UNIVERSITY OF CALIFORNIA, MERCED

The role of chromatin in regulating heritable cell fate decisions

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Quantitative and Systems Biology

by

Mohammad N. Qasim

Committee in charge:

Professor Ramen Saha, Chair of Advisory Committee

Professor Clarissa J. Nobile

Professor Richard Bennett

Professor Michael Cleary

Professor Aaron Hernday, Supervisor

2022

The dissertation of Mohammad Qasim, "**The role of chromatin in regulating heritable cell fate decisions***"*, is approved, and is acceptable in quality and form for publication on microfilm and electronically:

_____ Date _____

Professor Clarissa J. Nobile

_____ Date _____

Professor Richard Bennett

_____ Date _____

Professor Michael Cleary

Supervisor _____ Date _____

Professor Aaron Hernday

Chair _____ Date _____

Professor Ramen Saha

University of California, Merced

2022

Table of Contents

I. List of Figures

II. List of Tables

III. Acknowledgments

First and foremost, I am extremely grateful to my advisor, Dr. Aaron Hernday, for his invaluable guidance, continuous support, patience, and encouragement during graduate school. Your immense knowledge and plentiful experience have encouraged me in my academic research and daily life. You have helped me grow into the scientist that I am today, and thanks to you, I am confident in my ability to deal with and overcome the challenges that lie ahead. I have enjoyed my time in your lab, and I can't thank you enough for being a supportive advisor and mentor. Second, I would like to thank Dr. Clarissa Nobile, for being an exceptional collaborator, an awesome and encouraging committee member, and for all the helpful feedback on my work. Also, I would like to thank Hernday and Nobile lab members, many of whom I have had the pleasure of collaborating with on multiple research projects. There are too many to name, but I'd especially like to thank Namkha Nguyen, Dr. Ashley Arevalo Qasim, Dr. Akshay Paropkari, Deepika Gunasekaran, and Dr. Craig L. Ennis. Last but certainly not least, I would like to express my gratitude to my committee chair, Dr. Ramen Saha, and my committee members Dr. Richard Bennett, and Dr. Michael Cleary. You have all been invaluable to my learning and growth, and I appreciate your support and helpful feedback all these years.

IV. Vita for Mohammad N. Qasim

Education:

Doctoral Candidate, Quantitative and Systems Biology 2017 – 2022
University of California, Merced, CA

Bachelor of Science, Microbiology and Immunology 2009 – 2014
University of California, Merced, CA

Research Experiences:

Graduate Student Researcher, University of California, Merced 2017 – 2022
Investigated transcriptional networks and epigenetic mechanisms governing fungal cell-type switching using CRISPR-Cas9 genetic engineering and genome-wide sequencing approaches, including: ATAC-seq, CUT&RUN, ChIP-seq, MNase-seq and RNA-seq.

House Fellow, Living Learning Community 2018 – 2020

As a House Fellow for the Beyond the MD Living Learning Community (LLC) UCM, my primary responsibility was to build a safe and inclusive learning community for residents of the LLC, mentor first year residents, coordinate LLC-specific educational programs, facilitate campus referrals, coordinate impactful campus partnerships, and lead instruction of first year seminar for LLC scholars.

Senior Microbiology Technician, The Clorox Company 2014 – 2027
As a Microbiology Technician, I worked with scientists on a diverse portfolio of projects, often times working on multiple projects at a time to deliver impactful results. Being part of a team at The Clorox Company meant working with people from diverse ethnic and technical backgrounds and developing a diverse skillset.

List of Publications:

(* indicates co-first authors)

Qasim, M.N.*, Valle Arevalo, A.*, Paropkari, A.D.*, Sindi, S.S., Nobile, C.J., Hernday, A.D. (2022). Genome-wide profiling of transcription factor-DNA binding interactions in *Candida albicans*: a comprehensive CUT&RUN method and data analysis workflow. *Journal of Visual Experiments*.

Qasim, M.N., Valle Arevalo, A., Hernday, A.D., Nobile, C.J., (2021). The roles of chromatin accessibility in regulating the *Candida albicans* white-opaque phenotypic switch. *Journal of Fungi*.

Villa, S., Hamideh, M., Weinstock, A., **Qasim, M.N**., Hazbun, T.R., Sellam, A., Hernday, A.D. and Thangamani, S., 2020. Transcriptional control of hyphal morphogenesis in *Candida albicans*. *FEMS yeast research*.

V. Abstract

The role of chromatin in regulating heritable cell fate decisions

Mohammad Qasim
Doctor of Philosophy in Quantitative and Systems Biology
University of California, Merced
2022

The human body is made of many different cell types, the vast majority of which contain the same underlying genetic information. This singular genetic blueprint gives rise to unique transcriptional programs that specify patterns of gene expression and establish numerous distinct cell types, and chromatin plays an important role in this process. Research in the field of chromatin biology has historically separated the two main constituents of chromatin: nucleosomes, and the relatively smaller transcription factors (TFs). Recently, however, a growing body of evidence shows that it is the interplay between TFs and nucleosomes, two opposite but interconnected forces, that drives chromatin mediated mechanisms in a yin and yang fashion. Despite the progress made thus far, we still lack a clear understanding of the mechanisms through which chromatin regulates cell type establishment and heritability. In this dissertation, I examined how the interplay between TFs and nucleosomes regulates phenotypic switching between two distinct cell types referred to as white and opaque in the diploid polymorphic fungus *Candida albicans*. The white-opaque switch is observed in *C. albicans* and several closely related species but is absent in other common model yeasts like *Saccharomyces cerevisiae*. The two cell types are established and maintained by distinct transcriptional programs, leading to differences in metabolic preferences, mating abilities, cellular morphologies, responses to environmental signals, interactions with the host immune system and differential expression of ~20% of the transcriptome. Remarkably, the level of connectivity of the opaque cell transcriptional circuit closely resembles the circuits that control stem cell maintenance and differentiation in humans, indicating that the white-opaque switch represents an attractive and relatively “simple” model system to investigate how the interplay between TFs and chromatin structure regulates heritable cell fate decisions in higher eukaryotes. For this work, I adapted several next-generation sequencing (NGS) based assays to examine chromatin structure in *C. albicans*, one of which I describe in Chapter 3. Next, I used these genome-wide approaches to show that the opaque cell “master regulator” Wor1 establishes opaque-specific chromatin structures, some of which are required for heritable maintenance of the opaque transcriptional program (Chapter 2). All-in-all, I show that, in addition to Wor1, there are more TFs (Ndt80 and Wor2) with specialized ability to target many Nucleosome-Free Regions (NFRs)—compact chromatin structures that are observed to be inaccessible by ATAC-seq assays—and effect the expression of a large number of genes. Since this specialized ability to organize chromatin is observed throughout the genome, these types of chromatin organizing TFs are often recognized to serve as master regulators that are critical for developmental processes and heritable cell fate decisions. The work presented in this thesis will serve as a foundation to further investigate how the interplay between chromatin organizing TFs and dynamic nucleosomes regulates developmental processes and heritable cell fate decisions in *C. albicans*, and may lead to insights into how these processes regulate cellular differentiation in higher eukaryotes.

Chapter 1: Introduction — The roles of chromatin accessibility in regulating the *Candida albicans* white-opaque phenotypic switch

Mohammad N. Qasim, Ashley Valle Arevalo, Clarissa J. Nobile and Aaron D. Hernday
Journal of Fungi, 2021

1.1 Abstract

Candida albicans, a diploid polymorphic fungus, has evolved a unique heritable epigenetic program that enables reversible phenotypic switching between two cell types, referred to as “white” and “opaque”. These cell types are established and maintained by distinct transcriptional programs that lead to differences in metabolic preferences, mating competencies, cellular morphologies, responses to environmental signals, interactions with the host innate immune system, and expression of approximately 20% of genes in the genome. Transcription factors (defined as sequence specific DNA-binding proteins) that regulate the establishment and heritable maintenance of the white and opaque cell types have been a primary focus of investigation in the field; however, other factors that impact chromatin accessibility, such as histone modifying enzymes, chromatin remodelers, and histone chaperone complexes, also modulate the dynamics of the white-opaque switch and have been much less studied to date. Overall, the white-opaque switch represents an attractive and relatively “simple” model system for understanding the logic and regulatory mechanisms by which heritable cell fate decisions are determined in higher eukaryotes. Here we review recent discoveries on the roles of chromatin accessibility in regulating the *C. albicans* white-opaque phenotypic switch.

Keywords: white-opaque switching; *Candida albicans*; chromatin; transcriptional regulation; heritability; cell fate decisions; histone modifying enzymes; chromatin remodeling enzymes; histone chaperone complexes; epigenetics

1.2 Introduction

Multicellular organisms are comprised of many phenotypically and functionally distinct cell types, the vast majority of which contain the same primary genomic sequence. How a single set of genomic “instructions” can reliably yield many distinct and heritable phenotypic states is a fundamental question in biology. We have begun to understand that a single genome can support many transcriptional programs, which in turn specify unique cell type specific patterns of gene expression, and ultimately establish distinct phenotypes. These cell types are often heritably maintained in an epigenetic manner following each cell division, and it has become increasingly apparent that chromatin structure and accessibility play important roles in the transcriptional regulation of cell type specificity.

Candida albicans, a unicellular polymorphic fungus, has evolved the ability to establish two transcriptional programs that give rise to two distinct cell types called “white” and “opaque” based on their appearance at the single colony level. The white and opaque cell types are heritably maintained in an epigenetic manner through thousands of cell

divisions with no change to the primary sequence of the genome[1,2]. A growing body of literature has identified numerous similarities between the molecular mechanisms governing the *C. albicans* white-opaque switch and those that underlie heritable cell type differentiation in higher eukaryotes[3–7]. Since a similar heritable phenotypic switch is not observed in the classic model yeast *Saccharomyces cerevisiae*, *C. albicans* has emerged as a compelling "simple" and genetically tractable eukaryotic model system to study heritable transcriptional programs in higher eukaryotes.

The *C. albicans* white and opaque cell types are established and maintained by distinct transcriptional programs that lead to a wide range of phenotypic differences between the two cell types. These include differences in metabolic preferences, mating competencies, cellular morphologies, responses to environmental signals, interactions with the host innate immune system, and expression of ~20% of genes in the genome[3,4,8–15]. While certain environmental conditions can bias the switch in favor of the white or opaque cell type[1,2], phenotypic switching between the two cell types occurs stochastically at a frequency of approximately one switch event per 1,000-10,000 cell divisions under standard switch permissive growth conditions[16–19]. In other words, once established, each cell type is maintained through an epigenetic mechanism that is stably inherited over thousands of subsequent cell divisions.

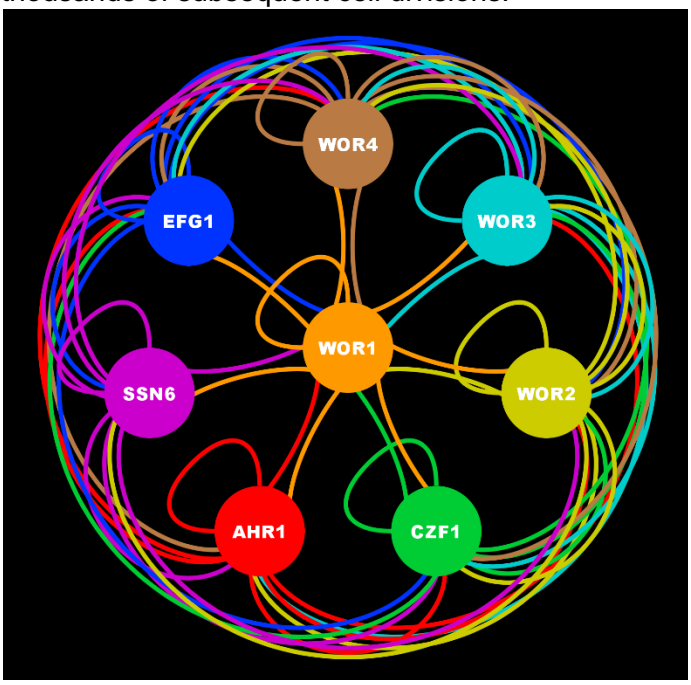


Figure 1-1. Core opaque transcriptional circuit. Color coded lines indicate direct binding interactions between each TF (same color as their circular node) and their respective target genes within the opaque circuit. Data to create this figure was obtained from[4,20,21,23].

The frequency of switching between the white and opaque cell types is controlled by a set of regulatory genes that encode seven sequence-specific DNA-binding proteins, i.e., transcription factors (TFs) (Wor1, Wor2, Wor3, Wor4, Ahr1, Czf1, Efg1), and one non-DNA-binding adapter protein (Ssn6). These eight switch regulators have been extensively characterized through genome-wide transcriptional profiling and chromatin immunoprecipitation (ChIP) experiments, as well as by genetic epistasis experiments [1,2,4,20–24]. Based on this work, these eight proteins have been shown to form the core of the transcriptional circuit that governs the establishment and heritable maintenance of the opaque cell type[4,9,20,21,23,25] (**Figure 1-1**). At the heart of this circuit lies Wor1, the "master regulator" of the opaque cell type, which is considered to be the key regulator involved in initiating the switch to, and heritable maintenance of, the opaque cell type[1,2,4,25]. *WOR1* expression, which is repressed

in white cells and upregulated in opaque cells [1,2], triggers the formation of a highly intertwined regulatory network – consisting of the core regulators and all of their directly bound target genes – that is responsible for the establishment and heritable maintenance of the opaque cell type[1–4]. Remarkably, the high degree of interconnectivity in the core opaque transcriptional circuit (**Figure 1-1**) is similar to that observed in transcriptional circuits controlling stem cell maintenance and differentiation in mammals[3,26–28], suggesting that similar regulatory architectures may be common across eukaryotes to control analogous heritable transcriptional programs and their associated phenotypic outputs. Subsequent to the identification of the core transcriptional circuit controlling white-opaque switching, an additional 108 genes that encode known or predicted sequence-specific DNA-binding proteins have been identified as "auxiliary" regulators of the white-opaque switch[1,2,9,20] (**Table 1-1**). This auxiliary designation indicates that while these genes are known to influence the frequency of switching, the regulators that they encode have yet to be incorporated into the white-opaque transcriptional regulatory network via genome-wide chromatin association studies.

Table 1-1. Genes encoding known or predicted TFs and a non-DNA-binding adapter protein with roles in white-opaque switching.

Core White-Opaque Transcriptional Regulators			
Gene Name	**Orf19 #**	**Known effect on white-opaque switch in mutant strain**	
		White to Opaque [1]	**Opaque to White** [1]
AHR1	Orf19.7381	1.96	0.13
CZF1	Orf19.3127	0.05	0.06
EFG1	Orf19.610	24.0	<0.02
SSN6*	Orf19.6798	N/A^{OP}	<0.04
WOR1	Orf19.4884	> 0.05	N/A^{WH}
WOR2	Orf19.5992	> 0.03	N/A^{WH}
WOR3	Orf19.467	0.42	0.26
WOR4	Orf19.6713	> 0.08	N/A^{WH}
Auxiliary White-Opaque Transcriptional Regulators			
Gene Name	**Orf19 #**	**Known effect on white-opaque switch in mutant strain**	
		White to Opaque [1]	**Opaque to White** [1]
AAF1	Orf19.7436	0.88	0.38
AFT2	Orf19.2272	0.36	0.59
ARG81	Orf19.4766	1.75	0.44
ASG1	Orf19.166	0.05	0.05
ASH1	Orf19.5343	0.83	0.04
BAS1	Orf19.6874	1.49	1.15
BCR1	Orf19.723	2.21	N/A^{WH}
BRG1	Orf19.4056	1.87	0.66
CAP1	Orf19.1623	0.67	0.23
CAS5	Orf19.4670	1.45	0.43
CPH1	Orf19.4433	0.46	0.39
CPH2	Orf19.1187	0.44	0.70
CRZ1	Orf19.7359	1.86	0.18
CSR1	Orf19.3794	1.02	2.56
CTA4	Orf19.7374	0.88	0.17
CTA7	Orf19.4288	2.37	0.48
CUP2	Orf19.5001	0.81	0.63
CUP9	Orf19.6514	4.74	0.07
DAL81	Orf19.3252	0.16	0.57
DPB4	Orf19.2088	0.32	0.41

ECM22	Orf19.2623	1.31	0.46
EFH1	Orf19.5498	1.69	0.61
FCR1	Orf19.6817	0.52	0.63
FGR15	Orf19.2054	0.06	4.78
FLO8	Orf19.1093	<0.04	N/A[WH]
GAL4	Orf19.5338	<0.04	0.77
GIS2	Orf19.3182	0.86	0.11
GRF10	Orf19.4000	1.37	0.14
GZF3	Orf19.2842	0.08	1.70
HAP2	Orf19.1228	<0.04	0.62
HAP3	Orf19.4647	0.34	1.30
HAP31	Orf19.517	0.04	0.79
HAP41	Orf19.740	0.10	1.14
HAP42	Orf19.1481	0.49	0.54
HAP5	Orf19.1973	0.15	0.33
HCM1	Orf19.4853	18.4	0.27
HFL1	Orf19.3063	<0.05	2.11
INO2	Orf19.7539	<0.04	0.28
INO4	Orf19.837.1	0.33	0.81
ISW2	Orf19.7401	3.40	2.89
KAR4	Orf19.3736	0.51	1.23
LYS143	Orf19.4776	7.25	0.88
LYS144	Orf19.5380	1.29	0.42
MAC1	Orf19.7068	0.05	0.63
MIG1	Orf19.4318	<0.03	1.37
MIG2	Orf19.5326	1.62	0.63
MSN4	Orf19.4752	1.88	0.21
NDT80	Orf19.2119	0.10	1.87
NTO1	Orf19.5910	2.80	0.66
OPI1	Orf19.1543	4.03	0.42
PTH2	Orf19.4231	2.96	0.24
RAP1	Orf19.1773	16.0	0.64
RBF1	Orf19.5558	N/A[OP]	<0.03
RCA1	Orf19.6102	0.11	0.53
REP1	Orf19.7521	0.68	2.43
RFG1	Orf19.2823	1.05	2.12
RFX1	Orf19.3865	1.78	1.67
RFX2	Orf19.4590	1.46	0.58
RME1	Orf19.4438	2.15	0.57
RPN4	Orf19.1069	18.2	0.67
RTG1	Orf19.4722	0.47	0.35
RTG3	Orf19.2315	0.36	0.46
SEF2	Orf19.1926	1.09	0.31
SFL1	Orf19.454	0.88	1.95
SKN7	Orf19.971	1.13	0.65
SKO1	Orf19.1032	0.56	0.42
STP2	Orf19.4961	9.18	0.15
STP4	Orf19.909	3.25	0.47
SWI4	Orf19.4545	0.22	1.02
TYE7	Orf19.4941	2.04	0.97
UGA33	Orf19.7317	0.99	0.60
UME6	Orf19.1822	0.62	2.02
UME7	Orf19.2745	0.50	1.44
UPC2	Orf19.391	0.94	3.13
WAR1	Orf19.1035	0.31	0.94
XBP1	Orf19.5210	0.16	0.85
ZCF16	Orf19.2808	1.47	1.19

ZCF17	Orf19.3305	1.31	2.22
ZCF2	Orf19.431	0.59	0.36
ZCF20	Orf19.4145	0.66	0.46
ZCF21	Orf19.4166	0.25	0.02
ZCF22	Orf19.4251	1.77	0.93
ZCF24	Orf19.4524	0.96	0.31
ZCF25	Orf19.4568	8.48	0.34
ZCF27	Orf19.4649	0.53	1.46
ZCF30	Orf19.5251	1.08	0.60
ZCF31	Orf19.5924	0.42	3.24
ZCF34	Orf19.6182	0.35	0.22
ZCF7	Orf19.1685	4.73	0.37
ZCF8	Orf19.1718	0.48	0.46
ZFU2	Orf19.6781	0.52	2.26
ZFU3	Orf19.6888	0.20	0.06
ZMS1	Orf19.5026	0.36	0.84
N/A	Orf19.1150	1.22	0.78
N/A	Orf19.1274	0.70	1.20
N/A	Orf19.1577	0.89	0.68
N/A	Orf19.1757	1.04	0.61
N/A	Orf19.217	0.60	0.57
N/A	Orf19.2476	1.91	2.49
N/A	Orf19.2612	2.38	1.40
N/A	Orf19.2961	7.02	2.05
N/A	Orf19.3928	5.71	0.23
N/A	Orf19.7098	7.77	1.10

[1]Fold changes in switch frequencies for each deletion mutant strain are calculated relative to an isogenic wildtype reference strain. **"N/A"** in the gene names column indicates that a given gene name does not yet exist for this open reading frame (ORF). "N/A^{OP}" indicates that the deletion mutant strain was opaque-locked and the white to opaque or opaque to white switch frequency could not be determined. "N/A^{WH}" indicates that the deletion mutant strain was white-locked, or failed to yield stable opaque colonies, and the opaque to white switch frequency could not be determined. **"*"** indicates a non-DNA-binding adaptor protein. Switch frequency data in this table was obtained from[9].

While much of the research on white-opaque switching to date has focused on regulatory TFs that bind directly to DNA in a sequence specific manner, an increasing body of literature has revealed important roles for factors that impact chromatin accessibility in regulating the dynamics of the switch[5,22,29–34]. Relatively little is known, however, at a mechanistic level, about how these factors influence switching. Since the ability of TFs to access their regulatory targets is substantially affected by chromatin landscapes[35], the roles that these chromatin accessibility factors play to influence white-opaque switching is an important avenue for future investigation.

Chromatin is composed of genomic DNA wrapped around four histone dimers, forming nucleosomes, as well as non-histone proteins that organize and stabilize chromatin structure. The regulation of chromatin is essential for cellular processes such as transcription, replication and repair, mitosis, and apoptosis[36]. The four histone dimers that make up the nucleosome core can be post-translationally modified either before or after their deposition into chromatin, adding an extra layer of information that is encoded on top of the primary sequence of the genome. These epigenetic modifications, which have unique context dependent functions, include acetylation, methylation, crotonylation, ubiquitinoylation, SUMOylation, and phosphorylation[37–43]. Although these modifications are

considered epigenetic marks, most of them are not heritably transmitted from one cellular or organismal generation to the next. Histone modification is a reversible process that is mediated by specific enzymes that are classified as “writers”, such as histone lysine methyltransferases, that catalyze the addition of chemical modifications, and “erasers”, such as histone deacetylases, that remove chemical modifications[44]. Often, the writers and erasers function within multiprotein complexes that also contain proteins with specialized domains, such as bromodomains, that are classified as “readers”. Readers recognize specific histone modifications and regulate the specificity and enzymatic activity of their associated writers and erasers[45]. Additional factors that influence chromatin structure and accessibility include remodelers, which actively translocate or evict nucleosomes, and histone chaperones, which deposit histones into chromatin[46–49].

28 genes that encode known or predicted writers (eleven genes), erasers (fourteen genes), chromatin remodelers (one gene) and histone chaperones (two genes) have been analyzed for their roles in white-opaque switching[5,22,29–31,33,34,50] (**Table 1-2**). Eighteen of these genes are involved in white-opaque switching since their deletion significantly affected the frequency of white-opaque switching relative to an isogenic wildtype strain (**Table 1-2**). These genes encode proteins that fall into six functional categories with respect to their roles in white-opaque switching: 1) stabilizers of the white cell type that do not affect opaque cell stability (Set1, Rpd31, Hst3, Hda1, Hda2, and Hda3); 2) destabilizers of the white cell type that do not affect opaque cell stability (Hst2); 3) stabilizers of the opaque cell type that do not affect white cell stability (Pho13); 4) destabilizers of the opaque cell type that do not affect white cell stability (Hst1); 5) stabilizers of the white cell type that also decrease opaque cell stability (Hat1, Swr1 and Yng2); and 6) destabilizers of the white cell type that also increase opaque cell stability (Hos2, Set3, Nat4, Rtt109, Rpd3 and Cac2) (**Figure 1-2**). Below we review the current knowledge of the roles of these writers, erasers, chromatin remodelers, and histone chaperones in regulating the white-opaque switch.

Table 1-2. Genes encoding known or predicted writers, erasers, chromatin remodelers and histone chaperones analyzed for their impacts on white-opaque switching.

Gene Name	ORF #	Protein Function	Known effect on white-opaque switch in mutant strain		Reference
			Wh --> Op[1]	Op --> Wh[1]	
YNG2	Orf19.878	**Histone Acetyltransferases (Writers)**	23.8	0.01	[5]
SPT10	Orf19.2361		no effect	no effect	[29]
HPA2	Orf19.6323		no effect	no effect	[29]
RTT109	Orf19.7491		0.10	6.98	[51]
NAT4	Orf19.4664		0.12	3.42	[29]
SAS2	Orf19.2087		no effect	no effect	[29]
HAT1	Orf19.779		7.60	0.13	[34]

ELP3	Orf19.7387		no effect	no effect	[29]
SET1	Orf 9.6009	**Histone Methyl Transferases (Writers)**	1.73	0.98[3]	[29]
SET2	Orf19.175		no effect	no effect	[29]
DOT1	Orf19.740		no effect	no effect	[29]
HDA1	Orf19.2606	**Histone Deacetylases (Erasers)**	2.73	1.06[3]	[29,31]*
HDA2	Orf19.6952		3.33	no data	[52]
HDA3	Orf19.7344		3.67	no data	[52]
RPD3	Orf19.2834		33.3	49.7	[31]
RPD31	Orf19.6801		2.85	1.23[3]	[30]
HST1	Orf19.4761		1.29[3]	0.37	[29]
HST2	Orf19.2580		0.04	1.86[3]	[29]
HST3[4]	Orf19.1934		6.00	no effect	[22]
HOS1	Orf19.4411		no effect	no effect	[29]
HOS2	Orf19.5377		0.13	2.29	[29]
HOS3	Orf19.2772		no effect	no effect	[29]
SET3	Orf19.7221		0.16	2.71	[29]
PHO13[2]	Orf19.4444	**Phosphatases (Erasers)**	0.93[3]	5.01	[29]
ORF19.4736	Orf19.4736		no effect	no effect	[29]
SWR1	Orf19.1871	**Chromatin Remodelers**	16.0	0.01	[5]
CAC2	Orf19.6670	**Histone Chaperones**	3.75	2.53	[33]
HIR1	Orf19.2099		no effect	no effect	[33]

[1]Fold changes in switch frequencies for each deletion mutant strain are calculated relative to an isogenic wildtype reference strain. [2]The protein encoded by *PHO13* has been shown to lack protein phosphatase activity and is instead believed to be involved in metabolism[53]. [3]The effect on white-opaque switching is not significant. [4]This strain is an *HST3/hst3* hemizygous mutant strain. *The *hda1/hda1* deletion mutant strain was investigated in both referenced articles with similar findings. All switch frequencies reported in this table originate from the indicated references.

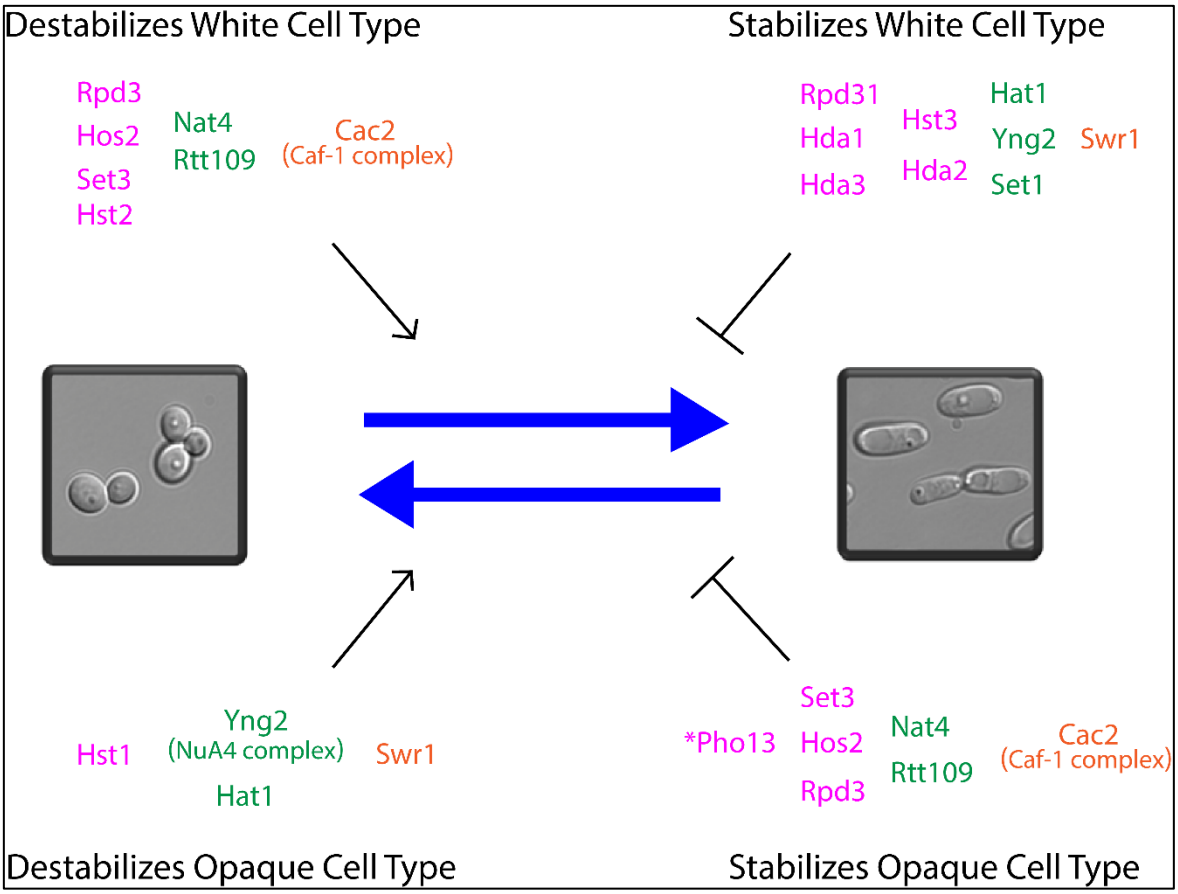


Figure 1-2. Roles of chromatin modifying enzymes in regulating the *C. albicans* white-opaque switch. White to opaque and opaque to white switching is indicated by blue arrows. Black arrows indicate proteins that act to promote switching in the white to opaque direction (upper left quadrant) or in the opaque to white direction (lower left quadrant), while black lines with crossbar indicate proteins that repress switching in the white to opaque direction (upper right quadrant) or in the opaque to white direction (lower right quadrant). Erasers are shown in pink, writers are shown in green, and chromatin remodelers and histone chaperones are shown in orange.
*Pho13 has been shown to lack protein phosphatase activity and is instead believed to be involved in metabolism[53].

1.3 Regulation of white-opaque switching by "writers"

Five of the histone modifiers that influence white-opaque switching are "writers", four of which are histone acetyltransferases (HATs) (Hat1, Rtt109, Nat4, Yng2) and one of which is a histone methyltransferase (HMT) (Set1). HATs modify histones by acetylating lysine residues at histone tails or at histone globular domains, while HMTs primarily modify histones by methylating lysine residues at histone tails. Each of the four HATs influence the stability of both the white and opaque cell types, with Hat1 and Yng2 playing opposing roles to Nat4 and Rtt109 (**Figure 2**)[5,22,34]. In contrast, the HMT Set1 specifically assists in

the establishment of the opaque cell type by increasing the white to opaque switch frequency but does not affect opaque cell maintenance. In the following sections, we review the current knowledge of how these writers regulate the establishment and maintenance of the white and opaque phenotypic states.

1.3.1 Regulation of white-opaque switching by the NuA4 histone acetyltransferase Yng2

Histone acetyltransferases (HATs) are characterized by their substrate and cellular localization. Type A HATs modify nucleosomal histones and are localized in the nucleus, while type B HATs modify histones before they are deposited into nucleosomes and are localized in the cytoplasm[54]. Yng2, the catalytic subunit of the NuA4 HAT complex, is the only type A HAT known to regulate the white-opaque switch[5] (**Table 2 and Figure 2**). Although the mechanism by which NuA4 regulates the white-opaque switch is unknown, some mechanistic insights can be gleaned from the knowledge of how NuA4 regulates the yeast to hyphal cell transition[55]. The NuA4 complex is recruited to the upstream intergenic regions of hyphal-specific genes by Efg1 and, upon filament induction, an increase in H4 acetylation levels is observed at these sites[55]. This acetylation event is required for the recruitment of the Swi/Snf complex, which activates the hyphal-specific genes through its chromatin remodeling activity. Given that Efg1 is also a key white-opaque switch regulator, and is bound upstream of many white- and opaque-specific genes, it is perhaps not surprising that genetic evidence suggests a similar process may be involved in regulating the white-opaque switch[55]. Notably, deletion of *YNG2* results in a nearly identical alteration of white-opaque switching as deletion of *EFG1* (~24-fold increase in white to opaque switching and ≥60-fold decrease in opaque to white switching) (**Table 1** and **Table 2**). These findings suggest that H4K5 acetylation at white and opaque regulated genes that are directly bound by Efg1 likely plays an important role in stabilizing the white cell type and destabilizing the opaque cell type. Since NuA4 regulates the expression of hyphal-specific genes indirectly, through H4K5 acetylation-dependent recruitment of the Swi/Snf chromatin remodeling complex, it seems plausible that Swi/Snf-dependent chromatin remodeling may ultimately play a role in modulating the stability of the white and opaque cell types.

