



# Kent Academic Repository

**Freire Benítez, Verónica (2016) *Epigenetics and genome stability in the human fungal pathogen *Candida albicans**. Doctor of Philosophy (PhD) thesis, University of Kent,.**

## Downloaded from

<https://kar.kent.ac.uk/57655/> The University of Kent's Academic Repository KAR

## The version of record is available from

## This document version

UNSPECIFIED

## DOI for this version

## Licence for this version

UNSPECIFIED

## Additional information

## Versions of research works

### Versions of Record

If this version is the version of record, it is the same as the published version available on the publisher's web site. Cite as the published version.

### Author Accepted Manuscripts

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding. Cite as Surname, Initial. (Year) 'Title of article'. To be published in *Title of Journal*, Volume and issue numbers [peer-reviewed accepted version]. Available at: DOI or URL (Accessed: date).

## Enquiries

If you have questions about this document contact [ResearchSupport@kent.ac.uk](mailto:ResearchSupport@kent.ac.uk). Please include the URL of the record in KAR. If you believe that your, or a third party's rights have been compromised through this document please see our [Take Down policy](https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies) (available from <https://www.kent.ac.uk/guides/kar-the-kent-academic-repository#policies>).

**Epigenetics and genome stability**  
**in the human fungal pathogen**  
***Candida albicans***

Thesis submitted to the University of Kent for the Degree of Ph.D. in Genetics



Verónica Freire Benítez

# 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.

Verónica Freire Benéitez

July 2016

*“There is a driving force more powerful than steam, electricity and atomic energy: the will”.*

*Albert Einstein*

# Acknowledgements

I want to start saying thank you to my supervisor Dr Alessia Buscaino. Not only for the outstanding supervision she has given me during my Ph.D but also, for her encourage in hard moments and mentorship. For her advice in which steps to follow in my early career when things became difficult. I will always be thankful for the opportunity and trust she gave me to pursue this Ph.D and joining research. She is a great scientist, leader and a model to follow.

I also want to say thanks to all members of the KFG, for the great output received after every meeting and support. Within the KFG you belong to a family. I cannot forget Eleana and Morena for being great friends and supportive not only at work but with every single small and big problem.

Thank you also to my parents and brother for their support during this Ph.D. Especially my mother who kindly understood my holidays spent working and stress moaning attitude.

Finally, thanks to the University of Kent for funding my Ph.D, otherwise it would have not been possible.

# Table of contents

Declaration .....	ii
Acknowledgements .....	iv
Table of contents .....	v
Abstract .....	x
Abbreviations.....	xi
Chapter 1. Introduction.....	1
1. Epigenetics .....	2
2. Chromatin and the histone code.....	3
2.1. Heterochromatin functions .....	7
2.1.1. Heterochromatin and transcriptionally silencing .....	7
2.1.2. Heterochromatin and inhibition of recombination .....	8
3. <i>Saccharomyces cerevisiae</i> as a model system to study chromatin structure and function.....	9
3.1. Euchromatin.....	9
3.2. Heterochromatin .....	11
3.2.1. Sir2: Master silent regulator.....	12
3.2.2. Sir2 heterochromatin at mating-type .....	13
3.2.3. Sir2 heterochromatin at telomeres .....	16

3.2.4. Sir2 heterochromatin at <i>rDNA</i> .....	18
4. The role of heterochromatin in promoting genome stability .....	22
4.1. Heterochromatin and genome variability in eukaryotic pathogens ...	22
4.2. Plasticity at the human fungal pathogen <i>Candida albicans</i> .....	23
4.3. <i>Candida albicans</i> genome instability .....	25
4.4. <i>Candida albicans</i> genomic repeats .....	28
4.4.1. Centromeres.....	28
4.4.2. <i>MRS</i> .....	31
4.4.3. <i>rDNA</i> .....	32
4.4.4. Telomeres and subtelomeric regions.....	33
4.5. Chromatin factors and pathogenicity in <i>C. albicans</i> .....	34
5. Thesis aim .....	36
6. Thesis organization .....	39
Chapter 2. Chromatin status at <i>C. albicans</i> DNA repeats .....	40
1. Summary .....	41
2. Main article .....	43
3. Supplementary information.....	56
Chapter 3. Chromatin status at <i>C. albicans</i> centromeres. ....	77
1. Summary .....	78
3. Main Article.....	80
4. Supplementary material.....	92

Chapter 4. Heterochromatin and genome stability in <i>C. albicans</i> .....	98
1. Summary .....	99
2. Main article .....	101
3. Supporting information .....	116
Chapter 5. Discussion .....	133
1. Characterization of the chromatin state associated with <i>C. albicans</i> DNA repeats.....	136
1.1. Heterochromatin at the <i>rDNA</i> locus (Freire-Benítez et al., 2016a)	137
1.2. Plastic heterochromatin at telomeres (Freire-Benítez et al., 2016a)	139
1.3. Intermediate heterochromatin at Centromere and <i>MRS</i> (Freire-Benítez et al., 2016a, 2016b).....	140
2. Function of heterochromatin in promoting genome stability (Freire-Benítez et al., 2016c) .....	143
2.1. Heterochromatin does not promote genome stability at the <i>rDNA</i> locus	143
2.2. Heterochromatin promotes genome stability at <i>C. albicans</i> subtelomeric genes.....	145
2.3. Stress conditions promote genome instability independently of Sir2	148
3. Future work .....	148
6. Materials and Methods .....	150

1. Microbiology techniques .....	151
1.1. Yeast media .....	151
1.1.1. YPAD: yeast extract, peptone, adenine dextrose media .....	151
1.1.2. SC/SC DROP-OUT –His/ -Arg/ -Uri media .....	151
1.1.3. LB media .....	152
1.2. Yeast growth .....	153
1.3. Yeast strain construction.....	153
1.4. Silencing assay .....	155
2. Molecular biology techniques .....	155
2.1. Genomic DNA extraction.....	155
2.2. Southern blot.....	156
2.3. Southern probe hybridization .....	157
2.4. RNA extraction and cDNA synthesis.....	159
2.5. Quantitative Chromatin ImmunoPrecipitation (qChIP). .....	159
2.6. Positive yeast transformants PCR .....	161
2.7. qPCR reactions.....	161
2.8. Fluctuation analysis .....	162
2.9. Marker Gene Loss Assay.....	163
2.10. SNP-RFLP analysis .....	163
2.11. TRE plasmid construction.....	165
2.12. Western blot detection .....	167
2.13. High- throughput RNA sequencing .....	169

2.14. Microscopy .....	170
2.15. Bioinformatic tools .....	170
7. References .....	171

## Abstract

Large blocks of DNA repeats are commonly assembled into heterochromatin. Heterochromatic regions impose a transcriptionally repressive environment that can propagate over long distances stochastically silencing native genes as well as reporter genes inserted at these regions independently of DNA sequence. Heterochromatin, associated with repressive histone marks, not only down-regulates transcription but also inhibits recombination at repetitive elements. This epigenetic regulation could be a key regulatory step in organisms, such as microbial pathogens, that have to adapt rapidly to different environments. Here for the first time, we analysed the chromatin state of *C. albicans* repetitive elements and addressed whether and how this epigenetic state controls *C. albicans* genome stability.

The results show that classic SIR2-dependent heterochromatin is assembled at the *rDNA* and telomeres. Interestingly, heterochromatin at telomeres is plastic and remodelled upon environmental changes. Pericentromeric regions and *MRS* (Major Repeated Sequences) are assembled into permissive chromatin bearing features of both heterochromatin and euchromatin. Surprisingly, SIR2-dependent heterochromatin does not control recombination at the *rDNA*. However, it inhibits recombination at the *TRE* (TLO Recombination Element) sequence associated with some *TLO* genes at subtelomeric regions. These results show that epigenetic factors promote differential genome stability at different loci.

## Abbreviations

CEN	Centromere
CENcore	Centromere core
BIR	Non-reciprocal break induced events
bp	Base pair
BTS	Bermude triangle sequence
ChIP	Chromatin immunoprecipitation
Chr	Chromosome
CR	Caloric restriction
DBS	Double strand break
dH <sub>2</sub> O	Distilled water
DIG	Digoxigenin- 11- dUTP
DNA	Deoxyribonucleic acid
EDTA	Ethylenediaminetetraacetic acid
ERC	Extra chromosomal rDNA circle
FOA	5-Fluorotic acid
FluR	Fluconazole resistance
HAT	Histone acetyltransferase

HDAC	Histone deacetylase
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HMT	Histone methyltransferase
IP	Immunoprecipitate
Kb	Kilobase
LOH	Loss of heterozytosity
LTR	Long terminal repeats
<i>MRS</i>	Major repeated sequences
NAM	Nicotinamide
N/S	Non selective
NTS	Non transcribed sequence
PCR	Polymerase chain reaction
PEV	Position effect variegation
Pol I	Polymerase I
Pol II	Polymerase II
Pol III	Polymerase III
siRNA	Small interference RNA
SNC	Supernumerary chromosome

<i>rDNA</i>	Ribosomal DNA
RFB	Replication fork barrier
RITS	RNA-induced transcriptional silencing
RNA	Ribonucleic acid
RNAs	Double strand ribonucleic acid
SC	Synthetic complete
<i>S. cerevisiae</i>	<i>Saccharomyces cerevisiae</i>
<i>S. pombe</i>	<i>Schizosaccharomyces pombe</i>
TE	Tris(hydroxymethyl)aminomethane-EDTA
TPE	Transcription position effect
TRE	<i>TLO</i> recombination element
ROS	Reactive oxygen species
rpm	Revolutions per minute
RT-qPCR	Real time quantitative polymerase chain reaction
Uri	Uridine
UV	Ultraviolet
YPAD	Yeast extract peptone adenine dextrose



# Chapter 1. Introduction

## 1. Epigenetics

Conrad Waddington introduced the term Epigenetics in the early 1940s. In his definition epigenetics was described as “the branch of biology which studies the causal interactions between genes and their products which brings the phenotype into being” (Waddington, 1942). The most established definition says that it refers to heritable changes within a cell that do not involve changes of the DNA sequence. Nowadays the term epigenetics covers a whole range of chemical modifications and molecules, ranging from DNA, RNA and non-coding DNA sequences to proteins and self-replicating prions (Bird, 2013). Epigenetics was firstly observed in the fruit fly *Drosophila melanogaster* which displays variegated expression of eye pigment due to the spread of repressive DNA organization into adjacent regions of the genome (Muller, 1930). Years later, around 1980, the field was established as an emerging discipline using genetics in this organism (Henikoff, 1990). Epigenetics have been studied in many other organisms ranging from yeast and plants to humans. One famous example showing the influence of epigenetics is the Dutch Hunger Winter of 1944 during World War II. The Hunger Winter resulted from a blockade of food and fuel shipments into western Holland from September 1944 until the liberation of the country from Nazi rule in May 1945. The Hunger Winter was well documented and many studies of the effect of starvation on human health were recorded. Results showed a link between famine of the mother during pregnancy and the health of the offspring. Children were associated with diabetes, heart problems, obesity and schizophrenia compared to children of normally fed mothers. These illnesses were also observed in grandchildren of Hunger Winter mothers. This data suggested that the diet of the grandmother could affect the health of

several generations, which further implies that an adaptation to her environment produced a heritable trait. Surprisingly, these effects were observed not only in the offspring of mothers who suffered starvation but also they were associated with men who suffered famine. Hypomethylation of the insulin-dependent growth factor (IGF2) associated with male mass index was observed in the offspring of Hunger Winter individuals, and in subsequent generations. This shows how a heritable trait was acquired and inherited as a consequence of the nutritional state of the parents (Heijmans et al., 2008).

## **2. Chromatin and the histone code**

DNA is packaged with specific proteins which help to condense the DNA into a small volume of space. This compact DNA-protein complex is called chromatin. Histones are the proteins around which the DNA is wrapped in a repetitive manner forming nucleosomes (Kornberg, 1974, 1977) (**figure 1.1**). Around 147 base pairs of DNA are wrapped around an octamer of histones. The octamer is composed of two dimers of the histones H2A-H2B and the histone tetramer  $(H3)_2(H4)_2$ . Nucleosomes are organized into arrays linked by segments of linker DNA (Correll et al., 2012). The variation in length of the linker DNA between nucleosomes is associated with histone H1, this linker histone helps to promote folding of chromatin into a compact fibre (**figure 1.1**) (Kornberg, 1974).

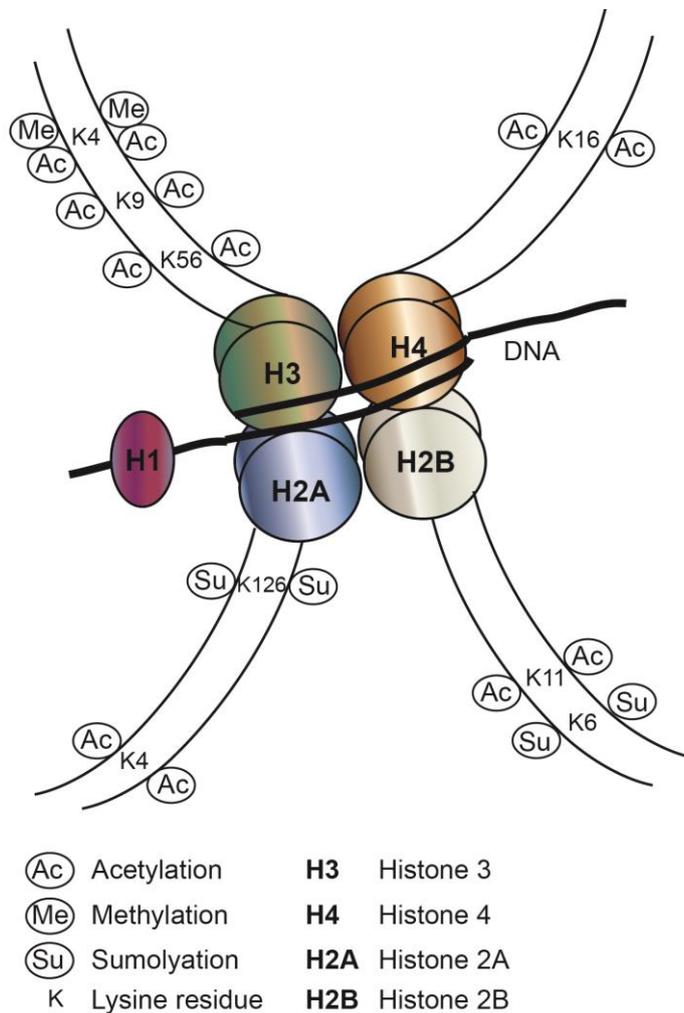
Histones are subject to post-translational modifications (PTMs). Highly basic histone amino (N)-terminal tails protrude from the nucleosome core and they can be post-translationally modified (**figure 1.1**). These modifications alter the chromatin structure to influence different processes such as gene expression, DNA repair, DNA replication or DNA recombination. These modifications are

thought to represent a code indicating the expression state of the underlying chromatin, in the so-called histone code hypothesis (Jenuwein and Allis, 2001; Strahl and Allis, 2000). This hypothesis states that multiple histone modifications acting in a combinatorial or sequential fashion on one or multiple histone tails specify unique downstream functions. This code is translated by effector proteins that bind specifically to particular modifications on the histones and, in turn, lead to further chromatin modifications that ultimately dictate particular expression states. Histone modifications also reinforce one another, with one modification affecting the likelihood of other modifications.

The histone code in a given cellular context is brought about by a series of writing and erasing events carried out by histone-modifying enzymes. The writer of histone modification refers to an enzyme (for example, a histone methyltransferase) that catalyzes a chemical modification of histones in a residue-specific manner, and the eraser of histone modification refers to an enzyme (for example, a histone demethylase) that removes a chemical modification from histones. The reading of the histone code is accomplished by reader or effector proteins that specifically bind to a certain type or a combination of histone modification and translate the histone code into a meaningful biological outcome such as transcriptional activation, transcriptional repression or silencing and other cellular responses (Chi et al., 2010).

There are several types of histone modifications: methylation, acetylation, phosphorylation, ubiquitylation, sumoylation, deimination,  $\beta$ -N-acetylglucosamination, histone proline isomerization, histone tail clipping and ADP-ribosylation (Bannister and Kouzarides, 2011; Cosgrove and Wolberger, 2005; Weinhold, 2006). The histone code hypothesis provides a framework for

understanding the role of histone modifications in chromatin assembly into euchromatin or heterochromatin (Rusche et al., 2003).



**Figure 1.1. Histones wrapping DNA to form the nucleosome. (A)** Nucleosome composition. Histones H3 and H4 assembled to form a heterotetramer containing two copies of H3 and two copies of H4, while histones H2A and H2B form heterodimers and then assemble into a tetramer. The  $(H3)_2(H4)_2$  heterotetramer and the H2A-H2B heterodimer bind to DNA, wrapped in a left-manner to form the nucleosome. Histone H1, a linker histone, binds to the linker DNA that connects successive nucleosomes, facilitating the compaction of the chromatin. (N)-terminal tails of histones are subject to post-translational modifications (PTMs) which are required to compact or to relax chromatin architecture modifying its biological function.

Euchromatic regions are an open state of the DNA. The DNA wrapped around the histones is loose so it remains accessible and allows different processes such as transcription or DNA replication. In general euchromatin is associated with histone acetylation and certain types of methylation that are associated with actively transcribed genes. Some examples are methylation of lysine 4 and lysine 36 at histone 3 (H3K4, H3K36) which are associated with (Polymerase II) Pol II transcribed genes (van Steensel, 2011).

Heterochromatin is a closed state of the DNA. Two types of heterochromatin can be distinguished: facultative and constitutive. Facultative heterochromatin is associated with a chromatin state that can change in response to cellular signals and gene activity. It is associated with cell differentiation and consists of regions of the genome that were expressed and that become silenced. The histone marks associated with this silence state are perpetuated to new generations so facultative heterochromatin is maintained. One example is the inactive X-chromosome present within female cells. One X chromosome is packed as facultative heterochromatin and silenced, while the other X chromosome is euchromatic and actively expressed (Bannister and Kouzarides, 2011). Constitutive heterochromatin contains silenced genes in genomic regions common in all cells of the same species such as centromeres and telomeres (Bannister and Kouzarides, 2011; Grewal and Jia, 2007).

Heterochromatin is associated in general with histones that are hypoacetylated and hypomethylated on specific histone residues. For example, heterochromatic regions contain low levels of methylation at lysines 4 and 79 of histone 3 (H3K4, H3K79). In fission yeast, fruit flies and higher eukaryotes, heterochromatin is not only hypoacetylated and hypomethylated at H3K4 and H3K79 but also

highly methylated at lysine 9 histone 3 (H3K9). Methylation at H3K9 is introduced by the histone methyltransferase (HMT) SU(VAR)3-9 (Clr4 in fission yeast). Histone methylation by SU(VAR)3-9 helps to recruit the chromodomain factor HP1. This is a good example of the histone code hypothesis. A histone modifier enzyme or writer (SU(VAR)3-9) creates an histone mark that subsequently tethers the effector reader (HP1). HP1 creates a platform for binding other chromatin-modifiers promoting compaction of the chromatin that is accompanied by gene repression (Bannister and Kouzarides, 2011; Bi, 2012; Kornberg, 1977).

## **2.1. Heterochromatin functions**

Heterochromatin governs different processes; such as gene expression, imprinting dosage compensation, chromosome condensation and recombination which are crucial for genome integrity and stability (Reik & Walter 1998; Cavalli 2002; Grewal & Jia 2007; Politz et al. 2013). In this thesis concepts such as transcriptional silencing and inhibition of recombination mediated by heterochromatin, are explained in detail.

### **2.1.1. Heterochromatin and transcriptionally silencing**

Heterochromatin influences gene expression independently of the DNA sequence. Heterochromatin usually nucleates at specific sites, termed nucleation sites, and spreads across domains leading to transcriptional silencing of associated genes (Grewal and Jia, 2007). Silencing is a stochastic process that can be conveniently monitored by insertion of marker genes into heterochromatic regions. At these regions marker genes show reduced transcriptional expression compared to an euchromatic position (Allshire et al.,

1994; Bannister and Kouzarides, 2011; Freeman-Cook et al., 2005; Fritze et al., 1997; Rusche et al., 2003; Thon and Verhein-Hansen, 2000; Weinhold, 2006). Silenced regions are less accessible to restriction nucleases and DNA methylases and display regularly spaced nucleosomes with a compact topology (van Steensel, 2011). While some histone modifiers act promoting silencing at specific gene promoters, heterochromatin assembly promotes a silenced state comprising a region of the genome. This compact state of the chromatin can be maintained and inherited to next generations during replication due to incorporation of epigenetically modified histones and newly synthesized histones which will be modified (Li and Zhang, 2012; Rusche et al., 2003). Heterochromatic regions also cause position effect variegation (PEV) or silencing of genes inserted close to these heterochromatic loci (Allshire et al., 1995a; Gottschling et al., 1990; Schoeftner and Blasco, 2009; Stavenhagen and Zakian, 1998; Tham and Zakian, 2002; Yankulov, 2013).

### **2.1.2. Heterochromatin and inhibition of recombination**

Heterochromatin can also repress recombination, which protects genome integrity by prohibiting illegitimate recombination between dispersed repetitive DNA sequences (Heringa, 1998). Heterochromatin protects genome integrity by rendering repetitive structures inert to recombination. At the *rDNA*, heterochromatin prevents an increase of recombination rates associated with the repeated nature of the cluster promoting genome stability (Gottlieb and Esposito, 1989a). Heterochromatin is also assembled at pericentromeric regions. At these positions heterochromatin prevents recombination events close to the centromere core which could impair correct kinetochore assembly and chromosome segregation (Ellermeier et al., 2010). At telomeres

heterochromatin protects them from telomere degradation and recombination events which could lead to end chromosomal fusions (Schoeftner and Blasco, 2009). It also promotes silencing of parasitic transposable elements. Active transposons are highly mutagenic causing chromosome breakage, illegitimate recombination and genome rearrangements (Grewal and Jia, 2007; Henikoff, 1990; Slotkin and Martienssen, 2007).

### **3. *Saccharomyces cerevisiae* as a model system to study chromatin structure and function**

Budding yeast are an outstanding system for epigenetic studies by the use of the well characterized model organism: *Saccharomyces cerevisiae*. This yeast is an outstanding model system to investigate chromatin structure and function due to the presence of silencing regions at the mating type, telomeres and the *rDNA*. Genetic manipulation in yeast is easy and cheap and many epigenetic factors from higher eukaryotes have homologs in *S. cerevisiae* facilitating their study.

#### **3.1. Euchromatin**

As mentioned before in this introduction, euchromatin is associated with high levels of acetylation and methylation. Histone acetylation is introduced by histone acetyltransferases (HATs) and histone methylation by histone methyltransferases (HMTs). Euchromatin in *S. cerevisiae* is associated with, among others, acetylation at lysine 16 histone 4 (H4K16), at lysines 9 and 36 histone 3 (H3K9, H3K36) and at lysine 12 histone 4 (Bannister and Kouzarides, 2011; Ryu and Ahn, 2014). H4K16 is acetylated by the HAT Sas2 (Choy et al., 2011; Oppikofer et al., 2011). H3K36 and H3K9 are acetylated by the HAT

Gcn5 (Cieniewicz et al., 2014) and H4K12 is acetylated by the HAT Esa1 (Chang and Pillus, 2009). Common methylation marks associated with euchromatin are methylation at H3K36, and lysine 79 histone 3 (H3K79). H3K36 is methylated by the HMT Set2 and H3K79 by HMT Dot1 (van Leeuwen et al., 2002; Morris et al., 2006; Ng et al., 2002). These marks are associated with active gene transcription. Of special interest in this thesis is methylation at H3K4. Set1 is the HMT introducing methylation at H3K4. The SET catalytic domain has high homology to *Drosophila* genes: Su(var) 3-9, Enhancer of zeste, and tritrhorax (Jenuwein, 2001) from which it took its name. In *S. cerevisiae* each modified lysine can exist in mono- (H3K4me), di- (H3K4me<sub>2</sub>), or tri-methylated (H3K4me<sub>3</sub>) form. Set1 deletion abolishes all H3K4 types of methylation. While H3K4me<sub>3</sub> peaks at the beginning of the transcribed portions of genes, H3K4me<sub>2</sub> is most enriched in the middle of genes, and H3K4me is found predominantly at the end of genes (Dehé and Géli, 2006; Ng et al., 2003).

However, Set1 action remains controversial. Studies showed that Set1 is also required for repression of Pol II transcribed genes placed within the *rDNA* locus (Briggs et al., 2001; Ng et al., 2003; Williamson et al., 2013). A reduction in H3K4 methylation is associated with a loss of silencing not only at the *rDNA* locus but also at telomeres and the mating cassette (Dehé and Géli, 2006; Fingerman et al., 2005; Krogan et al., 2002; Nislow et al., 1997). One model suggests that this could be due to the different localization of heterochromatin factors associated with low methylation levels. Absence of methylation could produce the redistribution of silencing factors being responsible of the silencing defects observed at the different heterochromatic loci (Dehé and Géli, 2006).

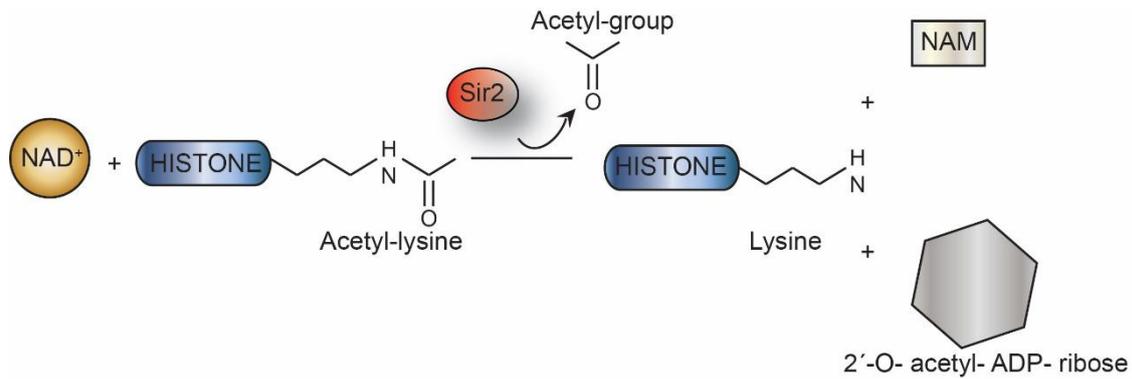
### 3.2. Heterochromatin

In *S. cerevisiae* heterochromatin assembles at different repetitive regions of chromosomes: telomeres, ribosomal DNA locus (*rDNA*) and the mating type locus (Bühler and Gasser, 2009; Cam, 2010; Cam et al., 2005). Heterochromatin represses the cryptic mating type to maintain haploid cell identity. At the *rDNA* heterochromatin suppresses unbalanced recombination of the tandem repeated *rDNA* genes via association of cohesin and at the end of the chromosomes it regulates subtelomeric gene expression (Rusche et al., 2003).

The main structural proteins of silenced chromatin in budding yeast are known as silent information regulators or Sir proteins. The hallmark of the Sir protein family or Sirtuin family is a domain of approximately 260 amino acids that has a high degree of sequence similarity in all sirtuins. The family is divided into five classes (I-IV and U) on the basis of phylogenetic analysis of different organisms. Sirtuins have been found from bacteria to eukaryotes. Class I and class IV are further divided into three and two subgroups, respectively. The U-class sirtuins are found only in Gram-positive bacteria. The *S. cerevisiae* genome encodes five sirtuins, Sir2 and four additional proteins termed 'homologs of sir two' (Hst1-4). Within this family, Sir2 is the main master silent regulator. Silenced domains consist of continuous distributions of Sir proteins that, together with hypoacetylated nucleosomes, are thought to form an ordered, compact structure restrictive to transcription (Freeman-Cook et al., 2005; North et al., 2004; Rusche et al., 2003).

### 3.2.1. Sir2: Master silent regulator

Sir2 (silent information regulator) is a histone deacetylase (HDAC) class III, member of a large protein family (Sirtuin family) that is conserved from bacteria to humans. Studies on the *SIR2* bacterial homolog, *cobB*, led to the conclusion that this gene could substitute for another bacterial gene, *cobT*, in the pathway of cobalamin synthesis. *cobT* was known to encode an enzyme that transferred ribose-phosphate from nicotinic acid mononucleotide to dimethyl benzimidazole (Guarente, 2000). Class I HDACs share similarity in their catalytic cores. Class III HDACs are dependent on nicotinamide adenine dinucleotide ( $\text{NAD}^+$ ) and have no sequence similarity to class I and II HDACs (North et al., 2004). Sir2 couples a deacetylation reaction to the hydrolysis of  $\text{NAD}^+$ . In this reaction one molecule of  $\text{NAD}^+$  is cleaved into ADP-ribose and nicotinamide (NAM) per acetyl group removed. The acetyl group is transferred to ADP-ribose, forming the product 2'-O-acetyl-ADP-ribose (Imai et al., 2000) (**figure 1.2**). Sir2 is the most widespread and well conserved of the Sir proteins. In fact, unlike the rest of Sir proteins, which are restricted to budding yeast, Sir2 has homologs among all domains of life including eubacteria and archaea (Hickman et al., 2011; Rusche et al., 2003; Sauve et al., 2005; Yang and Seto, 2003).



**Figure 1.2.** Sir2 deacetylation activity. Sir2 in the presence of  $\text{NAD}^+$  removes acetyl groups from lysines at the (N)-terminus of histone tails. This reaction deacetylates lysine tails and releases the acetyl-group, to form NAM and 2'-O-acetyl-ADP-ribose.

### 3.2.2. Sir2 heterochromatin at mating-type

*S. cerevisiae* haploid cells are of one of two mating types, a or  $\alpha$  or a/  $\alpha$  diploid by mating of opposite haploid types. The mating type is determined by the allele of the mating type locus *MAT*. The *MAT* locus consists of two alleles: *MATa* and *MAT $\alpha$* . Cells expressing the genes at *MATa* mate as a-type cells, cells expressing genes at *MAT $\alpha$*  mate as  $\alpha$ -type cells, and cells expressing the genes at both loci are unable to mate but are competent to go through meiosis. *MATa* encodes for a single protein, Mata1, a homeodomain protein, and *MAT $\alpha$*  codes for another homeodomain protein (Mata $\alpha$ 2) and an alpha-domain protein (Mata $\alpha$ 1). The three transcriptional regulators, Mata1, Mata $\alpha$ 1, and Mata $\alpha$ 2, control mating type via a simple combinatorial circuit. In haploid  $\alpha$  cells, Mata $\alpha$ 1 turns on  $\alpha$ -specific genes while Mata $\alpha$ 2 represses a-specific genes, directing the cell to mate as an  $\alpha$  cell. a-specific genes are constitutively expressed in the absence of Mata $\alpha$ 2 and  $\alpha$ -specific genes are not transcribed in the absence of Mata $\alpha$ 1. Cells that have mated express both Mata1 and Mata $\alpha$ 2 proteins, and these form a heterodimer that represses many of the mating genes, thereby blocking

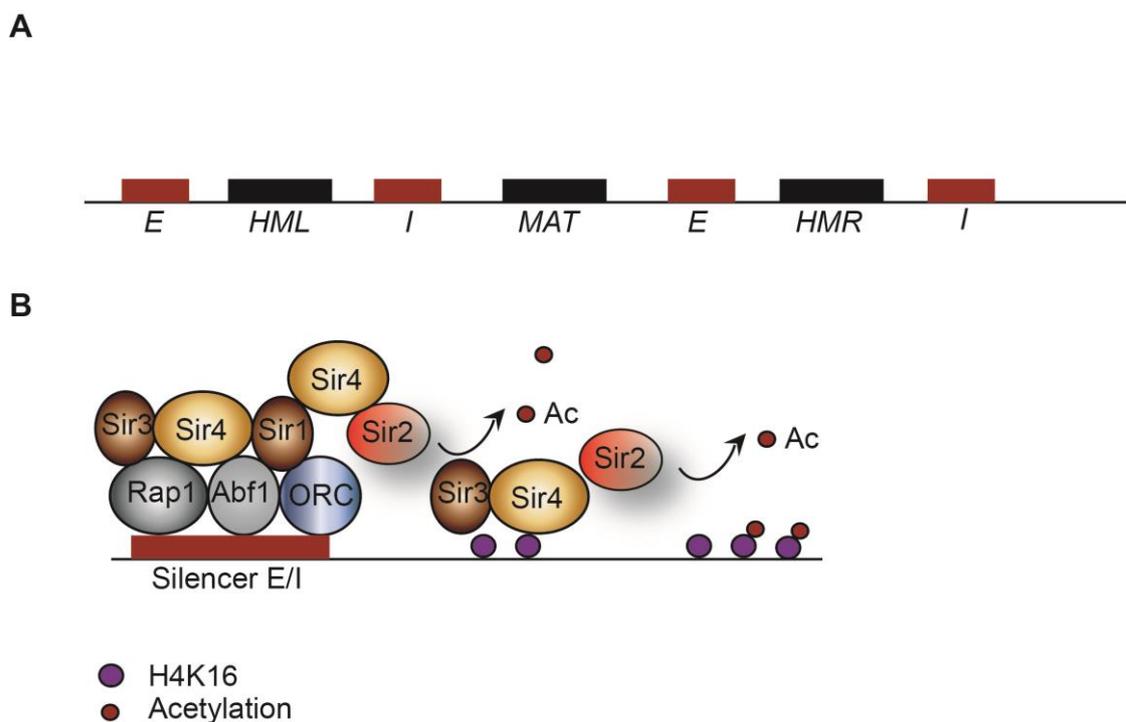
mating by the  $a/\alpha$  cell but permitting meiosis under favorable environmental conditions.

*S. cerevisiae* has the ability to convert some cells in a colony from one haploid mating type to the other. This process is called homothallism and the subsequent mating of the cells of the opposite mating type enables these homothallic organisms to self-diploidise. The diploid state provides evolutionary advantages, as the possibility to undergo meiosis and spore formation under nutritionally limiting conditions (Haber, 2012).

Mating conversion is possible thanks to the presence of unexpressed copies of mating type genes at other loci: the hidden mating type left (*HML*) and the hidden mating type right (*HMR*) (**figure 1.3 A**). The two silent mating-type loci, *HMR* and *HML*, donate information to the *MAT* locus, and are responsible for the mating-type switch. The conversion of one mating type to the other involves the replacement at the *MAT* locus of a or  $\alpha$  type cells by a gene conversion induced by a double strand break (DSB) by a HO endonuclease. The DSB will be repaired using the genetic information from the *HML* and the *HMR*. The HO endonuclease cannot cleave its recognition sequence at either *HML* or *HMR*, as these sites are occluded by nucleosomes in silenced DNA.

The *HM* loci are flanked by 150 bp of cis-acting elements, the silencers *E* and *I*, both of which are located around 1 kb from the genes they regulate. The silencers function to initiate the assembly of the Sir complex. Silencers contain binding sites for the origin replication complex (ORC), Rap1 and the autonomous replication sequence binding factor (Abf1). Silencing at the *HM* loci requires the Sir complex, which is composed of Sir1, Sir2, Sir3 and Sir4. Sir3

and Sir4 are recruited by Rap1. The Sir complex is just present at *S.cerevisiae*. Sir4 is also recruited by Sir1 which binds ORC. Once bound to chromatin, the Sir complex is believed to spread over nucleosomes. Sir4 tethers Sir2 to deacetylate H4K16. Sir3 has high affinity for low H4K16. This process spreads the whole complex away from the nucleation site and silences the *HML* and *HMR* copies (**figure 1.3 B**) (Bi, 2014; Huang, 2002; Rusche et al., 2003).



**Figure 1.3.** The assembly of Sir proteins at *S. cerevisiae* mating type locus. **(A)** Mating type cassette containing the *MAT* locus encoding for one of the mating types and the hidden mating cassettes encoding for silenced copies of both mating types. **(B)** Sir proteins are recruited by Rap1, Abf1 and ORC to silencers E/I. Sir2 deacetylates H4K16. Once nucleated, heterochromatin spreads via Sir3 and Sir4 binding with low acetylated H4K16 to silence the *HMR* and *HML*.

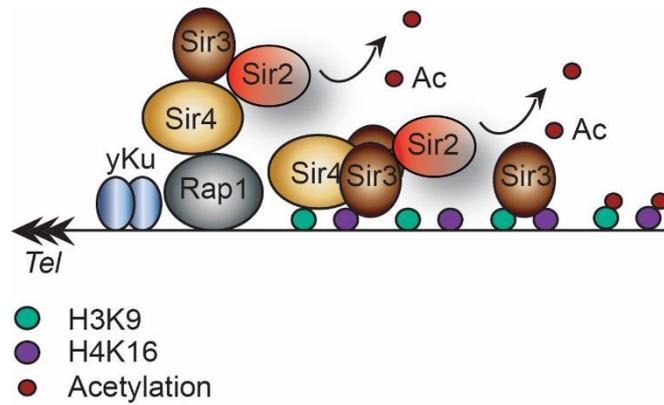
### 3.2.3. Sir2 heterochromatin at telomeres

*S. cerevisiae* telomeres contain TG-rich repeats which have adjacent conserved subtelomeric sequences called core X and Y' elements. While core X are present at all *S.cerevisiae* subtelomeres, Y' elements are found in half of *S.cerevisiae* subtelomeres. Insertion of a *URA3*<sup>+</sup> marker gene at subtelomeres was subject to transcriptional silencing (Marvin et al., 2009). Cells were able to grow in the presence of the drug 5-fluorotic acid (FOA), and when replica plated to media lacking uracil most cells were also able to grow. This reversible repression of *URA3*<sup>+</sup> transcription is termed telomere position effect TPE (Gottschling et al., 1990; Tham and Zakian, 2002). However at *S.cerevisiae*, silencing is observed at some, but not all chromosome ends. Not only that, TPE is independent of ploidy number. Haploid cells showed as strong transcriptional silencing at telomeres as diploid cells. However, haploid cells showed lower level of telomeric reporter proteins compared to diploid cells (McLaughlan et al., 2012). Moreover, silencing at subtelomeric regions is discontinuous. While there is strong repression at core X elements and it is also observed at telomeric TG-rich repeats, silencing at Y' elements is weak (Marvin et al., 2009).

At budding yeast subtelomeres, the SIR complex that contains Sir2, Sir3 and Sir4 is recruited by Rap1 and the yKu heterodimer. Every Rap1 molecule provides a binding site for Sir4. Sir2 reduces the acetylation levels of H3K9ac and H4K16ac, creating a high affinity binding site for Sir3. This results in the spreading of the complex outward from the nucleation start site. Sir4 enhances affinity of the SIR complex through its DNA binding domain while Sir3 specifically recognizes hypoacetylated histones contributing to the spreading of nucleosomes via dimerization (Bühler and Gasser, 2009; Moretti et al., 1994;

Schoeftner and Blasco, 2009) (**figure 1.4**). Deacetylation of histone tails promotes the association of extra copies of Sir3, allowing Sir4 and Sir2 to bind successively. Sir4 is necessary for the recruitment of the entire SIR complex, although once nucleated, excess Sir3 can propagate along nucleosomes without Sir4. The core region of heterochromatin contains all three Sir proteins whereas surrounding regions contain mainly Sir3 (**figure 1.4**). Spreading of heterochromatin eventually stops due to the increase of acetylation levels in adjacent euchromatin at telomeres (Bühler and Gasser, 2009; Grewal and Moazed, 2003; Rusche et al., 2003; Xu et al., 2007).

Deletion of members of the SIR complex abolishes silencing associated with TPE, causes shortening of telomeric repeats and mitotic instability associated with shortened chromosomes. Genome stability is kept in conjunction between the yKu heterodimer and Sir-dependent silencing. Interestingly, binding of the yKu heterodimer is Sir-independent at the core X element. However, stronger binding of the yKu heterodimer is associated with subtelomeres subject to silencing. For instance, the silenced subtelomere at the left arm of chromosome XI showed stronger yKu binding compared to the non-silenced subtelomere present at the right arm of chromosome III (Marvin et al., 2009). The yKu heterodimer binds the TG-rich repeats and the core X element mediating a protective fold-back structure at chromosome ends (Marvin et al., 2009; McLaughlan et al., 2012). Thus, subtelomeric heterochromatin promotes genome stability and integrity at telomeres (Huang, 2002; Tham and Zakian, 2002).



**Figure 1.4.** The assembly of Sir proteins at *S. cerevisiae* telomeres. Sir proteins are recruited by Rap1 and the yKu heterodimer to telomeres. Sir2 deacetylates H3K9 and H4K16 in the heterochromatin assembly process. Deacetylated histones serve to recruit Sir3 and its dimerization. Sir3 dimers in complex with Sir2-4 result in spreading of heterochromatin from nucleation sites.

### 3.2.4. Sir2 heterochromatin at *rDNA*

*S. cerevisiae rDNA* array consist of tandem repeated units of 9.1 kb in size each. There are 100 to 200 copies located on chromosome XII. Each *rDNA* repeat contains the 35S *rDNA* transcribed by Polymerase I (Pol I), which is processed into the 25S, 18S and 5.8S. It also contains the 5S which is transcribed by Polymerase III (Pol III). 35S and 5S are separated by non-transcribed spacers, NTS1 and NTS2 (**figure 1.5 A**). Pol II marker genes inserted at the NTS1 are transcriptionally silenced (Bryk et al., 1997; Fritze et al., 1997; Smith and Boeke, 1997; Smith et al., 1999). Silencing at the *rDNA* is SIR2-dependent. Sir2 forms part of the nucleolar RENT complex at budding yeast *rDNA* (Straight et al., 1999). The RENT complex is formed by Net1 required for Sir2 binding to *rDNA*. It is also formed by Cdc14, a phosphatase required for mitotic exit (Cioci et al., 2003; Gottlieb and Esposito, 1989b;

Straight et al., 1999). It is not clear yet, the mechanism by which Sir2 regulates silencing at *rDNA*.

Silent chromatin at the *rDNA* is essential for repression of genetic recombination via inhibition of Pol II non-coding transcripts arising from NTS1 independently of silencing at Pol I and Pol III *rDNA* genes (Ganley et al., 2005; Kobayashi, 2011). On the other hand, just a fraction of 35S and 5.8S genes are transcribed by Pol I and Pol III respectively at a given time, where inactive repeats seem to be associated heterochromatin (Grummt and Pikaard, 2003; Nazar, 2004). Psoralen cross-linking and electron microscopy imaging studies reveal that the coding regions of the active repeats are largely devoid of nucleosomes, whereas the inactive repeats are associated with nucleosomes (Rusche et al. 2003; Bhargava & Reese 2013; Hamperl et al. 2013). Surprisingly, reporter genes are more strongly silenced when inserted at *rDNA* genes in an active state associated with high Pol I transcription (Cioci et al., 2003). This could inhibit Pol II transcription of endogenous transposable elements at the *rDNA* in favor of Pol I transcribed genes (Rusche et al., 2003).

One of the main functions of SIR2- dependent heterochromatin is inhibition of recombination. The highly repetitive nature of the *rDNA* makes them highly prone to recombination events. This leads to the loss of copies after deleterious recombination events among repeats. Recombination between repeats of sister chromatids leads to the formation of a secondary structure which could block the DNA replication process. DNA replication blockage is repaired by recombination losing *rDNA* copies in the form of extra chromosomal *rDNA* circles (ERCs) (**figure 1.5**). These dynamic copy number variations make the *rDNA* locus a very fragile part of the genome. Reduced numbers of non-

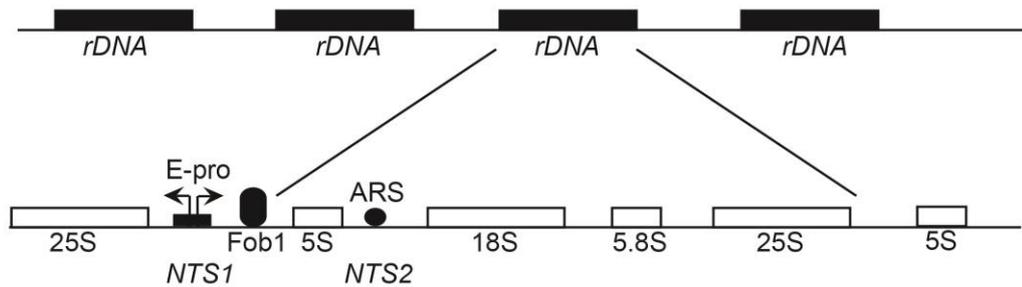
transcribed *rDNA* regions in low copy strains leads to improper chromatid segregation due to a lack of binding sites for cohesin (Ganley and Kobayashi, 2011; Kobayashi, 2011).

A counteracting recombination mechanism is well studied in *S. cerevisiae*. During S phase, replication starts from replication origins and is inhibited at the replication fork barrier site (RFB) by the fork blocking protein, Fob1. This stalled replication fork makes the single-stranded region of the blocked structure a target for endonuclease activity. This leads to the formation of a double strand break (DSB). This DSB can be repaired by homologous recombination with a sister chromatid. Due to the repetitive nature of the *rDNA*, the broken end of a repeat may be also recombined with a neighbor repeat which increases the number of copies in one of the sister chromatids (Huang et al., 2006b; Wang et al., 2006).

The amplification mechanism is mediated by E-pro, a bidirectional promoter directing transcription of non-coding transcripts. These transcripts may interfere with cohesin molecules and its ability to connect sister chromatids (Chan et al., 2011; Gartenberg, 2009; Hirano, 2012; Kobayashi and Ganley, 2005). This impairment promotes unequal sister chromatid recombination and promotes recombination with neighboring repeats increasing the number of *rDNA* units (Ganley and Kobayashi, 2013; Salminen and Kaarniranta, 2009; Sinclair and Guarente, 1997). When the repeat number is around wild-type level, Sir2 may play its role binding E-pro promoter, promoting heterochromatin assembly and repressing non-coding transcript formation. This allows cohesin molecules to bind to *rDNA* repeats promoting correct chromatin recombination and *rDNA*

stability (Li et al., 2006a) (**Figure 1.5 B**). How a cell monitors *rDNA* copy number is still unclear (Ganley and Kobayashi, 2011; Kobayashi, 2011).

A



B



**Figure 1.5.** Sir2 inhibition of recombination at *S. cerevisiae* *rDNA*. **(A)** *rDNA* tandem repeated units. Each unit is composed of the *rDNA* genes (18S, 5.8S, 25 S and 5S) and two non-transcribed regions (NTS1 and NTS2) flanking the 5S. NTS1 contains the bidirectional promoter E-pro and the replication fork barrier protein Fob1. NTS2 contains an autonomous replication origin. **(B)** In the presence of Sir2, non-coding transcripts arising from E-pro are stopped. This allows cohesin to bind to the *rDNA* units and promote correct sister chromatid segregation. In the absence of Sir2, E-pro bidirectional transcription produces non-coding transcripts impeding the correct binding of cohesin and promoting recombination events leading to copy number variations of the *rDNA* units and the appearance of ERCs.

## **4. The role of heterochromatin in promoting genome stability**

### **4.1. Heterochromatin and genome variability in eukaryotic pathogens**

Eukaryotic pathogens must adapt to environmental stimuli to successfully propagate within the host. Within the host, they need to avoid immune resistance and ensure transmission to the next host. This process often requires dramatic morphological changes, expression of virulence genes or DNA repair genes to repair possible DNA damage incurred during the host defense response. Many of these events involve chromatin alterations in order to adapt to the changing environmental stimuli (Haldar et al., 2006; Lopes da Rosa and Kaufman, 2012; Steinert, 2014). One well studied example of epigenetics and pathogenicity are subtelomeric regions of the eukaryotic pathogen *Plasmodium falciparum*. These regions are assembled into heterochromatin. They contain around 60 copies of VAR genes required for antigenic variation. Antigenic variation of VAR is an epigenetic survival mechanism. Only one VAR gene is transcribed in individual parasites, while all other members are silenced. Silenced VAR genes are associated with H3K9me3, low levels of acetylation associated with Sir2, low levels of H3K4me, and they are enriched in HP1. Active VAR genes contain high levels of acetylation and H3K4. Mitotic inheritance of these transcriptionally silent states prevents premature presentation of the full antigenic repertoire by the growing population. Furthermore, infrequent switches in mutually exclusive VAR genes give rise to subpopulations that can proliferate in the presence of adaptive immune responses (Voss et al., 2014).

#### **4.2. Plasticity at the human fungal pathogen *Candida albicans***

*Candida albicans* is the most prevalent human fungal pathogen. It normally resides as a harmless commensal within the gut, genito- urinary tract and skin. It is an opportunistic pathogen as it is virulent in individuals with reduced immune systems like AIDS patients, or in situations of an imbalance of competing bacterial microflora (Berman, 2012). It can cause mucosal infections, such as oral thrush or vaginitis. More severe are blood stream infections or candidemia which are associated with high mortality rates. Due to its ability to develop drug resistance it is one of the most common causes of nosocomial infections (Berman, 2012; Pfaller and Diekema, 2007).

*C. albicans* has different mechanisms to survive as a pathogen. One of which is morphological switching. It can grow in at least three different morphologies: yeast, pseudohyphae and hyphae. While yeast form is the classical morphology of budding yeast, pseudohyphae consists of chains of elongated yeast cells with constrictions between adjacent compartments, hyphae form long tube-like structures. *C. albicans* hyphae formation is found during tissue invasion culture experiments, and cells that do not readily form hyphae have reduced virulence. Nevertheless, yeast form is also crucial for dissemination of the pathogen through the blood stream (Berman, 2012; Berman and Sudbery, 2002). *C. albicans* also undergoes phenotypic switching. Yeast cells normally grow as smooth, white dome-shaped colonies known as white state but they can also switch to darker, flat colonies known as opaque. The importance of the opaque state resides in *C. albicans* ability to mate. White cells are unable to mate whereas opaque cells are mating-competent (Bennett and Johnson, 2005). It is thought that the presence of “pimples” in the wall of opaque cells,

but not in white cells, facilitate the cell-cell interactions that occur during mating and the reception of pheromones in a mating quorum sensing response (Berman, 2012; Miller and Johnson, 2002). Frequency of switching between white and opaque states is affected by chromatin modifiers such as the HDAC, Hda1. Deletion of *HDA1* increases the frequency of switching from white to opaque cells (Klar et al., 2001). Likewise, loss of the Set3/Hos2 HDAC complex and the HMT Set1 promotes the white phase. Set3/Hos2 HDAC complex is a key regulator of the master opaque switching *WOR1*. Increasing levels of Wor1 drive the transition from the white to the opaque phase (Hnisz et al., 2009).

Although being an asexual yeast and obligate diploid, *C. albicans* can mate at low frequencies when one of the mating alleles is present in homozygosis (Hull et al., 2000; Magee and Magee, 2000). Recombinants formed by the mating process contain the genetic material of both parents. It requires the parental strains to be homozygous and different at the *MTL* locus. The *MAT* locus encodes proteins similar to the three transcriptional regulators encoded by *S. cerevisiae* *MTL* locus (*MAT $\alpha$* , *MAT $\alpha$ 1* and *MAT $\alpha$ 2*) although *C. albicans* has a fourth gene (*MAT $\alpha$ 1*, *MAT $\alpha$ 2*, *MAT $\alpha$ 1* and *MAT $\alpha$ 2*). An ortholog of  $\alpha$ 2 is missing in *S. cerevisiae*. An important observation is that *MAT $\alpha$*  and *MAT $\alpha$*  are present in single copies ruling out the possibility of a silencing-like mating cassette (Bennett and Johnson, 2005; Hull et al., 2000). Most clinical isolates are heterozygous for the mating-type locus but a small proportion of them are homozygous, suggesting a potential to mate (Berman, 2012).

Cells must undergo a phenotypic switch from the predominant white form to the opaque form to become mating competent. In *S. cerevisiae* diploid cells, the products  $\alpha$ 1 and  $\alpha$ 2 form a heterodimer that represses *MAT $\alpha$*  or *MAT $\alpha$*  specific

genes blocking mating. In *C. albicans* diploid cells, the a1/α2 heterodimer blocks mating by preventing switching from white to opaque cells. a2 activates the a-specific genes, α1 activates the α-specific genes, and a1–α2 act together to repress haploid-specific genes, including mating genes and genes required for the white–opaque transition such as *WOR1* (Huang et al., 2006a). Mating of diploid *C. albicans* cells generates tetraploid cells that must undergo a reducing DNA division back to the diploid state. In most fungi this process occurs by meiosis. However, in *C. albicans* this process occurs through chromosome loss during rounds of mitotic division, thereby forming a parasexual cycle. During mating process *C. albicans* switches from diploid state to a transitional tetraploid state. This tetraploid state reverts to a more stable diploid stable state. It is very frequent that aneuploidies and recombination events occur during the reversion from tetraploid to diploid. This process is known as parasexual cycle and it could generate the diversity that facilitates *C. albicans* adaptation within the host (Bennett and Johnson, 2005; Bennett et al., 2014; Berman and Sudbery, 2002).

#### **4.3. *Candida albicans* genome instability**

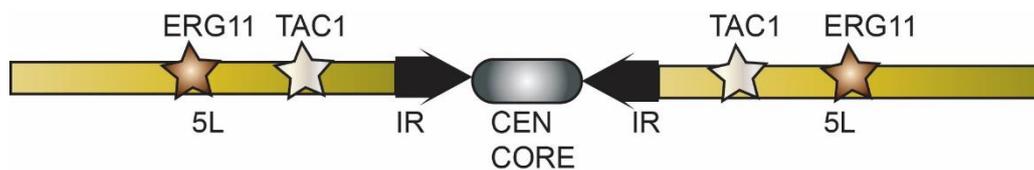
The *C. albicans* genome is composed of 16 Mb of DNA (32 Mb diploid) divided into eight chromosomes. *C. albicans* belongs to the CTG fungi clade. Members of this clade reassigned the CUG codon from leucine to serine through decoding ambiguity which helps to generate genome evolution, phenotypic variation and ecological adaptation. *C. albicans* genome contains a large number of repeats which could contribute to the high genome instability associated to this human fungal pathogen (Bezerra et al., 2013; Santos et al., 1999).

*C. albicans* is a highly successful pathogen due in great part to its genome plasticity and adaptability (Rustchenko, 2007; Selmecki et al., 2006, 2008, 2010, 2009). *C. albicans* genetic diversity is driven by different events including translocations and chromosome rearrangements, chromosome truncations, supernumerary chromosomes (SNCs), whole-chromosome aneuploidy, loss of heterozygosity (LOH) and ploidy changes (Selmecki et al., 2010). Some examples of these events are chromosome truncations like the one observed in the extensively used laboratory strain BWP17 which suffered a chromosome 5 break during disruption of *HIS1*, resulting in two separable size homologs (Noble and Johnson, 2005).

The formation of aneuploidy is a rapid and flexible mechanism. It can confer a selective advantage under conditions of severe stress within one or two divisions. Once the stress condition is not exerted the extra chromosomes are often readily lost in a quick single step. Appearance of isochromosome 5 after fluconazole exposure is a well-studied example. Isochromosome 5, i(5L), is composed by two identical left chromosome arms flanking the centromere (**figure 1.6**). This isochromosome formation confers fluconazole resistance (FluR) and it does not have fitness cost in untreated cells (Ford et al., 2015; Selmecki et al., 2008, 2009).

Fluconazole affects the production of ergosterol that is required for cell wall integrity. Fluconazole target is the ergosterol biosynthetic enzyme lanosterol demethylase encoded by *ERG11*. Fluconazole blocks the production of ergosterol and this blockage leads to the formation of secondary toxic sterols exerting severe membrane stress. *ERG11* is located on Chr5L arm. In this same chromosome arm is located *TAC1*, its product is a transcriptional factor

that activates the transcription of the efflux pump genes *CDR1* and *CDR2* located on chromosome 3. Cdr1 and Cdr2 are multidrug transporters that pump fluconazole and other drugs out the cell. Formation of i(5L) results in an increased copy number of *ERG11* and *TAC1*. As a result, ergosterol biosynthesis is not completely blocked due to extra copies of *ERG11* and fluconazole is pumped out more efficiently due to an increase in copies of *TAC1* (Shapiro et al., 2011) (**Figure 1.6**).

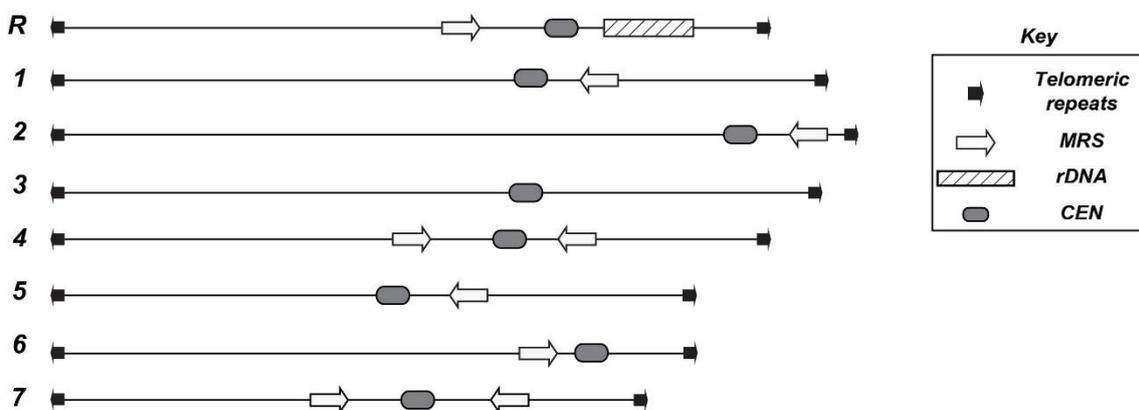


**Figure 1.6.** Isochromosome i(5L) in *C. albicans*. One left arm replaces the 5R arm conferring an increase in the number of copies of certain genes included in the 5L, being of special importance *ERG11* and *TAC1*.

Loss of heterozygosity (LOH) is a very frequent mechanism promoting genetic diversity in *C. albicans*. Due to its diploid state, the major source of genetic variability is provided by the maintenance of heterozygous alleles. LOH occurs mainly through non-reciprocal break induced events (BIR) and one example of its biological consequences is increasing azole resistance due to homozygosis of the hyperactive alleles of *ERG11* and the transcription factors for efflux pumps *TAC1* and *MRR1* (Forche et al., 2011; Rosenberg, 2011; Selmecki et al., 2010).

#### 4.4. *Candida albicans* genomic repeats

The *C. albicans* genome contains repeated DNA sequences: *rDNA*, telomeres, centromeres and MRS (Major Repeated Sequences) (van het Hoog et al., 2007) (**figure 1.7**). Repeats are hotspots for genomic instability driven events. Across eukaryotes, heterochromatin assembled over repeats prevents recombination events as mentioned earlier in the introduction. Nothing was known about the chromatin status at *C. albicans* repeats and how chromatin could control recombination at this fungal pathogen.



