CHROMATIN ORGANIZATION DURING YEAST MEIOSIS

A Dissertation in
Cell and Developmental Biology
by
Liye Zhang

© 2012 Liye Zhang

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Doctor of Philosophy

May 2012
The dissertation of Liye Zhang was reviewed and approved* by the following:

B. Franklin Pugh  
Willaman Chair in Molecular Biology and  
Professor of Biochemistry and Molecular Biology  
Dissertation Co-Advisor  
Co-Chair of Committee

Hong Ma  
Distinguished Professor in Biology and  
Professor and Dean of Life Sciences at Fudan University  
Dissertation Co-Advisor  
Co-Chair of Committee

Joseph Reese  
Professor of Biochemistry and Molecular Biology

Yu Zhang  
Assistant Professor of Statistics

Zhi-Chun Lai  
Professor of Biology and Biochemistry and Molecular Biology  
Chair, Intercollege Graduate Degree Program in Cell and Developmental Biology

*Signatures are on file in the Graduate School
ABSTRACT

All eukaryote genomes are packaged with DNA-binding proteins into a structure called chromatin. Chromatin plays crucial roles in regulating gene expression and recombination. There are multiple levels of chromatin folding, and current studies focus on primary chromatin structure—positioning and occupancy of nucleosome across a genome. It has been shown that there is a general theme of nucleosome organization on the genes, and such patterns are conserved among several model organisms from yeast to human. Most of the current studies on primary chromatin structure have utilized unsynchronized cells during vegetative growth, meaning there is a gap in our knowledge regarding the changes of primary chromatin structure and histone modifications during meiotic cell cycle.

To study primary chromatin structure during the meiotic cell cycle, we performed chromatin immunoprecipitation (ChIP) on samples from several stages of yeast meiosis to study nucleosome, modified histones, linker histone (Hho1p), and a partial nucleosome called tetrasome. These results revealed a few general principles regarding meiosis and transcription regulation, which may also apply to plant and metazoan. 1) Nucleosome positioning and occupancy is largely maintained during the meiotic cell cycle; 2) Repositioning of the +1 nucleosome correlates with transcription regulation in subsets of stress response genes; 3) The histone modification data suggest that antisense transcripts repress meiotic specific genes during vegetative growth; 4) linker histone Hho1p is enriched at linker DNA and promoter regions; 5) Globally elevated incorporation of Hho1p into chromatin may contribute to chromatin compaction in mature yeast spores; 6) The majority of tetrasome dyads are shifted approximately 25-50 bp away from nucleosome dyads and may be mainly positioned by DNA sequence. Taken together, the work presented in this thesis shows multiple aspects of primary chromatin architecture during stages of meiosis, and such a dynamic and comprehensive picture of primary
chromatin structure will be able to provide a framework to integrate results from other chromatin associated factors, thus revealing the interplay between cellular events and chromatin structure.
TABLE OF CONTENTS

LIST OF FIGURES ........................................................................................................ viii
LIST OF TABLES ............................................................................................................. x
ACKNOWLEDGEMENTS .............................................................................................. xi

Chapter 1 Introduction to chromatin.............................................................. 1

1.1 Multiple levels of chromatin structure ...................................................... 1
  1.1.1 Primary level of chromatin structure ................................................. 1
  1.1.2 Secondary and tertiary level of chromatin structure ................. 3
1.2 Histone variants and modifications .......................................................... 4
  1.2.1 Functions of histone variants ....................................................... 4
  1.2.2 Functions of histone modifications ............................................. 5
1.3 Nucleosome landscape and transcription .............................................. 7
1.4 Nucleosome dynamics during transcription and DNA replication ........ 11
1.5 Chromatin structure during yeast meiosis ............................................ 12
  1.5.1 Overview of chromatin structure in meiosis ............................. 12
  1.5.2 Chromatin structure in recombination sites ............................ 14
1.6 Transcription regulation of sporulation in yeast .................................. 15
  1.6.1 Repression of sporulation specific genes during vegetative growth 15
  1.6.2 Activation of sporulation specific genes ................................ 16
1.7 Genome-wide experimental approach to study chromatin ................. 17
  1.7.1 Yeast sporulation as model for meiosis ................................ 17
  1.7.2 MNase digestion and ChIP-seq .................................................. 18
  1.7.3 Overview of next-gen sequencing technology ......................... 18

Chapter 2 Stable and dynamic nucleosome states during a meiotic developmental process . 20

2.1 Summary ........................................................................................................... 20
2.2 Introduction ....................................................................................................... 21
2.3 Results ............................................................................................................... 23
  2.3.1 The primary organization of chromatin remains fixed during meiosis and
       upon compaction .............................................................................. 23
  2.3.2 Selective re-positioning of the +1 nucleosome over the TSS is linked to
       gene regulation ................................................................................ 34
  2.3.3 Transcription-linked histone modification marks exist at most genes and
       are selective for nucleosome positions ........................................ 38
  2.3.4 Histone marks around recombinational hotspots and promoters are similar .. 46
  2.3.5 Antisense repression of meiotic gene ........................................ 49
  2.3.6 Enrichment of H3K36me3 levels in transcriptionally inactive rDNA
       region ............................................................................................... 50
  2.3.7 Global increase of linker DNA length at stationary stage (0 hr) ........ 54
2.4 Discussion .......................................................................................................... 56
  2.4.1 Chromatin organization is a constant, but with locus-specific plasticity .... 56
  2.4.2 Regulation of transcription initiation by +1 nucleosome repositioning vs
       antisense ...................................................................................... 58
Chapter 3 Genome-wide histone H1 mapping upon sporulation program

3.1 Summary ................................................................. 69
3.2 Introduction ............................................................. 70
3.3 Results ....................................................................... 72
   3.3.1 Hho1p binds to the linker DNA region globally .......... 72
   3.3.2 Hho1p binds to the promoter region in subset of genome 74
   3.3.3 Correlation between Hho1p occupancy and nucleosome .... 79
   3.3.4 The occupancy of Hho1p increases in the majority of yeast genome in spores compared to the initiation of meiosis ........... 82
   3.3.5 Highly expressed genes in spores have lower Hho1p occupancy ............ 83
3.4 Discussion ................................................................... 85
   3.4.1 Regulation of global level of Hho1p in mature spores .... 85
   3.4.2 Hho1 occupancy and histone density ......................... 87
   3.4.3 Hho1 distribution and transcription regulation ............. 87
   3.4.4 Hho1p enrichment at 5’ NFR ................................... 89
3.5 Material and Methods ................................................... 91
   3.5.1 Sporulation induction ............................................. 91
   3.5.2 Preparation of sequencing sample ............................ 91
   3.5.3 Yeast whole cell extract preparation and western blot analysis ..... 92
   3.5.4 Data analysis ....................................................... 93

Chapter 4 Tetrasome does not reside on original nucleosome dyad and may be positioned by DNA sequence ............................................. 94

4.1 Summary ................................................................... 94
4.2 Introduction ............................................................. 95
4.3 Results ....................................................................... 98
   4.3.1 Tetrasomes are relatively rare in vivo ....................... 98
   4.3.2 Tetrasomes at 5’ and 3’ NFR are sensitive to MNase digestion 101
   4.3.3 Transcription regulates the stability of fragile tetrasomes 102
   4.3.4 Tetrasomes are not over-digested nucleosomes ............ 106
   4.3.5 Tetrasome dyads shift away from the canonical dyad locations of nucleosomes ........................................... 107
   4.3.6 Positioning of tetrasomes is independent of nucleosomes 113
4.3.7 Tetrasomes occupancy level is not positively correlated with transcription frequency ................................................................. 116
4.4 Discussion ...................................................................................................................................................................................... 117
  4.4.1 Dynamic picture of nucleosome assembly and disassembly .............................................. 117
  4.4.2 Genic tetrasome level and transcription frequency ......................................................... 118
  4.4.3 Unstable tetrasome in the NFR regions ............................................................................. 119
4.5 Material and Methods ................................................................................................................................................................. 121
  4.5.1 Tetrasome isolation ........................................................................................................ 121
  4.5.2 Data analysis ................................................................................................................ 121

Chapter 5 Discussion .............................................................................................................................................................................. 123

  5.1 Summary of this study ......................................................................................................................... 123
    5.1.1 Stable and specific changes of primary nucleosome structure during yeast meiosis ........................................................................ 123
    5.1.2 Mysterious linker histone in yeast ...................................................................................... 124
    5.1.2 Tetrasome as the intermediate of nucleosome assembly and disassembly .... 124
  5.2 Future Direction ............................................................................................................................................................................. 125

References ........................................................................................................................................................................................................ 128
LIST OF FIGURES

Figure 1-1. The major structures in chromatin organization. ................................................................. 4
Figure 1-2. Nucleosome landscape on yeast genome. ............................................................................. 7
Figure 2-1. Assessment of progress through sporulation. ................................................................. 25
Figure 2-2. Nucleosome organization around genes throughout sporulation. .............................. 26
Figure 2-3. Isolation of nucleosomal core particle. ............................................................................... 27
Figure 2-4. Western blot to examine H3 tail cleavage during log-phase growth and sporulation. ........................................................... 28
Figure 2-5. Cluster view of bulk nucleosome and histone modifications during log-phase growth and sporulation. .................................................................................. 32
Figure 2-6. Nucleosome organization around genomic features throughout sporulation.............. 33
Figure 2-7. +1 nucleosome shift is linked with carbon starvation regulation. ......................... 37
Figure 2-8. Predicted nucleosome occupancy for shifted genes......................................................... 38
Figure 2-9. Quantification of histone levels and histone modifications density during sporulation.............................................................................................................. 40
Figure 2-10. Nucleosomal and histone modifications distribution around genomic features. ......................................................................................................................... 41
Figure 2-11. Distribution of histone and histone modifications states according to transcription frequency. ......................................................................................................................... 43
Figure 2-12. Distribution of histone and histone modifications states around meiotic genes during meiosis and sporulation. .................................................................................. 45
Figure 2-13. Distribution of nucleosome and histone modifications states around meiotic recombination hot and cold spots. .................................................................................. 47
Figure 2-14. Distribution of nucleosome and histone modifications states around meiotic recombination hot and cold spots. .................................................................................. 48
Figure 2-15. H3K4me3 provides a signature for repressive antisense transcription............... 50
Figure 2-16. Modification states around rDNA locus. ........................................................................... 53
Figure 2-17. Multiple factors set up the increase of nucleosomal spacing at stationary phase. ................................................................................................................................. 55
Figure 3-1. ChIP-PCR to evaluate the Hho1p ChIP efficiency. ............................................................. 73
Figure 3-2. Histone H1 distribution pattern around TSS/TES in the genome. .......................... 74
Figure 3-3. Histone H1 heatmap overlap between 0 hr and 48 hr around genes. ....................... 77
Figure 3-4. Histone H1 distribution at RP genes between 0 hr and 48 hr around TSS of genes. .......................................................................................... 78
Figure 3-5. Heatmap of Hho1p and H3 at RP genes at 0 hr ..................................................... 79
Figure 3-6. Histone H1 scatter plot with bulk histone and histone modifications ......................... 81
Figure 3-7. Genomic fragment enrichment of Hho1p in spores versus meiotic onset. ............. 83
Figure 3-8. Hho1p occupancy of highly expressed genes in spores ........................................ 84
Figure 4-1. PCR amplification of different size fragment ......................................................... 99
Figure 4-2. WC distance frequency distribution of sequenced fragment in all three conditions ................................................................................................................. 99
Figure 4-3. NFR tetrasomes are more sensitive to MNase treatment ...................................... 104
Figure 4-4. Heatmap of tetrasome distribution at 24 hr ......................................................... 105
Figure 4-5. Tetrasomes peaks are shifted away from nucleosome dyad .................. 108
Figure 4-6. Tetrasome tags distribution relative to nucleosome dyad in YPD ......................... 109
Figure 4-7. Dinucleotide distribution relative to nucleosome & tetrasome dyad in YPD ...... 110
Figure 4-8. Tetrasome tag distribution relative to nucleosome dyad in two MNase digestion levels .................................................................................................................. 112
Figure 4-9. Heatmap of tetrasome dyad relative to nucleosome dyad in ISW1 mutant .......... 114
Figure 4-10. Heatmap of tetrasome dyad relative to nucleosome dyad in wild type and ISW1 mutant .................................................................................................................. 115
Figure 4-11. Frequency distribution of tetrasome density in 4 groups of genes. ................. 116
# LIST OF TABLES

Table 2-1. Sequencing tag count summary

Table 2-2. Data matrix for Chi-squared test

Table 3-1. Sequencing read counts summary

Table 3-2. Correlation between Hho1p and other transcription factors in RP genes

Table 4-1. Sequencing tag count summary

Table 4-2. Chi-squared test for different types of nucleosome
ACKNOWLEDGEMENTS

I am grateful to many wonderful people who contributed to the work presented in this thesis.

Sequencing Experiments – I would like to thank Ho Sung Rhee, the Genomics Core Facility in Penn State University, Ka Yim (Anna) Chan to run the SOLiD/illumina sequencer. I also want to thank Elissa Ward, Matthew Schaner and Ka Yim Chan for organizing the sequencing run and sequencing sample. I thank Yunfei Li and Dr. Albert for handling the raw sequencing data and mapping reads to the genome.

Analysis Scripts and Pipelines – I benefited a lot from the scripts written by Zhenhai Zhang and Cizhong Jiang. I would also thank Pindi Albert for the ChIP-exo pipeline scripts and Gue Su Chang for the scripts and help in processing the pair-end sequencing run.

Data chapters – I would like to thank Microscopy and Cytometry Facility in University Park for the assistance in flow cytometry experiments in chapter 2. Jessica Bryant in Shelly Berger’s lab in University of Pennsylvania harvested the Hho1-FLAG tag yeast cells and performed the western blot experiments for Hho1p in chapter 3. I thank Dr. Joe Reese for providing the ISW1 deletion strain.

Center for Gene Regulation – I thank the members in Gene regulation for providing helpful discussion. I will definitely miss the lively discussion in the megameeting.

Thesis committee – I would like to thank the members in my committee for their support and critical suggestions, especially from Joe Reese.
Pugh Lab – I would like to thank Christine Walsh and Travis Mavrich for teaching me the microarray experiment during my rotation. And the efforts of Christine and other technicians make it much easier and productive to work in the Pugh lab, and I appreciate that very much. I would like to thank all the lab members in the lab – Kuangyun Yen, Tom Koerber, Bryan Venter, Sujana Ghosh, Megha Walsh, Kiran Batta, Celine Ham, Matt Rossi, Vinesh Vinayachandran, Shinichiro Wachi for all your help and support. You make Pugh lab a wonderful and lively place to work.

Ma Lab – I thank the continuing support from all the members in the Ma Lab, Ma Lab is really like a big family to me.

Advisors – I feel really grateful to have both Frank and Hong as my advisors. I thank Frank for the continuing support and guidance. In short, Frank is a paragon of scientist for me to learn. As busy as he was, Hong is always a great scientist and a mentor. I can feel that Hong really care about students and try to share all his experience to help his student to mature and develop a successful career path.

Family – I would like to thank continuing support from my family and Zhenzhen. I would not be able to do this without you.
Chapter 1

Introduction to chromatin

1.1 Multiple levels of chromatin structure

In eukaryotic cells, DNA is packed into chromatin in order to compress the genome into a much smaller volume, and to precisely control the accessibility of DNA during transcription and recombination. Multiple levels of chromatin structure have been proposed (Woodcock and Dimitrov 2001) and each level of chromatin structure has its own function in the cellular processes. I will describe each level of chromatin structure in the following sections.

1.1.1 Primary level of chromatin structure

Around 147 base pair (bp) of DNA is wrapped around a mono-nucleosome. A mono-nucleosome consists of two copies of H2A, H2B, H3 and H4 histone proteins. The crystal structure shows that an (H3-H4)2 tetramer binds to the central portion of this 147 bp region of DNA, while two H2A-H2B dimers localize at the boundaries (Luger et al. 1997). Nucleosomes are normally separated by a fixed length of linker DNA. The linker histone H1 contributes to the packing of chromatin into 30 nm chromatin fiber by binding to both the nucleosome octamer and linker DNA (Pruss et al. 1996; Li et al. 2010). However, the spacing between two neighboring nucleosomes (length of linker DNA) differs significantly among different species. The length of linker DNA is around 18 bp in S. cerevisiae (Lee et al. 2007; Mavrich et al. 2008a), 28 bp in D. melanogaster (Mavrich et al. 2008b) and C. elegan (Valouev et al. 2008), and 38 bp in human (Barski et al. 2007; Valouev et al. 2011). Such bead-on-a-string organization of nucleosome and
linker DNA is defined as the primary level of chromatin structure (Woodcock and Dimitrov 2001).

Multiple factors regulate the linear arrangement of nucleosome on DNA, including DNA itself. Nucleosome formation favors DNA sequences with higher GC content in the nucleosome core, as previous studies suggested that CG/CG, CC, GG dinucleotide steps are more flexible than AT/TA, TT, AA steps (Packer et al. 2000; Kaplan et al. 2009; Valouev et al. 2011). In addition to GC content in the dyad, 10 bp periodicity of AT/TA, AA, TT also serves as strong nucleosomal positioning sequences due to their tendency to expand the major groove of DNA (Jiang and Pugh 2009). A 10 bp periodicity for GC/CG, CC, GG dinucleotides is offset by 5 bp compared to AT/TA, AA, TT dinucleotides, leading to a GC/CG, CC, GG dinucleotides enrichment in the minor groove of DNA (Jiang and Pugh 2009). Consistent with this observation, the relative flexible GC/CG, CC, GG dinucleotides are more likely to accommodate the sharp local bending of DNA in minor groove that are imposed by histone proteins (Tolstorukov et al. 2007). Such constraint in the bending of nucleosomal DNA may also explain why poly (dA:dT) sequences disfavor nucleosome deposition (Kaplan et al. 2009).

In addition to nucleosome positioning by DNA sequence, there are multiple cellular factors that influence nucleosome positioning in vivo. In vitro reconstitution assays showed that the DNA sequence itself is insufficient to recapitulate nucleosome positioning in vivo. Instead, ATP-dependent chromatin remodeling activity is required to restore nucleosomes to their in vivo positions (Gkikopulos et al. 2011; Zhang et al. 2011b). Several sequence specific transcriptional factor, such as Reb1p and Abf1p, have been shown to be essential in maintaining a subset of 5’ nucleosome free regions (NFR) by working in concert with RSC chromatin remodeling complex (Hartley and Madhani 2009). In fact, chromatin-remodeling complexes can directly regulate nucleosome positioning and occupancy levels (Cairns 2009; Tirosh et al. 2010; Zhang et al. 2011b). The positioning and occupancy level of nucleosomes regulate the accessibility of DNA to
other DNA-binding proteins, thereby controlling transcription, recombination and other DNA-templated processes (Li et al. 2007a). Moreover, recent studies suggested that nucleosome organization also influences evolution (Warnecke et al. 2008; Zhang and Pugh 2011).

### 1.1.2 Secondary and tertiary level of chromatin structure

Higher levels of chromatin structure are formed to compact the DNA into nucleus. The secondary level of chromatin structure is referred as “structures formed by interactions of nucleosomes” (Woodcock and Dimitrov 2001). With the addition of linker histone, a linear nucleosome array with “beads-on-a-string” structure is coiled into helical structure with a diameter of 30 nm (Figure 1-1). The majority of recent studies on secondary chromatin structure focused on “30 nm” chromatin fiber in vitro (Figure 1-1). Previous in vitro results also showed that both histone H1 and deacetylation of core histones are essential to reconstitute fully compacted 30 nm chromatin fiber (Li et al. 2010). Loss of histone H1 incorporation into chromatin in Iswi mutant of Drosophila leads to dramatic alterations in the high-order chromatin structure in mitotic chromosomes (Corona et al. 2007).

The secondary structures can be further packed into the tertiary level of chromatin structure, such as metaphase chromosomes (Figure 1-1). In addition, in vitro reconstituted 120-170 nm chromatin fibers were observed under the electron microscope (Kireev et al. 2008). The tertiary structures also include chromosome long-range interaction, such as enhancer and insulator looping. Several studies clearly demonstrated that chromosome long-range interaction plays an important role in transcription regulation (Montavon et al. 2011; de Wit and de Laat 2012). Recent development and advances of Chromosome Conformation Capture (3C) technology have revealed the nuclear organization at an unprecedented resolution and coverage in multiple model organisms (Duan et al. 2010; Umbarger et al. 2011; Sexton et al. 2012).
1.2 Histone variants and modifications

1.2.1 Functions of histone variants

Nucleosomes regulate the accessibility of DNA to other DNA binding proteins. However, the property of nucleosomes concerning their binding affinity to DNA and other proteins is not always identical. In fact, the property of a nucleosome is determined by the composition of histone variants and covalent post-translational modifications on the histone tails. Both histone variants and histone modifications act in diverse cellular processes including gene regulation and chromatin compaction (Santisteban et al. 2000; Krishnamoorthy et al. 2006; Kouzarides 2007).

All the histone proteins are highly conserved from yeast to human. In metazoan genome, there are gene families for each histone subunit. The histone variants were evolved to carry out more specialized functions, especially for histone H3 and H2A (Henikoff and Ahmad 2005). In yeast, there are two classes of histone H3: histone H3 and centromeric histone H3-like protein CSE4, which is a homolog of human CENP-A. In addition to canonical H3 and centromeric H3 variant CENP-A, metazoan genomes normally encode a third class—H3.3 variant. The occupancy level of H3.3 variant is closely associated with active transcription (Ahmad and
Henikoff 2002; Schwartz and Ahmad 2005). In *Saccharomyces cerevisiae*, histone H2A has two variants—H2A.Z (Htz1) and H2A. The Htz1 containing nucleosomes are less stable compared to H2A containing nucleosomes, thus promoting transcription activation (Santisteban et al. 2000). Histone H1 is the most divergent histone among the histone proteins in mammals, including five somatic subtypes, one terminally differentiated expressed isoforms and two tissue specific variants (Talasz et al. 1996; Parseghian and Hamkalo 2001). Interestingly, only one linker histone gene (*HHT1*) was identified in the yeast genome (Patterton et al. 1998).

1.2.2 Functions of histone modifications

Each histone subunit can be modified at its amino-terminal tail. Such modifications include acetylation, methylation, phosphorylation and ubiquitylation. Various histone modifications function in diverse biological processes including transcription regulation, chromatin condensation and recombination (Kouzarides 2007).

The acetylation of histone tails, more specifically the side chain of lysine, neutralizes the positively charged ammonium group, thus decreasing the affinity between histone tail and negatively charged DNA. Consistent with a less stable histone-DNA interaction, acetylation on histone tails, such as H3 K9ac, H3 K14ac, is positively correlated with active transcription (Pokholok et al. 2005). Nevertheless, the acetylation on specific lysine residues renders distinct functions. The acetylation of H3 K56 facilitates the nucleosome assembly of newly synthesized histone (Li et al. 2008b); H4 K16 acetylation prevents high-order chromatin compaction *in vitro* (Shogren-Knaak et al. 2006).

