose, A., Cross, J., Ramage, G., 2008. Mixed *Candida albicans* and *Candida glabrata* populations associated with the pathogenesis of denture stomatitis. Oral Microbiol. Immunol. 23, 377–383.

Coste, A., Turner, V., Ischer, F., Morschhäuser, J., Forche, A., Selmecki, A., Berman, J., Bille, J., Sanglard, D., 2006. A Mutation in Tac1p, a Transcription Factor Regulating CDR1 and CDR2, Is Coupled With Loss of Heterozygosity at Chromosome 5 to Mediate Antifungal Resistance in *Candida albicans*. Genetics 172, 2139–2156.

Costeton, J.W., Lewandowski, Z., Caldwell, D.E., Korber, D.R., Appin-Scott, H.M., 1995. Microbial biofilms. Annu. Rev. Microbiol. 49, 711–745.

Da, A., Dantas, S., Patterson, M.J., Smith, D.A., Maccallum, D.M., Erwig, L.P., Morgan, B.A., Quinn, J., 2010. Thioredoxin Regulates Multiple Hydrogen Peroxide-Induced Signaling Pathways in *Candida albicans*. Mol. Cell. Biol. 30, 4550–4563.

da Silva Dantas, A., Lee, K.K., Raziunaite, I., Schaefer, K., Wagener, J., Yadav, B., Gow, N.A., 2016. Cell biology of *Candida albicans*–host interactions. Curr. Opin. Microbiol. 34, 111–118.

Dalle, F., Wächtler, B., L'Ollivier, C., Holland, G., Bannert, N., Wilson, D., Labruère, C., Bonnin, A., Hube, B., 2010. Cellular interactions of *Candida albicans* with human oral epithelial cells and enterocytes. Cell. Microbiol. 12, 248–271.

Darling, A.E., Miklós, I., Ragan, M.A., 2008. Dynamics of Genome Rearrangement in Bacterial Populations. PLOS Genet. 4, e1000128.

Doi, A.M., Pignatari, A.C.C., Edmond, M.B., Marra, A.R., Camargo, L.F.A., Siqueira, R.A., Da Mota, V.P., Colombo, A.L., 2016. Epidemiology and Microbiologic Characterization of Nosocomial Candidemia from a Brazilian National Surveillance Program. PLoS One 11, e0146909.

Douglas, L.J., 2003. *Candida* biofilms and their role in infection. Trends Microbiol. 11, 30–36.

Drinnenberg, I.A., Weinberg, D.E., Xie, K.T., Mower, J.P., Wolfe, K.H., Fink, G.R., Bartel, D.P., 2009. RNAi in budding yeast. Science (80-.). 326, 544–550.

Dunkel, N., Blaß, J., Rogers, P.D., Morschhäuser, J., 2008. Mutations in the multi-drug resistance regulator MRR1, followed by loss of heterozygosity, are the main cause of MDR1 overexpression in fluconazole-resistant *Candida albicans* strains. Mol. Microbiol. 69, 827–840.

Dunn, M.J., Anderson, M.Z., 2019. To Repeat or Not to Repeat : Repetitive Sequences Regulate Genome Stability in *Candida albicans*.

Ellepola, K., Liu, Y., Cao, T., Koo, H., Seneviratne, C.J., 2017. Bacterial GtfB Augments *Candida albicans* Accumulation in Cross-Kingdom Biofilms. J. Dent. Res. 96, 1129–1135.

Ene, I. V., Farrer, R.A., Hirakawa, M.P., Agwamba, K., Cuomo, C.A., Bennett, R.J., 2018. Global analysis of mutations driving microevolution of a heterozygous diploid fungal pathogen. Proc. Natl. Acad. Sci. U. S. A. 115, E8688–E8697.

Erlendson, A.A., Friedman, S., Freitag, M., 2017. A Matter of Scale and Dimensions: Chromatin of Chromosome Landmarks in the Fungi. Microbiol. Spectr. 5.

Espinel-Ingroff, A., Pfaller, M.A., Bustamante, B., Canton, E., Fothergill, A., Fuller, J., Gonzalez, G.M., Lass-Flörl, C., Lockhart, S.R., Martin-Mazuelos, E., Meis, J.F., Melhem, M.S.C., Ostrosky-Zeichner, L., Pelaez, T., Szeszs, M.W., St-Germain, G., Bonfietti, L.X., Guarro, J., Turnidge, J., 2014. Multilaboratory study of epidemiological cutoff values for detection of resistance in eight *Candida* species to fluconazole, posaconazole, and voriconazole. Antimicrob. Agents Chemother. 58, 2006–2012.

Fasullot, M.T., Davis, R.W., 1988. Direction of chromosome rearrangements in *Saccharomyces cerevisiae*

by use of his3 recombinational substrates. Mol. Cell. Biol. 8, 4370–4380.

Feri, A., Loll-Krippleber, R., Commere, P.H., Maufrais, C., Sertour, N., Schwartz, K., Sherlock, G., Bougnoux, M.E., d'Enfert, C., Legrand, M., 2016. Analysis of repair mechanisms following an induced double-strand break uncovers recessive deleterious alleles in the *Candida albicans* diploid genome. MBio 7.

Ferrari, S., Sanguinetti, M., Torelli, R., Posteraro, B., Sanglard, D., 2011. Contribution of CgPDR1-regulated genes in enhanced virulence of azole-resistant *Candida glabrata*. PLoS One 6.

Fischer, G., James, S.A., Roberts, I.N., Oliver, S.G., Louis, E.J., 2000. Chromosomal evolution in *Saccharomyces*. Nat. 2000 4056785 405, 451–454.

Fischer, G., Rocha, E.P.C., Brunet, F., Vergassola, M., Dujon, B., 2006. Highly Variable Rates of Genome Rearrangements between Hemiascomycetous Yeast Lineages. PLOS Genet. 2, e32.

Fleiss, A., O'Donnell, S., Fournier, T., Lu, W., Agier, N., Delmas, S., Schachere, J., Fischer, G., 2019. Reshuffling yeast chromosomes with CRISPR Cas9. PLoS Genet. 15, 1–26.