The NuA4 complex has also been implicated in the regulation of white-opaque switching through H2 and H4 acetylation-dependent recruitment of the Swr1 chromatin remodeling complex, which is responsible for depositing an H2A.Z histone variant into chromatin in exchange for a canonical H2A histone[56]. Deletion of *SWR1* results in an increase in white to opaque switching and a decrease in opaque to white switching, with similar fold changes in switch frequencies as observed for a *yng2* deletion mutant strain[5] (**Table 2**). The NuA4 HAT complex and the Swr1 chromatin remodeling complex share four subunits, and it was recently determined that the two complexes can merge to function as one unit depending on the morphological state of the *C. albicans* cell[57]. Taken together, these findings suggest that there is a complex interplay between histone modification and chromatin remodeling enzymes in regulating the stability of the white and opaque cell types. Studies in fungi and higher eukaryotes have shown that chromatin remodeling enzymes are often recruited to their target loci by a combination of factors, including histone modifications; however, the specific roles of these interactions in regulating cell type heritability is not understood in eukaryotes.

1.3.2. Regulation of white-opaque switching by the Rtt109 histone acetyltransferase

Rtt109 is a type B HAT known to acetylate lysine 56 within the globular domain of histone 3 (H3K56) before the histone monomer is deposited into nucleosomes. This acetylation mark is especially important because the addition of a negative charge within the globular domains of histones reduces the electrostatic attraction between DNA and nucleosomes and thus destabilizes the affected nucleosomes[37]. This increase in DNA accessibility due to H3K56 acetylation influences the repackaging of chromatin after replication and DNA damage repair[58–60], and also plays an important role in anti-silencing and transcription at heterochromatic loci[61,62]. Interestingly, an increase in the levels of H3K56 acetylation is correlated with an increase in the rate of histone turnover at certain developmentally regulated genomic loci in higher eukaryotes[63]. This increase in the rate of histone turnover is, in turn, correlated with an increase in chromatin accessibility, and thus TFs (both activating and repressing) are more likely to bind to, and regulate the expression of, genes at these loci.

Rtt109 has been shown to play an important role in enabling the white to opaque transition and in the heritable maintenance of the opaque cell type[22]. In a *rtt109* deletion mutant strain, white cells were found to switch to the opaque cell type at a tenfold lower frequency than that of wildtype cells, and the resulting opaque cells were highly unstable[22]. In fact, opaque colonies of an *rtt109* deletion mutant strain were found to consist of a mixed population of both white and opaque cells, and only a subset of the elongated cells that resembled the opaque phenotype were found to express high levels of *WOR1*[22]. These findings suggest that the H3K56 acetylation mark plays a critical role in the stability of *WOR1* expression in opaque cells. Locus specific chromatin immunoprecipitation experiments by ChIP-qPCR revealed that H3K56 acetylation marks were differentially deposited at multiple loci upstream of genes that are differentially expressed between white and opaque cell types[22]. Indeed, H3K56 acetylation was found to be enriched within the upstream intergenic region of *WOR1* in opaque cells, relative to white cells[22]. How H3K56 acetylation is deposited in a cell type specific manner at the upstream intergenic regions of these differentially expressed genes remains an open question. While we do not yet know the specific mechanisms by which Rtt109 regulates the white-opaque switch, evidence suggests that Rtt109 regulates Wor1 accessibility to its own upstream intergenic region[22]. For example, we know that opaque cells expressing an ectopic copy of *WOR1* in an *rtt109* deletion mutant strain failed to heritably maintain the opaque cell type after the ectopic copy of *WOR1 was* turned off[22], suggesting that Rtt109 activity and H3K56 acetylation enrichment within the *WOR1* upstream intergenic region are necessary for the stable maintenance of the *WOR1* positive feedback loop that is central to the heritability of opaque cells.

1.3.3. Regulation of white-opaque switching by the Hat1 histone acetyltransferase

Hat1, a type B HAT that is part of the NuB4 complex, acetylates histone 4 (H4) tails at two different lysine residues (H4K5 and H4K12), and mediates the incorporation of free histones into nucleosomes[64]. The role of Hat1 in the NuB4 complex was confirmed by showing that reducing H4 levels mimicked inactivation of the NuB4 complex[34]. In *C. albicans*, a *hat1* deletion mutant strain displayed both increased white to opaque switching and decreased opaque to white switching[34]. In other words, the NuB4 complex seems to bias the switch towards the white cell type by both stabilizing the white cell type and destabilizing the opaque cell type. While the mechanism by which Hat1 influences the switch is unknown, we speculate that Hat1 may regulate the white-opaque switch by

modulating H4 levels in chromatin, which would affect chromatin accessibility. In white cells, reduced H4 levels in a *hat1* deletion mutant strain could increase DNA accessibility for TFs binding to *WOR1* cis-regulatory elements, thus making it easier for Wor1 to initiate the autoregulatory positive feedback loop central to the white to opaque transition. In opaque cells, this increase in DNA accessibility could stabilize the *WOR1* positive feedback loop, thereby stabilizing the opaque state. Consistent with this idea, one would predict that white to opaque switching would be reduced in an *MTL* heterozygous *hat1* deletion mutant strain due to increased binding of the *MTL* **a**1/α2 heterodimer upstream of *WOR1*.

1.3.4. Regulation of white-opaque switching by the Nat4 histone acetyltransferase

Deletion of *NAT4*, which encodes an N-terminal acetyltransferase (NAT), has been found to reduce white to opaque switching and increase opaque to white switching[29]. Thus, Nat4 tilts the scales in favor of opaque cell formation by destabilizing the white cell type and by stabilizing the opaque cell type. In *S. cerevisiae,* Nat4 acetylates the N-terminal serine residues of H4 and H2A[65]. While not much is known about the function of Nat4 in *C. albicans*, we briefly discuss a few unique properties of the *S. cerevisiae* Nat4 to highlight why NATs could be of interest to study in *C. albicans*.

Five NAT types have been identified in *S. cerevisiae* (NatA, NatB, NatC, NatD, NatE)[66], and with the exception of NatD, they all have human orthologs. Nat4, the catalytic subunit of NatD is the only NAT shown to regulate the white-opaque switch. Unlike the other NATs, NatD does not contain auxiliary subunits, and therefore does not require interacting partners for its enzymatic activity[67]. Interestingly, NatD recognizes a significantly longer N-terminal sequence for acetylation than the other NATs[67]. These unique properties of NatD suggest that this enzyme likely regulates the white-opaque switch by regulating the acetylation levels of H4 and H2A.

1.3.5. Regulation of white-opaque switching by the Set1 histone methyltransferase

Set1 deposits methylation marks at lysine 4 of H3 (H3K4) via its SET domain-containing methyltransferase[68]. It is the only methyltransferase in *C. albicans* that modifies H3K4, and thus deletion of *SET1* in *C. albicans* results in a complete loss of H3K4 methylation[69]. In *S. cerevisiae* and higher eukaryotes, H3K4 methylation marks are hallmarks of transcriptionally active chromatin sites[68]. Studies in *S. cerevisiae* have shown that *SET1* is required for transcriptional silencing of the silent mating-type loci and telomeres[70]; however, the specific relationship between H3K4 methylation and transcription has not been investigated in *C. albicans*. A *C. albicans set1* deletion mutant strain has been shown to display increased white to opaque switching relative to the wildtype strain[29]. Unlike the HATs, which influence both white to opaque switching and opaque cell stability, the Set1 histone methyltransferase, however, does not affect opaque cell stability[29]. The mechanism by which Set1-dependent H3K4 methylation specifically influences white cell stability, without affecting opaque cell stability, is an intriguing area of interest for future studies.

1.4. Regulation of white-opaque switching by “erasers”

Histone deacetylases remove histone acetylation marks and thus are often associated with a repressive function because of their roles in forming a condensed or

"closed" chromatin state that restricts DNA accessibility[71]. Ten histone deacetylases (HDACs) (Rpd3, Rpd31, Hda1, Hda2, Hda3, Set3, Hos2, Hst1, Hst2, and Hst3) have been shown to influence white-opaque switching[22,29,31,52]. Three of these HDACs (Hda1, Hst3, and Rpd31) have no known roles in regulating the stability of the opaque cell type, and thus, similar to the Set1 HMT discussed above, appear to be white cell-specific modulators of the white-opaque switch[22,29,31]. These findings suggest that decreased chromatin accessibility, mediated by these HDACs either genome-wide or at specific regulatory loci, plays a role in maintaining cell type epigenetic heritability by modulating accessibility for TFs (activating and repressing). Below, we discuss the known roles of these HDACs in white-opaque switching.

Rpd3 is a histone deacetylase that acts on both H3 and H4, and has been shown to play a direct role in regulating *WOR1* expression in white cells [30]. Upon deletion of *RPD3*, H4 acetylation levels increase throughout the *WOR1* upstream intergenic region, and these elevated acetylation levels appear to directly influence white cell stability by increasing the accessibility of chromatin. Interestingly, the effect of *RPD3* deletion on white cell stability is dependent upon the mating type of the cell. In *MTL* heterozygous cells, where the *MTL***a**1/α2 heterodimer stabilizes the white cell type through direct repression of *WOR1*[14,25], deletion of *RPD3* results in an increase in *MTL***a**1/α2 binding upstream of *WOR1* and a decrease in *WOR1* expression[30]. Conversely, in an *MTL* homozygous (**a**/**a**) strain, where the *MTL***a**1/α2 heterodimer is not present, deletion of *RPD3* results in a decrease in white cell stability[30]. This decrease is presumably due to an increase in Wor1 accessibility to the *WOR1* upstream intergenic region, which would ultimately facilitate activation of the *WOR1* positive feedback loop that is central to the formation and stabilization of the opaque cell type. These results establish that white cell stability can be regulated through the modulation of interactions between regulatory TFs and their cis-regulatory target sites via histone deacetylation. Alternatively, HDACs could also remove histone acetylation marks from within open reading frames, thus changing their chromatin accessibility. Set3 and Hos2 were recently shown to form part of an HDAC complex in *C. albicans* that regulates expression of metabolic and morphogenesis related genes via this type of mechanism[72], and a similar mechanism has been proposed to explain how Rpd31 represses white to opaque switching[30]. While Set3, Hos2 and Rpd31 have been shown to regulate the white-opaque switch, their specific mechanisms of action on the switch have yet to be elucidated.

The Set3 complex (Set3C), which includes the HDACs Set3, Hos2, and Hst1, has been shown to be intertwined at a genetic level with the transcriptional regulatory circuit that controls white-opaque switching. Genetic epistasis studies have highlighted an intriguing interaction between Set3C and a key repressor of the white to opaque switch, Efg1[29]. While deletion of *EFG1* alleviates Efg1-mediated repression of *WOR1*, and thus stimulates white to opaque switching and stabilizes the opaque state, deletion of either *SET3* or *HOS2* suppresses the *efg1* deletion phenotype and restores switching frequencies to near wildtype levels[29]. This suggests that the activation and sustained expression of *WOR1* that is central to opaque cell formation and stability is dependent not only on alleviation of Efg1 repression, but also on Set3C activity. Since Set3C has been shown to be a negative regulator of the protein kinase A (PKA) pathway[73], and Efg1 is believed to be a major regulatory target of the PKA pathway[74–76], this suggests a potential mechanism for the genetic interactions observed between *SET3*, *HOS2*, and *EFG1*.

Additional epistasis experiments revealed that deletion of *SET1*, encoding a methyltransferase, and the resulting loss of H3K4 histone methylation, suppresses the effect of *SET3* or *HOS2* deletion on white-opaque switching, revealing a complex interaction between these chromatin modifiers and the modulation of white and/or opaque cell stability. Although *SET3*, *HOS2* and *HST1* have all been shown to function as part of Set3C, deletion studies indicate that Hst1 may act independently of Set3C when regulating white-opaque switching[29,77]. Specifically, while Set3 and Hos2 both promote white to opaque switching and opaque cell stability, Hst1 appears to have no effect on white cell stability and acts to decrease opaque cell stability[29]. One potential explanation for this result could be that Hst1 may instead regulate opaque cell stability as a component of the Sum1/Rfm1/Hst1 complex, which has been shown to function as a repressor of sporulation-specific genes in *S. cerevisiae*[78] and a repressor of drug and oxidative stress resistance genes in *Candida glabrata*[79].

Similar to *SET3* and *HOS2* deletion phenotypes, deletion of *HST2* was found to result in a decrease in white to opaque switching, yet no alterations in opaque cell stability were detected[29]. This finding suggests that Hst2 is likely involved in destabilizing the white cell type. In contrast to *SET3* and *HOS2*, which are epistatic to *EFG1*, the *HST2* deletion phenotype is suppressed when *EFG1* is also deleted[29], suggesting that *HST2* may destabilize the white cell type by inhibiting or antagonizing *EFG1*. Hda1, which also plays roles in stabilizing the white cell type without influencing opaque cell stability, may also act through *EFG1*. *EFG1* expression is reduced in white cells when *HDA1* is deleted; however, the mechanism by which Hda1 influences *EFG1* expression has yet to be elucidated. Furthermore, *HDA1* expression is overall lower in opaque cells, relative to white cells[31], perhaps explaining why deletion of *HDA1* does not affect opaque cell stability. Hda1 has been proposed to form a complex along with Hda2 and Hda3[52], potentially explaining the similarity in phenotypes between *HDA1*, *HDA2*, and *HDA3* deletion strains. In addition, deletion of any one of these three genes has been reported to result in at least a fivefold increase in *WOR1* expression[52], possibly due to a reduction in Efg1-dependent repression of *WOR1* transcription. Although Hda1 has been shown to promote sustained hyphal development via deacetylation of Yng2, leading to eviction of the NuA4 HAT complex from hyphal promoters[80,81], it is unclear whether a similar mechanism may be involved in white-opaque switching, as both Hda1 and Yng2 contribute to stabilizing the opaque cell type. Hst3 is a sirtuin class HDAC that removes H3K56 acetylation and stabilizes the white cell type[22], presumably by counteracting the white cell destabilizing effects of Rtt109, which facilitates sustained *WOR1* expression though increased H3K56 acetylation. It is interesting to note that Hst3 levels are reduced in response to genotoxic stress[22], providing a possible explanation for the up to eightfold increase in white to opaque switching observed when white cells are cultured in the presence of the cytotoxic drugs methyl methanesulfonate or hydroxyurea[22]. Understanding how these HDACs regulate the expression of key white-opaque transcriptional regulators, such as *WOR1* and *EFG1* to ultimately influence the relative stabilities of white and opaque cell types, is an important area of focus for future studies.

1.5. Potential roles of “readers” in regulating white-opaque switching

“Reader” enzymes recognize or “read” post-translationally modified histone residues and possess a histone binding module, such as a bromodomain[82,83], that recognize specific histone residues with modifications. Bromodomain modules, for example, recognize

histone lysine residues with acetylation marks[84]. While the roles of readers in regulating white-opaque switching have yet to be investigated, readers are involved in other important processes in *C. albicans*. For example, one reader of the lysine acetylation mark (Bdf1) and two readers of the crotonylation mark (Taf14 and Yaf9), have been shown to play important roles in *C. albicans* pathogenicity[85,86]. NuA4-dependent acetylation upstream of hyphal-specific genes has been shown to result in the recruitment of the Swi/Snf chromatin remodeling complex via its bromodomain during hyphal induction[55]. In other fungi and higher eukaryotes, many readers have been identified and studied[87–90], but their functions have yet to be investigated in *C. albicans*. Several histone modifying enzymes and chromatin remodeling enzymes, which affect the white-opaque switch, also affect cellular differentiation and heritability in higher eukaryotes[91–95]. Readers are often components within histone modifying enzyme complexes and chromatin remodeling complexes, and thus, assist these large complexes in finding their target loci within the genome[96]. Therefore, it seems likely that readers are involved in regulating the *C. albicans* white-opaque switch, and this is an important unexplored area of interest for future studies.

1.6. Regulation of white-opaque switching by chromatin remodeling complexes

Chromatin remodeling enzyme complexes modulate chromatin accessibility through the function of their ATPase-translocase domains. We can classify chromatin remodeling enzymes into four subfamilies, each of which carries out specialized functions[48]. ISWI and CHD complex subfamilies preferentially reduce chromatin accessibility by regulating the assembly and organization of nucleosomes. The Swi/Snf complex subfamily remodels chromatin by sliding or evicting nucleosomes, which generally increases chromatin accessibility[48]. The Ino80 complex subfamily modulates chromatin accessibility by replacing canonical histones with histone variants, specifically targeting nucleosomes that flank transcription start sites. The Swr1 complex, a member of the Ino80 subfamily, is the only subfamily of chromatin remodelers that regulates white-opaque switching in *C. albicans*[5] and is discussed in more detail below.

1.6.1. Regulation of white-opaque switching by the Swr1 chromatin remodeling complex

SWR1 encodes a chromatin remodeling enzyme that is responsible for the deposition of the histone variant H2AZ. The Swr1 complex, which is an ortholog of the human SRCAP complex, is a multiprotein complex responsible for replacing canonical histone H2A-H2B dimers with the histone variant H2A.Z-H2B dimers without disassembling the H3/H4 tetramer from DNA[56,97]. H2A.Z is a highly conserved variant of H2A that is found throughout all eukaryotes [98]. Developmentally regulated genomic loci show increased enrichment of H2A.Z relative to non-developmentally regulated loci[99]. H2A.Z is deposited specifically into the two nucleosomes that flank transcription start sites[100], and is essential in several higher eukaryotic organisms, but not in fungi[101,102]. In *C. albicans*, H2A.Z is enriched in white cells, relative to opaque cells, within the upstream intergenic region of *WOR1*[5]. The complex responsible for depositing this histone variant appears to play a role in stabilizing the white cell type and destabilizing the opaque cell type, as deletion of *SWR1* causes a significant increase in the white to opaque switch frequency and in the heritable maintenance of opaque cells [5]. Since H2A.Z variant enriched sites have been shown to correlate with slightly increased chromatin accessibility relative to canonical histones[103], it is conceivable that higher levels of H2A.Z inhibit expression of *WOR1* by facilitating the binding of a repressor protein within the upstream intergenic region of *WOR1*.

A similar phenotype is observed upon disruption of the NuA4 complex, which is known to recruit and/or promote chromatin-related enzymatic activities of the Swr1 complex[55,104]. Therefore, it is likely that NuA4 regulates the white-opaque switch by modulating the recruitment or enzymatic activity of Swr1, which in turn results in decreased H2A.Z deposition throughout the genome. The nucleosome editing function of the Swr1 complex is also controlled through H3K56 acetylation, which is catalyzed by Rtt109. High levels of H3K56 acetylation lead to decreased levels of H2A.Z deposition genome-wide[105]. This is notable as H3K56 acetylation itself has been implicated in altering histone turnover rates[106], which consequently alters genome-wide chromatin accessibility. It remains an open question whether H3K56 acetylation regulates the white-opaque switch by modulating the enzymatic activity of the Swr1 complex, or whether H3K56 acetylation directly regulates the white-opaque switch by modulating histone turnover rates.

1.7. Regulation of white-opaque switching by histone chaperone complexes

The highly basic amino acid composition of histones makes them predisposed to aggregation and promiscuous histone-DNA interactions, thus necessitating a diverse network of histone chaperones to orchestrate the assembly and integration of histones into chromatin[46,47,107]. Below, we focus our discussion on the evolutionarily conserved histone chaperone complexes HIR (HIRA in humans) and CAF-1, and their roles in regulating the white-opaque switch in *C. albicans*. CAF-1 primarily assembles nucleosomes in a replication dependent manner[108,109], whereas HIR functions independent of replication[110,111]. Importantly, the replication coupled nucleosome assembly function of CAF-1 is conserved in humans[108,109]. Both chaperone complexes are essential in higher eukaryotes[112,113], which has complicated efforts to investigate their functions in cell type formation and maintenance. The *C. albicans* white-opaque switch provides a unique and robust alternative system to investigate the functions of these highly conserved chaperone complexes in higher eukaryotes.

Studies in both *S. cerevisiae* and human HeLa cells have revealed that the HIR and CAF-1 complexes modulate nucleosome dynamics[114], which in turn affect chromatin accessibility. Other than their replication dependent functions, these two enzymes have also been shown to have several overlapping functions that are unrelated to replication. Recent work in *C. albicans* has shown that they function similarly to their orthologs in *S. cerevisiae*. Deletion of *C. albicans HIR1,* a subunit of the HIR complex*,* had no effect on white-opaque switching, while deletion of CAC2, a subunit of CAF-1 complex, resulted in an overall increase in switching in both directions[33]. On the other hand, deletion of a subunit of both chaperone complexes in *C. albicans* has been shown to lead to reduced opaque cell stability, as evidenced by wildtype levels of white to opaque switching and a sixfold increase in opaque to white switching[33]. These results alone do not definitively point to a specific chaperone complex responsible for regulating opaque cell stability; however, they do reveal that nucleosome dynamics can significantly affect cell type maintenance in the context of the white-opaque switch. Modulating nucleosome dynamics has a significant effect on chromatin accessibility[35], and recent studies have acknowledged the impact of chromatin accessibility on cell type specification and maintenance[63,94,115–117]. It is possible that opaque cells, more so than white cells, depend on increased chromatin accessibility to maintain their cell type specific transcriptional program, which could explain why deleting subunits of the HIR and CAF-1 complexes have dramatic effects on opaque cell stability.

1.8. Conclusion

Most chromatin research has focused on the roles of chromatin in regulating cellular processes such as transcription, replication, repair, mitosis, and apoptosis[35]. Recently, this focus has shifted to investigating how chromatin and chromatin modifiers regulate cell type specification and heritability. This avenue of research has led to significant insights into how chromatin regulates these fundamental biological processes. Many of these insights come from studies in higher eukaryotes; however, the inherent complexity of myriad possible cell type lineages and large genome sizes has slowed progress in the field.

The white-opaque switch in *C. albicans* is not hindered by the same challenges as higher eukaryotes, and thus represents an attractive alternative model system for investigating the mechanisms by which chromatin dynamics regulate cell type specification and heritability. We have reviewed the roles of chromatin regulating proteins in modulating white-opaque switching and the heritability of white and opaque cell types in *C. albicans*. Most of these proteins have been shown to also affect cellular differentiation and heritability in higher eukaryotes, thus supporting the overall relevance of this research. Future studies on chromatin regulating proteins in *C. albicans* will certainly lead to significant insights into the mechanisms by which chromatin regulates cellular differentiation and heritability across eukaryotes.

Chapter 2: Interplay between chromatin organizing transcription factors and dynamic nucleosomes regulate cell fate decisions and developmental processes in *Candida albicans*

Mohammad N. Qasim and Aaron D. Hernday
In preparation

2.1 Abstract

Sequence-specific DNA-binding transcription factors (TFs) work within the malleable framework of chromatin structure to establish unique transcriptional programs that underlie heritably maintained cellular differentiation. One such transcriptional program enables *Candida albicans*—a diploid polymorphic fungus—to reversibly switch between two distinct and heritable cell types, referred to as “white” and “opaque”. A diverse group of TFs work together to establish the transcriptional circuit for each cell type, and the highly interwoven structure of the opaque circuit bears a striking resemblance to transcriptional circuits that control developmental programs in higher eukaryotes. Although TFs that participate in the white-opaque transcriptional circuits have been extensively characterized, very little is known about how the interplay between these TFs and chromatin structure regulates chromatin organization and reorganization events critical for the establishment and maintenance of the white and opaque cell types. Here, we report that among the diverse group of TFs known to regulate the white-opaque switch, some have specialized abilities to organize chromatin structure by establishing wide and accessible nucleosome-depleted regions (NDRs) flanked by well-positioned nucleosomes. Furthermore, we find that this chromatin organizing function is critical for regulating the establishment and maintenance of the white-opaque switch. This report provides the first systematic genome-wide view of the distinct epigenetic chromatin landscape established in white and opaque cells, and our results show that the white-opaque switch represents an attractive and relatively “simple” model system to investigate how the interplay between TFs and chromatin structure regulates heritable cell fate decisions in higher eukaryotes.

2.2 Introduction

Eukaryotes from fungi to mammals rely on sequence-specific DNA-binding transcription factors (TFs) to establish unique transcriptional programs that underlie heritably maintained cellular differentiation. Chromatin structure is the malleable framework within which transcriptional regulation is controlled by TFs, in part through nucleosome-mediated mechanisms[35,118–121]. Nucleosomes—the fundamental unit of chromatin structure—by wrapping up DNA and altering the availability of TF binding sites, can play important roles in helping to suppress or enable expression of nearby genes. For TFs, gaining access to occluded sites is often critical for establishing transcriptional programs involving heritable cell fate decisions[115,121–126]. This critical function— gaining access to occluded sites by displacing nucleosomes—is carried out by pioneer transcription factors (PTFs)[115,122,123,127–130] in higher eukaryotes, and general transcription factors (GTFs) have been found to perform similar activities in fungi[131–133]. Collectively,

these TFs have a specialized ability to target many locations in the genome and effect the expression of a large number of genes, and therefore often serve as “master regulators” which invade compact chromatin to increase its accessibility[124,128,130,134–137]. Interplay between these master regulators and the underlying chromatin is critical to developmental programs in higher eukaryotes[115,124,130,138–141], and assessing this interplay is key to understanding the role of chromatin organization in regulating these processes.

Candida albicans, a diploid polymorphic fungus, has evolved a unique transcriptional network that enables reversible phenotypic switching between two cell types, referred to as “white” and “opaque”[142,143]. These two cell types differ in their expression of approximately 20% of genes in the genome, leading to distinct metabolic preferences, mating competencies, cellular morphologies, responses to environmental signals, and interactions with the host innate immune system[1,3,4,8,11,13–15,17,144–148]. Each cell type is established and heritably maintained by distinct transcriptional circuits, and the highly interwoven structure of the opaque circuit bears a striking resemblance to transcriptional circuits that control developmental programs in higher eukaryotes[3,4,7]. Within the opaque transcriptional circuit Wor1 is known as the “master regulator”[25,149]. It is highly expressed only in opaque cells and ectopic expression of *WOR1* in white cells is sufficient to trigger white to opaque switching[1,2,4]. Once the opaque transcriptional program is established, a minimal threshold level of Wor1 is needed for heritable maintenance of the opaque cell type[2,149]. The function of Wor1 in opaque cells has been extensively characterized; Its genome-wide binding sites identified through ChIP-chip experiments[1,4,24,150], and its DNA-binding sequence motif (“TTAAAGTTT”) validated experimentally through mutational analysis[151] and MITOMI[4]—a micro-fluidics-based approach used to probe DNA-binding specificities of TFs[152,153]. Although the regulatory TFs that participate in the white and opaque-specific transcriptional regulatory circuits have been extensively characterized through transcriptional profiling and genome-wide chromatin immunoprecipitation (ChIP)[1–4], little is known about the role of chromatin organization in white and opaque cells.

Recently, numerous chromatin-modifying enzymes—many of which are highly conserved across all eukaryotes—have also been found to modulate white-opaque switch dynamics[5,6,22,29,31–34,50,52,119,154,155], indicating that chromatin organization and reorganization play important roles in the maintenance of, and interconversion between, the white and opaque transcriptional programs. Some studies have begun to explore the interplay between TF binding and chromatin organization in the context of white-opaque switching[29,154], yet to date none have taken a systematic genome-wide view of this important aspect of white-opaque switching. Research exploring similar interplay between TFs and chromatin structure in higher eukaryotes have been largely hindered because deleting these chromatin modulating transcription factors often cause global and pleiotropic effects[156,157]. In *C. albicans* however, the white-opaque switch represents an attractive and relatively “simple” model system to investigate how the interplay between TFs and chromatin structure regulates heritable cell fate decisions in higher eukaryotes.

In this study we find that the opaque cell “master regulator” establishes opaque-specific chromatin structures, some of which are required for heritable maintenance of opaque cell identity. In addition to Wor1, we find that two more TFs (Ndt80 and Wor2) possess this specialized ability to target many Nucleosome-Free Regions (NFRs)—compact chromatin structures that are observed to be inaccessible in ATAC-seq assay—

and effect the expression of a large number of genes. Because these TFs organize chromatin throughout the genome, they are often recognized to serve as "master regulators" that play critical roles in developmental processes and heritable cell fate decisions. The work presented in this chapter will serve as a foundation to further investigate how the interplay between chromatin organizing TFs and dynamic nucleosomes regulates developmental processes and heritable cell fate decisions in *C. albicans,* and may lead to insights that inform how these processes regulate cellular differentiation in higher eukaryotes.

2.3 Results

2.3.1 White and Opaque cells display significantly altered chromatin structures

To examine the chromatin structure differences in wild-type white and opaque cells we assessed chromatin accessibility using the Assay for Transposase-Accessible Chromatin (ATAC-seq) approach[158,159]. We also determined nucleosome occupancy using micrococcal nuclease digestion followed by sequencing (MNase-seq)[160]. Comparing the resulting nucleosome occupancy and chromatin accessibility profiles at a few key white-opaque regulatory loci reveals significant differences between the chromatin structure of white and opaque cells (**Fig. 2-1A-D**). Most strikingly, accessibility of the intergenic region upstream of *WOR1*—the master regulator of opaque cells—is significantly increased in opaque cells relative to white cells (**Fig. 2-1A**). Notably, altered chromatin accessibility and nucleosome occupancy is observed at loci which are overlapping or proximal to locations of Wor1 binding in opaque cells (**Fig. 2-1A-D, orange arrows**), suggesting a potential direct relationship between Wor1 binding and chromatin reorganization.

We also observed altered chromatin accessibility and nucleosome occupancy at intergenic regions which are farther away from Wor1 binding sites, indicating that changes in chromatin structure upon white to opaque switching are likely extensive and genome wide. To further explore these cell type-specific changes in chromatin accessibility we carried out differential accessibility (DA) analysis. We note that normalization was carried out as described by Reske et al[161] to reduce any biases due to variability in transposition reaction efficiency observed between different cell types. For DA analysis, we quantified the ATAC-seq signal at intergenic regions upstream of all genes in *C. albicans* genome and compared between white and opaque cells using DESeq2 analysis pipeline[162]. Specifically, we examined the degree to which changes in chromatin accessibility occur within Wor1-bound intergenic regions as opposed to intergenic regions that are not bound by Wor1. The resulting volcano plot (**Fig. 2-1E**) shows that changes in chromatin accessibility occur within intergenic regions bound by Wor1 (**Fig. 2-1E**, **dots with a label**) as well as those not bound by Wor1 (**Fig. 2-1E**, **dots without a label**). We note that decreased accessibility within a specific intergenic region is strongly correlated with decreased expression from the proximal gene downstream of the intergenic region, suggesting that formation of a compacted chromatin structure is a likely mechanism for repressing these genes. However, there are some white enriched genes (*PTH2, RMS1, and RME1*) which have increased accessibility in opaque cells, suggesting that this opened chromatin structure is providing access for binding of regulatory TFs that are actively repressing these genes. These DA results highlight that the changes in chromatin structure upon white to opaque switching are genome-wide, including many Wor1-bound intergenic regions.