In contrast with the more transient acetylation marks, methylation marks are much more stable and long-lived. The addition of a methyl group does not alter the charge of lysine and arginine residues. Instead of changing the charge of an amino acid, the methylation of histone
affects histone-protein interaction. The “reader” proteins with PHD or chromo domain bind to methylated histones to assist the recruitment of other components of the transcription machinery, which eventually may facilitate transcription initiation (Hung et al. 2009). H3 lysine 4 and H3 lysine 36 tri-methylation are considered to be active marks for transcription; whether H3 Lysine 79 tri-methylation is an active histone mark is still in debate (Kouzarides 2007). In the meantime, repressive histone methylation marks recruit other proteins to repress transcription. HP1 proteins bind to H3 K9 methylation to form heterochromatin; H3 K27 tri-methylation facilitates the binding of Polycomb proteins to HOX gene promoters; and H4 K20 methylation mark plays an essential role in recruiting checkpoint protein 53BP1/Crb2 to DNA break point (Bannister et al. 2001; Cao et al. 2002; Botuyan et al. 2006).

In fact, up to two methyl groups can be added to an arginine side-chain, while lysine can be methylated up to three times. Importantly, each methylation state on lysine may represent a distinct biological signal. Both H3 K4me2 and H3 K4me3 are associated with active transcription, while H3 K4me1 is enriched in the enhancer region (Barski et al. 2007). Genome-wide studies suggested that mono-methylation of “repressive” H3 K9 is positively correlated with the levels of expression, while di-methylation and tri-methylation of H3 K9 serve as repressive marks (Barski et al. 2007; Rosenfeld et al. 2009). Interestingly, all three repressive methylation marks (H3 K9, H3 K27 and H4 K20 methylation) are absent in budding yeast (Kouzarides 2007).

The phosphorylation of a protein can regulate enzyme activity and protein-protein interaction, thereby playing a crucial role in signal transduction pathway. Little is known about the role of histone phosphorylation in transcription regulation. Nevertheless, multiple histone phosphorylation marks are associated with chromatin compaction. Phosphorylation on H3 S10 and H3 T3 facilitate mitotic/meiotic chromosome condensation, while H4 S1 phosphorylation contributes to post-meiotic nuclear compaction in multiple model organisms (Dai et al. 2005; Fischle et al. 2005; Krishnamoorthy et al. 2006).
Both ubiquitylation and sumoylation are large modifications, and they are mainly found on H2A and H2B histone subunits. H2B K123 ubiquitylation is associated with transcriptional activation and required for histone methylation at H3K4 and H3K79 (Sun and Allis 2002).

In fact, there is extensive crosstalk between different histone modifications (Lee et al. 2010). These marks do not act alone in cells, therefore it will be important to understand how they function cooperatively.

1.3 Nucleosome landscape and transcription

Figure 1-2. Nucleosome landscape on yeast genome.
(Jiang and Pugh 2009)
One of the most important functions of chromatin is to regulate gene expression; chromatin can modulate transcription in multiple levels including the primary chromatin structure and the subnuclear localization in the nucleus. This thesis mainly focuses on the primary nucleosome structure, which is where nucleosomes, modified histones or other transcription factors bind in the genome.

Nucleosomes are not randomly distributed along the gene body; instead, there is a highly conserved organizational pattern across the eukaryotic species (Figure 1-2). There is a Nucleosome Free Region (NFR) before the Transcription Start Site (TSS) of a gene, and “-1” nucleosome at the upstream of 5’ NFR region normally regulates the accessibility of cis-regulatory elements to different transcription factors (Albert et al. 2007). The “+1” nucleosome resides mostly at +60 bp relative to the TSS in Saccharomyces cerevisiae, while the “+1” nucleosome resides at +135 bp in fruit fly (Mavrich et al. 2008b). Normally the “+1” nucleosome shows the tightest positioning compared to the downstream nucleosomes. The tight positioning of the “+1” nucleosome is driven by ATP-dependent packing mechanisms (Mavrich et al. 2008a; Zhang et al. 2011b). Following the “+1” nucleosome, a nucleosomal array is formed on the gene body with proper nucleosomal spacing. The fuzziness of nucleosomes increases with the distance to TSS. At the end of a gene, there is also a 3’ NFR, where Pol II terminates transcription (Figure 1-2). Unlike 5’ NFR, the nucleosome depletion is not intrinsic in 3’ NFR and relies on Pol II occupancy (Fan et al. 2010). Indeed, high resolution ChIP-exo experiments showed that Pol II is enriched after the Transcription End Site (TES) (Rhee and Pugh 2012).

The position, occupancy and fuzziness are three important parameters of a nucleosome. Moreover, the shifting of nucleosome arrays was observed between the paused and transcribed genes in Drosophila, which suggested that repositioning of nucleosome, especially the ‘+1’ nucleosomes, serves a general mechanism for transcription regulation (Mavrich et al. 2008b). A nucleosome positioned in the promoter of PHO5 in yeast can block the recruitment of pre-
initiation complex and represses transcription. The removal of the promoter nucleosome(s) of the PHO5 gene is required to activate transcription (Han and Grunstein 1988; Workman 2006).

Fuzziness is used to describe the stability of nucleosome positioning among the whole population of cells. Lower level of nucleosomal fuzziness means that the position of a nucleosome is highly fixed, while higher level of fuzziness indicates that a nucleosome has more freedom to move around (Mavrich et al. 2008a).

Not only are nucleosomes organized along genes, but different histone modifications on core histones also show distinct patterns in the context of genes. The active acetylation marks, such as H3 K9ac, are normally enriched at the 5’ end of a gene, which is written by various histone acetylases coupled with the transcription machinery (Kurdistani et al. 2004; Wang et al. 2009). Similarly, H3 K4 tri-methylation, which is also an active histone mark, concentrates at the 5’ end of a gene (Pokholok et al. 2005; Barski et al. 2007). Another active tri-methylation mark on H3 K36 shows enrichment at genic region and peaks at the 3’ end of a gene (Pokholok et al. 2005; Barski et al. 2007). Genome-wide mapping of H3 K79me3 in yeast showed that this mark is distributed evenly on the gene body (Pokholok et al. 2005).

In fact, most of active histone marks are written on the histone tails during active transcription. Previous ChIP-chip studies have already shown that the level of active histone modification marks is positively correlated with transcription frequency (Pokholok et al. 2005). Gcn5p is the histone acetyltransferase for histone H3 (K9, K14, K18, K23 and K27); it is also a shared subunit of ADA, SLIK/SALSA and SAGA complexes and functions as transcription co-activator (Grant et al. 1997). Although the general transcription machinery does not recruit Gcn5p directly, Gcn5p occupies most actively transcribed promoter (Robert et al. 2004; Imoberdorf et al. 2006). The mutation on the HAT domain of Gcn5p leads to changes in nucleosome positioning in promoter region, transcription regulation (Gregory et al. 1998). Nevertheless, histone deacetylases (HDACs) also control the acetylation pattern in the genome. Genome-wide studies
indicated that the occupancy of HDACs is not associated with transcription frequency. Similar to histone acetylases, HDACs are also recruited by sequence specific transcription factors or other histone methylation marks (Robert et al. 2004; Li et al. 2007b).

In *Saccharomyces cerevisiae*, there is only one methyltransferase Set1p for H3 K4, and Set1p containing complex (COMPASS) travels with transcription initiation complex via Paf1 complex (Briggs et al. 2001; Miller et al. 2001; Krogan et al. 2003a; Ng et al. 2003). The specific methyl-demethylase for lysine 4, Jhd2p, has been identified. Jhd2p can significantly lower H3 K4me3 levels during transcription shutdown (Ingvarsdottir et al. 2007; Liang et al. 2007). H3 K36 methyltransferase Set2p binds to C-terminal Domain (CTD) S2 phosphorylated form of RNA Polymerase II complex (Pol II) and moves together with transcription elongation complex (Schaft et al. 2003; Joshi and Struhl 2005). Consistently, the loss of the phosphorylation kinase of Pol II CTD or partial deletion of CTD leads to the selective loss of H3 K36 methylation mark (Xiao et al. 2003). Jhd1p has been identified as the H3 K36 specific demethylase in yeast, which is highly conserved across eukaryotic species (Tsukada et al. 2006).

So far, we focus on the active expressed mRNA genes, which adopt open chromatin structure—euchromatin. While transposons and other repressed regions in the genome maintain repressive chromatin structure—heterochromatin. Although all repressive histone methylation marks are absent in *Saccharomyces cerevisiae*, silent mating type loci (HMR, HML), rDNA and sub-telomeric regions are still packed into heterochromatin. Loss of heterochromatin in the genome causes genome instability and is associated with increased susceptibility of cancer and other diseases in human (Oberdoerffer and Sinclair 2007; Hahn et al. 2010).
1.4 Nucleosome dynamics during transcription and DNA replication

A nucleosome consists of 2 copies of H2A, H2B, H3 and H4; however, a histone octamer does not bind DNA all together to form nucleosome. Instead, (H3-H4)$_2$ tetramers normally associate with DNA first, followed by the assembly of 2 copies of H2A/H2B dimer (Almouzni et al. 1990; Annunziato 2005). A (H3-H4)$_2$ tetramer together with its interacting DNA is referred to a tetrasome. Similarly, H2A/H2B dimer also disassociates from DNA earlier than (H3-H4)$_2$ tetramer during nucleosome disassembly. Consistently, the binding affinity between (H3-H4)$_2$ tetramer and DNA is much higher compared to that between H2A/H2B dimer and DNA. During transcription and DNA replication, nucleosome structure is perturbed, nucleosome disassembly and assembly occur to allow DNA/RNA polymerase complex moving through gene and to form the proper nucleosome array afterwards (Kireeva et al. 2002). Therefore, understanding the intermediate (H3-H4)$_2$ tetrasome will deepen our understanding about the dynamic process of nucleosome assembly and disassembly process.

Several in vivo studies indicated that the majority of (H3-H4)$_2$ or (H3.1-H4)$_2$ tetramers remain bound to the DNA upon the transcription and DNA replication, while most of H2A/H2B dimers are removed and exchanged (Yamasu and Senshu 1990; Thiriet and Hayes 2005; Xu et al. 2010). Still, it is not known whether these tetrasomes maintain their original positions before nucleosome disassembly. In vitro reconstitution experiments showed that (H3-H4)$_2$ tetramer alone contributes to the positioning of core histone and could assemble on the DNA by itself (Dong and van Holde 1991; Hayes et al. 1991). Nevertheless, the conclusions from in vitro experiments were based on DNA template containing strong positioning sequence, which might not be representative for the whole genome in vivo. Other studies have focused on the nucleosome survival during Pol II transcription by in vitro transcription system. Moderate transcription generated hexasome, which is consisted of (H3-H4)$_2$ tetramer, one H2A/H2B dimer
and the associated DNA. Hexasomes on shorter template DNA (250-300 bp) remains at the original position during transcription (Kireeva et al. 2002), while the hexasomes on longer template DNA (658-3900 bp) translocate upon Pol II transcription (Hodges et al. 2009).

So far, no genome-wide in vivo mapping of tetrasomes has been reported. Such genome-wide mapping may shed light on the following questions—whether tetrasome positioning is DNA encoded and whether tetrasomes maintain their original positions during transcription and DNA replication.

1.5 Chromatin structure during yeast meiosis

1.5.1 Overview of chromatin structure in meiosis

In addition to transcription regulation, chromatin also plays important roles in chromatin condensation and meiotic recombination, both of which are crucial to ensure the proper progression of meiosis.

During meiotic cell cycle, a cell duplicates its genome and undergoes two consecutive rounds of chromosome segregations (meiosis I and meiosis II) to form four daughter cells, each containing half of the genetic material of the parental cell. Therefore, the fusion of two germ cells can restore the ploidy in the offspring. Meiosis is essential to the survival of the species and the generation of genetic diversity. For most of eukaryotes, the combination of both parents genetic information increases genetic diversity of their offspring. In addition, the meiotic recombination occurs between two homologous chromosomes, which results in gene conversion (copying sequence from homologous chromosomes) or crossover (exchange of regions between homologous chromosomes), thus adding another layer of genetics diversity to the next generation.
The DNA duplication between meiosis and mitosis is similar, but meiosis I is unique in that meiotic recombination occurs in high frequency and homologous chromosomes segregate but not sister chromatids. Meiosis I can be further divided into four stages: prophase I, metaphase I, anaphase I and telophase I. Chromosomes start to condense through the condensin complex during prophase I, which is important to achieve successful segregation (Klein et al. 1999). To ensure correct segregation of homologous chromosomes, pairing between homologous chromosomes occurs based on homology in DNA sequence (called synapsis), which also involves in the formation of the Synaptonemal Complex (SC). SC is consisted of two parallel filamentous elements and central elements that connect the two parallel filamentous elements. The gene conversion and crossover start with Double Strand Breaks (DSBs) in the genome, followed by resection of a single strand DNA from 5’ end through MRX complex. Heterotrimer RPA complex is then recruited to single strand DNA with 3’ overhang (Brill and Stillman 1991; Alani et al. 1992), the binding of Dmc1p and/or Rad51p is also necessary to form the nucleoprotein filament, which will be used in search for homologous DNA (Bishop et al. 1992; Sung 1994). After finding the sequence homology, a single stranded nucleoprotein filament moves into the paired DNA from homologous chromosomes. Such process is called strand invasion. Then the Holliday junction is formed; and decision between gene conversion and crossover is determined by how the Holliday junction is resolved. At least one crossover occurs between each pair of homologous chromosomes to maintain the association between two homologous chromosomes. In metaphase I, homologous pairs move to the equatorial plate. Then homologous pairs are pulled apart during anaphase I, and two nuclei form during telophase I. Then the cells enter meiosis II, in which the segregation of sister chromatid take places as mitosis.

Yeast sporulation is triggered by the lack of fermentable carbon source and nitrogen source (Freese et al. 1982). Both of these conditions are required for sporulation induction. If cells are induced with starvation of nitrogen source but not glucose, it will lead to pseudohyphal
growth but not sporulation (Gimeno et al. 1992). After induction of sporulation for the SK1 strain yeast cells, cells start to undergo the meiotic replication after two hours upon induction (Chu et al. 1998). And genome duplication would be completed at around five hour. Yeast chromosomes are then condensed and two rounds of the chromosome segregation occur within seven to nine hours after induction, followed by cell wall formation and spore maturation. Interestingly, post-meiotic chromatin compaction occurs in both sporulation and human spermatogenesis (Krishnamoorthy et al. 2006; Balhorn 2007). The compaction of spermatogenesis in human is facilitated by the replacement of histones by protamines (Jayaramaiah Raja and Renkawitz-Pohl 2005). But no homolog of protamines is found in the yeast genome, and the mechanism of post-meiotic compaction in yeast spores is still not clear.

1.5.2 Chromatin structure in recombination sites

One of the important questions in meiosis is where meiotic recombination occurs in the genome. The DSB sites in yeast are initially generated by Spo11p, which is a meiotic specific endonuclease (Keeney 2001). Nevertheless, Spo11p itself is not able to generate DSB; several other proteins are needed to successfully catalyze DNA cleavage (Prieler et al. 2005; Robine et al. 2007). Multiple genome-wide studies have shown that the meiotic recombination does not occur randomly in the yeast genome, instead, there are hotspots and coldspots of recombination on the chromosomes (Gerton et al. 2000a; Blitzblau et al. 2007; Buhler et al. 2007). It is also clear that the recombination hotspots tend to have lower nucleosome occupancy, while there is elevated level of nucleosome associated with the recombination coldspots (Ohta et al. 1994; Fan and Petes 1996; Berchowitz et al. 2009).

Spo11p does not bind to specific DNA motif. Still, it is not known what proteins recruit Spo11p to DNA and further dictate the frequency of recombination initiation on the genome.
Recent studies suggested that hotspots of recombination share similar histone modification patterns, such as enrichment of H3 K4me3 mark (Buard et al. 2009). The histone H3 K4 methyltransferase SET1 mutant strain of budding yeast exhibits divergent distribution pattern of recombination hotspots in the genome (Borde et al. 2009). Consistent with the association between Set1p and DSB, the recruitment of Set1p and elevated H3 K4me3 levels are observed in a newly created DSB in vivo. The transcription level of the gene that is closest to the DSB site does not change, suggesting that the recruitment of Set1p is by a DSB-dependent and transcription-independent manner (Faucher and Wellinger 2010). Moreover, the H3 K4 methyltransferase is also required during recombination initiation in mammals (Parvanov et al. 2010). In fact, several other marks (H3 K14ac, H3 K36me3 and H3 K79me3) have been shown to show correlation with DSB (Hansen et al. 2011). Most of histone modification data in yeast were generated by low-resolution microarray technology, in Chapter 2 we generated high-resolution ChIP-seq data of bulk nucleosome and several histone modification marks in order to obtain a high-resolution distribution pattern of histone modification marks around DSB hotspots and coldspots.

1.6 Transcription regulation of sporulation in yeast

1.6.1 Repression of sporulation specific genes during vegetative growth

Constant repression of meiotic/sporulation specific genes is important during vegetative growth for diploid cells. Ume6p is a sequence-specific DNA binding protein that associates with URS sequence in a number of early meiotic genes (Strich et al. 1994; Anderson et al. 1995). Ume6p recruits Isw2 and Rpd3-Sin3 complexes, and they function together to maintain repressive chromatin structure in the promoter region and inhibit the recruitment of general
transcription machinery (Kadosh and Struhl 1997; Goldmark et al. 2000). Upon sporulation induction, Ume6p is quickly degraded by Cdc20–directed APC/C to relieve transcription repression (Mallory et al. 2007). Consequently, the deletion of UME6 gene leads to 2 to 10 folds de-repression of these early meiotic genes (Strich et al. 1994). Sum1p is a sequence-specific transcription factor that represses the middle meiotic genes (Xie et al. 1999; Pierce et al. 2003). It has been shown that Sum1p recruits another histone deacetylase Hst1p by Rfm1p. The recruitment of Hst1p is essential to fully repress middle sporulation genes (McCord et al. 2003).

In haploid cells, there are additional mechanisms that prevent the expression of sporulation specific genes. IME4 gene is required for initiation of meiosis by activating IME1 transcription (Clancy et al. 2002). Interestingly, IME4 gene is repressed by antisense transcript in haploid but not in diploid cells (Hongay et al. 2006). Such antisense transcript can be repressed by a1/α2 heterodimer. The a1/α2 heterodimer is only produced in diploid cells, but not on haploid cells. Similarly, without the binding of a1/α2 heterodimer at the promoter of RME1 in haploid, Rme1p can be actively transcribed (Strathern et al. 1981; Covitz et al. 1991). Then Rme1p binds to two upstream cis-elements in IME1 promoter and recruits Sin4p and Rgr1p to form a repressive chromatin structure in haploid cells (Covitz et al. 1994; Shimizu et al. 1998). Consequently, haploid cells activate Rme1p in order to inhibit the expression of IME1 gene.

1.6.2 Activation of sporulation specific genes

The yeast sporulation program is tightly regulated by the nutrient. The strict regulation of sporulation induction is ensured by accurate transcriptional regulation of IME1 gene. ~2.1 kb IME1 promoter contains multiple cis-regulatory elements, which can integrate signals from various nutrient sources and fine-tune the transcription level of IME1 gene (Kassir et al. 1988; Kassir et al. 2003; Govin and Berger 2009). Ime1p binds to upstream URS element and induces
expression of most of the early meiotic genes (Kassir et al. 1988). Consistently, sporulation is absent in *IME1* null mutant (Kassir et al. 1988; Enyenihi and Saunders 2003).

The middle sporulation genes are upregulated by Ndt80p, which competes with Sum1p for the MSE element in the promoter (Hepworth et al. 1998; Pierce et al. 2003). In the *NDT80* null strain, cells are arrested at pachytene stage (prophase I) of meiosis (Hepworth et al. 1998). There are both URS and MSE elements in the promoter of *NDT80* gene, which together repress *NDT80* expression by recruiting Ume6p and Sum1p respectively (Pak and Segall 2002). In addition, both Ime1p and Ime2p are required to activate Ndt80 expression (Chu and Herskowitz 1998; Pak and Segall 2002). Such tight transcriptional cascade may ensure the coordination of two key processes during sporulation, meiotic cell cycle and gamete morphogenesis (Chu and Herskowitz 1998). Nevertheless, the transcription regulation of late sporulation genes is not well characterized.

**1.7 Genome-wide experimental approach to study chromatin**

**1.7.1 Yeast sporulation as model for meiosis**

To understand how the primary nucleosome structure changes during meiosis to regulate transcription regulation and recombination, we choose yeast sporulation as the model system due to its simplicity and similarity to metazoan.

Compared to other plant or metazoan, it is much easier to collect large number of synchronized meiotic cells. Around 90-95% of meiosis frequency can be achieved in the optimized SK1 strain for sporulation (Primig et al. 2000). While in most of the other organisms, it is extremely difficult and costly to obtain a number of meiotic cells in vivo. Given the existing transcriptome dataset on yeast sporulation process, it is possible to understand the changes in
primary nucleosome structure in the context of transcription regulation during sporulation (Chu et al. 1998; Primig et al. 2000).

1.7.2 MNase digestion and ChIP-seq

To study the genome-wide nucleosome position, occupancy and modifications, MNase ChIP-seq technique is applied. Micrococcal Nuclease (MNase) is an endo-exonuclease that cuts both single and double strand DNA (Axel 1975), and is widely used to generate mono-nucleosome, which is further used in nucleosome mapping in a variety of species (Albert et al. 2007; Mavrich et al. 2008b; Schones et al. 2008). MNase treatment is generally favored for its high resolution to map nucleosomal borders compared to traditional sonication method, although there is sequence bias in MNase digestion. MNase can digest unbound DNA until it encounters a nucleosome boundary, while the length of free DNA varies in the population of sonicated sample. In addition, chromatin immunoprecipitation for bulk histone or histone modification is performed to select the MNase digested DNA fragment bound by histone or specific modified histone. Finally, different samples were bar-coded (containing a unique DNA sequence identifier) and constructed into a DNA sequencing library and then sequenced on a “next-gen” sequencing platform, such as SOLiD and Illumina.

1.7.3 Overview of next-gen sequencing technology

The major advantage of next-gen sequencing technology compared to the traditional Sanger sequencing is that new technology parallelizes the sequencing process, which generates millions of sequencing output at once (Mardis 2008). The high-throughput greatly decreases the sequencing cost. The cost for sequencing 1 Mb DNA has dropped from $5,292 in September
2001 to $0.09 in October 2011 (Mardis 2011). Even though the read length of a “next-gen” sequencer (30-150 bp) normally is much shorter than that of a capillary sequencer, the read length is more than enough for mapping back into the assembled reference genome, especially for the ChIP-seq application in this study.