Flowers, S.A., Barker, K.S., Berkow, E.L., Toner, G., Chadwick, S.G., Gygax, S.E., Morschhäuser, J., David Rogers, P., 2012. Gain-of-function mutations in UPC2 are a frequent cause of ERG11 upregulation in azole-resistant clinical isolates of *Candida albicans*. Eukaryot. Cell 11, 1289–1299.

Forche, A., Abbey, D., Pisithkul, T., Weinzierl, M.A., Ringstrom, T., Bruck, D., Petersen, K., Berman, J., 2011. Stress alters rates and types of loss of heterozygosity in *Candida albicans*. MBio 2.

Forche, A., Alby, K., Schaefer, D., Johnson, A.D., Berman, J., Bennett, R.J., 2008. The Parasexual Cycle in *Candida albicans* Provides an Alternative Pathway to Meiosis for the Formation of Recombinant Strains. PLOS Biol. 6, e110.

Forche, A., Cromie, G., Gerstein, A.C., Solis, N. V., Pisithkul, T., Srifa, W., Jeffery, E., Abbey, D., Filler, S.G., Dudley, A.M., Berman, J., 2018. Rapid Phenotypic and Genotypic Diversification After Exposure to the Oral Host Niche in *Candida albicans*. Genetics 209, 725–741.

Forche, A., Magee, P.T., Selmecki, A., Berman, J., May, G., 2009. Evolution in *Candida albicans* Populations During a Single Passage Through a Mouse Host. Genetics 182, 799–811.

Fox, E.P., Nobile, C.J., 2012. A sticky situation. Untangling the transcriptional network controlling biofilm development in *Candida albicans* . Transcr. 3, 315–322.

Freire-Benéitez, V., Gourlay, S., Berman, J., Buscaino, A., 2016a. Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen *Candida albicans*. Nucleic Acids Res. 44, 9166–9179.

Freire-Benéitez, V., Price, R.J., Tarrant, D., Berman, J., Buscaino, A., 2016b. *Candida albicans* repetitive elements display epigenetic diversity and plasticity. Sci. Reports 2016 61 6, 1–12.

Frenkel, M., Mandelblat, M., Alastruey-Izquierdo, A., Mendlovic, S., Semis, R., Segal, E., 2016. Pathogenicity of *Candida albicans* isolates from bloodstream and mucosal candidiasis assessed in mice and *Galleria mellonella*. J. Mycol. Med. 26, 1–8.

Froyd, C.A., Kapoor, S., Dietrich, F., Rusche, L.N., 2013. The Deacetylase Sir2 from the Yeast Clavispora lusitaniae Lacks the Evolutionarily Conserved Capacity to Generate Subtelomeric Heterochromatin. PLOS Genet. 9, e1003935.

Fuchs, B.B., Mylonakis, E., 2009. Our Paths Might Cross: the Role of the Fungal Cell Wall Integrity Pathway in Stress Response and Cross Talk with Other Stress Response Pathways. Eukaryot. Cell 8, 1616.

Fuller, Z.L., Koury, S.A., Phadnis, N., Schaeffer, S.W., 2019. How chromosomal rearrangements shape

adaptation and speciation: Case studies in *Drosophila pseudoobscura* and its sibling species *Drosophila persimilis*. Mol. Ecol. 28, 1283–1301.

Galhardo, R.S., Hastings, P.J., Rosenberg, S.M., 2008. Mutation as a Stress Response and the Regulation of Evolvability. http://dx.doi.org/10.1080/10409230701648502 42, 399–435.

Gerstein, A.C., Lim, H., Berman, J., Hickman, M.A., 2017. Ploidy tug-of-war: Evolutionary and genetic environments influence the rate of ploidy drive in a human fungal pathogen. Evolution (N. Y). 71, 1025–1038.

Ghannoum, M.A., Jurevic, R.J., Mukherjee, P.K., Cui, F., Sikaroodi, M., Naqvi, A., Gillevet, P.M., 2010. Characterization of the Oral Fungal Microbiome (Mycobiome) in Healthy Individuals. PLOS Pathog. 6, e1000713.

Gomes, A.C., Miranda, I., Silva, R.M., Moura, G.R., Thomas, B., Akoulitchev, A., Santos, M.A.S., 2007. A genetic code alteration generates a proteome of high diversity in the human pathogen *Candida albicans*. Genome Biol. 8, 1–15.

Greenberg, R.G., Benjamin, D.K., 2014. Neonatal candidiasis: Diagnosis, prevention, and treatment. J. Infect. 69, S19–S22.

Guinea, J., 2014. Global trends in the distribution of *Candida* species causing candidemia. Clin. Microbiol. Infect. 20, 5–10.

Gulati, M., Nobile, C.J., 2016. *Candida albicans* biofilms: development, regulation, and molecular mechanisms. Microbes Infect. 18, 310.

Hamperl, S., Wittner, M., Babl, V., Perez-Fernandez, J., Tschochner, H., Griesenbeck, J., 2013. Chromatin states at ribosomal DNA loci. Biochim. Biophys. Acta - Gene Regul. Mech. 1829, 405–417.

Haran, J., Boyle, H., Hokamp, K., Yeomans, T., Liu, Z., Church, M., Fleming, A.B., Anderson, M.Z., Berman, J., Myers, L.C., Sullivan, D.J., Moran, G.P., 2014. Telomeric ORFs (TLOs) in *Candida* spp. Encode Mediator Subunits That Regulate Distinct Virulence Traits. PLOS Genet. 10, e1004658.

Hickman, M.A., Zeng, G., Forche, A., Hirakawa, M.P., Abbey, D., Harrison, B.D., Wang, Y.M., Su, C.H., Bennett, R.J., Wang, Y., Berman, J., 2013. The 'obligate diploid' *Candida albicans* forms mating-competent haploids. Nat. 2013 4947435 494, 55–59.

Hirakawa, M.P., Chyou, D.E., Huang, D., Slan, A.R., Bennett, R.J., 2017. Parasex generates phenotypic diversity de novo and impacts drug resistance and virulence in *Candida albicans*. Genetics 207, 1195–1211.