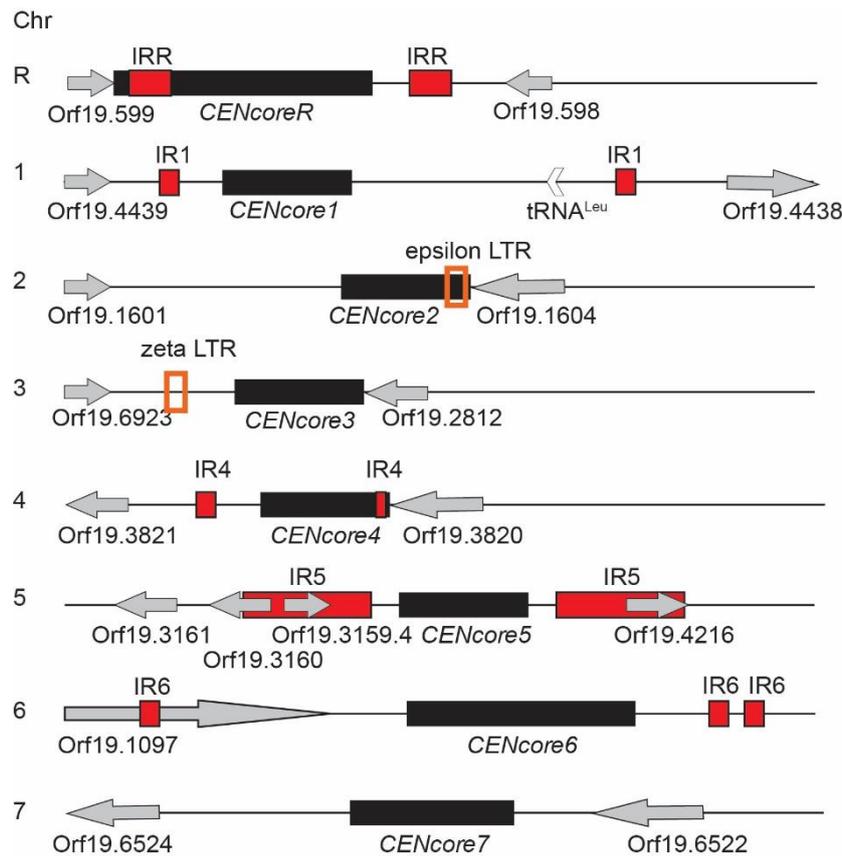
**Figure 1.7. Schematic of *Candida albicans* repeats.** The cluster of ribosomal DNA repeated genes (*rDNA*) is located on chromosome R. Major Repeated Sequences (*MRS*) are located in all chromosomes except chromosome 3. Telomeric repeats are present at the end of each chromosome (Freire-Benítez et al., 2016a).

##### 4.4.1. Centromeres

The *C. albicans* genome has an intermediate type of centromere between budding yeast point centromeres and fission yeast regional centromeres (Hegemann and Fleig, 1993; Henikoff and Henikoff, 2012; Steiner et al., 1993). Each of the eight centromeric DNA sequences is different and unique for every chromosome. They are enriched in the Histone 3 variant Cse4 (CnpA) across a

3 Kb core region (**Figure 1.8**). In most organisms, centromeric DNA is characterized by the presence of flanking DNA repeats. Seven of the eight chromosomes of *C. albicans* centromeres are near short repeat-like sequences (**figure 1.8**). Only centromere of Chr7 does not have any obvious repeated sequences nearby (Sanyal et al., 2004). This is in opposition to what is observed in *C. tropicalis*, where centromeres are flanked by inverted repeats (IR) whose sequences are conserved across centromeres and are determinant for CenpA loading. Repeat-associated centromeres in *C. tropicalis* share a high degree of sequence homology with each other (Chatterjee et al., 2016).

At *C. albicans* chromosome fragments containing naked centromeric DNA sequences are unable to load Cse4 and fail to create an artificial centromere. This suggests the requirement of epigenetic marks associated with centromeres rather than DNA sequence (Baum et al., 2006).



**Figure 1.8. Schematic of *C. albicans* centromeres.** On each chromosome, black squares indicate centromere core regions (*CENcore*). Red squares indicate inverted repeats (IR). Orange empty squares indicate Long Terminal Repeats (LTRs). Empty arrows indicate tRNA codons. Grey arrows indicate ORF and their transcription orientation (Freire-Benítez et al., 2016b).

The epigenetic nature of *C. albicans* centromeres is also supported by the assembly of new functional neocentromeres at Chr5 when a native centromere is depleted. Neocentromere formation is independent of DNA sequence (Ketel et al., 2009). Centromere core synteny is conserved among different *C. albicans* strains (Mishra et al., 2007; Padmanabhan et al., 2008). Cse4 enrichment at centromeres is proportional and dependent on the correct integrity of kinetochore formation, suggesting an underlying epigenetic mechanism

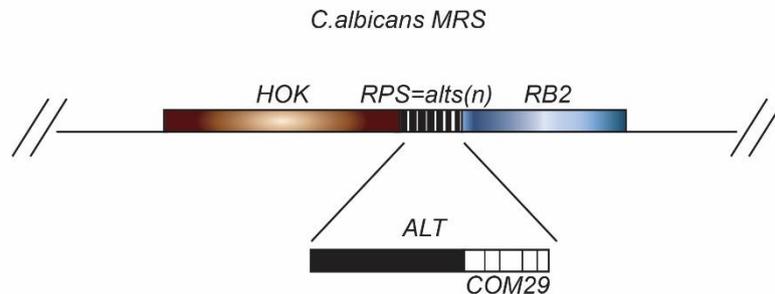
independent of DNA sequence for correct centromere function (Thakur and Sanyal, 2012).

#### **4.4.2. MRS**

The *C. albicans* genome contains a specific type of repeats, the Major Repeat Sequences, *MRS*, present in all chromosomes but chromosome 3. These repeats are also present in *C. dublinensis* but they are not observed in other members of the *Candida* family (van het Hoog et al., 2007; Jackson et al., 2009; Joly et al., 2002). Larger *MRS* affect chromosome stability increasing mitotic non-disjunction rates. A chromosome 5 homolog with a larger *MRS* has a higher rate of chromosome nondisjunction than the same chromosome 5 homolog with a shorter *MRS* (Lephart et al., 2005). *MRS* have also been shown to be hotspots for translocation events in *C. albicans* (Chibana et al., 2000; Iwaguchi et al., 2004).

*MRS* are composed by a region called *RPS* (around 2 Kb) flanked by *HOK* (8 kbp) and *RB2* (6 kbp) regions. The *RPS* are composed themselves by repeated sequences of around 172 bp called *alts*. *Alts* repeats contain at one end between 6-8 copies of a 29 bp conserved palindromic sequence called *COM29* and are very similar in all *MRS*. These repetitive sequences make them hotspots for interchromosomal recombination. As a result, *alts* repeats are responsible for *RPS* variability in length and *MRS* size expansion (**Figure 1.9**) (Chibana et al., 1994, 2000; Iwaguchi et al., 1992; Lephart et al., 2005). *MRS* contain the *FGR6* genes. These genes are present in several copies in the genome and are located within *MRS* repeats and are required for filamentous

growth (Uhl et al., 2003). Interestingly, *S. cerevisiae* lacks any possible ortholog suggesting a role specific for *C. albicans* pathogenicity.



**Figure 1.9. Schematic of *C. albicans* MRS.** *HOK*, *RPS* and *RB2* are the three units forming the *MRS*. The variable in size *RPS* unit, is composed by a repeated number (*n*) of *alts* subunits. Every *alt* subunit contains at one extreme the conserved palindromic sequence *COM29*.

#### 4.4.3. *rDNA*

The *rDNA* cluster is located on chromosome R. It is composed by around 55 copies of *rDNA* units. The size of each unit is around 12 Kb. Each unit contains the Pol I ribosomal gene 35S which is processed into 25S, 18S and 5.8S and the Pol III 5.8S. NTS1 and NTS2 separate the 5S from the 35S and the next *rDNA* unit. The overall identity of the mature *rRNA* sequences to *S. cerevisiae* and other organisms is very high due to the strong evolutionary pressure that has governed ribosome development (Pendrak and Roberts, 2011). In *S. cerevisiae* recombination between *rDNA* repeats leads to the appearance of *rDNA* extra chromosomal circles (ERCs). *C. albicans* accumulates linear and circular *rDNA* extrachromosomal plasmids (Ahmad et al., 2008; Huber and Rustchenko, 2001). The repetitive nature of the *rDNA* cluster makes it very variable in size in *C. albicans* both in clinical and laboratory strains (Hirakawa et

al., 2015; van het Hoog et al., 2007; Magee et al., 1987; Miletti-Gonzalez and Leibowitz, 2008; Rustchenko, 2007).

#### 4.4.4. Telomeres and subtelomeric regions

*C. albicans* telomeres are composed of a 23-bp tandem repeated element (Sadhu et al., 1991). Subtelomeres are composed of two segments with different levels of repetitiveness and divergence. Telomere-proximal sequences include short tandem repeats, whereas, telomere-distal domains encode unique genes, gene families, and repetitive elements of varying frequency (Pryde et al., 1997). Recombination operates extensively at subtelomeres in all eukaryote kingdoms (Louis et al., 1994).

*C. albicans* subtelomeres contain the telomere-associated (*TLO*) gene family. In *C. albicans* there are 14 annotated *TLO* genes and one *TLO* pseudogene, compared to only two *TLO* genes in the closely related oral pathogen *Candida dublinensis* and a single *TLO* in most other candida species (Anderson et al., 2012; Haran et al., 2014; van het Hoog et al., 2007). Tlo proteins function as transcriptional activators of the Mediator complex (Zhang et al., 2012). There are three clades of *TLO* genes:  $\alpha$ ,  $\beta$  and the LTR retrotransposon containing  $\gamma$  clade. The *TLO*  $\alpha$  clade has 6 members, the  $\beta$  clade has one member and the  $\gamma$  clade contains 6 members which encode transcripts with two possible RNA isoforms, either a single exon or a spliced transcript. All of them encode a predicted Med2 domain but not in the case of the *TLO* pseudogene. *TLO*  $\alpha$  clade has the highest levels of expression, while the *TLO*  $\gamma$  clade is much lower expressed. All Tlo proteins from the three clades are detected in the nucleus and the Tlo  $\gamma$  proteins also localize to mitochondria (Selmecki et al., 2009).

In *S. cerevisiae* genes at telomeres are subjected to telomere position effect (TPE). Marker genes inserted at telomere proximal positions are transcriptionally silenced (Allshire et al. 1995; Stavenhagen & Zakian 1998; Schoeftner & Blasco 2009). *C. albicans* subtelomeric *TLO* genes are subjected to TANGEN effect (telomere adjacent gene expression noise) dependent on the HDAC Sir2. In the absence of Sir2, *TLO* genes have more homogeneous gene expression. TANGEN effect at *TLO* genes is dependent on telomere proximity. *TLO* genes showed lower noise levels when they were inserted at an internal locus (Anderson et al., 2014). Through subtelomeric recombination, *TLO* genes generate genotypic diversity. Recombination events are observed at some *TLO* genes, especially due to homology at (C)-terminal coding sequence. The majority of *TLO* members contain at this position the BTS sequence (Bermuda Triangle Sequence) which is a hotspot for LOH or crossover events (Anderson et al., 2015).

#### **4.5. Chromatin factors and pathogenicity in *C. albicans***

As a human fungal pathogen, *C. albicans* counteracts many stresses which promote genotypic and phenotypic changes. One example is the effect of innate immune system of the host to clear *C. albicans* infections (Cheng et al., 2012). One host defense mechanism is the generation of reactive oxygen species (ROS) such as hydroxyl radicals ( $\cdot\text{OH}$ ), superoxide anions ( $\text{O}^{-2}$ ) and hydrogen peroxide ( $\text{H}_2\text{O}_2$ ) within the phagosome. In order to react to this attack, many *C. albicans* genes are up regulated such as oxidative response and repair genes. This happens along with a change in morphology from budding form to hyphal form to provoke mechanical rupture of the phagosome engulfing *C. albicans* cells (Krysan et al., 2014).

Histone post-translation modifications play a role in *C. albicans* pathogenicity. A notable example is acetylation of histone H3 at lysine 56 (H3K56ac), by the HAT *RTT109* (Repressor of Ty-1 Transposition 109). The decay of this mark depends on Sir2, Hst3 and Hst4 in *S. cerevisiae*. Acetylation of H3K56ac promotes histone deposition onto DNA at the replication fork after DSB-induced DNA repair and at promoters during transcriptional activity (Wurtele et al., 2010). Deletion of both alleles of *RTT109* in *C. albicans* and consequent loss of H3K56ac reduces virulence of this pathogen and mortality in mice subjected to systemic candidiasis (Lopes da Rosa et al., 2010). This reduced pathogenicity is due to *C. albicans* cells being more susceptible to reactive oxygen species (ROS) produced by macrophage response due to DNA damage (Lopes da Rosa and Kaufman, 2012; Lopes da Rosa et al., 2010).

White-opaque switching has been shown to be regulated by chromatin modifiers. Deletion mutants of the HMT *SET3*, the HDACs *HOS2* and *HST2* and the HAT *NAT4* show reduced switching frequencies from white to opaque form (Hnisz et al., 2009) while the HDAC *HDA1* deletion promotes the switching from white to opaque phase (Klar et al., 2001).

The master silent regulator HDAC *SIR2* and the HMT *SET1* also play a role in *C. albicans* phenotypic and morphological response to adaptation. Sir2 participates in the control of colony morphologies and karyotypic changes (Pérez-Martín et al., 1999) while Set1 is required for virulence in *C. albicans*. Disruption of Set1 resulted in complete loss of methylation of H3K4. This loss led *C. albicans* cells to hyperfilamentous growth in embedded agar, a reduction in negative cell surface charges, reduced adherence to endothelial cells and reduced virulence in murine infection models (Raman et al., 2006).

## 5. Thesis aim

The aim of this Ph.D thesis is to analyze for the first time the chromatin state associated with *C. albicans* DNA repeats and to understand whether, and how, this chromatin environment controls *C. albicans* genome plasticity. *C. albicans*, like *S. cerevisiae*, lacks an ortholog to SU(VAR)3-9 and most likely H3K9me. *S. cerevisiae* SIR2 (ScSIR2) has a paralog, HST1 (SchST1) (homolog of Sir two), which arose through gene duplication approximately 100 million years ago (Hickman et al., 2011). The mechanism of action of Hst1 differs from Sir2 as it represses single genes acting on promoters while Sir2 acts in larger scale promoting heterochromatin assembly (Mead et al., 2007). *C. albicans* Hst1 (CaHst1) and Sir2 (CaSir2), both share homology with ScSir2. Interestingly, CaHst1 has a higher degree of homology to ScSir2 than CaSir2 (46% vs. 42% amino acid identity respectively) (Freire-Benítez et al., 2016a). This raises the question as to whether CaHst1 and not CaSir2 could be the ortholog of ScSir2. Interestingly, *C. albicans* lacks any other members of the SIR complex of *S. cerevisiae*. The SIR complex acts at telomeres and the mating type locus via Sir2 deacetylation of H3K9 and H4K16 as mentioned in the introduction. The question at the beginning of this Ph.D was whether a homolog of Sir2 could promote heterochromatin assembly in *C. albicans*. On the other hand *C. albicans* genome could be devoid of heterochromatin (**figure 1.10**). The idea of *C. albicans* lacking heterochromatin is supported by the high genome instability associated with this organism.

Heterochromatin suppresses recombination among DNA repeats and promotes genome stability (Grewal and Jia, 2007). Heterochromatin at centromeres promotes proper sister chromatid segregation. At *rDNA*, telomeres and mating

locus of *S. cerevisiae* heterochromatin transcriptionally silences marker genes inserted at these positions. In budding yeast, Sir2 is the master silent regulator responsible for heterochromatin assembly at these three loci (Bühler and Gasser, 2009). Furthermore, SIR2-dependent heterochromatin promoted genome stability at *S. cerevisiae* *rDNA* locus (Kobayashi and Ganley, 2005). *C. albicans*' genome is plastic and undergoes numerous genome rearrangements in order to promptly adapt to environmental stimuli. Whether heterochromatin was assembled at *C. albicans* DNA repeats and whether this chromatin status controlled recombination was not known (**figure 1.10**).

We revealed that *C. albicans* *rDNA* and telomere repeats are assembled into transcriptionally silent heterochromatin. Heterochromatin at telomeres is plastic and remodeled upon environmental changes. Centromeres and MRS display an intermediate status of chromatin, bearing both features of euchromatin and heterochromatin. In addition, we show for the first time that SIR2-dependent heterochromatin is not required for *rDNA* stability whereas it has a major role inhibiting recombination at subtelomeric regions.



## 6. Thesis organization

The work of this Ph.D thesis is published in the form of three papers in peer reviewed journals. The thesis includes six chapters.

**Chapter 1: Introduction.** It includes key aspects of heterochromatin assembly in yeast and some of its roles. It also includes an analysis of *C. albicans* genome, genome instability and adaptability of pathogenic organisms.

**Chapter 2: Chromatin status at *C. albicans* DNA repeats. It contains the article:** Freire-Benítez, V., Price, R.J., Tarrant, D., Berman, J., and Buscaino, A. (2016a). *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.* 6, 22989.

**Chapter 3: Chromatin status at *C. albicans* centromeres. It contains the article:** Freire-Benítez, V., Price, R.J., and Buscaino, A. (2016b). The Chromatin of *Candida albicans* Pericentromeres Bears Features of Both Euchromatin and Heterochromatin. *Front. Microbiol.* 7.

**Chapter 4: Heterochromatin and genome stability at *C. albicans*. It contains the article:** Freire-Benítez, V., Gourlay, S., Berman, J., and Buscaino, A. (2016c). Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen *Candida albicans*. *Nucleic Acids Res.* gkw594.

**Chapter 5: Discussion.** It includes a critical discussion of the goals and findings of the work.

**Chapter 6: Methods.** It includes an extended version of all the methods used.

## Chapter 2. Chromatin status at *C. albicans* DNA repeats

**It contains the article:** Freire-Benéitez, V., Price, R.J., Tarrant, D., Berman, J., and Buscaino, A. (2016a). *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.* 6, 22989.

### Author contributions

VFB performed all the experimental work, new mutant construction, bioinformatics and result analysis show in figures 1 - 5 and supplementary figures 3 – 9. VFB also made supplementary tables.

JRP assisted with some H<sub>2</sub>O<sub>2</sub> silencing assays.

DT assisted in the second allele deletion of *set1Δ/Δ* strain.

JB contributed intellectually to the manuscript and kindly contributed with strains: 20, 44, 45, 70, 83

AB analyzed results, performed bioinformatics and wrote the manuscript.

## 1. Summary

DNA is packed in nucleosomes, complexes formed by DNA and histones, which constitute the so called chromatin. Chromatin can be divided into heterochromatin, a transcriptionally repressive environment normally assembled over repeats, or euchromatin, transcriptionally active chromatin. In budding yeast, heterochromatin is characterized by nucleosomes that are hypoacetylated and are hypomethylated on lysine 4 of histone 3 (H3K4), while nucleosomes at euchromatin are acetylated and are methylated at H3K4. In this chapter, we address how chromatin is assembled at the DNA repeats of *Candida albicans*, the most common human fungal pathogen.

We show that the *rDNA* locus and telomeric regions are assembled into transcriptionally silent heterochromatin. Sir2 is required for this silenced state by deacetylation of lysine 9 at histone 3 (H3K9) at *rDNA* locus and deacetylation of H3K9 together with lysine 16 at histone 4 (H4K16) at telomeres. RNA-sequencing analyses revealed that coding and non-coding transcripts located at the *rDNA* locus, telomeric and subtelomeric regions are repressed in a SIR2-dependent manner. In contrast, *MRS* (Major repeated sequences), a specific type of *C. albicans* DNA repeats, are assembled into transcriptionally permissive chromatin. This chromatin state is associated with high levels of acetylation and low levels of methylation. Silencing assays failed to detect transcriptional silencing at *MRS*. Only qRT-PCR assays detected low levels of transcription of a marker gene inserted at these repeats, defining a permissive state of the chromatin. Moreover, RNA-sequencing analyses confirmed that expression of genes located in proximity of *MRS* repeats is not regulated by Sir2.

One of the most remarkable findings showed that, SIR2-dependent heterochromatin at telomeres is plastic. Thermic switch from 30°C to 39°C enhances telomeric heterochromatin, reducing acetylation levels in a SIR2-dependent manner and increasing transcriptional silencing of a marker gene inserted at these regions.

In conclusion, in this chapter, we reveal the *C. albicans* repetitive elements are assembled into distinct chromatin states and that telomeric heterochromatin can be remodeled upon environmental changes.

## 2. Main article

# SCIENTIFIC REPORTS

OPEN

## *Candida albicans* repetitive elements display epigenetic diversity and plasticity

Received: 13 January 2016  
Accepted: 25 February 2016  
Published: 14 March 2016

Verónica Freire-Benítez<sup>1</sup>, R. Jordan Price<sup>1</sup>, Daniel Tarrant<sup>1</sup>, Judith Berman<sup>2</sup> & Alessia Buscaino<sup>1</sup>

Transcriptionally silent heterochromatin is associated with repetitive DNA. It is poorly understood whether and how heterochromatin differs between different organisms and whether its structure can be remodelled in response to environmental signals. Here, we address this question by analysing the chromatin state associated with DNA repeats in the human fungal pathogen *Candida albicans*. Our analyses indicate that, contrary to model systems, each type of repetitive element is assembled into a distinct chromatin state. Classical Sir2-dependent hypoacetylated and hypomethylated chromatin is associated with the rDNA locus while telomeric regions are assembled into a weak heterochromatin that is only mildly hypoacetylated and hypomethylated. Major Repeat Sequences, a class of tandem repeats, are assembled into an intermediate chromatin state bearing features of both euchromatin and heterochromatin. Marker gene silencing assays and genome-wide RNA sequencing reveals that *C. albicans* heterochromatin represses expression of repeat-associated coding and non-coding RNAs. We find that telomeric heterochromatin is dynamic and remodelled upon an environmental change. Weak heterochromatin is associated with telomeres at 30 °C, while robust heterochromatin is assembled over these regions at 39 °C, a temperature mimicking moderate fever in the host. Thus in *C. albicans*, differential chromatin states controls gene expression and epigenetic plasticity is linked to adaptation.

Large blocks of DNA repeats are commonly clustered at rDNA loci, telomeres and centromeres and are assembled into heterochromatin. Heterochromatic regions impose a transcriptionally repressive environment that can propagate over long distances (up to 50 kb) stochastically silencing native genes as well as reporter genes inserted at these regions independently of the underlying DNA sequence<sup>1–4</sup>. Transcriptionally repressive heterochromatin is distinguishable from transcriptionally active euchromatin by several epigenetic features where heterochromatin is characterised by nucleosomes that are hypoacetylated, hypomethylated on lysine 4 of histone H3 (H3K4) and methylated on lysine 9 of histone H3 (H3K9)<sup>5,6</sup>. Histone modifiers control the transcriptionally repressive state of heterochromatin regions via chromatin modifications. For example, the conserved histone deacetylase Sir2 controls the hypoacetylated state of heterochromatic regions while the histone methyltransferase Su(var)3–9 specifically methylates H3K9<sup>1,2,7–11</sup>. Heterochromatin has the ability to propagate thanks to specific silencing complexes. For example, in *Saccharomyces cerevisiae*, assembly of subtelomeric heterochromatin is mediated by the Sir silencing complex formed by Sir2, Sir3 and Sir4<sup>1</sup>. Likewise, heterochromatin assembly at the rDNA locus is dependent on the RENT (REgulator of Nucleolar silencing and Telophase) complex formed by Sir2, Net1 and Cdc14<sup>12–14</sup>. Heterochromatin modulation could be particularly important for organisms, such as microbial pathogens, that have to adapt rapidly to different environments. This is because heterochromatin can modulate gene expression without changes in DNA sequence. However, very little is known about heterochromatin-mediated transcriptional regulation in this group of organisms. Here we address this question by investigating the chromatin states associated with DNA repeats in the most common human fungal pathogen: *Candida albicans*. *C. albicans* normally lives as a commensal in humans but it can become virulent causing systemic life-threatening infections with mortality rates of up to 50%<sup>15</sup>. Upon environmental changes, *C. albicans* undergoes major morphological and genomic changes that can promote adaptation and survival<sup>1,16</sup>. It is unknown whether the chromatin structure of the *C. albicans* genome is also plastic and capable of remodeling upon environmental changes.

<sup>1</sup>University of Kent, School of Biosciences Canterbury Kent, CT2 7NJ. UK. <sup>2</sup>Department of Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, 69978, Israel. Correspondence and requests for materials should be addressed to A.B. (email: A.Buscaino@kent.ac.uk)

The *C. albicans* genome has 8 diploid chromosomes and contains 3 major classes of large blocks of repetitive DNA: the rDNA locus, Major Repeat Sequences (MRS) and telomeres<sup>17,18</sup> (Fig. S1A). The rDNA locus consists of a tandem array of a ~12 kb unit repeated 50 to 200 times on chromosome R. Each unit contains the two highly conserved 35 S and the 5 S rRNA genes that are separated by two Non-Transcribed Regions (NTS1 and NTS2), whose sequences are not conserved across eukaryotes (Fig. S1B)<sup>17,18</sup>. In other organisms, while the 35S and 5S rDNA genes are highly expressed, the NTS1 and NTS2 regions are assembled in transcriptionally silent heterochromatin<sup>12–14</sup>. At these locations, heterochromatin represses transcription of non coding RNAs promoting stability of the rDNA loci<sup>19,20</sup>. The MRS loci are long tracts (10–100 kb) of nested DNA repeats found on 7 of the 8 *C. albicans* chromosomes<sup>19,21</sup>. These repetitive domains are formed by large tandem arrays of 2.1 kb RPS unit flanked by non-repetitive HOK and RBP-2 elements (Fig. S1C). In addition, a RBP-2 element, but not an intact MRS, is found on chromosome 3<sup>22</sup>. Given their highly repetitive nature, MRS repeats are expected to be ideal substrates for heterochromatin assembly. *C. albicans* telomeres are composed of a terminal region composed of tandemly repeating 23 bp units and subtelomeric regions (Fig. S1D)<sup>23</sup>. Due to their size and repetitive nature, the sequences of subtelomeric regions remain poorly characterised<sup>18</sup>.

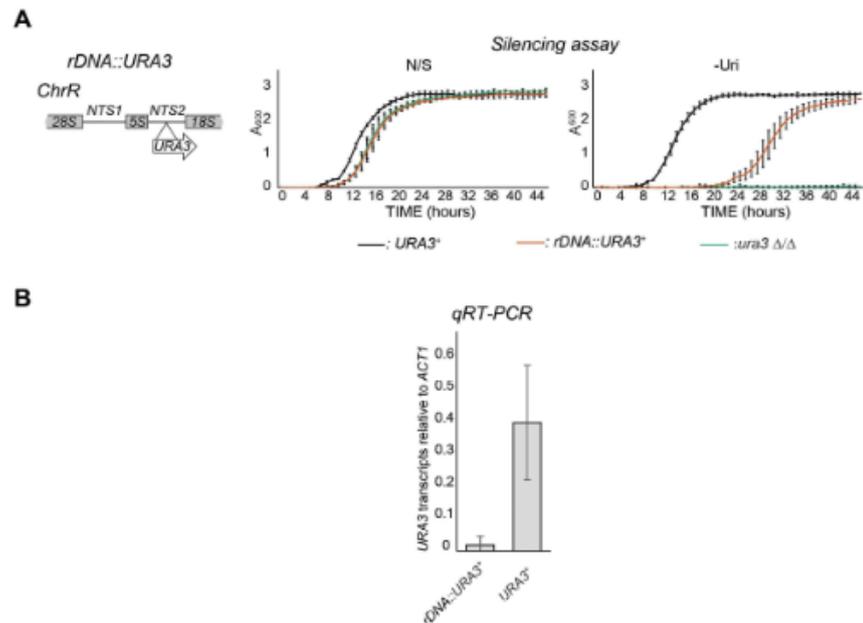
DNA repeats are central to *C. albicans* genome plasticity and pathogenicity as they play key roles in the regulation of genome organisation and structure in the host<sup>16</sup>. However, the chromatin state of these DNA elements is unknown. The *C. albicans* epigenome, like the *S. cerevisiae* epigenome, most probably lacks H3K9 methylation given that a Su(var) 3–9 orthologue cannot be identified in either *C. albicans* or *S. cerevisiae*. In contrast, the *C. albicans* genome encodes for two putative NAD-dependent histone deacetylases that resemble Sir2 (*orf19.1992* and *orf19.4762*)<sup>24</sup>. While it is possible to identify orthologues for RENT complex components, components of the Sir silencing complex are not apparent. Therefore, while heterochromatin might associate with the rDNA locus, it is possible that *C. albicans* telomeres lack heterochromatin as it has been recently shown for the yeast *Clavispora lusitanae*<sup>11</sup>.

In this study, we investigated the chromatin states associated with *C. albicans* repetitive elements under different environmental conditions. We found that the different types of repetitive elements are assembled in distinct chromatin states. Classical hypoacetylated heterochromatin is associated with the non-transcribed region of the rDNA locus. The histone deacetylase Sir2 (*orf19.1992*) is required to maintain this repressive epigenetic state via hypoacetylation of Lysine 9 of Histone H3. Heterochromatin associated with these regions represses non-coding RNA transcription. In contrast, MRS repeats are assembled into a transcriptionally permissive chromatin state bearing both heterochromatic and euchromatic histone marks. Finally, we find that, despite the apparent absence of the Sir silencing complex, telomeric regions are assembled into a Sir2-dependent hypoacetylated and hypomethylated heterochromatin. This chromatin state silences expression of endogenous transcripts as well as inserted marker genes. Telomeric heterochromatin is plastic and affected by environmental conditions with heterochromatin being more robust at higher temperature (39 °C) than at lower temperature (30 °C). Thus, the epigenetic state associated with telomeric repeats switches in response to an environmental change that are linked to *C. albicans* virulence and pathogenicity.

## Results

**Transcriptional silencing at the *C. albicans* rDNA locus.** Heterochromatin assembled onto repetitive DNA represses the transcription of marker genes inserted in their proximity<sup>25–28</sup>. To assess whether transcriptionally silent chromatin exists in *C. albicans*, the *URA3<sup>+</sup>* marker gene was integrated into the NTS2 region of the rDNA locus (*rDNA::URA3<sup>+</sup>*) (Fig. 1A). We investigated whether *URA3<sup>+</sup>* was transcriptionally silenced when present at this locus, by growing strains in non-selective (N/S) medium and in medium lacking uridine (–Uri) in which only cells expressing sufficient Ura3 protein are able to grow. Silencing of *URA3<sup>+</sup>* is expected to result in slower growth in –Uri medium compared to N/S. The *rDNA::URA3<sup>+</sup>* strain displayed a reduced growth rate in selective medium and it is, therefore, silenced (Fig. 1A). Consistent with the growth assays the level of *URA3* mRNA levels at the rDNA locus were significantly lower (20 fold) for *rDNA::URA3<sup>+</sup>* than for a *URA3<sup>+</sup>* gene expressed from its own euchromatic locus, as determined by quantitative reverse transcriptase analysis (qRT-PCR) (Fig. 1B). Therefore, in *C. albicans*, the non-transcribed region of the rDNA locus is assembled into a transcriptionally repressed state that is normally associated with repetitive heterochromatic regions.

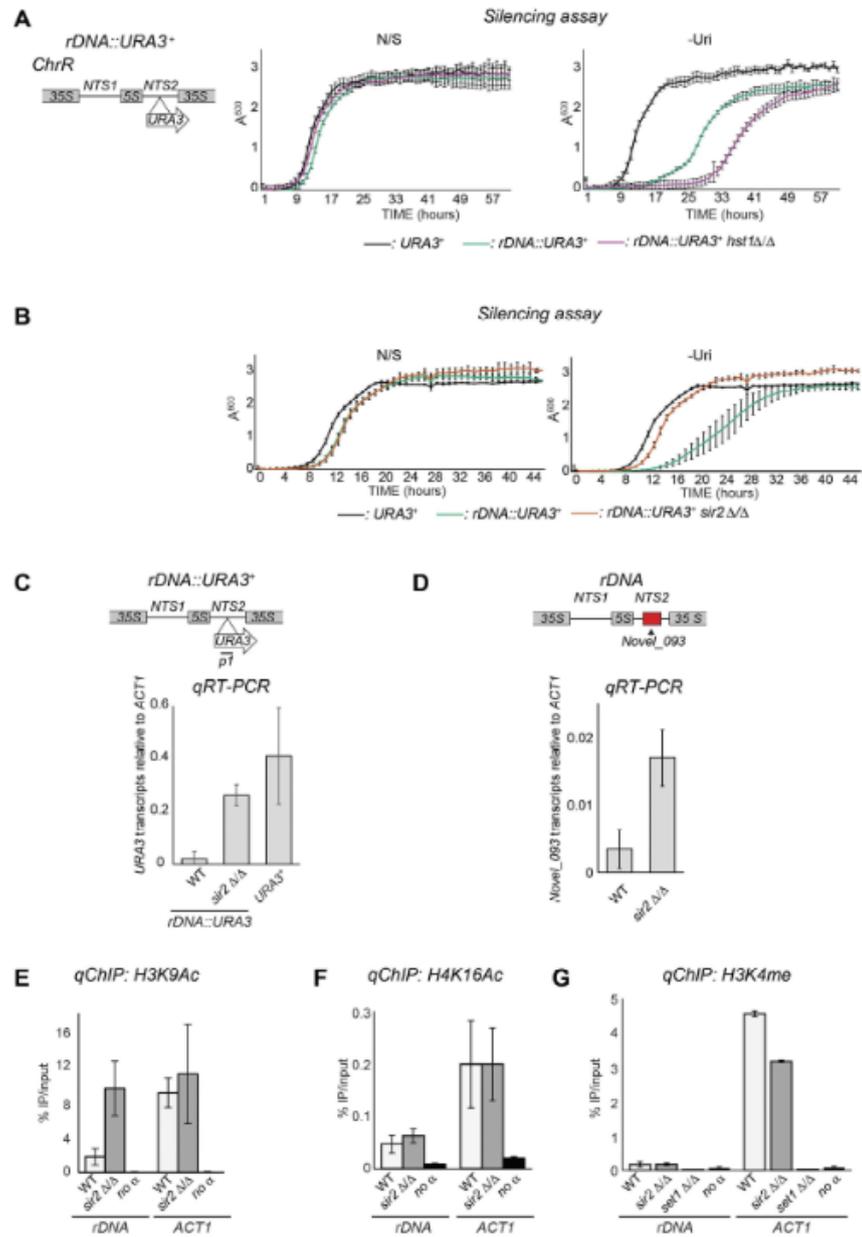
**Sir2-dependent heterochromatin at the rDNA locus.** The ability of *C. albicans* rDNA repeats to cause transcriptional silencing suggests that the rDNA NTS regions are assembled into heterochromatin. The histone deacetylase Sir2 is a key regulator of heterochromatin in all organisms studied<sup>1–29</sup>. BLAST analyses reveal that the *C. albicans* genome contains 5 genes encoding proteins with homology to *S. cerevisiae* SIR2 (Fig. S2). Among these proteins, *C. albicans* Hst1 (*orf19.4762*) and Sir2 (*orf19.1992*) share the highest homology with the *S. cerevisiae* Sir2 (46% and 42% identity respectively) (Fig. S2A,B). Given the high homology, it is possible that *C. albicans* Hst1, and not Sir2, is the true ortholog of *S. cerevisiae* Sir2 as it has been demonstrated for the yeast *Clavispora lusitanae*<sup>11</sup>. To determine whether Hst1 (*orf19.4762*) and/or Sir2 (*orf19.1992*) are required for maintaining the repressive state of the rDNA NTS region, we constructed *hst1*  $\Delta/\Delta$  and *sir2*  $\Delta/\Delta$  null mutants in strains carrying the *rDNA::URA3<sup>+</sup>* reporter and performed marker gene silencing assay. At the NTS region of the rDNA locus, silencing was not alleviated in *hst1*  $\Delta/\Delta$  cells (Fig. 2A) but it was strongly alleviated in *sir2*  $\Delta/\Delta$  cells (Fig. 2B). In agreement with these results, qRT-PCR analyses revealed that *URA3* mRNA levels in the *rDNA::URA3<sup>+</sup>* strain were significantly higher in *sir2* null cells relative to wild-type (WT) cells (Fig. 2C). Thus, the histone deacetylase Sir2, but not Hst1, is critical for the maintenance of the transcriptionally silent state associated with rDNA repeats in *C. albicans*. In *S. cerevisiae*, heterochromatin associated with the NTS region of the rDNA locus has been shown to repress transcription of a non-coding RNA<sup>20,30,31</sup>. In *C. albicans*, genome-wide analyses have identified an uncharacterised non-coding RNA (Novel\_Ca21Chr\_093) originating from the NTS region of the rDNA locus<sup>32</sup>. To test whether *C. albicans* heterochromatin represses transcription of this non-coding RNA, we isolated



**Figure 1. Transcriptional silencing at the *C. albicans* rDNA locus.** (A) *Left panel*: Schematic of *rDNA::URA3*<sup>+</sup> reporter strain. *Right panel*: silencing assay of the *rDNA::URA3*<sup>+</sup> reporter strain. *Ura*<sup>+</sup> (*URA3/URA3*) and *Ura*<sup>-</sup> (*ura3 $\Delta$ /ura3 $\Delta$* ) strains were included as controls. (B) *qRT-PCR* analyses to measure *URA3*<sup>+</sup> transcript levels of the *rDNA::URA3*<sup>+</sup> reporter strain relative to actin transcript levels (*ACT1*). Error bars in each panel: Standard deviation (SD) of three biological replicates.

RNA from wild-type and *sir2*  $\Delta/\Delta$  cells and performed *qRT-PCR* analyses. Deletion of the *SIR2* gene results in a clear upregulation of this non-coding transcript (Fig. 2D). Therefore, Sir2 is required to repress transcription of an endogenous transcript as well as inserted marker genes at the rDNA locus. To assess whether the NTS region of the rDNA locus is associated with heterochromatic pattern of histone marks (hypocetylation of H3K9/H4K16 and hypomethylation of H3K4), we monitored the presence of these histone modifications by quantitative Chromatin Immunoprecipitation (*qChIP*)<sup>33</sup>. The rDNA locus, but not the euchromatic *ACT1* locus, showed low enrichment for H3K9 acetylation, H4K16 acetylation and H3K4 methylation (Fig. 2E–G), a chromatin state typical of heterochromatic regions. Sir2 deacetylates lysine 9 on histone H3 and/or lysine 16 on histone H4<sup>8,34,35</sup>. Therefore, we compared the level of histone acetylation associated with the NTS rDNA region in WT and *sir2*  $\Delta/\Delta$  cells by performing *qChIP* analyses. In the *sir2* null strain we detected higher levels of H3K9Ac, but not of H4K16Ac (Fig. 2E,F) demonstrating that *C. albicans* Sir2 is required to maintain low levels of acetylated H3K9. This is consistent with the idea that disruption of H3K9 deacetylation is critical for maintaining the transcriptionally silent state associated with heterochromatic regions. We also compared the level of H3K4 methylation associated with the NTS region of the rDNA locus in WT and *sir2*  $\Delta/\Delta$  isolates. H3K4 methylation levels did not increase in *sir2*  $\Delta/\Delta$  isolates compared to wt cells (Fig. 2G). Therefore hypomethylation of H3K4 is maintained independently of the histone acetylation and transcriptional state of the NTS region. Taken together these observations demonstrate that heterochromatin exists in *C. albicans* and it is assembled over the NTS region of the rDNA locus. Histone modification by Sir2 is critical for the maintenance of the transcriptionally silent heterochromatic state associated with this locus.

**MRS repeats are assembled into transcriptionally permissive chromatin bearing euchromatic and heterochromatic histone modifications.** Having established that transcriptionally silent heterochromatin exists in *C. albicans* and is associated with the NTS region of the rDNA locus, we analysed the chromatin state associated with the MRS repeats by assessing silencing of a marker gene inserted into the tandem RPS repeats (*MRS::URA3*<sup>+</sup>) (Fig. 3A). As shown in Fig 3A, the *MRS::URA3*<sup>+</sup> is not silenced (Fig. 3A). If the MRS repeats are associated with Sir2-dependent heterochromatin, the genes close to these regions should be upregulated in *sir2*  $\Delta/\Delta$  isolates compared to WT cells. To address this question, we isolated RNA from WT and *sir2*  $\Delta/\Delta$  cells and performed RNA-seq analyses. FPKM (fragments per kilobase of exons per million mapped reads) were determined for all the genes proximal to MRS repeats and compared in *sir2*  $\Delta/\Delta$  and WT strains. Upon deletion of the *SIR2* gene, we did not observe any clear effect on expression of MRS associated (orf C) and proximal genes (orf L and R) as only 3 out of 24 genes were expressed at more than 2 fold of WT in



**Figure 2.** Sir2-dependent heterochromatin at the rDNA locus. (A) *Left panel:* Schematic of *rDNA::URA3<sup>+</sup>* reporter strain. *Right panel:* Silencing assay of the *rDNA::URA3<sup>+</sup>* reporter strain in WT and *hst1*  $\Delta/\Delta$  isolates. A *URA<sup>+</sup>* (*URA3<sup>+</sup>*) strain was included as a control. (B) Silencing assay of the *rDNA::URA3<sup>+</sup>* reporter strain WT and *sir2*  $\Delta/\Delta$  isolates. A *URA<sup>+</sup>* (*URA3<sup>+</sup>*) strain was included as a control. (C) *qRT-PCR* analyses to measure *URA3<sup>+</sup>* transcript levels relative to *ACT1* transcript levels in *rDNA::URA3<sup>+</sup>* WT and *sir2*  $\Delta/\Delta$  isolates. A *URA3<sup>+</sup>* strain is included as a control. (D) *qRT-PCR* analyses to measure levels of the rDNA non-coding RNA (*Novel\_Chr3\_093*) relative to *ACT1* in WT and *sir2*  $\Delta/\Delta$  isolates. (E,F) *qChIP* to detect H3K9Ac, H4K16Ac levels associated with the rDNA locus and *ACT1* in WT and *sir2*  $\Delta/\Delta$  isolates and (G) H3K4me2 levels associated with the rDNA locus and *ACT1* in WT, *sir2*  $\Delta/\Delta$  and *set1*  $\Delta/\Delta$  isolates. Error bars in each panel: Standard deviation (SD) of three biological replicates.



*sir2*  $\Delta/\Delta$  isolates (Fig. 3B and Table S5). Averaging the log<sub>2</sub> fold change of the MRS-associated genes confirms these results (average log<sub>2</sub> fold change = 0) (Fig. 3C). These data demonstrate that MRS repeats do not impose a Sir2-dependent transcriptionally repressive state.

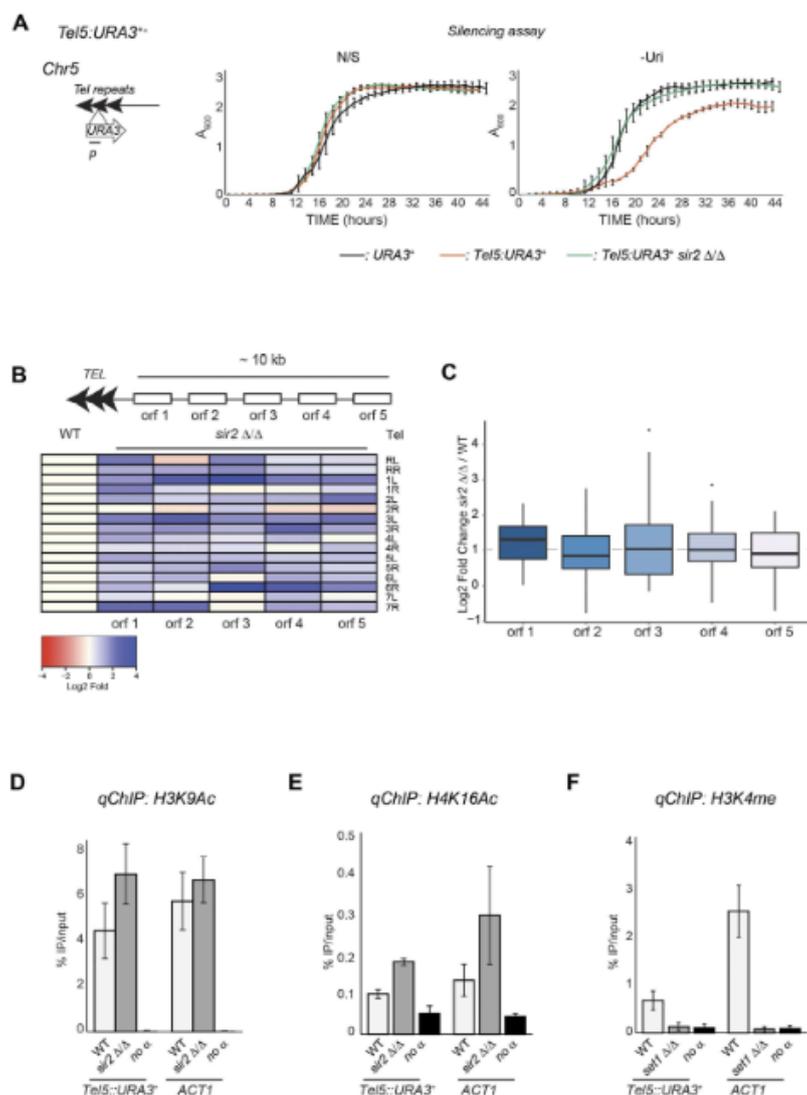
Acetylated H3K9 and H4K16 *qChIP* analyses demonstrated that MRSs are assembled into highly acetylated chromatin where H3K9 and H4K16 are acetylated to a level similar to the active and euchromatic locus actin 1 (*ACT1*) (Fig. 3D,E). Histone acetylation level associated with these regions are similar in WT and *sir2*  $\Delta/\Delta$  isolates (Fig. 3D,E). In contrast, levels of H3K4 methylation associated with MRS repeats are strikingly low compared to the euchromatic *ACT1* locus (Fig. 3F). Therefore, MRS repeats are not assembled into classical transcriptionally silent heterochromatin but they are associated with an intermediate chromatin state bearing features of euchromatin (high histone acetylation) and heterochromatin (H3K4 hypomethylation).

**Heterochromatin at *C. albicans* telomeric regions.** In many organisms, heterochromatin assembly over telomeric regions is dependent not only on Sir2 but also on the Sir silencing-complex<sup>1,2</sup>. With the exception of Sir2, BLAST analyses fail to identify orthologues of telomeric silencing proteins in *C. albicans*. It is therefore possible that telomeres are not assembled into heterochromatin. To test this hypothesis, we analysed silencing of a *URA3*<sup>+</sup> marker gene integrated into a telomeric region (*Tel5::URA3*<sup>+</sup>). The *Tel5::URA3*<sup>+</sup> marker gene displayed a small but reproducible reduction in growth rate on -Uri media compared to N/S media, indicative of weak silencing (Fig. 4A). Sir2 is required to maintain this transcriptionally repressive state as silencing is alleviated in *sir2*  $\Delta/\Delta$  cells (Fig. 4A). Consistent with the growth assays, the levels of *URA3*<sup>+</sup> mRNA levels are low in WT cells and dramatically increased in *sir2*  $\Delta/\Delta$  isolates (Fig. S3).

Telomeric heterochromatin has been shown to repress gene expression of proximal genes in a Sir2-dependent manner<sup>26–29</sup>. To assess whether the weak marker gene silencing associated with telomere 5 is a general property of all *C. albicans* telomeres and if it extends over subtelomeric regions, we analysed the transcriptional profile of coding and non-coding subtelomeric transcripts in WT and *sir2*  $\Delta/\Delta$  isolates. This analysis reveals that deletion of *SIR2* results in transcriptional upregulation of many subtelomeric genes (Fig. 4B and Table S6). Although not all the subtelomeric genes are upregulated to the same extent, on average deletion of Sir2 results in a 2 fold upregulation of telomeric-proximal genes compared to WT (Fig. 4C). We used *qRT-PCR* analyses with primers specific for each subtelomeric gene to validate the difference in gene expression between WT and *sir2*  $\Delta/\Delta$  isolates. All the genes tested were more highly expressed in *sir2*  $\Delta/\Delta$  compared to WT cells (Fig. S4). Therefore, telomeric heterochromatin silences expression of genes located in proximity (~10/15 kb) of telomeres. Telomeres are composed of a 23 bp unit tandemly repeated<sup>35</sup>. Their repetitive nature makes the design of suitable primers for *qChIP* analysis particularly challenging. Thus, we assessed the telomeric chromatin state by performing *qChIP* analyses with primers specific for the *Tel5::URA3*<sup>+</sup> marker gene. We found that telomeric chromatin is only mildly hypoacetylated on H3K9 and H4K16 and that histone acetylation levels are increased in *sir2*  $\Delta/\Delta$  null mutant compared to WT cells (Fig. 4D,E). In contrast, a low level of H3K4 methylation is associated with telomeric regions (Fig. 4F).

Taken together these observations demonstrate that telomeric repeats are assembled into weak heterochromatin. This chromatin state is dependent on Sir2 and is able to silence embedded marker genes as well as native proximal genes.

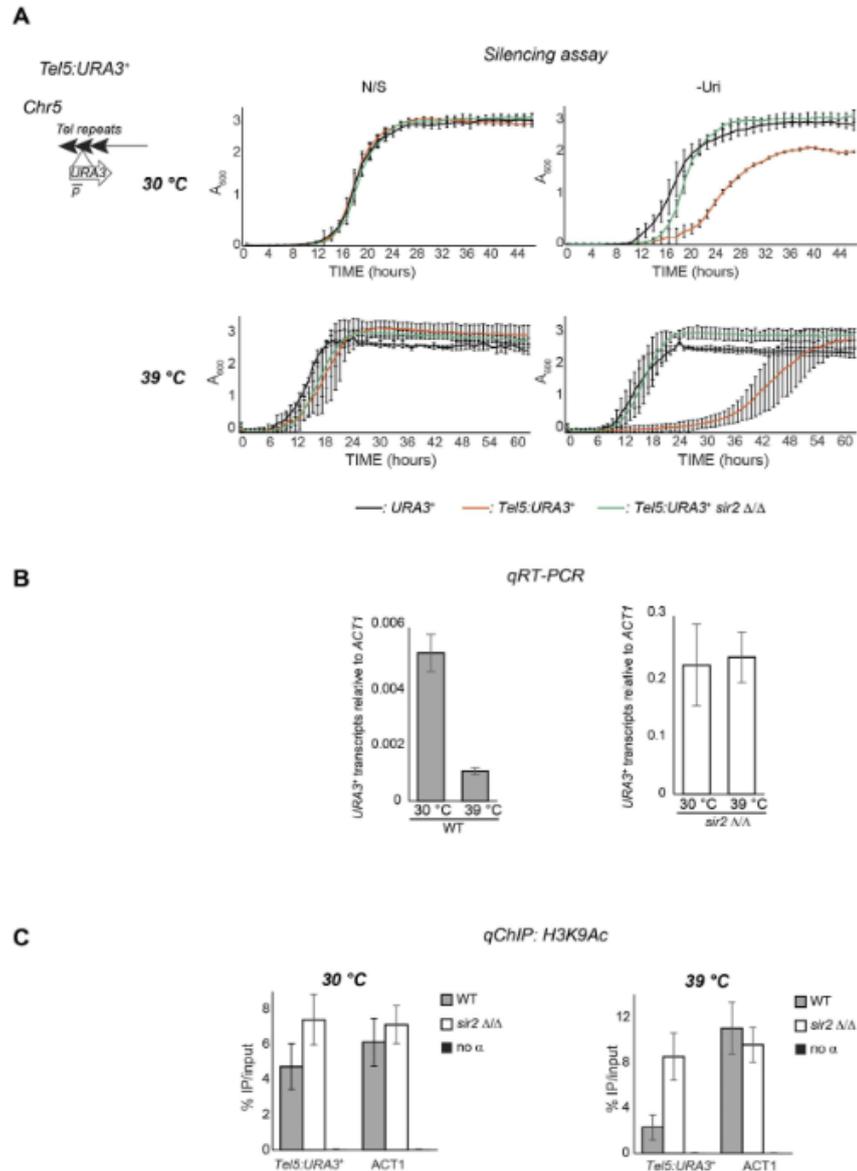
**Telomeric heterochromatin is plastic and remodelled upon environmental changes.** *C. albicans* is characterised by remarkable genomic and phenotypic plasticity that allow rapid adaptation to different environmental niches<sup>4,16</sup>. Therefore, we tested whether the chromatin state of *C. albicans* repetitive elements is also plastic and remodelled upon environmental changes. Remodelling could lead to dynamic chromatin structure where DNA repeats are assembled into robust transcriptionally repressive heterochromatin under specific environmental conditions and into weak heterochromatin under different environmental conditions. To test this hypothesis, we asked if physiologically-relevant stress conditions that *C. albicans* regularly encounters in the host or stress caused by antifungal agents affects the transcriptional states associated with the NTS region of the rDNA locus, telomeric repeats and MRS repeats. Silencing assays revealed that treatment with hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), mimicking the production of reactive oxygen species by the host's immune cells, or Fluconazole, the most widely used antifungal drug, does not change the transcriptional state associated with any of the loci tested (Figs S5, S6 and S7). Likewise we found that expression of the *URA3*<sup>+</sup> marker gene inserted at the rDNA locus and MRS repeats was not affected by high temperature (39 °C, mimicking moderate fever in the host) (Figs S5 and S6). In contrast, telomere-associated silencing was much stronger at 39 °C than 30 °C (Fig. 5A) indicating the telomeric chromatin is plastic. Growing cells at 39 °C did not strongly induce hyphal formation in the time frame of the experiment (Fig. S8). Therefore, the observed stronger silencing is not a consequence of a different morphology. To assess whether temperature-dependent silencing is a general feature of all *C. albicans* telomeres, we analysed silencing of a second reporter strain with a *URA3*<sup>+</sup> gene integrated at the telomeric repeats of chromosome 7 left arm (*Tel7::URA3*<sup>+</sup>). As with the telomere repeat tract on Chr5, silencing of telomeric repeats on Chr7 was much stronger at 39 °C compared to 30 °C (Fig. S8). Importantly, silencing was dependent on Sir2 as silencing growth rate in -Uri was increased in *sir2*  $\Delta/\Delta$  cells (Fig. 5A). These results were confirmed by analyses of *URA3*<sup>+</sup> mRNA levels by *qRT-PCR*: in WT, but not in *sir2*  $\Delta/\Delta$  cells, *Tel5::URA3*<sup>+</sup> RNA levels were lower at 39 °C compared to 30 °C (Fig. 5B). Consistent with the silencing assays, *qChIP* of the *TEL5::URA3*<sup>+</sup> marker revealed lower levels of H3K9 acetylation at 39 °C compared to 30 °C (Fig. 5C). Thus telomere repeats are silenced to a greater degree at febrile temperature than at 30 °C, a temperature considered relevant for growth of *C. albicans* on the skin. These results indicate that telomeric heterochromatin is dynamic and can be remodelled in a Sir2-dependent manner in response to an environmental change.



**Figure 4. Heterochromatin at *C. albicans* telomeric regions.** *Left panel:* Schematic of *Tel5::URA3<sup>+</sup>* reporter strain. *Right panels:* Silencing assay of the *Tel5::URA3<sup>+</sup>* reporter strain in WT and *sir2  $\Delta/\Delta$* . (A) *URA<sup>+</sup> (URA3<sup>+</sup>)* strain was included as a control. (B,C) RNA deep-sequencing of *sir2  $\Delta/\Delta$*  and WT isolates. (B) Normalised read counts (FPKM) of subtelomeric genes were calculated from RNA-seq data for WT and *sir2  $\Delta/\Delta$*  isolates. The heat-map depicts the log<sub>2</sub> fold ratio of FPKM data between *sir2  $\Delta/\Delta$*  and WT isolates. (C) Boxplot showing log<sub>2</sub> fold changes in transcriptional expression of subtelomeric genes between *sir2  $\Delta/\Delta$*  and WT isolates (D–F) qChIP to detect H3K9Ac and H4K16ac levels associated with *Tel5::URA3<sup>+</sup>* and *ACT1* in WT and *sir2  $\Delta/\Delta$*  strains. (C) H3K4me2 levels associated with *Tel5::URA3<sup>+</sup>* and *ACT1* in WT and *set1  $\Delta/\Delta$*  isolates. Error bars in each panel: Standard deviation (SD) of three biological replicates.

### Discussion

This study provides the first comprehensive analysis of the chromatin state of *C. albicans* DNA repeats. Our data demonstrate that, in *C. albicans*, differential chromatin states control gene expression and epigenetic plasticity is linked to adaptation to a specific environmental niche.



**Figure 5. Telomeric heterochromatin is plastic.** (A) *Left panel:* Schematic of *Tel5::URA3<sup>+</sup>* reporter strain. *Right panels:* Silencing assay assessing transcriptional silencing of the *Tel5::URA3<sup>+</sup>* reporter strain in WT and *sir2  $\Delta/\Delta$*  isolates at 30 °C, in the presence of 1 mM H<sub>2</sub>O<sub>2</sub>, 200 ng/μl fluconazole and at 39 °C. A *URA<sup>+</sup>* (*URA3<sup>+</sup>*) strain was included as a control. (B) qRT-PCR analyses to measure *Tel5::URA3<sup>+</sup>* transcript levels relative to *ACT1* at 30 °C and 39 °C in WT (left panel) and *sir2  $\Delta/\Delta$*  (right panel) strains. (C) qChIP to detect H3K9Ac levels associated with *Tel5::URA3<sup>+</sup>* and *ACT1* in WT and *sir2  $\Delta/\Delta$*  isolates at 30 °C (Left panel) and 39 °C (Right panel). Error bars in each panel: Standard deviation (SD) of three biological replicates.

We find that the NTS region of the *rDNA* locus maintains a *bona fide* silent heterochromatin state that represses coding and non-coding transcription and is marked by the histone modification pattern typical of heterochromatic regions: hypoacetylation of histone H3K9 and histone H4K16 together with hypomethylation of

H3K4 (Fig. 1). We identify Sir2 (*orf19.1992*) as one of the key enzymes necessary to maintain this chromatin state (Fig. 2). In *S. cerevisiae*, Sir2 targeting to the rDNA locus is dependent on the RENT complex<sup>13,14,49</sup>. BLAST analyses identify orthologues of the RENT complex components Net1 (*orf19.267*) and Cdc14 (*orf19.4192*). Therefore, it is very likely that, as observed in *S. cerevisiae*, a conserved RENT complex targets heterochromatin at the *C. albicans* rDNA locus. However, the mode of action of Sir2 in *C. albicans* differs from *S. cerevisiae*: while at *S. cerevisiae* rDNA locus, Sir2 specifically deacetylates K16 on histone H4, *C. albicans* Sir2, reduces H3K9 acetylation. Therefore, heterochromatin at the non-transcribed regions of the rDNA array seems ubiquitous but it can differ in its structure.