Two of the most popular platforms are SOLiD and Illumina sequencing platforms. The SOLiD platform is based on ligation-mediated sequencing method, and it also has unique two-base encoding system to gain further confidence in SNP calling (Valouev et al. 2008). The current SOLiD 4 system claims to deliver >100 Gb usable DNA data per run. The Illumina system adopts sequencing-by-synthesis approach (Metzker 2010) and is able to generate ~600 Gb per run in its HiSeq machine. The data in Chapter 2 were largely produced by SOLiD platform, while the data in Chapter 3 and Chapter 4 were generated by Illumina system. However, the results between two platforms are highly consistent with each other.

Probe design, dye labeling and probe hybridization are all eliminated during library preparation in next-gen sequencing. The sensitivity and dynamic range of sequencing technology also greatly outperform the microarray technology (Asmann et al. 2008). The major hurdle of sequencing technology is the complex downstream bioinformatics analysis. However, with rapid technology advances and the maturation of analytical pipelines, the next-generation sequencing are very likely to become the standard for genome wide ChIP study in the near future (Mardis 2008).
Chapter 2

Stable and dynamic nucleosome states during a meiotic developmental process

2.1 Summary

The plasticity of chromatin organization as chromosomes undergo a full compendium of transactions including DNA replication, recombination, chromatin compaction, and changes in transcription during a developmental program is unknown. We generated genome-wide maps of individual nucleosome organizational states, including positions and occupancy of all nucleosomes, and H3K9 acetylation and H3K4, K36, K79 tri-methylation, during meiotic spore development (gametogenesis) in Saccharomyces. Nucleosome organization was remarkably constant as the genome underwent compaction. However, during an acute meiotic starvation response, nucleosomes were repositioned to alter the accessibility of select transcriptional start sites. Surprisingly, the majority of the meiotic program did not employ this nucleosome repositioning but was dominated by antisense control. Histone modification states were also remarkably stable, being abundant at specific nucleosome positions at three-quarters of all genes, despite most genes being rarely transcribed. Our findings suggest that, during meiosis, the basic features of genomic chromatin organization are essentially a fixed property of chromosomes, but tweaked in a restricted and program-specific manner.
2.2 Introduction

Virtually all aspects of chromosome biology are rooted in chromatin structure, which includes nucleosome organization and modifications states. Yet, it is not known to what extent individual nucleosomes and modification states across a genome reorganize as chromosomes undergo a wide range of transactions from the unfurling of developmental transcription programs to DNA replication, recombination, chromosome segregation and genome compaction. Gametogenesis (spore formation) in budding yeast *Saccharomyces cerevisiae* is a model system to study fundamental cellular mechanisms of development, as well as large scale genomic transactions (Krishnamoorthy et al. 2006; Govin et al. 2010a; Handel and Schimenti 2010). Recent technological advances in high throughput DNA sequencing have allowed genome-wide nucleosome organization to be determined at the individual nucleosome level. We therefore examined nucleosome organization and modification states as yeast cells are reprogrammed from exponential growth in rich media to a starvation state and then allowed to proceed through meiosis and on to spore formation.

In mitotically growing haploid yeast cells, nucleosomes are typically spaced ~165 bp apart (Yuan et al. 2005). In a population of cells, nucleosomes tend to be well-positioned at the 5’ ends of genes and become less positioned (or “fuzzy”) towards the 3’ ends (Yuan et al. 2005; Mavrich et al. 2008a). Most genes have a similar chromatin architecture including a “-1” nucleosome upstream of the core promoter, a nucleosome-free promoter region (5’ NFR), a +1 nucleosome that is positioned to control access of the transcriptional start site (TSS), and an array of nucleosomes beginning at the +1 position and ending at a 3’ NFR (Jiang and Pugh 2009). Superimposed on this organizational state are a variety of histone modifications that tend to be enriched at specific regions along genes (Liu et al. 2005; Barski et al. 2007). A fundamental unanswered question is how plastic this basic organizational theme is.
Transcriptional responses to environmental signals typically entail nucleosome reorganization at reprogrammed genes (Hogan et al. 2006; Schones et al. 2008; Shivaswamy and Iyer 2008; Kaplan et al. 2009). Much of the remaining genome remains unaltered. However, the resolution of the employed methods and the limited perturbations to the system precluded definitive assessment as to whether the primary structure of chromatin can undergo massive reorganization when entire chromosomes are subjected to a full gamut of transactions, such as that occurring during meiosis and spore formation where the chromatin is compacted (Kleckner 1996). Although several studies on a few classical double-stranded break (DSB) hotspots indicate that the chromatin structure of these sites does not detectably change as cells progress through the meiosis (Ohta et al. 1994; Fan and Petes 1996; Borde et al. 1999), it is not known whether this is generalizable. Importantly, low-resolution assays are unable to track short-distance nucleosome repositioning that can have profound effects on the accessibility of specific DNA regulatory elements.

With both the 5’ and 3’ NFRs typically large enough to accommodate a nucleosome, and linker regions between nucleosomes being ~18 bp, approximately 20% of the genome is essentially nucleosome-free (Lee et al. 2007; Shivaswamy and Iyer 2008). Global transcription rates decrease during sporulation (Jona et al. 2000). Since nucleosome occupancy in NFRs has been attributed to transcriptional shutdown (Schones et al. 2008; Shivaswamy and Iyer 2008), NFRs might be expected to acquire nucleosomes during the sporulation program. In addition, the ensuing compaction might also involve nucleosome acquisition in NFRs and inter-nucleosomal spacing changes (Routh et al. 2008; Segal and Widom 2009). Compaction might also be expected to increase nucleosome positioning as the degrees of freedom for positioning become more restricted by inter-chromatin contacts. As far as we know, no study has addressed the issue of the primary nucleosome organizational state as chromosomes proceed through major
transactions during a developmental program such as sporulation, at a resolution sufficient to evaluate individual nucleosome positions, occupancy levels, and modification states.

2.3 Results

2.3.1 The primary organization of chromatin remains fixed during meiosis and upon compaction

Two biological replicates of MNase-treated nucleosomal DNA, prepared from cells grown in rich (YPD) medium, and from cells starved for fermentable carbon and nitrogen for 0, 2, 5, 7, and 24 hours were subjected to whole genome sequencing (ABI SOLiD), producing ~4.7 million nucleosomal sequence tags per sample (~79 tags per nucleosome position, Table 2-1). The progression from stationary phase (which is defined in this study as the conditions employed at 0 hr – see Methods), through meiosis, and on to spore formation was monitored by flow cytometry and DAPI staining analysis. Approximately 95% of the cells had formed spores by 24 hours (Figure2-1), indicating that the observed nucleosome organization (see Figure2-2A) was not due to unsporulated cells. Equivalent levels of mononucleosomes were extracted at each time point, including during spore formation, with very little remaining unextracted (see Figure2-3A), indicating that no significant sub-population was left behind. All mononucleosome preparations were dependent upon addition of exogenous MNase (Figure2-3B), which demonstrates that endogenous nucleases were not contributing to the nucleosome pattern. Moreover, since equivalent amounts of MNase applied to nucleosomes reconstituted on genomic DNA failed to produce the same pattern (Zhang et al. 2009), these patterns were not an artifact of sequence-specific cleavage preferences of MNase.
<table>
<thead>
<tr>
<th>Sample</th>
<th>Biological replicates</th>
<th>Total replicates</th>
<th>Uniquely mapped tags</th>
<th>Histone levels&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>H3 YPD</td>
<td>2</td>
<td>4</td>
<td>27,065,848</td>
<td>1.2</td>
<td>0.3</td>
</tr>
<tr>
<td>H3 0 hr</td>
<td>2</td>
<td>3</td>
<td>8,395,811</td>
<td>1.0</td>
<td>0.1</td>
</tr>
<tr>
<td>H3 2hr</td>
<td>3</td>
<td>5</td>
<td>4,908,104</td>
<td>1.4</td>
<td>0.1</td>
</tr>
<tr>
<td>H3 5hr</td>
<td>2</td>
<td>4</td>
<td>4,110,186</td>
<td>2.3</td>
<td>0.2</td>
</tr>
<tr>
<td>H3 7hr</td>
<td>2</td>
<td>3</td>
<td>5,078,304</td>
<td>3.0</td>
<td>0.2</td>
</tr>
<tr>
<td>H3 24hr</td>
<td>2</td>
<td>3</td>
<td>4,405,746</td>
<td>2.9</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K9ac YPD</td>
<td>2</td>
<td>3</td>
<td>7,152,913</td>
<td>1.9</td>
<td>0.8</td>
</tr>
<tr>
<td>H3K9ac 0 hr</td>
<td>2</td>
<td>3</td>
<td>1,635,189</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td>H3K9ac 2hr</td>
<td>2</td>
<td>3</td>
<td>838,426</td>
<td>1.3</td>
<td>0.3</td>
</tr>
<tr>
<td>H3K9ac 5hr</td>
<td>2</td>
<td>3</td>
<td>770,009</td>
<td>0.9</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K9ac 7hr</td>
<td>2</td>
<td>3</td>
<td>611,617</td>
<td>0.5</td>
<td>0.3</td>
</tr>
<tr>
<td>H3K9ac 24hr</td>
<td>2</td>
<td>4</td>
<td>3,363,267</td>
<td>0.7</td>
<td>0.4</td>
</tr>
<tr>
<td>H3K4me3 YPD</td>
<td>2</td>
<td>4</td>
<td>4,103,085</td>
<td>1.1</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K4me3 0 hr</td>
<td>2</td>
<td>3</td>
<td>5,622,171</td>
<td>1.0</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K4me3 2hr</td>
<td>2</td>
<td>3</td>
<td>7,963,774</td>
<td>0.8</td>
<td>0</td>
</tr>
<tr>
<td>H3K4me3 5hr</td>
<td>2</td>
<td>4</td>
<td>4,983,540</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K4me3 7hr</td>
<td>2</td>
<td>3</td>
<td>2,802,317</td>
<td>0.5</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K4me3 24hr</td>
<td>2</td>
<td>3</td>
<td>10,173,962</td>
<td>0.3</td>
<td>0</td>
</tr>
<tr>
<td>H3K36me3 YPD</td>
<td>2</td>
<td>3</td>
<td>3,682,378</td>
<td>0.9</td>
<td>0.2</td>
</tr>
<tr>
<td>H3K36me3 0 hr</td>
<td>2</td>
<td>3</td>
<td>4,818,470</td>
<td>1.0</td>
<td>0.2</td>
</tr>
<tr>
<td>H3K36me3 2hr</td>
<td>2</td>
<td>3</td>
<td>18,393,931</td>
<td>0.9</td>
<td>0.2</td>
</tr>
<tr>
<td>H3K36me3 5hr</td>
<td>2</td>
<td>3</td>
<td>8,547,837</td>
<td>0.7</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K36me3 7hr</td>
<td>2</td>
<td>3</td>
<td>5,911,048</td>
<td>0.7</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K36me3 24hr</td>
<td>3</td>
<td>4</td>
<td>14,373,748</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>H3K79me3 YPD</td>
<td>2</td>
<td>3</td>
<td>1,164,934</td>
<td>0.8</td>
<td>0.2</td>
</tr>
<tr>
<td>H3K79me3 0 hr</td>
<td>2</td>
<td>3</td>
<td>5,990,003</td>
<td>1.0</td>
<td>0.3</td>
</tr>
<tr>
<td>H3K79me3 2hr</td>
<td>2</td>
<td>3</td>
<td>2,851,785</td>
<td>0.7</td>
<td>0</td>
</tr>
<tr>
<td>H3K79me3 5hr</td>
<td>2</td>
<td>3</td>
<td>2,013,979</td>
<td>0.6</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K79me3 7hr</td>
<td>2</td>
<td>3</td>
<td>4,857,984</td>
<td>0.4</td>
<td>0.1</td>
</tr>
<tr>
<td>H3K79me3 24hr</td>
<td>2</td>
<td>3</td>
<td>4,420,846</td>
<td>0.4</td>
<td>0</td>
</tr>
</tbody>
</table>
Figure 2-1. Assessment of progress through sporulation.

(A) Samples were taken at indicated times and stained with DAPI (4’-6’ diamindino-2-phenylindole) by standard methods (2). Average percentage (n=3) of the cells completing meiosis I (binucleates + tri/tetranucleates, red), meiosis II (tri/tetranucleates, green) and spore maturation (asci, black) are shown. (B) The kinetics of DNA replication monitored by FACS analysis. The DNA content of the cells before replication and after replication is indicated as 2C and 4C, respectively.
Figure 2-2. Nucleosome organization around genes throughout sporulation.

(A) Composite nucleosome distribution traces for six time points in the meiosis/sporulation program are color-coded as indicated. Nucleosome midpoint density is represented by sequencing tag counts from crosslinked, MNase-digested, H3 immunoprecipitated, and gel-purified samples. Distributions were normalized such that the total tag count in each sample was equal. Tag counts are distributed about 6,676 transcript start (TSS) and end (TES) sites in 3 bp bins, and smoothed via a three-bin moving average. The percentage of regions analyzed is indicated by the black trace and covers a minimum of ±300 bp from the TSS or TES and a maximum of 300 bp from the next TSS or TES. Bin counts were normalized to the number of regions represent in each bin.

(B) Nucleosome fuzziness was taken to be the standard deviation of tag locations for each nucleosome. The average fuzziness per bin was determined, then plotted as described in panel A. The right panel shows a frequency distribution of nucleosome fuzziness at various meiotic time points. Color codes are as in panel A, except that YPD is represented by a black trace.

(C) Frequency distribution of nucleosomal widths (distance between the W/+ and C/- MNase digested borders).
Figure 2-3. Isolation of nucleosomal core particle.

A, Ethidium bromide stained agarose gel of MNase digested chromatin samples, from different meiotic stages. Each lane represents ~3.3 ml of cells. B, Ethidium bromide stained agarose gel of MNase-digested chromatin samples from YPD, 5 hr, and 24 hr sporulation induction. “No MNase” samples indicate that endogenous nucleases (which are likely to be inactivated by the formaldehyde) are not degrading the DNA to create artificial digestion patterns. “Supernates” are used for sequencing. “Pellets” demonstrate that very little chromatin remains unextracted. C, Distribution of read counts per nucleosome for the color-coded time points. The congruence of all distributions and the lack of leftward skewing of the 24 hr time point (purple), compared to all others, indicates that no subpopulation of nucleosomes were left behind in the residual chromatin pellets shown in panel B.
Since an H3 amino-terminal tail endopeptidase has been reported present in nuclear extracts prepared from stationary phase and sporulating cells (Santos-Rosa et al. 2009), we employed H3 antibodies directed against the C-terminal region of H3 in all experiments. However, we did not detect cleavage of H3 in vivo at any time (Figure 2-4).

**Western blot for H3**

![Western blot images](image-url)

Figure 2-4. Western blot to examine H3 tail cleavage during log-phase growth and sporulation. For each sporulation time point, the indicated OD of cells were electrophoresed through a 12% SDS-PAGE gel and probed for H3 by immunoblotting. Previous detection of H3 tail cleavage involved preparation and detection in native cell extracts (Santos-Rosa et al. 2009).

Sequencing tags from H3-immunopurified nucleosomes were plotted as a frequency distribution around all annotated transcription start sites (TSS), as a composite plot (Figure 2-2A) and as a cluster plot of individual genes. Tag counts were normalized to eliminate the bulk increase in histone occupancy levels that accompanied DNA replication at the start of meiosis. 6,576 transcription units were analyzed, covering virtually the entire yeast genome. The
resolution of the maps was sufficiently high to identify individual nucleosomes and linkers, and changes in their positions, occupancy levels, and modification states.

Strikingly, at no stage, from exponential growth in rich media to stationary phase in starvation media, through meiosis and sporulation, did the bulk of the NFRs acquire nucleosomes, as might be expected if any higher-order genome compaction was manifested at the primary structure of chromatin. This was evident in composite plots of all genes (Figure 2A) as well in cluster plots of individual genes (Figure 2-5A), and stands in contrast to a recent low resolution study purporting loss of 3’ NFRs upon carbon starvation (Fan et al. 2010). We cannot, however, rule out the possibility of partial or remodeled nucleosome assembly, which would not produce ~150 bp MNase-resistant fragments.

There was no widespread re-positioning of nucleosomal arrays near the TSS when compared to the pattern obtained from cells grown in rich media. Nucleosome fuzziness (standard deviation of tag distances from a consensus nucleosome midpoint) and width (distance between MNase-defined borders) were essentially unaltered from the pattern in rich media (Figure 2-2B, C). This was surprising because nucleosome fuzziness that increases towards the 3’ ends of genes might be expected to reflect translational fluidity of the nucleosomes on DNA, and that higher-order chromatin compaction would be expected to constrain such fluidity. However, such reduction in fuzziness was not observed.
Figure 2-5. Cluster view of bulk nucleosome and histone modifications during log-phase growth and sporulation.

Cluster view showing five cluster patterns within each panel of nucleosome organizations in vivo for mitotic growth in YPD, and five sporulation time points (each represented by a separate panel). 4792 genes (rows) were aligned by their TSS (David et al. 2006). Rows were then clustered by K means (K=5). The time point samples are color-coded. The intensity of each color reflects the nucleosome occupancy level, with white (in YPD) or black (all others) represent low nucleosome occupancy. (B-E) Cluster view showing five cluster patterns (arranged according to panel A) for each modification state: H3K9ac, H3K4me3, H3K36me3 and H3K79me3. The total tag count for each sample was normalized to reflect histone modification level obtained by western blot analysis.

We next focused on specific classes of genomic features that undergo significant changes in activity when cells go from a mitotically active to a quiescent state and then on through meiosis and sporulation. Hotspots and coldspots of double-strand breaks (DSB) in meiotic recombination have been identified on a genomic scale (Gerton et al. 2000b; Blitzblau et al. 2007; Buhler et al. 2007; Borde et al. 2009). Hotspots were constitutively depleted of nucleosome, and coldspots had constitutively high occupancy before, during, and after meiosis (Figure2-6A). This has been observed before at the ARG4, CYS3 locus (Ohta et al. 1994; Wu and Lichten 1994) and globally by FAIRE analysis (Berchowitz et al. 2009), but not previously documented on a
genomic scale with high resolution nucleosome maps. Thus, recombinational hotspots have a chromatin structure that is fixed and open at all time points tested in this study.

Figure 2-6. Nucleosome organization around genomic features throughout sporulation.
Nucleosomal midpoint tags were distributed around the indicated number of genomic features as described in Figure 2A, except bins were 15 bp.

DNA replication origins also displayed no change in nucleosome occupancy or positioning throughout the meiotic timecourse (Figure 2B). The RNA polymerase (Pol) III-transcribed tRNA and 5S rRNA genes, the Pol I transcribed rDNA locus (RDN37-1), and the 131 Pol II-transcribed ribosomal protein (RP) genes, which are all highly transcribed in rich media and down-regulated during stationary phase and meiosis (Chu et al. 1998), remained constitutively nucleosome-free in their promoter regions (Figure 2C, D, E). In most cases, the NFR was large enough to accommodate two nucleosomes (>300 bp), but filling-in was not observed.

Nevertheless, transcription-coupled nucleosome depletion within genes transcribed in rich media was returned to high occupancy states in accord with transcriptional shut-down. Reciprocally, meiotically-induced transcription units saw a decrease in nucleosome occupancy during the appropriate induction stage (Figure 2F). In many cases nucleosome depletion was transient (e.g., SPO74 and SPO75), while the associated mRNA was stable.

2.3.2 Selective re-positioning of the +1 nucleosome over the TSS is linked to gene regulation

Although global nucleosome positions did not noticeably change for the features or groups of features examined, genomic studies do suggest that factors cause nucleosome repositioning in vivo, and in some cases away from their intrinsically favored sites (Whitehouse et al. 2007; Schones et al. 2008; Shivashwamy et al. 2008; Kaplan et al. 2009; Tirosh et al. 2010). Recent studies suggest that the position of the +1 nucleosome may be cell cycle regulated and such shifts have consequences on gene expression (Kelly et al. 2010). We therefore sought out genes that had a shifted +1 nucleosome, as this position has been demonstrated to be moveable.
In Figure 2-7A, we calculated the shift of the +1 nucleosome (rows) at each time point (columns) in the sporulation program, relative to its position in YPD. One group shifted downstream in the direction of transcription (the magnitude of which is reflected by the intensity of yellow), and the other group shifted upstream (blue) during the stationary phase prior to entering meiosis, and remained shifted throughout meiosis and sporulation. Equivalent shifts were propagated along the genic nucleosomal arrays.

The group of genes in which the +1 nucleosome was shifted downstream in stationary phase were enriched with genes that have elevated expression during carbon starvation conditions (Figure 2-7B, based upon previous expression data (Bradley et al. 2009); chi-squared test P value=10^{-37}). In carbon-rich repressive media (YPD), their +1 nucleosome was positioned ~55 bp upstream of the canonical location. Thus the TSS is buried at the nucleosome dyad, where it is least accessible and thus likely to be most unable to initiate transcription. When starved for carbon, the nucleosome shifted to the canonical location. Consistent with transcriptional activation, the nucleosomes displayed reduced occupancy.

In contrast, the group of genes in which the +1 nucleosome was shifted upstream were enriched with genes that become repressed during carbon starvation (chi-squared test P value=10^{-7}). In stationary phase, the +1 nucleosomes of these genes shifted ~55 bp upstream of the canonical position to where the nucleosome dyad sits over the TSS. It and its associated nucleosomal array also increased in occupancy, consistent with lower expression levels. Re-inspection of the RP genes (Figure 2-6E), which are also down-regulated in stationary phase, showed an upstream shift in nucleosome positions as well during stationary phase.

The yeast cells experience carbon starvation and nitrogen starvation at 0 hr. Consistent with the regulation on carbon starvation genes, significant portion of downstream/upstream shifting genes overlap significantly with nitrogen starvation induced/repressed genes (Bradley et al. 2009). Taken together, these observations suggest that cellular mechanisms for gene control
involve repositioning nucleosomes over the TSS to make them more refractory to transcription initiation. This conclusion is consistent with reports that Isw2p may play a repressive role by repositioning genic nucleosomes towards the NFR (Whitehouse et al. 2007).

Conceivably, the +1 nucleosome positions in one medium vs. another may be DNA-encoded (Kaplan et al. 2009). To explore this possibility, we computed a DNA-encoded occupancy level for the two classes of nucleosome-shifted genes, using the algorithm of Kaplan et al (Kaplan et al. 2009). For the class of genes encountering a downstream shift in nucleosome positions in stationary phase, the algorithm predicted a depleted NFR at a position relatively downstream (Figure 2-8). For the class of genes experiencing an upstream nucleosome shift, a depleted NFR was predicted at a position relatively upstream. Thus, to the extent that nucleosome organization can be predicted by this algorithm, it suggests that nucleosome organization around the NFR for those genes that are differentially regulated during stationary phase is DNA-encoded during stationary phase. This results in an energy-requiring organizational state under energy abundant conditions, and an alternative organizational state at an energetic minimum when energy is limiting.
Figure 2-7. +1 nucleosome shift is linked with carbon starvation regulation.