Hirakawa, M.P., Martinez, D.A., Sakthikumar, S., Anderson, M.Z., Berlin, A., Gujja, S., Zeng, Q., Zisson, E., Wang, J.M., Greenberg, J.M., Berman, J., Bennett, R.J., Cuomo, C.A., 2015. Genetic and phenotypic intra-species variation in *Candida albicans*. Genome Res. 25, 413–425.

Hochrein, L., Mitchell, L.A., Schulz, K., Messerschmidt, K., Mueller-Roeber, B., 2018. L-SCRaMbLE as a tool for light-controlled Cre-mediated recombination in yeast. Nat. Commun. 2018 91 9, 1–10.

Holland, A.J., Cleveland, D.W., 2012. Chromoanagenesis and cancer: mechanisms and consequences of localized, complex chromosomal rearrangements. Nat. Med. 2012 1811 18, 1630–1638.

Hoot, S.J., Smith, A.R., Brown, R.P., White, T.C., 2011. An A643V amino acid substitution in Upc2p contributes to azole resistance in well-characterized clinical isolates of *Candida albicans*. Antimicrob. Agents Chemother. 55, 940–942.

Hull, C.M., Raisner, R.M., Johnson, A.D., 2000. Evidence for mating of the "asexual" yeast *Candida albicans* in a mammalian host. Science (80-.). 289, 307–310.

Hwang, G., Marsh, G., Gao, L., Waugh, R., Koo, H., 2015. Binding Force Dynamics of *Streptococcus*

mutans -glucosyltransferase B to *Candida albicans*. J. Dent. Res. 94, 1310–1317.

Iwaguchi, S.I., Homma, M., Chibana, H., Tanaka, K., 1992. Isolation and characterization of a repeated sequence (RPS1) of *Candida albicans*. J. Gen. Microbiol. 138, 1893–1900.

Iwaguchi, S.I., Suzuki, M., Sakai, N., Nakagawa, Y., Magee, P.T., Suzuki, T., 2004. Chromosome translocation induced by the insertion of the URA blaster into the major repeat sequence (MRS) in *Candida albicans*. Yeast 21, 619–634.

Jackson, A.P., Gamble, J.A., Yeomans, T., Moran, G.P., Saunders, D., Harris, D., Aslett, M., Barrell, J.F., Butler, G., Citiulo, F., Coleman, D.C., De Groot, P.W.J., Goodwin, T.J., Quail, M.A., McQuillan, J., Munro, C.A., Pain, A., Poulter, R.T., Rajandream, M.A., Renauld, H., Spiering, M.J., Tivey, A., Gow, N.A.R., Barrell, B., Sullivan, D.J., Berriman, M., 2009. Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. Genome Res. 19, 2231–2244.

Jacobsen, I.D., Wilson, D., Wächtler, B., Brunke, S., Naglik, J.R., Hube, B., 2014. *Candida albicans* dimorphism as a therapeutic target. http://dx.doi.org/10.1586/eri.11.152 10, 85–93.

Janssen, A., Colmenares, S.U., Karpen, G.H., 2018. Heterochromatin: Guardian of the Genome. Annu. Rev. Cell Dev. Biol. 34, 265–288.

Jiang, J., Zhang, L., Zhou, X., Chen, X., Huang, G., Li, F., Wang, R., Wu, N., Yan, Y., Tong, C., Srivastava, S., Wang, Y., Liu, H., Ying, Q.L., 2016. Induction of site-specific chromosomal translocations in embryonic stem cells by CRISPR/Cas9. Sci. Reports 2016 61 6, 1–9.

Joly, S., Pujol, C., Rysz, M., Vargas, K., Soll, D.R., 1999. Development and characterization of complex DNA fingerprinting probes for the infectious yeast *Candida dubliniensis*. J. Clin. Microbiol. 37, 1035–1044.

Joly, S., Pujol, C., Soll, D.R., 2002. Microevolutionary changes and chromosomal translocations are more

frequent at RPS loci in *Candida dubliniensis* than in *Candida albicans*. Infect. Genet. Evol. 2, 19–37.

Karamitros, T., Papadopoulou, G., Bousali, M., Mexias, A., Tsiodras, S., Mentis, A., 2020. SARS-CoV-2 exhibits intra-host genomic plasticity and low-frequency polymorphic quasispecies. J. Clin. Virol. 131, 104585.

Katragkou, A., Kruhlak, M.J., Simitsopoulou, M., Chatzimoschou, A., Taparkou, A., Cotten, C.J., Paliogianni, F., Diza-Mataftsi, E., Tsantali, C., Walsh, T.J., Roilides, E., 2010. Interactions between Human Phagocytes and *Candida albicans* Biofilms Alone and in Combination with Antifungal Agents. J. Infect. Dis. 201, 1941–1949.

Keeney, S., 2008. Spo11 and the formation of DNA double-strand breaks in meiosis. Genome Dyn. Stab. 2, 81–123.

Kelly, S.L., Lamb, D.C., Kelly, D.E., Manning, N.J., Loeffler, J., Hebart, H., Schumacher, U., Einsele, H., 1997. Resistance to fluconazole and cross-resistance to amphotericin B in *Candida albicans* from AIDS patients caused by defective sterol Δ5,6-desaturation. FEBS Lett. 400, 80–82.

Khoury, Z.H., Vila, T., Puthran, T.R., Sultan, A.S., Montelongo-Jauregui, D., Melo, M.A.S., Jabra-Rizk, M.A., 2020. The Role of *Candida albicans* Secreted Polysaccharides in Augmenting *Streptococcus mutans* Adherence and Mixed Biofilm Formation: In vitro and in vivo Studies. Front. Microbiol. 11, 307.

Kojic, E.M., Darouiche, R.O., 2004. *Candida* Infections of Medical Devices. Clin. Microbiol. Rev. 17, 255–267.

Kullberg, B.J., Arendrup, M.C., 2015. Invasive Candidiasis 1445–1456.

Lafayette, S.L., Collins, C., Zaas, A.K., Schell, W.A., Betancourt-Quiroz, M., Leslie Gunatilaka, A.A., Perfect, J.R., Cowen, L.E., 2010. PKC Signaling Regulates Drug Resistance of the Fungal Pathogen *Candida albicans* via Circuitry Comprised of Mkc1, Calcineurin, and Hsp90. PLOS Pathog. 6, e1001069.