We demonstrated that the MRS repeats are assembled into chromatin hypomethylated on H3K4, as observed in heterochromatic regions, but highly acetylated, as in euchromatic regions. Low levels of H3K4 methylation are not sufficient to create a transcriptionally repressive environment as a marker gene inserted at MRS repeats is not silenced. In addition, expression of MRS-proximal genes is not regulated by Sir2 (Fig. 3).

In many organisms, insertion of artificial DNA array is sufficient to seed heterochromatin<sup>41</sup>. Therefore, it is surprising that MRSs, being composed of very long tracts of nested repeats, are not assembled into classical heterochromatin. Why are MRSs not associated with heterochromatin? One primary function of heterochromatin is to inhibit recombination promoting genome stability. It is possible that hypomethylation on H3K4 is sufficient to block recombination or that an alternative mechanism promote genome stability at the MRSs.

Alternatively, genome instability at the MRSs could be beneficial for *C. albicans*, an organism lacking a canonical sexual cycle and meiosis<sup>42</sup>. Lack of heterochromatin at these loci could ensure high level of mitotic recombination, a key event to generate genomic diversity. In support of this hypothesis, analyses of clinical isolates suggest that MRSs might act as recombination hotspots as they can expand and contract and are known sites of translocations<sup>21,43,44</sup>. However it has been shown that, under standard laboratory growth conditions, recombination rate at the MRS repeats is not higher compared to a non-repetitive locus even though MRSs might have an effect on chromosome disjunction<sup>45,46</sup>. Our understanding of the biology and the function of the MRSs loci remains limited, making it very difficult to assess whether this epigenetic signature controls MRS function. Further studies will reveal whether and how the epigenetic state associated with these repetitive elements contributes to *C. albicans* biology.

Traditionally, telomeric heterochromatin has been described as a repressive chromatin structure that silences expression of subtelomeric genes<sup>2</sup>. Recent studies have challenged this hypothesis and highlighted the diversity of structure and functions of telomeric chromatin across organisms. For example, in *S. cerevisiae* the Sir silencing complex is responsible for the assembly of hypoacetylated telomeric heterochromatin<sup>3</sup>. However, this chromatin state has a limited ability to repress transcription of subtelomeric genes<sup>47</sup>. The yeast *C. lusitanae* appears to lack telomeric heterochromatin<sup>11</sup>. In the fungal pathogen *Cryptococcus neoformans* telomeric chromatin is methylated on K27 of histone H3 and silences expression of genes located in a 40 kb subtelomeric region<sup>48</sup>. We find that, *C. albicans* telomeres have also a specialised chromatin structure. Despite the apparent absence of Sir silencing proteins, *C. albicans* telomeres are associated with transcriptionally repressive heterochromatin (Fig. 5). Sir2 (*orf19.1992*) controls heterochromatin assembly at telomeres by deacetylation of histones. We hypothesise that an as yet unidentified protein complex targets Sir2 to telomeres. We demonstrate that telomeric heterochromatin transcriptionally silences subtelomeric genes as deletion of the *SIR2* gene causes their upregulation. This modulation is likely to have major impacts on the biology and the pathogenicity of *C. albicans* as many subtelomeric genes have key regulatory functions. For example, the subtelomeric *TLO* genes encode proteins with similarity to Med2, a component of the Mediator complex that regulates transcription by RNA polymerase II<sup>49</sup> and the subtelomeric gene *Nag4* encodes for a putative transporter<sup>50</sup>.

We find that telomeric heterochromatin is dynamic: the ability of telomere terminal repeats to repress the expression of an embedded *URA3* gene is affected by a physiologically-relevant stress condition with silencing much stronger at a temperature (39 °C) mimicking fever in the host than at lower temperatures (30 °C). This enhanced silencing is linked to changes in chromatin structure, as telomeres at 39 °C had lower levels of H3K9 acetylation and Sir2 was required for the increased silencing. This effect is specific for telomeric regions and for temperature shift as other loci are not affected by temperature changes and other physiologically relevant stresses do not lead to chromatin remodelling. Therefore, in *C. albicans*, adaptation to higher temperature is linked to chromatin remodelling at telomeric regions.

Dynamic heterochromatin is seen at telomeres in many species: For example, in *S. cerevisiae*, transcriptional silencing at higher temperature is enhanced at telomeres and weaker at the rDNA locus<sup>51</sup>. This could be due to a temperature-sensitive protein important for telomere function<sup>52</sup> or to changes in the levels of heat shock factors at different temperatures<sup>53</sup>. The dynamics of heterochromatin also affects virulence and pathogenesis in at least some microbial pathogens. For example, telomeric heterochromatin regulates the expression of the antigenic variation gene in parasites<sup>54</sup>. The biological consequences of telomere chromatin plasticity in *C. albicans* remain to be determined. It is possible that changes in telomeric heterochromatin correlate with changes in expression of subtelomeric genes. Alternatively, chromatin remodelling at telomeres could regulate genomic stability of these loci.

In conclusion, this study highlights the diversity and plasticity of chromatin states associated with DNA repeats in *Candida albicans*, the most common human fungal pathogen. We show that in *C. albicans* differential heterochromatin states control gene expression independently of the underlying DNA sequence and remodelling of heterochromatin is linked to adaptation in a stress condition.

## Methods

**Growth conditions.** Yeast cells were cultured in rich medium (YPAD) containing extra adenine (0.1 mg/ml) and extra uridine (0.08 mg/ml), complete SC medium (Formedium™) or SC Drop-Out media (Formedium™). Cells were grown at 30 °C or 39 °C as indicated.

**Yeast strain construction.** Strains are listed in Supplementary Table S1. Integration and deletion of genes were performed as previously described<sup>34</sup>. Oligonucleotides and plasmids used for strain constructions are listed in Supplementary Table S2 and Supplementary Table S4, respectively. Transformation was performed by electroporation (Gene Pulser™, Bio-Rad) using the protocol described in<sup>35</sup>. Correct integration events were checked by PCR and/or Southern blotting using primers listed in Supplementary Table S2 (Fig. S9).

**Silencing assay.** Growth analyses with *rDNA::URA3<sup>+</sup>*, *Tel5::URA3<sup>+</sup>* and *MRS::URA3<sup>+</sup>* strains were performed using a plate reader (SpectrostarNano, BMG labtech) in 24 well or 96 well plate format at 30 °C. When indicated Silencing assays were performed in the presence of 200 ng/μl of fluconazole (Sigma), 1 mM H<sub>2</sub>O<sub>2</sub> (Sigma) and 39 °C. For each silencing assay in a 24 well plate format, 1 ml of a starting culture was inoculated in SC or SC-Uri media to reach a concentration of 60 cells/μl. Growth was assessed by measuring A<sub>600</sub>, using the following conditions: OD600 nm, 3600 s cycle time, 30 flashes per well, 400 rpm shaking frequency, double orbital shaking mode, 850 s additional shaking time after each cycle, 0.5 s post delay, for 44 to 60 hours at 30 °C. For each silencing assay in 96 well plate format, 1:100 dilution of a starting culture was inoculated in a final volume of 95 μl of SC or SC-Uri media to reach a concentration of 60 cells/μl. Growth was assessed by measuring A<sub>600</sub>, using the following conditions: OD600 nm, 616 cycle time, 3 flashes per well, 700 rpm shaking frequency, orbital shaking mode, 545 s additional shaking time after each cycle 0.5 s post delay, for 44 hours.

Graphs represent data from three biological replicates. Error bars: standard deviations of three biological replicates. Data was processed using SpectrostarNano MARS software and Microsoft Excel.

**RNA extraction and cDNA synthesis.** All strains were grown in YPAD rich media. RNA extraction was performed using a yeast RNA extraction kit (E.Z.N.A.® Isolation Kit RNA Yeast, Omega Bio-Tek) following the manufacturer's instructions. RNA quality was checked by electrophoresis under denaturing conditions in 1% agarose, 1× HEPES, 6% Formaldehyde (Sigma). RNA concentration was measured using a NanoDrop ND-1000 Spectrophotometer. cDNA synthesis was performed using iScript™ Reverse Transcription Supermix for RT-qPCR (Bio-Rad) following manufacturer's instructions and a Bio-Rad CFXConnect™ Real-Time System.

**High-throughput RNA sequencing.** Strand-specific cDNA Illumina Barcoded Libraries were generated from 1 μg of total RNA extracted from WT and *sir2 Δ/Δ* and sequenced with an Illumina iSeq2000 platform. Illumina Library and Deep-sequencing was performed by the Genomics Core Facility at EMBL (Heidelberg, Germany). Raw reads were analysed following the RNA deep sequencing analysis pipeline described<sup>36</sup> using Galaxy (<https://usegalaxy.org/>) and Linux platform. Heatmaps and boxplot graphs were generated with R (<http://www.r-project.org/>). RNA sequencing data are deposited into ArrayExpress (accession number E-MTAB-4488).

**Quantitative Chromatin ImmunoPrecipitation (qChIP).** qChIP was performed as described<sup>33</sup> with the following modifications: 5 ml of an overnight culture grown in YPAD with extra uridine (0.08 mg/ml), diluted into fresh YPAD with extra uridine (0.08 mg/ml) and grown until OD600 nm of 1.4. Cells (50 ml/sample) were fixed with 1% Paraformaldehyde (Sigma) for 15 min at room temperature. Cells were lysed using acid-washed glass beads (Sigma) and a Disruptor genie™ (Scientific Industries) for 30 min at 4 °C. Chromatin was sheared to 500–1000 bp using a Bioruptor (Diagenode) for a total of 20 min (30 s ON and OFF cycle) at 4 °C. Immunoprecipitation was performed overnight at 4 °C using 2 μl of antibody anti-H3K4me2 (Active Motif- Cat Number: 39141), anti-H3K9ac (Active Motif- Cat Number: 39137), and anti-H4K16ac (Active Motif- Cat Number: 39167) and 25 μl of Protein G magnetic beads (Dyna - Invitrogen). DNA was eluted with a 10% slurry of Chelex 100-resin (Bio-Rad) using the manufacturer's instructions.

**qPCR reactions.** Primers used are listed in Supplementary Table S3. Real-time qPCR and RT-qPCR was performed in the presence of SYBR Green (Bio-Rad) on a Bio-Rad CFXConnect™ Real-Time System. Data were analysed with Bio-Rad CFX Manager 3.1 software and Microsoft Excel. Enrichments were calculated as the percentage ratio of specific IP over input for qChIP analysis and as enrichment over actin for RT-qPCR. Histograms represent data from three biological replicates. Error bars: standard deviation of three biological replicates.

**Southern blot.** Genomic DNA was extracted using glass acid beads (Sigma), phenol: chloroform: isoamyl alcohol (25:24:1) (Sigma) and RNase A treated (Fisher). Following centrifugation, pellets were precipitated with 0.05 mM Sodium Acetate (Sigma) and Ethanol (Fisher) at –20 °C during 30 minutes and resuspended in water. Genomic DNA was then digested with corresponding enzymes and run in 1% agarose gel. DNA was transferred to a nylon membrane (Zeta probe membranes, Bio-Rad), probed with DIG probes (Roche) and hybridized as described<sup>37</sup>.

**Microscopy.** Microscopy was carried out using an Olympus IX81 inverted microscope. Images were captured with a Hamamatsu photonics C4742 digital camera, with light excitation from an Olympus MT20 illumination system. Olympus CellR imaging software was used to control the apparatus.

## References

- Rusche, L. N., Kirchmaier, A. L. & Rine, J. The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* **72**, 481–516 (2003).
- Bühler, M. & Gasser, S. M. Silent chromatin at the middle and ends: lessons from yeasts. *EMBO J.* **28**, 2149–61 (2009).
- Saksouk, N., Simboeck, E. & Déjardin, J. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics Chromatin* **8**, 3 (2015).
- Kadosh, D. Shaping Up for Battle: Morphological Control Mechanisms in Human Fungal Pathogens. *PLoS Pathog.* **9**, 1–4 (2013).
- Kouzarides, T. Chromatin Modifications and Their Function. *Cell* **128**, 693–705 (2007).
- Strahl, B. D. & Allis, C. D. The language of covalent histone modifications. *Nature* **403**, 41–45 (2000).

7. Kueng, S., Oppikofer, M. & Gasser, S. M. SIR proteins and the assembly of silent chromatin in budding yeast. *Annu. Rev. Genet.* **47**, 275–306 (2013).
8. Shankaranarayana, G. D. *et al.* Sir2 Regulates Histone H3 Lysine 9 Methylation and Heterochromatin Assembly in Fission Yeast. *Curr. Biol.* **13**, 1240–1246 (2003).
9. Nakayama, J., Rice, J. C., Strahl, B. D., Allis, C. D. & Grewal, S. I. Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science* **292**, 110–3 (2001).
10. Rea, S. *et al.* Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**, 593–9 (2000).
11. Froyd, C. a., Kapoor, S., Dietrich, F. & Rusche, L. N. The deacetylase Sir2 from the yeast *Clavospora lusitanae* lacks the evolutionarily conserved capacity to generate subtelomeric heterochromatin. *PLoS Genet.* **9**, e1003935 (2013).
12. Huang, J. & Moazed, D. Association of the RENT complex with nontranscribed and coding regions of rDNA and a regional requirement for the replication fork block protein Fob1 in rDNA silencing. *Genes Dev.* **17**, 2162–76 (2003).
13. Straight, a. F. *et al.* Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. *Cell* **97**, 245–56 (1999).
14. Shou, W. *et al.* Exit from mitosis is triggered by Tem1-dependent release of the protein phosphatase Cdc14 from nucleolar RENT complex. *Cell* **97**, 233–44 (1999).
15. Pfaller, M. A. & Diekema, D. J. Epidemiology of invasive mycoses in North America. *Crit. Rev. Microbiol.* **36**, 1–53 (2010).
16. Selmecki, A., Forche, A. & Berman, J. Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryot. Cell* **9**, 991–1008 (2010).
17. Jones, T. *et al.* The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. USA* **101**, 7329–34 (2004).
18. Van het Hoog, M. *et al.* Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biol.* **8**, R52 (2007).
19. Kobayashi, T. & Ganley, A. R. D. Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science* **309**, 1581–4 (2005).
20. Li, C., Mueller, J. E. & Bryk, M. Sir2 represses endogenous polymerase II transcription units in the ribosomal DNA nontranscribed spacer. *Mol. Biol. Cell.* **17**, 3848–3859 (2006).
21. Chibana, H. *et al.* Diversity of tandemly repetitive sequences due to short periodic repetitions in the chromosomes of *Candida albicans*. *J. Bacteriol.* **176**, 3851–8 (1994).
22. Chibana, H. & Magee, P. T. The enigma of the major repeat sequence of *Candida albicans*. *Future Microbiol.* **4**, 171–9 (2009).
23. McEachern, M. J. & Hicks, J. B. Unusually large telomeric repeats in the yeast *Candida albicans*. *Mol. Cell. Biol.* **13**, 551–60 (1993).
24. Pérez-Martin, J., Uria, J. & Johnson, A. D. Phenotypic switching in *Candida albicans* is controlled by a SIR2 gene. *EMBO J.* **18**, 2580–92 (1999).
25. Henikoff, S. & Dreesen, T. D. Trans-inactivation of the *Drosophila brown* gene: evidence for transcriptional repression and somatic pairing dependence. *Proc. Natl. Acad. Sci. USA* **86**, 6704–8 (1989).
26. Gottschling, D. E., Aparicio, O. M., Billington, B. L. & Zakian, V. A. Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* **63**, 751–62 (1990).
27. Bryk, M. *et al.* Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. *Genes Dev.* **11**, 255–69 (1997).
28. Smith, J. S. & Boeke, J. D. An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev.* **11**, 241–254 (1997).
29. Greiss, S. & Gartner, A. Sirtuin/Sir2 phylogeny, evolutionary considerations and structural conservation. *Mol. Cells* **28**, 407–15 (2009).
30. Vasiljeva, L., Kim, M., Terzi, N., Soares, I. M. & Buratowski, S. Transcription Termination and RNA Degradation Contribute to Silencing of RNA Polymerase II Transcription within Heterochromatin. *Mol. Cell* **29**, 313–323 (2008).
31. Kobayashi, T. A new role of the rDNA and nucleolus in the nucleus - RDNA instability maintains genome integrity. *BioEssays* **30**, 267–272 (2008).
32. Bruno, V. M. *et al.* Comprehensive annotation of the transcriptome of the human fungal pathogen *Candida albicans* using RNA-seq. *Genome Res.* **20**, 1451–1458 (2010).
33. Pidoux, A., Mellone, B. & Allshire, R. Analysis of chromatin in fission yeast. *Methods* **33**, 252–259 (2004).
34. Imai, S., Armstrong, C. M., Kacherlein, M. & Guarente, L. Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* **403**, 795–800 (2000).
35. Tanny, J. C. & Moazed, D. Coupling of histone deacetylation to NAD breakdown by the yeast silencing protein Sir2: Evidence for acetyl transfer from substrate to an NAD breakdown product. *Proc. Natl. Acad. Sci. USA* **98**, 415–420 (2001).
36. Hansen, K. R., Ibarra, P. T. & Thon, G. Evolutionary-conserved telomere-linked helicase genes of fission yeast are repressed by silencing factors, RNAi components and the telomere-binding protein Tax1. *Nucleic Acids Res.* **34**, 78–88 (2006).
37. Freitas-Junior, L. H. *et al.* Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* **121**, 25–36 (2005).
38. Merrick, C. J. & Duraisingh, M. T. Heterochromatin-mediated control of virulence gene expression. *Mol. Microbiol.* **62**, 612–20 (2006).
39. Kaur, R., Domergue, R., Zupancic, M. L. & Cormack, B. P. A yeast by any other name: *Candida glabrata* and its interaction with the host. *Curr. Opin. Microbiol.* **8**, 378–384 (2005).
40. Huang, J., *et al.* Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer. *Genes Dev.* **20**, 2887–2901 (2006).
41. Dubarry, M., Lotodice, L., Chen, C. L., Thermes, C. & Taddei, A. Tight protein-DNA interactions favor gene silencing. *Genes Dev.* **25**, 1365–70 (2011).
42. Bennett, R. J. The parasexual lifestyle of *Candida albicans*. *Curr. Opin. Microbiol.* **28**, 10–17 (2015).
43. Chu, W. S., Magee, B. B. & Magee, P. T. Construction of an SfiI macrorestriction map of the *Candida albicans* genome. *J. Bacteriol.* **175**, 6637–51 (1993).
44. Iwaguchi, S. I. *et al.* Extensive chromosome translocation in a clinical isolate showing the distinctive carbohydrate assimilation profile from a candidiasis patient. *Yeast* **18**, 1035–46 (2001).
45. Lephart, P. R., Chibana, H. & Magee, P. T. Effect of the Major Repeat Sequence on Chromosome Loss in *Candida albicans*. *Eukaryot. Cell* **4**, 733–41 (2005).
46. Lephart, P. R. & Magee, P. T. Effect of the major repeat sequence on mitotic recombination in *Candida albicans*. *Genetics* **174**, 1737–44 (2006).
47. Ellahi, A., Thurtle, D. M. & Rine, J. The Chromatin and Transcriptional Landscape of Native *Saccharomyces cerevisiae* Telomeres and Subtelomeric Domains. *Genetics* **200**, 505–21 (2015).
48. Dumesic, P. A. *et al.* Product binding enforces the genomic specificity of a yeast polycomb repressive complex. *Cell* **160**, 204–18 (2015).
49. Haran, J. *et al.* Telomeric ORFs (TLOs) in *Candida* spp. Encode Mediator Subunits That Regulate Distinct Virulence Traits. *PLoS Genet.* **10**, e1004658 (2014).
50. Wendland, J., Schaub, Y. & Walther, A. N-acetylglucosamine utilization by *Saccharomyces cerevisiae* based on expression of *Candida albicans* NAG genes. *Appl. Environ. Microbiol.* **75**, 5840–5 (2009).
51. Bi, X., Yu, Q., Sandmeier, J. J. & Elizondo, S. Regulation of transcriptional silencing in yeast by growth temperature. *J. Mol. Biol.* **344**, 893–905 (2004).

52. Paschini, M. *et al.* A naturally thermolabile activity compromises genetic analysis of telomere function in *Saccharomyces cerevisiae*. *Genetics* **191**, 79–93 (2012).
53. Morano, K. a., Grant, C. M. & Moye-Rowley, W. S. The response to heat shock and oxidative stress in *Saccharomyces cerevisiae*. *Genetics* **190**, 1157–1195 (2012).
54. Wilson, R. B., Davis, D. & Mitchell, A. P. Rapid Hypothesis Testing with *Candida albicans* through Gene Disruption with Short Homology Regions. *J. Bacteriol* **181**, 1868–74 (1999).
55. De Backer, M. D. *et al.* Transformation of *Candida albicans* by electroporation. *Yeast* **15**, 1609–18 (1999).
56. Trapnell, C. *et al.* Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* **31**, 46–53 (2013).
57. Ketel, C. *et al.* Neocentromeres form efficiently at multiple possible loci in *Candida albicans*. *PLoS Genet.* **5**, e1000400 (2009).

#### Acknowledgements

We thank R. Allshire, A. Mitchell and A. Brown for reagents, strains and materials and A. Pidoux for critical reading of the manuscript. We thank the Gene Core Facility at EMBL (Heidelberg-Germany) for RNA Sequencing and M. Wass for support with the RNA Seq analyses. This work was supported by BBSRC (BB/L008041/1 to A.B. and D.T.), MRC (MR/M019713/1 to A.B. and R.J.P.), a Royal Society International Exchange Scheme (IE131376) grant (to A.B. and J.B.), a Royal Society Research Grant (RG130149) (to A.B.) and the NIAID R01 AI075096 (to J.B.), the People Programme (Marie Curie Actions) and the European Union's Seventh Framework Programme (FP7/2007-2013) under REA grant agreement number 303635 and an ERC Advanced Award, number 340087, RAPLODAPT (to J.B.).

#### Author Contributions

A.B., J.B. and V.F.B. conceived and designed the experiments. V.F.B., R.J.P. and D.T. conducted the experiments. A.B., V.F.B., R.J.P. and D.T. analysed the results. A.B. and J.B. wrote the manuscript.

#### Additional Information

**Supplementary information** accompanies this paper at <http://www.nature.com/srep>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Freire-Benítez, V. *et al.* *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.* **6**, 22989; doi: 10.1038/srep22989 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>

### **3. Supplementary information**

## Supplementary Information

### ***Candida albicans* repetitive elements display epigenetic diversity and plasticity**

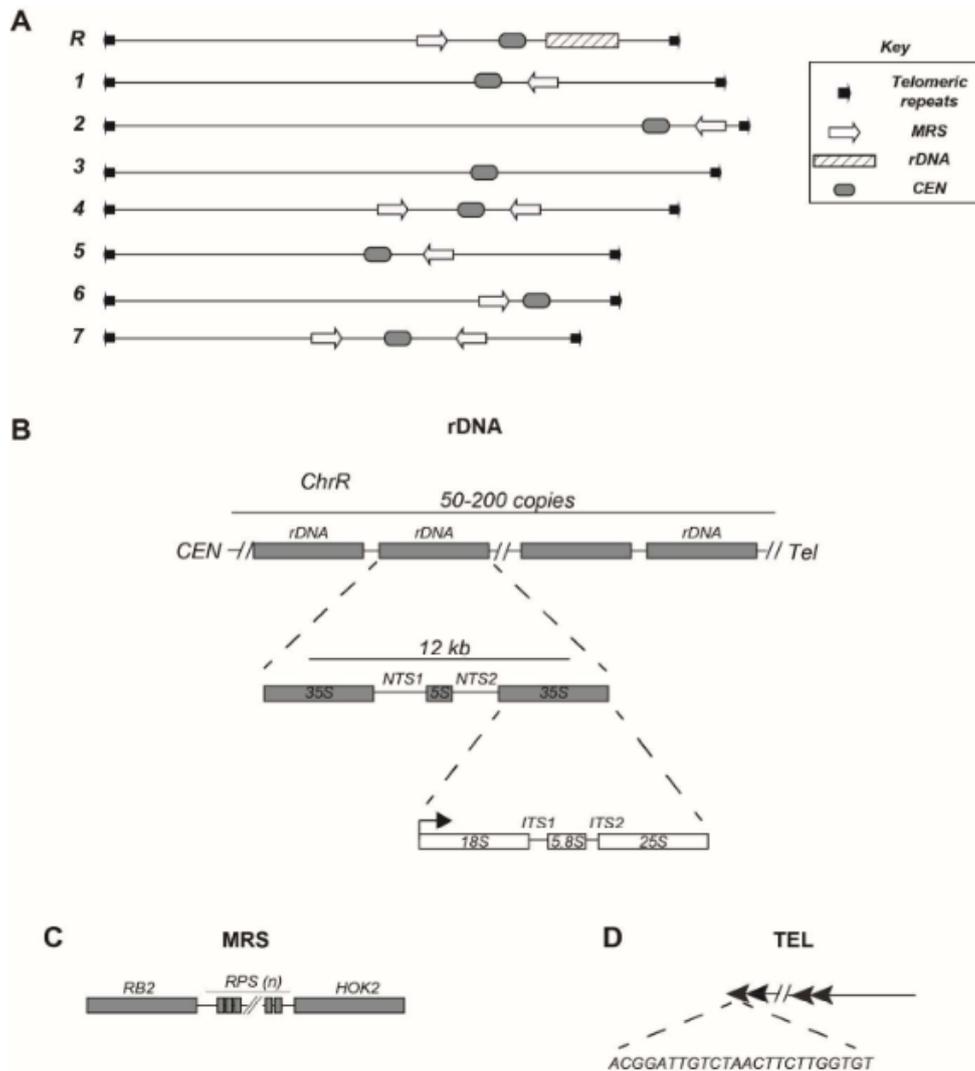
Verónica Freire-Benítez, Robert Jordan Price, Daniel Tarrant, Judith Berman and Alessia Buscaino

#### **Content:**

**Supplementary Figures S1 to S9 with Legends**

**Supplementary Tables S1 to S6**

Supplementary Figures



Supplementary Figure 1 *C. albicans* DNA repeats.

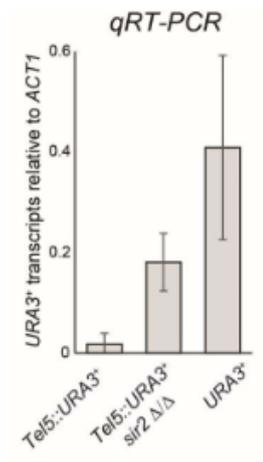
**(A)** Schematic of *Candida albicans* repeats. The cluster of ribosomal DNA repeated genes (rDNA) is located on chromosome R. Major Repeated Sequences (MRS) are located in all chromosomes except chromosome 3. Telomeric repeats are present at the end of each chromosome. **(B)** Organisation of the rDNA locus in *C. albicans*. The rDNA locus is formed by tandem arrays of a 12 kb unit repeated 50 to 200 times on chromosome R. Each unit

contains the 35 S and the 5S rRNA genes that are separated by two Non-Transcribed Regions (NTS1 and NTS2). The 35 S gene is the precursor of the 18S, 5.8S and 28S rRNAs that are separated by ITS (Internal Transcribed Spacer) elements **(C)** Organisation of MRS repeats. The MRS repeats are formed by tandem arrays of a 2.1 kb unit (RPS) flanked by HOK and RBP-2 elements. **(D)** Organisation of Telomeric repeats. Telomeric repeats are found at the end of each chromosome and are composed of tandemly repeating 23 bp unit. The sequence of the unit is indicated.



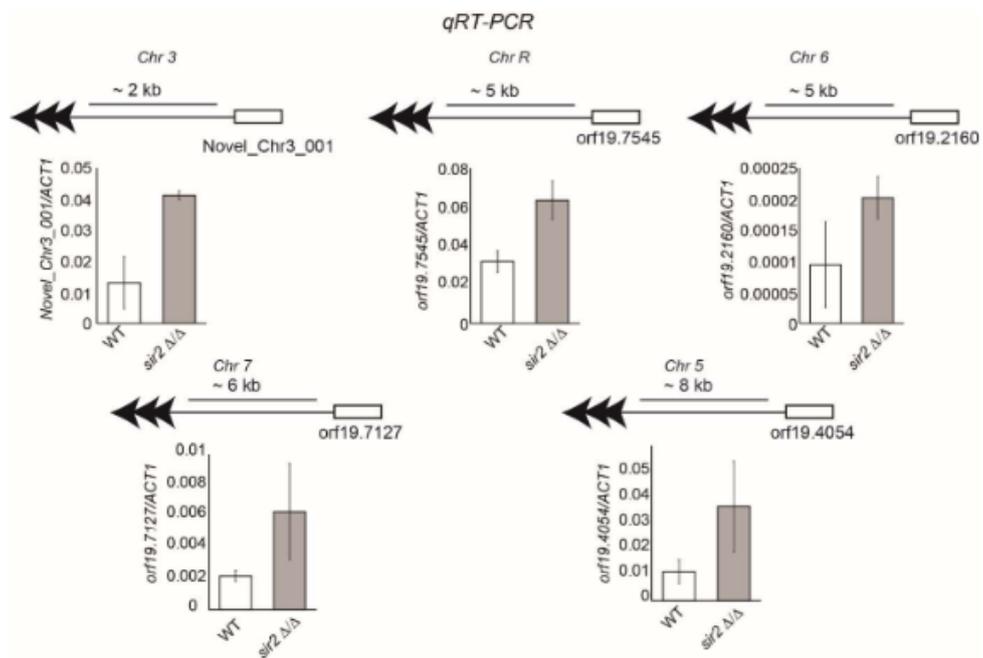
*Supplementary Figure 2. Sir2-like proteins in Candida albicans*

**(A-E)** *Left panels:* Diagram depicting protein length and domain organisation of the *C. albicans* Sir2-like proteins. The percentage (%) protein sequence identity between the *C. albicans* proteins and *S. cerevisiae* Sir2 is indicated. *Right panels:* Pairwise sequence alignments between *S. cerevisiae* Sir2 and *C. albicans* proteins with similarity to *S. cerevisiae* Sir2.

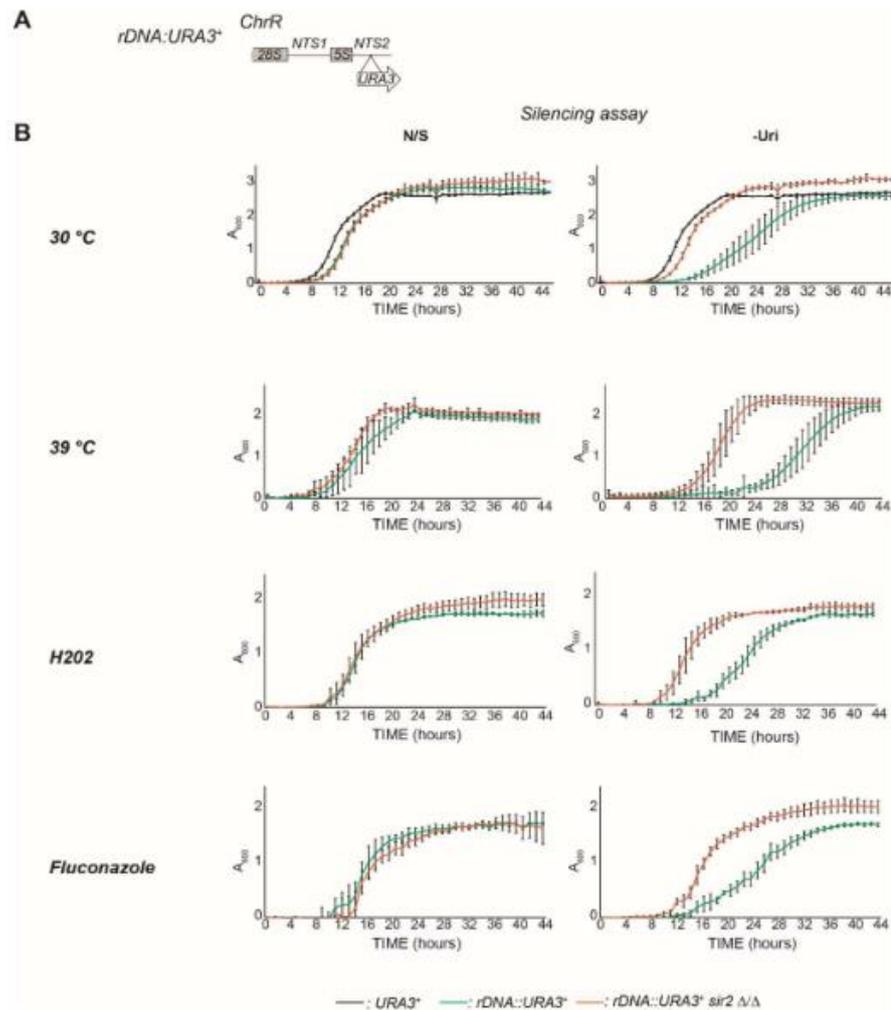


Supplementary figure 3. *Transcriptional Silencing at C. albicans telomeric repeats.*

qRT-PCR analyses to measure *URA3* transcript levels relative to actin transcript levels (*ACT1*) at 30°C in *Tel5::URA3* wild-type and *sir2 Δ/Δ* strains compared to *URA3* endogenous transcription levels. Error bars: SD of three biological replicates.

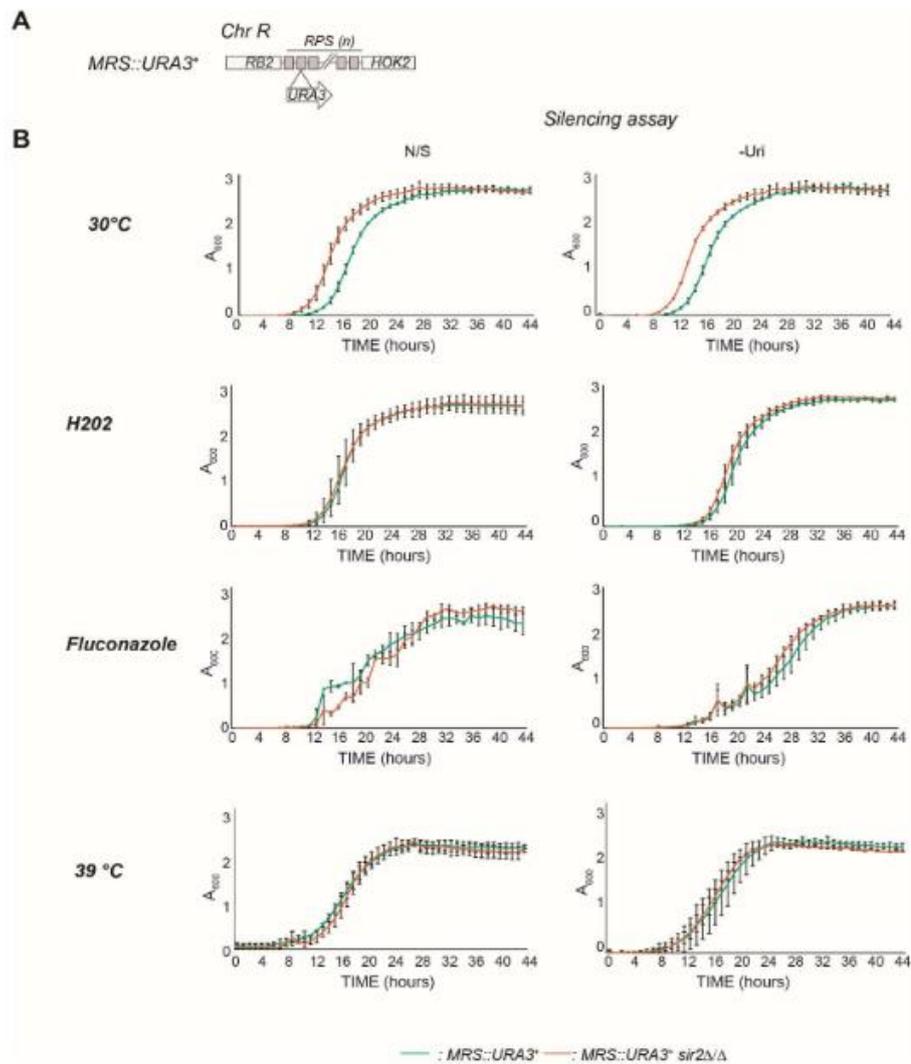


*Supplementary Fig 4.* qRT-PCR analyses to measure transcript levels of subtelomeric coding and non-coding transcripts (*Novel\_Chr3\_001*, *orf19.7545*, *orf19.2160*, *orf19.7127* and *orf19.4054*) relative to actin transcript levels (*ACT1*) in *Tel5::URA3* wild-type and *sir2*  $\Delta/\Delta$  strains compared to *URA3* endogenous transcription levels. Error bars in each panel: Standard deviation (SD) of three biological replicates.



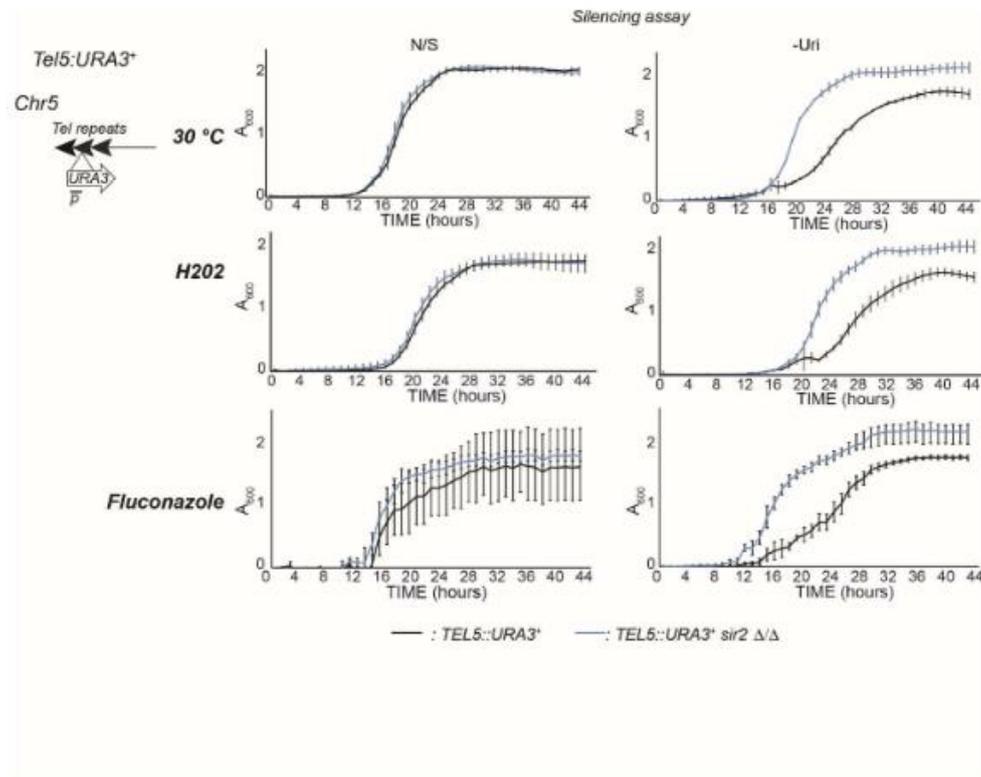
Supplementary figure 5. Environmental changes do not affect the transcriptional state of the *rDNA* locus

**(A)** Schematic of *rDNA::URA3* reporter strain. **(B)** Silencing assay assessing transcriptional silencing of the *rDNA::URA3* reporter strain in wild-type and *sir2* deletion mutant (*sir2*  $\Delta/\Delta$ ) at 30°C, 39°C, in the presence of 1 mM H<sub>2</sub>O<sub>2</sub>, and 200 ng/μl fluconazole. Cells were grown in non-selective (N/S) and media lacking uridine (-Uri) and  $A_{500}$  was measured every hour for 44 hours. A  $URA^+$  (*URA3*) strain was included as a control. Error bars: SD of three biological replicates.



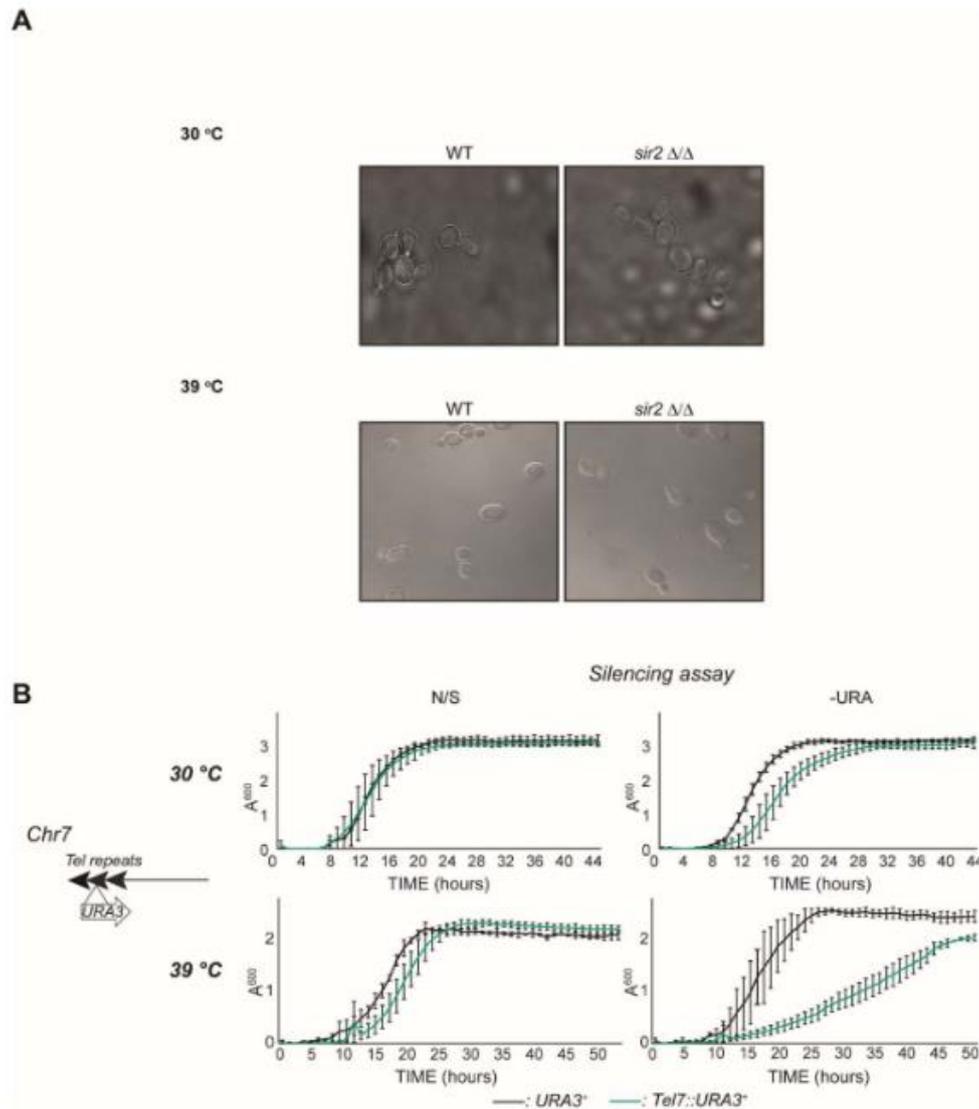
Supplementary figure 6. Environmental changes do not affect the transcriptional state of *MRS* repeats

**(A)** Schematic of *MRS::URA3<sup>+</sup>* reporter strain. **(B)** Silencing assay assessing transcriptional silencing of the *MRS::URA3<sup>+</sup>* reporter strain in WT and *sir2* deletion mutant (*sir2*  $\Delta/\Delta$ ) at 30°C, in the presence of 1 mM H<sub>2</sub>O<sub>2</sub>, and 200 ng/μl fluconazole and at 39°C. Cells were grown in non-selective (N/S) and media lacking uridine (-Uri) and A<sub>600</sub> was measured every hour for 44 hours. Error bars: SD of three biological replicates.



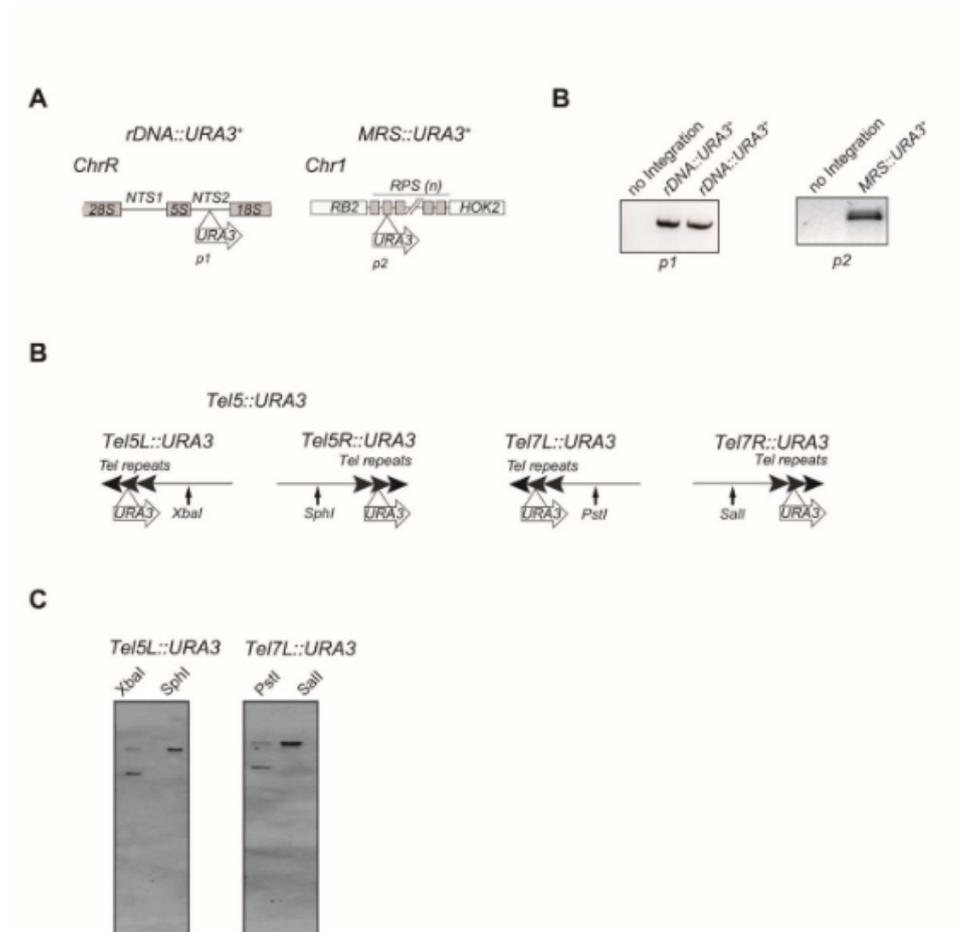
Supplementary Figure 7. H<sub>2</sub>O<sub>2</sub> treatment and Fluconazole treatment does not affect silencing at telomeric regions

**A)** Left panel: Schematic of *Tel5::URA3<sup>+</sup>* reporter strain. Right panels: Silencing assay assessing transcriptional silencing of the *Tel5::URA3<sup>+</sup>* reporter strain in WT and *sir2* deletion mutant (*sir2 ΔΔ*) at 30°C, in the presence of 1 mM H<sub>2</sub>O<sub>2</sub>, 200 ng/μl fluconazole. Cells were grown in non-selective (N/S) and media lacking uridine (-Uri) and A<sub>600</sub> was measured every hour for 44 to 60 hours.



Supplementary figure 8. Plastic Heterochromatin at telomeric region

**A)** Cell morphology of WT and *sir2* D/D strains grown at 30 °C 39 °C for 24 hours in YPAD medium. **B) Left panel:** Schematic of *Tel7::URA3<sup>+</sup>* reporter strain. **Right panels:** Silencing assay assessing transcriptional silencing of the *Tel7::URA3<sup>+</sup>* reporter strain at 30°C and at 39°C. Cells were grown in non-selective (N/S) and media lacking uridine (-Ura) and *A<sub>600</sub>* was measured every hour for 50 hours. Error bars: SD of three biological replicates.



Supplementary figure 9. *URA3<sup>+</sup>* marker gene reporter strains.

**(A)** Schematics of *rDNA::URA3<sup>+</sup>* and *MRS::URA3<sup>+</sup>*. **(B)** PCR analyses confirming integration of the *URA3<sup>+</sup>* marker gene at the *rDNA* locus (*rDNA::URA3<sup>+</sup>*) and the *MRS* repeats (*MRS::URA3<sup>+</sup>*). **(C)** Southern blot of *C. albicans* genomic DNA digested with arm specific enzymes (*Tel5L*: *XbaI*, *Tel5R*: *SphI*, *Tel7L*: *PstI*, *Tel7R*: *SalI*) demonstrating that the *URA3<sup>+</sup>* gene is inserted on Chromosome 5 left arm (*Tel5L::URA3*) and Chromosome 7 left arm (*Tel7L::URA3*). The blot was probed with a DIG probe targeting the *URA3<sup>+</sup>* gene.

**Supplementary Table**

*Supplementary Table 1: Strains used in this study*

<b>Strain number</b>	<b>Description</b>	<b>Genotype</b>
Bu_20	<i>sir2Δ/Δ</i>	<i>ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG sir2 Δ::HIS1/sir2 Δ::ARG4</i>
Bu_44	<i>Tel5:URA3</i>	<i>Tel5::URA3 ura3Δ::λimm434/ura3Δ::λimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG</i>
Bu_45	<i>Tel5:URA3 sir2Δ/Δ</i>	<i>Tel5::URA3 ura3Δ::λimm434/ura3Δ::λimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG sir2Δ::HIS1/sir2Δ::ARG4</i>
Bu_60	BWP17	<i>ura3Δ::λimm434/ura3Δimm434 HIS1::his1::hisG/his1::hisG ARG4::arg4::hisG/arg4::hisG</i>
Bu_70	<i>set1Δ/Δ</i>	<i>ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG set1 Δ::HIS1/set1 Δ::LEU2</i>
Bu_83	<i>Tel7:URA3</i>	<i>Tel7::URA3 ura3Δ::λimm434/ura3Δ::λimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG</i>
Bu_95	<i>rDNA:URA3</i>	<i>rDNA::URA3 ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG</i>
Bu_102	<i>rDNA:URA3 sir2Δ/Δ</i>	<i>rDNA::URA3 ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG sir2Δ::HIS1/sir2Δ::ARG4</i>
Bu_106	<i>rDNA:URA3 hst1Δ/Δ</i>	<i>rDNA::URA3 ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG hst1Δ::HIS1/hst1Δ::ARG4</i>
Bu_215	BWP17	<i>ura3Δ::λimm434/ura3Δ::λimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG</i>
Bu_227	<i>MRS:URA3</i>	<i>MRS::URA3 ura3Δ::λimm434/ura3Δimm434 HIS1::his1::hisG/his1::hisG ARG4::arg4::hisG/arg4::hisG</i>
Bu_244	<i>MRS:URA3 sir2Δ/Δ</i>	<i>MRS::URA3 ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG sir2Δ::HIS1/sir2Δ::NAT</i>

Supplementary table 2. Primers used in this study

Primer	Sequence	Figure	Description
Bu_91	TTACATCTTATGGCTATGA ATTGGCACTAGTTGGGTC CTTCTAGCATCGAACTGA CTGTGTCAGCCTTAGCGG AGTTAATTTACATTGTTTT CCCAGTCACGACGTT	Fig 3A, S6B	MRS::URA3
Bu_92	TTGAAGCTACAATTATGTA GAGTATTGGGTGTGAATTA GGCATGAATCGGATCAAA ATTGGTTGAGCTATTGAAG AAAACGTTTTCTCCGTGGA TGTGGAATTGTGAGCG	Fig 3A, S6B	MRS::URA3
Bu_98	CCTTAACTCCGTCTCCGTG T	Fig 3A, S6B	Primer to check MRS::URA3
Bu_121	GGAGTACTCCTATACTAAT AACAAACTCCACTATAA TTGGCAACCACAATTATGC CTGGGGAAATATTGTTTT CCAGTCACGACGTT	Fig 1A-B, 2A-C S5B	rDNA:URA3
Bu_122	AATGCACGTGACCCACAC AATTTTCAACCACCAACAA CACACCAAAAATGTATGTA CACTCGGAGTGAATGTG GAATTGTGAGCGGATA	Fig 1A-B, 2A- C, S5B	rDNA:URA3
Bu_131	GACTGGCCAATTATAAATG TGAAGG	Fig 1A-B, 2A- C, S5B	Primer to check rDNA:URA3
Bu_132	GTCTAAATTCCCCTTCCCC ATAC	Fig 1A-B, 2A- C, S5B	Primer to check rDNA:URA3
Bu_135	CTAGAAATCACTAGTGCG GCC	Fig 3A, S6B	Primer to check MRS:URA3
Bu_139	GAGTGAGTGAGTGGAGTA GCG	Fig2B-C, Fig3A, FigS5B, FigS6B	Primer to check <i>sir2Δ/Δ</i> deletion
Bu_152	CTGGAGAAAATATAACCAC GAGTCTAAGTTTCTTTATT ATATTGACGTTTCAGTTAT TTGAGAGAAATCCTCTAGT AGTTTTCCAGTCACGACG TT	Fig2B-C, Fig3A, FigS5B, FigS6B	<i>sir2Δ/Δ</i> deletion mutant
Bu_153	ATATATAAATATATAAATAT ATATATATAAAGAATTGA AAAGAAAAACATTAAGAC ACCAATATTAATTTAATGT GGAATTGTGAGCGGATA	Fig2B-C, Fig3A, FigS5B, FigS6B	<i>sir2Δ/Δ</i> deletion mutant
Bu-158	CTATCAAAACACTCACTTA GTTACATATATATTCTTATT CTTATCAATTACTAATA ACAAATAACAATCAATAGT TTTCCCAGTCACGACGTT	Fig2A	<i>hst1Δ/Δ</i> deletion mutant
Bu-159	ACGTCTATAGTTTATCTAT CGGGGCTTTCTCTTCTCT	Fig2A	<i>hst1Δ/Δ</i> deletion mutant

	TTGTCCTCGTTGTCCACTT TATCTTGTGTTTTGGCTCTTG TGGAATTGTGAGCGGATA		
Bu_164	CGGTCTGGTAAATGATTGA C	Fig2B-C, Fig3A, FigS5B, FigS6B	Primer to check <i>HIS1</i> integration
Bu_165	AGTGTGGAAGAAGAGAT GC	Fig2B-C, FigS5B	Primer to check <i>ARG4</i> integration
Pf_169	CACCACCACTTCTACCACT TC	Fig2A	Primer to check <i>hst1Δ/Δ</i> deletion
Bu_179	CTGTATCTATAAGCAGTAT CATCC	Fig 3A, FigS5B	Primer to check <i>NAT</i> integration
Bu_286	CTGGAGAAAATATAACCAC GAGTCTAAGTTTCTTTATT ATATTGACGTTTCAGTTAT TTGAGAGAAATCCTCTAGT AGTAAAACGACGGCCAGT GAATTC	Fig 3A, FigS5B	<i>sir2Δ/Δ</i> deletion mutant
Bu_287	ATATATAAATATATAAATAT ATATATATAAAAGAATTGA AAAGAAAACATTAAGAC ACCAATATTAATTTAATGC ATCAATTGACGTTGATACC AC	Fig 3A, FigS5B	<i>sir2Δ/Δ</i> deletion mutant

Supplementary table 3. qPCR primer used in this study

Primer	Sequence	Figure	Description
Bu_108	GGCACTAGTTGGGTCCTT CT	Fig3D-F	MRS: qChip
Bu_109	GGGCCGTTTTGAAGCTAC AA	Fig3D-F	MRS: qChip
Bu_129	GTTGTCTGACCATGGGTA TACCA	Fig2E-G	rDNA: qChip
Bu_138	CCAGGCATAATTGTGGTT GCC	Fig2E-G	rDNA :qChip
Bu_141	GTTGGGCAGATATTACCA ATG	Fig1B,2C, 5B-C, FigS3, FigS9	<i>URA3</i> probe, <i>Tel5-URA3</i> qChip, RT-qPCR
Bu-142	CCTTCACATTTATAATTGG CC	FigS9	<i>URA3</i> probe
Bu_174	CTACGTTTCCATTCAAGCT GTT	Fig1B, Fig2C-G, Fig3D-F, Fig4D- F, Fig5B-C FigS3, FigS4	<i>Act1</i> : qChip, RT- qPCR
Bu_176	AAACTGTAACCACGTTCA GACA	Fig1B, Fig2C-G, Fig3D-F, Fig4D- F, Fig5B-C FigS3, FigS4	<i>Act1</i> : qChip, RT- qPCR
Bu_204	CAAATTCCTTATCGGATTT AGC	Fig1B,2C, 5B-C, FigS3FigS3	<i>URA3</i> , <i>Tel5-URA3</i> qChip, RT-qPCR
Bu_430	GGCAGAGGAAGCGAAGA AG	FigS4	orf19.7127 RT- qPCR

Bu_431	CACTTGAACCTCCCTTCTAG	FigS4	orf19.7127 RT-qPCR
Bu_432	CTTGGACATGAACAACATACTTG	FigS4	orf19.4054 RT-qPCR
Bu_433	GTTGTAGAGTCGACTGACTCAAG	FigS4	orf19.4054 RT-qPCR
Bu_434	TGTCTGACCATGGGTATACCA	Fig2D	<i>Novel_ChrR_R093</i> RT-qPCR
Bu_435	CCGTAGCCCTAACCCCTAATT	Fig2D	<i>Novel_ChrR_R093</i> RT-qPCR
Bu_436	GACGCTAGAAGCTTGGTGTCTC	FigS4	orf19.2160 RT-qPCR
Bu_437	CGTAAACCAGATTCCAGGTC	FigS4	orf19.2160 RT-qPCR
Bu_438	AAATACGAGGGGACCAGAG	FigS4	orf19.7545 RT-qPCR
Bu_439	CTTCGATGTGGTGATTGCAC	FigS4	orf19.7545 RT-qPCR
Bu_440	CAGATGAAGAATGCAGTTGG	FigS4	<i>Novel_Chr3_001</i> RT-qPCR
Bu_441	TCTCCAGCACTGTTCACTCC	FigS4	<i>Novel_Chr3_001</i> RT-qPCR

Supplementary table 4. Plasmids used in this study

Plasmid	Description
pGEMURA3	<i>URA3</i> integration products (Wilson et al, 1999)
pGEMHIS1	<i>HIS1</i> substitution products (Wilson et al, 1999)
pRS-Arg4SpeI	<i>Arg4</i> substitution products (Wilson et al, 1999)
pHA_NAT	<i>NAT</i> substitution products (Gerami-Nejad et al, 2012)

Supplementary table 5. Gene expression profile of MRS-associated genes in *sir2*  $\Delta/\Delta$  versus wild-type isolates

Orf name	Gene name	FPKM ratio	Orf	Chr	Distance from MRS
orf19.1233	<i>ADE4</i>	0.123204	Orf L	1	2 kb
orf19.4712	<i>FGR6-3</i>	0.466022	Orf C	1	Inside MRS
orf19.4713	-	0.0576256	Orf R	1	1 kb
orf19.1742	<i>HEM3</i>	-0.0372117	Orf L	2	4 kb
orf19.3490	<i>FGR6-4</i>	0.0658554	Orf C	2	Inside MRS

orf19.5316	<i>FGR29</i>	1.31931	Orf R	2	1 kb
orf19.1801	<i>CBR1</i>	-0.202934	Orf L	4	6 kb
orf19.1234	<i>FGR6-10</i>	0.349738	Orf C	4	Inside MRS
orf19.1235	<i>HOM3</i>	0.273063	Orf R	4	1 kb
orf19.4349	-	0.307224	Orf L	5	6 kb
orf19.2655	<i>BUB3</i>	-0.454529	Orf R	5	1 kb
orf19.5773	-	-0.899313	Orf L	6	3 kb
NOVEL-Ca21chr6-037	-	-0.571963	Orf C	6	Inside MRS
orf19.1221	<i>ALG2</i>	0.0469119	Orf R	6	1 kb
orf19.7006	-	0.109973	Orf L	7	5 kb
NOVEL-Ca21chr7-016	-	0.228365	Orf C	7	Inside MRS
orf19.6898.1	-	0.122529	Orf R	7	3 kb
orf19.3695	-	-0.166999	Orf L	7	3 kb
orf19.5191	<i>FGR6-1</i>	0.795988	Orf C	7	Inside MRS
NOVEL-Ca21chr7-042	-	1.52733	Orf R	7	3 kb
orf19.3888	<i>PGI1</i>	-1.28362	Orf L	R	3 kb
NOVEL-Ca21chrR-061	-	1.49026	Orf C	R	Inside MRS
orf19.726	<i>PPZ1</i>	-0.0324678	Orf R	R	4 kb

*Supplementary Table 5.*

List of the MRS-proximal transcripts analysed in this study. The Log FPKM ratio of transcripts detected in *sir2*  $\Delta/\Delta$  versus wild-type isolates is indicated. Orf numbers, gene name, chromosomal position and distance of each transcripts to MRS repeats is indicated.

