IDENTIFICATION AND CHARACTERIZATION OF HISTONE H3 LYSINE 23 METHYLATION IN GERMLINE CHROMATIN

By
Romeo Papazyan

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

Baltimore, Maryland
February, 2014

ABSTRACT

Eukaryotic DNA and histone proteins are packaged into chromatin that can occur in two cytologically distinct states: heterochromatin, which is highly compacted and transcriptionally repressive; or euchromatin, which is relatively open and transcriptionally permissive. Increasing experimental evidence suggest that the structure and function of chromatin is governed, in part, by histone post-translational modifications (PTMs), but their characterization is generally protracted because of the biological complexities associated with higher eukaryotes. Moreover, recent studies have shown that euchromatin and heterochromatin represent the extremes of a continuum of physiological chromatin states, which may be encoded for by distinct combinations of histone PTMs. These states also have transcription-independent roles during interphase, mitosis, and meiosis that contribute to, among other activities, maintaining genomic stability.

To screen for novel heterochromatin associated histone PTMs, we took advantage of the nuclear dimorphism afforded by the single-celled ciliate *Tetrahymena thermophila*. This model organism has two physically separable nuclei with its cytoplasm: a heterochromatic micronucleus (germline genome), and a euchromatic macronucleus (somatic genome). In this dissertation, I present our work identifying and characterizing histone H3 lysine 23 methylation (H3K23me) in the micronucleus. We report that H3K23 trimethylation (H3K23me3) is a micronucleus-specific heterochromatin modification that becomes particularly enriched during meiosis. Loss of H3K23me3, through deletion of its putative methyltransferase Ezl3p, causes mislocalization of meiosis-induced DNA double-strand breaks to heterochromatin, and decreases progeny viability. We extend the

significance of H3K23me3 to higher eukaryotes by showing its meiotic upregulation is conserved in germlines of *C. elegans* and mice. In summary, our results suggest that the evolutionary conserved H3K23me3 plays an important role in protecting meiotic, germline heterochromatin from programmed DSBs.

DISSERTATION REFEREES

GRADAUTE ADVISOR: **SEAN D. TAVERNA, PH.D.**

ASSISTANT PROFESSOR

DEPARTMENT OF PHARMACOLOGY AND MOLECULAR SCIENCES

DISSERTATION READER: **PHILIP A. COLE, M.D., PH.D.**

DIRECTOR AND E.K. MARSHALL AND THOMAS H. MAREN PROFESSOR

DEPARTMENT OF PHARMACOLOGY AND MOLECULAR SCIENCES

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to all the wonderful people who have supported this work and my scientific pursuits, from my undergraduate research mentor Osvaldo Rey who inspired me to become a scientist, the Pharmacology Department for giving me this opportunity, the many collaborators for sharing their expertise and time, and the Taverna lab members, past and present, for their friendship and efforts throughout this journey.

I am grateful to Sean Taverna, my advisor, for his constant guidance and encouragement. Above all, he has taught me the importance of celebrating all scientific successes, however minor. I would also like to thank my thesis committee, Philip Cole, Jun Liu, Xin Chen, and Herschel Wade, for their comments and insights during our meetings. I must also recognize the generous group of collaborators that provided critical data for this work. In particular, I would like to thank Jessica Chapman for her invaluable contributions to the mass spectrometric identification of histone H3K23 methylation, which was instrumental in starting this project. Additionally, special thanks to Katya Voronina who provided all of the immunofluorescence data concerning the evolutionary conservation of H3K23me3 in worms and mice.

Finally, completion of this thesis would not have materialized without the love and support I have received from family and friends over the years.

TABLE OF CONTENTS

Chapter 3: Loss of Histone H3 Lysine 23 Methylation Leads to Developmental Defects in Tetrahymena

LIST OF FIGURES

LIST OF APPENDIX FIGURES AND TABLES

CHAPTER 1: GENERAL INTRODUCTION

1.1 EPIGENETICS AND CHROMATIN

A single human body contains roughly 250 distinct cell types, each virtually having the same DNA code.[1] Moreover, all human cell types are derived from a single, fertilized egg that after just two divisions produces four isogenic daughter cells with unique developmental trajectories.[2, 3] Understanding how multiple phenotypes arise from a single genome remains a central problem in biology and genetics.

In 1942, Conrad Waddington coined the term epigenetics to relate genes and gene products to development. A modern definition of epigenetics emerged in 1975, when two independent groups proposed that patterns of DNA methylation, a chemical modification of the cytosine-guanine dinucleotide (CpG), could affect gene expression and be inherited during cell division.[4, 5] Since then, epigenetics has been defined as the study of heritable changes in gene expression that are not caused by changes in DNA sequence. In addition to DNA methylation, the epigenome now includes histone proteins and their post-translational modifications, as well as several other factors such as non-coding RNA.[6]

Many decades before the modern definition of epigenetics took shape, its cellular basis was coming to light from cytological studies of eukaryotic DNA. Defined as the physiological form of the eukaryotic genome, chromatin is made up of DNA and histone proteins, as well as other proteins with chromatin modifying activities. Using dyes, Emil Heitz discovered in 1928 that chromatin can exist in structurally dense and loose forms, and he designated these regions as heterochromatin and euchromatin, respectively.[7]

Furthermore, he found that heterochromatin remains tightly packed throughout the entire cell cycle and, therefore, suggested that the underlying DNA was physically inaccessible and functionally inert. He supported this hypothesis by showing that heterochromatic DNA is deficient in genes.[8] Teleologically, these early studies suggested that chromatin structure could regulate gene transcription and other DNA-templated processes by controlling direct access to regulatory DNA elements (*e.g.* promoters and enhancers).

1.2 THE NUCLEOSOME AND HISTONES

While the existence of histones have been recognized since the late 19th century, it wasn't until the 1960s and 1970s that a molecular view chromatin began to unravel. In this period, several studies systematically uncovered the highly organized and structurally repetitive nature of the chromatin fiber.[9-12] Prompted by electron micrographs of highly purified chromatin strands, a model was proposed that likened the basic structure of chromatin to beads on a string.[13, 14] Early work revealed that each repetitive unit, called a nucleosome, is composed of histone proteins and around 200 base pairs (bp) of DNA.[15] Later, micrococcal nucleus digestions of chromatin identified the linker (~50 bp of DNA) and core particle ('bead'; ~150 bp of DNA and histones) regions of the nucleosome.[15] The structural dissection of nucleosomes finally culminated in 1997 with the x-ray crystallographic image of the core particle described by Luger *et al*.[16] The data revealed that roughly 146 bp of DNA wrap around (~1.7 turns) an octamer core of histone proteins containing two copies of the canonical histones H2A, H2B, H4, and H3 **(Figure 1)**. This structure not only provided an atomic view of how DNA is neatly packaged under physiological conditions, but it also provided the basis for the succeeding functional

Figure 1

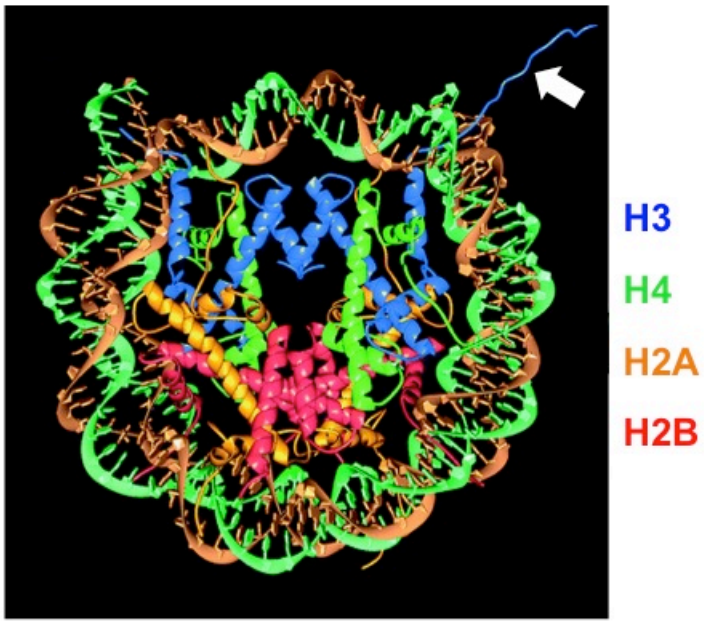


Figure 1. **Crystal structure of the nucleosome**. Shown here is the 2.8Å resolution structure of the nucleosome solved by Luger *et al.*[16] The eight histone proteins are labeled as: blue (H3), green (H4), orange (H2A), and red (H2B). The white arrow points to the N-terminal tail of histone H3 extending from the nucleosome.

studies of histones and their post-translational modifications.

The histone superfamily is categorized into two groups of highly conserved histone proteins: the linker histones and the core histones. Because they bind to the linker DNA region of nucleosomes, histone H1 and its variants constitute the linker histone group. Linker histones primarily play a structural role in chromatin, acting to form higher-ordered states of compaction. Furthermore, studies have shown that linker histones are enriched in heterochromatic regions of the genome and directly contribute to transcriptional silencing.[17, 18] The other group of histones contains the canonical histones and their many variants that make up the core particle of the nucleosome. While each of the core histones have distinct properties, collectively they contribute to processes involving genomic DNA (*e.g.* replication, transcription, and repair) and ultimately help to maintain genomic organization and stability.[19] In all, there are five major histones (H1, H2A, H2B, H4, and H3) and numerous minor variants within the human genome. These variants are less conserved, but they have specialized properties that help diversify chromatin function and structure.

1.3 HISTONE POST-TRANSLATIONAL MODIFICATIONS

In addition to the specific interactions between DNA and histone polypeptides within the nucleosome core, the crystal structure by Luger *et al.* showed that histone N-terminal tails extend beyond the nucleosome, interact with DNA on the outer portion of the nucleosome, and are relatively unstructured **(Figure 1, arrow)**.[16] This view suggested a functional role for histone tails, whose positively charged lysines and arginines closely interacted with the negatively charged backbone of DNA. Additionally, the unstructured

nature of the tails allowed for their extension beyond the nucleosome and exposure to nucleoplasmic proteins. Conforming to this paradigm, histone tails, particularly those of histone H3, are abundantly and dynamically modified with distinct types of PTMs **(Figure 2)**.[20] Anecdotally, most common types of PTMs have been detected on histones, and almost every nucleophilic center on the H3 tail has been modified in some cellular context.

The most well characterized classes of PTMs include acetylation, phosphorylation, and methylation.[20] As discussed below, histone PTMs are dynamically regulated by the balancing activities of enzymes that catalyze the addition of marks (*writers*) and their removal (*erasers*). How histone PTMs encode for biological outputs is not entirely understood, but there is a growing number of known effector proteins with chromatin modifying activities that can bind (*read*) specific types and combinations of histone PTMs. Their reader functions are also largely dependent on the position of PTMs on histone tails.[21] Therefore, specific combinations of histone PTMs have been proposed to constitute a "histone code" or signature that encode for distinct functional outputs.[22-24]

Figure 2

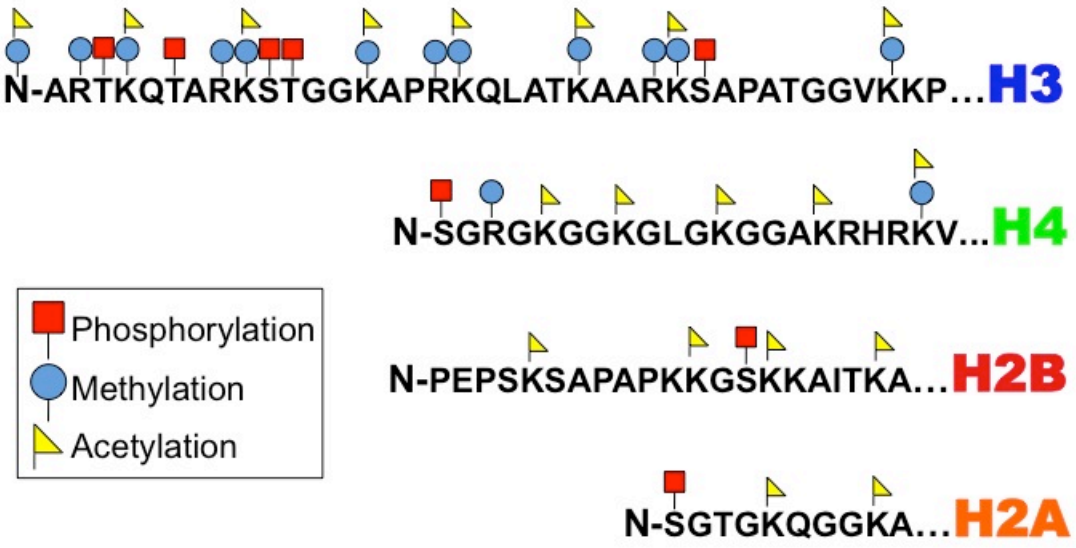


Figure 2. **Post-translational modifications of histone tails**. Histone tail regions are subject to a host of post-translational modification. Shown here is just a small subset of the types and locations of common histone PTMs. Methylated sites include mono-, di-, and tri-methylated forms of lysine residues.

1.3A LYSINE ACETYLATION

It has long been established that acetylation of lysines on histone tails neutralizes the positive charge on the ε-amino group and lowers the affinity of histones for DNA.[11, 25] Coupled with the insights from the Luger *et al.* nucleosome structure, these observations supported a model wherein dynamic acetylation of histone tails leads to the loosening of chromatin structure, facilitating histone eviction, transcription factor binding and gene transcription by RNA polymerase II.[26-28] Several lines of evidence have amassed that functionally link histone acetylation and gene activity. Some of the early examples include a seminal study showing that histones are largely devoid of acetylation in the transcriptionally inactive X chromosome and pericentric heterochromatin of mammals.[29] Another study found that antibodies raised against chemically acetylated histone H4 highly co-purify with chromatin containing the transcriptionally active gene, alpha D globin.[30] Although histone acetylation has strong ties to transcription, several studies have shown that acetylation alone is not sufficient to induce transcription.[31] For example, a recent study showed that pharmacological inhibition of histone deacetylase 3 (HDAC3) causes hyperacetylation, but not transcriptional activation, of target genes.[32] In addition to chromatin dynamics and transcription activity, histone acetylation is also involved in replication dependent and independent nucleosome assembly and histone deposition pathways.[33, 34]

As gene transcription states can range from ‘on’ to ‘off’ depending on environmental or developmental cues, it stands to reason that local histone acetylation is also reversibly regulated. This realization fueled the search for enzymes that catalyzed the acetylation and deacetylation of histones. The long awaited nuclear histone

acetyltransferase (HAT) was reported in 1996.[35] This nuclear HAT was identified in the model organism, *Tetrahymena thermophila,* and found to be a homolog of the yeast GCN5 protein, previously shown to be a transcriptional co-activator.[36] Coincidentally, another study found that the yeast Rap5 protein (a known repressor of transcription) has histone deacetylase (HDAC) activity.[37] The discovery of HATs and HDACs marked a key point in the field of epigenetics, as it provided a critical link between chromatin structure and transcriptional regulation.[38]

Although histone modifications, such as acetylation, can have direct effects on histone and chromatin function by virtue of their chemical properties, a wealth of structural and functional data now show that chromatin-associated proteins contain subdomains that can bind to their cognate histone PTM.[21] These findings represent a new mechanism of protein-protein interactions regulated by histone PTMs. The first histone PTM binding domain to be structurally characterized was the bromodomain, which was proposed to bind to acetyllysine.[39] Subsequent studies solidified the role of the evolutionary conserved bromodomain as a *reader* of acetylated histones.[40]

1.3B SERINE/THREONINE PHOSPHORYLATION

All core histones contain Serine/Threonine/Tyrosine amino acids that can be phosphorylated by kinases and dephosphorylated by phosphatases. Depending on the site, histone phosphorylation can either be involved in transcription, cellular division, or DNA damage recognition and repair.[41] Several proteins such as 14-3-3 and MDC (via the BRCT domain) have been shown to engage chromatin by binding to histones in a phosphorylation dependent manner.[21, 42] Like acetylation, phosphorylation neutralizes the

net positive charge of histone tails, although its effect on higher ordered chromatin packaging changes depending on the cell cycle and therefore is more complex than the effect of acetylation. There are at least 20 different sites known to be phosphorylated in core histones (see Rossetto *et al*.[41] for a detailed review), but here I will focus on two major sites that are relevant to this dissertation.

Phosphorylation of histone H3 serine 10 (H3S10ph) is an important and highly conserved modification that plays a role in chromosome condensation prior to mitosis. Although the mechanism is not clear, loss of H3S10ph leads to genomic instability and problems with segregation during mitosis.[43] H3S10ph also represents a wonderful example of the combinatorial crosstalk of histone PTMs occurring in *cis*. It was shown that heterochromatin protein 1 (HP1), which reads histone H3 lysine 9 methylation (described below), is extruded from chromatin if the flanking H3S10 is phosphorylated during mitosis.[44] As H3S10 is highly conserved across eukaryotic evolution, hyperphosphorylation of this site has generally been linked to mitosis in a wide range of organisms.[45] Notwithstanding, the presence of this mark during interphase has also been linked to transcriptional activation.[41, 44, 45]

Another well-known histone phosphorylation event occurs on the extended C-terminal tail of the variant H2A.X. This region is lacking in the canonical histone H2A. Specifically, the modification occurs on S139 in mammals and S129 in yeast, and is commonly referred to as γH2A.X. This modification is crucial for the recognition and repair signaling of DNA double-strand breaks (DSBs) that can be caused by exogenous agents, stalled replication forks, or meiotic recombination events.[46] Mice completely lacking the *H2A.X* gene contain unstable genomes that lead to infertility, growth

retardation, and increased sensitivity to radiation.[47] In line with these results, non-phosphorylatable mutations of S129 in budding yeast leads to hypersensitivity to chemically induced DSBs.[48, 49]

γH2A.X initiates directly next to DSBs and spreads in both directions over megabase pair regions in higher eukaryotes.[50-53] This spreading is thought to create a regulatory scaffold that integrates and recruits the activities of several cell cycle checkpoint and DNA repair protein complexes. One such protein is MDC1, which binds to γH2A.X directly through its BRCT domain.[42] As DSBs are especially toxic to cells and are associated with ageing and chemotoxicity, detecting them using antibodies against γH2A.X has had great clinical value.[54] Aside from disease states, γH2A.X is also present in meiotic germ cells, when massive amounts of programmed DSBs are catalyzed by the evolutionary conserved Spoll endonuclease to facilitate homolog pairing and DNA recombination. Details of the dynamics of γH2A.X in meiotic cells are described below in Chapter 4.

1.3C LYSINE METHYLATION

Whereas acetylation and phosphorylation primarily affect charge-based properties, methylation confers hydrophobic character to the positively charged lysines of histone tails. Adding to the complexity of methylation, lysines can be mono-, di-, or tri-methylated, and increasing experimental evidence show that each methylation state can be associated with different biological readouts. Indeed, it has been shown that several histone PTM readers can distinguish among methylation states of lysines.[21, 55] These readers not only contain cognate pockets that specifically interact with the PTMs, but can

also interact with regions flanking the modified amino acid (typically ±5 amino acids), underscoring the importance of the position of methylated lysines within polypeptides.[21] Genome-wide analyses have shown that methylation can be associated with either transcriptional activation or repression, depending on the position and extent of the methyl mark. As with most other histone PTMs, histone methylation is kept in balance with the opposing activities of lysine methyltransferases and demethylases. Within this context, the most diversely and dynamically methylated histone isoform is histone H3. The three major and most well characterized sites of methylation on H3 are lysines 4, 9, and 27, with each being involved in fundamentally different chromatin associated processes.

Methylation on H3K4 is a highly conserved chromatin marked correlated with euchromatin and active transcription.[56, 57] This mark is catalyzed by trithorax group proteins, including TRX in *Drosophila*, Set1p in *S. cerevisiae*, and myeloid or mixed lineage leukemia (MLL) in mammals.[58] H3K4 methylation, in particular trimethylation, positively regulates transcription by recruiting HATs and nucleosome remodeling enzymes to transcriptional start sites.[59] In light of these findings, it is especially surprising that yeast lacking Set1p (and therefore H3K4 methylation) are viable.[60] Nevertheless, this strain expresses growth defects and defective transcriptional silencing of the mating-type locus and telomeres. Thus, although H3K4 methylation is not required for viability in yeast, deposition of this mark is obligatory for maintaining chromatin homeostasis and achieving proper gene expression programs.

H3K9 methylation and H3K27 methylation are conserved PTMs involved in the formation and maintenance of heterochromatin. Despite their general association with

heterochromatin, it is becoming increasing clear that they contribute to a growing number of distinct biological mechanisms of transcriptional silencing. Indeed, H3K9me3 is enriched in *constitutive* heterochromatin, which is found at structural or highly repetitive stretches of the genome. Alternatively, H3K27me3 is enriched in *facultative* heterochromatin, which can be dynamically activated or silenced depending on developmental or environmental cues. A detailed description of the properties of H3K9 and H3K27methylation is presented below.

1.4 TYPES OF HETEROCHROMATIN

1.4A HISTONE H3 LYSINE 9 METHYLATION AND CONSTITUTIVE HETEROCHROMATIN

While heterochromatin was originally described as intensely stained parts of chromosomes, studies of H3K9 methylation and H3K27 methylation have helped uncover different varieties or sub-domains of condensed chromatin.[22, 27] *Constitutive* heterochromatin is found at structural or highly repetitive stretches of the genome like pericentric or subtelomeric regions, is enriched in Su(var) (suppressors of position effect variegation) proteins and H3K9me3.[61-64] The biological activities of H3K9 methylation are largely dependent on the methyl status of H3K9, which is regulated by a large number of putative mono-, di-, and tri-methyl H3K9 writers in the human genome.[65, 66]

Monomethylation of H3K9 has been biologically linked to the stability and deposition of H3K9me3.[67, 68] In this way, multi-step deposition of methyl groups on H3K9 by distinct writers could function to tightly regulate constitutive heterochromatin

formation. However, the roles of H3K9me1 do not entirely overlap with H3K9me3-associated biology. In fact, the subnuclear localization and regulation of H3K9me1 can be very distinct from those of H3K9me3 during development and cell growth.[65] Therefore, a comprehensive description of the biological activity of H3K9me1 is currently lacking, but in general terms it is a histone PTM associated with constitutive heterochromatin.