(A) +1 nucleosome midpoint distances from the position found in YPD were calculated for each gene (rows) at all time points (columns), and displayed as a cluster plot. Three distinct groups were identified by K-mean clustering. Distances shifted downstream in a positive direction are color-coded yellow. Upstream shifts are colored blue. Equivalent shifts were evident across the genic nucleosomal arrays.

(B) Venn diagram and chi-test are shown for the overlap of the indicated cluster of genes and those genes most up- or down-regulated (as indicated) upon carbon starvation (Bradley et al. 2009). Composite nucleosome distributions around the TSS are shown for the set of intersecting genes.
2.3.3 Transcription-linked histone modification marks exist at most genes and are selective for nucleosome positions

We next focused on transcription-linked acetylation and methylation states (Bernstein et al. 2002; Santos-Rosa et al. 2002; Ng et al. 2003; Kurdistani et al. 2004; Liu et al. 2005; Pokholok et al. 2005). Genome-wide fates of these modifications throughout a developmental program and its comprehensive localization to specific high-resolution nucleosome positions has not been previously determined. As expected, total chromosomal histone content increased and leveled off as cells progressed from stationary phase through DNA replication, meiosis, and onto spore formation (Figure 2-9A). Bulk histone modification densities (H3K9ac, H3K4me3, H3K36me3, and H3K79me3), which were normalized to bulk histone content, decreased over this time (Figure 2-9B-E). The rates of bulk loss of the methyl marks were consistent with a
general loss of de novo methylation and a simple dilution of existing modification states as new unmodified histones are incorporated during DNA replication. Modification losses were not due to endopeptidase cleavage of the H3 amino terminal tail (Santos-Rosa et al. 2009), as such cleavage was not detected (Figure 2-4).

We found H3K9ac predominantly at the +1 nucleosome (Figure 2-10A, top row of panels, and Figure 2-5B and Figure 2-12), with a graded decrease into the body of the gene. Its predominance at +1 has not been previously reported. As expected, more frequently transcribed genes had higher H3K9ac densities (defined as position-specific H3K9ac/H3 ratios) at the +1 position. However, the differential was rather small (Figure 2-11, H3K9ac vs. H3 panel), which was not due to high background in the ChIP assay or due to high levels of nonspecific acetylation, because even lowly transcribed genes had high levels of acetylation enriched specifically at +1. Approximately 77% of all genes were enriched with H3K9ac at the +1 nucleosome, compared to 36% expected by chance.
Histone H3 immunoblot band signal intensity was quantified using ImageJ. Median ratios of the biological replicates (n ≥ 6) for YPD media or different time points in sporulation media (SPM) are presented in the bar-graph, and normalized on a per cell basis (OD600). (B–E), Histone modification densities were obtained by dividing the total histone modification signal by the histone H3 signal at each time point. Shown is the median ratio from (n ≥ 6) biological replicates for the indicated time points (detailed procedures and designs are provided in the Methods). Standard deviations are shown.
Figure 2-10. Nuclsoeomal and histone modifications distribution around genomic features.

(A) Distributions of indicated H3 modification states (rows of panels) are plotted around the TSS for all genes, middle meiotic genes, and ribosomal protein genes (columns of panels). Traces reflecting different time points in the sporulation program are color-coded as indicated, and
further described in Figure 2-2A. Total tag counts in each sample were scaled to reflect the bulk distribution, measured by immunoblotting (Figure 2-9). Consequently, the plots reflect the level of modified nucleosomes, not the density of modification per H3 nucleosome (see diagram). The H3K79me3 antibody may have significant cross-reactivity with me2. Cluster plots for H3K4me3 (B) and H3K36me3 (D) changes on a gene-by-gene basis show both H3K4me3 and H3K36me3 are linked with transcription. In contrast to panel A, the total tag count for all samples (H3, H3K4me3 and H3K36me3) were normalized to be equal. Each row included gene-centered log2 transformed H3K4me3/H3 ratios (from 0-500 bp region relative to TSS for H3K4me3 and 300-1000 bp region relative to TSS for H3K36me3) for all six time points. Gene-centering means that the average of each row is set to zero. All genes were arranged by K mean clustering (K=5). The number of genes in each cluster is indicated. One cluster was omitted as it displayed no changes. Corresponding changes in mRNA levels (Primig et al. 2000; Williams et al. 2002) are shown. Levels of mRNA were gene-centered and log2-transformed. (C) Chi-tests between clusters of H3K4me3 and H3K36me3 suggest that two methylation marks are significantly coincident. Venn diagram relating the overlap of clusters 1-4 in panel B with the corresponding clusters in panel D. Values below the Venn diagram reflect log10 P-values (Chi-test) of the overlapping membership.

Nucleosomes at positions +1, +2, and +3 appeared to be the primary sites of H3K4 trimethylation (Figure 2-10A and Figure 2-12, H3K4me3 row of graphs), in accord with other studies (Liu et al. 2005; Barski et al. 2007). Unlike acetylation, the H3K4me3 levels at these positions were not diminished at the start of meiosis (red trace compared to gray fill and Figure2-10A H3K4me3 row), but like acetylation, the mark receded uniformly at each position and globally at each gene as the sporulation program unfolded (Figure 2-5C).

Higher H3K4me3 densities were linked to higher transcription frequencies in YPD, and gene induction during meiosis (Figure 2-10B and Figure 2-11, H3K4me3 vs. H3 panel), in accord with other studies (Liu et al. 2005; Barski et al. 2007). Remarkably, 74% of all genes had higher H3K4me3 density at nucleosomes +1,2,3 relative to nucleosomes in the middle of the same genes (Figure 2-5C).
Figure 2-11. Distribution of histone and histone modifications states according to transcription frequency.

Distributions of H3 and modification levels around the TSS for four color-coded subdivisions of genes classified according to transcription frequency in YPD media (Holstege et al. 1998). Histone modification densities can be evaluated by comparing modification plots to the corresponding H3 plot.
Figure 2-12. Distribution of histone and histone modifications states around meiotic genes during meiosis and sporulation.

(A) Expression time course for meiotic stage-specific gene groups during meiosis were replotted from Chu et al. (Chu et al. 1998).
(B) Distributions of the indicated H3 modification states (rows of panels) are plotted around the TSS for all genes and meiotic stage-specific genes (columns of panels). Traces reflecting different time points in the sporulation program are color-coded as indicated, and further described in Figure 2-2A. Total tag counts in a dataset were adjusted to reflect the corresponding bulk histone modification density determined by immunoblotting (Figure 2-9). The frames are also color-coded according to the group of meiotic genes from panel A being plotted.

When the H3K4me3-enriched genes were examined for gene expression dependency on this mark, they displayed lower relative expression levels in set1Δ or H3K4A mutants, compared to those genes that normally lack H3K4me3 (using expression data from (Venkatasubrahmanyam et al. 2007)). Set1 is the H3K4 methyltransferase. This suggests that the H3K4me3 co-transcriptional mark may function positively at most genes if we assume that those genes that lack the mark are not regulated by the mark.

The H3K36me3 mark is co-transcriptionally placed in the body of the gene(Krogan et al. 2003b). As with H3K4me3, the H3K36me3 pattern was enriched at most genes (Figure 2-5D), with substantially greater levels at highly expressed genes (Figure 2-11, H3K36me3 panel). Normalized changes in H3K36me3 were also linked to meiotic gene expression (Figure 2-10D). As expected of two co-transcriptional marks, changes in H3K4 and H3K36 tri-methylation were highly coincident (Figure 2-10C). The H3K36me3 mark was depleted at the +1 and +2 positions, but hyper-enriched at the +3 to +6 position in stationary phase and up to 2 hr into meiosis, which is just prior to DNA replication (Figure 2-10A and Figure 2-12, H3K36me3 row).

The transcription-linked H2BK123Ub mark is needed to generate H3K79 methylation (Shilatifard 2006). We found that H3K79me2,3 was essentially enriched at all genic canonical nucleosome positions in roughly equal levels (Figure 2-10A and Figure 2-5E, Figure 2-11, H3K79me3 panel), and generally was not correlated with expression levels in rich media or
during meiosis. As sporulation progressed, each nucleosome position lost the H3K79me3 mark monotonically as with the other methyl marks, and in accord with dilution by genome replication.

2.3.4 Histone marks around recombinational hotspots and promoters are similar

We also examined the distribution of histone marks around genome-wide locations of DSB recombination hotspots and coldspots (Figure 2-13), and found patterns that were consistent with what was observed at promoter regions. Indeed DSB hotspots were enriched in promoter (and termination) regions (Figure 2-14), which is consistent with previous suggestions of enrichment in intergenic regions or at a few selected promoters (Wu and Lichten 1994; Gerton et al. 2000b; Cromie et al. 2007), but has not before been attributed to promoter NFRs. Not all DSB hotspots are localized in the promoter regions, but the DSB hotspots outside of promoter regions also exhibit the similar patterns in nucleosome and histone modifications. This suggests that the patterns in the DSB hotspots and coldspots could be established by a transcription independent manner.

H3K9ac was concentrated closest to hotspots, followed nearby by H3K4me3. H3K36me3 and H3K79me2,3 were depleted near the hotspots, and enriched at more distal regions, including at recombinational coldspots. As previously shown with H3K4me3 (Borde et al. 2009), histone marks pre-existed around DSB hotspots well before meiotic DSBs are generated, and exist well after they are resolved. Thus, the chromatin surrounding DSBs have a clear organization of histone marks, and these marks appear to be a rather stable feature of the genome that demarcate where DSBs are to occur.
recombination hot and cold spots.

The distribution of nucleosomal tags and levels of the indicated modification states around double strand break (DSB) hotspots and cold spots (Borde et al. 2009) during the sporulation program are displayed in the first two columns of graphs. To assess nucleosome densities plots should be compared against Figure 2-6A. The third column displays tag distributions around the TSS as shown in Figure 2-2A, for comparison, but is highly smoothed to achieve the lower resolution of the DSB sites.
Figure 2-14. Distribution of nucleosome and histone modifications states around meiotic recombination hot and cold spots.

Composite distributions of DSB hotspot (gray fill) and coldspot (red trace) sites are plotted around the TSS and TES for all genes (columns of panels) (Borde et al. 2009).
2.3.5 Antisense repression of meiotic gene

Antisense transcription is emerging as an important means by which sense transcription is repressed (Camblong et al. 2007; Granovskaia et al. 2010). For instance, $IME4$, a proposed master regulator of the entry into meiosis, is repressed in mitotically growing cells via antisense transcription (Hongay et al. 2006). However, it is unclear whether repression of gene expression by antisense transcription is a wide-spread form of regulation for large classes of co-expressed genes. We therefore examined the extent to which meiotically-induced genes are repressed in mitotically growing cells by antisense transcription.

As an indicator of antisense transcription, we sought genes having more H3K4me3 in the 3’ versus 5’ region during mitotic growth. Approximately 10% of all genes (641) had higher 3’ H3K4 methylation (5’/3’ log2 ratio < 0). Analysis of existing transcriptome data confirms the presence of antisense transcription at these genes (Figure 2-15).

When a set of 575 meiotically-expressed genes were examined, more than 20% had greater 3’ H3K4me3 (lower left Venn diagram, $P = 10^{-30}$), with more than half of these having more antisense transcripts than sense transcripts in YPD identified in one RNA-seq study (Venn diagram in Figure 2-15). During the course of meiosis nearly all of these genes transitioned from having an enrichment of H3K4 methylation at the 3’ end to having K4 methylation enrichment at the 5’ end (cluster plot in Figure 2-15). Thus, we conclude that a large fraction of the yeast meiotic program is repressed by antisense transcription taking place during mitotic growth.
Figure 2-15. H3K4me3 provides a signature for repressive antisense transcription. The ratio of H3K4me3 density at the 5’ end to the 3’ end of every gene was calculated. Genes having log2 ratios that were positive at any time during the sporulation program were selected and K means (K=3) clustered (n=124). The cluster plot is shown on the left and is turned 90˚ from the normal orientation. Genes are columns, rows are time points, and blue/black/yellow color scale reflects the log2 5’/3’ ratio. The upper Venn diagram shows the overlap between these genes and those that produce greater antisense than sense transcription (Parkhomchuk et al. 2009). The lower Venn diagram shows the overlap with sporulation induced genes (Chu et al. 1998). Chi-test P-values are shown.

2.3.6 Enrichment of H3K36me3 levels in transcriptionally inactive rDNA region

DNA recombination is highly repressed in repetitive sequences in order to prevent non-allelic homologous recombination. The non-allelic homologous recombination is very likely to result in the loss or the increase of gene copy number within these regions. Considering the higher frequency of meiotic recombination, the repetitive sequence in yeast genome is at higher risk for the induction of genome instability (Sasaki et al. 2010). The number of crossover per chromosome (2 to 11) in yeast is much higher compared to that (1 to 3) in plant and mammals (Qi
et al. 2009; Lu et al. 2011). The most repetitive region in yeast genome is the rDNA sequence on chromosome 12. It contains 100-200 copies of tandem repeats of 9 kb DNA fragment. Consistent with the repression of recombination on DNA repeat sequence in the genome, the frequency of meiotic recombination in rDNA is ~100 folds lower compared to that of the genome average (Petes 1979). It is still not known that how meiotic recombination of rDNA region is repressed.

Previous study has shown that H3 K36me3 mark can recruit HDAC (Rpd3p) and repress initiation of cryptic transcription (Li et al. 2007b). Similarly, H3 K36me3 may facilitate the formation of stable nucleosome through HDAC in order to suppress recombination. Consistent with H3 K36me3’s role in suppressing recombination, the deletion of the histone methyltransferase SET2 for H3 K36 mark or histone deacetylase RPD3 greatly elevates meiotic recombination frequency in one classical meiotic hotspot (Merker et al. 2008). I also have observed the global depletion of H3 K36me3 mark in the DSB hotspots and slight enrichment of H3 K36me3 mark in the DSB coldspots (Figure 2-12). Given the evidence above, I suspect that H3 K36me3 is able to repress recombination.

Interestingly, rDNA in yeast genome shows drastic increase in the level of H3 K36me3 mark despite that the global H3K36me3 level decreases during sporulation time course (cyan/purple traces versus grey fill in H3 K36me3 panel of Figure 2-16). Such enrichment is very specific to rDNA, especially considering that even the H3 K36me3 levels at DSB coldspots decrease along with global H3 K36me3 level in yeast cells (Figure 2-12). It is possible that specific enrichment of H3 K36me3 during meiotic cell cycle is essential for strong repression of meiotic recombination in rDNA region.

One concern is that such increase of H3 K36me3 at rDNA region during later stages of sporulation is due to the decreased acetylation at H3 K36 upon transcription shutdown (Psathas et al. 2009). However, if the increase of H3 K36me3 at rDNA region were just due to decreased acetylation at H3 K36, mRNA genes should show similar level of increase in H3 K36me3.
However, such increase was not seen in the mRNA genes. On the contrary, almost all regions in the genome except rDNA region showed decreased H3 K36me3 level (Figure 2-10 and Figure 2-11). Given the evidence above, it is very unlikely that the decreased acetylation of H3 K36 at rDNA region leads to the specific enrichment of H3 K36me3 during later stages of sporulation.

rDNA transcription level is very low during sporulation, indicating that the enrichment of H3 K36me3 does not rely on active transcription (Chu et al. 1998). Consistently, specific enrichment of tri-methylation at rDNA is not observed at the other two methylation marks-H3K4 methylation and H3K79 methylation (Figure 2-16). Taken together, we propose that elevated levels of H3 K36me3 at later stages of yeast meiosis may be laid down by transcription-independent manner and facilitate meiotic recombination repression at the rDNA region.
Figure 2-16. Modification states around rDNA locus.

Distributions of H3 and the indicated H3 modification states (rows of panels) are plotted around the TSS for rDNA region. Traces reflecting different time points in the sporulation program are color-coded as indicated, and further described in Figure 2-2A. Total tag counts for H3 between different time points samples were adjusted to be equal, while total tag counts for modification marks between different time points samples were adjusted to reflect the bulk density ratio obtained by western blot (Figure 2-9).
2.3.7 Global increase of linker DNA length at stationary stage (0 hr)

Nucleosome spacing remained constant at 165 bp, with the exception of stationary phase (0 hr, red trace in Figure 2-2A). In this case, genic nucleosomes increased their spacing without changing the length of nucleosomal DNA (Figure 2-2C). This shift was reproducible as it was seen in individual replicates, as well as in the histone modification state maps (Figure 2-9A). The origins and functional significance of this stationary phase-specific shift predominating in the middle of genes is unclear. There are multiple conditions at stationary phase (0 hr) that may contribute to such global increase of linker DNA length. We tried to accurately characterize the minimal conditions that can trigger the global change by eliminating one of the four factors (OD 1.90 versus lower OD, YPA media versus YPD media, ~16 hr overnight incubation versus no incubation, ice incubation versus room temperature incubation) at a time. The global increase of linker DNA length was no longer observed when any one of them was eliminated (Figure 2-17 A-D). This suggests that the induction of such global alteration in linker DNA length is a complex process and require multiple environmental conditions.

ATP-dependent chromatin remodeling enzymes such as Isw1p have been shown to operate in the coding regions (Tirosh et al. 2010), and thus might contribute to altered spacing at 0 hr. To validate this assumption, I performed similar nucleosomes mapping in ISW1 null strain. Surprisingly, 0 hr still showed the increased nucleosomal spacing compared to YPD sample (Figure 2-17E compare black to red trace), indicating that Isw1p is not essential for this global increase in linker DNA length. The nucleosomal mapping for heat shock response was also plotting in Figure 2-17F, showing that such increase in nucleosomal spacing is specific to the 0 hr condition.
Figure 2-17. Multiple factors set up the increase of nucleosomal spacing at stationary phase.

The composite plots were generated similarly as Figure 2-2A, except that the bin size is 15 bp. (A-D) each of four factors were eliminated respectively. The original condition 0 hr (black trace) and the condition in that one factor is eliminated (red trace) were plotted for wild type cells. (E) The composite plots for *ISW1* mutant between YPD and 0 hr. (F) The composite plots for *ISW1* mutant between mock heat shock and heat shock in YPD.
2.4 Discussion

2.4.1 Chromatin organization is a constant, but with locus-specific plasticity

The constancy of the overall primary structure of chromatin as it proceeds through a full gamut of transactions that any chromosome would be expected to experience during its life time including DNA replication, recombination, transcription, chromosome segregation, and compaction is remarkable in that many of these transactions would be expected to alter chromatin structure. These in vivo findings contrast with one expectation from in vitro experiments that a ~165 bp nucleosome repeat length must increase by ~30 bp in order to undergo compaction (Routh et al. 2008). Indeed simple a priori expectations are that NFRs would acquire nucleosomes during compaction and become more closely spaced; nucleosomes that are not well positioned might have become more organized, but these outcomes were not observed.

The concept that these findings demonstrate is that during any transaction including compaction and packaging, the genome maintains its accessibility to cellular factors. In particular, an accessible state readies the cell for gene expression upon spore germination. Should NFRs be filled in, and linkers collapse, then the germinating spore would need to have evolved mechanisms that re-establish the functional state of chromatin. Particularly challenging would be specifying the removal of nucleosomes that occupy NFRs, and in having chromatin remodeling complexes gaining access to linker regions to re-establish proper spacing. Consistent with the importance of maintaining chromatin structure, histones are retained at developmentally important genes in mammalian gametogenesis, whereas protamines replace histones at most other locations (Hammoud et al. 2009). Thus, in yeast, higher-order chromosome compaction appears to occur without perturbation to the primary physical state of chromatin (positioning and occupancy), indicating that normal chromatin structure is compatible with a series of yeast
meiotic chromosomal events, including compaction. Such maintenance of primary chromatin may be applicable to the condensed chromatin states occurring in multi-cellular eukaryotes.

In fact, no significant change in fuzziness among samples in different stages of sporulation can be interpreted in two ways (Figure 2-2B). One possibility is that there is only one optimal nucleosome position, but nucleosome maintains similar level of fluidity during the whole process of sporulation, even with chromatin compaction. The other interpretation is that in a population of cells, nucleosomes may resides at various positions (within a certain distance relative to the most frequent positions in the whole population) in different cells, but once cells entering meiotic cell cycle, nucleosome positions are maintained and fixed due to high order chromatin compaction. The second assumption is more likely, but further experiments are needed to validate such assumption.

Specific locations related to gene expression do undergo local re-organization including decreased nucleosome occupancy levels in association with the passage of RNA polymerase during transcription, and +1 nucleosome repositioning which may regulate TSS access as seen during the cell cycle (Kelly et al. 2010). Occupancy of the NFR by nucleosomes may not appear to be a general mechanism to repress gene expression, and instead may be a mechanism that is restricted to certain types of genes. For example, TATA-containing genes, which constitute a small fraction of the genome, tend to have nucleosomes that encroach on the NFR and thus may be repressed by such a mechanism. In contrast nucleosome occupancy changes do occur within transcribed regions that is evident at high levels of transcription, and this may be related to transient displacement by the elongating transcription machinery.

The concept that nucleosome occupancy of NFRs does not represent a widespread mechanism of repression is also applicable to RNA polymerase III transcription units and recombinational hotspots in general, wherein these region remain constitutively nucleosome free.
Thus, factor recruitment rather than steric exclusion may be a key regulatory step that is common among these processes.

2.4.2 Regulation of transcription initiation by +1 nucleosome repositioning vs antisense

In *Saccharomyces* the position of the +1 nucleosome has largely been thought of as being fixed at a canonical distance from the TSS, which places the TSS on the edge of the nucleosome. This presumably is a suitable spot for transcription initiation for reasons that remain to be worked out. While a positionally displaced nucleosomes could occur either further upstream or downstream from its canonical location, we observed only placement were the TSS was buried near the nucleosome dyad, under conditions where the gene is repressed. This represents a clear example where nucleosome repositioning acts concertedly on a group of co-expressed and reciprocally-expressed genes (carbon-response genes), possibly revealing a general mechanism for transcriptional regulation that may also be applicable to other classes of genes.

Conceivably, under energy-rich growth conditions, nucleosomes may be actively placed in thermodynamically less favorable positions at carbon-responsive genes, a concept whose mechanism has been developed elsewhere (Whitehouse et al. 2007). Under comparatively lower energy states in carbon-depleted media, this energy consuming organization would not be maintained, thereby allowing those nucleosomes to move towards DNA-encoded positions (Figure 2-8). For carbon-utilization genes, the DNA-encoded positions correspond to the active state. For genes that are down-regulated in carbon-depleted media, the DNA-encoded positions correspond to the repressed state.

In contrast to the carbon starvation stress response, a similar nucleosomal shift was not evident at meiotically regulated genes at any time during meiosis. The temporal activation of meiotically regulated genes represents a timed developmental program, whereas response to
carbon starvation may represent more of an acute stress response. We speculate that nucleosome positioning control of the TSS may reflect a particular regulatory theme for acute/immediate transcriptional responses, whereas a program consisting of a series of responses may involve alternative types of regulation. Our analysis suggests that the latter could include negative regulation through antisense transcription, which is evident by a reversal of the 5′-3′ gradient of transcription-linked histone marks. As long-lived marks of transcription, such inverse gradients may be useful in identifying antisense transcription having unstable transcripts.

