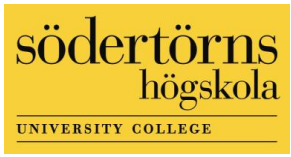
From Department of Biosciences and Medical Nutrition
Karolinska Institutet, Stockholm, Sweden
and
School of Life Science, Södertörns Högskola,
Huddinge, Sweden

Genome-wide patterns of histone modifications in fission yeast

Indranil Sinha



Stockholm 2010

Published by Karolinska Institutet. Printed by Larserics Digital Print AB.

ISBN 978-91-7409-747-4

Dedicated to my parents

বাবা ও মা কে

ABSTRACT

DNA is wrapped almost two times around a group of proteins called histones to form a chromosomal structure known as the nucleosome. Both DNA and histones can be modified with different chemical tags by several enzymes to activate or suppress a particular gene or group of genes. Histones can be covalently modified at several places. Among many different types of post-translational histone modifications, histone acetylation and methylation are two important modification types that are associated with transcriptional activation and repression. Histone acetylation and methylation can be added by histone acetyltransferases (HATs) and histone methyletransferases (HMTs), whereas these modifications can be removed by histone deacetylases (HDACs) and histone demethylases (HDMs). Histone modifications are not only involved in the regulation of gene expression, but also in DNA-based processes, such as replication, repair, and the formation and maintenance of heterochromatin.

Combinations of modified and unmodified states of histones can form distinct histone modification patterns. In many different genome-wide studies, it was observed that a distinctive pattern of histone modification in various organisms is important for gene regulation, DNA replication, chromosome segregation and heterochromatin-mediated silencing. In this thesis, we have conducted several genome-wide investigations to uncover different histone modification patterns and their roles in transcriptional control in fission yeast.

Our analysis of six different HDACs in fission yeast showed that Clr6 and Clr3 are mainly involved in keeping repressed genes silent; Sir2 and Hst2 repress non-expressed genes, and Hst4 acts globally to reduce gene expression, whereas Hos2 is required for the activation of gene expression. By investigating the influence of each HDAC on nucleosome density, we found that all sirtuins and Hos2 enzymes are required to maintain normal nucleosome density and distribution in the *S. pombe* genome.

We have reported that histone acetylation patterns show a 5' to 3' polarity, *i.e.*, the modification levels peak near the ATG and gradually decrease in the coding regions. We also found that histone acetylation patterns depend on gene expression but are independent of gene length. Comparing our data with other published datasets, we observed that different HDAC mutants affect acetylation in different parts of open reading frames (ORFs).

We have demonstrated that histone H4 acetylation proceeds in the direction from K16 to K5, consistent with a 'zip' model that may be involved in transcriptional control. Our analysis revealed antagonistic crosstalk between H3K36me2/me3 and H3K27ac at promoter regions. We observed that histone H3 K18, K27 and K9 acetylation positively correlate with gene expression, and a conserved pattern was also reported in other organisms.

Finally, we report that histone H4K20me1 is strongly linked to active genes, whereas H4K20me3 is associated with weakly expressed genes. Our analysis further shows that H4K20me1 modification levels peak at 3'UTR regions in active genes.

Thus, our analysis revealed many different aspects of histone modification patterns and their roles in transcriptional control in fission yeast.

ISBN 978-91-7409-747-4

LIST OF PUBLICATIONS

I. Wirén M*, Silverstein RA*, **Sinha I***, Walfridsson J, Lee HM, Laurenson P, Pillus L, Robyr D, Grunstein M and Ekwall K. Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. **EMBO J**. 2005 Aug 17;24(16):2906-18. Epub 2005 Aug 4.

II. Durand-Dubief M, **Sinha I**, Fagerström-Billai F, Bonilla C, Wright A, Grunstein M and Ekwall K. Specific functions for the fission yeast Sirtuins Hst2 and Hst4 in gene regulation and retrotransposon silencing. **EMBO J**. 2007 May 16;26(10):2477-88. Epub 2007 Apr 19.

III. **Sinha I**, Wirén M and Ekwall K. Genome-wide patterns of histone modifications in fission yeast. **Chromosome Res**. 2006;14(1):95-105.

IV. **Sinha I***, Buchanan L*, Bonilla C, Rönnerblad M, Shevchenko A, Grunstein M, Stewart F and Ekwall K. Genome wide mapping of histone modifications and mass spectrometry reveal a function for the histone H4 acetylation zip and a role for H3K36 methylation at gene promoters in fission yeast. Manuscript.

V. **Sinha I***, Jemt E*, Durand-Dubief M, Strålfors A, Sanders S and Ekwall K. Genome wide mapping suggests different roles for H4K20me1 and H4K20me3 in gene expression. Manuscript.

* These authors contributed equally to this work.

OTHER PUBLICATIONS

I. Zhu X, Wirén M, **Sinha I**, Rasmussen NN, Linder T, Holmberg S, Ekwall K and Gustafsson CM. Genome-wide occupancy profile of mediator and the Srb8-11 module reveals interactions with coding regions. Mol Cell. 2006 Apr 21;22(2):169-78.

II. Djupedal I, Durand-Dubief M, **Sinha I** and Ekwall K. Similarities and differences of Genome –associated RNA polymerase II and subunits Rpb4/7 of Fission yeast. Submitted to JBC.

Table of Contents

LIST OF ABBREVIATIONS

3'UTR	3’untranslated region
5'UTR	5’untranslated region
ANNOVA	analysis of variance
ATP	Adenosine triphosphate: the main energy storage and transfer molecule in the cell
Bbd	bar body deficient
Bp	base pare
BPMAP	Affymetrix sequence information file
cDNA	complementary DNA
CENP-A	Centromere protein A
CHD	Chromo-helicase/ATPase DNA binding
ChIP	chromatin immunoprecipitation
Clr	Cryptic Loci Regulator
DNA	Deoxyribonucleic Acid
DNMT	DNA methyltransferase
Dot	Disruptor of telomeric silencing
dsRNA	Double-stranded RNA
Gcn5	General control non-derepressible 5
GCOS	GeneChip® Operating Software from Affymetrix
GO	Gene Ontology
H2A	Histone H2A
H2B	Histone H2B
H3	Histone H3
H4	Histone H4
HAT	Histone acetyltransferases
Hda1	histone deacetylase 1
HDAC	histone deacetylase
HDM	histone demethylase
Hos2	Hda one similar 2
HP1	Heterochromatin protein 1
Hrp	Helicase related protein
HTML	Hyper Text Markup Language
IGB	Integrated Genomic Browser
IGR	Intergenic Region
ISWI	Imitation Switch
JmjC	Jumonji C-terminal domain
KMT	K(Lysine)-methyletransferases
MAT	Model-based Analysis of Tiling-array
miRNA	micro RNA
MLL	Mixed lineage leukemia
mRNA	messenger RNA
MS	Microsoft
MYST	MOZ-Ybf2/Sas3-Sas2-Tip60

NAD	Nicotinamide adenine dinucleotide
ODBC	Open Database connectivity
ORF	Open Reading Frame
PCA	principle component analysis
Pol	Polymerase
RDBMS	relational database management system
RNA	Ribonucleic acid
RNAi	RNA interference
rRNA	Ribosomal RNA
Sir2	Silent information regulator 2
siRNA	Short interfering RNA
SNF	Sucrose non fermenting
snRNA	Small nuclear RNA
snoRNA	Small nucleolar RNA
Su(var)	Suppressor of variegation
TAS	Tiling analysis software
Tip	Tat interactive protein
tRNA	transfer RNA
TSS	transcription start site

1. INTRODUCTION

DNA contains all the genetic instructions necessary to build the cells of an organism. DNA is wrapped around proteins known as histones to form the basic structure of a chromosome, called the nucleosome. Both DNA and histones can be modified with different chemical tags by several enzymes to activate or suppress a particular gene or a group of genes, which forms a second layer structure known as the Epigenome. In general, the genetic code in DNA remains fixed for life, but the Epigenome is variable for different cell types. Different signals from the outside world, such as environmental factors, diet, physical activity, stress, and exposure to toxins, can change our Epigenome. Both the histone and DNA modifications can be heritable. The study of these heritable modifications in gene functions or activity without changes in DNA sequences is defined as Epigenetics.

Histone modifications can regulate many different processes, such as gene silencing (Braunstein et al., 1993; Johnson et al., 1990; Rea et al., 2000), gene activity (Cheung et al., 2000; Martin and Zhang, 2005), DNA replication (Vogelauer et al., 2002), DNA damage (van Attikum and Gasser, 2005), and apoptosis (Ahn et al., 2005). The combination of several histone modifications can generate histone modification patterns that may lead to distinct biological outcomes. For example, the mapping of global histone acetylation patterns compared to gene expression revealed that several histone modification patterns were associated with groups of biologically related genes and gene activity in budding yeast (Kurdistani et al., 2004).

So far, we have conducted several genome-wide investigations to uncover different roles for fission yeast HDACs (histone deacetylases), including their roles in histone deacetylation and the regulation of gene expression described in papers I and II. The third study (paper III) focuses on the 5' to 3' distribution of different histone modifications and modification patterns, especially histone acetylation and methylation patterns over the average gene. In the fourth paper (paper IV), we used mass spectrometry analysis combined with the ChIP-chip method, which revealed a function for the histone H4 acetylation zip and a role for H3K36 methylation at gene promoters in *S. pombe*. Quantitative mass spectrometry (MS) analysis and ChIP-chip data also revealed crosstalk between histone acetylation and histone methylation. Finally, in the fifth paper (paper V), we examined different roles for histone H4K20 –

mono, –di, and –tri methylation in gene expression using a high-resolution tiling microarray in our genome-wide mapping study.

The section below will cover a general discussion regarding transcription, chromatin structure, epigenetic mechanisms and histone modification. Some details of the different methods used during this study will be presented, followed by the results and a discussion section.

2. TRANSCRIPTION AND CHROMATIN STRUCTURE

Deoxyribonucleic acid, or DNA, is the cellular library that contains all the genetic instructions required to build the cells of an organism. For the genetic continuity of a species, it is important to duplicate exact information from generation to generation. Transcription is the process by which genetic information stored in DNA is replicated into ribonucleic acid (RNA). The enzymes responsible for transcribing DNA into RNA are known as RNA polymerases. There are three different types of RNA polymerases: RNA pol I, II, and III, which can be found in all eukaryotes. RNA pol II is the enzyme responsible for transcribing protein-coding genes. To begin the transcription process, the RNA polymerase first locates an appropriate initiation site, also known as transcription start site (TSS), on duplex DNA. This is followed by the binding of the DNA, the separation of the double-stranded DNA and the generation of a new RNA strand. As the polymerase progresses along the DNA strand, it continues to separate the DNA duplex and generate the new RNA strand. It finally terminates the transcriptional process when it reaches the end of the protein-coding DNA template.

Chromosomes are made up of a complex known as chromatin, which is a combination of DNA and protein. Chromatin consists of a fundamental repeating unit, also known as the nucleosome. The major components of a nucleosome are DNA and histone proteins. Nucleosomes pack DNA into a stable, coiled form. Many nucleosomes together produce a flexible jointed chain, which is also known as chromatin fiber.

Chromatin is divided into two structurally and functionally distinct regions known as euchromatin and heterochromatin. However, other distinct genomic regions known as centromeres, Hda1-affected subtelomeric (HAST) domains, and Htz1-activated domains (HZADs) are also present in chromatin (Millar and Grunstein, 2006).

Several different enzymes can post-translationally modify histone proteins. From numerous studies, it has been established that these modifications affect gene transcription by activating or preventing transcription initiation and elongation (John et al., 2000; Krogan et al., 2003; Mizzen et al., 1996; Ng et al., 2003; Perez-Martin and Johnson, 1998; Winkler et al., 2002). Section 4.2 includes a detailed discussion of the role of chromatin modification in transcriptional control.

2.1 THE NUCLEOSOME

Nucleosomes are the basic unit of eukaryotic chromosome structures. In 1974, Roger Kornberg proposed the structure of the nucleosome. Four of the five canonical histones (H2A, H2B, H3, and H4), also known as core histones, form an eight-subunit structure. 146 base pairs of DNA are wrapped almost two times around this octameric structure to form a nucleosome (Luger et al., 1997) (Figure 1). The histone H1, also known as a linker histone, stabilizes the octameric core structure in higher-order structures (Luger et al., 1997). However, histone H1 is absent in fission yeast (Kobori et al., 2003). Many nucleosomes form a flexible joined chain, which can be observed under an electron microscope as a 'beads on a string' structure; nucleosomes are the beads and the DNA linking them is the string (Olins and Olins, 1974). The four core histones are relatively similar in structure and conserved through evolution. Amino acid sequences of the histone H3 and H4 are highly conserved in eukaryotes, whereas histones H2A and H2B are less conserved.

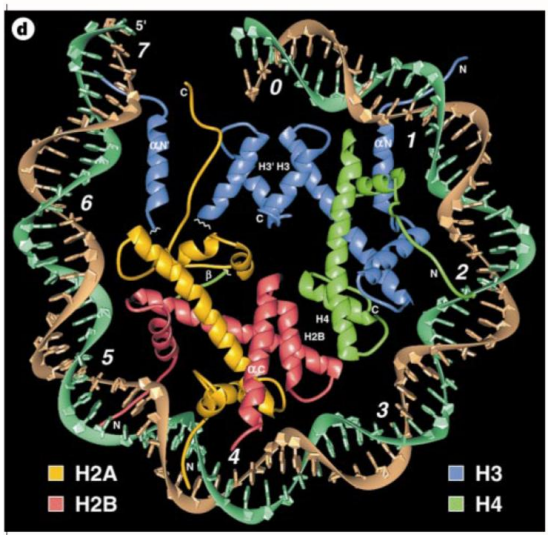


Figure 1. The nucleosome core particle: ribbon traces for the DNA phosphodiester backbones (brown and turquoise) and eight histone protein main chains (blue: H3; green: H4; yellow: H2A; red: H2B) are shown. Reprinted with permission from Luger et al., Nature, 1997. (Luger et al., 1997)

2.2 HETEROCHROMATIN AND EUCHROMATIN

Heterochromatin is the condensed form of chromatin, which is in general genetically inactive and remains relatively condensed throughout the cell cycle. A block of heterochromatin can control the gene expression and switch off a specific gene, group of genes, or even a whole chromosome (Turner, 2001). Generally, heterochromatin is found at the telomeres and domains near the centromeres in many different organisms, whereas the silent mating type loci in yeast and silent X chromosomes in female mammals are species-specific heterochromatic regions. There are two major subtypes of heterochromatin: constitutive and facultative. Chromosomal regions, such as centromeres and telomeres, are referred to as constitutive heterochromatins that contain a high density of repetitive DNA elements and transposable elements. Facultative heterochromatin refers to developmentally-regulated regions where chromatin states can be changed into a silenced state through a mechanism such as histone methylation, or through RNAi. Spreading of heterochromatin across domains causes repression of neighboring sequences, referred to as silencing.

The non-heterochromatic portion of the genome that includes both repressed and active genes is known as euchromatin. Euchromatin is less condensed than heterochromatin, and hence is often, but not always actively transcribed. Distinct histone modification patterns can be observed in heterochromatic and euchromatic regions (see section 4.2).

3. EPIGENETICS AND ITS MECHANISMS

Epigenetics describes the heritable modifications in gene functions or activity without changes in DNA sequence. In Greek, *Epi-* means *over* or *above*; this term literally refers to events that are over or above genetics. The history of epigenetics is connected with the study of evolution and development and was first defined by C. H. Waddington in 1953 (Waddington, 1953). However, over the past 50 years, research in this field has increased our understanding of the fundamental molecular mechanisms behind the regulation of gene expression (Felsenfeld, 2007). Robin Holliday proposed the present operational definition of epigenetics as "Nuclear inheritance, which is not based on differences in DNA sequence" (Holliday, 1994). Epigenetics is currently an emerging field. A PubMed search using "Epigenetics" as a keyword returns 32,992 entries as of November 19, 2009, certifying the popularity of the subject. Based on numerous studies, it is apparent that nucleosomes, the basic unit of chromatin, can transmit epigenetic information from one cell generation to the next.

3.1 DNA METHYLATION

DNA methylation was among the first epigenetic mechanisms discovered. DNA methylation can occur by adding a methyl group at the 5 position of the cytosine pyrimidine ring. It commonly occurs in CpG dinucleotides. In an early study, it was reported that 70% of all CpG sites were methylated in mammalian somatic tissues (Ehrlich et al., 1982). A sequence with a large no of CpG sites than expected is termed as CpG island. CpG islands are found upstream of many mammalian genes. In the early embryo the DNA methylation pattern is removed and then again restored at the time of implantation in mouse (Kafri et al., 1992; Monk et al., 1987). DNA methylation does not alter DNA sequences, but it contributes to gene silencing, X-chromosome inactivation, and genomic imprinting. It represents a mechanism of cell memory (Holliday and Pugh, 1975; Riggs, 1975). Tissue-specific gene expression patterns during cellular development can be maintained by DNA methylation, and in the absence of DNA methylation, the frequency of chromosome loss may increase (En Li, 2007). DNA methylation patterns are heritable through multiple cell divisions that allow a stable epigenetic marking of the genome (En Li,

2007). DNA methylation can occur by means of several different enzymes known as DNA methyltransferases (DNMTs).