Supplementary table 6. Gene expression profile of subtelomeric genes in *sir2*  $\Delta/\Delta$  versus wild-type isolates

Gene	Gene name	FPKM ratio	Orf #	Chr	Distance from Telomeric repeats
NOVEL_Ca21chr1-001	-	1.60282	Orf 1	1L	~2kb
NOVEL_Ca21chr1-002	-	2.75269	Orf 2	1L	~3kb
orf19.6115	-	3.76323	Orf 3	1L	~4kb
orf19.6114	-	1.98411	Orf4	1L	~4kb
NOVEL_Ca21chr1-003	-	1.94456	Orf5	1L	~5kb
orf19.7278	-	1.91179	Orf 1	1R	~700 bp
orf19.7276.1	<i>TLO4</i>	0.527333	Orf 2	1R	~2kb
orf19.7277	-	0.0468087	Orf 3	1R	~7kb
orf19.7275	<i>FGR24</i>	0.127306	Orf4	1R	~7kb
orf19.7274	-	0.542814	Orf5	1R	~7kb
NOVEL_Ca21chr2-001	-	1.37521	Orf 1	2L	~900 bp
NOVEL_Ca21chr2-002	-	0.659394	Orf 2	2L	~3 kb
orf19.1925	<i>TLO5</i>	1.06682	Orf 3	2L	~5 kb
orf19.1923	<i>RRN3</i>	1.1434	Orf4	2L	~7 kb
orf19.1920	-	2.10163	Orf5	2L	~8kb
NOVEL_Ca21chr2-097	-	0.029059	Orf 1	2R	~600 bp
orf19.5370	-	-0.417262	Orf 2	2R	~5 kb
orf19.5369	-	0.802781	Orf 3	2R	~6 kb
orf19.5368	-	-0.471883	Orf4	2R	~8 kb
orf19.5365.1	-	-0.705217	Orf5	2R	~10 kb
NOVEL_Ca21chr3-001	-	1.93103	Orf 1	3L	~2 kb
NOVEL_Ca21chr3-002	-	2.58399	Orf 2	3L	~3 kb
orf19.5474	-	1.71582	Orf 3	3L	~5 kb
orf19.5469	-	1.69472	Orf4	3L	~8 kb
orf19.5468	-	2.11177	Orf5	3L	~13 kb
NOVEL_Ca21chr3-067	-	1.29229	Orf 1	3R	~700 bp
orf19.6192	-	1.59438	Orf 2	3R	~4 kb
NOVEL_Ca21chr3-065	-	1.12003	Orf 3	3R	~5 kb
orf19.6191	<i>TLO8</i>	2.4063	Orf4	3R	~12 kb
orf19.6190	<i>SRB1</i>	1.37969	Orf5	3R	~15 kb
orf19.362	<i>TLO9</i>	1.32941	Orf 1	4L	~2 kb
NOVEL_Ca21chr4-001	-	0.750471	Orf 2	4L	~2 kb

orf19.364	-	0.395101	Orf 3	4L	~4 kb
orf19.366	-	0.950516	Orf4	4L	~5 kb
orf19.367	<i>CNH1</i>	0.0784385	Orf5	4L	~7 kb
NOVEL_Ca21chr4-075	-	0.416886	Orf 1	4R	1 kb
NOVEL_Ca21chr4-074	-	0.596045	Orf 2	4R	5 kb
orf19.3070	-	0.0911503	Orf4	4R	6 kb
orf19.3076	-	0.475355	Orf 3	4R	7 kb
orf19.3077	<i>VID21</i>	0.933247	Orf5	4R	9 kb
orf19.5700	<i>TLO11</i>	1.33724	Orf 1	5L	2 kb
orf19.5698	-	0.941254	Orf 2	5L	5 kb
orf19.5693	-	0.994588	Orf4	5L	6 kb
orf19.5694	-	0.779541	Orf 3	5L	7 kb
orf19.5691	<i>CDC11</i>	0.910976	Orf5	5L	10 kb
orf19.4055	-	0.705101	Orf 1	5R	2 kb
orf19.4054	<i>CTA24</i>	1.08844	Orf 2	5R	9 kb
NOVEL_Ca21chr5-051	-	1.76277	Orf4	5R	10 kb
orf19.4051	<i>HTS1</i>	0.996183	Orf 3	5R	13 kb
orf19.4048	<i>DES1</i>	0.656469	Orf5	5R	13 kb
orf19.6338	-	0.781793	Orf 1	6L	4 kb
orf19.6337	<i>TLO13</i>	1.14565	Orf 2	6L	6 kb
orf19.6336	<i>PGA25</i>	-0.155507	Orf4	6L	10 kb
orf19.6329	-	1.31487	Orf 3	6L	11 kb
orf19.6328	-	0.902296	Orf5	6L	12 kb
NOVEL_Ca21chr6-044	-	1.10283	Orf 1	6R	2 kb
orf19.2163	-	0.377088	Orf 2	6R	5 kb
orf19.2160	<i>NAG4</i>	4.41018	Orf4	6R	8 kb
orf19.2158	<i>NAG3</i>	2.84948	Orf 3	6R	10 kb
orf19.2157	<i>DAC1</i>	1.86075	Orf5	6R	12 kb
orf19.7125	-	0.454194	Orf 1	7L	1 kb
orf19.7124	<i>RVS161</i>	0.0573432	Orf 2	7L	2 kb
orf19.7123	-	-0.142634	Orf4	7L	3 kb
orf19.7121	-	1.02638	Orf 3	7L	3 kb
orf19.7119	<i>RAD3</i>	0.133712	Orf5	7L	6 kb
orf19.7127	<i>TLO16</i>	2.32205	Orf 1	7R	6 kb
orf19.7127.1	-	2.20881	Orf 2	7R	8 kb
orf19.7128	<i>SYS1</i>	0.110454	Orf4	7R	9 kb
orf19.7130	-	1.41063	Orf 3	7R	9 kb
orf19.7131	-	0.956586	Orf5	7R	10 kb
orf19.7545	-	2.06856	Orf 1	RR	5 kb
NOVEL_Ca21chrR-001	-	-0.770767	Orf 2	RR	7 kb
orf19.7544	<i>TLO1</i>	1.81897	Orf4	RR	9 kb
orf19.7539.1	-	0.471387	Orf 3	RR	12 kb
orf19.7539	<i>INO2</i>	0.600065	Orf5	RR	14 kb

orf19.7680	<i>CTA26</i>	1.26797	Orf 1	RL	900 bp
orf19.7678	<i>ATP16</i>	1.3481	Orf 2	RL	2 kb
orf19.7676	<i>XYL2</i>	1.66989	Orf4	RL	3 kb
orf19.7675	-	0.766025	Orf 3	RL	3 kb
orf19.7673	-	0.485425	Orf5	RL	4 kb

*Supplementary Table 6.*

List of the subtelomeric transcripts analysed in this study. The Log FPKM ratio of transcripts detected in *sir2*  $\Delta\Delta$  versus wild-type isolates is indicated. Orf numbers, gene name, chromosomal position and distance of each transcripts to telomeric repeats is indicated.

## **Chapter 3. Chromatin status at *C. albicans* centromeres.**

**It contains the article:** Freire-Benítez, V., Price, R.J., and Buscaino, A. (2016b). The Chromatin of *Candida albicans* Pericentromeres Bears Features of Both Euchromatin and Heterochromatin. *Front. Microbiol.* 7.

### **Author contributions**

VFB performed all the experimental work, new mutant construction, bioinformatics and result analysis show in figures 1 – 5A and 5B. VFB made all the figures, supplementary tables and wrote the first draft of the manuscript.

JRP performed the work show in figure 5C.

AB wrote the final manuscript and VFB made the final figures.

## 1. Summary

Centromeres at chromosomes provide the anchor for kinetochore assembly and spindle attachment. At centromeric regions, histone 3 (H3) is replaced by the histone variant CenpA. Centromeric regions are often associated with repetitive sequences assembled into transcriptionally silent heterochromatin. This heterochromatin is SIR2-dependent and it is associated with low levels of acetylation on lysine 9 histone 3 (H3K9) and lysine 16 histone 4 (H4K16) and also with low levels of methylation at lysine 4 histone 3 (H3K4). In *S. cerevisiae* methylation at H3K4 is introduced by the histone methyltransferase Set1 and removed by the histone demethylase Jhd2. *C. albicans* centromeres have been proposed to have an epigenetic nature independent of the underlying DNA sequence. Centromere core sequences are all different and unique among all chromosomes. Moreover, not all centromeres are associated with inverted repeats (*IR*) or even the same type of *IR*. In this chapter we address if *C. albicans* pericentromeres are assembled into heterochromatin.

We show that a *URA3*<sup>+</sup> marker gene inserted at pericentromeric regions, both containing and lacking repeats, is not silenced. Silencing assays failed to detect transcriptional silencing, while RT-PCR analyses detected a reduced expression of this marker gene. Furthermore, ChIP studies showed that all centromeres tested are associated with high levels of histone acetylation but low levels of methylation. RNA-sequencing analyses revealed that most proximal genes to the centromere core are expressed to a lesser extent than the genome average. In addition, Set1 is required to increase methylation levels of H3K4 at centromeres. Surprisingly, Jhd1, the demethylase in charge of

reducing those levels in other organisms, has no effect at *C. albicans* centromeres.

In conclusion, our analyses show an intermediate state of the chromatin associated with centromeres bearing both features of heterochromatin and euchromatin. This state reduces transcriptional levels associated with genes located in proximity to pericentromeric regions.

### **3. Main Article**



# The Chromatin of *Candida albicans* Pericentromeres Bears Features of Both Euchromatin and Heterochromatin

Verónica Freire-Benítez, R. Jordan Price and Alessia Buscaino\*

School of Biosciences Canterbury Kent, University of Kent, Canterbury, UK

## OPEN ACCESS

### Edited by:

Ana Traven,  
Monash University, Australia

### Reviewed by:

Laura Rusche,  
State University of New York  
at Buffalo, USA  
Kristin Scott,  
Duke University, USA

### \*Correspondence:

Alessia Buscaino  
a.buscaino@kent.ac.uk

### Specialty section:

This article was submitted to  
Fungi and Their Interactions,  
a section of the journal  
Frontiers in Microbiology

Received: 29 February 2016

Accepted: 05 May 2016

Published: 19 May 2016

### Citation:

Freire-Benítez V, Price RJ  
and Buscaino A (2016) The Chromatin  
of *Candida albicans* Pericentromeres  
Bears Features of Both Euchromatin  
and Heterochromatin.  
*Front. Microbiol.* 7:759.  
doi: 10.3389/fmicb.2016.00759

Centromeres, sites of kinetochore assembly, are important for chromosome stability and integrity. Most eukaryotes have regional centromeres epigenetically specified by the presence of the histone H3 variant CENP-A. CENP-A chromatin is often surrounded by pericentromeric regions packaged into transcriptionally silent heterochromatin. *Candida albicans*, the most common human fungal pathogen, possesses small regional centromeres assembled into CENP-A chromatin. The chromatin state of *C. albicans* pericentromeric regions is unknown. Here, for the first time, we address this question. We find that *C. albicans* pericentromeres are assembled into an intermediate chromatin state bearing features of both euchromatin and heterochromatin. Pericentromeric chromatin is associated with nucleosomes that are highly acetylated, as found in euchromatic regions of the genome; and hypomethylated on H3K4, as found in heterochromatin. This intermediate chromatin state is inhibitory to transcription and partially represses expression of proximal genes and inserted marker genes. Our analysis identifies a new chromatin state associated with pericentromeric regions.

**Keywords:** epigenetics, chromatin, centromere, *Candida albicans*, heterochromatin

## INTRODUCTION

The centromere is the *cis*-acting DNA site of kinetochore assembly and spindle attachment during chromosome segregation in mitosis and meiosis. Centromeric regions have a different organization across different species. Some organisms, such as the budding yeast *Saccharomyces cerevisiae*, have “point” centromeres while other organisms, such as the fission yeast, the fruit fly and human, have “regional” centromeres (Buscaino et al., 2010). Point centromeres are only ~125 bp long and include specific DNA binding sites necessary for centromere function (Westermann et al., 2007). Regional centromeres span large DNA domains (~10 to 10,000 kb) and do not contain a specific DNA sequence but are epigenetically specified by the presence of the histone H3 variant, CENP-A (also termed Cse4 and CENH3) (Buscaino et al., 2010). Regional centromeres are often associated with repetitive elements. The structure and organization of centromere-associated DNA repeats varies across organisms. For example, human centromeres are composed of tandem arrays of 171 alpha-satellite repeats and, in *Drosophila melanogaster*, centromeric DNA contains short repetitive elements interspersed with transposable elements (Buscaino et al., 2010). In the yeast *Schizosaccharomyces pombe* and the fungal pathogen *Candida tropicalis*, centromeres are organized in a CENP-A-containing central/mid core domain flanked by inverted repeats

(IRs) whose sequences are conserved across centromeres (Buscaino et al., 2010; Chatterjee et al., 2016). In both organisms, these IRs are important for *de novo* CENP-A deposition on a plasmid containing the central core sequence (Baum et al., 1994; Chatterjee et al., 2016). Pericentromeric regions are usually assembled into transcriptionally silent heterochromatin that is required for establishment of CENP-A chromatin and for faithful chromosome segregation (Bernard et al., 2001; Nonaka et al., 2002; Folco et al., 2008). At these locations, heterochromatin is hypoacetylated at Lysine 9 of Histone H3 (H3K9) and Lysine 16 of Histone H4 (H4K16). Heterochromatic regions are also hypomethylated at Lysine 4 of Histone H4 (H3K4) and methylated at H3K9 (Strahl and Allis, 2000; Kouzarides, 2007). Histone modifiers control this modification state: for example the histone deacetylase (HDAC) Sir2 deacetylates H3K9 and/or H4K16, the histone methyltransferase Set1 methylates H3K4 and the histone methyltransferase Su(var)3-9 methylates H3K9 (Shankaranarayanan et al., 2003; Wirén et al., 2005; Bühler and Gasser, 2009; Kueng et al., 2013). Although pericentromeric heterochromatin is usually associated with pericentromeric regions, this repressive chromatin state is not absolutely required for centromere function and faithful chromosome segregation. For example, in *Candida lusitanae* pericentromeric regions are not assembled into heterochromatin (Kapoor et al., 2015). Given the diversity of centromere structure across eukaryotes, it is important to analyze centromere organization in a variety of organisms.

*Candida albicans*, the most common human fungal pathogen, is an ideal system to investigate diversity and structure of centromeres because it possess regional centromeres that are much smaller and simpler than other regional centromeres (Sanyal et al., 2004; Baum et al., 2006; Mishra et al., 2007). Each of the 8 *C. albicans* diploid chromosomes has a relatively small regional centromere (2–4 kb) assembled into CENP-A chromatin (Baum et al., 2006). The organization and sequence of pericentromeric regions differs at each centromere (Figure 1A). Centromeres on chromosome 1, 4, 5, 6, and R are similar to centromeres of the fission yeast *Schizosaccharomyces pombe* where IRs flank the CENP-A containing domain. Contrary to *S. pombe*, the sequence of these repetitive elements is not conserved across centromeres. On chromosome 2 and 3 Long Terminal Repeats (LTRs) are found within ~3 kb of the CENP-A containing domain. Centromere on chromosome 7 does not have obvious repeats nearby (Mishra et al., 2007; Ketel et al., 2009). Therefore, 7 of the 8 pericentromeric regions are associated with DNA repeats. Several lines of evidence suggest that, in *C. albicans*, pericentromeric repeats are important for centromere function and/or establishing centromere identity. First of all, despite the lack of conservation in the DNA sequence, repetitive DNA is also associated with centromere of other *Candida* species such as *C. dubliniensis* (Padmanabhan et al., 2008). In addition, following deletions of endogenous centromeres, *C. albicans* neocentromeres form efficiently and are often assembled in proximity to DNA repeats (Ketel et al., 2009). However, the lack of repetitive elements surrounding the centromere on chromosome 7 argue against a role of DNA repeats in centromere function. In many eukaryotes,

pericentromeric regions are assembled into transcriptionally silent heterochromatin. It is possible that, despite the lack of a conserved DNA sequence and/or DNA feature, the common feature of *C. albicans* pericentromeric regions is a specific chromatin structure resembling heterochromatin. Here, we address this question.

The *C. albicans* epigenome, similarly to *S. cerevisiae*, lacks Su(var) 3–9 orthologs, hence H3 K9 methylation is absent in both *C. albicans* and *S. cerevisiae*.

We have recently shown that, in *C. albicans*, transcriptionally silent heterochromatin is assembled at the ribosomal DNA (*rDNA*) locus and at telomeres (Freire-Benítez et al., 2016). At these locations, heterochromatin is typified by nucleosomes that are hypoacetylated and hypomethylated on H3K4. The histone deacetylase Sir2 (*orf 19.1992*) is required to maintain this repressive epigenetic state via hypoacetylation of H3K9 and H4K16 (Freire-Benítez et al., 2016).

In this study, we investigate the chromatin state associated with *C. albicans* pericentromeric repeats. We find that pericentromeric regions are assembled into an intermediate chromatin state bearing features of both euchromatin (high histone acetylation) and heterochromatin (hypomethylation of H3K4). This intermediate chromatin state is associated with a weak transcriptionally silent environment that partially represses expression of proximal genes and inserted marker genes. Our analysis identifies a new chromatin state associated with pericentromeric regions.

## MATERIALS AND METHODS

### Growth Conditions

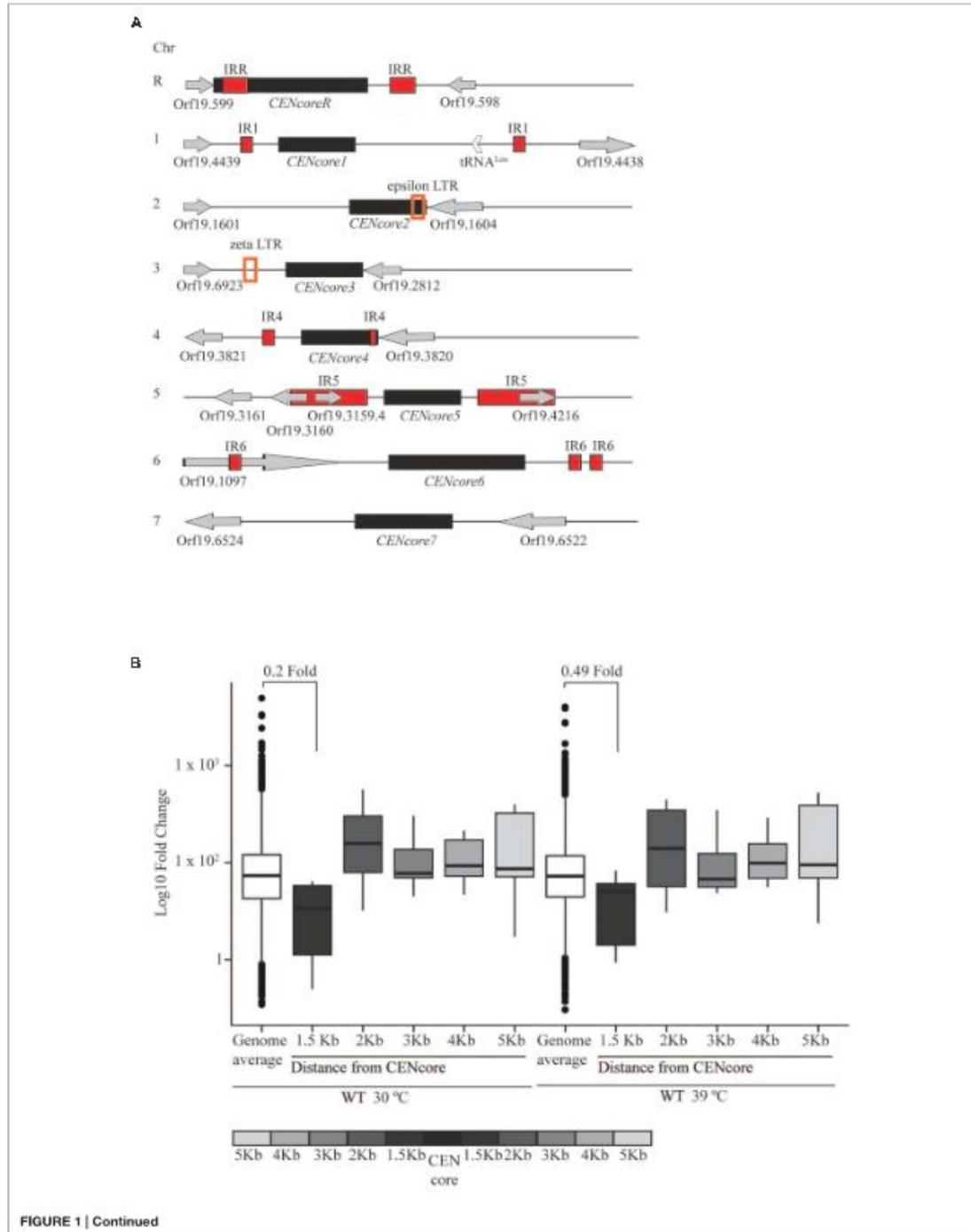
Yeast cells were cultured in rich medium (YPAD) containing extra adenine (0.1 mg/ml) and extra uridine (0.08 mg/ml), complete SC medium (Formedium™), or SC Drop-Out media (Formedium™). Cells were grown at 30 or 39°C as indicated.

### Yeast Strain Construction

Strains are listed in Supplementary Table S1. Integration and deletion of genes were performed using plasmids containing marker genes for substitution or integration at endogenous locus as previously described (Wilson et al., 1999). Transformation was performed by electroporation (Gene Pulser™, Bio-Rad) using the protocol described in (De Backer et al., 1999). *URA3* marker gene was used for silencing assays to determine *URA3* expression in complete SC medium (Formedium™) or SC URA Drop-Out media (Formedium™). *HIS1*, *ARG4*, and *NAT* marker genes were used to replace both copies of *SIR2*, *SET1*, and *JHD2* genes. PCR was used for screening of positive transformants. Oligonucleotides and plasmids used for strain constructions are listed in Supplementary Tables S2 and S3, respectively.

### Silencing Assay

Growth analyses were performed using a plate reader (SpectrostarNano, BMG labtech) in 96 well plate formats at 30°C. For each silencing assay in 96 well plate format, 1:100



**FIGURE 1 | Continued**

**Genes flanking centromeres display reduced expression. (A)** Schematic of *Candida albicans* centromeres. On each chromosome, black squares indicate centromere core regions (CENcore). Red squares indicate inverted repeats (IRs). Orange empty squares indicate Long Terminal Repeats (LTRs). Empty arrows indicate tRNA codons. Gray arrows indicate ORF and their transcription orientation. **(B)** RNA deep-sequencing of WT isolates at 30°C and 39°C. Boxplots represent normalized read counts (FPKM) of median genome expression compare to the median of all CENcore proximal genes from 1.5 to 5 Kb flanking both sides of CENcore regions on all chromosomes.

dilution of and overnight culture was inoculated in a final volume of 95  $\mu$ l of SC or SC-URA media to reach a concentration of 60 cells/ $\mu$ l. Growth was assessed by measuring  $A_{600}$ , using the following conditions: OD<sub>600nm</sub>, 616 cycle time, three flashes per well, 700 rpm shaking frequency, orbital shaking mode, 545 s additional shaking time after each cycle 0.5 s post delay, for 44 h. Graphs represent data from three biological replicates. Error bars: standard deviation (SD) of three biological replicates generated from three independent cultures of the same strain. Data was processed using SpectrostarNano MARS software and Microsoft Excel.

### RNA Extraction and cDNA Synthesis

RNA was extracted from Log<sub>2</sub> exponential cultures (OD<sub>600nm</sub> = 1.4) using a yeast RNA extraction kit (E.Z.N.A.™ Isolation Kit RNA Yeast, Omega Bio-Tek) following the manufacturer's instructions. RNA quality was checked by electrophoresis under denaturing conditions in 1% agarose, 1X HEPES, 6% Formaldehyde (Sigma). RNA concentration was measured using a NanoDrop ND-1000 Spectrophotometer. cDNA synthesis was performed using iScript™ Reverse Transcription Supermix for RT-qPCR (Bio-Rad) following manufacturer's instructions and a Bio-Rad CFXConnect™ Real-Time System.

### High-throughput RNA Sequencing

Strand-specific cDNA Illumina Barcoded Libraries were generated from 1  $\mu$ g of total RNA extracted from wt and *sir2 $\Delta$ / $\Delta$*  strains and sequenced with an Illumina iSeq2000 platform. Illumina Library and Deep-sequencing was performed by the Genomics Core Facility at EMBL (Heidelberg, Germany). Raw reads were analyzed using TopHat algorithm following the RNA deep sequencing analysis pipeline described (Trapnell et al., 2013) using Galaxy<sup>1</sup> and Linux platform. Heatmaps and boxplot graphs were generated with R<sup>2</sup>. RNA sequencing data are deposited into ArrayExpress (accession number E-MTAB-4622).

### Quantitative Chromatin ImmunoPrecipitation (qChIP)

Quantitative Chromatin ImmunoPrecipitation (qChIP) was performed as described (Pidoux et al., 2004) with the following modifications: 5 ml of an overnight culture grown in YPAD with extra uridine (0.08 mg/ml), diluted into fresh YPAD with extra uridine (0.08 mg/ml) and grown until OD<sub>600 nm</sub> of 1.4. Cells (50 ml/sample) were fixed with 1% Paraformaldehyde

(Sigma) for 15 min at room temperature. Cells were lysed using acid-washed glass beads (Sigma) and a Disruptor genie™ (Scientific Industries) for 30 min at 4°C. Chromatin was sheared to 500–1000 bp using a Bioruptor (Diagenode) for a total of 20 min (30 s ON and OFF cycle) at 4°C. Immunoprecipitation was performed overnight at 4°C using 2  $\mu$ L of antibody anti-H3K4me2 (Active Motif- Cat Number: 39141), anti-H3K9ac (Active Motif- Cat Number: 39137), and anti-H4K16ac (Active Motif- Cat Number: 39167) and 25  $\mu$ l of Protein G magnetic beads (Dyna – InVitrogen). DNA was eluted with a 10% slurry of Chelex 100-resin (Bio-Rad) using the manufacturer's instructions.

### qPCR Reactions

Primers used are listed in Supplementary Table S2. Real-time qPCR and RT-qPCR were performed in the presence of SYBR Green (Bio-Rad) on a Bio-Rad CFXConnect™ Real-Time System. Data was analyzed with Bio-Rad CFX Manager 3.1 software and Microsoft Excel. Enrichments were calculated as the percentage ratio of specific IP over input for qChIP analysis and as enrichment over actin for RT-qPCR. Histograms represent data from three biological replicates. Error bars: SD of three biological replicates generated from three independent cultures of the same strain.

## RESULTS

### Expression Level of Centromere-Proximal Genes Is Low

Heterochromatin represses expression of associated and proximal genes (Freitas-Junior et al., 2005; Kaur et al., 2005; Hansen et al., 2006; Merrick and Duraisingh, 2006). Therefore, if the pericentromeric regions of *C. albicans* were assembled into heterochromatin, genes in proximity to these regions would be poorly expressed. To test this hypothesis, we isolated RNA from wild-type (WT) cells grown at a temperature relevant for growth of *C. albicans* on the skin (30°C) and at a temperature mimicking fever in the host (39°C) and performed RNA-seq analyses. FPKM (fragments per kilobase of exons per million mapped reads) values were determined for each annotated gene. We then calculated the FPKM for genes in cumulative bins of 1 kb from the centromeres and compared these values to the genome-wide average (Figure 1B). This analysis reveals that genes in proximity to centromeres (0 to 1.5 kb) are less expressed compared with the genome-wide average ( $p$ -value =  $2.2 \times 10^{-16}$ ). These data suggest that pericentromeric and centromeric regions impose a weak transcriptionally repressive environment.

<sup>1</sup><https://usegalaxy.org/>

<sup>2</sup><http://www.r-project.org/>

## A Marker Gene Inserted at Pericentromeric Repeats Is Partially Repressed

Heterochromatin assembled onto repetitive DNA represses the transcription of marker genes inserted in their proximity (Henikoff and Dreesen, 1989; Gottschling et al., 1990; Bryk et al., 1997; Smith and Boeke, 1997). We have previously shown that, also in *C. albicans*, heterochromatin silences inserted marker genes (Freire-Benítez et al., 2016). Our RNA-seq analyses suggest that regions proximal to a centromere impose a weak transcriptional repressive environment (Figure 1B). To test whether *C. albicans* pericentromeres impose transcriptional silencing dependently or independently of the presence of DNA repeats, we integrated the URA3 marker gene into the pericentromeric regions of centromeres surrounded by DNA repeats (*peri-CEN4:URA3<sup>+</sup>* and *peri-CEN5:URA3<sup>+</sup>*) and into the pericentromeric regions of CEN7, lacking DNA repeats (*peri-CEN7:URA3<sup>+</sup>*). (Figure 2A). To investigate whether the URA3 marker gene is transcriptionally silenced when inserted at these locations, strains were grown in non-selective (N/S) medium and in medium lacking uridine (-Uri) in which only cells expressing sufficient Ura3 protein are able to grow. Silencing of URA3 is expected to result in slower growth in -Uri medium compared to N/S. However, none of the strains grew poorly in -Uri medium compared to N/S medium (Figure 2B). In contrast, quantitative reverse transcriptase analysis (qRT-PCR) reveals that the levels of URA3 mRNA for the *peri-CEN4:URA3<sup>+</sup>*, *peri-CEN5:URA3<sup>+</sup>*, and *peri-CEN7:URA3<sup>+</sup>* strains were significantly lower than the URA3 euchromatic gene (Figure 2C). Therefore, pericentromeric regions impose a weak transcriptional silencing that can only be detected at the RNA level independently of the presence of repetitive elements.

## Pericentromeric Repeats Are Assembled Into an Intermediate Chromatin State Bearing Features of Both Euchromatin and Heterochromatin

We have shown that, in *C. albicans*, heterochromatin regions are assembled into nucleosomes that are hypoacetylated on H3K9 and H4K16 and hypomethylated on H3K4 (Freire-Benítez et al., 2016). To assess whether pericentromeric regions are associated with heterochromatic histone marks, we monitored by *qChIP* the presence of H3K9Ac, H4K16Ac, and H3K4me at pericentromeric regions surrounding CEN4, CEN5, and CEN7 (Figure 3A). As a control, the chromatin state associated with the euchromatic *ACT1* locus was analyzed. We find that all pericentromeric regions analyzed are assembled into chromatin that is highly acetylated on H3K9 and H4K16 as levels of these two histone modifications are similar to levels detected at the active and euchromatic locus *ACT1* (Figures 3B,D, and F). High levels of H4K16 acetylation are also found at the central core region assembled into CENP-A chromatin (Figure 3). In contrast, pericentromeric chromatin is hypomethylated on H3K4, a chromatin state more similar to heterochromatic regions and different from the euchromatic *ACT1* locus

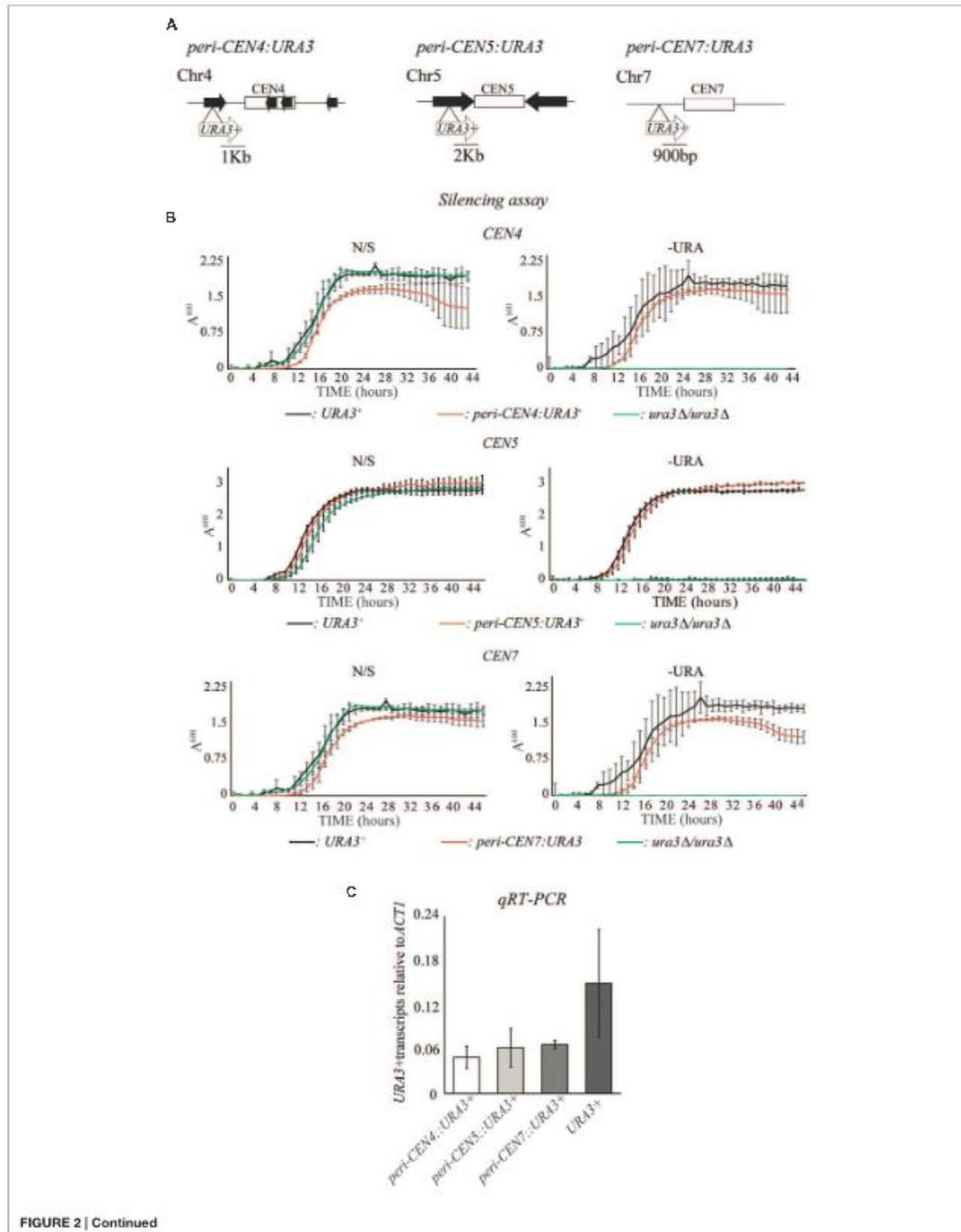
(Figures 3B,D, and F). Thus, pericentromeric regions have only one of the three marks (H3K4 hypomethylation) associated with classic heterochromatin. Importantly, this chromatin state marks pericentromeric regions independently of the presence of repeats as pericentromeres with DNA repeats (CEN4 and CEN5) and without (CEN7) are associated with a similar histone modifications pattern. We concluded that *C. albicans* pericentromeres are not assembled into classical transcriptionally silent heterochromatin but they are associated with an intermediate chromatin state bearing features of euchromatin (high histone acetylation) and heterochromatin (H3K4 hypomethylation).

## The Chromatin State Associated with Pericentromeric Regions Is Independent of the Histone Deacetylase Sir2

The HDAC Sir2 specifically deacetylates H3K9 and H4K16 and it is required for heterochromatin assembly across the eukaryotic kingdom (Rusche et al., 2003). We have shown that *C. albicans* Sir2 is necessary for heterochromatin integrity at the *rDNA* locus and telomeric regions via deacetylation of H3K9 and H4K16 (Freire-Benítez et al., 2016). To assess whether Sir2 contributes to the chromatin and transcriptional state of pericentromeric regions, we isolated RNA from WT and *sir2*  $\Delta/\Delta$  cells and performed RNA-seq analyses. FPKM values were determined for all genes proximal to CEN repeats and compared between *sir2*  $\Delta/\Delta$  and WT strains. Upon deletion of the *SIR2* gene, we did not observe any clear effect on expression of CEN proximal genes (Figure 4A). Only 2 out of the 12 genes located in proximity (<1.5 Kb) of centromeres were expressed more than 2 fold in *sir2*  $\Delta/\Delta$  isolates compared to WT cells (Table 1). Therefore Sir2 does not contribute to the poor expression of CEN-proximal genes. In agreement with these results, deletion of the *SIR2* gene does not increase the levels of H3K9 and H4K16 acetylation associated with pericentromeric region on chromosome five as revealed by *q-ChIP* analyses (Figure 4C).

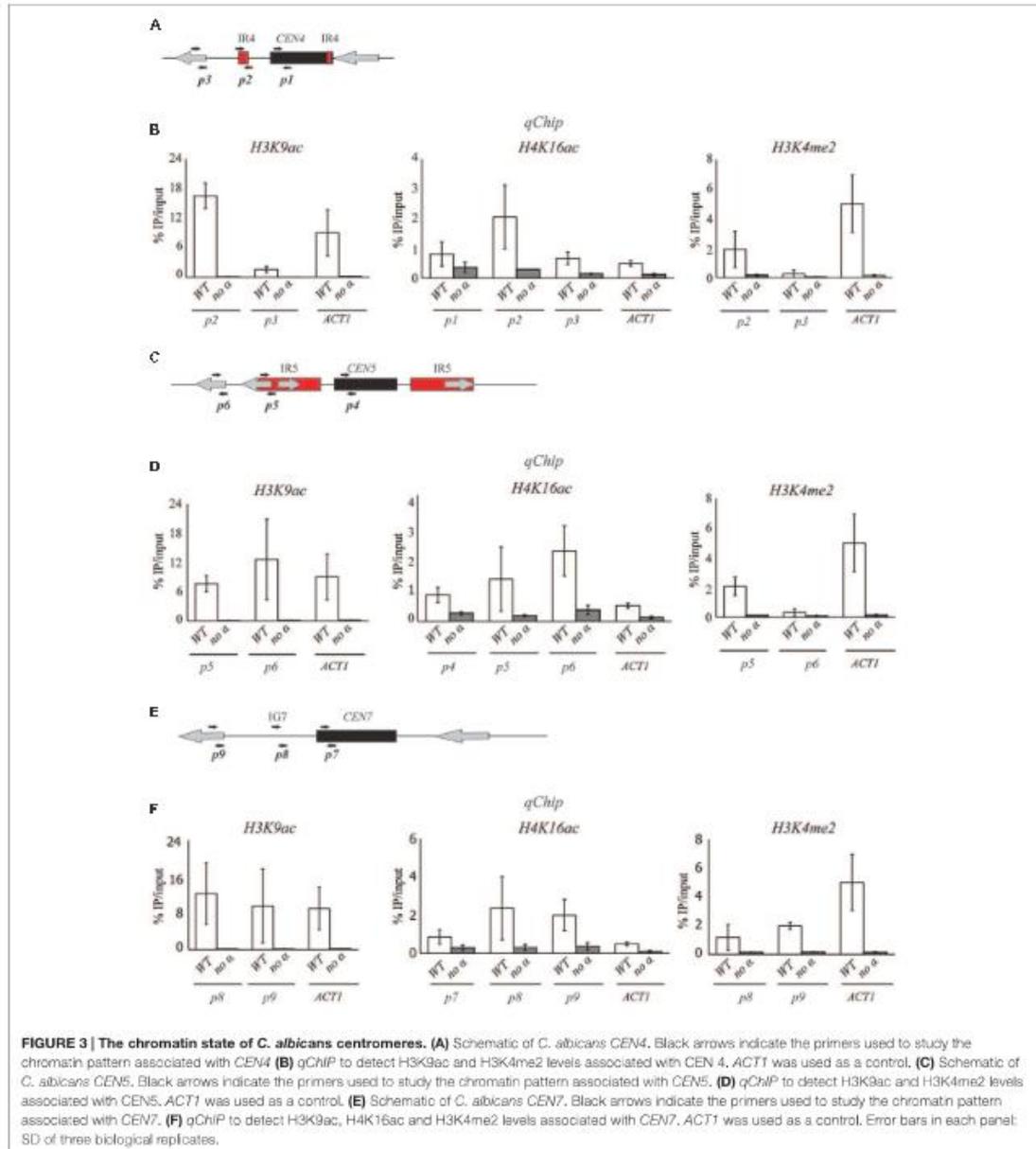
## The Histone Methyltransferase Set1, But Not the Histone Demethylase Jhd2, Contributes to the Chromatin State Associated with Pericentromeres

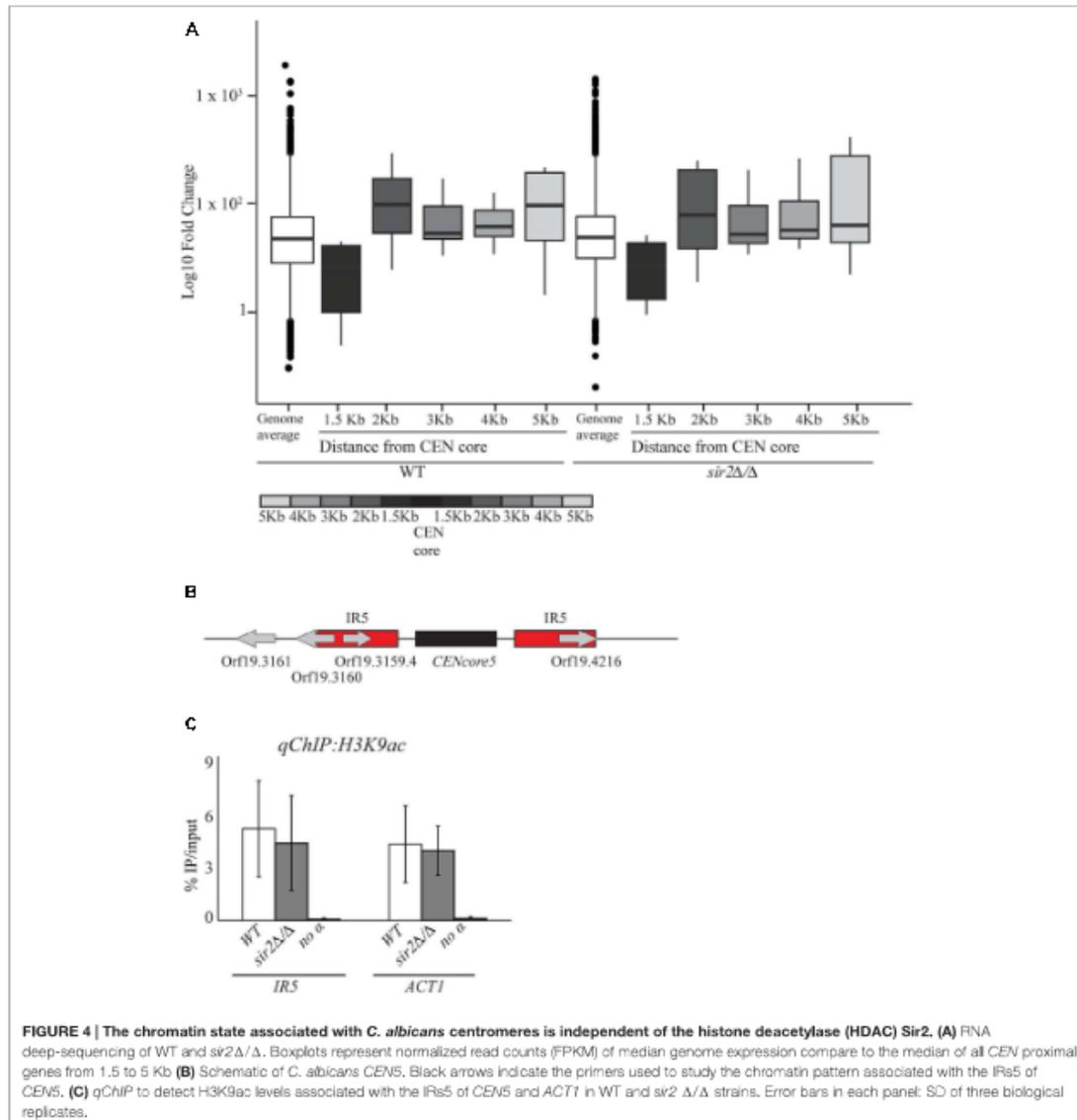
Coordination of activities between histone methyltransferases and histone demethylases ensures the right methylation level associated with euchromatic and heterochromatic loci. Therefore, specific histone methyltransferases and demethylases might be important for maintaining the H3K4 hypomethylated state associated with *C. albicans* pericentromeres. The *C. albicans* genome encodes for the H3K4 methyltransferase Set1 (Raman et al., 2006) and for the putative H3K4 demethylase Kmd5/Jhd2 (*orf19.5651*). In *S. cerevisiae*, both proteins have been implicated in transcriptional silencing and heterochromatin formation (Briggs et al., 2001; Ingvarsdottir et al., 2007; Ryu and Ahn, 2014). To assess whether Set1 and/or Jhd2 contribute to the chromatin state associated with *C. albicans* pericentromeric regions, we deleted both copies of the *SET1* and *JHD2* genes



**FIGURE 2 | Continued**

**Weak transcriptional silencing of a *URA3* marker gene integrated at pericentromeric region. (A)** Schematic of *peri-CEN4:URA3<sup>+</sup>*, *peri-CEN5:URA3<sup>+</sup>*, and *peri-CEN7:URA3<sup>+</sup>* reporter strains. Distance from the *URA3* marker gene to each *CEN* core is indicated in Kb. **(B)** Silencing assay of the *peri-CEN4:URA3<sup>+</sup>*, *peri-CEN5:URA3<sup>+</sup>*, and *peri-CEN7:URA3<sup>+</sup>* reporter strain. *URA3<sup>+</sup>* (*URA3/URA3*) and *Ura<sup>-</sup>* (*ura3Δ/ura3Δ*) strains were included as controls. **(C)** qRT-PCR analyses to measure *URA3<sup>+</sup>* transcript levels of the *peri-CEN4:URA3<sup>+</sup>*, *peri-CEN5:URA3<sup>+</sup>*, and *peri-CEN7:URA3<sup>+</sup>* reporter strain relative to actin transcript levels (*ACT1*). Error bars in each panel: standard deviation (SD) of three biological replicates.





from WT cells and quantified H3K4me2 levels by *qChIP* analyses. We find that *Set1* is necessary for maintaining the low levels of H3K4me associated with *CEN5* repeats (Figure 5A) as H3K4me dropped to background levels in *set1Δ/Δ* compared to WT cells (Figure 5B). In contrast, we find that *Jhd2* is not required for maintaining the hypomethylated state associated with *CEN5* repeats as H3K4me levels did not change between *jhd2Δ/Δ* and WT cells (Figure 5C).

## DISCUSSION

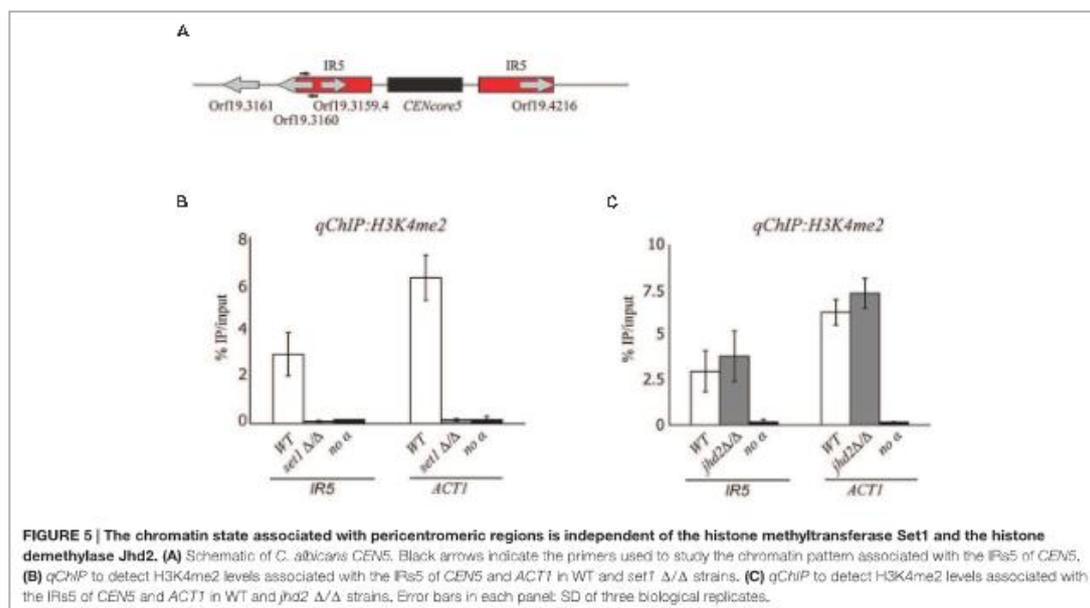
This study is the first analysis of the chromatin state associated with pericentromeric regions in the human fungal pathogen *C. albicans*.

In many organisms, regional centromeres have a conserved modular structure despite the lack of a conserved DNA sequence. At these locations, CENP-A domains, sites of kinetochore assembly, are flanked or interspersed by DNA repeats assembled

**TABLE 1 |** RNA deep-sequencing values of 1.5 Kb CEN proximal genes.

ORF	Gene	CENcore	bp distance to CENcore	WT FPKM	<i>sir2</i> Δ/Δ FPKM	( <i>sir2</i> Δ/Δ/WT) FPKM Fold	Log2(Fold)
<i>orf19.4439</i>	–	1	1500	6.60168	10.8237	1.639537	0.713284
<i>orf19.1604</i>	–	2	96	27.1566	36.3803	1.339649	0.421855
<i>orf19.2812</i>	–	3	100	29.8328	14.1075	0.472866	-1.08044
<i>orf19.3820</i>	–	4	51	0.913477	0.96385	1.055144	0.07744
<i>orf19.3159.4</i>	–	5	926	0.516172	2.34077	4.534864	2.18106
<i>orf19.1097</i>	<i>ALS2</i>	6	1498	9.68354	39.5448	4.083713	2.029882
<i>orf19.2123</i>	–	6	231	34.9161	47.2585	1.353487	0.43668
<i>orf19.2124</i>	–	6	765	20.2353	29.1543	1.440764	0.526838
<i>orf19.599</i>	–	R	102	1.24606	1.76585	1.417147	0.502991
<i>orf19.598</i>	–	R	1126	0.209874	1.269	6.046485	-1.16192

ORFs include all genes within 1.5 Kb of a centromere core. Values are indicated in FPKM for WT and *sir2* Δ/Δ strains, including FPKM fold increase and Log2 fold increase of *sir2* Δ/Δ strain over WT strain.



into heterochromatin (Buscaino et al., 2010). Both CENP-A chromatin and pericentromeric heterochromatin are inhibitory to transcription as illustrated by silencing of inserted marker genes (Allshire et al., 1994; Ketel et al., 2009).

The small *C. albicans* regional centromeres are composed of a CENP-A central core surrounded by pericentromeric regions with different sequence and organization. CENP-A chromatin represses transcription of embedded marker genes (Ketel et al., 2009). Here, we show that, despite the absence of canonical heterochromatin, *C. albicans* pericentromeric regions weakly repress transcription. These conclusions are supported by two observations. Firstly, RNA-sequence analysis highlights that genes located in proximity (<1.5 Kb) to centromeres are less expressed than the genome-wide average (Figure 1B). Secondly, a *URA3* marker gene inserted into the pericentromeric region of

chromosomes 4, 5, and 7 is weakly silenced. This weak silencing is observed when *URA3* RNA levels are measured by *qRT-PCR* but no silencing is detected by growing strains in -*URA* medium (Figure 2). This suggests that the reduced expression of the *URA3*<sup>+</sup> marker gene is sufficient to confer a *URA*<sup>+</sup> growth phenotype.

*Candida albicans* lacks the heterochromatic structure that is normally associated with regional centromeres because the *C. albicans* genome does not encode for *SuVar* 3-9, the H3K9-specific methyltransferase, or for HP-1, the chromodomain protein responsible for assembly and spreading of centromeric heterochromatin. We have recently shown that transcriptionally silenced heterochromatin exists in *C. albicans* and it is associated with the *rDNA* locus and telomeric regions (Freire-Benítez et al., 2016). This heterochromatic state,

similarly to *S. cerevisiae*, is characterized by hypoacetylated nucleosomes that are hypomethylated on H3K4 (Freire-Benítez et al., 2016). Here, we show that *C. albicans* pericentromeric regions are not associated with heterochromatin but with an intermediate chromatin state bearing features of both euchromatin and heterochromatin. Nucleosomes at pericentromeric regions are highly acetylated, as observed in euchromatin, and hypomethylated on H3K4, as observed in heterochromatin (Figure 3). Consistently, we find that deletion of the HDAC Sir2 does not perturb the chromatin state of pericentromeric regions (Figure 4).

Given the lack of canonical heterochromatin at pericentromeric regions, it is still unclear what drives the weak transcriptional silencing associated with these regions. We envisage two possible scenarios. It is possible that low levels of CENP-A, undetectable by ChIP analyses, associate with pericentromeric regions driving transcriptional silencing. In support of this hypothesis, it is well known that CENP-A chromatin is inhibitory to transcription. In addition, it has been observed that, following deletion of endogenous centromeric sequences, neocentromeres often form immediately adjacent to the site of the excised native centromeres (Ketel et al., 2009; Shang et al., 2013; Thakur and Sanyal, 2013). These data suggest that low levels of CENP-A might be present at pericentromeric regions and might be sufficient to nucleate new centromeres. In agreement with this hypothesis, it has been shown that a pool of free CENP-A accessory molecules is present in vicinity of centromeres. These accessory molecules do not nucleate kinetochore assembly but could allow for rapid incorporation of CENP-A in the event of eviction at the centromere (Haase et al., 2013). Therefore, it is possible that CENP-A accessory molecules drive the transcriptional silencing associated with pericentromeric repeats. Finally, it has been shown that, following neocentromere formation, assembly of CENP-A into transcribed regions is sufficient to repress gene expression (Shang et al., 2013).

## REFERENCES

- Allshire, R. C., Javerzat, J. P., Redhead, N. J., and Cranston, G. (1994). Position effect variegation at fission yeast centromeres. *Cell* 76, 157–169. doi: 10.1016/0092-8674(94)90180-5
- Baum, M., Ngan, V. K., and Clarke, L. (1994). The centromeric K-type repeat and the central core are together sufficient to establish a functional *Schizosaccharomyces pombe* centromere. *Mol. Biol. Cell* 5, 747–761. doi: 10.1091/mbc.5.7.747
- Baum, M., Sanyal, K., Mishra, P. K., Thaler, N., and Carbon, J. (2006). Formation of functional centromeric chromatin is specified epigenetically in *Candida albicans*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 14877–14882. doi: 10.1073/pnas.0606958103
- Bernard, P., Maure, J. F., Partridge, J. F., Genier, S., Javerzat, J. P., and Allshire, R. C. (2001). Requirement of heterochromatin for cohesion at centromeres. *Science* 294, 2539–2542. doi: 10.1126/science.1064027
- Briggs, S. D., Bryk, M., Strahl, B. D., Cheung, W. L., Davie, J. K., Dent, S. Y., et al. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev.* 15, 3286–3295. doi: 10.1101/gad.940201
- Bryk, M., Banerjee, M., Murphy, M., Knudsen, K. E., Garfinkel, D. J., and Curcio, M. J. (1997). Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. *Genes Dev.* 11, 255–269. doi: 10.1101/gad.11.2.255
- Bühler, M., and Gasser, S. M. (2009). Silent chromatin at the middle and ends: lessons from yeasts. *EMBO J.* 28, 2149–2161. doi: 10.1038/emboj.2009.185
- Buscaino, A., Allshire, R., and Pidoux, A. (2010). Building centromeres: home sweet home or a nomadic existence? *Curr. Opin. Genet. Dev.* 20, 118–126. doi: 10.1016/j.gde.2010.01.006
- Chatterjee, G., Sankaranarayanan, S. R., Guin, K., Thattikota, Y., Padmanabhan, S., Siddharthan, R., et al. (2016). Repeat-associated fission yeast-like regional centromeres in the ascomycetous budding yeast *Candida tropicalis*. *PLoS Genet.* 12:e1005839. doi: 10.1371/journal.pgen.1005839
- De Backer, M. D., Maes, D., Vandoninck, S., Logghe, M., Contreras, R., and Luyten, W. H. (1999). Transformation of *Candida albicans* by electroporation. *Yeast* 15, 1609–1618. doi: 10.1002/(SICI)1097-0061(199911)15:15<1609::AID-YEA485>3.3.CO;2-P
- Fischle, W., Wang, Y., and Allis, C. D. (2003). Histone and chromatin cross-talk. *Curr. Opin. Cell Biol.* 15, 172–183. doi: 10.1016/S0955-0674(03)00013-9
- Folco, H. D., Pidoux, A. L., Urano, T., and Allshire, R. C. (2008). Heterochromatin and RNAi are required to establish CENP-A chromatin at centromeres. *Science* 319, 94–97. doi: 10.1126/science.1150944

## FUNDING

This work was supported by BBSRC (BB/L008041/1 to AB), MRC (MR/M019713/1 to RJP and AB), and a Royal Society Research Grant (RG130149 to AB).