Lee, Y., Puumala, E., Robbins, N., Cowen, L.E., 2021. Antifungal Drug Resistance: Molecular Mechanisms in *Candida albicans* and beyond. Chem. Rev. 121, 3390–3411.

Legrand, M., Chan, C.L., Jauert, P.A., Kirkpatrick, D.T., 2007. Role of DNA mismatch repair and double-strand break repair in genome stability and antifungal drug resistance in *Candida albicans*. Eukaryot. Cell 6, 2194–2205.

Legrand, M., Jaitly, P., Feri, A., d'Enfert, C., Sanyal, K., 2019. *Candida albicans*: An Emerging Yeast Model to Study Eukaryotic Genome Plasticity. Trends Genet. 35, 292–307.

Legrand, M., Lephart, P., Forche, A., Mueller, F.M.C., Walsh, T., Magee, P.T., Magee, B.B., 2004. Homozygosity at the MTL locus in clinical strains of *Candida albicans*: karyotypic rearrangements and tetraploid formation†. Mol. Microbiol. 52, 1451–1462.

Lephart, Paul R, Chibana, H., Magee, P.T., 2005. Effect of the Major Repeat Sequence on Chromosome Loss in *Candida albicans*. Eukaryot. Cell 4.

Lephart, Paul R., Chibana, H., Magee, P.T., 2005. Effect of the major repeat sequence on chromosome loss in *Candida albicans*. Eukaryot. Cell 4, 733–741.

Lephart, P.R., Magee, P.T., 2006. Effect of the Major Repeat Sequence on Mitotic Recombination in *Candida albicans*. Genetics 174, 1737–1744.

Letscher-Bru, V., Herbrecht, R., 2003. Caspofungin: the first representative of a new antifungal class. J. Antimicrob. Chemother. 51, 513–521.

Li, W.S., Chen, Y.C., Kuo, S.F., Chen, F.J., Lee, C.H., 2018. The impact of biofilm formation on the persistence of candidemia. Front. Microbiol. 9, 1196.

Li, X., Yan, Z., Xu, J., 2003. Quantitative variation of biofilms among strains in natural populations of

Candida albicans. Microbiology 149, 353–362.

Liu, Z., Moran, G.P., Sullivan, D.J., MacCallum, D.M., Myers, L.C., 2016. Amplification of TLO Mediator Subunit Genes Facilitate Filamentous Growth in *Candida Spp*. PLOS Genet. 12, e1006373.

Lockhart, S.R., Etienne, K.A., Vallabhaneni, S., Farooqi, J., Chowdhary, A., Govender, N.P., Colombo, A.L., Calvo, B., Cuomo, C.A., Desjardins, C.A., Berkow, E.L., Castanheira, M., Magobo, R.E., Jabeen, K., Asghar, R.J., Meis, J.F., Jackson, B., Chiller, T., Litvintseva, A.P., 2017. Simultaneous emergence of multidrug-resistant *Candida auris* on 3 continents confirmed by whole-genome sequencing and epidemiological analyses. Clin. Infect. Dis. 64, 134–140.

Lockhart, S.R., Pujol, C., Daniels, K.J., Miller, M.G., Johnson, A.D., Pfaller, M.A., Soil, D.R., 2002. In *Candida albicans*, White-Opaque Switchers Are Homozygous for Mating Type. Genetics 162, 737–745.

Lortholary, O., Renaudat, C., Sitbon, K., Desnos-Ollivier, M., Bretagne, S., Dromer, F., 2017. The risk and clinical outcome of candidemia depending on underlying malignancy. Intensive Care Med. 43, 652–662.

Lundblad, V., Blackburn, E.H., 1993. An alternative pathway for yeast telomere maintenance rescues est1- senescence. Cell 73, 347–360.

MacCallum, D.M., Castillo, L., Nather, K., Munro, C.A., Brown, A.J.P., Gow, N.A.R., Odds, F.C., 2009. Property differences among the four major *Candida albicans* strain clades. Eukaryot. Cell 8, 373–387.

Madireddy, A., Gerhardt, J., 2017. Replication through repetitive DNA elements and their role in human diseases. Adv. Exp. Med. Biol. 1042, 549–581.

Magee,', B.B., Koltin, Y., Gorman, J.A., Magee', A.P.T., Kline, S., Laboratories, F., 1988. Assignment of

cloned genes to the seven electrophoretically separated *Candida albicans* chromosomes. Mol. Cell. Biol. 8, 4721–4726.

Magee, B.B., Magee, P.T., 1987. Electrophoretic karyotypes and chromosome numbers in *Candida* species. J. Gen. Microbiol. 133, 425–430.

Magee, B.B., Sanchez, M.D., Saunders, D., Harris, D., Berriman, M., Magee, P.T., 2008. Extensive chromosome rearrangements distinguish the karyotype of the hypovirulent species *Candida dubliniensis* from the virulent *Candida albicans*. Fungal Genet. Biol. 45, 338–350.

Maguire, S.L., ÓhÉigeartaigh, S.S., Byrne, K.P., Schröder, M.S., O'Gaora, P., Wolfe, K.H., Butler, G., 2013. Comparative Genome Analysis and Gene Finding in *Candida* Species Using CGOB. Mol. Biol. Evol. 30, 1281–1291.

Marton, T., Chauvel, M., Feri, A., Maufrais, C., D'enfert, C., Legrand, M., 2021. Factors that influence bidirectional long-tract homozygosis due to double-strand break repair in *Candida albicans*. Genetics 218.

Mathé, L., Van Dijck, P., 2013. Recent insights into *Candida albicans* biofilm resistance mechanisms. Curr. Genet. 59, 251–264.

Mba, I.E., Nweze, E.I., Eze, E.A., Anyaegbunam, Z.K.G., 2022. Genome plasticity in *Candida albicans*: A cutting-edge strategy for evolution, adaptation, and survival. Infect. Genet. Evol. 99, 105256.

McDonald, M.J., Rice, D.P., Desai, M.M., 2016. Sex speeds adaptation by altering the dynamics of molecular evolution. Nat. 2016 5317593 531, 233–236.