A confounding factor in assessing the biological significance of H3K9me1, is that its major writer, G9a, is also the major writer for H3K9me2, which increasing experimental data suggest is principally involved with the maintenance of heterochromatin at the nuclear periphery.[68-70] Compartmentalization of chromatin to the nuclear periphery is a major gene-silencing pathway during stem cell differentiation.[71] The nuclear periphery can also house gene-poor DNA such as centromeres and telomeres.[72, 73] The role of H3K9me2 appears to be uniquely different from those of H3K9me1 and H3K9me3, as its deposition and regulation primarily affects gene expression profiles rather than silencing of gene-poor DNA.[70, 73] The mechanisms of large-scale localization of chromatin to the nuclear periphery are multifaceted and unclear, but H3K9me2 appears to play an important role in anchoring heterochromatin to the nuclear periphery.[65, 67]

Trimethylation of H3K9 is highly correlated with transcriptional silencing of pericentric, repetitive, and telomeric DNA. Three elegant studies in the early 2000s helped solidify the importance of H3K9me3 in the formation of constitutively silent heterochromatin. The first was a genetic and biochemical study in fission yeast that linked the activities of Clr4 (a homolog of Suv3-9h), a known K9 methyltransferase, and

a heterochromatin protein 1 (HP1) isoform, known to colocalize with heterochromatin in *Drosophila*, together with H3K9me3.[61] The second was a structural and biochemical study of HP1 (via the chromodomain) binding to methylated H3K9 that showed the affinity of this interaction is in the low micromolar and physiologically relevant range.[64] Finally, a mouse study showed Su(var)3-9h is a H3K9 methyltransferase that targets H3K9me3 to pericentric heterochromatin and that its genetic disruption leads to genomic instability in somatic and germ cells.[63]

Curiously, HP1 isoforms can bind to H3K9me1/2/3 with similar affinities *in vitro*.[55] How these HP1 binding properties relate to the diversity of H3K9 methylation-related functions in the cells remain unknown, but the promiscuity of HP1 binding suggests the presence of higher levels of regulation *in vivo*. Notably, several recent studies have uncovered novel H3K9methyl readers, some of which are able to distinguish among different methylation states of H3K9.[55, 74] These studies also showed that HP1 isoforms can bind to mono-, di-, and tri-methylated H3K23 *in vitro*.[55] Although the levels of this mark appear to be very low (<1%) in the cells lines used in the studies, these finding point to a potential cross-talk or cross-regulation of H3K9 and H3K23 methylation. Intriguingly, H3K9 and H3K23 are flanked by a conserved $KQXAXK_{9/23}$ that may serve as the basis for some of the putative functional overlap of these residues.

1.4B HISTONE H3 LYSINE 27 METHYLATION AND FACULTATIVE HETEROCHROMATIN

Like constitutive heterochromatin, regions of *facultative* heterochromatin are also condensed; however, they are enriched in H3K27me3, and can become activated or silenced in response to different cellular environments. Much of what is known about

H3K27me3 is rooted in the early studies of Polycomb group proteins (PGP) in *Drosophila*. Long before it was shown that the PGP methyltransferase called Enhancer of Zeste (E[z]) writes H3K27me3[75] , several studies had linked this protein, together with several other genes, to the silencing of homeotic (Hox) genes in flies.[76] In general, it was known that PGP activity directly antagonizes trithorax group proteins. Subsequent studies helped define two major protein complexes that carried out Polycomb group silencing: Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2). PRC2 contains the H3K27me3 writer E(z) and at least three other proteins, including ESC, SU(Z)12, and NURF555 that that contribute in various ways to the activity of E(z).[76] PRC1 is generally composed of the chromodomain-containing protein Polycomb (PC), as well as PH, PSC, and RING.[77] PRC1 is targeted to chromatin via the specific interaction of the chromodomain of PC to H3K27me3.[76] These complexes are highly conserved in other eukaryotic organisms including mice, humans, plants, and the single-celled eukaryotes *Tetrahymena thermophila*. Moreover, PRCs occur in multiple versions (transient interactions or variants of the core proteins) that seem to have different activities *in vivo*.[78]

In mammals, the PGP silencing pathway is important for stem cell differentiation, X-chromosome inactivation in females, and cancer.[76] Loss of E(z) family proteins, as well as other PRC members, leads to embryonic lethality in mice and *Drosophila,* and sterility in *C. elegans*.[79-82] Genome-wide ChIP-seq studies have shown that H3K27me3, like H3K4me3, is primarily enriched in the nucleosome directly preceding the transcriptional start site of genes. Data from these studies confirm that H3K27me3 is highly correlated with silenced genes, whereas H3K4me3 is associated with actively

transcribed genes. Interestingly, both H3K4me3 and H3K27me3 can be detected in overlapping promoter nucleosomes of developmentally regulated genes in embryonic stems.[83] These so-called 'bivalent domains' are thought to regulate gene expression programs during stem cell differentiation. Despite these exciting advancements, how PGP proteins suppress transcription and how they are targeted to specific genes during development is still not clear. With respect to the latter, there is increasing evidence that small RNAs, in addition to H3K27me3, play a critical role in the genomic localization of PRC2.[84, 85]

Almost all of the studies concerning PGP silencing have focused on H3K27me3. However, other histone marks are also involved and are perhaps regulated, or stimulated, by the presence of H3K27me3. The most well studied example is mono-ubiquitination of the C-terminal tail of H2A, which is carried out by the RING domain containing E3 ligase protein present in PRC1.[76, 86, 87] Because PRC1 is thought to be recruited to chromatin by its interaction with H3K27me3, it is formally possible that its ubiquitinating activities are also downstream of PRC2-mediated H3K27me3 deposition.[76] However, several studies have shown that the deposition of these two modifications may be less co-dependent than the above model would suggest.[88, 89] Still, genome-wide studies have shown that there is significant overlap between H3K27me3 and H2A monoubiquitination, suggesting that they act coordinately to carry out PGP-mediated silencing. Consistent with their role in gene silencing, PRC2 activity and H3K27me3 deposition have also been linked to histone deacetylation. This effect is accomplished by the transient interaction of HDACs with EED (a conserved member of PRC2) at chromatin. Strengthening the potential existence of crosstalk between H3K27me3 and

histone deacetylation, it was also shown that EED contains a subdomain that functions as a reader of H3K27me3.[90]

1.5 THE BIOLOGICAL SIGNIFICANCE OF HISTONE PTMS

There is little doubt that certain histone PTMs mark distinct chromatin states, but their biological significance, and the existence of a strict 'histone code,' is currently under intense debate. The major point of contention converges on the issue of cause and effect. Namely, do histone PTMs underlie distinct chromatin states, or are they simply a consequence of the diverse processes associated with each type of chromatin? A myriad of studies have provided substantial data that collectively seem to support both viewpoints.

The histone code hypothesis states that specific combinations of histone PTMs encode distinct biological outcomes by recruiting reader proteins (or complexes) with chromatin modifying capabilities. This hypothesis stemmed in large part from the appreciation that histones are often multiply modified by specific PTM patterns within the same polypeptide and/or the same nucleosome (*e.g.* bivalent domains). This model is also reinforced by several studies showing that multiple histone PTMs are biochemically and genetically linked upstream of biological outputs. Moreover, chromatin associated proteins often contain multiple subdomains that can read specific classes of histone PTMs, underscoring their potential to engage chromatin through multivalent interactions with modified histones. Several studies have been published that provide *in vivo* and *in vitro* support for the multivalent character of chromatin associated proteins.[91-94]

Alternatively, many argue that a "histone code" hypothesis is overly complex and the biological influence of histone PTMs is overstated in the literature. They propose instead that the DNA accessibility/nucleosome positioning paradigm can sufficiently explain the growing complexities of chromatin and that histone PTMs could serve as allosteric activators (instead of recruiters) of chromatin modifying enzymes. Some of their main arguments, articulated in two perspective articles [95, 96], are the following: 1) Although there are hundreds of histone PTMs, genome-wide ChIP-seq studies have shown that they significantly overlap in the genome, suggesting that the seemingly infinite combinations of histone PTM manifest only as a few modifications states *in vivo*; 2) The presence of multiple binding modules within chromatin modifying proteins and complexes is irrelevant since most contain weak binding constants (*i.e.* ~100 uM) to their respective histone PTM and combining multiple histone PTMs in binding studies only yields modest increases in binding efficiency; 3) In the few studies where the data sufficiently supports multivalancy, the effects on transcription, or another chromatin function, are not entirely clear; 4) Loss of histone PTMs, either through mutation of the histone tail or deletion of a writer, is well-tolerated; and 5) In terms of histone acetylation, with the exception of H4K16, the position of the acetylation site (be it on histone H3 or histone H4) is negligible, instead the additive effect of the number of acetylations contributes to function.

Much of the arguments above are rooted in genetic and biochemical studies of yeast and genome-wide studies using ChIP-Chip or ChIP-seq. Although both study approaches have made a profound impact on the study of chromatin biology of higher eukaryotes, the data that they yield and their generalizability are not without any caveats.

Both *S. cerevisiae* and *S. pombe* chromatin are gene rich and relatively euchromatic[97], which does not recapitulate the complexity and the predominantly heterochromatic nature of higher eukaryotic chromatin. Even within the yeast clade, several profoundly distinct mechanisms of chromatin regulation have evolved. Take heterochromatin formation, for example, which is accomplished independently of known histone methylation sites in *S. cerevisiae* and almost exclusively with H3K9 methylation in *S. pombe*.[98] Furthermore, both yeast strains lack H3K27 methylation, which severely limits their utility as model organisms to study the facultative heterochromatin present in higher eukaryotes. Within this context, it is formally possible that the DNA accessibility/nucleosome positioning paradigm is more dominant in yeast than in higher eukaryotes, and the consequences of disrupting histone PTMs in each organism are different.

The pitfalls of interpreting antibody- or chip- based histone PTM data sets are well documented.[99] Namely, in a crowded landscape of histone PTMs, antibodies designed to recognize specific, singly modified epitopes on histone tails are often occluded by the other nearby PTMs. This causes a significant loss of information regarding the genomic localization of the histone PTM of interest. Moreover, most of the histone PTM antibodies developed and validated for ChIP are specific for methylation or acetylation, resulting in a dearth of information on the dozens of other histone PTM types known to occur in mammalian chromatin. Development of antibodies specific for multiply modified epitopes and other PTM classes will no doubt help resolve the molecular signatures associated with distinct chromatin states. Nevertheless, bioinformatic analyses utilizing the current data sets have already hinted at the existence

numerous chromatin states, which were demarcated by distinct combinations of histone PTMs and transcriptional states.

Central to this debate is the mechanism of epigenetic inheritance. Several studies have shown, from yeast to humans, that transcriptional programs can be inherited during the cell division. How is this information transmitted? To date, histone PTMs remain as the obvious mechanism underlying epigenetic memory and transmittance through mitosis or meiosis, although RNA seems to contribute as well in higher eukaryotes. Therefore, by inference, all epigenetic histone PTMs should have an inherent capability to signal for local chromatin remodeling in subsequent cell generations. Whether histone PTMs serve to recruit or allosterically activate chromatin-modifying pathways may be a rhetorical point given our limited, but growing, appreciation of the complexities of chromatin.[100]

1.6 GOING BEYOND A CYTOLOGICAL VIEW OF CHROMATIN

The **ENC**ylcopdieia **O**f **D**NA **E**lements (ENCODE) consortium recently published a series of papers that extensively described novel regulatory and functional elements of the human genome.[101] Among its many discoveries, they concluded that roughly 75% of the human genome is amenable to transcription, and about half of the resulting RNAs have a 'functional' role in the cell.[102] Although evidence of pervasive transcription has been previously reported in pilot, small scale studies of mammalian cells[103], these ENCODE studies were genome-wide and were cross compared with 15 different cell lines. Considering that the human genome is predominantly heterochromatic, these results have had a profound effect on our understanding of the definition and functions heterochromatin. Among many other implications, this emerging

paradigm challenges the notion that DNA accessibility is the major mode of transcriptional regulation in higher eukaryotes.

Evidence of pervasive transcription is also in line with the growing appreciation that RNA products are no longer exclusively dedicated to gene expression, protein synthesis machinery, or RNA processing machinery. In fact, one of the major products of transcription are non-coding RNAs (ncRNAs) that have regulatory capabilities, such as microRNAs and lncRNAs. Progress within this field has not only contributed to the understanding of the diverse roles of ncRNA, but also how they are regulated at the level of chromatin. Algorithms have shown that ncRNAs and other transcriptional outputs map very closely with unique combinations of chromatin factors, including histone PTMs and variants, readers, writers and erasers.[104] A major implication extending from these bioinformatic studies is that physiological chromatin exists in a myriad of states that presumably help to organize and regulate complex transcriptional programs. Several studies have already described dozens of chromatin states in a wide range of organisms using datasets limited to only small proportion of all the known histone PTMs and their associated enzymes.[104-107] Therefore, a more comprehensive view of histone PTM states is needed to help resolve the increasing intricacies of chromatin.

1.7 TETRAHYMENA AS A MODEL ORGANISM

The ciliated protozoan *Tetrahymena thermophila* is a very useful model organism for chromatin biology. Several seminal discoveries in biology have resulted from studies of *Tetrahymena,* including the initial characterization of telomeres and telomerase, the discovery of catalytic RNA, and the function of histone acetylation.[35,108-110] These

discoveries were facilitated in large part by the unique physiological characteristics of *Tetrahymena* that afforded unprecedented access to the underlying biology.

Much like metazoans, *Tetrahymena* have physically separate somatic and germline genomes.[111-113] However, in ciliates these genomes are compartmentalized in two physically separable nuclei found within a common cytoplasm (**Figure 3A)**. The micronucleus contains the germline genome, which is activated during sexual reproduction (conjugation) and functions to produce new somatic and germline genomes in the progeny **(Figure 3B)**. The macronucleus (MAC) contains the somatic genome, which is the site of gene expression in *Tetrahymena* during asexual (vegetative) growth and conjugation **(Figure 3B)**. The macronuclear genome does not contribute DNA to subsequent generations, although there is some evidence that it can have an epigenetic role in the formation of the progeny genome.[114]

The macronuclear genome undergoes several programmed DNA rearrangements and processing during conjugation that ultimately becomes significantly differentiated from the micronuclear genome. The genomes differ in the following ways. Firstly, the macronuclear genome lacks about 33% of the DNA sequence information of the micronuclear genome. The disparity occurs as a result of active DNA elimination of micronuclear-limited sequences (also known as internal eliminates sequences [IESs]).[115] Since these sequences are often intergenic, repetitive, and transposon–like sequences, the macronucleus is thought to be 'streamlined' for transcription, as it is highly enriched for genic and regulatory DNA sequences.[112] Another class of sequences targeted for elimination is centromeric DNA. Indeed, macronuclear chromosomes lack centromeres causing it to divide amitotically (a process akin to diffusion) during vegetative growth,

Figure 3

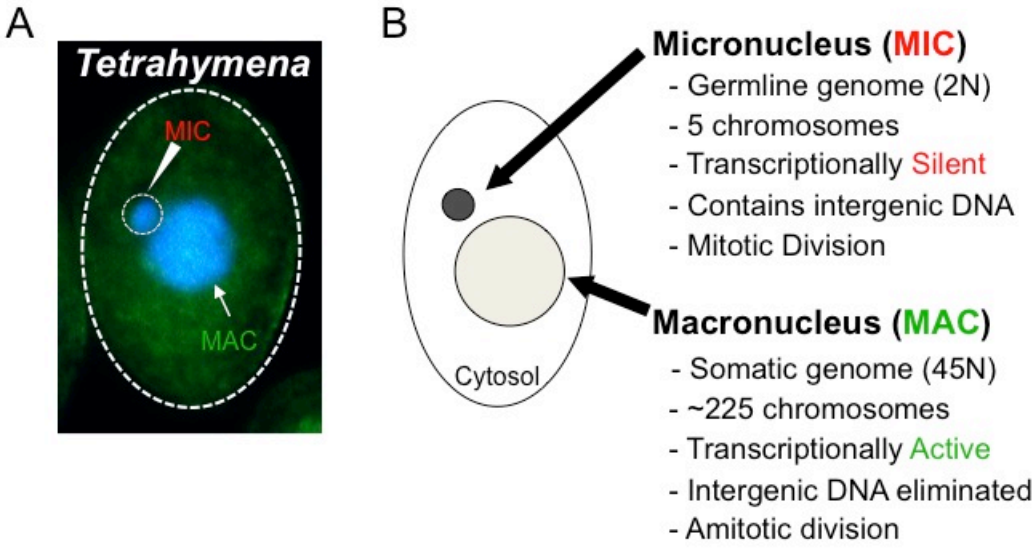


Figure 3. **Features of *Tetrahymena thermophila***. **(A)** A fluorescence image of *Tetrahymena* highlighting its nuclear dimorphism. The blue signal is DAPI stained DNA. **(B)** Genetic, functional and cytological disparities between micronuclei and macronuclei.

whereas micronuclei retain their centromeres and divide mitotically.[116, 117] Secondly, the macronuclear genome undergoes massive fragmentation at specific sites (~15 bp) known as chromosome breakage sequences (Cbs).[111] This fragmentation results in the formation of hundreds of chromosomes. By comparison, the micronucleus only contains 5 chromosomes. Finally, although the macronuclear genome lacks the full complement of the *Tetrahymena* genome, it is highly amplified, containing about 45 copies of each chromosome. Therefore, the macronuclear genome contains about 20X the chromatin content of the diploid micronuclear genome.

In addition to their genomic content, macronuclei and micronuclei greatly differ in their chromatin states. During vegetative growth, several studies have shown that micronuclei are transcriptionally silent.[118] In fact, micronuclei are dispensable for the viability of *Tetrahymena* during vegetative growth, strongly suggesting that micronuclei are functionally inert during this stage of the life cycle.[119] Consistent with these observations, several cytological studies using transmission electronic microscopy and heavy metal staining have shown that the micronuclear chromatin is relatively dense compared to macronuclei and macronuclear chromatin appears to be loosely associated, much like euchromatin.[120] Notably, macronuclear chromatin is not uniformly euchromatic since it also contains a punctate pattern of cytologically dense regions. Adding to these differences, macronuclei and micronuclei also differ in the types of histone variants present within each compartment. Namely, the linker protein histone H1 is present in macronuclei, but absent in micronuclei.[121] Instead, micronuclei contain four high-mobility-group-like proteins called alpha, beta, gamma, and delta.[122] Important to the findings described in the dissertation, macronuclei and micronuclei also contain

distinct histone H3 proteins. Macronuclei contain two H3 isoforms termed major and minor that correspond to human histone H3.2 and H3.3, respectively. Interestingly, minor H3 is absent in micronuclei, and the major H3 exists in two forms: a full-length form that is indistinguishable from macronuclear major H3; and a proteolytically processed ('clipped') form that lacks the first six amino acids of the N-terminal tail of major H3.[123] Because the clipped form is smaller and runs faster in electrophoretic assays, it is sometime referred to has H3 fast ($H3^F$) and the full-length formed is called H3 slow ($H3^S$). While the function of histone clipping in the transcriptionally silent micronucleus is not known, it is important to note that the clipped peptide includes H3K4, which when methylated could serve as a strong activator of transcription.

Are the structural and functional differences of macronuclei and micronuclei recapitulated in their histone PTMs states? Previous studies have shown that macronuclear histones contain the PTM hallmarks of active transcription; namely, they are highly acetylated and histone H3 contains H3K4me3.[57, 124] Antibody based studies have also suggested that macronuclei contain the heterochromatic H3K27me3 mark, although it is not clear if it colocalizes with the cytologically dense regions described above. Notably, it is known that micronuclei lack H3K4me3 and histone acetylation. Moreover, H3K27me3 is highly enriched in the micronucleus, in accordance with its transcriptionally 'off' status.[57] Thus, the functional dichotomy of the biochemically separable nuclei of *Tetrahymena* makes it an excellent model organism to study chromatin function.

CHAPTER 2: IDENTIFICATION OF HISTONE H3 LYSINE 23 METHYLATION IN GERMLINE HETEROCHROMATIN

2.1 INTRODUCTION

Recent advances in mass spectrometry and genome-scale analyses have suggested the existence of several chromatin states with unique functional, transcriptional, and molecular properties. Moreover, results from the ENCODE consortium have shown that the majority of the human genome, including constitutively heterochromatic regions, are transcribed to some degree. Within this context, the cytological definition of heterochromatin is rather simplistic, but a comprehensive molecular definition still remains elusive. Of particular importance is the array of histone PTMs that occur within these transcriptionally repressive environments. However, biochemical dissection of distinct chromatin types is problematic since they are often found on contiguous stretches of DNA in most eukaryotic cells. Coupled with the fact that all chromatin states are compartmentalized within a single nucleus, obtaining highly enriched samples of distinct chromatin types at scales necessary for the detection of their underlying histone PTMs is insurmountable at the present time.

Take, for example, the inactive X (Xi) chromosome of female mammals. It is well established through antibody-based approaches that H3K27me3 is enriched and histones are hypoacetylated on Xi, but whether other histone PTMs are contributing to silencing is not known.[125,126] Therefore, there is great interest in obtaining a

comprehensive list of the histone PTM states found with Xi, but no study to date has successfully developed a method that enriches for Xi chromatin to the extent required for PTM analysis. This is in large part due to the fact that the Xi, although not covalently linked to other chromosomes, is physically found intermingled with the other chromosomes within the same nucleus.

To obtain highly enriched heterochromatin and screen for novel histone PTM states, we took advantage of the nuclear dimorphism of *Tetrahymena*. In addition to the functional and physiological differences between macronuclei and micronuclei, previous studies have shown that they also contain disparate histone PTM states. In line with its transcriptional status, the macronucleus contains hyperacetylated core histones and H3K4 methylation.[57] Furthermore, antibody based and mass spectrometry approaches have shown that the macronucleus also contains H3K27 mono-, di-, and trimethylation.[57] In the cases of H3K27me1 and H3K27me2, these modifications can occur with H3K4me1 in *cis* on the same N-terminal peptide, suggesting the presence of bivalent-like transcriptional regulation in the macronucleus. The role and genomic localization of H3K27me3 in the macronucleus has not been studied to date. However, given that approximately half of the ~21,000 genes in *Tetrahymena* are not transcribed during vegetative growth, and the fact that no other heterochromatin associated histone methylation (such as H3K9me) is known to occur in the macronucleus, H3K27me3 likely mediates gene silencing programs in this compartment. In support of this idea, a homolog of HP1, called Hhp1p, in *Tetrahymena* has been shown to specifically localize to the macronucleus, bind to H3K27me3, and mediate starvation induced gene silencing.[127]

Similar mass spectrometry driven attempts to identify histone PTMs in

micronuclei have been lacking, in large part due to the difficulty of purifying micronuclei. Nevertheless, antibody based studies have helped identify some of the PTMs associated within this compartment. Notably, it was discovered that histone H3 serine 10 is phosphorylated (H3S10ph) in mitotic micronuclei.[128] Since its discovery in *Tetrahymena*, H3S10 hyperphosphorylation has been shown to be a universal marker of mitosis.[45] In contrast to macronuclear core histones, micronuclear core histones are hypoacetylated.[124,129] Finally, micronuclei are lacking in H3K4 methylation and are highly enriched for H3K27 trimethylation.[57] Aside from these modifications, no other modifications are known to occur on micronuclear H3.