## ACKNOWLEDGMENTS

We thank J. Berman and A. Mitchell for strains and materials and S. Gourlay for critical reading of the manuscript. We thank the Gene Core Facility at EMBL (Heidelberg-Germany) for RNA Sequencing and M. Wass for support with the RNA Seq analyses.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fmicb.2016.00759>

- Freire-Benítez, V., Price, R. J., Tarrant, D., Berman, J., and Buscaino, A. (2016). *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.* 6, 22989. doi: 10.1038/srep22989
- Freitas-Junior, L. H., Hernandez-Rivas, R., Ralph, S. A., Montiel-Condado, D., Ruvalcaba-Salazar, O. K., Rojas-Meza, A. P., et al. (2005). Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. *Cell* 121, 25–36. doi: 10.1016/j.cell.2005.01.037
- Gottschling, D. E., Aparicio, O. M., Billington, B. L., and Zakian, V. A. (1990). Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* 63, 751–762. doi: 10.1016/0092-8674(90)90141-Z
- Haase, J., Mishra, P. K., Stephens, A., Haggerty, R., Quammen, C., Taylor, R. M., et al. (2013). A 3D map of the yeast kinetochore reveals the presence of core and accessory centromere-specific histone. *Curr. Biol.* 23, 1939–1944. doi: 10.1016/j.cub.2013.07.083
- Hansen, K. R., Ibarra, P. T., and Thon, G. (2006). Evolutionary-conserved telomere-linked helicase genes of fission yeast are repressed by silencing factors, RNAi components and the telomere-binding protein Taz1. *Nucleic Acids Res.* 34, 78–88. doi: 10.1093/nar/gkj415
- Henikoff, S., and Dreesen, T. D. (1989). Trans-inactivation of the *Drosophila* brown gene: evidence for transcriptional repression and somatic pairing dependence. *Proc. Natl. Acad. Sci. U.S.A.* 86, 6704–6708. doi: 10.1073/pnas.86.17.6704
- Ingvarsdottir, K., Edwards, C., Lee, M. G., Lee, J. S., Schultz, D. C., Shilatifard, A., et al. (2007). Histone H3 K4 demethylation during activation and attenuation of GAL1 transcription in *Saccharomyces cerevisiae*. *Mol. Cell Biol.* 27, 7856–7864. doi: 10.1128/MCB.00801-07
- Kapoor, S., Zhu, L., Froyd, C., Liu, T., and Rusche, L. N. (2015). Regional centromeres in the yeast *Candida lusitanae* lack pericentromeric heterochromatin. *Proc. Natl. Acad. Sci. U.S.A.* 112, 12139–12144. doi: 10.1073/pnas.1508749112
- Kaur, R., Domergue, R., Zupancic, M. L., and Cormack, B. P. (2005). A yeast by any other name: *Candida glabrata* and its interaction with the host. *Curr. Opin. Microbiol.* 8, 378–384. doi: 10.1016/j.mib.2005.06.012
- Ketel, C., Wang, H. S. W., McClellan, M., Bouchonville, K., Selmecki, A., Lahav, T., et al. (2009). Neocentromeres form efficiently at multiple possible loci in *Candida albicans*. *PLoS Genet.* 5:e1000400. doi: 10.1371/journal.pgen.1000400
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693–705. doi: 10.1016/j.cell.2007.02.005
- Kueng, S., Oppikofer, M., and Gasser, S. M. (2013). SIR proteins and the assembly of silent chromatin in budding yeast. *Annu. Rev. Genet.* 47, 275–306. doi: 10.1146/annurev-genet-021313-173730
- Lachner, M., Sengupta, R., Schotta, G., and Jenuwein, T. (2004). Trilogies of histone lysine methylation as epigenetic landmarks of the eukaryotic genome. *Cold Spring Harb. Symp. Quant. Biol.* 69, 209–218. doi: 10.1101/sqb.2004.69.209
- Merrick, C. J., and Duraisingh, M. T. (2006). Heterochromatin-mediated control of virulence gene expression. *Mol. Microbiol.* 62, 612–620. doi: 10.1111/j.1365-2958.2006.05397.x
- Mishra, P. K., Baum, M., and Carbon, J. (2007). Centromere size and position in *Candida albicans* are evolutionarily conserved independent of DNA sequence heterogeneity. *Mol. Genet. Genomics* 278, 455–465. doi: 10.1007/s00438-007-0263-8
- Nonaka, N., Kitajima, T., Yokobayashi, S., Xiao, G., Yamamoto, M., Grewal, S. I. S., et al. (2002). Recruitment of cohesin to heterochromatic regions by Swi6/HP1 in fission yeast. *Nat. Cell Biol.* 4, 89–93. doi: 10.1038/ncb739
- Padmanabhan, S., Thakur, J., Siddharthan, R., and Sanyal, K. (2008). Rapid evolution of Cse4p-rich centromeric DNA sequences in closely related pathogenic yeasts, *Candida albicans* and *Candida dubliniensis*. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19797–19802. doi: 10.1073/pnas.0809770105
- Pidoux, A., Mellone, B., and Allshire, R. (2004). Analysis of chromatin in fission yeast. *Methods* 33, 252–259. doi: 10.1016/j.ymeth.2003.11.021
- Raman, S. B., Nguyen, M. H., Zhang, Z., Cheng, S., Jia, H. Y., Weisner, N., et al. (2006). *Candida albicans* SET1 encodes a histone 3 lysine 4 methyltransferase that contributes to the pathogenesis of invasive candidiasis. *Mol. Microbiol.* 60, 697–709. doi: 10.1111/j.1365-2958.2006.05121.x
- Rusche, L. N., Kirchmaier, A. L., and Rine, J. (2003). The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* 72, 481–516. doi: 10.1146/annurev.biochem.72.121801.161547
- Ryu, H.-Y., and Ahn, S. (2014). Yeast histone H3 lysine 4 demethylase Jhd2 regulates mitotic ribosomal DNA condensation. *BMC Biol.* 12:75. doi: 10.1186/s12915-014-0075-3
- Sanyal, K., Baum, M., and Carbon, J. (2004). Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11374–11379. doi: 10.1073/pnas.0404318101
- Shang, W. H., Hori, T., Martins, N. M. C., Toyoda, A., Misu, S., Monma, N., et al. (2013). Chromosome engineering allows the efficient isolation of vertebrate neocentromeres. *Dev. Cell* 24, 635–648. doi: 10.1016/j.devcel.2013.02.009
- Shankaranarayana, G. D., Motamedi, M. R., Moazed, D., and Grewal, S. I. S. (2003). Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. *Curr. Biol.* 13, 1240–1246. doi: 10.1016/S0960-9822(03)00489-5
- Smith, J. S., and Boeke, J. D. (1997). An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev.* 11, 241–254. doi: 10.1101/gad.11.2.241
- Strahl, B. D., and Allis, C. D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45. doi: 10.1038/47412
- Thakur, J., and Sanyal, K. (2013). Efficient neocentromere formation is suppressed by gene conversion to maintain centromere function at native physical chromosomal loci in *Candida albicans*. *Genome Res.* 23, 638–652. doi: 10.1101/gr.141614.112
- Trapnell, C., Hendrickson, D. G., Sauvageau, M., Goff, L., Rinn, J. L., and Pachter, L. (2013). Differential analysis of gene regulation at transcript resolution with RNA-seq. *Nat. Biotechnol.* 31, 46–53. doi: 10.1038/nbt.2450
- Westermann, S., Drubin, D. G., and Barnes, G. (2007). Structures and functions of yeast kinetochore complexes. *Annu. Rev. Biochem.* 76, 563–591. doi: 10.1146/annurev.biochem.76.052705.160607
- Wilson, R. B., Davis, D., and Mitchell, A. P. (1999). Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. *J. Bacteriol.* 181, 1868–1874.
- Wirén, M., Silverstein, R. A., Sinha, L., Walfridsson, J., Lee, H.-M., Laursen, P., et al. (2005). Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. *EMBO J.* 24, 2906–2918. doi: 10.1038/sj.emboj.7600758

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Freire-Benítez, Price and Buscaino. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

#### **4. Supplementary material**



*Supplementary Material*

**The chromatin state of *Candida albicans* pericentromeric repeats bears  
features of both euchromatin and heterochromatin**

Verónica Freire-Benítez, Robert Jordan Price and Alessia Buscaino\*

\* **Correspondence:** Alessia Buscaino : [A.Buscaino@kent.ac.uk](mailto:A.Buscaino@kent.ac.uk)

Supplementary table 1. Strains used in this study

Strain number	Description	Genotype
Bu_20	<i>sir2Δ/Δ</i>	<i>ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG sir2 Δ::HIS1/sir2 Δ::ARG4</i>
Bu_70	<i>set1Δ/Δ</i>	<i>arg4Δ/arg4Δ his1Δ/his1Δ leu2Δ/leu2Δ URA3/ura3Δ::λimm434 IRO1/iro1Δ::λimm434 set1Δ::C.d.HIS1/set1Δ::C.m.LEU2</i>
Bu_215	BWP17	<i>ura3Δ::λimm434/ura3Δ::λimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG</i>
Bu_225	<i>peri-CEN5:URA3<sup>+</sup></i>	<i>rpt-CEN5::URA3 ura3Δ::λimm434/ura3Δimm434 HIS1::his1::hisG/his1::hisG ARG4::arg4::hisG/arg4::hisG</i>
Bu_243	<i>peri-CEN5:URA3<sup>+</sup> sir2Δ/Δ</i>	<i>ura3Δ::λimm434/ura3Δ::λimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG CEN:URA3 sir2Δ::HIS1/sir2Δ::NAT</i>
Bu_380	<i>jhd2Δ/Δ</i>	<i>ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG jhd2 Δ::HIS1/jhd2 Δ::ARG4</i>
Bu_419	<i>peri-CEN4:URA3<sup>+</sup></i>	<i>rpt-CEN4::URA3 ura3Δ::λimm434/ura3Δimm434 HIS1::his1::hisG/his1::hisG ARG4::arg4::hisG/arg4::hisG</i>
Bu_420	<i>peri-CEN7:URA3<sup>+</sup></i>	<i>CEN7::URA3 ura3Δ::λimm434/ura3Δimm434 HIS1::his1::hisG/his1::hisG ARG4::arg4::hisG/arg4::hisG</i>

Supplementary table 2. Primers used in this study

Primer	Sequence	Figure	Description
Bu_113	TTTCTCCGGCGTCAGACATTTT GCAGTTTTCTATGGGATTTATGG TGTTTGTGCGAAAAAAACAAG ATTGTTTTCCAGTCACGACGT T	Fig 2B- C	<i>peri-CEN5:URA3<sup>+</sup></i>
Bu_114	TGTATATCTTCGAGGAATGGCA ACCTTTGCCCCCTCTCGAAAA ACAATATAAATAGAGTCAATTC TCTAGTAGAGGTAAATTCTTTGT GTGGAATTGTGAGCGGATA	Fig 2B- C	<i>peri-CEN5:URA3<sup>+</sup></i>
Bu_119	CCCTCTGCTTGTTCGGATTG	Fig Fig 2B-C 3D, 4C, 5B	Primer to check <i>peri-CEN5:URA3<sup>+</sup></i> , <i>p5: qChip</i>
Bu_120	GGCAAGGAACAAGTCACCAG	Fig 3D, 4C, 5B	<i>p5: qChip</i>
Bu_130	GTGTGTATGGGGTTGTTGCTC	Fig 2B- C	Primer to check <i>rpt- peri-CEN4:URA3<sup>+</sup></i> , <i>peri-CEN5:URA3<sup>+</sup></i> , <i>peri-CEN7:URA3<sup>+</sup></i>

Bu_139	GAGTGAGTGAGTGGAGTAGCG	Fig4C	Primer to check <i>HIS1/NAT</i> replacement of <i>SIR2</i>
Bu_141	GTTGGGCAGATATTACCAATG	Fig 2C	<i>URA3</i> , RT-qPCR
Bu_152	CTGGAGAAAATATAACCACGAG TCTAAGTTTCTTTATTATATTGAC GTTTCAGTTATTTGAGAGAAAT CCTCTAGTAGTTTCCAGTCA CGACGTT	Fig 4C	<i>sir2Δ/Δ</i> deletion mutant: <i>HIS1</i>
Bu_153	ATATATAAATATATAAATATATATA TATAAAAGAATTGAAAAGAAAA ACATTAAAGACACCAATATTAAT TTAATGTGGAATTGTGAGCGGA TA	Fig 4C	<i>sir2Δ/Δ</i> deletion mutant: <i>HIS1</i>
Bu_164	CGGTCTGGTAAATGATTGAC	Fig 4C, 5C	Primer to check <i>HIS1</i> integration
Bu_165	AGTGTGGAAAGAAGAGATGC	Fig 5C	Primer to check <i>ARG4</i> integration
Bu_174	CTACGTTTCCATTCAAGCTGTT	Fig 2C, 3B, 3D, 3F, 4C, 5B, 5C	<i>Act1</i> : qChip, RT-qPCR
Bu_176	AAACTGTAACCACGTTTCAGACA	Fig 2C, 3B, 3D, 3F, 4C, 5B, 5C	<i>Act1</i> : qChip, RT-qPCR
Bu_179	CTGTATCTATAAGCAGTATCATC C	Fig 4C	Primer to check <i>NAT</i> integration
Bu_204	CAAATTCCTTATCGGATTTAGC	Fig 2C	<i>URA3</i> , RT-qPCR
Bu_286	CTGGAGAAAATATAACCACGAG TCTAAGTTTCTTTATTATATTGAC GTTTCAGTTATTTGAGAGAAAT CCTCTAGTAGTAAAACGACGGC CAGTGAATTC	Fig 4C	<i>sir2Δ/Δ</i> deletion mutant: <i>NAT</i>
Bu_287	ATATATAAATATATAAATATATATA TATAAAAGAATTGAAAAGAAAA ACATTAAAGACACCAATATTAAT TTAATGCATCAATTGACGTTGAT ACCA	Fig 4C	<i>sir2Δ/Δ</i> deletion mutant: <i>NAT</i>
Bu_466	TATATTTATTGTGGTTGAAATTT TATATACGTCAAAGATTGCCAA ACGTTTGCTGTACACCAAGGTA TTCCAAGGTTTCCAGTCAC GACGTT	Fig5C	<i>sir2Δ/Δ</i> deletion mutant: <i>HIS1/ARG4</i>
Bu_467	CGTGCTCATAATTAATGAGAG	Fig5C	<i>sir2Δ/Δ</i> deletion

## Supplementary Material

	ACATCACTATGCTAATGTCGCA AAAAAACTGTATACATTTCACT ACTACACATTGTGGAATTGTGA GCGGATA		mutant: <i>HIS1/ARG4</i>
Bu_468	GATATAGTTATCCCTGGCTA	Fig5C	Primer to check <i>HIS1/ARG4</i> replacement of <i>JHD2</i>
Bu_517	CCTCCTGTCAACGACCAAC	Fig 3B	<i>p2</i> : qChip
Bu_518	AACAGCGGTGGTATAGACGT	Fig 3B	<i>p2</i> : qChip
Bu_553	AGAACTCAATGGCAGACTTAA ACCCACAACAAAAAGACAATT GAATAAATGATCCTCCTGTCAA CGACCAACGTTTTCCAGTCAC GACGTT	Fig 2B- C	<i>peri-CEN4:URA3<sup>+</sup></i>
Bu_554	CTTGAGTTCGGTGGTCCCTAT TTATATTGGTGACATTTCTCTTG GTGATATTTGTGAAGTTACAAAT TAAGGTGTGGAATTGTGAGCGG ATA	Fig 2B- C	<i>peri-CEN4:URA3<sup>+</sup></i>
Bu_555	GTCAAAGAGTATTACGAATATA AGGAAAAAACTCACGTATTATT CTACAGTAAGCCGAAACAGAA CAATAGTGTTTTCCAGTCACG ACGTT	Fig 2B- C	<i>peri-CEN7:URA3<sup>+</sup></i>
Bu_556	AATTACACTAATTGTGACTTAG ACAACAGTATACAGCAAGATTG CTAAACTCTATTATAGTTTTCA TCTAGCTGTGGAATTGTGAGCG GATA	Fig 2B- C	<i>peri-CEN7:URA3<sup>+</sup></i>
Bu_557	CTCCCAGTTCCAACACAATTC	Fig 2B- C	Primer to check <i>peri-CEN4:URA3<sup>+</sup></i>
Bu_558	CTGGTCCTCAAGTTGATGTGA	Fig 2B- C	Primer to check <i>peri-CEN7:URA3<sup>+</sup></i>
Bu_559	CATGAAGCGTTCACATGGT	Fig 3F	<i>p7</i> : qChip
Bu_560	CGACAGTGTCAGGTATGCTTAG	Fig 3F	<i>p7</i> : qChip
Bu_561	TGGAGATTCCCTACGATGGTA	Fig 3F	<i>p8</i> : qChip
Bu_562	AGATCACAGCCGACTTCAGT	Fig 3F	<i>p8</i> : qChip
Bu_563	GCACCTTGAGCTTGGAGTT	Fig 3F	<i>p9</i> :qChip
Bu_564	TCAGGTAATGAATTCAGTGGAG	Fig 3F	<i>p9</i> :qChip
Bu_565	AGGCAACCGAAAGAGCTTC	Fig 3B	<i>p1</i> : qChip
Bu_566	TAGATGGTAGAGTTCCTCCTGC	Fig 3B	<i>p1</i> : qChip
Bu_567	GTCATTAGGATATCGGGACACC	Fig 3B	<i>p3</i> :qChip
Bu_568	TCTATGTACAGCAGTAAGGGGT G	Fig 3B	<i>p3</i> :qChip
Bu_569	CGGCCATCATTGCACTATTA	Fig 3D	<i>p4</i> : qChip

Bu_570	CTGGTGACGATGTCACCTATGG	Fig 3D	<i>p4</i> :qChip
Bu_571	AGCTTGGCATCCGACTTATT	Fig 3D	<i>p6</i> :qChip
Bu_572	GGTTCACCGTATATGAAGC	Fig 3D	<i>p6</i> :qChip

Supplementary table 3. Plasmids used in this study

Plasmid	Description
pGEMURA3	<i>URA3</i> integration products (Wilson et al., 1999)
pGEMHIS1	<i>HIS1</i> substitution products (Wilson et al., 1999) (Wilson et al, 1999)
pHA_NAT	<i>NAT</i> substitution products (Gerami-nejad et al., 2012)

## References

- Gerami-nejad, M., Forche, A., McClellan, M., and Berman, J. (2012). Analysis of protein function in clinical *C. albicans* isolates. 5314. doi:10.1002/yea.
- Wilson, R. B., Davis, D., and Mitchell, A. P. (1999). Rapid Hypothesis Testing with *Candida albicans* through Gene Disruption with Short Homology Regions *J. Bacteriol* 181, 1868-74



## **Chapter 4. Heterochromatin and genome stability in *C. albicans*.**

**It contains the article:** Freire-Benítez, V., Gourlay, S., Berman, J., and Buscaino, A. (2016c). Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen *Candida albicans*. *Nucleic Acids Res.* gkw594.

### **Author contributions**

VFB performed all the experimental work, new mutant construction, bioinformatics and result analysis show in figures 1 – 8 and supplementary figures 1 -9. VFB made all the figures, supplementary tables and wrote the first draft of the manuscript.

SG assisted with media preparation and colony counting.

AB wrote the final manuscript and VFB made the final figures.

## 1. Summary

Repetitive DNA sequences are hotspots for recombination events and genome instability due to DNA homology. Heterochromatin is assembled at clusters of repeats to prevent unfaithful recombination events. However, some organisms such as microbial pathogens required genome variability in order to adapt to environmental changing conditions. *C. albicans* is a great example. It is a successful human fungal pathogen due in part to its high genome instability and variability. Previous studies showed that *C. albicans* genome undergoes loss of heterozygosity events (LOH) and whole chromosome aneuploidies under stress conditions. These mechanisms to generate genome variability are of crucial importance in an organism that cannot undergo meiosis.

The *rDNA* locus and telomeres are very variable regions in eukaryotic genomes despite the fact that heterochromatin is assembled at these positions to prevent genome instability. In chapter 2 we showed that heterochromatin is assembled at the *rDNA* locus and telomere repeats of *C. albicans* but it was not known whether heterochromatin promotes genome stability at these regions.

In this chapter, by using LOH fluctuation analysis and marker loss assays we demonstrated that Sir2 does not promote genome stability at the *rDNA* locus. This role is played by the Monopolin complex. In contrast, we have demonstrated that a group of genes located at subtelomeric regions, the *TLO* gene family, is associated with high recombination rates. We have identified a DNA element (*TRE*: Telomere Recombination Element) that is associated with a subset of *TLO* genes and it a hotspot for recombination. Interestingly *TRE* recombination rates are SIR2-dependent.

Stress conditions such as fluconazole and H<sub>2</sub>O<sub>2</sub> increase recombination rates at all loci tested without increasing whole chromosome aneuploidies. This increase of recombination bypasses the Sir2 control exerted at the *TRE* element.

In conclusion, we show different regulation of genome stability at *C. albicans* repeats. While the Monopolin complex controls recombination rates associated with the *rDNA* locus, Sir2 exerts its effect at the subtelomeric *TRE* element adjacent to some *TLO* members. This chapter highlights the differential control of genome variability associated with DNA repeats at *C. albicans*.

## 2. Main article

## Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen *Candida albicans*

Verónica Freire-Benítez<sup>1</sup>, Sarah Gourlay<sup>1</sup>, Judith Berman<sup>2</sup> and Alessia Buscaino<sup>1\*</sup>

<sup>1</sup>University of Kent, School of Biosciences, Canterbury, Kent CT2 7NJ, UK and <sup>2</sup>Department of Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, 69978, Israel

Received April 20, 2016; Accepted June 21, 2016

### ABSTRACT

DNA repeats, found at the ribosomal DNA locus, telomeres and subtelomeric regions, are unstable sites of eukaryotic genomes. A fine balance between genetic variability and genomic stability tunes plasticity of these chromosomal regions. This tuning mechanism is particularly important for organisms such as microbial pathogens that utilise genome plasticity as a strategy for adaptation. For the first time, we analyse mechanisms promoting genome stability at the *rDNA* locus and subtelomeric regions in the most common human fungal pathogen: *Candida albicans*. In this organism, the histone deacetylase Sir2, the master regulator of heterochromatin, has acquired novel functions in regulating genome stability. Contrary to any other systems analysed, *C. albicans* Sir2 is largely dispensable for repressing recombination at the *rDNA* locus. We demonstrate that recombination at subtelomeric regions is controlled by a novel DNA element, the TLO Recombination Element, TRE, and by Sir2. While the TRE element promotes high levels of recombination, Sir2 represses this recombination rate. Finally, we demonstrate that, in *C. albicans*, mechanisms regulating genome stability are plastic as different environmental stress conditions lead to general genome instability and mask the Sir2-mediated recombination control at subtelomeres. Our data highlight how mechanisms regulating genome stability are rewired in *C. albicans*.

### INTRODUCTION

Repetitive regions clustered at the *rDNA* locus and subtelomeric regions are often the most polymorphic and variable regions of eukaryotes genomes (1–4). Repetitive DNA sequences often undergo homologous recombination which

can instigate genomic instability. At these locations, moderate genetic variability is beneficial because it generates the genetic diversity driving evolution and allows adaptation to different environmental niches. However, excessive genome instability is deleterious and an optimum balance between genome integrity and instability is essential for ensuring fitness while permitting adaptation. This is particularly important for microbial pathogens that utilise genome plasticity as a strategy to rapidly and reversibly adapt to different environmental niches. Fungal pathogens are a leading cause of human mortality worldwide, especially in immunocompromised individuals (5). Among those, *Candida albicans*, the principal causal agent of mycotic death, is a highly successful pathogen due in great part to its genome plasticity (6). Natural isolates exhibit a broad spectrum of genetic and genomic variations including single nucleotide polymorphisms (SNPs), short and long range loss of heterozygosity (LOH) events and whole chromosome aneuploidy (7). Environmental stimuli, including exposure to the mammalian host, drug treatment and heat shock, alter the rate and type of chromosomal rearrangements that provide a selective growth advantage in specific environmental conditions (8,9). For example, under standard laboratory growth conditions, most chromosomal rearrangements are LOH events driven by break-induced replication (BIR) where a double stranded DNA break is repaired by invasion of the broken end into a homologous DNA sequence until it reaches the end of the chromosome (8). In contrast, treatment with fluconazole, the most used anti-fungal drug, or 39°C, mimicking host fever, triggers aneuploidy and long range LOH (6). Exposure to hydrogen peroxide (H<sub>2</sub>O<sub>2</sub>), mimicking reactive oxygen species by the host's immune cells, leads to high rates of short range LOH (8).

In many eukaryotes, genome stability at the repetitive *rDNA* locus is ensured by the assembly of a transcriptionally silent chromatin structure, heterochromatin, which suppresses unequal recombination events (1). Heterochromatin is characterised by a specific histone modification pattern controlled by histone modifiers (10,11). For example, heterochromatin in the budding yeast *Saccharomyces cerevisiae*

\*To whom correspondence should be addressed. Tel: +44 1227 824854; Fax: +44 1227 824034; Email: A.Buscaino@kent.ac.uk

© The Author(s) 2016. Published by Oxford University Press on behalf of Nucleic Acids Research.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

is marked by hypoacetylated nucleosomes. In this system, the histone deacetylase (HDAC) Sir2 deacetylates histone 4 on lysine 16 (H4K16) (10). In other systems, such as the fission yeast *Schizosaccharomyces pombe*, heterochromatin is marked by hypoacetylated nucleosomes that are methylated on lysine 9 of Histone H3 (H3K9me) (12). The histone methyltransferase SuVar3-9 specifically methylates H3K9 (13). The *S. cerevisiae* epigenome is devoid of H3K9 methylation as a SuVar3-9 orthologous is absent in this organism.

A role for heterochromatin in repression of *rDNA* recombination is well established in *S. cerevisiae* where Sir2-dependent hypoacetylated heterochromatin suppress unequal recombination events by repressing non coding transcription and ensuring high levels of cohesion (14–16). The *S. cerevisiae* Monopoliin complex, composed of the protein Csm1 and Lrs4, acts in parallel and independently of Sir2 to promote *rDNA* stability by aligning sister chromatids, ensuring silencing, and mediating perinuclear anchoring (17,18). Indeed, deletion of both *S. cerevisiae* Sir2 and Monopoliin components results in a dramatic increase of unequal *rDNA* recombination compared to single mutants (17).

Telomeric and subtelomeric regions are also assembled into transcriptionally silent heterochromatin. The role of heterochromatin in controlling genome stability at subtelomeres is still unclear. Indeed, although components of the Sir2-containing protein complex maintain subtelomeric DNA stability (18), mitotic recombination rate at subtelomeric regions is unaltered in cells deleted for Sir2 compared to wild-type (WT) cells (19).

Mechanisms ensuring genome stability at *C. albicans* repetitive elements are unknown. The *C. albicans* genome is composed of 8 diploid chromosomes (20). Similarly to *S. cerevisiae*, the *C. albicans rDNA* locus consists of a tandem array of a ~12 kb unit repeated 50–200 times on chromosome R. Each unit contains the two highly conserved 35S and the 5S rRNA genes that are separated by two Non-Transcribed Regions (NTS1 and NTS2) (20). The *C. albicans rDNA* locus is highly plastic as the number of *rDNA* units can vary among different *C. albicans* strains (7) with extra-chromosomal plasmids containing several *rDNA* units present in different isolates (21,22).

The 16 *C. albicans* telomeric regions are formed by a terminal element composed of 23 bp tandem repeats and subtelomeric regions containing transposons and subtelomeric genes (1). Among those, the telomere-associated *TLO* genes are a family of 14 closely related paralogues encoding a set of related Med2 subunits for the Mediator transcription regulation complex (23–25). All *TLO* genes are oriented similarly with transcription proceeding toward the centromeres. *TLO* genes can be subdivided in three clades ( $\alpha$ ,  $\beta$  and  $\gamma$ ) based on the presence of indels (25). There are 5 *TLO*  $\alpha$ , 1 *TLO*  $\beta$  and 7 *TLO*  $\gamma$  genes. These subtelomeric genes are organised in 3 regions: a 5' region, with homology to *MED2*, that is conserved across all the three clades, a middle region of variable length that contains gene-specific adenosine stretches and a 3' region that is conserved within a clade but not across clades (25). The presence of a *TLO* gene family is unique to *C. albicans* as most of the less pathogenic non-*albicans* species have only one *TLO* gene. *C. dubliniensis*, the closest relative to *C. albicans*, has two

*TLO* genes: an internally located *TLO1* gene and a subtelomeric *TLO2* gene (26). Due to the highly repetitive nature of subtelomeric regions, which precludes detailed analyses by next-generation sequencing, the genetic plasticity of *C. albicans TLO* genes is still poorly understood. However, targeted Sanger sequencing has shown that these regions exhibited hypervariation between clinical isolates (7) and movement of *TLO* genes via recombination occurs at detectable levels within lab-passaged cultures (27).

We have shown that, in *C. albicans*, the repetitive DNA sequences associated with the *rDNA* locus, telomeres and subtelomeric regions, are assembled into hypoacetylated and hypomethylated heterochromatin that lacks H3K9 methylation (28). At these loci, heterochromatin stochastically silences expression of endogenous transcripts as well as of inserted marker genes. The chromatin and transcriptional repressive state associated with heterochromatic regions is dependent on the HDAC Sir2 (28). It is unknown whether Sir2 promotes genome stability of these repetitive regions. In this study, we address this question.

Contrary to any other systems analysed, we find that Sir2 is largely dispensable for repressing recombination at the *rDNA* locus. At this location, the Monopoliin complex is the major regulator of recombination.

At subtelomeric regions, we have identified a previously undefined 300 bp DNA sequence, that we named *TLO* recombination element (TRE), as a site that can promote recombination of a subset of *TLO* genes. We find that under standard growth conditions, Sir2 represses recombination at the TRE sequence. Under stress conditions, such as treatment with the anti-fungal drug fluconazole or with H<sub>2</sub>O<sub>2</sub>, recombination rates are increased at different genomic loci. This increment in recombination rate is independent of Sir2 and masks the recombination control mediated by Sir2. Our data highlights how the role of the HDAC Sir2 in promoting genome stability has been rewired in *C. albicans*, the most important human fungal pathogen.

## MATERIALS AND METHODS

### Growth conditions

Yeast cells were cultured in rich medium (YPAD) containing extra adenine (0.1 mg/ml) and extra uridine (0.08 mg/ml), complete SC medium (Formedium™) or SC Drop-Out media (Formedium™). When indicated, media were supplemented with 5-fluorotic acid (5-FOA, Melford) at a concentration of 1 mg/ml, Nourseothricin (clonNAT, Melford) at a concentration of 100 µg/ml, Fluconazole (Sigma) at a concentration of 1 µg/ml or 0.5 mM H<sub>2</sub>O<sub>2</sub>. Cells were grown at 30 or 39°C as indicated.

### Yeast strain construction

Strains are listed in Supplementary Table S1. Integration and deletion of genes was performed as previously described (29) using long oligos-mediated PCR for gene deletion and tagging. Oligonucleotides and plasmids used for strain constructions are listed in Supplementary Table S2 and Supplementary Table S3, respectively. Transformation was performed by electroporation (Gene Pulser™, BioRad) using the protocol described in (30). The *URA3*

marker gene was used for silencing assay. *HIS1*, *ARG4* and *SAT1* marker genes were used to delete both copies of *SIR2* and *CSM1* genes. HA tag was used for *SIR2* tagging at the C-terminus. Correct integration events were checked by PCR using primers listed in Supplementary Table S2.

#### TRE plasmid construction

pTRE-URA3 was constructed using plasmid pGEMURA3 (29). The TRE (TLO Recombination Element) sequence located at 3' of *TLO $\alpha$ 10* was PCR amplified from *C. albicans* genomic DNA using oligos containing the restriction sites SacII (Supplementary Table S2). PCR purified TRE product and pGEMURA3 plasmid were digested with SacII (Promega), ligated and transformed in *DH5- $\alpha$*  *E. coli* cells. Positive transformants were confirmed by PCR and sequencing with primers listed in Supplementary Table S2.

#### Silencing assay in liquid media

Growth analyses were performed using a plate reader (SpectrostarNano, BMG labtech) in 96-well plate format at 30°C. For each silencing assay, 1:100 dilution of an overnight culture was inoculated in a final volume of 95  $\mu$ l of SC or SC-URA media to reach a concentration of 60 cells/ $\mu$ l. Growth was assessed by measuring  $A_{600}$ , using the following conditions: OD<sub>600 nm</sub>, 616 cycle time, three flashes per well, 700 rpm shaking frequency, orbital shaking mode, 545 s additional shaking time after each cycle 0.5 s post delay, for 32 or 44 h. Graphs represent data from three biological replicates. Error bars: standard deviations of three biological replicates. Data was processed using SpectrostarNano MARS software and Microsoft Excel.

#### Fluctuation analysis

Strains were first streaked on -Uri media to ensure the selection of cells carrying the *URA3*<sup>+</sup> marker gene. Fifteen parallel liquid cultures were pre-grown overnight from independent single colonies. Each culture was diluted in YPAD at a concentration of 100 cells/ $\mu$ l and grown for nine generations (18 h). When indicated, media were supplemented with fluconazole (Sigma) at a concentration of 1  $\mu$ g/ml or 0.5 mM H<sub>2</sub>O<sub>2</sub>. During thermic stress, cultures were grown at 39°C. LOH analyses were performed in three biological replicates. Cells were plated on SC plates containing 1 mg/ml 5-FOA (5-fluorotic acid, Sigma) and on non-selective SC plates and grown at 30°C. Colonies were counted after 2 days of growth and data were analysed using FALCOR (Fluctuation Analysis Calculator) software (31) based on the Lea-Coulson analysis of the median (32). Statistical differences between samples were calculated using Kruskal-Wallis test. Statistical analysis and violin plots were generated using R (<http://www.r-project.org/>). To quantify the number of colonies able to grow on FOA due to *URA3*<sup>+</sup> silencing, FOA resistant colonies were replica-plated, after fluctuation analysis, on complete SC plates for recovery. Single colonies were then replica-plated on SC plates supplemented with 1 mg/ml 5-FOA

and -Uri SC Drop-Out plates. After 24 hour growth single colonies were counted and analysed using Microsoft Excel. To analyse the presence of the SAT1 marker gene after fluctuation analysis, FOA resistant single colonies were replica-plated on complete SC plates for recovery. Single colonies were then replica-plated on YPAD plates containing extra adenine (0.1mg/ml), extra uridine (0.08mg/ml) and Nourseothricin (clonNAT, Melford) at a concentration of 100  $\mu$ g/ml. After 24 hour colonies were counted and analysed using Microsoft Excel.

#### Marker gene loss assay

Strains were first streaked on -Uri media to select for cells carrying the *URA3*<sup>+</sup> marker gene. 15 parallel liquid cultures were pre-grown overnight from independent single colonies. Each culture was diluted in YPAD at a concentration of 100 cells/ $\mu$ l and grown for 9 generations (18 hours). Cells were plated on SC plates containing 1 mg/ml 5-FOA (5-fluorotic acid, Sigma) and on non-selective SC plates and grown at 30°C. To distinguish between silencing and loss of the *URA3* marker gene, FOA resistant single colonies were streaked onto complete SC plates for recovery and then streaked onto -Uri SC Drop-Out plates. Colonies not able to grow on -Uri plates but FOA resistant were counted. Data were analysed using Microsoft Excel. Statistical differences between samples were tested using unpaired t-test using R (<http://www.r-project.org/>).

#### RNA extraction and cDNA synthesis

RNA was extracted from log<sub>2</sub> exponential cultures (OD<sub>600nm</sub> = 1.4) using a yeast RNA extraction kit (E.Z.N.A.<sup>®</sup> Isolation Kit RNA Yeast, Omega Bio-Tek) following the manufacturer's instructions. RNA quality was checked by electrophoresis under denaturing conditions in 1% agarose, 1X HEPES, 6% Formaldehyde (Sigma). RNA concentration was measured using a NanoDrop ND-1000 Spectrophotometer. cDNA synthesis was performed using iScript<sup>™</sup> Reverse Transcription Supermix for RT-qPCR (Bio-Rad) following manufacturer's instructions and a Bio-Rad CFXConnect<sup>™</sup> Real-Time System.

#### RT-qPCR reactions

Primers used are listed in Supplementary Table S2. RT-qPCR was performed in the presence of SYBR Green (Bio-Rad) on a Bio-Rad CFXConnect<sup>™</sup> Real-Time System. Data was analysed with Bio-Rad CFX Manager 3.1 software and Microsoft Excel. Enrichment was calculated over actin. Histograms represent data from three biological replicates. Error bars: standard deviation of three biological replicates generated from 3 independent cultures of the same strain.

#### Protein extraction and Western blotting

Yeast extracts were prepared as described (33) using 1  $\times$  10<sup>8</sup> cells from overnight cultures grown to a final OD<sub>600</sub> of 1.5–2. Protein extraction was performed in the presence of 2% SDS (Sigma) and 4 M acetic acid (Fisher) at

90°C. Proteins were separated in 2% SDS (Sigma), 40% acrylamide/bis (Biorad, 161-0148) gels and transfer into PVDF membrane (Biorad) by semi-dry transfer (Biorad, Trans Blot SD, semi-dry transfer cell). Western-blot antibody detection was used using antibodies from Roche Diagnostics Mannheim Germany (Anti-HA, mouse monoclonal primary antibody (12CA5 Roche, 5 mg/ml) at a dilution of 1:1000, and anti-mouse IgG-peroxidase (A4416 Sigma, 0.63 mg/ml) at a dilution of 1:5000, and Clarity™ ECL substrate (Bio-Rad).

#### Bioinformatics analysis

*Candida albicans* and *C. dubliensis* TLO genes and flanking sequences were downloaded from the *Candida* Genome Database (34). DNA alignment was performed using Muscle with default setting and visualised using Jalview (35). Motif finder analyses was performed using MEME SUITE using the MEME discovery programme in discriminative mode (36).

#### SNP-RFLP analysis

PCR Primers and Restriction Enzymes were chosen according to (37). PCRs were performed in a final volume of 15 µl using Taq DNA polymerase (VWR, 733–1364) using manufacturer's instructions. DNA was extracted from single colonies following NaOH heat extraction. PCR conditions were performed as follow: initial denaturation at 94°C for 7 min, 30 cycles each of denaturation at 94°C for 45 s, annealing at 55°C for 1 min, and extension at 72°C for 1 min, and a final extension at 72°C for 7 min. Each PCR product was digested overnight with the relevant restriction enzyme, AseI for SNP 1, TaqI for SNP 85 and DdeI for SNP135 (37). Enzymatic reactions were performed in a total volume of 15 µl with 1 µl RE, 10× restriction buffer, distilled water, and 5 µl of SNP amplified PCR. 15 µl of the digested PCR product was run on a 3% agarose gel (Melford) along with an undigested control PCR sample. Gels were stained with ethidium bromide and photographed. Genotypes were assigned based on banding patterns for each SNP marker as described in (37).

## RESULTS

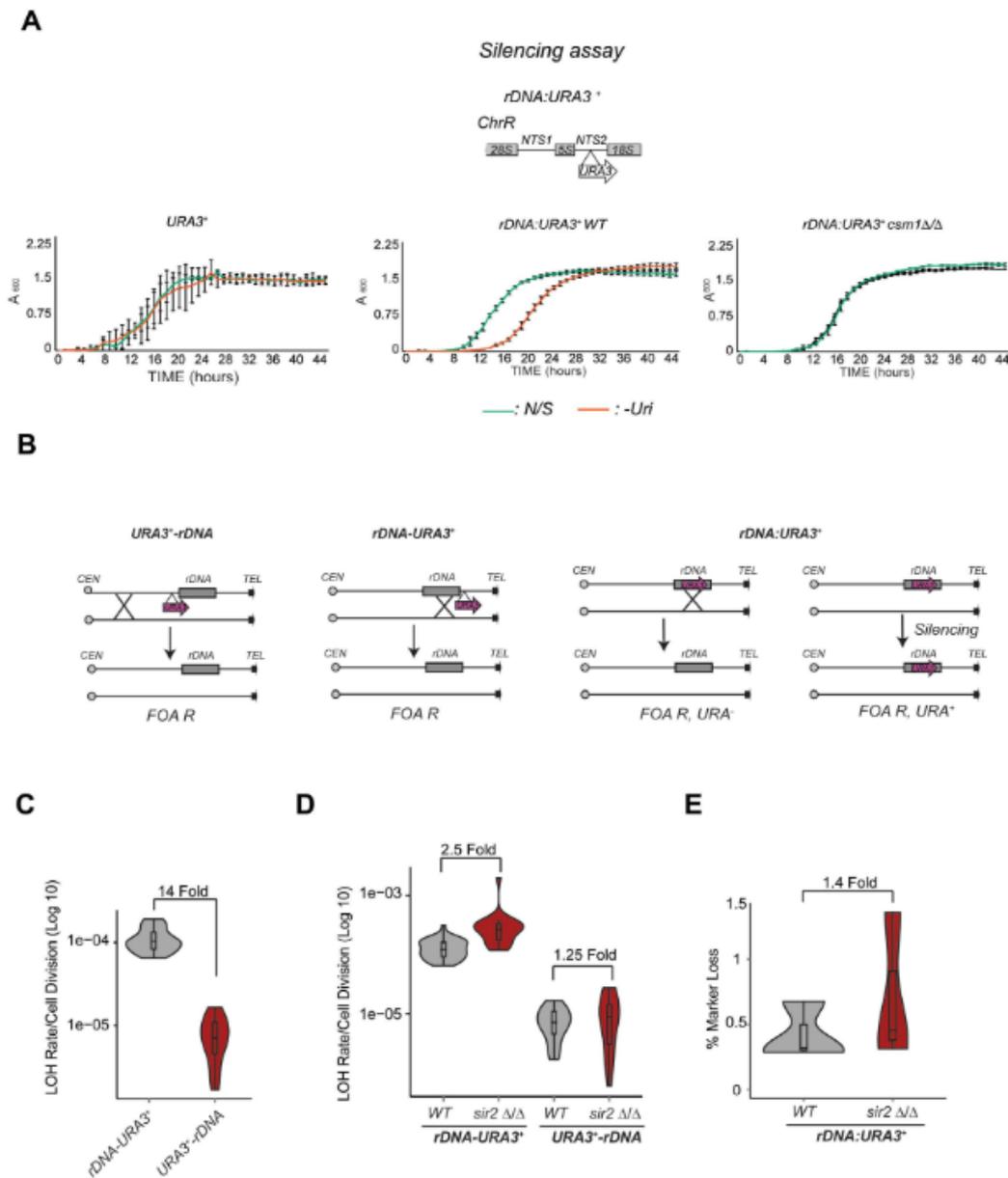
### The Monopolin complex, but not Sir2, promotes *rDNA* stability

In *S. cerevisiae*, it is well established that Sir2 acts in parallel to the Monopolin complex to suppress *rDNA* mitotic recombination (38). We have shown that the *C. albicans* NTS region of the *rDNA* locus is assembled into heterochromatin able to silence an embedded *URA3*<sup>+</sup> marker gene in a silencing reporter strain (*rDNA:URA3*<sup>+</sup>). The HDAC Sir2 maintains this silent state as deletion of Sir2 results in elevated expression of the *URA3*<sup>+</sup> marker gene (28). To assess whether the Monopolin complex contributes to the assembly of silent heterochromatin at the *rDNA* locus, we deleted the *CSMI* gene encoding for the Monopolin component Csm1 in the *rDNA:URA3*<sup>+</sup> reporter strain. Although cells lacking Csm1 grow slower on non-selective (N/S) media

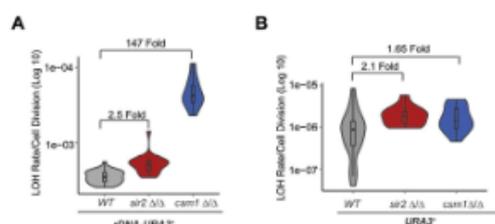
compared to WT cells, it is clear that deletion of *CSMI* results in alleviation of *rDNA* silencing (Figure 1A). Therefore, similarly to *S. cerevisiae*, the *C. albicans* Monopolin complex maintains the transcriptionally silenced state associated with the NTS region of the *rDNA* locus.

To assess whether *C. albicans* Sir2 and/or the Monopolin complex control mitotic recombination at the *rDNA* locus, we engineered two *C. albicans* strains where a *URA3*<sup>+</sup> heterozygous marker gene was integrated at centromere-proximal (*URA3*<sup>+</sup>-*rDNA*) and telomere-proximal side (*rDNA-URA3*<sup>+</sup>) of the *rDNA* locus (Figure 1B). Comparison of the *URA3*<sup>+</sup> loss of heterozygosity between the two strains gives a measure of recombination at the *rDNA* locus because loss of the *URA3*<sup>+</sup> in the *URA3*<sup>+</sup>-*rDNA* strain detects whole chromosome aneuploidy and/or recombination event upstream of the *rDNA* while loss of the *URA3*<sup>+</sup> marker in the *rDNA-URA3*<sup>+</sup> strain additionally detects recombination events at the *rDNA* locus (Figure 1B). We performed fluctuation analyses where 15 independent cultures were grown for 10 generations before plating on plates containing the *URA3* counter-selective drug FOA. Loss of the *URA3*<sup>+</sup> marker gene was determined by scoring the number of colonies that were able to grow on FOA-containing media compared to N/S media (Figure 1B and C). Importantly, growth of FOA resistant colonies is not a consequence of *URA3*<sup>+</sup> silencing because all the FOA resistant colonies have irreversibly lost the ability to grow on -Uri media (Supplementary Figure S1A and B). Fluctuation analyses in WT strains reveal that the *rDNA* locus is a hotspot of mitotic recombination as mitotic recombination rate of the *rDNA-URA3*<sup>+</sup> gene is 14 fold higher than the mitotic recombination rate of the *URA3*<sup>+</sup>-*rDNA* gene (Figure 1C). To assess whether *C. albicans* Sir2 controls recombination at the *rDNA* locus as it occurs in *S. cerevisiae*, we deleted both copies of *SIR2* gene at both *rDNA-URA3*<sup>+</sup> and *URA3*<sup>+</sup>-*rDNA* strains. Contrary to *S. cerevisiae*, we observed only a slight increase in recombination rate in *sir2* Δ/Δ compared to WT cells (Figure 1D). To assess whether Sir2 represses recombination events within the *rDNA* cluster, we measured, by marker gene loss assay, the number of FOA resistant colonies emerging from a strain with an integrated *URA3*<sup>+</sup> marker gene at the *rDNA* locus (Figure 1B and E). FOA resistant colonies could arise from loss of the *URA3*<sup>+</sup> marker gene following a recombination event or from silencing of the *URA3*<sup>+</sup> marker gene. The first event is irreversible and leads to FOA resistant colonies that are unable to grow on media lacking Uridine (-Uri). In contrast, silencing is reversible and leads to FOA resistant colonies able to grow on -Uri plates (Figure 1B). Therefore, to assess the role of Sir2 in controlling *rDNA* intra recombination, we counted the number of FOA resistant colonies that have irreversibly lost the *URA3*<sup>+</sup> marker gene. As shown in Figure 1E the number of FOA resistant colonies that have lost the *URA3*<sup>+</sup> marker gene following a recombination event is very similar in *sir2* Δ/Δ versus WT cells (1.4-fold difference) (Figure 1E). Therefore, *C. albicans* Sir2 is not a major contributor of *rDNA* stability.

To assess whether the Monopolin complex controls *rDNA* mitotic recombination, we deleted both copies of the *CSMI* gene in the *rDNA-URA3*<sup>+</sup> strain and performed fluctuation analyses. Similarly to *S. cerevisiae*, *C. albicans*



**Figure 1.** Sir2 does not control *rDNA* stability. (A) Upper panel: schematic of *rDNA:URA3<sup>+</sup>* reporter strain. Bottom panel: silencing assay of a *URA3<sup>+</sup>*, *rDNA:URA3<sup>+</sup>* reporter strains in WT and *csm1* Δ/Δ isolates in non-selective (N/S) and -uridine (-Uri) media. Error bars: standard deviation (SD) of three biological replicates. (B) Schematic *URA3<sup>+</sup>-rDNA*, *rDNA-URA3<sup>+</sup>* and *rDNA:URA3<sup>+</sup>* reporter strains and the mechanism leading to FOA resistance. (C) *URA3<sup>+</sup>-rDNA* and *rDNA-URA3<sup>+</sup>* fluctuation analyses. *p*-value calculated with Kruskal–Wallis statistical test is:  $2.877 \times 10^{-09}$ . (D) Fluctuation analyses in *rDNA-URA3* and *URA3<sup>+</sup>-rDNA* in WT, *sir2*Δ/Δ isolates. *p*-values = 1.302e-07 and 0.665, respectively. (E) Percentage (%) marker *URA3<sup>+</sup>* loss with in the *rDNA:URA3* reporter strain in WT and *sir2*Δ/Δ isolates. Violin plots represent all colonies that lost the *URA3<sup>+</sup>* after fluctuation analysis detected by lack of growth on -Uri media.



**Figure 2.** The Monopolin complex, but not Sir2, controls *rDNA* stability. (A) *rDNA-URA3* Fluctuation analyses in WT, *sir2* $\Delta/\Delta$  and *csm1* $\Delta/\Delta$  cells. *p*-values for WT versus *sir2* $\Delta/\Delta$  and *csm1* $\Delta/\Delta$  is respectively:  $1.30 \times 10^{-07}$  and  $2.727 \times 10^{-10}$ . (B) Fluctuation analyses of a *URA3*<sup>+</sup> heterozygous strain containing a heterozygous *URA3* at its endogenous locus in WT, *sir2* $\Delta/\Delta$  and *csm1* $\Delta/\Delta$  cells. *p*-values = 0.00355 and 0.02623, respectively.

Csm1 represses mitotic recombination rate at the *rDNA* locus as mitotic recombination was 147-fold higher in *csm1* $\Delta/\Delta$  strain compared to WT cells (Figure 2A). Importantly, Csm1 specifically represses recombination at the *rDNA* locus as LOH rate of a heterozygous *URA3*<sup>+</sup> gene at its endogenous locus (Chr 3) is similar in WT, *csm1* $\Delta/\Delta$  and *sir2* $\Delta/\Delta$  cells (2.1 Fold and 1.65 Fold, respectively) (Figure 2B). Therefore, we concluded that *C. albicans* Sir2 does not contribute to *rDNA* stability. Repression of mitotic recombination at this locus is solely dependent on the Monopolin complex that also contributes to transcriptional silencing.

#### Loss of heterozygosity is elevated at subtelomeric regions

Subtelomeric regions are often the most variable regions of the genome (39–41). Due the complexity of these regions, analysis of subtelomere plasticity is not amenable to genome-wide studies. To analyse *TLO* plasticity in a large population of cells, we set up a system where a heterozygous *URA3*<sup>+</sup> marker gene was integrated at the 3' end (centromere-proximal) of the *TLO*  $\alpha 10$ , *TLO*  $\alpha 12$  and *TLO*  $\gamma 6$  genes on chromosomes (Chr) 4, 5 and 7, respectively (Figure 3A). Silencing assay demonstrated that at this location the *URA3*<sup>+</sup> marker gene is not silenced as *TLO*  $\alpha 10$ -*URA3*<sup>+</sup> and *TLO*  $\alpha 12$ -*URA3*<sup>+</sup> strains grow as well as a *URA3*<sup>+</sup> strain in N/S and -Uri media (Supplementary Figure S2A and B). Fluctuation analyses revealed that FOA resistant colonies appeared more frequently (5–22-fold) when the *URA3*<sup>+</sup> marker gene was inserted centromere-proximal to the three different *TLO* genes than when *URA3*<sup>+</sup> is at its endogenous locus (Figure 3C). This is not the result of reversible silencing as the FOA resistant colonies have irreversibly lost the ability to grow on -Uri media (Supplementary Figure S2C and D). FOA resistant colonies could arise from a point mutation in the *URA3*<sup>+</sup> gene, from whole chromosome loss (with or without regain of the remaining chromosome) or from a telomere-proximal mitotic recombination event (followed by co-segregation of the homologous copies or by BIR) (Figure 3B). The possibility that a point mutation is the major contributor to the appearance of FOA resistance colonies was excluded by PCR analyses with primers specific for the *URA3*<sup>+</sup> gene, as none of the resistant colonies analysed (*n* = 65) retained the

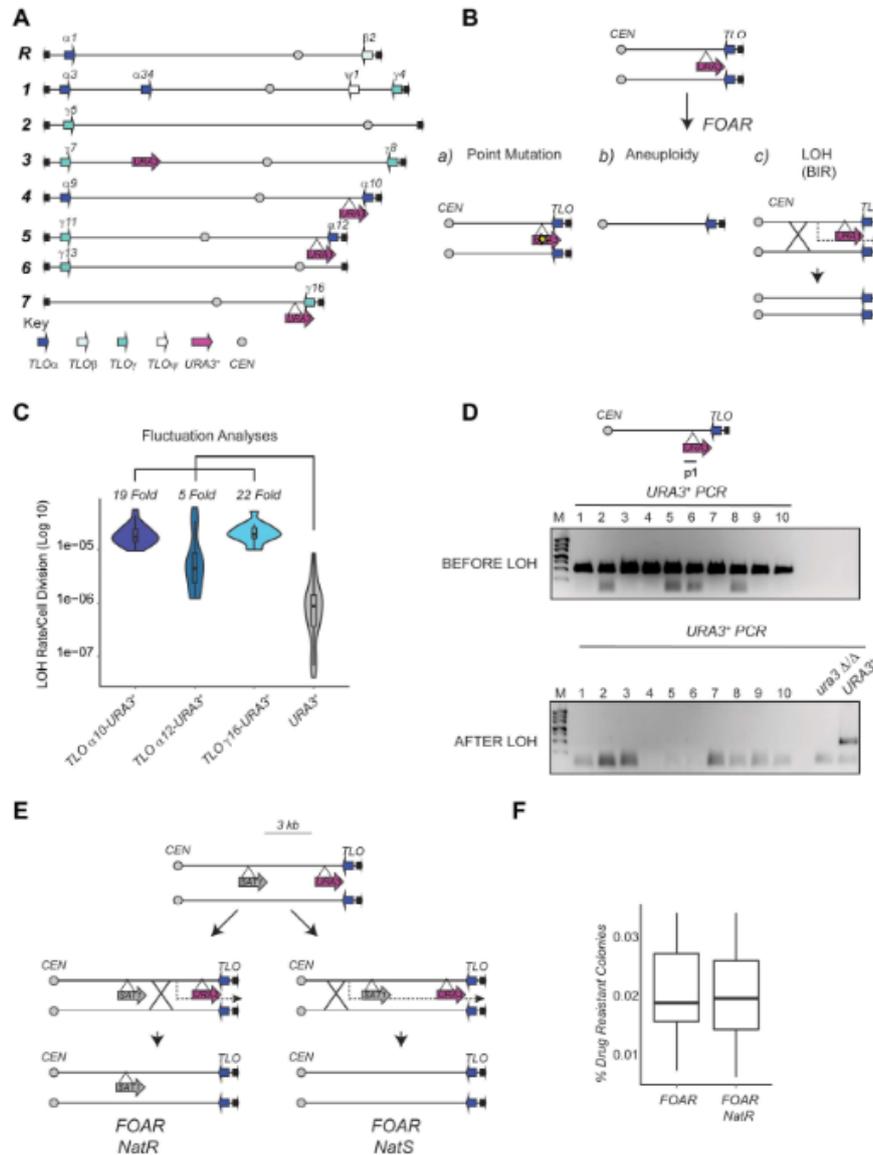
*URA3*<sup>+</sup> gene (Figure 3D and Supplementary Figure S3). These results are in agreement with previous observations establishing that in *C. albicans* mutation rate is much lower (~1000-fold) than the LOH rate (8). Furthermore, these data establish that, similarly to other systems, *C. albicans* subtelomeric genes are highly unstable and associated with high recombination rates. To distinguish between recombination events that were somewhat centromere-proximal or to whole chromosome events, we inserted a second heterozygous marker gene (*SAT1*) 3 kb upstream of the *URA3*<sup>+</sup> marker gene in the same homologous chromosome creating the reporter strain *TLO*  $\alpha 10$ -*URA3*<sup>+</sup>-*SAT1* (Figure 3E). The *SAT1* marker gene confers resistance to the antibiotic nourseothricin (NAT). If whole chromosome aneuploidy was the cause of the *URA3*<sup>+</sup> marker loss, then FOA resistant colonies should also have lost the *SAT1* marker gene and therefore be sensitive to the antibiotic NAT. On the other hand, if the *URA3*<sup>+</sup> marker was lost as a consequence of a mitotic recombination event near the *TLO* gene, FOA resistant colonies should retain the *SAT1* marker gene and therefore be able to grow on a medium containing the antibiotic NAT. As shown in Figure 3F, most of the FOA resistant colonies retained the *SAT1* marker gene and therefore were NAT resistant and not NAT sensitive (Figure 3F). These results indicate that the majority of the *URA3*<sup>+</sup> LOH events are due to cross-overs within 3 kb of the *URA3*<sup>+</sup> gene. Thus, *TLO* instability is largely due to mitotic recombination events occurring very close to the *TLO* genes. An important question is why this region is particularly prone to recombination.