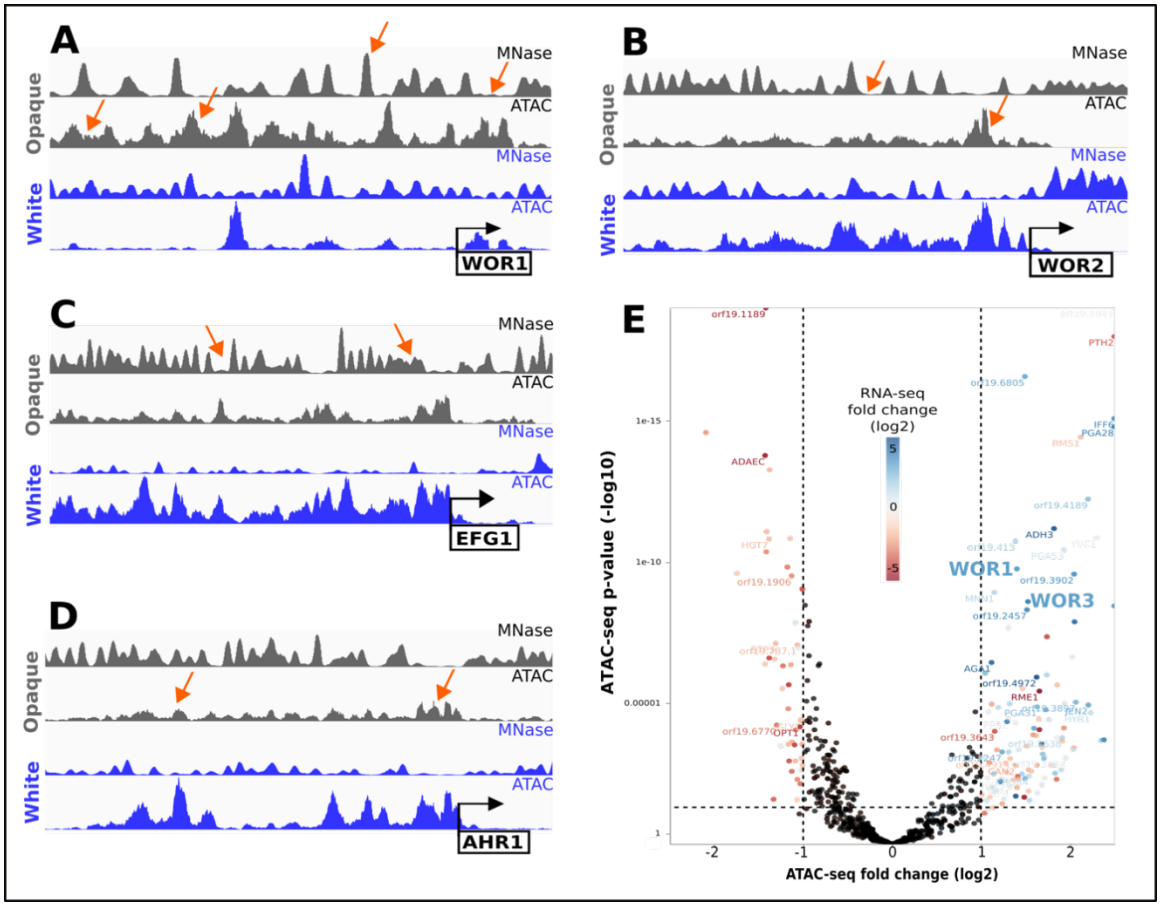


Figure 2-1. Chromatin structure differs between white and opaque cells.

(**A-D**) Intergenic regions upstream of white-opaque regulating genes show altered nucleosome occupancy and chromatin accessibility profiles in wild-type white (blue) and wild-type opaque (grey) cells. Four technical replicates were evaluated for MNase-seq experiments and a minimum of three biological replicates were evaluated for ATAC-seq experiments. (**E**) Volcano plot depicting the differential accessibility analysis of ATAC-seq peak signals (ATAC-seq reads shorter than 120 bp[159]) between wild-type opaque and white cells. The y-axis represents the negative log[10] adjusted p-value (false discovery rate, FDR) and the x-axis depicts the log2 fold-change between wild-type opaque and wild-type white cells. The horizontal dashed red lines indicate an FDR of 0.05. Each dot represents one intergenic region annotated to the proximal downstream gene, which is overlaid with a color gradient indicating the log2-fold-change (wild-type opaque vs wild-type white cells) from RNA-seq experiment. Colored dots represent changes in chromatin accessibility that is greater than 2-fold and p-value less than 0.05. Dots with a gene label indicate that Wor1 is bound to the intergenic region upstream of these genes.

2.3.2 Opaque cell "master regulator" Wor1 establishes opaque-specific chromatin structure

The opaque-specific transcriptional regulatory network encompasses 748 target genes that are bound by at least one of the eight switch regulatory TFs that have been characterized thus far by genome-wide binding assays[4,21,23]. Of these, a majority are bound by the opaque master regulator Wor1 (68%)[1,4]. Given that Wor1 is not only required for the establishment of the opaque transcriptional program, but also for heritable maintenance of the opaque cell type, we reasoned that a systematic examination of chromatin structures surrounding Wor1 binding sites should highlight chromatin organization and reorganization events critical for the maintenance of, and interconversion between, the white and opaque transcriptional programs.

To discern how Wor1 regulates the chromatin structure in opaque cells, we assessed nucleosome occupancy across all genomic sites where Wor1 consensus binding motifs overlapped with Wor1 enrichment in opaque cells (i.e., Wor1-occupied motif sites). These sites were identified by scanning Wor1 ChIP enriched loci for the Wor1 consensus DNA binding motif ("TTAAAGTTT") using "FIMO"[163], a computational pipeline that scans a set of sequences for a motif of interest, and we only retained matches with a p-value cutoff of 0.05. For this analysis, we used the position-specific weight matrix (PSWM) for Wor1 as input into FIMO, which allowed for some substitutions and alleviated concerns that the analysis maybe too narrow and restrictive. Indeed, the sites identified by FIMO consisted of some consensus hits and many more hits that had some minor substitutions. Examining nucleosome occupancy at these Wor1-occupied motif sites revealed that these loci are depleted of nucleosomes in both white and opaque cells (**Fig. 2-2A**). Interestingly, the width of these nucleosome depleted regions (NDRs) is expanded in opaque cells relative to white cells (**Fig. 2-2A**, **black arrow**), suggesting that Wor1 binding directly or indirectly impacts nucleosome positioning at loci immediately surrounding these Wor1 motifs. The noticeably wider NDRs overlapping Wor1-bound motifs in opaque cells suggested that the chromatin structure at these sites may be more accessible in opaque cells than in white cells. Indeed, we observed that chromatin accessibility is significantly increased at these sites in opaque cells relative to white cells. In contrast, the reduced NDR width observed in white cells correlated with a significant decrease in chromatin accessibility at these Wor1 binding sites (**Fig. 2-2B**).

DNA-binding TFs are thought to establish relatively wide NDRs, a critical factor in determining chromatin organization[125,164,165]. It has been proposed that TFs act as a barrier to nucleosome formation, establishing accessible chromatin and wide NDRs by directing nucleosomes into flanking positions[164,166,167]. Therefore, we hypothesized that the noticeably wider NDRs observed at Wor1-occupied motif sites should correlate with high TF occupancy in opaque cells, while the reduced NDR width observed at these sites in white cells should correlate with decreased TF occupancy. Although results from previous ChIP and MITOMI experiments indicate that these sites are likely bound by Wor1 in opaque cells[4], we note that neither technique can confirm Wor1 binding to discreet motif sequences in vivo. We tested our hypothesis by examining TF occupancy profiles, centered on Wor1-occupied motif sites, in wild-type white and opaque cells. For this, we used a MNase based approach that was recently developed to report TF occupancy throughout the genome[168,169]. Briefly, limited digestion of chromatin with MNase produces protected DNA fragments in the range of 10-50 bp lengths, which represent the length of DNA protected by a chromatin-bound TFs. Sequencing of these small, protected DNA

fragments after enriching for them using a PAGE gel provides a high-resolution occupancy profile for chromatin-bound TFs genome wide. Consistent with our hypothesis, the resulting TF occupancy profile showed a significant enrichment for TF occupancy in opaque cells, whereas these Wor1-occupied motif sites showed no evidence for TF occupancy in white cells (**Fig. 2-2C**). Together these results suggest that Wor1 acts like a chromatin organizing TF, binding to the upstream intergenic regions of opaque specific genes and establishing highly accessible and wide NDR chromatin structures that are flanked by well-positioned nucleosomes. This activity is similar to what is observed for PTFs and GTFs and suggests that the role of Wor1 in white-opaque switching may be to help "pioneer" the assembly of opaque-specific chromatin structures which ultimately enable assembly of opaque-specific transcriptional regulatory complexes. Alternatively,

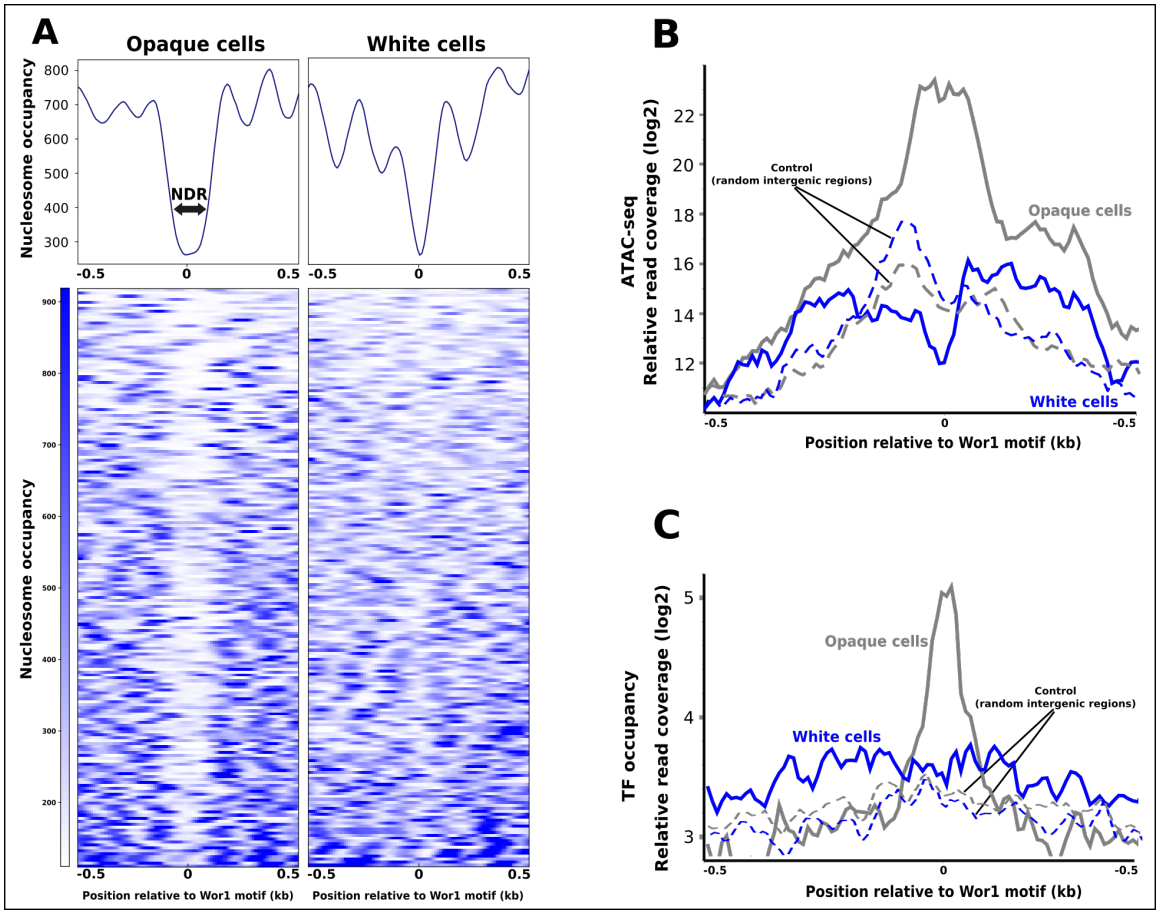


Figure 2-2. Wor1 establishes opaque-specific chromatin structures.

(**A**) Normalized nucleosome occupancy at Wor1-occupied motif sites (centered on Wor1 motif) in opaque (left) and white cells (right). The signal intensity (i.e., enrichment) as a profile plot is shown above the heatmaps. Profile plots depicting normalized chromatin accessibility (**B**), and TF occupancy (**C**) at Wor1-occupied motif sites (centered on Wor1 motif) in opaque and white cells. Blue colored lines indicate white cells and gray colored lines indicate opaque cells. Dashed lines show data for a control location set which represent an equal number of randomly chosen intergenic regions.

the role for Wor1 in white-opaque switching may be to modulate the chromatin structure at these sites by recruiting either chromatin remodeling complexes or histone modifying enzymes.

2.3.3 Opaque-specific alterations in the chromatin structure required for heritable maintenance of the opaque transcriptional program

The majority of Wor1-bound sites are co-enriched for additional regulators of the white-opaque switch, including Wor2 (52%), a key factor required for heritable maintenance of the opaque state. Deletion of *WOR2* results in cells that are locked in white cell type unless induced to switch to the opaque cell type through ectopic expression of *WOR1*[1]. These induced wor2 delta/delta opaque cells express wild-type opaque levels of *WOR1* and most other opaque-specific transcripts yet fail to maintain the opaque transcriptional program once ectopic *WOR1* expression is turned off[1]. In the absence of ectopic *WOR1* expression, these induced wor2 delta/delta opaque cells rapidly revert to the white cell-type, indicating that Wor2 is required for the heritable maintenance of opaque cells. Given that Wor1 binding correlates with opaque-specific increases in chromatin accessibility, and that Wor2 is often co-enriched with Wor1 and is essential for heritable maintenance of the opaque cell type, we hypothesized that Wor2 may effect changes in chromatin organization that are critical for opaque heritability.

To identify Wor2-dependent changes in opaque cell chromatin structure, we compared nucleosome occupancy in the induced wor2 delta/delta opaque cells (i.e., non-heritable opaque cells) against nucleosome occupancy in induced wild-type opaque cells (i.e., heritable opaque cells) at intergenic regions where Wor2 consensus binding motifs overlapped with Wor2 enrichment in opaque cells (i.e., Wor2-occupied motif sites). In the induced wor2 delta/delta opaque cells, these sites were enriched for nucleosome occupancy in the low-MNase digestion condition (**Fig. 2-3,** blue line with red circles), but not in the high-MNase digestion condition (**Fig. 2-3,** blue line). This observed pattern is highly indicative of “fragile” nucleosomes —nucleosomes known to be extremely sensitive to the extent of MNase digestion[125,164,170] — and suggests that this nucleosome eviction may be critical for heritability. Consistent with this, we did not observe fragile nucleosomes at Wor2-occupied motif sites in the induced wild-type opaque cells (**Fig. 2-3,** orange line). To address the possibility that the inducing condition itself does not alter the chromatin structure in wild-type opaque cells, we compared nucleosome occupancy at Wor2-occupied motif sites for opaque cells cultured under inducing and non-inducing conditions. As shown in **Figure 2-3**, nucleosome occupancy at these sites appears unaltered between inducing and the corresponding non-inducing conditions, which confirms that effects due to the inducing condition don't underlie the Wor2-dependent changes observed in the induced wor2 delta/delta opaque cells. Based on these results, we conclude that Wor2 is directly or indirectly involved in evicting fragile nucleosomes from specific sites in opaque cells, and that this function may be critical for heritability of the opaque state.

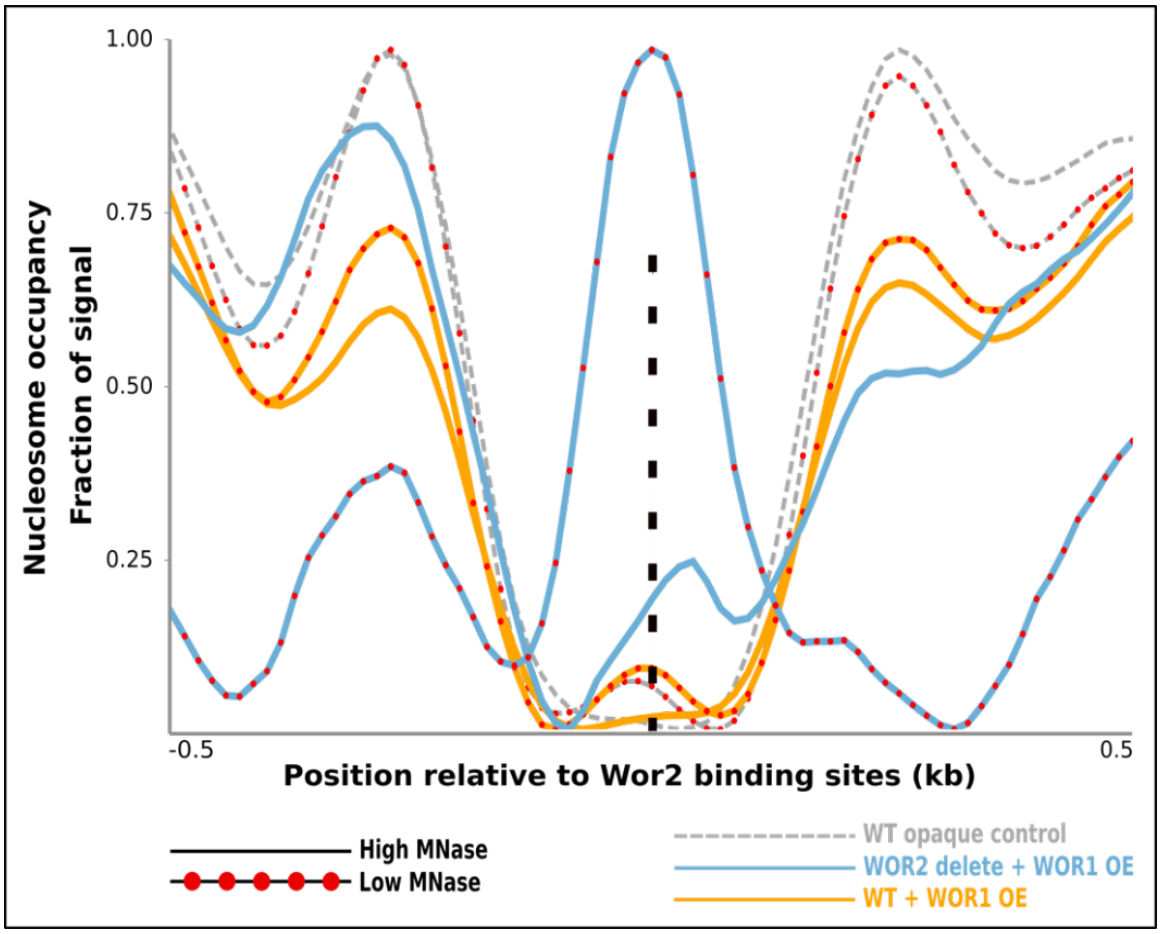


Figure 2-3. Wor2-dependent changes in chromatin structure is required for heritable maintenance of opaque cells.

Normalized nucleosome occupancy signal at Wor2-occupied sites for *C. albicans* chromatin digested under low (solid lines with red circles) and high (solid lines) MNase digestion conditions. Two technical replicates were evaluated for each digestion condition: 500 bp regions upstream (-0.5kb) and downstream (0.5kb) are displayed in the profile plot. Dashed grey lines show nucleosome occupancy at these same sites for wild-type opaque cells.

2.3.4 White-opaque switching frequencies are modulated by a chromatin organizing TF

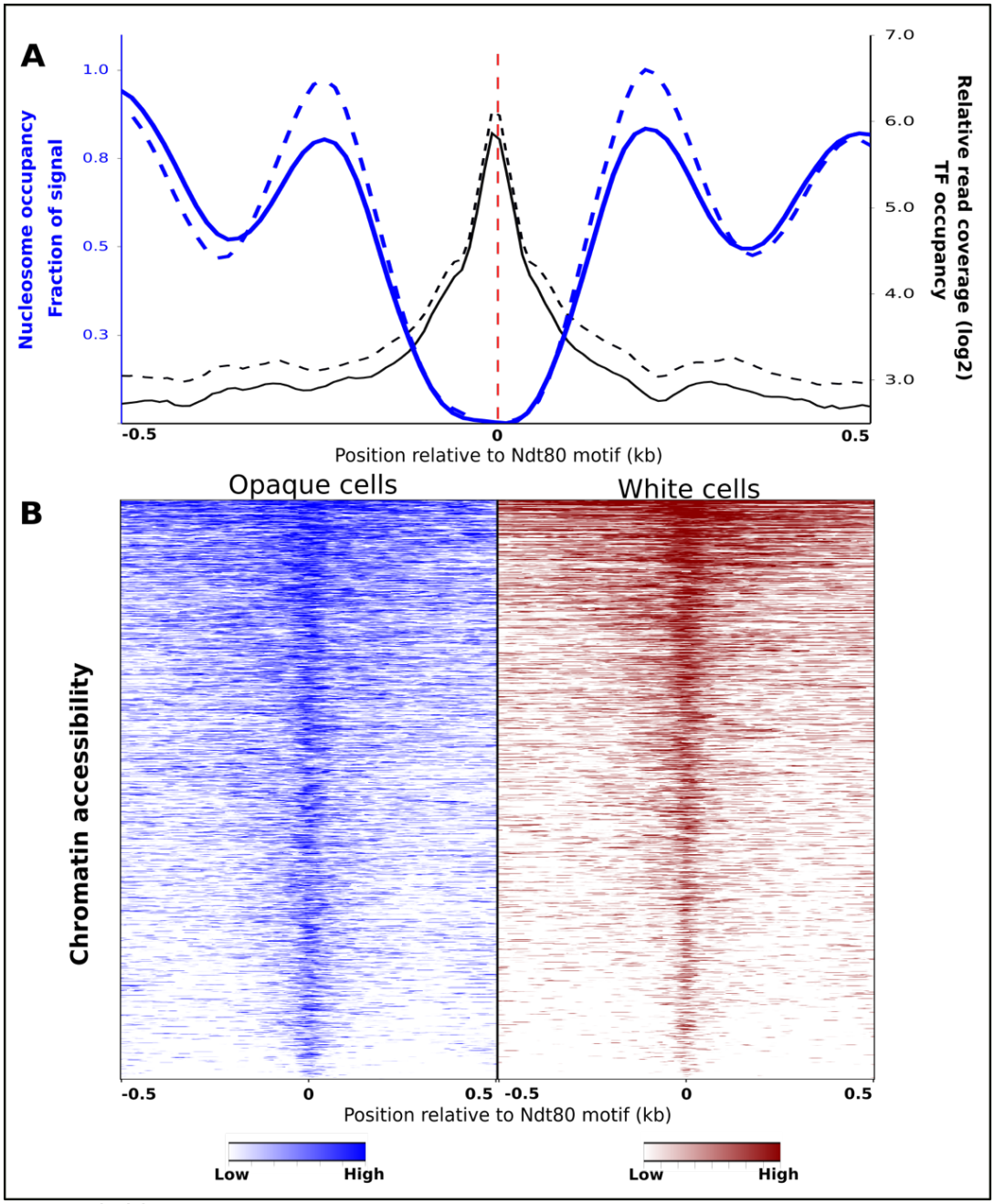
The changes in chromatin accessibility and nucleosome occupancy observed at motifs that are occupied by Wor1 and Wor2 appear to be correlated with, or necessary for, the establishment and maintenance of the opaque state. Therefore, we hypothesized that additional motifs which are enriched in loci that are co-occupied by Wor1 and Wor2 may play important roles in regulating the white-opaque switch. Using a de novo motif search approach on intergenic regions that are co-occupied by Wor1 and Wor2 revealed

substantial enrichment of a “CACAAA” motif. This motif is the consensus DNA binding sequence for Ndt80, a TF that is highly conserved across a large group of fungal species. While previous investigations have revealed important roles for Ndt80 in regulating diverse processes including white-opaque switching, biofilm formation, sexual development, filamentation, drug resistance, virulence, and the response to nutrient stress[171–176], none have reported a role for Ndt80 in regulating chromatin structure.

Ndt80 has previously been shown to modulate the frequency of white-opaque switching[9]. Deletion of NDT80 reduces white-to-opaque switching by 10-fold and increases opaque-to-white switching by ~2-fold[9], however the mechanism through which Ndt80 influences white-opaque switching is uncharacterized. Given the impact of Ndt80 on the white-opaque switch, and the high enrichment of Ndt80 motifs within Wor1 and Wor2-bound intergenic regions, we examined chromatin structure at Ndt80 motif sites. Because Ndt80 influences both opaque and white cells, we thought it was important to examine Ndt80 motifs genome wide so that we don’t bias our analysis by only examining loci that are known to be important for regulating opaque cells only (i.e., Wor1 and Wor2 binding sites). Here, our results showed that there is no change in the apparent nucleosome occupancy (**Fig. 2-4, A**), TF occupancy (**Fig. 2-4, B**), and chromatin accessibility (**Fig. 2-4, C**) profiles between white and opaque cells.

Given that genome-wide motif sites for Ndt80 appear to be associated with NDR chromatin structure in both white and opaque cells, it prompted us to ask if deletion of *NDT80* results in loss of NDR chromatin structure at these sites in white and opaque cells. Toward this end, we carried out MNase-seq experiments in ndt80 delta/delta white and ndt80 delta/delta opaque cells and then compared the resulting nucleosome occupancy profiles against their corresponding wild-type AHY135 control strain. Compiling nucleosome occupancy profiles at Ndt80 motif sites showed significant differences between ndt80 delta/delta and the corresponding wild-type cells in both white and opaque state (**Fig. 2-5A-B**). Specifically, nucleosome occupancy profiles revealed that the NDRs overlapping Ndt80 motif sites were relatively wide and flanked by well-positioned nucleosomes in the wild-type strain, whereas these same NDRs in the ndt80 delta/delta cells are narrower and flanked by poorly positioned or “fuzzy” nucleosomes that have many alternative positions nearby. To confirm the statistical significance of this reduction in NDR width, we quantified the width for all NDRs at these sites and then compared the NDR widths for the wild-type strains against those of the ndt80 delta/delta opaque and white cells using the Wilcoxon Rank Sum Test. Indeed, the reduction in NDR widths was statistically significant with a p-value of $2.04e^{-16}$ in opaque cells and p-value of $3.24e^{-15}$ in white cells. These results suggested a potential direct relationship between Ndt80 binding and chromatin organization in both white and opaque cells. To test this, we used a method called “Cleavage Under Targets & Release Using Nucleases” (CUT&RUN)[177] to identify genome-wide Ndt80 binding sites in both white and opaque cells. Results from CUT&RUN assay showed that Ndt80 binds to 1081 intergenic regions in opaque cells and 1714 intergenic regions in white cells **(Fig. 2-5, C**); 883 of these Ndt80-bound intergenic regions are shared in common between white and opaque cells. Additionally, de novo motif search under Ndt80 enrichment peaks identified the consensus binding motif for Ndt80 with p-value of $6.1e^{-21}$ Consistent with our hypothesis, compiling nucleosome occupancy profiles

at Ndt80-occupied motif sites showed significant differences between ndt80 delta/delta and the corresponding wild-type cells in both white and opaque state (**Fig. 2-5, top half of D**). As a control, we compiled nucleosome occupancy at a control location set (i.e., a number of randomly chosen intergenic regions) for wild-type and ndt80 delta/delta opaque cells and observed relatively wide NDRs in both wild-type and ndt80 delta/delta cells at this control location set **(Fig. 2-5, bottom half of D**). Based on these results, we conclude that Ndt80 functions like a chromatin organizing TF by establishing wide NDRs flanked by well-positioned nucleosomes, and that this function impacts the frequency of white-opaque switching.



<u>Figure 2-4</u>. Genome-wide Ndt80 motif loci are highly accessible in both white and opaque cells.

(**A**) Normalized nucleosome occupancy (colored blue) and TF occupancy (colored black) at Ndt80 motif sites genome wide in wild-type opaque (solid lines) and white cells (dashed lines). (**B**) The chromatin accessibility signals for all Ndt80 motif sites genome wide is shown as colored heatmaps (blue = opaque cells; dark red = white cells); colored bars indicate signal intensities for ATAC-seq); 500 bp regions upstream (-0.5kb) and downstream (0.5kb) are displayed in the heatmaps. Three biological replicates were evaluated for visualizing the chromatin accessibility signal in white and opaque cells.

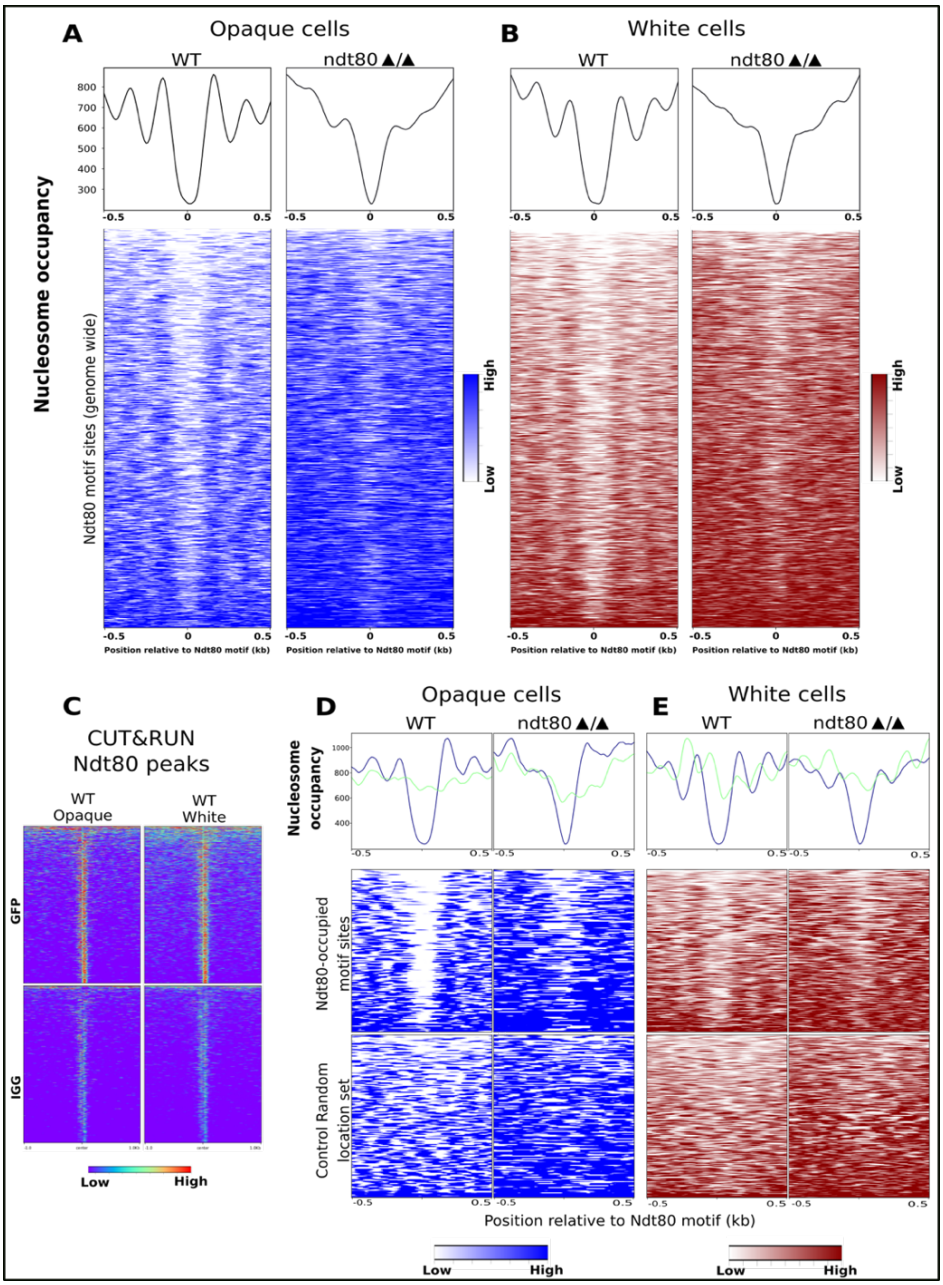


Figure 2-5. Ndt80 establishes ordered nucleosome arrays in white and opaque cells.

(**A**) Normalized nucleosome occupancy at Ndt80 motif sites genome wide in WT opaque strain (left), ndt80 delta/delta mutant strain in the opaque state (right), (**B**) WT white strain (left), and ndt80 delta/delta mutant strain in the white state (right). (**C**) The CUT&RUN signals for all binding events for Ndt80 is shown as colored heatmaps (top half = GFP signal; bottom half = background IGG signal; colored bar indicates signal intensities for GFP and IGG samples); 1,000 bp regions upstream (-1.0kb) and downstream (1.0kb) are displayed in the heatmaps. Four biological replicates for Ndt80 were evaluated for visualizing the CUT&RUN enrichment. Normalized nucleosome occupancy at Ndt80-occupied motif sites in WT strain (left) and ndt80 delta/delta mutant strain (right); cells in the opaque state (**D**) and the white state (**E**).

2.3.5 Ndt80 regulates chromatin structure in biofilm cells and planktonic cells

In addition to regulating the white-opaque switch, Ndt80 also regulates the ability of *C. albicans* to form biofilms. Biofilms are organized communities of microorganisms embedded in a matrix of extracellular polymers, where individual cells acquire properties, such as drug resistance, that are distinct from the free-floating planktonic cells observed in suspension cultures. In *C. albicans* biofilms, Ndt80 binds to ~637 intergenic regions, most of which are upstream of genes that are differentially expressed between biofilm and planktonic cells[178,179]. Cells that lack *NDT80* form highly deficient biofilms, indicating that Ndt80 is a key regulator of biofilm formation in *C. albicans*[179]. Considering that Ndt80 significantly impacts chromatin structure in white and opaque cells, we asked whether Ndt80 also regulates chromatin structure in *C. albicans* planktonic cells, biofilm cells, or both.