Mceachernt, M.J., Hicks, J.B., 1993. Unusually large telomeric repeats in the yeast *Candida albicans*. Mol. Cell. Biol. 13, 551–560.

Miller, M.G., Johnson, A.D., 2002. White-Opaque Switching in *Candida albicans* Is Controlled by Mating-Type Locus Homeodomain Proteins and Allows Efficient Mating. Cell 110, 293–302.

Min, K., Ichikawa, Y., Woolford, C.A., Mitchell, A.P., 2016. *Candida albicans* Gene Deletion with a Transient CRISPR-Cas9 System. mSphere 1.

Mixão, V., Gabaldón, T., 2020. Genomic evidence for a hybrid origin of the yeast opportunistic pathogen *Candida albicans*. BMC Biol. 18, 1–14.

Montoya, M.C., Moye-Rowley, W.S., Krysan, D.J., 2019. *Candida auris*: The Canary in the Mine of Antifungal Drug Resistance. ACS Infect. Dis. 5, 1487–1492.

Morii, D., Seki, M., Binongo, J.N., Ban, R., Kobayashi, A., Sata, M., Hashimoto, S., Shimizu, J., Morita, S., Tomono, K., 2014. Distribution of *Candida* species isolated from blood cultures in hospitals in Osaka, Japan. J. Infect. Chemother. 20, 558–562.

Morio, F., Loge, C., Besse, B., Hennequin, C., Le Pape, P., 2010. Screening for amino acid substitutions in the *Candida albicans* Erg11 protein of azole-susceptible and azole-resistant clinical isolates: new substitutions and a review of the literature. Diagn. Microbiol. Infect. Dis. 66, 373–384.

Morio, F., Pagniez, F., Lacroix, C., Miegeville, M., Le pape, P., 2012. Amino acid substitutions in the *Candida albicans* sterol Δ5,6-desaturase (Erg3p) confer azole resistance: characterization of two novel mutants with impaired virulence. J. Antimicrob. Chemother. 67, 2131–2138.

Morschhäuser, J., Barker, K.S., Liu, T.T., Blaß-Warmuth, J., Homayouni, R., Rogers, P.D., 2007. The Transcription Factor Mrr1p Controls Expression of the MDR1 Efflux Pump and Mediates Multidrug Resistance in *Candida albicans*. PLOS Pathog. 3, e164.

Moyes, D.L., Runglall, M., Murciano, C., Shen, C., Nayar, D., Thavaraj, S., Kohli, A., Islam, A., Mora-Montes, H., Challacombe, S.J., Naglik, J.R., 2010. A Biphasic Innate Immune MAPK Response

Discriminates between the Yeast and Hyphal Forms of *Candida albicans* in Epithelial Cells. Cell Host Microbe 8, 225–235.

Moyes, D.L., Wilson, D., Richardson, J.P., Mogavero, S., Tang, S.X., Wernecke, J., Höfs, S., Gratacap, R.L., Robbins, J., Runglall, M., Murciano, C., Blagojevic, M., Thavaraj, S., Förster, T.M., Hebecker, B., Kasper, L., Vizcay, G., Iancu, S.I., Kichik, N., Häder, A., Kurzai, O., Luo, T., Krüger, T., Kniemeyer, O., Cota, E., Bader, O., Wheeler, R.T., Gutsmann, T., Hube, B., Naglik, J.R., 2016. Candidalysin is a fungal peptide toxin critical for mucosal infection. Nat. 2016 5327597 532, 64–68.

Nébavi, F., Ayala, F.J., Renaud, F., Bertout, S., Eholié, S., Moussa, K., Mallié, M., De Meeûs, T., 2006. Clonal population structure and genetic diversity of *Candida albicans* in AIDS patients from Abidjan (Côte d'Ivoire). Proc. Natl. Acad. Sci. 103, 3663–3668.

Nett, J., Lincoln, L., Marchillo, K., Massey, R., Holoyda, K., Hoff, B., VanHandel, M., Andes, D., 2007. Putative role of β-1,3 glucans in *Candida albicans* biofilm resistance. Antimicrob. Agents Chemother. 51, 510–520.

Ng, H., Dean, N., 2017. Dramatic Improvement of CRISPR/Cas9 Editing in *Candida albicans* by Increased Single Guide RNA Expression. mSphere 2.

Nguyen, N., Quail, M.M.F., Hernday, A.D., 2017. An Efficient, Rapid, and Recyclable System for CRISPR-Mediated Genome Editing in *Candida albicans*. mSphere 2.

Nobile, C.J., Johnson, A.D., 2015. *Candida albicans* Biofilms and Human Disease. Annu. Rev. Microbiol. 69, 71.

Noble, S.M., Gianetti, B.A., Witchley, J.N., 2017. *Candida albicans* cell type switches and functional plasticity in the mammalian host. Nat. Rev. Microbiol. 15, 96.

Odds, F.C., Bougnoux, M.E., Shaw, D.J., Bain, J.M., Davidson, A.D., Diogo, D., Jacobsen, M.D., Lecomte,

M., Li, S.Y., Tavanti, A., Maiden, M.C.J., Gow, N.A.R., D'Enfert, C., 2007. Molecular phylogenetics of *Candida albicans*. Eukaryot. Cell 6, 1041–1052.

Odds, F.C., Brown, A.J.P., Gow, N.A.R., 2004. *Candida albicans* genome sequence: a platform for genomics in the absence of genetics. Genome Biol. 2004 57 5, 1–3.

Ostrosky-Zeichner, L., Casadevall, A., Galgiani, J.N., Odds, F.C., Rex, J.H., 2010. An insight into the antifungal pipeline: selected new molecules and beyond. Nat. Rev. Drug Discov. 2010 99 9, 719–727.

Pande, K., Chen, C., Noble, S.M., 2013. Passage through the mammalian gut triggers a phenotypic switch that promotes *Candida albicans* commensalism. Nat. Genet. 2013 459 45, 1088–1091.

Pappas, P.G., Lionakis, M.S., Arendrup, M.C., Ostrosky-Zeichner, L., Kullberg, B.J., 2018. Invasive candidiasis. Nat. Rev. Dis. Prim. 2018 41 4, 1–20.