Perhaps the most surprising aspect of histone PTMs in micronuclei and macronuclei is the fact that neither contains H3K9 methylation during vegetative growth. In the context of the macronucleus, the lack of H3K9 methylation is not so problematic because the DNA elements that are typically associated with this mark (namely repetitive, structural and pericentric DNA) are actively eliminated from the genome during conjugation. Interestingly, these eliminated elements are targeted for H3K9 methylation in new macronuclei during conjugation.[115] The lack of H3K9 methylation in the micronucleus is more difficult to explain. One implication stemming from this fact is that H3K27me3 in the micronucleus may target multiple different classes of genomic elements (intergenic vs. genic). However, the parsimonious explanation follows that there is at least another histone modification that underlies these eliminated regions within the micronucleus.

Advances in mass spectrometry have long supported the identification and study of histone PTMs. As the chromatin field moves from identifying single histone PTMs to

characterizing combinatorial interplay or "crosstalk" among histone PTM states, specialized techniques like ECD (electron capture dissociation) and ETD (electron transfer dissociation) have helped identify novel combinations of modifications in *cis* on relatively long histone peptides (>20 amino acids).[130-133] In this study, we employed an in-line liquid chromatography-mass spectrometry (LC-MS) strategy wherein a polyCAT column (cation exchange) and an organic mobile phase (acetonitrile) are directly coupled to an ETD enabled LTQ-FT mass spectrometer and an ion trap mass spectrometer (collectively called HILIC-MS/MS). This high-resolution method allows for the separation of large histone peptides that differ by charge and as few as one methyl group. Despite this capability, HILIC-MS/MS still requires front-end purification of complex samples to yield interpretable spectra. Therefore, samples from *Tetrahymena* are ideal for HILIC-MS/MS because they are relatively under modified compared to human histones, and upon successful purification of micronuclei would provide highly enriched heterochromatin histones.

2.2 METHODS

2.2A CELL CULTURE AND STRAINS

Wildtype *Tetrahymena* strains CU427 and CU428 were provided by the *Tetrahymena* Stock Center (http://tetrahymena.vet.cornell.edu/). Standard culture media and conditions were used.[113] All experiments using vegetative *Tetrahymena* involved cells growing at mid-log phase (~2.5 X 10^5 cells/mL). Conjugation was induced as previously described.[113, 134]

2.2B PREPARATION OF MICRONUCLEAR HISTONE H3

Micronuclei were isolated following suggested conditions from published protocols.[113, 135] Minor modifications to these protocols yielded purified micronuclei in a highly reproducible manner. Unless otherwise stated, centrifugation was done for 5 min at each step. Typically, the first fraction (pure macronuclei) was obtained by centrifugation at 610*g*. The next fraction was collected at 780*g* and it often contained pure macronuclei. Fractions 3-8, containing a mixture of MAC/MIC, were collected by centrifugation at the following increments: 970*g*, 1190*g*, 1430*g*, 1680*g*, 1970*g*, and 2100*g,* respectively. By fraction 8, the MAC to MIC ratio was typically 1:10. At this point, the *g* forces of each centrifugation step were increased by increments of 100*g* until macronuclei were completely clarified. This typically occurred at 2600*g*. If macronuclei were still present after centrifugation at 2600*g,* subsequent fractions were collected at 2600*g* using 5 min increments (i.e. 5 min, 10 min, and so on) until macronuclei were almost completely clarified. Once the macronuclei were removed, 3-4 fractions were collected (using 5 min increments per fraction) at 3500*g* to obtain highly purified micronuclei.

In optimizing these conditions, I learned of a few pitfalls that lessen the chances of obtaining pure micronuclei. As I have not had a chance to systematically test each of these recommendations, they are anecdotal in some respect. Nevertheless, the first and perhaps most important recommendation is that it is critical to follow the increments described above. Ramping up the centrifugal forces faster than recommended above in an effort to expedite the procedure often leads to incomplete removal of macronuclei. The next recommendation pertains to the total volume of medium B used during the

preparation. I have found that pelleting the nuclei in volumes less than 200 mL of medium B (in 250 mL Corning conical tubes) reduces micronuclei purity. Therefore, since the concentration of cells in medium B is 1.5 X 10^6 cells/mL, each experiment should minimally include 3 X 10^8 cells (or ~1.5 L of cells at mid-log phase). Finally, it is important to blend for one minute on high in between each step of centrifugation. This ensures that the nuclei are not clumped and are homogenously distributed in medium B prior to pelleting.

Once isolated, nuclei from each fraction were stained with DAPI and the MAC/MIC ratio of each fraction was determined by fluorescence microscopy. Alternatively, nuclei were stained with methyl green and observed by light microscopy. In all experiments, micronuclear fractions containing less than 0.1% macronuclear contamination were pooled. Then, total histones were extracted in 0.4 N H_2SO_4 and the histone isoforms were separated by reversed-phase HPLC using a C8 column (220 X 4.6 mm, Aquapore RP-300, Perkin Elmer) following previously described methods.[136] HPLC fractions were analyzed by SDS-PAGE and Coomassie blue staining. HPLC fractions containing micronuclear histone H3 were pooled, dried under vacuum, and their PTM states were analyzed by HILIC-MS/MS.

2.2C HILIC-MS/MS Analysis of Micronuclear H3 PTM States

N-terminal tails of purified micronuclear histone H3 species, $H3^S$(1-50) and $H3^F$(7-50), were obtained by GluC digestion and analyzed by hydrophilic liquid chromatography coupled with high resolution MS^1 and linear ion trap ETD MS/MS (HILIC-MS/MS). Dried RP-HPLC fractions containing micronuclear H3 (fractions 56-

58) were reconstituted in 100 mM ammonium acetate (pH 4.1) and digested with endoproteinase Glu-C (Roche Applied Science) at a ratio of 1:20 enzyme to protein for 4 hours at 37 C. Digestion was quenched by freezing the samples at -30 C.

An aliquot of the endoproteinase Glu-C digested H3 histone peptide solution (5 pmole) was dried to completion in a Savant SpeedVac® and reconstituted in solvent A (as described below) prior to pressure loading onto fused silica capillary analytical column (360 μm o.d. x 75 μm i.d.) containing 10 cm of PolyCAT A™ (300 Å pore size, 3 μm diameter, Poly LC Inc., Columbia, MD) equipped with an integrated electrosprayemitter tip. Peptides were gradient eluted off of the column using a linear gradient as previously described.[132] The eluent was directly introduced into a front-end electron transfer dissociation enabled LTQ-FT (Thermo Scientific) mass spectrometer. MS^1 spectra were acquired in the high-resolution FT mass analyzer and MS^2 spectra were acquired in the linear ion trap mass analyzer. Following each high resolution MS^1 spectrum, eight data-dependent low-resolution ETD MS^2 spectra were acquired. All MS^2 spectra were collected using the following instrument parameters: one microscan, 3 m/z isolation window, default charge state of +10, monoisotopic precursor selection "enabled". Full automated gain control targets were set to $1e^6$ for FTMS and $6e^5$ for ITMS. In addition dynamic exclusion was enabled and set for a repeat count of 2, repeat duration of 30 s, exclusion list of 150 masses, and exclusion duration of 20 s. Azulene was the electron transfer reagent.

All MS/MS spectra were searched against a *Tetrahymena* histone H3 protein database using the Open Mass Spectrometry Search Algorithm (OMSSA).[137] Prior to searching the data, MS^2 peak lists were generated using a Java-based program written in-

house (Dr. Donald Hunt, University of Virginia, Charlottesville, VA). All searches were completed using either the "no enzyme" or "Glu-C" digestion parameter. Precursor and peptide mass tolerances were set to +/- 0.05 and +/- 0.35, respectively. Search parameters included the following variable modifications: mono-, di-, and tri-methylation of Lys; acetylation of Lys; and phosphorylation of Ser, Thr, Tyr. OMSSA removed the reduced charge species from the ETD peak lists prior to searching the data. Results from the searches were used as a guide for data analysis. All data were interpreted manually in order to determine sequence coverage and modification sites.

2.2D PEPTIDE SYNTHESIS AND HISTONE H3 PEPTIDES USED IN THIS STUDY

Peptides were synthesized using a Prelude Synthesizer (Protein Technologies) using standard Fmoc-based solid phase peptide chemistry. Peptides were cleaved from the resin by Reagent K, precipitated in diethyl ether, and purified using Varian Dynamax Microsorb C18 preparative column (Agilent, Santa Clara, CA). The masses of HPLC-purified peptides were confirmed by MALDI-TOF mass spectrometry. Below is a list of peptides that we used in the dissertation. Those marked with and asterisk were synthesize in-house. The numbers in the parenthesis of each peptide name correspond to the residues of the H3 N-terminal tail. Peptides with an asterisk were synthesized in-house.

Peptides:

Unmod (18-36)- KQLASK_{23}AARK_{27}SAPATGGIK(biotin)*

K23me3 (18-36)- KQLASK_{23}(me3)AARK_{27}SAPATGGIK(biotin)*

K23me3 (19-28)- QLASK_{23}(me3)AARK_{27}C*

K27me3 (18-36)- KQLASK$_{23}$AARK$_{27}$(me3)SAPATGGIK(biotin)*

K23/27me3 (18-36)- KQLASK$_{23}$(me3)AARK$_{27}$(me3)SAPATGGIK(biotin)*

Unmod (1-20)- ARTK$_4$QTARK$_9$STGGKAPRKQLY (Gift from C.D. Allis)

K4me3 (1-15)- ARTK$_4$(me3)QTARK$_9$STGGKAY*

K9me3 (1-15)- ARTK$_4$QTARK$_9$(me3)STGGKAY*

K23me1 (20-28)- CLATK$_{23}$(me1)AARK$_{27}$S (Active Motif)

K23me2 (20-28)- CLATK$_{23}$(me2)AARK$_{27}$S (Active Motif)

2.2E GENERATION OF THE H3K23ME3 SPECIFIC POLYCLONAL ANTIBODY

α-H3K23me3 antibodies were generated by injecting rabbits (Cocalico Biologicals Inc.) with a synthetic peptide spanning amino acids 18-36 of histone H3 and containing a trimethylated lysine 23 modification (H3K23me3[18-36]).

H3K23me3-specific antibodies were affinity purified from rabbit serum as previously described.[138] Briefly, IgG from the serum were precipitated in 50% ammonium sulfate (saturated) and resuspended in 1 X PBS + 1 mM PMSF. The IgG were then passed through a gravity column containing SulfoLink beads (Thermo Scientific) coupled to the non-native cysteine of a short synthetic peptide, K23me3 (19-28), from the H3 N-terminus containing the H3K23me3 modification. Bound antibodies were eluted with 100 mM glycine pH 2.5 and the eluate was immediately neutralized to pH ~7 with 1 M Tris pH 8.0, flash frozen, and stored in -80 °C.

Specificity of the purified IgG was evaluated by ELISA. Greiner high-binding 96-well plates (Sigma) were incubated with peptides diluted in 0.05 M carbonate buffer pH 9.6 overnight, blocked with 1% BSA resuspended in PBS-T (1 X PBS pH 7.4 + 0.05%

Tween 20) for 1 h, incubated with affinity purified α-H3K23me3 diluted in PBS-T (1:100) for 1.5 h, and incubated with HRP-conjugated α-rabbit secondary antibody (GE Life Sciences) diluted in PBS-T (1:5,000) for 1.5 h. All incubation steps were carried out at 37 °C followed by 3X washes with PBS-T. To detect the ELISA signal, plates were treated with stable peroxide substrate buffer (Thermo Scientific) supplemented with OPD (Thermo Scientific) following manufacture's procedures and the absorbance at 492 nm was measured using a microplate reader. Peptides were bound in triplicate in each plate and all ELISA experiments were repeated at least three times. Representative results from a single experiment are presented and the error bars denote standard deviation of triplicate samples. Peptide competition experiments for ELISAs were carried out by incubating the indicated peptides at a concentration of 5 μg / mL (of PBS) with the antibodies for 1 h on ice prior to their use.

2.2F WESTERN BLOT AND INDIRECT IMMUNOFLUORESCENCE ANALYSES

SDS-PAGE was performed as described [57,139] or materials and procedures corresponding to Novex NuPAGE precast gel system were used (Life Technologies). The procedures for methanol fixing and indirect immunofluorescence staining of *Tetrahymena* cells are described previously.[140]

Rabbit polyclonal antibodies used for Western blotting included: α-H3 (*Tetrahymena* H3 #117; 1:5,000; gift from C.D. Allis), α-H3K23me3 (1:100; made in this study), α-H3S10p (1:25,000; gift from C.D. Allis), α-H3K9me3 (1:1,000; ab8898; Abcam), α-H3K27me3 (1:2,500; 39155; Active Motif), α-H3K23me1 (1:1,000; #39387; Active Motif), and α-H3K23me2 (1:1,000; #39653; Active Motif). HRP-conjugated α-

rabbit secondary antibodies (GE Life Sciences) were used at dilution of 1:5,000 and detected by ECL plus reagent (Amersham) following manufacturer's procedures.

For *Tetrahymena* immunofluorescence staining, cells were fixed in partial Schaudin's fixative as described.[140, 141] The primary antibodies used were: rabbit α-H3K27me3 (1:250; gift from T. Jenuwein), mouse α-H3K4me3 (1:250; ab1012; Abcam), rabbit α-H3K23me3 (1:100; made in this study), rabbit α-H2A (1:250; control stain for *Tetrahymena*), mouse α-H3K9me3 (1:100; 39285; Active Motif), rabbit α-H3K23me1 (1:250; Active Motif), and rabbit α-H3K23me2 (1:250; Active Motif). Secondary antibodies, used at a dilution of 1:2000, included Alex Fluor 647-conjugated goat α-mouse, Alexa Fluor 488-conjugated goat α-rabbit, and Alex Fluor 488-conjugated goat α-mouse (Molecular Probes). Slides were mounted in ProLong Gold antifade reagent with DAPI (P36931; Life Technologies) and stored overnight at room temperature prior to imaging. Digital images were captured using a Zeiss Axio Observer fluorescence microscope equipped with an Apotome for optical sectioning and AxioCam MRm camera. Raw images were exported as tiff files and processed using Adobe Photoshop software CS3.

2.3 RESULTS

2.3A PURIFICATION OF HISTONE H3 FROM THE MICRONUCLEUS

To highlight the different histone PTM signatures found within macronuclei and micronuclei, we performed co-immunofluorescence staining of *Tetrahymena* growing under vegetative conditions. As expected, the results show that the heterochromatic

micronucleus is relatively enriched for H3K27me3 while it completely lacks H3K4me3 **(Figure 4)**. The macronucleus is positive for both the H3K27me3 and H3K4me3 marks, although by our co-immunofluorescence staining they appear localized to distinct sub-nuclear domains **(Figure 4, merge)**. Although previous studies have also shown that the macronuclear genome is streamlined for gene expression, these results suggest that transcription is highly regulated by silencing/activating pathways (and perhaps distinct chromatin states) that have not yet been fully interrogated. Comparatively, the micronucleus appears to be less complex than the macronucleus in terms of histone PTM states and in overall function during vegetative growth.[57]

Two major properties of *Tetrahymena* physiology facilitate the biochemical purification of micronuclei. First, the micronucleus is a membrane-bound organelle that is separate from the macronucleus.[120] This feature allows micronuclei to remain intact *in vitro* after outer membrane lysis. Second, although the micronucleus is cytologicalldenser than the macronucleus, its genomic content (2N), and by extension its weight, is at least 20 times less than the polyploid macronucleus (45N).[112] Together, these factors allow for the separation of micronuclei from macronuclei using differential centrifugation.

Figure 4

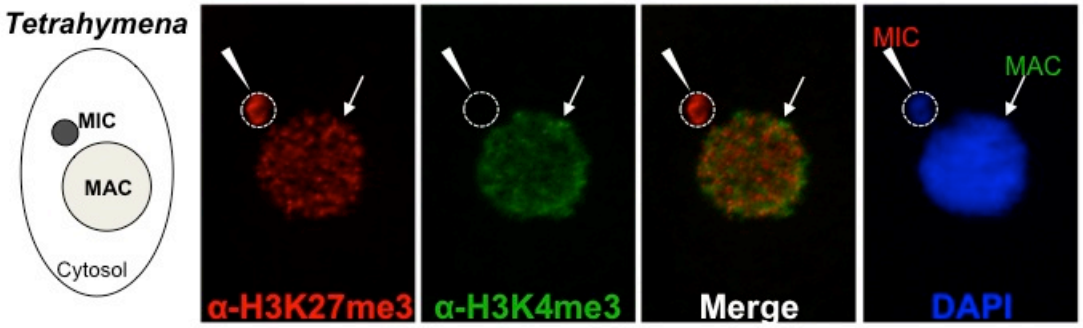


Figure 4. **Micronuclei and macronuclei are functionally distinct**. Cartoon depiction of *Tetrahymena* and co-immunofluorescence staining using α-H3K4me3 and α-H3K27me3 antibodies in *Tetrahymena*.

However, the decades-old protocol for nuclear preparation from *Tetrahymena* has been optimized and overwhelmingly used for the purification of macronuclei.[113] Purification of micronuclei is considerably more problematic because a single macronucleus has 20 times the contaminating effect than a micronucleus. For example, a fraction containing a 1:1 ratio of macronuclei to micronuclei actually contains a 20:1 ratio of macronuclear to micronuclear chromatin. Therefore, I initially had to optimize the nuclear preparation protocol to significantly minimize macronuclear contamination (the protocol is detailed in the **Methods** section of this chapter).

By optimizing this differential centrifugation approach, I was routinely able to obtain nuclear fractions that were >99.9% pure micronuclei **(Figure 5A)**. This ensured that macronuclear chromatin content was <2% of the micronuclear samples used during subsequent analyses. Once purified, micronuclear histones were acid-extracted and resolved by reverse-phase HPLC **(Figure 5B)**. SDS-PAGE separation and Coomassie staining of H3-containing HPLC fractions (56-58) showed that I was able to achieve an extremely high degree of purification for micronuclear H3, which occurs in two electrophoretically distinct forms *in vivo* **(Figure 5C).** The electrophoretically slow form of micronuclear H3, $H3^S$, is the same sequence as macronuclear H3, and the fast form, $H3^F$, is derived from $H3^S$ during a micronucleus-specific proteolytic processing step that irreversibly removes the first six amino acids from the $H3^S$ N-terminus.[123]

Figure 5

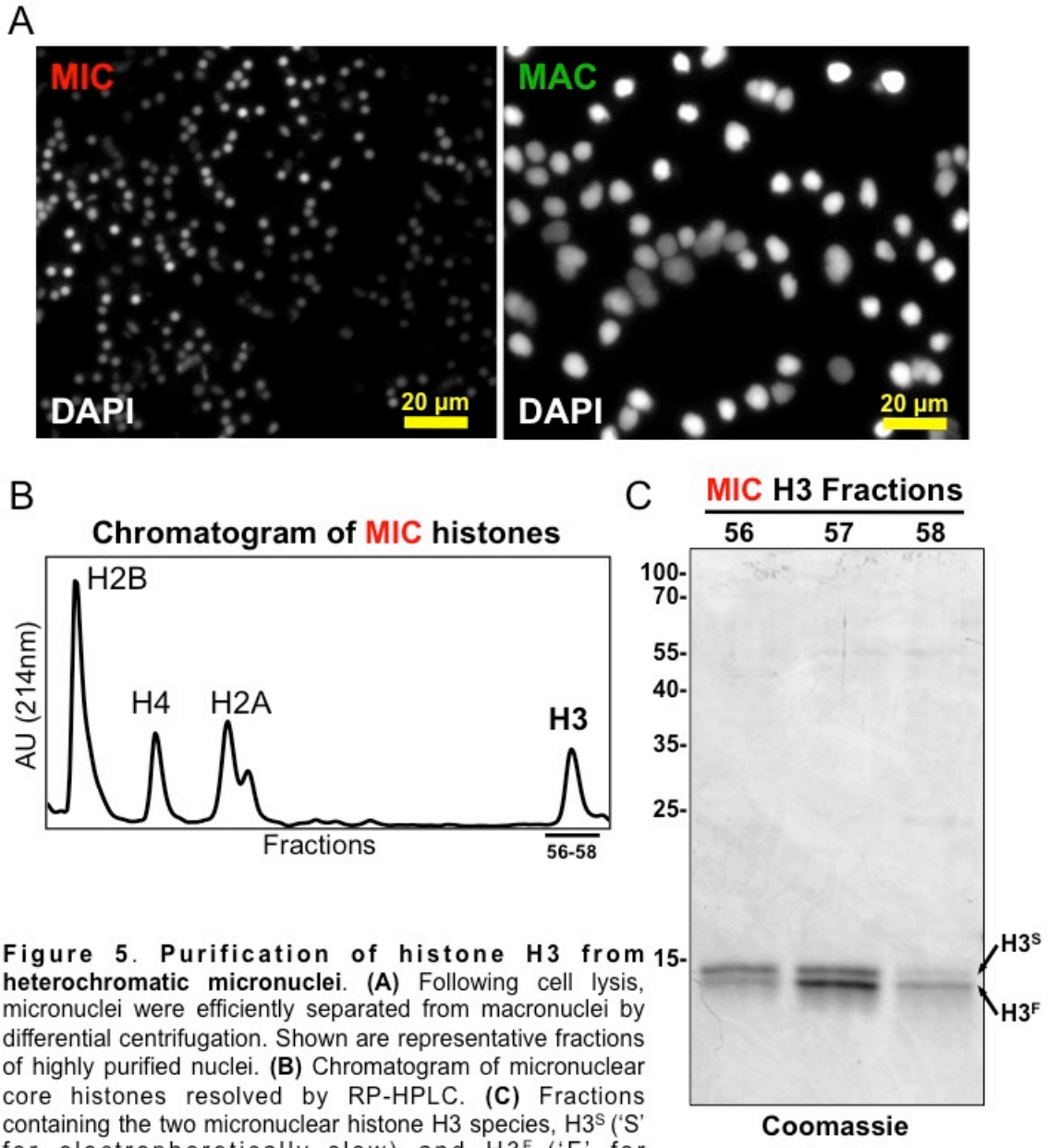


Figure 5. Purification of histone H3 from heterochromatic micronuclei. **(A)** Following cell lysis, micronuclei were efficiently separated from macronuclei by differential centrifugation. Shown are representative fractions of highly purified nuclei. **(B)** Chromatogram of micronuclear core histones resolved by RP-HPLC. **(C)** Fractions containing the two micronuclear histone H3 species, $H3^S$ ('S' for electrophoretically slow) and $H3^F$ ('F' for electrophoretically fast), which are stained with Coomassie following SDS-PAGE

2.3B COMBINATORIAL MODIFICATION STATES OF MIC HISTONE H3 TAILS

To determine the histone PTM states associated with micronuclear H3, highly purified $H3^S$ and $H3^F$ from micronuclei of vegetative cells were analyzed using HILIC-MS/MS. HILIC offered sufficient separation of the large, hydrophilic, highly modified H3 peptides. In fact, under these chromatographic conditions, N-terminal tails from $H3^F$ (amino acids 1-50) and $H3^S$ (amino acids 7-50) were separated into distinct fractions (data not shown). In addition, ETD provided complete to near complete fragment ion coverage allowing for interrogation of modifications that occur on single peptide molecules.