We examined the influence of Ndt80 on chromatin structure in biofilm and planktonic cells by comparing nucleosome occupancy at Ndt80-occupied motif sites in the wild-type SN250 strain against the corresponding loci in a ndt80 delta/delta mutant strain. The results at these sites, for both biofilm and planktonic conditions, revealed wide NDRs flanked by well-positioned nucleosomes in the wild-type strain (**Fig. 2-6**). In the corresponding ndt80 delta/delta strains however, the NDRs overlapping these Ndt80-occupied motif sites had shrunk significantly in width, and nucleosomes flanking these NDRs also seemed less organized and poorly positioned (**Fig. 2-6**). To confirm the statistical significance of this reduction in NDR width, we quantified the width for all NDRs at these sites and then compared the NDR widths for the wild-type strains against those of the ndt80 delta/delta cells using the Wilcoxon Rank Sum Test. Indeed, the reduction in NDR widths was statistically significant with a p-value of $1.002e^{-31}$. As a control, we compiled nucleosome occupancy data centered on a control location set (i.e., a number of randomly chosen intergenic regions) for biofilms and planktonic cells. As expected, results for the control location set showed that the two cell types did not exhibit any significant differences in chromatin structure at these control loci (data not shown). These results indicate that Ndt80 establishes relatively wide and accessible NDRs that are flanked by well-positioned nucleosomes in *C. albicans* cells, and that this function plays

an important role in regulating at least two developmental programs, including biofilm development and white-opaque switching.

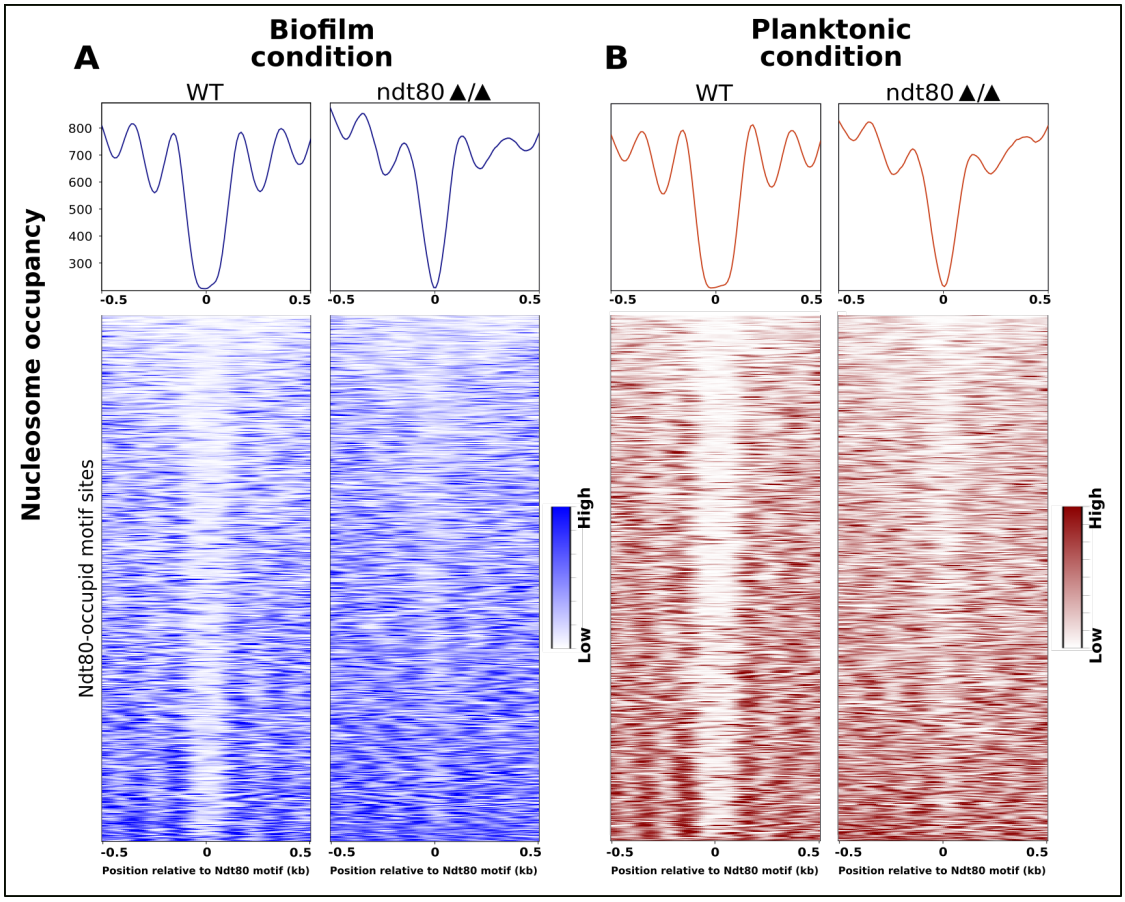


Figure 2-6. Ndt80 organizes chromatin in biofilm and planktonic cells.

(**A**) Normalized nucleosome occupancy at Ndt80-occupied motif sites in wild-type biofilm samples (left, blue colored plot), ndt80 delta/delta mutant defective biofilm samples (right, blue colored plot), (**B**) wild-type planktonic cells (left, dark red colored plot), and ndt80 delta/delta mutant planktonic cells (right, dark red colored plot).

2.4 Discussion

This study provides the first systematic genome-wide view of the interplay between TF binding and chromatin organization in *C. albicans*. We reveal that white-to-opaque switching involves extensive chromatin reorganization, and that three white-opaque switch regulatory TFs play distinct yet complementary roles in this process. We also reveal that the opaque cell "master regulator" Wor1, together with Wor2, establishes opaque-specific chromatin structures that are correlated with the establishment and heritability of opaque cells, respectively. We also find that Ndt80 acts as a chromatin organizing TF, establishing ordered arrays of stably positioned nucleosomes flanking Ndt80-bound motifs. This chromatin organizing function impacts two distinct developmental processes in *C. albicans*; white-opaque switching and biofilm formation. Generally, these results indicate that chromatin organization plays important roles in regulating both the white-opaque switch and biofilm formation, and that these changes in the chromatin structure are chiefly modulated by TFs with specialized abilities to modulate chromatin organization. Below we present several lines of evidence that support these claims.

In this study, we show that white and opaque cells display significant alterations in their nucleosome occupancy, TF occupancy and chromatin accessibility profiles genome-wide. Importantly, we show that the observed alterations are especially striking at Wor1 motif encompassing intergenic regions that are bound by the opaque 'master regulator' Wor1. In opaque cells these sites display increased TF occupancy, increased chromatin accessibility, and noticeably wide NDRs—features consistent with a NDR chromatin structure. Given our results, combined with previous studies demonstrating the critical role for Wor1 in regulating the white-opaque switch, we conclude that Wor1 is likely involved in establishing the opaque-specific NDR chromatin structures observed at Wor1-occupied motif sites, and that this function is critical for the establishment and maintenance of opaque cells. We note that because white and opaque cells are grown under identical growth conditions, and both cells contain the same primary genome sequence, the observed changes in the chromatin structure are independent of environmental variables or any effects due to differences in the genetic composition. It is more likely that the observed effects are due to epigenetic mechanisms that alter chromatin structure.

The terms "NFR" and "NDR" have been used interchangeably often[120,125,180,181] even though they describe two distinct features of chromatin organization. Nucleosomes, the fundamental units of chromatin organization, are thought to be randomly deposited wherever a region of "free" DNA is greater than ~150 bp in length. Transcriptional activity subsequently establishes well-organized arrays of phased nucleosomes within gene bodies[169,182,183], while binding of TFs to the DNA within promoter regions form a "barrier" complex and the resulting steric hindrance between nucleosomes and the barrier complex establishes NDRs (greater than ~150 bp in length) flanked by phased nucleosomes on both sides[125]. In contrast, NFRs do not require any barrier complexes; they are less than ~150 bp in length and are thought to be genuinely free of nucleosomes because they show no evidence of protection even in limited MNase digestion conditions. We believe that a "barrier" model for TFs in regulating the opaque-specific NDR chromatin structure at Wor1-occupied motif sites is consistent with our results. In white cells, low concentrations of Wor1 protein don't reach the minimal threshold level needed to assemble a barrier complex at Wor1-occupied motif sites, thus the nucleosomes that are deposited at these intergenic regions are distributed randomly in an unorganized fashion. In opaque cells on the other hand, Wor1 concentration exceeds the minimal threshold level needed to

assemble the barrier complexes which in turn establishes the NDR chromatin structure at Wor1-occupied motif sites.

In agreement with our results, some recent studies in other fungi have demonstrated that NDRs are indeed bound by nonhistone "barrier" complexes[164,165]. However, these findings are contradicted by others who claim that NDRs are bound by "fragile" nucleosomes known to be extremely sensitive to the extent of MNase digestion[164,184,185]. Nevertheless, none of these studies have investigated the significance of these chromatin structures in regulating heritable cell fate decisions. In this study we show that the existence of opaque-specific NDR chromatin structures at Wor1-occupied motif sites correlate with high TF occupancy, indicating that the NDRs established at these sites are bound by nonhistone "barrier" complexes in wild-type heritable opaque cells. Remarkably, while heritable opaque cells are observed to be nucleosome free at Wor1-occupied motif sites, non-heritable opaque cells which lack *WOR2* show significant enrichment for fragile nucleosomes at these sites. It will be important to carry out MNase-ChIP-seq experiments to confirm that these sites are in fact enriched for histone proteins and not some other nonhistone proteins. Given that the opaque-specific NDR chromatin structures established at Wor1-occupied motif sites are greater than ~150 bp in length, meaning that these NDRs are wide enough to accommodate a nucleosomal particle, we speculate that Wor2 increases stability of the DNA-bound nonhistone "barrier" complexes and the resulting steric hinderance keeps these sites free of nucleosomes. Thus, Wor2 likely facilitates heritability in opaque cells by forming highly stable opaque-specific NDR chromatin structures that are subsequently propagated for hundreds of generations. Further work will be required to determine whether these TFs (Wor1, Wor2, and Ndt80) alone can function as chromatin organizing factors or whether they require help from chromatin remodeling enzymes, or other TFs, to perform their specialized functions as organizing factors. Nonetheless, we note that our report is the first to describe the chromatin organizing function of TFs regulating the white-opaque switch and biofilm developmental process in *C. albicans*.

2.5 Methods and Materials

Fungal strains and growth conditions

White and Opaque cells were grown at 25°C in synthetic complete dextrose (SCD) agar plates (supplemented with amino acids), and overnight cultures were grown at 25°C with shaking using SCD liquid media. For biofilms, wild-type SN250 and ndt80 delta/delta strains were grown at 30°C in YPD agar plates for 48 hours, and overnight cultures were grown at 30°C in YPD broth with shaking for 12-16 hours.

MNase-seq

DNA recovered from two independent chromatin samples digested for a short period of time (MNase-low) along with DNA from two samples digested for a longer period (MNase-high) were fractionated via polyacrylamide gel electrophoresis, with nucleosome sized fragments (100 – 300 bp) and TF sized fragments (less than 80 bp) recovered for subsequent library preparation. Sequencing ready libraries were prepared using NEBNext Ultra II library prep kit per the manufacturer's instructions. Briefly, after the end-repairing and dA-tailing reactions, sequencing adaptors were ligated to the MNase-digested DNA and then PCR amplified using indexing primers to generate the barcoded libraries. The barcoded libraries were then pooled together and sequenced on Illumina's NextSeq

platform in paired-end mode. White-opaque switch assay was carried out before all experiments to confirm that white and opaque cell populations were homogenous.

ATAC-seq

ATAC-seq in *C. albicans* was carried out as previously described by Shapiro et al[186]. Briefly, *C. albicans* cells were spheroplasted using zymolase for 30 mins at 25°C. Cells were then washed twice with 500 uL of sorbitol buffer (Supplementary File 1) supplemented with protease inhibitor cocktail. After the last wash, cells were pelleted at 2K rpm, and the supernatant removed. The cell pellet was then resuspended in the ATAC-reaction mixture (Supplementary File 1) and incubated at 25°C. DNA was recovered and purified using QIAGEN KIT per manufacturer's instructions. White-opaque switch assay was carried out before all experiments to confirm that white and opaque cell populations were homogenous.

CUT&RUN sample preparation and sequencing (See Chapter 3 for details)

CUT&RUN assay was performed as described in Qasim et al., 2020. Briefly, Ndt80 was tagged with an HA-tag at the N-terminal end of the protein in the AHY135 wildtype strain background. Cells are permeabilized to isolate intact nuclei. ConA beads are activated, and the intact nuclei are then bound to the activated ConA beads. Polyclonal anti-GFP antibody (Takara Bio, Cat No. 632592) was added to the bead-bound nuclei and incubated at 4 °C. Next, pAG-MNase is added and allowed to bind to the target antibody. After the addition of CaCl2, pAG-MNase is activated and targeted chromatin digestion proceeds until the addition of the chelating reagent to inactivate pAG-MNase. The pAG-MNase-bound antibody complex is allowed to diffuse out of the permeabilized nuclei, and the resulting DNA is extracted and cleaned using NEB DNA Cleanup kit (NEB, Cat No. T1030L). Sequencing libraries are prepared from the CUT&RUN enriched DNA fragments using NEBNext Ultra II library prep kit (NEB, Cat No. E7645L) per the manufacturer's instructions. The resulting libraries were then run on a 10% PAGE gel to separate and remove contaminating adapter dimers prior to sequencing. Final DNA clean-up was performed using bead-based size selection with SPRIselect (Beckman Coulter, Cat No. B23317). Illumina paired-end sequencing was performed on the NextSeq platform to obtain 5-10 million reads per sample. White-opaque switch assay was carried out before all experiments to confirm that white and opaque cell populations were homogenous.

Bioinformatics Data Analysis

Paired-end sequencing reads were mapped to the *C. albicans* genome and then normalized for sequencing depth before generating the normalized read density profiles. To call peaks of significant enrichment for TF occupancy, we used MACS2, a peak calling algorithm, to identify regions in the genome that show significant enrichment for short reads (smaller than 50 bp) resulting from protection by TFs. MACS2 was also used to find chromatin accessible regions—i.e., regions where ATAC-seq reads shorter than 120 bp are significantly enriched relative to the surrounding regions. Differential accessibility (DA) analysis was performed using a computational pipeline designed to take specialized precautions and accurately normalize for differences in the tagmentation reaction efficiency between different cell types. Nucleosome occupancy analysis was performed using Dynamic Analysis of Nucleosome and Protein Occupancy by Sequencing (DANPOS)[187], a computational analysis specifically designed for analyzing nucleosome occupancy data generated through MNase-seq approach. Motif analysis was performed using "FIMO" (Find Individual Motif Occurrences) pipeline and MEME software[188] with default parameters.

Chapter 3: Genome-wide profiling of transcription factor–DNA binding interactions in *Candida albicans*

Mohammad N. Qasim, Ashley Valle Arevalo, Akshay D. Paropkari, Craig L. Ennis, Suzanne S. Sindi, Clarissa J. Nobile, Aaron D. Hernday
Journal of Visualized Experiments, 2022

3.1 Abstract

Regulatory transcription factors control many important biological processes, including cellular differentiation, responses to environmental perturbations and stresses, and host–pathogen interactions. Determining the genome-wide binding of regulatory transcription factors to DNA is essential to understanding the function of transcription factors in these often complex biological processes. Cleavage under targets and release using nuclease (CUT&RUN) is a modern method for genome-wide mapping of in vivo protein–DNA binding interactions that is an attractive alternative to the traditional and widely used chromatin immunoprecipitation followed by sequencing (ChIP-seq) method. CUT&RUN is amenable to a higher-throughput experimental setup and has a substantially higher dynamic range with lower per-sample sequencing costs than ChIP-seq. Here, a comprehensive CUT&RUN protocol and accompanying data analysis workflow tailored for genome-wide analysis of transcription factor–DNA binding interactions in the human fungal pathogen *Candida albicans* are described. This detailed protocol includes all necessary experimental procedures, from epitope tagging of transcription factor-coding genes to library preparation for sequencing; additionally, it includes a customized computational workflow for CUT&RUN data analysis.

KEYWORDS: *Candida albicans*, CUT&RUN, transcription factors, genome-wide binding methods, transcriptional regulation

3.2 Introduction

Candida albicans is a clinically relevant, polymorphic human fungal pathogen that exists in a variety of different modes of growth, such as the planktonic (free-floating) mode of growth and as communities of tightly adhered cells protected by an extracellular matrix, known as the biofilm mode of growth[189–191]. Similar to other developmental and cellular processes, biofilm development is an important *C. albicans* virulence trait that is known to be controlled at the transcriptional level by regulatory transcription factors (TFs) that bind to DNA in a sequence-specific manner[179]. Recently, chromatin regulators and histone modifiers have also emerged as important regulators of *C. albicans* biofilm formation[192] and morphogenesis[119] by mediating DNA accessibility. To understand the complex biology of this important fungal pathogen, effective methods to determine the genome-wide localization of specific TFs during distinct developmental and cellular processes would be valuable.

Chromatin immunoprecipitation followed by sequencing (ChIP-seq) is a widely used method to investigate protein–DNA interactions in *C. albicans*[119,192] and has largely replaced the more classical chromatin immunoprecipitation followed by microarray (ChIP-

chip)[193] method. Both ChIP-seq and ChIP-chip methods, however, require a large number of input cells[194], which can be a complicating factor when investigating TFs in the context of specific samples and modes of growth, such as biofilms collected from patients or animal models of infection. In addition, the chromatin immunoprecipitation (ChIP) assay often yields a significant amount of background signal throughout the genome, requiring a high level of enrichment for the target of interest to sufficiently separate signal from noise. While the ChIP-chip assay is largely outdated today, the sequencing depths necessary for ChIP-seq make this assay prohibitively expensive for many researchers, particularly those studying multiple TFs and/or chromatin-associated proteins.

Cleavage under targets and release using nuclease (CUT&RUN) is an attractive alternative to ChIP-seq. It was developed by the Henikoff lab in 2017 to circumvent the limitations of ChIP-seq and chromatin endogenous cleavage followed by sequencing ChEC-seq[195,196], another method to identify protein–DNA interactions on a genome-wide level, while providing high-resolution, genome-wide mapping of TFs and chromatin-associated proteins[197]. CUT&RUN relies on the targeted digestion of chromatin within permeabilized nuclei using tethered micrococcal nucleases, followed by sequencing of the digested DNA fragments[193,194]. As DNA fragments are specifically generated at the loci that are bound by a protein of interest, rather than being generated throughout the genome via random fragmentation as in ChIP assays, the CUT&RUN approach results in greatly reduced background signals and, thus, requires one-tenth of the sequencing depth compared to ChIP-seq[195,197,198]. These improvements ultimately lead to significant reductions in sequencing costs and reductions in the total number of input cells needed as starting material for each sample.

Here, a robust CUT&RUN protocol is described that has been adapted and optimized for determining the genome-wide localization of TFs in *C. albicans* cells isolated from biofilms and planktonic cultures. A thorough data analysis pipeline is also presented, which enables the processing and analysis of the resulting sequence data and requires users to have minimal expertise in coding or bioinformatics. Briefly, this protocol describes epitope tagging of TF-coding genes, harvesting of biofilm and planktonic cells, isolation of intact permeabilized nuclei, incubation with primary antibodies against the specific protein or epitope-tagged protein of interest, tethering of the chimeric A/G-micrococcal nuclease (pAG-MNase) fusion proteins to the primary antibodies, genomic DNA recovery after chromatin digestion, and preparation of genomic DNA libraries for sequencing.

The experimental CUT&RUN protocol is followed by a purpose-built data analysis pipeline, which takes raw DNA sequencing reads in FASTQ format and implements all required processing steps to provide a complete list of significantly enriched loci bound by the TF of interest (targeted by the primary antibody). Note that multiple steps of the described library preparation protocol have been specifically adapted and optimized for CUT&RUN analysis of TFs (as opposed to nucleosomes). While the data presented in this manuscript were generated using TF-specific adaptations of a commercial CUT&RUN kit, these protocols have also been validated using individually sourced components (i.e., pAG-MNase enzyme and magnetic DNA purification beads) and in-house-prepared buffers, which can significantly reduce experimental cost. The comprehensive experimental and data analysis protocols are described in detail below in a step-by-step format. All reagents and critical equipment, as well as buffer and media recipes, are listed in the **Table of Materials** and **Supplementary File 1**, respectively.

3.3 Results

This robust CUT&RUN protocol was adapted and optimized for investigating the genome-wide localization of specific TFs in *C. albicans* biofilms and planktonic cultures (see **Figure 3-2** for an overview of the experimental approach). A thorough data analysis pipeline is also included to facilitate analysis of the resulting CUT&RUN sequencing data and requires users to have minimal expertise in coding or bioinformatics (see **Figure 3-3** for an overview of the analysis pipeline). Contrary to the ChIP-chip and ChIP-seq methods, CUT&RUN is carried out using intact, permeabilized nuclei prepared from a significantly reduced number of input cells, without formaldehyde crosslinking. Isolating intact nuclei from *C. albicans* spheroplasts is a critical step in the protocol. Efficient spheroplasting via the digestion of the *C. albicans* cell wall using lyticase can be challenging as the enzymatic digestion reaction conditions must be optimized for each cell type. Thus, to ensure a successful CUT&RUN experiment with high-quality sequencing results, an early quality control step is included, and the presence of intact nuclei is verified using a standard fluorescence microscope.

Cell wall digestion and nuclear integrity are regularly assessed by visualizing both control intact cells and isolated nuclei stained with fluorescent cell wall and nucleic acid strains. In contrast to the isolated intact nuclei, where cell wall staining is not observed, both nuclei and the cell walls are fluorescently labeled in the intact control cells (**Figure 3-4A**). Lastly, prior to sequencing, fragment size distribution of CUT&RUN libraries is evaluated using a capillary electrophoresis instrument. This quality control step is a reliable measure in assessing the quality of CUT&RUN libraries. As seen in the electrophoresis trace on the top panel of **Figure 3-4B**, successful libraries generated for experiments investigating TFs show high enrichment for fragments smaller than 280 bp. The electrophoresis trace in the bottom panel of **Figure 3-4B** represents results from an unsuccessful CUT&RUN experiment.

Here, the peak at ~2,000 bp arose mainly from undigested DNA released from nuclei extensively damaged or broken during the CUT&RUN experiment. Assessing the fragment size distribution of the final pooled libraries is also recommended to confirm the complete removal of contaminating adapter dimers (**Figure 3-4C**). Typically, 5–10 million paired-end reads (~40 base read length) per library provide sufficient sequencing depth for most TF CUT&RUN experiments in *C. albicans*. Alternative sequencing read lengths, either paired-end or single-end, can be used to sequence CUT&RUN libraries. For example, paired-end read lengths between 25 and 150 bp should not affect the quality of the results. Single-end reads greater than or equal to 150 bp should also work for TF CUT&RUN experiments. However, the accompanying data analysis pipeline must be modified to accommodate single-end reads.

This CUT&RUN protocol and accompanying data analysis pipeline were validated by investigating two TFs, Ndt80 and Efg1, which control *C. albicans* biofilm formation. As shown in **Figure 3-4D**, Ndt80 is bound at intergenic regions (highlighted by the black bars indicating significantly enriched Ndt80 ChIP-chip loci) upstream of the EFG1 ORF. This intergenic region upstream of EFG1 was previously shown to be highly enriched for Ndt80 binding during biofilm formation by ChIP-chip4. However, CUT&RUN experiments identified significantly more peaks within this region than ChIP-chip experiments (10 peaks vs 4 peaks). Ndt80 DNA binding motifs are enriched across all the Ndt80-bound loci

identified by CUT&RUN (**Supplementary Figure 3-1**), indicating that the additional peaks identified by this methodology are likely bona fide Ndt80-bound sites. A systematic comparative analysis indicated that this presented CUT&RUN protocol successfully identified the majority of the previously known binding events for Ndt80 and Efg1 during biofilm formation4 (**Figure 3-5**).

Furthermore, many new TF binding events that were not captured in the previously published ChIP-chip experiments were identified using CUT&RUN (**Figure 3-5**). Overall, both Ndt80 and Efg1 bound to loci overlapping with previously published ChIP-chip data and to loci identified only using the CUT&RUN method (overlaps between these CUT&RUN data and previously published ChIP-chip data for Ndt80 and Efg1 are summarized in the Venn diagrams in **Figure 3-5**). Furthermore, the fraction of reads within significantly called peaks (FRiP scores) were consistently higher in GFP-tagged samples than in their control IgG samples (**see Supplementary Figure 3-2**). In summary, these results show that the CUT&RUN protocol described here is a robust method optimized to investigate *C. albicans* TF-DNA binding interactions from low-abundance samples.

3.4 Figure and Figure Legends

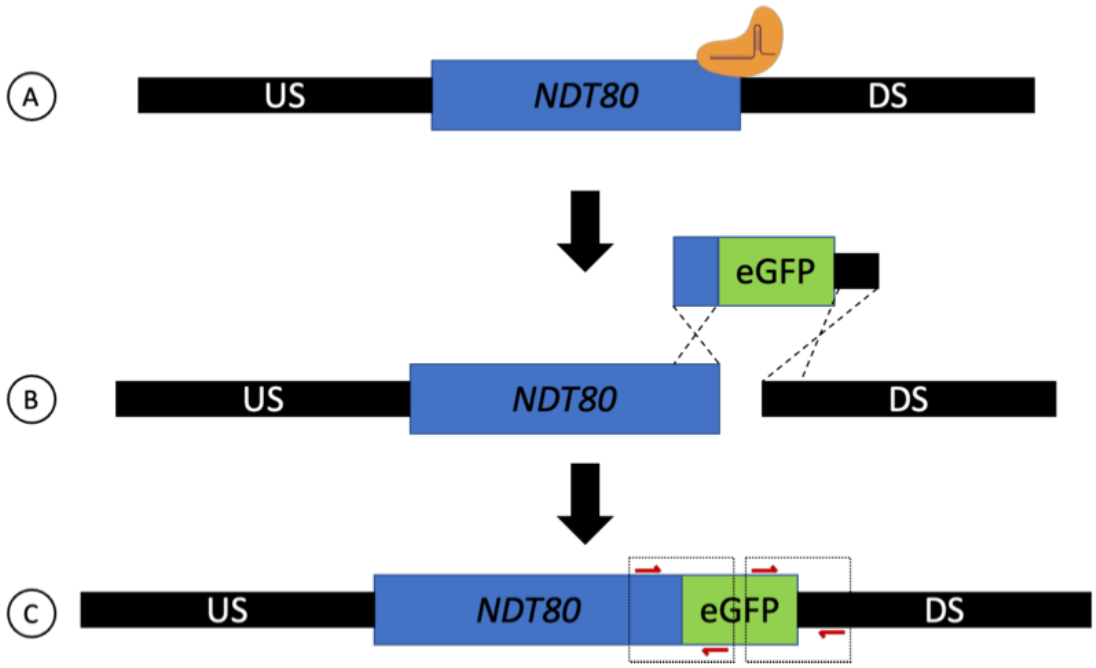


Figure 3-1: Epitope tagging of *C. albicans* TFs.

(**A**) Identify a TF gene of interest (*NDT80* in this case) and select a gRNA to target the cutting of Cas9 (represented by the orange shape) at the 3' end of the ORF. (**B**) *Candida* clade-optimized eGFP with >50 bp homology upstream and downstream from the cut site will provide the dDNA to repair the DSB. (**C**) Use colony PCR to screen individual colonies to confirm the intended integration. Primers are indicated by red arrows, and amplicons are denoted by dashed boxes. This figure was created using BioRender.com. Abbreviations: TF = transcription factor; gRNA = guide RNA; ORF = open reading frame; eGFP = enhanced green fluorescent protein; DSB = double-stranded break; US = upstream; DS = downstream.

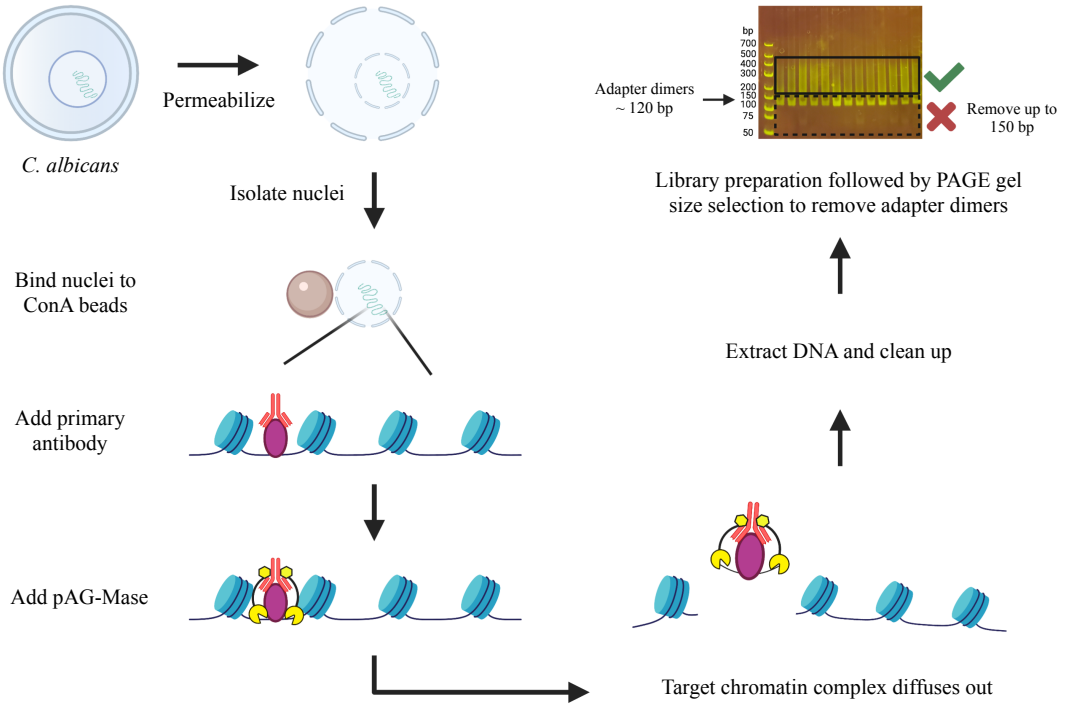


Figure 3-2: Schematic overview of the CUT&RUN experimental protocol.

C. albicans cells collected from biofilm or planktonic culture conditions are permeabilized to isolate intact nuclei. ConA beads are activated, and the intact nuclei are then bound to the activated ConA beads. The antibody of interest is added to the bead-bound nuclei and incubated at 4 °C. Next, pAG-MNase is added and allowed to bind to the target antibody. After the addition of $CaCl_2$, pAG-MNase is activated, and targeted chromatin digestion proceeds until the addition of the chelating reagent to inactivate pAG-MNase. The pAG-MNase-bound antibody complex is allowed to diffuse out of the permeabilized nuclei, and the resulting DNA is extracted and cleaned. Sequencing libraries are prepared from the CUT&RUN-enriched DNA fragments, and the resulting libraries are then run on a 10% PAGE gel to separate and remove contaminating adapter dimers prior to sequencing. This figure was created using BioRender.com. Abbreviations: CUT&RUN = cleavage under targets and release using nuclease; ConA = concanavalin A; PAGE = polyacrylamide gel electrophoresis.