A large fraction of meiotically expressed genes are repressed in rich media, and this repression is associated with antisense transcription. Antisense transcription has been demonstrated to play gene-specific roles in repression (Hongay et al. 2006), but has not before been demonstrated to be broadly associated with a large set of co-regulated genes. Studies suggest that there may be multiple mechanistic role of antisense transcription in sense inhibition. The extent to which antisense contributes to repression of meiotic genes and whether the same mechanism is involved at all of these genes remains to be determined. I suspect that the purpose of antisense transcription at meiotic genes provides an added level of repression that already exists at the promoters of these genes. This redundancy may ensure that little or no basal meiotic gene expression occurs in rich media, which otherwise might be detrimental to the cell.

2.4.3 Transcription-linked nucleosome modification states mark nucleosome positions at most genes

The study presented here provides a comprehensive high-resolution genome-wide map of histone modification states that occur along a developmental pathway, and expands upon earlier lower resolution and lower coverage modification state maps in YPD (Liu et al. 2005; Pokholok et al. 2005). The concentration of H3K9ac at the +1 nucleosome, H3K4me3 at +1,2,3,
H3K36me3 at all genic positions except +1,2, and H3K79me3 at all genic locations, results in a specific marking system in which the first four nucleosomes downstream of the TSS may be distinguished from each other. This may be important for events surrounding transcription initiation and early elongation, where nucleosome-binding regulatory proteins selectively interact with nucleosomes in a position-specific manner (Koerber et al. 2009). However, there is no evidence that the widespread occurrence of marks as detected here is linked to an equivalent widespread binding of cognate factors, and so additional regulatory determinants are likely involved.

Most genes contain position-specific transcription-linked histone marks, but paradoxically most genes are infrequently transcribed (i.e., at a basal level). On the time frame of each transcription cycle, these marks may be quite long-lived. Thus, the marks may be constitutively present to a large extent, and any that are lost over time might be regenerated with each transcription event. In stationary phase and throughout the stages of sporulation, these marks dissipate (acetylation more rapidly than methylation), possibly owing to a substantial drop in basal transcription. The persistence of these marks at the active transcribed genes, particularly at +1, presents a conundrum in that the +1 nucleosomes of highly expressed genes are relatively dynamic (Dion et al. 2007; Rufiange et al. 2007). Conceivably, the same modified histone may be preferentially returned to the same site once it is evicted perhaps owing to a negligible local pool of free histones, or some mechanism must exist to restore pre-existing modification states.

2.4.4 H3K36me3 mark maintain repressive chromatin structure and repress recombination in rDNA region

It has been shown that H3 K36me3 is specifically enriched at rDNA region during sporulation. I suspect the enrichment of H3 K36me3 is correlated with heterochromatin
formation, thus greatly suppressing the meiotic recombination in rDNA region. In fact, there is additional indirect evidence to support such assumption.

In fact, H3 K36me3 also play an important role in the formation of heterochromatin in mammalian genome (Chantalat et al. 2011). Consistent with transcription-independent deposition of H3 K36me3 at rDNA region in yeast, the deposition of H3 K36me3 at heterochromatin in mouse genome is not correlated with transcription events neither.

The role of forming heterochromatin may also explain the more severe sporulation defects for H3K36A mutant strain. Sporulation induction was performed for the whole histone point mutation library, and H3K36A mutant shows severe reduction in sporulation frequency (~60%) versus the H3K4A (~90%) and H3K9A (~95%). However, it is unclear why H3K36A mutant strain has the most severe phenotype in sporulation frequency among these three strains (Govin et al. 2010a). This cannot be explained by defects in transcription, since there is no evidence to believe that H3K36A mutation will have greater defect on transcription compared to the other two. It is possible that chromatin structure of rDNA region is more open without H3K36me3-mediated heterochromatin formation. Consequently, aberrant meiotic recombination at rDNA region occurs at higher frequency among the whole population of diploid cells. And the cells with aberrant meiotic recombination may be arrested at meiotic combination checkpoint, leading to a significant lower frequency of sporulation in H3K36A cells.

In fact, such H3K36me3-mediated heterochromatin formation at rDNA may also apply to the inactive copies of rDNA during vegetative cells. Although global enrichment of H3 K36me3 at rDNA region is not reported during vegetative growth, previous studies also indicated the assumption that H3 K36me3 (or SET2) might facilitate the heterochromatin formation in vegetative cells. The deletion of histone methyltransferase SET1 for H3 K4 methylation did not increase rRNA expression, while the deletion of SET2 increased Polymerase I transcription (Hontz et al. 2009). This also suggested that the H3 K36 methylation may be involved in
sustaining repressive chromatin structure at rDNA region. Taken together, I suggest that H3K36me3 mark may facilitate heterochromatin formation at rDNA region in order to repress recombination and transcription, and such mechanism may be generalized to other model organisms.

2.4.5 Global increase in linker DNA length

The average length of linker DNA ranges from 5 bp to 73 bp among different organisms (Perisic et al. 2010). Even during a developmental program of the same organism, nucleosomal spacing can be altered significantly (Jaeger and Kuenzle 1982). In addition, the genome-wide mapping of nucleosomes in primary human cells revealed that regions in the genome with different transcription levels show up to 27 bp differences regarding the length of linker DNA. The length of linker DNA is negatively correlated with transcription frequency (Valouev et al. 2011). However, the global increase in linker DNA length at 0 hr is not solely due to the lower transcription level. Yeast spores (24 hr) have even lower level of transcription compared to cells at 0 hr and more compacted chromatin (Chu et al. 1998), but the linker DNA length at 24 hr is identical to that of vegetative cells in YPD media (Figure 2-2A). Thus it suggests that the length of linker DNA may be affected by transcription frequency, but other factors also play important roles.

In budding yeast, several chromatin remodelers regulate the nucleosomal spacing. The ISW1 deletion strain shows decreased linker DNA length in the coding region (Tirosh et al. 2010). Consistent with ISW1 complex’s role in regulating the linker DNA length, the deletion of IOC2 gene, another essential component of ISW1b complex, leads to the similar decrease in linker DNA length (Kuang’s unpublished data). In addition, such decrease in linker DNA length is observed in other chromatin remodeler deletion strains such as ARP5 deletion strain (INO80
complex) (Kuang’s unpublished data). This indicates that multiple chromatin remodelers are involved in the regulation of nucleosomal spacing and they may have redundant functions. The functional redundancy in chromatin remodelers may explain why the global increase in linker DNA length can still be observed in ISWI single deletion strain (Figure 2-17E).

Interestingly, linker histone does not normally affect the linker DNA length in yeast. Nucleosomal spacing does not change in the HHO1 null strain (Freidkin and Katcoff 2001). The increased nucleosomal spacing in 0 hr is not due to increased histone H1 in chromatin neither. As histone H1 level is ~5 folds higher in mature spores (24 hr) compared to 0 hr sample (data in chapter 3), but still no changes in linker DNA length is seen at 24 hr.

The functional impact of increased linker DNA length is not known. Previous modeling studies indicate that higher order nucleosome packing couples with alteration in linker DNA length (Wong et al. 2007). Therefore, the global increased linker DNA length is possibly due to the specific higher order nucleosome packing at 0 hr. In this case, I suspect that specific higher order chromatin conformation at 0 hr may serve as a “poised” state for yeast cells, thus repressing the global transcription and enabling cells to quickly respond to the changes of environmental clues in the same time. Interestingly, such global alteration is completely reversed within 2 hours (compare red with orange traces Figure 2-2A). However, it is not clear why multiple factors are needed to set up such poised state, and much work is needed to elucidate the mechanisms of such global alteration on nucleosomal spacing.
2.5 Materials and Methods

2.5.1 Cell Harvest.

All results were obtained from the DSY1089 diploid strain (Raithatha and Stuart 2005), which is an SK1 background \((MATa/\alpha ho::LYS2/ ho::LYS2 lys2/ lys2 leu2::hisG/ leu2::hisG trp1::hisG/ trp1::hisG ura3/ ura3 arg4-BglII/arg4-NspI his4B/his4X)\). The sporulation protocol is as described (Primig et al. 2000). All growth and sporulation procedures were carried out at 30°C. Colonies from YPG plates were incubated in 5 ml YPA media overnight, then inoculated into 500 ml YPA at 275 rpm, using one flask per time point. Cells were grown to a density of \(8.0\times10^7\) cells/ml (~OD 1.9), then 450 ml were harvested by centrifugation, washed with distilled water and stored at 4°C overnight. Cells were then resuspended in pre-warmed 450 ml SPM media and incubated at 30°C at 275 rpm for 0, 2, 5, 7, 9, 11 and 24 hours, then fixed with 1% formaldehyde for 15 minutes, then quenched with 0.125 M glycine for 5 minutes. Samples were taken to perform DAPI staining and flow cytometry at the same time as harvesting. The cells grown to log phase (OD=0.8-1.0) in YPD were processed similarly. Each nucleosome map was produced by averaging two independent biological replicates. Maps derived from each sample were indistinguishable from their biological replicate.

2.5.2 Monitoring of sporulation progression.

The sporulation efficiency was evaluated by FACS and DAPI staining. The cells for FACS analysis were fixed with 70% ethanol and stained with Propidium iodide at the Flow Cytometry Facility at Penn State University. The cells for DAPI staining were also fixed with
70% ethanol, resuspended in TE and sonicated for 30 seconds. 5 ul DAPI was added to the sample and cells counted at 100X zoom under a microscope.

2.5.3 Evaluation of mononucleosome extraction.

Since the cell walls of mature spores (24 hr samples) resist disruption, we added an additional sonication step before cell lysis. After chromatin preparation, the pellet was checked under the light microscope, no intact spores were observed.

NUC1 is a mitochondrial endonuclease that is released upon cellular disruption. Studies on DSB break repair and meiotic chromatin commonly use a nuc1 mutant so as to eliminate nonspecific chromosomal cleavage (de Massy et al. 1995; Buttner et al. 2007). However, the formaldehyde crosslinking that we employ in vivo typically inactivate proteins including NUC1. Consistent with this, chromatin pellets prior to MNase digestion revealed no degradation of the chromatin (Figure 2-3B), which would need to occur if NUC1 activity were creating the observed chromatin patterns.

2.5.4 Preparation of sequencing samples.

Cells were collected and disrupted by bead beating. Chromatin was prepared and washed with FA lysis buffer (as described (Albert et al. 2007)). Mono-nucleosomes were solubilized with 40 units of MNase in 600 ul of NP-S buffer (0.5 mM Spermidine, 0.075% IGEPAL, 50 mM NaCl, 10 mM Tris-Cl (pH = 7.5), 5 mM MgCl2, 1 mM CaCl2, 1 mM β-mercaptoethanol). The mononucleosomes were incubated with primary antibody overnight at 4°C, then immunoprecipitated with Protein A sepharose for 1.5 hours. Stringent washes were applied to
minimize non-specific binding (Albert et al. 2007). Samples were then prepared for DNA sequencing using Applied Biosystems SOLiD genome sequencer.

2.5.5 Antibodies.

The following antibodies were obtained from Abcam (cat. No.): H3 (ab1791), H3K4me3 (ab8580), H3K79me3 (ab2621, may have some cross reactivity with K79me2), H3K36me3 (ab9050), beta-actin (ab8224), H3K9ac (Millipore 07-352). Since H3K79me3 pattern looked like H3, we confirmed its specificity by western blotting with a H3K79A mutant.

2.5.6 Immunoblot normalization.

The cells used for immunoblotting were harvested and processed as described above, only without the formaldehyde cross-linking. This includes chromatin pellet washes to remove unbound histones, and thus produces a measure of bulk chromosomal histone levels. The same OD equivalents of cells were loaded on an SDS-PAGE gel to measure the relative bulk H3 level ratio between time points. To calculate the modified histone density ratio between different time point samples, the modified histones were normalized to the bulk H3 histone level. Actin was used as a loading control. The band intensity was quantified using ImageJ software. At least six replicates were included in the ratio calculation.

2.5.7 Data Analysis.

Sequencing tags were aligned to the genome using SHRiMP software package (Rumble et al. 2009). The coordinate of the 5’ end of each sequencing tag was shifted 73 bp towards the 3’
direction to reflect the location of the nucleosome midpoint (dyad). Tag counts located between ±4 kb of 6557 annotated TSSs were collected in 1 bp bins. The TSS/TES annotation was compiled from existing TSS/TES resources (David et al. 2006; Miura et al. 2006; Nagalakshmi et al. 2008; Xu et al. 2009). Each H3 dataset was scaled to set total counts to be equal. Similarly for histone modification states, each dataset was scaled such that the total tag counts reflected bulk modification densities, as shown in Figure 2-9. These binned distributions of modification states do not take into account local changes in nucleosome occupancy levels.

Nucleosome calls were made by GeneTrack software (Albert et al. 2008a). The fuzziness and width calculations are described elsewhere (Mavrich et al. 2008a). Processed nucleosome data including nucleosome midpoints, tag counts and fuzziness are available online (http://genome.cshlp.org/content/21/6/875/suppl/DC1).

The DSB hotspot coordinates were obtained from published studies (Borde et al. 2009), the 1013 sites with lowest DSB ratio were selected as the coldspot sites.

The statistical significance of nucleosome shifting in Figure 2-7A was evaluated by the Student’s t test. The W strand tags were shifted 73 bp downstream and the C strand tags were shifted 73 bp upstream. Then the shifted W and C strand tags were combined. Nucleosome calls were performed using Genetrack on shifted tags (sigma 10 and exclusion zone 147 bp). The sequence reads within 73 bp to the peak middle point were selected. Then standard deviation for each nucleosome was calculated on these selected tags. For two condition A and B, the total tag counts for each sample was denoted as \( n_A \) and \( n_B \), the standard deviation for each sample was denoted as \( s_A \) and \( s_B \), and nucleosome middle points were denoted as \( l_A \) and \( l_B \). The standard error (SE) was calculated by \( \frac{(s_A^2/n_A + s_B^2/n_B)^{0.5}}{\sqrt{n_A+n_B-2}} \) ts was calculated by \( (l_A - l_B)/SE \). The degree of freedom was \( n_A + n_B -2 \). The one-tailed p value was then obtained.

The statistical significance of the overlapping between two groups of genes was evaluated by the Chi-squared test. The two groups of genes were designated as A and B. The
The number of genes in each group was \( n_A \) and \( n_A \). The number of overlapping genes was \( n_{AB} \), the total number gene in yeast genome was denoted as \( n_{all} \) (I used 6000 in the analysis). The matrix for expected result was calculated (shown in Table 2-2). Then the \( P \) value of Chi-squared test for the overlapping was calculated by using CHITEST function in Microsoft Office Excel 2008 and chisquare function in Scipy package for Python respectively. \(<0.01 \) \( P \) value was considered to be statistically significant.

Table 2-2. Data matrix for chi-square Test.

<table>
<thead>
<tr>
<th></th>
<th>Real data</th>
<th></th>
<th>Expected result</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Not A</td>
<td>A</td>
<td>Not A</td>
</tr>
<tr>
<td>B</td>
<td>( n_{AB} )</td>
<td>( n_{B} - n_{AB} )</td>
<td>( n_A * n_B / n_{all} )</td>
</tr>
<tr>
<td>Not B</td>
<td>( n_A - n_{AB} )</td>
<td>( n_{all} - n_B - n_A + n_{AB} )</td>
<td>( (n_{all} - n_B)^* n_A / n_{all} )</td>
</tr>
</tbody>
</table>

To calculate a histone modification 5’/3’ ratio for an individual gene, each ORF was divided into equal length 5’ and 3’ ends. The H3 and H3 modification tags mapped to the 5’ and 3’ ends for each gene was counted, the ratio of 5’/3’ was calculated to be the following: \( (5’ \) histone modification tags / 5’ H3 tags) / (3’ histone modification tags / 3’ H3 tags).
Chapter 3

Genome-wide histone H1 mapping upon sporulation program

3.1 Summary

Similar nucleosome positioning and occupancy are observed between mature spores and vegetative cells of *Saccharomyces cerevisiae*, however, transcription is silenced in mature spores. Therefore, the high order chromatin compaction, instead of primary chromatin structure, may play an important role in genome-wide transcriptional repression. Histone H1 is involved in the formation of high order chromatin structure and gene regulation. Previous studies in *Saccharomyces cerevisiae* showed that linker histone (Hho1p) level peaked in mature spores and the nuclear volume of yeast spores was enlarged in the *HHO1* mutant strain, suggesting that Hho1p may facilitate high-order chromatin compaction in mature spores. ChIP-seq was applied to map the localization of Hho1p across the yeast genome during the onset (0 hr) and completion (48 hr) of sporulation. I found that Hho1p is enriched in the linker DNA region, and spreads across the genome. In addition, Hho1p localizes in the middle of 5’ NFRs but not 3’ NFRs. The positioning of linker histone is largely consistent between the start and the end of sporulation process. The majority of the genome shows similar level of increase in Hho1p occupancy after normalizing with protein level of *HHO1* gene, except the highly expressed genes in mature spores. Taken together, the ChIP-seq results suggest that the occupancy level of Hho1p increases in mature spores in order to form higher order chromatin compaction and to repress transcription, while the positioning of Hho1p is largely unaltered.
3.2 Introduction

Histone H1 (linker histone) binds to the linker DNA outside of the core nucleosome. *In vitro* study suggested that histone H1 is involved in packing chromatin into high order 30nm chromatin fiber (Li et al. 2010). The *Saccharomyces* genome encodes only one copy of histone H1 protein called Hho1p. Hho1p contains two globular domains, while canonical metazoan H1 consists of one globular domain flanked by the N and C terminal regions.

Although H1 is essential during development of mouse and *Drosophila*, the *HHO1* null mutant in yeast is still viable and exhibits minor defects (Hellauer et al. 2001; Fan et al. 2003; Lu et al. 2009). The function of Hho1p is similar to its homolog in the higher eukaryotes, which involves in the compaction of chromatin (Schafer et al. 2008; Georgieva et al. 2011b). The linear relationship between the Histone H1 content and Nucleosomal Repeat Length (NRL) has been observed in the mouse (Woodcock et al. 2006). Nevertheless, nucleosomal spacing and the sensitivity to MNase do not change in the *HHO1* deletion strain in yeast (Freidkin and Katcoff 2001).

Although *HHO1* is considered to repress transcription due to its role in stabilizing the chromatin, so far studies have shown that very few yeast genes are significantly repressed by Hho1p except for ribosomal DNA in vegetative cells (Levy et al. 2008). *HHO1* deletion strain shows decreased expression of specific mRNA genes (Hellauer et al. 2001). Yet, no differential binding of Hho1p at these down-regulated genes was found by ChIP-PCR assay (Freidkin and Katcoff 2001). However, I cannot rule out the possibility that the positioning pattern of Hho1p may be important for transcription regulation, and therefore, the high resolution ChIP-seq approach may provide more accurate information on the positioning of Hho1p, not only the occupancy level of Hho1p.
The mature yeast spores form thick and protective cell walls with two additional layers (chitosan and dityrosine layer), which enable yeast cells to survive various environmental stresses (Smits et al. 2001). In addition to a protective cell wall, chromatin is compacted and global transcription rate greatly decreases in mature spores (Jona et al. 2000). It is not known which protein facilitates the higher order chromatin compaction in yeast spores and whether the compacted chromatin represses global transcription in spores.

The protein level of HHO1 peaked in the mature spores and dropped precipitately after germination; in addition, chromatin was de-condensed in the spores of the HHO1 deletion strain (Shelly Berger’s lab unpublished data). Therefore, it is reasonable to assume that Hho1p may be involved in higher order chromatin compaction in mature spores. The HHO1 null mutant shows minor decrease in sporulation frequency (60% sporulation efficiency in hho1Δ/hho1Δ strain compared with 83% in wild type diploid cells); however, the role of Hho1p during yeast sporulation is unclear (Patterton et al. 1998; Deutschbauer et al. 2002).

To understand the whether the localization of Hho1p changes in mature spores and how the localization can be linked with chromatin structure and transcription regulation, I applied the genome-wide ChIP-seq approach to study Hho1p localization on the onset and the end of yeast sporulation cells. We intended to answer the following questions: 1) Where does Hho1p bind on the genome? 2) Is the occupancy level of Hho1p correlated with expression regulation and chromatin compaction? 3) What factors/proteins facilitate the recruitment of Hho1p into chromatin?
3.3 Results

3.3.1 Hho1p binds to the linker DNA region globally

The yeast strain with FLAG tag on HHO1 was used in the current study. To confirm the ChIP efficiency of Hho1p, we performed ChIP-PCR before preparing ChIP-seq sequencing library. We chose three Hho1p binding regions described in previous literatures, 25rDNA, LEA1 and SPB1 linker regions (Albert et al. 2007). For LEA1 and SPB1 linker regions, I observed significant enrichment of ChIP sample versus mock IP samples (compare 0hA, 48hA, 0hB, 48hB samples with No-TAG control C lane in Figure 3-1). Slight enrichment of ChIP versus mock IP samples was seen in the 25rDNA region. The ChIP-PCR results suggested that linker DNA is indeed enriched by the ChIP approach, therefore, I continued to the ChIP-seq sample preparation.
Figure 3-1. ChIP-PCR to evaluate the Hho1p ChIP efficiency.

Same OD equivalent of cells were loaded in the PCRs for all three regions. The 25rDNA was selected as Hho1p is enriched in rDNA region (Li et al. 2008a). The other two linker DNA regions were selected from previous published literature (Albert et al. 2007). “I” represents the Input sample, while “C” represents the ChIPed sample. NTC stands for None Template Control (just water), while No-Tag control is ChIPed PCR product from the strain without the FLAG tag on *HHO1* gene. Two biological replicates of 0 hr and 48 hr were denoted as A and B respectively.

The total tag count for each sample was presented in Table 3-1, above 20 million sequencing reads were obtained for each sample, therefore enough coverage was obtained for this study (>3X coverage). Although Hho1p is a linker histone, no one to our knowledge has shown that the genome-wide mapping of histone H1 is enriched at the linker DNA region. As shown in Figure 3-2, both in 0 hr and 48 hr, Hho1p peaks at the linker region—the valley between the nucleosomal peaks (nucleosomal organization is plotted as grey fill in the background from Chapter 2). To compare the 0 hr and 48 hr datasets, I scaled the data (multiplied by a constant) to reflect the ratio of *HHO1* protein levels present at 0 and 48 hrs, as measured by immunoblotting. The peak to valley ratio is not as pronounced as nucleosome data, which is largely due to the DNA fragmentation heterogeneity in sonication.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Total tag count</th>
<th>Paired tag count</th>
<th>Percentage of unique paired tag</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 hr replicate A</td>
<td>18,842,933</td>
<td>12,739,747</td>
<td>67.61%</td>
</tr>
<tr>
<td>0 hr replicate B</td>
<td>23,590,983</td>
<td>16,186,552</td>
<td>68.60%</td>
</tr>
<tr>
<td>48 hr replicate A</td>
<td>20,538,276</td>
<td>13,412,563</td>
<td>65.30%</td>
</tr>
<tr>
<td>48 hr replicate B</td>
<td>20,139,464</td>
<td>11,201,175</td>
<td>55.61%</td>
</tr>
<tr>
<td>0 hr mock IP control</td>
<td>15,486,983</td>
<td>8,562,703</td>
<td>55.29%</td>
</tr>
</tbody>
</table>
To address whether Hho1p binds to genome globally or to specific genes, I generated the heatmap for Hho1p binding sites. The heatmap suggests that Hho1p binds to the majority of the genome (Figure 3-3). And Hho1p distribution pattern is almost identical between both conditions, as the overlapping plot is pretty similar to individual heatmap pattern.