Histone modification and DNA methylation pathways are dependent on one another. Histone modification has a role in establishing the DNA methylation pattern whereas DNA methylation is important for maintaining histone modification patterns (Cedar and Bergman, 2009). Recent studies suggests that during early development histone H3K4 methylation patterns might form at CpG islands across genome before *de novo* DNA methylation (Ooi et al., 2007). Next, during DNA methylation, DNMTs may recognize H3K4 methylated CpG islands and thus prevent DNA methylation at that site whereas it takes place at the majority of other CpG sites in the genome (Cedar and Bergman, 2009; Ooi et al., 2007). The anti-correlation between DNA methylation and H3K4 methylation in many cell types (Meissner et al., 2008; Mohn et al., 2008; Okitsu and Hsieh, 2007; Weber et al., 2007) may confirm the above-mentioned model. Using chromatin immunoprecipitation, it was observed that the acetylated histones were mainly assembled with unmethylated DNA (Eden et al., 1998; Hashimshony et al., 2003), which confirms the importance of DNA methylation in maintaining histone modification patterns.

Growing evidence has emerged that aberrant DNA methylation can lead to cancer (En Li, 2007). One DNMT known as Dnmt3B leads to immune deficiency in humans. A lack of the methyl-CpG-binding protein MeCP2 causes Rett syndrome (Amir et al., 1999). Hence, errors in DNA methylation contribute to human diseases.

3.2 HISTONE TAIL MODIFICATIONS

Two classes of proteins, histones and non-histones, bind to DNA to form eukaryotic chromosomes. Histones are the principle component of chromatin. They are involved in regulating processes such as replication, repair, and transcription in eukaryotes. A large amount of histone proteins can be found in eukaryotic cells. The weight of histones in the cell nucleus is approximately equal to the weight of DNA (Alberts and National Center for Biotechnology Information (U.S.), 2002).

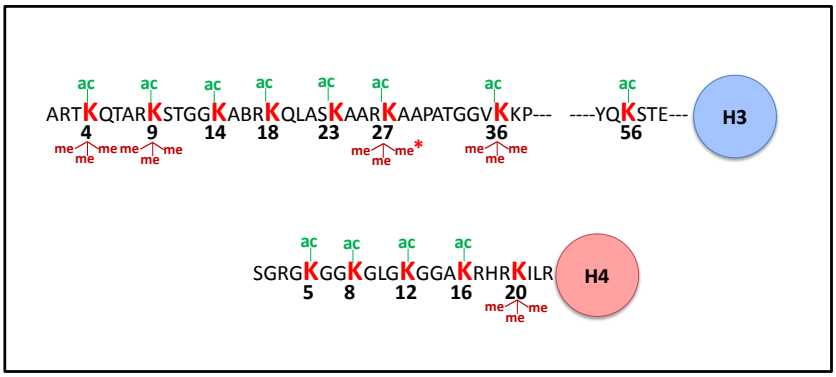


*Figure 2. Sites of acetylation and methylation on histones H3, and H4 in fission yeast. (*H3K27 methylation not reported in fission yeast)*

Each of the core histones contains N-terminal and C-terminal tails, which project from the core of the nucleosomes. N-terminal tails are made up of several amino acids that are highly conserved through evolution from yeast to human. Histone tails can be covalently modified at several places (Figure 2). N-terminal tails can be post-translationally modified by a variety of enzymes to be acetylated, deacetylated, methylated, phosporylated (Iizuka and Smith, 2003), ubiquitinated and sumoylated (Gill, 2004). These modifications are a good source of epigenetic information, which will be further discussed in section 4.

3.3 CHROMATIN REMODELING

In addition to the posttranslational covalent modifications of nucleosomes that were discussed in the previous section, nucleosomes can also be modified in non-covalent ways. Several different enzymes can modify nucleosomes and are known as chromatin remodeling factors. Chromatin remodeling factors modify nucleosomes in an ATP-dependent manner. Remodeling factors are highly conserved in eukaryotes but also found in bacteria and archaea (Flaus et al., 2006). These enzymes, termed SNF2 ATP-dependent chromatin remodeling factors, are divided into distinct subfamilies of proteins. Different helicase-related domains and additional specific domains distinguish subfamilies of remodeling factors. For example, the SNF2 subfamily contains bromodomains, CHD remodelers contain two chromodomains, the Mi-2 subfamily has PHD and chromodomains, the ISWI subfamily has a SANT domain, and the Rad5/16 family has a ring finger domain (Flaus et al., 2006).

Chromatin remodeling families are involved in many different biological processes, such as chromatin formation (Alen et al., 2002; Tsukiyama et al., 1999), transcriptional regulation (Alen et al., 2002) and centromere function (Walfridsson et al., 2005). It was reported that the ISWI complex is involved in the repression of transcription of early meiotic genes during mitotic growth in fission yeast (Goldmark et al., 2000). ISWI ATPases are also involved in chromatin assembly during replication in mammals. In the absence of the ISWI complex, a structural defect on the male X-chromosome was observed in *Drosophila* (Deuring et al., 2000). SNF2 family members are involved in transcriptional gene activation in Drosophila (Armstrong et al., 2002), although these members are associated with both transcriptional activation and repression in budding yeast (Ng et al., 2002). The Ino80 ATPase family protein has a role in DNA damage repair and transcription in *S. cerevisiae* (Shen et al., 2000). CHD modeling factors are involved in transcriptional termination in *S. pombe* and *S. cerevisiae* (Alen et al., 2002) . Hrp1 and Hrp3 CHD modeling factors in *S. pombe* have roles in regulating centromere assembly and function (Walfridsson et al., 2005; Walfridsson et al., 2007).

Thus, chromatin remodeling factors have various roles in many different biological processes, and these proteins influence chromatin organization.

3.4 HISTONE VARIANTS

Histones are the basic unit of nucleosomes. However, variant genes encode histones with different nucleotide sequences known as histone variants. There are four different variants of histone H3: CENP-A, H3.1, H3.2 and H3.3. There are four different H2A variants: H2A.X, H2A.Z, macroH2A and H2A-bar-body-deficient ($H2A^{BBD}$). However, no H4 variants have been reported so far. Most histone variants are conserved over evolution.

Histone H3 variant CENP-A localizes exclusively to centromeres (Palmer et al., 1991). Even though all the different core histones were lost during spermatogenesis, CENP-A remained stable at centromeres (Palmer et al., 1991). CENP-A is a conserved histone variant found in all eukaryotes (Malik and Henikoff, 2003). Two distinct chromatin structures can be found in fission yeast centromeres, CENP-A containing kinetochore chromatin and heterochromatin. In fission yeast, heterochromatic features can be displayed at pericentric regions by H3-containing nucleosomes whereas a group of CENP-A proteins occupy the central core regions

(Amor et al., 2004). Remarkably, a lack of DNA sequence specificity can be observed in centromeric nucleosomes (Henikoff, 2007), although their function remains constant. For example, although budding yeast centromeres are very small (125-bp sequences) and contain the Cse4 centromeric variant, the homolog of Cse4 can structurally and functionally replace human CENP-A (Wieland et al., 2004).

The histone variant H3.3 is similar to histone H3, differing by only four amino acids. H3.3 can be found at transcriptionally-active chromatin (Henikoff and Ahmad, 2005). Histone modifications such as H3 K9, K18, and K23 acetylation and H3K4 and K79 methylation, known to be enriched on active chromatin, are also enriched on H3.3 (Henikoff, 2007; Sarma and Reinberg, 2005).

The H2A histone variant H2A.X is a conserved histone variant found in many eukaryotes (Henikoff, 2007). H2A.X is involved in DNA double-strand break repair. When a double-strand break occurs during the DNA replication process, H2A.X becomes phosphorylated and forms a large chromatin domain near the break. After the ds break repair, the phosporylation mark of H2A.X eventually disappears (Henikoff, 2007).

In addition to H2A.X another H2A variant, H2A.Z is an essential histone in many organisms such as ciliated protozoans and mammals. However, a mutation of this gene can produce viable cells in fission and budding yeasts (Henikoff, 2007). H2A.Z is involved in regulation of gene expression (Kamakaka and Biggins, 2005), and also plays a role as an antisilencing factor (Henikoff, 2007). Studies in *Drosophila* and budding yeast have revealed that H2A.Z is widely distributed along chromosomes (Leach et al., 2000).

MacroH2A and $H2A^{Bbd}$ (H2A Barr body deficient) are two other vertebrate-specific histone H2A variants involved in gene transcription. $H2A^{Bbd}$ –containing nucleosomes play a role during active transcription, whereas macroH2A-containing nucleosomes are less mobile and resist active transcription.

Thus, histone variants are involved in gene transcription, and the most fundamental level of differentiation of chromatin can be provided by histone variants.

3.5 RNA INTERFERENCE

RNA interference (RNAi) is an RNA-dependent gene silencing process. RNAi can regulate endogenous gene expression. microRNA (miRNA) and small interfering RNA (siRNA) are two types of small RNA molecules fundamental to

RNAi. RNA interference was first described in *C. elegans* (Fire et al., 1998). The RNAi mechanism initiates in cells when they encounter a double stranded RNA (dsRNA) molecule, either by external insertion or from an endogenous source. In this mechanism, double stranded RNA molecules can be recognized and cleaved into small 21-25 bp fragments, with a few unpaired overhang bases on each end, by an enzyme called dicer. These short double stranded fragments are known as small interfering RNAs (siRNAs). These siRNAs can form a ribo nucleic protein complex called RISC (RNAi Silencing Complex). RISC first unwinds the siRNA duplex into single stranded siRNA and binds to the target mRNA in a sequence-specific manner, which cleaves the target mRNA with the catalytic component of the RISC complex, Argonaute. The cleaved mRNA is aberrant for cells and can be degraded, thus preventing the translation of the target mRNA.

The RNA interference pathway works as a defense mechanism inside a cell that can prevent unwanted genes from being expressed. The RNAi pathway is found in many organisms, such as animals, plants and single cellular eukaryotes like *S. pombe*. The formation of heterochromatin at centromeres, telomeres and mating-type locus can be mediated by the RNAi machinery in *S. pombe* (Hall et al., 2002). The heterochromatin assembly mechanism mediated by RNAi machinery is also conserved in plants, *Drosophila, C. elegans* and mammals (Fukagawa et al., 2004; Grishok et al., 2005; Kanellopoulou et al., 2005; Pal-Bhadra et al., 2004; Zilberman et al., 2003).

4. HISTONE MODIFICATION

Histone tails are normally positively charged, which helps them interact with negatively-charged DNA. Histone acetylation neutralizes the positive charges of histones, which decreases the histone-DNA interaction and allows chromatin expansion. On the other hand, histone methylation does not neutralize the positive charge of histones (Figure 3). Although many different types of post-translational histone modifications are described in numerous articles, this discussion is restricted to histone acetylation and methylation.

4.1 HISTONE MODIFYING ENZYMES

Histone acetylation and methylation can be added by histone acetyltransferases (HATs) and histone methyletransferases (HMTs), whereas these modifications can be removed by histone deacetylases (HDACs) and histone demethylases (HDMs). According to the new nomenclature, HATs are known as KATs (K-acetyltransferases), HMTs are known as KMTs and HDMs are known as KDMs (Allis et al., 2007). However, arginine methyltransferases are represented by HMTs but not by KMTs. Similarly, arginine demethylases are symbolized by HDMs but not by KDMs. Lysine acetylation can occur by adding single acetyl groups to the ε-amino group, whereas mono-, di- and trimethylation occur by adding up to three methyl groups to the lysine residue (Figure 3). Acetylation can change the histone-DNA and nucleosome-nucleosome interactions by neutralizing the charge of lysine (Zheng and Hayes, 2003), whereas methylation does not alter the charge. However, both modifications can create binding sites for chromosomal proteins (Millar and Grunstein, 2006). Histone modification is not only involved in the regulation of gene expression, but also in DNA-based processes such as replication, repair, and the formation and maintenance of heterochromatin.

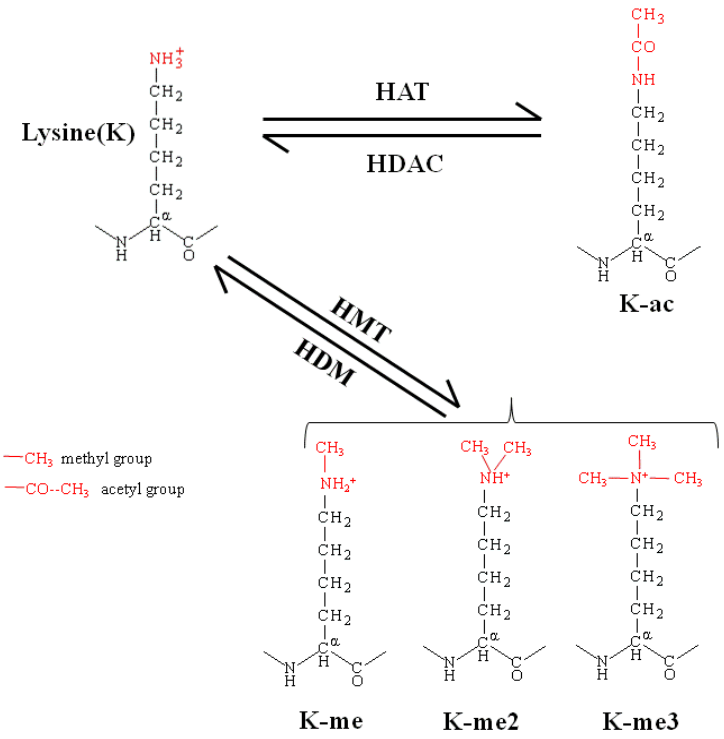


Figure 3. Histone lysine acetylation, deacetylation, methylation and demethylation.

4.1.1 HATs

Histone acetyltransferases, also known as HATs, are enzymes that acetylate lysine amino acids on histone proteins. Most HATs act on several sites in different histones. HATs can be grouped into two families based on their catalytic domains: the GNAT family (Gcn5-related N-AcetylTransferase) and the MYST family, although another protein, p300/CBP (CREB-binding protein), also shows HAT activity. Gcn5^{KAT2} is the founding member of the GNAT family, whereas Morf^{KAT6B}, Ybf2 (Sas3^{KAT6}), Sas2^{KAT8} and Tip60^{KAT5} are founding members of the MYST family (Kimura et al., 2005; Lee and Workman, 2007). Gcn5^{KAT2}, one of the enzymes from the HAT family (Kuo et al., 1998), acetylates histone H3 and H2B sites (Suka et al., 2001), whereas another HAT Esa1^{KAT5} preferentially acetylates histone H4 (Allard et al., 1999), H2A, and H2B sites (Suka et al., 2001). HATs are also responsible for

their specialized roles in chromosome decondensation, DNA-damage repair and the modification of non-histone substrates (Lee and Workman, 2007).

4.1.2 HDACs

Histone deacetylases, also known as HDACs, are divided into three main classes (Figure 1). Class I contains Rpd3- and Hos2-like enzymes, class II contains Hda1-like enzymes and class III contains NAD+ dependent enzymes like Sir2, also known as sirtuins. Although HDACs act on several sites, these enzymes have distinct site preferences. For example, in *S. cerevisiae*, the deacetylation of histone H4 at lysines 5, 8 and 12 require HDACs like Rpd3 and Hos2 *in vivo*, whereas for histone H4 at lysine 16, the deacetylation only requires Hos2 (Wang et al., 2002).

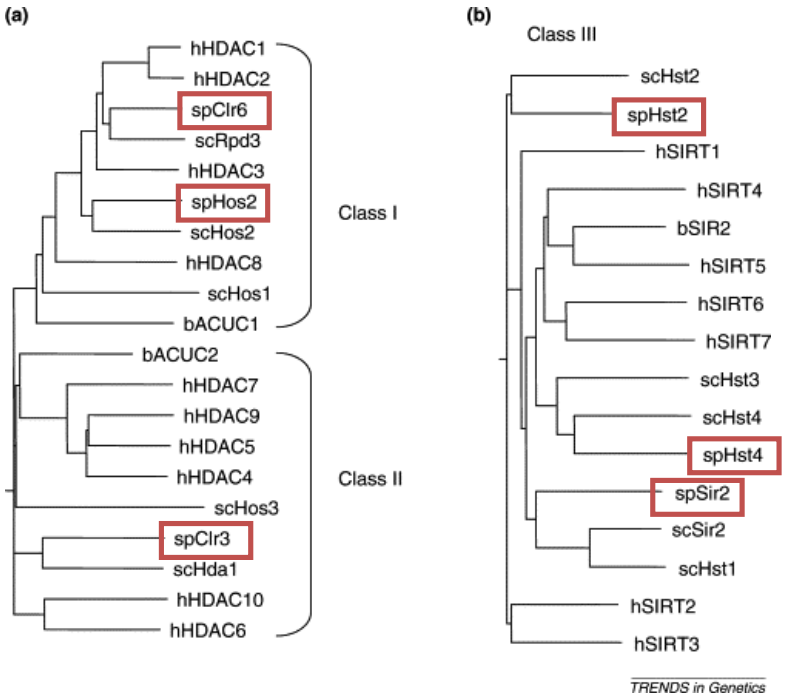


Figure 4. Different classes of HDACs. Reprinted and slightly modified with permission from Ekwall et al., Trends in Genetics, 2005 (Ekwall, 2005).