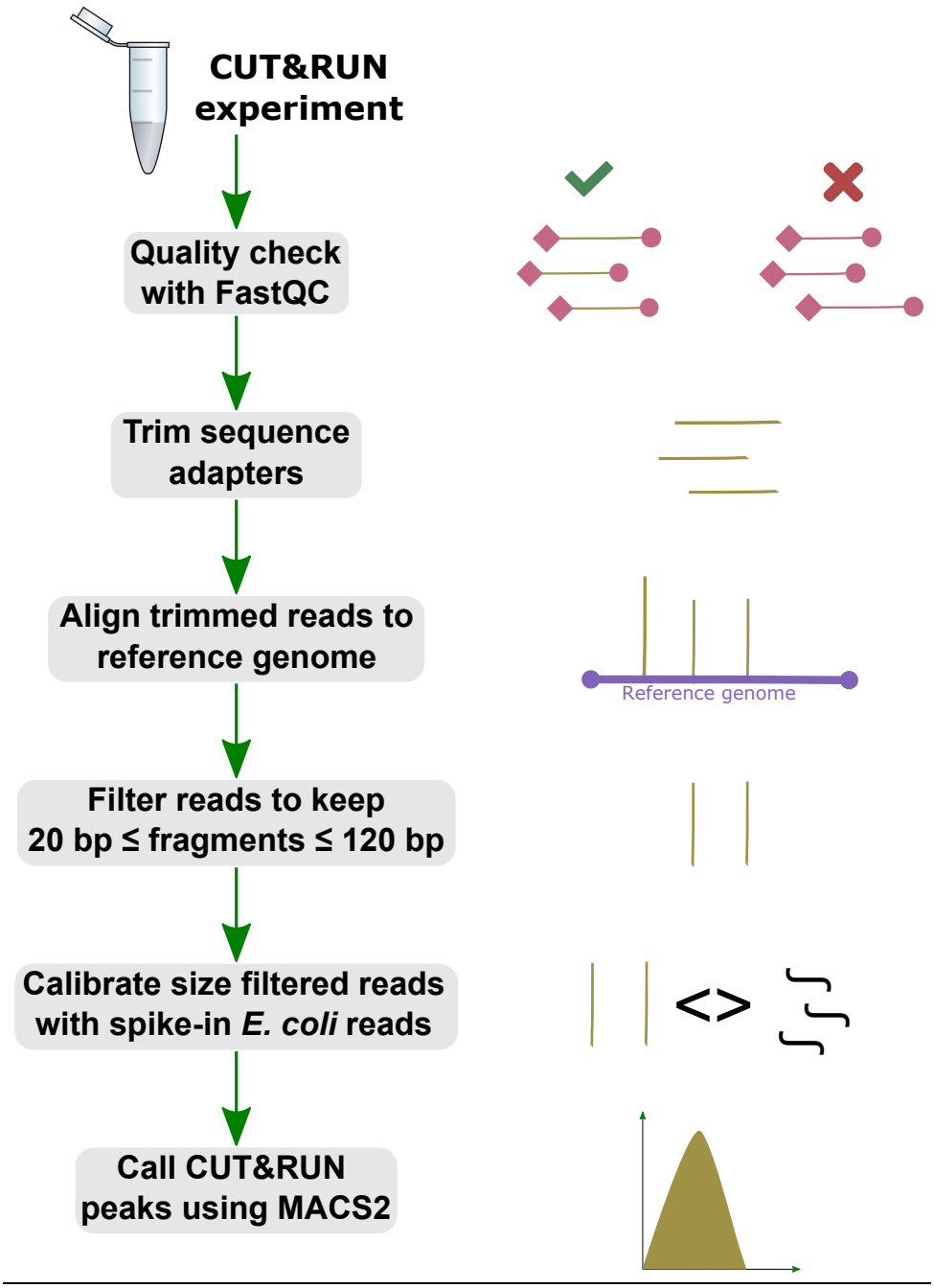


Figure 3-3: Schematic overview of the CUT&RUN data analysis pipeline.

The workflow starts by performing quality checks on the raw FASTQ files using FastQC, followed by trimming to remove sequencing adapters. The trimmed reads are then aligned to the reference genome, and the aligned reads are filtered based on their size to enrich for TF-sized binding signals (20 bp ≤ aligned read ≤ 120 bp). Size-selected reads are then calibrated against spike-in *Escherichia coli* reads, and calibrated reads are used to call peaks using MACS2. Abbreviations: CUT&RUN = cleavage under targets and release using nuclease; TF = transcription factor.

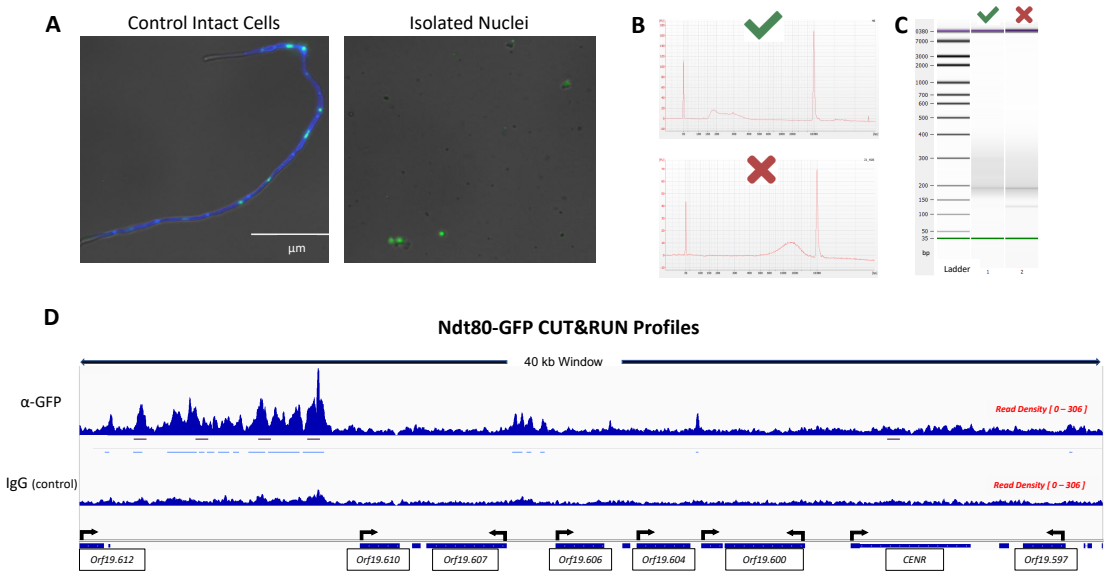
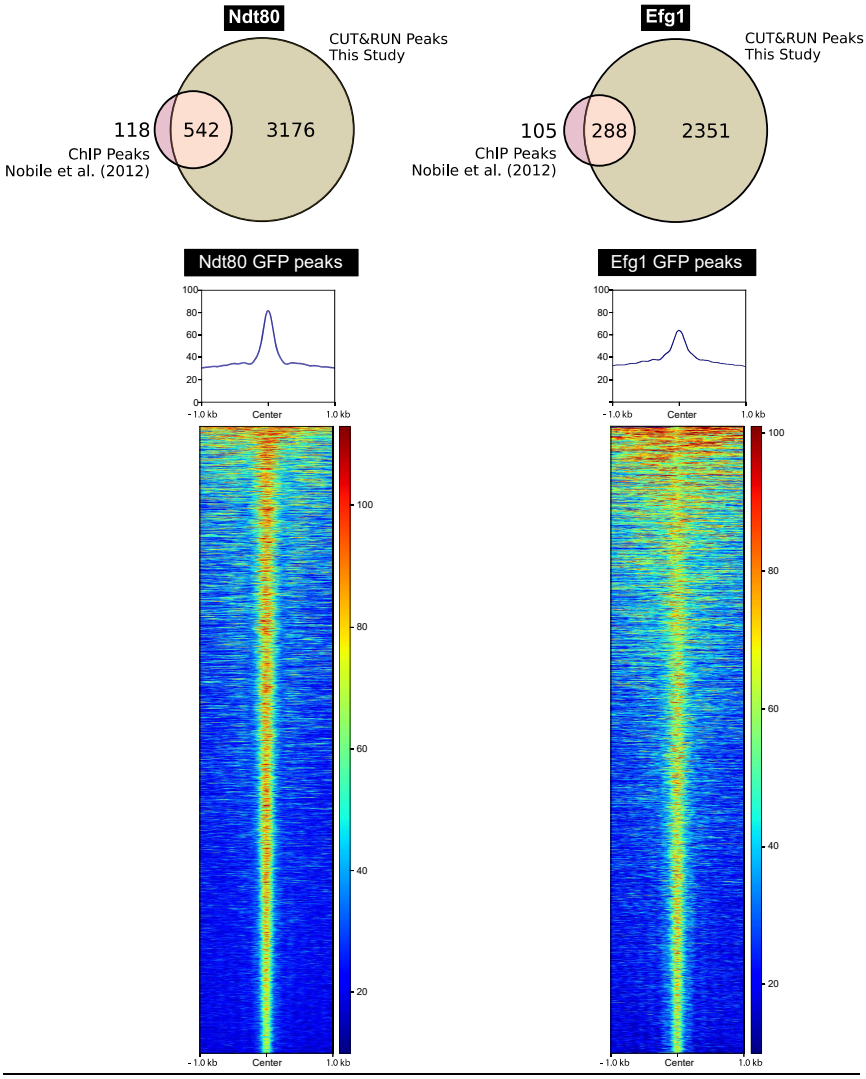


Figure 3-4: Quality control steps critical for successful CUT&RUN experiments.

(**A**) Cells are stained with fluorescent cell wall (blue) and nucleic acid (green) stains before and after nuclei isolation and visualized using a fluorescence microscope. (**B**) CUT&RUN TF libraries are analyzed using a capillary electrophoresis instrument. Successful CUT&RUN TF libraries (indicated by the green checkmark) are enriched for short fragments smaller than 200 bp. Suboptimal CUT&RUN TF libraries (indicated by the "X") show enrichment for large DNA fragments. (**C**) 48 CUT&RUN libraries are pooled together and analyzed using a capillary electrophoresis instrument. High-quality, pooled libraries (indicated by the green checkmark) are free of adapter dimers (lane 1), while low-quality pooled libraries (indicated by the red "X") retain small amounts of adapter dimers (lane 2). (**D**) Representative IGV tracks from CUT&RUN datasets (including the Ndt80 IgG control, bottom track) showing significant enrichment for Ndt80 binding (top track) at the intergenic region upstream of the *EFG1* ORF (Orf19.610). The black bars represent significantly enriched Ndt80 binding peaks identified by previously published ChIP-chip experiments[179]. The blue bars represent significantly enriched Ndt80 binding peaks identified in the CUT&RUN experiments presented here. Scale bars = 50 µm (**A**). Abbreviations: CUT&RUN = cleavage under targets and release using nuclease; TF = transcription factor; GFP = green fluorescent protein; ORF = open reading frame.



<u>Figure 3-5</u>: Evaluation of Ndt80 and Efg1 enriched peaks identified using the presented CUT&RUN protocol and data analysis pipeline on *Candida albicans* biofilms.

The Venn diagrams in the top row illustrate the degree of overlap between Ndt80 and Efg1 binding sites identified via CUT&RUN with previously published ChIP-chip data[179]. The CUT&RUN signals for all binding events for Ndt80 and Efg1 are shown in the bottom row as colored heatmaps (red = high peak signal; blue = low/no peak signal; colored bar indicates IgG signal subtracted from GFP signal); 1,000 bp regions upstream (-1.0 kb) and downstream (+1.0 kb) are displayed in the heatmaps. The signal intensity (i.e., enrichment) as a profile plot is shown above the heatmaps. Three biological replicates for Efg1 and Ndt80 were evaluated for visualizing the CUT&RUN enrichment. The <> symbol is a visual representation to indicate that *C. albicans* read counts were calibrated using *Escherichia coli* read counts for each sample. Abbreviations: ChIP = chromatin immunoprecipitation; CUT&RUN = cleavage under targets and release using nuclease; GFP = green fluorescent protein.

3.5 Discussion

This protocol presents a comprehensive experimental and computational pipeline for genome-wide localization of regulatory TFs in *C. albicans*. It is designed to be highly accessible to anyone with standard microbiology and molecular biology training. By leveraging the high dynamic range and low sample input requirements of the CUT&RUN assay and including optimizations for the localization of TF–DNA binding interactions in *C. albicans* biofilm and planktonic cultures, this protocol presents a powerful and low-cost alternative to traditional ChIP-seq approaches. Compared to ChIP-seq, CUT&RUN yields significantly higher sensitivity, with a higher proportion of sequencing reads mapped to bound peaks, is more amenable to high-throughput, requires substantially lower input cell numbers, does not require the use of toxic crosslinking agents, and requires tenfold fewer sequencing reads per sample to produce high-quality results[197–201]. To further reduce the per-sample cost of this protocol, buffer and media recipes are included along with a detailed reagent list to facilitate the in-house preparation of all necessary buffers and media, as well as the bulk sourcing of essential reagents. As *C. albicans* biofilm formation, phenotypic switching, and commensalism are all regulated by complex interwoven transcriptional networks[202], this robust, facile, and cost-effective CUT&RUN protocol provides a powerful new tool for understanding these and many other cellular processes in this important human fungal pathogen.

TFs are not as abundant as histones or other chromatin-associated proteins, creating a unique challenge for investigating TF–DNA binding interactions via CUT&RUN. To address this challenge, critical adjustments and optimizations were made to the standard CUT&RUN experimental protocol. As most successful CUT&RUN experiments targeting TFs yield a small amount of DNA that is too dilute to quantify and is often enriched for fragments smaller than 150 bp[197,200], the library preparation reaction conditions were also optimized to favor these smaller fragments[203]. Even with this optimization step, the resulting PCR-amplified libraries contain a significant proportion of adapter dimers, which are not completely removed using magnetic bead-based DNA size selection methods. To address this issue, a PAGE gel size-selection step was included to generate final sequencing-ready libraries that are largely devoid of adapter dimers. This is a critical step of the protocol, as removing adapter dimers while retaining the smaller TF-derived CUT&RUN fragments is essential for obtaining high-quality results.

Furthermore, the detailed computational pipeline filters the sequencing data to focus on the smaller reads derived from TF–DNA binding interactions in the CUT&RUN assay. Due to these TF-specific adjustments, this protocol is not recommended for profiling large chromatin-associated complexes such as nucleosomes. While it is theoretically possible to adapt the protocol for this purpose by following the standard library preparation protocol included with the commercially available library prep kit, the user would need to adjust the post sequencing size selection included in the computational pipeline. Specifically, in the size-filtering section in the code file cut_n_run_pipeline.sh, users would need to replace the current value of “14400” (120 bp * 120 bp) with the square of the desired fragment length to enable the analysis pipeline to analyze the sequencing results generated for other types of chromatin–DNA binding interactions.

Another key step in a successful CUT&RUN experiment includes choosing optimal post sequencing data analysis parameters. While most of the computational pipeline is designed to be standardized and applicable to the study of any regulatory TF of interest

in *C. albicans*, there are two important considerations that the user should evaluate while running the pipeline. The first consideration is whether to include or remove duplicate reads from the sequencing data prior to the identification of bound target sites. As low-abundance TFs will typically yield sequencing data containing a significant percentage of reads that are derived from PCR duplication during the library amplification step, removing PCR duplicates can have a significantly negative impact on the results. However, with highly abundant TFs or chromatin-associated proteins, PCR duplicates typically represent a smaller portion of the total number of reads and are often removed to suppress background noise in the data. Ultimately, this decision to keep or remove PCR duplicates is dependent on the TF of interest and the depth of the sequencing data obtained. Thus, the pipeline automatically generates independent output files for data derived with or without PCR duplicate reads, so the user can decide which output files yield the best results for each experiment.

The second consideration is whether to identify and remove problematic loci that yield significant, yet highly variable, enrichment in both experimental (antibody against the protein of interest) and negative control (IgG) samples. The peak-calling algorithm uses MACS2 to identify significantly enriched loci in both the experimental and control samples and excludes those that appear in both. While this approach typically eliminates most problematic loci, some occasionally appear as significant peaks in certain experiments, even though prior experience indicates that they are unlikely to be true positive sites of TF enrichment. Thus, an optional filtering step is provided to remove these problematic loci, referred to as “blocklisted” loci. The list of blocklisted loci primarily contains highly repetitive sequence elements and regions such as telomeric repeats and centromeres that have historically yielded false-positive results in previous genome-wide binding assays. This is a very conservative list of loci assigned as problematic with high confidence. However, each user should evaluate whether this filter is appropriate for their experiment(s) on a case-by-case basis. A potential alternative to the IgG negative control would be to perform CUT&RUN against a nuclear-localized protein that does not bind DNA. This approach has been shown to be an ideal control for ChIP experiments[204], and a similar approach is worth considering for CUT&RUN.

CUT&RUN has become a popular choice for investigating protein–DNA interactions in higher eukaryotes and the model yeast Saccharomyces cerevisiae. Here, it has been successfully adapted to investigate genome-wide TF–DNA binding interactions in the clinically relevant fungal pathogen *C. albicans*. This protocol provides detailed methods for all necessary experimental and computational procedures, from the engineering of strains that express epitope-tagged TFs, through to the computational analysis of the resulting CUT&RUN sequencing data. Overall, this protocol and the accompanying data analysis pipeline produce robust TF–DNA binding profiles, even when using complex multimorphic populations of cells isolated from low-abundance biofilm samples and provides superior data quality at a lower overall cost than ChIP-seq methodologies.

3.6 Methods and Materials

1. Epitope Tagging of *C. albicans* Strains

1.1. Upload the gene of interest, along with its 1 kb upstream and downstream flanking sequences, from the *Candida* Genome Database to the primer design tool (see the **Table of Materials**). Design a guide RNA (gRNA) by highlighting 50 bp upstream and downstream from the stop codon and click the **gRNA selection**

tool on the right. Select **Design and Analyze Guides**. Use the Ca22 (*Candida albicans* SC5314 Assembly 22 (diploid)) genome and an NGG (SpCas9, 3' side) protospacer adjacent motif (PAM) for the guide parameters and click **Finish**.

NOTE: **Figure 3-1** describes the workflow for epitope tagging a *C. albicans* gene of interest with enhanced green fluorescent protein (eGFP).

1.1.1. On the subsequent page, confirm the target region for gRNA design and press the green button. Sort gRNAs by the **On-Target Score**.

NOTE: The primer design tool computes on-target and off-target scores to quantify the specificity of the gRNAs. An ideal guide has an on-target score of >60, an off-target score of ~33, and overlaps the stop codon. This enables high gRNA specificity while ablating gRNA targeting after GFP integration. A gRNA with an off-target score of ~50 indicates allelic variation; thus, only one allele will be recognized by the gRNA.

1.1.2. Add the sequences (5'-CGTAAACTATTTTTAATTTG-3') and (5'-GTTTTAGAGCTAGAAATAGC-3') to the 5' and 3' ends, respectively, of the 20 bp gRNA target sequence, creating a 60 bp primer/oligonucleotide. Alternatively, copy the 20 bp sequence to the gRNA calculator supplied by Nguyen et al.[205]. Order the 60 bp custom gRNA oligonucleotide.

1.1.3. Amplify the “universal A fragment” with 100 mM AHO1096 (5'-GACGGCACGGCCACGCGTTTAAACCGCC-3') and 100 mM AHO1098 (5'-CAAATTAAAAATAGTTTACGCAAG-3') and the “unique B fragment” with the custom 60 bp gRNA oligonucleotide (100 mM) from step 1.1.2. and 100 mM AHO1097 (5'-CCCGCCAGGCGCTGGGGTTTAAACACCG-3') using pADH110 (plasmid repository ID# 90982) and pADH139 (plasmid repository ID# 90987), respectively, as template DNA. Use the PCR reaction and cycling conditions provided in **Table 3-1**.

NOTE: pADH139 is specific to strains that carry the heterologous *Candida maltosa LEU2* marker. If using a strain with a single copy of the *C. albicans LEU2* gene, substitute pADH119 (plasmid repository ID# 90985) in place of pADH139

1.1.4. Confirm successful amplification by checking 5 µL of the PCR on a 1% agarose gel. Look for ~1 kb A and B fragment amplicons.

1.1.5. Mix 1 µL each of A and B fragments and stitch them together using the PCR reaction and cycling conditions provided in **Table 3-2** to create a full-length C fragment.

1.1.6. Add 0.5 µL of 100 mM AHO1237 (5'-AGGTGATGCTGAAGCTATTGAAG-3') and 0.5 µL of 100 mM AHO1453 (5'-ATTTTAGTAACAGCTTCGACAATCG-3') to each PCR reaction, mix well by pipetting, and complete the cycling conditions listed in **Table 3-3**.

NOTE: If using pADH119 in place of pADH139, substitute AHO1238 (5'-TGTATTTTGTTTTAAAATTTTAGTGACTGTTTC-3') in place of AHO1453.

1.1.7. Confirm proper stitching and amplification of the C fragment by checking 5 µL of the PCR on a 1% agarose gel. Look for a ~2 kb amplicon. Store the C fragment at -20 °C until ready for use.

NOTE: If stitching and amplification results in multiple, nonspecific bands or smearing, perform a PCR cleanup of the A and B fragments and repeat from step 1.1.5.

1.2. Add the entire CTG-optimized monomeric eGFP with linker sequence (RIPLING)[206] (pCE1, plasmid repository ID# 174434) immediately upstream of the stop codon of the gene of interest using the primer design tool, creating a C-terminal translational fusion. Use this construct to design oligonucleotides for amplifying the donor DNA (dDNA) from pCE1.

1.2.1. Design a forward oligonucleotide with 18–22 bp homology to the linker sequence and >50 bp homology to the 3' end of the open reading frame (ORF).

NOTE: The 18–22 bp homology creates an annealing temperature for amplification between 55 °C and 58 °C. If the full-length oligonucleotide forms primer dimers, adjust homology to the linker sequence/GFP or the genome accordingly.

1.2.2. Create a reverse oligonucleotide with 18–22 bp homology to the 3' end of GFP and >50 bp homology to the downstream noncoding sequence of the ORF to be tagged.

1.2.3. Order these oligonucleotides and amplify the dDNA using the provided touchdown PCR cycling conditions in **Table 3-4**.

1.3. Design two sets of colony PCR (cPCR) oligonucleotides for confirming the integration of GFP by amplifying across the flanking integration sites. First, select the forward dDNA oligonucleotide using the primer design tool and click the **Primer** button on the right.

1.3.1. Click **Create Primers | Wizard | Tm Param** and confirm that the algorithm is set to **SantaLucia 1998**. Click **Use Selection** to input the coordinates of the forward oligonucleotide designed in 1.2.1 as the target sequence.

1.3.2. Set the optimal primer temperature to 55 °C and the maximum amplicon size to 900 bp, and click the **Generate Primers** button at the top right.

1.3.3. Select the oligonucleotide pair with the lowest penalty score and confirm that the primers amplify across the 5' integration site. Ensure that the forward cPCR primer lies upstream of the forward dDNA primer sequence and the reverse cPCR primer fully within the eGFP tag or the linker sequence.

1.3.4. Repeat steps 1.3–1.3.3. with the reverse dDNA oligonucleotide to create the second set of cPCR oligonucleotides that amplify across the 3' integration site. Ensure that the forward cPCR primer lies entirely within the eGFP tag or the linker sequence and the reverse cPCR primer downstream of the reverse dDNA primer sequence.

1.3.5. Order these oligonucleotides.

1.4. Digest 2,500 ng of pADH140, which contains Cas9 (plasmid repository ID# 90988), with restriction enzyme for each gene to be GFP-tagged. Set the total volume of each digestion at 15 µL; adjust the volume of water accordingly based on the pADH140 plasmid concentration. Use the digestion conditions specified in **Table 3-5**. Store the digested plasmid at -20 °C until ready for use.

NOTE: If transforming a strain with a single copy of the *C. albicans LEU2* gene, instead of the heterologous *C. maltosa LEU2* marker, substitute pADH137 (plasmid repository ID# 90986) in place of pADH140.

1.5. Denature 12 µL of 10 mg/mL salmon sperm DNA for each gene that will be GFP-tagged at 99 °C for 10 min and rapidly cool to ≤4 °C. Store at -20 °C until ready for use.

1.6. Streak a *C. albicans LEU2* hemizygous nourseothricin-sensitive strain onto yeast peptone dextrose (YPD) plates and incubate at 30 °C for two days.

1.7. Select a single colony and transfer it to 4 mL of liquid YPD. Incubate for 12–16 h at 30 °C with shaking at 250 rpm.

1.8. Measure the optical density at 600 nm (OD_{600}) of the overnight (12–16 h) culture in a spectrophotometer using a disposable cuvette (1 mL, 1 cm path length).

1.9. Dilute the overnight culture into an Erlenmeyer flask to an OD_{600} of 0.1 in YPD. Account for 5 mL per reaction and include an additional 5 mL for checking the OD_{600} later.

NOTE: The volume of the culture depends on the number of transformation reactions.

1.10. Incubate the diluted overnight culture in a shaking incubator at 30 °C with shaking at 250 rpm until it reaches an OD_{600} of 0.5–0.8.

1.11. Centrifuge at 4,000 × *g* at room temperature for 5 min; remove and discard the supernatant.

1.12. Resuspend the cell pellet in 1 mL of sterile water via gentle pipette mixing with filter tips and transfer to a sterile 1.5 mL microfuge tube.

1.13. Pellet the cells by centrifuging at 4,000 × *g* at room temperature for 1 min;

remove and discard the supernatant. Resuspend in 1 mL of sterile water and repeat for a total of two washes.

1.14. Resuspend the pellet in 1/100th of the volume used in step 1.10. For example, if 15 mL was used, resuspend the pellet in 150 µL of sterile water.

1.15. In a separate tube for each transformation reaction, mix 50 µL of C fragment, 50 µL of dDNA, 2,500 ng of restriction enzyme-digested pADH140, and 10 µL of denatured salmon sperm DNA.

1.16. Add 50 µL of the cell slurry from step 1.14 and mix by pipetting.

NOTE: Tap the bottoms of the tubes while inverting to dislodge any remaining liquid.

1.17. Make a stock of the plate mix (**Supplementary File 3-1**) for n + 1 transformations.

1.18. Add 1 mL of the plate mix to the cell/DNA mixture and mix by inverting 5 times.

1.19. Place the mixture in an incubator at 30 °C overnight (12–16 h) without shaking.

1.20. Heat-shock the cells for 15 min at 44 °C in a water bath.

1.21. Centrifuge the 1.5 mL microfuge tubes at 5,000 × *g* at room temperature for 2 min.

1.22. Remove the PLATE mix by vacuum aspiration using sterile pipette tips, being careful to avoid disturbing the cell pellet.

1.23. Resuspend the cell pellet in 1 mL of YPD, pellet by centrifugation at 4,000 × *g* at room temperature for 1 min and remove and discard the supernatant. Repeat for a second wash, resuspend the cell pellet in 1 mL of YPD, and transfer the suspension to a 10 mL round-bottom, disposable culture tube containing an additional 1 mL of YPD (2 mL final volume). Recover the cells at 30 °C with shaking at 250 rpm for 5 h.

1.24. Centrifuge the tubes at 4,000 × *g* at room temperature for 5 min; remove and discard the supernatant.

1.25. Resuspend the cell pellet in 100 µL of sterile water and plate on YPD supplemented with 200 µg/mL nourseothricin (NAT200). Incubate at 30 °C for 2–3 days.

1.26. Aliquot 100 µL of 20 mM NaOH into the wells of a 96-well PCR plate, with each well corresponding to an individual colony that grew on the NAT200 plates. Using a sterile toothpick or pipette tip, pick individual transformed colonies, patch them onto a new NAT200 plate, and swirl the remaining cells into a well with 20 mM NaOH. Repeat for the remaining colonies to create the cell lysate used as the

DNA template for the cPCR reaction.

1.27. Seal the PCR plate and incubate for 10 min at 99 °C in a thermocycler with a heated lid.

1.28. Set up two cPCR reactions with the oligonucleotides designed in steps 1.3–1.3.5. Scale up the number of reactions as needed. Perform the PCR reaction with the cell lysate prepared in step 1.26 following cycling conditions and PCR reaction mixtures from **Table 3-6**. Run 10–20 µL from each well on a 1% agarose gel. Look for colonies with amplification of the two cPCR primer sets indicating properly incorporated GFP dDNA.

1.29. Restreak colonies that incorporated GFP on synthetic complete (SC) media lacking leucine. Incubate in a 30 °C incubator for 2–3 days. Pick individual colonies and patch onto YPD and YPD supplemented with 400 mg/mL nourseothricin (NAT400) plates. Identify colonies that fail to grow on NAT400 plates after 24 h as those that have successfully lost the CRISPR components.

1.30. Confirm that the GFP tag is retained by repeating steps 1.25–1.28 using cells from the YPD patch plate. If the correct bands are present, inoculate into 4 mL of YPD and grow overnight (12–16 h), as described in step 1.7.

1.31. Mix the overnight culture of the new GFP-tagged strain with filter-sterilized 50% glycerol in a 1:1 ratio in a sterile cryotube. Store at -80 °C and restreak onto YPD plates as needed.

NOTE: It is recommended to validate the GFP-tagged strains by confirming nuclear localization of the tagged TF via fluorescent microscopy and confirming a wild-type phenotype in an appropriate phenotypic assay.

2. Sample preparation of biofilm cultures

2.1. Streak *C. albicans* GFP-tagged strain(s) onto YPD agar plates and incubate at 30 °C for 2–3 days. Using a single isolated colony from the agar plate, inoculate into 4 mL of YPD liquid medium. Incubate at 30 °C with shaking overnight (12–16 h). Determine the OD_{600} of the overnight culture(s).

NOTE: It is recommended to use three biological replicates per sample for the CUT&RUN experiments.

2.2. Inoculate a sterile 12-well untreated cell culture plate with the overnight culture to a final OD_{600} of 0.5 (equivalent to 2×10^7 cells/mL) in Roswell Park Memorial Institute (RPMI)-1640 medium to a final volume of 2 mL. Incubate for 90 min at 37 °C in a microplate incubator with shaking at 250 rpm.

NOTE: It is recommended to use one 12-well cell culture plate per strain with one well uninoculated as a medium-alone contamination control. This protocol has been successfully applied using as little as $1/10^{th}$ of one 12-well cell culture plate well (or as few as 5 million cells). Using a higher number of cells increases total DNA yields, which typically results in high-quality sequencing libraries.

2.3. Remove unadhered cells by aspiration using sterile pipette tips attached via flexible plastic tubing to a vacuum trap apparatus. Wash the adhered cells once with 2 mL of sterile 1x phosphate-buffered saline (PBS). Add 2 mL of fresh RPMI-1640 medium to the wells and incubate for 24 h at 37 °C with shaking at 250 rpm.

NOTE: Change pipette tips between wells of different strains and/or conditions. Do not scrape the bottom of the well with the tip while aspirating.

2.4. At the end of the 24 h incubation, collect and pool the liquid and biofilm material from each of the 11 inoculated wells into a single, sterile 50 mL conical tube. Repeat as necessary with independent pools if processing more than one strain or growth condition concurrently.

NOTE: Scrape the bottoms and edges of each well with a pipette filter tip to dislodge cells that remain adhered to the surface. Use the pipette to homogenize the biofilms.

2.5. Pellet samples by centrifuging at 4,000 × *g* at room temperature for 5 min. Decant as much of the supernatant as possible, taking care to minimize disruption of the pellet. Snap-freeze the pellet in liquid nitrogen and store at -80 °C immediately after collection or continue directly to step 4 (isolation of nuclei).

3. Sample preparation of planktonic cultures

3.1. Streak *C. albicans* GFP-tagged strain(s) onto YPD agar plates and incubate at 30 °C for 2–3 days. Using a single isolated colony from the agar plate, inoculate into 4 mL of YPD liquid medium. Incubate at 30 °C with shaking overnight (12–16 h). Determine the OD_{600} of the overnight culture(s).

3.2. Back-dilute overnight cultures to OD_{600} of 0.1 in 50 mL of RPMI-1640 liquid medium and incubate at 30 °C with shaking at 225 rpm for 2–5 h until OD_{600} is between 0.5 and 0.8.

NOTE: Cells should go through at least two doublings before being harvested. Conditions used for planktonic cultures can be adjusted as needed.

3.3. Pellet the samples by centrifuging at 4,000 × *g* at room temperature for 5 min. Decant as much of the supernatant as possible, taking care to minimize disruption of the pellet. Snap-freeze the pellet in liquid nitrogen and store at -80 °C immediately after collection or continue directly to step 4 (isolation of nuclei).

4. Isolation of nuclei

4.1. Resuspend pellet(s) in 1 mL of room-temperature Resuspension Buffer and transfer to a sterile 1.5 mL microfuge tube. Pellet at 2,000 × *g* at room temperature for 2 min in a table-top centrifuge and remove the supernatant.

4.2. Resuspend the pellet(s) in 200 µL of room-temperature Resuspension Buffer. From the resuspended pellet, transfer a 5 µL aliquot into a new PCR tube and store it at 4 °C for use later.

NOTE: This aliquot will be used as a control during a subsequent quality control step to evaluate the quality of the isolated nuclei.

4.3. Pellet at 2,000 × *g* for 2 min and remove and discard the supernatant using a pipette. Repeat the wash step twice using 200 µL of Resuspension Buffer.

4.4. Centrifuge at 2,000 × *g* at room temperature for 2 min and remove the supernatant. Add 300 µL of Resuspension Buffer and 10 µL of lyticase solution (50 mg/mL, see the **Table of Materials**). Incubate for 30 min at 30 °C in a heat block.

NOTE: Alternatively, a water bath heated to 30 °C can also be used instead of a heat block. The spheroplasting conditions used here have been optimized to be effective for both yeast and hyphal cells of *C. albicans*. It is recommended to optimize the spheroplasting conditions when applying this protocol to *C. albicans* cells with mutations that impact cell wall integrity or other cellular morphologies. During this 30 min incubation step, the user has the option to complete step 5 (Concanavalin A Bead Activation) ahead of time to save time.