Pathak, A.K., Sharma, S., Shrivastva, P., 2012. Multi-species biofilm of *Candida albicans* and non-*Candida albicans Candida* species on acrylic substrate. J. Appl. Oral Sci. 20, 70–75.

Pellestor, F., 2019. Chromoanagenesis: cataclysms behind complex chromosomal rearrangements. Mol. Cytogenet. 2019 121 12, 1–12.

Peng, J.C., Karpen, G.H., 2006. H3K9 methylation and RNA interference regulate nucleolar organization and repeated DNA stability. Nat. Cell Biol. 2006 91 9, 25–35.

Pfaller, M.A., Diekema, D.J., 2007. Epidemiology of invasive candidiasis: A persistent public health problem. Clin. Microbiol. Rev. 20, 133–163.

Pfaller, M.A., Moet, G.J., Messer, S.A., Jones, R.N., Castanheira, M., 2011. Geographic Variations in Species Distribution and Echinocandin and Azole Antifungal Resistance Rates among *Candida*

Bloodstream Infection Isolates: Report from the SENTRY Antimicrobial Surveillance Program (2008 to 2009). J. Clin. Microbiol. 49, 396–399.

Ponde, N.O., Lortal, L., Ramage, G., Naglik, J.R., Richardson, J.P., 2021. *Candida albicans* biofilms and polymicrobial interactions. Crit. Rev. Microbiol. 47, 91–111.

Price, R.J., Weindling, E., Berman, J., Buscaino, A., 2019. Chromatin profiling of the repetitive and nonrepetitive genomes of the human fungal pathogen *Candida albicans*. MBio 10.

Pujol, C., Joly, S., Nolan, B., Srikantha, T., Soll, D.R., 1999. Microevolutionary changes in *Candida albicans* identified by the complex Ca3 fingerprinting probe involve insertions and deletions of the full-length repetitive sequence RPS at specific genomic sites. Microbiology 145, 2635–2646.

Ramage, G., Bachmann, S., Patterson, T.F., Wickes, B.L., López-Ribot, J.L., 2002. Investigation of multidrug efflux pumps in relation to fluconazole resistance in *Candida albicans* biofilms. J. Antimicrob. Chemother. 49, 973–980.

Raut, J.S., Chauhan, N.M., Shinde, R.B., Karuppayil, S.M., 2013. Inhibition of planktonic and biofilm growth of *Candida albicans* reveals novel antifungal activity of caffeine. J. Med. Plants Res. 7, 777–782.

Revie, N.M., Iyer, K.R., Robbins, N., Cowen, L.E., 2018. Antifungal drug resistance: evolution, mechanisms and impact. Curr. Opin. Microbiol. 45, 70–76.

Rhodes, J., Fisher, M.C., 2019. Global epidemiology of emerging *Candida auris*. Curr. Opin. Microbiol. 52, 84–89.

Robbins, N., Wright, G.D., Cowen, L.E., 2016. Antifungal Drugs: The Current Armamentarium and Development of New Agents. Microbiol. Spectr. 4.

Roemer, T., Krysan, D.J., 2014. Antifungal drug development: challenges, unmet clinical needs, and new approaches. Cold Spring Harb. Perspect. Med. 4.

Ropars, J., Maufrais, C., Diogo, D., Marcet-Houben, M., Perin, A., Sertour, N., Mosca, K., Permal, E., Laval, G., Bouchier, C., Ma, L., Schwartz, K., Voelz, K., May, R.C., Poulain, J., Battail, C., Wincker, P., Borman, A.M., Chowdhary, A., Fan, S., Kim, S.H., Le Pape, P., Romeo, O., Shin, J.H., Gabaldon, T., Sherlock, G., Bougnoux, M.E., D'Enfert, C., 2018. Gene flow contributes to diversification of the major fungal pathogen *Candida albicans*. Nat. Commun. 2018 91 9, 1–10.

Rustchenko-bulgac, E.P., Rustchenko-Bulgac, E.P., Sherman, F., Hicks, J.B., Bacteriol, J., 1991. Variations of *Candida albicans* electrophoretic karyotypes. J. Bacteriol. 173, 6586–6596.

Rustchenko, E.P., Curran, T.M., Sherman, F., 1993. Variations in the number of ribosomal DNA units in morphological mutants and normal strains of *Candida albicans* and in normal strains of *Saccharomyces cerevisiae*. J. Bacteriol. 175, 7189–7199.

Santos, M.A.S., Keith, G., Tuite, M.F., 1993. Non-standard translational events in *Candida albicans* mediated by an unusual seryl-tRNA with a 5'-CAG-3' (leucine) anticodon. EMBO J. 12, 607–616.

Schönherr, F.A., Sparber, F., Kirchner, F.R., Guiducci, E., Trautwein-Weidner, K., Gladiator, A., Sertour, N., Hetzel, U., Le, G.T.T., Pavelka, N., D'Enfert, C., Bougnoux, M.E., Corti, C.F., LeibundGut-Landmann, S., 2017. The intraspecies diversity of C. *albicans* triggers qualitatively and temporally distinct host responses that determine the balance between commensalism and pathogenicity. Mucosal Immunol. 2017 105 10, 1335–1350.

Selmecki, A., Forche, A., Berman, J., 2010. Genomic plasticity of the human fungal pathogen *Candida albicans*. Eukaryot. Cell 9, 991–1008.

Selmecki, A., Forche, A., Berman, J., 2006. Aneuploidy and Isochromosome Formation in Drug-Resistant

Candida albicans. Science (80-.). 313, 367–370.

Selmecki, A., Gerami-Nejad, M., Paulson, C., Forche, A., Berman, J., 2008. An isochromosome confers drug resistance in vivo by amplification of two genes, ERG11 and TAC1. Mol. Microbiol. 68, 624–641.

Selmecki, A.M., Dulmage, K., Cowen, L.E., Anderson, J.B., Berman, J., 2009. Acquisition of Aneuploidy Provides Increased Fitness during the Evolution of Antifungal Drug Resistance. PLOS Genet. 5, e1000705.

