

**DYNAMIC HISTONE MODIFICATIONS COORDINATE
TRANSCRIPTION, DNA REPAIR AND RECOMBINATION
ACROSS DISTINCT METABOLIC STATES**

by
Zheng Kuang

A dissertation submitted to Johns Hopkins University in conformity with the
requirements for the degree of Doctor of Philosophy

Baltimore, Maryland
June, 2014

© 2014 Zheng Kuang
All Rights Reserved

Abstract

Eukaryotic genomes are packed in the form of chromatin, a complex of DNA and protein. The basic unit is called a nucleosome and consists of 147 bp of DNA wrapped around a histone octamer consisting of two molecules each of core histones H2A, H2B, H3 and H4. Histone proteins are subject to diverse types of post-translational modifications (PTMs), which exert dramatic influence on almost all DNA-associated processes, such as transcription, DNA replication and repair. The pervasiveness of biological dynamics suggests that temporal interrogation of chromatin function is warranted. By using the yeast metabolic cycle (YMC) as the system model, we examined the genome-wide transcription and chromatin states at unprecedented temporal resolution by RNA-seq and ChIP-seq. We revealed a “just in time supply chain” by which specific cellular processes such as ribosome biogenesis are coordinated in time with remarkable precision. We identified distinct chromatin and splicing patterns associated with different gene categories and determine the relative timing of chromatin modifications to maximal transcription. We examined the dynamic occupancy of chromatin modifiers and revealed subtly distinct spatial and temporal patterns compared to the modifications themselves. Genetic analyses support a potentially cooperative role of histone modifications in the YMC. Additionally, We identified ~1000 H3K56ac peaks in the YMC of a haploid prototrophic yeast strain which spatially correlate with meiotic DSB hotspots. Spo11 and other meiotic DSB proteins are actively regulated and contribute to the formation of H3K56ac peaks under stress conditions. Spo11 is required for mitotic DNA recombination at meiotic DSB hotspots and facilitates yeast evolution under stress conditions. Furthermore, we developed a computational pipeline to predict the dynamic activity of transcription factors via integrating transcription factor motif sites with time-course ChIP-seq data of histone modification. We screened 177 transcription factors in the YMC and found that 55 of them are enriched at specific stages. We

further validated the binding and function of eight transcription factors in this process. This method is a valuable tool to study the functions of transcription factors during dynamic processes.

Jef D. Boeke, Ph.D. (Thesis Advisor and Reader)

Professor, Department of Molecular Biology and Genetics

Johns Hopkins University School of Medicine

Sean D. Taverna, Ph.D. (Thesis Reader)

Assistant Professor, Department of Pharmacology and Molecular Sciences

Johns Hopkins University School of Medicine

Acknowledgements

First and foremost, I would like to express my sincere appreciation and gratitude to my advisor, Dr. Jef Boeke. I feel extremely lucky to have Jef as my PhD advisor. He is very supportive and helpful in almost everywhere. I have been exposed to all necessary scientific resources and a stimulating environment supported by Jef. He is very encouraging and has given me freedom to explore various projects. Not only in the guidance of my research, he constantly motivated me to better communicate it to the public. He always encouraged me to express my ideas in all circumstances and helped me improve writing skills. I can never thank Jef enough.

Secondly, I would like to thank my committee members, Brendan Cormack, Sean Taverna, Joel Bader, Benjamin Tu and Hongkai Ji. They have given constructive suggestions and guidance over the years and I really appreciate their valuable input. I especially want to thank Benjamin Tu for the collaboration and the continuous support in the yeast metabolic cycle projects. I also want to thank Hongkai Ji for the guidance of my MHS thesis project in biostatistics. In addition, I would like to thank Sean Taverna for reading my thesis and giving me valuable suggestions.

I am grateful to the BCMB program for all the support given to me. The BCMB program is a warm home and a very helpful platform. It provides great opportunities for research and other possibilities. The director Carolyn Machamer and the administrative staff Margie, Sharon, Lesley, Arhonda and Jessica have helped me through the PhD study smoothly.

I would like to express my sincere gratitude to my collaborators. I couldn't complete this thesis without their support. I want to thank Ling Cai and Lei Shi for their expertise in the yeast metabolic cycle and their efforts in ChIP-seq experiments. I would like to thank Xuekui Zhang for his help in the analysis of ChIP-seq data. I appreciate Govind Rao and Shaunak Uplekar for their help in setting up fermentation. I am grateful to Amar Klar for sharing his *S.pombe* and *S.japonicus* strains and I want to thank Stefan Canzar for his efforts in the analysis of the Pacbio sequencing data.

I would also like to thank past and present members in the Boeke lab for their valuable discussion and their efforts in making the lab a great place for research. I want to specially thank Pamela Meluh, Lisa Huang, Junbiao Dai, Lixin Dai, Marty Taylor, Yizhi Cai and Anna Sliva for their ideas and reagents to facilitate my study and making my life in the lab much more enjoyable and colorful. I wish to single out Yu-yi Lin and express my sincere gratitude to him for being a great mentor. It was a pleasure to work with him and I miss him very much.

I have made special friendship during my graduate study inside and outside of Johns Hopkins campus, which makes my life in Baltimore much more enjoyable. I am very grateful for that. Finally, I would like to express my sincere thanks to my family who always support me. My parents, Yiqun Kuang and Meiqing Qiu, have been giving me endless love and encouraging me through the tough times. I would like to thank my wife, Yulian Zhou. We met each other at Hopkins and we have gone through every piece of happiness and sadness together. I am sincerely grateful for all she has done for me. I will forever be indebted to my family.

Table of Contents

Title Page	i
Abstract	ii
Acknowledgements	iv
Table of Contents	vi
List of Tables	ix
List of Figures	x
Chapter 1: Introduction	1
Histone modifications and modifiers	4
Combinatorial histone modifications	5
Spatial distribution of histone modifications and modifiers	8
Temporal pattern of histone modifications and modifiers	11
Genome-wide analysis of dynamic histone modifications	17
Summary	19
Chapter 2: High temporal resolution view of transcription and chromatin states across distinct metabolic states in budding yeast	25
Abstract	26
Introduction	27
Results	29
Gene expression patterns in the YMC	29
Accumulation of pre-mRNAs in the OX phase	30
Dynamics of histone modifications across YMC	32
Periodic transcription and histone modification are not systematically correlated	34
Landscape of chromatin-modifying enzyme occupancy in the YMC	36
Histone mutant analysis	38

Discussion	39
Methods and Materials	42
Chapter 3: Genome-wide location of H3K56ac reveals active roles of "meiosis-specific"	
Spo11 in mitosis	78
Abstract	79
Introduction	80
Results	81
Spatial and temporal patterns of H3K56ac in the yeast metabolic cycle	81
H3K56ac correlates with meiotic DSB hotspots spatially	82
Spo11 is required for H3K56ac peaks	83
Spo11 is activated in mitosis of haploid yeast	83
Spo11 regulates mitotic recombination and heatshock resistance	84
H3K56ac ensures the completion of gametogenesis	85
Discussion	86
Methods and Materials	88
Chapter 4: Predicting dynamic activity of transcription factors through time-course	
ChIP-seq data of histone modifications	103
Abstract	104
Introduction	105
Results	106
Prediction Pipeline	106
Motif mapping	107
Peak calling	108
Filtering	108
Clustering	110
Confirmation	111

Discussion	112
Methods and Materials	113
Reference	133
CURRICULUM VITAE	146

List of Tables

Table 1.1 Common histone modifications in <i>S.cerevisiae</i> and the enzymes that add or remove them.	21
Table 2.1 Yeast strains used in chapter 2.	43
Table 3.1 Yeast strains used in chapter 3.	88
Table 4.1 Summary of numbers of H3K9ac peaks detected by different peak callers.	119
Table 4.2 Numbers of motif sites picked by these 4 peak callers.	120
Table 4.3 Numbers of motif sites picked by combinatorial peak callers.	121
Table 4.4 Percentages of motif sites uniquely picked by each peak caller.	122
Table 4.5 Numbers of motif sites after Naïve Bayes classifier filtering.	123
Table 4.6 TFs randomly picked by ChIP-qPCR validation.	124

List of Figures

Figure 1.1 Native and synthetic histone genes in <i>S.cerevisiae</i> .	23
Figure 1.2 Dynamic histone modifications during activation of the yeast genes.	24
Figure 2.1 High temporal resolution analysis of gene expression reveals meticulous temporal compartmentalization in yeast.	55
Figure 2.2 Dynamics of gene expression in YMC.	57
Figure 2.3 RNA-seq analysis at introns reveals transient accumulation of pre-mRNAs during OX phase.	59
Figure 2.4 Temporal patterns of RNA-seq signals at RP introns.	61
Figure 2.5 Dynamic chromatin states across the YMC.	63
Figure 2.6 Dynamic chromatin modifications across the YMC.	65
Figure 2.7 Temporal patterns of chromatin modifications at OX genes.	66
Figure 2.8 Temporal patterns of H3K56ac.	67
Figure 2.9 Temporal association between chromatin states and gene expression.	68
Figure 2.10 Temporal combinatorial patterns of chromatin states and gene expression define functionally related genes.	70
Figure 2.11 Similar but not identical spatial and temporal patterns of chromatin modifiers and corresponding modifications.	72
Figure 2.12 Dynamic analysis of chromatin modifiers and corresponding modifications.	74
Figure 2.13 O ₂ consumption traces of histone point mutants.	76
Figure 2.14 Analysis of histone mutants in O ₂ oscillation.	77
Figure 3.1 Spatial and temporal patterns of H3K56ac in YMC.	94

Figure 3.2 H3K56ac peaks correlate with meiotic DSB hotspots but not mitotic DSB hotspots.	95
Figure 3.3 H3K56ac peaks correlate with meiotic DSB hotspots at example genes.	96
Figure 3.4 Spo11 is required for H3K56ac peaks.	97
Figure 3.5 Spo11 is activated in the RC phase of YMC.	98
Figure 3.6 Deletion of <i>SPO11</i> causes decrease of homologous recombination.	99
Figure 3.7 LOH assay mapping individual homologous recombination.	100
Figure 3.8 Evolution of heat shock tolerance in S288C and CEN.PK hybrid background.	101
Figure 3.9 H3K56ac ensures normal spore viability.	102
Figure 4.1 Pipeline for prediction of spatio-temporal TF binding in dynamic processes.	125
Figure 4.2 Distribution of H3K9ac read counts.	127
Figure 4.3 Distribution of motif sites at TSSs.	128
Figure 4.4 Temporal patterns of H3K9ac in YMC.	129
Figure 4.5 Numbers of targeted genes in filtered motif sites and control motif sites.	130
Figure 4.6 Validation of spatial and temporal recruitment of TFs.	132

Chapter 1

Introduction

Eukaryotic genomes are packed in the form of chromatin, a complex of DNA and protein. The primary structure of chromatin is a 10 nm “beads-on-a-string” fiber, which can be further packaged into more condensed configuration (Olins and Olins, 1974). The basic unit, or “bead”, is called a nucleosome and consists of 147 bp of DNA wrapped around a histone octamer consisting of two molecules each of core histones H2A, H2B, H3 and H4 (Luger et al., 1997). Chromatin structure not only controls the compaction state of DNA, but also exerts dramatic influence on almost all DNA-associated processes, such as transcription, DNA replication and repair.

Transcription is an essential process in biology to transfer genetic information from DNA to RNA by RNA polymerases. For decades, drastic progress has been made to understand the regulatory mechanisms of accuracy and specificity of transcription by cis- (i.e. DNA sequences) and trans- (i.e. DNA binding proteins) factors (Venters and Pugh, 2009). Among them, chromatin is now recognized not only as a simple passive barrier to transcription but can also transform dynamically to an active platform to help ensure spatial and temporal control of transcription. Nucleosome dynamics, chromatin remodeling, histone variants and histone modifications are the major contributors to the regulation of transcription (Rando and Winston, 2012). In general, they function in opening the chromatin structure to facilitate the passage of RNA polymerases and recruiting gene-specific transacting factors to their proper functional sites. An excellent recent review by Rando and Winston systematically describes how these factors regulate transcription in detail (Rando and Winston, 2012). Here, we will focus on the dynamical aspects of histone modifications in transcription.

The histone genes are distinctively conserved across eukaryotic species (Luger et al., 1997). Probably due to their essentiality, histone genes usually have multiple copies, 2 copies in *S. cerevisiae* and dozens to hundreds of copies in higher eukaryotes (Hereford et al., 1979; Lifton et

al., 1978; Marzluff et al., 2002; Smith and Andr sson, 1983). The yeast histone genes are arranged in four divergently transcribed pairs at different genomic loci: HTA1-HTB1 and HTA2-HTB2 for H2A and H2B; HHT1-HHF1 and HHT2-HHF2 for H3 and H4 (Eriksson et al., 2012) (Fig. 1.1A). Deletion of one copy of any histone gene does not cause lethality but may modulate the behaviors of yeast in growth and responses to different stimuli (Norris and Osley, 1987; Norris et al., 1988; Rykowski et al., 1981; Smith and Stirling, 1988). There are also three other types of histones in yeast, which are linker histone H1 encoded by HHO1 (Patterton et al., 1998), the removal of which has surprisingly little phenotypic effect, a centromere-specific H3 variant named Cse4 (Meluh et al., 1998) and the H2A variant H2A.Z encoded by *HTZ1* (Santisteban et al., 2000). Expression levels of histone genes are carefully regulated by transcription factors, histone chaperones and post-transcriptional mechanisms (Eriksson et al., 2012). Perturbation of histone levels can alter gene expression, probably by reducing global nucleosome occupancy (Clark-Adams et al., 1988). For example, the yeast *PHO5*, *GAL1*, *CYCI*, *CUP1* and *HIS3* promoter activities are increased by repression of histone H4 synthesis (Durrin et al., 1992; Han and Grunstein, 1988). The fact that *S. cerevisiae* histone genes have just two copies each and the fact that they always come in pairs greatly facilitate genetic analysis of histone functions. N-terminal tails of core histone proteins have been found to have defects in activating or repressing gene transcription (Durrin et al., 1991; Lenfant et al., 1996; Wan et al., 1995). A number of lysine codon mutants were also found to affect transcription of representative genes or a set of genes by genome-wide expression analyses (Dion et al., 2005; Hirschhorn et al., 1995; Jin et al., 2009; Mann and Grunstein, 1992). Recent systematic collections of histone gene mutants provide a versatile tool to interrogate functions of specific histone residues in transcription and other processes (Fig. 1.1B) (Dai et al., 2008; Kawano et al., 2011; Matsubara et al., 2007; Nakanishi et al., 2008).

Histone modifications and modifiers

Histone proteins are subject to diverse types of post-translational modifications (PTMs), including lysine acetylation, methylation, ubiquitylation, sumoylation, arginine methylation, serine and threonine phosphorylation and others (Table 1.1). A majority of the modifications are located on the N-terminal tails that extend out from the globular nucleosome core and these have been studied more intensively than the modifications on the core region. Interestingly, nearly all these PTMs are linked to transcription through two kinds of mechanisms. Acetylation neutralizes the positive charges on lysine and phosphorylation attaches negative charges on serine and threonine. Alternation of the charge distribution on histones can disrupt electrostatic contact between histone and DNA and loosen the compact chromatin structure, therefore facilitating remodeling of nucleosomes and improving accessibility of transcription factors (Cosgrove et al., 2004). H4K16ac directly inhibits chromatin higher order folding (Shogren-Knaak et al., 2006). However, other PTMs like methylation don't have obvious electrostatic effects. Another common mechanism is that histone PTMs alter the binding properties of nucleosomes to help recruit specific regulatory proteins (also called "histone readers") that modulate transcription activities (Taverna et al., 2007). A well-studied reader of histone acetylation is the bromodomain, which exists in many complexes regulating transcription in yeast (Table 1.1). For example, chromatin-remodeling complexes Swi/Snf and RSC both contain bromodomains, and mutations of these domains either decrease remodeling activity or cause lethality (Cairns et al., 1999; Hassan et al., 2001; Haynes et al., 1992; Kasten et al., 2004). Thus histone acetylation may effect chromatin behavior both by "reader" interactions and charge effects. Additional domains such as PHD and chromodomains recognize methylated lysine and regulate transcription (Table 1.1). For example, H3K4me3 recruits the Yng1 subunit of NuA3 (Nucleosome acetyltransferase of H3 complex) via its PHD domain to keep hyper-acetylation at H3K14 and active transcription of Yng1 targeted

genes (Taverna et al., 2006). Comprehensive discussion of histone readers can be found in the following reviews (Taverna et al., 2007; Yun et al., 2011).

Besides the histone readers, many histone modifying enzymes have been found which can add (writer) or remove (eraser) modifications on histones (as well as other proteins), strengthening the notion that histone modifications may be dynamically regulated. Acetylation is catalyzed by histone acetyltransferases (HATs) and removed by deacetylases (HDACs) (Table 1.1), also known as KATs and KDACs. Hat1 was the first HAT identified in yeast and it functions in DNA repair and telomeric silencing rather than gene activation (Kelly et al., 2000; Kleff et al., 1995; Qin and Parthun, 2002; Travis et al., 1984). However, many HATs and HDACs were identified in proteins previously known as transcription factors. For example, Gcn5, originally known as a co-activator, was proven to possess HAT activity (Brownell et al., 1996). Rpd3, an HDAC previously recognized as a transcriptional repressor, is another example (Taunton et al., 1996; Vidal and Gaber, 1991). HATs and HDACs are usually functional at multiple residues on different histones (Table 1.1). Gcn5 can acetylate H3K9/K14/K18/K23/K27 and H2BK16 while Esa1 can acetylate for H4K5/8/12/16 and H2AK5/K8. Histone methyltransferases (KMTs) transfer methyl groups onto histones and demethylases (KDMs) remove methylations. Interestingly, KMTs and KDMs are seemingly much more residue-specific than HATs and HDACs. In yeast, four KMTs are identified, Set1, Set2, Set5 and Dot1, which are responsible for H3K4, H3K36, H4K5,K8,K12 and H3K79 separately (Table 1.1). Although the residues are specific, these KMTs can catalyze mono-, di- and tri- methylation on the same residue. Histone methylations are also dynamically regulated but KDMs in yeast were identified more recently (Liang et al., 2007; Seward et al., 2007; Tsukada et al., 2006; Tu et al., 2007).

Combinatorial histone modifications

Because of the many different PTMs identified on histones, it was naturally tempting to propose the hypothesis that histone modifications in different spatiotemporal combinations could be recognized by combination-specific “readers” and promote different downstream effects, or “Histone code” theory (Strahl and Allis, 2000; Turner, 2000). Many discoveries were made inspired by different aspects of this hypothesis, like the interaction between histone PTMs, covalent PTMs binding, combinatorial histone mutants, genomic localization patterns and sequential histone modification pathways. I will briefly introduce some examples of the first three aspects but focus primarily on spatiotemporal analysis of histone modifications.

Modified histone residues can display “cross-talk”. One modification can either promote or impede another modification. For example, H2BK123 monoubiquitination is required for both H3K4 and H3K79 methylation through Set1 and Dot1 methyltransferases (Briggs et al., 2002; Dover et al., 2002; Ng et al., 2002; Sun and Allis, 2002; Wood et al., 2003). H3K14 and H3K4 may be able to communicate with each other given the evidences that H3K14A mutant abolishes formation of H3K4me3 (Nakanishi et al., 2008) and furthermore, H3K4me3 recruits the NuA3 complex to facilitate H3K14ac (Taverna et al., 2006). On the other hand, asymmetrically dimethylated H3R2 negatively regulates H3K4me3 by excluding the binding of Spp1, a subunit of COMPASS (Set1C), which is required for trimethylation (Kirmizis et al., 2007). H3K4 may also be acetylated, and this acetylation is limited by methylation at the same residue (Guillemette et al., 2011). Fpr4 isomerization of H3 proline 38, a noncovalent modification, regulates Set2 methylation of H3K36 (Nelson et al., 2006). It will be very interesting to examine the physiological consequences of these specific pairs of histone modifications.

Multivalent binding may be important to connect different histone PTMs or interpret specific “histone codes”. As discussed above, different histone PTMs are recognized by different protein domains. Actually some proteins or complexes contain multiple domains and therefore obtain the

capacity to bind combinatorial histone PTMs. The human TAF1 has two bromodomains which recognize H4K5acK12ac and the binding affinity decreases when only one lysine is acetylated (Jacobson et al., 2000). Furthermore, peptide arrays containing combinatorial histone modifications were developed to examine specificity of histone PTMs antibodies and reader proteins (Fuchs et al., 2011). A few mammalian proteins, such as RAG2, BPTF and CHD1 recognize H3K4me3 specifically, and binding is affected by neighboring modifications.

To directly interrogate the importance of specific histone PTMs, histone mutants or mutants of histone-interacting protein can be examined in perturbation of transcription. Given such an abundant pool of histone PTMs, it is not surprising that individual mutants rarely display strong effects. For example, both H3K4 mutant and the *set1Δ* mutant cause surprisingly little perturbation of genome-wide gene expression, suggesting significant redundancy in the system (Guillemette et al., 2011; Lenstra et al., 2011). Gcn5 is generally localized to almost all active promoters (Robert et al., 2004) but the *gcn5Δ* mutant only alters the expression levels of a subset of genes (Lee et al., 2000), suggesting redundancy of histone acetylation. Combinatorial histone mutants were examined for expression and other phenotypes. Acetylated lysines on H3 or H4 tails were systematically mutated and a simple cumulative effect on growth and transcription was observed except H4K16, which regulates a unique set of genes (Dion et al., 2005; Martin et al., 2004). However, combinatorial mutants of methylated lysines and acetylated lysines on H3 show more complex patterns (Jin et al., 2009). Altered expression of many genes in the H3K[AcMe]G mutant (H3K4,9,14,18,23,27,36,79G) was simply due to the additive effect of the H3K[Ac]G (K9,14,18,23,27G) and the H3K[Me]G (K4,36,79G) mutants. However, a few other genes were differentially regulated only in the H3K[Me]G or H3K[Ac]G mutants. For these genes, H3K[AcMe]G mutant either causes no more effect than H3K[Me]G or H3K[Ac]G only or brings the expression levels back to WT level. Therefore, H3 acetylated lysines and methylated lysines can contribute to gene expression in both an additive and interacting manner. It will be very

interesting to test how complex the combinations can be by extending the types of PTMs and residues across different histone proteins.

Spatial distribution of histone modifications and modifiers

From this point, I will discuss the complexity of combinatorial histone PTMs from spatial and temporal patterns. The technology of chromatin immunoprecipitation (ChIP) with antibodies specific to modified peptides derived from histones enables detection of *in vivo* levels of various histone PTMs at particular locations. In recent years, ChIP followed by DNA microarray (ChIP-chip) or deep sequencing (ChIP-seq) has allowed the interrogation of genome-wide patterns of histone PTMs. There are limitations of ChIP that would lead to misinterpretation. First, ChIP depends on the quality of antibodies such that antibodies with low specificity would cause misinterpretation such as false positive behaviors of other modifications. Second, sites that are modified but occluded by a reader might falsely suggest that no modification is occurring in a region. Third, adjacent modifications may block antibody access to the residue of interest.

Most of the earlier genome-wide studies reported a static map from actively growing cells (log phase) in rich medium. These studies consistently show that histone modifications are enriched at specific genome features (Millar and Grunstein, 2006; Rando, 2007). Many histone acetylations like H3K9/14/18ac, H4K5/12ac, H2AK7ac, are located at the promoter and the 5' portion of the open reading frame (ORF), often called the +1 nucleosome, and positively correlated with transcription. H4K16ac and H2BK16ac are located in the coding regions and are uncorrelated or anti-correlated with transcription. H4K16ac is also hypoacetylated near telomere regions where genes are silenced. H2AZ and H2AZK14ac are both localized to the promoter regions; the former is anti-correlated with transcription but the latter is positively correlated. H3K56ac is found at some promoters like histone gene promoters (Xu et al., 2005) but it is also strongly associated

with novel nucleosome incorporation in DNA replication and DNA damage and repair hotspots (Chen et al., 2008; Kaplan et al., 2008). Another PTM which is associated with DNA damage hotspots is H2AS129ph (Szilard et al., 2010). H3K4me3 is enriched at 5' end of genes while H3K4me2 and me1 are located at the middle and 3' end of coding regions. H3K36me3 and me2 are enriched within ORFs. However only tri-methylation is correlated with transcription level; di-methylation is not. H3K79me3 is also enriched in the ORF regions but it has little correlation with transcription. Some novel PTMs such as H3K42me2 and H3K4ac are also located at the 5' end of genes and they both regulate transcription (Guillemette et al., 2011; Hyland et al., 2011). H2AQ105 in yeast is methylated by Nop1 and it is enriched specifically at the 35S rDNA locus, suggesting its association with Pol I transcription. The Q105A mutant reduces histone incorporation and increase rDNA transcription (Tessarz et al., 2013).

Given so many different combinations of histone PTMs differentially distributed across the genome, it seems quite reasonable to evaluate the complexity of the spatial combinatorial pattern of PTMs, and whether specific PTM combinations generate specific biological outputs. The Grunstein group examined 11 different acetylation sites on all histones by CHIP-chip and identified 53 different clusters of patterns of acetylation at promoter regions although many of the acetylations were highly correlated (Kurdistani et al., 2004). More interestingly, genes in certain of these clusters are co-expressed across many different conditions. A higher spatial resolution study was done by Liu et al. and 12 different acetylations and methylations were examined, mainly on chromosome III (Liu et al., 2005). By principle component analysis, they found that two components explained most of the variance among the 12 modifications. The first component is the overall level of histone modification and the second component is the relative levels of two groups of histone modifications: positive correlated modifications over coding regions and hypo-acetyl nucleosomes at TSS. Similar analysis and outcomes were recaptured in higher eukaryotes (Bernstein et al., 2005). Recently, with increasing CHIP-seq datasets, the idea of combinatorial

histone modifications has been revisited by updated computational methods. For example, Cui et al. analyzed the dataset generated in Rando's lab and found two major groups of PTMs, transcription enhancing modifications (H2AK7ac, H3K14ac, H3K9ac, H4K5ac, H3K18ac, H4K12ac and H3K4me3) and transcription repressing/insignificant modifications (H3K4me1, H2BK16ac, H3K4me2, H4K16ac and H4K8ac) (Cui et al., 2011). Then they constructed Bayesian networks to reveal the causal relationships among histone PTMs and 23 clusters were identified. Some of these had been experimentally validated previously, such as H3K14ac is required for H3K4me3. Lai and Buck integrated data from 12 histone PTMs, H3, H2AZ and nucleosome ChIP-chip or ChIP-seq datasets and identified 18 clusters of nucleosome types in 6 different major groups (Lai and Buck, 2013). Interestingly, the 6 groups not only differed in histone PTM composition, but also exhibited different biological functions. They distinguish the promoters recognized by Pol II and Pol III, coding regions with different expression levels, and other genomic features such as telomeres and double strand break hotspots. Clearly, the spatial combinatorial complexity is much lower than the potential combinatorial complexity, probably because of the prevalence of correlation between histone modification types.

The genome-wide location of chromatin modifiers has also been studied in yeast. HATs such as Gcn5 and Esa1 are generally recruited to promoter regions of actively transcribed genes and are recruited to inactive genes upon activation (Robert et al., 2004). Occupancy of Gcn5 and Esa1 correlates with transcription levels. Sas3 occupancy is also positively correlated with transcription rate but others, like Sas2, Elp3, Nut1 and Hpa2 are not (Rosaleny et al., 2007). Gcn5 and Esa1 are also associated with the coding regions to facilitate nucleosome eviction in elongation (Ginsburg et al., 2009; Govind et al., 2007). Gcn5 is also re-localized to the coding regions of long genes during stress response (Xue-Franzén et al., 2013). In contrast to the general roles of HATs, HDACs function in different classes of genes. Hst1 is enriched in the promoter regions of mid-sporulation and kynurenine (*de novo* nicotinamide) pathway genes for Hst1 but Rpd3 is enriched

at cell cycle genes (Robert et al., 2004). Rpd3 is complex and comes in at least two forms called Rpd3S and Rpd3L. Rpd3L is also actively recruited to the promoters of ribosomal protein genes upon cold shock to repress transcription. Rpd3L also positively regulates expression of specific genes including heat-shock responsive genes, DNA-damage induced genes and *FLO11* (Barrales et al., 2012; Ruiz-Roig et al., 2010; Sharma et al., 2007). Hos2, another HDAC, is located in the ORF region and required for gene activation (Wang et al., 2002). Rpd3S (contains Rpd3) and Set3C (contains Hos2 and Hst1), two HDAC complexes, are recruited to the coding regions by Pol II CTD (C-terminal domain) and histone methylation (Carrozza et al., 2005; Kim et al., 2012). A recent study by Li et al. examined genome-wide targets of sirtuins, NAD⁺ dependent HDACs (Li et al., 2013). They confirmed known targets like silent mating-type loci, rDNA repeats, telomere and tRNA genes and revealed new targets like the ORF regions of diauxic shift responsive genes and condensin/cohesin binding sites. Set1 is recruited to the 5' end of ORF regions by phosphorylated serine 5 on Pol II CTD and occupancy is correlated with transcription level (Ng et al., 2003). Set2 is recruited to the coding regions by Ser2-phosphorylated CTD and also correlates with transcription (Krogan et al., 2003; Li et al., 2003).

Relative to histone PTMs themselves, histone modifiers are seemingly localized more flexibly across transcription units. Set1 and Set2 can move through coding regions, presumably coincident with RNAPII movement/interaction and HATs/HDACs are also recruited to both promoter and ORF regions. This reflects the multiple ways in which these modifiers interact with histones and the diverse roles of histone modifiers during transcription.

Temporal pattern of histone modifications and modifiers

Besides the spatial pattern of histone modifications and modifiers, the temporal pattern is also very important for understanding the functions. The intact transcription cycle can be divided into

several stages, repression-initiation-elongation-termination-repression, and histone modifications and modifiers are seemingly involved kinetically in all of the stages (Smolle and Workman, 2013). Sequential appearance/disappearance of modifications and association/disassociation of modifiers, or histone modification pathways change dynamically, consistent with the idea that they actually drive the transitions between different transcriptional phases. To study the dynamic property of histone PTMs in transcription, an experimental setting in which gene expression is switched from repression to activation or the reverse order would be very informative. In yeast, there are many classic scenarios to exploit, such as nutrient conditions, heat shock and other stress treatments. Below I will summarize the known behavior of some specific and common histone modification pathways at model genes or the genome (Fig 1.2).

PHO5

PHO5, encoding an acid phosphatase, is repressed in high phosphate condition and activated in low phosphate condition (Svaren and Hörz, 1997). Two transcription factors, Pho2 and Pho4, drive the activation process. Besides the trans-factors, *PHO5* is a very early example showing the importance of nucleosomes in transcription. There are 4 well-positioned nucleosomes at the promoter region of *PHO5* and they are disrupted upon activation by the remodeling activity of SWI/SNF complex. Disruption of the SWI/SNF complex delays the kinetics of nucleosome depletion at *PHO5* promoter region (Reinke and Hörz, 2003). The histone chaperone Asf1 is also required for chromatin disassembly upon induction of *PHO5* (Adkins et al., 2004). Instead of chromatin remodeling complex and histone chaperones, here we will focus on histone PTMs.

Under repressive condition, Esa1/NuA4 is recruited to the promoter of *PHO5* in a Pho2-dependent manner and acetylates H4 tails (Nourani et al., 2004). Hyperacetylation of H4 itself is not sufficient to induce *PHO5* but it “presets” the *PHO5* promoter for later activation. H4 tail mutants or *esa1* ts mutants significantly reduce the rate and level of *PHO5* activation. However,

after induction, Esa1 is dispensable and the binding of NuA4 is significantly decreased upon activation. The hyperacetylation status of H4 during repression probably allows the promoter to be relatively accessible for TFs and remodeling complexes in subsequent activation. The Rpd3L complex is recruited to the *PHO5* promoter before induction in a Set1-mediated H3K4me dependent manner and keeps H3 hypoacetylated under repression (Wang et al., 2011). Deletion of *RPD3* or *SET1* derepresses or increases the induction rate of *PHO5*. Upon induction, Gcn5/SAGA is recruited to the promoter and acetylates H3, which facilitates nucleosome remodeling and activation (Barbaric et al., 2003; Reinke and Hörz, 2003). The acetyl groups are rapidly removed along with the eviction of nucleosomes. Gcn5 deletion doesn't affect the final activation level of *PHO5* but the mutant significantly decreases the rate of activation (Barbaric et al., 2001). H3K56ac also increases after induction and functions in promoting chromatin disassembly (Williams et al., 2008). H3K56R mutant or *asf1Δ* delays the chromatin disassembly and decreases the activation level of *PHO5*. Together, these findings reveal a fine-tuned temporal orchestration of histone acetylations, methylation and modifiers from repression to activation stages of *PHO5*.

CUP1

CUP1 encodes a metallothionein, which binds Cu⁺⁺ ions and remove this ion from the cell. *CUP1* is activated upon addition of Cu⁺⁺ (Rutherford and Bird, 2004). Interestingly, *CUP1* gene is naturally duplicated in many copies to offer sufficient resistance to copper (Fogel and Welch, 1982). Nucleosomes are also involved in its activation. Instead of the “eviction” seen at *PHO5* gene, nucleosomes are re-localized from clusters of specific positions to distribution across the entire gene (Shen et al., 2001). The key activator, Ace1, binds copper and becomes active. It is recruited to the promoter, followed by remodeling factors Ino80 and Snf2, leading to nucleosome rearrangement (Wimalarathna et al., 2013). When enough metallothionein is generated and

sequesters free copper, Ace1 is competed for copper binding and as it loses copper, *CUP1* expression is reduced. During the process, histone PTMs are dynamically regulated.

Gcn5 is not required for *CUP1* activation but deletion of another HAT, *SPT10* decreases the maximum level of *CUP1* expression (Shen et al., 2002). Histone acetylations are observed during activation but they depend on nucleosome remodeling and transcription machinery assembly, suggesting that they are not critical in early initiation. More accurately, H3K9/14ac appear within 10 min of copper induction, together with the peak of *CUP1* expression (Kuo et al., 2005). They also decrease together with the down regulation of *CUP1* expression. On the other hand, H4 acetylation decreases at the first 10 min but then significantly increases at 30 min when *CUP1* is tuned down. This change is another example showing different functions of H3 and H4 acetylations during transcription. Why the temporal pattern of H3/H4ac occurs at *CUP1* is not fully understood, but *Spt10* was shown to regulate both activation and inactivation of *CUP1*. Synthetic mutations of H2A and *SPT10* inhibit down-regulation of *CUP1*. Additionally, H2AS129 phosphorylation at the promoter region is coincident with down-regulation of *CUP1*. Together, the dynamics of histone PTMs coincide with the auto-regulation of *CUP1*, probably contributing to the fine-tuned level of expression corresponding to the copper level. Interestingly, *CUP1* is also induced by heat shock through Hsf1. After a 20 min heat shock, both H3 and H4 acetylations increase at *CUP1* (Deckert and Struhl, 2001). A high temporal resolution time-course experiment may be necessary to decipher whether the dynamics of H4 acetylation differ upon different induction conditions. It is interesting to test the idea that the dynamics of histone PTMs are differentially regulated at the same gene by different regulators.

MET16

MET16 encodes a 3'-phosphoadenylylsulfate reductase which is involved in sulfate assimilation and methionine biosynthesis (Morillon et al., 2005). It has basal transcription level when

methionine is available but it is fully activated under methionine starvation. Besides TFs like Cbf1, Met4 and Met28 (Kuras et al., 1996), chromatin modifiers and remodeling factors also regulate its induction. For example, the basal level of transcription doesn't change in *gcn5Δ* or *ada2Δ* strains but the induced level decrease a lot in the two mutants (Ferreiro et al., 2006). Isw1 modulates 5' end chromatin structure of *MET16* in a Set1-mediated H3K4me manner (Santos-Rosa et al., 2003). Deletion of *ISWI*, a chromatin-remodeling factor, increases the expression rate. Fpr4-mediated proline isomerization inhibits the level of H3K36me3 at inactive *MET16* and *FPR4* mutants delay *MET16* induction (Nelson et al., 2006). The Mellor lab characterized the dynamics of transcription, Pol II and histone modifications at very high temporal resolution, which will be discussed below (Morillon et al., 2005).

MET16 RNA is detected 10 min after induction and becomes stable at 20-30 min while nucleosome remodeling occurs within the first 7 min. Pol II binding signal appears at the promoter 0.5 min upon induction, disappears and appears again at 5 min. The signal starts to be detected in the coding region after 7 min, therefore dividing the entire activation into two stages, regulatory phase within the first 7 min and elongation phase after. H3K4me1 is detected under repression condition and drops quickly upon induction. Simultaneously, H3K4me3 appears at the promoter and 5' end of *MET1* from 0.5 min of induction, together with H3K36me3 and me2. Instead of the relative persistence of H3K4me3 and H3K36me2, H3K36me3 ChIP signals drops rapidly 3 min after induction. H3K36me3 increases again when entering into the elongation phase but the localization is shifted from promoter region to the 3' end of ORF. A similar pattern holds true for H3K36me2. H3K4me2 and H3K79me2 signals increases also only in the elongation phase after 10 min of induction. Besides methylations, Esa1 and Esa1-mediated H4K8ac both appear immediately after induction, coincident with H3K4me3 and H3K36me2/3 peaks, and then decrease quickly in the elongation phase. These transient signals depend on both Set1 and Set2 functions. Esa1 is required for *MET1* activation by antagonizing the negative function of Isw1.

GALI

GALI is a galactose responsive gene which encodes a galactokinase. It is suppressed under glucose condition and activated in the presence of galactose. Because of its very high induction ratio, this gene has been widely used to study transcription regulation by transcription factors, chromatin remodeling, histone modifiers and modifications. Given the well-understood mechanisms and easy manipulation, it has also been used a lot as a molecular tool for ectopic expression of various genes. Here we will focus on the histone modification pathway which occurs at the *GALI* locus.