#### Sir2 suppresses recombination of *TLO* $\alpha 10$ and *TLO* $\alpha 12$ genes but not *TLO* $\gamma 16$

We have shown that telomeric regions are assembled into Sir2-dependent heterochromatin (28). To assess whether Sir2 represses mitotic recombination at *TLO* genes, we deleted both copies of the *SIR2* gene in the *TLO* recombination-tester strains on Chr 4, 5 and 7 (*TLO*  $\alpha 10$ -*URA3*<sup>+</sup>, *TLO*  $\alpha 12$ -*URA3*<sup>+</sup>, *TLO*  $\gamma 16$ -*URA3*<sup>+</sup>) (Figure 4A) and performed fluctuation analyses. In WT cells recombination rate associated with these *TLO* genes is similar (Figure 4B). However, LOH rate for *TLO*  $\alpha 10$ -*URA3*<sup>+</sup> and *TLO*  $\alpha 12$ -*URA3*<sup>+</sup> increased in *sir2* $\Delta/\Delta$  compared to WT cells (23 and 90 fold respectively) (Figure 4B). In contrast, *TLO*  $\gamma 16$ -*URA3*<sup>+</sup> recombination rate was increased only marginally, by 1.8-fold, in *sir2* $\Delta/\Delta$  cells (Figure 4B). We concluded that Sir2 represses mitotic recombination at *TLO*  $\alpha 10$  and *TLO*  $\alpha 12$  but not *TLO*  $\gamma 16$ .

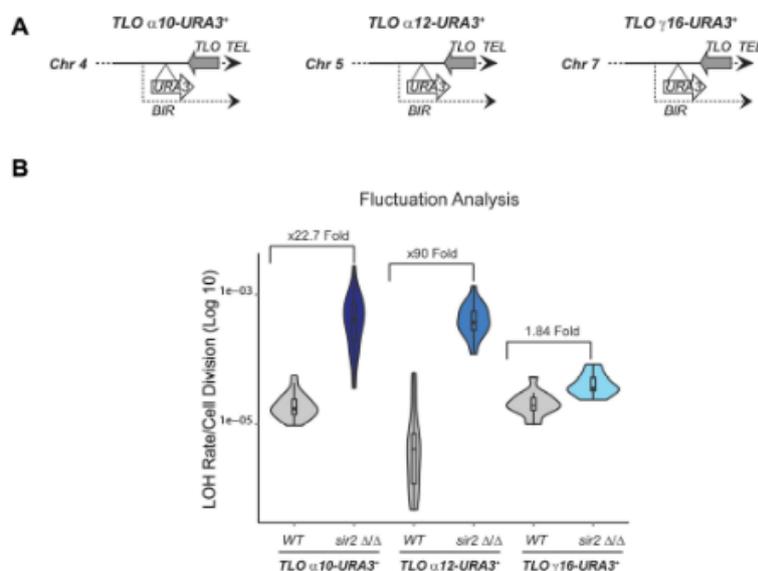
#### Sir2 represses mitotic recombination at *TLO* genes via a 300 bp *TLO* recombination element

Recombination occurs in a 3 kb telomeric-distal region of the *TLO*  $\alpha 10$  and *TLO*  $\alpha 12$  genes and it is dependent on the HDAC Sir2. In contrast, Sir2 does not repress recombination at *TLO*  $\gamma 16$  gene. We hypothesised that a *cis*-acting DNA element promotes recombination of *TLO*  $\alpha 10$  and *TLO*  $\alpha 12$ , but not of the *TLO*  $\gamma 16$ , and that Sir2 acts on this element to repress recombination. To identify this putative



**Figure 3.** Loss of Heterozygosity is elevated at subtelomeric regions. (A) Schematic of *Candida albicans* chromosome organisation. The locations of *TLO*  $\alpha$ ,  $\beta$  and  $\gamma$  genes are indicated with blue arrows, the locations of the integrated *URA3<sup>+</sup>* marker genes are indicated with magenta arrows. (B) Schematic of possible mechanisms leading to FOA resistance. Point mutation: the *URA3<sup>+</sup>* marker is non-functional due to a mutation in the gene sequence. Aneuploidy: the *URA3<sup>+</sup>* marker gene is lost due to a whole-chromosome loss event. Loss of Heterozygosity (LOH): a break-induced recombination (BIR) event leads to loss of the *URA3<sup>+</sup>* marker gene. (C) Fluctuation analysis for LOH Rates in *TLO $\alpha$ 10-URA3<sup>+</sup>*, *TLO $\alpha$ 12-URA3<sup>+</sup>* and *TLO $\gamma$ 16-URA3<sup>+</sup>* compared to the *URA3/ura3 $\Delta$*  endogenous heterozygous strain (*URA3<sup>+</sup>*). *P*-values, calculated with the Kruskal-Wallis statistical test, are  $2.877 \times 10^{-09}$  for *TLO $\alpha$ 10-URA3<sup>+</sup>*, 0.003892 for *TLO $\alpha$ 12-URA3<sup>+</sup>* and  $2.035 \times 10^{-07}$  for *TLO $\gamma$ 16-URA3<sup>+</sup>*. (D) *URA3<sup>+</sup>* PCR analyses with primers specific for the *URA3<sup>+</sup>* marker genes was performed with 10 colonies before the Fluctuation analyses (BEFORE LOH) and with 10 FOA resistant colonies (AFTER LOH). A *URA3<sup>+</sup>* and a *ura3 $\Delta/\Delta$*  strains was included as a positive and negative control. (E) Schematics of the *TLO $\alpha$ 10-URA3<sup>+</sup>-SAT1* strain. A breaking point between the *SAT1* gene and *URA3<sup>+</sup>* marker gene would produce FoA resistant (FOAR) and NAT resistant (NATR) colonies. A breaking point upstream of both marker genes would produce FOAR colonies that are sensitive to NAT. (F) Percentage (%) of drug resistant colonies after fluctuation analyses.

Downloaded from <http://nar.oxfordjournals.org/> by guest on July 5, 2016



**Figure 4.** *Sir2* suppresses recombination of *TLOα10* and *TLOα12* genes but not *TLOγ16*. (A) Schematic *TLOα10-URA3+*, *TLOα12-URA3+* and *TLOγ16-URA3+* reporter strains. (B) for *TLOα10-URA3+*, *TLOα12-URA3+* and *TLOγ16-URA3+* fluctuation analyses in WT and *sir2Δ/Δ* cells. p-value for each *TLO* gene in WT versus *sir2Δ/Δ* is respectively:  $4.846 \times 10^{-10}$ ,  $2.035 \times 10^{-17}$  and 0.0001475.

control region, we aligned all *TLO* genes and their downstream sequences. This alignment reveals that a 300 bp region downstream of the *TLO* stop codon, that we named TRE, is conserved among all chromosome ends except for the subtelomeric region containing *TLOγ16* gene (Chr 7R) and *TLOβ2* (Chr RR) genes (Supplementary Figure S4). Blast analysis reveals that the TRE element is not present in any other locations in the *C. albicans* genome and it is poorly conserved at the 3' region of the subtelomeric *C. dubliniensis TLO2* gene (Supplementary Figure S5A). This finding raises the possibility that the TRE element is important for the *TLO* gene expansion observed in *C. albicans*. Motif finder analysis identifies a 50 nt long motif present in all *TLO* genes except *TLOβ2* and *TLOγ16* (Supplementary Figure S5B). This motif corresponds to the previously identified BTS (Bermuda Triangle Sequence), a site of recombination detected during *C. albicans* evolution experiments (27).

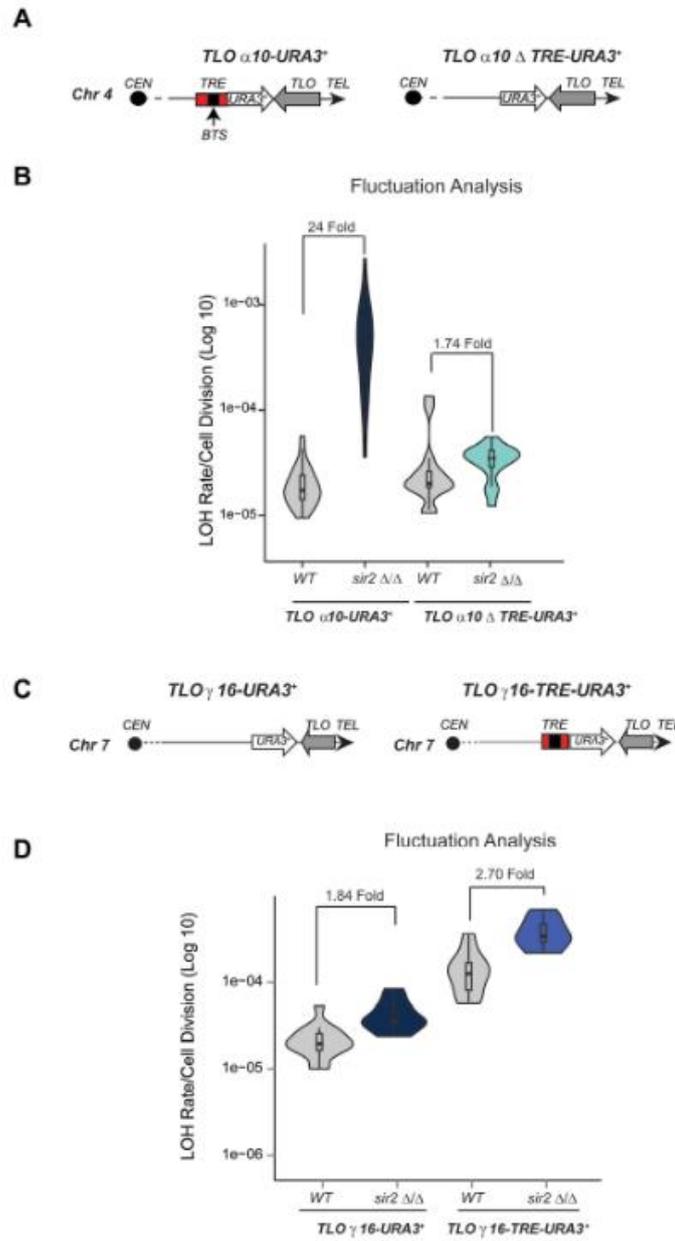
To determine whether the TRE element is necessary to promote recombination in the absence of *Sir2*, we replaced it with a *URA3+* marker gene creating the *TLOα10 ΔTRE-URA3+* reporter strain lacking the TRE element and containing a heterozygous *URA3+* gene in its place (Figure 5A). Fluctuation analyses reveal that recombination rate at *TLOα10-ΔTRE-URA3+* did not significantly increase in *sir2Δ/Δ* compared to WT cells (Figure 5B). Therefore, we concluded that TRE element is necessary to promote recombination and that *Sir2* acts via TRE to repress recombination of *TLO* genes that have an adjacent TRE.

To test whether the TRE element is sufficient to induce recombination in the absence of *Sir2*, we integrated the TRE

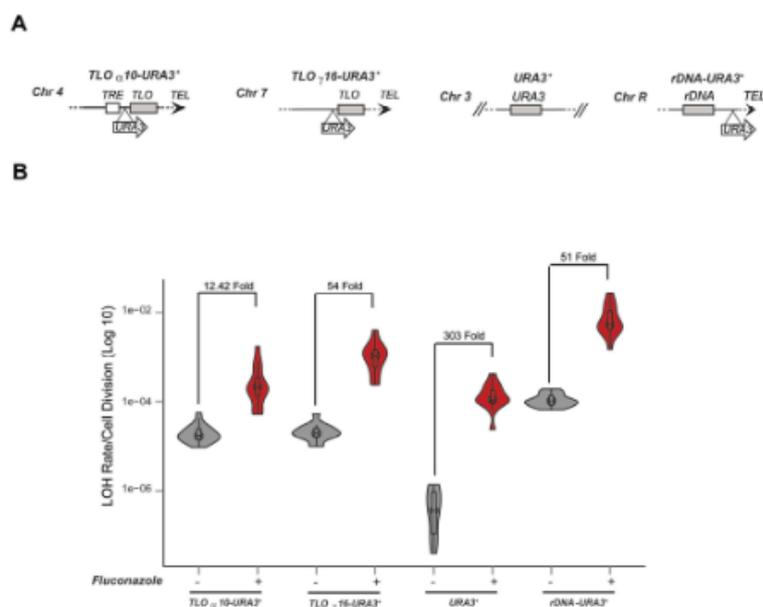
element together with a *URA3+* marker gene downstream of the *TLOγ16* gene, which normally lacks the TRE and does not show *Sir2*-dependent recombination repression (Figure 5C). Fluctuation analyses revealed that the ectopic TRE partially increases LOH rate (~6-fold) in WT cells and that deletion of *SIR2* resulted in an additional increase of ~3-fold at the *TLOγ16* locus (Figure 5D). Taken together our data demonstrate that the TRE element promotes high levels of recombination and that *Sir2* represses recombination of *TLO* genes containing the TRE element.

#### Stress conditions trigger repeats-associated instability independently of *Sir2*

A range of stress conditions (high temperature, fluconazole treatment and  $H_2O_2$  treatment) has been reported to induce *C. albicans* genome instability (8). To assess the effect of stress conditions on the stability of *C. albicans* subtelomeric regions, we asked if treatment with fluconazole, the most common and widely used antifungal drug, affects LOH and/or aneuploidy rate at *TLOα10* (Chr4) and *TLOγ16* (Chr7) (Figure 6A). As a control, we measured LOH rate at the *URA3+* endogenous locus (Chr3) and the *rDNA* locus (ChrR) (Figure 6A). Consistent with previous results (8), fluctuation analyses showed that fluconazole treatment results in a dramatic increase of LOH rates at all loci tested indicating that fluconazole leads to general genome instability including subtelomeric regions (Figure 6B). SNP-RFLP analysis of FOA Resistant colonies reveals that fluconazole treatment does not increase whole chromosome aneuploidy of two different chromosomes (Supplementary Fig-



**Figure 5.** *Sir2* represses mitotic recombination at *TLO* genes via a 300 bp *TLO* Recombination Element. (A) Schematic of *TLO*α10-*URA3*<sup>+</sup> and *TLO*α10Δ*TRE-URA3*<sup>+</sup> reporter strains. The TRE region and the BTS regions are highlighted. (B) *TLO*α10-*URA3*<sup>+</sup>, *TLO*α10 Δ *TRE-URA3*<sup>+</sup> fluctuation analysis in WT and *sir2*Δ/Δ cells. (C) Schematic of *TLO*γ16-*URA3*<sup>+</sup>, *TLO*γ16-*TRE-URA3*<sup>+</sup>. (D) *TLO*γ16-*URA3*<sup>+</sup>, *TLO*γ16-*TRE-URA3*<sup>+</sup> fluctuation analyses in WT and *sir2*Δ/Δ cells. *p*-value = 3.067 × 10<sup>-10</sup>.



**Figure 6.** Stress conditions increase LOH associated with all genomic loci tested. (A) Schematic of *TLO<sub>α10</sub>-URA3<sup>+</sup>*, *TLO<sub>γ16</sub>-URA3<sup>+</sup>*, *URA3<sup>+</sup>* and *rDNA-URA3<sup>+</sup>*. (B) *TLO<sub>α10</sub>-URA3<sup>+</sup>*, *TLO<sub>γ16</sub>-URA3<sup>+</sup>*, *URA3<sup>+</sup>* and *rDNA-URA3<sup>+</sup>* fluctuation analysis without (–) and with (+) fluconazole treatment. The calculated *p*-values in the presence and absence of fluconazole for each strain are respectively  $2.035 \times 10^{-07}$ ,  $1.125 \times 10^{-05}$ ,  $3.697 \times 10^{-07}$  and  $2.035 \times 10^{-07}$ .

ure S6). To assess whether, fluconazole increases long LOH tracts, we performed fluctuation analyses with and without fluconazole in the reporter strain *TLO<sub>α10</sub>-URA3<sup>+</sup>-SAT1* containing the SAT1 marker gene 3 Kb upstream of the *URA3<sup>+</sup>* marker gene in the same homologous chromosome (Supplementary Figure S7A). Fluconazole treatment does not increase long LOH tracts as all the FOA resistant colonies were also resistant to the antibiotic NAT and therefore loss of the *URA3<sup>+</sup>* marker gene is due to recombination in proximity of telomeres (Supplementary Figure S7B).

To assess whether fluconazole treatment impacts on the Sir2-mediated control of recombination, we measured LOH rate of the TRE-containing subtelomere gene *TLO<sub>α10</sub>* in WT and *sir2*  $\Delta/\Delta$  cells. While in untreated cells, LOH rate dramatically increases in *sir2*  $\Delta/\Delta$  compared to WT cells (23-fold) (Figure 7B), in the presence of fluconazole, recombination rate only slightly increases in *sir2*  $\Delta/\Delta$  compared to WT cells (0.64-fold, Figure 7B). Although less dramatic, H<sub>2</sub>O<sub>2</sub> treatment leads to similar results. H<sub>2</sub>O<sub>2</sub> treatment leads to an increase in LOH rate without affecting whole chromosome aneuploidy or long LOH tracts (Figure 7C and Supplementary Figure S8). In addition, following treatment with H<sub>2</sub>O<sub>2</sub> the increase in LOH rate at *TLO<sub>α10</sub>-URA3<sup>+</sup>* is much smaller in *sir2*  $\Delta/\Delta$  cells compared to WT cells (from 23-fold to 4.6-fold) (Figure 7C). Not all stress conditions have the same effect as growing *C. albicans* cells at high temperature (39°C), a temperature mimicking fever in the host, does not abolish the Sir2-mediated recombi-

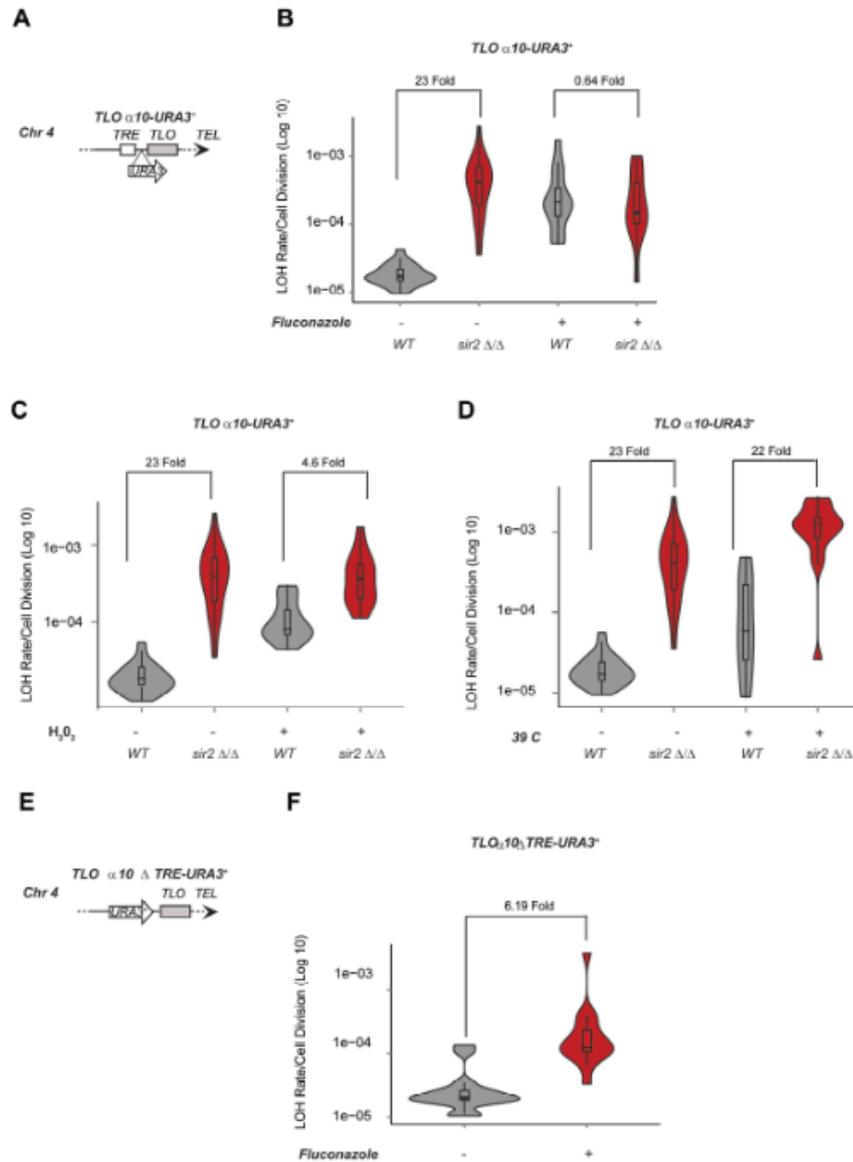
nation control or affects whole chromosome aneuploidy or long LOH tracts (Figure 7D and Supplementary Figure S8). Therefore, the high genome instability instigated by stress conditions (fluconazole and H<sub>2</sub>O<sub>2</sub>) masks the recombination control mediated by Sir2. Importantly, fluconazole increases LOH rate independently of the TRE element, as recombination rate at *TLO<sub>α10</sub>-ΔTRE-URA3<sup>+</sup>*, a reporter strain lacking the TRE element, was still higher following fluconazole treatment (Figure 7E and F). Given that fluconazole does not affect Sir2 RNA and protein level (Supplementary Figure S9), we suggest that stress conditions increase recombination rates independently of Sir2.

## DISCUSSION

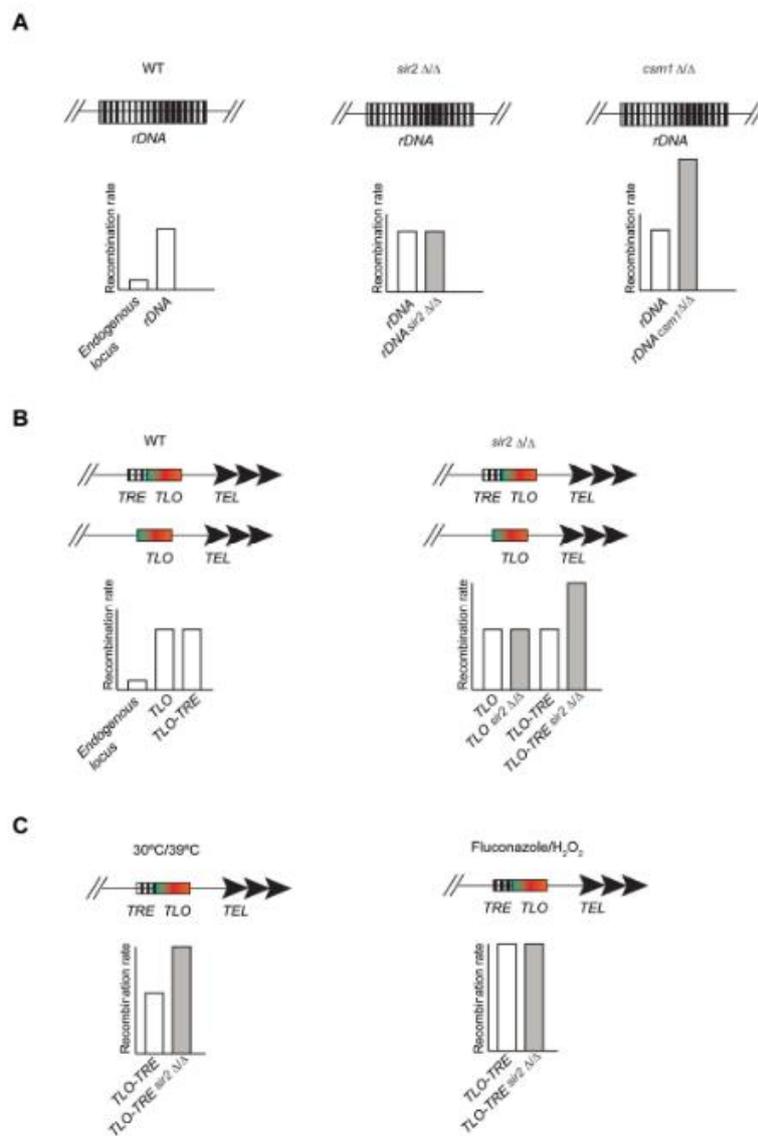
This study highlights how the HDAC Sir2 has acquired novel roles in the regulation of repeats-associated genome stability in *C. albicans*, the most common human fungal pathogen.

### The Monopolin complex, but not Sir2, promotes rDNA stability

In most eukaryotes, the *rDNA* locus is very plastic: the number of *rDNA* units changes in response to nutrients availability (42). However, excessive plasticity is deleterious and regulatory mechanisms have been evolved to ensure *rDNA* integrity. The regulatory network promoting *rDNA* stability is well established in *S. cerevisiae* where the HDAC Sir2



**Figure 7.** Instability of *TLO* genes triggered by stress conditions masks the action of *Sir2*. (A) Schematic of *TLOα10-URA3<sup>+</sup>*. (B) *TLOα10-URA3<sup>+</sup>* fluctuation analysis in WT and *sir2Δ/Δ* cells without (-) and with (+) fluconazole treatment, *p*-values =  $4.846 \times 10^{-09}$  and 0.6591, respectively. (C) Fluctuation analysis for LOH Rates in *TLOα10-URA3<sup>+</sup>* in WT and *sir2Δ/Δ* cells without (-) and with (+) H<sub>2</sub>O<sub>2</sub> treatment, *p*-values =  $4.846 \times 10^{-09}$  and 0.0001475, respectively. (D) *TLOα10-URA3<sup>+</sup>* in WT and *sir2Δ/Δ* cells fluctuation analysis at 30°C or 39°C, *p*-values =  $4.846 \times 10^{-09}$  and  $3.065 \times 10^{-05}$ , respectively. (E) Schematic of *TLOα10Δ TRE-URA3<sup>+</sup>*. (F) *TLOα10Δ TRE-URA3<sup>+</sup>* fluctuation analysis without (-) and with (+) fluconazole treatment, *p*-values =  $3.065 \times 10^{-05}$ .



**Figure 8.** Regulation of genome plasticity in *C. albicans*. (A) Schematic model to represent recombination control at the *rDNA* locus in WT, *sir2* $\Delta/\Delta$  and *csm1* $\Delta/\Delta$  cells. (B) Schematic model to represent recombination control at subtelomeric regions in WT and *sir2* $\Delta/\Delta$  cells. (C) Schematic model to represent recombination rates in control or under stress conditions (30/39°C, fluconazole and H<sub>2</sub>O<sub>2</sub> treatment) in WT, *sir2* $\Delta/\Delta$  cells.

and the Monopolin complex act in parallel to protect *rDNA* repeats integrity (43,44). Here, we demonstrate that *C. albicans* Sir2 has lost the ability to promote *rDNA* stability. In *C. albicans*, *rDNA* stability is ensured only by the Monopolin complex that also contributes to transcriptional repression (Figure 8A). This observation is surprising and in striking contrast with all the other organisms analysed to date. We propose that heterochromatin assembled at the *rDNA* locus has lost the ability to repress mitotic recombination in order to facilitate karyotype diversity. Indeed, in several *C. albicans* clinical isolates the chromosomal regions distal to the *rDNA* locus are largely homozygous (12), presumably due to a high recombination rate at the *rDNA* array. This strategy could be particularly important for *C. albicans* because it lacks a meiotic cycle and therefore must generate genetic diversity through mitotic events (13,14).

#### Recombination control of *C. albicans* subtelomeres

Here, we have analysed mechanisms governing genome stability at *C. albicans* subtelomeres (Figure 8). We found that all subtelomeres are unstable regions of the *C. albicans* genome as recombination rate associated with these regions is much higher than recombination rate associated with an internal not-repetitive region (Figure 8B). High genomic instability is associated with subtelomeres in many organisms and it is thought to play a key role in adaptation and evolution (2). This could be a critical regulatory mechanism in *C. albicans*. This is because *TLO* genes encode proteins with similarity to Med2, a component of the Mediator complex that regulates transcription by RNA polymerase II (45). Recombination between non-allelic *TLO* genes has the potential to introduce primary sequence changes into *Tlo* proteins encoded by the recombined ORFs. Since the resulting *Tlo* proteins all have different primary sequences (25), and since mediator complexes include only one Med2 subunit (23), changes in the identity and/or stoichiometry of the *TLO* genes is expected to alter the composition of the mediator complex driving changes in global transcriptional patterns. Consistent with this hypothesis, strains with different *TLO* organisation have altered fitness levels (27).

In many organisms, subtelomeric regions are assembled into transcriptionally silent heterochromatin that stochastically silences gene expression of associated genes (46). In previous work we found that Sir2-dependent heterochromatin assembles over *C. albicans* telomeric regions, where it stochastically represses expression of nearby genes, including *TLO* genes (28,47). In this study, we found that Sir2 controls the recombination dynamics of *C. albicans* subtelomeric regions via TRE, a novel recombination control element located at the 3' region of a subset of *TLO* genes. Our data demonstrate that the TRE element has the potential to mediate high levels of recombination and that Sir2 tempers recombination at all *TLO* genes that have an adjacent TRE (Figure 8B). Mechanisms underlying the TRE-recombination control are still unknown. However, it is possible that, similarly to what has been observed at the *S. cerevisiae* *rDNA* locus, the TRE element could lead to high level of recombination via inducing a replication fork stress and the Sir2 could promote genome stability by repressing non-

coding transcription ensuring high levels of cohesion (14–16).

Finally, we found that specific stress conditions, such as Fluconazole and H2O2 treatment, increase genome instability across the *C. albicans* genome (Figure 8C). Mechanisms that increase genome instability in specific stresses (e.g. fluconazole and H<sub>2</sub>O<sub>2</sub>) operate independently of Sir2 but they can mask the Sir2-dependent recombination control. This is likely because these reagents directly drive chromosome missegregation and/or DNA breaks and Sir2 is not involved in the effect of either fluconazole or H<sub>2</sub>O<sub>2</sub> on DNA integrity.

In summary, we show that, while *C. albicans* Sir2 does not promote *rDNA* stability, Sir2 ensures stability of subtelomeric genes via the *cis*-acting DNA element TRE. The contribution of Sir2-dependent recombination is independent of mechanisms triggering genomic instability in fluconazole or H<sub>2</sub>O<sub>2</sub>, but appears to be temperature sensitive (Figure 8). This study highlights another layer of complexity in the regulation of DNA repeats plasticity.

#### SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

#### FUNDING

Medical Research Council [MR/M019713/1]; Biotechnology and Biological Sciences Research Council [BB/L008041/1]. Funding for open access charge: University of Kent.

Conflict of interest statement. None declared.

#### REFERENCES

- Peng, J.C. and Karpen, G.H. (2008) Epigenetic regulation of heterochromatic DNA stability. *Curr. Opin. Genet. Dev.*, **18**, 204–211.
- Rando, O.J. and Verstrepen, K.J. (2007) Timescales of genetic and epigenetic inheritance. *Cell*, **128**, 655–668.
- Lazzerini-Denchi, E. and Sfeir, A. (2016) Stop pulling my strings – what telomeres taught us about the DNA damage response. *Nat. Rev. Mol. Cell Biol.*, **17**, 364–378.
- Larsen, D.H. and Stueck, M. (2016) Nucleolar responses to DNA double-strand breaks. *Nucleic Acids Res.*, **44**, 538–544.
- Pfaffel, M.A. (2012) Antifungal drug resistance: mechanisms, epidemiology, and consequences for treatment. *Am. J. Med.*, **125**, S3–S13.
- Selmecki, A., Forche, A. and Berman, J. (2010) Genomic plasticity of the human fungal pathogen *Candida albicans*. *Eukaryot. Cell*, **9**, 991–1008.
- 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. et al. (2015) Genetic and phenotypic intra-species variation in *Candida albicans*. *Genome Res.*, **25**, 413–425.
- Forche, A., Abbey, D., Pisithkul, T., Weinzierl, M.A., Ringstrom, T., Bruck, D., Petersen, K. and Berman, J. (2011) Stress alters rates and types of loss of heterozygosity in *Candida albicans*. *MBio*, **2**. doi:10.1128/mBio.00129-11.
- Forche, A., Magee, P.T., Selmecki, A., Berman, J. and May, G. (2009) Evolution in *Candida albicans* populations during a single passage through a mouse host. *Genetics*, **182**, 799–811.
- Rusche, L.N., Kirchmaier, A.L. and Rine, J. (2003) The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.*, **72**, 481–516.
- Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. et al. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, **406**, 593–599.

12. Allshire, R.C. and Ekwall, K. (2015) Epigenetic Regulation of Chromatin States in *Schizosaccharomyces pombe*. *Cold Spring Harb Perspect. Biol.*, **7**, a018770.
13. Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D. and Grewal, S.I. (2001) Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science*, **292**, 110–113.
14. Kobayashi, T. (2006) Strategies to maintain the stability of the ribosomal RNA gene repeats—collaboration of recombination, cohesion, and condensation. *Genes Genet. Syst.*, **81**, 155–161.
15. Kobayashi, T., Horiuchi, T., Tongaonkar, P., Vu, L. and Nomura, M. (2004) SIR2 regulates recombination between different rDNA repeats, but not recombination within individual rDNA genes in yeast. *Cell*, **117**, 441–453.
16. Kobayashi, T. and Ganley, A.R.D. (2005) Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science*, **309**, 1581–1584.
17. Huang, J., Brito, I.L., Villén, J., Gygi, S.P., Amon, A. and Moazed, D. (2006) Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer. *Genes Dev.*, **20**, 2887–901.
18. Chan, J.N.Y., Poon, B.P.K., Salvi, J., Olsen, J.B., Emili, A. and Mekhail, K. (2011) Perinuclear cohesin complexes maintain replicative life span via roles at distinct silent chromatin domains. *Dev. Cell*, **20**, 867–879.
19. Marvin, M.E., Griffin, C.D., Eyre, D.E., Barton, D.B.H. and Louis, E.J. (2009) In *Saccharomyces cerevisiae*, yKu and subtelomeric core X sequences repress homologous recombination near telomeres as part of the same pathway. *Genetics*, **183**, 441–451.
20. van het Hoog, M., Rast, T.J., Martchenko, M., Grindle, S., Dignard, D., Hogues, H., Cuomo, C., Berriman, M., Scherer, S., Magee, B.B. *et al.* (2007) Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biol.*, **8**, R52.
21. Huber, D. and Rustchenko, E. (2001) Large circular and linear rDNA plasmids in *Candida albicans*. *Yeast*, **18**, 261–272.
22. Rustchenko, E.P., Curran, T.M. and 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*. **175**, 7189–7199.
23. Zhang, A., Petrov, K.O., Hyun, E.R., Liu, Z., Gerber, S.A. and Myers, L.C. (2012) The Tlo proteins are stoichiometric components of *Candida albicans* mediator anchored via the Med3 subunit. *Eukaryot. Cell*, **11**, 874–84.
24. Haran, J., Boyle, H., Hokamp, K., Yeomans, T., Liu, Z., Church, M., Fleming, A.B., Anderson, M.Z., Berman, J., Myers, L.C. *et al.* (2014) Telomeric ORFs (TLOs) in *Candida* spp. encode mediator subunits that regulate distinct virulence traits. *PLoS Genet.*, **10**, e1004658.
25. Anderson, M.Z., Baller, J.A., Dulmage, K., Wigen, L. and Berman, J. (2012) The three clades of the telomere-associated TLO gene family of *Candida albicans* have different splicing, localization and expression features. *Eukaryot. Cell*, **11**, 612–625.
26. Jackson, A.P., Gamble, J.A., Yeomans, T., Moran, G.P., Saunders, D., Harris, D., Aslett, M., Barrell, J.F., Butler, G., Citiulo, F. *et al.* (2009) Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. *Genome Res.*, **19**, 2231–2244.
27. Anderson, M.Z., Wigen, L.J., Burack, L.S. and Berman, J. (2015) Real-time evolution of a subtelomeric gene family in *Candida albicans*. *Genetics*, **200**, 907–919.
28. Freire-Benítez, V., Price, R.J., Tarrant, D., Berman, J. and Buscaino, A. (2016) *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.*, **6**, 22989.
29. Wilson, R.B., Davis, D. and Mitchell, A.P. (1999) Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. *J. Bacteriol.*, **181**, 1868–1874.
30. De Backer, M.D., Maes, D., Vandoninck, S., Logghe, M., Contreras, R. and Luyten, W.H. (1999) Transformation of *Candida albicans* by electroporation. *Yeast*, **15**, 1609–1618.
31. Hall, B.M., Ma, C.-X., Liang, P. and Singh, K.K. (2009) Fluctuation analysis Calculator: a web tool for the determination of mutation rate using Luria-Delbrück fluctuation analysis. *Bioinformatics*, **25**, 1564–1565.
32. Lea, D.E. and Coulson, C.A. (1949) The distribution of the numbers of mutants in bacterial populations. *J. Genet.*, **49**, 264–285.
33. von der Haar, T. (2007) Optimized protein extraction for quantitative proteomics of yeasts. *PLoS One*, **2**, e1078.
34. Inglis, D.O., Arnaud, M.B., Binkley, J., Shah, P., Skrzypek, M.S., Wymore, F., Binkley, G., Miyasato, S.R., Simson, M. and Sherlock, G. (2012) The *Candida* genome database incorporates multiple *Candida* species: multispecies search and analysis tools with curated gene and protein information for *Candida albicans* and *Candida glabrata*. *Nucleic Acids Res.*, **40**, D667–D674.
35. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M. and Barton, G.J. (2009) Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, **25**, 1189–1191.
36. Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W. and Noble, W.S. (2009) MEME SUITE: tools for motif discovery and searching. *Nucleic Acids Res.*, **37**, W202–W208.
37. Forche, A., Steinbach, M. and Berman, J. (2009) Efficient and rapid identification of *Candida albicans* allelic status using SNP-RFLP. *FEMS Yeast Res.*, **9**, 1061–1069.
38. Huang, J., Brito, I.L., Villén, J., Gygi, S.P., Amon, A. and Moazed, D. (2006) Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer. *Genes Dev.*, **20**, 2887–2901.
39. Winzeler, E.A., Castillo-Davis, C.I., Oshiro, G., Liang, D., Richards, D.R., Zhou, Y. and Hartl, D.L. (2003) Genetic diversity in yeast assessed with whole-genome oligonucleotide arrays. *Genetics*, **163**, 79–89.
40. Cuomo, C.A., Güldener, U., Xu, J.-R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.-J., Baker, S.E., Rep, M. *et al.* (2007) The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen specialization. *Science*, **317**, 1400–1402.
41. Brown, C.A., Murray, A.W. and Verstrepen, K.J. (2010) Rapid expansion and functional divergence of subtelomeric gene families in yeasts. *Curr. Biol.*, **20**, 895–903.
42. Jack, C.V., Cruz, C., Hull, R.M., Keller, M.A., Ralsler, M. and Houseley, J. (2015) Regulation of ribosomal DNA amplification by the TOR pathway. *Proc. Natl. Acad. Sci. U.S.A.*, **112**, 9674–9679.
43. McMurray, M.A. and Gottschling, D.E. (2003) An age-induced switch to a hyper-recombinational state. *Science*, **301**, 1908–1911.
44. Andersen, M.P., Nelson, Z.W., Hetrick, E.D. and Gottschling, D.E. (2008) A genetic screen for increased loss of heterozygosity in *Saccharomyces cerevisiae*. *Genetics*, **179**, 1179–1195.
45. Batta, K., Zhang, Z., Yen, K., Goffman, D.B. and Pugh, B.F. (2011) Genome-wide function of H2B ubiquitylation in promoter and genic regions. *Genes Dev.*, **25**, 2254–2265.
46. Bühler, M. and Gasser, S.M. (2009) Silent chromatin at the middle and ends: lessons from yeasts. *EMBO J.*, **28**, 2149–2161.
47. Anderson, M.Z., Gerstein, A.C., Wigen, L., Baller, J.A. and Berman, J. (2014) Silencing is noisy: population and cell level noise in telomere-adjacent genes is dependent on telomere position and Sir2. *PLoS Genet.*, **10**, e1004436.

### **3. Supporting information**

## Supporting Information

### **Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen**

#### ***Candida albicans***

Verónica Freire-Benítez<sup>1</sup>, Sarah Gourlay<sup>1</sup>, Judith Berman<sup>2</sup> and Alessia Buscaino<sup>1</sup>

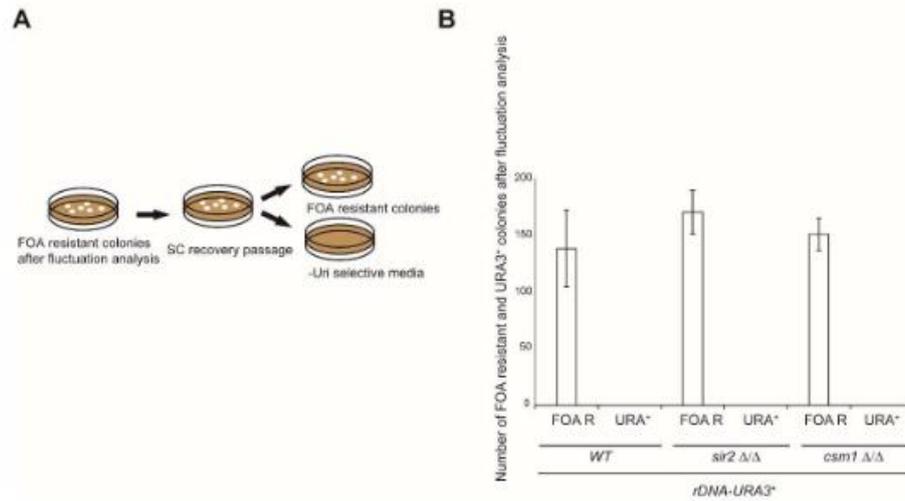
<sup>1</sup> University of Kent, School of Biosciences Canterbury Kent, CT2 7NJ. UK

<sup>2</sup> Department of Microbiology and Biotechnology, George S. Wise Faculty of Life Sciences, Tel Aviv University, Ramat Aviv, 69978, Israel

#### **Content:**

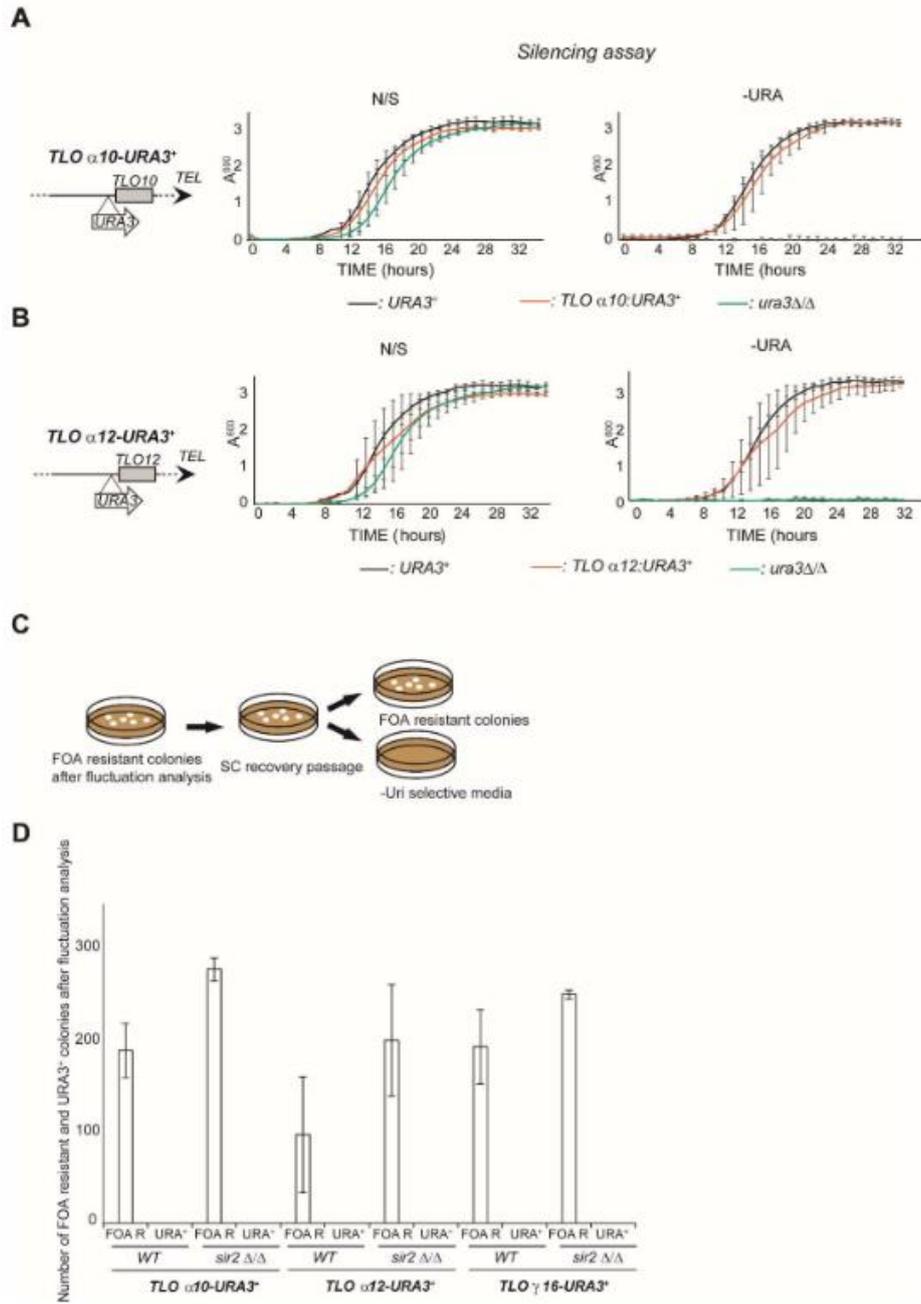
**Supplementary Fig S1 and S9**

**Supplementary Table S1 to S3**



**Supplementary figure 1.**

**(A)** Schematic of experimental procedure to detect silencing. Following fluctuation analysis, FOA resistant single colonies were streaked on SC plates and grown for 24 hour. The day after, single colonies were streaked on FOA selective plates and -Uri plates. **(B)** Histogram shows the number of FOA resistant (URA3<sup>-</sup>) and URA3<sup>+</sup> colonies following fluctuation analyses of rDNA-URA3<sup>+</sup> strain in WT, *sir2* Δ/Δ and *csm1* Δ/Δ strain. Error bars: Standard Deviation of three biological replicates.



**Supplementary figure 2.**

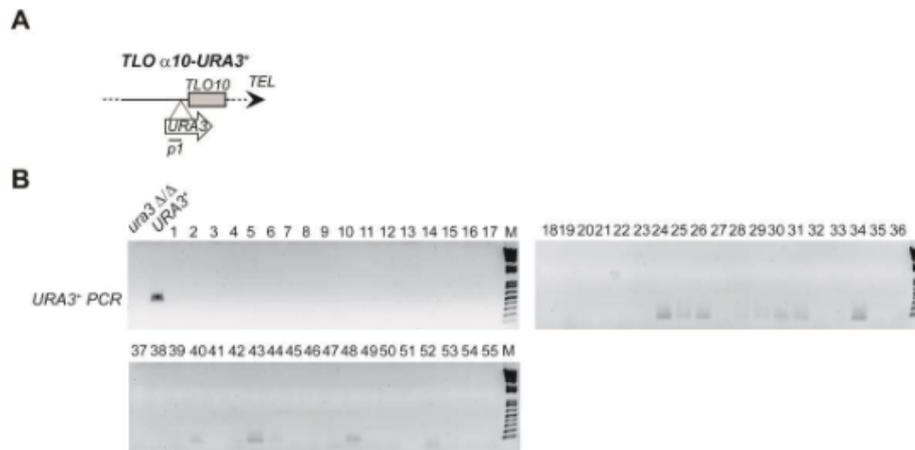
**(A)** Left panel: Schematic of *TLO $\alpha$ 10-URA3<sup>+</sup>* strain. Right panel: silencing assay of the *TLO $\alpha$ 10-URA3<sup>+</sup>* strain. Ura<sup>+</sup> (URA3/URA3) and Ura<sup>-</sup> (*ura3 $\Delta$ /ura3 $\Delta$* ) strains were included as controls.

**(B)** Left panel: Schematic of *TLO $\alpha$ 12-URA3<sup>+</sup>* strain. Right panel: silencing assay of the *TLO $\alpha$ 12-URA3<sup>+</sup>* strain. *URA3<sup>+</sup>* (URA3/URA3) and *ura<sup>-</sup>* (*ura3 $\Delta$ /ura3 $\Delta$* ) strains were included as controls.

Error bars in each panel: Standard deviation (SD) of three biological replicates. **(C)** Schematic of experimental procedure to detect silencing following fluctuation analysis.

**(D)** Histogram shows the number of FOA resistant (URA3<sup>-</sup>) and URA3<sup>+</sup> colonies following fluctuation analyses of *TLO $\alpha$ 10-URA3<sup>+</sup>*, *TLO $\alpha$ 12-URA3<sup>+</sup>* and *TLO $\gamma$ 16-URA3<sup>+</sup>* strains in WT and *sir2  $\Delta/\Delta$*  strains.

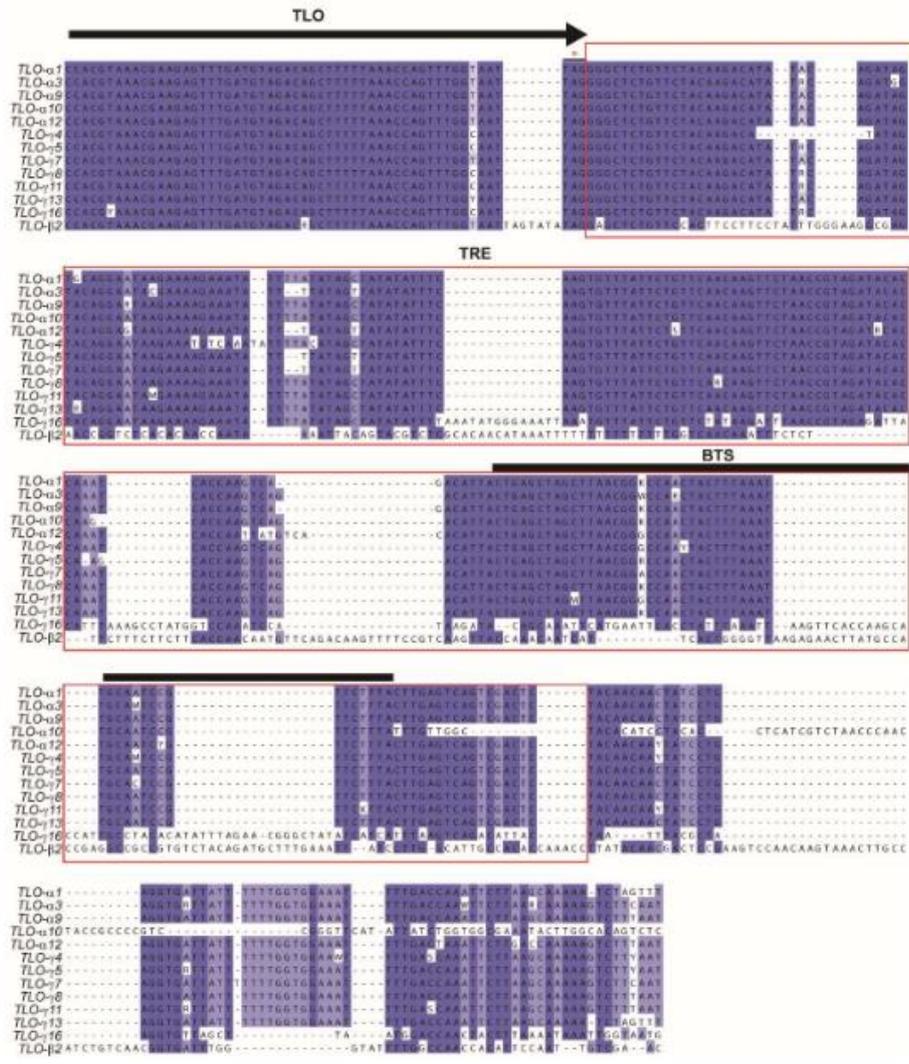
Error bars: Standard Deviation of three biological replicates.



**Supplementary figure 3.**

(A) Upper panel: Schematic of *TLO $\alpha$ 10-URA3<sup>+</sup>* strain (B) *URA3<sup>+</sup>* PCR analyses with primers specific for the *URA3<sup>+</sup>* marker gene was performed with 55 FOA resistant colonies (after fluctuation analysis). A *URA3<sup>+</sup>* and a *ura3  $\Delta/\Delta$*  strains were included as a positive and negative control.

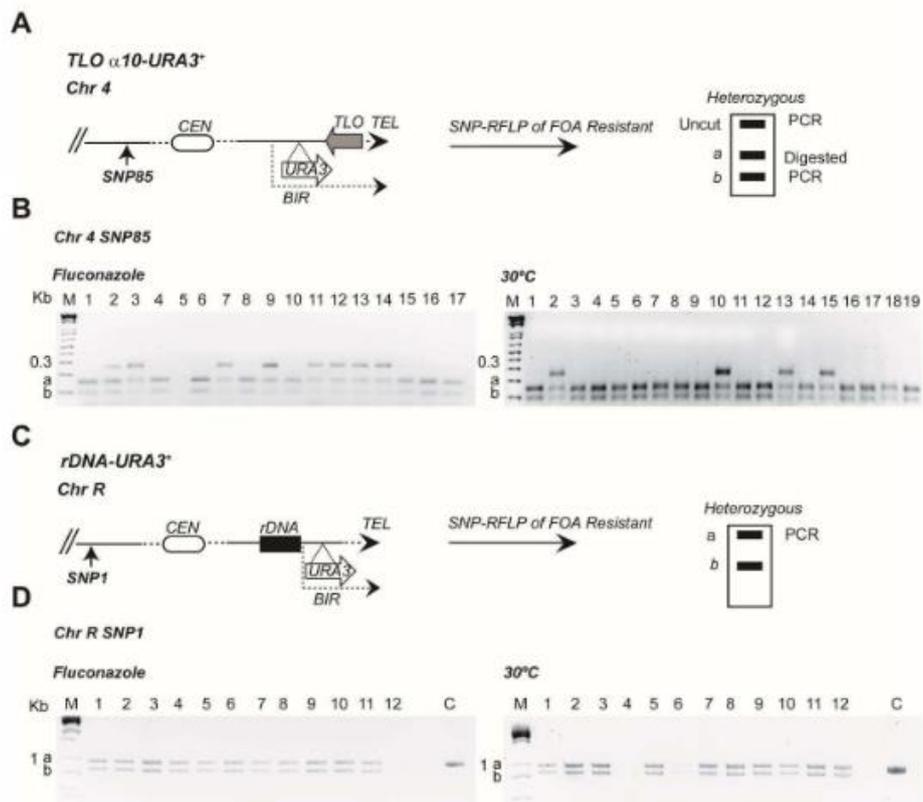
Freire-Benítez et al\_ Fig S4



Supplementary figure 4.

Sequence alignment of *TLO* genes and downstream sequences. A red square highlights the 300bp TRE element and a black solid line the 50 bp BTS sequence.

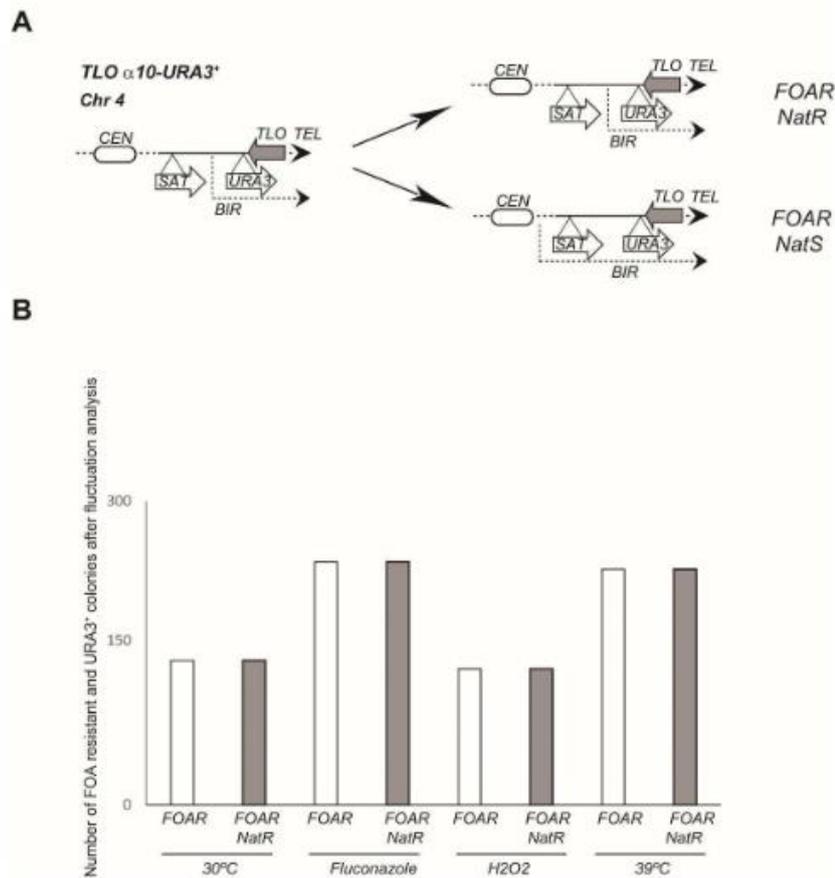




**Supplementary figure 6**

(A) Schematic of *TLOα10-URA3+* strain and SNP85 located on the left arm of chromosome 4 and predicted digestion pattern after SNP-RFLP analysis. A different digestion pattern between the two different homologues chromosome (a and b) allows to detect homozygosity or heterozygosity (B) Following fluctuation analyses in the presence of absence of 1 µg/ml fluconazole, PCR analyses were performed on 17-19 FOA resistant following by TaqI enzymatic digestion. The digested product was run into a 1.5 % agarose gel. M: 100 bp ladder, the two homologues chromosome (a and b) are indicated (C) Schematic of *TLOα12-URA3+* strain and SNP1 located on the left arm of chromosome R and predicted digestion pattern after SNP-RFLP analysis. A different digestion pattern between the two different homologues chromosome (a and b) allows to detect homozygosity or heterozygosity (D) Following fluctuation analyses in the presence of absence of 1 µg/ml fluconazole, PCR analyses were performed on 12 FOA

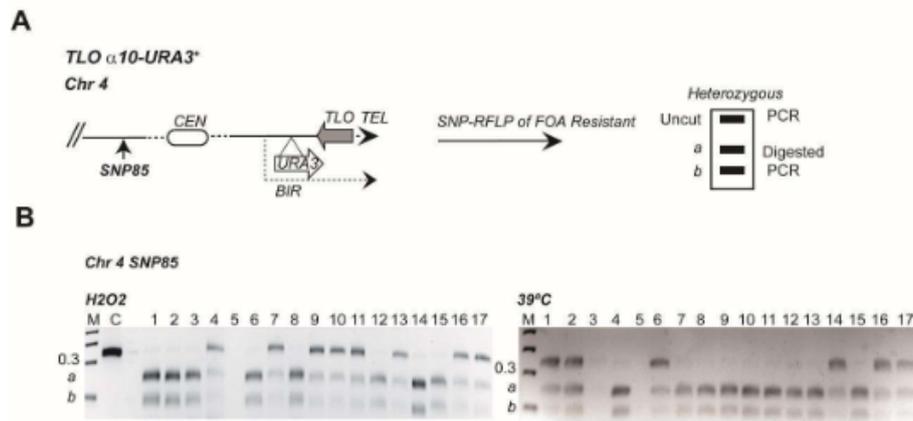
resistant colonies, the PCR product was digested with AseI. The digested product was run into a 1.5 % agarose gel. A non-digested SNP 1 PCR product was included as a control. M: 100 bp ladder, the two homologues chromosome (a and b) are indicated



**Supplementary figure 7**

**(A)** Schematics of the *TLO α10-URA3<sup>+</sup>-SAT1* strain to confirm LOH range at *TLO α10* locus in the presence of stress conditions. A breaking point between the *SAT1* gene and *URA3<sup>+</sup>* marker gene would produce FOA resistant (FoAR) and NAT resistant (NATR) colonies. A breaking point upstream of both marker genes would produce FOAR colonies that are sensitive to NAT

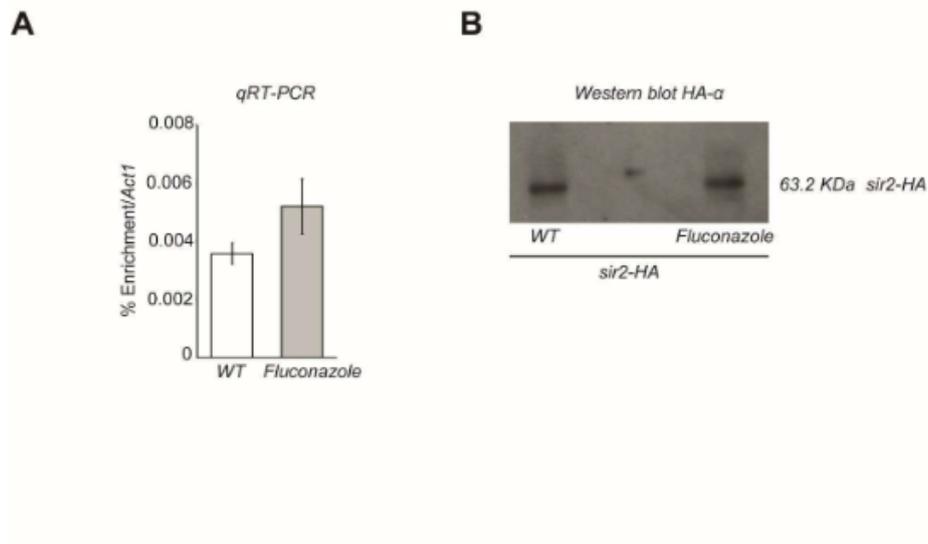
**(B)** Schematic of experimental procedure to test whether, following fluctuation analyses, colonies that have lost the *URA3<sup>+</sup>* marker gene (FOAR) retain the *SAT1* marker gene and, therefore, are NAT resistant. **(E)** Histograms show that most of the FOA resistant colonies analysed after fluctuation analysis in all stress conditions are resistant to the antibiotic NAT.