Our mass spectrometric analysis of micronuclear H3 PTMs is represented in schematic form **(Figure 6)**. This analysis revealed that, similar to condensed chromatin in higher eukaryotes, $H3^S$ and $H3^F$ were collectively hypoacetylated.[29] The few multiply acetylated states on $H3^S$ and $H3^F$ are devoid of methylation marks in *cis* and likely represent newly deposited histone H3 in micronuclei.[142] H3K4me1, which is commonly associated with euchromatin and histone deposition[143], was also detected on micronuclear $H3^S$, but was absent on $H3^F$, as H3K4 is missing on this proteolytically processed form. As expected from our previous studies, H3K9 methylation was not detected on any isoforms of micronuclear H3, while H3K27me3 and the mitosis associated PTM H3S10ph were only detected on $H3^F$.[57, 115, 128, 144] Surprisingly, we detected methylation on H3K23, which was specific to $H3^F$ and mostly occurred in combination with H3K27me. To validate the presence of H3K23 methylation within *Tetrahymena*, we obtained existing antibodies reported to recognize H3K23me1 and H3K23me2, and showed that these PTMs are detectable on micronuclear H3 **(Figure 7A-B)**. We also

Figure 6

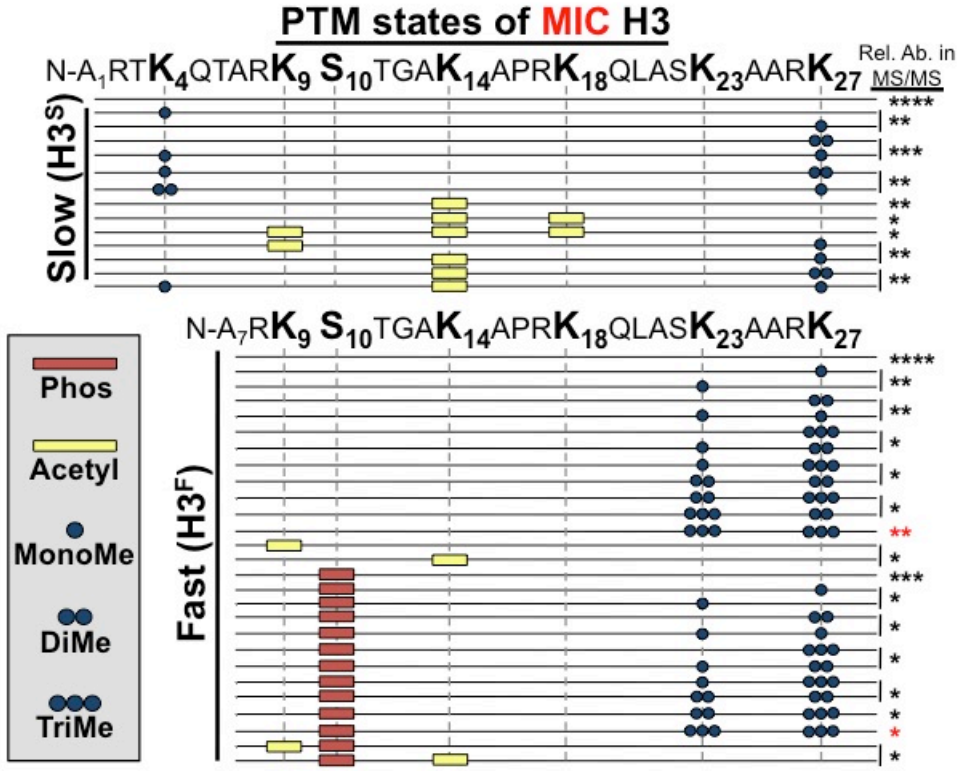


Figure 6. **Micronuclear H3 PTM states**. Summary of modified forms of micronuclear histone H3 N-terminal peptides (residues 1-50 for $H3^S$ and 7-50 for $H3^F$) from vegetatively growing *Tetrahymena* as detected by HILIC-MS/MS. Sequences of $H3^S$ and $H3^F$ are shown up to K27 since no modifications were detected on residues 28-50. Each row in the table represents a single, discrete, and uniquely modified H3 N-terminal peptide, and each dotted horizontal line beneath a bolded residue represents a site where the specified modification was identified. The relative abundance of each peptide, or multiple isobaric peptides identified within the same spectrum (grouped by brackets on the right), is highlighted by the number of asterisks. For example, peptides with four asterisks are more abundant than peptides with one asterisk. Forms of the $H3^S$(1-50) peptide with four, five, and six methyls were detected, but due to the isobaric character of each group of peptides and the low levels of these species it was not possible to determine the locations of the methyl modifications. Abbreviations: MonoMe, monomethylation; DiMe, dimethylation; TriMe, trimethylation.

Figure 7

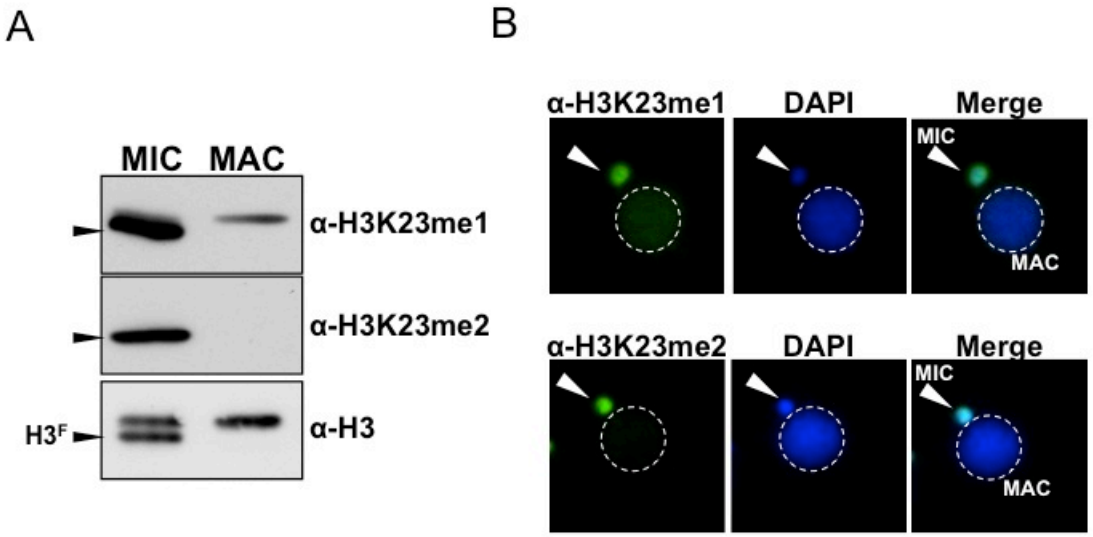


Figure 7. Presence of H3K23me1 and H3K23me2 in *Tetrahymena*. **(A)** Western blot analysis of micronuclear and macronuclear extracts and **(B)** indirect immunofluorescence staining using commercially available antibodies that recognize mono- or di-methylated H3K23. Black arrowheads (in part A) point to $H3^F$.

detected H3K23me1 in macronuclei, but this signal is likely non-specific given that α-H3K23me1 binds to acetylated histones to an appreciable degree.[55]

Although a few studies have identified H3K23 methylation through Edman degradation, mass spectrometry, and available commercial antibodies (α-H3K23me1 and α-H3K23me2 only), a biological understanding of this modification has been lacking.[55, 145, 146] There is *in vitro* evidence showing that HP1 can bind to H3K23 methylated peptides, thereby linking this mark to heterochromatin.[55] Curiously, levels of H3K23 mono, di, and tri-methylation are exceedingly low, and in most cases H3K23me3 is undetectable in the somatic cell lines studied to date.[55, 74, 145] In this work, we focused on characterizing H3K23me3 since this PTM was highly associated with H3K27me3, and the H3K23me3/H3K27me3 PTM state was most abundant out of all the dually methylated H3K23/H3K27 states **(Figure 6, red asterisks)**. To investigate the function of this under-characterized histone PTM, we first had to develop an antibody specific for H3K23me3 as none were commercially available at that time.

2.3C GENERATION OF A POLYCLONAL ANTIBODY SPECIFIC FOR H3K23ME3

Antibodies against H3K23me3 were generated using a synthetic H3 peptide containing residues 18-36 of the H3 tail and the H3K23me3 modification **(Figure 8A, underlined in red)**. Our initial ELISA analysis revealed that the serum was unable to distinguish between unmodified H3 (18-36) peptide and H3K23me3 (18-36) modified peptide (data not shown). As detailed in the **Methods** section above, we obtained H3K23me3 specific antibodies using affinity purification and a smaller bait peptide (H3K23me3[19-27]). ELISA analysis of the affinity-purified antibody was then carried

Figure 8

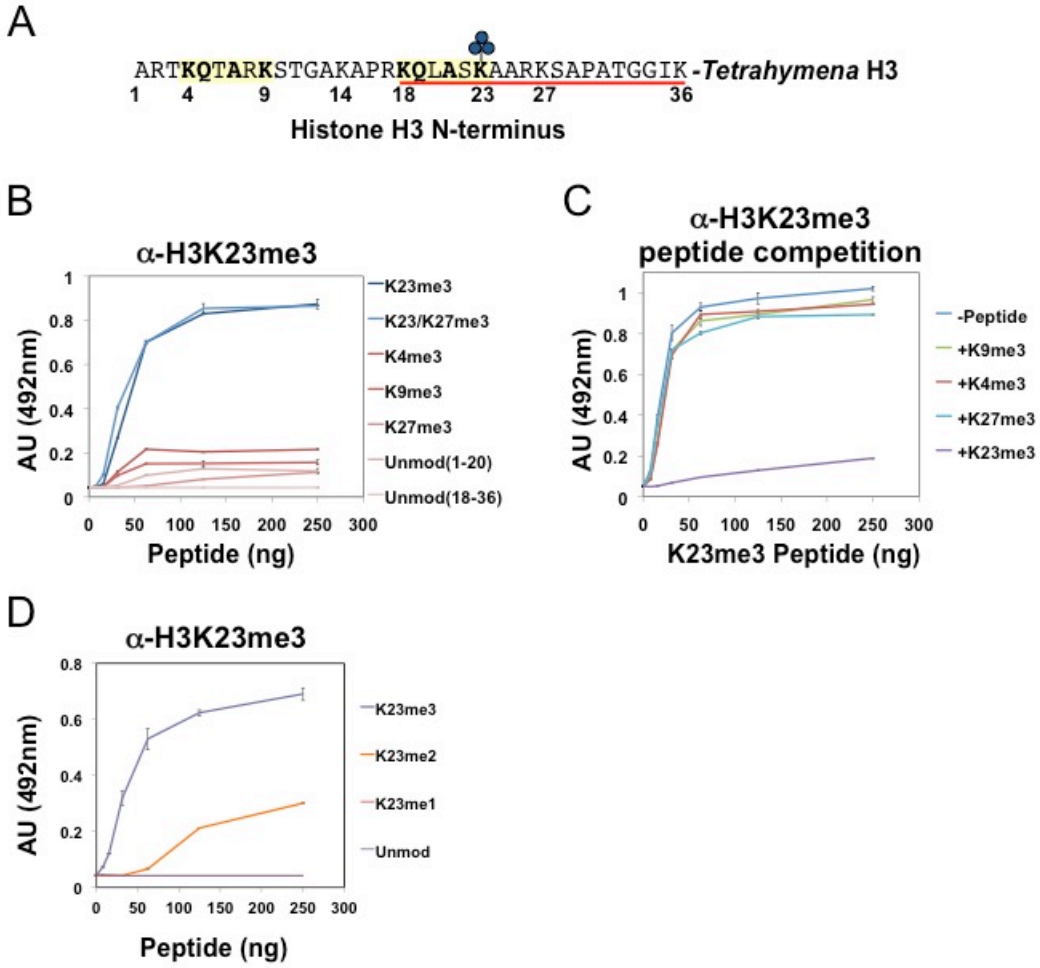


Figure 8. **Development of an H3K23me3-specific antibody**. **(A)** The first 36 amino acids of the N-terminal tail of *Tetrahymena* histone H3 (major isoform). Yellow boxes and bolded residues denote the conserved KQXAXK motif flanking H3K9 and H3K23. The epitope used to generate antibodies is underlined in red. **(B)** ELISA was used to test the specificity of α-H3K23me3 against the indicated peptide epitopes. **(C)** ELISA analysis of α-H3K23me3 binding to the H3K23me3 epitope in the presence of excess competing peptides. **(D)** ELISA analysis α-H3K23me3 binding to distinct methylation states of H3K23. (B-D) Data points represent mean ± SD of triplicate samples from a representative experiment.

out to test for specificity. As shown in **Figure 8B**, the antibody was highly specific for peptides containing the H3K23me3 mark. Furthermore, preincubation of the antibody with excess H3K23me3 peptide, but not H3K4me3, H3K9me3, or H3K27me3 peptides, competed away immunoreactivity for H3K23me3 **(Figure 8C)**. The antibody was also tested against distinctly methylated forms of H3K23 and only the H3K23me2 peptide was weakly recognized **(Figure 8D)**. Notably, the antibody did not bind to the H3K9me3 peptide, even though the neighboring sequence of H3K9 is very similar to that of H3K23 **(Figure 8A, yellow boxes)**. Thus, the antibody is highly specific for H3K23me3, and importantly, the presence of H3K27me3 did not occlude it from binding to the dual H3K23me3/H3K27me3 PTM state we identified in *Tetrahymena*. This antibody, referred to as "α-H3K23me3" hereafter, was used in all of the subsequent studies described below.

2.3D H3K23ME3 IS A GERMLINE SPECIFIC HISTONE PTM

To confirm the utility of α-H3K23me3 as a probe for following H3K23me3 in *Tetrahymena*, we tested the antibody in Western blots using acid extracted histones from highly purified micronuclei and macronuclei. In line with our HILIC-MS\MS analysis, the α-H3K23me3 antibody specifically recognized $H3^F$ of micronuclear histone H3 **(Figure 9A)**. The antibody also failed to recognize recombinant H3 indicating that H3K23me3 is required for effective epitope recognition. Finally, we tested the efficacy of this antibody in immunofluorescence assays. α-H3K23me3 specifically stained the micronucleus and was completely excluded from the macronucleus and the cytosol **(Figure 9B)**. Our results confirm that H3K23me3 is present *in vivo*, and exclusively

Figure 9

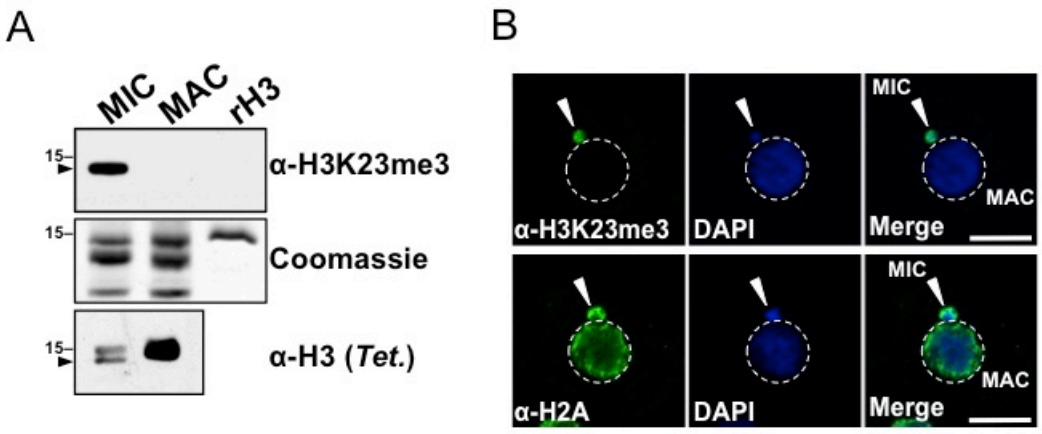


Figure 9. **H3K23me3 is a germline specific PTM in *Tetrahymena*. (A)** Western blot analysis of α-H3K23me3 specificity using micronuclear and macronuclear extracts, and recombinant human H3 (rH3). Black arrowhead points to $H3^F$. **(B)** Vegetatively growing *Tetrahymena* were examined by indirect immunofluorescence using either α-H3K23me3, or α-H2A as a control (white arrowheads indicate micronuclei). Bar, 10 μm.

Figure 10

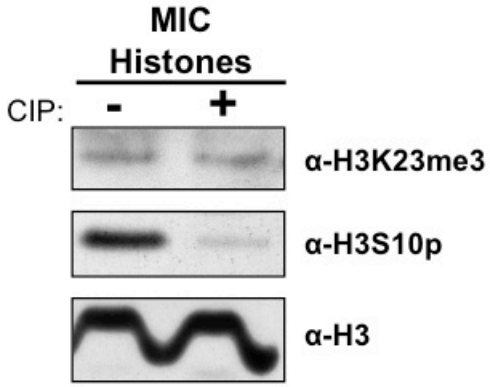


Figure 10. α-H3K23me3 binding after phosphatase treatment. α-H3K23me3 is not blocked by nearby potential phosphorylations (such as S22) in micronuclear H3 of vegetatively growing *Tetrahymena*. Abbreviation: CIP, calf intestinal phosphatase.

localized to $H3^F$ isoform in micronuclei of *Tetrahymena*. Furthermore, calf intestinal phosphatase treatment of micronuclear H3 did not increase the levels of H3K23me3, suggesting that neither H3S10 phosphorylation nor other unknown phosphorylation events on micronuclear H3 do not occlude the antibody form binding to H3K23me3 during vegetative growth **(Figure 10)**. Coupled with the fact that we did not detect H3K23me3 in our previous ETD analysis of macronuclear H3, these observations suggest H3K23me3 is exclusive to the *Tetrahymena* germline.

2.3E H3K23ME3 IS SIGNIFICANTLY UPREGULATED DURING MEIOSIS

The aforementioned biological experiments were carried out using *Tetrahymena* maintained in an asexual (or vegetative) growth cycle. Since we found that H3K23me3 was restricted to the germline micronucleus during vegetative growth, we wanted to explore the possibility that H3K23me3 would play an important role during conjugation, when micronuclear function is most evident. The developmental stages of conjugation include pairing of opposite mating types, meiotic division of micronuclei, fertilization, mitotic division of zygotic nuclei, and micronuclear differentiation (to form new somatic macronuclei) while the parental macronuclei degrade **(Figure 11)**.[111] Importantly, conjugation proceeds in a synchronous manner, allowing the study of histone PTMs at specific developmental windows.

Whole-cell-extracts were collected at specific time points during conjugation, and probed with α -H3K23me3 as well as other antibodies routinely used to gauge the progression of conjugation. Western blot analysis revealed that, compared to H3K23me3 levels detected in vegetative or starved cell extracts, H3K23me3 levels greatly increased

Figure 11

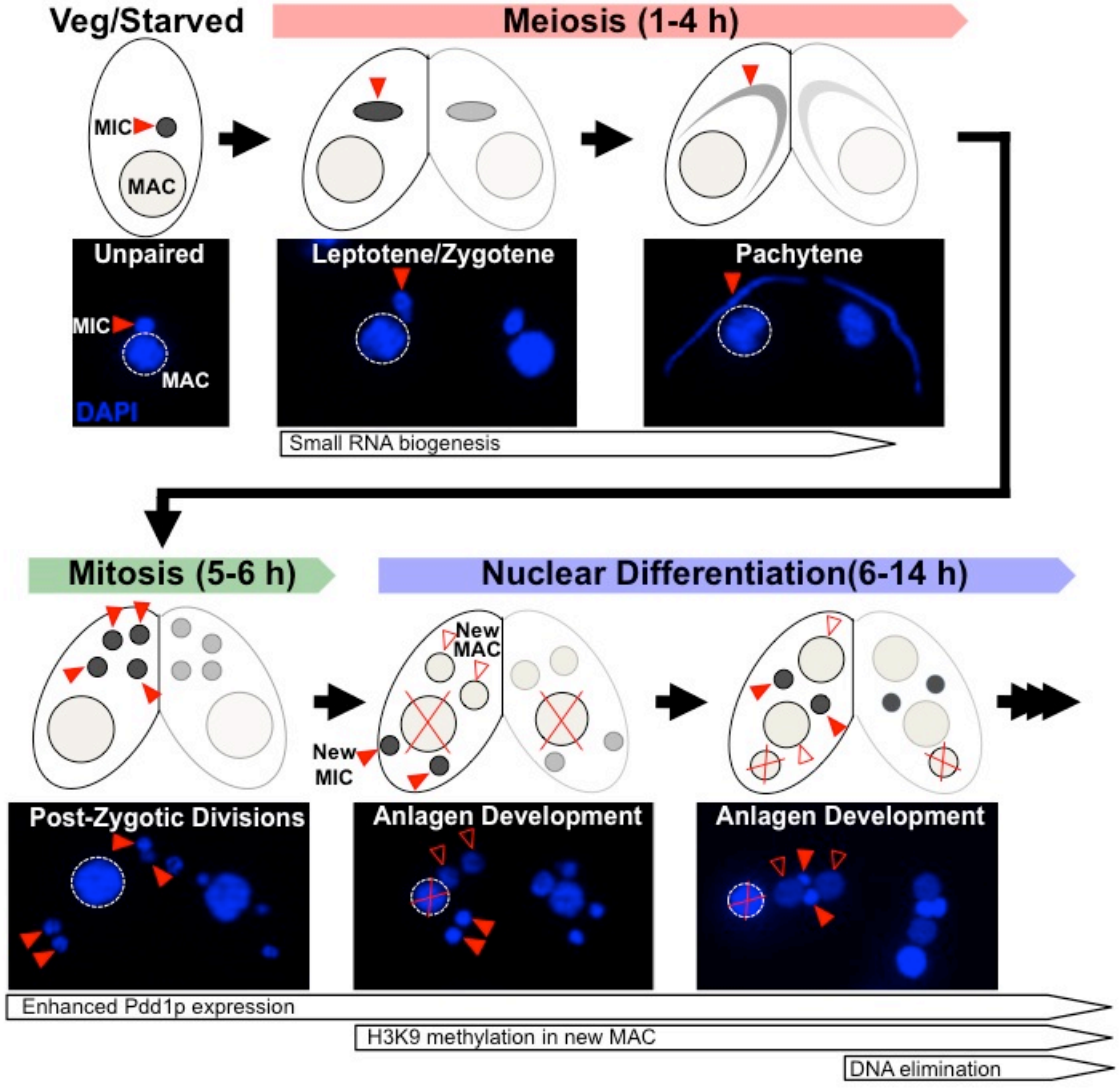


Figure 11. **Developmental stages of *Tetrahymena* conjugation**. During conjugation the MAC undergoes degradation (red "X") and the MIC divide mitotically and meiotically to give rise to four pluripotent nuclei of which two are selected to differentiate into new MAC. During this process several dramatic nuclear reorganization event takes place, including DNA elimination of MIC-limited, repeat DNA. These sequences are targeted by H3K27 and H3K9 trimethylation at hour 9 of conjugation. Conjugation yields four progeny that are genetically distinct from their parents.

early in conjugation, peaking around hour 4, and diminishing to pre-conjugation levels by hour 8 **(Figure 12A)**. Interestingly, the increase in H3K23me3 signal occurred before H3S10 phosphorylation **(Figure 12B)**, which is a conserved histone PTM marker for mitotic division. This pattern suggests that H3K23me3 is increased during meiosis, the developmental stage that precedes micronuclear mitosis.