Upon galactose induction, Rad6, an E2 ubiquitin-conjugating enzyme, is recruited to the *GALI* promoter in 30 min, and ubiquitylates H2BK123 immediately together with Bre1, the E3 ligase (Henry et al., 2003; Kao et al., 2004). The binding and activity of Rad6/Bre1 depend on a series of events, Ser5 phosphorylation of Pol II CTD and recruitment of Paf1 and Bur1/Bur2 complex (Larabee et al., 2005; Ng et al., 2003; Wood et al., 2005; Wood et al., 2003; Xiao et al., 2005). H2BK123ub causes subsequent methylation at H3K4 and H3K79. SAGA is also recruited to the promoter, slightly after the binding of Rad6, and it remains bound when Rad6 is disassociated. The binding of SAGA, which is independent of H2B ubiquitylation, is critical for the subsequent binding of TBP at *GALI* promoter and transcription. It is also required for the deubiquitylation of H2BK123 through its component, Ubp8, the ubiquitin-specific protease. The deubiquitylation enables the Ser2 phosphorylation of Pol II CTD by Ctk1 kinase, followed by Set2 recruitment and H3K36 methylation. Additionally, H3 and H4 acetylations behave differently at *GALI* locus. H4 acetylation decreases about 3 fold after 8 hours of galactose induction but H3 acetylation is not affected at this time point (Deckert and Struhl, 2001).

The apparent sequential order of histone modifications at *GALI* actually reflects the key role of Pol II CTD during transcription regulation. The YSPTSPS peptide on Pol II is conserved through eukaryotes although the number of repeats varies. The alternative phosphorylation on serine 5 and 2 play a critical role in coordinating different events through initiation, elongation and termination. A lot of factors, besides histone modifiers discussed above, are recruited by specific CTD phosphorylation status, and fulfill particular roles at different stages of transcription. An interesting idea of the Pol II CTD code was proposed and discussed in the cited reviews (Buratowski, 2003; Venters and Pugh, 2009).

Genome-wide analysis of dynamic histone modifications

In recent years, high-throughput technologies such as microarray and next-generation sequencing have been widely applied in chromatin biology research. These tools enable scientists to extend their research targets from single genes to the genome, which greatly facilitates the study of diversity and generality of dynamic histone modifications and modifiers. Below I will introduce some genome-wide studies in yeast.

An early study by Iyer's lab utilized RNA microarray and CHIP-chip to analyze dynamics of transcription, nucleosome occupancy and H4 acetylation under two common stress conditions, heat shock and stationary-phase stress (Shivaswamy and Iyer, 2008). Histone acetylation is usually positively correlated with transcription. However, their study displayed distinct temporal patterns of H4 acetylation at the same set of genes under different stress conditions. Ribosomal proteins (RP) genes are repressed both in stationary phase and with heat shock. Surprisingly, H4 acetylation level at RP genes is decreased after 15 min heat shock but increased in stationary phase. H4K5, K8 and K12 acetylations but not K16 acetylation contributed to the increment in stationary phase. In contrast, H3K9/18ac at RP genes was reduced from log phase to stationary

phase, suggesting that increased acetylation is not uniform on all histones. Caution is warranted in interpreting the contrasting phenomena, given that heat shock is a transient stress treatment while stationary phase is a relatively long-term and slow process. Although the time scales are distinct, the observation provides an example of versatile dynamics of histone acetylation on different histones or different residues at the cluster of RP genes.

Another study was performed to examine the dynamics of SAGA and TFIID in response to DNA damage (Ghosh and Pugh, 2011). Yeast genes can be roughly divided into two classes, TATA-containing genes, which are dependent on SAGA, and TATA-less genes dependent only on TFIID. SAGA and TFIID are not mutually exclusive but collaborate on transcription of many genes. During MMS (a DNA damaging molecule) treatment, SAGA is recruited to MMS-induced genes first (< 1 h), followed by decrease of SAGA binding and increased TFIID recruitment (2 h). This sequential pattern is different at heat shock activated genes, where SAGA and TFIID are concomitantly recruited immediately after heat shock (5-15 min). The discovery of distinct timing patterns suggests that yeast uses distinct mechanistic strategies for gene activation upon different stresses.

Sporulation is a developmental process when yeast switches from active growth to starvation and undergoes meiosis and gametogenesis. Transcription is reprogrammed to repress growth genes and induce meiosis-specific genes. Zhang et al. (Zhang et al., 2011) examined the dynamics of chromatin structures during sporulation. Although the nucleosome free region and the majority of nucleosome positions don't change much during sporulation, the positions of +1 nucleosomes are dynamically regulated, probably associated with transcription of genes sensitive to carbon starvation. By analyzing average modification levels across all genes, Zhang et al. showed that H3K9ac is rapidly decreased upon starvation but H3 methylations are relatively stable at the beginning and reduce globally later. H3K4me3 and H3K36me3 also positively correlate with

transcription of meiosis-specific genes. A significant percent of meiosis-induced genes contain antisense transcripts, which repress their expression in rich media. A chromatin signature of that is enriched H3K4me3 at 3' of genes instead of 5'. During sporulation, the position of H3K4me3 enrichment is shifted from the 3' to the 5' end of these meiosis specific genes.

Besides mapping the dynamics of histone modifications or histone modifiers, analyzing the kinetics of transcription in various mutants of histone PTMs or modifiers is also important in understanding the roles of chromatin factors because some mutations don't affect the end points of transcription level but the induction kinetics. Weiner et al. systematically investigated the dynamics of 170 transcripts in 83 histone mutants and 119 deletion mutants of chromatin regulators upon diamide stress in yeast (Weiner et al., 2012). The kinetics of transcription is more sensitive to chromatin mutations than the end points of transcription. Mutants which exhibit similar patterns of transcription induction/repression were shown to be in the same complexes or pathways or potentially interact with each other. Together with time-course analysis of histone modifications, they found that *set1Δ* mutants inhibit the repression of RP and ribosomal biogenesis (ribi) genes. Repression of ribi genes is associated with increased H3S10ph and modest elevation of H3K4me3 but decrease of H3K14ac. Given the evidence that Rpd3L mutants also inhibit ribi repression, they suggested that ribi repression is regulated by Set1-mediated H3K4me3 followed by Rpd3L recruitment and function. In contrast, repression of RP genes is associated with strong elevation of both H3K4me3 and H3S10ph but no H3K14 deacetylation, and Rpd3L mutants has no effects on RP gene repression, suggesting that repression of RP genes is mediated through another pathway which is also dependent on Set1.

Summary

Histone modifications play an essential role in transcriptional regulation. Various modifications function cooperatively from both spatial and temporal perspectives, leading to complex but specific consequences across the entire transcription cycle. It should be emphasized that the temporal analysis of histone modifications reveals the importance of transient histone modifications and transient effects on transcription. Temporal studies have been done in other dynamic processes in mammalian systems, such as the circadian clock, T cell differentiation, cardiac development and so on (Koike et al., 2012; Paige et al., 2012; Wamstad et al., 2012; Zhang et al., 2012), which also illustrate the prevalence of sequential histone modifications in controlling these dynamic processes. Chapter 2 describes the dynamics of transcription, histone modifications and modifiers associated with distinct metabolic states in yeast, providing some insights into the temporal chromatin structure in regulation of the timely orchestration of transcription. Chapter 3 documents a nice example of how analysis of dynamic H3K56ac leads to the surprising discovery of meiosis-specific endonuclease Spo11 functioning in mitosis. Chapter 4 further exhibits an application of dynamic analysis of histone modifications. It describes a strategy of predicting the timing of transcription factor binding/activation through analyzing the timing of H3K9ac in the same process. Life lies in motion. The entire thesis aims to display the beauty of dynamics in histone modifications and beyond.

Modification	Residue	Writer	Eraser	Reader
Acetylation	H2AK5	Esa1	Rpd3	
	H2AK8	Esa1, Hat1	Rpd3	
	H2AZK3	Esa1		
	H2AZK8	Esa1		
	H2AZK10	Esa1		
	H2AZK14	Esa1, Gcn5		
	H2BK11	Esa1, Gcn5	Rpd3, Hda1	
	H2BK16	Gcn5, Esa1	Rpd3, Hda1	
	H3K4	Rtt109, Gcn5	Hst1, Sir2	
	H3K9	Gcn5, rtt109	Rpd3, Hos2, Hda1	Bromo
	H3K14	Gcn5, Hpa2, Esa1, Elp3, Sas2, Sas3	Rpd3, Hos2, Hda1	PHD, Bromo
	H3K18	Gcn5	Rpd3, Hos2, Hda1	
	H3K23	Gcn5, Sas3	Rpd3, Hos2, Hda1	
	H3K27	Gcn5, Rtt109	Rpd3, Hos2, Hda1	
	H3K36	Gcn5		
	H3K56	Rtt109	Hst3, Hst4	PH, Snf5
	H4K5	Esa1	Rpd3, Hos2	Bromo
	H4K8	Esa1, Elp3, Gcn5	Rpd3, Hos2	Bromo
	H4K12	Esa1, Hat1, Hpa2	Rpd3, Hos2	Bromo
	H4K16	Esa1, Sas2	Sir2, Hos2, Hst1	Bromo
	H4K20	Esa1, Sas2	Sir2, Hos2, Hst1	
Methylation	H3R2			
	H3K4	Set1	Jhd2	PHD, Chromo, WD40, ADD, Tudor, MBT, Zf-CW
	H3K36	Set2	Rph1, Jhd1, Gis1	Chromo, PHD, PWWP, Tudor
	H3K42			
	H3K79	Dot1		Tudor
	H4R3	Rmt1		Tudor, ADD
	H4K5	Set5		
	H4K8	Set5		
	H4K12	Set5		
	H4K31			

Phosphorylation	H2AS122			
	H2AT126			
	H2AS129	Mec1, Tel1	Pph3	BRCT
	H2BS10	Ste20		
	H3S10	Snf1, Ipl1	Glc7	Gcn5, 14-3-3
	H4S1	CK2		
Ubiquitylation	H2BK123	Rad6, Bre1	Ubp8, Ubp10	Cps35
	H4 N- terminal tail	Ubc9		
Sumoylation	H2AK126	Ubc9, Siz1, Siz2		
	H2BK6/K7	Ubc9, Siz1, Siz2		

Table 1.1 Common histone modifications in *S. cerevisiae* and the enzymes that add or remove them. Domains recognizing histone modifications are not limited to *S. cerevisiae*. The information for this table were collected from Rando and Winston (2012), Smolle *et al.* (2013) and Yun *et al.* (2011). Methylations on H4 K5, K8, K12 by Set1 was from Green *et al.* (2012) (Green *et al.*, 2012; Rando and Winston, 2012; Smolle and Workman, 2013; Yun *et al.*, 2011).

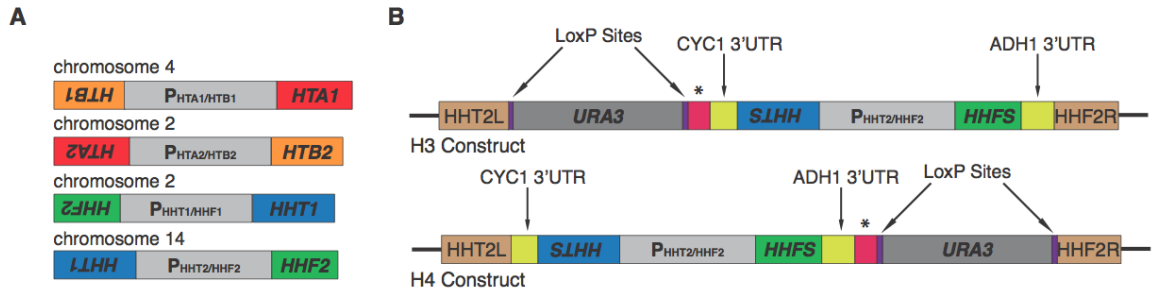


Figure 1.1 Native and synthetic histone genes in *S.cerevisiae*. (A) Divergent pairs of H2A, H2B, H3 and H4. (B) Synthetic histone H3 and H4 mutant construct from the collection by Dai *et al.* (2008). The constructs contain synthetic H3 and H4 with *HHT2/HHF2* promoter. The constructs also contain *CYC1* and *ADH1* 3'UTR, barcodes marked by “*”, a *URA3* gene flanked by LoxP sites and native *HHT2-HHF2* flanking regions.

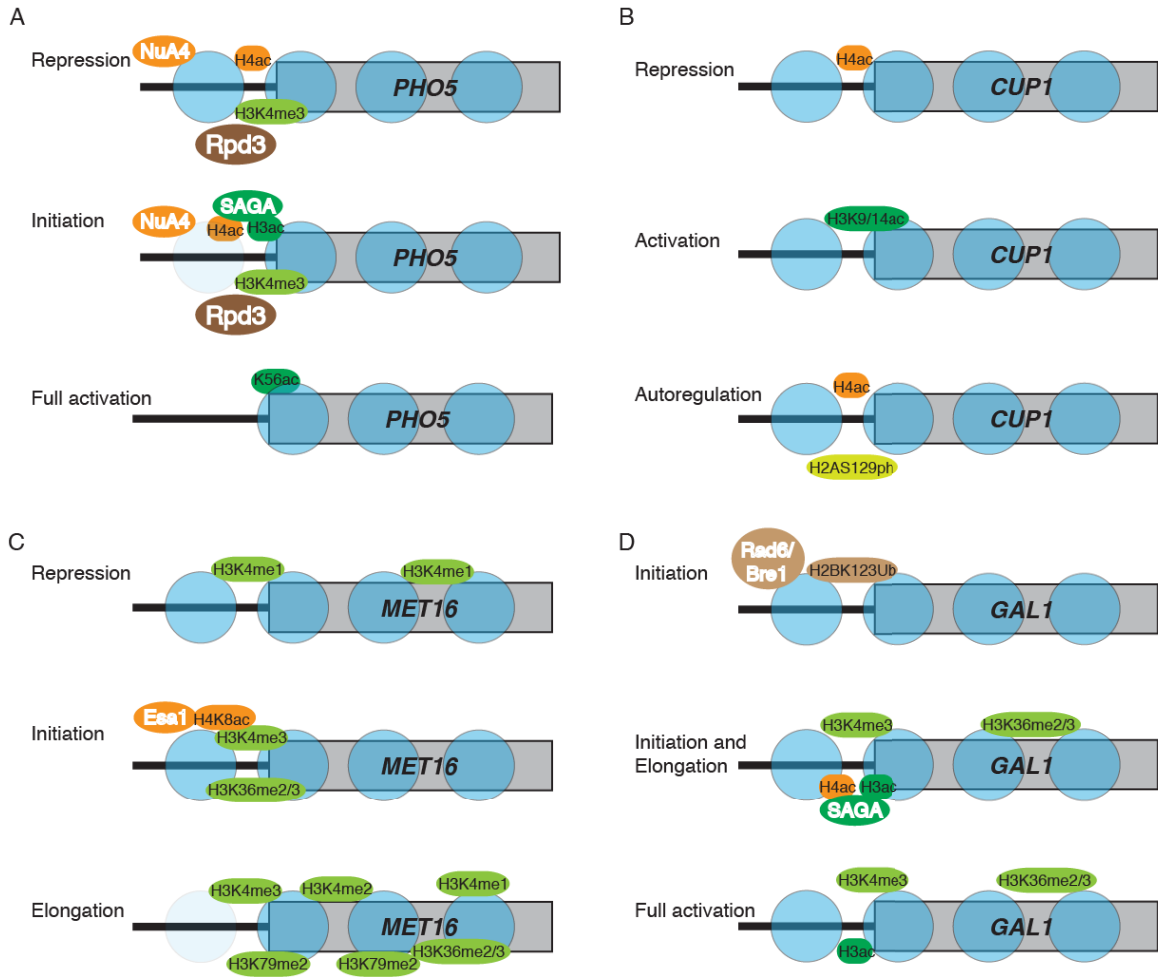


Figure 1.2 Dynamic histone modifications during activation of the yeast genes. *PHO5* (A), *CUP1* (B), *MET16* (C), *GAL1* (D). The black lines represent the promoter regions and the boxes represent the ORF regions. The blue circles represent the nucleosomes. From top to bottom are sequential stages of transcription activation but the timelines are not same for different genes (Barbaric et al., 2003; Henry et al., 2003; Kao et al., 2004; Kuo et al., 2005; Laribee et al., 2005; Morillon et al., 2005; Ng et al., 2003; Nourani et al., 2004; Reinke and Hörz, 2003; Wang et al., 2011; Williams et al., 2008; Wood et al., 2005; Wood et al., 2003; Xiao et al., 2005).

Chapter 2

**High temporal resolution view of transcription and chromatin states
across distinct metabolic states in budding yeast**

Abstract

Under continuous, glucose-limited conditions, budding yeast exhibit robust metabolic cycles associated with major oscillations of gene expression and metabolic state. However, how such fluctuations might be coordinately linked to changes in chromatin status is less well understood. Here, we examine the correlated genome-wide transcription and chromatin states across the yeast metabolic cycle (YMC) at unprecedented temporal resolution, revealing a “just in time supply chain” by which specific cellular processes such as ribosome biogenesis are coordinated in time with remarkable precision. We identify distinct chromatin and splicing patterns associated with different gene categories and determine the relative timing of chromatin modifications to maximal transcription. There is surprising variation in the chromatin modification/expression relationship, with acetylation peaks occurring with varying timing and “sharpness” relative to RNA expression both within and between cycle phases. Chromatin modifier occupancy reveals subtly distinct spatial and temporal patterns compared to the modifications themselves. Genetic analyses support a potentially cooperative role of histone modifications in the YMC.

Introduction

Chromatin plays fundamental roles in DNA-related processes, including transcription, replication, recombination and repair (Kouzarides, 2007). For example, most histone acetylations and certain methylations (eg. H3K4me3, H3K36me3) are correlated with active transcription while deacetylation and other methylations are correlated with repression (Berger, 2007; Pokholok et al., 2005). The pervasiveness of biological dynamics suggests that temporal interrogation of chromatin function is warranted. Compensatory and homeostatic effects almost certainly limit understanding of chromatin functions from steady-state studies (Weiner et al., 2012). The discovery of histone “readers” and “writers” further suggests the significance of dynamics of chromatin (Brownell et al., 1996; Millar and Grunstein, 2006), as do studies of bulk acetylation upon depletion of acetyl-CoA using a temperature-sensitive mutant (Takahashi et al., 2006). Examination of chromatin state and RNA expression level at single genes upon activation reveals many dynamic changes, such as the sequential appearance of H2BK123ub and H3K4me3 at *GALI* during galactose induction (Krebs, 2007). Genome-wide ChIP-seq studies reveal temporally dynamic chromatin patterns in circadian rhythm, heart development, yeast meiosis and other dynamic processes. These patterns have been correlated with transcription factor binding and RNA expression (Koike et al., 2012; Paige et al., 2012; Zhang et al., 2011) and support a generally dynamic role for chromatin in gene regulation. However, no prior studies have sufficient temporal resolution to dissect the specific roles of individual histone modifications during distinct steps in transcription *in vivo*, and the scope of genes regulated in these paradigms on a fast time scale has been limited.

Here we exploit a uniquely informative dynamical system, the yeast metabolic cycle (YMC), to examine regulation of >3,000 genes in *Saccharomyces cerevisiae* (Tu et al., 2005). Under continuous glucose-limited conditions, yeast exhibits a ~4-5 hour respiratory cycle with >3,000

transcripts and dozens of metabolites oscillating at the same pace (Tu et al., 2007). Although nearly all cycling genes share the same period of oscillation, the YMC can be broadly divided into three phases based on the defining expression profile: OX (oxidative), RB (reductive/building) and RC (reductive/charging). The overall logic of the YMC is well-defined: growth genes, such as ribosomal and amino acid biosynthesis genes are activated in OX; mitochondria and cell cycle genes are expressed in RB; and genes responding to starvation, stress and survival are elevated in RC. These divergent expression profiles imply that the dynamic interaction between chromatin modifications and transcription across distinct biological processes can be queried in a single system. Additionally, acetyl-CoA, a most dynamically oscillating metabolite, is a critical metabolic signal driving the YMC via histone (and perhaps other proteins) acetylation (Cai et al., 2011). Studies on acetyl-CoA, S-adenosylmethionine and other metabolites directly involved in histone modification indicate that chromatin could play key roles in coordinating metabolism and gene expression (Kaochar and Tu, 2012; Takahashi et al., 2006). Given such prevalent orchestration of expression and metabolites in the YMC, investigating the dynamic patterns of chromatin state could be useful for gaining a systems-level understanding of the YMC and beyond, such as how cells respond to a wide variety of nutrient and other biological stimuli through chromatin modifications. Discoveries made in the YMC could also extend our understanding of other cycling systems, such as circadian clocks.

Here, we reveal the genome-wide landscape of 7 histone modifications across the YMC and compare it to a high-resolution RNA-seq profile and dynamic localization of 3 critical chromatin modifiers, Gcn5p, Esa1p and Set1p. High sampling frequency allowed definition of clear-cut sub-phases of gene expression in each phase, such as the sequential activation of ribosome biogenesis (ribi) and ribosomal protein (RP) genes during the growth phase. We also defined 7 distinct chromatin patterns associated with genes of different functional classes, revealing the diversity of histone modification pathways during transcription. We also observed that binding of certain

chromatin modifiers does not always correlate with the appearance of the corresponding histone modifications, suggesting sophisticated temporal regulation. Inspired by the pervasiveness of the changes on the H3 and H4 N-termini during the YMC, we evaluated mutants in H3 and H4 modifiable lysines and showed that only certain multi-lysine mutants abolish the YMC, implying high system redundancy. Together, these data provide a comprehensive and high temporal resolution view of the dynamics of transcription and chromatin states. As such, they reveal how transcription and chromatin modifications are coordinated for different classes of genes during distinct phases of the life of a yeast cell.

Results

Gene expression patterns in the YMC

Microarray-based transcriptome analysis revealed three superclusters of genes defined by expression pattern in the YMC (Tu et al., 2005). Because of the short duration of the OX growth phase (~0.5 h), the previous sampling strategy used, (evenly-spaced every ~25 min and spread over 3 cycles) may have missed important transient regulatory events. To improve temporal resolution, we collected 16 samples across one cycle for RNA-seq by increasing the density of time points taken in the OX phase (Fig. 2.1A). The same three superclusters of genes were observed in excellent agreement with the prior data (Fig. 2.2A). We also identified an additional ~500 genes with a clear expression peak in the OX phase from the “non-periodic” genes defined previously (Fig. 2.2B). Gene Ontology (GO) analysis indicated enrichment of ribosome biogenesis genes. Many showed hints of periodic expression in the microarray dataset but evaded detection by the periodicity algorithm due to less frequent sampling.

Thanks to high temporal resolution, we resolved several temporal subclusters associated with distinct biological functions within previously defined OX, RB and RC super-clusters (Fig. 2.1B-

D). OX phase genes are subdivided into OX1, defining an earlier and more transient peak of expression compared to OX2 (Fig. 2.1B, Fig. 2.2C). OX1 contains genes associated with ribosome biogenesis (ribi) and rRNA processing. Ribosomal subunit (RP), amino acid metabolism (aa) and cytoplasmic translation genes are enriched in OX2, and peak ~5 min later. The subtle differences between ribi and RP genes were confirmed by RT-qPCRs (Fig. 2.2F). This observation suggests precise temporal control of the energetically-demanding process of ribosome biogenesis. Ribosomal assembly factors and RNA processing factors are expressed just before ribosomal proteins and translational factors, providing a possible example of “just-in-time” orchestration of a critical process in the YMC leading to increased translational capacity. Three temporal subclusters of RB and RC genes are also observed (Fig. 2.1C,D, Fig. 2.2C). Cell cycle and mitochondrial genes are readily distinguished although the two processes both occur under similar metabolic states. Collectively, the RNA-seq results not only recapitulate known gene expression patterns, but also display elaborate and fine-grained temporal control of coordinated sequential processes.

Accumulation of pre-mRNAs in the OX phase

Besides the temporal compartmentalization of gene expression, RNA-seq analysis also revealed an unexpected pattern in intron sequence accumulation. As discussed previously, RP genes are activated in OX, peaking at t5. Surprisingly, we observed a dynamic intron signal in several intron-containing RP genes consistently peaking at t4 (Fig. 2.3A). Several other intron-containing OX genes show similar patterns (Fig. 2.3A).

We evaluated RNA-seq reads across exon and intron regions separately for all intron-containing genes. Intron, exon and full-length signals were displayed in a heat map, identifying 6 clusters (Fig. 2.3B). Among them, only intron1 showed significant temporal enrichment of intron signals

in the OX phase. 89 of the 103 genes in intron1 are RP genes. Both the intron signals and the ratio of intron/exon signals peak at t4 (Fig. 2.3B and Fig. 2.4A).

During RNA splicing, intron 5' ends are cleaved and join to the branchpoint in the intron, forming a lariat. To distinguish intron accumulation as lariats or pre-mRNA, we examined exon-exon and intron-exon junctional reads of the 89 RP genes. As expected, exon-exon junctional reads are highest at t5, consistent with the timing of exon signals (Fig. 2.3C). However, 5' intron junctional reads, representing the pre-mRNA form, peaked at t4 (Fig. 2.3D). Additionally, we examined individual RP genes and discovered that the majority of the RP genes exhibit the same pattern (Fig. 2.4B-D). RT-qPCR of 4 RP genes show that the peaks of introns are consistently ahead of the peaks of exon signals by a mean time of 7.3 min (p value = 0.0286) (Fig. 2.4F).

Interestingly, t4 marks a strikingly reproducible inflection point in the O₂ curve of unknown functional significance occurring near mid-OX (Fig. 2.1A). These data suggest exquisite temporal control of RP gene splicing – RP genes appear to be transcribed at t4, but the corresponding mature spliced RP gene transcripts do not emerge until 5-7 min later at t5. Moreover, prior studies demonstrated accumulation of RP pre-mRNAs under stress conditions in yeast (Bergkessel et al., 2011; Pleiss et al., 2007). This delay in RP gene mRNA splicing in the YMC suggests that yeast might suffer a certain level of amino acid or other stresses during early OX (i.e., t4).

Alternatively, RP mRNAs may be spliced only on cue, following proper assembly of the ribosome biogenesis machinery. These may lead to a transient pause in O₂ consumption and the transient accumulation of RP pre-mRNAs precisely at this time point. Casein Kinase 2 complex (CK2) is required for RP gene pre-mRNA accumulation (Bergkessel et al., 2011); we observed peak expression of CK2 components at t4/t5 (Fig. 2.4E).

Dynamics of histone modifications across YMC

Given the transcriptome dynamics and close dependence of transcription on chromatin modification, we next investigated chromatin states across the YMC. A similar set of 16 samples across one cycle were collected for ChIP-seq (Fig. 2.5A) and 7 histone modifications were assessed, including H3K4me3, H3K36me3, H3K9ac, H3K14ac, H3K56ac, H4K5ac, H4K16ac, and H3 as control. Most of these are associated with active transcription (Kurdistani et al., 2004; Pokholok et al., 2005). H3K56ac also functions in DNA replication and repair (Avvakumov et al., 2011) and H4K16ac is involved in silencing maintenance and lifespan (Dang et al., 2009; Dion et al., 2005; Moazed, 2001). The two time series were computationally aligned by performing ChIP-seq for H3K9ac on both set of samples used for RNA and ChIP-seq (Methods and Materials).

Figure 2.5C shows the spatial and temporal signals at the *RMT2* locus, encoding an arginine methyltransferase for ribosomal protein Rpl12 (Chern et al., 2002). Its expression peaks in mid-OX phase (Fig. 2.5B). Mapping of ChIP-seq signals across *RMT2* reveals a typical spatial pattern (Millar and Grunstein, 2006) where H3K9ac, H3K14ac, H4K5ac and H3K4me3 signals peak in the 5' end whereas the H3K36me3 signal is located in the coding region (Fig. 2.5C). Although all positively correlated with transcription, the kinetics and relative timing are quite distinct at this locus. The H3K9ac signal is most dynamic, with a transient peak at t4. H3K14ac and H4K5ac signals also increase in OX with temporally broader peaks than H3K9ac. Instead, the H3K4me3 signal is much more stable than these acetylation marks, with clear-cut peaks across all 16 time points. It starts to increase only in late OX at a modest level and peaks in early RB, well after acetylation peaks. No other modifications were obviously dynamic at *RMT2*.

Expanding examination to broader regions revealed another unexpected difference between modifications. Most H3K9ac peaks are observed in OX, despite the presence of nearby RB and RC genes, consistent with previous study showing that H3K9ac appears preferentially at OX

growth genes (Fig. 2.6). RC/stress response genes are less dependent on H3K9ac (Cai et al., 2011). Instead, H3K14ac, H4K5ac and H3K4me3 peaks are found at most genes indiscriminately. Strikingly, H3K9ac peaks seemingly appear and disappear completely coupled with transcription of neighboring genes while H3K14ac, H4K5ac and H3K4me3 only display an adjustment of amplitude. The data suggests that H3K9ac is a very active modification that regulates mainly OX growth genes, possibly by directly sensing metabolic stimuli. Others, like H3K14ac, H4K5ac and H3K4me3, have a more general role in transcription regulation. For more examples see Fig. 2.7.

Next, we sought to determine average genome-wide chromatin patterns for various gene classes. First, we averaged the spatial signals of these modifications relative to the transcription start site (TSS) across the genome (Fig. 2.9A). Gene-specific signals were calculated based on their localization. A heat map was used to display the genome-wide interactive temporal patterns (Fig. 2.9B). Generally, acetylation signals were more dynamic and less noisy than methylation signals, consistent with acetylation shifts reflecting the immediate chromatin response under continuous glucose-limitation. Fourteen clusters were defined, suggesting complex dynamic relationships among these modifications across the YMC.

Intriguingly, H3K56ac exhibits very dynamic but noticeably different temporal patterns compared to other histone modifications. By comparing the results to previous work (Kaplan et al., 2008), we clearly observed cell cycle-related dynamics of H3K56ac (Fig. 2.8). H3K56ac is strongly elevated in early S phase; similarly, YMC signals peak at t8, the middle of RB phase, corresponding to S-phase entry. Correspondingly, DNA replication genes (RNA_RB1) are also expressed highest at t8 (Fig. 2.1C). The expression level of H3K56 acetyltransferase *RTT109*, peaks at t8 whereas RNA levels of H3K56 deacetylase *HST3*, begins to increase after t8 (Fig. 2.2D). These data corroborate known involvement of H3K56ac in DNA replication in RB and support the relationship between cell cycle and YMC.

Periodic transcription and histone modification are not systematically correlated

An important function for histone modifications is transcriptional regulation. With this in mind, we examined the genome-wide temporal relationships between each chromatin modification and transcription. Given the correlation analysis (Methods and Materials) and previous studies (Pokholok et al., 2005), we focused on five modifications most clearly linked to transcription (Fig. 2.10). We discovered 7 clusters of cycling genes (Fig. 2.10A). Are there different histone modification pathways associated with transcription of different gene clusters? To tackle this, we estimated relative timing of each modification peak relative to its corresponding RNA peak and examined GO enrichment on all clusters.

In clusters RNA_H1 and 2, the H3K9ac peak notably appears just minutes before gene expression peak (9/6.8 minutes in cluster 1/2) whereas the H3K4me3 and H3K36me3 peak 10-30 minutes after the peak of RNA (Fig. 2.10B, C). Surprisingly, H3K14ac and H4K5ac signals appear concomitantly with peak gene expression in RNA_H 2 (6.2 minutes of H3K14ac and 10.8 minutes of H4K5ac), but in RNA_H 1 they are advanced by one hour relative to RNA (48 minutes for H3K14ac and 75 minutes for H4K5ac). aa genes are enriched in RNA_H 1, whereas RNA_H 2 contains mostly ribi and RP genes. Combined with the RNA-seq analysis (Fig. 2.1), we could define three distinct chromatin-transcriptional “programs” among OX growth genes. Ribosome biogenesis (ribi) and ribosomal protein (RP) genes share the same chromatin pattern, distinct from that of aa genes. However, peak mRNA levels of RP genes and amino acid genes appear later and last longer than those of the ribi genes. As discussed earlier, regulated splicing of RP genes may cause a delay of the peak level of mature RP mRNA (Fig. 2.2).

Among the RB clusters, RNA_H 3 contains primarily cell cycle genes and exhibits similar chromatin patterns to OX clusters relative to RNA dynamics (Fig. 2.10A). In contrast, the genes

encoding mitochondrial proteins (RNA_H 4) behaved quite differently. Although they have a similar peak expression time, timing of the modification signals is out of phase. H3K14ac and H4K5ac increase in late RC and H3K9ac increases in late RC through OX. A motif sequence recognized by RNA-binding protein Puf3 was found in the 3' untranslated regions (3' UTR) of nearly all mitochondrial ribosomal genes (Tu et al., 2005). This discrepancy between RNA level and chromatin states is consistent with the hypothesis that this gene class may be subject to post-transcriptional rather than transcriptional regulation.

The quantification of the relative timing of chromatin signals summarizes different histone modification programs during transcription in yeast. In general, acetylation marks appear first, followed by RNA and then methylation marks, consistent with an overall initiation role for acetylation and an elongation role for methylation. Intriguingly, the fine-grained timing of different histone acetylation modifications could be a key factor to distinguish different transcription initiation mechanisms (Fig. 2.10D). In RNA_H 1, H4K5ac appears at first, followed by H3K9ac and then RNA. This general sequence suggests that H4K5ac functions at an early stage of transcription initiation in these sets of genes, perhaps to “pre-set” their activation, while H3K9ac, which peaks much more sharply, may play a critical role in triggering phase transitions, and may regulate a later initiation step or a transition to elongation. In RNA_H 2, H4K5ac and H3K9ac peak together immediately before RNA peak, suggesting dispensability of a pre-setting stage during transcription of these genes.

The distinct patterns between chromatin and gene expression raised the question whether specific transcription factors (TFs) are associated with these clusters. TF motif analysis reveals that binding sites for TFs involved in amino acid metabolism, like Gcn4p, Bas1p, Cbf1p, Leu3p, are enriched in RNA_H 1, whereas in RNA_H 2 we observe enrichment primarily of TF motifs related to RP gene expression, such as Stb3p, Sfp1p, Rap1p, Pbf1p, Pbf2p, Dot6p, Tod6p.

RNA_H 3 contains several common cell-cycle related TF motifs including those for Xbp1p, Mbp1p, Fkh1/2p, Swi4p. Many TF motifs associated with stress response are enriched in RNA_H 6, a typical RC cluster, such as Gis1p, Msn2/4p, Adr1p, Crz1p. Collectively, these results reveal a complex, dynamic and interactive regulatory network of chromatin states, TF binding sites, and post-transcriptional regulation.

Landscape of chromatin-modifying enzyme occupancy in the YMC

To further understand the dynamic patterns of histone modifications, we analyzed genome-wide occupancy of several chromatin-modifying enzymes in the YMC, specifically Gcn5p, Esa1p and Set1p. Gcn5p is the major histone H3 tail acetyltransferase and Esa1p acetylates H4 tails, although other HATs also acetylate H3 and H4. Set1p is the catalytic subunit of COMPASS complex H3K4 methyltransferase.

To simultaneously evaluate all three enzymes, we constructed a doubly-tagged GCN5-FLAG, SET1-3HA strain employing an anti-native Esa1p antibody. This allowed ChIP-seq of Gcn5p, Esa1p and Set1p in one set of samples (Fig. 2.11A). First, we examined spatiotemporal patterns of these enzymes by displaying ChIP-seq signals at sample regions; temporal patterns were quite distinct (Fig. 2.11B). All three enzymes are recruited to RP gene *RPS27A* (*YKL156W*) in OX but with subtly distinct kinetics. Gcn5p occupancy was detected only at t5, corresponding to t4 in the 16 time point ChIP-seq of histone modifications. However, Esa1p and Set1p signals preceded this peak, spanning mostly t2/t3, although Set1p occupancy lasted much longer, extending through early RB. Only Gcn5p and Set1p signals were detected at *GPM1*(RC) and *SDHI*(RB) loci, in a manner correlated with RNA levels (Fig. 2.11B). Interestingly, the signal of Set1p appeared 1-2 time points (5~20 minutes) before Gcn5p signal at *GPM1*, similar to *RPS27A*.

Given the complex spatiotemporal relationships between the three enzymes, we sought to assess the interactive patterns and their relationship to corresponding histone modifications and gene expression from a genome-wide perspective. We found 2369 Gcn5p peaks, 2225 Esa1 peaks and 2279 Set1p peaks, including 1874/1810/1629 with neighboring annotated genes (Fig. 2.12A). Among these, ~1750 genes are bound by at least two modifiers and 1035 genes are bound by all three modifiers, suggesting that these enzymes are not necessary for expression of every cycling gene. Spatial analysis shows that Gcn5p and Esa1p are localized 200-300 bp upstream of the TSS while Set1p is localized within the transcribed regions (Fig. 2.11C,D,E). This supports the well-accepted hypothesis that Gcn5p and Esa1p function in transcriptional initiation whereas Set1p is associated with elongation (Li et al., 2007). Intriguingly, the locations of the modifiers did not overlap completely with corresponding histone modifications. Although Gcn5p and Esa1p locate upstream of the TSS, H3K9ac and H4K5ac are shifted downstream. This suggests that Gcn5p and Esa1p can “reach across” the TSS to act on the downstream nucleosome. The Set1p signal is slightly shifted downstream of the H3K4me3 peak, presumably because Set1p also produces mono- and di-methylation on H3K4 during transcriptional elongation.