Figure 3-2. Histone H1 distribution pattern around TSS/TES in the genome.
Scaled tag counts for 0, 48 hr, and mock-IP datasets were binned in 15 bp intervals relative to the TSS (left panel) or TES (right panel) of 6,576 genes. These values were then normalized to the mock-IP values, and plotted with a 3-bin moving average. Also shown are occupancy levels of nucleosomal midpoints measured in the same genetic background (SK1).

3.3.2 Hho1p binds to the promoter region in subset of genome

Strikingly, Hho1p was also detected in 5’ nucleosome-free promoter regions (5’ NFRs), but not at 3’ NFRs at the ends of genes (Figure 3-2). There is also a black stripe in the heatmap (Figure 3-3) suggesting the Hho1p enrichment in the NFR occurs in significant portion of the yeast genome in both time points. The histone H1 generally associate with nucleosome and linker DNA, however, the majority of promoter regions are free of nucleosome. Therefore, the
enrichment in 5’ NFRs seems to be in conflict with the association between linker histone and nucleosome.

The 5’ NFR regions in most of the genes are around 130 bp in length (Xu et al. 2009), it is possible that Hho1p in the middle of the 5’ NFR binds to neighboring nucleosomes, such as “+1” and/or “-1” nucleosome. Therefore, we examined the RP genes, which have much larger 5’ NFR regions. There is a clear Hho1p peak in the 5’ NFR of the RP genes (Figure 3-4), and yet there is no mono-nucleosome around the Hho1p peak (Figure 3-4 nucleosomal grey fill background), suggesting that Hho1p can associate with promoter region even when the neighboring nucleosomes are further apart (>150 bp).

I then generated the heatmap for all RP genes in order to examine whether the enrichment of Hho1p can be observed in each individual RP gene. As shown in Figure 3-5, Hho1p binds to subset but not all RP genes. Such pattern is consistent in 48 hr sample as well. I separated the RP genes into two groups, the ones showing enrichment of Hho1p in the 5’ NFR (G3, 55 genes) and the ones (G1, G2 and G4) without Hho1p in the 5’ NFR.

Binding of histone H1 in the promoter region has been reported to repress transcription in mouse (El Gazzar et al. 2009). I want to evaluate whether Hho1p plays repressive role in the RP genes through its binding to the promoters. If this is the case, higher level of transcription repression are expected in these genes with Hho1p in the promoter. To test such assumption, fold changes of mRNA level between vegetative growth condition and 0 hr for both groups of the RP genes were obtained (Primig et al. 2000). Both groups showed a median value of 2 folds decrease in mRNA level and similar mRNA levels. Thus, I did not see clear evidence that the binding of Hho1p in the promoter represses transcription, at least in the mRNA level. However, this does not rule out the possibility that Hho1p can repress transcription.

If the enrichment of Hho1p in the promoter region is not due to transcription shutdown, then Hho1p has to be recruited to the promoter through other factors in a transcriptional-
independent manner. High mobility group protein in human cells colocalizes with histone H1 at the promoter region (El Gazzar et al. 2009). In yeast, high mobility group family HMO1 protein also associates with promoters of many RP genes and plays an important role in transcription regulation of RP (Berger et al. 2007). Therefore, I tested the correlation between Hmo1p and Hho1p occupancy at RP promoters. We extracted the significantly Hmo1p bound RP genes from previous ChIP-on-chip study (Hall et al. 2006). The chi-square tests indicated that the Hho1p promoter bound genes showed significant overlap with Hmo1p bound RP genes, while the Hho1 promoter unbound genes were significantly under-presented in Hmo1p bound genes (P-value 8.4 x 10^{-5}). Such enrichment is unique to Hmo1p, but not other transcription factors that bind to RP genes (Table 3-2) (Rudra et al. 2005; Kasahara et al. 2007). The Rap1p and Fhl1p bound RP gene lists were obtained from Sujana in the lab (unpublished). Taken together, this suggests that Hho1p may be recruited to the promoter of the RP genes through Hmo1p.

Table 3-2. Correlation between Hho1p and other transcription factors in RP genes.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Hho1 bound genes</th>
<th>Hho1 unbound genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hmo1p bound genes</td>
<td>Over-representations 8.4 x 10^{-5}</td>
<td>under-representations 8.4 x 10^{-5}</td>
</tr>
<tr>
<td>Rap1p bound genes</td>
<td>Random 0.52</td>
<td>Random 0.52</td>
</tr>
<tr>
<td>Fhl1p bound genes</td>
<td>Random 0.06</td>
<td>Random 0.06</td>
</tr>
</tbody>
</table>
Figure 3-3. Histone H1 heatmap overlap between 0 hr and 48 hr around genes. The read counts were color-coded and heatmap was generated as Figure 2-5. The 0 hr time point Hho1p data (red color) was overlaid with the 48 hr Hho1p data (green color).
Figure 3-4. Histone H1 distribution at RP genes between 0 hr and 48 hr around TSS of genes.

The composite for RP genes was generated similarly as Figure 3-2.
The heatmap for RP genes was generated similarly as Figure 2-5, the bin size is 5 bp with moving window of 3. The heatmap was in Hho1p data was clustered by K-mean clustering (K=4) and the H3 data was arranged to show the same order. The read count was color-coded.

3.3.3 Correlation between Hho1p occupancy and nucleosome

Although I have shown that Hmo1p may recruit Hho1p to the RP promoters, such mechanism will not apply to the linker DNA region as Hmo1p is enriched mainly in the promoter region. In addition, Hho1p is not evenly distributed in the linker DNA region, as shown in Hho1p panel on Figure 3-5. Therefore, I want to understand what factors cause such uneven distribution pattern of Hho1p on the nucleosomal array.

Since histone modifications alter the property of nucleosome, especially the binding with other proteins. It is possible that the histone modifications affect the binding affinity of Hho1p. I then examined the correlation between Hho1p data with nucleosomal and histone modifications datasets in Chapter 2.
I generated the scatter plot between Hho1 density (per gene) and bulk histone and histone modification density and performed the linear regression on all the dataset. Bulk H3 histone density is positively correlated with Hho1 density (Figure 3-6, A and B). I then evaluated correlation between five histone modification marks with Hho1p at meiotic sporulation onset, no strong correlation was observed for any of these histone modification marks, none of the $R^2$ values is above 0.50 (Figure 3-6, C to G). This suggests that the five histone modification marks examined does not facilitate nor prevent the association of Hho1p into the chromatin.
Figure 3-6. Histone H1 scatter plot with bulk histone and histone modifications.

The total read counts were calculated for every gene, the density of bulk histone and Hho1 was obtained by dividing the total tag count per gene by the gene length. The histone modification density was calculated through the ratio of histone modification tag count versus total bulk H3 tag counts. The linear regressions were performed to calculate the correlation coefficient.
3.3.4 The occupancy of Hho1p increases in the majority of yeast genome in spores compared to the initiation of meiosis

Considering the chromatin is more compacted in 48 hr, I wondered whether the occupancy of Hho1p changes between 0 hr and 48 hr. After normalizing with the western result from purified chromatin (unpublished data from Shelly Berger’s lab), the majority of fragments in the genome showed significant increase in Hho1p occupancy from the scatter plot (Figure 3-7). Such global increase of Hho1p occupancy in the genome is also confirmed by ChIP-PCR experiments (unpublished data from Shelly Berger’s lab).

However, a second population, consisting of 262 genes, was extracted. These genes did not experience a significant increase in Hho1 occupancy. Remarkably, these genes overlapped significantly with genes related to gluconeogenesis (GO:0006094, P-value 1.7 x 10^-6) and stress response (GO:0006950, P-value 3.4 x 10^-6), both of which are highly expressed in spores (Joseph-Strauss et al. 2007). Thus, increased Hho1 binding throughout the yeast genome in mature spores, except at spore specific expressed genes, may contribute to a global level of transcriptional repression.
Figure 3-7. Genomic fragment enrichment of Hho1p in spores versus meiotic onset.

The read count between two conditions were first normalized so that the total read counts in two samples ratio reflect the western blot signal intensity of Hho1 between two conditions. The genome is chopped into 100bp fragments, and total normalized tag counts of all 100-bp fragment were calculated for both time points and shown in the scatter plot. The 1:1 ratio diagonal line was shown in the figure as well.

3.3.5 Highly expressed genes in spores have lower Hho1p occupancy

Since Hho1p is involved in compacting the genome in the 48 hr spores, I suspect that the highly expressed genes in spores may adopt more open chromatin structure, and therefore less Hho1p is bound in these genes compared to the genome average. I retrieved the highly expressed genes from previous study and compared it with all genes level (Figure 3-8) (Joseph-Strauss et al. 2007). Indeed, significant depletion of Hho1p is seen between the highly expressed genes and genome average (compare red traces between the left and right panel in Figure 3-8). The bulk
nucleosome occupancy level was also examined, no significant changes is observed between 0 hr and 48 hr.

Figure 3-8. Hho1p occupancy of highly expressed genes in spores.

The composite for highly expressed genes in 0 hr and 48 hr composites were generated similarly as Figure 3-2. The composite plot for all genes was put on the left as a comparison.
3.4 Discussion

3.4.1 Regulation of global level of Hho1p in mature spores

In the Chapter two, I have shown that the primary structure of chromatin along the gene body hardly changes during the whole sporulation process. But the transcription level differs greatly upon sporulation induction, and this may be in part due to the alteration of higher order structure. The total protein level of HHO1 increased about 5 folds in the mature spore (unpublished data from Shelly Berger’s lab), however, no drastic changes in distribution pattern were observed between the two conditions in Hho1p ChIP-seq (Figure 3-2). This suggests that Hho1p occupancy in mature spores increases globally in the genome, not restricted to certain region in the genome (Figure 3-7).

Such global increase in Hho1p binding to chromatin has also been reported during the stationary stage (Schafer et al. 2008). However, unlike in mature spores, the protein level of HHO1 does not change in stationary stage. The authors also showed that Hho1p does not function as a general transcriptional repressor during stationary stage (Schafer et al. 2008). It is interesting in that yeast linker histone has distinct properties compared to the linker histone in mammalian cells. The linker histones in other model organisms were reported to repress transcription (Roche et al. 1985; Sera and Wolffe 1998). One of the major differences between yeast linker histone and mammalian linker histone is the linker histone/nucleosome ratio. A previous study showed that during mitotic growth condition, there is one histone H1 per 37 nucleosome (Freidkin and Katcoff 2001) or one histone H1 per 4 nucleosome (Downs et al. 2003), while the ratio is
normally around 1:1 in mammalian cells (Woodcock et al. 2006). Therefore, it is possible that histone H1 to nucleosome ratio has to be met in order for histone H1 to function properly.

Hho1p to nucleosome ratio may be too low in vegetative yeast cells to carry out its function. Consistent with such low level of Hho1p, yeast chromosomes are less compacted compared with other higher eukaryotes, and therefore increasing Hho1p level in the cell to achieve ~1:1 ratio may be adopted by yeast cells to compact the genome and repress global transcription. In addition, high expression of a heterogeneous H1 in yeast leads to transcription repression in the reporter gene (Linder and Thoma 1994).

Such drastic increase of Hho1p level is quite unique in mature spores and has not been observed in any other conditions reported so far. Because histone H1 to nucleosome ratio is already close to saturation in the mammalian system, such mechanism will not apply to mammalian cells. Similarly, if histone H1 to nucleosome ratio drops significantly in mammalian system, it is expected that the transcription repression by linker histone would also be compromised. Indeed, previous study suggested that a few folds decrease in protein level of histone H1 by siRNA would relieve the histone H1-mediated transcription repression significantly (El Gazzar et al. 2009).

In addition to bulk histone H1 amount, the modification states of core histones have also been suggested to affect the high order chromatin compaction. Recent studies showed that although Hho1p itself is able to condense chromatin to some extent, both Hho1p and deacetylation of core histone are required to fully reconstitute 30nm chromatin fiber in vitro (Li et al. 2010). Indeed, the global H3 K9 acetylation level of cells goes down during yeast sporulation (Figure 2-8). Interestingly, the acetylation levels of H4 increase in later stages of sporulation (Govin et al. 2010b). It will be interesting to examine the effect of H3 and H4 acetylation separately on the histone H1-mediated chromatin compaction.
3.4.2 Hho1 occupancy and histone density

Although I was able to show the enrichment of histone H1 in linker DNA region, it is still not known how linker protein is recruited to chromatin. I then evaluated whether the linker protein is recruited by a histone octamer. In that case, very strong correlation between Hho1p and nucleosome occupancy level is expected. However, I did not observe a strong correlation ($R^2$ value less than 0.5) between linker histone level with histone octamer level (Figure 3-6 A, B and Figure 3-5). This suggests that additional mechanisms are involved in recruiting linker histones to chromatin. I continued to test the hypothesis that whether particular histone modifications is involved in recruiting Hho1p. For the five histone modification marks examined so far, none of them showed significant correlation with Hho1p occupancy (Figure 3-6 C-G). Yet it is still possible that the other histone marks, histone variants or the combination of histone modification mark play a role in recruiting/evicting Hho1p. In addition, histone chaperons may also play an important role in recruiting Hho1p to chromatin (Ridgway and Almouzni 2001; Shintomi et al. 2005). Much work remains to be done to elucidate the incorporation of Hho1p into chromatin.

3.4.3 Hho1 distribution and transcription regulation

Previous expression studies showed that the HHO1 mutant strain is mainly defective in up-regulating a subset of mRNA genes. Nevertheless, these genes don’t have higher level of Hho1p occupancy compared to the genome average (Freidkin and Katcoff 2001). I intended to address whether these genes maintain a unique Hho1p distribution pattern based on the dataset I had generated. I plotted the composite plot of Hho1p distribution pattern for these genes at 0 hr and 48 hr, but no specific enrichment of Hho1p were identified neither in promoter nor genic region. Nevertheless, because I am not looking into the same vegetative growth condition, I
cannot rule out the possibility that the patterns of these genes in vegetative cells are drastically different from the two conditions presented here. It is also possible that down-regulation of mRNA in these genes is the downstream effect of upregulated genes in the *HHO1* deletion strain.

I have also shown that the highly expressed genes in spores have lower Hho1p occupancy, while most of the repressed genes have higher level of Hho1p occupancy (Figure 3-8). This suggests that higher occupancy of Hho1p may repress transcription in mature spores. However, there is no evidence supporting that Hho1p repress the transcription of mRNA genes directly *in vivo*. There might be several possibilities to explain why Hho1p in mature spores can repress mRNA expression, but not in vegetative yeast cells.

One possibility is that the amount of Hho1p in vegetative cells is low, which has been discussed in the 3.4.1 section. In fact, it is also possible that the modification states of histone H1 differ greatly between two conditions. The phosphorylation state of histone H1 in *Tetrahymena* greatly influences its function in gene regulation, and constitutive phosphorylation of H1 mimics the H1 depletion phenotype (Takami and Nakayama 1997; Dou et al. 1999; Dou and Gorovsky 2000). Interestingly, the phosphorylation of histone H1 is also coupled with cell cycle regulation. Histone H1 is highly phosphorylated in growing cells, but de-phosphorylated during starvation (Dou et al. 2005). Such changes of phosphorylation states between different conditions may also apply to *Saccharomyces* in order to regulate H1 function. Indeed, the mass spectrometry experiments identified the phosphorylated peptide of Hho1p in yeast (Li et al. 2007c) and it is possible that different phosphorylation states of Hho1 exhibit distinct functions in yeast as well. Therefore, it will be important to measure the types and quantity of covalent modifications on Hho1p *in vivo* among different conditions and then to dissect the function of these modified forms of Hho1p separately.
3.4.4 Hho1p enrichment at 5’ NFR

A previous study showed that upon heat shock, histone H1 is depleted from the promoter of repressed genes, but whose expression is not up-regulated (Zanton and Pugh 2006). Other studies on human cell lines suggested that histone H1 in the promoter region regulates the chromatin states and transcription (Koop et al. 2003). These together seem to indicate a potential regulatory role of histone H1 in the promoter region. Indeed, an Hho1p peak in the 5’ NFR region was observed in both conditions (Figure 3-2).

Such results suggested that linker histone could be enriched in the NFR, especially in the 5’ NFR of a subset RP genes, as shown in Figure 3-5. This may suggest that Hho1p can associate with DNA directly, or Hho1p can associate with other factors instead of nucleosome in the 5’ NFR region. So far, no one to our knowledge has reported that histone H1 can bind to promoter region without nucleosome in vivo. It has been shown that histone H1 preferentially bind to DNA with higher compaction (Ivanchenko et al. 1997). Therefore, it will be interesting to examine the higher-order organization of promoter DNA for the genes with Hho1p occupancy in the 5’ NFR, compared to the genes without Hho1p in the promoter region.

In addition, there is slight higher level of occupancy of Hho1p binding at the promoter region at 48 hr compared with 0 hr sample (Figure 3-2 compare red trace with black trace in the promoter region). It is possible that such increase level of Hho1p in 5’ NFR may be due to decrease occupancy of general transcription machinery, which needs to be verified by further experiments.

Consistent with co-localization between the intracellular high mobility group box 1 protein (HMGB1) and histone H1 at TNF-α promoter after severe systemic inflammation (El Gazzar et al. 2009), Hho1p are positively correlated with high mobility group family Hmo1p (Table 3-2). Similar to the repressive function of HMGB1 in mouse, Hmo1p is required for the
inhibition of RP expression through TOR-dependent pathway (Berger et al. 2007). Interestingly, during vegetative growth, neither mRNA level or rapid reduction of mRNA level of RP genes differs significantly between HMO1 deletion and wild-type strains (Hall et al. 2006). It will also be interesting to examine the Hho1p binding in vegetative cells and whether the binding of Hho1p is essential for the inhibition by Hmo1p. This suggests that the transcription repression is not turned on in growing cells and such repressive function of high mobility group protein and linker histone may require activation through TOR-dependent pathway.

In fact, the TOR pathway is the major regulatory pathway that responds to nutrient-deprived signal in yeast (Rohde et al. 2008). TOR pathway regulates the nuclear localization of nutrient-regulated transcription factors to control nutrient metabolism. The nutritional signals also induce sporulation through TOR pathway (van Werven and Amon 2011). More specifically, the RP genes are repressed through TOR pathway in response to nutrient deprivation (Rohde and Cardenas 2003). Taken together, this suggests that TOR pathway activated co-repression by histone H1 and high mobility group gene may also function in yeast.
3.5 Material and Methods

3.5.1 Sporulation induction

For sporulation, diploid yeast cells were grown in pre-sporulation media (2% Bacto Peptone, 1% Yeast Extract, 2% K Acetate) for approximately 15 hours to an OD$_{600}$ of approximately 1. Yeast were then washed with water and resuspended in sporulation media (2% K Acetate supplemented with auxotrophic amino acids) at OD$_{600} = 2$. Samples were taken for western blot or RT-qPCR analysis at the indicated time points. Progression through sporulation was determined by fixing cells equivalent to 1 OD$_{600}$ in 50% ethanol at various time points post induction of sporulation.

3.5.2 Preparation of sequencing sample

Hho1-3XFLAG and WT (mock ChIP) diploid cells (MATa/MATa his3-200/ his3-200 leu2-1/ leu2-1 ura3-52/ ura3-52 trp1::hisG/trp1::hisG HHO1-3xFLAG/ HHO1-3xFLAG SK1 background) were grown in pre-sporulation media (“0 hr”) or sporulated to maturity for 48 hours (“48 hr”). Approximately 50 OD of cells were cross-linked in 1% formaldehyde for 10 minutes and washed with water. Cells were resuspended in FA lysis buffer [50mM Hepes pH 7.5, 150mM NaCl, 2mM EDTA, 0.1% Triton X-100, 0.2% SDS, CPI Mini EDTA free (Roche)]. 1mL Zirconia/silica beads were added to each tube and cells were disrupted 3 minutes X 2 for the “0 hr” sample and 3minutes X 4 for “48 hr” sample at maximum speed with intermediate incubation at -20°C for 3 minutes (Mini-Beadbeater 96, Biospec). Chromatin was washed twice with FA lysis buffer and sonicated for 30 cycles (30 seconds ON at High level and 30 seconds
OFF per cycle) (Bioruptor, Diagenode). Cellular debris were removed by centrifugation at 14,000 rpm for 15 minutes at 4°C. Extracts were diluted with 3 volumes of FA lysis buffer without SDS (final SDS concentration of 0.05%) and incubated overnight with 200µL of anti-FLAG M2 Resin slurry (Sigma). The resin was washed to eliminate non-specific binding except that the deoxycholic acid is eliminated from all washing buffer and resuspended in 100µL M2 Elution buffer (50mL NaCl, 10mM Tris-HCl, 1mM EDTA) (Albert et al., 2007). 10µL 15mg/mL 3X FLAG peptide was added to elute Hho1-3XFLAG protein for 60 minutes at 4°C. Reverse cross-linking was performed at 65°C overnight (no longer than 16 hours) with 2µL Proteinase K (20mg/mL). The diploid DSY1089 cells were used as mock IP control and prepared similarly as the real samples.

3.5.3 Yeast whole cell extract preparation and western blot analysis

Yeast cells were harvested by centrifuging cultures at ~4000 rpm for 2 minutes at 4°C. The cells were washed with ice-cold water once and broken in lysis buffer (50mM Tris-HCl at pH 7.4; 300mM NaCl; 0.5% NP-40; 10% glycerol; 1mM EDTA pH 8; protease inhibitors that included pepstatin, leupeptin, aprotinin, and 0.5mM PMSF; phosphatase inhibitor cocktail from Sigma) using silica beads and bead beater (Bio-Spec). The lysate was sonicated for 5 minutes and cleared at 14,000 rpm at 4°C for 15 minutes. The supernatant was collected and stored at ~80°C. Protein estimation was determined using Bradford dye. Extracts were resolved on SDS-PAGE gels and images of the resultant gels or western blots were taken with a Fujifilm LAS-4000 imager. Contrast of images was adjusted with ImageJ software.
3.5.4 Data analysis

ChIP samples were subjected to paired-end sequencing using an Illumina GAII in accordance with the manufacturer’s instructions. The paired-end sequencing reads pass the filtering criterion only when the pair-ended reads are less than 300 bp apart. The midpoint coordinate of pair tags demarcated the binding location of Hho1p. *HHO1*-3XFLAG ChIP signal was normalized to the WT (untagged) ChIP signal. For scaling the ChIP-seq data as described in (Zhang et al. 2011a), I used ImageJ to quantify Hho1p and H4 band intensities from their respective western blots. Hho1 signal was normalized to the amount of H4 present (loading control), and we determined that there was a five-fold increase of Hho1 between the 0 hr and 48 hr time points. Two biological replicates of each time point were merged to demonstrate that the resulting patterns were reproducible, and the relative tag counts at 0 hr were set to be one-fifth of that at 48 hr, based on the western blot results.