In total, there are six HDACs present in fission yeast (Figure 4). Clr6 and Hos2 belong to class I, Clr3 belongs to class II, and Sir2, Hst2 and Hst4 belong to

class III. Historically, histone deacetylation has been associated with gene repression. For example, the deacetylation of histones H3 and H2B by histone deacetylase 1 (Hda1) is caused by the general repressor Tup1 (Wu et al., 2001). In *S. pombe*, we observed that the class III HDACs Hst2 and Hst4 act as repressors of gene expression (Durand-Dubief et al., 2007). It is also clear that deacetylation is important for gene activity. Hos2 deacetylates histone H4K16 during gene activity in *S. cerevisiae* (Wang et al., 2002). Interestingly, we also found that the Hos2 enzyme is responsible for removing histone H4K16 acetylation at highly active genes in *S. pombe* (Wiren et al., 2005).

HDACs are also recruited in a global manner. HDACs such as Rpd3 and Hda1 function globally in *S. cerevisiae*. Both of them deacetylate large regions in euchromatin, where they are not targeted by DNA-binding proteins (Vogelauer et al., 2000). Rpd3 also represses genes which are involved in a number of biological pathways, such as meiosis, metabolism, and transport (Shahbazian and Grunstein, 2007).

Clr3 and Sir2 enzymes are required for efficient silencing of rDNA, centromeres, telomeres, and the mating-type region (Bjerling et al., 2002; Freeman-Cook et al., 2005; Shankaranarayana et al., 2003). These enzymes act jointly at all four of the silent regions (Wiren et al., 2005). Mutation of *clr3* and *clr6* reduces the silencing (Ekwall, 2005), whereas the *hos2* mutation increases silencing at centromeres, telomeres, and mating-type regions (Olsson et al., 1998).

Recently, it was reported that a member of the sirtuin family, SIRT6, linked histone H3K9 deacetylation to NF-κB-dependent gene expression in mammals (Kawahara et al., 2009). In another observation, it was reported that the inhibition of SIRT1 (another member of the mammalian class III HDAC family) in breast and colon cancer cells increased H4K16 and H3K9 acetylation at endogenous promoters and gene activation (Pruitt et al., 2006).

HDACs are evolutionarily-ancient enzymes that are well conserved (Ekwall, 2005). All three classes of HDACs constitute important potential drug targets (Kramer et al., 2001). Class I and II enzymes have been linked to cancer (Lin et al., 1998; Yang et al., 2005), whereas the class III enzyme Sir2 has been implicated in cancer and aging (Chen and Guarente, 2007; Pruitt et al., 2006). As HDACs are structurally well-conserved, their basic functions can be studied in the less-complex

yeast model organism. In papers I and II, we have successfully analyzed the genome-wide roles of all six HDACs present in fission yeast, including the *in vivo* enzymatic specificity. It remains to be seen to what extent the HDACs functions are conserved between human and yeast.

4.1.3 HMTs

Histones can be methylated either on their lysine or arginine residues (DeLange et al., 1969; Murray, 1964). SET domain or non-SET domain containing lysine histone methyltransferases ($HMTs^{KMTs}$) are responsible for methylation of histone lysine residues. There are three forms of lysine methylation: mono-, di-, and trimethylation. Arginine residues of histones can only be mono- or dimethylated. Histone trimethylation is much more stable than mono- or dimethylation.

$HMTs^{KMTs}$ enzymes are divided into three classes: (a) $HMTs^{KMTs}$ containing SET domains that are involved in histone H3 lysines 4, 9, 27, and 36 methylation, and histone H4 lysine 20 methylation; (b) $HMTs^{KMTs}$ containing non-SET domains are involved in histone H3 lysine 79 methylation; and (c) arginine methyltransferases involved in histone H3 arginine 2, 17, and 26 methylation, and in histone H4 arginine 3 methylation (Table 1).

Type of HMTs	***Sites***
Containing SET domains	*H3 K(4, 9, 27, 36)*
	H4 K20
Containing non-SET domain	*H3 K79*
Arginine methyltransferases	*H3 R(2, 17, 26)*
	H4 R3

Table 1: Different classes of HMTs.

Site specificity	***S. cerevisiae***	***S. pombe***	***D. melanogaster***	***Human***
H3K4	*Set1*^{KMT2}	*Set1* ^{KMT2}	*Trx* ^{KMT2A}, *Ash1*^{KMT2H}	*MLL1-5*$^{KMT2A-E}$, *Set7/9* ^{KMT7}, *ASH1*^{KMT2H}, *hSET1A*^{KMT2F}
H3K9	-	*Clr4*^{KMT1}	*Su(var)3-9* ^{KMT1}	*SUV39H1* ^{KMT1A}, *SUV39H2* ^{KMT1B}, *SETDB1* ^{KMT1E}, *RIZ1*^{KMT8}
H3K27	-	-	*E(Z)* ^{KMT6}	*EZH2*^{KMT6}
H3K36	*Set2* ^{KMT3}	*Set2* ^{KMT3}	-	*SET2* ^{KMT3A}, *NSD1*^{KMT3B}, *SYMD2* ^{KMT3C}
H3K79	*Dot1* ^{KMT4}	-	-	*DOT1L* ^{KMT4}
H4K20	-	*Set9* ^{KMT5}	*PR-Set7*^{KMT5A}, *SUV4-20* ^{KMT5B}	*PR-SET7/8* ^{KMT5A}, *SUV4-20H1* ^{KMT5B}

Table 2: Site specific HMTs^{KMTs} *in different organisms (Allis et al., 2007).*

Histone methylations are linked to either gene activation or repression. Histone H3K4me2/me3, H3K27me1, and H3K36me2/me3 are enriched in euchromatic regions and play a role in transcriptional activation, whereas histone H3K9me2/me3 and H4K20me3 are enriched in heterochromatic regions and are responsible for gene repression (Agger et al., 2008; Kouzarides, 2007). However, a recent study indicates that H4K20me3 does not show any association with silent promoters in human cells, nor does it correlate with gene repression (Barski et al., 2007). In fission yeast we also observed no H4K20me3 enriched at pericentric regions, although the modification mark was linked to gene repression (Sinha et al., 2010).

Histone H3K4 methylation *in vivo* occurs in three states: mono-, di- and trimethyl. In budding yeast, Set1^{KMT2} mediates histone H3K4 methylation. Yeast cells that lack *set1* are slow in growth and have defects in telomeres and rDNA silencing (Fingerman et al., 2005). The deletion of *set1* eliminates histone H3K4 methylation in fission yeast (Noma and Grewal, 2002). In budding yeast, histone H3K4

trimethylation is required for proper silencing, whereas mono- and dimethylation are required for cell growth (Fingerman et al., 2005). However, in another study it was reported that histone dimethylation occurs at inactive and active euchromatic genes, whereas trimethylation only occurs at active genes in *S. cerevisiae* (Santos-Rosa et al., 2002). In humans, the MLL (Mixed lineage leukemia) SET domain protein is a histone H3K4-specific methyltransferase whose activity is associated with *Hox* (homeobox) gene activation (Milne et al., 2002). The expression of stress-responsive genes increases as a result of the loss of monomethylation in histone H3K4 in *S. cerevisiae* (Seol et al., 2006).

Histone H3K9 methylation is involved in the initiation and maintenance of heterochromatic silencing (Nakayama et al., 2001). The histone H3K9me3 mark is in most cases negatively correlated with transcription. HP1, a chromodomain containing protein (Lachner et al., 2001), can specifically recognize histone H3K9 methylation (Jacobs and Khorasanizadeh, 2002). This recognition by HP1 is required for the establishment and maintenance of heterochromatin (Shilatifard, 2006). In fission yeast, Clr4^{KMT1} is responsible for methylation of H3K9 at heterochromatin-associated regions (Nakayama et al., 2001). The localization of *Drosophila* HP1 homolog Swi6 at heterochromatic regions is functionally dependent on H3K9 methylation in fission yeast.

Histone H3K36 methylation is conserved in eukaryotes. Set2^{KMT3}, which is a transcriptional repression factor in budding yeast, shows HMT^{KMT} activity. The HMT activity of Set2 ^{KMT3} is specific for histone H3K36 methylation, and both play a role in the repression of gene transcription (Strahl et al., 2002). The methylation of histone H3K36 by Set2^{KMT3} also influences the transcriptional elongation in *S. cerevisiae* (Schaft et al., 2003). In *Arabidopsis thaliana*, the loss of the function of SET domains (a homolog of yeast Set2) reduces the dimethylation of histone H3K36 (Zhao et al., 2005). The deletion of the *S. pombe set2* gene results in complete elimination of histone H3K36 methylation (Morris et al., 2005). K36 methylation is mediated by Set2^{KMT3}, which recruits Rpd3S; a mechanism that is required to suppress the initiation of cryptic transcription in coding regions in budding yeast (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). In various histone modification studies, it was observed that di- and trimethylation of H3K36 are enriched in the coding regions in chickens, *Drosophila,* and *Arabidopsis* (Bannister et al., 2005; Bell et al., 2007; Xu et al., 2008).

Histone H4K20 methylation occurs in a variety of organisms, from humans to fission yeast. It can be detected in fission yeast, but it is absent in budding yeast (Schotta et al., 2004). Set8^{KMT5A} specifically methylates histone H4K20 in humans (Fang et al., 2002). In fission yeast, Set9^{KMT5} is responsible for histone H4K20 methylation (Sanders et al., 2004). The methylation of histone H4K20 has been described as a marker of heterochromatin and gene silencing in humans (Nishioka et al., 2002; Schotta et al., 2004). Surprisingly, fission yeast histone H4K20 methylation does not have any heterochromatin function, whereas the modification has a role in responses to DNA damage (Sanders et al., 2004). A recent study also suggests that the loss of histone H4K20 trimethylation is a common hallmark of human cancer (Fraga et al., 2005).

Thus, different histone methylations have different roles. Histone methylation is involved in various activities, such as cell growth, DNA damage response, the initiation of heterochromatin silencing, the repression of gene transcription and human cancer.

4.1.4 HDMs

Histone methylation is stable inside the cell, and until recently, there was no known histone lysine demethylase responsible for removing methyl groups from histones. In 2004, Shi et al. reported that LSD1^{KDM1}, one of the HDMs responsible for removing mono- and dimethyl groups from histone H3K4, was involved in gene repression (Shi et al., 2004). Tsukada et al. reported in 2006 that JmjC (jumonji) domain-containing proteins were also involved in histone demethylase activity (Tsukada et al., 2006). Since then, many JmjC-domain proteins have been found, leading to the identification of multiple demethylases. JmjC-domain proteins are able to remove mono-, di-, and trimethylated histone marks, while LSD1^{KDM1} is unable to remove tri-methyl groups from lysine. A complete list of demethylases can be found in Table 3.

Although many different enzymes are responsible for removing different methylation marks at different histones, to date there is no known enzyme, which can remove histone H4K20 methylation marks. Recently, it was reported that removing LSD1^{KDM1} *in vivo* (in fission yeast Swm1^{KDM1}) increases the global level of histone H3K4 as well as H3K9 dimethylation in *S. pombe* (Gordon et al., 2007; Lan et al., 2007; Opel et al., 2007). To repress gene expression, LSD1^{KDM1} acts with the HDAC

Clr6 and the chromatin remodeler Hrp1 (Opel et al., 2007). JHDM1^{KDM2} demethylates histone H3K36 -me1 and -me2, whereas JHDM2A^{KDM3A} demethylates histone H3K9 -me1 and -me2 (Tsukada et al., 2006; Yamane et al., 2006). PLU-1, a transcription repressor involved in breast cancer cell proliferation, is a histone H3K4 demethylase (Yamane et al., 2007). UTX^{KDM6A} and JMJD3^{KDM6B} demethylases are responsible for removing the repressive marks of H3K27me3/me2, and appear to function as transcriptional activators (Agger et al., 2007; De Santa et al., 2007; Lee et al., 2007a). Thus, histone demethylation is involved in gene transcription and repression in several different organisms.

Family	***Substrate***	***Human***	***S. cerevisiae***	***S. pombe***	***D. melanogaster***
LSD1^{KDM1}	*H3K4me2/1, H3K9me2/1*	*LSD1*	-	*Lsd1/ Swm1/ Saf110*	*CG17149*
JHDM1^{KDM2}	*H3K36me2, H3K4me3*	*JHDM1A JHDM1B*	*Jhd1*	-	*CG11033*
JHDM2^{KDM3}	*H3K9me2*	*JMJD1A JMJD1B JMJD1C*	-	-	*CG8165(NR*)*
JMJD2^{KDM4}	*H3K9me3/2, H3K36me3/2*	*JMJD2A JMJD2B JMJD2C JMJD2D*	*Rph1*	-	*CG15835 CG33182(NR*)*
JARID^{KDM5}	*H3K4me3/2*	*JARID1A JARID1B JARID1C JARID1D*	*Jhd2*	*Jmj2*	*Lid CG3654*
JMJD3^{KDM6}	*K27me3/2*	*UTX UTY JMJD3*	-	-	*CG5640(NR*)*
JMJD6	*H3R2, H4R3*	*PTDSR*	-	-	-

Table 3: HDMs in different organisms. NR indicates that the function of the gene is not reported although homolog to human HDM (Allis et al., 2007; Lan et al., 2008).*

Taken together, these results indicate that histone methylation is the result of an equilibrium between opposing chromatin-modification activities, while histone demethylases are required to maintain a precise level of methylation (Agger et al., 2008).

4.2 THE ROLE OF HISTONE MODIFICATION IN TRANSCRIPTIONAL CONTROL

Histones are not to be considered simple 'DNA-packaging' proteins; they should be recognized as dynamic regulators of gene activity. Histone modifications and DNA methylation have a critical role in arranging nuclear architecture, which is involved in regulating transcription and other nuclear processes (Esteller, 2007). Several histone modifications affect gene expression by increasing or preventing transcriptional initiation and elongation. By analysis of different published genome-wide data, it was observed that distinct chromatin regions display a distinctive pattern of histone modification in budding yeast (Kurdistani et al., 2004; Liu et al., 2005; Pokholok et al., 2005), in fission yeast (Durand-Dubief et al., 2007; Sinha et al., 2006; Wiren et al., 2005) and in human (Barski et al., 2007; Wang et al., 2008). Different modifications are distributed in different locations in the genome and thus often illustrate distinct modification patterns within the core promoter regions, ORF regions, 5' and 3' ends of the ORF regions, and downstream of the 3' end or 3'UTR regions.

Acetylation of histone H3, H4 and H3K4 –di or –tri methylation are the common modification marks associated with active transcription and thus referred to as euchromatin modifications (Li et al., 2007). However, not all H4 acetylation marks are associated with active transcription. For example, different genome-wide modification studies have revealed that histone H4K8ac and H4K16ac modification sites within ORF regions negatively correlate with gene expression in budding yeast (Kurdistani et al., 2004; Liu et al., 2005). However, these modification marks are positively correlated with transcription in the human genome (Wang et al., 2008). Histone H3K9me and H3K27me modification marks are often associated with inactive genes or regions and thus termed as heterochromatin modifications (Li et al., 2007). We have conducted several histone modification studies in fission yeast, and a detailed discussion about the role of different histone modifications in transcriptional

control and a comparative analysis among various organisms can be found in paper IV (see section 7.4).

4.4 THE HISTONE CODE HYPOTHESIS

Numerous studies have confirmed that histone modifications affect gene expression. Two models that have been proposed are the "histone code" model and the "quantitative" model. According to the histone code model, modification patterns function like codes, which can be short-term or long-term and even epigenetically heritable (Jenuwein and Allis, 2001; Turner, 1993). The quantitative model indicates that the chromatin structure is affected by histone modifications as a result of the net charges of the histones (Wade et al., 1997).

By comparing different datasets of modifications, it was found that some modifications occur together, and their similar patterns of modification is present in biologically-related genes (Kurdistani et al., 2004).. Many enzymes have unique site preferences, and many act only on certain chromosomal regions.

One question may arise as to why there are so many modifications. Many modifications correlate with gene expression, while others correlate with different DNA-based processes. The histone code hypothesis suggests that distinct histone modifications on histone tails act sequentially or in concert to form a "histone code", which can be read by a protein and cause downstream events (Strahl and Allis, 2000). Different combinations of modifications on histone tails may function as a histone "language" or a molecular bar code, which may read by other proteins or protein modules that create a distinct chromatin state. Thus, single or combined histone tail modifications build a code that influences the transcriptional state of the gene (Spotswood and Turner, 2002).

The histone code hypothesis describes two groups of biochemical activities: i) enzymes are responsible for writing the code which can add or remove specific modifications at specific sites in the histones, and ii) proteins that are responsible for reading the code, which can produce subsequent functional effects by interacting with the appropriate histone modification patterns (Dutnall, 2003). Many researchers seek to crack the complete histone code so they can "read" the functional status of any region of the genome by its histone modification patterns as easily as they can translate the DNA sequence into protein sequences using the genetic code (Dutnall, 2003). Several genomic studies indicate that histone modifications tend to co-occur in

groups, which may suggest that the purpose of histone modifications is to generate a very simple code (Rando, 2007). The histone code hypothesis has been extended to the epigenetic code hypothesis. According to this hypothesis, the code should be heritable; the information should be passed from one cell generation to the next via a stable transcriptional state (Turner, 2002).