Next, we assessed temporal patterns of chromatin-modifiers, focusing on annotated genes associated with all three modifiers. We defined 4 superclusters of temporal patterns by K-means clustering (Fig. 2.12B). The results suggest that all three modifiers are functional in all three phases and recruitment of modifiers is roughly in phase with corresponding gene expression.

However, chromatin modifiers, chromatin states and gene expression are not simply correlated. Here we focused on cell growth genes, including ribi, RP and aa genes (Jorgensen et al., 2004). Ribi and RP genes individually exhibit one pattern while aa genes show three distinct patterns (Fig. 2.11F). Since RP genes show the cleanest signal, we examined relative timing within this group (Fig. 2.11G). As described above, H3K9ac and H4K5ac appear with the RNA peak

whereas H3K4me3 exhibits a delay. Surprisingly, Set1p is recruited ~10 minutes prior to the binding of Gcn5p, and remains bound until dissociation of Gcn5p. This pattern is consistent at *RPS27A*, *GPM1* (described above) and many other loci. The relative timing of histone modifications and modifiers at ribosome and aa genes were further confirmed by ChIP-qPCR in similar YMC experiment using Gcn5-3FLAG, 6HA-Set1 (Fig. 2.12D,E). The temporal discrepancy between modifiers and modifications suggests that Set1p is recruited prior to transcription in a state that is non-functional, at least for tri-methylation, until binding of Gcn5p occurs (Fig. 2.11H), whereas Esa1p is also recruited early but is immediately active.

Histone mutant analysis

Given the complex patterns of histone modifications and modifiers in YMC, we investigated the criticality of individual lysine residue modifications. Gcn5p mediates entry into the growth phase, as a mutant lacking it does not exhibit metabolic cycles (Cai et al., 2011). To broaden our understanding of histone modifications in the YMC, we examined the dO₂ oscillation phenotype of all histone H3 and H4 lysine mutants corresponding to residues at which acetylation or methylation have been reported. All single lysine to alanine or arginine mutants, including H3K4, H3K9, H3K14, H3K18, H3K23, H3K27, H3K36, H3K56, H3K79, H4K5, H4K8, H4K12, H4K16, exhibited normal cycles, suggesting that no single lysine modification on H3 or H4 is critical for the YMC (data not shown). Next, we combined lysine to arginine mutations in H3 or H4 tails, including H3K(9,14)R, H3K(9,14,18)R, H3K(9,14,18,23)R, H3K(9,14,18,23,27)R, H4K(5,8)R, H4K(5,8,12)R. Notably, we observed decreased growth on YPD with increasing numbers of lysine to arginine mutations, consistent with published reports (Fig. 2.13A) (Megee et al., 1995). As the number of K to R mutations increased, the amplitude of the dO₂ oscillation decreased (Fig. 2.13B-I). Strikingly, H3K(9,14,18,23,27)R, and H4K(5,8,12)R mutants abolished dO₂ oscillation. These results support the hypothesis that PTMs such as acetylation on the H3 and

H4 tails are required for normal cycling and support a role for acetyl-CoA in establishing the YMC, although there is no essential role in the YMC for any individual lysine residue.

Discussion

In this study, we describe the dynamic landscape of transcription and chromatin modifications across distinct metabolic states in yeast. We uncovered distinct combinatorial patterns associated with different functional groups of genes, emphasizing high-resolution timing of expression of distinct groups of genes and relative timing of histone modifications and modifiers within each group.

A “Just in time” supply chain?

Cellular processes are well organized spatiotemporally to maximize fitness especially in the face of limited nutrients. One interesting example is the sequential activation of promoters in bacterial aa biosynthesis pathways (Zaslaver et al., 2004). YMC transcriptional profiling reveals several examples of staged biosynthesis consistent with a “just in time” program. Ribi transcripts appear just before those encoding RP and aa (OX phase) genes (Fig. 2.1B), supporting a “just in time supply chain” that maximizes efficiency of ribosome biogenesis and translation. The short pulse of ribi genes may also play a critical role in helping them achieve an efficient ramp-down of protein synthesis at the end of OX. The observation of transient RP gene pre-mRNA accumulation suggests the possibility of elaborate control of splicing in regulation of ribosome biogenesis. Alternatively, this staging may reflect the fact that ribi genes can be much larger and therefore take longer to translate than RP proteins, which are universally short. The translation of the longest ribi gene might require 3-4 minutes more than the longest RP gene, consistent with the relative timing of ribi and RP expression. It is likely that the sequential activation of ribi and RP leads to concurrent synthesis of all ribosome biogenesis proteins. Likewise, cell cycle-related

genes and mitochondrial genes, the two major groups that have an elevated RNA level in the RB phase, also exhibit subtle timing disparities (Fig. 2.1C), perhaps indicating two distinct YMC subpopulations - one that will enter the CDC, and one which will not (Laxman et al., 2010). Collectively, the evidence suggests a complex and precise orchestration of macromolecular biosynthesis processes in yeast during the YMC from a transcriptional perspective. It will be interesting to examine whether protein factors are also organized temporally.

Different histone modification programs associated with distinct categories of genes

We defined seven different combinatorial clusters between histone modifications and gene expression, revealing several different regimes of chromatin states in transcription (Fig. 2.10A). Interestingly, different histone acetylation marks behave distinctly in association with different groups of genes. H3K9ac, H3K14ac and H4K5ac all appear coincident with the peak of gene expression of ribosomal genes. However, H3K14ac and H4K5ac are shifted earlier at aa (OX phase_aa2) genes (1 hour before the RNA peak). We hypothesize that H3K14ac/H4K5ac function in “pre-setting” the promoters, perhaps by chromatin remodeling. However, RP genes do not show evidence of this pre-activation phase from the perspective of chromatin states, perhaps because they are permanently pre-activated. RP genes could be immediate sensors of critical metabolic/environmental stimuli, e.g. intracellular acetyl-CoA levels. Conversely, aa2 genes appear to have a preparatory phase.

Dynamics of histone modifiers increase the complexity of chromatin states. In many cases, Set1p is recruited earlier than Gcn5p, in seeming contradiction to the order of appearance of H3K4me3 and H3K9ac (Fig. 2.11H). We hypothesize that Set1p may catalyze mono- or dimethylation in promoter regions to preset TSSs for activation. It is known that H3K4me1 accumulates at the *MET16* promoter prior to induction in yeast (Morillon et al., 2005). Additionally, H3K4me1/me2 precede transcription in cardiac development and T cell differentiation, suggesting that wiring of

at least some chromatin-based programs are phylogenetically conserve (Wamstad et al., 2012; Zhang et al., 2012). It may also be possible that the rapid on/off cycling rates of the modifications cause the contradiction.

Roles of histone modifications in the YMC

Various histone modifications may function “solo” or in a combinatorial or cumulative manner (Fischle et al., 2003; Strahl and Allis, 2000). These modifications either change the charge status of histone tails or provide platforms to recruit transcriptional regulatory complexes. The very different patterns between H3K9ac and H4K5ac suggest distinct functions during transcription. Histone mutant analysis reveals functional redundancy within each histone tail. The H3-5KR quintuple and H4-3KR triple mutant each totally abolished the oscillation, partially corroborating the evidence that the *gcn5Δ* strain does not support oscillation. Surprisingly however, H3-5KA and H4-3KA mutants grow better than corresponding KR mutants and still exhibit O₂ oscillation phenotypes (Fig. 2.14), suggesting that charge status of H3 and H4 may play a fundamental role in the YMC. Maintaining the positive charges on these histone tails may “lock” growth genes into a repressed state, resulting in inability to undergo bursts of growth characteristic of the YMC and cells in general.

Similar dynamics of chromatin states between YMC and mammalian circadian cycle

Remarkably, circadian rhythm in mammals and this yeast ultradian rhythm show several noteworthy similarities. Through systematic analysis of the transcriptome and chromatin state as a function of circadian clock in mouse liver, Koike *et al.* revealed similar temporal patterns of histone modifications (Koike et al., 2012), although the genes under regulation are completely distinct and the time scales of the cycles are very different, suggesting phylogenetically conserved

roles of histone modifications during transcription. More intriguingly, beyond the oscillatory phenomenon of transcriptome, metabolism and chromatin states in these two systems, the metabolism-epigenome-transcriptome loop is consistent in both systems, suggesting that the circular interactions among the three players actually sustain oscillations. Our findings on dynamics of chromatin modifications and temporal associations between epigenome and transcriptome suggest an active metabolism-epigenome-transcriptome loop in YMC. The observation that histone acetylation mutants abolish oscillatory respiration in yeast directly supports the indispensability of histone modification in sustaining metabolic cycles and by extension perhaps the circadian clock.

We have revealed a network of gene expression, chromatin state, and chromatin modifiers across the YMC as highly synchronized yeast cells transition between metabolic states akin to growth, division, and quiescence, at very high temporal resolution. Our study not only demonstrates how chromatin states are related to and directing the timing of gene expression, but also provides examples of how these processes are coordinated in a dynamic fashion to achieve optimal growth, metabolic efficiency, and presumably fitness. This yeast study may help illuminate the exquisite temporal control of various dynamic biological processes.

Methods and Materials

Media

YPD medium contains 1% yeast extract, 2% bacto-peptone, 2% dextrose and 1.6 mM tryptophan. YPD supplemented with 200 µg/ml G418 or 100 µg/ml ClonNat or 300 µg/ml Hygromycin were used for drug resistance selection. The YMC medium is a minimal medium consisting of 5 g/L $(\text{NH}_4)_2\text{SO}_4$, 2 g/L KH_2PO_4 , 0.5 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.1 g/L $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.02 g/L $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$,

0.01 g/L ZnSO₄•7H₂O, 0.005 g/L CuSO₄•5H₂O, 0.001 g/L MnCl₂•4H₂O, 1 g/L yeast extract, 10 g/L glucose, 0.5 mL/L 70% H₂SO₄, and 0.5 mL/L Antifoam 204 (Sigma).

Plasmids and histone mutagenesis

Plasmids carrying histone mutations were obtained from the histone mutant library described previously (Dai et al., 2008). pJD233 and pJD234, each containing a *hygMX* selectable marker cassette, were used for making *HHT1* and *HHF1* mutants respectively. pZK8 was generated from pJD154 by replacing the *URA3* cassette with a *natMX* cassette for mutagenesis on *HHT2* and *HHF2*. Plasmids carrying multiple lysine mutations were generated by site-directed mutagenic fusion PCR and subcloning in the indicated backbones, pJD233, pJD234 and pZK8. Briefly, primers carrying the desired mutations were used to generate two PCR fragments containing the overlapping mutation regions. Fusion PCR was used to generate a single fragment and the product was digested and ligated to the indicated backbones, pJD233, pJD234 and pZK8 and verified by DNA sequencing.

Strains and methods

All yeast strains were generated from the previously described strain CEN.PK and manipulated by standard budding yeast protocols.

Yeast strains used in this study:

Table 2.1. Yeast strains used in chapter 2.

Name	Background	Genotype
BY5764	CEN.PK	<i>MATa</i>
BY5765	CEN.PK	<i>MATa</i>
ZKY328	CEN.PK	<i>MATa GCN5-FLAG::natMX</i>
ZKY329	CEN.PK	<i>MATa SPT7-FLAG::natMX</i>
ZKY396	CEN.PK	<i>MATa SET1-3HA::KanMX6</i>
ZKY428	CEN.PK	<i>MATa GCN5-FLAG::natMX,SET1-3HA::KanMX6</i>
ZKY242	CEN.PK	<i>MATa HHT1-HHF1-WT::hygMX, HHT2-HHF2-WT::natMX</i>
ZKY340	CEN.PK	<i>MATa HHT1-HHF1-H3K9R::hygMX, HHT2-HHF2-</i>

		<i>H3K9R::natMX</i>
ZKY344	CEN.PK	<i>MATa HHT1-HHF1-H3K9,14R::hygMX, HHT2-HHF2-H3K9,14R::natMX</i>
ZKY348	CEN.PK	<i>MATa HHT1-HHF1-H3K9,14,18R::hygMX, HHT2-HHF2-H3K9,14,18R::natMX</i>
ZKY252	CEN.PK	<i>MATa HHT1-HHF1-H3K9,14,18,23R::hygMX, HHT2-HHF2-H3K9,14,18,23R::natMX</i>
ZKY356	CEN.PK	<i>MATa HHT1-HHF1-H3K9,14,18,23,27R::hygMX, HHT2-HHF2-H3K9,14,18,23,27R::natMX</i>
ZKY390	CEN.PK	<i>MATa HHT1-HHF1-H4K5R::hygMX, HHT2-HHF2-H4K5R::natMX</i>
ZKY391	CEN.PK	<i>MATa HHT1-HHF1-H4K5,8R::hygMX, HHT2-HHF2-H4K5,8R::natMX</i>
ZKY392	CEN.PK	<i>MATa HHT1-HHF1-H4K5,12R::hygMX, HHT2-HHF2-H4K5,12R::natMX</i>
ZKY388	CEN.PK	<i>MATa HHT1-HHF1-H4K5,8,12R::hygMX, HHT2-HHF2-H4K5,8,12R::natMX</i>
ZKY482	CEN.PK	<i>MATa HHT1-HHF1-H3K9,14,18,23,27A::hygMX, HHT2-HHF2-H3K9,14,18,23,27A::natMX</i>
ZKY484	CEN.PK	<i>MATa HHT1-HHF1-H3K9,14,18,23,27Q::hygMX, HHT2-HHF2-H3K9,14,18,23,27Q::natMX</i>
ZKY488	CEN.PK	<i>MATa HHT1-HHF1-H4K5,8,12A::hygMX, HHT2-HHF2-H4K5,8,12A::natMX</i>
ZKY490	CEN.PK	<i>MATa HHT1-HHF1-H4K5,8,12Q::hygMX, HHT2-HHF2-H4K5,8,12Q::natMX</i>
ZKY741	CEN.PK	<i>MATa 6HA-SET1 GCN5-3FLAG::KanMX6</i>

Gene knockout strains were generated by homologous recombination using PCR products containing a drug cassette (*kanMX6*) and 40 bp sequences flanking the target gene. Tagged-protein strains were generated similarly by integrating a cassette containing a protein tag and a drug resistance cassette at C terminus. Briefly, *MATa GCN5-FLAG::natMX* (ZKY328) was obtained from a previous study. A PCR product consisting of pFA6a-3HA-kanMX6 with 40 bp sequences flanking the *SET1* stop codon at both ends was transformed into *MATa* CEN.PK strain (BY5765) and was selected for G418 resistance. Integration of the 3HA tag and marker gene were validated by PCR using ~500 bp flanking primers. *MATa GCN5-FLAG::natMX* (ZKY328) and *MATa SET1-3HA::kanMX6* (ZKY396) were mated and diploids were sporulated and dissected to select for haploids containing both tagged-protein. 6HA-SET1 was generated by transforming a PCR fragment containing kanMX flanking by loxP sites and 6-HA and integrating after start codon. KanMX marker was removed by expressing Cre-Recombinase and confirmed

by colony PCR. Histone mutant strains were generated following the protocol described previously (Dai et al., 2008). All histone mutations were introduced into both the *HHT1-HHF1* and *HHT2-HHF2* loci prior to assaying phenotypes. *HHT1-HHF1* mutations were introduced into the genomic locus of the *MATa* strain (BY5764) and selected on YPD + 300 $\mu\text{g/ml}$ Hygromycin. *HHT2-HHF2* mutations were introduced into the genomic locus of the *MATa* strain (BY5765) and selected on YPD + 100 $\mu\text{g/ml}$ ClonNat. Haploid strains were mated and diploids were sporulated and dissected to select for haploids containing both *HHT1-HHF1*, *HHT2-HHF2* mutations. Fitness was examined on YPD plates at 30°C. Log-phase cultures were diluted to an OD of A_{600} of 0.1 and plated on the first row, and 10-fold serial dilutions were plated on the following rows. Photos were taken after 2 days.

Metabolic cycles

Metabolic cycle experiments were performed as previously described (Tu et al., 2005) except for the timing of sampling for RNASeq and CHIP-Seq; samples were intentionally taken unevenly to more deeply sample the very rapidly changing OX phase, and less densely outside the OX phase. Fermentors were from New Brunswick Scientific (BioFlo 110 or BioFlo 3000). 10 mL overnight-saturated culture was inoculated to start each YMC run. Depending on the model, YMC runs were operated at an agitation speed of 400 rpm (BioFlo 110) or 475 rpm (BioFlo 3000), an aeration rate of 1 L/min, a temperature of 30°C, and a pH of 3.4 in 1 L YMC medium. Once the batch culture was saturated, at least 4 hour of starvation was performed. After starvation, fresh medium was added continuously at a dilution rate of $\sim 0.09\text{--}0.1\text{ h}^{-1}$.

RT-qPCR

2 OD BY5764 cells from the cycle were collected and flash frozen. RNA was extracted using the Qiagen RNeasy Mini kit (QIAGEN, 74104, Valencia, CA) with the standard protocol. First strand

cDNA was synthesized using SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen, 18080-051, Grand Island, NY). Oligo(dT)₂₀ primer was used for reverse-transcription. Fast SYBR Green Master Mix (Applied Biosystems, 4385612, Foster City, CA) was used for real-time PCR and experiments were done on the platform of StepOnePlus Real-Time PCR System (Applied Biosystems, 4385612, Foster City, CA).

RNA-seq

1 mL BY5765 cells from the cycle (about 15 OD/mL) were collected and flash frozen. The pellet was resuspended in 200 µl RNA stat-60 (Tel-Test, Friendswood, TX) and disrupted in a Mini-Beadbeater-16 (Biospec, Bartlesville, OK) broken up three times with 100 µl glass beads in 1.5 ml conical screw cap tube (USA SCIENTIFIC, 1415-8700, Ocala, FL). Supernatant was collected into a new microcentrifuge tube and 800 µl RNA stat-60 was added to repeat bead beating one more time. The supernatants were combined and sit at room temperature for 10 min. 200 µl chloroform was then added and the tube was vortexed for 15 sec and left at room temperature for 10 min. The tubes were centrifuged at 4°C, 16000 rcf for 15 min and the supernatant was transferred to a new tube. 500 µl isopropanol was added and mixed by inverting the tube several times. The mixture was incubated at room temperature for 10 min and centrifuged at 4°C, 16000 rcf for 15 min. The pellet was washed with 70% ethanol twice, air dried and resuspended in 100 µl H₂O. Library construction and sequencing were performed using the Hiseq platform supervised by the UTSW Microarray Core Facility. Briefly, Samples were run on the Agilent 2100 Bioanalyzer to ensure that high quality RNA was used. 4 µg total RNA was then prepared with the TruSeq™ RNA Sample Prep Kit (Illumina, San Diego, CA). mRNA was purified and fragmented before cDNA synthesis. cDNA was then end repaired and A-tailed. After adapter ligation, samples were PCR amplified and purified with AmpureXP beads (Agencourt, A63880, Brea, CA), then validated again on the Agilent 2100 Bioanalyzer. Before being run on the

Illumina Hiseq 2000 samples were quantified by qPCR. Primers were within the adaptor sequence, which are as below:

P1: 5' AAT GAT ACG GCG ACC ACC GA 3'

P2: 5' CAA GCA GAA GAC GGC ATA CGA 3'.

Differential gene expression analysis was performed with TopHat and Cuffdiff following a standard protocol. Briefly, the TopHat parameters used was: --bowtie1 -i 40 --genome-read-mismatches 4 --no-coverage-search -G

Saccharomyces_cerevisiae_Ensembl_EF2/Saccharomyces_cerevisiae/Ensembl/EF2/Annotation/Archives/archive-2012-03-09-08-22-49/Genes/genes.gtf

Saccharomyces_cerevisiae_Ensembl_EF2/Saccharomyces_cerevisiae/Ensembl/EF2/Sequence/BowtieIndex/genome. K-means clustering and heatmaps were produced using an R. Previously defined three super-clusters of genes were obtained from a previous study (Tu et al., 2005). A subset of genes from genes previously identified as non-cycling show elevated expression in the OX phase by hierarchical clustering using the dChIP package (Li and Wong, 2003). Previously defined and newly identified OX phase genes were combined. FPKM values were imported into R and the list was intersected with the three superclusters. The values of each superclusters were then subject to K-means clustering using the function "kmeans". Genes inside each subcluster were further ordered based on the timing of peak expression. 10 clusters were calculated by K-means clustering and ordered by the timing of peak expression from visualization. Genes inside clusters were further ordered by weighted mean time of peak of expression. Weights were calculated as the percentage of RNA-seq reads at each time point over total reads across 16 time points. The weighted mean for each gene was calculated as the sum of weight x time at each time point. Intron-containing genes were defined from the SGD annotation file of 20100102 corresponding to the mapping index.

(http://downloads.yeastgenome.org/curation/chromosomal_feature/). To display RNA signals in the CisGenome Browser (Ji et al., 2008), RNAseq data was first mapped to Saccar2 by bowtie (-5

3 -3 8 –best saccar2). Results were imported into CisGenome browser package and were displayed in 20 bp windows in the browser. Signals at each window were first normalized by total number of mappable reads across 16 time points and then displayed as the percentage of the maximum value of the 16 time points. Exon-exon junctional reads were directly called from the files of junctions.bed by TopHat. Intron junction reads were identified using mappable reads of 39 bp (50 bp-3 bp-8 bp) from bowtie. Each end of a junction read should extend at least 4 bp away from the junction site. Junction reads spanning the 5' and 3' ends of 89 RP gene introns at 16 time points were counted and hierarchical clustering and heatmaps were generated in R using the heatmap.2 function with default parameters. Boxplots were drawn using the boxplot function at every time point. The sums of the 16 time point junction reads were calculated and displayed using the barplot function.

ChIP-seq

Chromatin immunoprecipitation was performed as described previously (Cai et al. 2011). ~5 OD ZKY329 cycling cells per time point were collected for ChIP of single histone modification and ~50 OD ZKY428 cycling cells (*GCN5-FLAG::natMX*, *SET1-3HA::kanMX6*) per time point were collected for ChIP of each chromatin modifying enzyme. Antibodies used are list as following: H3K9ac (Millipore 06-942), H3K14ac (Millipore 07-353), H3K56ac (07-677), H4K5ac (Millipore 07-327), H4K16ac (Millipore 07-329), H3 (Millipore 05-928), H3K4me3 (Millipore CS200580), H3K36me3 (Abcam ab9050), Esa1 (Abcam ab4466), FLAG M2 (Sigma, F1804), HA 12CA5 (Roche, 11583816001). 3 µg primary antibody was used per ChIP experiment. Briefly, cells were first fixed in 1% formaldehyde at 25°C for 15 min and quenched in 125mM glycine at 25°C for 10 min. Cells were pelleted and washed twice with buffer containing 100 mM NaCl, 10 mM Tris-Cl pH 8.0, 1 mM EDTA, 1 mM PMSF, 1 mM benzamidine•HCl before freezing. The frozen pellet was resuspended in 0.45 ml ChIP lysis buffer (50 mM HEPES•KOH

pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% deoxycholate (DOC), 0.1% SDS, 1 mM PMSF, 10 μ M leupeptin, 5 μ M pepstatin A, Roche protease inhibitor cocktail) and lysed by bead beating. Lysate from 50 OD cells were split into two tubes each containing 280 μ l lysate and sonicated for 16 cycles (30 sec on, 1 min off, high output) using a Bioruptor (Diagenode, Denville, NJ or Tosho Denki, Japan). The supernatant of the sonicated lysate was pre-cleared. 50 μ l lysate was saved as input. For ChIP with histone antibodies, 50 μ l whole cell extract (WCE) was diluted 1:10 and used for each ChIP. For ChIP of histone modifier proteins, 500 μ l WCE and 3 μ g antibody was used. After incubation overnight, 50 μ l protein G magnetic beads (Invitrogen, Grand Island, NY, 10003D) or protein A Sepharose beads (GE Healthcare Life Sciences, Piscataway, NJ, 17-5280-01) resuspended in ChIP lysis buffer was added and incubated for 1.5 h at 4 °C. Beads were washed twice with ChIP lysis buffer, twice with DOC buffer (10 mM Tris•Cl pH 8.0, 0.25 M LiCl, 0.5% deoxycholate, 0.5% NP-40, 1 mM EDTA) and twice with TE. 125 μ l of TES buffer (TE pH8.0 with 1% SDS, 150 mM NaCl, and 5 mM dithiothreitol) was added to resuspend the beads. Supernatant was collected after incubation at 65°C for 10 min. A second round of elution was performed and the eluates were combined. Reverse crosslinking was performed by incubation for 6 h at 65°C. An equal volume of TE containing 1.25 mg/ml proteinase K and 0.4 mg/ml glycogen was added to the samples after reverse crosslinking and samples were incubated for 2 h at 37°C. Samples were extracted twice with an equal volume of phenol and once with 25:1 chloroform:isoamyl alcohol. DNA was precipitated in 0.1 volume 3.0 M sodium acetate (PH 5.3) and 2.5 volume of 100% ice-cold ethanol at -20°C overnight. Pellets were washed once with cold 70% ethanol and resuspended in 20 μ l TE. Library construction and sequencing were performed following the Illumina protocol. Briefly, DNA was end repaired and A-tailed. Barcoded adaptors were ligated and DNA was run in 2% agarose gel. DNA fragments from 150 bp to 300 bp long were excised from gel and used for PCR. PCR products were gel-extracted again and quantified on an Agilent Bioanalyzer. Sequencing was performed on an

Illumina GXII or Hiseq 2000 or Solid. Raw reads were mapped to the reference genome (sacCer2) by bowtie and peaks were visualized by the CisGenome Browser as described above.

Distribution of ChIP-seq signals relative to TSSs

Information on TSS locations was adopted from a previous study (Xu et al., 2009). TSSs of genes of interest were aligned at zero of X-axis and ChIP-seq reads were counted for each 10 bp window from -1000 bp to 1000 bp relative to the closest TSS. Number of reads in each window were summed and plotted in R using the plot function.

ChIP-seq signals at genes

ChIP signals of histone modifications at genes were evaluated by counting reads overlapping defined regions of every gene using the CisGenome package. Briefly, a -100 bp to +400 bp window spanning TSS was used for H3K9ac, H3K14ac, H3K56ac, H4K5ac and H3K4me3. The H3K36me3, H4K16ac and H4 signals were calculated from TSS to TES. Peaks of chromatin modifiers were called by MACS peak-calling package with p value= 10^{-02} (Zhang et al., 2008) and annotated by the nearest TSS. ChIP values of chromatin modifiers were calculated similarly by counting reads -250 bp to +250 bp windows spanning the center of peaks or -100 bp to +400 bp window spanning TSS. Ribosomal assembly (ribi) factors were described by Jorgensen et al.,. The ribosomal protein (RP gene) list is from the website, <http://ribosome.med.miyazaki-u.ac.jp/rpg.cgi?mode=orglist&org=Saccharomyces%20cerevisiae>. Amino acid metabolism genes are from the GO term “cellular amino acid metabolic process” of the SGD annotation file. ChIP values were normalized according to the total number of aligned reads. Time point 13 of H3K14ac and H4K5ac, time point 1 of H4K16ac and time point 5/7 of H3K36me3 show certain level of fluctuation. We tried to delete these samples or average these samples with neighboring samples and observed no significant differences in the downstream analysis. Based on the

continuity of time course experiments, we smoothed these samples by averaging the signals with neighboring samples for visualization. K-means clustering and heatmaps were performed in R as described above. Hierarchical clustering was performed by dChIP package (Li and Wong, 2003).

Data Pre-processing

For each sample, we normalized the observed value by the total number of aligned reads. For each modification, we then standardized the 16 observed values (\log_2 of read counts plus 1) within each gene. We were primarily interested in the trend of change over time, rather than the relative magnitude of expression among genes. So we standardized the 16 values within each gene, i.e. we subtract the mean from the 16 values and then divide by their standard deviation. Then we pooled the data for all RNA-seq and ChIP-seq datasets together to get a data matrix, in which each row is a gene and each column is a time point.

Notations

For a given marker, we denote the matrixes as $Y = [Y_{gj}]$ as processed gene expression levels, where $g = 1, \dots, N$ is the gene index, and $j = 1, \dots, 16$ is the time point index. We denote t_j as observation times.

Correlation between chromatin modifications and gene expression

A gene-specific correlation coefficient was calculated between 16-time-point gene expression values and each histone modification value. A 2-time-point shift was allowed to achieve the maximum correlation value. Basically, we calculated gene-specific correlation coefficient with time point shift of -2, -1, 0, 1, 2 and took the maximum correlation value as the output correlation coefficient. We note that some modifications are shifted forward or backward relative to RNA level and this could help get the accurate correlation value. A density distribution of correlation

values was plotted across the whole genome or across subsets of interest (OX, RB, RC). We observed a positive correlation in the distribution of H3K4me3, H3K36me3, H3K9ac, H3K14ac and H4K5ac with RNA level (data not shown). Interestingly, the OX phase genes always showed the best correlation with expression, while RB phase genes showed a relatively poor correlation. We also noticed that H4K16ac and total H3 exhibit a slight negative correlation and H3K56ac shows no obvious correlation (data not shown).

Curve fitting

For each gene, we observed gene expression levels at 16 time points. To convert these discrete data into a continuous profile/function, we fit every set of 16 points into a smooth curve using a penalized B-spline. We denote function $f_g(t)$ as fitted continuous profile from data $[Y_{g1}, \dots, Y_{g16}]$ and $[t_1, \dots, t_{16}]$. $f_g(t)$ is a linear combination of B-spline basis functions. Estimation of the coefficient of each basis function is obtained by a penalized optimization problem.

Time alignment (Batch effect removal)

We assume the dynamics of gene expression levels changes according to a biologically specified clock or time system. Such time systems can be affected by unknown or uncontrolled fluctuations in experimental conditions such as minute variations in medium composition, temperature, etc. When data are generated in multiple batches, the results can be biased by the different time systems.

RNA-seq and ChIP-seq of histone modifications were generated in two batches. To minimize batch-to-batch variations, the histone modification H3K9ac, which is the most periodic, was independently measured in both batches. We used this information to align or “time warp” the two batches into a single time system. We denote $f_g^1(t)$ and $f_g^2(t)$ as the fitted gene profiles

for two different batches. Theoretically, $f_g^1(t)$ and $f_g^2(t)$ should be the same under the identical time system. Thus we seek a mathematical transformation $s:[t1, \dots, t16] \rightarrow [s(t1), \dots, s(t16)]$, such that $f_g^1(t)$ and $f_g^2(s(t))$ are most alike for all genes. Mathematically, we need to minimize

$$E(f^1, f^2) = \sum_{g=1}^N (f_g^1(t) - f_g^2(s(t)))^2 dt$$

We obtain the final result $[s(t1), \dots, s(t16)]$ using a simulated annealing algorithm.

As a result, the observation times of the raw RNA-seq data are replaced by $[s(t1), \dots, s(t16)]$, and we assume that the batch effect is minimized as much as possible given the current data.

Clustering genes

We cluster the genes using the processed gene expression levels. The K-means algorithm is used for its speed and robustness of dealing with large data. The K-means algorithm is affected by the choice of initial cluster centers. To minimize the dependency on initial cluster number, we repeated the K-means algorithm 50 times using different randomly chosen initial cluster centers, and kept the result with the smallest total within-cluster sum of squares. To choose the number of clusters, we tried different options range from 1 to 30. Finally we selected 7 clusters since it had the best Bayesian Information Criterion (BIC) value: BIC is one of the most popular criteria for model selection. Similarly we selected 14 clusters of 8 histone modifications in Figure 2.9.

Finding the profile peak location of each cluster

For each cluster of genes we want to know when the expression profile reaches its peak value.

For each gene, we can find the peak (highest point) of profile from the fitted smooth curve. The cluster peak location is defined as the average of peak locations for all genes in a cluster.

Similarly we can calculate the standard deviation of the cluster peak location.

GO analysis

Gene ontology analysis was performed by the hypergeometric test using the SGD annotation file (<http://www.yeastgenome.org/download-data/curation#.UQFqKhyLFXy>). P values were adjusted by the False Discovery Rate (FDR).

Analysis of H3K56ac

The analysis of H3K56ac in the cell cycle was performed following procedure similar to one described previously (Kaplan et al., 2008). We adopted the same regions of nucleosomes in chromosome *III* and calculated H3K56ac signals at these nucleosomes as described above. We sorted nucleosomal H3K56ac signals by genomic location and fit our data with those of the previous study by generating heat maps in R.

Enrichment analysis of motifs associated with clusters

We took advantage of three motif databases, TRANSFAC, JASPAR and UNIPROBE. We first identified the genomic coordinates of each motif based on its weight matrix using the CisGenome package. Since we focused on transcriptional initiation, we filtered these binding sites by selecting sequences located between -600 bp to +400 bp of any TSS. Next, we examined genome-wide whether any motif sequence exists at the 5' end of the gene. Subsequently, we calculated the percentage of genes with certain motif sequence in all RNA_H clusters and performed the Wald test to determine whether the percentage in each cluster was significantly higher than the genome-wide percentage. We applied this procedure to all motifs in the three databases.

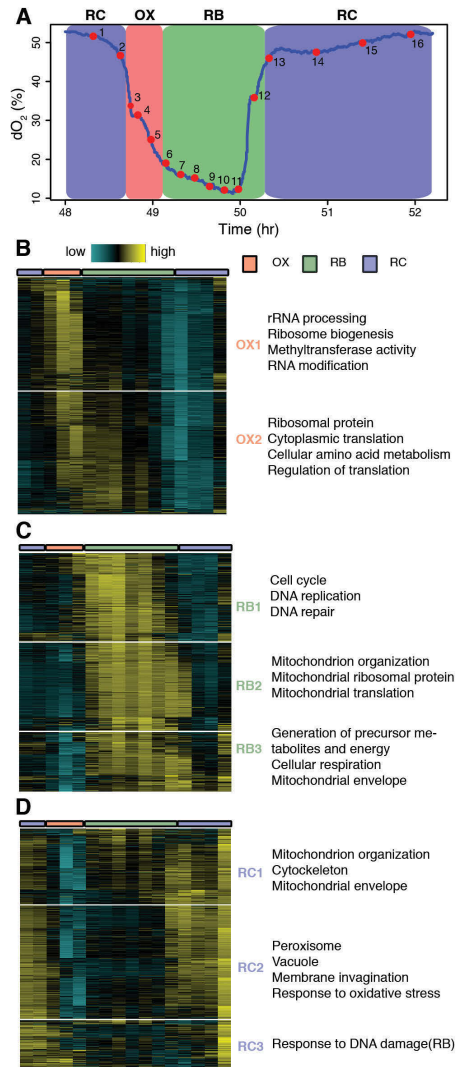


Figure 2.1 High temporal resolution analysis of gene expression reveals meticulous

temporal compartmentalization in yeast. (A) RNA was extracted across 16 time points of one cycle at the indicated time points and RNA-seq was performed. With this strategy, the shortest interval between two OX phase time points is < 5 minutes, and across the entire cycle the samples were taken at intentionally uneven intervals. The same color scheme is used for labeling the three phases throughout the paper: Red = OX phase, Green = RB phase, Blue = RC phase. (B-D) K-means clustering of OX, RB and RC phase genes separately reveals subtly distinct temporal patterns. Enriched GO terms are listed on the right of each sub-cluster. Each row of the heat maps represents one gene and columns represent the 16 consecutive time points from left to right. 3-color bar on top marks the three phases. The blue-yellow gradient represents RNA signals normalized to read depth using Cuffdiff and centered to a mean of zero across 16 time points. (B) shows 2 subclusters of OX phase genes; (C) shows 3 subclusters of RB phase genes; (D) shows 3 subclusters of RC phase genes.

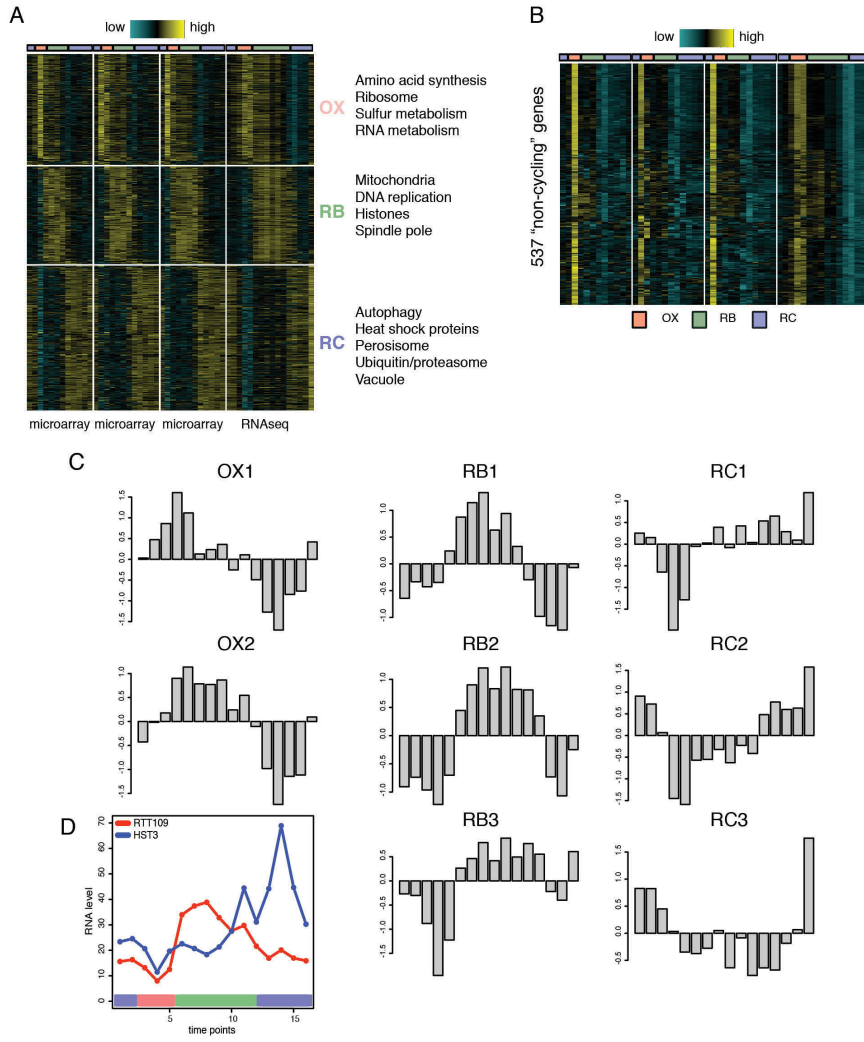


Figure 2.2 Dynamics of gene expression in YMC. (A, B) RNA-seq experiment reveals the same three clusters (OX, RB and RC) discovered by previous microarray experiments. 3 cycles of microarray data are plotted in the first three joint columns, each containing 12 time points. 16 time-point RNA-seq data are plotted in the last joint column. (A) Enriched GO terms are labeled on the right. (B) 537 “non-cycling” genes defined previously show clearly elevated expression level in OX phase. Algorithm previously used to define cycling genes required highest expression of 2 or more time points and highest expression in all three cycle periods. As can be seen, these 537 genes form very sharp peaks with only one high time point and/or are represented in only one or two cycles. (C) Average 16 time-point RNA-seq data across genes in subclusters from Figure 2.1 (B-D). (D) Expression level of *RTT109* and *HST3* across 16 time points extracted from the RNA-seq experiment.