The normalized reads were mapped to different genomic features such as the TSS and TES as described in previous chapter. In the scatter plot of Figure 3-7, the genome was fragmented into thousands pieces of 100 bp fragment. X-axis and Y-axis of each dot in the scatter plot represents the total normalized tag count for 0 hr and 48 hr respectively.

The correlation analysis of Hho1p and bulk histone and histone modifications were also represented in the scatter plot of Figure 3-7. The bulk nucleosome/Hho1p density on each gene equaled to the total tag count divided by the gene length, while the total histone modifications tag divided by total bulk nucleosome tag in each gene represented the histone modification density. Linear regression was performed in Kaleidagraph 4.1 to test the correlation and $R^2$ value was reported.
Chapter 4

Tetrasome does not reside on original nucleosome dyad and may be positioned by DNA sequence

4.1 Summary

Tetrasomes consisting of histones H3/H4 tetramers, but lacking H2A/H2B dimers, are stable intermediates in nucleosome assembly in vitro. However, their in vivo existence has not been established. Here I employed MNase ChIP-seq to map tetrasomes across the yeast genome under a number of environmental conditions. In all conditions, the tetrasomes in the NFR were sensitive to MNase digestion. Interestingly, the majority of tetrasomes were repositioned 25-50 bp from canonical nucleosome dyads, suggesting that they could have distinct physiological properties from intact nucleosomes. In addition, the CG dinucleotides are enriched in the dyads of tetrasome, which is consistent with in vitro nucleosome reconstitution results and patterns in the nucleosomal DNA in other model organisms. To determine whether tetrasomes retained fixed distance to nucleosome dyad, I mapped tetrasomes in ISW1 deletion strain and found that tetrasomes largely maintained their original positions though nucleosomes were shifted. Taken together, this suggests that the positions of tetrasomes were largely encoded by DNA sequence and independent of positioning of nucleosome.
4.2 Introduction

Eukaryotic genomes are packaged into nucleosomes, which consist of ~147 bp of DNA and a histone octamer containing two copies each of histones H2A, H2B, H3, and H4. Nucleosomes have one assembly/disassembly (H3-H4)$_2$ tetrasome intermediate that protects ~70 bp DNA. Here I explicitly refer to the former as a nucleosome and the latter as a tetrasome. The tetrasomes have been identified in vivo at the mouse mammary tumor virus (MMTV) promoter (Spangenberg et al. 1998). The nucleosome to tetrasome transition has been shown to facilitate the binding of transcriptional factors (Tse et al. 1998; Vicent et al. 2004). It is still not known whether tetrasomes represent physiologically relevant entities in yeast.

During DNA replication and moderate transcription, (H3-H4)$_2$ tetramers generally are stable and don’t split into two (H3-H4) dimers (Yamasu and Senshu 1990). But the splitting of (H3-H4)$_2$ tetramer can be observed at actively transcribed genes in yeast (Katan-Khaykovich and Struhl 2011). In addition to transcription level, the composition of (H3-H4)$_2$ tetramer also affects its stability. The H3.3 containing nucleosomes are more sensitive to salt-dependent disruption compared to H3.1 containing nucleosomes (Jin and Felsenfeld 2007). (H3.1-H4)$_2$ tetramers do not disassociate during DNA replication dependent nucleosome assembly, while large number of (H3.3-H4)$_2$ tetramers split during DNA replication (Xu et al. 2010).

Fluorescence Resonance Energy Transfer (FRET) assay suggested that nucleosome disassembly starts with opening of (H3-H4)$_2$ tetramer/ (H2A-H2B) dimer interface, then the release of (H2A-H2B) dimers and finally the removal of (H3-H4)$_2$ tetramer from DNA (Bohm et al. 2011). Consistently, point mutation strains on histone H4 that disrupt the interaction between histone tetramer and H2A-H2B dimer are lethal, while temperature-sensitive strains on one allele fail to express key genes for G1 cell cycle transition (Santisteban et al. 1997).
Several studies have indicated that histone chaperon is involved in tetrasome formation during DNA replication and transcription. Previous study also showed that histone chaperon Nap1p can bind to (H3-H4)₂ tetramer \textit{in vitro} and \textit{in vivo}, and facilitate the deposition of histone tetramer to DNA, suggesting that tetrasome may exist as an intermediate to form nucleosome (Bowman et al. 2011). Vps75p promotes H2A/H2B dimer disassociation from mono-nucleosome \textit{in vitro} (Selth et al. 2009). It has been suggested that Asf1p can facilitate (H3-H4)₂ deposition onto DNA, however, the removal of (H3-H4)₂ by Asf1p requires additional factors or post-translational modifications (Donham et al. 2011).

As an intermediate during nucleosome assembly and disassembly, one of the key questions is whether the tetramer positioning in nucleosome is maintained in the tetrasome. Previous \textit{in vitro} reconstitution study suggested that the (H3-H4)₂ tetramer can protect around 70bp DNA sequence (Dong and van Holde 1991), and the histone tetramer itself is able to find the DNA sequence to bind through salt gradient dialysis. However, it is not known where intermediate tetrasomes reside \textit{in vivo}. Previous \textit{in vitro} studies on nucleosome survival after transcription have shown controversial results on the positioning of hexamer ( (H3-H4)₂ tetramer with one copy of H2A/H2B dimer) intermediates: hexasome remains at original position during \textit{in vitro} transcription on short (250-300 bp) template, while hexamer is shifted away from octameric dyad on longer \textit{in vitro} transcription template (658-3,900 bp) (Kireeva et al. 2002; Kulaeva and Studitsky 2010). The author also proposed that short template represents the case where no significant nucleosome loss occurs, while longer template pattern will be observed in the case when the majority of the nucleosomes are depleted (Kulaeva and Studitsky 2010).

The other important question is, whether tetrasome positioning and occupancy is influenced by DNA sequence. Considering the crucial role of DNA sequence in positioning nucleosome described in Chapter one and \textit{in vitro} studies (Dong and van Holde 1991), it seems very likely that the DNA sequence also regulates tetrasome assembly. In addition, it has been
suggested that the depletion of H2A/H2B dimer depends on DNA sequence (Vicent et al. 2004; Kelbauskas et al. 2008). In MMTV promoter of mouse, SWI/SNF complex specifically displaces two H2A/H2B dimers on one nucleosome but not the other one in the promoter region. Based on this and some other pieces of evidence, the authors concluded that it must be the DNA sequence that determines the outcome of remodeling complex (Vicent et al. 2004). The FRET assay on H2A/H2B depletion also showed sequence-dependent variations from three well-characterized promoters (Kelbauskas et al. 2008).

Here I reported on the genome-wide in vivo mapping of tetrasomes in yeast genome, and intended to address the following questions: Do tetrasomes appreciably exist in vivo? Do they occupy the same positions as their nucleosome counterparts? Are tetrasome positioning and occupancy determined by DNA sequence? Are they linked to gene regulation?
4.3 Results

4.3.1 Tetrasomes are relatively rare in vivo

Nucleosomes are operationally defined as ~120-180 bp DNA fragments that arise after MNase treatment of chromatin. MNase digests linker DNA but generally leaves nucleosomal DNA intact. Tetrasomes are operationally defined as producing ~30-105 bp MNase-resistant fragments. In undertaking a genome-wide study of tetrasomes, it was first necessary to demonstrate that tetrasomes were not simply nucleosomes that were particularly sensitive to MNase. By definition, tetrasomes lack H2A/H2B. However, in vivo a given nucleosome position may have a mixture of nucleosomes and tetrasomes in a population of cells. If positions exist that contain predominantly tetrasomes and a low percentage of nucleosomes, then their presence can be discerned as having H3/H4 but little or no H2A/H2B in a chromatin immunoprecipitation (ChIP) experiment. However, if positions exist that contain predominantly nucleosomes with few tetrasomes, then their H2A/H2B levels may be not be measurably different from H3/H4 levels. A lack of difference could be erroneously construed as a lack of tetrasomes. However, tetrasomes may be quite transient, for example in the wake of transcription or DNA replication, and thus represent a small fraction of total nucleosomes at any position. Therefore, detection of such tetrasomes would require that they possess a measurable property that is distinct from its nucleosome counterpart. Here we use these alternative criteria to validate different types of putative tetramers found across the genome.
Figure 4-1. PCR amplification of different size fragment.

The different size range ChIPed DNA fragments from the same sample were size-selected on the PAGE gel and prepared for illumina sequencing library. The ~140-220 bp range fragment includes the illumina adaptor sequence (108bp) and tetrasome-sized DNA fragment (~30-105 bp). The 220-298 bp range fragment represents the mono-nucleosome sized DNA. The 340-500 bp range fragment corresponds to di-nucleosome DNA fragment. The 12 and 15 cycles PCR amplification products were run on 2% agarose gel respectively.

Figure 4-2. WC distance frequency distribution of sequenced fragment in all three conditions.
The peak calling was performed on W and C strand separately using genetrack (Albert et al. 2008a), then W and C strand peak were paired based on the WC distance mode value. No tag count filtering was involved. The frequency distribution plots of WC distance among all peak pairs in the genome were generated for low (black trace) and high (red trace) level of MNase digestion in all three conditions (YPD, 0 hr and 24 hr).

To examine the genome-wide organization of putative tetrasomes and their changes in response to environmental perturbation, yeast cells were either grown exponentially in rich (YPD) media, or over night in media (YPA) lacking a fermentable carbon source, or in media (SPM) lacking nitrogen for 24 hr. The latter resulted in >95% of the cells differentiating into spores. Tetrasomes were isolated from yeast cells by first subjecting cells to 1% formaldehyde crosslinking, followed by MNase digestion of isolated chromatin, H3 immunoprecipitation, and gel-purification of DNA fragments in the size range of 30-105 bp. The use of H3 antibodies ensured that the isolated DNA fragments were not simply histone-free DNA fragments that were intrinsically resistant to MNase. Nucleosomes were similarly isolated by size-selection in the 120-180 bp range. The nucleosome band can normally be visible by amplify 9-12 PCR cycles, while the tetrasomes from equal cell equivalent can only be detected after 15 or more PCR cycles (Figure 4-1). Thus, tetrasomes are rare in vivo in comparison to nucleosomes, although the gel-quantification should only be considered approximate in that over-digested octamer nucleosomes in this assay could be misconstrued as tetramers. Nevertheless, rare tetrasomes do not necessarily indicate a lack of physiological relevance; as such tetrasomes might be highly localized in the genome and/or represent transient intermediates in nucleosome assembly/disassembly.

The resulting size-selected DNA fragments were subjected to deep sequencing, with clustered 5’ ends defining the two borders of each tetrasome; the sequencing read count summary was shown on Table 4-1. Confirming the presence of properly sized fragments, the mode distances between the two borders of a tetrasome are around 70 bp (Figure 4-2). Only those nucleosomes with appropriate size (30-105 bp) were chosen for further study.
Table 4-1. Sequencing tag count summary.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Biological replicates</th>
<th>Total replicates</th>
<th>Uniquely mapped tags</th>
</tr>
</thead>
<tbody>
<tr>
<td>YPD low</td>
<td>2</td>
<td>2</td>
<td>15,514,374</td>
</tr>
<tr>
<td>YPD high</td>
<td>1</td>
<td>1</td>
<td>12,867,929</td>
</tr>
<tr>
<td>0h low</td>
<td>2</td>
<td>2</td>
<td>12,671,662</td>
</tr>
<tr>
<td>0h high</td>
<td>1</td>
<td>1</td>
<td>14,442,480</td>
</tr>
<tr>
<td>24h low</td>
<td>3</td>
<td>3</td>
<td>18,505,323</td>
</tr>
<tr>
<td>24h high</td>
<td>2</td>
<td>2</td>
<td>9,742,353</td>
</tr>
</tbody>
</table>

Low suggests the low MNase digestion level, while high represents the higher MNase digestion level for sample treatment.

4.3.2 Tetrasomes at 5’ and 3’ NFR are sensitive to MNase digestion

To eliminate the over-digestion concern of MNase treatment, I tested different MNase digestion levels. The tetrasomes at 5’ NFR were sensitive to MNase digestion. The tetrasome peaks in the NFR regions were completely eliminated at higher MNase digestion condition in all three conditions (Compare red trace high MNase with black track low MNase Figure 4-3), similar pattern was observed for the 3’ NFR as well. This result was consistent with previous nucleosome data that nucleosome in the NFR is sensitive to high MNase digestion (Weiner et al. 2010). Therefore, I used low MNase data in the following analysis. Despite the drastic difference observed at NFR tetrasomes, the occupancy and positioning of genic tetrasomes, especially the +2 and even downstream tetrasomes, is pretty similar between high and low MNase digestion.

Considering the lower transcription level in mature spores (Chu et al. 1998), lower levels of genic tetrasomes are expected. Indeed, this was observed at 24 hr dataset (Figure 4-3). Noticeably, the tetrasome peaks in mature spores (24 hr) are much higher compared to the other two conditions. The functional significance of the increased tetrasomes at NFR is not clear.

To examine whether tetrasomes bind to the promoter of all genes, heatmaps of tetrasome from -1000 bp to 1000 bp relative to TSS were generated and clustered by K mean (K=5). The
black stripe of NFR region in nucleosome heatmap (Figure 2-5) is not observed in tetrasome heatmap (Figure 4-4), suggesting that tetrasomes bind to the promoters of most genes.

Interestingly, subset of the genes showed much stronger tetrasome occupancy in the promoter compared to other genes. 745 genes showing highest level of tetrasome in the promoter of mature spores (24 hr) were extracted (Figure 4-4). I then performed the GO analysis on these 745 genes. The chi-square test indicated that RNA processing (RNA metabolic process $P$ value $4.3\times10^{-10}$) and cellular metabolism related groups (cellular metabolic process $P$ value $1.4\times10^{-9}$) are significantly enriched. Both groups of genes are highly repressed in mature spores. Therefore, it is possible that the transcription repression will lead to enrichment of tetrasomes in the promoter region in mature spores.

### 4.3.3 Transcription regulates the stability of fragile tetrasomes

The fragile nucleosomes are defined as the nucleosomes showing high sensitivity to MNase digestion. Clearly, such fragile tetrasomes were observed in Figure 4-3. A previous study has showed that fragile nucleosomes are enriched in upregulated genes upon heat shock (Xi et al. 2011). To study the correlation between the fragile nucleosome and transcription, I calculated the fragility index for each gene in each condition. Low fragility index means that tetrasomes are stable in the NFRs, while genes with high fragility index have unstable tetrasomes in the NFRs.

In vegetative cells, the genes that have lowest fragility index overlap very significantly with lowly expressed genes ($P$ value $6\times10^{-17}$). This suggests that tetrasomes in lowly expressed genes are more stable. In addition, general repressor Rap1p and Tup1p are enriched in these genes with lower fragility index ($P$ value $2\times10^{-15}$ and $P$ value $1\times10^{-18}$ respectively). On the contrary, the genes with highest fragility are correlated with lower level of Rpd3p ($P$ value $1\times10^{-5}$), which also involves in forming repressive chromatin. Taken together, the factors involved in
the establishment of repressive chromatin can stabilize the tetrasome as well. Noticeably, Swr1p bound genes tend to have unstable tetrasomes in the promoter regions ($P$ value $1 \times 10^{-4}$), suggesting that chromatin remodeler may function to destabilize the tetrasomes.

To further demonstrate that the stability of tetrasome is correlated with transcription, I examined the changes of the fragility index between two conditions: YPD and stationary stage (0 hr). The up-regulated genes showed significant overlap with the genes whose stability goes down the most ($P$ value $1.6 \times 10^{-5}$). This again supports the idea that higher level of transcription normally leads to lower tetrasome stability in the promoter and vice versa.
Figure 4-3. NFR tetradsomes are more sensitive to MNase treatment.

The composite plots were generated as Figure 2-2A, total read count between two MNase digestion conditions were normalized to be equal. The black trace represents the low level of MNase digestion while the red trace represents the high level of MNase digestion.
Figure 4. Heatmap of tetrasome distribution at 24 hr.

The heatmaps were generated similarly as Figure 2-5. K means cluster was performed on the 24 hr dataset (K=5). When ordered according to one replicate after clustering, two biological replicates of 24 hr sample showed almost identical patterns. The first cluster (G1) showed highest level of tetrasome in the promoter region. The tag counts were color-coded.
4.3.4 Tetrasomes are not over-digested nucleosomes

To validate that tetrasomes are not just due to over-digestion, I compared the tetrasomes being identified in growing cells with other H2A, Htz1, H2B containing nucleosomes (ChIP-exo data generated by Ho Sung Rhee in the lab) in vegetative growth condition. Because H2A, Htz1 and H2B bound nucleosome data were generated by ChIP-exo, which is an MNase independent approach, there is no MNase bias in the ChIP-exo dataset for H2A/H2B subunits (Rhee and Pugh 2011). Therefore, H2A/H2B would not be detected in the tetrasomes if the tetrasomes that I identified were real. If we could detect H2A/H2B in the tetrasomes, it suggests that the tetrasomes we identified is just due to over-digestion of MNase.

Considering that the tetrasomes and nucleosomes can coexist in a population of cells, there are potentially three types of nucleosome. Therefore, I classified all nucleosomes into three types of nucleosome, tetramer-only nucleosome (I), tetramer and octamer co-existing nucleosome (II) and octamer-only nucleosome (III). To assign the reference nucleosome into these three categories, I mapped the dyads of tetrasomes and nucleosome to all reference nucleosomes (Jiang and Pugh 2009). If the dyads of tetrasome/nucleosome fall into the -73 to +73bp of reference nucleosome dyad, I assigned the mapped tetrasome and/or nucleosome to that reference nucleosome. Similarly I applied the same approach to H2A, Htz1 and H2B subunit as well.

As expected, the tetramer-only nucleosomes showed under-representative overlapping with the nucleosomes containing H2A, H2B or Htz1, while the octamer-only nucleosomes showed over-representative overlapping with H2A/H2B containing nucleosome (Table 4-2). This result suggests that the tetrasomes I identified are indeed depleted of H2A/H2B subunits. Noticeably, the H2A containing nucleosomes show more significant correlation with the octamer-only nucleosomes (0.08 versus 0.17), while the Htz1 containing nucleosomes exhibited higher correlation with the tetramer-octamer co-existing nucleosomes (10^{43} versus 10^{6}). This may
suggest that for the moderately expressed genes, the H2A.Z/H2B dimers may preferentially associate with tetrasomes to form newly assembled nucleosomes after partial nucleosome disassembly by transcription. While the stable nucleosomes, which are not frequently disrupted by transcription, tend to contain H2A/H2B dimer compared to H2A.Z/H2B.

<table>
<thead>
<tr>
<th>Chi-square test</th>
<th>Htz1-containing nucleosomes</th>
<th>H2A-containing nucleosomes</th>
<th>H2B-containing nucleosomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tetramer-only Nucleosomes</td>
<td>(close to 0) under</td>
<td>10^{-189} under</td>
<td>10^{-309} under</td>
</tr>
<tr>
<td>Tetramer-Octamer Nucleosomes</td>
<td>10^{-33} over</td>
<td>0.17 over</td>
<td>10^{-20} over</td>
</tr>
<tr>
<td>Octamer-only nucleosomes</td>
<td>10^{-6} over</td>
<td>0.08 over</td>
<td>10^{-5} over</td>
</tr>
</tbody>
</table>

4.3.5 Tetrasome dyads shift away from the canonical dyad locations of nucleosomes

An interesting property of genic tetrasomes is found in the composite plot of their distribution around the TSS in low MNase level (black traces in Figure 4-3 left panel). Particularly evident at the +1 position is a bimodal distribution. However, such pattern is not seen in the high MNase level (red traces in Figure 4-3 left panel). This observation suggests that tetrasomes may not reside at the same dyad positions in the nucleosome.

To test this hypothesis, the distribution of all tetrasomes was plotted around the canonical midpoint/dyad location of resident nucleosomes, and the bimodal distribution became more evident (Figure 4-5). And such bimodal distribution can be seen all conditions, YPD and 0 hr samples were shown in Figure 4-5.
Figure 4-5. Tetrasomes peaks are shifted away from nucleosome dyad.
The peak calling was performed for tetrasome samples using genetrack (sigma=10, D=70), and the distances between the tetrasome dyad and nucleosome dyad were calculated. The direction of transcription is also considered in this graph, the positive value represents downstream shift, and vice versa.

To confirm the pattern observed in Figure 4-5, I also generated the heat map plotting the unfiltered tetrasomes tag (globally shifted to the center of tetrasomes) relative to the nucleosome dyad (Figure 4-6). Apparently, the majority of the tetrasomes do not reside at the same location as their nucleosome counterparts.
Figure 4-6. Tetrasome tags distribution relative to nucleosome dyad in YPD.

All the tetrasome tags were shifted by half value of mode WC distance, which is 33 bp. The raw shifted read count was plotted relative to nucleosome dyad defined in chapter 2 (-80 to 80 bp). On the right panel, the composite plot for each cluster was generated from two biological replicates (red and black traces).
Tetrasomes shift away from the dyad position in either direction by ~25-50 bp (Figure 4-5). This raises an interesting question as to whether such repositioning represents a thermodynamic move towards energy minimization, or whether tetrasomes are handled differently than nucleosomes by chromatin remodelers, which are expected to operate on nucleosomes at the 5’ ends of genes.

Normally positioned nucleosomes display biased distributions of dinucleotides across the nucleosomal sequences. Such biases of dinucleotides are thought to contribute to nucleosome positioning and are also supported by in vitro reconstitution results. Tetrasomes also displayed a dinucleotide biased or periodic arrangement across the tetrasomal DNA (Figure 4-7), but the pattern was opposite to that in nucleosomes. In fact, the displacement of dinucleotide pattern is consistent with the ~25-50 bp positional offset between tetrasomes and nucleosomes (Figure 4-5). The enrichment of CG dinucleotide in tetrasomes dyad is consistent with in vitro experiments, and such CG dinucleotide enrichment in the nucleosome dyad was observed in other species such as C. elegan and human (Kaplan et al. 2009; Valouev et al. 2011). To conclude, this suggests that tetrasomes reside at thermodynamically favored positions.

Figure 4-7. Dinucleotide distribution relative to nucleosome & tetrasome dyad in YPD.
The dinucleotide frequency plots were generated relative to nucleosome dyads from wild type cells. The dinucleotide distribution of AT/CG of nucleosome was plotted in black traces, while the dinucleotide distribution of AT/CG of tetrasome was plotted in red traces.