The dynamic regulation of H3K23me3 levels early in conjugation, as revealed by Western blot analysis of whole-cell-extracts, suggested that this mark is involved in early stages of meiotic division. Thus, we performed immunofluorescence assays on whole cells to characterize the sub-cellular localization of H3K23me3 during conjugation **(Figure 13)**. We observed that H3K23me3 was exclusively localized to micronuclei throughout conjugation. While increased H3K23me3 coincided with meiotic entry, the PTM was still present, although much reduced, during stages of mitosis as evidenced by our Western blot, immunofluorescence, and HILIC-MS/MS data **(Figures 12B, 13, and 14)**. A small amount of H3K23me3 was detected on $H3^{S}$ during conjugation **(Figure 12A)**, but its absence in macronuclei (as determined by immunofluorescence) suggests $H3^{S}$K23me3 is likely micronuclear in origin and at levels below the detection limit of HILIC-MS/MS **(Figures 13 and 14)**. Taken together, our findings suggest that H3K23me3 may play an important role during meiotic division, as its levels are specifically elevated during this critical developmental stage.

The physical linkage between H3K23me3 and H3K27me3 in vegetative and post-meiotic chromatin prompted us to examine the dynamic changes to $H3^{F}$K27me3 during meiosis. While macronuclear $H3^{S}$K27me3 levels remain constant throughout conjugation, we found that $H3^{F}$K27me3 levels in the micronucleus increased, peaked

Figure 12

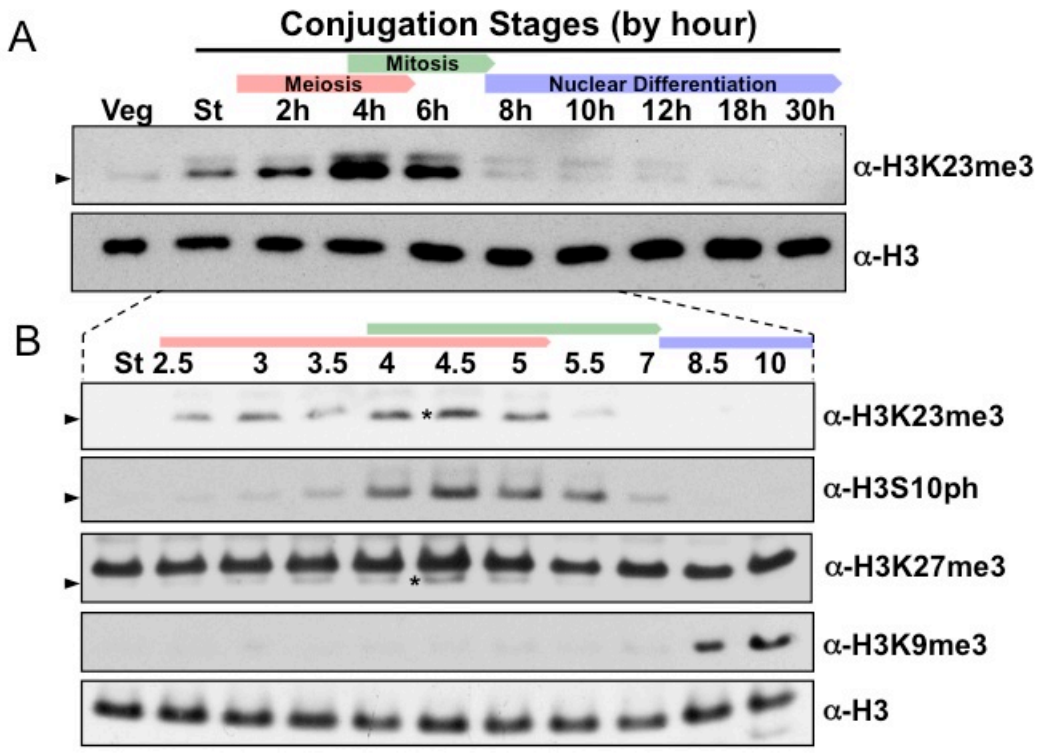


Figure 12. H3K23me3 levels are dramatically increased during early meiosis. (A-B) Whole-cell-extracts of wildtype cells were collected at the indicated time points during conjugation, resolved on SDS-PAGE gels, and analyzed by Western blotting. Color-coded arrows highlight the period of each specified developmental stage. Black arrowhead points to $H3^{F}$. Asterisks point to the peak of H3K23me3 and H3K27me3 in $H3^{F}$.

Figure 13

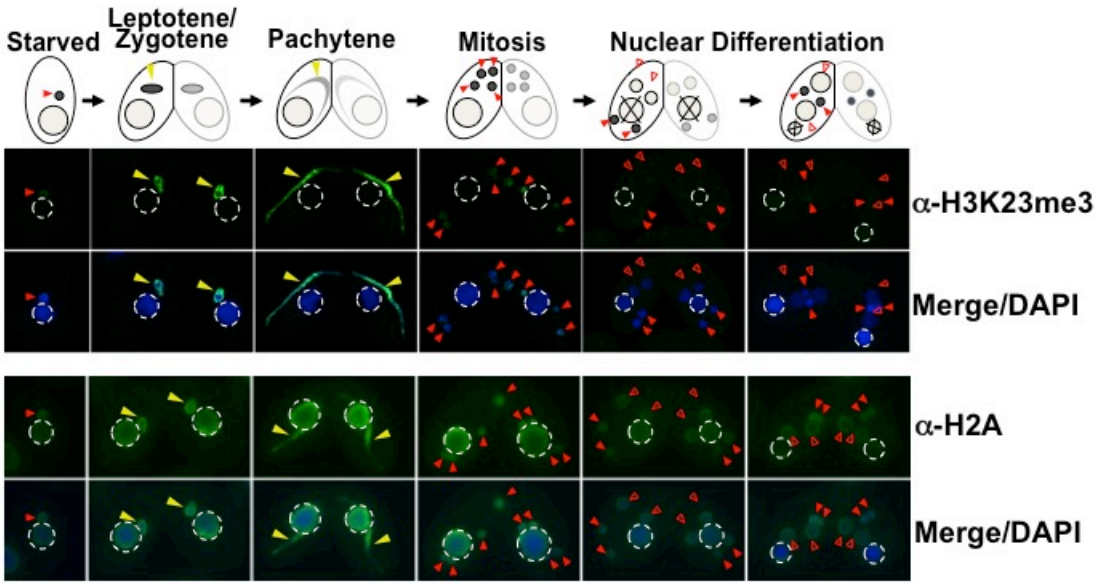


Figure 13. H3K23me3 is a germline specific PTM in conjugating *Tetrahymena*. Indirect immunofluorescence analysis of conjugating *Tetrahymena* using α-H3K23me3. Red arrowheads point to micronuclei. Yellow arrowheads point to meiotic micronuclei. During nuclear differentiation, the parental macronucleus (labeled with 'X' in the cartoon) degrades while two post-zygotic nuclei begin to differentiate to new macronuclei (open red arrowheads). α-H2A is a control stain for all nuclei.

Figure 14

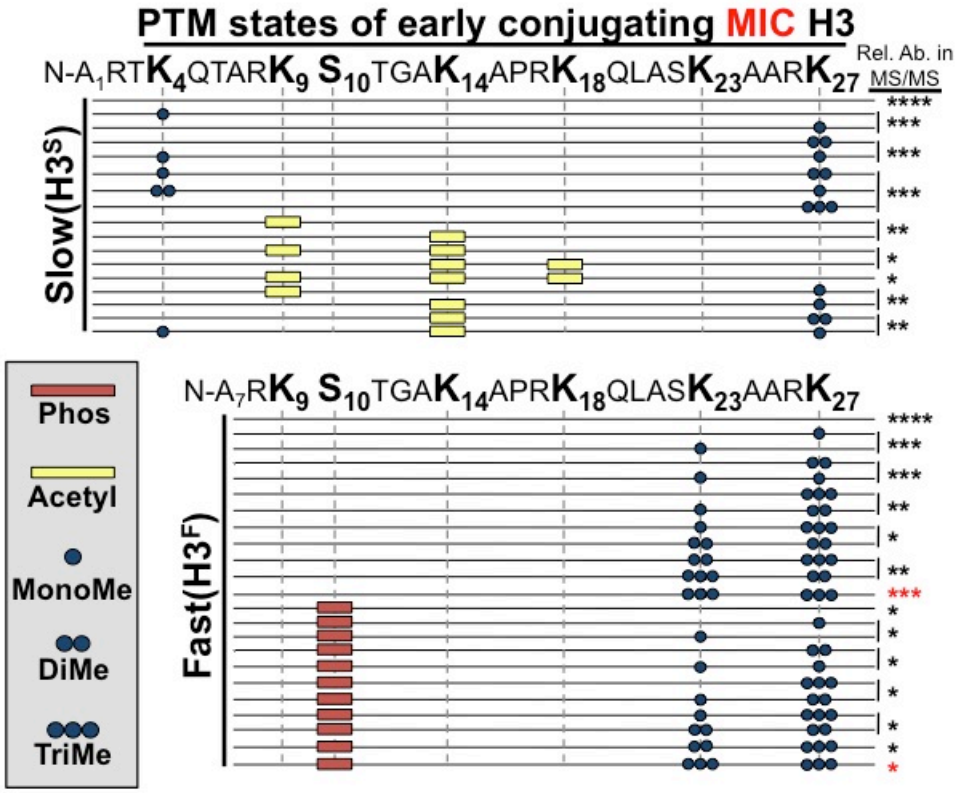


Figure 14. PTM states of H3 from early conjugating cells. Combinatorial modification states of micronuclear H3 from conjugating *Tetrahymena*. Summary of modified forms of histone H3 N-terminal peptides, as detected by HILIC-MS/MS, from *Tetrahymena* micronuclei of hour 5 conjugating cells. Each row represents a uniquely modified H3 N-terminal peptide species (residues 1-50 for $H3^S$ and 7-50 for $H3^F$), and each dotted horizontal line beneath a bolded residue represents a site where the specified modification was identified. No modifications were detected on residues 28-50 on $H3^S$ and $H3^F$. The relative abundance of each peptide, or multiple isobaric peptides identified within the same spectrum (grouped by brackets on the right), is highlighted by the number of asterisks. Forms of the $H3^S$(1-50) and $H3^F$(7-50) peptides were detected, but due to the isobaric character of each group of peptides and the low levels of these species it was not possible to determine the locations of the methyl modifications. Abbreviations: MonoMe, monomethylation; DiMe, dimethylation; TriMe, trimethylation; Acetyl, acetylation; Phos, phosphorylation.

Figure 15

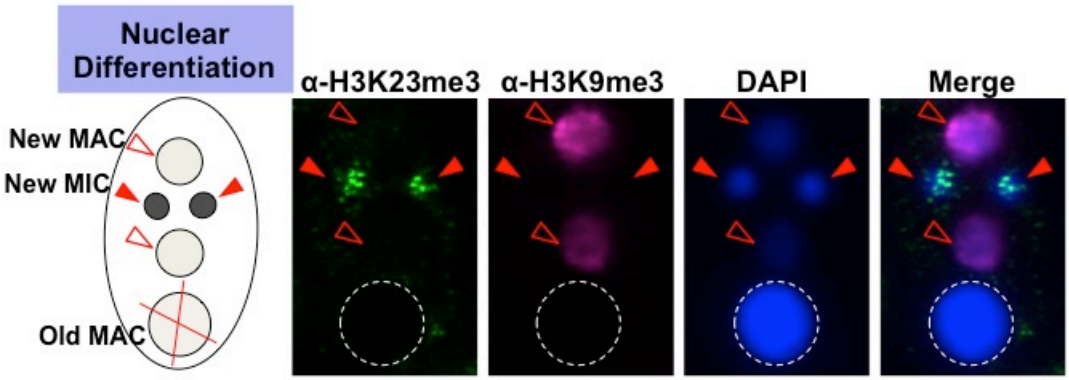


Figure 15. H3K9me3 and H3K23me3 do not overlap in *Tetrahymena*. Co-immunofluorescence staining of conjugating *Tetrahymena* (at nuclear differentiation stage) using α-H3K23me3 and α-H3K9me3 antibodies. Red arrowheads point to micronuclei and open red arrowheads point to the new developing macronuclei. The parental macronucleus (Old MAC) is marked by an 'X' in the cartoon and by dotted circles in the immunofluorescence images.

(Figure 12B, asterisks) and concurrently diminished with H3K23me3 levels. In contrast to the coordinate increase of H3K23me3 and H3K27me3 observed during meiosis, H3K9me3, which specifically emerges during nuclear differentiation, was only detected after levels of H3K23me3 diminished **(Figure 12B)**. The fact that low levels of H3K23me3 detected during nuclear differentiation are present in micronuclei, and not the H3K9me3-enriched developing macronuclei **(Figure 15)**, further defines distinct functions within heterochromatin for these PTMs.

2.4 DISCUSSION

2.4A GERMLINE SPECIFIC FUNCTIONS OF H3K23ME3

We used ETD (electron transfer dissociation) mass spectrometry to screen for combinatorial histone PTMs of micronuclear H3. We identified H3K23 methylation as a micronuclear specific PTM on $H3^F$ and used an antibody against H3K23me3 to demonstrate that this mark becomes highly enriched during early stages of meiosis. Furthermore, we showed that H3K23me3 is completely lacking in the somatic macronucleus throughout all stages of *Tetrahymena* development, providing strong evidence that H3K23me3 is a germline specific histone PTM in *Tetrahymena*. The increase of H3K23me3 in meiosis is noteworthy since little is known about the role of histone PTMs in governing the structure and function of meiotic germline chromatin. With the exception of γH2A.X (detailed in Chapter 4), no other histone PTM has been shown to be dramatically upregulated during meiosis, to the best of our knowledge. Previous studies have shown, through mass spectrometry, that H3K23 methylation levels are exceedingly rare in somatic cells and in particular, H3K23me3 and H3K23me2 levels

were not detected in most cases. Our observations in *Tetrahymena* suggest that H3K23me3 may have an evolutionary conserved role in germline heterochromatin and therefore may be enriched in germ cells.

Supporting a functional and specific role of H3K23me3 in germline chromatin, we noticed that this mark appeared to be actively demethylated from the two pronuclei that are destined to differentiate into new macronuclei at around hour 6 of conjugation (data not available). The loss of H3K23me3 in these nuclei preceded the emergence of H3K9 methylation. It is tempting to speculate that the dual H3K23me3/H3K27me3 signature is characteristic of germline chromatin in *Tetrahymena,* and that loss of H3K23me3 in micronuclei helps to trigger their differentiation into new macronuclei. Another implication from these observations is that H3K23me3 could help protect the underlying micronuclear genome from DNA damaging programs associated with DNA elimination in new macronuclei.

In addition to the dramatic upregulation of H3K23me3 during meiosis, this mark is present in micronuclear chromatin during vegetative growth, when micronuclei divide mitotically. Indeed, our ETD analysis showed that H3K23me3 and H3S10ph can occur on the same peptide indicating that H3K23me3 is also present in mitotic chromosomes. This suggests that H3K23me3 can be epigenetically inherited through mitotic divisions. What is the function of H3K23me3 in vegetative cells? Our immunofluorescence and mass spectrometric data indicated that H3K23me3 only marks a subset of micronuclear chromatin, indicating that it localizes to specific regions (*i.e.* telomeric, satellite, intergenic, *etc*.) of the micronuclear genome. Once the cells enter meiosis, the H3K23me3 associated chromatin could act as nucleation sites for the spreading and

propagation of H3K23me3, which contributes to the regulation of the underlying DNA elements. The levels of H3K23me3 in the new micronuclei are returned to baseline during late conjugation and again propagated through mitotic divisions in the progeny as detected by immunofluorescence (data not shown). Thus, H3K23me3 in vegetative cells may help to poise the underlying DNA for downstream processing associated with germline development during conjugation.

2.4B HISTONE PTM CROSSTALK HIGHLIGHTS THE MANY FACETS OF H3K27ME3

Our observations suggest that H3K27me3 in *Tetrahymena* occurs in many distinct combinations of PTMs during vegetative growth and conjugation that may function to silence, index, or otherwise poise the underlying DNA elements for specific biological processes. In the micronucleus, H3K27me3 is present on H3 simultaneously with other well-known PTMs associated with repression and chromatin condensation. This heterochromatic 'signature' includes methylation on H3K23, hypoacetylation, and proteolytic clipping of H3. However, despite the co-occurrence of H3K27me3 and H3K23me3 on a significant subset of micronuclear H3 peptides, the deposition of H3K27me3 does not always overlap with that of H3K23me3. While H3K27me3 is only found on $H3^F$ in the micronucleus, it can also occur on $H3^S$ in the macronucleus. Furthermore, during conjugation, H3K27me3 coordinates with H3K9me3 in new macronuclei to target intergenic sequences for heterochromatin formation and eventual elimination. Thus, the co-occurrence of clipping, H3K23me3, or H3K9me3 may underlie the many putative functions of H3K27me3 in *Tetrahymena*.

Several lines of evidence support the above hypothesis. Firstly, the *Tetrahymena* chromodomain containing protein Hhp1p[121], shown to bind H3K27me3, localizes to macronucleus but not the micronucleus. Thus, either Hhp1p is not imported into micronuclei or the occurrence of H3K23me3/clipping in *cis* with H3K27me3 affects its binding in micronuclear chromatin. Supporting the latter, we found that the presence of H3K23me3 in *cis* severely lowers the ability of Hhp1p to bind to H3K27me3 peptides **(Figure 16)**. Secondly, another *Tetrahymena* chromodomain protein, Pdd1p, that binds to H3K27me3 upstream of heterochromatin formation and DNA elimination is affected by the presence of H3K23me3 in *cis* **(Figure 16)**. Thus, at least two known H3K27me3 readers are negatively affected by the presence of H3K23me3. By comparison, the as yet identified H3K23me3 reader would likely have to bind to the binary H3K23me3\H3K27me3 mark since they are predominantly found to occur simultaneously in *Tetrahymena* during vegetative growth and early conjugation.

2.4C MACRONUCLEI LIKELY CONTAIN TRANSCRIPTIONALLY OPPOSING SUBDOMAINS

As detailed in Chapter 1 of this dissertation, the macronucleus is the site of gene expression during vegetative growth. Still, not all of its ~21,000 genes are expressed during this life cycle and, of the expressed genes, most are either up- or down-regulated depending environmental or life cycle changes (*e.g.* cell density, nutrients in media, and induction of conjugation). A previous study identified at least 1,158 genes that were specifically silenced during vegetative growth and another 2,369 genes that were significantly up-regulated during starvation and conjugation.[147] Our observation that

Figure 16

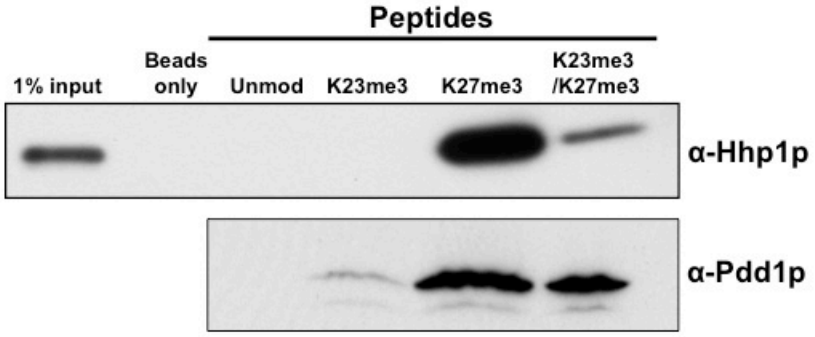


Figure 16. H3K23me3 affects H3K27me3 reader binding. Peptide pull downs from *Tetrahymena* extracts (for Hhp1p) and MBP-tagged Pdd1p. Western blot analysis of two known H3K27me3 readers, Hhp1p and Pdd1p following peptide pull downs.

H3K27me3 and H3K4me3 occupy distinct sub-nuclear domains within the macronucleus suggests that transcription may be regulated, in part, by the localization of genes to transcriptionally repressive and permissive environments, respectively. Nevertheless, our previous study characterizing histone PTM states of macronuclear H3 showed an enrichment of H3K4me3 and acetylation, as well as H3K27me1 and H3K27me2, but not H3K27me3.[57]

CHAPTER 3: LOSS OF HISTONE H3 LYSINE 23 METHYLATION LEADS TO DEVELOPMENTAL DEFECTS IN TETRAHYMENA

3.1 INTRODUCTION

3.1A THE ENHANCER OF ZESTE FAMILY OF HISTONE METHYLTRANSFERASES

SET domain-containing proteins are enzymes that can be involved in the positive or negative regulation of gene expression. Named for their occurrence in ***S****uppressor of variegation 3-9*, ***E****nhance of zeste (E[z])* and ***T****rithorax* genes in *Drosophila*, SET domains are proposed to have methyltransferase activity towards lysine residues on histones and other targets.[148] E(z) is a highly conserved member of the Polycomb group proteins important for regulating homeotic genes during *Drosophila* development.[76] Biochemical studies have shown that E(z) is a member of the polycomb repressive complex 2 (PRC2) that writes H3K27me3, which is then read by the chromodomain containing Polycomb (Pc) protein of PRC1.[76] Similar repressive complexes have been described in mammals, invertebrates, plants and *Tetrahymena*.[76, 149, 150]

3.1B E(Z) FAMILY OF METHYLTRANSFERASES ARE GUIDED BY RNA

A important commonality among the E(z) family of methyltransferases is that their heterochromatin forming activity is thought to be guided by non-coding RNAs (ncRNA) *in vivo*. In higher eukaryotes, several studies have shown that PRC2 binds to different classes of ncRNA, including long ncRNA (*e.g.* HOTAIR) and small ncRNA

(<150bp). A recent study also showed that Ezh2 (the mammalian homolog of E[z]) can promiscuously bind to RNA in a sequence independent manner to regulate the genomic localization of PRC2.[84] Of note, most studies focusing on E(z) homologs and ncRNAs have been conducted in somatic cells.