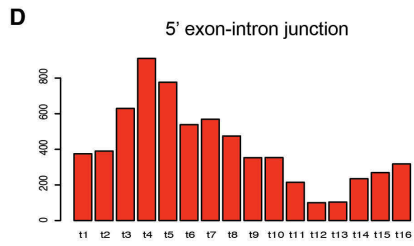
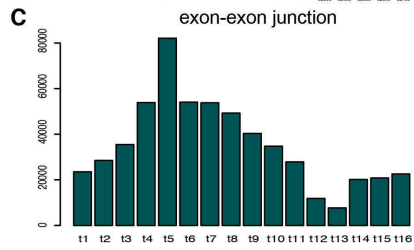
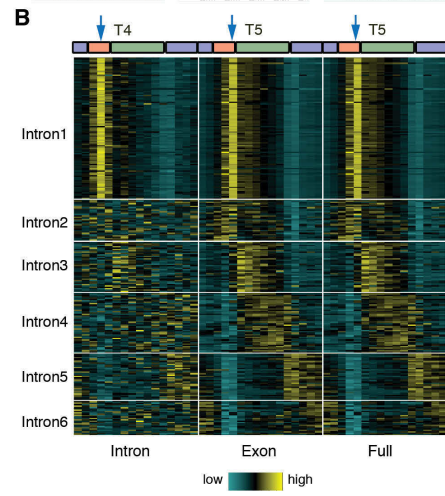
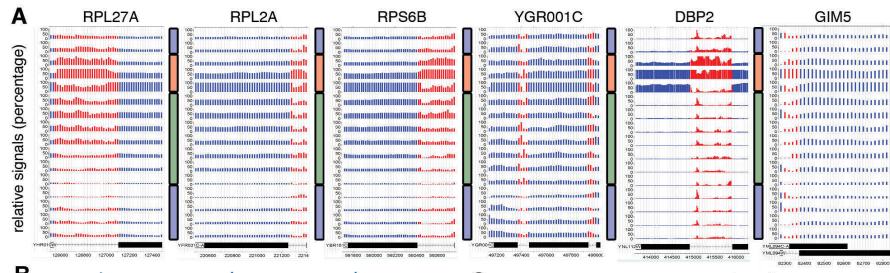


Figure 2.3 RNA-seq analysis at introns reveals transient accumulation of pre-mRNAs

during OX phase. (A) Relative RNA signals are displayed at intron-containing genes. Each track represents relative RNA levels at one of 16 time points, ordered sequentially from top to bottom. Signals are displayed as percentage of maximum value of the 16 time points at each 20 bp window. Red regions represent introns and blue regions represent exons. 3-color bars represent the 3 YMC phases vertically. Examples include genes in this cluster with the longest (*DBP2*), shortest (*GIM5*) intron and two introns (*YGR001C*). (B) A total of 280 intron-containing genes are divided into 6 groups based on their temporal RNA signals at introns, exons, and intron+exon (full) regions, ordered by the phase of peak full-length RNA expression. 16 time-point signals (individual columns) of each region are grouped together consecutively as one joint column. Three regions are ordered as “Intron”, “Exon”, “Full” and separated by white lines. Each single row represents signals of one gene standardized to a mean of zero. Arrows mark the peak time point. (C) The total number of reads spanning the exon-exon junctions of RP genes from Intron1 cluster are displayed across 16 time points in the bar plot. (D) The total numbers of junction reads spanning the 5’ end of RP introns from Intron1 cluster are displayed across 16 time points in the bar plot.

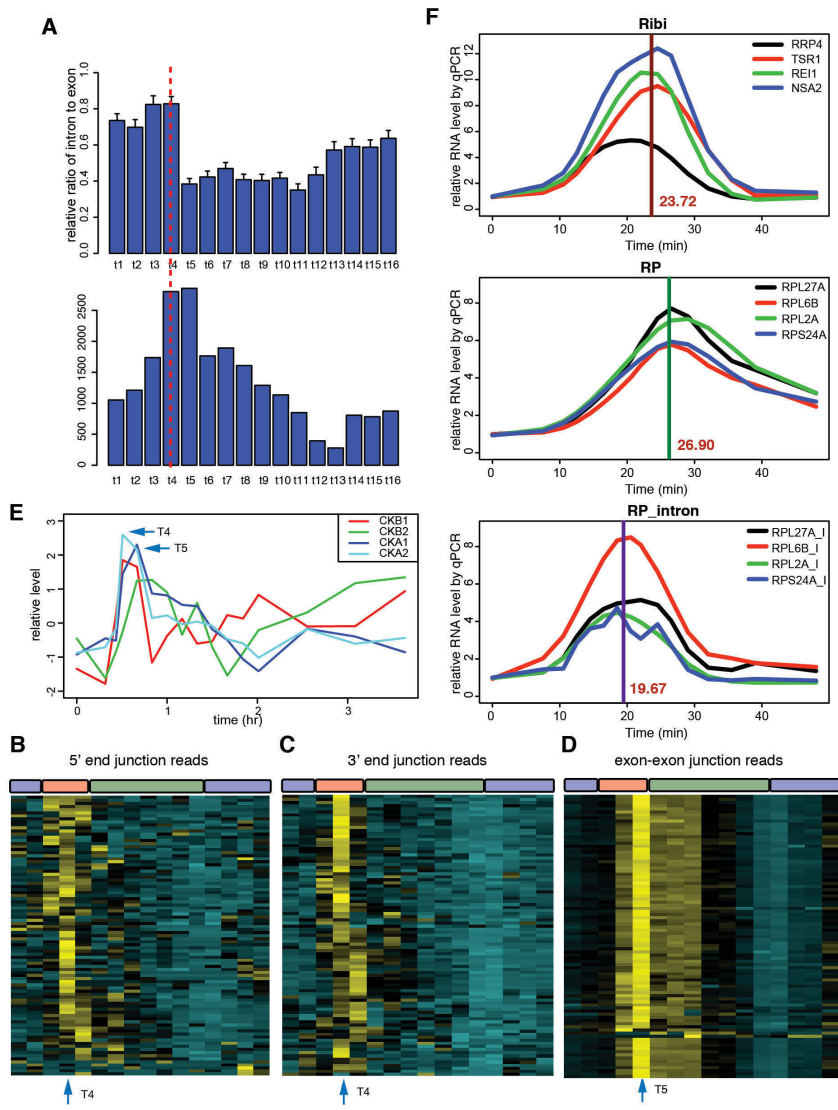


Figure 2.4 Temporal patterns of RNA-seq signals at RP introns. 89 RP genes from Intron1 cluster were included in the analysis. (A) The ratio of intron to exon was calculated for every gene in the cluster and the means are displayed across 16 time points in the top panel. The bottom panel is total numbers of junction reads spanning the 3' end of RP introns at 16 time points. Note that the numbers of 3' junction reads are consistently higher than the numbers of 5' junction reads (Figure 3C), consistent with an arrest of splicing at the "2/3 intermediate" stage of splicing. Heat maps show the temporal patterns of junction reads spanning the 5' end (B) and the 3' end (C) of RP introns and exon-exon junctions (D) of RP genes. Data is centered to mean zero. "T4" or "T5" mark the time point where signals reach the peak. (E) The RNA levels of CK2 subunits, catalytic subunits *CKA1* and *CKA2*, and regulatory subunits *CKB1* and *CKB2*, are plotted across the 16 time points. Time point 4 and 5 are labeled by "T4" and "T5". (F) RT-qPCRs of 4 Rib1 (top panel), 4 RP (middle panel) and 4 RP intron (bottom panel) transcripts across 16 ultra-dense OX time points (1.5~2 min intervals). Signals are normalized by the level at time point 1. 16 values of each transcript are plotted in a smoothed line. The vertical lines in each panel mark the mean peak time of the 4 transcripts and the estimation values are labeled at the right side of the lines.

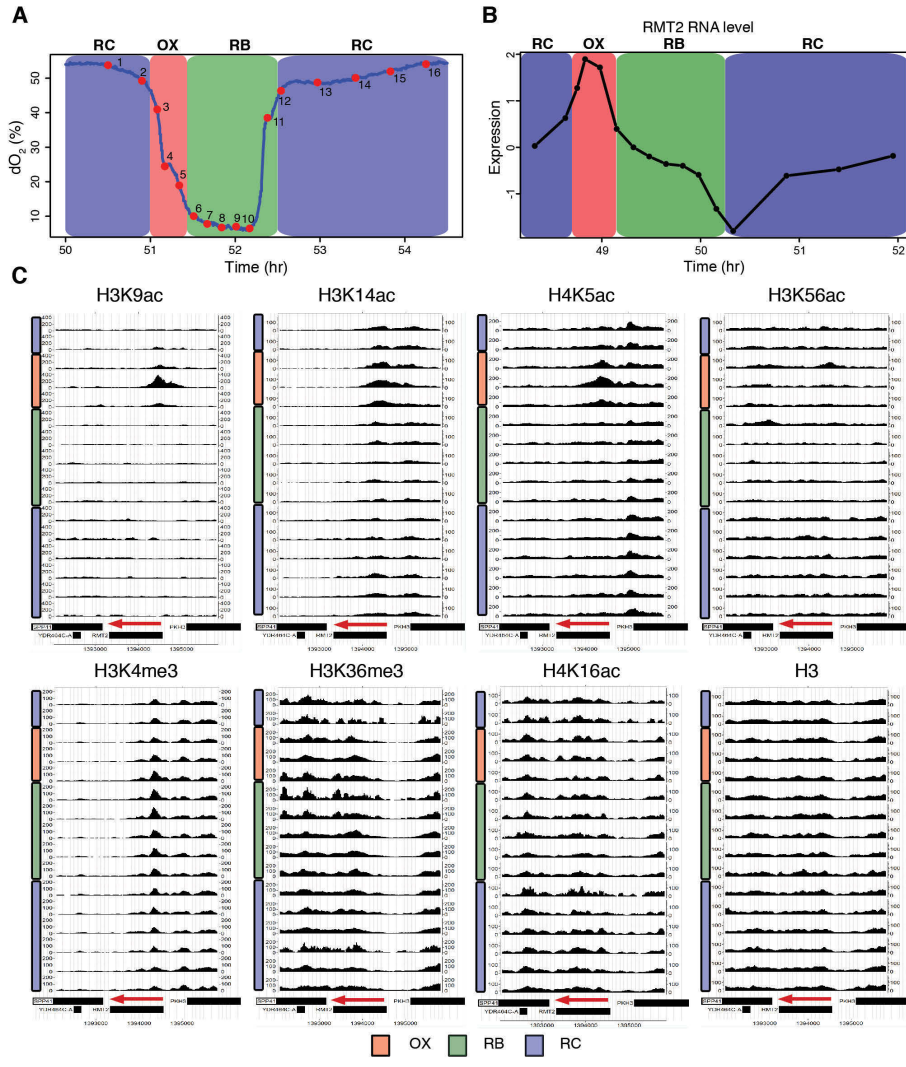


Figure 2.5 Dynamic chromatin states across the YMC. (A) Cycling cells were collected at 16 intentionally uneven time points over one cycle for ChIP-seq. (B-C) Temporal relationship between RNA level and histone modifications at the *RMT2* locus. (B) The RNA level of *RMT2* is elevated in OX phase, with maximum value at time point 4. Normalized RNA-seq data are plotted. (C) CisGenome browser views of H3K9ac, H3K14ac, H4K5ac, H3K56ac, H3K4me3, H3K36me3, H4K16ac and H3. 16 tracks represent 16 time points from top to bottom consecutively. *RMT2* is encoded on the Crick strand, labeled by the red arrows pointing to the left.

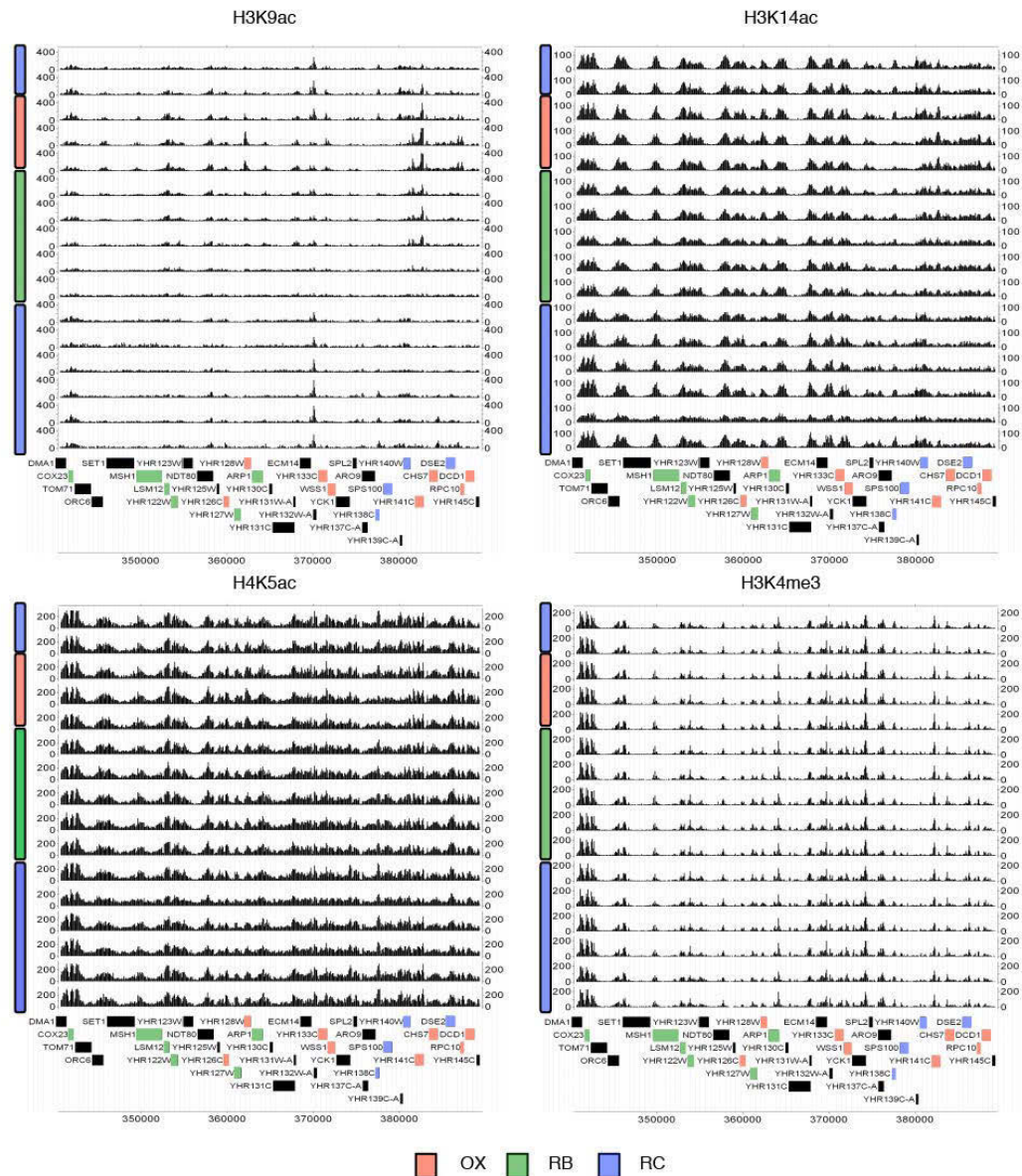


Figure 2.6 Dynamic chromatin modifications across the YMC. H3K9ac, H3K14ac, H4K5ac and H3K4me3 ChIP-seq signals at chrVII: 340669-389335 are displayed in CisGenome browser. 16 tracks represent 16 time points from top to bottom consecutively. 3-color bars on the left of each panel mark the three phases. The boxes with the same 3 colors in the last track mark the phases of expression of corresponding genes. Black boxes represent non-cycling genes.

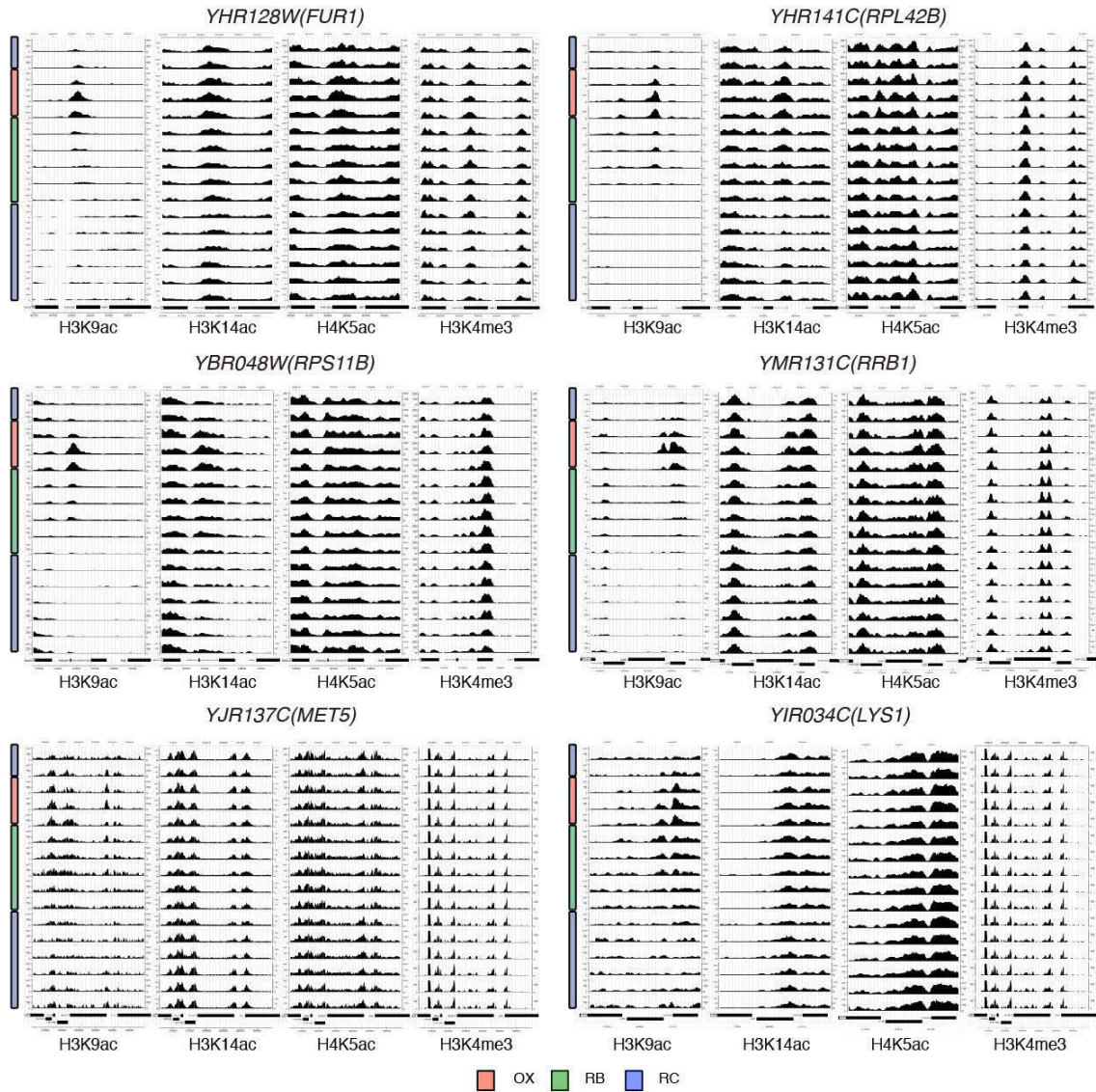


Figure 2.7 Temporal patterns of chromatin modifications at OX genes. H3K9ac, H3K14ac, H4K5ac and H3K4me3 ChIP-seq signals at *FUR1*, *RPL42B*, *RPS11B*, *RRB1*, *MET5* and *LYS1* are displayed in CisGenome browser screenshots. 16 tracks represent 16 time points from top to bottom consecutively. 3-color bars on the left of each panel mark the three phases.

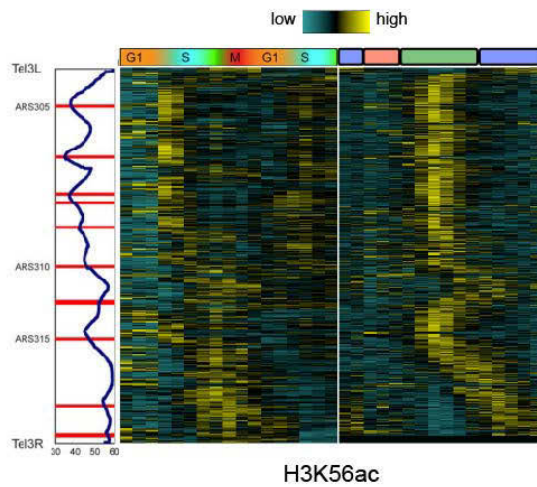


Figure 2.8 Temporal patterns of H3K56ac. Comparison of H3K56ac incorporation between the standard cell cycle and YMC. The left panel shows the ChIP-chip data of H3K56ac across the cell cycle from the study of Kaplan *et al.* 2008, and the right panel is the new ChIP-seq data of H3K56ac across the YMC. The H3K56ac signal was calculated for every nucleosome on chromosome III, sorted by chromosomal location and standardized to a mean of zero. The vertical panel, adapted from the previous study, shows the DNA replication time and red lines represent origins of replication (ARS).

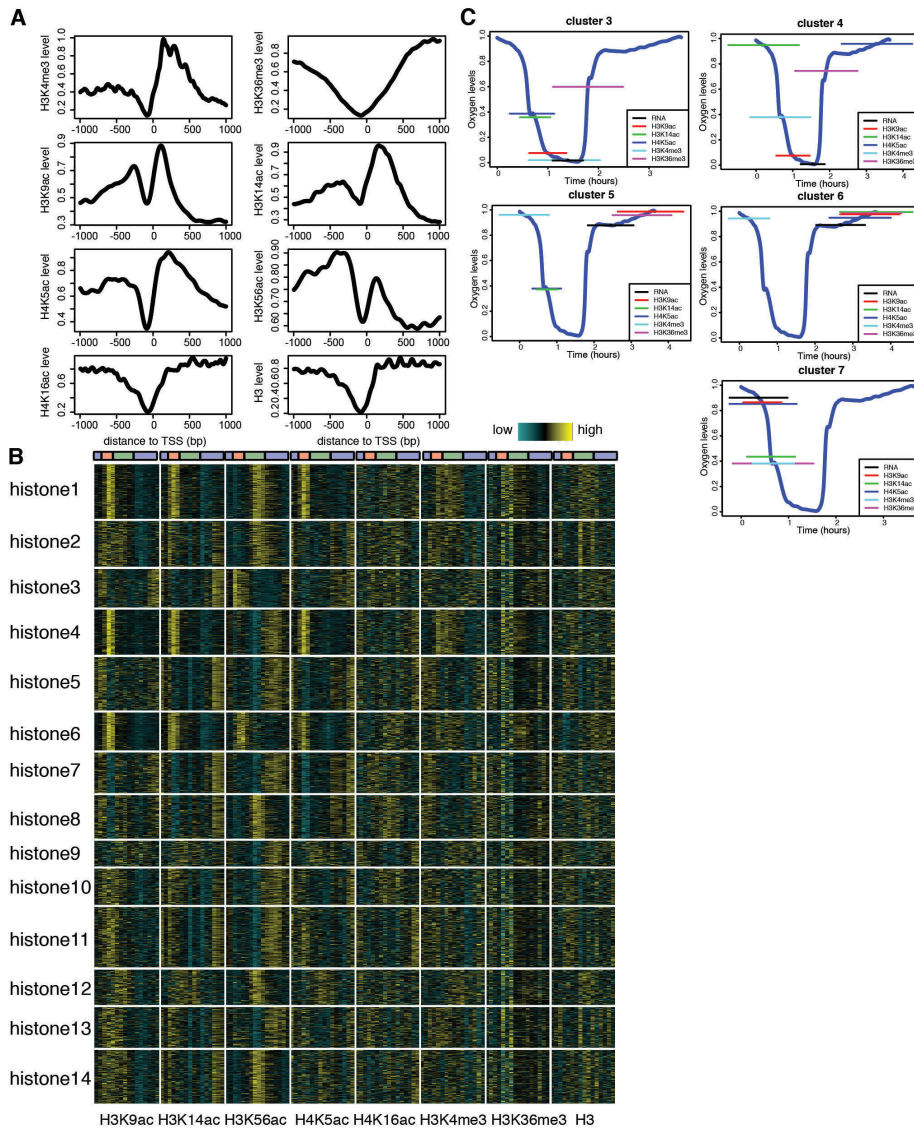


Figure 2.9 Temporal association between chromatin states and gene expression. (A) Spatial distribution of chromatin markers. ChIP-seq reads are averaged from -1000 to +1000 bp of TSSs across the genome. (B) 14 clusters were determined for the whole genome, revealing complex combinatorial temporal patterns of H3K9ac, H3K14ac, H4K5ac, H3K56ac, H4K16ac, H3K4me3, H3K36me3 and H3. Each row represents centered signals of one gene and each column represents one time-point ChIP-seq corresponding to the indicated chromatin mark. We calculated reads at the 5' end as gene-specific signals of H3K4me3, H3K9ac, H3K4ac, H4K5ac and H3K56ac, and reads across the full gene as signals of H3K36me3, H4K16ac and H3. (C) Estimation of time in YMC when RNA and histone modifications in “RNA_H3”, “RNA_H4”, “RNA_H5”, “RNA_H6”, “RNA_H7” reach the maximum. The middle of the horizontal colored lines on the O₂ curves represent the mean peak value of genes in each cluster and the lengths represent standard deviations for genes in these clusters.

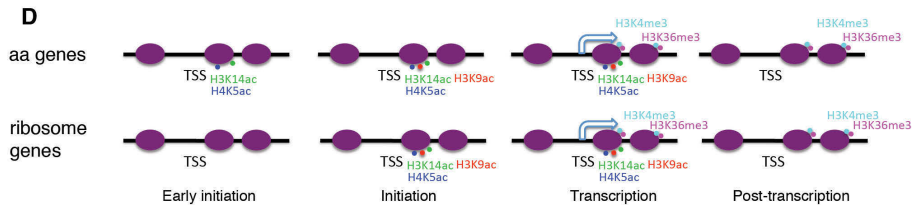
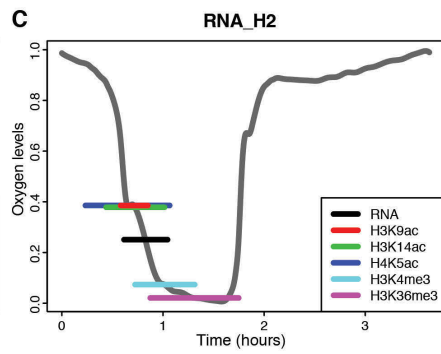
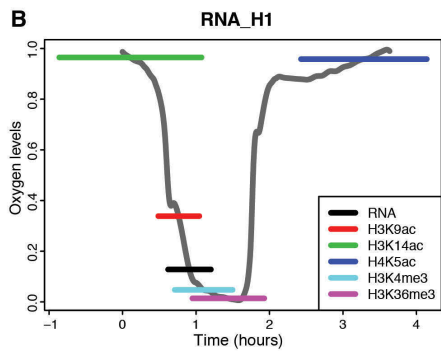
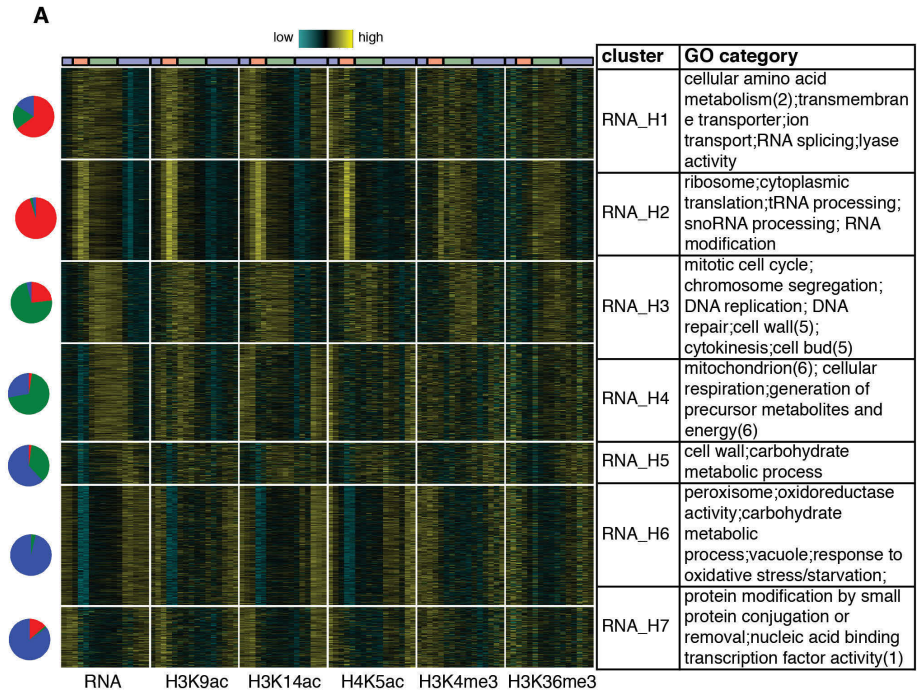


Figure 2.10 Temporal combinatorial patterns of chromatin states and gene expression define functionally related genes. (A) 7 clusters of cycling genes (combining OX, RB, RC phase genes) were defined by K-means clustering and ordered by the time of peak gene expression, revealing distinct combinatorial patterns of RNA and chromatin states. RNA, H3K9ac, H3K14ac, H4K5ac, H3K4me3 and H3K36me3 are included in this analysis and each is displayed in a joint column, separated by white lines. The data is standardized to a mean of zero. Pie charts on the left indicate composition of OX (Red), RB (Green), RC (Blue) phase genes in each cluster. RNA_H1 and 2 consist primarily of OX phase genes; RNA_H3-4 consist mostly of RB phase genes; RNA_H5 is a mixture of RB and RC genes; RC genes dominate RNA_H 6-7. Enriched GO categories are listed on the right. (B-C) Estimation of time in YMC when RNA and histone modifications in “RNA_H1” (B) and “RNA_H2” (C) reach the maximum. The middle of the horizontal colored lines on the O2 curves represent the mean peak value of genes in each cluster and the lengths represent standard deviations for genes in these clusters. (D) Models of histone modification pathways during the transcription of aa and ribosomal genes. Purple ovals represent -1, +1 and subsequent ORF nucleosomes for the indicated gene. Hooked arrows represent transcription and spots represent histone modifications.

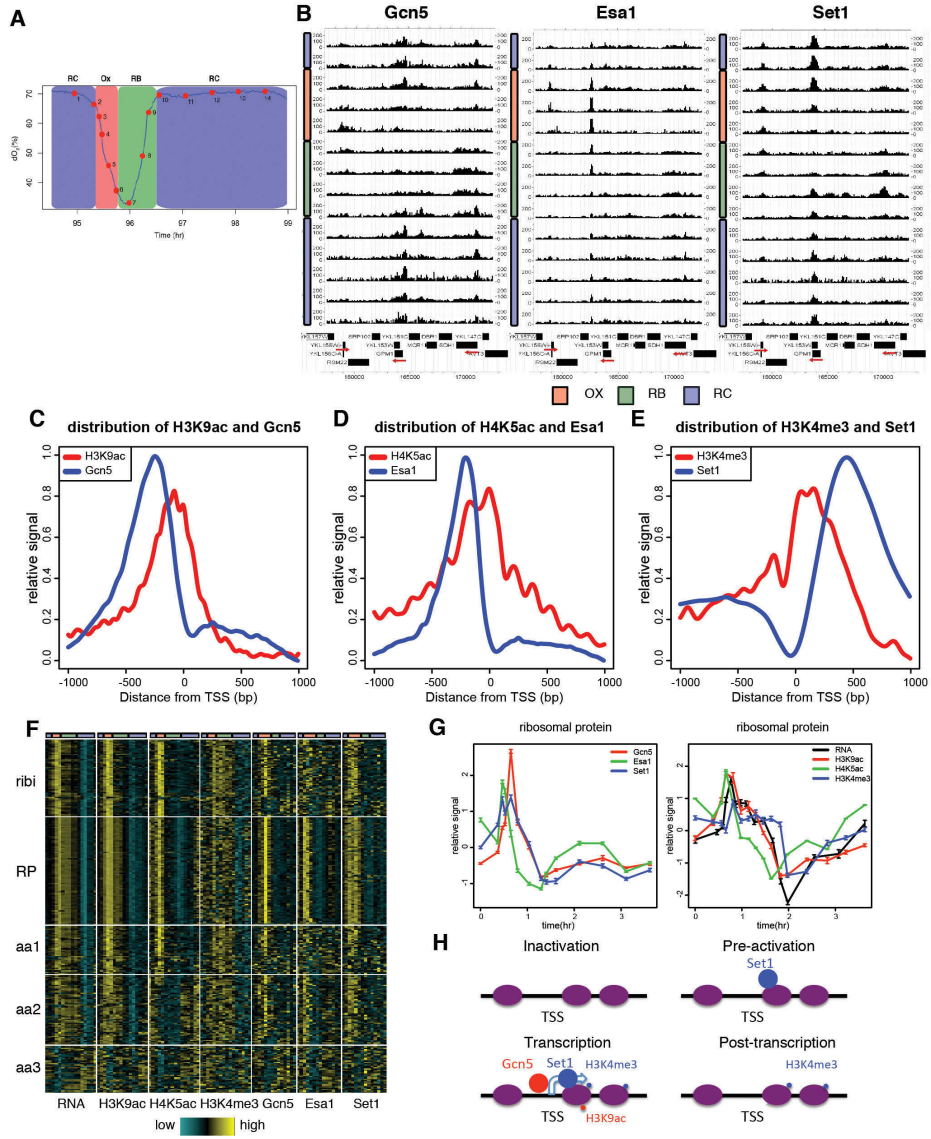


Figure 2.11 Similar but not identical spatial and temporal patterns of chromatin modifiers and corresponding modifications. (A) Cycling cells of Gcn5-FLAG, Set1-3HA were collected at 14 intentionally unevenly spaced time points for ChIP-seq. The strain displayed a YMC, albeit with a slightly shorter RB phase. (B) CisGenome browser views of the temporal patterns of Gcn5p, Set1p and Esa1p binding at chromosome *XI*: 157000-173000. The 14 tracks represent 14 consecutive but unevenly spaced time points from top to bottom. Red arrows label the direction of interesting genes: from left to right are *YKL156W (RPS27A)*, *YKL152C (GPM1)*, and *YKL148C (SDHI)*. (C-E) The binding sites of chromatin modifiers do not overlap directly with the corresponding modifications. ChIP signals were averaged from -1000 to +1000 bp of TSSs (transcription start site) derived from MACS detected peaks. Signals of each modifier are plotted together with its corresponding modification normalized by H3. (F) Comparison of the temporal patterns of RNA, histone modifications and corresponding modifiers for ribosome biogenesis genes, ribosomal protein genes and amino acid metabolism genes. ChIP values were calculated by counting reads spanning TSSs and centered to a mean of zero. Three different groups of amino acid genes were defined by K-means clustering. (G) Averaged signals of histone modifications and modifiers for ribosomal protein genes across the YMC. (H) A model of Set1p/H3K4me3 and Gcn5p/H3K9ac during transcription. Purple ovals represent -1, +1 and subsequent ORF nucleosomes for the indicated gene. The hooked arrow represents transcription and small/big spots represent histone modifications/histone modifiers.