**Supplementary figure 8**

(A) Schematic of *TLO* $\alpha$ 10-*URA3*<sup>\*</sup> strain and SNP85 located on the left arm of chromosome 4 and predicted digestion pattern after SNP-RFLP analysis. A different digestion pattern between the two different homologues chromosome (a and b) allows to detect homozygosity or heterozygosity (B) SNP 85 (Forche) PCR analyses coupled with TaqI enzymatic digestion were performed with 17 FOA resistant colonies after LOH in the presence of 0.5mM H<sub>2</sub>O<sub>2</sub> and at 39°C.

Following fluctuation analyses, SNP 85 PCR analyses coupled with TaqI enzymatic digestion were performed on 17 resistant colonies in the presence and absence of 1  $\mu$ g/ml fluconazole and run into a 1.5 % agarose gel. M: 100 bp ladder, the two homologues chromosome (a and b) are indicated



**Supplementary figure 9**

**(A)** qRT-PCR analyses to measure *SIR2* transcript levels relative to actin transcript levels (*ACT1*) at 30°C in the presence and absence of fluconazole. Error bars: SD of three biological replicates. **(B)** Western blot to measure Sir2 protein levels in the presence and absence of fluconazole.

**Supplementary table 1.** *C. albicans* strains used in this study.

Strain number	Description	Genotype
Bu_54	SN152. URA3 heterozygous (1)	MTL a/alpha ura3Δ-iro1Δ::imm434/URA3-IRO1 his1Δ/his1Δ arg4Δ/arg4Δ leu2Δ/leu2Δ
Bu_95	rDNA:URA3* (2)	rDNA::URA3 ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG
Bu_97	URA3-rDNA	URA3-rDNA ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG
Bu_102	rDNA:URA3* sir2Δ/Δ (2)	rDNA ::URA3 ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG sir2Δ::HIS1/sir2Δ::ARG4
Bu_117	URA3-rDNA sir2Δ/Δ	URA3-rDNA ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG sir2Δ::HIS1/sir2Δ::ARG4
Bu_133	SN152. URA3 heterozygous sir2Δ/Δ	MTL a/alpha ura3Δ-iro1Δ::imm434/URA3-IRO1 his1Δ/his1Δ arg4Δ/arg4Δ leu2Δ/leu2Δ sir2Δ::HIS1/sir2Δ::ARG4
Bu_147	TLO α10-URA3* (3)	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG TLOα10/ TLOα10-GFP-URA3-tADH
Bu_148	TLO α 12-URA3* (3)	ura3Δ::imm434/ura3Δ03::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG TLOα12/ TLOα12-GFP-URA3-tADH
Bu_152	TLO α 12-URA3* sir2Δ/Δ (3)	ura3Δ::imm434/ura3Δ03::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG TLOα12/ TLOα12-GFP-URA3-tADH sir2Δ::HIS1/sir2 Δ::ARG4
Bu_153	TLO α10-URA3* sir2Δ/Δ (3)	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG TLOα10/ TLOα10-GFP-URA3-tADH sir2Δ::HIS1/sir2 Δ::ARG4
Bu_185	rDNA-URA3	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG ERG13/ERG13-GFP-URA3
Bu_202	rDNA-URA3 sir2Δ/Δ	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG ERG13/ERG13-GFP-URA3 sir2Δ::HIS1/sir2 Δ::NAT
Bu_234	rDNA-URA3* csm1Δ/Δ	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG ERG13/ERG13-GFP-URA3 csm1Δ::HIS1/csm1Δ::NAT
Bu_253	TLO γ16-URA3* [3]	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG TLO γ16/TLO γ16-GFP-URA3
Bu_272	Sir2-HA	MTL a/alpha tel::URA3 ura3Δ::Δimm434/ura3Δ::Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG SIR2-HA
Bu_277	TLO γ16-URA3* sir2Δ/Δ	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG TLO γ16/TLO γ16-GFP-URA3 sir2Δ::HIS1/sir2Δ::NAT
Bu_304	NAT- TLO α10-URA3*	ura3Δ::imm434 ura3Δ::imm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG VID21:NAT TLOα10/ TLOα10-GFP-URA3-tADH
Bu_318	TLO α10 ΔTRE-URA3*	MTL a/alpha ura3Δ::Δimm434/ura3Δ::Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG bts10Δ::URA3/bts10Δ::NAT
Bu_323	TLO α10 ΔTRE-URA3* sir2Δ/Δ	MTL a/alpha ura3Δ::Δimm434/ura3Δ::Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG bts10Δ::URA3/bts10Δ::NAT sir2 Δ::HIS1/sir2 Δ::ARG4
Bu_365	rDNA:URA3 csm1Δ/Δ	rDNA::URA3 ura3Δ::λimm434/ura3Δimm434 his1::hisG/his1::hisG arg4::hisG/arg4::hisG csm1Δ::ARG4/csm1Δ::NAT
Bu_450	SN152. URA3 heterozygous csm1Δ/Δ	MTL a/alpha ura3Δ-iro1Δ::imm434/URA3-IRO1 his1Δ/his1Δ arg4Δ/arg4Δ leu2Δ/leu2Δ csm1Δ::NAT/csm1Δ::ARG4

**Supplementary table 2.** Primers used in this study.

Primer	Sequence	Figure	Description
Bu_131	GACTGGCCAATTATAAATGTGAAGG	Fig5 A-B	Primer to check TRE deletion
Bu_139	GAGTGAGTGAGTGGAGTAGCG	Fig 1 D-E, Fig2, 4B, Fig5, FigS1B, FigS2D	Primer to check <i>sir2Δ/Δ</i> deletion
Bu_141	GTTGGGCAGATATTACCAATG	Fig5 A-B, FigS3, Fig3D	Primer to check URA3 <sup>+</sup> , TRE deletion
Bu_142	CCTTCACATTTATAATTGGCC	FigS3, Fig3D	Primer to check URA3 <sup>+</sup>
Bu_143	GACACTATATAGTGAGAATAAAATAA TTCATTTGAAAAAAATTGATAGTAAA AGATTTGGAGTAAATATTTTGGTTTTTC CCAGTCACGACGTT	Fig1C-D	URA3 <sup>+</sup> -rDNA
Bu_144	GTAAAGCACATATGCATAATACTGGT CGAGAATTGAAAGATTGTTATTTTTTT CTAACAAAAAAGAAACGCCATATGT GGAATTGTGAGCGGATA	Fig1C-D	URA3 <sup>+</sup> -rDNA
Bu_147	GTTCTAGCTTCTTCGCTAGTG	Fig1C-D	Primer to check URA3 <sup>+</sup> -rDNA
Bu_148	CAATTCGCCCTATAGTGAG	Fig1C-D	Primer to check URA3 <sup>+</sup> -rDNA
Bu_152	CTGGAGAAAATATAACCACGAGTCTA AGTTTCTTTATTATATTGACGTTTCAG TTATTTGAGAGAAATCCTCTAGTAGT TTTCCCAGTCACGACGTT	Fig 1 D-E, Fig2, 4B, Fig5, FigS1B, FigS2D	<i>sir2Δ/Δ</i> deletion mutant
Bu_153	ATATATAAATATATAAATATATATAT AAAAGAATTGAAAAGAAAACATTAA AGACACCAATATTAATTTAATGTGGA ATTGTGAGCGGATA	Fig 1 D-E, Fig2, 4B, Fig5, FigS1B, FigS2D	<i>sir2Δ/Δ</i> deletion mutant
Bu_164	CGGTCTGGTAAATGATTGAC	Fig1A, Fig1D, Fig2A, Fig4B, Fig5B, Fig5D, FigS1B, FigS2D	Primer to check <i>HIS1</i> integration
Bu_165	AGTGTGGAAAGAAGAGATGC	Fig1D-E, Fig5B, Fig2, FigS1B	Primer to check <i>ARG4</i> integration
Bu_179	CTGTATCTATAAGCAGTATCATCC	Fig1A, Fig1D, Fig2A-B, Fig4B, Fig5D, FigS1B, FigS2D	Primer to check <i>NAT</i> integration

Bu_274	CGTCGTTTTTTTTTTCTTTTTGTTATA TCCCCTTTAACTATCAGCACATACA CACCAACACCATTGTTTTCCAGTCA CGACGTT	g2SD Fig1A,Fi g2, FigS1B	<i>csm1Δ/Δ</i> deletion mutant
Bu_275	CAAAAACTTAAAGAAATTCTCAAGA CATTGAAAATTACCAAAGTAGATGGC AAAATCGTTATAGAATAATAATGTGG AATTGTGAGCGGATA	Fig1A,Fi g2, FigS1B	<i>csm1Δ/Δ</i> deletion mutant
Bu_276	GGGAGATCGAGACAATGGAT	Fig1A,Fi g2, FigS1B	Primer to check <i>csm1Δ/Δ</i> deletion
Bu_286	CTGGAGAAAATATAACCACGAGTCTA AGTTTCTTTATTATATTGACGTTTCAG TTATTTGAGAGAAATCCTCTAGTAGT AAAACGACGGCCAGTGAATTC	Fig1A, Fig1D,Fi g2A-B, Fig4B,Fi g5B, Fig5D,Fi GS1B,Fi G2SD	<i>sir2Δ/Δ</i> deletion mutant
Bu_287	ATATATAAATATATAAATATATATAT AAAAGAATTGAAAAGAAAACATTAA AGACACCAATATTAATTTAATGCATC AATTGACGTTGATACCAC	Fig1A, Fig1D,Fi g2A-B, Fig4B,Fi g5B, Fig5D,Fi GS1B,Fi G2SD	<i>sir2Δ/Δ</i> deletion mutant
Bu_288	CGTCGTTTTTTTTTTCTTTTTGTTATA TCCCCTTTAACTATCAGCACATACA CACCAACACCATTGAAAACGACGG CCAGTGAATTC	Fig2A- B, FigS1B	<i>csm1Δ/Δ</i> deletion mutant
Bu_289	CAAAAACTTAAAGAAATTCTCAAGA CATTGAAAATTACCAAAGTAGATGGC AAAATCGTTATAGAATAATAATGCAT CAATTGACGTTGATACCAC	Fig2A- B, FigS1B	<i>csm1Δ/Δ</i> deletion mutant
Bu_315	GATGAATTTAAATTATCAAACTTAAA AATGGTGATTGGGAAATTGTCAAGAA ATCAACTTCGACAAAAAACGGATCC CCGGGTTAATTA	FigS8B	<i>SIR2-HA</i> tagging
Bu_316	ATATATAAATATATATATAAAAGAA TTGAAAAGAAAAACATTAAGACACC AATATTAATTTAATCAGTAAAACGAC GGCCAGTGAATTC	FigS8B	<i>SIR2-HA</i> tagging
Bu_317	GATAGCCCGCATAGTCAGGAAC	FigS8B	Primer to check BTS HA tagged <i>SIR2-HA</i> mutant
Bu_356	GGCTTTGACATCAACGACA	Fig5 A- B	Primer to check TRE $\Delta$ deletion mutant
Bu_370	GCACATTTGGTGCCCACTCAAGCAC TACACGCCTTGCTGTGCCCTGAC CCTGCCGCAAGTGGAAGAGCCGTAA AACGACGGCCAGTGAATTC	Fig3E-F, FigS7A- B	<i>NAT- TLO <math>\alpha</math>10-URA3<sup>+</sup></i>
Bu_371	CTAGCAACTCCTTTCCCAATAATTC ATTACCTCCTCGTCGGCAACATCTAT TGTTTTACCTCTGGGATAATGCATCA ATTGACGTTGATACCAC	Fig3E-F, FigS7A- B	<i>NAT- TLO <math>\alpha</math>10-URA3<sup>+</sup></i>
Bu_372	CTAACCAATGTCAGACAACG	Fig3E-F, FigS7A- B	Primer to check <i>NAT- TLO <math>\alpha</math>10-URA3<sup>+</sup></i> construction

Bu_400	CGTAATCTGGAACGTCATATGG	Fig5 A-B	Primer to check TRE $\Delta$ deletion mutant
Bu_401	TACTGTTCTGCAGAAGCCT	Fig5 A-B	Primer to check TRE $\Delta$ deletion mutant
Bu_402	TGCTCGACAATGGCGACCACGTAAA CGAAGAGTTTGATGTAGACAGCTTTT TAAACCAGTTTGGTAATCCGCGGCG GATCCCCGGGTTAATTA	Fig5 A-B	TRE $\Delta$ with URA3/NAT
Bu_403	TTGAGTGGCGGGTCATTCTTTAGAGT GTTTGCAGGGTTGGTGGTGGTTTGGC ATTGGGTTGGCGTATGTGGTTCTAGA AGGACCACCTTTGATTG	Fig5 A-B	TRE $\Delta$ with URA3
Bu_404	TTGAGTGGCGGGTCATTCTTTAGAGT GTTTGCAGGGTTGGTGGTGGTTTGGC ATTGGGTTGGCGTATGTGGTGAAAA CGACGGCCAGTGAATTC	Fig5 A-B	TRE $\Delta$ with NAT
BU_606	TGCCCAAATGTCTTCCGAT	Fig6C-D	SNP1
BU_607	GAGGTAAGGGTTCAAGTCCA	Fig6C-D	SNP1
Bu_612	CTCGGGAGAATAAGCTTACCATCTG	Fig6A-B, Fig8	SNP85
Bu_613	TTACTTGTGGGAAATCTGAACAGC	Fig6A-B, Fig8	SNP85

**Supplementary table 3.** Plasmids used.

Plasmid	Description
pGEMURA3	URA3 integration products (4)
pGEMHIS1	HIS1 substitution products (4)
pRS-Arg4Spel	Arg4 substitution products (4)
PHA_URA3	HA-URA3 substitution products
pHA_NAT	NAT substitution products (5)

## REFERENCE

1. Noble,S.M. and Johnson,A.D. (2005) Strains and Strategies for Large-Scale Gene Deletion Studies of the Diploid Human Fungal Pathogen *Candida albicans*. *Eukaryotic Cell*, **4**, 298.
2. Freire-Benítez,V., Price,R.J., Tarrant,D., Berman,J. and Buscaino,A. (2016) *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.*, **6**, 22989.
3. Anderson,M.Z., Gerstein,A.C., Wigen,L., Baller,J. a and Berman,J. (2014) Silencing Is Noisy: Population and Cell Level Noise in Telomere-Adjacent Genes Is Dependent on Telomere Position and Sir2. *PLoS Genet.*, **10**, e1004436.
4. Wilson,R.B., Davis,D. and Mitchell,A.P. (1999) Rapid Hypothesis Testing with *Candida*

albicans through Gene Disruption with Short Homology Regions These include : Rapid Hypothesis Testing with Candida albicans through Gene Disruption with Short Homology Regions. **181.**

5. Gerami-nejad,M., Forche,A., Mcclellan,M. and Berman,J. (2012) Analysis of protein function in clinical C . albicans isolates. **5314.**

## Chapter 5. Discussion

Understanding the function and structure of the chromatin state associated with the genome of a large variety of organisms is crucial in order to discover how epigenetic regulation modulates the function of eukaryote genomes. In this thesis, I have investigated the chromatin state associated with DNA repeats in *C. albicans*, the most common human fungal pathogen. *C. albicans* genome is highly plastic in stressful conditions, undergoing a number of important genome rearrangements, such as loss of heterozygosity (LOH) events, aneuploidies, and the formation of isochromosomes (Berman and Sudbery, 2002; Rustchenko, 2007; Selmecki et al., 2008). Heterochromatin promotes genome stability by preventing unfaithful recombination events between DNA repeated sequences (Grewal and Jia, 2007). *C. albicans* is an outstanding model to study the effect of heterochromatin on genome stability due to the plastic nature of its genome. My results show that heterochromatin is assembled differentially at *C. albicans* DNA repeats and that epigenetic factors differentially control genome stability associated with different loci.

## 1. Characterization of the chromatin state associated with *C. albicans* DNA repeats.

The first goal of my thesis was to characterize the chromatin state associated with *C. albicans* DNA repeats. The hypothesis was that, as observed in the model systems *S. cerevisiae* and *S. pombe*, repetitive domains were assembled into heterochromatin. However as in the case of *S. cerevisiae*, the *C. albicans* epigenome does not encode for a SuVar 3-9 orthologue. Therefore its epigenome is almost certainly devoid of H3K9me, a histone mark associated with heterochromatin in higher eukaryotes. *Candida albicans* has five members for the Sirtuin family. Within these members gene dosage of *HST1* correlates with the replicative life span of *C. albicans*. Cells with both gene copies of *HST1* have longer life span. These observations are consistent with the role of *ScSIR2* preventing aging in *S. cerevisiae* (Hickman et al., 2011). On the other hand, deletion of *CaSIR2* lowers the frequency of phenotypic switching from the opaque state to the white state, a requisite for the mating process (Miller and Johnson, 2002). However, while BLAST analyses identify an ortholog for *SIR2*, it is not possible to identify any ortholog for the *S. cerevisiae* SIR complex members (*SIR1*, *SIR3* and *SIR4*). Therefore, one possibility was that *C. albicans* epigenome was devoid of heterochromatin. My analyses demonstrated that heterochromatin exists in *C. albicans* and it is assembled at the *rDNA* locus and telomeres and it is *SIR2*-dependent. In addition, I have demonstrated that heterochromatin at telomeres is plastic. Furthermore, an intermediate chromatin state, bearing features of both heterochromatin and euchromatin, is assembled at the *MRS* and pericentromeric repeats (**figure 5.1**).

### 1.1. Heterochromatin at the *rDNA* locus (Freire-Benítez et al., 2016a)

My results showed that the *rDNA* is assembled into classical silent heterochromatin associated with low acetylation levels at H3K9 and H4K16 and low methylation levels at H3K4. Surprisingly, at the *rDNA* locus *C. albicans* Sir2 deacetylates H3K9 but not H4K16. This is in contrast with what is observed in *S. cerevisiae* where the major substrate of Sir2 is H4K16 (Margie T. Borra et al., 2004). This shows differential association of epigenetic marks at the same locus in relatively evolutionary distant organisms. Thus, the same epigenetic factors assemble heterochromatin differentially among organisms.

The HDAC Hst1 at *C. albicans* (CaHst1) has higher amino acid identity to *S. cerevisiae* Sir2 (ScSir2) than *C. albicans* Sir2 (CaSir2). However, *C. albicans* heterochromatin assembly is CaSir2-dependent. Indeed, deletion of both copies of *HST1* did not abolish the transcriptional silencing of the *URA3*<sup>+</sup> marker gene inserted at repeats. In addition, *hst1Δ/Δ* null mutant increased transcriptional silencing in a strain carrying a *rDNA:URA3*<sup>+</sup> reporter. This effect could be due to an indirect effect of *HST1* deletion. Contrarily to Sir2 that promotes regional repression of big regions of DNA, in other organism such as *S. cerevisiae*, Hst1 acts on specific gene promoters and it is involved in gene specific repression (Mead et al., 2007). Similarly, our results show that Hst1 at *C. albicans* does not promote silencing of big blocks of repeats. Deletion of *HST1* and loss of its gene specific repression could lead to an increase of Sir2 molecules to repress the transcription of the *URA3*<sup>+</sup> reporter gene inserted at the *rDNA*.

In *S. cerevisiae*, Sir2 is part of the RENT protein complex (NET1, CDC14 and SIR2) (Straight et al., 1999). BLAST analyses reveal the presence of *NET1* and

*CDC14* homologs in *C. albicans*. ScSir2 participates within the RENT complex promoting gene silencing and inhibition of recombination at the *rDNA* (Fritze et al., 1997; Huang et al., 2006b; Kobayashi, 2011; Straight et al., 1999). At *C. albicans* Sir2 could also have a different mechanism of action where Sir2 performs transcriptional silencing associated with other proteins. This is supported by differences observed in the amino acid identity between CaSir2 and ScSir2 and also by results showing that Sir2 does not control recombination at *C. albicans rDNA* contrarily to ScSir2. This finding will be discussed later in this thesis.

In this thesis introduction, it was explained how at *S. cerevisiae* Sir2 represses transcription of non-coding DNA at the *rDNA* independently of Pol I and Pol III transcription (Li et al., 2006b). Our results revealed that Sir2 also represses Pol II transcription of non-coding RNAs arising from *C. albicans rDNA* locus. This means that Sir2 role of silencing is conserved. This is consistent with the idea that SIR2-dependent silencing is observed in other organisms which lack the RENT complex, as it is the case of *S. pombe* or mammalian cells (Imai et al., 2000; Salminen and Kaarniranta, 2009; Shankaranarayana et al., 2003). *rDNA* repeats are then a hotspot for heterochromatin assembly among different organisms, including pathogens as *C. albicans*. This shows the importance of silencing certain blocks of genomic repeats with crucial functions. The *rDNA* locus is implied in the ribosome biogenesis crucial for cell survival and accounts for more than 50% transcription of a cell (Russell et al., 2005). Despite the high genome plasticity associated with pathogens, silencing could ensure genome integrity at specific loci required for cell viability.

## 1.2. Plastic heterochromatin at telomeres (Freire-Benítez et al., 2016a)

My results show that SIR2-dependent heterochromatin is also assembled at telomeres. While Sir2 deacetylates H3K9 at the *rDNA* locus, at telomeric repeats Sir2 deacetylates both H3K9ac and H4K16ac. This is in agreement with what has been observed in *S. cerevisiae* where H4K16 hypoacetylation is required for heterochromatin assembly at telomeres (Bi, 2014). Apart from Sir2, no other components of the SIR complex can be detected in *C. albicans*. It is possible that other proteins tether Sir2 to telomeric regions in a complex which has not yet been identified. This is supported by the idea that a Rap1 orthologue is encoded by *C. albicans* genome (Yu et al., 2010). In *S. cerevisiae* Rap1 is required for subtelomeric silencing via tethering the SIR complex to telomeric regions (Moretti et al., 1994). Although the SIR complex is absent in *C. albicans* it will be interesting to determine whether Rap1 or other proteins tether Sir2 to subtelomeric regions.

Our RNA-sequencing analysis has demonstrated that subtelomeric coding and non-coding transcripts are increased in *sir2*  $\Delta/\Delta$  cells compared to WT cells. As described in this thesis introduction, *C. albicans* subtelomeric genes are subject to TANGEN or transcriptional noise associated with subtelomeric genes. TANGEN effect is Sir2-dependent and *sir2*  $\Delta/\Delta$  null mutants reduced the transcriptional gene noise associated with subtelomeres (Anderson et al., 2014). In *S. cerevisiae* and *S. pombe*, Sir2-dependent transcriptional position effect (TPE) is also observed at subtelomeric regions (Smith et al., 1999; Stavenhagen and Zakian, 1998; Tham and Zakian, 2002). Our results support the idea that, also in *C. albicans*, Sir2-dependent transcriptional silencing spans from telomeres to subtelomeric regions. That could explain the loss of TANGEN

noise observed in *SIR2* null mutants. Silencing is a stochastic process; some of the genes in one region can be transcriptionally silenced in one cell while the same genes in another cell can be silenced to a lesser extent. Deletion of *SIR2* could erase the stochastic effect of silencing at subtelomeric regions. Therefore in the absence of Sir2, subtelomeric genes could be homogeneously transcribed.

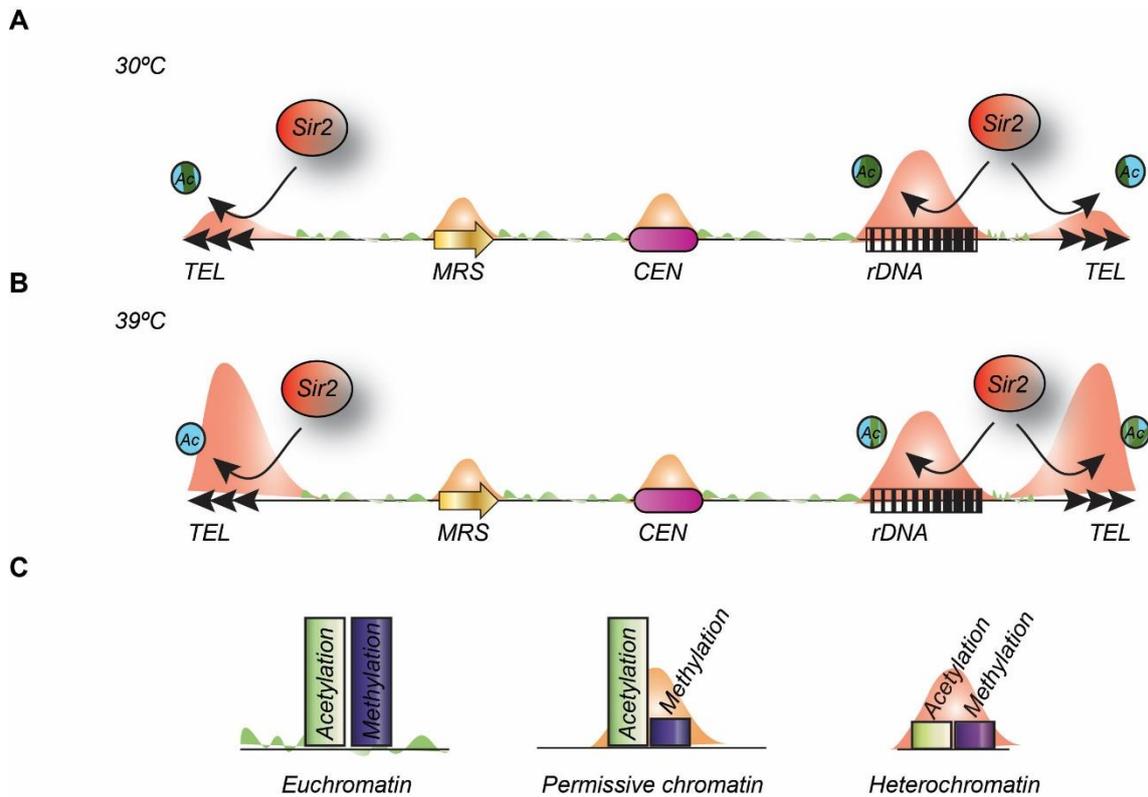
One of the most striking results is that telomeric heterochromatin is plastic. Thermal increase from 30°C to 39°C, mimicking fever within the host, led to an increase of transcriptional silencing of a *URA3<sup>+</sup>* marker gene inserted at telomeres. Interestingly, this effect seems to be specific to telomeric regions as no effect was observed at the *rDNA* locus, MRS or centromeres. Moreover, other stress conditions such as fluconazole or H<sub>2</sub>O<sub>2</sub> did not have any silencing effect at any of the loci tested. Plastic heterochromatin is also observed in *S. cerevisiae*, where increased thermal shift reduces silencing observed at *rDNA* locus in favour of telomere silencing (Bi et al., 2004). A redistribution of Sir2 proteins could be responsible for this effect. *C. albicans* subtelomeres contain the *TLO* genes which act as transcriptional regulators (their characteristics and influence on subtelomeric regions will be discussed in depth later). Stronger transcriptional silencing at 39°C compared to 30°C could modulate expression of these transcriptional regulators which in turn could affect the transcriptional level of genes required for heat shock response and adaptability.

### **1.3. Intermediate heterochromatin at Centromere and MRS (Freire-Benítez et al., 2016a, 2016b).**

In contrast to the *rDNA* locus and telomeric regions that are assembled into classical heterochromatin, centromeres and *MRS* repeats are associated to an

intermediated chromatin state bearing features of both heterochromatin and euchromatin. A *URA3*<sup>+</sup> marker gene inserted at these positions is transcriptionally silenced to a lesser extent than a *URA3*<sup>+</sup> endogenous locus. Lower *URA3*<sup>+</sup> transcript levels are detectable only by RT-qPCR. However, silencing assays revealed that a strain with an inserted *URA3*<sup>+</sup> marker gene at *MRS* or *CEN* pericentromeric repeats showed the same growth as a WT *URA3*<sup>+</sup> endogenous strain in media lacking uridine. This could be due to the fact that, a minimal amount of Ura3 proteins is sufficient to confer the Ura3<sup>+</sup> phenotype and therefore a weaker silencing is only detectable by a more sensitive assay such as RT-qPCR. RNA-sequencing analysis revealed that both, pericentromeric and *MRS* repeats were associated with high levels of H3K9ac and H4K6ac. However, peri-centromeres and *MRS* are associated with low H3K4me levels. It is possible that low methylation level is an epigenetic mark required for preventing excessive recombination at *MRS* and pericentromeric regions. Low methylation could be the epigenetic mark required to anchor protein factors that inhibit recombination and promote genome stability at these locations. In support of this hypothesis, it has been observed that *MRS* are hotspots of mitotic recombination (Iwaguchi et al., 2004, 2008; Lephart et al., 2005; Magee et al., 2008).

In summary, we have established the chromatin state associated with DNA repeats at *C. albicans* (**figure 5.1**). While classic heterochromatin is assembled at the *rDNA* locus and telomeres, an intermediate state of the chromatin is associated with centromeres and *MRS*. In addition, heterochromatin at telomeres is plastic under increasing temperatures (**figure 5.1**).



**Figure 5.1. Model of the heterochromatin status at *C. albicans* DNA repeats.**

(A) The *rDNA* locus and telomere repeats are associated with classic assembly of heterochromatin refractory to transcription. Heterochromatin regions are associated with low levels of acetylation and methylation. This heterochromatin status is dependent on the HDAC Sir2. Sir2 effect spans from telomere repeats to subtelomeric regions, controlling the transcriptional profiles of genes embedded within these regions. *MRS* and centromeres display a permissive chromatin status associated with low levels of methylation but high levels of acetylation. (B) Heterochromatin at telomeres is increased when temperature is switched from 30°C to 39°C, mimicking fever within the host. (C) Heterochromatin and permissive chromatin are repressive states of transcription in opposition to euchromatin and active transcription.

## **2. Function of heterochromatin in promoting genome stability (Freire-Benítez et al., 2016c)**

Sir2-dependent heterochromatin is assembled at *C. albicans* *rDNA* and telomeres (Freire-Benítez et al., 2016a). One of the main functions of heterochromatin is inhibition of recombination at repeats which are potential hotspots for recombination (Grewal and Jia, 2007). Loss of heterozygosity (LOH) is one of the mechanisms driving diversity at *C. albicans* (Rosenberg, 2011). LOH ensures variability as this organism cannot undergo meiosis. LOH after whole chromosome aneuploidy and long or short gene conversions is observed when *C. albicans* is grown under stress conditions. The selection of a recombination mechanism is stress-dependent. Fluconazole and high temperature promote whole chromosome aneuploidies and long tract recombination events, while H<sub>2</sub>O<sub>2</sub> promotes short track gene conversions (Forche et al., 2011). We investigated how Sir2 inhibited recombination at subtelomeres and at the *rDNA* locus. My results show that Sir2 controls genome stability at subtelomeric regions but has lost this ability at the *rDNA* locus. Stress conditions increase recombination rates independently of SIR2 control of recombination.

### **2.1. Heterochromatin does not promote genome stability at the *rDNA* locus**

*C. albicans* *rDNA* locus is assembled in Sir2-dependent heterochromatin (Freire-Benítez et al., 2016a). In order to test if Sir2 promotes *rDNA* stability we deleted both copies of *SIR2* in a system that allowed us to measure LOH rates. We inserted the *URA3*<sup>+</sup> marker gene before, within and after the *rDNA*

locus. These experiments demonstrated that Sir2 does not control genome stability at the *rDNA* locus. This is of striking contrast to *S. cerevisiae* where deletion of *SIR2* increased recombination rates up to 20 fold at the *rDNA* locus (Gottlieb and Esposito, 1989a). Surprisingly, we observed that Csm1, a member of the Monopolin complex, controls *rDNA* genome stability. Deletion of *CSM1* led to a 140 fold increase in recombination rates. In *S. cerevisiae* Csm1 is member of the monopolin complex. The *S. cerevisiae* Monopolin complex, composed of the protein Csm1 and Lrs4, promotes *rDNA* stability by aligning sister chromatids, ensuring silencing, and mediating perinuclear anchoring. *CSM1* deletion acts in a parallel to *SIR2* deletion at *S. cerevisiae* where deletion of both genes results in up to 40 fold increase of recombination rates in opposition to 20 fold increase in a *sir2Δ/Δ* deletion mutant (Huang et al., 2006b). Our results showed that Csm1 alone inhibits recombination at *rDNA* in *C. albicans*. This could indicate that the parallel pathway of Sir2 controlling recombination at *rDNA* is lost in *C. albicans*. This is supported by our results showing that Csm1 is required for *rDNA* silencing at this locus, indicative of an independent *SIR2* pathway (**figure 5.2**). One question is how Csm1 inhibits recombination independently of Sir2.

At *S. cerevisiae* *rDNA* locus, Sir2 inhibits non-coding transcripts arising from a bidirectional promoter allowing cohesin ring loading. Cohesin is required for correct sister chromatid segregation promoting genome stability at the cluster (Kobayashi, 2011; Kobayashi and Ganley, 2005). In the absence of Sir2, transcription of non-coding transcripts impairs cohesin loading. Lack of cohesin molecules results in unfaithful recombination events within the *rDNA* units. This increases the formation of extra copies of *rDNA* and formation of

extrachromosomal *rDNA* circles (ERC) (Ganley and Kobayashi, 2013; Kobayashi, 2011; Kobayashi and Ganley, 2005). It is worth noting that, an amplification system similar to the one described for *S. cerevisiae* has not been described in *C. albicans* before. There is not known homolog to FOB1 in *C. albicans*. As described in the introduction, this gene encodes for a replication fork barrier protein that stops the replication machinery forming a DSBs that induces recombination for repair (Huang et al., 2006b; Kobayashi, 2011). At *C. albicans rDNA*, Csm1 could act simply to promote proper sister chromatid segregation helping to load cohesin independently of an amplification unit machinery, independently of Sir2.

## **2.2. Heterochromatin promotes genome stability at *C. albicans* subtelomeric genes**

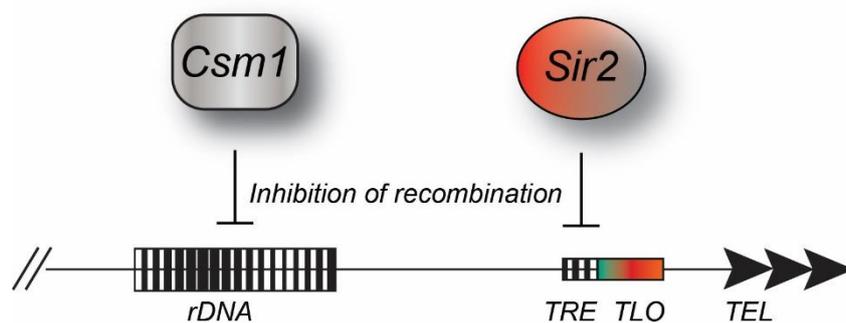
Subtelomeres are highly variable regions of the genome (Louis et al., 1994). In *C. albicans* a family of subtelomeric genes, the *TLO* gene family, is one example of this variability. *TLO* genes encode for Med2, a subunit of the Mediator complex which acts as a transcriptional coactivator in eukaryotes (Kelleher et al., 1990). There are 14 *TLO* genes in *C. albicans*, 2 in *C. dublinensis* and one copy in most of *Candida* species. *C. albicans* and *C. dublinensis* diverged at same evolutionary distance to the other *Candida* members (Haran et al., 2014; Jackson et al., 2009). Therefore, expansion of the *TLO* gene family is a relative recent event. Interestingly, *C. albicans* is more virulent than *C. dublinensis*, suggesting that *TLO* gene expansion might be responsible for increasing virulence.

*TLO* gene family is not the only family of genes which have undergone a gene copy number expansion in *C. albicans*. Other examples include agglutinin-like sequence genes (*ALS*). Products of these genes promote adhesion between cells and host tissue invasion. Other examples are gene families coding for aspartyl proteases and lipases which promote tissue invasion or ferric reductases which promote iron acquisition within the host (Berman, 2012; Berman and Sudbery, 2002). Expansion of subtelomeric genes is observed at other organisms such as *S. cerevisiae* where subtelomeric regions harbor multicopy genes involved in carbon source metabolism (Louis, 1995). Gene expansion at subtelomeres is also observed in Plasmodium parasites where *VAR* genes located at subtelomeres on different chromosomes undergo recombination at frequencies much higher than that expected for ectopic recombination (Freitas-Junior et al., 2000). This suggests that telomere variability is important for adaptation to new environments (Carreto et al., 2008). As mentioned before, the Mediator complex is required for Pol II transcription (Kelleher et al., 1990). *C. albicans* is the most virulent and successful pathogen of the Candida clade. The *TLO* gene expansion observed in *C. albicans* but not in other candida members could be one of the determinants for this. Recombination between different *TLO* genes could lead to different protein products which differentially modulate the Med2-dependent transcription of virulence and/or adaptation genes. In *C. albicans*, telomeric and subtelomeric regions are assembled into Sir2-dependent heterochromatin (Freire-Benítez et al., 2016a). Importantly, we have shown that Sir2 represses recombination of a subset of *TLO* genes containing a 300 bp DNA element, which we named *TRE* and contains the BTS sequence. Translocation events associated with the BTS

at *TLO* sharing homology were observed before in long term evolution experiments (Anderson et al., 2015) (**figure 5.2**).

One question to address is how *TRE* promotes recombination and how Sir2 represses it. Our results showed that the *TRE* is a hotspot for recombination events. This could be facilitated by the expansion of *TLO* genes. Multiple copies of *TLO* genes with homology at the *TRE* may increase the frequency of recombination events among *TLO* at different chromosomes. Sir2 binds the *TRE* sequence independently of heterochromatin formation, as our results showed that marker genes inserted at this position are not silenced. Sir2 could tether other molecules, such as cohesin to the *TRE* region, thus preventing recombination among different *TLO* genes (**figure 5.2**).

Results at both *rDNA* and subtelomeric regions showed for the first time, dual function of Sir2: assembly of heterochromatin independently of inhibition of recombination at *rDNA* and contrarily inhibition of recombination without heterochromatin assembly at subtelomeric regions.



**Figure 5.2. LOH control at *rDNA* locus and *TLO* genes of *C. albicans*.** *Csm1* inhibits LOH recombination events and promotes *rDNA* stability, while *Sir2* inhibits LOH recombination at a 300 bp DNA region (*TRE*) located downstream of most *TLO* gene members promoting genome stability at subtelomeric regions.

### **2.3. Stress conditions promote genome instability independently of Sir2**

As described before stress conditions increased recombination rates at *C. albicans* genome (Forche et al., 2011). We investigated the effect of stress environments on Sir2 control of recombination. 39°C, H<sub>2</sub>O<sub>2</sub> and fluconazole treatments led to an increase in recombination rates, as expected. Interestingly, mutation rates at *TLO* in a *sir2Δ/Δ* null mutant were not increased in the presence of fluconazole and in the presence of H<sub>2</sub>O<sub>2</sub>. This is in opposition to what observed at 30°C standard growth conditions where *sir2Δ/Δ* null mutants showed an increase in recombination rates at *TRE*-containing *TLO* compared to WT cells. 39°C still showed Sir2-dependency to control LOH rates at *TRE*-containing *TLO* compared to that observed at 30°C. This shows that some environmental stress conditions increase recombination rates independently of Sir2 inhibition of recombination. High recombination rates overpassing inhibition of recombination mechanisms could lead to a prompt variability required to survive stress conditions. This is of crucial importance at *C. albicans* as this organism does not undergo meiosis. A large amount of genetic and phenotypic variation appears over very short periods of time during passage in a mammalian host where LOH and aneuploidies were observed in mouse blood infection model (Forche et al., 2005).

### **3. Future work**

Many questions about the chromatin status and genome stability at *C. albicans* still remain to be elucidated. For instance, the mechanism behind how Sir2 acts at *rDNA* locus and telomere repeats remains an intriguing question. We have

shown that Sir2 is the HDAC in charge of reducing H3K9ac and H4K16ac levels and Set1 is the HMT keeping H3K4me levels at *C. albicans* epigenome. But nothing is known about the counterparts of these proteins in *C. albicans*. Furthermore, much more needs to be revealed about the sophisticated Sir2-dependent mechanism controlling recombination rates at the downstream region of the *TLO* gene family. It is important to know why contrarily to what observed in other organisms, Sir2 barely affects recombination rates at the *rDNA* locus, where the Monopolin complex is fully in charge of this mission. Which proteins participate together with Sir2 in these processes or the existence of different specialized complexes remains to be elucidated.

Finally, *MRS* and centromeres have low levels of methylation, but the demethylase in charge of keeping this low methylation level remains to be discovered. This histone demethylase could be of special importance to maintain genome stability associated with pericentromeric regions and *MRS*. Moreover, the biological consequences of the assembly of permissive chromatin at *MRS* remains an interest subject of further studies, due in great part to the presence of virulent genes like the *FGR-6* genes located within the *MRS* repeats.

## **6. Materials and Methods**

Due to the brief description of methods presented in the published papers of this thesis, in this chapter techniques are described in more detail.

## 1. Microbiology techniques

### 1.1. Yeast media

All media was prepared using Ultra pure deionised water (dH<sub>2</sub>O) produced by Thermo Scientific Barnstead NanoPure Diamond system. Sterilisation of media, dH<sub>2</sub>O and other liquids was carried out at 121°C in a bench top Prestige Medical autoclave.

#### 1.1.1. YPAD: yeast extract, peptone, adenine dextrose media

Yeast extract (Melford)	10% w/v
Bacto peptone (Becton, Dickinson and company Sparks)	20% w/v
Agar (Melford) in solid media	20% w/v
Glucose (Fisher)	20% w/v
Adenine (Sigma)	0.1 mg/ml
Uridine (Sigma)	0.08 mg/ml

#### 1.1.2. SC/SC DROP-OUT –His/ -Arg/ -Uri media

Synthetic aminoacid complete media or synthetic complete media lacking the specific aminoacids respectively: histidine for DROP-OUT –His, arginine for DROP-OUT –Arg and uracil for DROP-OUT –Uri.

SC complete drop-out (Formedium™)	0.2 % w/v
SC drop-out: histidine (Formedium™)	0.2 % w/v
SC drop-out: arginine (Formedium™)	0.2 % w/v
SC drop-out: uracil (Formedium™)	0.2 % w/v
Yeast nitrogen base without aminoacids (Fisher)	0.67 % w/v
Agar (Melford) in solid media	2 % w/v
Glucose (Fisher)	2 % w/v
Uridine (Sigma) (Absent in DROP-OUT – Uri)	0.08 mg/ml

### 1.1.3. LB media

Bacto Tryptone (Becton, Dickinson and company Sparks)	1 % w/v
NaCl (Melford)	1 % w/v
Yeast extract (Melford)	0.5 % w/v
Agar (Melford) in solid media	2 % w/v

## **1.2. Yeast growth**

Yeast cells were cultured in rich medium (YPAD) containing extra adenine (0.1mg/ml) and extra uridine (0.08mg/ml) complete SC medium (Formedium™) or SC Drop-Out media (Formedium™). When indicated, media were supplemented with 5-Fluorotic acid (5-FOA, Melford) at a concentration of 1 mg/ml, Nourseothricin (clonNAT, Melford) at a concentration of 100 µg/ml, fluconazole (SIGMA) at a concentration of 1 µg/ml or 0.5 mM H<sub>2</sub>O<sub>2</sub>. Cells were grown at 30°C or 39°C as indicated.

## **1.3. Yeast strain construction**

Strains, oligonucleotides and plasmids are listed in each corresponding chapter. Integration and deletion of genes was performed as previously described (Wilson et al., 1999) using long oligos-mediated PCR for gene deletion and gene tagging. Long oligos share around 75 nucleotides of homology with previous and rear sequence flanking the target gene and 20 nucleotides of homology to a plasmid containing a marker gene for substitution. In case of gene tagging, a plasmid containing a HA epitope tagged and a marker gene was used. PCR using long oligos was PCR BIO HiFi Polymerase (PCR BIOSYSTEMS) using the next reaction: 10 µl of 10X PCR BIO reaction buffer containing MgCl<sub>2</sub>, 2 µl of 10 µM of primer forward, 2 µl of 10 µM of primer reverse, 0.5 µl of 10 mM dNTPs, 500 ng of DNA, 0.5 µl (1 unit) of PCR BIO HiFi polymerase, up to 50 µl dH<sub>2</sub>O. PCR conditions were performed as follow: initial denaturation at 95°C for 1 min, 30 cycles each of denaturation at 95 °C for 15 s, annealing at 58°C for 15 s, and extension at 72°C for 1 min, and a final extension at 72°C for 7 min. 5 µl of PCR product were checked in a 1 %

agarose gel (Melford). Gels were stained with ethidium bromide (Sigma) (3  $\mu$ g in 100 ml) and visualized with UV light and photographed in a Syngene G-box apparatus.

Transformation was performed by electroporation described before (De Backer et al., 1999). Overnight cultures were grown to stationary phase in YPAD at 30°C and diluted into 100 ml of fresh YPAD media. Cultures were grown until OD<sub>600nm</sub> of 1.4. Cells were collected by centrifugation at 4000 rpm for 2 minutes. Pellets were washed with sterile dH<sub>2</sub>O and collected again by centrifugation. Cells were incubated in rotation at room temperature with 25 ml TELiDTT buffer [(10mM Tris-HCl (Melford), pH 8, 1mM EDTA (Fisher), 100 mM LiAc (Sigma), 10mM DTT (Sigma) fresh added] during 1 hour. Cells were pelleted at 4000 rpm for 2 minutes at 4°C, and washed with ice cold sterile dH<sub>2</sub>O. After cell collection by 4°C centrifugation step, cells were washed with ice cold sterile 1 M sorbitol (Sigma) and collected again at 4°C as previous steps. Cells were resuspended in 100  $\mu$ l ice cold sterile 1 M sorbitol. 5  $\mu$ l of a PCR product with the amplified fragment to transform were incubated with 40  $\mu$ l of ice cold cells resuspended in 1 M sorbitol during 10 minutes on ice and transferred into a transformation cuvette (0.2 cm Gene Pulser/Micro Pulser Electroporation Cuvettes Bio-Rad). Electroporation was performed by electroporator (Gene Pulser TM, Bio-Rad). The cuvette was placed in the chamber and an electric pulse of 1.5 kV, 25  $\mu$ F, 201  $\Omega$  was applied. After that, cuvettes were quickly placed on ice and 1ml of ice cold sterile 1 M Sorbitol was added. Suspension was plated in the appropriate selective media based on auxotrophic selection: SC - His, SC - Arg, SC - Uri or YPAD containing Nourseothricin (clonNAT, Melford) at a concentration of 100  $\mu$ g/ml.

#### **1.4. Silencing assay**

Growth analyses were performed using a plate reader (SpectrostarNano, BMG labtech) in 24 well or 96 well plate format at 30 °C. When indicated, silencing assays were performed in the presence of 200 ng/μl of fluconazole (Sigma), 1 mM H<sub>2</sub>O<sub>2</sub> (Sigma) and 39 °C. For each silencing assay in a 24 well plate format, 1 ml of a starting culture was inoculated in SC or SC-Uri media to reach a concentration of 60 cells/μl. Growth was assessed by measuring A600, using the following conditions: OD600 nm, 3600 s cycle time, 30 flashes per well, 400 rpm shaking frequency, double orbital shaking mode, 850 s additional shaking time after each cycle, 0.5 s post delay, for 44 to 60 hours at 30 °C. For each silencing assay in 96 well plate format, 1:100 dilution of an starting culture was inoculated in a final volume of 95 μl of SC or SC-Uri media to reach a concentration of 60 cells/μl. Growth was assessed by measuring A600, using the following conditions: OD600 nm, 616 cycle time, 3 flashes per well, 700 rpm shaking frequency, orbital shaking mode, 545 s additional shaking time after each cycle 0.5 s post delay, for 44 hours. Graphs represent data from three biological replicates. Error bars: standard deviations of three biological replicates. Data was processed using SpectrostarNano MARS software and Microsoft Excel.

## **2. Molecular biology techniques**

### **2.1. Genomic DNA extraction**

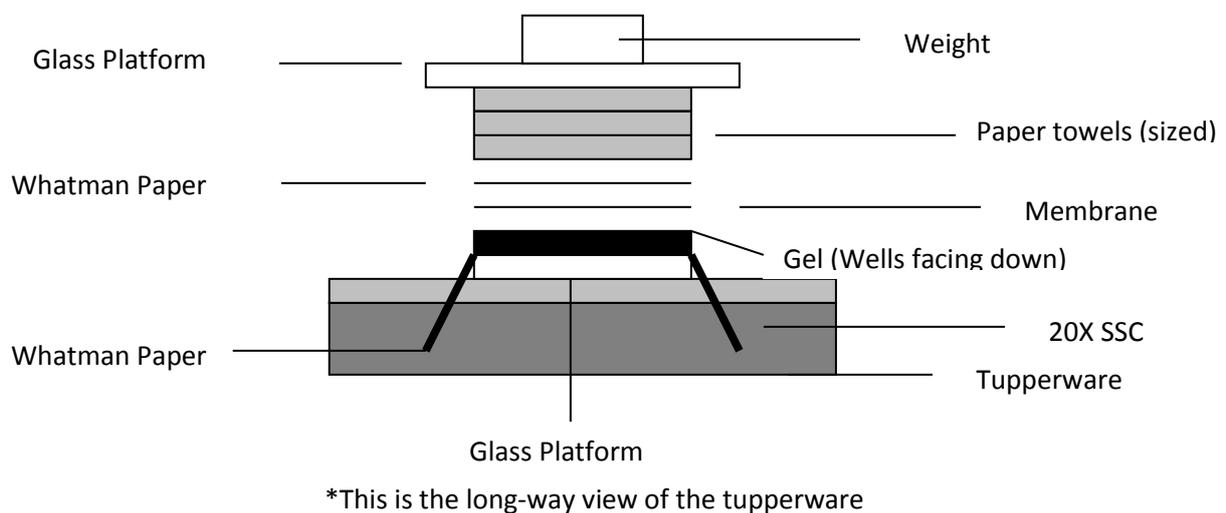
Strains were grown to stationary phase in YPAD media containing extra adenine (0.1 mg/ml) and extra uridine (0.08 mg/ml). Cells were collected by centrifugation at 2000 rpm at room temperature. Cell pellets were washed with

sterile dH<sub>2</sub>O and resuspended in lysis buffer [10% SDS (Sigma), 2% Triton X-100 (Fisher), 1 M NaCl (Melford), 10 mM Tris pH 8 (Melford), 1 mM EDTA (Fisher)] containing 0.4 grams of acid glass beads (Sigma). After short vortex, 200 µl of phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma) were added. Samples were vortex at 4°C during 30 minutes followed by centrifugation at 13000 rpm 5 minutes. The aqueous layer was transferred to a new tube and an equal amount of chloroform (Sigma) was added. After short vortex time, samples were centrifuged at 13000 rpm 5 minutes. Pellets were precipitated with ethanol (Fisher) at -20°C during 1 hour followed by centrifugation at 13000 rpm 5 minutes. Pellets were treated with 3 µg of RNase A (Fisher) in 400 µl of 1X TE [(10mM Tris-HCl (Melford), pH 8, 1 mM EDTA (Fisher))] at 37°C during 30 minutes followed by DNA precipitation with 1 ml 95% Ethanol (Fisher) and 50 µM sodium acetate (Sigma) at -20°C during 30 minutes. Finally genomic DNA was collected by centrifugation at 13000 rpm 5 minutes. Pellets were air-dry and resuspended in 40 µl of dH<sub>2</sub>O water.

## **2.2. Southern blot**

Genomic DNA was digested with corresponding enzymes and run in 1% agarose (Sigma) gel. After that, gel was treated with 2 volumes of 0.25 M HCl (Fisher) for 15 minutes and denatured with 1.5M NaCl (Melford), 0.5M NaOH (Fisher) for 30 minutes. Following washes with dH<sub>2</sub>O, the gel was neutralized with 1M Ammonium Acetate (Melford) 0.02 M NaOH (Fisher) for 40 minutes. The gel was rinsed in 20X SSC [3 M NaCl (Melford), 0.3 M Sodium Citrate (Melford)] for 10 minutes. A nylon membrane (Zeta probe membranes, Bio-Rad), was moistened with dH<sub>2</sub>O and rinsed with 20X SSC. The gel was flipped and assembled into capillarity transfer apparatus. Two bridge size and two gel

size 20X SSC wet Whatman papers were in contact to 20X SSC buffer for capillarity transfer. Gel was flipped and situated over wet Whatman paper. 20X SSC moistened membrane was situated on the gel and bubbles removed. To perform capillarity force, two gel-sized dry Whatman papers were situated on the membrane together with dry paper towel and a weight (**figure 6.1**). Capillarity was performed overnight. Membranes were cross-linked at 80°C during 2 hours and stored at room temperature.



**Figure 6.1 Southern blot apparatus.** DNA was transferred from DNA agarose gel to a nylon membrane by capillarity transfer.

### 2.3. Southern probe hybridization

Cross-linked nylon membranes were probed with DIG probes (Roche) following manufacturer's instructions. In a classical PCR reaction, DIG-dUTP is incorporated instead of dTTP while a specific region of the DNA is amplified as a probe. Hybridization was performed as described before by Ketel et al., 2009. PCR amplification was performed as follows using VWR Taq polymerase: 1 µg

genomic DNA, 2  $\mu$ l of 10  $\mu$ M of primer forward, 2  $\mu$ l of 10  $\mu$ M of primer reverse, 5  $\mu$ l of 10X VWR buffer containing  $MgCl_2$ , 2.5 $\mu$ l of 2.5mM dNTPs or DIG dNTP mix (2mM dCTP, 2mM dGTP, 2mM dATP, 1.3 mM dTTP, 0.7mM DIG-11-dUTP) and 0.5  $\mu$ l of VWR Taq polymerase. PCR conditions were performed as follow: initial denaturation at 94°C for 5 min, 30 cycles each of denaturation at 94 °C for 30 s, annealing at 58°C for 45 s, and extension at 72°C for 1 min, and a final extension at 72°C for 5 min. 3  $\mu$ l of PCR product was checked in a 1% agarose gel (Melford). Gels were stained with ethidium bromide (Sigma) (3  $\mu$ g in 100 ml) and visualized with UV light and photographed in a Syngene G-box apparatus.

The cross-linked membrane containing digested DNA was washed with 2X SSC 0.1% SDS 15 minutes followed by a second wash of 0.05x SSC 0.1% SDS at 65°C. Membrane was blocked 1x blocking solution (Roche) in maleic acid buffer [100 mM maleic acid (Melford), 150 mM NaCl (Melford), pH = 7.5] during 3 hours in gentle agitation. After that, blocking solution was removed and the DIG-PCR probe was added to new 1x blocking solution (Roche) in maleic acid buffer and pour onto the membrane with gentle agitation during 1 hour. After hybridization, membrane was washed with wash buffer [maleic acid buffer, 0.3% Tween 20 (Sigma)] during 30 minutes in gentle agitation. Wash buffer was removed and replaced by detection buffer [10mM Tris base (Sigma), 100mM NaCl, pH = 9.5] for 12 minutes in gentle agitation. CDP- Star drops (Roche) were added onto the membrane for the luminescence detection. In dark room a sheet of film (Hyperfilm ECL, Amersham) was exposed to the membrane for varying lengths of time depending on signal intensity, ranging from a few seconds to 30 minutes. After exposing the film, it was developed using a Compact X4 Automatic Processor (Xograph Healthcare).

## **2.4. RNA extraction and cDNA synthesis**

RNA was extracted from 2-5 ml of Log<sub>2</sub> exponential cultures (OD<sub>600nm</sub> = 1.4) using a yeast RNA extraction kit (E.Z.N.A.® Isolation Kit RNA Yeast, Omega Bio-Tek) following manufacturer's instructions. RNA quality was checked by electrophoresis under denaturing conditions in 1% agarose, 1X HEPES (Melford), 6% Formaldehyde (Sigma). RNA concentration was measured using a NanoDrop ND-1000 Spectrophotometer getting yields around 1 µg/µl. cDNA synthesis was performed using iScript™ Reverse Transcription Supermix for RT-qPCR (Bio-Rad) and a Bio-Rad CFXConnect™ Real-Time System as follows: 500 ng RNA, 2 µl of 5x iScript reaction mix, 0.5 µl iScript reverse transcriptase, dH<sub>2</sub>O up to 10 µl. PCR conditions were as follows: 42°C for 30 min, 85°C for 10 min, cold down at 12°C.