Shi-Kunne, X., Faino, L., van den Berg, G.C.M., Thomma, B.P.H.J., Seidl, M.F., 2018. Evolution within the fungal genus *Verticillium* is characterized by chromosomal rearrangement and gene loss. Environ. Microbiol. 20, 1362–1373.

Shin, J.H., Bougnoux, M.E., D'Enfert, C., Kim, S.H., Moon, C.J., Joo, M.Y., Lee, K., Kim, M.N., Lee, H.S., Shin, M.G., Suh, S.P., Ryang, D.W., 2011. Genetic diversity among Korean *Candida albicans* bloodstream isolates: Assessment by multilocus sequence typing and restriction endonuclease analysis of genomic DNA by use of BssHII. J. Clin. Microbiol. 49, 2572–2577.

Simonetti, A., Ottaiano, E., Diana, M. V., Onza, C., Triassi, M., 2013. Epidemiology of hospital-acquired infections in an adult intensive care unit: results of a prospective cohort study. Ann. Ig. 25, 281–289.

Singh-Babak, S.D., Babak, T., Fraser, H.B., Johnson, A.D., 2021. Lineage-specific selection and the evolution of virulence in the *Candida* clade. Proc. Natl. Acad. Sci. U. S. A. 118.

Sitterlé, E., Maufrais, C., Sertour, N., Palayret, M., d'Enfert, C., Bougnoux, M.E., 2019. Within-Host Genomic Diversity of *Candida albicans* in Healthy Carriers. Sci. Reports 2019 91 9, 1–12.

Skrzypek, M.S., Binkley, J., Binkley, G., Miyasato, S.R., Simison, M., Sherlock, G., 2017. The *Candida*

Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. Nucleic Acids Res. 45, D592–D596.

Slutsky, B., Staebell, M., Anderson, J., Risen, L., Pfaller, M., Soll, D.R., 1987. "White-opaque transition": a second high-frequency switching system in *Candida albicans*. J. Bacteriol. 169, 189–197.

Soll, D.R., 2007. Mating in *Candida albicans* and Related Species. Biol. Fungal Cell 195–217.

Soll, D.R., 1997. Gene regulation during high-frequency switching in *Candida albicans*. Microbiology 143, 279–288.

St-Germain, G., Laverdière, M., Pelletier, R., René, P., Bourgault, A.M., Lemieux, C., Libman, M., 2008. Epidemiology and antifungal susceptibility of bloodstream *Candida* isolates in Quebec: Report on 453 cases between 2003 and 2005. Can. J. Infect. Dis. Med. Microbiol. 19, 55–62.

Staab, J.F., White, T.C., Marr, K.A., 2011. Hairpin dsRNA does not trigger RNA interference in *Candida albicans* cells. Yeast 28, 1–8.

Strollo, S., Lionakis, M.S., Adjemian, J., Steiner, C.A., Prevots, D.R., 2017. Epidemiology of Hospitalizations Associated with Invasive Candidiasis, United States, 2002–2012. Emerg. Infect. Dis. 23, 7.

Suzuki, T., Ueda, T., Watanabe, K., 1997. The 'polysemous' codon—a codon with multiple amino acid assignment caused by dual specificity of tRNA identity. EMBO J. 16, 1122–1134.

Tao, L., Du, H., Guan, G., Dai, Y., Nobile, C.J., Liang, W., Cao, C., Zhang, Q., Zhong, J., Huang, G., 2014. Discovery of a "White-Gray-Opaque" Tristable Phenotypic Switching System in *Candida albicans*: Roles of Non-genetic Diversity in Host Adaptation. PLOS Biol. 12, e1001830.

Teoh, F., Pavelka, N., 2016. How Chemotherapy Increases the Risk of Systemic Candidiasis in Cancer

Patients: Current Paradigm and Future Directions. Pathog. 2016, Vol. 5, Page 6 5, 6.

Teresa Avelar, A., Perfeito, L., Gordo, I., Godinho Ferreira, M., 2013. Genome architecture is a selectable trait that can be maintained by antagonistic pleiotropy. Nat. Commun. 2013 41 4, 1–10.

Tollemar, J., 2001. Need for alternative trial designs and evaluation strategies for therapeutic studies of invasive mycoses. Clin. Infect. Dis. 33, 95–106.

Tomaska, L., Nosek, J., Makhov, A.M., Pastorakova, A., Griffith, J.D., 2000. Extragenomic double-stranded DNA circles in yeast with linear mitochondrial genomes: potential involvement in telomere maintenance. Nucleic Acids Res. 28, 4479–4487.

Torres, R., Martin, M.C., Garcia, A., Cigudosa, J.C., Ramirez, J.C., Rodriguez-Perales, S., 2014. Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–Cas9 system. Nat. Commun. 2014 51 5, 1–8.

Torssander, J., Morfeldt-Månson, L., Biberfeld, G., Karlsson, A., Putkonen, P.O., Wasserman, J., 2009. Oral *Candida albicans* in HIV Infection. http://dx.doi.org/10.3109/00365548709018473 19, 291–295.

Tsao, S., Rahkhoodaee, F., Raymond, M., 2009. Relative contributions of the *Candida albicans* ABC transporters Cdrlp and Cdr2p to clinical azole resistance. Antimicrob. Agents Chemother. 53, 1344–1352.

Tsui, C., Kong, E.F., Jabra-Rizk, M.A., 2016. Pathogenesis of *Candida albicans* biofilm 74, 18.

Tumbarello, M., Fiori, B., Trecarichi, E.M., Posteraro, P., Losito, A.R., de Luca, A., Sanguinetti, M., Fadda, G., Cauda, R., Posteraro, B., 2012. Risk Factors and Outcomes of Candidemia Caused by Biofilm-Forming Isolates in a Tertiary Care Hospital. PLoS One 7, e33705.

Turner, S.A., Butler, G., 2014. The *Candida* Pathogenic Species Complex. Cold Spring Harb. Perspect. Med. 4.

Uhl, M.A., Biery, M., Craig, N., Johnson, A.D., 2003. Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen *C.albicans*. EMBO J. 22, 2668–2678.

Usher, J., 2019. The Mechanisms of Mating in Pathogenic Fungi—A Plastic Trait. Genes 2019, Vol. 10, Page 831 10, 831.