5. MATERIALS AND METHODS

5.1 THE MODEL ORGANISM *SCHIZOSACCHAROMYCES POMBE*

Schizosaccharomyces pombe, also known as fission yeast, has a relatively small genome of 12.5 Mb and 3 chromosomes that are fully sequenced. Fission yeast cells are rod-shaped, they grow by tip elongation, and the cells divide through medial fission. Fission yeast diverged from *Saccharomyces cerevisiae* and humans more than 1000 million years ago (Heckman et al., 2001). *S. pombe* is a simply organized unicellular organism, having a typical eukaryotic cell cycle. Fission yeast cells grow quickly, and it is easy to manipulate their genetic modifications in the laboratory, making this an advantageous model system to study.

In this thesis, we have used several different genome-wide methods to investigate different roles for fission yeast HDACs. Using the ChIP-chip (chromatin immunoprecipitation, see section 5.2.3) method, we revealed the distribution of different histone modifications over an average gene in fission yeast. Global- and locus-specific analysis of several different histone modifications by ChIP-chip and mass spectrometry methods revealed different roles of histone modification patterns in transcriptional control in fission yeast. Using high-resolution tiling microarrays, we further investigated various roles of histone H4K20 methylation in gene expression in fission yeast.

5.2 GENOME WIDE METHODS IN MOLECULAR BIOLOGY

5.2.1 DNA microarray

A microarray is a slide that consists of thousands of microscopic spots of DNA oligonucleotides. In general, each oligonucleotide spot originates from a section of a gene or other DNA element and is typically referred to as a probe. The microarray technology is based on the hybridization properties of nucleic acid. Typically, the solid surface, a glass slide or silicon chip, contains different complementary molecules attached in each probe, used to determine the amount of specific nucleic acid transcripts that are present in a sample. The experimental samples are labeled, hybridized and scanned through a specialized scanner to determine the amount of hybridized targets at each probe.

Based on the data generated by different microarray platforms, microarray technology can be divided into two main classes: the two-color spotted platform and

the high-density oligonucleotide array platform. The two-color platform delivers two sets of probe-level data per microarray, typically red and green channels, whereas the other platform produces one set of probe-level data per microarray to measure specific and non-specific binding. For example, in my thesis work, we have used Eurogentec custom designed *S. pombe* microarrays and Affymetrix *S. pombe* gene-chip microarrays. The Eurogentec array is a two-color spotted microarray platform, whereas the Affymetrix array is a high-density oligonucleotide array platform.

Before using any valuable information from microarray experiments, several different tasks are usually required, and these are known as preprocessing steps. There are 6 different steps in preprocessing: image analysis, data import, background adjustment, normalization, summarization and quality assessment (Huber et al., 2005).

After scanning the microarray slide through a special scanner, probe-level data can be obtained from pixel intensity in the *image analysis* preprocessing step. Previously, this step was crucial, but now, automated systems and computer software can easily and efficiently carry out the task.

Flexible *data import* steps are required to extract and organize microarray data. Microarray data come from various platforms and manufacturers, which often create many different data formats and files. Sometimes, it is very difficult to compare them with each other.

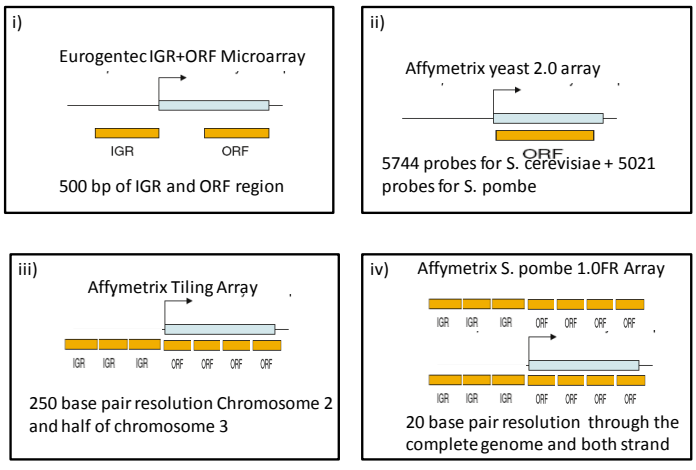


Figure 5.A schematic representation of different types of microarrays used in this thesis

For example, we have used four distinct types of microarrays for different projects (Figure 5). i) Eurogentec two-color spotted IGR+ORF microarray contains information for IGR and ORF spots. Each intergenic region (IGR) is represented by a 500 bp probe immediately upstream of the ATG, and, for each coding region (ORF), by a 500 bp probe immediately upstream of the stop codon. ii) The GeneChip® Yeast Genome 2.0 Array from Affymetrix contains probe sets to identify transcripts from both *S. cerevisiae* and *S. pombe*. The array covers approximately 5,744 probe sets for 5,841 of the 5,845 genes present in *S. cerevisiae*, and 5,021 probe sets for all 5,031 genes present in *S. pombe*. iii) The Gene Chip *S. pombe* Tiling Array from Affymetrix is a high-density oligonucleotide microarray with 250 bp resolution, and it includes chromosome II and half of chromosome III of the *S. pombe* genome. iv). The *S. pombe* 1.0FR Array from Affymetrix contains 1.2 million perfect match/mismatch probe pairs tiled through the complete *S. pombe* genome. Probes are tiled for both strands of the genome at an average of 20 base pair resolution, as measured from the central position of adjacent 25-mer oligos, creating an overlap of approximately 5 base pairs on adjacent probes. For my thesis, I have designed several in-house scripts to compare different microarray datasets. For instance, to compare Eurogentec IGR+ORF microarray with the Affymetrix 1.0FR tiling array, I designed a script that can extract all the fragments from a tiling array, corresponding to IGR+ORF data. An outline of the process can be found in Figure 6.

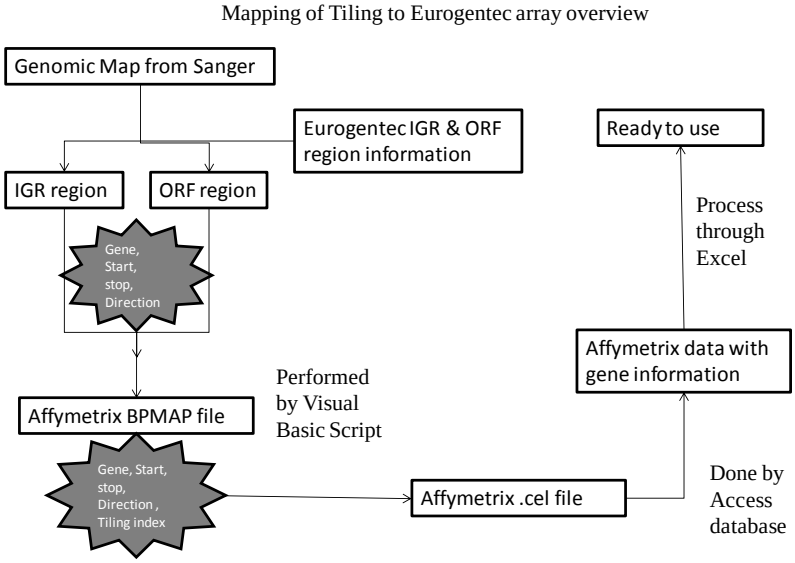


Figure 6. Mapping of Affymetrix tiling array to Eurogentec array.

Genomic information from the Sanger database and Eurogentec IGR-ORF region information was combined and matched with the Affymetrix sequence information file (BPMAP) through a Visual basic script and loaded into a Microsoft Access database. The script was able to match all the probes present in the tiling data with Eurogenetic data, based on each probe's starting and ending coordinates. As Eurogentec probes are 500 bp, whereas tiling probes are 20 bp, for each Eurogentec probe, 24 to 25 tiling probes were collected and grouped, based on genomic information. Experimental data were collected from .cel file(s) and loaded into MS Access. A query was generated through MS Access, and the resulting file was formatted through MS Excel.

To measure non-specific hybridization, the *background adjustment* step is used. In the two-color spotted platform, background levels are measured from the area of the glass slide, where no probes are present. On the contrary, in the high-density oligonucleotide array platform, non-specific binding can be measured from mismatched probe sets.

In microarray data, *normalization* is an important preprocessing step. *Normalization* is a technique that eliminates data redundancy and reduces a complex data structure into the simplest and most stable structure possible. Different

normalization methods can be applied on different datasets to obtain meaningful information. For example, typically for the two-color array, the 'lowess normalization' method is applied, but for the high-density oligonucleotide Affymetrix tiling array, the 'quantile normalization' method is used. 'Lowess normalization,' also known as non-linear normalization, is used to remove dye-related artifacts in two-color experiments. For all the expression profiles used in papers I and II, we have employed the 'lowess normalization' method. However, for the high-resolution tiling array used in paper V, we applied the 'quantile normalization' method. To measure the differences in highly similar datasets, the variance normalization method can be applied. For example, in paper IV, we analyzed 12 different histone modification patterns in fission yeast, using the Eurogentec IGR+ORF microarray. Many different histone modification marks, especially acetylation marks, show similar patterns, but to investigate differences among these sites, we applied the 'variance normalization' method.

In a microarray, multiple probes can represent one transcript. In the *summarization* step for each transcript, a single value can be calculated from multiple probes' background-adjusted intensities and normalized intensities.

Quality assessment is a step during which different spots are compared, and flawed measurements can be discarded, based on the acceptable level of random fluctuations. These data are represented as flagged and generally not used for further analysis.

There are several prerequisites essential to carrying out the abovementioned tasks. The experimental design includes the selection of the array platform, selection of samples, choice of controls, and selection of sample extraction, amplification, labeling and hybridization procedures. The technical, logistic and financial boundary conditions should be taken into account by the experimental design. The collection of samples includes the types of samples (cell lines or tissues exposed to different conditions) and biological and technical replicates. Replication of the biological samples is crucial for extracting conclusions from the experiment, and technical replicates help to ensure precision. Highly developed skills and expertise are also required to conduct these experiments.

5.2.2 Gene expression data

The use of a set of different DNA samples in arrays for expression profiling was first described in 1987 (Kulesh et al., 1987), and the first spotting of a cDNA microarray for gene expression profiling was reported in 1995 (Schena et al., 1995). Since then, the technology has become an excellent, powerful tool in several areas of biomedical research. Measuring gene expression in a genome-wide scale is the most common use of microarray technology. Typically, to generate gene expression data, mRNA is extracted from cells and converted into cDNA by reverse transcription. The resulting cDNA is then labeled with a fluorescence dye and hybridized to the microarray. The slide is scanned and all preprocessing steps are followed. In this thesis, we have used the two-color platform from Eurogentec to obtain gene expression data used in papers I, II and III. We have also generated gene expression data using the Affymetrix Yeast 2.0 and 1.0FR oligonucleotide arrays platforms that were featured in paper IV and V respectively.

5.2.3 Genome-wide location data

A few years later, when gene expression microarray was developed, a group of scientists created a microarray method that revealed the genome-wide location of DNA-bound proteins (Ren et al., 2000). The method is popularly known as 'Chip-chip' or the 'ChIP on chip' method and it combines chromatin immunoprecipitation ("ChIP") with microarray technology ("chip"). Chromatin immunoprecipitation is a method through which one can determine the location of DNA binding sites on the genome for a protein of interest. Combining this method with microarray technology, one can determine this on a genome-wide scale. Several experimental steps are involved in the ChIP-chip method. In brief, the fixation of protein-DNA interaction by formaldehyde cross-linking, followed by sonication, is used to separate the DNA into smaller fragments. Next, a specific antibody is used to immunoprecipitate the protein of interest. The resulting DNA fragments are reverse crosslinked to obtain only DNA and then hybridized to a microarray. The slide is scanned and all pre-processing steps are followed. In this thesis, we have applied the ChIP-chip method in all of the papers with several microarray platforms.

5.2.4 Annotation of the *S.pombe* genome

As discussed above, it is possible to map the entire fission yeast genome using the *S. pombe* 1.0FR array from Affymetrix. Strand-specific information from this array technology is also available. However, when we started to use the platform, we realized that the annotation file available for the *S. pombe* genome was incomplete. For example, using the earlier annotation file, one can obtain information about the ORF regions. However, information about the non-coding regions was missing from the annotation file. To analyze the entire genome tiling array (as described in paper V), I designed a new *S. pombe* annotation file that contains the relevant information about ORF, 5'UTR (5'untranslated region), 3'UTR (3'untranslated region), repeat regions, repeat units (such as the centromeric region), tRNA, rRNA, snRNA, snoRNA, introns, promoters and IGRs. All information regarding the *S. pombe* genome (except IGRs) was based on the Sanger database and other published datasets (Dutrow et al., 2008; Wilhelm et al., 2008). The data from the Sanger database were collected from the Chromosome contigs directory (at ftp://ftp.sanger.ac.uk/pub/yeast/pombe/, June 2009) using Artemis (Rutherford et al., 2000) and processed with an in-house script. For IGRs, I calculated approximately 500 bp of intergenic regions, based on their Sanger positions and directions, and added these to the annotation file. The file is designed in GFF (General Feature Format) format and can be accepted by many different microarray data analysis software.

5.2.5 Microarray data structure and data analysis using different software

Two different approaches are typically used to perform microarray data analysis: the stepwise and integrated approaches. In the *stepwise* approach, different tasks are separated into a set of instructions. This approach takes probe level data as input, sequentially executes the tasks and produces an expression matrix data table as output. However, integrated approaches use combined evaluation processes to solve specific problems.

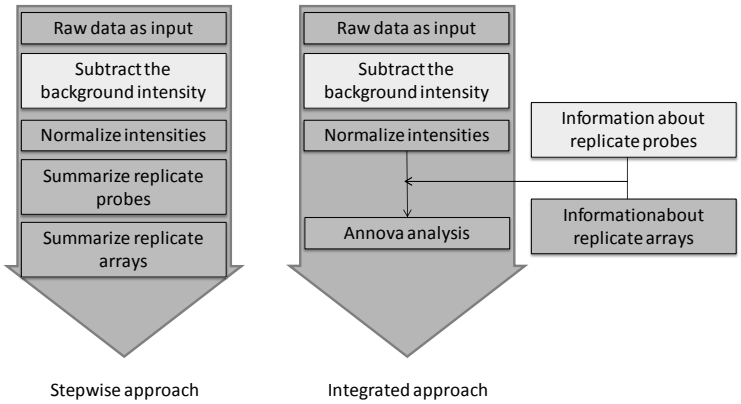


Figure 7. Microarray data preprocessing by different approaches.

An example of a set of sequential instruction and integrated approaches can be found in Figure 7.

5.2.5.1 Data Types

Microarray data preprocessing handles several basic data types, such as probe and background intensities, array layout, probe annotations and sample annotations. Typically, these data types come in a flat file format or in specific file formats.

The probe and background intensity data are derived from the image processing. The array layout information is typically obtained from array venders.

The array layout indicates the physical position of each probe in a particular array. For the two-color array, the information can be designed in block, column and row format, where each block represents a subsection of the array, and the row and column represent the exact coordinates of each probe in that subsection. In the high-resolution tiling array, the array layout information is typically obtained from the –x and –y coordinates.

The organization that produces the array generally provides the probe annotation information. Typically, probe annotation information contains information about the probe sequence and its location in the genome. For the high-resolution microarray, several different probes can be found for a particular gene transcript, whereas for the two-color array, one or two probes for each transcript can be found.

The sample annotations are produced by the researcher and include technical information and processing protocols.

5.2.5.2 Data analysis software

One of the most challenging jobs of DNA microarray is analyzing data relating to intensity information with biological interpretations. There are many different ways to analyze microarray data: free software, commercially available software and in-house scripts.

Many freely distributed software types are available in the market, and they can be download and used free of charge. However, as these are freely available, it can become difficult to obtain updated versions or to receive technical assistance. To operate some software, basic computational skills are also required.

Commercially available software is easy to use, and various types of software are available in the market. However, lack of flexibility and certain statistical limitations are common problems for existing software types. A yearly license fee is also applied for this kind of software.

One of the significant advantages of in-house scripts is that the user can consider a specific problem and design the script to solve it. However, this increases the time required to complete the workflow. Specific computational skills are also required to perform the task.

I believe that combining the above three types of analysis produces the most efficient output.

In this thesis, we have used two different commercially available software: Genespring from Agilent Technology and Partek Genomics Suite from Partek. Genespring is an excellent tool designed to analyze two-color microarray data. The software features one database in which one can store all the microarray data and make gene lists for gene of interest. A number of genomic views can be found in Genespring. Several different statistical analyses tools, such as ANNOVA, clustering, and principle component analysis (PCA), are integrated within the software. A useful feature of the software is that it can produce similar gene lists, based on similarly produces p-values. Although it has several positive features, one of the disadvantages of this software is its inability to analyze high-resolution tiling array data. Another disadvantage of the software is that it occasionally functions poorly due to its background database workflow.

Partek Genomics Suite is another excellent software type for microarray data analysis. Unlike Genespring, it can handle both two-color and high-resolution tiling arrays. Even the software is designed to handle and analyze high throughput

sequencing data. A number of statistical features are included inside the software, which is also very fast and user-friendly. Strong technical support is also available. The main disadvantage of the software is that there is no database situated inside the software, so similar list analyses cannot be performed. The visualization features of the software are also unimpressive.