## **2.5. Quantitative Chromatin Immunoprecipitation (qChIP).**

qChIP was performed as described before by Pidoux et al., 2004 as follows: 5 ml of an overnight culture grown in YPAD with extra uridine (0.08 mg/ml), diluted into fresh YPAD with extra uridine (0.08 mg/ml) and grown until OD<sub>600 nm</sub> of 1.4. Cells (50 ml/sample) were fixed with 1% Paraformaldehyde (Sigma) for 10 min at room temperature. Cells were pelleted at 4000 rpm at 4°C, followed by two washes of ice cold sterile 1x PBS (Sigma). Pellets were collected at 4000 rpm at 4°C and store at -80°C. Cells were lysed using acid-washed glass beads (Sigma) and lysis buffer [50mM HEPES-KOH (Sigma) pH7.5, 140mM NaCl (Melford), 1mM EDTA (Fisher), 1% Triton X-100 (Sigma), 0.1% Sodium deoxycholate (Sigma)] and a Disruptor genie™ (Scientific Industries) for 30 min at 4 °C. Chromatin was sheared to 500–1000 bp using a

Bioruptor (Diagenode) for a total of 20 min (30 s ON and OFF cycle) at 4°C. Cells were pelleted at 13000 rpm at 4°C twice. Immunoprecipitation was performed overnight at 4 °C using 300 µl of sample and 2 µL of antibody anti-H3K4me2 (Active Motif- Cat Number: 39141), or anti-H3K9ac (Active Motif- Cat Number: 39137), or anti-H4K16ac (Active Motif- Cat Number: 39167) and 25 µl of Protein G magnetic beads (Dynal - InVitrogen). By using a magnetic rack, the liquid phase was removed and beads were washed at 4°C as follows: 1 min in 1ml of lysis buffer [50mM HEPES-KOH (Sigma) pH 7.5, 250 mM NaCl (Melford), 1 mM EDTA (Fisher), 1% Triton X-100 (Sigma), 0.1% Sodium deoxycholate (Sigma)], 10 min with 1 ml of 0.5 M NaCl lysis buffer [50 mM HEPES-KOH (Sigma) pH7.5, 500mM NaCl (Melford), 1mM EDTA (Fisher), 1% Triton X-100 (Sigma), 0.1% Sodium deoxycholate (Sigma)], 10 min in 1 ml of wash buffer [10 mM Tris-HCl (Melford), pH 8, 0.25 M LiCl (Sigma), 0.5% NP-40 (Melford), 0.5% Sodium deoxycholate (Sigma), 1 mM EDTA (Fisher)], 1 min in 1X TE [(10mM Tris-HCl (Melford), pH 8, 1mM EDTA (Fisher)]. DNA from immunoprecipitate and input samples was eluted with a 10% slurry of Chelex 100-resin (Bio-Rad) and boiled at 100°C during 10 min. Following that, beads were treated with 25 µg of proteinase K (Sigma) during 30 min at 55°C. After that, samples were boiled at 100°C during 10 min. Samples were centrifuged at room temperature at 4000 rpm during 3 min twice, supernatants were recovered avoiding any residual Chelex 100-resin in the final samples. For immunoprecipitated samples 1:10 dilution was used for analysis, while 1:1000 was used for input control samples for qPCR reactions explained in section **2.7**.

## **2.6. Positive yeast transformants PCR**

Positive colonies were checked by PCR in a final volume of 15 µl using Taq DNA polymerase (VWR, 733-1364) as follows: 0.6 µl of 10 µM primer forward, 0.6 µl of 10 µM primer reverse, 0.5 µl of 10 mM dTNPs, 1.5 µl 10X VWR buffer containing MgCl<sub>2</sub>, 0.13 µl VWR Taq DNA polymerase, 1 µl DNA and dH<sub>2</sub>O up to 15 µl. DNA was extracted from single colonies following NaOH heat extraction: 10 min boiling step in 0.02M NaOH followed by 10 min on ice. PCR conditions were performed as follow: initial denaturation at 94°C for 7 min, 30 cycles each of denaturation at 94 °C for 45 s, annealing at 55°C for 1 min, and extension at 72°C for 1 min, and a final extension at 72°C for 7 min. PCR results were analysed in 1% agarose (Melford) electrophoresis gels. Gels were stained with ethidium bromide (Sigma) (3 µg in 100 ml) and visualized with UV light and photographed in a Syngene G-box apparatus.

## **2.7. qPCR reactions**

Real-time qPCR and RT-qPCR was performed in the presence of SYBR Green (Bio-Rad) on a Bio-Rad CFXConnect™ Real-Time System. qPCR reaction was performed as follows: 2 µl of DNA for ChIP reactions of 1:10 dilution Immunoprecipitate and 1:1000 matching input sample, and 2 µl of 500 ng cDNA for RT-qPCR, 5 µl of iTaq™ Universal SYBR® Green (BioRad), 0.5 µl of 10 µM of primer forward, 0.5 µl of 10 µM of primer reverse and 2 µl of dH<sub>2</sub>O up to a total volume of 10 µl. For both qPCR and RT-qPCR conditions were as follows: initial denaturation at 95°C for 2 min, 45 cycles each of denaturation at 95 °C for 20 s, annealing at 55°C for 20 s, extension at 70°C for 20 s, and melt curve at 55°C to 95°C, with 0.5 °C increment for 0.05 s. Histograms represent data from

three biological replicates. Data were analysed with Bio-Rad CFX Manager 3.1 software and Microsoft Excel. Enrichments were calculated as the percentage ratio of specific IP over input for qChIP analysis and as enrichment over actin for RT-qPCR using Microsoft excel. Error bars: standard deviation of 3 biological replicates generated from 3 independent cultures of the same strain.

## **2.8. Fluctuation analysis**

Strains were first streaked on –Uri media to ensure the selection of cells carrying the *URA3*<sup>+</sup> marker gene. 15 parallel liquid cultures were pre-grown overnight from independent single colonies. Each culture was diluted in YPAD at a concentration of 100 cells/μl and grown for 9 generations (18 hours). When indicated, media were supplemented with fluconazole (SIGMA) at a concentration of 1 μg/ml or 0.5 mM H<sub>2</sub>O<sub>2</sub>. During thermic stress, cultures were grown at 39°C. Every culture was plated on SC plates containing 1 mg/ml FOA (5-Fluorotic acid, Sigma) at a cell density depending on strain (ranging from 10<sup>4</sup> to 10<sup>7</sup> cells/plate) and on non-selective SC plates at a cell density of 100 cells/plate and grown at 30°C. Colonies were counted after 2 days of growth and data were analysed using FALCOR (Fluctuation Analysis Calculator) software (Hall et al., 2009) based on the Lea-Coulson analysis of the median (Lea and Coulson, 1949). Statistical differences between samples were calculated using Kruskal-Wallis test (Kruskal and Wallis, 2012). Statistical analysis and violin plots were generated using R (<http://www.r-project.org/>). To quantify the number of colonies able to grow on FOA due to *URA3*<sup>+</sup> silencing, between 100 and 200 FOA resistant colonies were replica-plated after fluctuation analysis on complete SC plates for recovery. Between 100 and 200 single colonies were then replica-plated on SC plates supplemented with 1

mg/ml FOA and –Uri SC Drop-Out plates. After 24 hour growth single colonies were counted and analysed using Microsoft Excel. To analyse the presence of the SAT1 marker gene after fluctuation analysis, between 100 and 200 FOA resistant single colonies were replica-plated on complete SC plates for recovery. Single colonies were then replica-plated on YPAD plates containing extra adenine (0.1mg/ml), extra uridine (0.08mg/ml) and Nourseothricin (clonNAT, Melford) at a concentration of 100 µg/ml. After 24 hour colonies were counted and analysed using Microsoft Excel.

### **2.9. Marker Gene Loss Assay**

Strains were first streaked in –Uri media to select for cells carrying the *URA3* marker gene. 15 parallel liquid cultures were pre-grown overnight from independent single colonies. Each culture was diluted in YPAD at a concentration of 100 cells/µl and grown for 9 generations (18 hours). Cells were plated on SC plates containing 1 mg/ml 5-FOA (5-Fluorotic acid, Sigma) and on non-selective SC plates and grown at 30°C. To distinguish between silencing and loss of the *URA3*<sup>+</sup> marker gene, around 500 FOA resistant single colonies were streaked onto complete SC plates for recovery and then streaked onto –Uri SC Drop-Out plates. Colonies not able to grow on –Uri plates but FOA resistant were counted. Data were analysed using Microsoft Excel. Statistical differences between samples were tested using unpaired T-test using R (<http://www.r-project.org/>).

### **2.10. SNP-RFLP analysis**

PCR primers and restriction enzymes were chosen according to Forche et al., 2009 and primers used are described in **chapter 4** (Freire-Benítez et al.,

2016c). PCR in a final volume of 15  $\mu$ l using Taq DNA polymerase (VWR, 733-1364) was performed as follows: 0.6  $\mu$ l of 10  $\mu$ M primer forward, 0.6  $\mu$ l of 10  $\mu$ M primer reverse, 0.5  $\mu$ l of 10 mM dTNPs, 1.5  $\mu$ l 10X VWR buffer containing  $MgCl_2$ , 0.13  $\mu$ l VWR Taq DNA polymerase, 1  $\mu$ l DNA and  $dH_2O$  up to 15  $\mu$ l. DNA was extracted from single colonies following NaOH heat extraction: 10 min boiling step in 0.02M NaOH followed by 10 min on ice. PCR conditions were performed as follows: initial denaturation at 94°C for 7 min, 30 cycles each of denaturation at 94 °C for 45 s, annealing at 55°C for 1 min, and extension at 72°C for 1 min, and a final extension at 72°C for 7 min. PCR results were analysed in 1% agarose (Melford) electrophoresis gels. Gels were stained with ethidium bromide (Sigma) (3  $\mu$ g in 100 ml) and visualized with UV light and photographed in a Syngene G-box apparatus. Each PCR product was digested overnight with the corresponding restriction enzyme, AseI for SNP 1, TaqI and for SNP 85 (Forche et al., 2009). Enzymatic reactions were performed in a total volume of 15  $\mu$ l with 1  $\mu$ l RE, 1.5  $\mu$ l of 10 X restriction buffer, 5  $\mu$ l of SNP amplified PCR and  $dH_2O$  up to 15  $\mu$ l. The digested PCR product was run on a 3% agarose gel (Melford) along with an undigested control PCR sample. Gels were stained with ethidium bromide visualized with UV light and photographed in a Syngene G-box apparatus. Genotypes were assigned based on banding patterns for each SNP marker as described (Forche et al., 2009). Every PCR product digested gives a pattern of two bands of different size if the allele is heterozygous, one band if the allele is homozygous. The schematics of expected bands for each enzyme are described in **Chapter 4 supplementary figures 6 and 8**.

## 2.11. TRE plasmid construction

pTRE-URA3 was constructed using plasmid pGEMURA3 (Wilson et al., 1999). The *TRE* (*TLO* Recombination Element) sequence located at 3' of *TLO $\alpha$ 10* was PCR amplified from *C. albicans* genomic DNA using oligos containing the restriction sites *SacII* (Promega) (**Supplementary Table S2, chapter 4**) (Freire-Benítez et al., 2016c). PCR set up was as follows: 10  $\mu$ l of 10X PCRBIO reaction buffer containing  $MgCl_2$ , 2  $\mu$ l of 10  $\mu$ M of primer forward, 2  $\mu$ l of 10  $\mu$ M of primer reverse, 0.5  $\mu$ l of 10 mM dNTPs, 500 ng of DNA, 0.5  $\mu$ l (1 unit) of PCRBIO HiFi polymerase (PCRBIO SYSTEMS) and  $dH_2O$  up to 50  $\mu$ l. PCR conditions were performed as follows: initial denaturation at 95°C for 1 min, 30 cycles each of denaturation at 95 °C for 15 s, annealing at 58°C for 15 s, and extension at 72°C for 1 min, and a final extension at 72°C for 7 min. PCR product was purified using E.Z.N.A.® Cycle-Pure Kit following manufacturer's instructions. PCR purified *TRE* flanked by *SacII* restriction sites product was digested as follows: 10  $\mu$ l of purified PCR, 3  $\mu$ l of *SacII* (Promega) and 6  $\mu$ l of restriction enzyme buffer,  $dH_2O$  up to 60  $\mu$ l. pGEMURA3 (Wilson et al., 1999) plasmid was digested as follows: 3  $\mu$ l of pGEMURA3 plasmid, 3  $\mu$ l *SacII* (Promega) and 6  $\mu$ l of restriction enzyme buffer,  $dH_2O$  up to 60  $\mu$ l. Both digestions were incubated at 37°C overnight and after that purified using E.Z.N.A.® Cycle-Pure Kit following manufacturer's instructions. Ligation of both products was performed using DNA ligation kit (Roche) as follows: 0.5  $\mu$ l of digested and purified pGEMURA3, 10  $\mu$ l of digested and purified PCR product, 2  $\mu$ l of 10 X ligase buffer, 1  $\mu$ l of ligase and  $dH_2O$  up to 20  $\mu$ l. Ligation was incubated at 25°C overnight. Ligated products were used to transform competent *DH5- $\alpha$*  *Escherichia coli* cells as follows: 10  $\mu$ l of ligation products and

40  $\mu$ l *DH5- $\alpha$  Escherichia coli* competent cells were incubated 30 min on ice, 1 min at 42°C and 1 min on ice. After this, cells were recovered in 300  $\mu$ l of LB medium 45 min at 37°C in agitation. Following this, cells were plated on LB plates containing 100  $\mu$ g/ml Ampicillin (Sigma) and grown at 37°C overnight. Positive transforming colonies were confirmed by colony PCR as follows: single colonies were touched with a yellow tip and mix in an ice cold PCR reaction containing 0.6  $\mu$ l of 10  $\mu$ M primer forward, 0.6  $\mu$ l of 10  $\mu$ M primer reverse, 0.5  $\mu$ l of 10 mM dTNPs, 1.5  $\mu$ l 10 X VWR buffer containing MgCl<sub>2</sub>, 0.13  $\mu$ l VWR Taq DNA polymerase and dH<sub>2</sub>O up to 15  $\mu$ l. PCR conditions were performed as follows: initial denaturation at 94°C for 7 min, 30 cycles each of denaturation at 94 °C for 45 s, annealing at 55°C for 1 min, and extension at 72°C for 1 min, and a final extension at 72°C for 7 min. PCR results were analysed in 1% agarose (Melford) electrophoresis gels. Gels were stained with ethidium bromide (Sigma) (3  $\mu$ g in 100 ml) and visualized with UV light and photographed in a Syngene G-box apparatus. Plasmid extraction of positive PCR transformants was performed using E.Z.N.A.® Plasmid Mini Extraction Kit following manufacturer's instructions. Plasmids were confirmed by SacII (Promega) digestion as follows: 1  $\mu$ l of plasmid, 2  $\mu$ l of restriction enzyme, 2  $\mu$ l of 10 X restriction enzyme buffer and dH<sub>2</sub>O up to 20  $\mu$ l and incubated at 37°C overnight. Products were visualized in 1% agarose (Melford) electrophoresis gels. Gels were stained with ethidium bromide (Sigma) (3  $\mu$ g in 100 ml) visualized with UV light and photographed in a Syngene G-box apparatus. Finally, correct plasmids were sequenced with primers listed in **Supplementary Table S2 (chapter 4)** as indicated by GATC Biotech facility.

## 2.12. Western blot detection

Yeast extracts were prepared as described by von der Haar, 2007 using  $1 \times 10^8$  cells from overnight cultures grown to a final OD<sub>600</sub> of 1.5–2. Cells were resuspended in 200  $\mu$ l of lysis buffer [0.1 M NaOH (Fisher), 0.05 M EDTA (Sigma), 2% SDS (Sigma), 2%  $\beta$ -mercaptoethanol (Sigma)] and incubated at 90°C during 10 minutes. After that 5  $\mu$ l of 4M acetic acid (Fisher) were added and samples were vortex and incubated 10 minutes at 90°C. After that, 50  $\mu$ l of loading buffer were added [0.25 M Tris-HCL (Melford) pH = 6.8, 50% glycerol (Fisher), 0.05 % bromophenol blue (Sigma)]. Samples were centrifuged at 13000 rpm for 10 minutes and loaded in 10 % SDS polyacrylamide gel for electrophoresis (SDS-PAGE). Premade gel cassettes (Invitrogen) were used for electrophoresis process. First, 8 ml of the resolving layer [2ml of Acrylamide 40% (BioRad), 2 ml of 1.5 M Tris (Melford) pH 8.8, 4 ml of dH<sub>2</sub>O, 5  $\mu$ l of TEMED (Sigma) and 35  $\mu$ l of 10 % APS (Fisher)] and 1 ml of isopropanol (Fisher) was pipetted on top of resolving layer to flatten the surface. Gel was set during 1 hour. After that the isopropanol was poured off and the stacking layer [375  $\mu$ l of Acrylamide 40% (BioRad), 750  $\mu$ l of 0.5M Tris (Melford) pH = 6.8, 1.875 ml of dH<sub>2</sub>O, 3.5  $\mu$ l of TEMED (Sigma) and 35  $\mu$ l of 10 % APS (Fisher)] was poured on top, a comb was inserted and the gel was settle for at least 30 minutes. After that, 10 to 20  $\mu$ l of each sample was loaded along with 5  $\mu$ l of protein marker (Precision Plus Protein Standards, BioRad) and gel was run in a chamber full of of TGS buffer [(25mM Tris (Melford) pH = 8.3, 0.19 M glycine and 0.1% SDS (Sigma)] and run at 90 V until samples entered the resolving gel. After that gel was run at 120 V during approximately 2 hours. After electrophoresis gels were transfer onto a membrane following a semi-dry transfer method. Gels were

placed in transfer buffer [0.0029% glycine (Sigma), 0.0058% Tris- base (Sigma), 0.00004% SDS (Sigma), 20% methanol (Fisher)]. PVDF membrane (BioRad) was cut to gel-size and pre-wet in methanol and then transfer buffer during 10 minutes. A piece of Western blotting paper (Invitrogen) soaked on transfer buffer was placed on the anode plated of a Trans-Blot Semi-Dry Transfer Cell (BioRad) followed by the PVDF membrane, gel and a second piece of Western blotting paper (Invitrogen) soaked on transfer buffer. The cathode was placed on top and the transfer was performed at 25 Volts during 30 minutes.

To perform western blotting detection the membrane was transferred into blocking buffer [5% dried milk powder (Marvel) in PBST (PBS (Sigma) 0.2% Tween-20 (Sigma))] and incubated at room temperature with shaking for at least 1 hour. The membrane was rinsed with PBST [PBS (Sigma), 0.1 % Tween-20 (Sigma)] twice followed by 2 washes of 15 minutes each. After that, blocking buffer containing a 1:1000 dilution of primary antibody Anti-HA, mouse monoclonal primary antibody (12CA5 Roche, 5 mg/ml) was added and incubated at 4°C over-night. The membrane was rinsed with PBST twice followed by 2 washes of 15 minutes each and blocking solution containing secondary antibody, anti-mouse IgG-Peroxidase (A4416 Sigma, 0.63 mg/ml) at a dilution of 1:5000, and Clarity™ ECL substrate (Biorad) was added. After that, the membrane was rinsed with PBST twice followed by 2 washes of 15 minutes each. In Dark room, equal volumes of ECL developing solution (BioRad) were added to the membrane during 1 minute. After that, the membrane was wrapped in Saran wrap and expose to film (Hyperfilm ECL, Amersham) in a cassette. Exposure times were adapted depending on the intensity of the bands

and the film was developed using a Compact X4 Automatic Processor (Xograph Healthcare).

### **2.13. High- throughput RNA sequencing**

Strand-specific cDNA Illumina Barcoded Libraries were generated from 1 µg of total RNA extracted from WT and *sir2*  $\Delta/\Delta$  corresponding background strains and sequenced with an Illumina iSeq2000 platform. Illumina Library and Deep-sequencing was performed by the Genomics Core Facility at EMBL (Heidelberg, Germany). Raw reads were analyzed following the RNA deep sequencing analysis pipeline described by (Trapnell et al., 2012) using Galaxy (<https://usegalaxy.org/>) and Linux platform. The pipeline includes initial steps of Fastqc quality controls of the raw data generated (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). TopHat algorithm (Trapnell et al., 2009) was used to map sequencing reads to the *C. albicans* reference genome (Jones et al., 2004; Odds et al., 2004). Cufflink software (Trapnell et al., 2010) was run to assemble transcripts, estimates their abundances, and tests for differential expression and regulation of the different RNA-Sequencing samples. Cuffmerge scripts (Trapnell et al., 2010) merge all the assemblies generated by Cufflinks. Finally, differential expression was analysed using Cuffdiff algorithm (Trapnell et al., 2010). Heatmaps and boxplot graphs of differential gene expression output data were generated with R (<http://www.r-project.org/>). RNA sequencing data are deposited into ArrayExpress (accession number E-MTAB-4488 for wt and *sir2*  $\Delta/\Delta$  strains at 30°C and accession number E-MTAB-4622 for for wt and *sir2*  $\Delta/\Delta$  strains at 39°C).

## **2.14. Microscopy**

Liquid cultures of (YPAD) containing extra adenine (0.1mg/ml) and extra uridine (0.08mg/ml) at 30°C or 37°C were grown overnight. After that, 20 µl of cells were deposited on a slide (Thermo scientific). Microscopy was carried out using an Olympus IX81 inverted microscope. Images were captured with a Hamamatsu photonics C4742 digital camera, with light excitation from an Olympus MT20 illumination system. Olympus CellR imaging software was used to control the apparatus.

## **2.15. Bioinformatic tools**

DNA sequences were obtained from the Candida Genome Database (CDG, <http://www.candidagenome.org/>) and engineered using APE (<http://biologylabs.utah.edu/jorgensen/wayned/ape/>) and SnapGene (<http://www.snapgene.com/>). RNA-sequencing reads were visualized using IGV (<https://www.broadinstitute.org/igv/>). DNA alignments were performed using Blast (NCBI) or Muscle with default settings and visualised using Jalview (<http://www.jalview.org/>). Motif finder analyses were performed using MEME SUITE using the MEME discovery programme (<http://meme-suite.org/>)

## 7. References

Ahmad, A., Kabir, M.A., Kravets, A., Andaluz, E., Larriba, G., and Rustchenko, E. (2008). Chromosome instability and unusual features of some widely used strains of *Candida albicans*. *Yeast* 25, 433–448.

Allshire, R.C., Javerzat, J.P., Redhead, N.J., and Cranston, G. (1994). Position effect variegation at fission yeast centromeres. *Cell* 76, 157–169.

Allshire, R.C., Nimmo, E.R., Ekwall, K., Javerzat, J.P., and Cranston, G. (1995a). Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev.* 9, 218–233.

Allshire, R.C., Nimmo, E.R., Ekwall, K., Javerzat, J.P., and Cranston, G. (1995b). Mutations derepressing silent centromeric domains in fission yeast disrupt chromosome segregation. *Genes Dev.* 9, 218–233.

Anderson, M.Z., Baller, J.A., Dulmage, K., Wigen, L., and Berman, J. (2012). The three clades of the telomere-associated TLO gene family of *Candida albicans* have different splicing, localization, and expression features. *Eukaryot. Cell* 11, 1268–1275.

Anderson, M.Z., Gerstein, A.C., Wigen, L., Baller, J.A., and Berman, J. (2014). Silencing Is Noisy: Population and Cell Level Noise in Telomere-Adjacent Genes Is Dependent on Telomere Position and Sir2. *PLoS Genet.* 10, e1004436.

Anderson, M.Z., Wigen, L.J., Burrack, L.S., and Berman, J. (2015). Real-Time Evolution of a Subtelomeric Gene Family in *Candida albicans*. *Genetics* 200,

907–919.

De Backer, M.D., Maes, D., Vandoninck, S., Logghe, M., Contreras, R., and Luyten, W.H. (1999). Transformation of *Candida albicans* by electroporation. *Yeast* 15, 1609–1618.

Bannister, A.J., and Kouzarides, T. (2011). Regulation of chromatin by histone modifications. *Cell Res.* 21, 381–395.

Baum, M., Sanyal, K., Mishra, P.K., Thaler, N., and Carbon, J. (2006). Formation of functional centromeric chromatin is specified epigenetically in *Candida albicans*. *Proc. Natl. Acad. Sci. U. S. A.* 103, 14877–14882.

Bennett, R.J., and Johnson, A.D. (2005). Mating in *Candida albicans* and the Search for a Sexual Cycle. *Annu. Rev. Microbiol* 59, 233–255.

Bennett, R.J., Forche, A., and Berman, J. (2014). Rapid mechanisms for generating genome diversity: whole ploidy shifts, aneuploidy, and loss of heterozygosity. *Cold Spring Harb. Perspect. Med.* 4.

Berman, J. (2012). *Candida albicans*. *Curr. Biol.* 22, R620–R622.

Berman, J., and Sudbery, P.E. (2002). *Candida Albicans*: a molecular revolution built on lessons from budding yeast. *Nat. Rev. Genet.* 3, 918–930.

Bezerra, A.R., Simões, J., Lee, W., Rung, J., Weil, T., Gut, I.G., Gut, M., Bayés, M., Rizzetto, L., Cavalieri, D., et al. (2013). Reversion of a fungal genetic code alteration links proteome instability with genomic and phenotypic diversification. *Proc. Natl. Acad. Sci. U. S. A.* 110, 11079–11084.

Bhargava, P., and Reese, J.C. (2013). Transcription by Odd Pols. *Biochim.*

Biophys. Acta - Gene Regul. Mech. 1829, 249–250.

Bi, X. (2012). Functions of chromatin remodeling factors in heterochromatin formation and maintenance. *Sci. China. Life Sci.* 55, 89–96.

Bi, X. (2014). Heterochromatin structure: lessons from the budding yeast. *IUBMB Life* 66, 657–666.

Bi, X., Yu, Q., Sandmeier, J.J., and Elizondo, S. (2004). Regulation of transcriptional silencing in yeast by growth temperature. *J. Mol. Biol.* 344, 893–905.

Bird, A. (2013). Epigenetics: Discovery. *New Sci.* 217, ii – iii.

Briggs, S.D., Bryk, M., Strahl, B.D., Cheung, W.L., Davie, J.K., Dent, S.Y., Winston, F., and Allis, C.D. (2001). Histone H3 lysine 4 methylation is mediated by Set1 and required for cell growth and rDNA silencing in *Saccharomyces cerevisiae*. *Genes Dev.* 15, 3286–3295.

Bryk, M., Banerjee, M., Murphy, M., Knudsen, K.E., Garfinkel, D.J., and Curcio, M.J. (1997). Transcriptional silencing of Ty1 elements in the RDN1 locus of yeast. *Genes Dev.* 11, 255–269.

Bühler, M., and Gasser, S.M. (2009). Silent chromatin at the middle and ends: lessons from yeasts. *EMBO J.* 28, 2149–2161.

Cam, H.P. (2010). Roles of RNAi in chromatin regulation and epigenetic inheritance. *Epigenomics* 2, 613–626.

Cam, H.P., Sugiyama, T., Chen, E.S., Chen, X., FitzGerald, P.C., and Grewal, S.I.S. (2005). Comprehensive analysis of heterochromatin- and RNAi-mediated

epigenetic control of the fission yeast genome. *Nat. Genet.* 37, 809–819.

Carreto, L., Eiriz, M.F., Gomes, A.C., Pereira, P.M., Schuller, D., Santos, M.A., Biddenne, C., Blondin, B., Dequin, S., Vezinhet, F., et al. (2008). Comparative genomics of wild type yeast strains unveils important genome diversity. *BMC Genomics* 9, 524.

Cavalli, G. (2002). Chromatin as a eukaryotic template of genetic information. *Curr. Opin. Cell Biol.* 14, 269–278.

Chan, J.N.Y., Poon, B.P.K., Salvi, J., Olsen, J.B., Emili, A., and Mekhail, K. (2011). Perinuclear cohibin complexes maintain replicative life span via roles at distinct silent chromatin domains. *Dev. Cell* 20, 867–879.

Chang, C.S., and Pillus, L. (2009). Collaboration between the essential Esa1 acetyltransferase and the Rpd3 deacetylase is mediated by H4K12 histone acetylation in *Saccharomyces cerevisiae*. *Genetics* 183, 149–160.

Chatterjee, G., Sankaranarayanan, S.R., Guin, K., Thattikota, Y., Padmanabhan, S., Siddharthan, R., and Sanyal, K. (2016). Repeat-Associated Fission Yeast-Like Regional Centromeres in the Ascomycetous Budding Yeast *Candida tropicalis*. *PLoS Genet.* 12, e1005839.

Cheng, S.-C., Joosten, L.A.B., Kullberg, B.-J., and Netea, M.G. (2012). Interplay between *Candida albicans* and the Mammalian Innate Host Defense. *Infect. Immun.* 80, 1304–1313.

Chi, P., Allis, C.D., and Wang, G.G. (2010). Covalent histone modifications — miswritten, misinterpreted and mis-erased in human cancers. *Nat. Rev. Cancer* 10, 457–469.

Chibana, H., Iwaguchi, S., Homma, M., Chindamporn, A., Nakagawa, Y., and 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., Beckerman, J.L., and Magee, P.T. (2000). Fine-resolution physical mapping of genomic diversity in *Candida albicans*. *Genome Res.* *10*, 1865–1877.

Choy, J.S., Acuña, R., Au, W.-C., and Basrai, M.A. (2011). A role for histone H4K16 hypoacetylation in *Saccharomyces cerevisiae* kinetochore function. *Genetics* *189*, 11–21.

Cieniewicz, A.M., Moreland, L., Ringel, A.E., Mackintosh, S.G., Raman, A., Gilbert, T.M., Wolberger, C., Tackett, A.J., and Taverna, S.D. (2014). The bromodomain of Gcn5 regulates site specificity of lysine acetylation on histone H3. *Mol. Cell. Proteomics* *13*, 2896–2910.

Cioci, F., Vu, L., Eliason, K., Oakes, M., Siddiqi, I.N., and Nomura, M. (2003). Silencing in yeast rDNA chromatin: reciprocal relationship in gene expression between RNA polymerase I and II. *Mol. Cell* *12*, 135–145.

Correll, S.J., Schubert, M.H., and Grigoryev, S.A. (2012). Short nucleosome repeats impose rotational modulations on chromatin fibre folding. *EMBO J.* *31*, 2416–2426.

Cosgrove, M.S., and Wolberger, C. (2005). How does the histone code work? *Biochem. Cell Biol.* *83*, 468–476.

Dehé, P.-M., and Géli, V. (2006). The multiple faces of Set1. *Biochem. Cell Biol.*

84, 536–548.

Ellermeier, C., Higuchi, E.C., Phadnis, N., Holm, L., Geelhood, J.L., Thon, G., and Smith, G.R. (2010). RNAi and heterochromatin repress centromeric meiotic recombination. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 8701–8705.

Fingerman, I.M., Wu, C.-L., Wilson, B.D., and Briggs, S.D. (2005). Global loss of Set1-mediated H3 Lys4 trimethylation is associated with silencing defects in *Saccharomyces cerevisiae*. *J. Biol. Chem.* *280*, 28761–28765.

Forche, A., May, G., and Magee, P.T. (2005). Demonstration of loss of heterozygosity by single-nucleotide polymorphism microarray analysis and alterations in strain morphology in *Candida albicans* strains during infection. *Eukaryot. Cell* *4*, 156–165.

Forche, A., Steinbach, M., and Berman, J. (2009). Efficient and rapid identification of *Candida albicans* allelic status using SNP-RFLP. *FEMS Yeast Res.* *9*, 1061–1069.

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

Ford, C.B., Funt, J.M., Abbey, D., Issi, L., Guiducci, C., Martinez, D.A., Delorey, T., Li, B.Y., White, T.C., Cuomo, C., et al. (2015). The evolution of drug resistance in clinical isolates of *Candida albicans*. *Elife* *4*, e00662.

Freeman-Cook, L.L., Gómez, E.B., Spedale, E.J., Marlett, J., Forsburg, S.L., Pillus, L., and Laurenson, P. (2005). Conserved locus-specific silencing functions of *Schizosaccharomyces pombe* sir2+. *Genetics* *169*, 1243–1260.

Freire-Benítez, V., Price, R.J., Tarrant, D., Berman, J., and Buscaino, A. (2016a). *Candida albicans* repetitive elements display epigenetic diversity and plasticity. *Sci. Rep.* 6, 22989.

Freire-Benítez, V., Price, R.J., and Buscaino, A. (2016b). The Chromatin of *Candida albicans* Pericentromeres Bears Features of Both Euchromatin and Heterochromatin. *Front. Microbiol.* 7.

Freire-Benítez, V., Gourlay, S., Berman, J., and Buscaino, A. (2016c). Sir2 regulates stability of repetitive domains differentially in the human fungal pathogen *Candida albicans*. *Nucleic Acids Res.* gkw594.

Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. (2000). Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. *Nature* 407, 1018–1022.

Fritze, C.E., Verschueren, K., Strich, R., and Easton Esposito, R. (1997). Direct evidence for SIR2 modulation of chromatin structure in yeast rDNA. *EMBO J.* 16, 6495–6509.

Ganley, A.R.D., and Kobayashi, T. (2011). Monitoring the rate and dynamics of concerted evolution in the ribosomal DNA repeats of *Saccharomyces cerevisiae* using experimental evolution. *Mol. Biol. Evol.* 28, 2883–2891.

Ganley, A.R.D., and Kobayashi, T. (2013). Ribosomal DNA and cellular senescence: new evidence supporting the connection between rDNA and aging. *FEMS Yeast Res.*

Ganley, A.R.D., Hayashi, K., Horiuchi, T., and Kobayashi, T. (2005). Identifying

gene-independent noncoding functional elements in the yeast ribosomal DNA by phylogenetic footprinting. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 11787–11792.

Gartenberg, M. (2009). Heterochromatin and the cohesion of sister chromatids. *Chromosome Res.* *17*, 229–238.

Gottlieb, S., and Esposito, R.E. (1989a). A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA. *Cell* *56*, 771–776.

Gottlieb, S., and Esposito, R.E. (1989b). A new role for a yeast transcriptional silencer gene, SIR2, in regulation of recombination in ribosomal DNA. *Cell* *56*, 771–776.

Gottschling, D.E., Aparicio, O.M., Billington, B.L., and Zakian, V.A. (1990). Position effect at *S. cerevisiae* telomeres: reversible repression of Pol II transcription. *Cell* *63*, 751–762.

Grewal, S.I.S., and Jia, S. (2007). Heterochromatin revisited. *Nat. Rev. Genet.* *8*, 35–46.

Grewal, S.I.S., and Moazed, D. (2003). Heterochromatin and epigenetic control of gene expression. *Science* *301*, 798–802.

Grummt, I., and Pikaard, C.S. (2003). Epigenetic silencing of RNA polymerase I transcription. *Nat. Rev. Mol. Cell Biol.* *4*, 641–649.

Guarente, L. (2000). Sir2 links chromatin silencing, metabolism, and aging. *Genes Dev.* *14*, 1021–1026.

von der Haar, T. (2007). Optimized Protein Extraction for Quantitative

Proteomics of Yeasts. PLoS One 2, e1078.

Haber, J.E. (2012). Mating-type genes and MAT switching in *Saccharomyces cerevisiae*. *Genetics* 191, 33–64.

Haldar, K., Kamoun, S., Hiller, N.L., Bhattacharje, S., and van Ooij, C. (2006). Common infection strategies of pathogenic eukaryotes. *Nat. Rev. Microbiol.* 4, 922–931.

Hall, B.M., Ma, C.-X., Liang, P., and Singh, K.K. (2009). Fluctuation analysis CalculatOR: a web tool for the determination of mutation rate using Luria-Delbruck fluctuation analysis. *Bioinformatics* 25, 1564–1565.

Hamperl, S., Wittner, M., Babl, V., Perez-Fernandez, J., Tschochner, H., and Griesenbeck, J. Chromatin states at ribosomal DNA loci. *Biochim. Biophys. Acta* 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., et al. (2014). Telomeric ORFs (TLOs) in *Candida* spp. Encode Mediator Subunits That Regulate Distinct Virulence Traits. *PLoS Genet.* 10, e1004658.

Hegemann, J.H., and Fleig, U.N. (1993). The centromere of budding yeast. *BioEssays* 15, 451–460.

Heijmans, B.T., Tobi, E.W., Stein, A.D., Putter, H., Blauw, G.J., Susser, E.S., Slagboom, P.E., and Lumey, L.H. (2008). Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. U. S. A.* 105, 17046–17049.

Henikoff, S. (1990). Position-effect variegation after 60 years. *Trends Genet.* 6, 422–426.

Henikoff, S., and Henikoff, J.G. (2012). “Point” centromeres of *Saccharomyces* harbor single centromere-specific nucleosomes. *Genetics* 190, 1575–1577.

Heringa, J. (1998). Detection of internal repeats: how common are they? *Curr. Opin. Struct. Biol.* 8, 338–345.

Hickman, M.A., Froyd, C.A., and Rusche, L.N. (2011). Reinventing heterochromatin in budding yeasts: Sir2 and the origin recognition complex take center stage. *Eukaryot. Cell* 10, 1183–1192.

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., et al. (2015). Genetic and phenotypic intra-species variation in *Candida albicans*. *Genome Res.* 25, 413–425.

Hirano, T. (2012). Condensins: universal organizers of chromosomes with diverse functions. *Genes Dev.* 26, 1659–1678.

Hnisz, D., Schwarzmüller, T., and Kuchler, K. (2009). Transcriptional loops meet chromatin: a dual-layer network controls white-opaque switching in *Candida albicans*. *Mol. Microbiol.* 74, 1–15.

van het Hoog, M., Rast, T.J., Martchenko, M., Grindle, S., Dignard, D., Hogues, H., Cuomo, C., Berriman, M., Scherer, S., Magee, B.B., et al. (2007). Assembly of the *Candida albicans* genome into sixteen supercontigs aligned on the eight chromosomes. *Genome Biol.* 8, R52.

Huang, Y. (2002). Transcriptional silencing in *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*. *Nucleic Acids Res.* *30*, 1465–1482.

Huang, G., Wang, H., Chou, S., Nie, X., Chen, J., and Liu, H. (2006a). Bistable expression of WOR1, a master regulator of white-opaque switching in *Candida albicans*. *Proc. Natl. Acad. Sci.* *103*, 12813–12818.

Huang, J., Brito, I.L., Villén, J., Gygi, S.P., and Amon, A. (2006b). Inhibition of homologous recombination by a cohesin-associated clamp complex recruited to the rDNA recombination enhancer. *2*, 2887–2901.

Huber, D., and Rustchenko, E. (2001). Large circular and linear rDNA plasmids in *Candida albicans*. *Yeast* *18*, 261–272.

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

Imai, S., Armstrong, C.M., Kaeberlein, M., and Guarente, L. (2000). Transcriptional silencing and longevity protein Sir2 is an NAD-dependent histone deacetylase. *Nature* *403*, 795–800.

Iwaguchi, S., Homma, M., Chibana, H., and 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., and 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.

Iwaguchi, S.-I., Suzuki, M., Sakai, N., Yokoyama, K., and Suzuki, T. (2008). The loss of parts of chromosome 7 followed by the insertion of URA cassette into RB2 on MRS in *Candida albicans* strain CAI-4. *Med. Mycol.* *46*, 655–663.

Jackson, A.P., Gamble, J.A., Yeomans, T., Moran, G.P., Saunders, D., Harris, D., Aslett, M., Barrell, J.F., Butler, G., Citiulo, F., et al. (2009). Comparative genomics of the fungal pathogens *Candida dubliniensis* and *Candida albicans*. *Genome Res.* *19*, 2231–2244.

Jenuwein, T. (2001). Re-SET-ting heterochromatin by histone methyltransferases. *Trends Cell Biol.* *11*, 266–273.

Jenuwein, T., and Allis, C.D. (2001). Translating the histone code. *Science* *293*, 1074–1080.

Joly, S., Pujol, C., and 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.

Jones, T., Federspiel, N.A., Chibana, H., Dungan, J., Kalman, S., Magee, B.B., Newport, G., Thorstenson, Y.R., Agabian, N., Magee, P.T., et al. (2004). The diploid genome sequence of *Candida albicans*. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 7329–7334.

Kelleher, R.J., Flanagan, P.M., Kornberg, R.D., Berger, S.L., Cress, W.D., Cress, A., Triezenberg, S.J., Guarente, L., Bradford, M.M., Bram, R.J., et al. (1990). A novel mediator between activator proteins and the RNA polymerase II transcription apparatus. *Cell* *61*, 1209–1215.

Ketel, C., Wang, H.S.W., McClellan, M., Bouchonville, K., Selmecki, A., Lahav,

T., Gerami-Nejad, M., and Berman, J. (2009). Neocentromeres form efficiently at multiple possible loci in *Candida albicans*. *PLoS Genet.* 5, e1000400.

Klar, A.J., Srikantha, T., and Soll, D.R. (2001). A histone deacetylation inhibitor and mutant promote colony-type switching of the human pathogen *Candida albicans*. *Genetics* 158, 919–924.

Kobayashi, T. (2011). Regulation of ribosomal RNA gene copy number and its role in modulating genome integrity and evolutionary adaptability in yeast. *Cell. Mol. Life Sci.* 68, 1395–1403.

Kobayashi, T., and Ganley, A.R.D. (2005). Recombination regulation by transcription-induced cohesin dissociation in rDNA repeats. *Science* 309, 1581–1584.

Kornberg, R.D. (1974). Chromatin structure: a repeating unit of histones and DNA. *Science* 184, 868–871.

Kornberg, R.D. (1977). Structure of Chromatin. *Ann. Rev. Biochem* 46, 931–954.

Krogan, N.J., Dover, J., Khorrami, S., Greenblatt, J.F., Schneider, J., Johnston, M., and Shilatifard, A. (2002). COMPASS, a histone H3 (Lysine 4) methyltransferase required for telomeric silencing of gene expression. *J. Biol. Chem.* 277, 10753–10755.

Kruskal, W.H., and Wallis, W.A. (2012). Use of Ranks in One-Criterion Variance Analysis. *J. Am. Stat. Assoc.*

Krysan, D.J., Sutterwala, F.S., Wellington, M., Seider, K., Heyken, A., Luttich,

A., Miramon, P., Hube, B., Lionakis, M., Netea, M., et al. (2014). Catching Fire: *Candida albicans*, Macrophages, and Pyroptosis. *PLoS Pathog.* *10*, e1004139.

Lea, D.E., and Coulson, C.A. (1949). The distribution of the numbers of mutants in bacterial populations. *J. Genet.* *49*, 264–285.

van Leeuwen, F., Gafken, P.R., and Gottschling, D.E. (2002). Dot1p Modulates Silencing in Yeast by Methylation of the Nucleosome Core. *Cell* *109*, 745–756.

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

Li, Q., and Zhang, Z. (2012). Linking DNA replication to heterochromatin silencing and epigenetic inheritance. *Acta Biochim. Biophys. Sin. (Shanghai)*. *44*, 3–13.

Li, C., Mueller, J.E., and Bryk, M. (2006a). Sir2 represses endogenous polymerase II transcription units in the ribosomal DNA nontranscribed spacer. *Mol. Biol. Cell* *17*, 3848–3859.

Li, C., Mueller, J.E., and Bryk, M. (2006b). Sir2 represses endogenous polymerase II transcription units in the ribosomal DNA nontranscribed spacer. *Mol. Biol. Cell* *17*, 3848–3859.

Lopes da Rosa, J., and Kaufman, P.D. (2012). Chromatin-mediated *Candida albicans* virulence. *Biochim. Biophys. Acta* *1819*, 349–355.

Lopes da Rosa, J., Boyartchuk, V.L., Zhu, L.J., and Kaufman, P.D. (2010). Histone acetyltransferase Rtt109 is required for *Candida albicans* pathogenesis. *Proc. Natl. Acad. Sci. U. S. A.* *107*, 1594–1599.

Louis, E.J. (1995). The chromosome ends of *Saccharomyces cerevisiae*. *Yeast* 11, 1553–1573.

Louis, E.J., Naumova, E.S., Lee, A., Naumov, G., and Haber, J.E. (1994). The chromosome end in yeast: its mosaic nature and influence on recombinational dynamics. *Genetics* 136, 789–802.

Magee, B.B., and Magee, P.T. (2000). Induction of mating in *Candida albicans* by construction of MTL $\alpha$  and MTL $\alpha$  strains. *Science* 289, 310–313.

Magee, B.B., D'Souza, T.M., and Magee, P.T. (1987). Strain and species identification by restriction fragment length polymorphisms in the ribosomal DNA repeat of *Candida* species. *J. Bacteriol.* 169, 1639–1643.

Magee, B.B., Sanchez, M.D., Saunders, D., Harris, D., Berriman, M., and 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.

Margie T. Borra, ‡, Michael R. Langer, §., James T. Slama, || and, and John M. Denu\*, ‡ (2004). Substrate Specificity and Kinetic Mechanism of the Sir2 Family of NAD<sup>+</sup>-Dependent Histone/Protein Deacetylases†.

Marvin, M.E., Becker, M.M., Noel, P., Hardy, S., Bertuch, A.A., and Louis, E.J. (2009). The association of yKu with subtelomeric core X sequences prevents recombination involving telomeric sequences. *Genetics* 183, 453–467, 1SI – 13SI.

McLaughlan, J.M., Liti, G., Sharp, S., Maslowska, A., and Louis, E.J. (2012). Apparent ploidy effects on silencing are post-transcriptional at HML and

telomeres in *Saccharomyces cerevisiae*. PLoS One 7, e39044.

Mead, J., McCord, R., Youngster, L., Sharma, M., Gartenberg, M.R., and Vershon, A.K. (2007). Swapping the gene-specific and regional silencing specificities of the Hst1 and Sir2 histone deacetylases. *Mol. Cell. Biol.* 27, 2466–2475.

Miletti-Gonzalez, K.E., and Leibowitz, M.J. (2008). Molecular Characterization of Two Types of rDNA Units in a Single Strain of *Candida albicans*. *J. Eukaryot. Microbiol.* 55, 522–529.

Miller, M.G., and 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.

Mishra, P.K., Baum, M., and Carbon, J. (2007). Centromere size and position in *Candida albicans* are evolutionarily conserved independent of DNA sequence heterogeneity. *Mol. Genet. Genomics* 278, 455–465.

Moretti, P., Freeman, K., Coodly, L., and Shore, D. (1994). Evidence that a complex of SIR proteins interacts with the silencer and telomere-binding protein RAP1. *Genes Dev.* 8, 2257–2269.

Morris, S.A., Rao, B., Garcia, B.A., Hake, S.B., Diaz, R.L., Shabanowitz, J., Hunt, D.F., Allis, C.D., Lieb, J.D., and Strahl, B.D. (2006). Identification of Histone H3 Lysine 36 Acetylation as a Highly Conserved Histone Modification. *J. Biol. Chem.* 282, 7632–7640.

Muller, H.J. (1930). Types of visible variations induced by X-rays in *Drosophila*. *J. Genet.* 22, 299–334.

Nazar, R.N. (2004). Ribosomal RNA processing and ribosome biogenesis in eukaryotes. *IUBMB Life* 56, 457–465.

Ng, H.H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., and Struhl, K. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev.* 16, 1518–1527.

Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* 11, 709–719.

Nislow, C., Ray, E., and Pillus, L. (1997). SET1, a yeast member of the trithorax family, functions in transcriptional silencing and diverse cellular processes. *Mol. Biol. Cell* 8, 2421–2436.

Noble, S.M., and Johnson, A.D. (2005). Strains and strategies for large-scale gene deletion studies of the diploid human fungal pathogen *Candida albicans*. *Eukaryot. Cell* 4, 298–309.

North, B.J., Verdin, E., Kouzarides, T., Ruijter, A. de, Gennip, A. van, Caron, H., Kemp, S., Kuilenburg, A. van, Verdin, E., Dequiedt, F., et al. (2004). Sirtuins: Sir2-related NAD-dependent protein deacetylases. *Genome Biol.* 5, 224.

Odds, F.C., Brown, A.J., Gow, N.A., Jones, T., Federspiel, N., Chibana, H., Dungan, J., Kalman, S., Magee, B., Newport, G., et al. (2004). *Candida albicans* genome sequence: a platform for genomics in the absence of genetics. *Genome Biol.* 5, 230.

Oppikofer, M., Kueng, S., Martino, F., Soeroes, S., Hancock, S.M., Chin, J.W.,

Fischle, W., and Gasser, S.M. (2011). A dual role of H4K16 acetylation in the establishment of yeast silent chromatin. *EMBO J.* 30, 2610–2621.

Padmanabhan, S., Thakur, J., Siddharthan, R., and Sanyal, K. (2008). Rapid evolution of Cse4p-rich centromeric DNA sequences in closely related pathogenic yeasts, *Candida albicans* and *Candida dubliniensis*. *Proc. Natl. Acad. Sci. U. S. A.* 105, 19797–19802.

Pendrak, M.L., and Roberts, D.D. (2011). Ribosomal RNA processing in *Candida albicans*. *RNA* 17, 2235–2248.

Pérez-Martín, J., Uría, J.A., and Johnson, A.D. (1999). Phenotypic switching in *Candida albicans* is controlled by a SIR2 gene. *EMBO J.* 18, 2580–2592.

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

Pidoux, A., Mellone, B., and Allshire, R. (2004). Analysis of chromatin in fission yeast. *Methods* 33, 252–259.

Politz, J.C.R., Scalzo, D., and Groudine, M. (2013). Something silent this way forms: the functional organization of the repressive nuclear compartment. *Annu. Rev. Cell Dev. Biol.* 29, 241–270.

Pryde, F.E., Gorham, H.C., and Louis, E.J. (1997). Chromosome ends: all the same under their caps. *Curr. Opin. Genet. Dev.* 7, 822–828.

Raman, S.B., Nguyen, M.H., Zhang, Z., Cheng, S., Jia, H.Y., Weisner, N., Iczkowski, K., and Clancy, C.J. (2006). *Candida albicans* SET1 encodes a histone 3 lysine 4 methyltransferase that contributes to the pathogenesis of

invasive candidiasis. *Mol. Microbiol.* 60, 697–709.

Reik, W., and Walter, J. (1998). Imprinting mechanisms in mammals. *Curr. Opin. Genet. Dev.* 8, 154–164.

Rosenberg, S.M. (2011). Stress-induced loss of heterozygosity in *Candida*: a possible missing link in the ability to evolve. *MBio* 2.

Rusche, L.N., Kirchmaier, A.L., and Rine, J. (2003). The establishment, inheritance, and function of silenced chromatin in *Saccharomyces cerevisiae*. *Annu. Rev. Biochem.* 72, 481–516.

Russell, J., Zomerdijk, J.C.B.M., Grummt, I., Moss, T., Stefanovsky, V.Y., Reeder, R.H., Hannan, K.M., al., et, Comai, L., Moss, T., et al. (2005). RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem. Sci.* 30, 87–96.

Rustchenko, E. (2007). Chromosome instability in *Candida albicans*. *FEMS Yeast Res.* 7, 2–11.

Ryu, H.-Y., and Ahn, S. (2014). Yeast histone H3 lysine 4 demethylase Jhd2 regulates mitotic rDNA condensation. *BMC Biol.* 12, 75.

Sadhu, C., McEachern, M.J., Rustchenko-Bulgac, E.P., Schmid, J., Soll, D.R., and Hicks, J.B. (1991). Telomeric and dispersed repeat sequences in *Candida* yeasts and their use in strain identification. *J. Bacteriol.* 173, 842–850.

Salminen, A., and Kaarniranta, K. (2009). SIRT1 regulates the ribosomal DNA locus: epigenetic candles twinkle longevity in the Christmas tree. *Biochem. Biophys. Res. Commun.* 378, 6–9.

Santos, M.A., Cheesman, C., Costa, V., Moradas-Ferreira, P., and Tuite, M.F. (1999). Selective advantages created by codon ambiguity allowed for the evolution of an alternative genetic code in *Candida* spp. *Mol. Microbiol.* *31*, 937–947.

Sanyal, K., Baum, M., and Carbon, J. (2004). Centromeric DNA sequences in the pathogenic yeast *Candida albicans* are all different and unique. *Proc. Natl. Acad. Sci. U. S. A.* *101*, 11374–11379.

Sauve, A.A., Moir, R.D., Schramm, V.L., and Willis, I.M. (2005). Chemical activation of Sir2-dependent silencing by relief of nicotinamide inhibition. *Mol. Cell* *17*, 595–601.

Schoeftner, S., and Blasco, M.A. (2009). A “higher order” of telomere regulation: telomere heterochromatin and telomeric RNAs. *EMBO J.* *28*, 2323–2336.

Selmecki, A., Forche, A., and Berman, J. (2006). Aneuploidy and isochromosome formation in drug-resistant *Candida albicans*. *Science* *313*, 367–370.

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

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

Selmecki, A.M., Dulmage, K., Cowen, L.E., Anderson, J.B., and Berman, J. (2009). Acquisition of aneuploidy provides increased fitness during the evolution of antifungal drug resistance. *PLoS Genet.* *5*, e1000705.

Shankaranarayana, G.D., Motamedi, M.R., Moazed, D., and Grewal, S.I.S. (2003). Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. *Curr. Biol.* *13*, 1240–1246.

Shapiro, R.S., Robbins, N., and Cowen, L.E. (2011). Regulatory circuitry governing fungal development, drug resistance, and disease. *Microbiol. Mol. Biol. Rev.* *75*, 213–267.

Sinclair, D.A., and Guarente, L. (1997). Extrachromosomal rDNA circles--a cause of aging in yeast. *Cell* *91*, 1033–1042.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. *Nat. Rev. Genet.* *8*, 272–285.

Smith, J.S., and Boeke, J.D. (1997). An unusual form of transcriptional silencing in yeast ribosomal DNA. *Genes Dev.* *11*, 241–254.

Smith, J.S., Caputo, E., and Boeke, J.D. (1999). A genetic screen for ribosomal DNA silencing defects identifies multiple DNA replication and chromatin-modulating factors. *Mol. Cell. Biol.* *19*, 3184–3197.

Stavenhagen, J.B., and Zakian, V.A. (1998). Yeast telomeres exert a position effect on recombination between internal tracts of yeast telomeric DNA. *Genes Dev.* *12*, 3044–3058.

van Steensel, B. (2011). Chromatin: constructing the big picture. *EMBO J.* *30*, 1885–1895.

Steiner, N.C., Hahnenberger, K.M., and Clarke, L. (1993). Centromeres of the fission yeast *Schizosaccharomyces pombe* are highly variable genetic loci. *Mol.*

Cell. Biol. 13, 4578–4587.

Steinert, M. (2014). Pathogen intelligence. *Front. Cell. Infect. Microbiol.* 4.

Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* 403, 41–45.

Straight, A.F., Shou, W., Dowd, G.J., Turck, C.W., Deshaies, R.J., Johnson, A.D., and Moazed, D. (1999). Net1, a Sir2-associated nucleolar protein required for rDNA silencing and nucleolar integrity. *Cell* 97, 245–256.

Thakur, J., and Sanyal, K. (2012). A coordinated interdependent protein circuitry stabilizes the kinetochore ensemble to protect CENP-A in the human pathogenic yeast *Candida albicans*. *PLoS Genet.* 8, e1002661.

Tham, W.-H., and Zakian, V.A. (2002). Transcriptional silencing at *Saccharomyces* telomeres: implications for other organisms. *Oncogene* 21, 512–521.

Thon, G., and Verhein-Hansen, J. (2000). Four chromo-domain proteins of *Schizosaccharomyces pombe* differentially repress transcription at various chromosomal locations. *Genetics* 155, 551–568.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111.

Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515.

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578.

Uhl, M.A., Biery, M., Craig, N., Johnson, A.D., Altschul, S., Madden, T., Schaffer, A., Zhang, J., Zhang, Z., Miller, W., et al. (2003). Haploinsufficiency-based large-scale forward genetic analysis of filamentous growth in the diploid human fungal pathogen *C.albicans*. *EMBO J.* 22, 2668–2678.

Voss, T.S., Bozdech, Z., and Bártfai, R. (2014). Epigenetic memory takes center stage in the survival strategy of malaria parasites. *Curr. Opin. Microbiol.* 20C, 88–95.

Wang, B.-D., Butylin, P., and Strunnikov, A. (2006). Condensin function in mitotic nucleolar segregation is regulated by rDNA transcription. *Cell Cycle* 5, 2260–2267.

Weinhold, B. (2006). Epigenetics: the science of change. *Environ. Health Perspect.* 114, A160–A167.

Williamson, K., Schneider, V., Jordan, R.A., Mueller, J.E., Henderson Pozzi, M., and Bryk, M. (2013). Catalytic and functional roles of conserved amino acids in the SET domain of the *S. cerevisiae* lysine methyltransferase Set1. *PLoS One* 8, e57974.

Wilson, R.B., Davis, D., and Mitchell, A.P. (1999). Rapid hypothesis testing with *Candida albicans* through gene disruption with short homology regions. *J. Bacteriol.* 181, 1868–1874.

Wurtele, H., Tsao, S., Lépine, G., Mullick, A., Tremblay, J., Drogaris, P., Lee, E.-H., Thibault, P., Verreault, A., and Raymond, M. (2010). Modulation of histone H3 lysine 56 acetylation as an antifungal therapeutic strategy. *Nat. Med.* *16*, 774–780.

Xu, F., Zhang, Q., Zhang, K., Xie, W., and Grunstein, M. (2007). Sir2 Deacetylates Histone H3 Lysine 56 to Regulate Telomeric Heterochromatin Structure in Yeast. *Mol. Cell* *27*, 890–900.

Yang, X.-J., and Seto, E. (2003). Collaborative spirit of histone deacetylases in regulating chromatin structure and gene expression. *Curr. Opin. Genet. Dev.* *13*, 143–153.

Yankulov, K. (2013). Dynamics and stability: epigenetic conversions in position effect variegation. *Biochem. Cell Biol.* *91*, 6–13.

Yu, E.Y., Yen, W.-F., Steinberg-Neifach, O., and 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.

Zhang, A., Petrov, K.O., Hyun, E.R., Liu, Z., Gerber, S.A., and Myers, L.C. (2012). The Tlo Proteins Are Stoichiometric Components of *Candida albicans* Mediator Anchored via the Med3 Subunit. *Eukaryot. Cell* *11*, 874–884.