Recent work by our group has shed light on the connectivity of E(z) homologs and ncRNAs in the germline of *Tetrahymena*.[150] Specifically, our study showed that Ezl1p, which catalyzes H3K27/H3K9 trimethylation of IESs in developing macronuclei, is controlled by the upstream activities of piRNA-like small RNA (called scanRNA) during conjugation. Very early in conjugation (during leptotene/zygotene of meiotic prophase), certain sections of micronuclear genome are transcribed by RNA polymerase II. Although the mechanism of RNA polymerase II targeting is not known, these transcripts are enriched for micronuclear-limited sequences, which include intergenic, repetitive, centromeric, and transposon-like sequences.[151] Next, the transcripts are processed into scanRNA (28-32 bp) by a dicer-like protein, called Dcl1p. The scanRNAs are then bound by an argonaute homolog, Twi1p, in the micronucleus. These scanRNA-Twi1p complexes then get shuttled to the parental macronucleus (at hour ~5 of conjugation) where they presumably 'scan' the genome for complementarity. It is thought that this process removes the complexes bearing complementary scanRNAs to the parental macronuclear genome. The remaining complexes are then shuttled into the new developing macronucleus (at hour ~9 of conjugation) where they target the Ezl1p-catalyzed trimethylation of H3K27 and H3K9 on chromatin bearing IESs. Finally, the conjugation specific chromodomain proteins, Pdd1p and Pdd3p, bind to methylated histone H3 to form heterochromatin, and to recruit DNA eliminating enzymes. This

process takes about 16 h to complete. Studies have shown that disrupting the generation of scanRNA in micronuclei during early conjugation leads to aberrations in heterochromatin formation in developing macronuclei during late conjugation.[150, 152, 153]

3.1C Dynamics of Meiotic Chromatin

Our observations in Chapter 2 indicated that the upregulation of H3K23me3 occurs at time points consistent with the leptotene, zygotene, and pachytene stages of meiotic prophase. These early stages mark critical periods of germline development and structural changes to chromatin that contribute to gametogenesis, genetic diversity, and genomic stability in all eukaryotes. During leptotene, the chromosomes begin to condense and form thread-like structures. Additionally, sister chromatids, which are tightly bound, begin to form the synaptonemal complexes necessary for homolog pairing and proper segregation. By zygotene, the chromatids are completely paired and programmed DNA double strand breaks begin to occur. These sites are marked by γH2A.X. Pachytene stage includes the recombination and repair of DNA double-strand breaks (DSBs). Additionally, the homologs are fully paired and synapsed by pachytene stage, which marks the peak of γH2A.X and the formation of DSBs. Notably, although several hundreds to thousands of DSBs are catalyzed during meiotic prophase, only a few break sites per cell actually end up recombining.[154] As a requisite process of meiotic division, DSBs are precisely controlled by cellular checkpoints, and dysregulation of these pathways can have deleterious consequences such as genomic instability and loss of progeny viability.[154-156]

Chromatin structure is a key factor in controlling programmed DSBs catalyzed by

the opportunistic, sequence-independent, and highly conserved endonuclease Spo11 during early meiotic prophase. The relatively open structural configuration of euchromatic regions is thought to make them more susceptible to Spo11 activity. Indeed, several recent reports have described a transcription-independent role for H3K4me3 in promoting DNA recombination "hot spots".[155]

Although these recent reports suggest that Spo11 activity is primarily guided by euchromatic marks, several reports have shown that proper heterochromatin formation is also a critical factor in regulating Spo11 activity. For example, loss of H3K9me3 at pericentric heterochromatin in *Suv39h* dn mice leads to genomic instability in both germ and somatic cells.[47] DSBs at pericentric heterochromatin are difficult to properly resolve because their repetitive nature complicates the repair mechanisms that utilize homologous recombination. Therefore, repair of DSBs at these sites are more prone to errors, which give rise to phenotypes associated with genomic instability. In the germline, these errors can manifest as phenotypes associated with progeny viability and sterility. In that regard, *Tetrahymena* is an interesting organism to study meiotic DSBs since meiotic micronuclei lack H3K4me3 and H3K9me3, and H3K23me3 is simultaneously upregulated with the occurrence of γH2A.X.

3.2 METHODS

3.2A CELL CULTURE AND STRAINS

Tetrahymena strains were maintained and conjugation was induced as described above (Chapter 2, Methods). *ΔEZL2 and ΔEZL*3 cells were generated following the same procedures used to obtain *ΔEZL*1 cells as previously described.[150]

3.2B WESTERN BLOT AND INDIRECT IMMUNOFLUORESCENCE ANALYSES

Immune-based assays were performed as described above (Chapter 2 *Methods*). Rabbit polyclonal antibodies used in Western blots in the chapter include: α-H3 (1:5,000; gift from C.D. Allis), α-H3K23me3 (1:100; made in this study), α-H3S10p (1:25,000; gift from C.D. Allis), α-H3K27me3 (1:2,500; #39155; Active Motif), α-alpha linker (1:5,000; gift from C.D. Allis), α-H4 penta acetyl (1:5,000; gift from C.D. Allis), α-Pdd1p (1:5,000; ab5338; Abcam), α-H3K27me1 (1:1,000; #07-448; Millipore), α-H3K27me2 (1:500; #07-452; Millipore), α-H3K23me1 (1:1,000; 39387; Active Motif), and α-H3K23me2 (1:1,000; 39653; Active Motif). HRP-conjugated α-rabbit secondary antibodies (GE Life Sciences) were used at dilution of 1:5,000 and detected by ECL plus reagent (Amersham) following manufacturer's procedures.

Tetrahymena immunofluorescence staining was performed in cells fixed in partial Schaudin's fixative as previously described.[140, 141] The primary antibodies used were: rabbit α-H3K27me3 (1:250; gift from T. Jenuwein), mouse α-H3K4me3 (1:250; ab1012; Abcam), rabbit α-H3K23me3 (1:100), rabbit α-H2A (1:250; control stain for *Tetrahymena*), mouse α-H3K9me3 (1:100; #39285; Active Motif), mouse α-γH2AX (1:250; #613402; BioLegend), rabbit α-Cna1p (1:5000; Gift of Harmit Malik, Fred Hutchinson Cancer Center, Seattle WA), rabbit α-H3K23me1 (1:250; Active Motif), and rabbit α-H3K23me2 (1:250; Active Motif). Secondary antibodies, used at a dilution of 1:2000, included Alex Fluor 647-conjugated goat α-mouse, Alexa Fluor 488-conjugated goat α-rabbit, and Alex Fluor 488-conjugated goat α-mouse (Molecular Probes). Slides

were mounted in ProLong Gold antifade reagent with DAPI (P36931; Life Technologies) and stored overnight at room temperature prior to imaging. Digital images were captured using a Zeiss Axio Observer fluorescence microscope equipped with an Apotome for optical sectioning and AxioCam MRm camera. Raw images were exported as tiff files and processed using Adobe Photoshop software CS3.

3.2C OBSERVATION OF SMALL RNAS IN *TETRAHYMENA*

Total RNA was extracted with Trizol reagent (Life Technologies) according to manufacturer's procedures and precipitated in ethanol. RNA pellet was dissolved in water and the OD_{260nm} was measured. Total RNA (15 μg) from conjugating cells were diluted in 2X loading buffer (95% formamide, 18 mM EDTA, 0.025% SDS, and Bromophenol Blue) and resolved in a denaturing acrylamide gel (1X TBE, 7 M Urea, 15% acrylamide) according to previously described methods[150] (see also "Gel purification of miRNA from total RNA" protocol from Life Technologies). Gels were visualized by soaking in 1.5 μg ethidium bromide per mL of TAE buffer. [152]

3.2D PCR

To check the expression status of *EZL3* in *EZL* knockouts, cells were grown to mid-log phase, their total RNA was extracted with Trizol reagent (Life Technologies) and precipitated in ethanol. Samples were resuspended in water and treated with TURBO DNase (Life Technologies) following the manufacturer's procedures. Reverse strand synthesis was carried out using poly dT primers, procedures and reagents from the Superscript II RT kit (Life Technologies). *EZL3* specific primers (Forward -

AATGGTCTATGCTAAAGGTGG; Reverse - TTTGAGAGTTCCTTGTCCATC) were used to assess transcript levels and primers that target the ribosomal protein RPL21 (Forward – AAGTTGGTTATCAACTGTTGCGTT and Reverse - GGGTCTTTCAAGGACGACGTA; sequence obtained from Aronica et al. [153]) were used as controls in end-point PCR assays.

Quantitative PCR was employed to assess the levels of *EZL3* expression during conjugation. Total RNA of conjugating cells were harvested at the indicated time points during conjugation and processed following procedures described above. qPCR was performed using an Applied Biosystems StepOne Plus Real-Time PCR System (Applied Biosystems, 4376600) with StepOne Software v2.1 and Power SYBR Green (Applied Biosystems, 4367659) to quantify relative mRNA abundance. *EZL3* specific primers used are shown above. mRNA levels were normalized to time point zero. Samples were analyzed in duplicate.

3.2E PROGENY VIABILITY ASSAY

For mating survival and progeny production tests, the *EZL3* gene deletion was reproduced as a somatic knockout[157] in the genetic background of strains CU428 and B2086. Three independent lines in each background were established by cloning single cells from drug-resistant transformants. Each line was starved, mated to each of the others and 44 single mating pairs isolated into separate drops of standard growth medium.[158] After three days growth, surviving cultures were replicated to medium containing 6-methyl purine. The 6MP-resistance allele is found in the silent germ line nucleus of CU428, but not its somatic nucleus; therefore, resistance indicates the

production of true genetic progeny. Survival and progeny production were scored and compared with parallel matings of the parental strains.

3.3 RESULTS

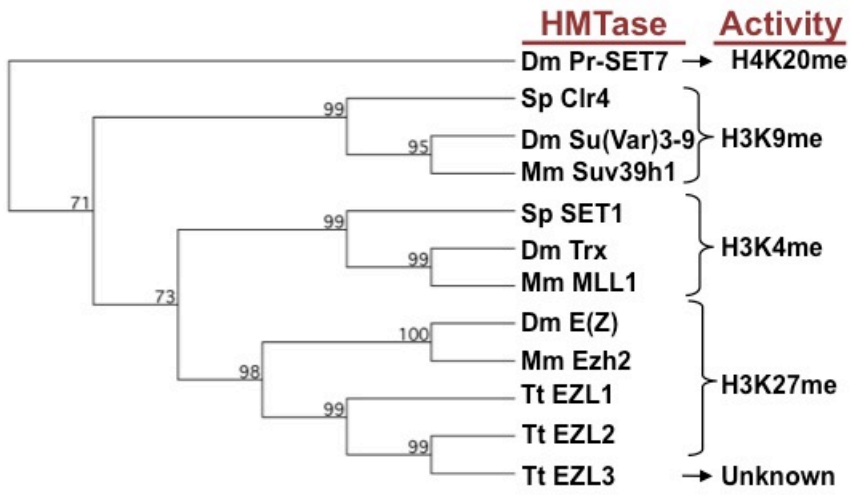
3.3A AN E(Z) HOMOLOG IS REQUIRED FOR H3K23 METHYLATION IN TETRAHYMENA

We sought to identify the methyltransferase responsible for H3K23 methylation to obtain better insight into the biological role of this modification in *Tetrahymena*. Using a-H3K23me3, we screened for the presence of H3K23me3 in cell lines deleted for various HMTs. Initially, we focused the highly conserved E(z) family of SET-domain containing proteins that catalyze the di- and tri-methylation of H3K27. We reasoned that disrupting H3K27me3 in *Tetrahymena* may influence H3K23me3 levels, since they co-occur on H3 tails.

Tetrahymena possess three E(z) homologs: Ezl1p (for Enhancer of zeste-like), Ezl2p, and Ezl3p **(Figure 17A)**. Previous studies have shown that Ezl1p and Ezl2p are required for H3K27 trimethylation during late and early conjugation, respectively.[150, 159] Moreover, Ezl2p is required for both $H3^{F}K27me3$ in the micronucleus and $H3^{S}K27me3$ in the macronucleus during vegetative growth. Ezl1p is not expressed during vegetative growth. While Ezl3p is expressed during vegetative growth and conjugation, its histone or non-histone targets are not known **(Figure 17B)**.

Figure 17

A



B

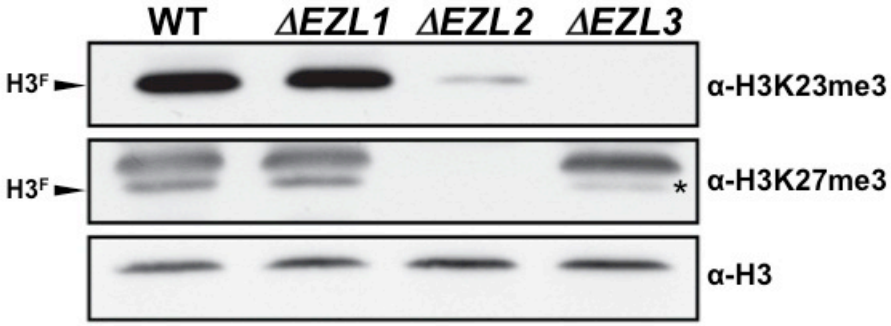
	Expression		Putative Activity	
	Veg	Conj	Veg	Conj
EZL1	-	+	-	H3K27me, late conjugation
EZL2	+	+	H3K27me	H3K27me, early conjugation
EZL3	+	+	Described in this thesis	Described in this thesis

Figure 17. **E(z) homologs in *Tetrahymena***. **(A)** Phylogenetic analysis of putative E(z) homologs of *Tetrahymena* with other known methyltransferases of higher eukaryotes. Abbreviations: Dm, *Drosophila melanogaster*; Sp, *Schizosaccharomyces pombe*; Mm, *Mus musculus*; and Tt, *Tetrahymena thermophila*. **(B)** Expression and activity of Ezl proteins during the *Tetrahymena* life cycle. This table was generated based on previous studies of our laboratory and others (references 150 and 159).

EZL knockouts (*ΔEZL1*, *ΔEZL2*, or *ΔEZL3)* growing vegetatively were screened for changes in H3K23me3 using α-H3K23me3 in Western blot and immunofluorescence assays. While H3K23me3 was present at wildtype levels in *ΔEZL1* cells, its levels were significantly reduced in *ΔEZL2* (lacking H3K27me3), and completely undetectable in *ΔEZL3* cells **(Figure 18A-B)**. These results suggest that expression of Ezl3p is required for H3K23me3 (as well as H3K23me1/2) **(Figure 19A-B)**. Supporting a role for Ezl3p in meiosis and H3K23 methylation, its expression was strongly upregulated during early conjugation **(Figure 20)**. Interestingly, H3K27me3 was moderately reduced on $H3^F$ in *ΔEZL3* cells **(Figure 18A, asterisk)**, while H3K27me1/me2 remained at or slightly above wildtype levels **(Figure 21)**, suggesting a functional link between the trimethylated forms of H3K27 and H3K23. Since H3K23me3 was most dramatically impacted in *ΔEZL3* cells, we proceeded to analyze *ΔEZL3* cells for phenotypes to better understand biological roles of H3K23me3 *in vivo*.

Figure 18

A



B

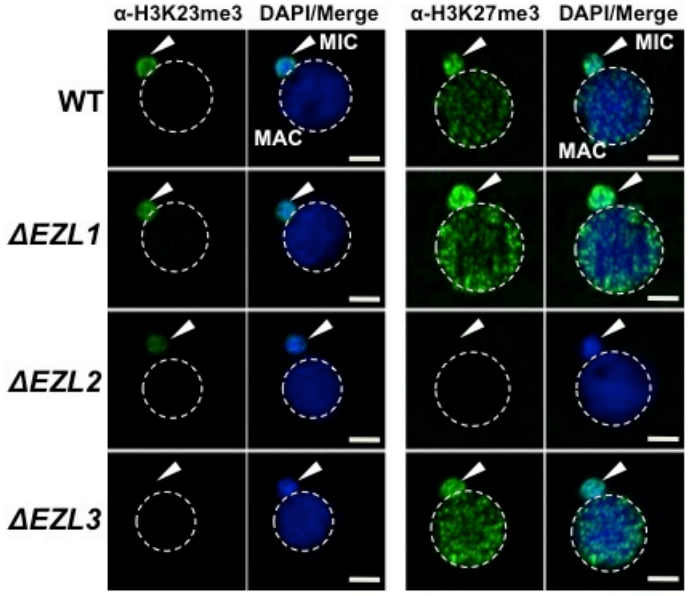


Figure 18. Ezl3p is required for H3K23 methylation in *Tetrahymena*. (A) Western blot analysis of whole-cell-extracts and **(B)** indirect immunofluorescence staining of wildtype (WT), *ΔEZL1*, *ΔEZL2*, and *ΔEZL3* cells grown under vegetative conditions. Asterisks mark $H3^F$. White arrowheads point to micronuclei. Bar, 5 μm.

Figure 19

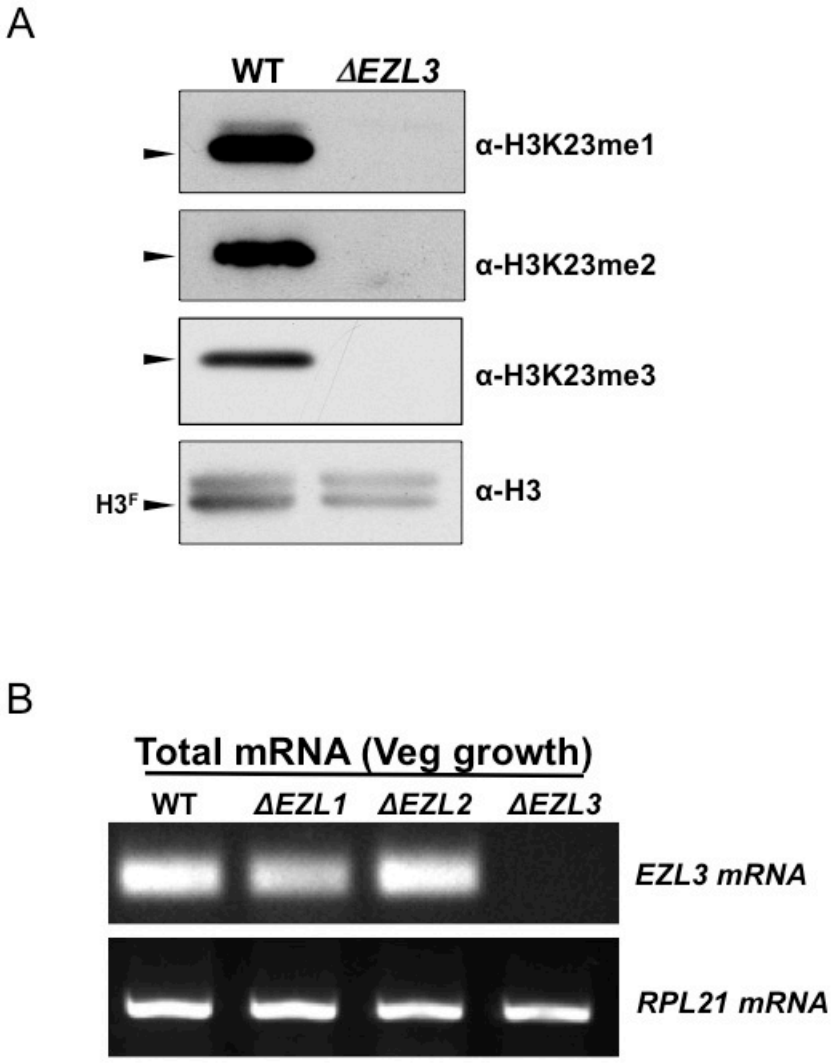


Figure 19. Ezl3p is required for mono-, di-, and tri-methylation of H3K23. **(A)** Western blot analysis of micronuclear acid extracts using antibodies that recognize mono-, di-, and tri-methylated H3K23. Black arrowheads (in part A) point to $H3^F$. **(B)** Total mRNA from the indicated strains growing vegetatively were purified and analyzed for the presence of *EZL3* mRNA transcripts. *RPL21* is a constitutively transcribed ribosomal gene used as a positive control.

Figure 20

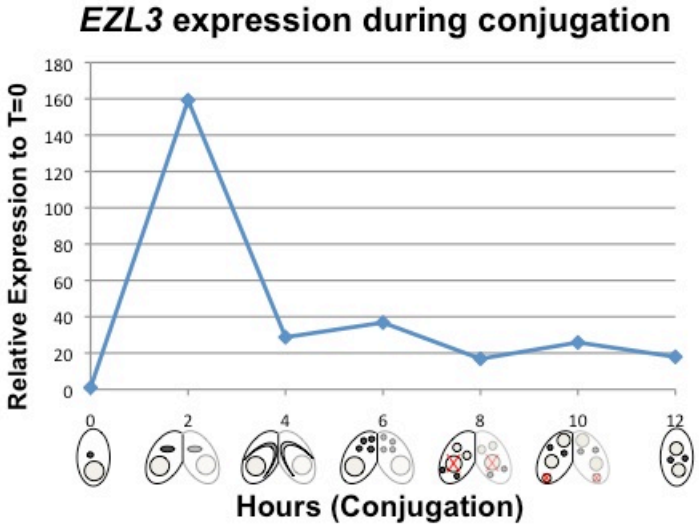


Figure 20. *EZL3* expression during conjugation. *EZL3* expression during conjugation as determined by qPCR. Normalization was carried out by comparison to *RPL21* expression, which is constitutive throughout conjugation. *EZL3* transcript levels were plotted relative to their levels at time 0. This trend is similar to the one reported in the *Tetrahymena* gene expression database (http://tfgd.ihb.ac.cn/; gene ID# TTHERM_00499660).

Figure 21

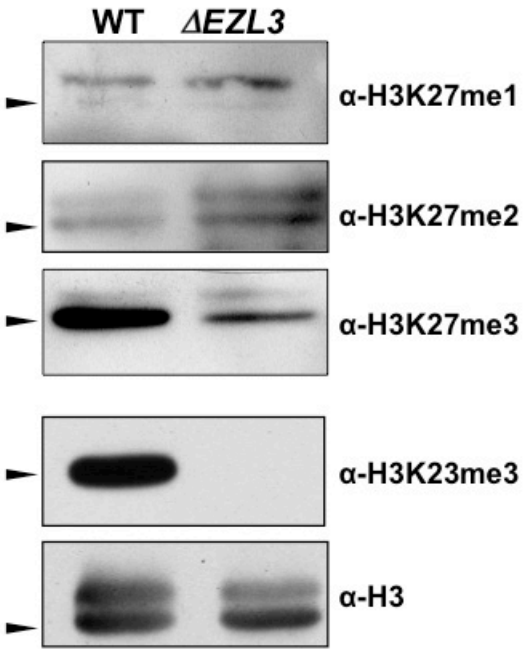


Figure 21. Loss of H3K23me3 leads to loss of H3K27me3 but not H3K27me1 or H3K27me2 levels. Western blot analysis of micronuclear acid extracts from wildtype and *ΔEZL3* strains growing vegetatively. Black arrow heads indicate the location of $H3^F$.