Among freely distributed, open source software, Bioconductor R (http://www.bioconductor.org/) is used commonly by microarray analysts. R is a development software project for the analysis and comprehension of genomic data. A large community is involved in this project, and many researchers internationally are constantly improving upon it. A few useful books regarding Bioconductor R are also available to the microarray community and can be helpful for data analysis (Gentleman et al., 2005; Hahne et al., 2008).

Affymetrix provides freely available software for tiling array analysis, known as TAS (Tiling analysis software). The product is user-friendly and can perform some preprocessing tasks, such as data import, background adjustment, normalization, summarization and quality assessment. Affymetrix also provides a software called GCOS (GeneChip® Operating Software), which is useful for array scanning and several preprocessing steps. Another popular Affymetrix product is Integrated Genomic Browser (IGB), designed for the visualization and exploration of genomes and corresponding annotations from multiple data sources.

Tm4 Microarray Software Suite and MAT (Model-based Analysis of Tiling-array) are additional types of popular open source software available to the microarray community.

Except microarray analysis software, two other products from Microsoft, MS Excel and MS Access, are useful for microarray analysis. MS Excel is very popular among scientists, and there is no need to describe it here. Although scientists use MS Access less commonly, I found this to be an excellent tool for managing microarray data. Microsoft Access is a relational database management system (RDBMS) from Microsoft that joins the relational Microsoft Jet Database Engine with a graphical user interface and software development tools. Access stores data in its own format, but various data types, such as Excel, Text, Outlook, HTML and any kind of ODBC (Open Database connectivity), can be exported, imported or linked to Access databases. MS Access is a user-friendly database system, and no specific computational skills are required to use the software. Users can create their own

simple database solutions by designing tables, queries, forms and reports, and connect them together with macros.

One of the main advantages of MS Access is that it can store unlimited numbers of rows in a data table, which is very useful for microarray data, especially for high-resolution data containing millions of data points. Another advantage of the software is that it can create simple quarry related to microarray data. For example, two different microarray platforms can be matched together by a simple Access quarry.

6. AIMS

The aim of this thesis was to study different HDACs functions and histone modification patterns using different microarray technologies in fission yeast. A list of more detailed aims is given below.

- To analyze the genome-wide roles of all six HDACs present in fission yeast;
- To study genome-wide histone modifications in fission yeast using different microarray approaches;
- To develop a method for comparing different microarray platforms; and
- To study histone modifications in both global and locus-specific scales.

7. RESULTS AND DISCUSSION

7.1 PAPER I

Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast

HDAC enzymes are associated with gene regulation. In addition, they are involved in several DNA-related processes, such as DNA replication and DNA repair. In this study, we sought to determine the enzyme specificity of different HDACs in fission yeast. Furthermore, we investigated genome-wide binding targets and acetylation patterns of HDACs.

We have conducted a genome-wide analysis to investigate functions of four different HDACs in fission yeast: Hos2, Clr6, Clr3, and Sir2. To determine the *in vivo* HDAC activity on a genome-wide scale we constructed an IGR+ORF spotted microarray. We used the ChIP-chip method to determine enzyme activity maps for four different HDACs. A set of five specific antibodies for H3K9ac, H3K14ac, H4K5ac, H4K12ac and H4K16ac were used for ChIP-chip. To control the nucleosome occupancy, we used an antibody (H3-cter) against the histone H3 C-terminal region. Expression profiling experiments were conducted with wild type and different HDAC mutants to determine the division of labor between HDACs. We established expression-profiling experiments for four different HDACs, compared them with wild type expression data and established high and low expressed gene lists. Chromosomal binding maps for epitope-tagged Sir2 and Clr3 proteins were produced using the ChIP-chip method.

Our nucleosome occupancy analysis and acetylation ratios in HDAC mutants, as compared to wild type expression data, revealed that Sir2 and Hos2 play a major role in histone assembly by preventing nucleosome loss. Furthermore, we observed that the *clr6-1* mutant showed the strongest and broadest increase in acetylation compared to wt.

HDACs have previously been reported to be involved in gene repression. From our expression analysis, it is clear that all HDACs in fission yeast are involved in gene repression. However, a significant number of down-regulated genes in the *hos2* mutant suggest that the enzyme also plays a role in gene activation. Our analysis revealed that in general, different HDACs show non-overlapping roles in gene expression, although Sir2 and Clr3 displayed a significant overlap in gene repression.

We also discovered that Sir2 and Clr3 act together throughout the silent regions. Both the enzymes share many target genes at silent heterochromatin regions of *S. pombe*. Various studies have reported that both enzymes are needed for efficient rDNA, centromere, telomere and mating type locus silencing (Bjerling et al., 2002; Freeman-Cook et al., 2005; Shankaranarayana et al., 2003). Our ChIP-chip binding maps indicated that these two enzymes target significant amounts of fragments from all four silent regions.

Comparing histone acetylation patterns with gene expression data, we observed a tendency for reduced histone content in the IGR and ORF of highly expressed genes. Similar to studies in budding yeast (Kurdistani et al., 2004; Liu et al., 2005), we also observed that histone H3K9ac positively correlates with gene expression, whereas H4K16ac negatively correlates with gene expression at IGR and ORF regions in fission yeast. Furthermore, we observed that in *hos2* mutants, the level of H4K16ac positively correlates with gene expression; this further supports our observation about its involvement in promoting gene transcription. Furthermore, our hypergeometric probability distribution analysis revealed that Hos2 is required for high expression of growth-related genes, and we also show that hos2 mutated cells grow slowly in comparison to wild type cells, which indicates that the enzyme is involved in promoting rapid growth.

7.2 PAPER II

Specific functions for the fission yeast Sirtuins Hst2 and Hst4 in gene regulation and retrotransposon silencing

There are six HDACs present in fission yeast. In our previous study, we analyzed the different cellular roles of Clr6, Clr3, Sir2 and Hos2 HDACs in fission yeast. However, analyses of two additional HDACs, namely, Hst2 and Hst4, were missing from that study. Hst2 and Hst4, together with Sir2, belong to the class III HDAC family, also known as Sirtuins. In this paper, we explored the cellular functions of all three Sirtuins in fission yeast.

We used ChIP-chip, expression profiling and phenotypic analyses to investigate different class III histone deacetylase (HDAC) functions in fission yeast. To reveal different functions for Sirtuins in the regulation of gene expression, we performed expression profiling of *hst2Δ* and compared our results with previously

published data for the other HDACs (Fagerstrom-Billai et al., 2007; Wiren et al., 2005). To investigate the role of Sirtuins in genomic nucleosome density, we performed the ChIP-chip experiment using the Eurogentec IGR+ORF spotted microarray. H3-cter antibody was used to control and calculate the nucleosome occupancy. To determine the cellular localization of Hst2, we performed Immunofluorescence (IF). We also evaluated the phenotypes of single and double mutant combinations of Sirtuins to understand the functional relationship of the enzymes.

From our investigation, it was observed that although all the Sirtuins share some common functions, each of them has specific genomic targets. Our analysis revealed that Hst4 is required for RNA processing and retrotransposon silencing, whereas Hst2 is required for heterochromatin silencing. By investigating the influence of all class III HDACs on nucleosome density, we found that the three Sirtuin enzymes are required to maintain normal nucleosome density and distribution in the genome of *S. pombe*. From expression profiling data, it was also observed that each Sirtuin has a different function in the regulation of gene expression. Sir2 and Hst2 repress non-expressed genes, whereas Hst4 acts globally to reduce gene expression. We also found that for highly expressed genes, the level of acetylation is changed by *hst4Δ* whereas *sir2Δ* shows increased histone acetylation levels within the ORFs of repressed genes. *In vivo* experiments suggest that among all Sir2 binding sites, histone H3K9 acetylation is the most significant hyper-acetylated modification in *sir2Δ,* in both IGR and ORF regions, in the presence of other Sirtuins (Hst2 and Hst4). However, in the absence of the remaining two Sirtuins (*i.e.,* for Sir2 exclusively in ORF regions), histone H4K16 acetylation is the most significant hyperacetylated site in *sir2Δ.*

7.3 PAPER III

Genome-wide patterns of histone modifications in fission yeast

In our previous studies, we systematically investigated different histone deacetylase (HDAC) functions in fission yeast. However, in those studies, we used the Eurogentec IGR+ORF spotted microarray, which contains 500 bp IGR and ORF fragments, and this restricted our analysis to that resolution. Since then, Affymetrix has launched the gene chip *S. pombe* tiling array. We have taken advantage of the

high-resolution array and initiated a study of the genome-wide patterns of histone modifications in fission yeast.

In this study, we used the Affymetrix oligonucleotide tiling array to establish genome-wide, high-resolution maps for fission yeast. The ChIP-chip method was used to obtain the acetylation patterns of five histone modification marks at *S. pombe* IGR and ORF regions. To measure and control nucleosome occupancy, we used an antibody (H3 cter) against histone H3. When we began using the tiling microarray, which lacks proper gene annotation, we realized that a proper array annotation could provide us with more accurate output data. To compare the tiling array data with our previously used Eurogentec spotted microarray data, we designed a new nomenclature for almost all the tiling fragments of the Gene Chip *S. pombe* tiling array. To investigate the acetylation patterns across the fission yeast genome, we computed the median histone acetylation value (corrected with histone H3 occupancy) for the different 5' to 3' positions of the tiling fragments. To establish the correlation between histone acetylation patterns and gene expression, we compared our acetylation maps with our previously published expression data. The *S. pombe* spotted microarray platform was also used to validate the tiling microarray data. Finally, we compared our tiling array data with previously published datasets on HDACs to investigate how different HDAC mutants affect histone acetylation in various parts of the fission yeast genome.

Our analysis revealed that independent of gene length, a typical fission yeast gene shows a 5' to 3' polarity, *i.e.*, the histone acetylation levels peak near the ATG and gradually decrease in the coding regions. To test the statistical significance of the observed patterns, we produced several ranking lists of high and low acetylation levels at IGR and ORF regions. Using the hypergeometric distribution test, we concluded that the level of histone acetylation is relatively high within 0-2 kb of IGR regions. At ORF regions, however, histone acetylation decreases between 250 to 500 bp regions. Comparing acetylation patterns with gene expression data, we also found that histone acetylation patterns depend on gene expression but are independent of gene length. From different studies, it was observed that H3K9ac positively correlates with gene expression, whereas H4K16ac negatively correlates with gene transcription in budding and fission yeast (Kurdistani et al., 2004; Wiren et al., 2005). We observed that histone H3K9ac produces a stronger peak near the ATG region and gradually decreases at the coding regions, whereas the level of H4K16ac is generally

low at ATG regions, as well as at coding regions of highly expressed genes. Comparing our results with other published datasets, we observed that different HDAC mutants affect acetylation in different parts of ORFs. When comparing high and low resolution microarray data, it was clear that different HDACs are responsible for different positions within the ORF regions. We observed that Hos2 acts in the 5' region, Sir2 and Clr6 act in the middle region, and Clr6 acts in the 3' region.

7.4 PAPER IV

Genome wide mapping of histone modifications and mass spectrometry reveal a function for the histone H4 acetylation zip and a role for H3K36 methylation at gene promoters in fission yeast

Posttranslational modifications of histones alter their interactions with DNA and nuclear proteins or engage effector molecules, thus regulating DNA-dependent processes, such as transcription, replication and repair. 'ChIP-chip' is an excellent technique to investigate combinations of histone modification in a global way whereas by the top-down mass spectrometry approach, combinations of histone modification can be studied in a locus-specific manner.

To characterize histone modifications at both global and locus-specific scales, we carried out ChIP-chip analyses of histone acetylation and methylation and combined this approach with a proteomic Mass Spectrometry survey of histone modifications in fission yeast. This combined study enabled us to identify histone modifications, to measure their distribution across the fission yeast genome and to characterize the relationships between histone modifications on single histone peptides and at co-regulated gene clusters. For the ChIP-chip experiments, we used 12 different histone acetylation and methylation modification sites. To investigate the role of individual histone modifications at gene promoters and gene expression, we compared each modification site to all other modified lysines and with gene expression data. Gene clustering analysis, followed by GO analysis, was performed to investigate how histone modification patterns are linked to gene function. The distribution of lysine acetylation at the histone H4 tail is often biased toward the nucleosome core, and has been termed as the 'acetylation zip' (Turner et al., 1989; Zhang et al., 2002). We applied both ChIP-chip and MS approaches to investigate the existence of the zip model in fission yeast. Finally, we examined the relationship

between histone H3K36 methylation and histone H3K27 acetylation using the MS and Chip-chip techniques.

Several genome-wide studies in *Saccharomyces cerevisiae* have suggested that gene transcription correlates positively with acetylation of most individual lysines in histone H3 and H4 tails (Kurdistani et al., 2004; Pokholok et al., 2005).

In *S. pombe,* we observed that histone H3K9 acetylation in intergenic regions and open reading frame regions positively correlates with gene expression (paper I) (Wiren et al., 2005). Genome-wide studies in *Saccharomyces cerevisiae* have shown that transcription correlates positively with histone H3K9, H3K18, and H3K27 acetylation at promoter regions (Kurdistani et al., 2004; Liu et al., 2005) (Table 4). In human T-cells, it was also observed that histone H3 K9, K18, and K27 acetylation positively correlate with gene expression (Wang et al., 2008) (Table 4). Here, we found that H3 K9, K18, and K27 acetylation in promoter regions positively correlate with gene expression in *S. pombe.* Comparing *S. cerevisiae, S. pombe,* and human T-cell data at promoter regions, we noticed that H3K23ac positively correlates with transcription in *S. pombe* and human T-cells, but not in *S. cerevisiae.* H4K8ac negatively correlates with gene expression in both yeasts, but not in human T-cells. Thus, H3K9, K18, and K27ac are positively correlated with transcription in both yeasts and human T-cells (Table 4). The conserved correlation from yeast to human may suggest that these histone H3 acetylation sites have special importance for gene activity.

When comparing acetylation levels at ORF regions, transcription correlates positively with histone H3K9, H3K18, and H3K27 acetylation in *S. cerevisiae* (Kurdistani et al., 2004) (Table 4). Although H3K27ac was not included in the single-nucleosome mapping study, Liu et al. observed that histone H3K9 and H3K18 acetylation at promoter regions positively correlates with transcription. In our study, we also observed that H3K9, H3K18, H3K23, and H3K27 acetylation at ORF regions positively correlate with transcription in *S. pombe*. In human T-cells, all 18 histone H3 and H4 acetylation sites were tested, including H3K9, H3K18, H3K27, and H3K23, at coding regions, demonstrating positive correlation with gene expression (Wang et al., 2008).

In budding yeast, H4K16 and H4K8 acetylation provide relatively poor correlations with transcription (Kurdistani et al., 2004). Using single-nucleosome mapping, Liu et al. observed that H4K8ac and H4K16ac negatively correlate with

Table 4: A comparison of histone modification patterns in IGR and ORF regions in different species

A comparison of histone modification patterns in gene promoters in different species						
Gene expression	**Human CD4+ T-cells CHIP-seq; Single single nucleosome mapping (Wang et al.)**	**Human ES cells CHIP–chip (Guenther et al.)**	**Flies CHIP–chip (Schubeler et al.)**	**Budding yeast CHIP–chip (Kurdistani et al.)**	**Budding yeast CHIP–chip; Single single nucleosome mapping (Liu et al.)**	**Fission yeast ChIP-chip (this study)**
High	H2AK5ac, H2AK9ac, H2BK5ac, H2BK12ac, H2BK20ac, H2BK120ac, H3K4ac, H3K9ac, H3K14ac, H3K18ac, H3K23ac, H3K27ac, H3K36ac, H4K5ac, H4K8ac, H4K12ac, H4K16ac and H4K91ac, H3K4(me1,me2,me3),H3K9me1, H3K27me1,H3K36me1, H3K36me3, H3K79(me1, me2,me3), H4K20me1	H3K9ac, H3K14ac, H3K4me3	Hyperacetylated for H3 and H4 and H3K4me, H3K79me	H3K9ac, H3K18ac, H3K27ac	H3K18ac, H4K12ac, H3K9ac, H3K14ac H4K5ac H3K4me3	H3K9ac, H3K14ac, H3K18ac, H3K23ac, H3K27ac, H3K56ac, H4K5ac, H4K12ac, H4K16ac, H3K4me2, H3K36me2
Low	H3K9me2, H3K9me3, H3K27me2, H3K27me3, H4K20me3	H3K4me3	Inactive genes being hypomethylated and deacetylated at the same residues	H4K8ac, H4K12ac, H4K16ac, H3k14ac, H3K23ac, H2AK7ac, H2BK11ac, H2BK16ac	H4K16ac, H2BK16ac, H4K8ac	H4K8ac

A comparison of histone modification patterns in ORF regions in different species						
Gene expression	**Human CD4+ T-cells CHIP-seq; Single single nucleosome mapping (Wang et al.)**	**Human ES cells CHIP–chip (Guenther et al.)**	**Flies CHIP–chip (Schubeler et al.)**	**Budding yeast CHIP–chip (Kurdistani et al.)**	**Budding yeast CHIP–chip; Single single nucleosome mapping (Liu et al.)**	**Fission yeast ChIP-chip (this study)**
High	**H3K9ac, H3K18ac, H3K23ac, H3K27ac**,H4K5ac, **H4K8ac**, H4K12ac, **H4K16ac**, H2BK12ac, H2BK20ac, H2BK120ac, H3K4ac, H3K4(me1, me2, me3), H3K36me3, H3K27me1, H3K9me1, H4K20me1, H2BK5me1		H3K4me2, H3K4me3, H3K79me2	**H3K9ac**, H3K14ac, **H3K18ac, H3K23ac, H3K27ac**, H4K12ac	**H3K9ac**, H3K14ac, **H3K18ac**, H4K5ac, H4K12ac, H2AK7ac, H3K4me2, H3K4me3	H3K4me2, **H3K9ac, H3K18ac, H3K23ac, H3K27ac**,H3K56ac, H4K5ac
Low	H3K27me2, H3K27me3, H3K9me2, H3K9me3, H3K79me3			**H4K8ac, H4K16ac**, H2AK7ac, H2BK11ac, H2BK16ac	**H4K8ac, H4K16ac**	H3K14ac, **H4K8ac**, H4K12ac, **H4K16ac**, H3K36me2

transcription in *S. cerevisiae* (Liu et al., 2005). In *S. pombe,* we noticed that H4K8ac and H4K16ac are also negatively correlated with transcription. However, in human T-cells, H4K8ac and H4K16ac are positively correlated with transcription at ORF (Wang et al., 2008). Thus, H3K9ac, H3K18ac, and H3K27ac are positively correlated with transcription in both yeasts (*S. cerevisiae* and *S. pombe*) and human T-cells. These modifications may be referred to as transcription-dependent modifications, as suggested by Liu et al. (Liu et al., 2005). H4K8 and K16ac are negatively correlated with transcription in both yeasts (*S. cerevisiae* and *S. pombe*), but positively correlated in humans, indicating that species-specific modification patterns exist (Lennartsson and Ekwall, 2009).