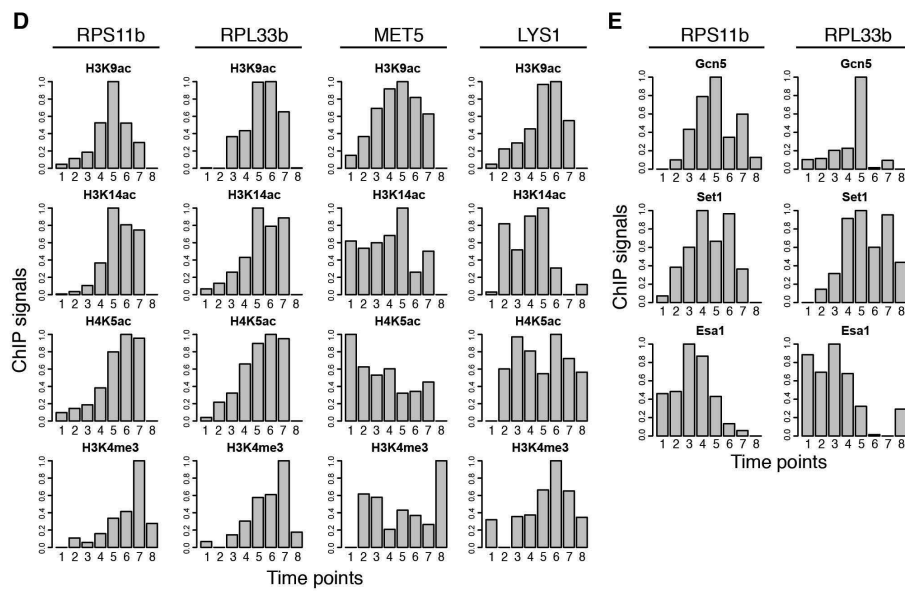
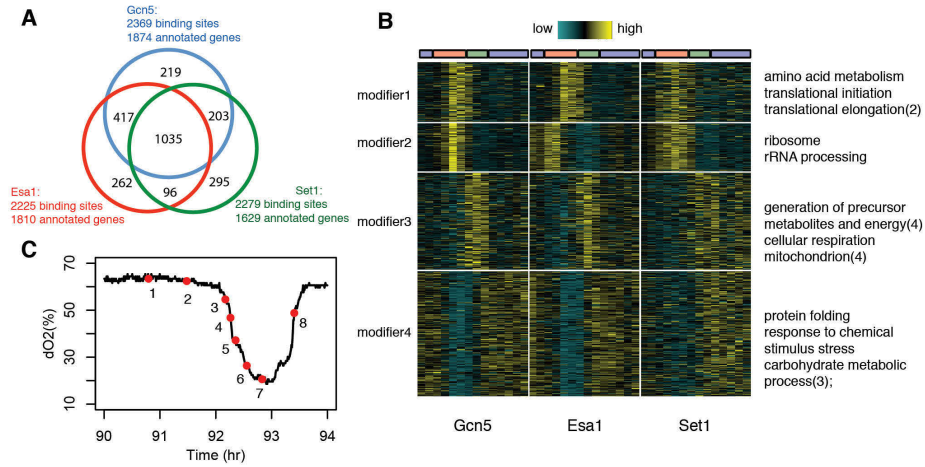


Figure 2.12 Dynamic analysis of chromatin modifiers and corresponding modifications. (A) shows the numbers of binding sites and annotated genes by MACS for Gcn5, Esa1 and Set1. Areas of overlap among the three enzymes and numbers of annotated genes in each subset are labeled inside the Venn diagram. (B) 4 distinct patterns of the 1035 genes to which all three enzymes bind show the temporal combinatorial recruitment of the chromatin modifiers. Overrepresented GO terms for each cluster are listed on the right. (C) Gcn5-3FLAG 6HA-Set1 cells were collected at 8 time points from late RC to RB phase in YMC for ChIP-qPCR. (D) ChIP-qPCR of H3K9ac, H3K14ac, H4K5ac and H3K4me3 at examples of OX phase genes. RPS11b and RPL33b are two RP genes and MET5 and LYS1 are aa genes. H3K14ac and H4K5ac were detected earlier than H3K9ac on aa genes but not on RP genes. (E) ChIP-qPCR of Gcn5, Set1 and Esa1 at RPS11b and RPL33b. Esa1 and Set1 binding increase before Gcn5 binding.

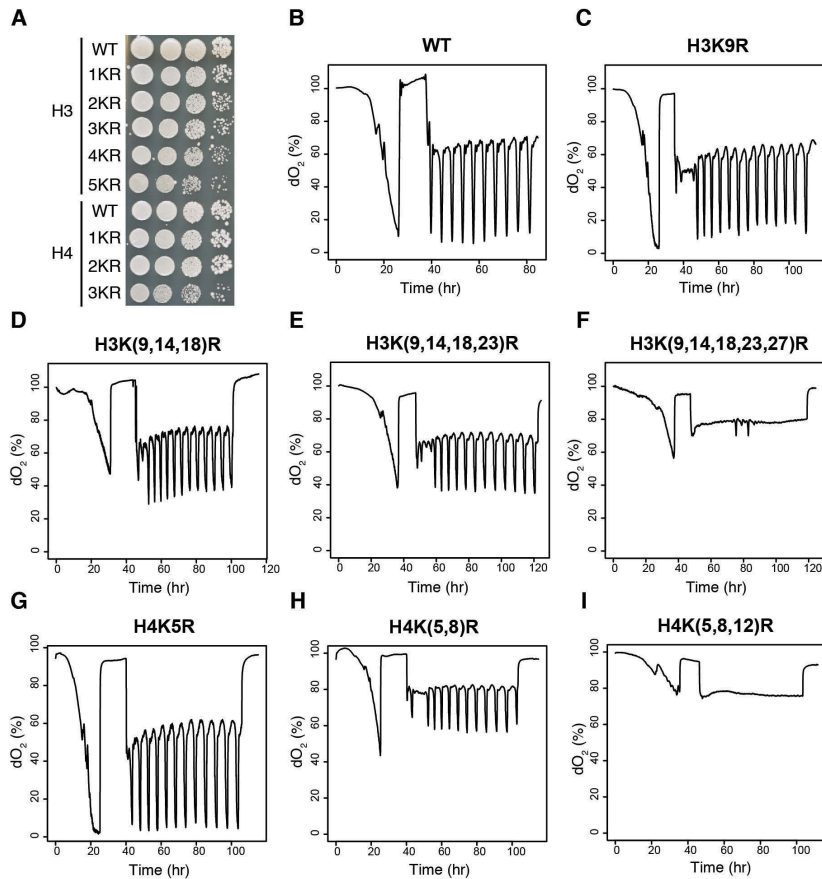


Figure 2.13 O₂ consumption traces of histone point mutants. (A) Strains were plated (1:10 serial dilutions) on YPD plates and grown at 30°C. H3 mutants represent H3K9R, H3K(9,14)R, H3K(9,14,18)R, H3K(9,14,18,23)R, H3K(9,14,18,23,27)R. H4 mutants represent H4K5R, H4K(5,8)R, H4K(5,8,12)R. (B-I) Metabolic cycling is disrupted with increasing number of lysine to arginine mutations on H3 or H4.

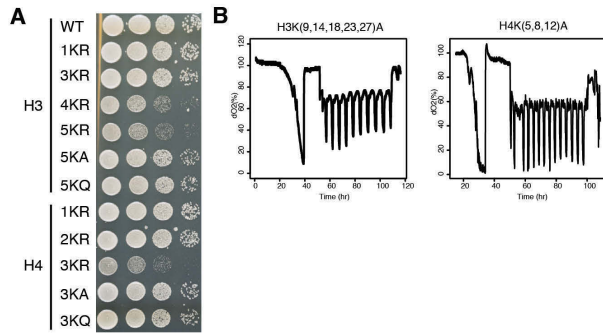


Figure 2.14 Analysis of histone mutants in O₂ oscillation. (A) Strains were plated (1:10 serial dilutions) on YPD plates and grown at 30°C. H3 mutants represent H3K9R, H3K(9,14,18)R, H3K(9,14,18,23)R, H3K(9,14,18,23,27)R, H3K(9,14,18,23,27)A, H3K(9,14,18,23,27)Q. H4 mutants represent H4K5R, H4K(5,8)R, H4K(5,8,12)R, H4K(5,8,12)A, H4K(5,8,12)Q. (B) O₂ oscillation curves of H3K(9,14,18,23,27)A and H4K(5,8,12)A.

Chapter 3

Genome-wide location of H3K56ac reveals active roles of “meiosis-specific” Spo11 in mitosis

Abstract

Spo11 is a meiotic-specific endonuclease that makes double strand breaks (DSBs) and initiates recombination in meiosis. We identified ~1000 H3K56ac peaks in a haploid prototrophic yeast strain which spatially correlate with meiotic DSB hotspots. Spo11 and other meiotic DSB proteins are actively regulated and contribute to the formation of H3K56ac peaks under stress conditions. Spo11 is required for mitotic DNA recombination at meiotic DSB hotspots and facilitates yeast evolution under stress conditions. Thus, yeast may hijack the meiotic DSB pathway to stimulate programmed DSBs and DNA recombination in response to certain stresses. Additionally, H3K56ac is required for the recovery of gametes after sporulation.

Introduction

DNA damage is harmful to the integrity of genomes and it is usually repaired by different routes such as homologous recombination and non-homologous end joining, which also facilitate genome diversity and evolution (Andersen and Sekelsky, 2010). Two types of DNA recombination exist, meiotic and mitotic recombination. Meiotic recombination is initiated by programmed DSBs at non-randomly defined hotspots (Borde and de Massy, 2013). Chromatin structure, histone modifications and protein factors are thought to determine the distribution of meiotic DSB hotspots whereas the meiosis-specific endonuclease, Spo11, actually catalyzes the cleavage (Keeney, 2008). Meiotic DSBs and recombination are called “programmed” because they are activated by endogenous factors and are required for physical association between homologous chromosomes which thereby directs accurate segregation. In contrast, mitotic recombination is mainly used for repair of accidental DNA damage associated with DNA replication or other spontaneous and induced scenarios (Andersen and Sekelsky, 2010). The types of DNA damage are not limited to DSBs and the locations are more diverse than the meiotic ones, although some genomic regions are inherently more fragile, such as tRNA genes and DNA replication origins (Szilard et al., 2010).

A few histone modifications are strongly associated with DNA damage response and regulate DNA repair. Phosphorylation of H2A (γ -H2A) is an early response of DNA damage and it is very important for efficient DNA repair by recruiting repair machinery, cohesins and chromatin remodeling complexes to damage sites (Vidanes et al., 2005). H3K56ac is another histone modification required in DNA damage response (Chen et al., 2008). It functions in late stages of repair, which ensures the completion of DNA repair by facilitating de novo chromatin reassembly. It may be also required for shutting off the DNA damage checkpoints and continuation of cell cycle.

In this study, we identified ~1000 H3K56ac peaks from a haploid prototrophic yeast strain which resemble the distribution of meiotic DSB hotspots. Spo11 is actively regulated and required for the dynamic H3K56ac peaks in a nutrient limited stress condition. We have also discovered a role of Spo11 in mitotic DNA recombination and heat shock resistance. It suggests yeast could utilize the meiotic DSB machinery to induce recombination and increase genetic diversity intentionally in face of stresses. H3K56ac is also required for the completion of gametogenesis.

Results

Spatial and temporal patterns of H3K56ac in the yeast metabolic cycle

Under continuous nutrient limited conditions, the prototrophic CEN.PK strain exhibits a respiratory cycle of ~4-5 hours with transcription and metabolites oscillating at the same pace, called the yeast metabolic cycle (YMC) (Tu et al. 2005). The cycle is divided into three phases: OX phase represents an active growth phase when translation machinery is assembled and proteins are synthesized, coupled with high assumption of oxygen; RB phase is consist of cell division cycle and mitochondrion biogenesis; RC phase resembles a stationary phase when catabolism and stress responses occur.

The previous study (Chapter 2) reveals dynamics of histone modifications associated with transcription through 16 time point ChIP-seq experiments across one YMC. Particularly, we observed hundreds of H3K56ac peaks with unique temporal pattern (Fig. 3.1A). They start to increase in the middle of RC phase, reach a maximum in early OX phase and decrease in RB phase. 1166 H3K56ac peaks were detected and they exhibit the same temporal pattern (Fig. 3.1B). These peaks are not correlated with transcription of the nearest genes. Genes with different expression timing (OX/RB/RC/non-cycling) are indiscriminately associated with these H3K56ac

peaks (Fig. 3.1C). Interestingly, H3K56ac peaks are significantly associated with several types of genomic features, such as autonomously replicating sequences (ARS), Introns, noncoding-exons, telomere and tRNA genes (Fig. 3.1D). Many of them correspond to DNA fragile regions identified by ChIP of DNA damage markers or genetic assays monitoring DNA recombination (Charles and Petes, 2013; Szilard et al., 2010). H3K56ac is relatively excluded from regions of coding sequences and retrotransposons (Fig. 3.1D).

H3K56ac is involved in transcription, DNA replication and repair (Costelloe and Lowndes, 2010). Chapter 2 reported association between H3K56ac and DNA replication in the YMC, occurring in the RB phase. Here we excluded the possibility of these H3K56ac peaks reflect phase-specific transcription. Instead, the association with DNA fragile regions suggests that these H3K56ac peaks could be involved in DNA damage and repair.

H3K56ac correlates with meiotic DSB hotspots spatially

We first examined the spatial correlation between H3K56ac peaks and DNA damage hotspots. Two datasets of DSB hotspots were included, meiotic DSB hotspots represented by Spo11 immunoprecipitated DNA oligos (Spo11 oligos) from meiotic culture (Pan et al., 2011) and mitotic DSB hotspots represented by γ H2A ChIP-chip in active growing haploid cells (Szilard et al., 2010). Surprisingly, we observed a strong correlation between H3K56ac peaks with the meiotic DSB hotspots but not the mitotic DSB hotspots, although the H3K56ac peaks were generated from haploid cells (see Fig. 3.2A for examples and Fig. 3.3 for zoomed-in examples). Next, we examined the genome-wide correlation between H3K56ac peaks and meiotic DSB hotspots. The distance between H3K56ac peaks and Spo11 hotspots is significantly smaller than that of scrambled H3K56ac peaks and also significantly more H3K56ac peaks overlap with Spo11 hotspots (Fig. 3.2B). By mapping H3K56ac reads on all Spo11 hotspots, we observed the enrichment of H3K56ac close to the center of hotspots but this enrichment is specifically

occluded right at the center of the peaks (Fig. 3.2C) and a set of double peaks is formed that peak ~200-250 bp from the cut sites. However, the intensity of H3K56ac signals is not correlated with Spo11 oligos (Fig. 3.2D). It is unknown whether the intensities of H3K56ac and Spo11 are unlinked or the activity of meiotic DSB machinery across the genome is different in YMC from that in meiosis. Take together, the spatial correlation between H3K56ac and Spo11 oligos suggests possible action of the meiotic DSB pathway in the YMC of haploid cells.

Spo11 is required for H3K56ac peaks

Spo11 is the catalytic component required DSBs in meiosis (Bergerat et al., 1997) and H3K56ac marks the completion of DNA repair (Chen et al., 2008). To test the causality of Spo11 for H3K56ac peaks in YMC, we deleted *SPO11* in the CEN.PK strain. This strain retained the ability to cycle normally in the YMC. We then examined the H3K56ac level in the YMC by ChIP-qPCR. From late RC to RB phase, H3K56ac peaked at time point 2 (early OX phase) at *WHI5*, an example of well-defined meiotic DSB hotspots (Pan et al., 2011), consistent with ChIP-seq results (Fig. 3.4A,B). In the *spo11Δ* strain, we observed a significant (but incomplete) decrease of H3K56ac at time point 2, suggesting that Spo11 is required for the H3K56ac peaks seen in the wild-type YMC (Fig. 3.4B). A few more loci of typical meiotic DSB hotspots (Borde et al., 2000) were examined at time point 2 and showed similar decreases (Fig. 3.4C).

Spo11 is activated in mitosis of haploid yeast

Spo11 is known to function in meiosis; it is very interesting to investigate its function in mitosis given the observations above. From previous RNA-seq data, we note that *SPO11* and a lot of other genes in the meiotic DSB pathway (Prieler et al., 2005) are activated in the RC phase (Fig. 3.5A), suggesting a functional co-expression pattern. Spo11 protein is also detected by immunoprecipitation (Fig. 3.5B). To test the binding of Spo11 in the YMC, we performed ChIP

of Spo11 and H3K56ac across the YMC. qPCR of the *WHI5* locus shows that Spo11 is transiently recruited at time point 9 in the middle of RC phase (Fig. 3.5C). Simultaneously, H3K56ac at the *WHI5* locus starts to increase. These data fit the model that Spo11 makes the DSBs and H3K56ac marks the completion of repair.

Spo11 regulates mitotic recombination and heatshock resistance

Both mitotic and meiotic DSBs are associated with DNA recombination and exchange of genetic information between homologous chromosomes. To examine the physiological role of Spo11 in mitosis, we first examined DNA recombination by a LOH assay (Fig. 3.6A) in a *spo11Δ* mutant. A *URA3* gene is integrated at one homolog close to a DSB site (supported by both Spo11 oligos and H3K56ac ChIP-seq data) in an *ura3Δ0/ura3Δ0* diploid strain, making the strain sensitive to 5-FOA. If a DSB and the following recombination occur, the *URA3* gene is transferred between homologous chromosomes and in 50% of the cases one of the daughter cells will be 5-FOA resistant. Fig. 3.6B shows an example where we integrated *URA3* near the *ERGI* locus. We note that the LOH frequency in the CEN.PK background is higher than that in the S288C background. *SPO11* deletion causes a decrease of LOH in the CEN.PK strain but not in the S288C background. We further examined individual DSBs by integrating two markers flanking a DSB site (Fig. 3.7A). We consistently observed a higher frequency of LOH in the CEN.PK strain than the S288C strain; LOH frequency decreases in the CEN.PK *spo11Δ* strain but not in the S288C *spo11Δ* strain consistent with the model that Spo11 promotes this excess mitotic recombination in CEN.PK (Fig. 3.7B, C).

To examine whether the recombination contributed by Spo11 is physiologically and perhaps evolutionarily beneficial, we made a S288C and CEN.PK hybrid diploid so that numerous genetic variations which might provide a selective advantage be shuffled mitotically. We tested pure and

hybrid strains for heatshock resistance by incubating log phase yeast cells at 50°C for 30 min and calculating the survival rate. The CEN.PK strain has a much lower intrinsic heatshock resistance than the S288C strain (Fig. 3.8B). It may be related to the observation that RC stress response genes are expressed relatively higher in S288C than CEN.PK in log phase (Fig. 3.8A).

Interestingly, without heat tolerance training, the hybrid WT strain has a lower heatshock resistance than the hybrid *spo11Δ* strain probably because OX growth genes are expressed relative more highly in the CEN.PK strain and hybrid strain is adapted to the “CEN.PK_log” style under optimal growth condition. After growing at 39°C for ~200 division to try to select for increased resistance to heat shock, both hybrid WT and *spo11Δ* strains have a dramatic increase of heatshock resistance and fewer differences of heatshock resistance were observed (Fig. 3.8B).

H3K56ac ensures the completion of gametogenesis

H3K56ac is important for sporulation. The bulk level increases dramatically in the premeiotic S phase and both K56R and K56Q mutants decrease the sporulation efficiency (Recht et al., 2006). We found that H3K56ac is also critical for the completion of gametogenesis in CEN.PK strain, i.e. most of the spores failed to germinate. None of the H3K56R spores are viable after dissection of the H3K56R/+ heterozygote tetrads whereas both WT and H3K56Q spores are viable (Fig. 3.9A). Transforming a plasmid carrying H3WT or H3K56Q but not H3K56R into the H3K56R/+ heterozygotes rescues the inviable spore defect (Fig. 3.9B). We note that the inviable spores either fail to divide entirely or divide a few times and then arrest. These observations suggest that H3K56ac is not only important in the S phase of meiosis but also required from post meiosis to completion of gametogenesis. Given that the haploid K56R mutant grows just fine, the defect of inviable spores may be due to failure of repair of the DSBs by Spo11 in meiosis.

Discussion

In this study, we identified nonrandomly distributed H3K56ac peaks in a prototrophic yeast strain under nutrient limited condition. The mitotic H3K56ac sites highly resemble meiotic DSB hotspots. The meiotic-specific Spo11 contributes the H3K56ac peaks and it is shown to have an important role in DNA recombination and stress adaptation. We also discovered an important role of H3K56ac in gametogenesis.

H3K56ac marks DSB hotspots

The programmed meiotic DSB hotspots are well defined spatially in yeast (Buhler et al., 2007; Pan et al., 2011). A few histone modifications are associated with meiotic DSB hotspots, such as H3K4me3, H3K36me3, H3K79me3, H3K14ac and also nucleosome occupancy (Hansen et al., 2011). Some of the correlations were analyzed at low resolution and may be based doubtful correlations by coincidental localization to promoters (Tischfield and Keeney, 2012). The high level of correlation between H3K56ac and DSB hotspots from this study suggests that H3K56ac is a prominent marker of DSB hotspots genome-wide. It is easy to understand their causal relationship since it is well known that H3K56ac is involved in DNA damage response and ensures chromatin reassembly after DNA repair. The relative timing of Spo11 binding and H3K56ac reinforces the view of how and when H3K56ac functions in DNA damage repair.

Meiotic-specific DSB pathway contributes in mitosis

What is more surprising is the high similarity between the mitotic H3K56ac peaks and the meiotic DSB hotspots rather than mitotic ones in active growth. It is very likely that the meiotic DSB machinery and similar chromatin structure direct the distribution of DSBs in the RC phase of the YMC in the CEN.PK strain. Many factors are shared between meiotic and mitotic DSB repair pathways (Andersen and Sekelsky, 2010), supporting this hypothesis. This raises the

question whether the presumably meiosis-specific Spo11 also functions in mitotic DSB formation. *spo11* mutants have been previously shown to affect DNA recombination and reciprocal crossover in mitosis (Bruschi and Esposito, 1983; Lee et al., 2009). We also observed a significant decrease of LOH frequency in the CEN.PK strain (Fig. 3.6 and 3.7), reinforcing a role of Spo11 in mitosis under certain genetic and metabolic conditions. *SPO11* and other components of the meiotic DSB pathway are co-expressed in the RC phase and Spo11 protein is also detected (Fig. 3.5). CHIP of Spo11 and H3K56ac also supports an active role of Spo11 in mitosis (Fig. 3.5). However, many questions remain to be understood. What factors shape the meiosis-like DSB hotspots in mitosis? How strong is the DSB activity compared to meiotic DSBs? Is Spo11 the only or primary cutter for the programmed mitotic DSBs?

Do programmed DSBs and attendant recombination benefit yeast survival under stress conditions?

Compared to the spontaneous or induced DSBs in mitosis, meiotic DSBs are programmed to ensure accurate chromosome segregation. Interestingly, the DSBs in YMC seem to be programmed in the RC phase, a time when a starvation-like program is in effect. Meiosis could be viewed as an extreme case of starvation. We know that genetic information is reorganized dramatically during sexual reproduction. Is yeast programmed to “hijack” the meiotic DSB machinery to generate DSBs and recombination intentionally in face of stress conditions? Aged yeast has a higher recombination frequency than “young” yeast in the replicative lifespan paradigm (McMurray and Gottschling, 2003). Spo11 seems to be active under stress conditions and might regulate heatshock resistance under heat stress condition. We postulate that yeast uses the meiotic DSB machinery to increase genetic diversity in the face of stress conditions. The meiotic DSB machinery may be preferred because it is programmed and relatively controlled.

However, our preliminary investigations did not yet provide support for faster generation of heat resistant derivatives in a *SPO11*-dependent way.

Methods and Materials

Media

YPD medium contains 1% yeast extract, 2% bacto-peptone, 2% dextrose and 1.6 mM tryptophan. YPD supplemented with 200 µg/ml G418 or 100 µg/ml ClonNat or 300 µg/ml Hygromycin were used for drug resistance selection. The YMC medium is a minimal medium consisting of 5 g/L (NH₄)₂SO₄, 2 g/L KH₂PO₄, 0.5 g/L MgSO₄•7H₂O, 0.1 g/L CaCl₂•2H₂O, 0.02 g/L FeSO₄•7H₂O, 0.01 g/L ZnSO₄•7H₂O, 0.005 g/L CuSO₄•5H₂O, 0.001 g/L MnCl₂•4H₂O, 1 g/L yeast extract, 10 g/L glucose, 0.5 mL/L 70% H₂SO₄, and 0.5 mL/L Antifoam 204 (Sigma).

Strains and methods

Yeast strains were generated from CEN.PK background or S288C background including FY4 or BY4741, BY4742 and BY4743, and manipulated by standard budding yeast protocols.

Table 3.1. Yeast strains used in chapter 3.

Name	Background	Genotype
BY5764	CEN.PK	<i>MATa</i>
BY5765	CEN.PK	<i>MATa</i>
ZKY242	CEN.PK	<i>MATa HHT1-HHF1::hygMX, HHT2-HHF2::natMX</i>
ZKY497	CEN.PK	<i>MATa HHT1-HHF1 H3K56R::hygMX</i>
ZKY498	CEN.PK	<i>MATa HHT1-HHF1 H3K56Q::hygMX</i>
ZKY502	CEN.PK	<i>MATa HHT2-HHF2 H3K56R::natMX</i>
ZKY503	CEN.PK	<i>MATa HHT2-HHF2 H3K56Q::natMX</i>
ZKY531	CEN.PK	<i>MATa/a HHT1-HHF1 WT/H3K56R:hygMX, HHT2-HHF2 WT/H3K56R:natMX,ura3Δ0/ura3Δ0, [pJD154-HHT2-HHF2::URA3/CEN]</i>
ZKY532	CEN.PK	<i>MATa/a HHT1-HHF1 WT/H3K56R:hygMX, HHT2-HHF2 WT/H3K56R:natMX,ura3Δ0/ura3Δ0, [pJD154-HHT2-HHF2-H3K56R::URA3/CEN]</i>
ZKY533	CEN.PK	<i>MATa/a HHT1-HHF1 WT/H3K56R:hygMX, HHT2-HHF2 WT/H3K56R:natMX,ura3Δ0/ura3Δ0, [pJD154-HHT2-HHF2-H3K56Q::URA3/CEN]</i>

ZKY538	CEN.PK	<i>MATa spo11Δ::KanMX6</i>
ZKY542	CEN.PK	<i>MATa SPO11-3HA::kanMX6</i>
ZKY598	BY4742	<i>MATa ERG1_R::URA3, trp1Δ::hygMX, his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0</i>
ZKY599	BY4742	<i>MATa ERG1_R::URA3, spo11Δ::kanMX6, trp1Δ::hygMX, his3Δ0 leu2Δ0 lys2Δ0 ura3Δ0</i>
ZKY600	CEN.PK	<i>MATa ERG1_R::URA3, trp1Δ::hygMX, ura3Δ0</i>
ZKY601	CEN.PK	<i>MATa ERG1_R::URA3, spo11Δ::kanMX6, ura3Δ0, trp1Δ::hygMX</i>
ZKY630	BY4741	<i>MATa CCT6_L::TRP1, CCT6_R::URA3 trp1Δ::natMX, his3Δ0 leu2Δ0 met15Δ0 ura3Δ0</i>
ZKY631	BY4741	<i>MATa CCT6_L::TRP1, CCT6_R::URA3 spo11Δ::kanMX6, trp1Δ::natMX, his3Δ0 leu2Δ0 met15Δ0 ura3Δ0</i>
ZKY632	CEN.PK	<i>MATa CCT6_L::TRP1, CCT6_R::URA3 trp1Δ::natMX, ura3Δ0</i>
ZKY633	CEN.PK	<i>MATa CCT6_L::TRP1, CCT6_R::URA3 spo11Δ::kanMX6, ura3Δ0, trp1Δ::natMX</i>
ZKY662	BY4743	<i>MATa/a</i> WHI5_R::URA3/+, WHI5_L::TRP1/+, trp1Δ::hygMX/trp1Δ::natMX, his3Δ0/his3Δ0, leu2Δ0/leu2Δ0, lys2Δ0/+, ura3Δ0/ura3Δ0, met15Δ0/+
ZKY663	By4743	<i>MATa/a</i> WHI5_R::URA3/+, WHI5_L::TRP1/+, spo11Δ::kanMX6/spo11Δ::kanMX6, trp1Δ::hygMX/trp1Δ::natMX, his3Δ0/his3Δ0, leu2Δ0/leu2Δ0, lys2Δ0/+, ura3Δ0/ura3Δ0, met15Δ0/+
ZKY664	CEN.PK	<i>MATa/a</i> WHI5_R::URA3/+, WHI5_L::TRP1/+, trp1Δ::hygMX/trp1Δ::natMX, ura3Δ0/ura3Δ0
ZKY665	CEN.PK	<i>MATa/a</i> WHI5_R::URA3/+, WHI5_L::TRP1/+, spo11Δ::kanMX6/spo11Δ::kanMX6, trp1Δ::hygMX/trp1Δ::natMX, ura3Δ0/ura3Δ0
ZKY714	FY4xCEN.PK	<i>MATa/a SPO11-3HA::hygMX/SPO11-3HA::KanMX</i>
ZKY735	CEN.PK	<i>MATa SPO11-3FLAG::kanMX</i>
ZKY738	FY4xCEN.PK	<i>MATa/a spo11Δ::hygMX6/spo11Δ::kanMX6</i>

Gene knockout strains were generated by homologous recombination using PCR products containing a drug cassette (*kanMX* or *hygMX*) and 40 bp sequences flanking the target gene. Tagged-protein strains were generated similarly by integrating a cassette containing a protein tag and a drug resistance cassette at C terminus. A PCR product consisting of pFA6a-3HA-kanMX6 or pFA6a-3HA-*hygMX* with 40 bp sequences flanking *SPO11* at both ends was transformed into *MATa* or *a* CEN.PK strain and was selected for G418 or Hygromycin resistance. Integration of the tags and marker gene were validated by PCR using ~500 bp flanking primers. For the LOH assay, diploids were generated by mating a WT strain with appropriately marked strains. Histone

mutant strains were generated following the protocol described previously. *HHT1-HHF1* mutations were introduced into the genomic locus of the *MATa* strain (BY5764) and selected on YPD + 300 µg/ml Hygromycin. *HHT2-HHF2* mutations were introduced into the genomic locus of the *MATa* strain (BY5765) and selected on YPD + 100 µg/ml ClonNat. Haploid strains were mated and diploids were sporulated and dissected to select for haploids containing both *HHT1-HHF1*, *HHT2-HHF2* mutations.

Metabolic cycles

Metabolic cycle experiments were performed as previously described (Tu et al., 2005). The fermenter was from New Brunswick Scientific (BioFlo 3000). 10 mL overnight saturated culture was inoculated to start each YMC run. YMC runs were operated at an agitation speed of 475 rpm (BioFlo 3000), an aeration rate of 1 L/min, a temperature of 30°C, and a pH of 3.4 in 1 L YMC medium. Once the batch culture was saturated, at least 4 hour of starvation was performed. After starvation, fresh medium was added continuously at a dilution rate of $\sim 0.09\sim 0.1\text{ h}^{-1}$.

Immunoprecipitation

100 OD yeast cells were lysed in 1 ml lysis buffer (100 mM Tris-Cl pH 7.5, 100 mM NaCl, 50 mM NaF, 1 mM EDTA, 1 mM EGTA, 0.1% Tween-20, 10% glycerol, 1 mM PMSF, 5 µM pepstatin A, Roche protease inhibitor cocktail, 14 mM BME) by bead-beating on a paint shaker. Lysate was spun to get rid of cell debris and supernatant was incubated with 10 µl magnetic beads (Invitrogen) conjugated to 2 µg Flag M2 antibody (Sigma) for 3 hours at 4°C. Beads were washed 4 times with the same buffer and proteins were eluted in LDS sample buffer at 70°C for 10 minutes. Proteins were separated in 4-12% Bis-Tris gels followed by western blot against FLAG M2 antibody.

Chromatin Immunoprecipitation

Chromatin immunoprecipitation was performed as described previously (Cai et al., 2011). ~5 OD cells per time point were collected for ChIP of histone modification and ~50 OD cells per time point were collected for ChIP of Spo11. Antibodies used are list as following: H3K56ac (07-677), FLAG M2 (Sigma, F1804), HA 12CA5 (Roche, 11583816001). 3 µg primary antibody was used per ChIP experiment. Briefly, cells were first fixed in 1% formaldehyde at 25°C for 15 min and quenched in 125 mM glycine at 25°C for 10 min. Cells were pelleted and washed twice with buffer containing 100 mM NaCl, 10 mM Tris-Cl pH 8.0, 1 mM EDTA, 1 mM PMSF, 1 mM benzamidine•HCl before freezing. The frozen pellet was resuspended in 0.45 ml ChIP lysis buffer (50 mM HEPES•KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% deoxycholate (DOC), 1 mM PMSF, 5 µM pepstatin A, Roche protease inhibitor cocktail) and lysed by bead beating. Lysate from 50 OD cells were split into two tubes each containing 280 µl lysate and sonicated for 16 cycles (30 sec on, 1 min off, high output) using a Bioruptor (Diagenode, Denville, NJ). The supernatant of the sonicated lysate was pre-cleared. 50 µl lysate was saved as input. For ChIP with histone antibodies, 50 µl whole cell extract (WCE) was diluted 1:10 and used for each ChIP. For ChIP of Spo11, 500 µl WCE and 3 µg antibody was used. After incubation overnight, 50 µl protein G magnetic beads (Invitrogen, Grand Island, NY, 10003D) resuspended in ChIP lysis buffer was added and incubated for 1.5 h at 4°C. Beads were washed twice with ChIP lysis buffer, twice with DOC buffer (10 mM Tris•Cl pH 8.0, 0.25 M LiCl, 0.5% deoxycholate, 0.5% NP-40, 1 mM EDTA) and twice with TE. 125 µl of TES buffer (TE pH8.0 with 1% SDS, 150 mM NaCl, and 5 mM dithiothreitol) was added to resuspend the beads. Supernatant was collected after incubation at 65°C for 10 min. A second round of elution was performed and the eluates were combined. Reverse crosslinking was performed by incubation for 6 h at 65°C. An equal volume of TE containing 1.25 mg/ml proteinase K and 0.4 mg/ml glycogen was added to the samples after reverse crosslinking and samples were incubated for 2 h at 37°C. Samples were extracted twice with an equal volume of phenol and once with 25:1

chloroform:isoamyl alcohol. DNA was precipitated in 0.1 volume 3.0 M sodium acetate (PH 5.3) and 2.5 volume of 100% ice-cold ethanol at -20°C overnight. Pellets were washed once with cold 70% ethanol and resuspended in 20 µl TE. Fast SYBR Green Master Mix (Applied Biosystems, 4385612, Foster City, CA) was used for real-time PCR and experiments were done on the platform of StepOnePlus Real-Time PCR System (Applied Biosystems, 4385612, Foster City, CA).

RT-qPCR

2 OD cells from the cycle were collected and flash frozen. RNA was extracted using the Qiagen RNeasy Mini kit (QIAGEN, 74104, Valencia, CA) with the standard protocol. First strand cDNA was synthesized using SuperScript III First-Strand Synthesis System for RT-PCR (Invitrogen, 18080-051, Grand Island, NY). Oligo(dT)₂₀ primer was used for reverse-transcription. qPCR was performed as described above.

RNA-seq

15 OD cells were collected and flash frozen RNA was extracted using the Qiagen RNeasy Mini kit (QIAGEN, 74104, Valencia, CA) with the standard protocol. Library construction and sequencing were performed using the Hiseq platform supervised by the JHU HIT Microarray Core Facility. Briefly, Samples were run on the Agilent 2100 Bioanalyzer to ensure that high quality RNA was used. 1 µg total RNA was then prepared with the TruSeq™ RNA Sample Prep Kit (Illumina, San Diego, CA). mRNA was purified and fragmented before cDNA synthesis. cDNA was then end repaired and A-tailed. After adapter ligation, samples were PCR amplified and purified with AmpureXP beads (Agencourt, A63880, Brea, CA), then validated again on the Agilent 2100 Bioanalyzer. Differential gene expression analysis was performed with TopHat and Cuffdiff following a standard protocol (Trapnell et al., 2012). Briefly, the TopHat parameters

used were: --bowtie1 -i 40 --genome-read-mismatches 4 --no-coverage-search -G

Saccharomyces_cerevisiae_Ensembl_EF2/Saccharomyces_cerevisiae/Ensembl/EF2/Annotation/
Archives/archive-2012-03-09-08-22-49/Genes/genes.gtf

Saccharomyces_cerevisiae_Ensembl_EF2/Saccharomyces_cerevisiae/Ensembl/EF2/Sequence/Bo
wtieIndex/genome.

LOH assay

Diploid strains were first generated by mating a haploid strain with markers and a haploid strain without markers. Diploid strains were cultured in YPD medium for 2 days and plated on selected plates and YPD plates. Number of colonies were counted and rates were calculated by dividing the number of colonies on selected plates by the number of colonies on YPD plates.

Heatshock resistance

Log phase cells in YPD medium were incubated at 50°C for 30 minutes and then plated on YPD plates. The same cells without heatshock were plated on YPD plates for normalization. Numbers of colonies were counted and rates were calculated by dividing the number of colonies with heatshock by the number of colonies without heatshock. Heat tolerance evolution was done by incubation of yeast cells at 39°C in YPD medium. Every two days, saturated cultures were diluted into fresh YPD medium to an A_{600} of 0.1. To test heat shock resistance of cells during heat tolerance evolution experiments, cells at 39°C were first inoculated in YPD medium at an A_{600} of 0.1 and incubated at 30°C for 24 hours. The saturated cultures were inoculated again in YPD medium at an A_{600} of 0.1 and entered into log phase after 4-5 hours and then heat shocked at 50°C for 30 min.