4.4.1. **CRITICAL STEP**: After the 30 min incubation step, transfer a 5 µL aliquot into a new PCR tube. To the 5 µL of isolated nuclei and the 5 µL aliquot of intact cells stored at 4 °C from step 4.2, add 1 µL calcofluor white (a fluorescent cell wall dye) and 1 µL of SYTO 13 (a nucleic acid stain). Incubate at 30 °C in the dark for 30 min.

4.4.2. Visually inspect the integrity and purity of the isolated nuclei using a fluorescence microscope. Look for isolated nuclei that show prominently stained intact nuclei (using a 488–509 nm excitation filter) and ensure that there is no cell wall staining by the calcofluor white dye (using a 390-420 nm excitation filter). In the intact control cells, look for prominent cell wall staining by the calcofluor white dye (using a 390-420 nm excitation filter) and stained intact nuclei (using a 488–509 nm excitation filter).

4.5. Centrifuge at 2,000 × *g* at 4 °C for 5 min and remove the supernatant. Resuspend the pellet in 500 µL of ice-cold Resuspension Buffer using 1 mL filter tips by pipetting gently up and down 5 times. Centrifuge at 2,000 × *g* at 4 °C for 5 min, and remove the supernatant using a 1 mL pipette. Resuspend the pellet with 1 mL of freshly made ice-cold Ficoll Buffer.

NOTE: Keep the samples and buffers on ice from this point forward.

4.6. Centrifuge the samples at 5,000 × *g* at 4 °C for 10 min and remove the supernatant. Resuspend the pellet in 500 µL of ice-cold SPC Buffer.

NOTE: From this point onward, handle the nuclei extremely gently to avoid damaging them.

4.7. Centrifuge the samples at 5,000 × *g* at 4 °C for 10 min and remove as much of the supernatant as possible without disrupting the pellet. Place the tubes

containing the pelleted nuclei on ice and proceed to step 5. If step 5 was already completed ahead of time in step 4.4, proceed to step 6 or snap-freeze the pellets in liquid nitrogen and store them at -80 °C immediately after collection.

5. Concanavalin A bead activation

NOTE: This is a critical step. From this point forward, users have the option to continue with the protocol using a commercially available CUT&RUN kit or source key components individually and prepare buffers in-house. If using the commercial kit, all buffers and reagents used below are included in the kit unless otherwise noted. Individual catalog numbers for sourcing reagents independently are also provided in the **Table of Materials**. Chill all buffers on ice before use. Once step 5 is completed, it is recommended to proceed to step 6 immediately. Avoid multiple freeze-thawing of isolated nuclei as it is known to increase DNA damage and could lead to poor-quality results.

5.1. Gently resuspend the concanavalin A (ConA) beads using a pipette. Transfer 22 µL of ConA bead suspension per sample to be processed in a single 1.5 mL microfuge tube. Place the tube on a magnetic rack until the bead slurry is clear; remove and discard the supernatant using a pipette.

NOTE: When performing CUT&RUN for a total of 10 samples, for example, transfer 220 µL of the ConA bead suspension to a 1.5 mL microfuge tube.

5.2. Remove the tube containing the ConA beads from the magnetic rack and immediately add 200 µL of ice-cold Bead Activation Buffer and gently mix using a pipette. Place the tube on the magnetic rack until the bead slurry is clear; remove and discard the supernatant using a pipette. Repeat this step for a total of two washes.

5.3. Resuspend the beads in 22 µL of ice-cold Bead Activation Buffer per sample of nuclei to be processed. Keep the beads on ice until needed.

NOTE: The throughput of the subsequent steps is dependent on the number and capacity of magnetic tube racks available. Processing 32 samples in two 16-well tube racks is a manageable number for most users of this protocol. However, higher throughput is possible for more experienced users, or if robotic liquid handling systems are available.

6. Binding nuclei to activated beads

NOTE: Chill all buffers on ice before use. All buffers supplemented with protease inhibitors should be prepared fresh on the day of the experiment. It is recommended to use 0.2 mL strip tubes in the subsequent steps.

6.1. Resuspend the pelleted nuclei from step 4 in 100 µL of ice-cold SPC Buffer and transfer to a new 8-tube 0.2 mL strip. Add 20 µL of the activated beads to each sample and gently pipette to mix. Incubate at room temperature for 10 min without agitation.

6.2. Place the tubes on the magnetic rack until the slurry is clear; remove and discard the supernatant using a pipette. Remove the tubes from the magnetic rack, and add 200 µL of ice-cold Wash Buffer to each sample. Resuspend the beads by gently pipetting up and down 5 times. Transfer 100 µL aliquots from each sample into a new 8-tube 0.2 mL strip.

CRITICAL STEP: Divide each CUT&RUN sample into two separate aliquots. Use one of the aliquots for the negative control antibody (e.g., IgG negative control antibody) and the other for the target antibody against the protein of interest (e.g., anti-GFP antibody).

NOTE: Both samples are required for the computational pipeline to accurately identify enrichment signals specific to the TF of interest. An additional control using anti-GFP antibodies with an untagged strain can also be performed. This control has shown results comparable to the use of IgG antibodies in a GFP-tagged strain. Therefore, for simplicity, it is recommended to use the standard IgG control for all experiments.

7. Primary antibody binding

NOTE: pAG-MNase fusion protein binds well to rabbit, goat, donkey, guinea pig, and mouse IgG antibodies[199]. Generally, most commercial ChIP-seq-certified commercial antibodies are compatible with CUT&RUN procedures. The amount of primary antibody used depends on the efficiency of the antibody, and titration of the antibody (e.g., 1:50, 1:100, 1:200, and 1:400 final dilution) may be necessary if the antibody of interest has not been previously tested in ChIP or CUT&RUN experiments. Chill all buffers on ice prior to use. All buffers used for antibody binding steps should be prepared fresh on the day of the experiment.

7.1. Place the tubes on a magnetic rack and wait until the slurry is completely clear; remove and discard the supernatant using a pipette. Add 50 µL of the Antibody Buffer and gently mix by pipetting.

7.2. Add 3 µL of the anti-GFP polyclonal antibody (or 0.5 µg if using an untested antibody). Incubate the tubes on a nutating mixer at 4 °C for 2 h.

NOTE: Some CUT&RUN protocols report increased yield by adding a secondary antibody prior to pAG-MNase addition[198]; however, no significant improvement was observed using this added step, and thus, it is not included in this protocol.

7.3. Briefly centrifuge the tubes at 100 × *g* at room temperature for 5 s, place the tubes on a magnetic rack, and once the slurry is clear, remove and discard the supernatant using a pipette. While the tubes containing the beads are still on the magnetic rack, add 200 µL of ice-cold Cell Permeabilization Buffer directly onto the beads. Remove and discard the supernatant using a pipette. Repeat for a total of two washes with the ice-cold Cell Permeabilization Buffer.

7.4. Add 50 µL of ice-cold Cell Permeabilization Buffer to each tube and gently mix by pipetting.

NOTE: Beads are often aggregated at this point but can easily be dispersed by mixing gently using a 200 µL pipette.

8. Binding of pAG-MNase to antibody

8.1. Add 2.5 µL of the pAG-Mnase (20x stock) to each sample and gently mix by pipetting. Place the samples (slightly elevated at ~45° angle) on a nutator at 4 °C. Turn on the nutator and incubate the samples for 1 h.

8.2. Briefly centrifuge the strip tubes at 100 × *g* at room temperature for 5 s, place the tubes on a magnetic rack and, once the slurry is clear, remove and discard the supernatant using a pipette.

NOTE: This step is critical. Carryover antibody remaining in the cap or sides of the tubes after this step will significantly increase the amount of background signal.

8.3. While the tubes containing the beads are still on the magnetic rack, add 200 µL of ice-cold Cell Permeabilization Buffer, allow the slurry to clear, and remove and discard the supernatant using a pipette. Repeat this step for a total of two washes with the Cell Permeabilization Buffer.

8.4. Add 100 µL of the ice-cold Cell Permeabilization Buffer to the samples and gently pipette up and down 5 times.

9. Targeted chromatin digestion and release

9.1. Incubate the tubes containing the sample(s) in a wet ice bath for 5 min. Add 3 µL of 100 mM $CaCl_2$ into each sample using a multichannel pipette. Gently pipette up and down 5 times, immediately return the tubes to the wet ice bath, and incubate for 30 min.

9.2. Add 66 µL of the Stop Buffer to each sample and gently vortex to mix. Incubate samples for 10 min at 37 °C in a dry bath.

NOTE: It is recommended to add 1.5 pg of heterologous E. coli spike-in DNA per sample in the Stop Buffer. The addition of 1.5 pg of E. coli spike-in DNA results in 1,000-10,000 mapped spike-in reads for 1–10 million mapped experimental reads[198]. The spike-in DNA is used to calibrate the sequencing depth and is especially important for comparing samples in a series. The addition of spike-in E. coli is highly recommended but not essential. The commercial CUT&RUN kit includes E. coli spike-in DNA, but it can also be purchased separately.

9.3. Place the tubes on the magnetic rack and transfer 160 µL of the supernatant into a 1.5 mL microfuge tube. Transfer 80 µL of the sample into a new 2 mL microfuge tube and store at -20 °C in the event that a backup sample is needed. Proceed to step 10 with the 80 µL sample.

10. **Cleanup of collected DNA samples**

10.1. Vortex DNA Purification Beads to homogenize the bead suspension. Add 50 µL (~0.6x sample volume) of the resuspended beads to each sample. Pipette-mix and incubate the samples on a nutator for 5 min at room temperature.

NOTE: The ratio of DNA purification beads to sample used is critical. Using 0.6x volume of DNA Purification Bead solution relative to the sample allows the magnetic beads to bind to large DNA fragments released from damaged nuclei. CUT&RUN-enriched DNA fragments are much smaller than these large DNA fragments and are thus retained in the supernatant at this step.

10.2. Place the tubes on a magnetic rack and transfer 130 µL of the supernatant containing the DNA to a 0.2 mL 8-tube strip. Add an additional 30 µL of DNA Purification Beads to the sample(s) (the total volume is 160 µL).

10.3. Add 170 µL (~1x sample volume) of ice-cold 100% isopropanol, mix well by pipetting up and down 10 times, and incubate on ice for 10 min.

NOTE: It is critical that 100% ice-cold isopropanol is used for this step for the DNA purification beads to efficiently capture the CUT&RUN-enriched small fragments.

10.4. Place the tubes on the magnetic rack, and once the slurry has cleared, carefully remove and discard the supernatant using a pipette.

10.5. While the tubes are on the magnetic rack, add 200 µL of freshly prepared, room-temperature 80% ethanol to the tubes and incubate at room temperature for 30 s. Carefully remove and discard the supernatant using a pipette. Repeat this step for a total of two washes with 80% ethanol.

10.6. Quickly spin the tubes at 100 × *g*, place the tubes back on the magnetic rack, and remove any residual ethanol using a pipette after the slurry has cleared. Air-dry the beads for 5 min while the tubes remain on the magnetic rack with the lid open.

NOTE: Do not exceed 5 min of drying time as this can significantly reduce the final DNA yield.

10.7. Remove the tubes from the magnetic rack and elute the DNA from the beads by adding 17 µL of 0.1x Tris-EDTA (TE) at pH 8. Mix well and incubate the tubes for 5 min at room temperature.

10.8. Place the tubes on the magnetic rack until the slurry becomes clear. Once the slurry has cleared, carefully transfer 15 µL of the supernatant to a sterile 0.2 mL PCR tube.

10.9. Measure the concentration of the collected DNA using a fluorometer following the manufacturer’s protocol.

NOTE: Typically, the concentration of the collected DNA is ~1 ng/µL. Sometimes, the concentration of the collected DNA is too low to quantify using a fluorometer. This is not an indicator of a failed experiment. Proceed with the library preparation regardless of the concentration of the collected DNA.

10.10. Proceed to step 11 or store the samples at -20 °C until ready.

11. Library preparation for sequencing

NOTE: The following steps use a commercially available library prep kit. When performing steps using the Ligation Master Mix, minimize touching the tubes and always keep them on ice.

11.1. Using 0.1x TE at pH 8, bring up the total volume of the CUT&RUN DNA to 50 µL. Make a master mix of 3 µL of End Prep Enzyme Mix and 7 µL of End Prep Reaction Buffer per sample. Add 10 µL of master mix to the CUT&RUN DNA and mix thoroughly by pipetting up and down 5 times.

11.2. Perform a quick spin at 100 × g to collect all liquid from the sides of the tube. Place the tubes in a thermocycler with the heated lid set to ≥75 °C and run the cycling conditions in **Table 3-7**.

NOTE: Depending on the starting input DNA concentrations collected from step 10, follow the required adapter dilution from **Table 3-8**.

11.3. Add 2.5 µL of Adapter per sample and mix thoroughly by pipetting up and down 10 times.

NOTE: It is critical that the adapter is added to the sample and mixed thoroughly before the ligation master mix is added.

11.4. Make a master mix of 30 µL of Ligation Master Mix and 1 µL of Ligation Enhancer. Add 31 µL of the master mix to the sample(s). Mix thoroughly by pipetting up and down 10 times.

11.5. Incubate at 20 °C for 15 min in a thermocycler with the heated lid off.

NOTE: It is critical that samples be kept on ice and transferred to the thermocycler only after the thermocycler has reached 20 °C.

11.6. Perform a quick spin at 100 × g to collect all liquid from the sides of the tube, add 3 µL of Uracil Excision Enzyme, and incubate the tubes in the thermocycler at 37 °C for 15 min with the heated lid set to ≥47 °C.

NOTE: This is a safe stopping point; store the samples at -20 °C or continue directly to step 11.7. If continuing directly to step 11.7, incubate the DNA Purification Beads at room temperature for 30 min before use.

11.7. Add 154.4 µL (~1.6x sample volume) of the DNA Purification Beads to the Adapter Ligation reaction from step 11.6. Pipette-mix and incubate the samples for 5 min at room temperature.

11.8. Place the tubes on the magnetic rack and once the slurry has cleared, carefully remove and discard the supernatant using a pipette.

11.9. Add 200 µL of freshly prepared, room-temperature 80% ethanol to the tubes and incubate at room temperature for 30 s. Carefully remove and discard the supernatant using a pipette and repeat this step for a total of two washes with 80% ethanol.

11.10. Spin the tubes briefly at 100 × *g*. Place the tubes back on the magnetic rack and remove any residual ethanol using a pipette. Air dry the beads for 5 min while the tubes remain on the magnetic rack with the lid open.

NOTE: Do not exceed 5 min of drying time as this can significantly reduce the final DNA yield.

11.11. Remove the tubes from the magnetic rack and elute the DNA from the beads by adding 17 µL of 0.1x TE at pH 8. Mix well and incubate for 5 min at room temperature.

11.12. Place the tubes on the magnetic rack until the slurry becomes clear. Once the slurry has cleared, carefully transfer 15 µL of the supernatant to a sterile 0.2 mL PCR tube.

11.13. Make a master mix of 25 µL of DNA Polymerase Master Mix and 5 µL of Universal Forward Library Amplification Primer (10 µM) per sample.

NOTE: Prepare one extra sample of master mix to account for pipetting losses.

11.14. Add 30 µL of the master mix to the 15 µL of Adapter-ligated DNA sample. Add 5 µL of Reverse Uniquely Indexed Library Amplification Primer (10 µM) to each sample to bring the final volume to a total of 50 µL. Mix thoroughly by pipetting up and down 10 times. Perform the PCR cycling conditions in **Table 3-9**.

NOTE: Incubate the DNA Purification Beads at room temperature for 30 min before use.

11.15. Vortex the DNA Purification Beads to resuspend. Add 35 µL (~0.7x sample volume) of the resuspended beads to the PCR-amplified DNA samples. Mix and incubate the samples on a nutator for 5 min at room temperature.

11.16. Place the tubes on the magnetic rack and once the slurry is clear, transfer the supernatant containing the DNA to a new 0.2 mL 8-well PCR strip tube.

11.17. Add 119 µL (~1.4x sample volume) of beads to the sample, and mix by pipetting up and down 5 times. Incubate the samples on a nutator for 5 min at room

temperature.

11.18. Place the tubes on the magnetic rack and once the slurry has cleared, carefully remove and discard the supernatant using a pipette.

11.19. Add 200 µL of freshly prepared, room-temperature 80% ethanol to the tubes and incubate at room temperature for 30 s. Carefully remove and discard the supernatant using a pipette; repeat the step for a total of two washes with 80% ethanol.

11.20. Spin the tubes briefly at 100 × *g*. Place the tubes back on the magnetic rack and remove any residual ethanol using a pipette. Air-dry the beads for 5 min while the tubes remain on the magnetic rack with the lid open.

NOTE: Do not exceed 5 min of drying time as this can significantly reduce the final DNA yield.

11.21. Remove the tubes from the magnetic rack and elute the DNA from the beads by adding 14 µL of 0.1x TE at pH 8. Mix well and incubate for 5 min at room temperature.

11.22. Place the tubes on the magnetic rack until the slurry becomes clear. Once the slurry has cleared, carefully transfer 13 µL of the supernatant to a sterile 0.2 mL PCR tube.

11.23. Prepare fresh 1x Tris-borate-EDTA (TBE) and insert premade, commercial 10% acrylamide TBE gel into the gel electrophoresis apparatus filled with 1x TBE.

11.24. In the first well, add 2 µL of Low-range DNA ladder. Mix 3 µL of 6x loading dye with 13 µL of the sample previously collected from step 11.22. Carefully add 15 µL into each well of the gel. Run the gel for 90 min at 70 V.

NOTE: It is recommended to leave one well in the gel empty between each sample, as this reduces the likelihood of sample cross contamination. Experienced users may find it appropriate to use all wells while carefully avoiding cross contamination, particularly when processing a large number of samples.

11.25. Remove the gel cast from the gel box. Open the gel cast per the manufacturer's instructions, gently remove the gel from the gel cast, and place it inside a gel holding tray containing 100 mL of 1x TBE.

NOTE: Make sure to gently remove the gel from the gel cast to avoid ripping the gel as it is thin and fragile. It is critical to prewet gloves and the gel with 1x TBE whenever handling the gel. The gel holding tray should be slightly larger than the size of the gel (approximately 0.5" on each side). The plastic lids provided with 96-well PCR-tube storage boxes are convenient holding trays for standard mini gel sizes.

11.26. Add 10 µL of the nucleic acid gel stain to the tray and gently swirl. Cover with

foil to protect from light and incubate statically at room temperature for 10 min.

11.27. Rinse the gel twice with 100 mL of deionized tap water. Image the gel under blue light illumination using an amber filter cover (**Figure 3-2**).

NOTE: Successful libraries show a smear between 100 and 500 bp. There will also be a prominent ~125 adapter dimer band. The presence of adapter dimers is not an indicator of poor library quality. This amount of adapter dimers is unavoidable for CUT&RUN experiments performed on low-abundance TFs and is a consequence of the low amount of input material used to prepare these libraries. Do not use ultraviolet light, which can damage the DNA.

11.28. As shown in **Figure 3-2**, for each library, cut the gel slightly above the ~125 bp prominent adapter dimer band (making sure to avoid touching the adapter dimer band) and below the 400 bp ladder mark.

NOTE: It is critical to avoid the ~125 adapter dimer band. Even tiny amounts of adapter dimers will significantly reduce the library quality.

11.29. Puncture the bottom of a 0.65 mL tube using a 22 G needle and place the punctured tube inside a sterile 2 mL microfuge tube. Transfer the gel slice to the punctured tube inside the 2 mL microfuge tube.

11.30. Centrifuge the 2 mL microfuge tube containing the 0.65 mL punctured tube and sample at 10,000 × *g* at room temperature for 2 min to collect the gel slurry inside the 2 mL microfuge tube.

NOTE: The punctured tube should now be empty and can be discarded. If the punctured tube still has any gel remaining inside, place the punctured tube back inside the 2 mL microfuge tube and centrifuge again at 10,000 × g at room temperature for an additional 2 min.

11.31. To the gel slurry inside the 2 mL microfuge tube, add 300 µL of ice-cold gel elution Buffer and mix on a nutator at room temperature for a minimum of 3 h or overnight (12–16 h).

11.32. Transfer all liquid and gel slurry to a 0.22 µm filter column. Centrifuge at 10,000 × *g* at room temperature for 1 min; the collected volume should be ~300 µL.

11.33. Add 450 µL (~1.5x sample volume) of DNA Purification Beads, incubate at room temperature for 5 min on a nutator, and then place the sample on the magnetic rack until the slurry is clear.

NOTE: Incubate the DNA Purification Beads at room temperature for 30 min before use.

11.34. Remove and discard 500 µL of the supernatant, making sure not to disrupt the beads.

11.35. Remove the sample from the magnetic rack and mix the beads by pipetting up and down 5 times. Transfer 200 µL of the sample into a new PCR strip tube.

11.36. Place the strip tube on the magnetic rack, and once the slurry has cleared, carefully remove and discard the supernatant using a pipette.

11.37. Add 200 µL of freshly prepared, room-temperature 80% ethanol to the tubes and incubate at room temperature for 30 s. Carefully remove and discard the supernatant using a pipette; repeat this step for a total of two washes with 80% ethanol.

11.38. Spin the tubes briefly at 100 × *g*. Place the tubes back on the magnetic rack and remove any residual ethanol using a pipette. Air-dry the beads for up to 5 min while the tubes remain on the magnetic rack with the lid open.

NOTE: Do not exceed 5 min of drying time as this can significantly reduce the final DNA yield.

11.39. Remove the tubes from the magnetic rack and elute the DNA from the beads by adding 17 µL of 0.1x TE at pH 8. Mix well and incubate the tubes for 5 min at room temperature.

11.40. Place the tubes on the magnetic rack until the slurry becomes clear. Once the slurry has cleared, carefully transfer 15 µL of the supernatant to a sterile 0.2 mL PCR tube.

11.41. Measure the final library quantity using the fluorometer; use this final library for sequencing.

NOTE: Up to 48 libraries can be pooled and sequenced together in a single lane using a sequencing platform that provides at least 300 million 40 bp or longer paired-end reads.

12. **CUT&RUN sequence analysis**

NOTE: This section presents the computational protocol used to analyze the CUT&RUN sequence data. The protocol begins with setting up the computational virtual environment and walks users through executing the commands on their local machine. This protocol will work on all computational resources, such as local machines, virtual cloud servers, and high-performance computing clusters. All CUT&RUN data presented in this paper can be accessed at NCBI GEO under accession number GSE193803.

12.1. Download the source code for the CUT&RUN analysis from https://github.com/akshayparopkari/cut_run_analysis.

NOTE: The workflow will work best on a MacOS or Linux OS system. Windows users can run the workflow using GitBash (see the **Table of Materials**).

12.2. Directly download the code from the GitHub page by clicking on the green **Code** button | **Download ZIP** option. Unzip the folder to a relevant location on the local machine.

NOTE: This workflow uses the Conda command line tool environment to install all required software and tools.

12.3. Install Conda environment (see the **Table of Materials**) and run only once.

12.4. Once Conda is installed (run only once), create a virtual environment using the **Supplementary File 3-2** provided with the following command:
conda create --name <env> --file Supplementary_File_2.txt

12.5. Activate the virtual environment every time this workflow is to be executed using:
conda activate <env>

12.6. Organize the input raw FASTQ files into a single folder, ideally one folder per CUT&RUN experiment.

13. **Generation of the genome file for alignment**
13.1. Generate the genome file for alignment (run only once for each genome file). Create a folder to save all *C. albicans* genome files, such as: *mkdir ca_genome_files* nnnn

14. **Downloading *C. albicans* genome assembly 21**
14.1. Download *C. albicans* genome assembly 21 from the *Candida* Genome Database using either wget or curl tools (see the **Table of Materials**)[207].

NOTE: *C. albicans* assembly 21 was used here to compare CUT&RUN results with previously published ChIP-chip results, which were aligned to Assembly 21. Users can download other assembly versions and run similar commands to generate relevant genome files for their alignment needs.

15. **Generate a Bowtie 2 index database (database name: ca21):**
15.1. Use the following:
bowtie2-build C_albicans_SC5314_A21_current_chromosomes.fasta.gz ca21
bowtie2-inspect -s ca21

16. **Run the CUT&RUN analysis pipeline.**
16.1. Read the help section to become familiar with the parameters of the pipeline.
bash cut_n_run_pipeline.sh -h

17. **Execute the cut_n_run_pipeline.sh file with relevant parameters**.
17.1. Execute the script:
bash cut_n_run_pipeline.sh /path/to/input/folder 4 y y y y y > /path/to/output.log 2>&1

NOTE: The relevant parameters with detailed descriptions on lines 19–36 of the code are described on the GitHub page https://github.com/akshayparopkari/cut_run_analysis/blob/main/cut_n_run_pipeline.sh.

18. **Organize output files.**
    18.1. Merge significant peaks from all replicates called by MACS2 located in /path/to/input/folder/peakcalling/macs2 using the BedTools merge function[208].
    cat /path/to/input/folder/peakcalling/macs2/all_replicate_files sort -k1,1 -k2,2n | mergeBed -c 4,5,6,7,8,9 -o last,mean,first,mean,mean,mean > /path/to/merged_output.bed

    NOTE: For additional information and best practices on assessing overlapping peaks in replicate samples, please refer to Landt et al.[209]. and Boyd et al.[210].

19. **Remove matches to blocklisted genomic regions using the BedTools subtract function.**
    19.1 Use the following: *subtractBed -a /path/to/merged_output.bed -b /path/to/Supplementary_File_3.bed -A > /path/to/merged_output_no_blocklist_hits.bed*

    NOTE: The blocklisted regions in the *C. albicans* genome are provided as a .bed file in **Supplementary File 3-3**. The list consists primarily of highly repetitive sequence elements and regions such as telomeric repeats and centromeres that commonly yield false-positive results in *C. albicans* CUT&RUN, ChIP-seq, and ChIP-chip datasets. Hence, it is recommended to remove the blocklisted regions. However, for certain protein targets, it may be inappropriate or undesirable to exclude these loci. Users can skip this step to retain signals contained within these blocklisted regions or create their own blocklisted regions.

20. **Merge BigWig files from replicates using the UCSC bigWigMerge function[211].**
    20.1. Use the following: *bigWigMerge /path/to/input/folder/bigwig/all_final_bw_files /path/to/input/folder/bigwig/ all_final_bdg_file*

    20.2. Convert the BedGraph output from bigWigMerge to BigWig using the UCSC bedGraphToBigWig function.
    bedGraphToBigWig /path/to/input/folder/bigwig/ all_final_bdg_file /path/to/input/folder/bigwig/ all_final_bw_file

3.7 List of Tables and Supplementary Figures

Table 3-1: PCR reaction mix and PCR cycling conditions to amplify universal A and unique B fragments.

	Universal A Fragment	Unique B Fragment
dH_2O	75.5 µL	75.5 µL
DNA Polymerase Buffer	20 µL	20 µL
10 mM dNTP	2 µL	2 µL
Forward Primer (100 µM)	0.5 µL (AHO1096)	0.5 µL unique gRNA
Reverse Primer (100 µM)	0.5 µL (AHO1098)	0.5 µL (AHO1097)
DNA (1 ng/µL)	1 µL pADH110	1 µL pADH139
DNA Polymerase	0.5 µL	0.5 µL

Temp (°C)	Time	Cycles
98	30 s	1
98	20 s	30
58	20 s	
72	30 s	
72	15 s	1
8	hold	1

Table 3-2: PCR reaction mix and PCR cycling conditions to stitch A and B fragments to create the full-length C fragment.

C Fragment PCR Reaction	
dH_2O	74.5 µL
DNA Polymerase Buffer	20 µL
10 mM dNTPs	2 µL
Universal A Fragment	1 µL
Unique B Fragment	1 µL
DNA Polymerase	0.5 µL

Cycling Condition 1		
Temp (°C)	Time	Cycles
98	30 s	1
98	10 s	5
58	20 s	
72	60 s	

Table 3-3: PCR cycling conditions to PCR amplify the full-length C fragment.

Cycling Condition 2		
Temp (°C)	Time	Cycles
98	30 s	1
98	10 s	30
66	20 s	
72	60 s	
72	30 s	1
8	hold	1

Table 3-4: PCR reaction mix and PCR cycling conditions to amplify the donor DNA GFP tag.

GFP Amplification PCR Reaction	
dH_2O	37.25 µL
DNA Polymerase Buffer	10 µL
10 mM dNTPs	1 µL
pCE1 (1 ng/µL)	1 µL
Forward Primer (100 µM)	0.25 µL
Reverse Primer (100 µM)	0.25 µL
DNA Polymerase	0.25 µL

Temp (°C)	Time	Cycles	
98	30 s	1	
98	10 s	10	Δ -1 °C / cycle
65	20 s		
72	30 s		
98	10 s	25	
57	20 s		
72	30 s		
72	15 s	1	
8	hold	1	

Table 3-5: Plasmid digestion reaction mix and reaction conditions for the plasmid digestion reactions.

Plasmid Digestion Reaction	
dH_2O	Variable
pADH140	Variable
Restriction Enzyme Buffer	1.5 µL
Restriction Enzyme	0.8 µL

Temp (°C)	Time
37	1 h
65	15 m
8	hold

Table 3-6: PCR reaction mix and PCR cycling conditions for colony PCR.

cPCR Reaction	
dH_2O	11.66 µL
cPCR DNA Polymerase Buffer	2.2 µL
5M Betaine	4.4 µL
$MgCl_2$	0.44 µL
10 mM dNTP	0.44 µL
cPCR DNA Polymerase	0.22 µL
Forward Primer (100 mM)	0.22 µL
Reverse Primer (100 mM)	0.22 µL
Lysate	2.2 µL

Temp (°C)	Time	Cycles
94	30 s	1
94	10 s	35
55	30 s	
72	1 min	
72	15 s	1
8	hold	1

Table 3-7: PCR cycling conditions for the end repair step of the library in preparation for sequencing.

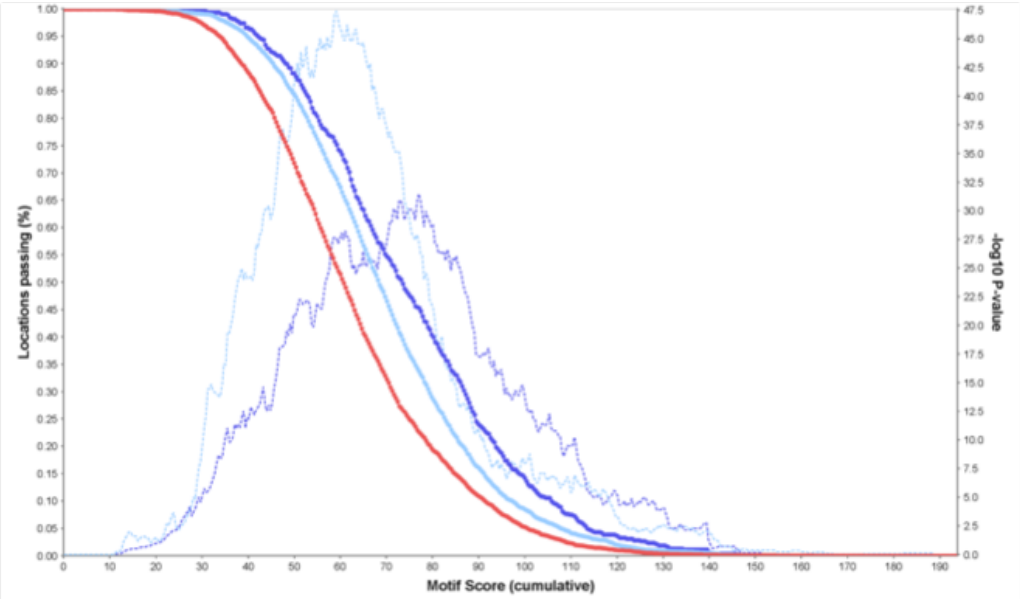
Temp (°C)	Time	Total Number of Cycles
20	30 min	1
50	60 min	1
4	Hold	1

Table 3-8: Adapter dilution recommendations for the input DNA to prepare sequencing libraries.