Our MS and ChIP-chip clustering analyses suggest that the H4 acetylation zip model is present in fission yeast and may be involved in gene expression. From our analysis, we speculate that the Clr6 HDAC at the coding region is involved in generating the H4 acetylation zip, and that inefficient deacetylation of H4K16 by Clr6 reduces gene expression level.

Our combined analysis reveals antagonistic crosstalk between histone H3K36me2/Me3 and H3H27ac at many *S. pombe* gene promoters. From our cluster analysis, it was observed that in many gene clusters H3K36 methylation and H3 acetylation sites showed an antagonistic modification pattern, and this was confirmed by our MS analysis. From MS data, it was also concluded that H3K27 acetylation increases in the absence of $Set2^{KMT3}$. Our analysis shows that Clr6, the Rpd3 HDAC orthologue, targets H3K27 for deacetylation. These findings suggest antagonistic crosstalk between H3K36me2/me3 and H3K27ac at gene promoters involving two enzymes, $Set2^{KMT3}$ and Clr6 HDAC.

7.5 PAPER V

Genome wide mapping suggests different roles for H4K20me1 and H4K20me3 in gene expression

In this genome-wide study, we have mapped *S. pombe* histone H4K20 methylation modification marks to reveal their distinct roles in gene expression. Our aim was to determine the exact functions of histone H4K20 –mono, -di and –tri methylation in different chromosomal locations in fission yeast.

We have used the Affymetrix oligonucleotide whole genome tiling array to establish genome-wide, high-resolution histone H4K20 maps for fission yeast. The ChIP-chip method was used to obtain the H4K20 -mono, -di and -tri methylation patterns in wild type cells. In fission yeast, $Set9^{KMT5}$ is the sole enzyme responsible for establishing all three H4K20 methylation marks (Sanders et al., 2004). We also obtained three different modification patterns, using the *set9Δ* strain as a control for our experiment. We used an antibody against histone H4 to measure and control nucleosome occupancy. Applying principle component analysis (PCA), we compared the ChIP enrichment of each histone modification at each individual lysine to all other modified lysines. To establish the correlation between histone methylation patterns and gene expression, we compared our methylation maps with expression data. We also designed an annotation file for the Gene Chip *S. pombe* tiling array to gain detailed information about where in the genome these modifications are located.

Our data revealed that H4K20me1 is strongly associated with active genes and is enriched at 3'UTRs of actively transcribed genes in fission yeast. This result is different from that obtained with the human genome, where it was observed that H4K20me1 peaks at the TSS (Transcription Start Site) and is enriched in the coding region of active human genes (Barski et al., 2007). These observations indicate that H4K20me1 is most likely involved during the initiation and elongation of transcription in the human genome, whereas the mark is involved in transcription termination in fission yeast. Another striking difference in our analysis was the involvement of H4K20me3 with weakly expressed genes. Genome-wide mapping of histone H4K20me3 in the human genome did not indicate any connection between repressed genes and a high level of H4K20me3 (Barski et al., 2007).

As reported earlier, our methylation maps in the *set9Δ* strain also reveal that $Set9^{KMT5}$ is the single methyltransferase (HMT^{KMT}) responsible for all three methylation states of H4K20 in fission yeast.

We also observed that unlike in mammals, H4K20me3 was absent at the pericentric regions in fission yeast. However, moderate levels of H4K20me1 were observed in the same regions. Histone H4 occupancy data at pericentric and central core regions revealed that a histone H4 depleted region between these two centromeric domains exists in fission yeast.

The most striking observation in our datasets was the correlation between the 3'UTR regions of actively transcribed genes and histone H4K20me1. Using a hypergeometric distribution test, we observed that these regions are correlated with stress-repressed genes. Chen et al. reported that different stresses can trigger down-regulation of many genes that are associated with a number of cellular functions, such as transcription (RNA polymerase I, II, and III subunits) (Chen et al., 2003). In our recent study, we observed that Rpb4, one of the core subunits of RNA polymerase II, is not enriched at stress-inducible genes. Furthermore, the global binding map of Rpb4 reveals that enrichment of Rpb4 peaks near 3'UTR regions (Djupedal et al., 2010). From these data, we speculate that highly transcribed genes that are repressed by stress have some special mechanism for transcription termination. Histone H4K20 methylation, especially mono methylation by $Set9^{KMT5}$, occurs at the 3'UTR regions of these genes and may be recognised by the RNA polymerase II subunit Rpb4, in order to terminate the transcription.

8. FUTURE PERSPECTIVES

In terms of the philosophy of science, genome-wide studies on histone modification patterns, as described in this thesis, often follow a model called the hypothetico-deductive model or method. According to this model, an experiment such as gene expression microarray or ChIP-chip can be designed based on a hypothesis. Next, the experiment is performed to generate a large amount of data. Different data analysis procedures can be applied, which ultimately produce a new set of hypotheses. Thus, a genome-wide study can be expanded in a number of directions. Based on the studies in this thesis, here are some examples.

In papers I and II, we conducted genome-wide investigations to reveal the different roles of the six HDACs present in fission yeast. All the microarray experiments were performed using the low-resolution microarray. For example, we used the Eurogentec, custom-designed IGR+ORF microarray for these studies, which contains 500 bp of IGR and 500 bp of 3' bias ORF fragments. Using the high-resolution microarray or high throughput sequencing techniques may produce much more meaningful and accurate data.

In paper III, we studied histone modification patterns in fission yeast on a genome-wide scale, using the high-resolution tiling array. However, the version of the array used contains only chromosome 2 and half of chromosome 3 fragments. A new version of the tiling array would produce results that are more accurate. Our analysis also proposed that histone acetylation patterns show 5' to 3' polarity effects. Further experiments are required to test this hypothesis.

In paper IV, we performed an epigenomic study of histone modifications, applying ChIP-chip and mass spectrometry approaches. Using the ChIP-chip method, followed by clustering analysis, we identified characteristic histone modification patterns in promoters of groups of co-regulated genes. One of the limitations of this project is that we used a low-resolution microarray, which provided data at a 500 bp resolution. In general, 500 bp genomic DNA contains around three nucleosomes. It would be better to use a mono-nucleosomal resolution study, which would offer a more thorough view of modification patterns at the nucleosome level. Recent studies show that mapping nucleosome positions on a genome-wide scale using microarrays is possible in budding yeast (Lee et al., 2007b; Whitehouse et al., 2007; Yuan et al.,

2005). Recently, the method for genome-wide mapping of nucleosome positions in *S. pombe* was also established in our lab, in collaboration with Philipp Korber's lab (Lantermann et al., 2009). To extend this study, one could possibly use micrococcal nuclease digestion, followed by the ChIP-chip method through the high-resolution tiling microarray, which can produce a more accurate, nucleosome-level modification map for fission yeast.

Our histone modification study, both at global and locus-specific scales, enabled us to reveal many different histone modification patterns in fission yeast. These findings lead to an open question regarding whether the observed patterns involved with some downstream events are "histone code." According to the histone code hypothesis, distinct histone modifications on the same or different histone tails act sequentially or in combination to form a "histone code" that can be read by proteins to cause downstream events (Strahl and Allis, 2000). Our findings indicate that the observed patterns involved with some downstream events are probable candidates for a histone code. Future research should explore the histone modification profile at different time points in the cell cycle and the recognition of histone modification patterns during DNA replication, which may further support the histone code hypothesis.

In paper V, we conducted a genome-wide study using the high-resolution tiling microarray to map histone H4K20 methylation sites in fission yeast. Our analysis revealed that the histone H4K20me1 is enriched at 3'UTR regions of active genes. From our results, we speculate that histone methyl transferase $Set9^{KMT5}$ may be involved in transcription termination of active genes. Further experiments should be performed to test this hypothesis. Our methylation maps at centromeric regions revealed that a histone H4-free region between two centromeric domains contains $tRNA^{Ala}$. We proposed that the H4-free region might prevent heterochromatin spreading into the centromeric chromatin. Further experiments should be conducted in order to evaluate this hypothesis.

9. ACKNOWLEDGEMENTS

This work is financially supported by The Swedish Cancer Society, Swedish Research Council (VR), 'The Epigenome' NoE network, Karolinska Institutet and Södertörn University. My thanks to Karolinska Institutet for providing me with excellent international scientific and multi-cultural environment.

First and foremost I wish to express my most deep and sincere gratitude to my supervisor **Prof. Karl Ekwall** for welcoming me to join his research group in 2004. My PhD studies would not have been possible without the expert guidance and professional support of my esteemed advisor. You have given me great freedom to carry independent work and continuously encouraged me to develop as an independent researcher. Thanks for your honest comments, which were extremely helpful in improving me scientifically.

I would like to thank my co-supervisor **Dr. Karin Dahlman-Wright** for her valuable discussions and important suggestions.

I wish to thank **Dr. Jan Larsson** for accepting our invitation to act as my opponent and **Prof. Claes Wadelius**, **Dr. Chandrasekhar Kanduri** and **Dr. Ola Hermanson** to serve as members of my examination board.

I would like to thank all our past and present **KEK** group members. Special thanks to **Ingela** and **Carolina**. Not only for sharing the office but also for all the scientific and non-scientific discussions (specially the gossips!!☺), answering all my silly and serious questions and many more. Thanks a lot. I will miss you so much. **Ingela**, you are always so kind and helpful. Thanks for the collaboration and always helping me with all the doctoral degree application related issues. Your critical comments and suggestions improved my thesis a lot. I really appreciate your help. **Carolina**, you were always there when I needed help. Thanks for your genuine and practical suggestions, Swedish to English translation, collaborations and correcting our manuscripts. Thanks to **Marianna** for all the collaborations and introducing me to the ChIP-chip technique in the lab. Still I can remember the day when we started the experiment. Thanks to **Annelie** for being a great friend. You are so kind and helpful. Thanks for your help for finishing my 5th manuscript. I could not have managed this without your help. I really hope we can stay in contact. **Andreas**, you were always very kind and helpful to me. Thanks for correcting my thesis. **Michelle,** thanks for critically reading and commenting on our manuscript. I enjoyed Friday fikas you used to plan. I will miss it now. **Jenna,** you are so nice and helpful. Thanks for critically reading and commenting on the thesis in a short period of time and for arranging the 'Thanksgiving party' every year. **Ulrika,** I should learn how to keep clean and organized lab bench and office desk from you! Thanks for inviting me during your birthday celebration. That was a different kind of party, which I enjoyed a lot. Thanks to **Agata, Babett, Peter, Falko** and **Martin,** you all are very helpful and nice persons. All former members of the KEK group: thanks to **Pernilla, Julian, Rebecca, Hangmao, Hasan, Niklas, Ann,** for always being so helpful. **Mickaël,** thanks for your scientific and non-scientific discussions and introducing me to the

excellent French cheeses. I really enjoyed a lot when working with you. **Elisabeth,** thanks for your nice and excellent work, which helped me a lot to prepare our manuscript.

I wish to thank all the people that I have collaborated with throughout the years, especially **Dr. Luke Buchanan**, **Dr. Xuefeng Zhu**, **Prof. A. Francis Stewart**, **Prof. Michael Grunstein** and **Dr. Steven Sanders.**

Thanks for all the help I received from IT & Administrative departments of Södertörn University and Karolinska Institute, **Sam, Jhohan, Asim, Erik, Rickard, Jonas, Anders and Therese.** Special thanks to **Marie Granroth** and **Monica Ahlberg**. Thanks to **Marika Rönnholm**, **Fredrik Fagerström-Billai** and **David Brodin** from Bioinformatics and Expression Analysis core facility.

AWR group: I would like to thank **Prof. Anthony Wright** for his help and suggestions. **Helmi, Yongtao, Chiounan, Anna, Azadeh** and **Monica,** thanks for your help.

Out of the group: **Satish, Gayathri, Prasad, Swati, Sagar da, Moumita di**, **Kaustuv da and Madhumita,** thanks for your friendship, Indian foods and scientific & non-scientific discussions. It would have been a very lonely life in Sweden without you people. **Jiyoti and Rahul:** I enjoyed your company**, Jiyoti's** special 'gajar ka halwa' (a special Indian Sweet dish) and ABC of SLR camera from **Rahul.** Thanks to **Nalini, Ullasa, Neha** and **Abhiman** for your friendship.

PPS group: Indian Institute of Science (IISc), Bangalore: I owe my deepest gratitude to **Dr. Parag Sadhale** for giving me the opportunity to work in his group. Thanks for introducing me with the microarray technology. Thanks to **Vijender, Maya, Aruna, Binu, Sunanda, Arti** and **Punit** for being nice colleagues and friends. I will not forget those good old days in IISc.

Special thanks to **Angshuman** and **Sagarika** at IISc Bangalore for their help whenever I needed. **Sagarika**, I will not forget your special Indian dish ("data chachari") which I am missing in Sweden.

NCST, Bangalore India: It is an honor for me to thank **Amrit Raj Misra**, **Dr. K.S. Selvakumar** and **R.K.V.S. Raman** for introducing me to different software techniques and especially to bioinformatics. All my friends at NCST, especially to **Mahesh, Nischaya, Akshay (id1), Dhanashri, Akshay, Nahid, Vandana, Pradnya, Subhasish, Sourabh** and **Shikhar.** I will not forget those busy project submission days when we even spent several nights to complete our projects.

Kolkata, India: I would like to thank all my teachers and friends in Kolkata; **Snehanshu, Abir, Sandeep, Sudip, Sucharita, Anirban, Chiradip.** Special thanks to **Sarita,** for her kind help during my MSc studies and for her true friendship.

My hometown Deoghar, India: It is a pleasure to thank my all teachers from school, art school and college, all friends and all well-wishers. Special thanks to **Dr. Aurobinda Chatterjee** for teaching chemistry and his guidance. I would like to

thank **Saumendra Mukherjee** for his true and sensible suggestions. Thanks to **Prof. Chitta Mitra** and **Annapurna Mitra** for their true guidance and good wishes.

This thesis would not have been possible without the love and encouragement from my dear family. I wish to thank all my family members at Hooghly Ghat, Kolkata, Krishnanagar, Baharampur, Pune and Kolhapur. Special thanks to **Atanu Rakshit & family, Nandita Sinha & family, Asit Choudhury & family, Saurendranath De & family** and **Prashanta Sinha & family**. My uncle **Shakti Prasad Sinha** and aunt **Alpana Sinha** for always being supportive. **Subham, Sohoni** and **Subhasish** for entertaining me**.** Thanks to **Shashikant Mirajkar, Ratnaprabha Mirajkar, Abhijit Mirajkar** & **Dr. Ulka Bhokare** for your endless support and care.

I would like to thank my late grandmother **Raj Lakshmi Sinha** for her love and nice stories.

Thanks to my brother **Partha Sinha** (Dada) and my sister-in-law **Vrinda Deshpande** for their support and motivations. I always found you standing by me. Very special thanks to my mother (Ma) **Jharna Sinha** and my father (Baba) **Sunil Sinha** for their love, support, encouragement and giving me the freedom to follow my own choice in life. My wife **Netra,** your unconditional love, affection and great care always encouraged me. Without your support, I could never have achieved this.

10. REFERENCES

Agger, K., Christensen, J., Cloos, P.A. and Helin, K. (2008) The emerging functions of histone demethylases. *Curr Opin Genet Dev*, **18**, 159-168.

Agger, K., Cloos, P.A., Christensen, J., Pasini, D., Rose, S., Rappsilber, J., Issaeva, I., Canaani, E., Salcini, A.E. and Helin, K. (2007) UTX and JMJD3 are histone H3K27 demethylases involved in HOX gene regulation and development. *Nature*, **449**, 731-734.