Figure 22

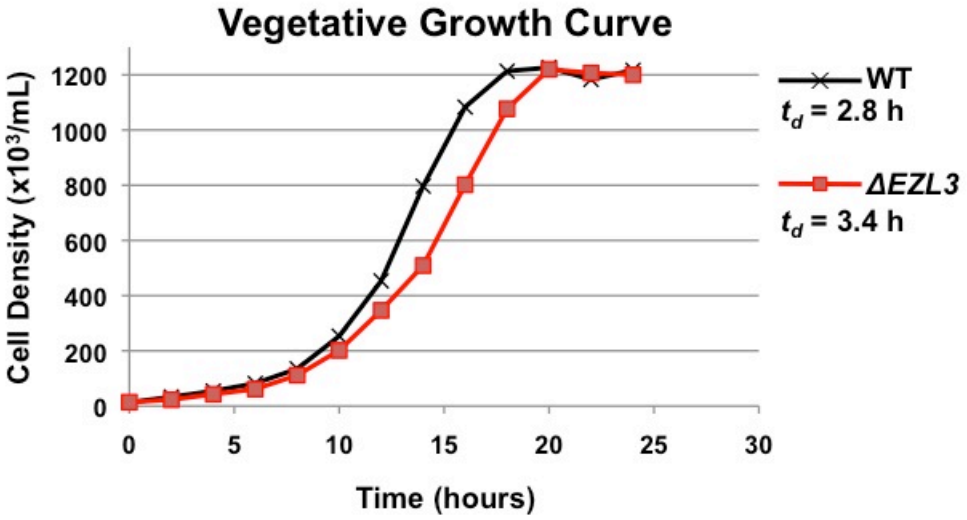


Figure 22. Vegetatively growing *ΔEZL3* cells have a slow growth phenotype. Growth curve of wildtype and *ΔEZL3* cells grown under vegetative conditions in SPP media at 30 ◦C. Doubling time (t_d) was calculated by measuring the growth rate of each strain at mid-log phase. Experiments were performed in triplicate and similar doubling times were obtained in each case.

3.3B DEVELOPMENTAL LAGS IN TETRAHYMENA LACKING H3K23ME3

The vegetative growth rate of cells missing Ezl3p and H3K23me3 were about 20% slower than wildtype **(Figure 22)**. However, since H3K23me3 was particularly enriched during early meiosis, we induced conjugation in opposite mating pairs of *ΔEZL3* cells and analyzed them for phenotypes. Whole-cell-extracts of conjugating wildtype or *ΔEZL3* cells were collected at indicated time points and analyzed by Western blot. H3S10ph (mitotic marker) and Pdd1p levels were used as proxies for progression of early and late conjugation, respectively **(Figure 23A)**. A significant lag was detected in *ΔEZL3* during early conjugation, exemplified by the delayed onset of H3S10 phosphorylation. Moreover, the lag time was amplified at later time points of conjugation, during nuclear differentiation, highlighted by the α-Pdd1p immunoblot. Cytologic comparison of wildtype and *ΔEZL3* mating pairs showed that the initial lag during conjugation began at the leptotene/zygotene stages of meiotic prophase **(compiled in Figure 23B)**. *ΔEZL3* cells also showed delayed formation of piRNA-like scanRNA **(Figure 23C)**, a class of non-coding RNAs intricately involved with heterochromatin formation and DNA elimination at the nuclear differentiation stage of conjugation. Despite a pronounced lag in the completion of meiosis, nuclear differentiation, and scanRNA formation, *ΔEZL3* cells apparently recover and proceed to terminal stages of conjugation **(Figure 23B, 48h)**.

Figure 23

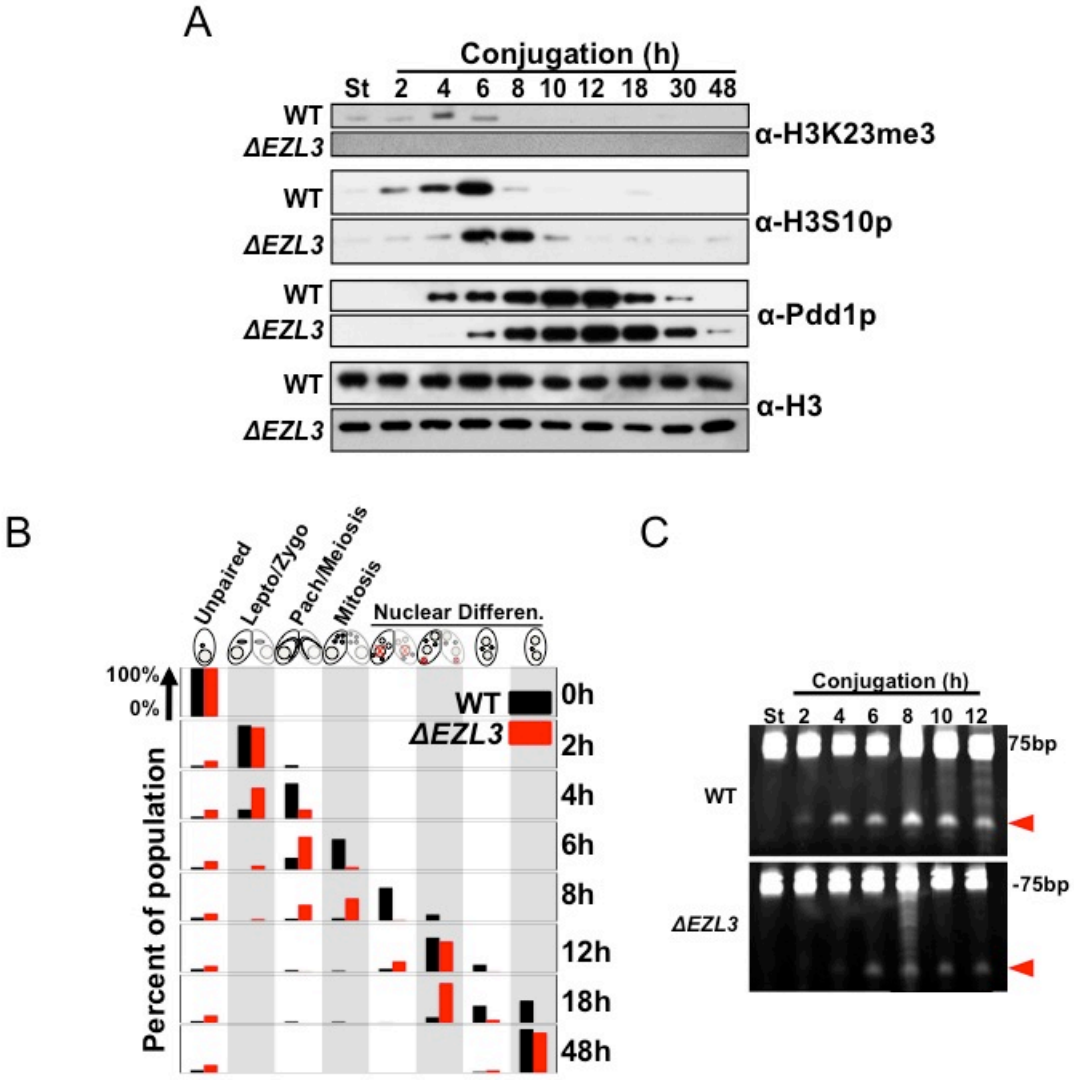


Figure 23. Meiotic progression is disrupted in *Tetrahymena* lacking Ezl3p and H3K23 methylation. **(A)** Western blot analysis of whole-cell-extracts from wildtype and *ΔEZL3* strains progressing through conjugation. H3S10ph and Pdd1p are markers of early and late conjugation, respectively. **(B)** Developmental progression of conjugating wildtype and *ΔEZL3* cells. Cells were fixed at the indicated time points. The population of cells at each time point were examined and tallied with respect to cytological hallmarks associated with developmental stages. **(C)** Total RNA was extracted from wildtype and *ΔEZL3* cells at the indicated times during conjugation, resolved on a denaturing acrylamide gel, and stained with ethidium bromide. Red arrowheads point to the location of piRNA-like scanRNA.

Figure 24

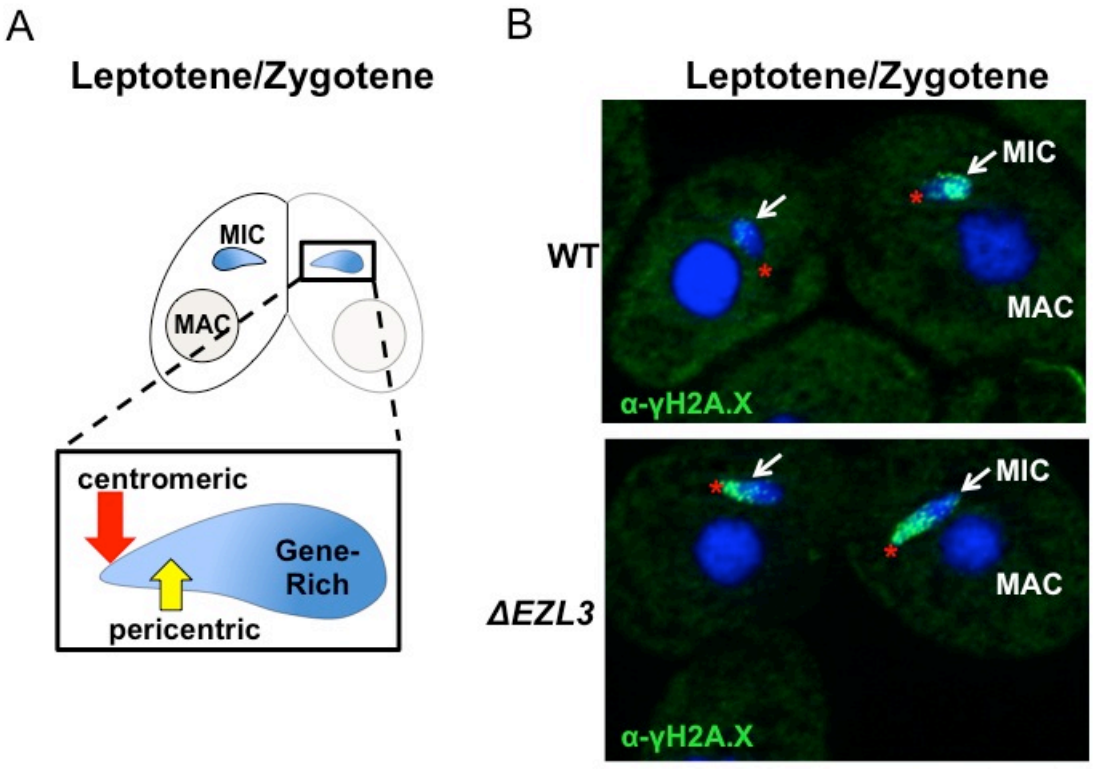


Figure 24. Localization of DSBs in meiotic micronuclei. (A) The structure and chromatin of early meiotic micronuclei are highly polarized. Centromere-proximal chromatin is gene- and DAPI-poor (yellow arrow). This region is flanked by a DAPI- and gene-rich region. **(B)** Zoomed out images of wildtype and *ΔEZL3* pairs undergoing early meiosis and stained with α-γH2A.X. Red asterisks point to the centromeric pole of meiotic micronuclei. Note that there is no DSBs occurring in macronuclei.

3.3C LOSS OF EZL3P AND H3K23ME3 LEADS ECTOPIC DNA DOUBLE-STRAND BREAKS

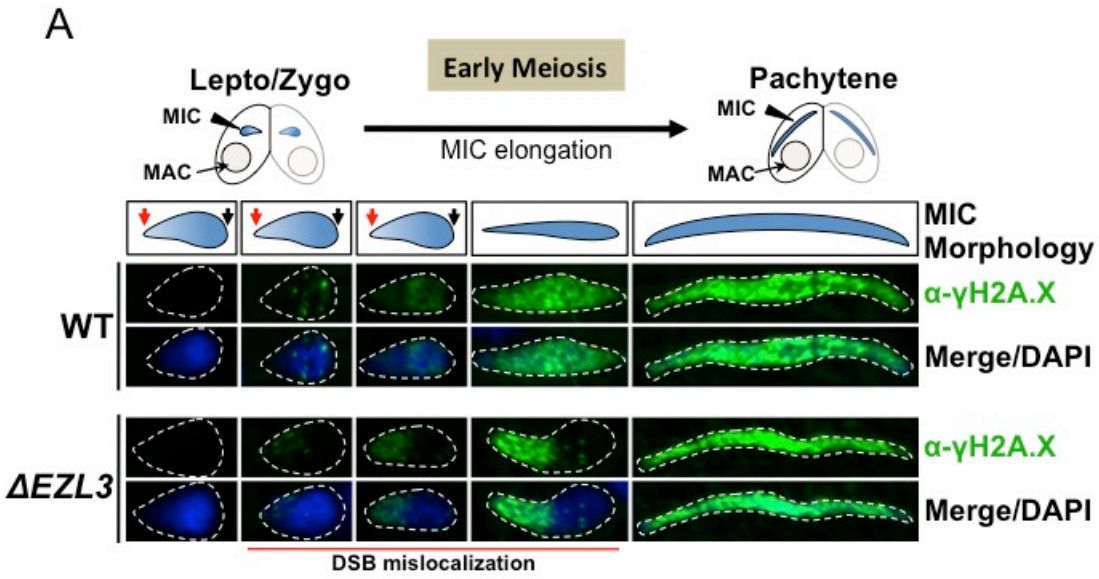
Given the significant early meiotic lag (~1.5 h) observed in conjugating *ΔEZL3* *Tetrahymena* strains, we next asked whether meiotic DNA damage or repair pathways were affected in these mutants. It has previously been reported that in the early stages of meiosis, micronuclear chromosomes in *Tetrahymena* become highly organized - centromeres and telomeres segregate to opposite poles of the elongating micronuclei, while genic (DAPI-rich) and intergenic (DAPI-poor) regions are suggested to be more centrally localized.[160, 161] **(Figure 24A)**. Therefore, we used α-γH2A.X to investigate the spatial dynamics of DSBs in meiotic *ΔEZL3 Tetrahymena* cells **(Figure 24B)**.

DSBs in wildtype micronuclei initiated early in meiosis (leptotene/zygotene) within the DAPI-rich genic region, becoming distributed throughout the micronucleus by pachytene **(Figures 25A)**. In striking contrast, DSBs in meiotic *ΔEZL3* cells initiated in the DAPI-poor, centromeric pole of micronuclei and remained localized to that half of the micronucleus for the majority of the lag, although γH2A.X localization resembled wildtype distribution when cells entered pachytene **(Figure 25A)**. Similar positioning of centromeric H3 (α-Cna1p) in both wildtype and *ΔEZL3* cells showed that DSBs in *ΔEZL3* cells directly flanked centromeres, rather than the shift of DNA damage being a result of massive genome reorganization **(Figure 25B)**.[116] Interestingly, DSBs were not mislocalized during leptotene/zygotene in *ΔEZL2* cells (lacking H3K27me3) **(Figure 26)**, suggesting that only the *ΔEZL3* cell line completely lacking H3K23me3 has mislocalized meiotic DNA damage.

3.3D LOSS OF EZL3P AND H3K23ME3 LEADS TO REDUCED PROGENY VIABILITY

Misregulation of DSB pathways within the germline reduces genomic stability, the production of viable gametes, and progeny viability.[162] Therefore, we tested whether loss of H3K23me3 and ectopic DSB formation translated to reduced progeny viability. Although we previously showed that *ΔEZL3* cells reach terminal cytological stages of conjugation **(Figure 23B, 48h)**, the progeny viability was substantially impaired **(Figures 27A-C)**. Furthermore, fertility was negatively affected when surviving *ΔEZL3* progeny were backcrossed to wildtype cells (data not shown). The decrease in progeny viability for the *ΔEZL3* cells missing H3K23me3 is similar in scope to the reported infertility and increased spermatocyte apoptosis in male mice lacking both copies of the H3K9me3 methyltransferases, Suv39h1/2.[63]

Figure 25



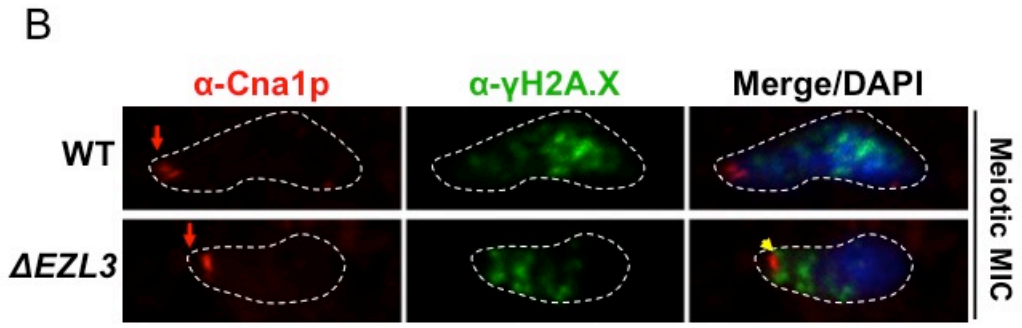


Figure 25. Loss of K3K23me3 during meiosis leads to increased DSBs in centromere proximal chromatin. **(A)** Cartoon depiction of the morphological changes to micronuclei during meiotic prophase. At leptotene/ zygotene, chromatin is spatially organized within the micronucleus with a DAPI-rich pole black arrow), a DAPI-poor pericentric region, and a centromeric pole (red arrow). By pachytene stage, micronuclei are elongated and their structural polarization diminishes. Below the illustrations are indirect immunofluorescence images of DSBs (marked by α-γH2A.X) in meiotic micronuclei of wildtype and *ΔEZL3* cells. Representative images are shown from at least 3 independent experiments. **(B)** Zoomed images of leptotene/ zygotene micronuclei from wildtype and *ΔEZL3* cells co-stained with α-Cna1p (centromeric H3) and α-γH2A.X.

Figure 26

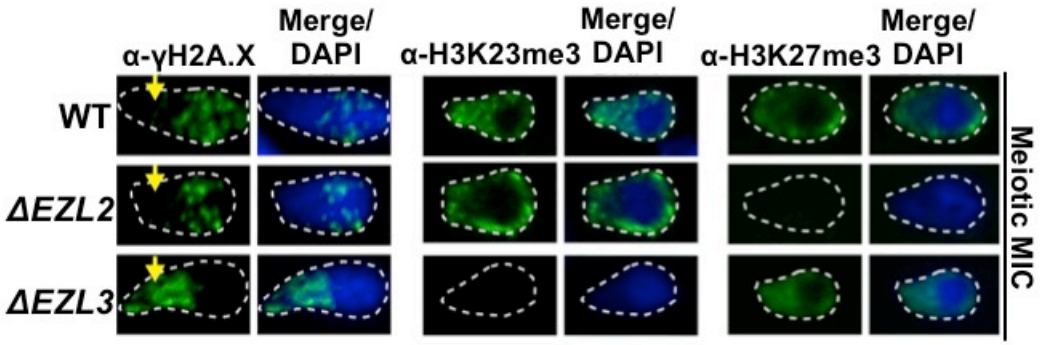
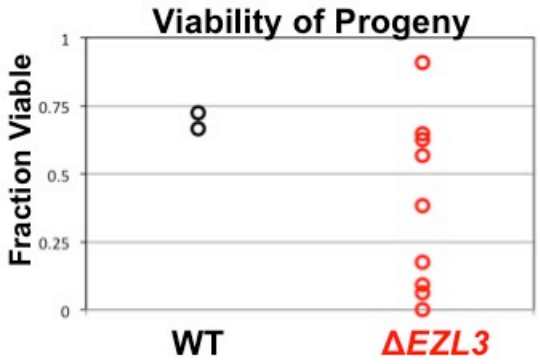


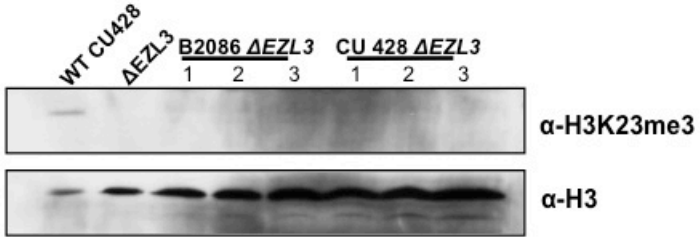
Figure 26. **Loss of K3K27me3 does not lead to ectopic DSBs in centromere proximal heterochromatin**. Indirect immunofluorescence analysis of leptotene/zygotene micronuclei stained with α-γH2A.X, α-H3K23me3 and α-H3K27me3. Yellow arrows point to the pericentric region of meiotic micronuclei.

Figure 27

A



B



C

Proportion of viable progeny

		CU428 Background				
		WT #1	WT #2	ΔEZL3 #1	ΔEZL3 #2	ΔEZL3 #3
B2086 Background	WT #1	21/29				
	WT #2		24/36			
	ΔEZL3 #1			17/30	15/24	3/17
	ΔEZL3 #2			22/34	30/33	2/22
	ΔEZL3 #3			1/16	5/13	0/19

Figure 27. Loss of H3K23me3 is detrimental to progeny viability. (A) Percentage of wildtype and *ΔEZL3* cells producing viable progeny. **(B-C)** Western blot analysis of H3K23me3 in somatic knockouts of *EZL3*. Three different *ΔEZL3* mutants were generated in 2 strains (B2086 vs. CU428) that are compatible for mating. Then, each *ΔEZL3* mutant was allowed to conjugate with the 3 other *ΔEZL3* mutants from the opposite mating strain. The proportions of viable progeny resulting from 2 wildtype and 9 *ΔEZL3* mating experiments are shown here and graphed part A.

3.4 DISSCUSION

3.4A A MODEL FOR THE ROLE H3K23ME3 IN MEIOTIC HETEROCHROMATIN

Our findings suggest that H3K23me3 is a heterochromatic mark that localizes to centromere-proximal and DNA poor regions of the micronuclear genome. The striking absence of H3K23me3 in macronuclei may be related to the loss of centromeres and intergenic/micronuclear-limited sequences in macronuclear genome that occurs during conjugation. We showed that levels of H3K23me3 are dramatically upregulated in micronuclei during meiotic prophase, at a developmental stage concurrent with the increase of DSB formation and γH2A.X levels. As noted above, meiotic micronuclei lack the H3K4me3 and H3K9me3 marks known to regulate meiotic recombination/DSB pathways in higher eukaryotes. Our findings that H3K23me3 is upregulated during early conjugation and is required for proper DSB localization suggest the existence of novel pathway, dominated by H3K23me3, that contributes to the regulation of meiotic DSB formation. We propose that H3K23me3 mediates the formation of a higher-ordered chromatin structure to protect pericentric heterochromatin from programmed DSB formation in *Tetrahymena* **(Figure 28)**. Supporting this conclusion, the loss of H3K27me3, which occurs in *cis* with H3K23me3 and is the only other heterochromatin-associated mark in micronuclei, did not lead to ectopic DSB damage **(Figure 26)**.