Input DNA	Adapter Dilution	Working Adapter Concentration
101 ng-1 µg	No dilution	15 µM
5-100 ng	10-fold	1.5 µM
< 5 ng	25-fold	0.6 µM

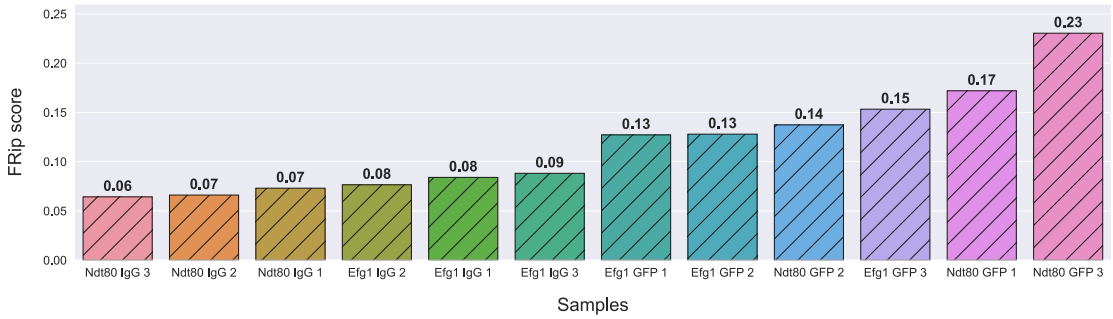
Table 3-9: PCR cycling conditions to PCR amplify the adapter-ligated DNA.

	Temp (°C)	Time	Total Number of Cycles
Initial Denaturation	98	45 s	1
Denaturation	98	15 s	14
Annealing/Extension	65	10 s	
Final Extension	65	5 min	1
Hold	4	Hold	1



Supplementary Figure 1: Ndt80 DNA binding sequence motif enrichment. Cumulative Ndt80 motif enrichment scores for Ndt80-bound loci identified in biofilm ChIP-chip data (dark blue dotted line) or CUT&RUN data (light blue dotted line) compared against random intergenic loci (red dotted line). The lefthand Y-axis indicates the percentage of total loci for each dataset that contains a corresponding cumulative motif

score on the X-axis. Dashed lines indicate the significance of the motif score enrichment (-log10 P-value, righthand Y-axis) at each point on the X-axis, relative to the random intergenic control loci. Cumulative motif scores and P-values were determined using MochiView[200] (see **Table of Materials**). All Ndt80-bound loci were sampled as 500 bp windows centered under each peak of Ndt80 binding within the ChIP-chip and CUT&RUN datasets and compared against randomly selected 500 bp intergenic loci to control for differences in length. Abbreviations: ChIP = chromatin immunoprecipitation; ChIP-chip = chromatin immunoprecipitation followed by microarray; CUT&RUN = cleavage under targets and release using nuclease.



Supplementary Figure 2: Fraction of reads within a peak score per sample. Bar plot depicting the FRiP score for each replicate. FRiP scores are calculated as the number of mapped reads within a peak divided by the total number of mapped reads in a CUT&RUN sample. The different bar colors represent individual biological replicate samples, as indicated along the X-axis. As expected, FRiP scores of positive GFP samples are consistently higher than those for negative IgG samples. All samples show acceptable FRiP thresholds (>1%), as per the recommendations set by Landt et al.[209]. Abbreviations: FRiP = fraction of reads within a peak; CUT&RUN = cleavage under targets and release using nuclease.

Supplementary File 1: Recipes for all buffers and solutions used.

Supplementary File 2: Bioinformatic tools required for the data analysis pipeline.

Supplementary File 3: Recommended blocklisted regions in the *C. albicans* genome. This list of blocklisted loci primarily contains highly repetitive sequence elements and regions such as telomeric repeats and centromeres that have historically yielded false-positive results in previous genome-wide binding experiments.

Chapter 4: Discussion, new hypotheses, and future directions

In this thesis we have examined how the interplay between transcription factors (TFs) and chromatin structure regulates developmental processes and heritable cell fate decisions in *Candida albicans*—a diploid polymorphic fungus. Our results show, for the first time, that pioneer-like TFs modulate the chromatin structure to regulate important biological processes in *C. albicans*. In this chapter we provide a brief overview of the white-opaque switch, summarize several novel findings regarding the interplay between TFs and the chromatin structure in *C. albicans*, and discuss new hypotheses generated considering these findings. Together these findings indicate that the white-opaque switch represents an attractive and relatively "simple" model system to investigate how the interplay between TFs and chromatin structure regulates heritable cell fate decisions in higher eukaryotes.

Higher eukaryotes are generally comprised of numerous cells and many different cell types, the vast majority of which contain the same underlying genomic DNA sequences. These distinct cell types are established and maintained through transcriptional programs which specify specific patterns of gene expression for a particular cell type. To maintain cellular identity, these transcriptional programs must be re-established each time a cell divides, and chromatin structure plays an important role in this process[92]. *C albicans* differentiates into two distinct cell types known as white and opaque. These two cell types are established and heritably maintained by distinct transcriptional programs, leading to differences in metabolic preferences, mating abilities, cellular morphologies, responses to environmental signals, interactions with the host immune system and differential expression of ~20% of the transcriptome[1–4,10–15,17,18,144,146–148,212–215]. Twenty sequence-specific TFs are known to effect establishment and heritability of the two cell types. Eight of these TFs have been extensively characterized, leading to identification of the core transcriptional circuits of each cell[1,2,4,9,21,23]. Remarkably, the level of connectivity of the opaque cell transcriptional circuit closely resembles the circuits that control stem cell maintenance and differentiation in humans[3,26–28]. Recent research revealed that chromatin structure modulating enzymes also play an important role in the differentiation and maintenance of white and opaque cells[1,5,22,29,33,34,119,154]. However, little is known about the role of chromatin organization, or reorganization, in the process of establishing and maintaining the white and opaque cell types (details and references in Chapter 1).

The white-opaque switch is observed in *C. albicans* and several closely related species[216,217] but is absent in other model yeasts like *Saccharomyces cerevisiae* and *Schizosacharomyces pombe*. Therefore, *C. albicans* represents a unique model system to examine the interplay between TFs and chromatin structure in a simple unicellular eukaryote. Importantly, given the close resemblance between the transcriptional circuit in opaque cells and those that control developmental programs in higher eukaryotes, we believe that findings from this model may have broad implications in understanding the establishment and maintenance of cell identity in higher eukaryotes.

Nucleosomes are a fundamental unit of chromatin. The interplay between nucleosomes and TFs regulates chromatin organization and reorganization events that play important roles in the establishment and maintenance of cell identity[125,137,218–224]. In this work, we first showed that the chromatin structure in opaque cells is significantly altered relative to white cells. Next, we showed that the opaque cell "master regulator" Wor1 establishes opaque-specific chromatin structures, some of which are required for heritable maintenance of the

opaque transcriptional program. Additionally, we discovered that Wor2, which is co-localized at 52% of Wor1-bound sites[4], establishes opaque-specific chromatin structures that are required for heritable maintenance of the opaque cell type. Furthermore, we discovered that Ndt80, a TF that is highly conserved across a large group of fungal species[173,174,225], establishes highly accessible chromatin structures throughout the *C. albicans* genome in multiple cell types. Ndt80 is also required for *C. albicans* ability to form biofilms[179]—an organized community of cells embedded in an extracellular polymer matrix, in which the cells have acquired properties resistant to drugs and physical perturbations[179,189,191,226]. Here, we discovered that the interplay between Ndt80 and chromatin structure is critical for biofilm development as well as white-opaque switching (Chapter 2), indicating that Ndt80 is likely a general chromatin organizing factor in *C.* albicans—i.e., a TF that regulates changes in chromatin structure required for most of the cellular processes.

The model I propose here is that the interplay between TFs and nucleosomes regulates the transitions between "nucleosome free" and "nucleosome depleted" regions (NFRs and NDRs), and that this process is central to controlling chromatin organization and gene regulation. Here, I will summarize the similarities and differences between NFR and NDR chromatin structures, describe the proposed "**see-saw NFR-NDR**" model (**Fig. 4-1**), and lay out the evidence generated in the course of this thesis that supports this model. NFR chromatin structure is characterized by regions of "free" DNA less than 150 bp (i.e., the minimum length of DNA required to form a nucleosome) that are thought to be genuinely free of nucleosomes because they show no evidence of protection from MNase even in limited MNase digests[185]. Another distinguishing feature of NFRs is that the flanking nucleosomes are not constrained by other DNA-bound proteins, and thus these nucleosomes are poorly positioned and appear "fuzzy" or randomly distributed within a population of cells[164,167,185]. In contrast, NDR chromatin structure is characterized by a region of "free" DNA more than 150 bp. Nucleosomes flanking NDRs are thought to be constrained by "barrier" complexes which impose order on surrounding nucleosomes, leading to the appearance of sharp, well-separated nucleosome occupancy peaks in MNase-seq data[125,164,167,185].

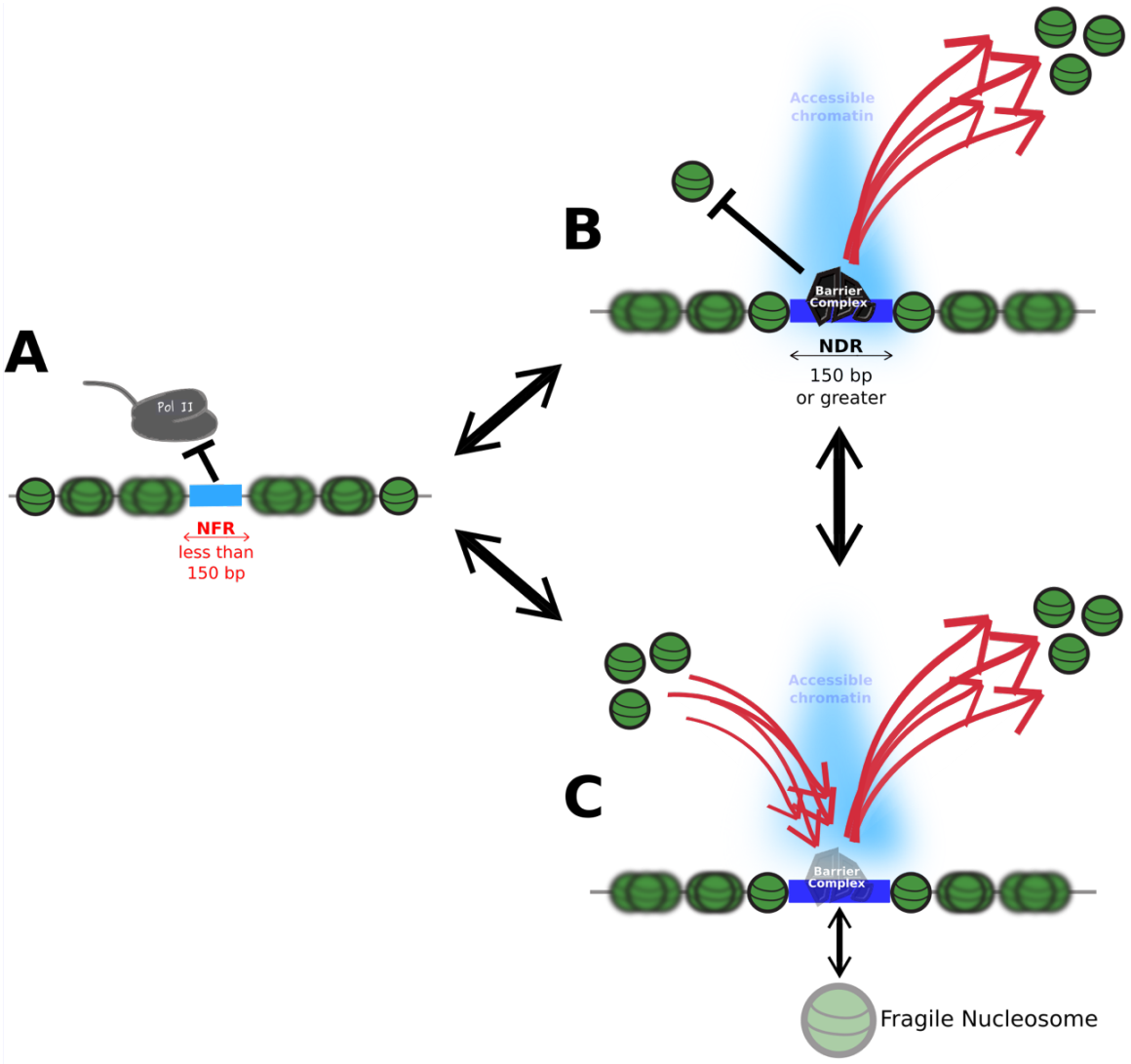


Figure 4-1. The "see-saw NFR-NDR" chromatin model.

(A) NFRs are generally in an inaccessible chromatin state. (B) Chromatin organizing TFs bind within NDRs to form a stable barrier complex and establish the highly accessible and organized chromatin structure. (C) In the absence of some critical TFs, the barrier complex doesn't stably bind to DNA (translucent barrier complex) within the NDRs, and thus nucleosomes are not evicted from NDRs as efficiently.

The transition from NFR to NDR chromatin structure is an active process that is driven by the stable binding of a barrier complex to NDRs[125,164,170,184,185]. Multiple TFs and general regulatory factors are thought to comprise the barrier complex; however, the exact composition is still actively debated[120,125,164,165,227,228]. Consistent with this, we find that NDRs are significantly enriched for short reads in our MNase-seq assay, indicating that they are likely bound by small TF-sized proteins (see results in Chapter 2). Additionally, our results showed that some NDR-bound TFs are likely critical for the stable binding of

barrier complexes within the NDRs, and their presence is correlated with the stable eviction of nucleosomes from these NDRs. In our model, we speculate that in the absence of TFs that stabilize the barrier-complex, nucleosomes are not stably evicted and consequently, these sites teeter between nucleosome-bound and barrier-complex-bound states. This teetering between the two states is evidenced in the MNase-seq data as nucleosomes that are "fragile" (see Fig.2-3 in Chapter 2) because of their extreme sensitivity to the extent of MNase digestion likely resulting from their dynamic competition with the barrier complexes for binding to the DNA. However, further work is required to test this model. To summarize, I propose a model (described in Fig. 4-2) whereby the interplay between TFs and nucleosomes regulates the transition from NFR to NDR chromatin structure that then teeters between a fragile-nucleosome-bound state or barrier-complex-bound state. This transition from NFR to NDR chromatin structure is likely initiated by chromatin organizing TFs with a specialized ability to target many NFRs—compact chromatin structures that are inaccessible in our ATAC-seq assay—and effect the expression of a large number of genes. Since this specialized ability to organize the chromatin is observed throughout the genome, these TFs are often recognized to serve as "master regulators" that are critical for developmental processes and heritable cell fate decisions.

In this body of work, we discovered several chromatin organizing TFs and showed evidence in support of their specialized ability to promote the transition from NFR to NDR chromatin structures. One TF that possesses this specialized ability to organize the chromatin structure is the opaque "master regulator" Wor1. Our results showed that hundreds of Wor1-occupied motif sites exhibit NFR chromatin structures in white cells, which subsequently are transformed into the NDR chromatin structures in opaque cells when these sites are bound by Wor1. The NDR chromatin structure observed in opaque cells is characterized by high chromatin accessibility and the presence of a barrier complex that is flanked by well-positioned nucleosomes. Additionally, we also showed that Wor2, a TF co-enriched at Wor1-bound sites (52%), is required for stable eviction of nucleosomes from Wor2-bound sites, indicating that Wor2 increases stability of the barrier complexes bound to opaque-specific NDRs that are subsequently propagated for hundreds of generations to facilitate heritability in opaque cells. In the absence of Wor2, the barrier complexes appear to form, but are unable to stably maintain the NDR chromatin structure associated with heritable opaque cells and thus the cells fail to propagate the opaque transcriptional program from one generation to the next. The transition that we observe from NFR to NDR chromatin structure at Wor1-bound sites during white to opaque switching—which is also associated with changes in gene expression critical for the switch—clearly illustrates that the interplay between TFs and nucleosomes plays an important role in regulating heritable cell fate decisions in *C. albicans*.

4.1 Future Directions

Thus far, we have shown that chromatin organizing TFs establish NDR chromatin structures which play important roles in regulating the white-opaque switch and biofilm formation in *C. albicans.* The TFs observed to possess this specialized ability to organize the chromatin structure include Wor1, Wor2 and Ndt80—TFs with important roles in the white-opaque switch and/or biofilm formation. There are several other TFs known to be critical for the white-opaque switch[9] or biofilm formation[229] and one future direction worth exploring is to examine the impact of these TFs on the chromatin structure. Some of my preliminary work shows that Ssn6, a TF critical for the white-opaque switch[23], also has an

impact on the chromatin structure of opaque cells, though the impact is observed to be less extensive compared to Wor1, Wor2 and Ndt80. Among the white-opaque regulators, Rap1 has already been shown to have a significant impact on the chromatin structure in *Saccharomyces cerevisiae*[133,230] and it will be interesting to assess whether Rap1 possesses similar chromatin organizing activity in the context of *C. albicans* white-opaque switching as well. Additionally, examining the chromatin organizing roles for TFs such as Efg1, Flo8, and Gal4—TFs critical for both the white-opaque switch and biofilm formation[9,229,231]—would likely elucidate whether the chromatin organizing activities of these TFs are limited to specific developmental process. Given that NDRs impose order on surrounding nucleosomes, another future direction worth exploring would be to examine the influence of two proximal NDRs on the overall chromatin organization in the proximal intergenic region. In Chapter 2, we showed that Wor1-occupied motif sites are wide and highly accessible NDRs and that Wor1 and Wor2 are critical for establishing these NDRs. Interestingly, there are multiple occurrences of Wor1-dependent NDRs (i.e., NDRs overlapping Wor1-occupied motif sites) proximal to one another in the intergenic region upstream of *WOR1*. Systematic deletion of these Wor1-occupied motif sites, individually and in combinations, followed by an examination of the resulting impact on the local nucleosome occupancy using MNase-qPCR could shed some light on whether the combined influence of two NDRs in close proximity leads to increased or decreased overall nucleosome organization.

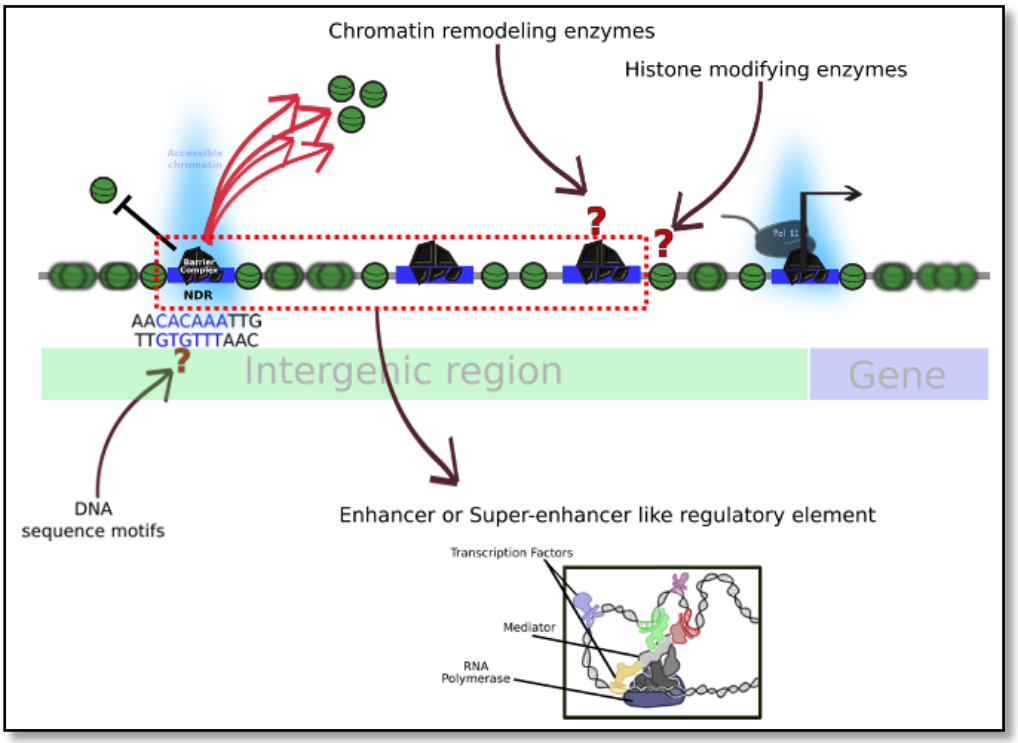


Figure 4-2. Future directions.

The proposed “see-saw NFR-NDR” model suggests that TFs are directly responsible for the changes in the chromatin structure observed at Wor1, Wor2, and Ndt80-bound loci. Alternatively, TFs could be indirectly modulating the chromatin structure by recruiting chromatin remodeling or histone modifying enzymes. Therefore, as shown in **Figure 4-2**, another future direction worth exploring is to examine whether Wor1, Wor2 or Ndt80 recruit any chromatin remodeling or histone modifying enzymes. A mass spectroscopy-based approach could help identify proteins associated with Wor1, Wor2, and Ndt80.

Lastly, the chromatin features observed at NDRs, like high chromatin accessibility and TF occupancy, are highly similar to those observed at enhancers in higher eukaryotic organisms. Among enhancers, some are enriched for TFs, chromatin remodelers or mediator-complex subunits significantly more than the average enhancer and consequently these are classified as “super-enhancers” (**Figure 4-2**). Super-enhancers have been observed to phase-separate and it has been shown that protein-protein interactions between the TFs that contain intrinsically disordered low-complexity domains plays critical role in this process. Recently, work by the Bennett Lab showed that white-opaque regulating TFs also undergo phase separation in vivo and in vitro, and that the ability of Wor1 to undergo phase separation is critical to the establishment of the opaque cell type. The Bennett Lab study, along with our results (see Chapter 2), suggest that a cluster of proximal NDRs, which are bound by many TFs that are shown to have the ability to form condensates, are likely to form super-enhancer like regulatory elements to regulate gene expression. It would be worth it to explore whether any specific TFs bound within these proximal NDRs have a special role in condensate formation and if these proximal NDRs influence one another and regulate gene expression in a combinatorial manner.

I hope that the work presented in this thesis will serve as a foundation for continued use of the white-opaque switch as a model system to better understand how the interplay between chromatin organizing TFs and chromatin structure regulates developmental processes and heritable cell fate decisions in higher eukaryotes.

References

1. Zordan, R. E., Miller, M. G., Galgoczy, D. J., Tuch, B. B. & Johnson, A. D. Interlocking transcriptional feedback loops control white-opaque switching in Candida albicans. *PLoS Biol* **5**, 2166–2176 (2007).

2. Zordan, R. E., Galgoczy, D. J. & Johnson, A. D. Epigenetic properties of white-opaque switching in Candida albicans are based on a self-sustaining transcriptional feedback loop. *Proceedings of the National Academy of Sciences* **103**, 12807–12812 (2006).

3. Tuch, B. B. *et al.* The transcriptomes of two heritable cell types illuminate the circuit governing their differentiation. *PLoS Genet* **6**, (2010).

4. Hernday, A. D. *et al.* Structure of the transcriptional network controlling white-opaque switching in Candida albicans. *Mol Microbiol* **90**, 22–35 (2013).

5. Guan, Z. & Liu, H. Overlapping Functions between SWR1 Deletion and H3K56 Acetylation in Candida albicans. *Eukaryot Cell* **14**, 578–587 (2015).

6. Anderson, M. Z. *et al.* A Multistate Toggle Switch Defines Fungal Cell Fates and Is Regulated by Synergistic Genetic Cues. *PLoS Genet* **12**, 1–28 (2016).

7. Frazer, C. *et al.* Epigenetic cell fate in Candida albicans is controlled by transcription factor condensates acting at super-enhancer-like elements. *Nat Microbiol* **5**, 1374–1389 (2020).

8. Takagi, J., Singh-Babak, S. D., Lohse, M. B., Dalal, C. K. & Johnson, A. D. Candida albicans white and opaque cells exhibit distinct spectra of organ colonization in mouse models of infection. *PLoS One* **14**, e0218037 (2019).

9. Lohse, M. B. *et al.* Systematic genetic screen for transcriptional regulators of the Candida albicans white-opaque switch. *Genetics* **203**, 1679–1692 (2016).

10. Solis, N. V. *et al.* Candida albicans whiteopaque switching influences virulence but not mating during oropharyngeal candidiasis. *Infect Immun* **86**, 1–14 (2018).

11. Craik, V. B., Johnson, A. D. & Lohse, M. B. Sensitivity of white and opaque Candida albicans cells to antifungal drugs. *Antimicrob Agents Chemother* **61**, 8–11 (2017).

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

13. Lohse, M. B. & Johnson, A. D. Differential phagocytosis of white versus opaque Candida albicans by Drosophila and mouse phagocytes. *PLoS One* **3**, (2008).

14. Miller, M. G. & Johnson, A. D. White-opaque switching in Candida albicans is controlled by mating-type locus homeodomain proteins and allows efficient mating. *Cell* **110**, 293–302 (2002).

15. Sasse, C., Hasenberg, M., Weyler, M., Gunzer, M. & Morschhäuser, J. White-opaque switching of Candida albicans allows immune evasion in an environment-dependent fashion. *Eukaryot Cell* **12**, 50–58 (2013).

16. Alby, K. & Bennett, R. J. To switch or not to switch?: Phenotypic switching is sensitive to multiple inputs in a pathogenic fungus. *Commun Integr Biol* **2**, 509–511 (2009).

17. Huang, G., Srikantha, T., Sahni, N., Yi, S. & Soll, D. R. CO2 Regulates White-to-Opaque Switching in Candida albicans. *Current Biology* **19**, 330–334 (2009).

18. Huang, G. *et al.* N-acetylglucosamine induces white to opaque switching, a mating prerequisite in Candida albicans. *PLoS Pathog* **6**, (2010).

19. Morrow, B., Anderson, J., Wilson, J. & Soll, D. R. Bidirectional stimulation of the white-opaque transition of Candida albicans by ultraviolet irradiation. *J Gen Microbiol* **135**, 1201–1208 (1989).

20. Lohse, M. B. *et al.* Identification and characterization of a previously undescribed family of sequence-specific DNA-binding domains. *Proceedings of the National Academy of Sciences* **110**, 7660–7665 (2013).

21. Lohse, M. B. & Johnson, A. D. Identification and characterization of Wor4, a new transcriptional regulator of white-opaque switching. *G3: Genes, Genomes, Genetics* **6**, 721–729 (2016).

22. Stevenson, J. S. & Liu, H. Regulation of white and opaque cell-type formation in Candida albicans by Rtt109 and Hst3. *Mol Microbiol* **81**, 1078–1091 (2011).

23. Hernday, A. D. *et al.* Ssn6 defines a new level of regulation of white-opaque switching in Candida albicans and is required for the stochasticity of the switch. *mBio* **7**, 1–9 (2016).

24. Alkafeef, S. S., Yu, C., Huang, L. & Liu, H. Wor1 establishes opaque cell fate through inhibition of the general co-repressor Tup1 in Candida albicans. *PLoS Genet* **14**, 1–25 (2018).

25. Srikantha, T. *et al.* TOS9 regulates white-opaque switching in Candida albicans. *Eukaryot Cell* **5**, 1674–1687 (2006).

26. Neph, S. *et al.* Circuitry and dynamics of human transcription factor regulatory networks. *Cell* **150**, 1274–1286 (2012).

27. Loh, Y.-H. *et al.* The Oct4 and Nanog transcription network regulates pluripotency in mouse embryonic stem cells. *Nat Genet* **38**, 431–440 (2006).

28. Chen, X. *et al.* Integration of External Signaling Pathways with the Core Transcriptional Network in Embryonic Stem Cells. *Cell* **133**, 1106–1117 (2008).

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

30. Xie, J., Jenull, S., Tscherner, M. & Kuchler, K. Roles in Regulating the White-Opaque Switch in the Fungal Pathogen Candida albicans. *mBio* **7**, e01807-16 (2016).

31. Srikantha, T., Tsai, L., Daniels, K., Klar, A. J. S. & Soll, D. R. The histone deacetylase genes HDA1 and RPD3 play distinct roles in regulation of high-frequency phenotypic switching in Candida albicans. *J Bacteriol* **183**, 4614–4625 (2001).

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

33. Stevenson, J. S. & Liu, H. Nucleosome Assembly Factors CAF-1 and HIR Modulate Epigenetic Switching Frequencies in an H3K56 Acetylation-Associated Manner in Candida albicans. *Eukaryot Cell* **12**, 591–603 (2013).

34. Tscherner, M., Stappler, E., Hnisz, D. & Kuchler, K. The histone acetyltransferase Hat1 facilitates DNA damage repair and morphogenesis in Candida albicans. *Mol Microbiol* **86**, 1197–1214 (2012).

35. Klemm, S. L., Shipony, Z. & Greenleaf, W. J. Chromatin accessibility and the regulatory epigenome. *Nat Rev Genet* **20**, 207–220 (2019).

36. Williamson, W. D. & Pinto, I. Histones and genome integrity. *Frontiers in Bioscience* **17**, 984–995 (2012).

37. Lawrence, M., Daujat, S. & Schneider, R. Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends in Genetics* **32**, 42–56 (2016).

38. Tessarz, P. & Kouzarides, T. Histone core modifications regulating nucleosome structure and dynamics. *Nat Rev Mol Cell Biol* **15**, 703–708 (2014).

39. Badeaux, A. I. & Shi, Y. Emerging roles for chromatin as a signal integration and storage platform. *Nat Rev Mol Cell Biol* **14**, 211–224 (2013).

40. Rando, O. J. & Winston, F. Chromatin and transcription in yeast. *Genetics* **190**, 351–387 (2012).

41. Kurdistani, S. K. & Grunstein, M. Histone acetylation and deacetylation in yeast. *Nat Rev Mol Cell Biol* **4**, 276–284 (2003).

42. Jaiswal, D., Turniansky, R. & Green, E. M. Choose Your Own Adventure: The Role of Histone Modifications in Yeast Cell Fate. *J Mol Biol* **429**, 1946–1957 (2017).

43. Sheikh, B. N. & Akhtar, A. The many lives of KATs — detectors, integrators and modulators of the cellular environment. *Nat Rev Genet* **20**, 7–23 (2019).

44. Marmorstein, R. & Trievel, R. C. Histone modifying enzymes: Structures, mechanisms, and specificities. *Biochim Biophys Acta Gene Regul Mech* **1789**, 58–68 (2009).

45. Marmorstein, R. & Zhou, M. M. Writers and readers of histone acetylation: Structure, mechanism, and inhibition. *Cold Spring Harb Perspect Biol* **6**, (2014).

46. Gurard-Levin, Z. A., Quivy, J.-P. & Almouzni, G. Histone Chaperones: Assisting Histone Traffic and Nucleosome Dynamics. *Annu Rev Biochem* **83**, 487–517 (2014).

47. Amin, A. D., Vishnoi, N. & Prochasson, P. A global requirement for the HIR complex in the assembly of chromatin. *Biochim Biophys Acta Gene Regul Mech* **1819**, 264–276 (2012).

48. Clapier, C. R., Iwasa, J., Cairns, B. R. & Peterson, C. L. Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* **18**, 407–422 (2017).

49. Venkatesh, S. & Workman, J. L. Histone exchange, chromatin structure and the regulation of transcription. *Nat Rev Mol Cell Biol* **16**, 178–189 (2015).

50. Pérez-Martín, J., Uría, J. A. & Johnson, A. D. Phenotypic switching in Candida albicans is controlled by a SIR2 gene. *EMBO J* **18**, 2580–92 (1999).

51. Stevenson, J., Krycer, J. R., Phan, L. & Brown, A. J. A practical comparison of ligation-independent cloning techniques. *PLoS One* **8**, e83888 (2013).

52. Peterson, M. R. *et al.* The fungal-specific Hda2 and Hda3 proteins regulate morphological switches in the human fungal pathogen Candida albicans. *bioRxiv* 340364 (2018) doi:10.1101/340364.

53. Kročová, E. *et al.* PHO15 genes of Candida albicans and Candida parapsilosis encode HAD-Type phosphatases dephosphorylating 2-phosphoglycolate. *FEMS Yeast Res* **19**, (2019).

54. Roth, S. Y., Denu, J. M. & Allis, C. D. Histone acetyltransferases. *Annu Rev Biochem* **70**, 81–120 (2001).

55. Sophie Mokas, J. R. M. *et al.* Uncoupling Stress Granule Assembly and Translation Initiation Inhibition. *Mol Biol Cell* **20**, 2673–2683 (2009).

56. Mizuguchi, G. *et al.* ATP-Driven Exchange of Histone H2AZ Variant Catalyzed by SWR1 Chromatin Remodeling Complex. *Science (1979)* **303**, 343–348 (2004).

57. Wang, X. *et al.* Merge and separation of NuA4 and SWR1 complexes control cell fate plasticity in Candida albicans. *Cell Discov* **4**, (2018).

58. Masumoto, H., Hawke, D., Kobayashi, R. & Verreault, A. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. *Nature* **436**, 294–298 (2005).