Ahn, S.H., Cheung, W.L., Hsu, J.Y., Diaz, R.L., Smith, M.M. and Allis, C.D. (2005) Sterile 20 kinase phosphorylates histone H2B at serine 10 during hydrogen peroxide-induced apoptosis in S. cerevisiae. *Cell*, **120**, 25-36.

Alberts, B. and National Center for Biotechnology Information (U.S.). (2002) *Molecular biology of the cell*. Garland Science, New York.

Alen, C., Kent, N.A., Jones, H.S., O'Sullivan, J., Aranda, A. and Proudfoot, N.J. (2002) A role for chromatin remodeling in transcriptional termination by RNA polymerase II. *Mol Cell*, **10**, 1441-1452.

Allard, S., Utley, R.T., Savard, J., Clarke, A., Grant, P., Brandl, C.J., Pillus, L., Workman, J.L. and Cote, J. (1999) NuA4, an essential transcription adaptor/histone H4 acetyltransferase complex containing Esa1p and the ATM-related cofactor Tra1p. *Embo J*, **18**, 5108-5119.

Allis, C.D., Berger, S.L., Cote, J., Dent, S., Jenuwien, T., Kouzarides, T., Pillus, L., Reinberg, D., Shi, Y., Shiekhattar, R., Shilatifard, A., Workman, J. and Zhang, Y. (2007) New nomenclature for chromatin-modifying enzymes. *Cell*, **131**, 633-636.

Amir, R.E., Van den Veyver, I.B., Wan, M., Tran, C.Q., Francke, U. and Zoghbi, H.Y. (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet*, **23**, 185-188.

Amor, D.J., Kalitsis, P., Sumer, H. and Choo, K.H. (2004) Building the centromere: from foundation proteins to 3D organization. *Trends Cell Biol*, **14**, 359-368.

Armstrong, J.A., Papoulas, O., Daubresse, G., Sperling, A.S., Lis, J.T., Scott, M.P. and Tamkun, J.W. (2002) The Drosophila BRM complex facilitates global transcription by RNA polymerase II. *Embo J*, **21**, 5245-5254.

Bannister, A.J., Schneider, R., Myers, F.A., Thorne, A.W., Crane-Robinson, C. and Kouzarides, T. (2005) Spatial distribution of di- and tri-methyl lysine 36 of histone H3 at active genes. *J Biol Chem*, **280**, 17732-17736.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I. and Zhao, K. (2007) High-resolution profiling of histone methylations in the human genome. *Cell*, **129**, 823-837.

Bell, O., Wirbelauer, C., Hild, M., Scharf, A.N., Schwaiger, M., MacAlpine, D.M., Zilbermann, F., van Leeuwen, F., Bell, S.P., Imhof, A., Garza, D., Peters, A.H. and Schubeler, D. (2007) Localized H3K36 methylation states define histone H4K16 acetylation during transcriptional elongation in Drosophila. *Embo J*, **26**, 4974-4984.

Bjerling, P., Silverstein, R.A., Thon, G., Caudy, A., Grewal, S. and Ekwall, K. (2002) Functional divergence between histone deacetylases in fission yeast by distinct cellular localization and in vivo specificity. *Mol Cell Biol*, **22**, 2170-2181.

Braunstein, M., Rose, A.B., Holmes, S.G., Allis, C.D. and Broach, J.R. (1993) Transcriptional silencing in yeast is associated with reduced nucleosome acetylation. *Genes Dev*, **7**, 592-604.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P. and Workman, J.L. (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell*, **123**, 581-592.

Cedar, H. and Bergman, Y. (2009) Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet*, **10**, 295-304.
Chen, D. and Guarente, L. (2007) SIR2: a potential target for calorie restriction mimetics. *Trends Mol Med*, **13**, 64-71.
Chen, D., Toone, W.M., Mata, J., Lyne, R., Burns, G., Kivinen, K., Brazma, A., Jones, N. and Bahler, J. (2003) Global transcriptional responses of fission yeast to environmental stress. *Mol Biol Cell*, **14**, 214-229.
Cheung, W.L., Briggs, S.D. and Allis, C.D. (2000) Acetylation and chromosomal functions. *Curr Opin Cell Biol*, **12**, 326-333.
De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G. and Natoli, G. (2007) The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. *Cell*, **130**, 1083-1094.
DeLange, R.J., Fambrough, D.M., Smith, E.L. and Bonner, J. (1969) Calf and pea histone IV. II. The complete amino acid sequence of calf thymus histone IV; presence of epsilon-N-acetyllysine. *J Biol Chem*, **244**, 319-334.
Deuring, R., Fanti, L., Armstrong, J.A., Sarte, M., Papoulas, O., Prestel, M., Daubresse, G., Verardo, M., Moseley, S.L., Berloco, M., Tsukiyama, T., Wu, C., Pimpinelli, S. and Tamkun, J.W. (2000) The ISWI chromatin-remodeling protein is required for gene expression and the maintenance of higher order chromatin structure in vivo. *Mol Cell*, **5**, 355-365.
Djupedal, I., Durand-Dubief, M., Sinha, I., Varga-Weiz, P. and Ekwall, K. (2010) Differential Genome-wide Occupancies of RNA PolymeraseII Subunits Rpb4 and Rpb7 in Fission Yeast. *In press*.
Durand-Dubief, M., Sinha, I., Fagerstrom-Billai, F., Bonilla, C., Wright, A., Grunstein, M. and Ekwall, K. (2007) Specific functions for the fission yeast Sirtuins Hst2 and Hst4 in gene regulation and retrotransposon silencing. *Embo J*.
Dutnall, R.N. (2003) Cracking the histone code: one, two, three methyls, you're out! *Mol Cell*, **12**, 3-4.
Dutrow, N., Nix, D.A., Holt, D., Milash, B., Dalley, B., Westbroek, E., Parnell, T.J. and Cairns, B.R. (2008) Dynamic transcriptome of Schizosaccharomyces pombe shown by RNA-DNA hybrid mapping. *Nat Genet*, **40**, 977-986.
Eden, S., Hashimshony, T., Keshet, I., Cedar, H. and Thorne, A.W. (1998) DNA methylation models histone acetylation. *Nature*, **394**, 842.
Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A. and Gehrke, C. (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res*, **10**, 2709-2721.
Ekwall, K. (2005) Genome-wide analysis of HDAC function. *Trends Genet*, **21**, 608-615.
En Li, A.B. (2007) DNA Methylation in Mammals. In C. David Allis, T.J., Danny Reinberg (ed.), *EPIGENETICS*. Cold Spring Harbor Laboratory Press, NY.
Esteller, M. (2007) Cancer epigenomics: DNA methylomes and histone-modification maps. *Nat Rev Genet*, **8**, 286-298.
Fagerstrom-Billai, F., Durand-Dubief, M., Ekwall, K. and Wright, A.P. (2007) Individual subunits of the Ssn6-Tup11/12 corepressor are selectively required for repression of different target genes. *Mol Cell Biol*, **27**, 1069-1082.
Fang, J., Feng, Q., Ketel, C.S., Wang, H., Cao, R., Xia, L., Erdjument-Bromage, H., Tempst, P., Simon, J.A. and Zhang, Y. (2002) Purification and functional characterization of SET8, a nucleosomal histone H4-lysine 20-specific methyltransferase. *Curr Biol*, **12**, 1086-1099.
Felsenfeld, G. (2007) A Brief History of Epigenetics. In C. David Allis, T.J., Danny Reinberg (ed.), *EPIGENETICS*. Cold Spring Harbor Laboratory Press, NY, p. 16.
Fingerman, I.M., Wu, C.L., Wilson, B.D. and Briggs, S.D. (2005) Global loss of Set1-mediated H3 Lys4 trimethylation is associated with silencing defects in Saccharomyces cerevisiae. *J Biol Chem*, **280**, 28761-28765.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E. and Mello, C.C. (1998) Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. *Nature*, **391**, 806-811.
Flaus, A., Martin, D.M., Barton, G.J. and Owen-Hughes, T. (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res*, **34**, 2887-2905.
Fraga, M.F., Ballestar, E., Villar-Garea, A., Boix-Chornet, M., Espada, J., Schotta, G., Bonaldi, T., Haydon, C., Ropero, S., Petrie, K., Iyer, N.G., Perez-Rosado, A., Calvo, E., Lopez, J.A., Cano, A., Calasanz, M.J., Colomer, D., Piris, M.A., Ahn, N., Imhof, A., Caldas, C., Jenuwein, T. and Esteller, M. (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nat Genet*, **37**, 391-400.
Freeman-Cook, L.L., Gomez, E.B., Spedale, E.J., Marlett, J., Forsburg, S.L., Pillus, L. and Laurenson, P. (2005) Conserved locus-specific silencing functions of Schizosaccharomyces pombe sir2+. *Genetics*, **169**, 1243-1260.
Fukagawa, T., Nogami, M., Yoshikawa, M., Ikeno, M., Okazaki, T., Takami, Y., Nakayama, T. and Oshimura, M. (2004) Dicer is essential for formation of the heterochromatin structure in vertebrate cells. *Nat Cell Biol*, **6**, 784-791.
Gentleman, R., Carey, V., Huber, W., Irizarry, R. and Dudoit, S. (2005) *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer.
Gill, G. (2004) SUMO and ubiquitin in the nucleus: different functions, similar mechanisms? *Genes Dev*, **18**, 2046-2059.
Goldmark, J.P., Fazzio, T.G., Estep, P.W., Church, G.M. and Tsukiyama, T. (2000) The Isw2 chromatin remodeling complex represses early meiotic genes upon recruitment by Ume6p. *Cell*, **103**, 423-433.
Gordon, M., Holt, D.G., Panigrahi, A., Wilhelm, B.T., Erdjument-Bromage, H., Tempst, P., Bahler, J. and Cairns, B.R. (2007) Genome-wide dynamics of SAPHIRE, an essential complex for gene activation and chromatin boundaries. *Mol Cell Biol*, **27**, 4058-4069.
Grishok, A., Sinskey, J.L. and Sharp, P.A. (2005) Transcriptional silencing of a transgene by RNAi in the soma of C. elegans. *Genes Dev*, **19**, 683-696.
Hahne, F., Huber, W., Gentleman, R. and Falcon, S. (2008) *Bioconductor Case Studies*. Springer.
Hall, I.M., Shankaranarayana, G.D., Noma, K., Ayoub, N., Cohen, A. and Grewal, S.I. (2002) Establishment and maintenance of a heterochromatin domain. *Science*, **297**, 2232-2237.
Hashimshony, T., Zhang, J., Keshet, I., Bustin, M. and Cedar, H. (2003) The role of DNA methylation in setting up chromatin structure during development. *Nat Genet*, **34**, 187-192.
Heckman, D.S., Geiser, D.M., Eidell, B.R., Stauffer, R.L., Kardos, N.L. and Hedges, S.B. (2001) Molecular evidence for the early colonization of land by fungi and plants. *Science*, **293**, 1129-1133.
Henikoff, S. and Ahmad, K. (2005) Assembly of variant histones into chromatin. *Annu Rev Cell Dev Biol*, **21**, 133-153.
Henikoff, S., Smith,MM. . (2007) Histone Variants and Epigenetics. In C. David Allis, T.J., Danny Reinberg (ed.), *EPIGENETICS*. Cold Spring Harbor Laboratory Press, NY.
Holliday, R. (1994) Epigenetics: an overview. *Dev Genet*, **15**, 453-457.
Holliday, R. and Pugh, J.E. (1975) DNA modification mechanisms and gene activity during development. *Science*, **187**, 226-232.
Huber, W., Irizarry, R. and Gentleman, R. (2005) Preprocessing Overview. In Gentleman, R., Carey, V., Huber, W., Irizarry, R. and Dudoit, S. (eds.), *Bioinformatics and Computational Biology Solutions Using R and Bioconductor*. Springer.
Iizuka, M. and Smith, M.M. (2003) Functional consequences of histone modifications. *Curr Opin Genet Dev*, **13**, 154-160.

Jacobs, S.A. and Khorasanizadeh, S. (2002) Structure of HP1 chromodomain bound to a lysine 9-methylated histone H3 tail. *Science*, **295**, 2080-2083.
Jenuwein, T. and Allis, C.D. (2001) Translating the histone code. *Science*, **293**, 1074-1080.
John, S., Howe, L., Tafrov, S.T., Grant, P.A., Sternglanz, R. and Workman, J.L. (2000) The something about silencing protein, Sas3, is the catalytic subunit of NuA3, a yTAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex. *Genes Dev*, **14**, 1196-1208.
Johnson, L.M., Kayne, P.S., Kahn, E.S. and Grunstein, M. (1990) Genetic evidence for an interaction between SIR3 and histone H4 in the repression of the silent mating loci in Saccharomyces cerevisiae. *Proc Natl Acad Sci U S A*, **87**, 6286-6290.
Joshi, A.A. and Struhl, K. (2005) Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell*, **20**, 971-978.
Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H. and Razin, A. (1992) Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes Dev*, **6**, 705-714.
Kamakaka, R.T. and Biggins, S. (2005) Histone variants: deviants? *Genes Dev*, **19**, 295-310.
Kanellopoulou, C., Muljo, S.A., Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M. and Rajewsky, K. (2005) Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev*, **19**, 489-501.
Kawahara, T.L., Michishita, E., Adler, A.S., Damian, M., Berber, E., Lin, M., McCord, R.A., Ongaigui, K.C., Boxer, L.D., Chang, H.Y. and Chua, K.F. (2009) SIRT6 links histone H3 lysine 9 deacetylation to NF-kappaB-dependent gene expression and organismal life span. *Cell*, **136**, 62-74.
Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., Boone, C., Emili, A., Weissman, J.S., Hughes, T.R., Strahl, B.D., Grunstein, M., Greenblatt, J.F., Buratowski, S. and Krogan, N.J. (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell*, **123**, 593-605.
Kimura, A., Matsubara, K. and Horikoshi, M. (2005) A decade of histone acetylation: marking eukaryotic chromosomes with specific codes. *J Biochem*, **138**, 647-662.
Kobori, T., Yoshino, T., Sugiyama, S. and Ohtani, T. (2003) Hierarchical chromatin structure of Schizosaccharomyces pombe revealed by atomic force microscopy. *Curr Microbiol*, **47**, 404-407.
Kouzarides, T. (2007) Chromatin modifications and their function. *Cell*, **128**, 693-705.
Kramer, O.H., Gottlicher, M. and Heinzel, T. (2001) Histone deacetylase as a therapeutic target. *Trends Endocrinol Metab*, **12**, 294-300.
Krogan, N.J., Dover, J., Wood, A., Schneider, J., Heidt, J., Boateng, M.A., Dean, K., Ryan, O.W., Golshani, A., Johnston, M., Greenblatt, J.F. and Shilatifard, A. (2003) The Paf1 complex is required for histone H3 methylation by COMPASS and Dot1p: linking transcriptional elongation to histone methylation. *Mol Cell*, **11**, 721-729.
Kulesh, D.A., Clive, D.R., Zarlenga, D.S. and Greene, J.J. (1987) Identification of interferon-modulated proliferation-related cDNA sequences. *Proc Natl Acad Sci U S A*, **84**, 8453-8457.
Kuo, M.H., Zhou, J., Jambeck, P., Churchill, M.E. and Allis, C.D. (1998) Histone acetyltransferase activity of yeast Gcn5p is required for the activation of target genes in vivo. *Genes Dev*, **12**, 627-639.
Kurdistani, S.K., Tavazoie, S. and Grunstein, M. (2004) Mapping global histone acetylation patterns to gene expression. *Cell*, **117**, 721-733.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K. and Jenuwein, T. (2001) Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature*, **410**, 116-120.
Lan, F., Nottke, A.C. and Shi, Y. (2008) Mechanisms involved in the regulation of histone lysine demethylases. *Curr Opin Cell Biol*, **20**, 316-325.
Lan, F., Zaratiegui, M., Villen, J., Vaughn, M.W., Verdel, A., Huarte, M., Shi, Y., Gygi, S.P., Moazed, D., Martienssen, R.A. and Shi, Y. (2007) S. pombe LSD1 homologs regulate heterochromatin propagation and euchromatic gene transcription. *Mol Cell*, **26**, 89-101.
Lantermann, A., Strålfors, A., Fagerstrom-Billai, F., Korber, P. and Ekwall, K. (2009) Genome-wide mapping of nucleosome positions in S. pombe. *In press*.
Leach, T.J., Mazzeo, M., Chotkowski, H.L., Madigan, J.P., Wotring, M.G. and Glaser, R.L. (2000) Histone H2A.Z is widely but nonrandomly distributed in chromosomes of Drosophila melanogaster. *J Biol Chem*, **275**, 23267-23272.
Lee, K.K. and Workman, J.L. (2007) Histone acetyltransferase complexes: one size doesn't fit all. *Nat Rev Mol Cell Biol*, **8**, 284-295.
Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.P., Reinberg, D., Di Croce, L. and Shiekhattar, R. (2007a) Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science*, **318**, 447-450.
Lee, W., Tillo, D., Bray, N., Morse, R.H., Davis, R.W., Hughes, T.R. and Nislow, C. (2007b) A high-resolution atlas of nucleosome occupancy in yeast. *Nat Genet*, **39**, 1235-1244.
Lennartsson, A. and Ekwall, K. (2009) Histone modification patterns and epigenetic codes. *Biochim Biophys Acta*.
Li, B., Carey, M. and Workman, J.L. (2007) The role of chromatin during transcription. *Cell*, **128**, 707-719.
Lin, R.J., Nagy, L., Inoue, S., Shao, W., Miller, W.H., Jr. and Evans, R.M. (1998) Role of the histone deacetylase complex in acute promyelocytic leukaemia. *Nature*, **391**, 811-814.
Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N. and Rando, O.J. (2005) Single-nucleosome mapping of histone modifications in S. cerevisiae. *PLoS Biol*, **3**, e328.
Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F. and Richmond, T.J. (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature*, **389**, 251-260.
Malik, H.S. and Henikoff, S. (2003) Phylogenomics of the nucleosome. *Nat Struct Biol*, **10**, 882-891.
Martin, C. and Zhang, Y. (2005) The diverse functions of histone lysine methylation. *Nat Rev Mol Cell Biol*, **6**, 838-849.
Meissner, A., Mikkelsen, T.S., Gu, H., Wernig, M., Hanna, J., Sivachenko, A., Zhang, X., Bernstein, B.E., Nusbaum, C., Jaffe, D.B., Gnirke, A., Jaenisch, R. and Lander, E.S. (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature*, **454**, 766-770.
Millar, C.B. and Grunstein, M. (2006) Genome-wide patterns of histone modifications in yeast. *Nat Rev Mol Cell Biol*, **7**, 657-666.
Milne, T.A., Briggs, S.D., Brock, H.W., Martin, M.E., Gibbs, D., Allis, C.D. and Hess, J.L. (2002) MLL targets SET domain methyltransferase activity to Hox gene promoters. *Mol Cell*, **10**, 1107-1117.
Mizzen, C.A., Yang, X.J., Kokubo, T., Brownell, J.E., Bannister, A.J., Owen-Hughes, T., Workman, J., Wang, L., Berger, S.L., Kouzarides, T., Nakatani, Y. and Allis, C.D. (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell*, **87**, 1261-1270.
Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M. and Schubeler, D. (2008) Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. *Mol Cell*, **30**, 755-766.