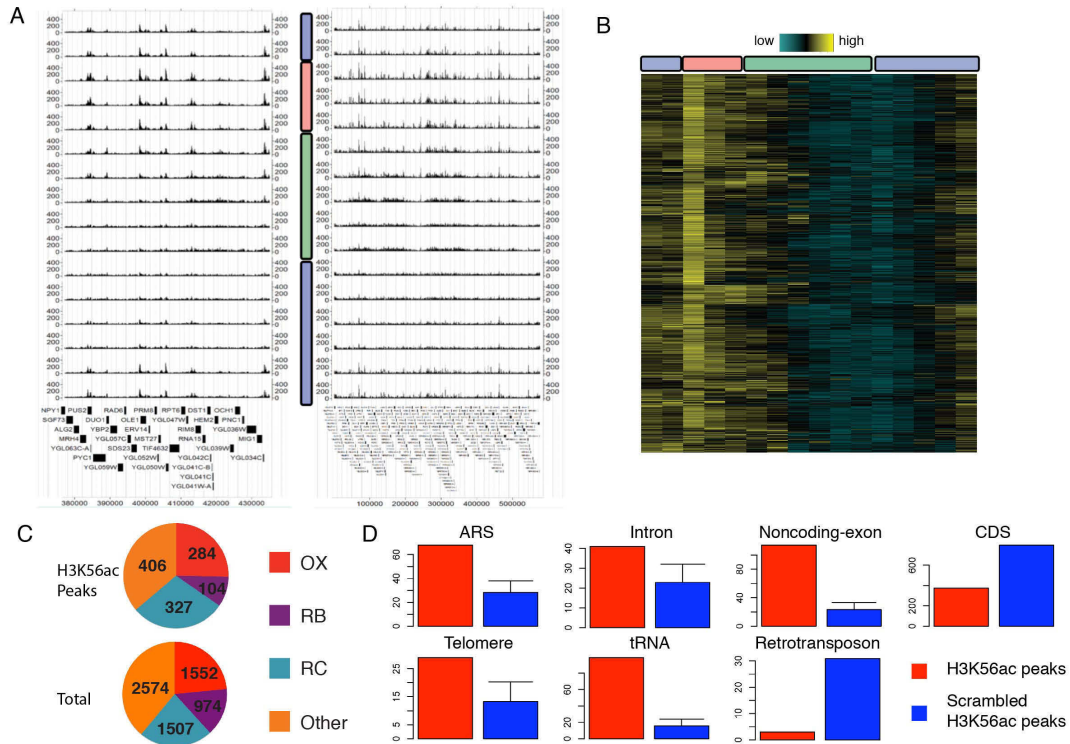


Figure 3.1 Spatial and temporal patterns of H3K56ac in YMC. (A) 16 time point ChIP-seq of H3K56ac is displayed in CisGenome browser (Ji et al., 2008) from top to bottom. The left is chrVII: 377000-436000 and the right is the entire chrV. The vertical bar represents the three phases: Red = OX phase, Green = RB phase, Blue = RC phase. (B) 1166 H3K56ac peaks are detected and the temporal pattern is displayed in a heat map. The blue-yellow gradient represents H3K56ac peak signals normalized to total H3. (C) Genes closest to the H3K56ac peaks are categorized based on their expression peaks in the three phases of YMC. Total yeast genes are categorized based on expression peaks as control. The numbers of genes expressed in each phase or not cycling are included in the pie chart. (D) Genomic features overlapping H3K56ac peaks or scrambled peaks are counted and compared in barplots. Scrambled peaks were generated by randomly picking regions in the genome with the same length of H3K56ac peaks. Scrambling was repeated 100 times and the mean and sd were calculated. Error bars represent mean+1.96*sd.

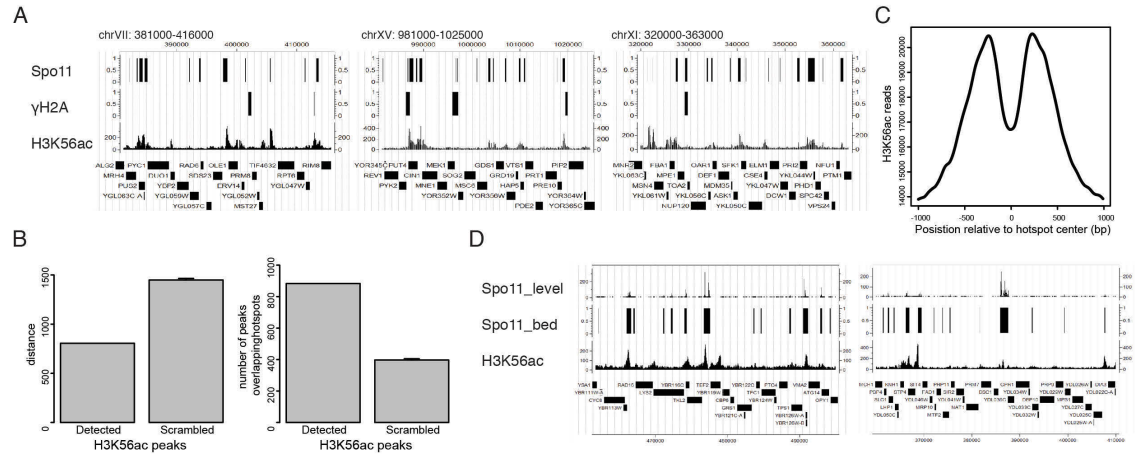


Figure 3.2 H3K56ac peaks correlate with meiotic DSB hotspots but not mitotic DSB hotspots. (A) Spatial distribution of meiotic and mitotic DSB hotspots and H3K56ac at three regions. Meiotic hotspots are from mapping Spo11-bound oligos in meiosis and mitotic hotspots are from mapping γ H2A ChIP-seq reads in mitotic cells. H3K56ac t3 ChIP-seq result is shown here. (B) H3K56ac peaks correlate with Spo11-bound hotspots. Distance between H3K56ac peaks and Spo11 hotspots is shown in the left bar graph and number of H3K56ac peaks overlapping hotspots is shown in the right bar graph. (C) Distribution of H3K56ac forms double peaks around the centers of Spo11-associated hotspots. (D) The intensity of H3K56ac doesn't correlate with Spo11-bound oligo counts. Spo11_level track shows the sequencing counts of Spo11-bound oligos. Spo11_bed track shows the derived hotspots. H3K56ac track shows ChIP-seq signals at t3.

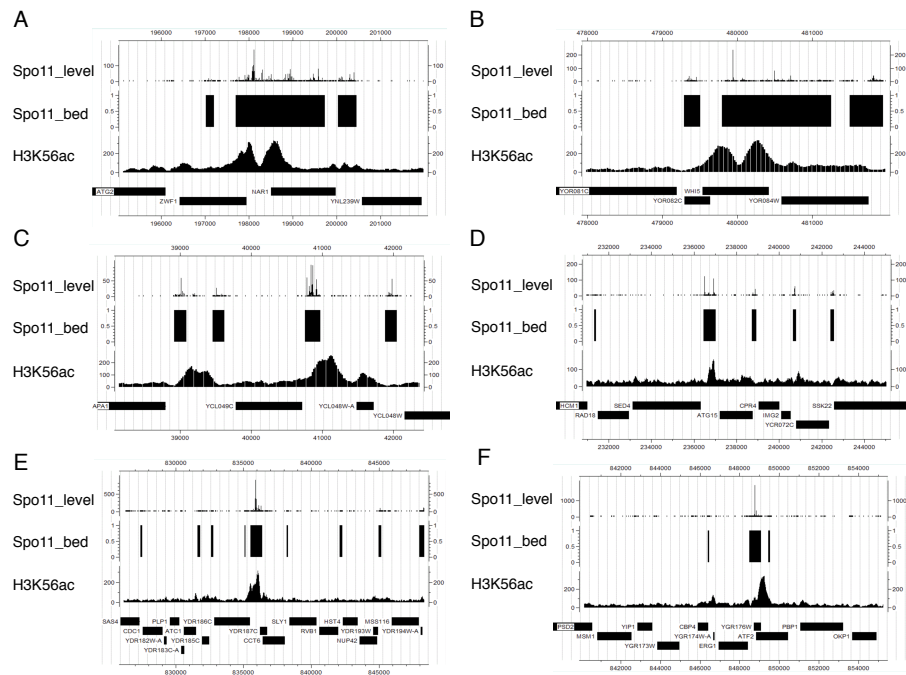


Figure 3.3 H3K56ac peaks correlate with meiotic DSB hotspots at example genes. (A-F) are *ZWF1*, *WHI5*, *YCL049C*, *ATG15*, *CCT6*, *ERG1*.

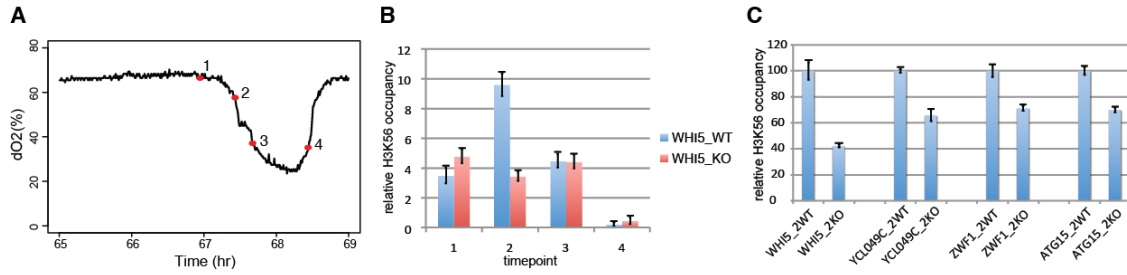


Figure 3.4 Spo11 is required for H3K56ac peaks. (A) 4 time points of WT and *spo11del* cells were collected from late RC to RB phase for H3K56ac ChIP. (B) ChIP-qPCR of H3K56ac at *WHI5* locus. H3K56ac is significantly decreased at t2 in *spo11del*(KO) cells. (C) H3K56ac is significantly decreased at *WHI5*, *YCL049C*, *ZWF1*, *ATG15* in *spo11del*(KO) cells. ChIP-qPCR results at t2 were shown here.

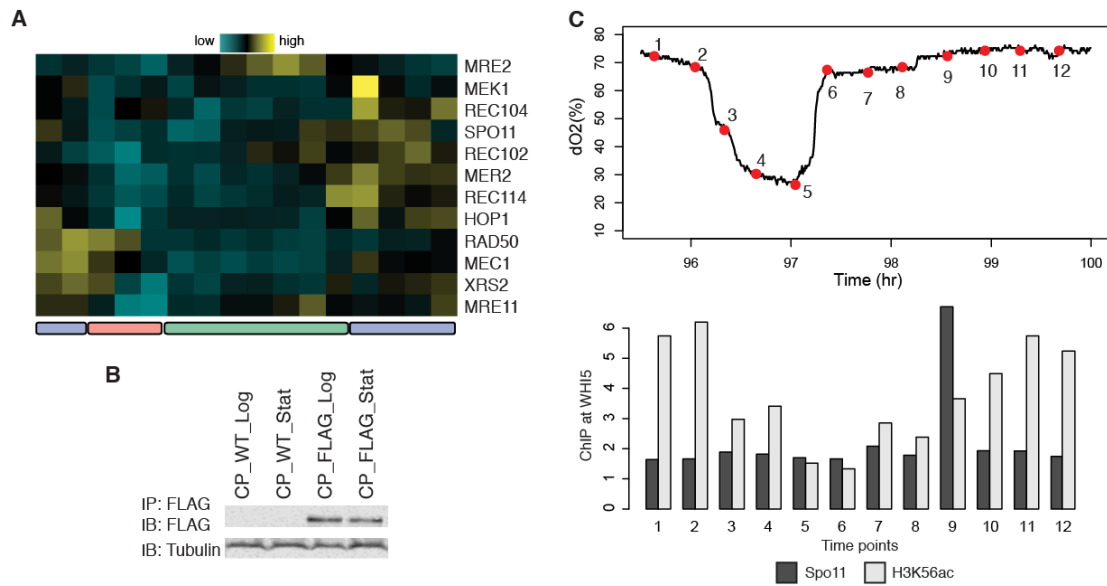


Figure 3.5 Spo11 is activated in the RC phase of YMC. (A) *SPO11* and other components in the DSB pathways are co-expressed in YMC. FPKM values from previous RNA-seq were displayed in the heat map. (B) Immunoprecipitation of Spo11-3FLAG in log and stationary phase of CEN.PK (CP) strain. (C) 12 time points of Spo11-3HA cells were collected across one YMC for ChIP of HA and H3K56ac. ChIP-qPCR results at the *WHI5* locus are shown in a bar graph.

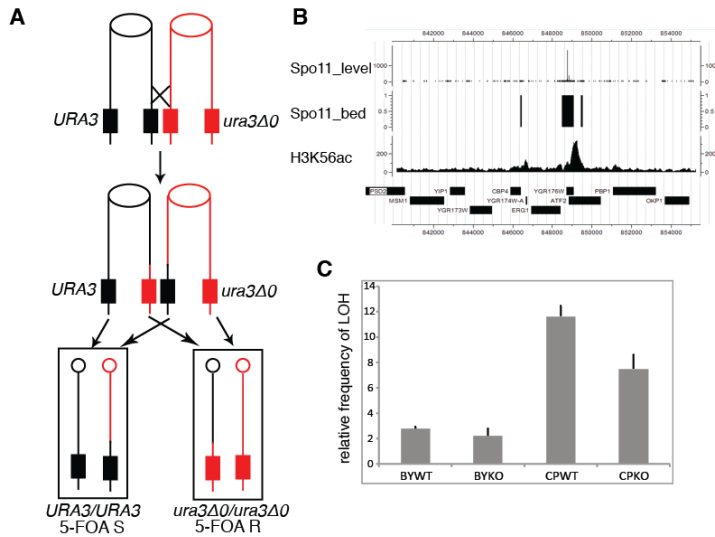


Figure 3.6 Deletion of *SPO11* causes decrease of homologous recombination. (A) Scheme of the LOH assay. In WT or *spo11/spo11* diploid strain, *URA3* gene is integrated at one homolog chromosome, close to a hotspot (distal to the centromere). The strains are 5-FOA sensitive. After 2 days' batch culture, cells were plated on SC+5-FOA plates and 5-FOA resistant colonies were counted. (B) *URA3* was integrated at chrVII: 850646, ~1.6 kb away from the *ERG1* hotspot (distal to CENVII) and LOH was examined in BY4743 or CEN.PK WT and *spo11/spo11* diploid strains (mating products of ZKY598, ZKY599, ZKY600, ZKY601 with corresponding WT strains). Top panel shows the intensity of Spo11 oligos and H3K56ac at *ERG1* and the bottom panel shows the relative frequency of 5-FOA resistant colonies. Three replicates per strain were averaged and error bars represent s.d.

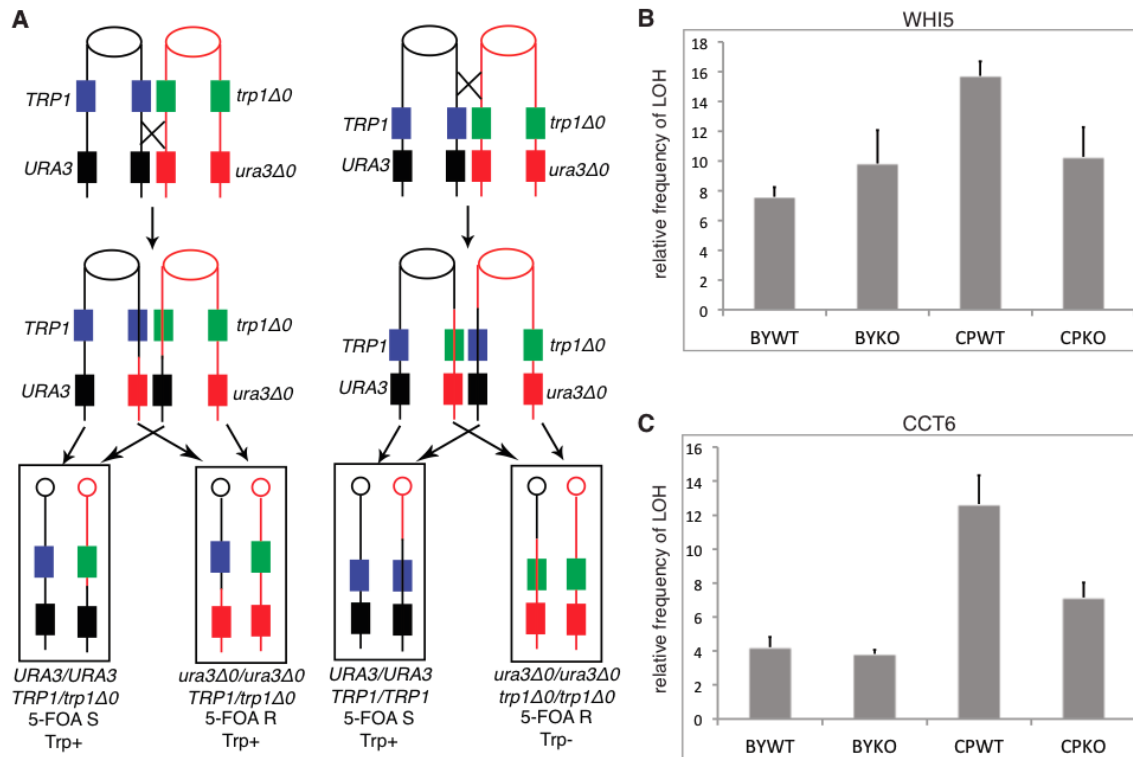


Figure 3.7 LOH assay mapping individual homologous recombination. (A) Scheme of the LOH assay. In WT or *spo11/spo11* diploid strain, *URA3* and *TRP1* are integrated at one homolog chromosome close to a hotspot (*URA3* is distal to the centromere and *TRP1* is proximal). The strains are 5-FOA sensitive. After 2 days' batch culture, cells were plated on SC-trp+5-FOA plates and numbers of 5-FOA resistant colonies measure the recombination frequency due to the targeted hotspot. (B) LOH assay at *WHI5*. Strains were ZKY662, ZKY663, ZKY664, ZKY665. *TRP1* was integrated at chrXV: 476865, ~3 kb from the DSB site; *URA3* was integrated at chrXV: 487090, ~7 kb from the DSB site. (C) LOH assay at *CCT6*. Strains were ZKY630, ZKY631, ZKY632, ZKY633. *TRP1* was integrated at chrIV: 830397, ~5.6 kb from the DSB site; *URA3* was integrated at chrIV: 855491, ~19 kb from the DSB site.

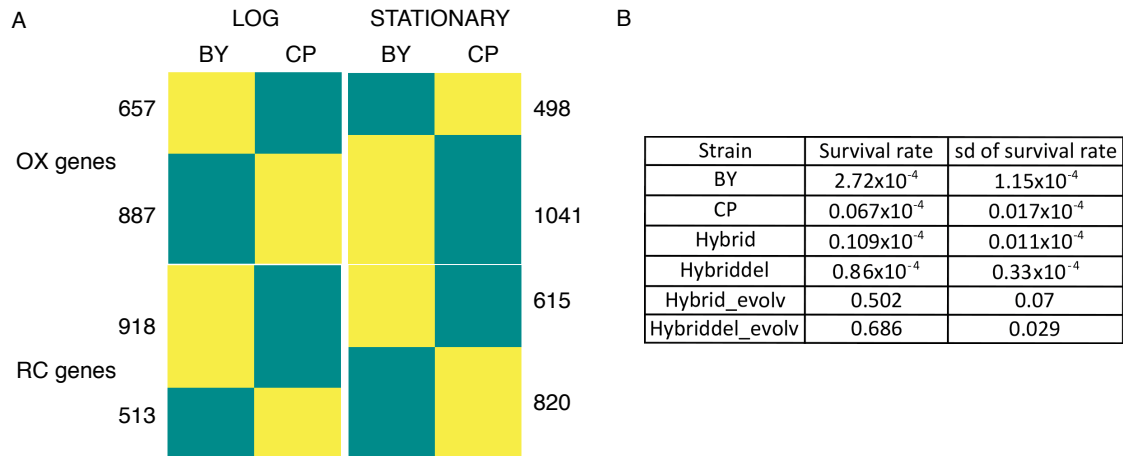


Figure 3.8 Evolution of heat shock tolerance in S288C and CEN.PK hybrid background.

(A) OX phase growth genes and RC phase stress response genes are differentially expressed in log and stationary phase of BY4741 and CEN.PK strains. Yellow means higher expression level and cyan means lower expression level relative to the overall expression of BY4741 and CEN.PK strains. Numbers of genes in each category are listed besides the heat maps. (B) Survival rates after 30 minutes' heat shock at 50°C using log phase yeast cells. Evolution was performed under 39°C for 60 divisions.

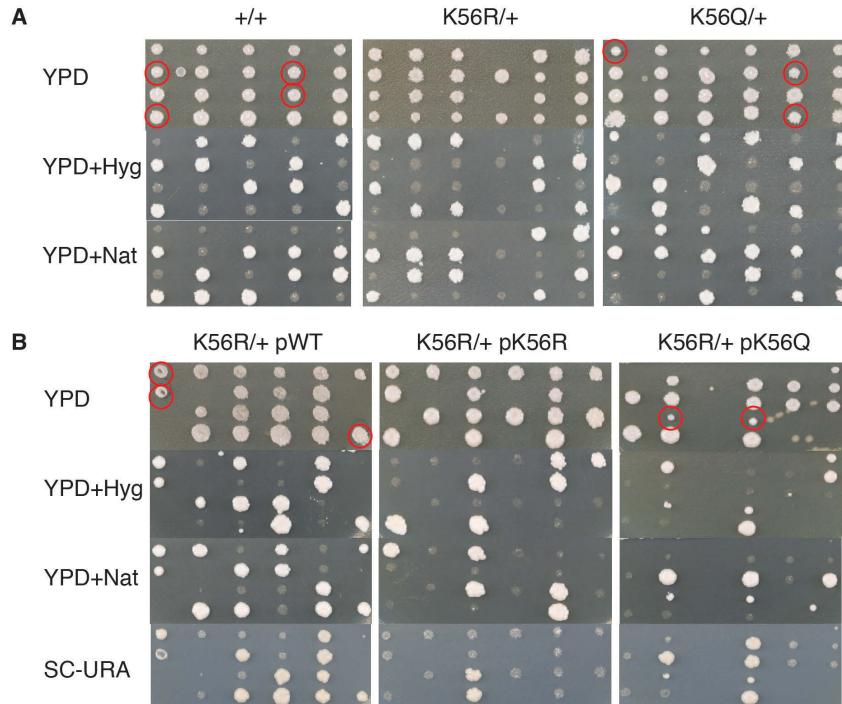


Figure 3.9 H3K56ac ensures normal spore viability. (A) H3K56R mutants don't germinate in Cen.PK strain. Tetrads of heterozygous diploids (mating products of ZKY242xBY5765, ZKY497xZKY502, ZKY498xZKY503) were dissected and replicated on YPD+Hygromycin and YPD+clonNat plate. The diploids contain a native copy and a synthetic copy of *HHT1-HHF1* and *HHT2-HHF2* genes. Synthetic *HHT1-HHF1* WT/H3K56R/H3K56Q are associated with a *hygMX* cassette and synthetic *HHT2-HHF2* WT/H3K56R/H3K56Q are associated with a *natMX* cassette. Spores with both hygromycin and clonNAT resistance are marked by red circles. (B) Lethality of H3K56R spores is rescued by expressing WT or H3K56Q protein. A *CEN* plasmid carrying either WT, H3K56R, or H3K56Q cassette was transformed into the H3K56R heterozygous diploid strain (ZKY531, ZKY532, ZKY533) and tetrad dissection was performed.

Chapter 4

Predicting dynamic activity of transcription factors through time-course

ChIP-seq data of histone modifications

Abstract

To gain an in-depth understanding of the role of transcription factors, it is crucial to study their dynamics during various biological processes. Chromatin immunoprecipitation is widely used to determine the transcription factor-DNA binding, but it is tedious to interrogate hundreds of these in a dynamic process. Here, we developed a computational pipeline to predict the dynamic activity of transcription factors via integrating transcription factor motif sites with time-course ChIP-seq data of histone modification. We screened 177 transcription factors in the yeast metabolic cycle and found that 55 of them are enriched at specific stages. We further validated the binding and function of eight transcription factors in this process. This method is a valuable tool to study the function of transcription factors during dynamic processes.

Introduction

Transcription factors (TFs) are a class of proteins that bind to the promoter region of targeted genes, and regulate their expression. Hundreds of TFs have been identified from yeast to mammals. The accuracy and specificities of transcription regulation are achieved through combinatorial and dynamic interactions between TFs and chromatin and among different TFs and their co-regulators (Hughes and de Boer, 2013; Lemon and Tjian, 2000). The functions of most TFs are context-dependent (Chua et al., 2004). For example, when phosphate is sufficient, Pho4 is retained in the cytoplasm. However, when phosphate level is low, Pho4 accumulates in the nucleus and up-regulates the expression of phosphate-responsive genes (O'Neill et al., 1996). Mutations in TFs may dramatically affect gene expression upon certain specific perturbations, though the effects under a steady state may be minimal (Aragon et al., 2008; Chua et al., 2006). Therefore, the dynamics and context of TF binding and activity are important for its physiological function. TF binding dynamics can be determined by coupling time course chromatin immunoprecipitation (ChIP) with PCR or deep sequencing (ChIP-seq) (Odom, 2011; Robertson et al., 2007). However, screening hundreds of TFs by ChIP is laborious and costly.

TFs usually recognize certain patterns of DNA sequences, known as TF binding motifs (Hughes and de Boer, 2013). One way to characterize a binding motif is through consensus sequence, which corresponds to the most probably nucleotide sequence, usually 6-30 bp, bound by a TF. Another way to describe a motif is through a probability or position-specific weight matrix, which describes the probability or likelihood each nucleotide occurs at each position within the motif. Often these show symmetry of binding reflecting the fact that many TFs bind as dimers or higher order multimers (Hellauer et al., 1996; Thukral et al., 1991). If the binding motif of a TF is known, TF binding sites can be predicted computationally by mapping the motif to the genome (Ji and Wong, 2006). This method, however, often has high false discovery rate and its predictions

are static instead of dynamic. Both of these issues may be tackled by incorporating data on chromatin states and structures, including histone occupancy, histone modifications and DNase sensitivity, etc. (Cheng et al., 2011; Pique-Regi et al., 2011; Wei et al., 2013). To the best of our knowledge, this approach currently is primarily used to predict TF binding in a static condition or pairwise comparisons of different conditions.

Here, we propose to test the strategy in a dynamic temporal process. We use the yeast metabolic cycle (YMC) as a model system. In YMC, synchronized yeast culture undergoes three distinct metabolic phases, oxidative (OX), reductive building (RB) and reductive charging (RC), defined by the oxygen consumption rate (Figure 4.4A) (Tu et al., 2005). More than half of the genes in the *Saccharomyces cerevisiae* genome are expressed periodically in the YMC, with distinct functional categories of genes activated at different phases. This suggests that multiple TFs are involved in controlling the specificity and accurate timing of gene expression. As an initial step to test whether we could predict which TFs are involved in this process and how each TF dynamically binds to its target sequences in a time course, we integrated 177 TF motifs from the JASPER database (Sandelin et al., 2004) and a time course ChIP-seq dataset of H3K9ac in the YMC with high temporal resolution. This approach identified 55 TFs with motif sites enriched at specific stages of the YMC. For each TF, its binding sites and temporal binding patterns are also predicted. Using ChIP-qPCR, we validated our predictions for 8 TFs at 12 motif binding sites per TF.

Results

Prediction Pipeline

The pipeline of this method is illustrated in the flowchart shown in Fig 4.1. Position Specific Frequency Matrices (PSFM) or consensus sequences of TFs are mapped to the genome using the

CisGenome software package (Ji et al., 2008) to obtain TF motif sites. Histone modification ChIP-seq peaks are detected by four different peak-calling methods, BayesPeak, iSeq, PICS and SPP (Kharchenko et al., 2008; Mo, 2012; Spyrou et al., 2009; Zhang et al., 2011). There are many other peak callers (Ji et al., 2008; Qin et al., 2010; Rozowsky et al., 2009; Valouev et al., 2008; Zhang et al., 2008). We chose the four because we wanted to deliver our pipeline through an R package and these peak callers are available in R at the time when we developed our approach. Peak regions at different time points are consolidated into one comprehensive list by merging overlapping peak regions into one peak region. Motif sites at this step contain a large proportion of false positive hits because the short motif sequences are usually not specific enough, given a large and complex genome. Many histone modifications are enriched at promoter regions and mark active transcription (Kouzarides, 2007; Li et al., 2007). Therefore, motif sites are first filtered with a requirement for overlapping histone modification peaks. Given the assumption that TF binding/activity correlates with histone modifications, we further filter motif sites using a Naïve Bayes classifier. A Naïve Bayes classifier is a probabilistic model to categorize a variable (motif sites here) with a few outcomes (background or real binding sites), conditional on other features (the distributions of histone modification signals at motif sites and background histone modification signals here). Filtered motif sites of all TFs are pooled and K-means clustering of histone modification signals at these motif sites is performed to capture different temporal patterns during the dynamic process. Based on the second assumption that important TFs are dynamically regulated in the process, a statistical test is performed to select TFs with motif sites enriched in any cluster produced by the previous K-means clustering. Additionally, if an expression profile is available, correlations between each cluster and TF expression is calculated to help predict the temporal pattern of TF activity. Details of the algorithm can be found in the Methods and Materials section.

Motif mapping

Multiple databases of yeast TF motifs exist, such as TRANSFAC, JASPAR, UniPROBE (Matys et al., 2003; Newburger and Bulyk, 2009; Sandelin et al., 2004). JASPAR was used here as an example because it contains a non-redundant binding profile for each TF. 177 PSFMs corresponding to 177 TFs were mapped to the yeast genome; 175 of these had at least one motif site except Pho2 and Fzfl. Therefore, we proceeded with the remaining 175 TFs.

Peak calling

In a previous study, we collected 16 time points ChIP-seq data of histone acetylations and methylations across one yeast metabolic cycle (Figure 4.3A and Chapter 2). Here we used H3K9ac as an example because it has been found to localize at only a subset of genes and shown higher dynamics than other modifications. We hoped to identify TFs dynamically associated with H3K9ac. Four peak-callers were used and peaks detected from 16-time point ChIP-seq experiments were pooled into a single list for each peak-caller. Numbers of peaks called by each peak caller are shown in Table 4.1. Assuming that roughly half of yeast genes (of ~6000 genes in total) have H3K9ac peaks, BayesPeak and SPP detected comparable numbers of peaks.

Filtering

The first filtering step was to select motif sites that overlap with H3K9ac peaks. We tried to count overlap with either motif sites directly or 200 bp windows centered at motif sites. If a motif site overlaps with a H3K9ac peak detected by peak caller A, we can say that this motif site is recognized by peak caller A. Table 4.2 summarized the numbers of motif sites recognized by each peak caller. We noticed that extending motif sites to 200 bp windows increased the numbers of motif sites selected. iSeq is most sensitive to such extension and PICS is least sensitive, probably reflecting the property of peaks defined by them. Another criterion is whether a motif site is recognized by multiple peak callers, which is summarized in Table 4.3. About 10% of

motif sites are recognized by at least 2 peak callers and only ~1% of total motif sites are recognized by all 4 peak callers, suggesting that the similarity of output peaks decreases very fast with increasing number of peak callers. To further address the issue of similarity among peak callers, we examined the percentage of motif sites uniquely recognized by each peak caller, which is shown in Table 4.4. If a motif site is picked by only one peak caller, it tends to be less conserved across different peak-calling methods. The very different performance of these peak callers shown above is probably related to how they handle the H3K9ac data. It is notable that different peak-calling methods may have variable capacities of dealing with different types of ChIP-seq data (TF/Histone modification) (Kharchenko et al., 2008; Mo, 2012; Park, 2009; Spyrou et al., 2009; Zhang et al., 2011).

Next, we examined the distribution of H3K9ac signals at selected motif sites. Taking Abf1 as an example, the distribution of H3K9ac at all Abf1 motif sites is almost the same as the distribution of the genome background and it is significantly right-shifted after filtering of overlapping H3K9ac peaks (Figure 4.2). It is also clear that different peak callers or combinations lead to different levels of distribution separation. By recognition of at least 2 peak callers, the distribution is separated from the background distribution.

To determine an automatic cutoff to separate H3K9ac at selected motif sites and genomic background, a Naïve Bayes classifier was implemented. Based on the classifier, a working False discovery rate (FDR) can be computed. The working FDR is on average 5%-10% for different TFs by applying the Naïve Bayes classifier. The numbers of motif sites obtained by such filtering are shown in Table 4.5.

A majority of TFs bind to the promoter regions that lie near (i.e. within 500 bp usually in yeast) transcriptional start sites (TSSs). To examine whether the two-step filtering restricts motif sites to

the promoters, we overlaid the positions of motif sites near the closest TSSs. Filtering by each peak caller confines the selected motif sites near the TSSs at certain level and combining peak callers increase the level of the enrichment (Figure 4.3A). We note that each step of filtering differentially contributes to the enrichment of motif sites depending on how H3K9ac peaks were obtained (Figure 4.3B). When the first step led to a proper level of enrichment (BayesPeak, 2 or 3 peak callers), the second step didn't change much. When the enrichment was lower after the first step, the second step increased the enrichment (iSeq, PICS, 1 peak caller).

Together, these motif sites were strongly associated with H3K9ac level and promoter regions, and therefore used for downstream analysis of temporal patterns.

Clustering

Motif sites from the “two or more peak caller” method were used for K-means clustering to capture the temporal patterns of H3K9ac at motif sites. A few statistics were used to estimate the number of clusters (see Methods and Materials) and most of them generated two or three clusters. Figure 4.4B shows the temporal pattern of 3 clusters. Clear OX and RC patterns were readily observed in clusters 1 and 3.

If one TF is dynamically regulated, it is very likely its motif sites would be enriched in particular clusters. Statistic tests were performed for all TFs to identify dynamically regulated TFs. 54 out of 525 TF_cluster combinations (175 TFs multiplied by 3 clusters) were enriched significantly ($FDR < 0.05$). Selected motif sites and clustering information were produced for downstream confirmation. In addition, correlation between time-course H3K9ac in each TF_cluster pair and TF expression was provided for reference. Consistent with the observation that TF expression doesn't correlate with expression of target genes, out of 31 cluster-enriched TFs ($FDR < 0.01$), only 7 show modest correlation in the same cluster (correlation > 0.3).

Confirmation

We applied both computational and experimental methods to confirm the predicted motif sites. An independent dataset of 203 TF ChIP-chip in yeast was used for cross validation, although the conditions were not same (log phase/YMC) (Venter et al., 2011). Overlapping TFs from the two studies were screened and for most TFs (27 out of 32 TFs), ChIP-chip validated target genes were significantly enriched in the selected motif sites (Figure 4.5).

To directly confirm the predicted motif sites and temporal recruitments/activities, we randomly selected 8 TFs, as shown in Table 4.6. 4 of these 8 were enriched in certain clusters and the others were not. For these 8 TFs, 9 motif sites passing the filtering steps and 3 motif sites not passing were tested. The eight TFs were tagged with 3xHA or 3xFLAG at the C-terminus of their ORFs. Individual or combinatorial tagged strains were run in YMC and cells were collected at 6 time points across one cycle for ChIP. Surprisingly, TFs enriched in clusters show various phenotypic defects. C-terminal tagged Msn2 could not be constructed in this strain background. Xbp1-3HA failed to cycle in the YMC, as did the double-tagged strain, Swi5-3HA Gcr1-3FLAG, although Swi5-3HA and Gcr1-3FLAG individually could cycle. None of the unenriched TFs showed any defects in YMC. Tagging a protein is a mild modulation and could perturb the protein's function at certain level. Tagging of the cluster-enriched TFs didn't affect the growth but disrupted cycling capacity, suggesting that they have critical roles in driving the YMC. Cluster-enrichment analysis could potentially predict additional key TFs which exert particular roles in dynamic processes like the YMC.

For the cycling strains, ChIP-qPCR was performed to examine the spatiotemporal binding/activities. 4 out of 9 motif sites show higher Swi5 binding than control motif sites (Figure 4.6A). Either one or zero motif sites show increased binding of Arg81, Oaf1 or Cin5 compared to

control sites. The 4 motif sites with higher Swi5 binding were further examined in a time course (Figure 4.6B). Target 2 is close to the promoter of *HOM2*, an OX gene, consistent with the H3K9ac pattern in cluster 1. However, Swi5 recruitment peaks at late RC/early OX phase, which is slightly ahead of the transcription peak. Similarly, target 9 is close to *ALD4*, a RC gene, but Swi5 is recruited at RB and early RC phase, also preceding the transcription peak. The order is consistent with the previous findings (Featherstone, 2002). Targets 4 and 5 are located at dubious ORFs, which are not discussed here. The observation suggests that the temporal pattern of H3K9ac could be a good indicator of expression timing of TF's targets but the recruitment of TF itself may be shifted due to its particular function or regulatory mechanism.

Discussion

In this study, we presented a method to identify important TFs and their binding sites, and to predict their temporal recruitment/activity during dynamic processes. By using high-temporal-resolution H3K9ac ChIP-seq data across YMC, we identified 54 of 175 yeast TFs that were possibly regulated at specific phases. Computational cross validation and direct experiments confirmed good performance of this method.

Many chromatin surrogates have been used to identify confident TF binding sites from motif sites over the genome with differential prediction power (Wei et al., 2013). Some surrogates are unbiased, such as DNase I sensitivity, DNA methylation changes, and pol II occupancy. Others, like histone modification and its combination can be more specific to different groups of genes. Both are very useful for prediction: unbiased surrogates generate a more complete list whereas specific surrogates offer hints of mechanisms. Choosing proper surrogates will greatly facilitate the analysis.

Most TFs seemingly have few critical functions under steady state conditions – perhaps this is the result of redundancy (Hughes and de Boer, 2013). They are recruited and activated only upon specific stimuli or under particular dynamic processes like development and differentiation. Dynamic analysis of TF binding and activity at motif sites is very useful to understand their functions (Ni et al., 2009). Similar to static analysis, dynamic chromatin surrogates can be used to infer the dynamics of TFs. The specific roles of different TFs affect how to infer the temporal pattern of TF from the pattern of surrogates. Whether a TF is a repressor, or it functions in presetting the promoter, initiation or elongation could shift its temporal pattern forward or backward. Temporal patterns of different surrogates are different too, especially when high temporal resolution is available. They could appear at any stage of the transcriptional cycle. Careful interpretation is critical when making predictions of this kind.