59. Xu, F., Zhang, K. & Grunstein, M. Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell* **121**, 375–385 (2005).

60. Voichek, Y. *et al.* Epigenetic Control of Expression Homeostasis during Replication Is Stabilized by the Replication Checkpoint. *Mol Cell* **70**, 1121-1133.e9 (2018).

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

62. Värv, S. *et al.* Acetylation of H3 K56 Is Required for RNA Polymerase II Transcript Elongation through Heterochromatin in Yeast. *Mol Cell Biol* **30**, 1467–1477 (2010).

63. Hu, G. *et al.* H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. *Cell Stem Cell* **12**, 180–192 (2013).

64. Haigney, A., Ricketts, M. D. & Marmorstein, R. Dissecting the molecular roles of histone chaperones in histone acetylation by type B histone acetyltransferases (HAT-B). *Journal of Biological Chemistry* **290**, 30648–30657 (2015).

65. Song, O. K., Wang, X., Waterborg, J. H. & Sternglanz, R. An Nα-acetyltransferase responsible for acetylation of the N-terminal residues of histones H4 and H2A. *Journal of Biological Chemistry* **278**, 38109–38112 (2003).

66. Ree, R., Varland, S. & Arnesen, T. Spotlight on protein N-terminal acetylation. *Exp Mol Med* **50**, 1–13 (2018).

67. Polevoda, B., Hoskins, J. & Sherman, F. Properties of Nat4, an Nα-Acetyltransferase of Saccharomyces cerevisiae That Modifies N Termini of Histones H2A and H4. *Mol Cell Biol* **29**, 2913–2924 (2009).

68. Freitag, M. Histone Methylation by SET Domain Proteins in Fungi. *Annu Rev Microbiol* **71**, 413–439 (2017).

69. Raman, S. B. *et al.* Candida albicans SET1 encodes a histone 3 lysine 4 methyltransferase that contributes to the pathogenesis of invasive candidiasis. *Mol Microbiol* **60**, 697–709 (2006).

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

71. Downs, J. A. *et al.* Binding of chromatin-modifying activities to phosphorylated histone H2A at DNA damage sites. *Mol Cell* **16**, 979–990 (2004).

72. Hnisz, D. *et al.* A Histone Deacetylase Adjusts Transcription Kinetics at Coding Sequences during Candida albicans Morphogenesis. *PLoS Genet* **8**, e1003118 (2012).

73. Hnisz, D., Majer, O., Frohner, I. E., Komnenovic, V. & Kuchler, K. The set3/Hos2 histone deacetylase complex attenuates camp/pka signaling to regulate morphogenesis and virulence of candida albicans. *PLoS Pathog* **6**, 1–18 (2010).

74. Kornitzer, D. Regulation of Candida albicans Hyphal Morphogenesis by Endogenous Signals. *Journal of Fungi* **5**, 21 (2019).

75. Bockmühl, D. P., Krishnamurthy, S., Gerads, M., Sonneborn, A. & Ernst, J. F. Distinct and redundant roles of the two protein kinase A isoforms Tpk1p and Tpk2p in morphogenesis and growth of Candida albicans. *Mol Microbiol* **42**, 1243–1257 (2001).

76. Bockmüh, D. P. & Ernst, J. F. A potential phosphorylation site for an A-Type kinase in the Efgl regulator protein contributes to hyphal morphogenesis of Candida albicans. *Genetics* **157**, 1523–1530 (2001).

77. Pim Pijnappel, W. W. M. *et al.* The S. cerevisiae SET3 complex includes two histone deacetylases, Hos2 and Hst1, and is a meiotic-specific repressor of the sporulation gene program. *Genes Dev* **15**, 2991–3004 (2001).

78. Xie, J. *et al.* Sum1 and Hst1 repress middle sporulation-specific gene expression during mitosis in Saccharomyces cerevisiae. *EMBO Journal* **18**, 6448–6454 (1999).

79. Orta-Zavalza, E. *et al.* Local silencing controls the oxidative stress response and the multidrug resistance in Candida glabrata. *Mol Microbiol* **88**, 1135–1148 (2013).

80. Lu, Y., Su, C. & Liu, H. A GATA transcription factor recruits Hda1 in response to reduced Tor1 signaling to establish a hyphal chromatin state in Candida albicans. *PLoS Pathog* **8**, e1002663 (2012).

81. Lu, Y., Su, C., Wang, A. & Liu, H. Hyphal Development in Candida albicans Requires Two Temporally Linked Changes in Promoter Chromatin for Initiation and Maintenance. *PLoS Biol* **9**, e1001105 (2011).

82. Sanchez, R. & Zhou, M.-M. The role of human bromodomains in chromatin biology and gene transcription. *Curr Opin Drug Discov Devel* **12**, 659–65 (2009).

83. Dhalluin, C. *et al.* Structure and ligand of a histone acetyltransferase bromodomain. *Nature* **399**, 491–496 (1999).

84. Simithy, J. *et al.* Characterization of histone acylations links chromatin modifications with metabolism. *Nat Commun* **8**, 1–13 (2017).

85. Wang, Q. *et al.* The YEATS Domain Histone Crotonylation Readers Control Virulence-Related Biology of a Major Human Pathogen. *Cell Rep* **31**, 107528 (2020).

86. Mietton, F. *et al.* Selective BET bromodomain inhibition as an antifungal therapeutic strategy. *Nat Commun* **8**, (2017).

87. Yap, K. L. & Zhou, M.-M. Keeping it in the family: diverse histone recognition by conserved structural folds. *Crit Rev Biochem Mol Biol* **45**, 488–505 (2010).

88. Musselman, C. A., Lalonde, M.-E., Côté, J. & Kutateladze, T. G. Perceiving the epigenetic landscape through histone readers. *Nat Struct Mol Biol* **19**, 1218–1227 (2012).

89. Taverna, S. D., Li, H., Ruthenburg, A. J., Allis, C. D. & Patel, D. J. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nat Struct Mol Biol* **14**, 1025–1040 (2007).

90. Andrews, F. H., Strahl, B. D. & Kutateladze, T. G. Insights into newly discovered marks and readers of epigenetic information. *Nat Chem Biol* **12**, 662–668 (2016).

91. Buschbeck, M. & Hake, S. B. Variants of core histones and their roles in cell fate decisions, development and cancer. *Nat Rev Mol Cell Biol* **18**, 299–314 (2017).

92. Xu, Q. & Xie, W. Epigenome in Early Mammalian Development: Inheritance, Reprogramming and Establishment. *Trends Cell Biol* **28**, 237–253 (2018).

93. Alabert, C. & Groth, A. Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* **13**, 153–167 (2012).

94. Serra-Cardona, A. & Zhang, Z. Replication-Coupled Nucleosome Assembly in the Passage of Epigenetic Information and Cell Identity. *Trends Biochem Sci* **43**, 136–148 (2018).

95. Kaufman, P. D. & Rando, O. J. Chromatin as a potential carrier of heritable information. *Curr Opin Cell Biol* **22**, 284–290 (2010).

96. Swygert, S. G. & Peterson, C. L. Chromatin dynamics: Interplay between remodeling enzymes and histone modifications. *Biochim Biophys Acta Gene Regul Mech* **1839**, 728–736 (2014).

97. Luk, E. *et al.* Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. *Cell* **143**, 725–736 (2010).

98. Zlatanova, J. & Thakar, A. H2A.Z: View from the Top. *Structure* **16**, 166–179 (2008).

99. Creyghton, M. P. *et al.* H2AZ Is Enriched at Polycomb Complex Target Genes in ES Cells and Is Necessary for Lineage Commitment. *Cell* **135**, 649–661 (2008).

100. Raisner, R. M. *et al.* Histone variant H2A.Z Marks the 5′ ends of both active and inactive genes in euchromatin. *Cell* **123**, 233–248 (2005).

101. Jackson, J. D. & Gorovsky, M. A. Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. *Nucleic Acids Res* **28**, 3811–3816 (2000).

102. Liu, X., Li, B. & GorovskyMA. Essential and nonessential histone H2A variants in Tetrahymena thermophila. *Mol Cell Biol* **16**, 4305–4311 (1996).

103. Brunelle, M. *et al.* The histone variant H2A.Z is an important regulator of enhancer activity. *Nucleic Acids Res* **43**, 9742–9756 (2015).

104. Altaf, M. *et al.* NuA4-dependent acetylation of nucleosomal histones H4 and H2A directly stimulates incorporation of H2A.Z by the SWR1 complex. *Journal of Biological Chemistry* **285**, 15966–15977 (2010).

105. Watanabe, S., Radman-Livaja, M., Rando, O. J. & Peterson, C. L. A histone acetylation switch regulates H2A.Z deposition by the SWR-C remodeling enzyme. *Science (1979)* **340**, 195–199 (2013).

106. Rufiange, A., Jacques, P. É., Bhat, W., Robert, F. & Nourani, A. Genome-Wide Replication-Independent Histone H3 Exchange Occurs Predominantly at Promoters and Implicates H3 K56 Acetylation and Asf1. *Mol Cell* **27**, 393–405 (2007).

107. Elsässer, S. J. & D'Arcy, S. Towards a mechanism for histone chaperones. *Biochim Biophys Acta Gene Regul Mech* **1819**, 211–221 (2012).

108. Kaufman, P. D., Kobayashi, R. & Stillman, B. Ultraviolet radiation sensitivity and reduction of telomeric silencing in Saccharomyces cerevisiae cells lacking chromatin assembly factor-I. *Genes Dev* **11**, 345–357 (1997).

109. Stillman, B. Chromatin assembly during SV40 DNA replication in vitro. *Cell* **45**, 555–565 (1986).

110. Green, E. M. *et al.* Replication-independent histone deposition by the HIR complex and Asf1. *Current Biology* **15**, 2044–2049 (2005).

111. Prochasson, P., Florens, L., Swanson, S. K., Washburn, M. P. & Workman, J. L. The HIR corepressor complex binds to nucleosomes generating a distinct protein/DNA complex resistant to remodeling by SWI/SNF. *Genes Dev* **19**, 2534–2539 (2005).

112. Houlard, M. *et al.* CAF-1 is essential for heterochromatin organization in pluripotent embryonic cells. *PLoS Genet* **2**, 1686–1696 (2006).

113. Roberts, C. *et al.* Targeted Mutagenesis of the Hira Gene Results in Gastrulation Defects and Patterning Abnormalities of Mesoendodermal Derivatives Prior to Early Embryonic Lethality. *Mol Cell Biol* **22**, 2318–2328 (2002).

114. Ray-Gallet, D. *et al.* Dynamics of Histone H3 Deposition In Vivo Reveal a Nucleosome Gap-Filling Mechanism for H3.3 to Maintain Chromatin Integrity. *Mol Cell* **44**, 928–941 (2011).

115. Zaret, K. S. & Mango, S. E. Pioneer transcription factors, chromatin dynamics, and cell fate control. *Curr Opin Genet Dev* **37**, 76–81 (2016).

116. Schulz, V. P. *et al.* A Unique Epigenomic Landscape Defines Human Erythropoiesis. *Cell Rep* **28**, 2996-3009.e7 (2019).

117. Li, D. *et al.* Chromatin Accessibility Dynamics during iPSC Reprogramming. *Cell Stem Cell* **21**, 819-833.e6 (2017).

118. Chen, H. & Pugh, B. F. What do Transcription Factors Interact With? *J Mol Biol* **433**, 166883 (2021).

119. Qasim, M. N., Valle Arevalo, A., Nobile, C. J. & Hernday, A. D. The Roles of Chromatin Accessibility in Regulating the Candida albicans White-Opaque Phenotypic Switch. *Journal of Fungi* **7**, 37 (2021).

120. Brahma, S. & Henikoff, S. Epigenome Regulation by Dynamic Nucleosome Unwrapping. *Trends Biochem Sci* **45**, 13–26 (2020).

121. Chereji, R. v. & Clark, D. J. Major Determinants of Nucleosome Positioning. *Biophys J* **114**, 2279–2289 (2018).

122. Zaret, K. S. & Carroll, J. S. Pioneer transcription factors: Establishing competence for gene expression. *Genes Dev* **25**, 2227–2241 (2011).

123. Iwafuchi-Doi, M. & Zaret, K. S. Pioneer transcription factors in cell reprogramming. *Genes Dev* **28**, 2679–2692 (2014).

124. Soufi, A. *et al.* Pioneer transcription factors target partial DNA motifs on nucleosomes to initiate reprogramming. *Cell* **161**, 555–568 (2015).

125. Kubik, S., Bruzzone, M. J. & Shore, D. Establishing nucleosome architecture and stability at promoters: Roles of pioneer transcription factors and the RSC chromatin remodeler. *BioEssays* **39**, 1–10 (2017).

126. Henikoff, S. & Ramachandran, S. Pioneers Invade the Nucleosomal Landscape. *Mol Cell* **71**, 193–194 (2018).

127. Zaret, K. S. Pioneering the chromatin landscape. *Nat Genet* **50**, 167–169 (2018).

128. Mayran, A. & Drouin, J. Pioneer transcription factors shape the epigenetic landscape. *Journal of Biological Chemistry* **293**, 13795–13804 (2018).

129. Zaret, K. S. Pioneer Transcription Factors Initiating Gene Network Changes. *Annu Rev Genet* **54**, 367–385 (2020).

130. Tanaka, H. *et al.* Interaction of the pioneer transcription factor GATA3 with nucleosomes. *Nat Commun* **11**, 4136 (2020).

131. Rossi, M. J., Lai, W. K. M. & Pugh, B. F. Genome-wide determinants of sequence-specific DNA binding of general regulatory factors. *Genome Res* **28**, 497–508 (2018).

132. Kubik, S. *et al.* Sequence-Directed Action of RSC Remodeler and General Regulatory Factors Modulates +1 Nucleosome Position to Facilitate Transcription. *Mol Cell* **71**, 89-102.e5 (2018).

133. Ganapathi, M. *et al.* Extensive role of the general regulatory factors, Abf1 and Rap1, in determining genome-wide chromatin structure in budding yeast. *Nucleic Acids Res* **39**, 2032–2044 (2011).

134. Mayran, A. *et al.* Pioneer and nonpioneer factor cooperation drives lineage specific chromatin opening. *Nat Commun* **10**, 3807 (2019).

135. Meers, M. P., Janssens, D. H. & Henikoff, S. Pioneer Factor-Nucleosome Binding Events during Differentiation Are Motif Encoded. *Mol Cell* **75**, 562-575.e5 (2019).

136. Johnson, T. A. *et al.* Conventional and pioneer modes of glucocorticoid receptor interaction with enhancer chromatin in vivo. *Nucleic Acids Res* **46**, 203–214 (2018).

137. Miao, L. *et al.* The landscape of pioneer factor activity reveals the mechanisms of chromatin reprogramming and genome activation. *Mol Cell* **82**, 986-1002.e9 (2022).

138. Iwafuchi-Doi, M. *et al.* The Pioneer Transcription Factor FoxA Maintains an Accessible Nucleosome Configuration at Enhancers for Tissue-Specific Gene Activation. *Mol Cell* **62**, 79–91 (2016).

139. King, H. W. & Klose, R. J. The pioneer factor OCT4 requires the chromatin remodeller BRG1 to support gene regulatory element function in mouse embryonic stem cells. *Elife* **6**, (2017).

140. Iwafuchi, M. *et al.* Gene network transitions in embryos depend upon interactions between a pioneer transcription factor and core histones. *Nat Genet* **52**, 418–427 (2020).

141. Swinstead, E. E., Paakinaho, V., Presman, D. M. & Hager, G. L. Pioneer factors and ATP-dependent chromatin remodeling factors interact dynamically: A new perspective. *BioEssays* **38**, 1150–1157 (2016).

142. Lohse, M. B. & Johnson, A. D. White-opaque switching in Candida albicans. *Curr Opin Microbiol* **12**, 650–654 (2009).

143. Palková, Z. & Váchová, L. Yeast cell differentiation: Lessons from pathogenic and non-pathogenic yeasts. *Semin Cell Dev Biol* **57**, 110–119 (2016).

144. Ramírez-Zavala, B., Reuß, O., Park, Y. N., Ohlsen, K. & Morschhäuser, J. Environmental induction of white-opaque switching in Candida albicans. *PLoS Pathog* **4**, (2008).

145. Sun, Y. *et al.* PH regulates white-opaque switching and sexual mating in Candida albicans. *Eukaryot Cell* **14**, 1127–1134 (2015).

146. Rao, K. H., Ruhela, D., Ghosh, S., Abdin, M. Z. & Datta, A. N-acetylglucosamine kinase, HXK1 contributes to white-opaque morphological transition in Candida albicans. *Biochem Biophys Res Commun* **445**, 138–144 (2014).

147. Geiger, J., Wessels, D., Lockhart, S. R. & Soll, D. R. Release of a Potent Polymorphonuclear Leukocyte Chemoattractant Is Regulated by White-Opaque Switching in Candida albicans. *Infect Immun* **72**, 667–677 (2004).

148. Ene, I. v. *et al.* Phenotypic profiling reveals that Candida albicans opaque cells represent a metabolically specialized cell state compared to default white cells. *mBio* **7**, 1–20 (2016).

149. Huang, G. *et al.* Bistable expression of WOR1, a master regulator of white-opaque switching in Candida albicans. *Proc Natl Acad Sci U S A* **103**, 12813–12818 (2006).

150. Zhang, S., Zhang, T., Yan, M., Ding, J. & Chen, J. Crystal structure of the WOPR-DNA complex and implications for Wor1 function in white-opaque switching of Candida albicans. *Cell Res* **24**, 1108–1120 (2014).

151. Lohse, M. B., Zordan, R. E., Cain, C. W. & Johnson, A. D. Distinct class of DNA-binding domains is exemplified by a master regulator of phenotypic switching in Candida albicans. *Proc Natl Acad Sci U S A* **107**, 14105–14110 (2010).

152. Fordyce, P. M. *et al.* De novo identification and biophysical characterization of transcription-factor binding sites with microfluidic affinity analysis. *Nature Biotechnology 2010 28:9* **28**, 970–975 (2010).

153. Maerkl, S. J. & Quake, S. R. A systems approach to measuring the binding energy landscapes of transcription factors. *Science (1979)* **315**, 233–237 (2007).

154. Xie, J., Jenull, S., Tscherner, M. & Kuchler, K. The Paralogous Histone Deacetylases Rpd3 and Rpd31 Play Opposing Roles in Regulating the White-Opaque Switch in the Fungal Pathogen Candida albicans. *mBio* **7**, (2016).

155. Rai, L. S., Singha, R., Brahma, P. & Sanyal, K. Epigenetic determinants of phenotypic plasticity in Candida albicans. *Fungal Biol Rev* **32**, 10–19 (2018).

156. Takahashi, K. & Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell* **126**, 663–676 (2006).

157. Lee, C. S., Friedman, J. R., Fulmer, J. T. & Kaestner, K. H. The initiation of liver development is dependent on Foxa transcription factors. *Nature 2005 435:7044* **435**, 944–947 (2005).

158. Buenrostro, J. D., Wu, B., Chang, H. Y. & Greenleaf, W. J. ATAC-seq: A method for assaying chromatin accessibility genome-wide. *Curr Protoc Mol Biol* **2015**, 21.29.1-21.29.9 (2015).

159. Buenrostro, J. D., Giresi, P. G., Zaba, L. C., Chang, H. Y. & Greenleaf, W. J. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat Methods* **10**, 1213–1218 (2013).

160. Wal, M. & Pugh, B. F. Genome-Wide Mapping of Nucleosome Positions in Yeast Using High-Resolution MNase ChIP-Seq. *Methods Enzymol* **513**, 233–250 (2012).

161. Reske, J. J., Wilson, M. R. & Chandler, R. L. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. *Epigenetics Chromatin* **13**, 22 (2020).

162. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, (2014).

163. Grant, C. E., Bailey, T. L. & Noble, W. S. FIMO: scanning for occurrences of a given motif. *Bioinformatics* **27**, 1017 (2011).

164. Chereji, R. v., Ocampo, J. & Clark, D. J. MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers. *Mol Cell* **65**, 565-577.e3 (2017).

165. Brahma, S. & Henikoff, S. RSC-Associated Subnucleosomes Define MNase-Sensitive Promoters in Yeast. *Mol Cell* **73**, 238-249.e3 (2019).

166. Ramani, V., Qiu, R. & Shendure, J. High Sensitivity Profiling of Chromatin Structure by MNase-SSP. *Cell Rep* **26**, 2465-2476.e4 (2019).

167. Mavrich, T. N. *et al.* A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res* **18**, 1073–1083 (2008).

168. Rao, S., Ahmad, K. & Ramachandran, S. Cooperative binding between distant transcription factors is a hallmark of active enhancers. *Mol Cell* **81**, 1651-1665.e4 (2021).

169. Ramachandran, S., Ahmad, K. & Henikoff, S. Transcription and Remodeling Produce Asymmetrically Unwrapped Nucleosomal Intermediates. *Mol Cell* **68**, 1038-1053.e4 (2017).

170. Knight, B. *et al.* Two distinct promoter architectures centered on dynamic nucleosomes control ribosomal protein gene transcription. *Genes Dev* **28**, 1695–1709 (2014).

171. Min, K., Biermann, A., Hogan, D. A. & Konopka, J. B. Genetic Analysis of NDT80 Family Transcription Factors in Candida albicans Using New CRISPR-Cas9 Approaches. *mSphere* **3**, (2018).

172. Sopko, R. & Stuart, D. T. Purification and characterization of the DNA binding domain of Saccharomyces cerevisiae meiosis-specific transcription factor Ndt80. *Protein Expr Purif* **33**, 134–144 (2004).

173. Montano, S. P. *et al.* Crystal structure of the DNA-binding domain from Ndt80, a transcriptional activator required for meisis in yeast. *Proc Natl Acad Sci U S A* **99**, 14041–14046 (2002).

174. Fingerman, I. M. Characterization of critical interactions between Ndt80 and MSE DNA defining a novel family of Ig-fold transcription factors. *Nucleic Acids Res* **32**, 2947–2956 (2004).

175. Sellam, A. *et al.* Role of Transcription Factor CaNdt80p in Cell Separation, Hyphal Growth, and Virulence in Candida albicans. *Eukaryot Cell* **9**, 634–644 (2010).

176. Wang, J. S. *et al.* The DNA-binding domain of CaNdt80p is required to activate CDR1 involved in drug resistance in Candida albicans. *J Med Microbiol* **55**, 1403–1411 (2006).

177. Qasim, M. N. *et al.* Genome-wide Profiling of Transcription Factor-DNA Binding Interactions in Candida albicans: A Comprehensive CUT&RUN Method and Data Analysis Workflow. *Journal of Visualized Experiments* **2022**, (2022).

178. Mancera, E. *et al.* Evolution of the complex transcription network controlling biofilm formation in candida species. *Elife* **10**, (2021).

179. Nobile, C. J. *et al.* A recently evolved transcriptional network controls biofilm development in Candida albicans. *Cell* **148**, 126–138 (2012).

180. Yan, C., Chen, H. & Bai, L. Systematic Study of Nucleosome-Displacing Factors in Budding Yeast. *Mol Cell* **71**, 294-305.e4 (2018).

181. Bai, L., Ondracka, A. & Cross, F. R. Multiple Sequence-Specific Factors Generate the Nucleosome-Depleted Region on CLN2 Promoter. *Mol Cell* **42**, 465–476 (2011).

182. Lai, W. K. M. & Pugh, B. F. Understanding nucleosome dynamics and their links to gene expression and DNA replication. *Nat Rev Mol Cell Biol* **18**, 548–562 (2017).

183. Hainer, S. J., Pruneski, J. A., Mitchell, R. D., Monteverde, R. M. & Martens, J. A. Intergenic transcription causes repression by directing nucleosome assembly. *Genes Dev* **25**, 29–40 (2011).

184. Chereji, R. v. *et al.* Genome-wide profiling of nucleosome sensitivity and chromatin accessibility in Drosophila melanogaster. *Nucleic Acids Res* **44**, 1036–1051 (2016).

185. Kubik, S. *et al.* Nucleosome Stability Distinguishes Two Different Promoter Types at All Protein-Coding Genes in Yeast. *Mol Cell* **60**, 422–434 (2015).

186. Schep, A. N. *et al.* Structured nucleosome fingerprints enable high-resolution mapping of chromatin architecture within regulatory regions. *Genome Res* **25**, 1757–1770 (2015).

187. Chen, K. *et al.* DANPOS: Dynamic analysis of nucleosome position and occupancy by sequencing. *Genome Res* **23**, 341–351 (2013).

188. Bailey, T. L. & Elkan, C. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. (1994).

189. Gulati, M. & Nobile, C. J. Candida albicans biofilms: development, regulation, and molecular mechanisms. *Microbes Infect* **18**, 310–321 (2016).

190. Lohse, M. B., Gulati, M., Johnson, A. D. & Nobile, C. J. Development and regulation of single-and multi-species Candida albicans biofilms. *Nat Rev Microbiol* **16**, 19–31 (2018).

191. Nobile, C. J. & Johnson, A. D. Candida albicans Biofilms and Human Disease. *Annu Rev Microbiol* **69**, 71–92 (2015).

192. Iracane, E., Vega-Estévez, S. & Buscaino, A. On and Off: Epigenetic Regulation of C. albicans Morphological Switches. *Pathogens* **10**, 1463 (2021).

193. Hernday, A. D., Noble, S. M., Mitrovich, Q. M. & Johnson, A. D. Genetics and molecular biology in Candida albicans. *Methods Enzymol* **470**, 737–758 (2010).

194. Lee, T. I., Johnstone, S. E. & Young, R. A. Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat Protoc* **1**, 729–748 (2006).

195. Zentner, G. E., Kasinathan, S., Xin, B., Rohs, R. & Henikoff, S. ChEC-seq kinetics discriminates transcription factor binding sites by DNA sequence and shape in vivo. *Nat Commun* **6**, 8733 (2015).

196. Schmid, M., Durussel, T. & Laemmli, U. K. ChIC and ChEC: Genomic mapping of chromatin proteins. *Mol Cell* **16**, 147–157 (2004).

197. Skene, P. J. & Henikoff, S. An efficient targeted nuclease strategy for high-resolution mapping of DNA binding sites. *Elife* **6**, (2017).

198. Skene, P. J., Henikoff, J. G. & Henikoff, S. Targeted in situ genome-wide profiling with high efficiency for low cell numbers. *Nat Protoc* **13**, 1006–1019 (2018).

199. Meers, M. P., Bryson, T. D., Henikoff, J. G. & Henikoff, S. Improved CUT&RUN chromatin profiling tools. *Elife* **8**, (2019).

200. Homann, O. R. & Johnson, A. D. MochiView: Versatile software for genome browsing and DNA motif analysis. *BMC Biol* **8**, 1–8 (2010).

201. Hainer, S. J. & Fazzio, T. G. High-Resolution Chromatin Profiling Using CUT&RUN. *Curr Protoc Mol Biol* **126**, e85 (2019).

202. Rodriguez, D. L., Quail, M. M., Hernday, A. D. & Nobile, C. J. Transcriptional Circuits Regulating Developmental Processes in Candida albicans. *Front Cell Infect Microbiol* **10**, 752 (2020).

203. Zhu, Q., Liu, N., Orkin, S. H. & Yuan, G.-C. CUT&RUNTools: a flexible pipeline for CUT&RUN processing and footprint analysis. *Genome Biol* **20**, 192 (2019).

204. Teytelman, L., Thurtle, D. M., Rine, J. & Van Oudenaarden, A. Highly expressed loci are vulnerable to misleading ChIP localization of multiple unrelated proteins. *Proc Natl Acad Sci U S A* **110**, 18602–18607 (2013).

205. Nguyen, N., Quail, M. M. F. & Hernday, A. D. An Efficient, Rapid, and Recyclable System for CRISPR-Mediated Genome Editing in Candida albicans. *American Scociety For Microbiology* **2**, 1–10 (2017).

206. Ennis, C. L., Hernday, A. D. & Nobile, C. J. A Markerless CRISPR-Mediated System for Genome Editing in Candida auris Reveals a Conserved Role for Cas5 in the Caspofungin Response. *Microbiol Spectr* **9**, (2021).

207. Skrzypek, M. S. *et al.* The Candida Genome Database (CGD): incorporation of Assembly 22, systematic identifiers and visualization of high throughput sequencing data. *Nucleic Acids Res* **45**, D592–D596 (2017).

208. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).

209. Landt, S. G. *et al.* ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. *Genome Res* **22**, 1813–1831 (2012).

210. Boyd, J., Rodriguez, P., Schjerven, H. & Frietze, S. ssvQC: an integrated CUT&RUN quality control workflow for histone modifications and transcription factors. *BMC Res Notes* **14**, 1–7 (2021).

211. Kent, W. J., Zweig, A. S., Barber, G., Hinrichs, A. S. & Karolchik, D. BigWig and BigBed: enabling browsing of large distributed datasets. *Bioinformatics* **26**, 2204–2207 (2010).

212. Soll, D. R., Lockhart, S. R. & Zhao, R. Relationship between switching and mating in Candida albicans. *Eukaryot Cell* **2**, 390–397 (2003).

213. Frazer, C., Hernday, A. D. & Bennett, R. J. Monitoring Phenotypic Switching in Candida albicans and the Use of Next-Gen Fluorescence Reporters. *Curr Protoc Microbiol* **53**, 1–27 (2019).

214. Si, H., Hernday, A. D., Hirakawa, M. P., Johnson, A. D. & Bennett, R. J. Candida albicans White and Opaque Cells Undergo Distinct Programs of Filamentous Growth. *PLoS Pathog* **9**, e1003210 (2013).

215. Xie, J. *et al.* White-Opaque Switching in Natural MTLa/α Isolates of Candida albicans: Evolutionary Implications for Roles in Host Adaptation, Pathogenesis, and Sex. *PLoS Biol* **11**, (2013).

216. Yue, H. *et al.* Discovery of the gray phenotype and white-gray-opaque tristable phenotypic transitions in Candida dubliniensis. *Virulence* **7**, 230–242 (2016).

217. Bentz, M. L., Joseph Sexton, D., Welsh, R. M. & Litvintseva, A. P. Phenotypic switching in newly emerged multidrug-resistant pathogen Candida auris. *Med Mycol* **57**, 636–638 (2019).

218. Ramachandran, S., Ahmad, K. & Henikoff, S. Capitalizing on disaster: Establishing chromatin specificity behind the replication fork. *BioEssays* **39**, 1–8 (2017).

219. Stadhouders, R. *et al.* Transcription factors orchestrate dynamic interplay between genome topology and gene regulation during cell reprogramming. *Nat Genet* **50**, 238–249 (2018).

220. Mirny, L. A. Nucleosome-mediated cooperativity between transcription factors. *Proceedings of the National Academy of Sciences* **107**, 22534–22539 (2010).

221. Ramachandran, S. & Henikoff, S. Transcriptional Regulators Compete with Nucleosomes Post-replication. *Cell* **165**, 580–592 (2016).

222. Hug, C. B., Grimaldi, A. G., Kruse, K. & Vaquerizas, J. M. Chromatin Architecture Emerges during Zygotic Genome Activation Independent of Transcription. *Cell* **169**, 216-228.e19 (2017).

223. Knaupp, A. S. *et al.* Transient and Permanent Reconfiguration of Chromatin and Transcription Factor Occupancy Drive Reprogramming. *Cell Stem Cell* **21**, 834-845.e6 (2017).

224. Luo, H. *et al.* Cell identity bookmarking through heterogeneous chromatin landscape maintenance during the cell cycle. *Hum Mol Genet* **26**, 4231–4243 (2017).

225. Katz, M. E. Nutrient sensing-the key to fungal p53-like transcription factors? *Fungal Genetics and Biology* **124**, 8–16 (2019).

226. Imbert, C. *Fungal Biofilms and related infections: Advances in Microbiology*. vol. 3 (2016).

227. Kubik, S., Bruzzone, M. J., Albert, B. & Shore, D. A Reply to “MNase-Sensitive Complexes in Yeast: Nucleosomes and Non-histone Barriers,” by Chereji et al. *Mol Cell* **65**, 578–580 (2017).

228. Chereji, R. v., Bryson, T. D. & Henikoff, S. Quantitative MNase-seq accurately maps nucleosome occupancy levels. *Genome Biol* **20**, 198 (2019).

229. Noble, S. M., French, S., Kohn, L. A., Chen, V. & Johnson, A. D. Systematic screens of a Candida albicans homozygous deletion library decouple morphogenetic switching and pathogenicity. *Nat Genet* **42**, 590–598 (2010).

230. Mivelaz, M. *et al.* Chromatin Fiber Invasion and Nucleosome Displacement by the Rap1 Transcription Factor. *Mol Cell* **77**, 488-500.e9 (2020).

231. Fox, E. P. *et al.* An expanded regulatory network temporally controls Candida albicans biofilm formation. *Mol Microbiol* **96**, 1226–1239 (2015).