Had these tetrasomes simply reflected over-digested nucleosomes, I would expect that the enrichment on both side of nucleosome dyad (Figure 4-5 and Figure 4-6) become more pronounced in the over-digested samples. However, when I shifted all tetrasome tags to their dyads and mapped them to the -100 bp to +100 bp region of each nucleosome dyads. The two peaks that are ~30 bp away from nucleosome dyad did not get sharper in the over-digested condition (Figure 4-8). Instead, the over-digested tetrasome tags were more enriched in the dyad region, which suggests that the over-digested nucleosomes are more likely to reside at the dyad of the nucleosomes (Figure 4-8). In the other words, the over-digested nucleosomes tend to localize at the original dyad locations. This also suggests that the tetrasomes I identified are not due to the over-digestion by MNase treatment.
Figure 4-8. Tetrasome tag distribution relative to nucleosome dyad in two MNase digestion levels.

The shifted tags were mapped to the nucleosome dyad, and the composite plots were generated relative to the nucleosome dyads.
4.3.6 Positioning of tetrasomes is independent of nucleosomes

Since the CG dinucleotide is enriched in the dyads of tetrasomes, it is possible that the tetrasome position is encoded by DNA sequence and independent of the positioning of nucleosome. If so, the enrichment of tetrasome dyad at ±35 bp away from nucleosome dyad will not been maintained when nucleosomes shift. Given the nucleosomes in the coding region shift significantly upstream in the ISW1 mutant (Tirosh et al. 2010), I mapped the tetrasomes in the ISW1 mutant strain in order to test this hypothesis.

Consistent with my assumption, no clear connection between tetrasome dyads and nucleosome dyads was observed in Figure 4-9. The tetrasome data (Figure 4-9B) in ISW1 mutant were ordered according to the shifted distance of nucleosome dyad in ISW1 mutant relative to wild type cells (Figure 4-9A). On the contrary, when tetrasome data in ISW1 mutant were ordered according to wild type tetrasome data, it is clear that most of tetrasomes maintain their positions in the ISW1 mutant (Figure 4-10). I did observed slight increase in tetrasomes located at upstream region. However, the origin and functional significance of such newly detected tetrasome is not clear.
Figure 4-9. Heatmap of tetrasome dyad relative to nucleosome dyad in *ISW1* mutant.

The peak calling was performed by genetrack (Albert et al. 2008a), nucleosomes with WC distance ranges from 120 to 180 bp were selected and tetrasomes with WC distance ranges from 30 to 105 bp were used in further analysis and the middle points of these nucleosome/tetrasome peaks were plotted as dot. The *ISW1* nucleosome data was ordered according to the distance between *ISW1* mutant dyad and wild type dyad. The tetrasome data of *ISW1* mutant on the right is ordered to be the same as nucleosome (red and blue represented two biological replicates).
Figure 4-10. Heatmap of tetrasome dyad relative to nucleosome dyad in wild type and ISW1 mutant.

The peak calling was performed by genetrack (Albert et al. 2008a), tetrasome with WC distance ranges from 30 to 105 bp were used in further analysis and the middle points of these tetrasome peaks were plotted as dot. The wild type data was ordered according to the distance between upstream tetrasome dyad and nucleosome dyad. The ISW1 mutant data (blue) is ordered to be the same as wild type (red).
4.3.7 Tetrasomes occupancy level is not positively correlated with transcription frequency

If tetrasomes are the intermediate products depleted of H2A/H2B subunit, it is likely that there is higher tetrasome density in actively transcribed genes. To validate such assumption, I have plotted the tetrasome density on all genes, genes with top 5% transcription frequency, genes with top 10% transcription frequency and genes with bottom 10% transcription frequency (Holstege et al. 1998). However, no significant difference was found among these four groups (Figure 4-11A). The tetrasome density is higher in highly expressed genes when normalized by the nucleosome levels (Figure 4-11B), but such enrichment is largely due to the lower bulk nucleosome density in actively transcribed genes.

![Figure 4-11](image)

Figure 4-11. Frequency distribution of tetrasome density in 4 groups of genes.

The transcriptional frequency data is retrieved from previous study, 4 group of genes were selected: all genes (black trace), genes with top 10% transcription frequency (orange trace), genes with top 5% transcription frequency (red trace) and genes with bottom 10% transcription frequency (blue trace). The total tetrasome tags mapped to each gene were summed and divided by the gene length and log 2 transformed in (A), while the total tetrasome tags were summed and divided by the total nucleosome tags and log 2 transformed in (B).
4.4 Discussion

4.4.1 Dynamic picture of nucleosome assembly and disassembly

Both the depletion of H2A-H2B dimer from nucleosome and de novo nucleosome assembly can generate tetrasomes, but I am not able to distinguish between the two by the approach applied in this thesis. As intermediate products, they will be eventually reconstituted to form mono-nucleosome. Our data suggested that a (H3-H4)$_2$ tetramer shifted its location between tetrasome and nucleosome states, which fits with the previous model that the nucleosome either shifts upstream or downstream after the Pol II transcribes through the nucleosome (Kireeva et al. 2002). Therefore, I propose that to shift the (H3-H4)$_2$ tetramer back to its nucleosomal position, H2A/H2B dimer binds to (H3-H4)$_2$ tetramer first, and then the whole re-assembled nucleosome is shifted by chromatin remodeler to return to the original mono-nucleosome position.

This model is consistent with previous in vitro result showing that tetrasomes are poor substrate for SWI/SNF complex compared to mono-nucleosome (Boyer et al. 2000). In addition, this model also fits with ISW1 nucleosome and tetrasome data. The positioning of tetrasomes in the cells lacking ISW1 does not change much, while most of the nucleosomes in coding region of ISW1 mutant have shifted away from their canonical positions in wild type cells (Figure 4-9 and 4-10).

The next question is what determines the positioning of tetrasome? One possibility is that tetrasome preferentially binds to DNA favored position. This assumption is consistent with the fact that tetramer itself determine the nucleosome positioning in vitro (Dong and van Holde 1991). In addition, I also observed that tetrasomes tend to have CG enriched in the dyad compared to AT enriched for nucleosome in yeast. Interestingly, only the yeast genome shows AT enrichment in the nucleosome dyad, while fly, human, C. elegan genomes all show CG
enrichment in the dyad (Mavrich et al. 2008b; Valouev et al. 2011). Moreover, the enrichment of CG versus AT is also seen in the Salt Gradient Dialysis (SGD) sample as well (Zhang et al. 2011b), which fits with the current concept that CG dinucleotide in the dyad is more energetically favorable for nucleosome to reside on. Therefore, tetrasomes may just reside on the DNA favored positions after the polymerase move along the gene body. Upon the active reconstitution and other chromatin remodeler activity, a nucleosome is reconstituted and positioned back by the chromatin remodeler complex to maintain proper nucleosome positioning and spacing.

To further confirm the tetrasome positioning does not rely on nucleosome, I mapped tetrasomes in the ISW1 mutant. If tetrasomes retained fixed distance to nucleosome, it is expected that the positioning of tetrasome is linked with nucleosome. However, positioning of tetrasomes does not follow the shifting of nucleosomes (Figure 4-9), instead majority of tetrasomes maintain their original positions. In conclusion, tetrasome positions in ISW1 mutant strain supported the assumption that DNA sequence dictates tetrasome position.

The thermodynamic favored positions may be more efficient in energy compared to the model that tetrasomes maintain the same location. In that case, tetrasomes have to be repositioned back each time RNA polymerases move through the nucleosome. It is a waste of energy to constantly reposition these intermediate tetrasomes when they will be continuously perturbed by Pol II transcription.

4.4.2 Genic tetrasome level and transcription frequency

I did not see the positive correlation between the tetrasome and transcription frequency, however, when normalized tetrasome occupancy to nucleosome level in the same gene, the normalized tetrasome density is positively correlated with transcription level (Figure 4-11B). This does not necessarily disconnect tetrasome formation and transcription. In the highly expressed
genes, there are multiple RNA polymerases moving along the genes at the same time. Multiple RNA polymerase complexes constantly move through the nucleosome, which will lead to complete nucleosome eviction (Kulaeva et al. 2010). Therefore, although the rate of tetrasome generation in highly expressed genes is high, the tetrasome eviction rate is also higher. This may explain why the highly expressed genes do not have higher-level tetrasome in Figure 4-11A.

Presumably, the incorporation of H2A/H2B requires tetrasome, and therefore the tetrasome level will limit the rate of H2B incorporation. Indeed, similar level of tetrasome density is consistent with equal level of H2B incorporation among actively and inactively transcribed genes (Jamai et al. 2007).

4.4.3 Unstable tetrasome in the NFR regions

It is very clear that the tetrasomes in the NFRs are easily digested upon high level of MNase (Figure 4-3). In fact, easily digested nucleosomes have also been reported in the NFR regions (Weiner et al. 2010). The nucleosome fragility in the promoter poises gene for a swift up-regulation upon environmental stress (Xi et al. 2011). The unstable tetrasomes in the NFR regions may also be the intermediate products of the unstable nucleosomes in the NFR regions.

Compared to the genic tetrasomes, they have similar dinucleotide pattern in the tetrasomal DNA. The stability of these tetrasomes may be due to the constant removal by the chromatin remodeler and/or other transcription factors in the promoter region. However, such assumption does not agree with 24 hr dataset (Figure 4-3). Considering the global transcription shutdown, the activity of chromatin remodeler may be much lower in the 24 hr. This will possibly lead to the increased stability of tetrasome in 24 hr samples, which is not seen in the data (Figure 4-3). Therefore, the DNA-tetrasome interaction in the NFR regions may differ from that in the genic regions. Considering the abundant of transcription factors in the promoter, it is possible that
the tetrasomes in the NFR regions are “factor-interacting” tetrasomes. The association of factors to tetrasome may change the DNA-tetrasome interaction in promoter regions. In fact, the Bdf1-interacting nucleosomes have been shown to have lower H3 turnover rate (Koerber et al. 2009). Considering the potential functional role of such fragile nucleosome/tetrasome, these fragile nucleosomes/tetrasomes can be utilized to uncover the novel promoter region and even to predict the promoter activity.
4.5 Material and Methods

4.5.1 Tetrasome isolation

The cell lysis, MNase treatment and sequencing library preparation were following almost the same protocol as Chapter 2, *ISW1* mutant strain (BY4741 background, *MAT a his3D1 leu2D0 met15D0 ura3D0 isw1D*), except that tetrasome size DNA fragment was selected (140 – 220 bp instead of 240 – 300 bp). Each sample was treated with both high and lower MNase treatment. Most of the product in higher MNase treatment is mono-nucleosome, while di-nucleosomes and tri-nucleosomes can be observed in lower MNase treatment condition. Also the libraries were multiplexed and PCR amplified from 15 to 21 cycles. And all the samples were sequenced in the illumina sequencing platform.

4.5.2 Data analysis

The tag mapping and peak calling was performed by genetrack similarly as Chapter 2, except that sigma 10 and exclusion zone 70 bp were used in tetrasome peak calling (Albert et al. 2008b). In addition, the W strand tags were shifted ~35 bp downstream (half of the mode CW distance), and the C strand tags were shifted ~35 bp upstream. The mapping of globally shifted tags was done not only on TSS/TES but also on nucleosome dyads (such as all nucleosome dyads, +5 and +6 nucleosome dyads). The middle points of nucleosome W-C peak pair of YPD data in Chapter 2 were used as the nucleosome dyads.

Only the tetrasomes with WC distance range from 30 to 105 bp were selected, similarly nucleosomes with WC distance ranging from 120 to 180 bp were selected for downstream
analysis. The dyads of filtered tetrasomes and nucleosomes were mapped to -80 to +80 bp range of each reference nucleosome (Jiang and Pugh 2009). Then the reference nucleosomes with tetrasome mapped only were classified as tetramer-only nucleosome, the reference nucleosomes with both tetrasomes and nucleosomes mapped were defined as tetramer octamer coexisting nucleosome and the reference nucleosome with only nucleosomes mapped were considered to be octamer-only nucleosomes.

To represent the stability of fragile nucleosome in the 5’ NFR regions, I calculated a fragility index by using the low and high MNase levels data for each condition. The promoter total tag count (-300 bp to 0 bp relative to TSS) and middle genic region total tag count (middle one-third of the mRNA transcripts) were obtained. The density of tag for promoter (Dp) and middle genic regions (Dm) were calculated by normalizing total tag count through the DNA length. The fragility index is then calculated by $\log_2 (Dp/Dm)_{\text{low MNase}} - \log_2 (Dp/Dm)_{\text{high MNase}}$. The higher the fragility index is, the higher occupancy changes will be seen at the promoters between two MNase conditions. Considering the tetrasomes in the ORF are generally stable (Figure 4-3), the middle genic regions were used as an internal control between two conditions.

The total count of each dinucleotide (AT, GC and so on) at each position (-100 to +100 bp) relative to nucleosome dyads were summed and plotted. The nucleosome dyads were obtained from data in Chapter 2. Both W strand and C strand DNA was considered in calculating the dinucleotide frequency.
Chapter 5
Discussion

5.1 Summary of this study

5.1.1 Stable and specific changes of primary nucleosome structure during yeast meiosis

In Chapter 2, I have mapped bulk H3 nucleosome and several histone modifications during several stages of yeast meiosis. The data showed that primary chromatin organization and histone modification patterns were largely maintained (Figure 2-2 and Figure 2-5). It indicates that the high order chromatin compaction during meiosis does not necessarily alter the nucleosome positioning and occupancy, which has never been demonstrated before by genome-wide data. Drastic decrease of histone modification levels upon global transcription repression is not observed during meiosis, instead histone modification marks are also stably inherited into yeast spores without replication (Figure 2-9 and 2-10). The inheritance of histone modification marks during meiotic cell cycle may also be applied to other high eukaryotes. Both nucleosome and histone modification data supports the notion that the primary chromatin structure is stable during meiosis.

However, specific changes in nucleosome occupancy and positioning occur to allow proper sporulation progression (Figure 2-6 and Figure 2-7). In addition, histone modification results also suggested that anti-sense transcription might repress meiotic gene expression during mitotic cell cycle (Figure 2-15). To conclude, during yeast meiosis, the genome basically maintains the primary chromatin structure and only makes specific changes to allow the meiotic progression.
5.1.2 Mysterious linker histone in yeast

Unlike the well-characterized core histone proteins, the function of linker protein is poorly characterized. In Chapter 3, I generated a genome-wide ChIP-seq mapping for linker histone Hho1p during sporulation onset (0 hr) and completion (48 hr), and tried to understand the function of Hho1p in chromatin compaction and transcription regulation. Western blot quantification and ChIP-seq data together showed that there is a global increase in Hho1p recruitment to chromatin at in mature spores, but not specific to small portion of the genome (Figure 3-7). The deletion of \textit{HHO1} leads to abnormal high order chromatin structure in vegetative cells (Georgieva et al. 2011a) and enlarged nuclei volume in yeast spores. Therefore, the global increased Hho1p incorporation to chromatin may play a crucial role in forming more condensed high order chromatin structure in mature spores. Noticeably, highly transcribed genes in 48 hr have lower level of Hho1p occupancy (Figure 3-8), suggesting that the lower level of transcription repression in these genes may be, in part, due to less compacted high-order chromatin structure. Looking further into the positioning of linker protein, both conditions showed that Hho1p was enriched in linker DNA region (Figure 3-2) and could be detected in the promoter region as well (Figure 3-5). The binding of Hho1p in RP promoters is closely associated with the Hmo1p binding (Table 3-2). It is not clear how Hho1p is recruited to linker DNA region, except the fact that Hho1p occupancy is weakly correlated with nucleosome occupancy in the coding region (Figure 3-6).

5.1.2 Tetrasome as the intermediate of nucleosome assembly and disassembly

The (H3-H4)$_2$ tetrasome is a stable intermediate product of nucleosome assembly and disassembly product during transcription and DNA replication. Applying the size-selection on H3
ChIPed DNA, I sequenced the H3 bound tetrasome-sized DNA fragment in several conditions. The NFR tetrasomes are very sensitive to MNase digestion (Figure 4-3). Our result also shown that majority of tetrasome dyads did not reside at nucleosome dyads (Figure 4-5 and Figure 4-6), and tetrasomes show CG dinucleotide enrichment at the dyad (Figure 4-7). Such nucleosome independent positioning of tetrasome is confirmed by the tetrasome data of ISW1 mutant, the tetrasome position did not shift upstream in ISW1 mutant strain, but largely retained its original position (Figure 4-9 and 4-10). Therefore, I speculated that tetrasomes reside at thermodynamic favored positions. No positive correlation between tetrasome level and transcription frequency was seen in our data (Figure 4-11), which might be explained by the interplay between tetrasome formation and depletion.

5.2 Future Direction

The thesis mainly focused on characterizing the primary chromatin structure from different perspectives in yeast genome, and it has lead to some interesting observations to follow up.

Although nucleosome patterns are almost identical to each other during sporulation time course, I did observe that at 0 hr during sporulation, nucleosome spacing increases globally, and such pattern has never been reported in previous studies (red trace compared to other traces in Figure 2-2A). It is not known whether such global increase in spacing is involved in transcription regulation. Several experiments can be done to further study such interesting nucleosome reorganization. Firstly, the spacing increase phenotype can be examined in the mutant library from Saccharomyces Genome Deletion Project (http://www-sequence.stanford.edu/group/yeast_deletion_project/deletions3.html) to identify essential
proteins, whose null mutants lose the increased nucleosomal spacing phenotype at 0 hr. Secondly, an RNA-seq approach can be taken to compare the steady state mRNA level between the identified mutants and wild type cells at 0 hr. In addition, the transcription activities can also be measured by the genome-wide nuclear run-on experiments (Core et al. 2008; McKinlay et al. 2011). Thus, it is possible to understand the effect on transcription of such increase in nucleosomal spacing. Similarly, genes that are responsible for restoring the proper nucleosome spacing from 0 hr to 2 hr should also be characterized and analyzed. Thirdly, the chromosome conformation capture experiments can be performed to examine whether high order structure show distinct patterns at the 0 hr time point (Duan et al. 2010).

I also found distinct H3K36me3 mark pattern in rDNA region during sporulation process (Figure2-16), as discussed in Chapter 2. The assumption is that these H3K36me3 marks repress meiotic recombination and facilitate the formation of heterochromatin in rDNA region. Therefore, the first experiment should measure the frequency of meiotic recombination at rDNA region in SET2 mutant or H3K36A mutant strain in yeast. If elevated recombination frequency is observed in these two mutants, it will support the assumption. Then it will be interesting to figure out how Set2p is recruited without active transcription.

Despite the well-studied core histone subunits, linker histone is still poorly understood. It will be important to identify the proteins (such as histone chaperons) that recruit Hho1p into the chromatin. The Hho1p interacting proteins should be identified first, and genome-wide Hho1p mapping in the mutants of these candidate proteins can be performed to identify the mutant that affects Hho1p distribution in the genome. To address further whether histone modifications are involved in recruiting Hho1p, the global recruitment of Hho1p into chromatin and positioning can be examined in the histone point mutant library and tail-deletion strains. Moreover, considering the effect of Hho1p in the formation of high order chromatin structure, the chromosome conformation capture can also be used to study the mutants that show redistribution of Hho1p.
This may shed light on the understanding of the interplay among Hho1p occupancy, high order chromatin compaction and transcriptional regulation.

To confirm that tetrasome positioning is largely determined by DNA sequence and independent from histone remodeler and nucleosome, it may be important to carry out in vitro tetramer histone reconstitution assay to demonstrate that in vivo tetrasome positioning is similar to that reconstituted by the salt gradient dialysis experiments (Zhang et al. 2011b). Adding whole cell extract and ATP should not appreciably affect the tetrasome positioning, otherwise, it may suggest that the histone remodeler in vivo also reposition the tetrasomes.

The ChIP-seq technique studied a pool of cells and obtained the average among the cell population. With the development of single cell genomic technology (Kalisky and Quake 2011), it will be interesting to perform the nucleosome mapping on several individual cells. In addition to primary nucleosome structure, high order chromatin structure plays crucial role in transcription regulation and other cellular events as well (Duan et al. 2010). Therefore, it will be interesting to characterizing the high order structure changes during meiosis.

With the effort from different directions and rapid technology advances, the science community is getting closer to a holist view of chromatin biology, transcription regulation, translation and other areas in life science as well. As the computer science and the Internet have transformed the society in last decade, I believe that the basic research in life science will translate into fruitful applications in various fields of society, bring personal medicine into reality and greatly improve the health care system in the near future.
References


Anderson SF, Steber CM, Esposito RE, Coleman JE. 1995. UME6, a negative regulator of meiosis in *Saccharomyces cerevisiae*, contains a C-terminal Zn2Cys6 binuclear cluster that binds the URS1 DNA sequence in a zinc-dependent manner. *Protein Sci* 4: 1832-1843.


landscape of DNA double-strand breaks in Saccharomyces cerevisiae. PLoS Biol 5:
e324.

Buttner S, Eisenberg T, Carmona-Gutierrez D, Ruli D, Knauer H, Ruckenstuhl C, Sigrist

Cairns BR. 2009. The logic of chromatin architecture and remodelling at promoters. Nature

stabilization induces transcriptional gene silencing via histone deacetylation in S.

298: 1039-1043.

H3 trimethylation at lysine 36 is associated with constitutive and facultative


Chu S, Herskowitz I. 1998. Gametogenesis in yeast is regulated by a transcriptional cascade
dependent on Ndt80. Mol Cell 1: 685-696.

Clancy MJ, Shambaugh ME, Timpte CS, Bokar JA. 2002. Induction of sporulation in
Saccharomyces cerevisiae leads to the formation of N6-methyladenosine in mRNA:


Georgieva M. 2011b. Hho1p, the linker histone of Saccharomyces cerevisiae, is important for the proper chromatin organization in vivo. *Biochim Biophys Acta*.


transcriptional silencing at the ribosomal DNA. *Proc Natl Acad Sci U S A* 105: 11703-11708.


Venkatasubrahmanyam S, Hwang WW, Meneghini MD, Tong AH, Madhani HD. 2007. Genome-wide, as opposed to local, antisilencing is mediated redundantly by the euchromatic factors Set1 and H2A.Z. Proc Natl Acad Sci U S A 104: 16609-16614.


VITA

Liye Zhang

Education:

**PhD in Cell and Developmental Biology**
The Pennsylvania State University, University park, PA
*Dissertation Title*: Chromatin organization of yeast meiosis
*Dissertation Advisors*: Dr. B. Franklin Pugh, Dr. Hong Ma

**BS in Biotechnology**, May 2007
Shanghai Jiaotong University, China
*Dissertation Title*: cis-regulatory element regulation of p63 promoter
*Dissertation Advisor*: Dr. Jingde Zhu, Shanghai Cancer Institute

Publications:

Jessica Bryant, Jenome Govin, Liye Zhang, Greg Donahue, Frank Pugh, Shelley Berger. The linker histone plays a dual role during gametogenesis in S. cerevisiae. (submitted)

Zhang Liye, Hong Ma, Frank Pugh. (2011) Stable and dynamic nucleosome states during a meiotic developmental process. Genome Res. 21: 875-884.


Honors and Awards:

Braddock Research Award (2010)