Methods and Materials

Datasets

Yeast TF PSSM were downloaded from JASPAR website (Sandelin et al., 2004). H3K9ac ChIP-seq and RNA-seq data was generated from my previous study (Chapter 2). 203 yeast TF ChIP-chip data were taken from a study by Pugh’s group (Venter et al., 2011).

Motif mapping

TF PSSMs were mapped to the genome by CisGenome package (Ji et al., 2008) using “motifmap_matrixscan_genome” function. Other tools can also be used to generate the coordinates of motif sites.

Peak calling

4 peak calling methods were used in this study, BayesPeak, iSeq, PICS, SPP (Kharchenko et al., 2008; Mo, 2012; Park, 2009; Spyrou et al., 2009; Zhang et al., 2011). All of them have R packages which can be directly incorporated in our pipeline. 16 time points ChIP-seq reads were independently used for peak calling and peak regions were combined into one list by merging overlapping peaks into one single peak.

Filtering motif sites that overlap H3K9ac peaks

Motif sites were either expanded from the centers symmetrically to 200 bp windows or not extended (ori). Then we retained motif sites that have at least 1 bp overlap with the H3K9ac peaks.

Naïve Bayes classification

The whole genome was divided into concatenated 200 bp windows and the numbers of H3K9ac ChIP-seq reads were calculated at each window. The 60806 H3K9ac read counts form the background distribution. Total or peak-overlapping motif sites for each TF were extended to 200 bp windows and numbers of H3K9ac ChIP-seq reads were calculated at each window. The probably density functions of the background distribution and the distribution of read counts at filtered motif sites were determined by kernel density estimation. The kernel density estimator is

$$\hat{f}(x) = \frac{1}{n} \sum_{i=1}^n K_h(x - x_i).$$

The normal distribution function was used here for the kernel function K_h . The bandwidth parameter h was estimated by the mean integrated squared error:

$$\text{MISE}(h) = E \int (\hat{f}_h(x) - f(x))^2 dx.$$

With the background density function $\hat{f}_0(x)$ and density function of filtered motif sites $\hat{f}(x)$, the read count classifier $N_{\text{classifier}}$ for individual TF at each time point was calculated as the smallest integer N which satisfies the criterion below:

$$N_{motif} \hat{f}(N) \geq N_0 \hat{f}_0(N)$$

N_{motif} is the number of motif sites which overlap H3K9ac peaks from previous step and N_0 is the total number of genomic windows, which is 60806 in this study. Motif sites with read counts more than $N_{classifier}$ for each time point were combined across all time points for the downstream clustering.

K-means clustering

Motif sites passing the Naïve Bayes classifier were combined together across all TFs and read counts at 16 time points were implemented for K-means clustering. Number of clusters were determined by trying different types of information criteria: Akaike information criterion (AIC), Bayesian information criterion (BIC), average silhouette width (asw) or Calinski-Harabasz (ch).

RNA correlation

RNA-seq data with the same time course as H3K9ac was incorporated from the previous study (Chapter 2). H3K9ac read counts at motif sites from the same cluster were averaged and correlations between H3K9ac and TF RNA level were calculated for each cluster of each TF.

Validation with TF binding map

Three probes for each gene corresponding to upstream activating sequences (UASs), transcriptional start sites (TSSs) and 3' end of open reading frames (ORFs) were assayed by 203 TF ChIP-chips at 25°C and 37°C in YPD (Venter et al., 2011). Although the conditions used were not same, it could still serve as a validation of our predictions. Genes passing the 5% FDR cutoff thresholds at 25°C and 37°C of UAS and TSS were combined for each TF. Genes closest to the predicted motif sites with distances of TSSs and motif sites less than 500 bp (predicted genes) were used for intersection with genes reported from ChIP-chip. Total motif sites were randomly

sampled 100 times and the closest genes (control genes) were intersected with targeted genes. Numbers of ChIP-chip genes overlapping with predicted genes or control genes were compared by bar graphs and 95% confidence intervals for the control category were plotted.

Yeast protein tagging

Tagged-protein strains were generated by integrating a cassette containing a protein tag and a drug resistance cassette at the C terminus. PCR products consisting of pFA6a-3HA-kanMX with 40 bp sequences flanking the *SWI5*, *XBPI*, *ARG80* and *OAF1* stop codons at both ends were transformed into MATa CEN.PK strain and were selected for G418 resistance. Integration of the 3HA tag and marker gene were validated by PCR using ~500 bp flanking primers. Similarly, PCR products of pFA6a-3FLAG-hygMX were transformed into HA tagged strains or WT strain and selected for Hygromycin resistance, followed by PCR validation.

Yeast Metabolic Cycle

Metabolic cycle experiments were performed as previously described (Tu et al., 2005). Fermentors were from New Brunswick Scientific (BioFlo 3000). A 10 mL overnight saturated culture was inoculated to start each YMC run. YMC runs were operated at an agitation speed of 475 rpm in a Bioflo 3000 fermenter, an aeration rate of 1 L/min, a temperature of 30°C, and a pH of 3.4 in 1 L YMC medium. Once the batch culture was saturated, at least 4 hour of starvation was performed. After starvation, fresh medium was added continuously at a dilution rate of ~0.09~0.1 h⁻¹.

ChIP

Chromatin immunoprecipitation was performed as described previously. ~50 OD A₆₀₀ cells per time point were collected for ChIP of each TF. Antibodies used are FLAG M2 (Sigma, F1804),

HA 12CA5 (Roche, 11583816001). 3 µg primary antibody was used per ChIP experiment. Briefly, cells were first fixed in 1% formaldehyde at 25°C for 15 min and quenched in 125mM glycine at 25°C for 10 min. Cells were pelleted and washed twice with buffer containing 100 mM NaCl, 10 mM Tris-Cl pH 8.0, 1 mM EDTA, 1 mM PMSF, 1 mM benzamidine•HCl before freezing. The frozen pellet was resuspended in 0.45 ml ChIP lysis buffer (50 mM HEPES•KOH pH 7.5, 500 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% deoxycholate (DOC), 0.1% SDS, 1 mM PMSF, 10 µM leupeptin, 5 µM pepstatin A, Roche protease inhibitor cocktail) and lysed by bead beating. Lysate from 50 OD cells were split into two tubes each containing 280 µl lysate and sonicated for 16 cycles (30 sec on, 1 min off, high output) using a Bioruptor (Diagenode, Denville, NJ). The supernatant of the sonicated lysate was pre-cleared, and 50 µl lysate was saved as input. After incubation overnight, 50 µl protein G Sepharose beads (GE Healthcare Life Sciences, Piscataway, NJ, 17-0618-01) resuspended in ChIP lysis buffer was added and incubated for 1.5 h at 4 °C. Beads were washed twice with ChIP lysis buffer, twice with DOC buffer (10 mM Tris•Cl pH 8.0, 0.25 M LiCl, 0.5% deoxycholate, 0.5% NP-40, 1 mM EDTA) and twice with TE. 125 µl of TES buffer (TE pH8.0 with 1% SDS, 150 mM NaCl, and 5 mM dithiothreitol) was added to resuspend the beads. Supernatant was collected after incubation at 65°C for 10 min. A second round of elution was performed and the eluates were combined. Reverse crosslinking was performed by incubation for 6 h at 65°C. An equal volume of TE containing 1.25 mg/ml proteinase K and 0.4 mg/ml glycogen was added to the samples after reverse crosslinking and samples were incubated for 2 h at 37°C. Samples were extracted twice with an equal volume of phenol and once with 25:1 chloroform:isoamyl alcohol. DNA was precipitated in 0.1 volume 3.0 M sodium acetate (PH 5.3) and 2.5 volume of 100% ice-cold ethanol at -20°C overnight. Pellets were washed once with cold 70% ethanol and resuspended in 20 µl TE.

qPCR

Fast SYBR Green Master Mix (Applied Biosystems, 4385612, Foster City, CA) was used for real-time PCR and experiments were done on the platform of StepOnePlus Real-Time PCR System (Applied Biosystems, 4385612, Foster City, CA). 3 motif sites from each cluster and 3 from filtered-out motif sites were used to design primers. 0.1 μ l DNA was added per reaction.

Peak callers	BayesPeak	iSeq	PICS	SPP
Number of peaks	4629	14907	807	2761

Table 4.1 Summary of numbers of H3K9ac peaks detected by different peak callers.

Number of motif sites	Total	Total picked	BayesPeak	iSeq	PICS	SPP
200	1,070,043	517,456	239,332	288,714	211,303	98,948
ori	1,070,043	377,166	165,948	95,866	200,598	51,274

Table 4.2 Numbers of motif sites picked by these 4 peak callers. Count numbers of motif sites overlapping with detected peaks. 200 means extends motif sites to 200 bp windows and ori means no extension.

Number of motif sites	Picked by only one peak caller	Picked by two peak callers	Picked by three peak callers	Picked by four peak callers
200	308,028	116,006	75,431	17,991
ori	278,512	65,895	27,652	5,107

Table 4.3 Numbers of motif sites picked by combinatorial peak callers.

	BayesPeak	iSeq	PICS	SPP
200_uniquePicked	26.57	39.43	60.63	2.49
ori_uniquePicked	47.84	42.32	78.01	3.98

Table 4.4 Percentages of motif sites uniquely picked by each peak caller.

	Peak-overlapping	Naïve Bayes classifier
BayesPeak	165,948	32,152
iSeq	95,866	15,246
PICS	200,598	6,672
SPP	51,274	19,645
4 callers	5,107	2,408
3 callers	27,652	13,801
2 callers	65,895	25,981
1 callers	377,166	34,661

Table 4.5 Numbers of motif sites after Naïve Bayes classifier filtering. Motif sites are not extended.

TFs with motif sites enriched in clusters	Swi5, Msn2, Xbp1, Gcr1
TFs with no motif sites enriched	Arg81, Cin5, Oaf1, Cha4

Table 4.6 TFs randomly picked for ChIP-qPCR validation.

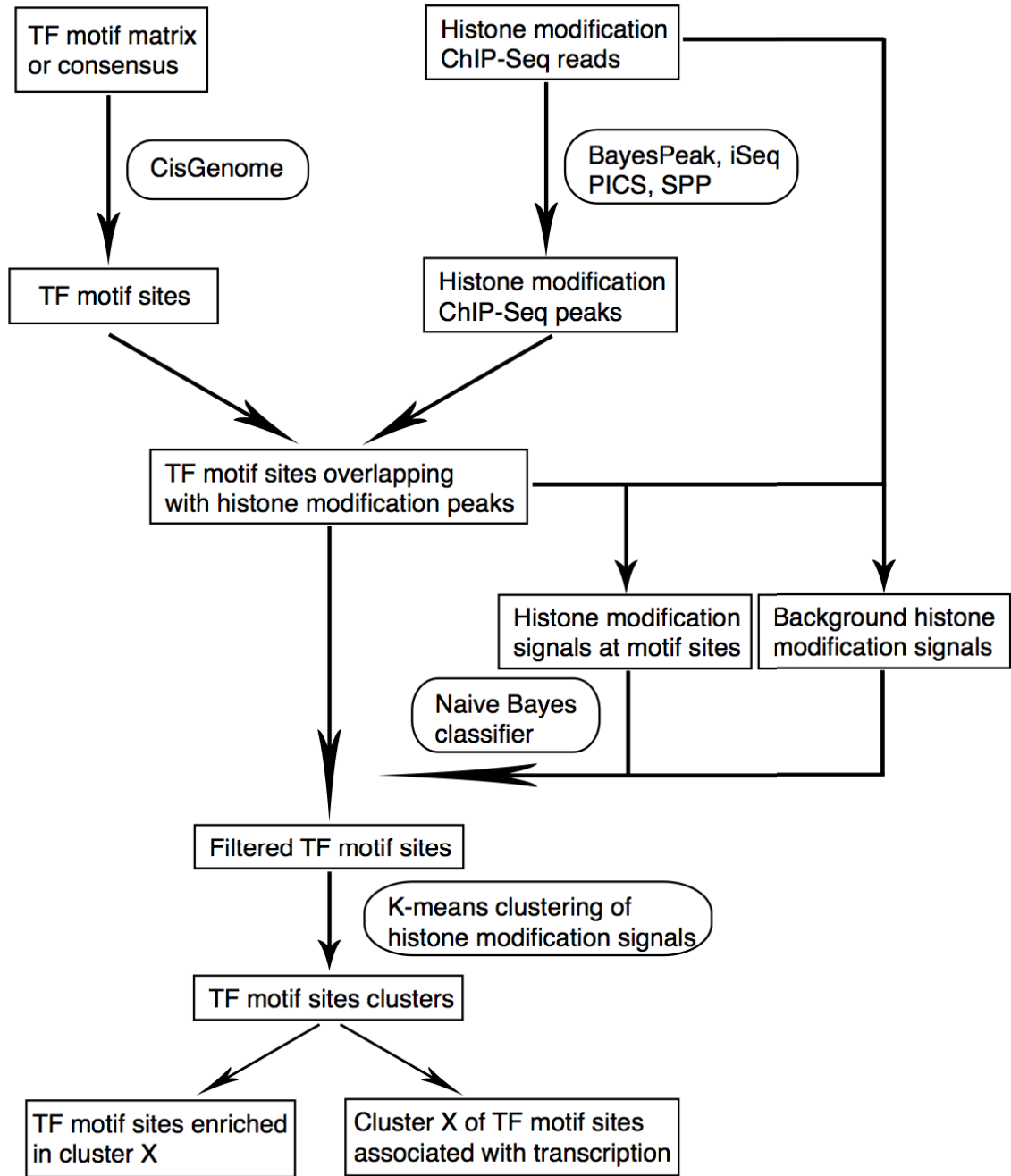


Figure 4.1 Pipeline for prediction of spatiotemporal TF binding in dynamic processes. See main text, methods and materials for details. Briefly, motif sites are obtained by mapping PSFMs or consensus sequences to the genome. Histone modification peaks are detected by peak callers. Motif sites are filtered by overlapping histone modification peaks first and the Naïve Bayes classifier second. Filtered motif sites are combined for K-means clustering and TFs with motif sites enriched in clustering are output. Correlation between histone modification at motif sites and TF transcription is calculated also if expression data are available.

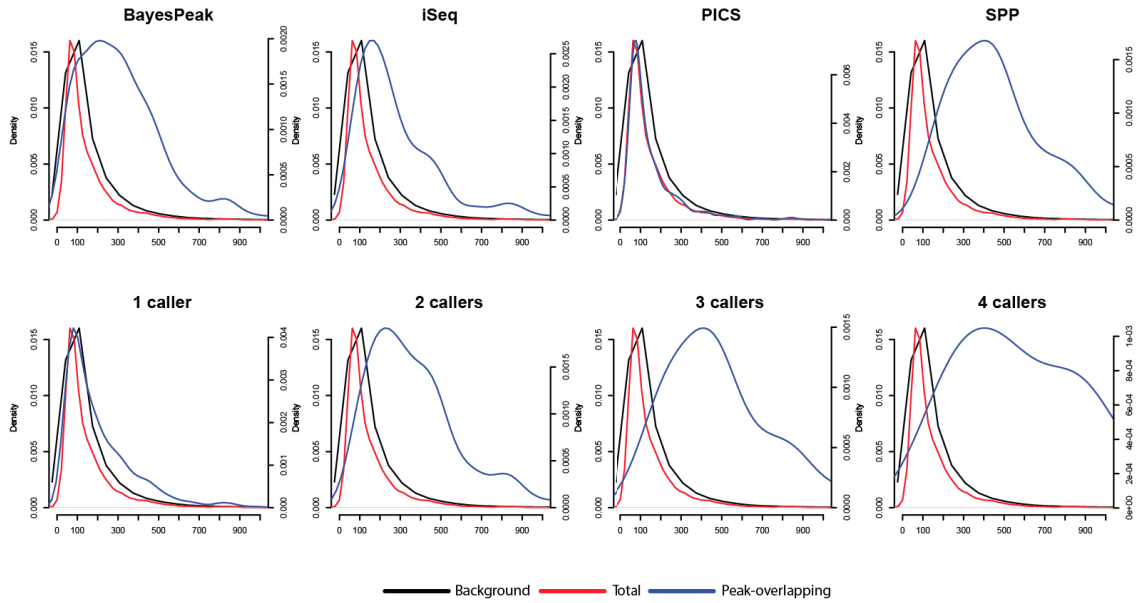


Figure 4.2 Distribution of H3K9ac read counts. ABF1 motif sites are plotted here for illustration. Black lines represent distribution of genomic background. Red lines represent distribution of total motif sites. Blue lines represent distribution of peak-overlapping motif sites.

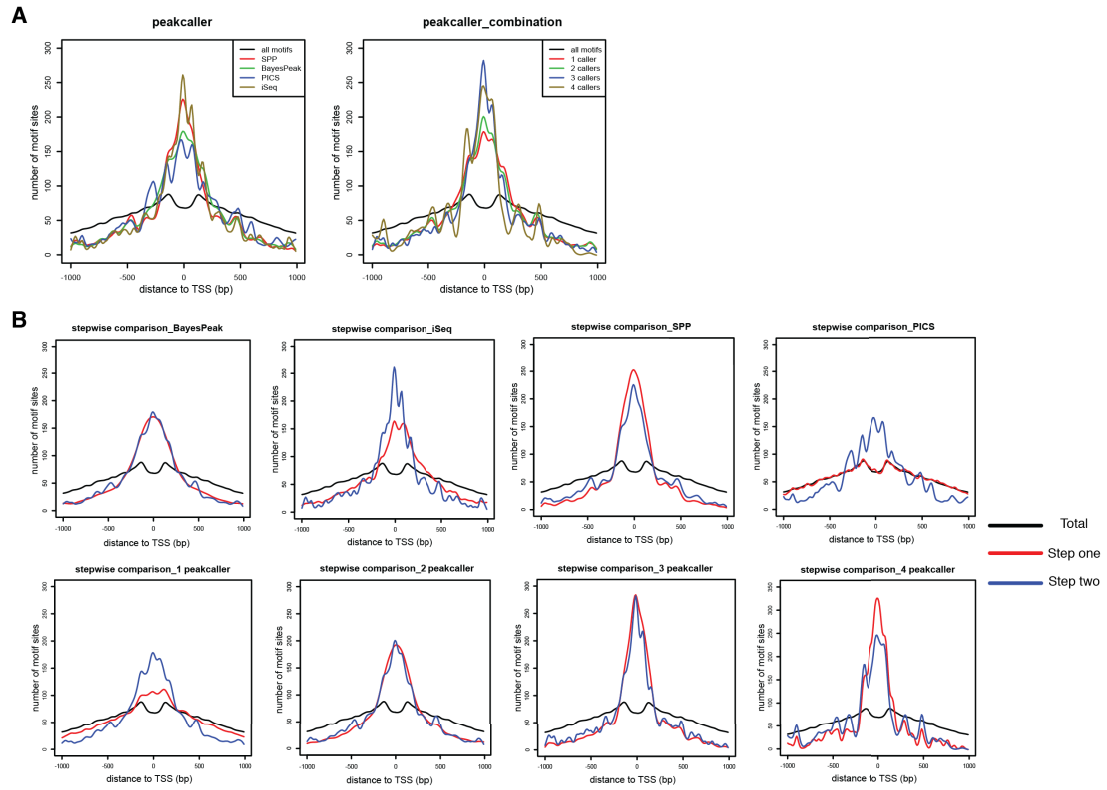


Figure 4.3 Distribution of motif sites at TSSs. (A) Left panel shows the motif sites filtered by individual peak caller and right panel shows the motif sites filtered by multiple peak callers. (B) Distributions of motif sites at TSSs after first and second steps of filtering.

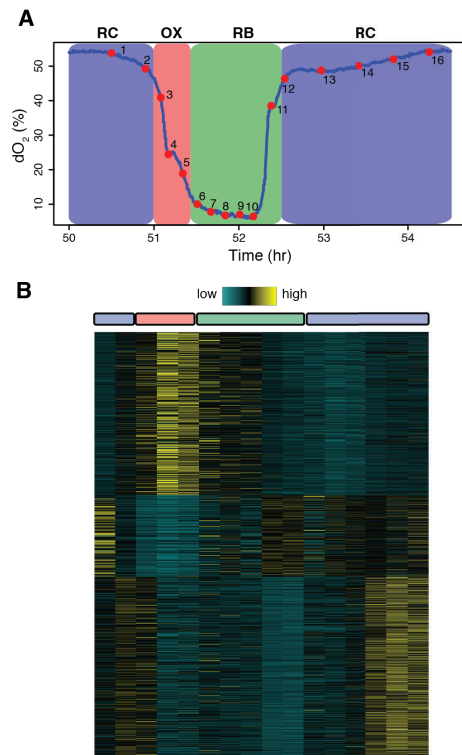


Figure 4.4 Temporal pattern of H3K9ac in YMC. (A) 16 time points sampling in YMC. OX phase is marked by red, RB phase is marked by green and RC phase is marked by blue. (B) K-means clustering of H3K9ac read counts at filtered motif sites of all TFs reveals three different temporal patterns. Each row of the heat map represents one motif site and columns represent the 16 consecutive time points from left to right. 3-color bar on top marks the three phases. The cyan-yellow gradient represents H3K9ac signals.

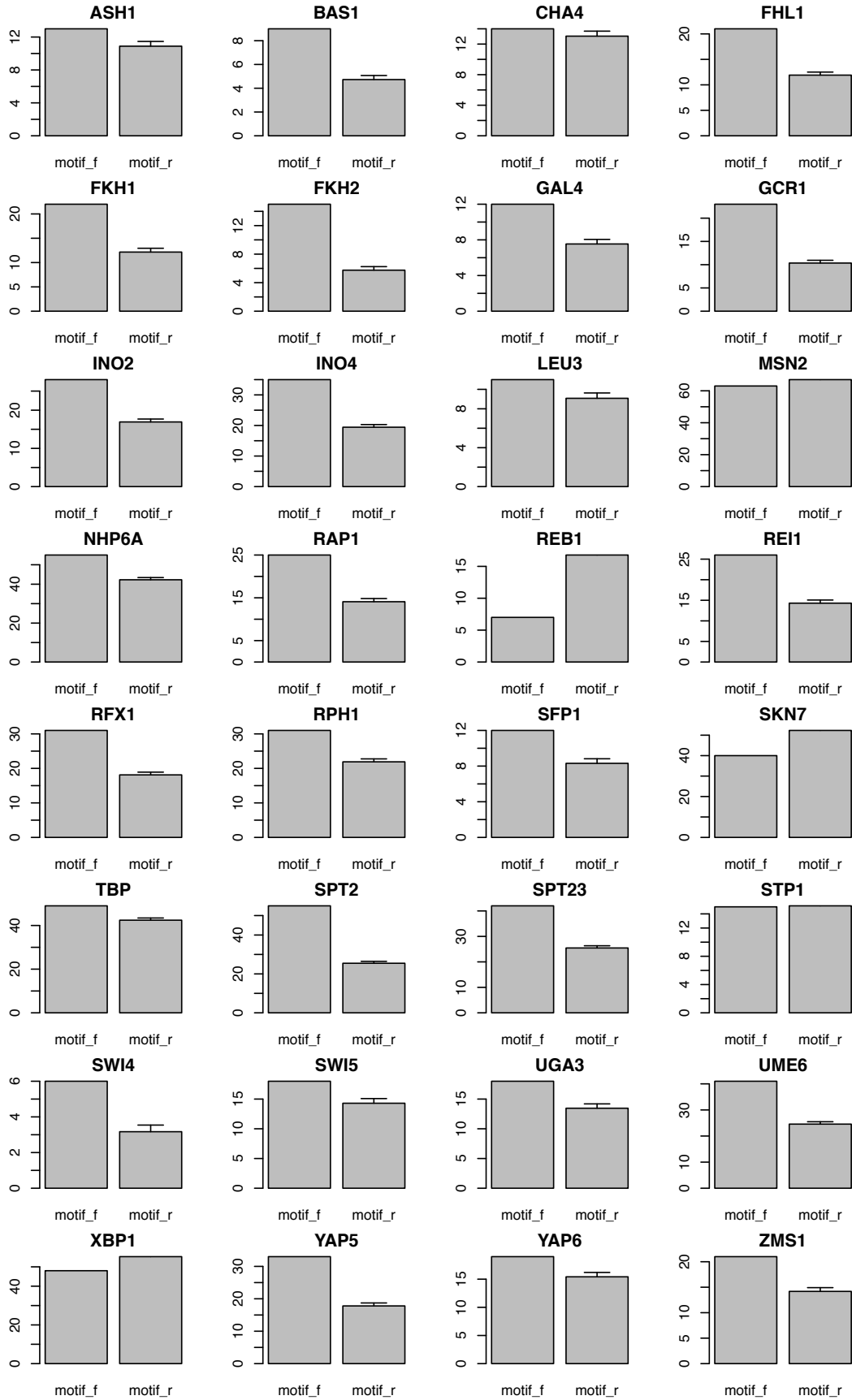


Figure 4.5 Numbers of targeted genes in filtered motif sites and control motif sites. TFs with more than 5 genes predicted and confirmed by the study of Venter et al. are listed. Control groups are generated by randomly sampling from total motif sites 100 times.

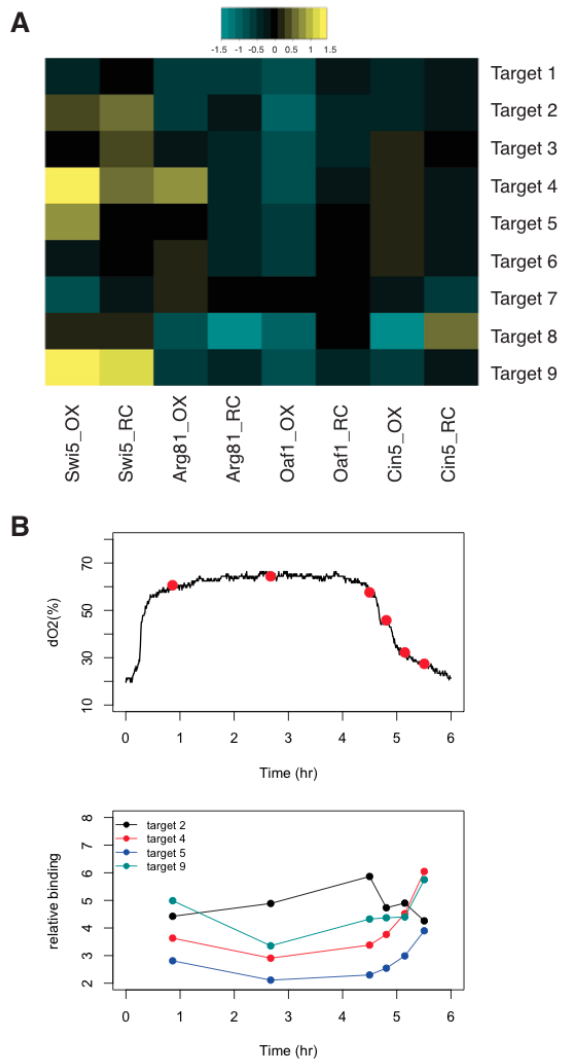


Figure 4.6 Validation of spatial and temporal recruitment of TFs. (A) A heat map shows the ratio of TF binding at nine selected motif sites compared to control motif sites. Binding in OX and RC phase are tested for each TF. (B) Temporal patterns of Swi5 at four motif sites. Top panel shows the 6 time points in YMC and bottom panel shows the time-course ChIP-qPCR signals.

References

- Adkins, M.W., Howar, S.R., and Tyler, J.K. (2004). Chromatin Disassembly Mediated by the Histone Chaperone Asf1 Is Essential for Transcriptional Activation of the Yeast *PHO5* and *PHO8* Genes. *Mol. Cell* *14*, 657-666.
- Andersen, S.L., and Sekelsky, J. (2010). Meiotic versus mitotic recombination: Two different routes for double-strand break repair. *Bioessays* *32*, 1058-1066.
- Aragon, A.D., Rodriguez, A.L., Meirelles, O., Roy, S., Davidson, G.S., Tapia, P.H., Allen, C., Joe, R., Benn, D., and Werner-Washburne, M. (2008). Characterization of differentiated quiescent and nonquiescent cells in yeast stationary-phase cultures. *Mol. Biol. Cell* *19*, 1271-1280.
- Avvakumov, N., Nourani, A., and Côté, J. (2011). Histone chaperones: modulators of chromatin marks. *Mol. Cell* *41*, 502-514.
- Barbaric, S., Walker, J., Schmid, A., Svejstrup, J.Q., and Hoerz, W. (2001). Increasing the rate of chromatin remodeling and gene activation—a novel role for the histone acetyltransferase Gcn5. *EMBO J.* *20*, 4944-4951.
- Barbaric, S., Reinke, H., and Horz, W. (2003). Multiple mechanistically distinct functions of SAGA at the *PHO5* promoter. *Mol. Cell. Biol.* *23*, 3468-3476.
- Barrales, R.R., Korber, P., Jimenez, J., and Ibeas, J.I. (2012). Chromatin modulation at the *FLO11* promoter of *Saccharomyces cerevisiae* by HDAC and Swi/Snf complexes. *Genetics* *191*, 791-803.
- Berger, S.L. (2007). The complex language of chromatin regulation during transcription. *Nature* *447*, 407-412.
- Bergerat, A., de Massy, B., Gabelle, D., Varoutas, P., Nicolas, A., and Forterre, P. (1997). An atypical topoisomerase II from Archaea with implications for meiotic recombination.
- Bergkessel, M., Whitworth, G.B., and Guthrie, C. (2011). Diverse environmental stresses elicit distinct responses at the level of pre-mRNA processing in yeast. *RNA* *17*, 1461-1478.
- Bernstein, B.E., Kamal, M., Lindblad-Toh, K., Bekiranov, S., Bailey, D.K., Huebert, D.J., McMahon, S., Karlsson, E.K., Kulbokas III, E.J., and Gingeras, T.R. (2005). Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* *120*, 169-181.
- Borde, V., and de Massy, B. (2013). Programmed induction of DNA double strand breaks during meiosis: setting up communication between DNA and the chromosome structure. *Curr. Opin. Genet. Dev.* *23*, 147-155.
- Borde, V., Goldman, A.S., and Lichten, M. (2000). Direct coupling between meiotic DNA replication and recombination initiation. *Science* *290*, 806-809.
- Briggs, S.D., Xiao, T., Sun, Z., Caldwell, J.A., Shabanowitz, J., Hunt, D.F., Allis, C.D., and Strahl, B.D. (2002). Gene silencing: trans-histone regulatory pathway in chromatin. *Nature* *418*, 498-498.

- Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* *84*, 843-851.
- Bruschi, C.V., and Esposito, M.S. (1983). Enhancement of spontaneous mitotic recombination by the meiotic mutant spo11-1 in *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. U. S. A.* *80*, 7566-7570.
- Buhler, C., Borde, V., and Lichten, M. (2007). Mapping meiotic single-strand DNA reveals a new landscape of DNA double-strand breaks in *Saccharomyces cerevisiae*. *PLoS Biology* *5*, e324.
- Buratowski, S. (2003). The CTD code. *Nature Structural & Molecular Biology* *10*, 679-680.
- Cai, L., Sutter, B.M., Li, B., and Tu, B.P. (2011). Acetyl-CoA induces cell growth and proliferation by promoting the acetylation of histones at growth genes. *Mol. Cell* *42*, 426-437.
- Cairns, B.R., Schlichter, A., Erdjument-Bromage, H., Tempst, P., Kornberg, R.D., and Winston, F. (1999). Two functionally distinct forms of the RSC nucleosome-remodeling complex, containing essential AT hook, BAH, and bromodomains. *Mol. Cell* *4*, 715-723.
- Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W., Anderson, S., Yates, J., and Washburn, M.P. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* *123*, 581-592.
- Charles, J.S., and Petes, T.D. (2013). High-resolution mapping of spontaneous mitotic recombination hotspots on the 1.1 Mb arm of yeast chromosome IV. *PLoS Genetics* *9*, e1003434.
- Chen, C., Carson, J.J., Feser, J., Tamburini, B., Zabaronick, S., Linger, J., and Tyler, J.K. (2008). Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. *Cell* *134*, 231-243.
- Cheng, C., Shou, C., Yip, K.Y., and Gerstein, M.B. (2011). Genome-wide analysis of chromatin features identifies histone modification sensitive and insensitive yeast transcription factors. *Genome Biol.* *12*, R111.
- Chern, M.K., Chang, K.N., Liu, L.F., Tam, T.C., Liu, Y.C., Liang, Y.L., and Tam, M.F. (2002). Yeast ribosomal protein L12 is a substrate of protein-arginine methyltransferase 2. *J. Biol. Chem.* *277*, 15345-15353.
- Chua, G., Robinson, M.D., Morris, Q., and Hughes, T.R. (2004). Transcriptional networks: reverse-engineering gene regulation on a global scale. *Curr. Opin. Microbiol.* *7*, 638-646.
- Chua, G., Morris, Q.D., Sopko, R., Robinson, M.D., Ryan, O., Chan, E.T., Frey, B.J., Andrews, B.J., Boone, C., and Hughes, T.R. (2006). Identifying transcription factor functions and targets by phenotypic activation. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 12045-12050.
- Clark-Adams, C.D., Norris, D., Osley, M.A., Fassler, J.S., and Winston, F. (1988). Changes in histone gene dosage alter transcription in yeast. *Genes Dev.* *2*, 150-159.
- Cosgrove, M.S., Boeke, J.D., and Wolberger, C. (2004). Regulated nucleosome mobility and the histone code. *Nature Structural & Molecular Biology* *11*,
- Costelloe, T., and Lowndes, N.F. (2010). Chromatin assembly and signalling the end of DNA repair requires acetylation of histone H3 on lysine 56. In *Genome Stability and Human Diseases*, (Springer) pp. 43-54.

- Cui, X., Li, H., and Liu, G. (2011). Combinatorial patterns of histone modifications in *Saccharomyces cerevisiae*. *Yeast* 28, 683-691.
- Dai, J., Hyland, E.M., Yuan, D.S., Huang, H., Bader, J.S., and Boeke, J.D. (2008). Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants. *Cell* 134, 1066-1078.
- Dang, W., Steffen, K.K., Perry, R., Dorsey, J.A., Johnson, F.B., Shilatifard, A., Kaeberlein, M., Kennedy, B.K., and Berger, S.L. (2009). Histone H4 lysine 16 acetylation regulates cellular lifespan. *Nature* 459, 802-807.
- Deckert, J., and Struhl, K. (2001). Histone acetylation at promoters is differentially affected by specific activators and repressors. *Mol. Cell. Biol.* 21, 2726-2735.
- Dion, M.F., Altschuler, S.J., Wu, L.F., and Rando, O.J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. *Proc. Natl. Acad. Sci. U. S. A.* 102, 5501-5506.
- Dover, J., Schneider, J., Tawiah-Boateng, M.A., Wood, A., Dean, K., Johnston, M., and Shilatifard, A. (2002). Methylation of histone H3 by COMPASS requires ubiquitination of histone H2B by Rad6. *J. Biol. Chem.* 277, 28368-28371.
- Durrin, L.K., Mann, R.K., Kayne, P.S., and Grunstein, M. (1991). Yeast histone H4 N-terminal sequence is required for promoter activation in vivo. *Cell* 65, 1023-1031.
- Durrin, L.K., Mann, R.K., and Grunstein, M. (1992). Nucleosome loss activates CUP1 and HIS3 promoters to fully induced levels in the yeast *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 12, 1621-1629.
- Eriksson, P.R., Ganguli, D., Nagarajavel, V., and Clark, D.J. (2012). Regulation of histone gene expression in budding yeast. *Genetics* 191, 7-20.
- Featherstone, M. (2002). Coactivators in transcription initiation: here are your orders. *Curr. Opin. Genet. Dev.* 12, 149-155.
- Ferreiro, J.A., Powell, N.G., Karabetsou, N., Mellor, J., and Waters, R. (2006). Roles for Gcn5p and Ada2p in transcription and nucleotide excision repair at the *Saccharomyces cerevisiae* MET16 gene. *Nucleic Acids Res.* 34, 976-985.
- Fischle, W., Wang, Y., and Allis, C.D. (2003). Histone and chromatin cross-talk. *Curr. Opin. Cell Biol.* 15, 172-183.
- Fogel, S., and Welch, J.W. (1982). Tandem gene amplification mediates copper resistance in yeast. *Proc. Natl. Acad. Sci. U. S. A.* 79, 5342-5346.
- Fuchs, S.M., Krajewski, K., Baker, R.W., Miller, V.L., and Strahl, B.D. (2011). Influence of combinatorial histone modifications on antibody and effector protein recognition. *Current Biology* 21, 53-58.
- Ghosh, S., and Pugh, B.F. (2011). Sequential recruitment of SAGA and TFIID in a genomic response to DNA damage in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* 31, 190-202.
- Ginsburg, D.S., Govind, C.K., and Hinnebusch, A.G. (2009). NuA4 lysine acetyltransferase Esa1 is targeted to coding regions and stimulates transcription elongation with Gcn5. *Mol. Cell. Biol.* 29, 6473-6487.