Valentine, M., Benadé, E., Mouton, M., Khan, W., Botha, A., 2019. Binary interactions between the yeast *Candida albicans* and two gut-associated *Bacteroides* species. Microb. Pathog. 135, 103619.

Van het Hoog, M., Rast, T.J., Martchenko, M., Grindle, S., Dignard, D., Hogues, H., Cuomo, C., Berriman, M., Scherer, S., Magee, B.B., Whiteway, M., Chibana, H., Nantel, A., Magee, P.T., 2007. Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. Genome Biol. 8, 1–11.

Vincent, B.M., Lancaster, A.K., Scherz-Shouval, R., Whitesell, L., Lindquist, S., 2013. Fitness Trade-offs Restrict the Evolution of Resistance to Amphotericin B. PLOS Biol. 11, e1001692.

Vincent, J.-L., Rello, J., Marshall, J., Silva, E., Anzueto, A., Martin, C.D., Moreno, R., Lipman, J., Gomersall, C., Sakr, Y., Reinhart, K., Investigators, for the E.I.G. of, 2009. International Study of the Prevalence and Outcomes of Infection in Intensive Care Units. JAMA 302, 2323–2329.

Volpe, T.A., Kidner, C., Hall, I.M., Teng, G., Grewal, S.I.S., Martienssen, R.A., 2002. Regulation of heterochromatic silencing and histone H3 lysine-9 methylation by RNAi. Science (80-.). 297, 1833–1837.

Vyas, V.K., Barrasa, M.I., Fink, G.R., 2015. A *Candida albicans* CRISPR system permits genetic engineering

of essential genes and gene families. Sci. Adv. 1.

Vyas, V.K., Bushkin, G.G., Bernstein, D.A., Getz, M.A., Sewastianik, M., Barrasa, M.I., Bartel, D.P., Fink, G.R., 2018. New CRISPR Mutagenesis Strategies Reveal Variation in Repair Mechanisms among Fungi. mSphere 3.

Wang, H., Yin, H., Jiao, C., Fang, X., Wang, G., Li, G., Ni, F., Li, P., Su, P., Ge, W., Lyu, Z., Xu, S., Yang, Y., Hao, Y., Cheng, X., Zhao, J., Liu, C., Xu, F., Ma, X., Sun, S., Zhao, Y., Bao, Y., Liu, C., Zhang, J., Pavlicek, T., Li, A., Yang, Z., Nevo, E., Kong, L., 2020. Sympatric speciation of wild emmer wheat driven by ecology and chromosomal rearrangements. Proc. Natl. Acad. Sci. U. S. A. 117, 5955–5963.

Wang, J.M., Bennett, R.J., Anderson, M.Z., 2018. The genome of the human pathogen *Candida albicans* is shaped by mutation and cryptic sexual recombination. MBio 9.

Whaley, S.G., Berkow, E.L., Rybak, J.M., Nishimoto, A.T., Barker, K.S., Rogers, P.D., 2017. Azole antifungal resistance in *Candida albicans* and emerging non-*albicans Candida* Species. Front. Microbiol. 7, 1–12.

White, T.C., 1997. Increased mRNA levels of ERG16, CDR, and MDR1 correlate with increases in azole resistance in *Candida albicans* isolates from a patient infected with human immunodeficiency virus. Antimicrob. Agents Chemother. 41, 1482–1487.

Williams, D.W., Jordan, R.P.C., Wei, X.-Q., Alves, C.T., Wise, M.P., Wilson, M.J., Lewis, M.A.O., 2013. Interactions of *Candida albicans* with host epithelial surfaces. Journal of oral Microbiology, 5, p.22434.

Wirsching, S., Michel, S., Morschhäuser, J., 2000. Targeted gene disruption in *Candida albicans* wild-type strains: the role of the MDR1 gene in fluconazole resistance of clinical *Candida albicans* isolates. Mol. Microbiol. 36, 856–865.

Wisplinghoff, H., Bischoff, T., Tallent, S.M., Seifert, H., Wenzel, R.P., Edmond, M.B., 2004. Nosocomial bloodstream infections in US hospitals: Analysis of 24,179 cases from a prospective nationwide surveillance study. Clin. Infect. Dis. 39, 309–317.

Wu, W., Lockhart, S.R., Pujol, C., Srikantha, T., Soll, D.R., 2007. Heterozygosity of genes on the sex chromosome regulates *Candida albicans* virulence. Mol. Microbiol. 64, 1587–1604.

Yadav, V., Sun, S., Coelho, M.A., Heitman, J., 2020. Centromere scission drives chromosome shuffling and reproductive isolation. Proc. Natl. Acad. Sci. U. S. A. 117, 7917–7928.

Yang, F., Todd, R.T., Selmecki, A., Jiang, Y., Cao, Y., Berman, J., 2021. The fitness costs and benefits of trisomy of each *Candida albicans* chromosome. Genetics 218.

Yu, E.Y., Yen, W.-F., Steinberg-Neifach, O., Lue, N.F., 2010. Rap1 in *Candida albicans* : an Unusual Structural Organization and a Critical Function in Suppressing Telomere Recombination . Mol. Cell. Biol. 30, 1254–1268.

Zarnowski, R., Westler, W.M., Lacmbouh, G.A., Marita, J.M., Bothe, J.R., Bernhardt, J., Sahraoui, A.L.H., Fontainei, J., Sanchez, H., Hatfeld, R.D., Ntambi, J.M., Nett, J.E., Mitchell, A.P., Andes, D.R., 2014. Novel entries in a fungal biofilm matrix encyclopedia. MBio 5, 1–13.

Zhang, A., Petrov, K.O., Hyun, E.R., Liu, Z., Gerber, S.A., Myers, L.C., 2012. The Tlo proteins are stoichiometric components of *Candida albicans* Mediator anchored via the Med3 subunit. Eukaryot. Cell 11, 874–884.

Zhang, J., Hollis, R.J., Pfaller, M.A., 1997. Variations in DNA subtype and antifungal susceptibility among clinical isolates of *Candida tropicalis*. Diagn. Microbiol. Infect. Dis. 27, 63–67.