Monk, M., Boubelik, M. and Lehnert, S. (1987) Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development*, **99**, 371-382.
Morris, S.A., Shibata, Y., Noma, K., Tsukamoto, Y., Warren, E., Temple, B., Grewal, S.I. and Strahl, B.D. (2005) Histone H3 K36 methylation is associated with transcription elongation in Schizosaccharomyces pombe. *Eukaryot Cell*, **4**, 1446-1454.
Murray, K. (1964) The Occurrence of Epsilon-N-Methyl Lysine in Histones. *Biochemistry*, **3**, 10-15.
Nakayama, J., Rice, J.C., Strahl, B.D., Allis, C.D. and Grewal, S.I. (2001) Role of histone H3 lysine 9 methylation in epigenetic control of heterochromatin assembly. *Science*, **292**, 110-113.
Ng, H.H., Robert, F., Young, R.A. and Struhl, K. (2002) Genome-wide location and regulated recruitment of the RSC nucleosome-remodeling complex. *Genes Dev*, **16**, 806-819.
Ng, H.H., Robert, F., Young, R.A. and Struhl, K. (2003) Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Mol Cell*, **11**, 709-719.
Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., Lis, J.T., Allis, C.D. and Reinberg, D. (2002) PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Mol Cell*, **9**, 1201-1213.
Noma, K. and Grewal, S.I. (2002) Histone H3 lysine 4 methylation is mediated by Set1 and promotes maintenance of active chromatin states in fission yeast. *Proc Natl Acad Sci U S A*, **99 Suppl 4**, 16438-16445.
Okitsu, C.Y. and Hsieh, C.L. (2007) DNA methylation dictates histone H3K4 methylation. *Mol Cell Biol*, **27**, 2746-2757.
Olins, A.L. and Olins, D.E. (1974) Spheroid chromatin units (v bodies). *Science*, **183**, 330-332.
Olsson, T.G., Ekwall, K., Allshire, R.C., Sunnerhagen, P., Partridge, J.F. and Richardson, W.A. (1998) Genetic characterisation of hda1+, a putative fission yeast histone deacetylase gene. *Nucleic Acids Res*, **26**, 3247-3254.
Ooi, S.K., Qiu, C., Bernstein, E., Li, K., Jia, D., Yang, Z., Erdjument-Bromage, H., Tempst, P., Lin, S.P., Allis, C.D., Cheng, X. and Bestor, T.H. (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature*, **448**, 714-717.
Opel, M., Lando, D., Bonilla, C., Trewick, S.C., Boukaba, A., Walfridsson, J., Cauwood, J., Werler, P.J., Carr, A.M., Kouzarides, T., Murzina, N.V., Allshire, R.C., Ekwall, K. and Laue, E.D. (2007) Genome-Wide Studies of Histone Demethylation Catalysed by the Fission Yeast Homologues of Mammalian LSD1. *PLoS ONE*, **2**, e386.
Pal-Bhadra, M., Leibovitch, B.A., Gandhi, S.G., Rao, M., Bhadra, U., Birchler, J.A. and Elgin, S.C. (2004) Heterochromatic silencing and HP1 localization in Drosophila are dependent on the RNAi machinery. *Science*, **303**, 669-672.
Palmer, D.K., O'Day, K., Trong, H.L., Charbonneau, H. and Margolis, R.L. (1991) Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone. *Proc Natl Acad Sci U S A*, **88**, 3734-3738.
Perez-Martin, J. and Johnson, A.D. (1998) Mutations in chromatin components suppress a defect of Gcn5 protein in Saccharomyces cerevisiae. *Mol Cell Biol*, **18**, 1049-1054.
Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., Zeitlinger, J., Lewitter, F., Gifford, D.K. and Young, R.A. (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell*, **122**, 517-527.
Pruitt, K., Zinn, R.L., Ohm, J.E., McGarvey, K.M., Kang, S.H., Watkins, D.N., Herman, J.G. and Baylin, S.B. (2006) Inhibition of SIRT1 reactivates silenced

cancer genes without loss of promoter DNA hypermethylation. *PLoS Genet*, **2**, e40.
Rando, O.J. (2007) Global patterns of histone modifications. *Curr Opin Genet Dev*, **17**, 94-99.
Rea, S., Eisenhaber, F., O'Carroll, D., Strahl, B.D., Sun, Z.W., Schmid, M., Opravil, S., Mechtler, K., Ponting, C.P., Allis, C.D. and Jenuwein, T. (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature*, **406**, 593-599.
Ren, B., Robert, F., Wyrick, J.J., Aparicio, O., Jennings, E.G., Simon, I., Zeitlinger, J., Schreiber, J., Hannett, N., Kanin, E., Volkert, T.L., Wilson, C.J., Bell, S.P. and Young, R.A. (2000) Genome-wide location and function of DNA binding proteins. *Science*, **290**, 2306-2309.
Riggs, A.D. (1975) X inactivation, differentiation, and DNA methylation. *Cytogenet Cell Genet*, **14**, 9-25.
Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.A. and Barrell, B. (2000) Artemis: sequence visualization and annotation. *Bioinformatics*, **16**, 944-945.
Sanders, S.L., Portoso, M., Mata, J., Bahler, J., Allshire, R.C. and Kouzarides, T. (2004) Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell*, **119**, 603-614.
Santos-Rosa, H., Schneider, R., Bannister, A.J., Sherriff, J., Bernstein, B.E., Emre, N.C., Schreiber, S.L., Mellor, J. and Kouzarides, T. (2002) Active genes are tri-methylated at K4 of histone H3. *Nature*, **419**, 407-411.
Sarma, K. and Reinberg, D. (2005) Histone variants meet their match. *Nat Rev Mol Cell Biol*, **6**, 139-149.
Schaft, D., Roguev, A., Kotovic, K.M., Shevchenko, A., Sarov, M., Shevchenko, A., Neugebauer, K.M. and Stewart, A.F. (2003) The histone 3 lysine 36 methyltransferase, SET2, is involved in transcriptional elongation. *Nucleic Acids Res*, **31**, 2475-2482.
Schena, M., Shalon, D., Davis, R.W. and Brown, P.O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science*, **270**, 467-470.
Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D. and Jenuwein, T. (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev*, **18**, 1251-1262.
Seol, J.H., Kim, H.J., Yang, Y.J., Kim, S.T., Youn, H.D., Han, J.W., Lee, H.W. and Cho, E.J. (2006) Different roles of histone H3 lysine 4 methylation in chromatin maintenance. *Biochem Biophys Res Commun*, **349**, 463-470.
Shahbazian, M.D. and Grunstein, M. (2007) Functions of Site-Specific Histone Acetylation and Deacetylation. *Annu Rev Biochem.*
Shankaranarayana, G.D., Motamedi, M.R., Moazed, D. and Grewal, S.I. (2003) Sir2 regulates histone H3 lysine 9 methylation and heterochromatin assembly in fission yeast. *Curr Biol*, **13**, 1240-1246.
Shen, X., Mizuguchi, G., Hamiche, A. and Wu, C. (2000) A chromatin remodelling complex involved in transcription and DNA processing. *Nature*, **406**, 541-544.
Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., Casero, R.A. and Shi, Y. (2004) Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell*, **119**, 941-953.
Shilatifard, A. (2006) Chromatin Modifications by Methylation and Ubiquitination: Implications in the Regulation of Gene Expression. *Annu Rev Biochem.*
Sinha, I., Jemt, E., Durand-Dubief, M., Strålfors, A., Sanders, S. and Ekwall, K. (2010) Genome wide mapping suggests different roles for H4K20me1 and H4K20me3 in gene expression. *In press*.
Sinha, I., Wiren, M. and Ekwall, K. (2006) Genome-wide patterns of histone modifications in fission yeast. *Chromosome Res*, **14**, 95-105.
Spotswood, H.T. and Turner, B.M. (2002) An increasingly complex code. *J Clin Invest*, **110**, 577-582.

Strahl, B.D. and Allis, C.D. (2000) The language of covalent histone modifications. *Nature*, **403**, 41-45.
Strahl, B.D., Grant, P.A., Briggs, S.D., Sun, Z.W., Bone, J.R., Caldwell, J.A., Mollah, S., Cook, R.G., Shabanowitz, J., Hunt, D.F. and Allis, C.D. (2002) Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Mol Cell Biol*, **22**, 1298-1306.
Suka, N., Suka, Y., Carmen, A.A., Wu, J. and Grunstein, M. (2001) Highly specific antibodies determine histone acetylation site usage in yeast heterochromatin and euchromatin. *Mol Cell*, **8**, 473-479.
Tsukada, Y., Fang, J., Erdjument-Bromage, H., Warren, M.E., Borchers, C.H., Tempst, P. and Zhang, Y. (2006) Histone demethylation by a family of JmjC domain-containing proteins. *Nature*, **439**, 811-816.
Tsukiyama, T., Palmer, J., Landel, C.C., Shiloach, J. and Wu, C. (1999) Characterization of the imitation switch subfamily of ATP-dependent chromatin-remodeling factors in Saccharomyces cerevisiae. *Genes Dev*, **13**, 686-697.
Turner, B.M. (1993) Decoding the nucleosome. *Cell*, **75**, 5-8.
Turner, B.M. (2001) *Chromatin and gene regulation : mechanisms in epigenetics*. Blackwell Science, Oxford ; Malden, MA.
Turner, B.M. (2002) Cellular memory and the histone code. *Cell*, **111**, 285-291.
Turner, B.M., O'Neill, L.P. and Allan, I.M. (1989) Histone H4 acetylation in human cells. Frequency of acetylation at different sites defined by immunolabeling with site-specific antibodies. *FEBS Lett*, **253**, 141-145.
Waddington, C.H. (1953) Epigenetics and evolution. *Symp. Soc. Exp. Biol.*, **7**, 186-199.
Wade, P.A., Pruss, D. and Wolffe, A.P. (1997) Histone acetylation: chromatin in action. *Trends Biochem Sci*, **22**, 128-132.
Walfridsson, J., Bjerling, P., Thalen, M., Yoo, E.J., Park, S.D. and Ekwall, K. (2005) The CHD remodeling factor Hrp1 stimulates CENP-A loading to centromeres. *Nucleic Acids Res*, **33**, 2868-2879.
Walfridsson, J., Khorosjutina, O., Matikainen, P., Gustafsson, C.M. and Ekwall, K. (2007) A genome-wide role for CHD remodelling factors and Nap1 in nucleosome disassembly. *Embo J*, **26**, 2868-2879.
van Attikum, H. and Gasser, S.M. (2005) The histone code at DNA breaks: a guide to repair? *Nat Rev Mol Cell Biol*, **6**, 757-765.
Wang, A., Kurdistani, S.K. and Grunstein, M. (2002) Requirement of Hos2 histone deacetylase for gene activity in yeast. *Science*, **298**, 1412-1414.
Wang, Z., Zang, C., Rosenfeld, J.A., Schones, D.E., Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Peng, W., Zhang, M.Q. and Zhao, K. (2008) Combinatorial patterns of histone acetylations and methylations in the human genome. *Nat Genet*, **40**, 897-903.
Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M. and Schubeler, D. (2007) Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. *Nat Genet*, **39**, 457-466.
Whitehouse, I., Rando, O.J., Delrow, J. and Tsukiyama, T. (2007) Chromatin remodelling at promoters suppresses antisense transcription. *Nature*, **450**, 1031-1035.
Wieland, G., Orthaus, S., Ohndorf, S., Diekmann, S. and Hemmerich, P. (2004) Functional complementation of human centromere protein A (CENP-A) by Cse4p from Saccharomyces cerevisiae. *Mol Cell Biol*, **24**, 6620-6630.
Wilhelm, B.T., Marguerat, S., Watt, S., Schubert, F., Wood, V., Goodhead, I., Penkett, C.J., Rogers, J. and Bahler, J. (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature*, **453**, 1239-1243.
Winkler, G.S., Kristjuhan, A., Erdjument-Bromage, H., Tempst, P. and Svejstrup, J.Q. (2002) Elongator is a histone H3 and H4 acetyltransferase important for normal histone acetylation levels in vivo. *Proc Natl Acad Sci U S A*, **99**, 3517-3522.

Wiren, M., Silverstein, R.A., Sinha, I., Walfridsson, J., Lee, H.M., Laurenson, P., Pillus, L., Robyr, D., Grunstein, M. and Ekwall, K. (2005) Genomewide analysis of nucleosome density histone acetylation and HDAC function in fission yeast. *Embo J*, **24**, 2906-2918.

Vogelauer, M., Rubbi, L., Lucas, I., Brewer, B.J. and Grunstein, M. (2002) Histone acetylation regulates the time of replication origin firing. *Mol Cell*, **10**, 1223-1233.

Vogelauer, M., Wu, J., Suka, N. and Grunstein, M. (2000) Global histone acetylation and deacetylation in yeast. *Nature*, **408**, 495-498.

Wu, J., Suka, N., Carlson, M. and Grunstein, M. (2001) TUP1 utilizes histone H3/H2B-specific HDA1 deacetylase to repress gene activity in yeast. *Mol Cell*, **7**, 117-126.

Xu, L., Zhao, Z., Dong, A., Soubigou-Taconnat, L., Renou, J.P., Steinmetz, A. and Shen, W.H. (2008) Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in Arabidopsis thaliana. *Mol Cell Biol*, **28**, 1348-1360.

Yamane, K., Tateishi, K., Klose, R.J., Fang, J., Fabrizio, L.A., Erdjument-Bromage, H., Taylor-Papadimitriou, J., Tempst, P. and Zhang, Y. (2007) PLU-1 is an H3K4 demethylase involved in transcriptional repression and breast cancer cell proliferation. *Mol Cell*, **25**, 801-812.

Yamane, K., Toumazou, C., Tsukada, Y., Erdjument-Bromage, H., Tempst, P., Wong, J. and Zhang, Y. (2006) JHDM2A, a JmjC-containing H3K9 demethylase, facilitates transcription activation by androgen receptor. *Cell*, **125**, 483-495.

Yang, X., Li, J., Qin, H., Yang, H., Li, J., Zhou, P., Liang, Y. and Han, H. (2005) Mint represses transactivation of the type II collagen gene enhancer through interaction with alpha A-crystallin-binding protein 1. *J Biol Chem*, **280**, 18710-18716.

Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J. and Rando, O.J. (2005) Genome-scale identification of nucleosome positions in S. cerevisiae. *Science*, **309**, 626-630.

Zhang, K., Williams, K.E., Huang, L., Yau, P., Siino, J.S., Bradbury, E.M., Jones, P.R., Minch, M.J. and Burlingame, A.L. (2002) Histone acetylation and deacetylation: identification of acetylation and methylation sites of HeLa histone H4 by mass spectrometry. *Mol Cell Proteomics*, **1**, 500-508.

Zhao, Z., Yu, Y., Meyer, D., Wu, C. and Shen, W.H. (2005) Prevention of early flowering by expression of FLOWERING LOCUS C requires methylation of histone H3 K36. *Nat Cell Biol*, **7**, 1256-1260.

Zheng, C. and Hayes, J.J. (2003) Intra- and inter-nucleosomal protein-DNA interactions of the core histone tail domains in a model system. *J Biol Chem*, **278**, 24217-24224.

Zilberman, D., Cao, X. and Jacobsen, S.E. (2003) ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. *Science*, **299**, 716-719.

Parts of the cover picture were reprinted and modified with permission from Cedar et al., Nature Reviews, 2009.