- Govind, C.K., Zhang, F., Qiu, H., Hofmeyer, K., and Hinnebusch, A.G. (2007). Gen5 promotes acetylation, eviction, and methylation of nucleosomes in transcribed coding regions. *Mol. Cell* 25, 31-42.
- Green, E.M., Mas, G., Young, N.L., Garcia, B.A., and Gozani, O. (2012). Methylation of H4 lysines 5, 8 and 12 by yeast Set5 calibrates chromatin stress responses. *Nature Structural & Molecular Biology* 19, 361-363.
- Guillemette, B., Drogaris, P., Lin, H.H.S., Armstrong, H., Hiragami-Hamada, K., Imhof, A., Bonneil, É, Thibault, P., Verreault, A., and Festenstein, R.J. (2011). H3 lysine 4 is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. *PLoS Genetics* 7, e1001354.
- Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. *Cell* 55, 1137-1145.
- Hansen, L., Kim, N., Mariño-Ramírez, L., and Landsman, D. (2011). Analysis of biological features associated with meiotic recombination hot and cold spots in *saccharomyces cerevisiae*. *PloS One* 6, e29711.
- Hassan, A.H., Neely, K.E., and Workman, J.L. (2001). Histone acetyltransferase complexes stabilize swi/snf binding to promoter nucleosomes. *Cell* 104, 817-827.
- Haynes, S.R., Dollard, C., Winston, F., Beck, S., Trowsdale, J., and Dawid, I.B. (1992). The bromodomain: a conserved sequence found in human, *Drosophila* and yeast proteins. *Nucleic Acids Res.* 20, 2603.
- Hellauer, K., Rochon, M.H., and Turcotte, B. (1996). A novel DNA binding motif for yeast zinc cluster proteins: the Leu3p and Pdr3p transcriptional activators recognize everted repeats. *Mol. Cell. Biol.* 16, 6096-6102.
- Henry, K.W., Wyce, A., Lo, W.S., Duggan, L.J., Emre, N.C., Kao, C.F., Pillus, L., Shilatfard, A., Osley, M.A., and Berger, S.L. (2003). Transcriptional activation via sequential histone H2B ubiquitylation and deubiquitylation, mediated by SAGA-associated Ubp8. *Genes Dev.* 17, 2648-2663.
- Hereford, L., Fahrner, K., Woolford Jr, J., Rosbash, M., and Kaback, D.B. (1979). Isolation of yeast histone genes H2A and H2B. *Cell* 18, 1261-1271.
- Hirschhorn, J.N., Bortvin, A.L., Ricupero-Hovasse, S.L., and Winston, F. (1995). A new class of histone H2A mutations in *Saccharomyces cerevisiae* causes specific transcriptional defects in vivo. *Mol. Cell. Biol.* 15, 1999-2009.
- Hughes, T.R., and de Boer, C.G. (2013). Mapping yeast transcriptional networks. *Genetics* 195, 9-36.
- Hyland, E.M., Molina, H., Poorey, K., Jie, C., Xie, Z., Dai, J., Qian, J., Bekiranov, S., Auble, D.T., Pandey, A., and Boeke, J.D. (2011). An evolutionarily 'young' lysine residue in histone H3 attenuates transcriptional output in *Saccharomyces cerevisiae*. *Genes Dev.* 25, 1306-1319.
- Jacobson, R.H., Ladurner, A.G., King, D.S., and Tjian, R. (2000). Structure and function of a human TAFII250 double bromodomain module. *Science* 288, 1422-1425.
- Ji, H., and Wong, W.H. (2006). Computational biology: toward deciphering gene regulatory information in mammalian genomes. *Biometrics* 62, 645-663.
- Ji, H., Jiang, H., Ma, W., Johnson, D.S., Myers, R.M., and Wong, W.H. (2008). An integrated software system for analyzing ChIP-chip and ChIP-seq data. *Nat. Biotechnol.* 26, 1293-1300.

- Jin, Y., Rodriguez, A.M., and Wyrick, J.J. (2009). Genetic and genomewide analysis of simultaneous mutations in acetylated and methylated lysine residues in histone H3 in *Saccharomyces cerevisiae*. *Genetics* *181*, 461-472.
- Jorgensen, P., Rupeš, I., Sharom, J.R., Schneper, L., Broach, J.R., and Tyers, M. (2004). A dynamic transcriptional network communicates growth potential to ribosome synthesis and critical cell size. *Genes Dev.* *18*, 2491-2505.
- Kao, C.F., Hillyer, C., Tsukuda, T., Henry, K., Berger, S., and Osley, M.A. (2004). Rad6 plays a role in transcriptional activation through ubiquitylation of histone H2B. *Genes Dev.* *18*, 184-195.
- Kaochar, S., and Tu, B.P. (2012). Gatekeepers of chromatin: Small metabolites elicit big changes in gene expression. *Trends Biochem. Sci.*
- Kaplan, T., Liu, C.L., Erkmann, J.A., Holik, J., Grunstein, M., Kaufman, P.D., Friedman, N., and Rando, O.J. (2008). Cell cycle- and chaperone-mediated regulation of H3K56ac incorporation in yeast. *PLoS Genet.* *4*, e1000270.
- Kasten, M., Szerlong, H., Erdjument-Bromage, H., Tempst, P., Werner, M., and Cairns, B.R. (2004). Tandem bromodomains in the chromatin remodeler RSC recognize acetylated histone H3 Lys14. *EMBO J.* *23*, 1348-1359.
- Kawano, A., Hayashi, Y., Noguchi, S., Handa, H., Horikoshi, M., and Yamaguchi, Y. (2011). Global analysis for functional residues of histone variant Htz1 using the comprehensive point mutant library. *Genes to Cells* *16*, 590-607.
- Keeney, S. (2008). Spo11 and the formation of DNA double-strand breaks in meiosis. In *Recombination and meiosis*, Springer) pp. 81-123.
- Kelly, T.J., Qin, S., Gottschling, D.E., and Parthun, M.R. (2000). Type B histone acetyltransferase Hat1p participates in telomeric silencing. *Mol. Cell. Biol.* *20*, 7051-7058.
- Kharchenko, P.V., Tolstorukov, M.Y., and Park, P.J. (2008). Design and analysis of ChIP-seq experiments for DNA-binding proteins. *Nat. Biotechnol.* *26*, 1351-1359.
- Kim, T., Xu, Z., Clauder-Münster, S., Steinmetz, L.M., and Buratowski, S. (2012). Set3 HDAC mediates effects of overlapping noncoding transcription on gene induction kinetics. *Cell* *150*, 1158-1169.
- Kirmizis, A., Santos-Rosa, H., Penkett, C.J., Singer, M.A., Vermeulen, M., Mann, M., Bähler, J., Green, R.D., and Kouzarides, T. (2007). Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* *449*, 928-932.
- Kleff, S., Andrulis, E.D., Anderson, C.W., and Sternglanz, R. (1995). Identification of a gene encoding a yeast histone H4 acetyltransferase. *J. Biol. Chem.* *270*, 24674-24677.
- Koike, N., Yoo, S.H., Huang, H.C., Kumar, V., Lee, C., Kim, T.K., and Takahashi, J.S. (2012). Transcriptional architecture and chromatin landscape of the core circadian clock in mammals. *Science* *338*, 349-354.
- Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* *128*, 693-705.
- Krebs, J.E. (2007). Moving marks: dynamic histone modifications in yeast. *Mol. Biosyst* *3*, 590-597.

- Krogan, N.J., Kim, M., Tong, A., Golshani, A., Cagney, G., Canadien, V., Richards, D.P., Beattie, B.K., Emili, A., and Boone, C. (2003). Methylation of histone H3 by Set2 in *Saccharomyces cerevisiae* is linked to transcriptional elongation by RNA polymerase II. *Mol. Cell. Biol.* *23*, 4207-4218.
- Kuo, H.C., Moore, J.D., and Krebs, J.E. (2005). Histone H2A and Spt10 cooperate to regulate induction and autoregulation of the CUP1 metallothionein. *J. Biol. Chem.* *280*, 104-111.
- Kuras, L., Cherest, H., Surdin-Kerjan, Y., and Thomas, D. (1996). A heteromeric complex containing the centromere binding factor 1 and two basic leucine zipper factors, Met4 and Met28, mediates the transcription activation of yeast sulfur metabolism. *EMBO J.* *15*, 2519-2529.
- Kurdistani, S.K., Tavazoie, S., and Grunstein, M. (2004). Mapping global histone acetylation patterns to gene expression. *Cell* *117*, 721-733.
- Lai, W.K., and Buck, M.J. (2013). An integrative approach to understanding the combinatorial histone code at functional elements. *Bioinformatics* *29*, 2231-2237.
- Larabee, R.N., Krogan, N.J., Xiao, T., Shibata, Y., Hughes, T.R., Greenblatt, J.F., and Strahl, B.D. (2005). BUR kinase selectively regulates H3 K4 trimethylation and H2B ubiquitylation through recruitment of the PAF elongation complex. *Current Biology* *15*, 1487-1493.
- Laxman, S., Sutter, B.M., and Tu, B.P. (2010). Behavior of a Metabolic Cycling Population at the Single Cell Level as Visualized by Fluorescent Gene Expression Reporters. *PloS One* *5*, e12595.
- Lee, P.S., Greenwell, P.W., Dominska, M., Gawel, M., Hamilton, M., and Petes, T.D. (2009). A fine-structure map of spontaneous mitotic crossovers in the yeast *Saccharomyces cerevisiae*. *PLoS Genetics* *5*, e1000410.
- Lee, T.I., Causton, H.C., Holstege, F.C., Shen, W., Hannett, N., Jennings, E.G., Winston, F., Green, M.R., and Young, R.A. (2000). Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* *405*, 701-704.
- Lemon, B., and Tjian, R. (2000). Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* *14*, 2551-2569.
- Lenfant, F., Mann, R.K., Thomsen, B., Ling, X., and Grunstein, M. (1996). All four core histone N-termini contain sequences required for the repression of basal transcription in yeast. *EMBO J.* *15*, 3974-3985.
- Lenstra, T.L., Benschop, J.J., Kim, T., Schulze, J.M., Brabers, N.A., Margaritis, T., van de Pasch, Loes AL, van Heesch, S.A., Brok, M.O., and Groot Koerkamp, M.J. (2011). The specificity and topology of chromatin interaction pathways in yeast. *Mol. Cell* *42*, 536-549.
- Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* *128*, 707-719.
- Li, C., and Wong, W.H. (2003). DNA-chip analyzer (dChip). *The Analysis of Gene Expression Data: Methods and Software* 120-141.
- Li, M., Valsakumar, V., Poorey, K., Bekiranov, S., and Smith, J.S. (2013). Genome-wide analysis of functional sirtuin chromatin targets in yeast. *Genome Biol.* *14*, R48.

- Li, B., Howe, L., Anderson, S., Yates, J.R., 3rd, and Workman, J.L. (2003). The Set2 histone methyltransferase functions through the phosphorylated carboxyl-terminal domain of RNA polymerase II. *J. Biol. Chem.* *278*, 8897-8903.
- Liang, G., Klose, R.J., Gardner, K.E., and Zhang, Y. (2007). Yeast Jhd2p is a histone H3 Lys4 trimethyl demethylase. *Nature Structural & Molecular Biology* *14*,
- Lifton, R.P., Goldberg, M.L., Karp, R.W., and Hogness, D.S. (1978). The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. *Cold Spring Harb. Symp. Quant. Biol.* *42 Pt 2*, 1047-1051.
- Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N., and Rando, O.J. (2005). Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biology* *3*, e328.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* *389*, 251-260.
- Mann, R.K., and Grunstein, M. (1992). Histone H3 N-terminal mutations allow hyperactivation of the yeast GAL1 gene in vivo. *EMBO J.* *11*, 3297-3306.
- Martin, A.M., Pouchnik, D.J., Walker, J.L., and Wyrick, J.J. (2004). Redundant roles for histone H3 N-terminal lysine residues in subtelomeric gene repression in *Saccharomyces cerevisiae*. *Genetics* *167*, 1123-1132.
- Marzluff, W.F., Gongidi, P., Woods, K.R., Jin, J., and Maltais, L.J. (2002). The human and mouse replication-dependent histone genes. *Genomics* *80*, 487-498.
- Matsubara, K., Sano, N., Umehara, T., and Horikoshi, M. (2007). Global analysis of functional surfaces of core histones with comprehensive point mutants. *Genes to Cells* *12*, 13-33.
- Matys, V., Fricke, E., Geffers, R., Gossling, E., Haubrock, M., Hehl, R., Hornischer, K., Karas, D., Kel, A.E., Kel-Margoulis, O.V., *et al.* (2003). TRANSFAC: transcriptional regulation, from patterns to profiles. *Nucleic Acids Res.* *31*, 374-378.
- McMurray, M.A., and Gottschling, D.E. (2003). An age-induced switch to a hyper-recombinational state. *Science* *301*, 1908-1911.
- Megee, P.C., Morgan, B.A., and Smith, M.M. (1995). Histone H4 and the maintenance of genome integrity. *Genes Dev.* *9*, 1716-1727.
- Meluh, P.B., Yang, P., Glowczewski, L., Koshland, D., and Smith, M.M. (1998). Cse4p Is a Component of the Core Centromere of *Saccharomyces cerevisiae*. *Cell* *94*, 607-613.
- Millar, C.B., and Grunstein, M. (2006). Genome-wide patterns of histone modifications in yeast. *Nature Reviews Molecular Cell Biology* *7*, 657-666.
- Mo, Q. (2012). A fully Bayesian hidden Ising model for ChIP-seq data analysis. *Biostatistics* *13*, 113-128.
- Moazed, D. (2001). Enzymatic activities of Sir2 and chromatin silencing. *Curr. Opin. Cell Biol.* *13*, 232-238.

- Morillon, A., Karabetsov, N., Nair, A., and Mellor, J. (2005). Dynamic lysine methylation on histone H3 defines the regulatory phase of gene transcription. *Mol. Cell* *18*, 723-734.
- Nakanishi, S., Sanderson, B.W., Delventhal, K.M., Bradford, W.D., Staehling-Hampton, K., and Shilatifard, A. (2008). A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. *Nature Structural & Molecular Biology* *15*, 881-888.
- Nelson, C.J., Santos-Rosa, H., and Kouzarides, T. (2006). Proline isomerization of histone H3 regulates lysine methylation and gene expression. *Cell* *126*, 905-916.
- Newburger, D.E., and Bulyk, M.L. (2009). UniPROBE: an online database of protein binding microarray data on protein-DNA interactions. *Nucleic Acids Res.* *37*, D77-82.
- Ng, H.H., Robert, F., Young, R.A., and Struhl, K. (2003). Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol. Cell* *11*, 709-719.
- Ng, H.H., Dole, S., and Struhl, K. (2003). The Rtf1 component of the Paf1 transcriptional elongation complex is required for ubiquitination of histone H2B. *J. Biol. Chem.* *278*, 33625-33628.
- Ng, H.H., Feng, Q., Wang, H., Erdjument-Bromage, H., Tempst, P., Zhang, Y., and Struhl, K. (2002). Lysine methylation within the globular domain of histone H3 by Dot1 is important for telomeric silencing and Sir protein association. *Genes Dev.* *16*, 1518-1527.
- Ni, L., Bruce, C., Hart, C., Leigh-Bell, J., Gelperin, D., Umansky, L., Gerstein, M.B., and Snyder, M. (2009). Dynamic and complex transcription factor binding during an inducible response in yeast. *Genes Dev.* *23*, 1351-1363.
- Norris, D., Dunn, B., and Osley, M.A. (1988). The effect of histone gene deletions on chromatin structure in *Saccharomyces cerevisiae*. *Science* *242*, 759-761.
- Norris, D., and Osley, M.A. (1987). The two gene pairs encoding H2A and H2B play different roles in the *Saccharomyces cerevisiae* life cycle. *Mol. Cell. Biol.* *7*, 3473-3481.
- Nourani, A., Utley, R.T., Allard, S., and Côté, J. (2004). Recruitment of the NuA4 complex poises the PHO5 promoter for chromatin remodeling and activation. *EMBO J.* *23*, 2597-2607.
- Odom, D.T. (2011). Identification of Transcription Factor–DNA Interactions In Vivo. In *A Handbook of Transcription Factors*, Springer) pp. 175-191.
- Olins, A.L., and Olins, D.E. (1974). Spheroid chromatin units (v bodies). *Science* *183*, 330-332.
- O'Neill, E.M., Kaffman, A., Jolly, E.R., and O'Shea, E.K. (1996). Regulation of PHO4 nuclear localization by the PHO80-PHO85 cyclin-CDK complex. *Science* *271*, 209-212.
- Paige, S.L., Thomas, S., Stoick-Cooper, C.L., Wang, H., Maves, L., Sandstrom, R., Pabon, L., Reinecke, H., Pratt, G., Keller, G., *et al.* (2012). A temporal chromatin signature in human embryonic stem cells identifies regulators of cardiac development. *Cell* *151*, 221-232.
- Pan, J., Sasaki, M., Kniewel, R., Murakami, H., Blitzblau, H.G., Tischfield, S.E., Zhu, X., Neale, M.J., Jasin, M., and Socci, N.D. (2011). A hierarchical combination of factors shapes the genome-wide topography of yeast meiotic recombination initiation. *Cell* *144*, 719-731.

- Park, P.J. (2009). ChIP-seq: advantages and challenges of a maturing technology. *Nature Reviews Genetics* *10*, 669-680.
- Patterson, H.G., Landel, C.C., Landsman, D., Peterson, C.L., and Simpson, R.T. (1998). The biochemical and phenotypic characterization of Hho1p, the putative linker histone H1 of *Saccharomyces cerevisiae*. *J. Biol. Chem.* *273*, 7268-7276.
- Pique-Regi, R., Degner, J.F., Pai, A.A., Gaffney, D.J., Gilad, Y., and Pritchard, J.K. (2011). Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data. *Genome Res.* *21*, 447-455.
- Pleiss, J.A., Whitworth, G.B., Bergkessel, M., and Guthrie, C. (2007). Rapid, transcript-specific changes in splicing in response to environmental stress. *Mol. Cell* *27*, 928-937.
- Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., and Herbolsheimer, E. (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* *122*, 517-527.
- Prieler, S., Penkner, A., Borde, V., and Klein, F. (2005). The control of Spo11's interaction with meiotic recombination hotspots. *Genes Dev.* *19*, 255-269.
- Qin, Z.S., Yu, J., Shen, J., Maher, C.A., Hu, M., Kalyana-Sundaram, S., Yu, J., and Chinnaiyan, A.M. (2010). HPeak: an HMM-based algorithm for defining read-enriched regions in ChIP-Seq data. *BMC Bioinformatics* *11*, 369.
- Qin, S., and Parthun, M.R. (2002). Histone H3 and the histone acetyltransferase Hat1p contribute to DNA double-strand break repair. *Mol. Cell Biol.* *22*, 8353-8365.
- Rando, O.J., and Winston, F. (2012). Chromatin and transcription in yeast. *Genetics* *190*, 351-387.
- Rando, O.J., and Winston, F. (2012). Chromatin and transcription in yeast. *Genetics* *190*, 351-387.
- Rando, O.J. (2007). Global patterns of histone modifications. *Curr. Opin. Genet. Dev.* *17*, 94-99.
- Recht, J., Tsubota, T., Tanny, J.C., Diaz, R.L., Berger, J.M., Zhang, X., Garcia, B.A., Shabanowitz, J., Burlingame, A.L., Hunt, D.F., Kaufman, P.D., and Allis, C.D. (2006). Histone chaperone Asf1 is required for histone H3 lysine 56 acetylation, a modification associated with S phase in mitosis and meiosis. *Proc. Natl. Acad. Sci. U. S. A.* *103*, 6988-6993.
- Reinke, H., and Hörz, W. (2003). Histones Are First Hyperacetylated and Then Lose Contact with the Activated *PHO5* Promoter. *Mol. Cell* *11*, 1599-1607.
- Robert, F., Pokholok, D.K., Hannett, N.M., Rinaldi, N.J., Chandy, M., Rolfe, A., Workman, J.L., Gifford, D.K., and Young, R.A. (2004). Global position and recruitment of HATs and HDACs in the yeast genome. *Mol. Cell* *16*, 199-209.
- Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., and Delaney, A. (2007). Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nature Methods* *4*, 651-657.
- Rosaleny, L.E., Ruiz-García, A.B., García-Martínez, J., Pérez-Ortín, J.E., and Tordera, V. (2007). The Sas3p and Gcn5p histone acetyltransferases are recruited to similar genes. *Genome Biol.* *8*, R119.

- Rozowsky, J., Euskirchen, G., Auerbach, R.K., Zhang, Z.D., Gibson, T., Bjornson, R., Carriero, N., Snyder, M., and Gerstein, M.B. (2009). PeakSeq enables systematic scoring of ChIP-seq experiments relative to controls. *Nat. Biotechnol.* *27*, 66-75.
- Ruiz-Roig, C., Viéitez, C., Posas, F., and De Nadal, E. (2010). The Rpd3L HDAC complex is essential for the heat stress response in yeast. *Mol. Microbiol.* *76*, 1049-1062.
- Rutherford, J.C., and Bird, A.J. (2004). Metal-responsive transcription factors that regulate iron, zinc, and copper homeostasis in eukaryotic cells. *Eukaryot. Cell.* *3*, 1-13.
- Rykowski, M.C., Wallis, J.W., Choe, J., and Grunstein, M. (1981). Histone H2B subtypes are dispensable during the yeast cell cycle. *Cell* *25*, 477-487.
- Sandelin, A., Alkema, W., Engstrom, P., Wasserman, W.W., and Lenhard, B. (2004). JASPAR: an open-access database for eukaryotic transcription factor binding profiles. *Nucleic Acids Res.* *32*, D91-4.
- Santisteban, M.S., Kalashnikova, T., and Smith, M.M. (2000). Histone H2A. Z regulates transcription and is partially redundant with nucleosome remodeling complexes. *Cell* *103*, 411-422.
- Santos-Rosa, H., Schneider, R., Bernstein, B.E., Karabetsou, N., Morillon, A., Weise, C., Schreiber, S.L., Mellor, J., and Kouzarides, T. (2003). Methylation of histone H3 K4 mediates association of the Isw1p ATPase with chromatin. *Mol. Cell* *12*, 1325-1332.
- Seward, D.J., Cubberley, G., Kim, S., Schonewald, M., Zhang, L., Tripet, B., and Bentley, D.L. (2007). Demethylation of trimethylated histone H3 Lys4 in vivo by JARID1 JmjC proteins. *Nature Structural & Molecular Biology* *14*, 240-242.
- Sharma, V.M., Tomar, R.S., Dempsey, A.E., and Reese, J.C. (2007). Histone deacetylases RPD3 and HOS2 regulate the transcriptional activation of DNA damage-inducible genes. *Mol. Cell. Biol.* *27*, 3199-3210.
- Shen, C.H., Leblanc, B.P., Alfieri, J.A., and Clark, D.J. (2001). Remodeling of yeast CUP1 chromatin involves activator-dependent repositioning of nucleosomes over the entire gene and flanking sequences. *Mol. Cell. Biol.* *21*, 534-547.
- Shen, C.H., Leblanc, B.P., Neal, C., Akhavan, R., and Clark, D.J. (2002). Targeted histone acetylation at the yeast CUP1 promoter requires the transcriptional activator, the TATA boxes, and the putative histone acetylase encoded by SPT10. *Mol. Cell. Biol.* *22*, 6406-6416.
- Shivaswamy, S., and Iyer, V.R. (2008). Stress-dependent dynamics of global chromatin remodeling in yeast: dual role for SWI/SNF in the heat shock stress response. *Mol. Cell. Biol.* *28*, 2221-2234.
- Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* *311*, 844-847.
- Smith, M.M., and Andrésson, ÓS. (1983). DNA sequences of yeast H3 and H4 histone genes from two non-allelic gene sets encode identical H3 and H4 proteins. *J. Mol. Biol.* *169*, 663-690.
- Smith, M.M., and Stirling, V.B. (1988). Histone H3 and H4 gene deletions in *Saccharomyces cerevisiae*. *J. Cell Biol.* *106*, 557-566.
- Smolle, M., and Workman, J.L. (2013). Transcription-associated histone modifications and cryptic transcription. *Biochimica Et Biophysica Acta (BBA)-Gene Regulatory Mechanisms* *1829*, 84-97.

- Spyrou, C., Stark, R., Lynch, A.G., and Tavaré, S. (2009). BayesPeak: Bayesian analysis of ChIP-seq data. *BMC Bioinformatics* *10*, 299.
- Strahl, B.D., and Allis, C.D. (2000). The language of covalent histone modifications. *Nature* *403*, 41-45.
- Sun, Z., and Allis, C.D. (2002). Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* *418*, 104-108.
- Svaren, J., and Hörz, W. (1997). Transcription factors vs nucleosomes: regulation of the PH05 promoter in yeast. *Trends Biochem. Sci.* *22*, 93-97.
- Szilard, R.K., Jacques, P., Laramée, L., Cheng, B., Galicia, S., Bataille, A.R., Yeung, M., Mendez, M., Bergeron, M., and Robert, F. (2010). Systematic identification of fragile sites via genome-wide location analysis of [gamma]-H2AX. *Nature Structural & Molecular Biology* *17*, 299-305.
- Takahashi, H., McCaffery, J.M., Irizarry, R.A., and Boeke, J.D. (2006). Nucleocytoplasmic acetyl-coenzyme A synthetase is required for histone acetylation and global transcription. *Mol. Cell* *23*, 207-217.
- Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* *272*, 408-411.
- Taverna, S.D., Ilin, S., Rogers, R.S., Tanny, J.C., Lavender, H., Li, H., Baker, L., Boyle, J., Blair, L.P., and Chait, B.T. (2006). Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol. Cell* *24*, 785-796.
- Taverna, S.D., Li, H., Ruthenburg, A.J., Allis, C.D., and Patel, D.J. (2007). How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature Structural & Molecular Biology* *14*, 1025-1040.
- Tessarz, P., Santos-Rosa, H., Robson, S.C., Sylvestersen, K.B., Nelson, C.J., Nielsen, M.L., and Kouzarides, T. (2013). Glutamine methylation in histone H2A is an RNA-polymerase-I-dedicated modification. *Nature*
- Thukral, S.K., Eisen, A., and Young, E.T. (1991). Two monomers of yeast transcription factor ADR1 bind a palindromic sequence symmetrically to activate ADH2 expression. *Mol. Cell. Biol.* *11*, 1566-1577.
- Tischfield, S.E., and Keeney, S. (2012). Scale matters. *Cell Cycle* *11*, 1496-1503.
- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D.R., Pimentel, H., Salzberg, S.L., Rinn, J.L., and Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nature Protocols* *7*, 562-578.
- Travis, G.H., Colavito-Shepanski, M., and Grunstein, M. (1984). Extensive purification and characterization of chromatin-bound histone acetyltransferase from *Saccharomyces cerevisiae*. *J. Biol. Chem.* *259*, 14406-14412.
- Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P., and Zhang, Y. (2006). Histone demethylation by a family of JmjC domain-containing proteins. *Nature* *439*, 811-816.
- Tu, B.P., Kudlicki, A., Rowicka, M., and McKnight, S.L. (2005). Logic of the yeast metabolic cycle: temporal compartmentalization of cellular processes. *Science* *310*, 1152-1158.

- Tu, B.P., Mohler, R.E., Liu, J.C., Dombek, K.M., Young, E.T., Synovec, R.E., and McKnight, S.L. (2007). Cyclic changes in metabolic state during the life of a yeast cell. *Proceedings of the National Academy of Sciences* *104*, 16886-16891.
- Turner, B.M. (2000). Histone acetylation and an epigenetic code. *Bioessays* *22*, 836-845.
- Valouev, A., Johnson, D.S., Sundquist, A., Medina, C., Anton, E., Batzoglou, S., Myers, R.M., and Sidow, A. (2008). Genome-wide analysis of transcription factor binding sites based on CHIP-Seq data. *Nature Methods* *5*, 829-834.
- Venters, B.J., and Pugh, B.F. (2009). How eukaryotic genes are transcribed. *Crit. Rev. Biochem. Mol. Biol.* *44*, 117-141.
- Vidal, M., and Gaber, R.F. (1991). RPD3 encodes a second factor required to achieve maximum positive and negative transcriptional states in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* *11*, 6317-6327.
- Vidanes, G.M., Bonilla, C.Y., and Toczyski, D.P. (2005). Complicated tails: histone modifications and the DNA damage response. *Cell* *121*, 973-976.
- Wamstad, J.A., Alexander, J.M., Truty, R.M., Shrikumar, A., Li, F., Eilertson, K.E., Ding, H., Wylie, J.N., Pico, A.R., Capra, J.A., *et al.* (2012). Dynamic and coordinated epigenetic regulation of developmental transitions in the cardiac lineage. *Cell* *151*, 206-220.
- Wan, J.S., Mann, R.K., and Grunstein, M. (1995). Yeast histone H3 and H4 N termini function through different GAL1 regulatory elements to repress and activate transcription. *Proc. Natl. Acad. Sci. U. S. A.* *92*, 5664-5668.
- Wang, A., Kurdistani, S.K., and Grunstein, M. (2002). Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science* *298*, 1412-1414.
- Wang, S.S., Zhou, B.O., and Zhou, J.Q. (2011). Histone H3 lysine 4 hypermethylation prevents aberrant nucleosome remodeling at the PHO5 promoter. *Mol. Cell. Biol.* *31*, 3171-3181.
- Wei, Y., Wu, G., and Ji, H. (2013). Global Mapping of Transcription Factor Binding Sites by Sequencing Chromatin Surrogates: a Perspective on Experimental Design, Data Analysis, and Open Problems. *Statistics in Biosciences* *5*, 156-178.
- Weiner, A., Chen, H.V., Liu, C.L., Rahat, A., Klien, A., Soares, L., Gudipati, M., Pfeffner, J., Regev, A., and Buratowski, S. (2012). Systematic dissection of roles for chromatin regulators in a yeast stress response. *PLoS Biology* *10*, e1001369.
- Williams, S.K., Truong, D., and Tyler, J.K. (2008). Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 9000-9005.
- Wimalarathna, R.N., Pan, P.Y., and Shen, C. (2013). Co-dependent recruitment of Ino80p and Snf2p is required for yeast CUP1 activation. *Biochemistry and Cell Biology* *92*, 1-7.
- Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2005). The Bur1/Bur2 complex is required for histone H2B monoubiquitination by Rad6/Bre1 and histone methylation by COMPASS. *Mol. Cell* *20*, 589-599.

- Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *J. Biol. Chem.* *278*, 34739-34742.
- Wood, A., Schneider, J., Dover, J., Johnston, M., and Shilatifard, A. (2003). The Paf1 complex is essential for histone monoubiquitination by the Rad6-Bre1 complex, which signals for histone methylation by COMPASS and Dot1p. *J. Biol. Chem.* *278*, 34739-34742.
- Xiao, T., Kao, C.F., Krogan, N.J., Sun, Z.W., Greenblatt, J.F., Osley, M.A., and Strahl, B.D. (2005). Histone H2B ubiquitylation is associated with elongating RNA polymerase II. *Mol. Cell. Biol.* *25*, 637-651.
- Xu, F., Zhang, K., and Grunstein, M. (2005). Acetylation in histone H3 globular domain regulates gene expression in yeast. *Cell* *121*, 375-385.
- Xu, Z., Wei, W., Gagneur, J., Perocchi, F., Clauder-Münster, S., Camblong, J., Guffanti, E., Stutz, F., Huber, W., and Steinmetz, L.M. (2009). Bidirectional promoters generate pervasive transcription in yeast. *Nature* *457*, 1033.
- Xue-Franzén, Y., Henriksson, J., Bürglin, T.R., and Wright, A.P. (2013). Distinct roles of the Gcn5 histone acetyltransferase revealed during transient stress-induced reprogramming of the genome. *BMC Genomics* *14*, 479.
- Yun, M., Wu, J., Workman, J.L., and Li, B. (2011). Readers of histone modifications. *Cell Res.* *21*, 564-578.
- Zaslaver, A., Mayo, A.E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M.G., and Alon, U. (2004). Just-in-time transcription program in metabolic pathways. *Nat. Genet.* *36*, 486-491.
- Zhang, J.A., Mortazavi, A., Williams, B.A., Wold, B.J., and Rothenberg, E.V. (2012). Dynamic transformations of genome-wide epigenetic marking and transcriptional control establish T cell identity. *Cell* *149*, 467-482.
- Zhang, L., Ma, H., and Pugh, B.F. (2011). Stable and dynamic nucleosome states during a meiotic developmental process. *Genome Res.* *21*, 875-884.
- Zhang, X., Robertson, G., Krzywinski, M., Ning, K., Droit, A., Jones, S., and Gottardo, R. (2011). PICS: Probabilistic Inference for ChIP-seq. *Biometrics* *67*, 151-163.
- Zhang, Y., Liu, T., Meyer, C., Eeckhoute, J., Johnson, D., Bernstein, B., Nusbaum, C., Myers, R., Brown, M., and Li, W. (2008). Model-based Analysis of ChIP-Seq (MACS). *Genome Biol.* *9*, R137.

CURRICULUM VITAE

Name: Zheng Kuang

Date and Place of Birth: September 26, 1987
Suichuan, Jiangxi Province, China

Mailing Address (home): 929 North Wolfe Street, Apt. 1803, Baltimore, MD 21205

Mailing Address (work): Department of Molecular Biology and Genetics
Johns Hopkins University School of Medicine
733 N. Broadway, 320 MRB
Baltimore, MD 21205

Telephone: (410) 497-2497 (home)
(410) 955-2474 (work)

Email: kuangzheng04@gmail.com

Education:

August 2008 – June 2014: Johns Hopkins University School of Medicine
Baltimore, MD
Ph.D. (BCMB – Molecular Biology and Genetics)

April 2012 – June 2014: Johns Hopkins University School of Public Health
Baltimore, MD
MHS (Biostatistics)

August 2004 – July 2008: Tsinghua University, Beijing, China
B.S. (Biological Sciences and Biotechnology)

Research Experience:

- July 2009 – June 2014: “Dynamic histone modifications coordinate transcription, DNA repair and recombination across distinct metabolic states” in the laboratory of Dr. Jef D. Boeke, Johns Hopkins University School of Medicine
- April 2012 – June 2014: “Predicting dynamic activity of transcription factors through time-course ChIP-seq data of histone modifications” in the laboratory of Dr. Hongkai Ji, Johns Hopkins University School of Public Health
- August 2006 – June 2008: “Heart-specific isoform of tropomyosin4 is essential for heartbeat in zebrafish embryos” in the laboratory of Dr. Anming Meng, Tsinghua University

Honors and Awards:

- 2005: First Class Scholarship for Academic Excellence
- 2006: Third Class Scholarship for Academic Excellence

Publications:

1. **Kuang Z**, Cai L, Zhang XK, Ji HK, Tu BP, Boeke JD. (2013) High temporal resolution view of transcription and chromatin states across distinct metabolic states in budding yeast. (Submitted)
2. Annaluru N, Muller H, Mitchell LA, Ramalingam S, Stracquadanio G, Richardson SM, Dymond JS, **Kuang Z**, Scheifele LZ, Cooper EM, et al. (2014). Total synthesis of a functional designer eukaryotic chromosome. *Science* 344, 55-58.
3. Zhao L, Zhao X, Tian T, Lu Q, Skrbol-Larsen N, Wu D, **Kuang Z**, Zheng X, Han Y, Yang S, Zhang C, and Meng A. (2008). Heart-specific isoform of tropomyosin4 is essential for heartbeat in zebrafish embryos. *Cardiovasc. Res.* 80, 200-208.

Manuscripts in preparation:

1. Kuang Z, Boeke JD. Genome-wide location of H3K56ac reveals active roles of “meiosis-specific” Spo11 in mitosis.
2. Kuang Z, Boeke JD, Ji HK. Predicting dynamic activity of transcription factors through time-course ChIP-seq data of histone modifications.

Conference Abstracts:

1. **Kuang Z**, Cai, L, Zhang XK, Ji HK, Tu BP, Boeke JD. (2013) A high-resolution view of transcription and chromatin states across distinct metabolic states in budding yeast. TCNP All Hands meeting, Bethesda, MD. Talk Presentation.
2. **Kuang Z**, Zhang XK, Cai, L, Sliva A, Ji HK, Tu BP, Boeke JD. (2012) Oscillation of histone modifications in yeast metabolic cycle reveals complex interaction between metabolism, chromatin and transcription. TCNP All Hands meeting, Bethesda, MD. Poster Presentation.
3. Sliva A, **Kuang Z**, Zhang H, Tu BP, Boeke JD. (2012) FACS-based genetic screens in *S.cerevisiae* to identify genes and histone residues involved in dynamic processes. TCNP All Hands meeting, Bethesda, MD. Poster Presentation.
4. **Kuang Z**, Zhang XK, Cai, L, Ji HK, Tu BP, Boeke JD. (2012) Oscillation of histone modifications in yeast metabolic cycle reveals complex interaction between metabolism, chromatin and transcription. Yeast Genetics and Molecular Biology Meeting, Princeton, NJ. Poster Presentation.
5. **Kuang Z**, Zhong XG, Cai, L, Sliva A, Ji HK, Tu BP, Boeke JD. (2011) Oscillation of histone modifications in yeast metabolic cycle reveals complex interaction between metabolism, chromatin and transcription. TCNP All Hands meeting, Rockville, MD. Poster Presentation.
6. **Kuang Z**, Lin YY, Ji HK, Tu BP, Boeke JD. (2010) Oscillation of histone methylation in yeast metabolic cycle reveals complex interaction between metabolism, chromatin and transcription. Yeast Genetics and Molecular Biology Meeting, Vancouver, BC Canada. Poster Presentation.

References:

Jef D. Boeke, Ph.D.

Professor, Department of Molecular Biology and Genetics

Johns Hopkins University School of Medicine

733 N. Broadway, 339 MRB

Baltimore, MD 21205

Phone: (410) 955-0398

E-mail: jboeke@jhmi.edu

Sean D. Taverna, Ph.D.

Assistant Professor, Department of Pharmacology and Molecular Sciences

Johns Hopkins University School of Medicine

855 N. Wolfe Street, 575 Rangos

Baltimore, MD 21205

Phone: (410) 502-0790

E-mail: staverna@jhmi.edu