Figure 28

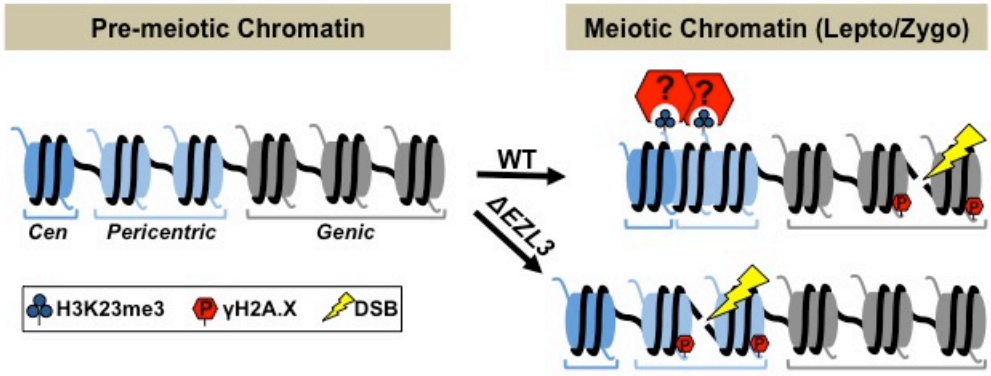


Figure 28. **A role for H3K23me3 in meiotic heterochromatin**. A model proposing that H3K23me3 mediates the formation of higher-ordered heterochromatin to limit DSBs at pericentric chromatin during early meiosis.

3.4B THE GENIC TO PERICENTRIC SHIFT IN DSBS OF EZL3P MUTANTS

Our data and model highlight a peculiar shift in the formation of DSBs from gene-rich regions to pericentric DNA in leptotene/zygotene cells lacking H3K23me3. We submit that this shift is not a consequence of a change in the substrate specificity of Spo11, which is a promiscuous DNA endonuclease, but rather is a reflection of mechanisms of homolog pairing and spatial/temporal regulation of Spo11. Although the pairing pathways have not been fully elucidate in eukaryotes, there is general consensus that most species, including *Tetrahymena*, initiate pairing through the formation of a 'bouquet conformation'.[160, 163] This meiosis-specific organization of chromosomes occurs as a result of early interactions of homologous centromeres and telomeres. Pairing and DSB formation in the genic regions of chromosome occur subsequent to these early interactions, as synapses iteratively form along the chromosome arms in a zipper-like fashion.[164] Importantly, although the major role of Spo11 is to catalyze DSBs in euchromatic regions of chromatin, it is present and required for the initial pairing of telomeric and centromeric DNA.[165] Therefore, heterochromatin formation functions to protect the underlying gene-poor repetitive DNA from Spo11 activity at early stages of homolog pairing. Given that H3K9me3 is absent from micronuclei, and that the loss of H3K27me3 does not alter DSB sites in *Tetrahymena*, we argue that H3K23me3 underlies early meiotic heterochromatin formation, which functions to inhibit DSBs in gene-poor DNA.

3.4C LOSS OF H3K23ME3 EFFECTS SMALL RNA-GUIDED HETEROCHROMATIN FORMATION

The bulk of the work in this Chapter focused on the role of H3K23me3 in meiosis, which was motivated by our detection of a ~1.5h lag in early conjugating ΔEZL3 cells. We also detected a second lag during late conjugation that started with the development of new macronuclei and persisted through hour 48 of conjugation. Compared to wildtype mating rates, this second lag lasted about 6 hours early in the nuclear differentiation stages and extended to at least 18 h by the end of conjugation. Cytologically, the lag appeared to occur before the formation of heterochromatic bodies in anlagen that directly precede DNA elimination **(Figure 23B)**. Despite the severe lag in the formation of heterochromatic bodies, the levels of Pdd1p, H3K27me3 and H3K9me3 were normal in ΔEZL3 cells. Therefore, our results suggest that the second lag in ΔEZL3 mating cells occurs during heterochromatin formation in new, developing macronuclei.

What is the nature of this second lag? Although there is no detectable H3K23me3 in new macronuclei at any stage of their development, loss of this mark could affect scanRNA activity, which are transcribed in meiotic micronuclei and are critically important for macronuclear differentiation. Since the formation of heterochromatic bodies are dependent on the proper deposition of H3K9me3 and H3K27me3, as well as the engagement of their reader proteins (Pdd3p and Pdd1p, respectively), it is tempting to speculate that H3K23me3 contributes to new macronuclei development indirectly by affecting scanRNA generation or by regulating Pdd1p, which localizes to micronuclei early in conjugation. A thorough comparison of scanRNA from wildtype and ΔEZL3 cells would provide the necessary evidence to support or to oppose the above hypothesis. Moreover, whether ΔEZL3 cells can successfully carry out DNA elimination has not been addressed in this work.

A major caveat to the above interpretations is that Ezl3p, although required for H3K23 methylation in micronuclei, could have other non-histone targets that affect anlagen dynamics later in conjugation. Several studies have recently emerged highlighting the ability of so-called *histone* methyltransferases to modify non-histone proteins. Of particular importance to this work, Ezh2 was recently shown to target methylation of several non-histone proteins in the cytoplasm of mammalian cells.[166] Therefore, a comprehensive understanding of Ezl3p activity and localization is required to effectively parse out its contribution to the phenotypes described above. Notably, our efforts to make and antibody against Ezl3p yielded some success for Western blotting **(Figure 29)**, but not immunofluorescence.

Figure 29

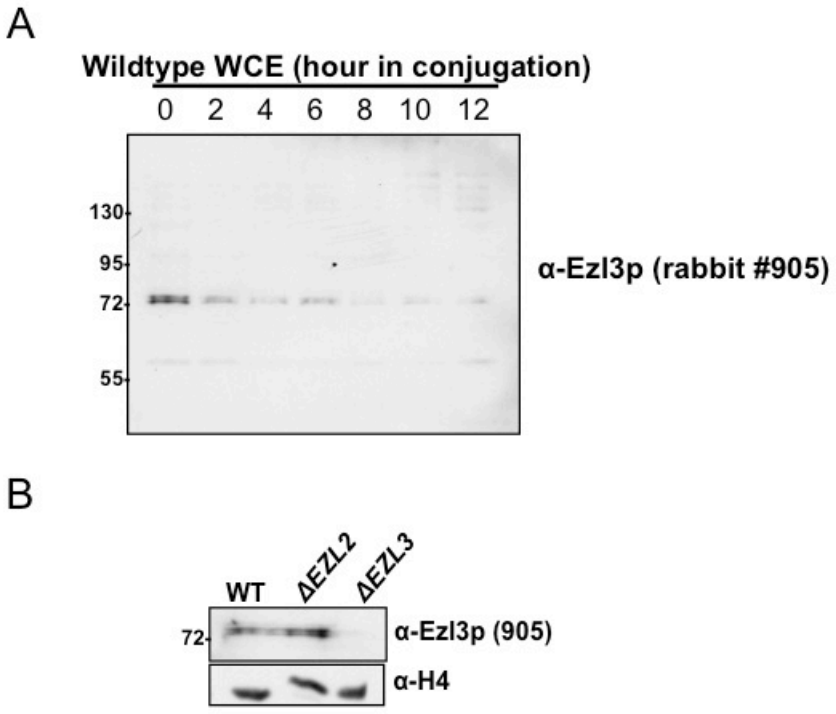


Figure 29. **Development of α-Ezl3p**. **(A-B)** Specificity of α-Ezl3p in Western blot assays. Antibody generated in rabbits injected with MBP-SET3 (SET domain of Ezl3p). Serum was not specific, however upon ammonium sulfate precipitation of IgG, an enriched signal at around 80 kDa was observed. Although this antibody appeared to be specific, it was highly unstable and degraded soon after the ammonium sulfate precipitation step. Further characterization was therefore not possible. The expected size of Ezl3p is 80 kDa.

CHAPTER 4: EVOLUTIONARY CONSERVATION OF HISTONE H3 LYSINE 23 TRIMETHYLATION

4.1 INTRODUCTION

Our observations using conjugating *Tetrahymena* suggested a potentially critical role for H3K23me3 during meiosis. To understand the biological significance of this modification (and since the H3 tail is highly conserved) in higher eukaryotic homologs of H3 **(Figure 30A, red underline)**, we sought to determine if H3K23me3 could be detected in mice and *C. elegans* during meiosis. We first turned to *C. elegans,* which is an excellent model organism to study germline chromatin. Like *Tetrahymena*, the nematode *C. elegans* is a genetically tractable model organism whose basic physiology and developmental stages are well characterized. Studies on the gonads of *C. elegans* have provided important insights into germline development and germ cell differentiation, as well as processes regulating mitotic and meiotic checkpoints.[156] In adult *C. elegans*, germ cells are essentially organized into two physically and functionally distinct regions Cells near the distal tip are maintained as undifferentiated mitotic precursors through the signaling of the Notch-like receptor GLP-1. As the cells migrate proximally the GLP-1 signal weakens triggering the differentiation of mitotic cells into meiotic cells, which occurs within a well-defined transition zone **(Figure 30B)**.[156, 167]

Figure 30

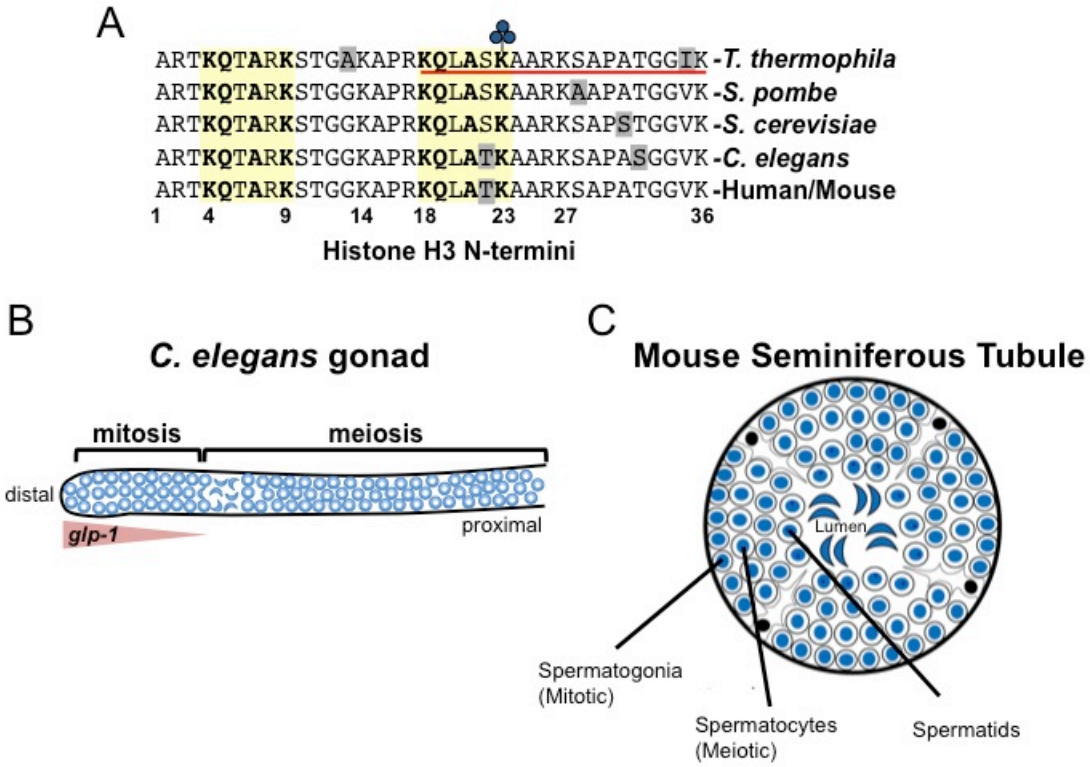


Figure 30. **Evolutionary conservation of H3 tails**. **(A)** Alignment of the first 36 residues of histone H3 from the indicated organisms. An H3 peptide spanning amino acids 18-36 (red underline) and trimethylated on K23 (blue clover) was synthesized and used as the immunogen for antibody generation in rabbits. Yellow boxes highlight the conserved KQXAXK motif associated with H3K9 and H3K23. **(B)** Schematic of *C. elegans* adult hermaphrodite gonad. Transition to meiosis is marked by crescent-shaped chromatin of leptotene/zygotene nuclei. **(C)** Schematic of mouse seminiferous tubule showing differentiation of pre-meiotic spermatogonia (in the periphery) to spermatids (near the lumen).

Antibody-based studies of *C. elegans* gonads have provided a wealth of data that speak to the localization of histone PTMs in germline chromatin. Notably, H3K27me3 and H3Kme3 are uniformly present along the distal-proximal germline axis. Curiously, H3K9me2 appears to specifically localized to the inactive X chromosome of females. Studies have also shown that loss of the only E(z) homolog of the worm, MES-2, completely abolishes H3K27me3 in the gonad. In line with other organisms, γH2A.X is detected at the leptotene/zygotene transition zone. To our knowledge, no other histone PTM is specifically upregulated at the transition zone.[168-170]

To further examine evolutionary connections between H3K23me3 upregulation and meiosis, we analyzed germ cells of mouse testes. Similar to the *C. elegans* gonad, mouse seminiferous tubules are structurally organized into distinct functional regions - peripherally located mitotic spermatogonia differentiate into meiotic spermatocytes near the lumen **(Figure 30C)**.[168] Post-meiotic haploid cells are called spermatids, which further differentiate into spermatozoa and get deposited into the lumen. This differentiation scheme also involves dramatic changes to chromatin structure, exemplified by the formation of a single, DAPI-rich, heterochromatic chromocenter containing centromeric and pericentric DNA in spermatids. Notably, chromocenters are enriched for H3K9me3 and HP1 isoforms.

4.2 METHODS

4.2A CELL CULTURE AND STRAINS

C. elegans strains were derived from *C. elegans* Bristol strain N2 and cultured using standard techniques.[171] Strain DG2389 *glp-1 (bn18lf)* was cultured at the permissive temperature (15 °C) and shifted to 25 °C for 18 hours to interfere with Notch signaling and induce meiotic differentiation of germ cells. Strain GC833 *glp-1(ar202gf)* was grown at 25 °C from L1 larvae to adulthood to continuously expose germ cells to high levels of Notch signaling. To analyze loss of the *C. elegans* Enhancer of zeste homolog, we used the *mes-2(bn11)* homozygous mutant. To analyze loss of *mes-4* function, wildtype *C. elegans* were treated with *mes-4(RNAi)* from L1 larvae to adulthood by a standard feeding protocol.[172]

The mouse tissues were obtained from wildtype C57BL/6 surplus male mice that were scheduled for termination under IACUC protocol and provided to our laboratory following euthanasia. This procedure was reviewed and approved by Johns Hopkins Animal Care and Use Committee and was performed in accordance with the NIH Guide for the Care and Use of Laboratory Animals.

4.2B WESTERN BLOT AND INDIRECT IMMUNOFLUORESCENCE ANALYSES

SDS-PAGE and Western blotting were performed as described above (Chapter 2, Methods). Rabbit polyclonal antibodies used in Western blots included: α-H3 (1:75,000; ab1791; Abcam) and α-H3K23me3 (1:100; made in this study).

Worm immunofluorescence staining - adult *C. elegans* hermaphrodites were washed in PBS and germlines were dissected on poly-L-lysine treated slides, covered

with a coverslip to ensure attachment to slide, and snap-frozen on aluminum blocks on dry ice. The samples were fixed for 1 minute in 100% methanol at -20 °C, followed by 5 minutes in 2% EM grade paraformaldehyde in 100 mM K_2HPO_4 pH 7.2 at room temperature. The samples were blocked for at least 30 minutes with PBS containing 0.1% BSA and 0.1% Tween-20 (PBTw) with 10% Normal Goat Serum. Primary antibodies were diluted in PBSTw as follows: rabbit α-H3K23me3 (1:70); mouse α-H3K9me3 (1:100; #39285; Active Motif); mouse α-H3K27me3 (1:100; #39536; Active Motif); mouse α-H3K4me3 (1:250; ab1012; Abcam). Secondary antibodies were goat α-rabbit or α-mouse conjugated with Cy3 or Alexa-488 diluted at 1:200 in PBTw (Jackson Immunoresearch and Invitrogen). Primary antibody incubations were overnight at 4 °C; secondary antibody incubations were for 2 hours at room temperature. For peptide competition experiments, α-H3K23me3 was incubated with 5 μg / mL of indicated peptides for 1 h on ice, and immune complexes were removed by centrifugation for 10 min at 14,000 rpm. Preabsorbed primary antibodies were applied to the tissue samples for 1 h at room temperature followed by secondary antibody incubation for 1 h at room temperature. Before imaging, DNA was counterstained with DAPI.

Mouse immunofluorescence staining: testes of 7 week old male C56BL/6 mice were dissected and detunicated in PBS, fixed in 2% paraformaldehyde in PBS for 3 hours at 4 C, equilibrated into 30% sucrose solution overnight at 4 °C, embedded in OCT, and then frozen. Seven micron-thick sections were cut, adhered to slides, baked at 50 C for 45 minutes, and permeablized in PBS/0.05% Triton-X100. Sections were blocked for 30 minutes in PBS containing 0.05% Triton-X100, 10% normal goat serum and 0.3% BSA.

Further immunostaining procedures were carried out as described above for the worm samples.

Microscopy for worm and mouse slides: confocal images were acquired with a Cascade QuantEM 512SC camera (Photometrics) attached to a Zeiss Axioimager microscope with Yokogawa spinning disk confocal scanner and Slidebook software (Intelligent Imaging Innovations). Maximal-intensity projections encompassing a single layer of nuclei were generated by the Slidebook software. Image processing was performed in Adobe Photoshop CS4.

4.2C HISTONE EXTRACTION FROM C. ELEGANS EMBRYO NUCLEI

Nuclei from embryos were isolated following a previously described method. [173] Once isolated, nuclei were pelleted and resuspended in 5X pellet volume 0.4 N H_2SO_4 and rotated for 2 h at 4 °C. Following centrifugation for 10 min at 10000 g, histones in the supernatant were acid-precipitated by 100% trichloroacetic acid (TCA) at a dilution factor of 1:4 (TCA: supernatant; final 20% TCA) for 1 h on ice. TCA precipitates were pelleted by centrifugation for 10 min at 10,000 g, washed 2X with ice-cold acetone, resuspended in water, and stored at -80 °C.

4.3 RESULTS

4.3A H3K23ME3 IS CONSERVED IN NEMATODES

In collaboration with Dr. Ekaterina Voronina (University of Montana, Missuoula, MT; formerly a postdoctoral fellow in Dr. Geraldine Seydoux's laboratory, Johns Hopkins University, Baltimore, MD), we initially examined nuclear extracts from whole

C. elegans embryos, since their nuclei could be easily purified with respect to adult worms, making possible a direct comparison of H3K23me3 levels on H3 between *C. elegans* and *Tetrahymena*. Comparative Western blot analysis of whole-worm nuclear extracts showed that H3K23me3 was indeed present on *C. elegans* H3, although at a much lower level than that found on *Tetrahymena* micronuclear $H3^F$ **(Figure 31A)**.

Given that H3K23me3 is present in *C. elegans*, we next wanted to determine if this mark is upregulated during meiosis. By indirect immunofluorescence, H3K23me3 was nearly absent in the mitotic region of the gonad, although a small number of mitotic nuclei in this region were positive for the PTM **(Figure 31B)**. In striking contrast, H3K23me3 staining became robust in meiotic cells within the gonad, beginning at the leptotene/zygotene transition zone that marks meiotic entry, and plateauing through the rest of meiotic prophase **(Figure 31B)**. Compared to H3K27me3, H3K9me3, and H3K4me3 distribution in the gonad, H3K23me3 was the only PTM that became enriched as germ cells entered the transition zone **(Figure 32A-C)**. Under high magnification, H3K23me3 and other PTMs also exhibited comparatively different sub-nuclear staining patterns within meiotic cells **(Figure 32A-C)**. Notably, H3K23me3 appeared relatively punctate across the entire nucleus rather than being enriched on either the autosomes (like H3K4me3) or the silenced Xs (like H3K27me3) **(Figure 32A,C).** [174, 175] Importantly, our peptide competitions showed that α-H3K23me3 is highly specific for H3K23me3 in the *C. elegans* gonad **(Figure 33)**.

Figure 31

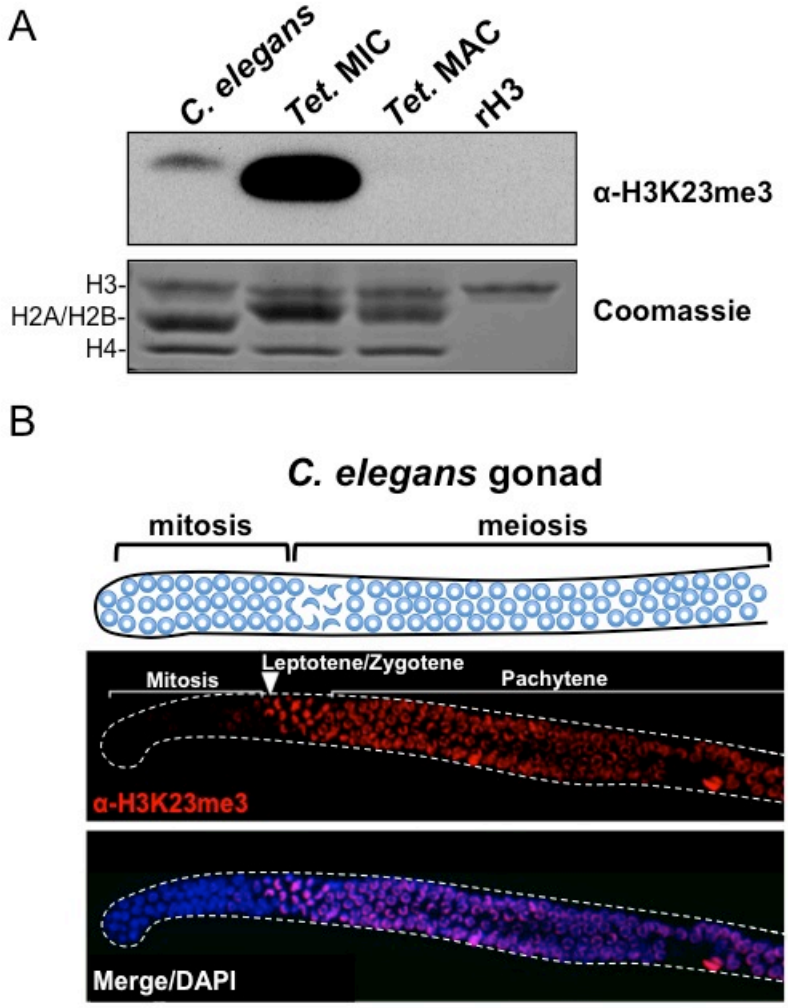


Figure 31. **H3K23me3 is present in the germline of *C. elegans***. **(A)** Acid-extracts of *C. elegans* embryonic nuclei, *Tetrahymena* micro- or macronuclei, and recombinant human histone H3 (rH3) were resolved with SDS-PAGE and blotted against α-H3K23me3. Histone samples were normalized by Coomassie staining. **(B)** Immunofluorescence of formaldehyde fixed *C. elegans* gonad stained with α-H3K23me3 and counterstained with DAPI. Scale bar, 10 μm. White arrowheads mark the mitotic/meiotic transition zone.

Figure 32

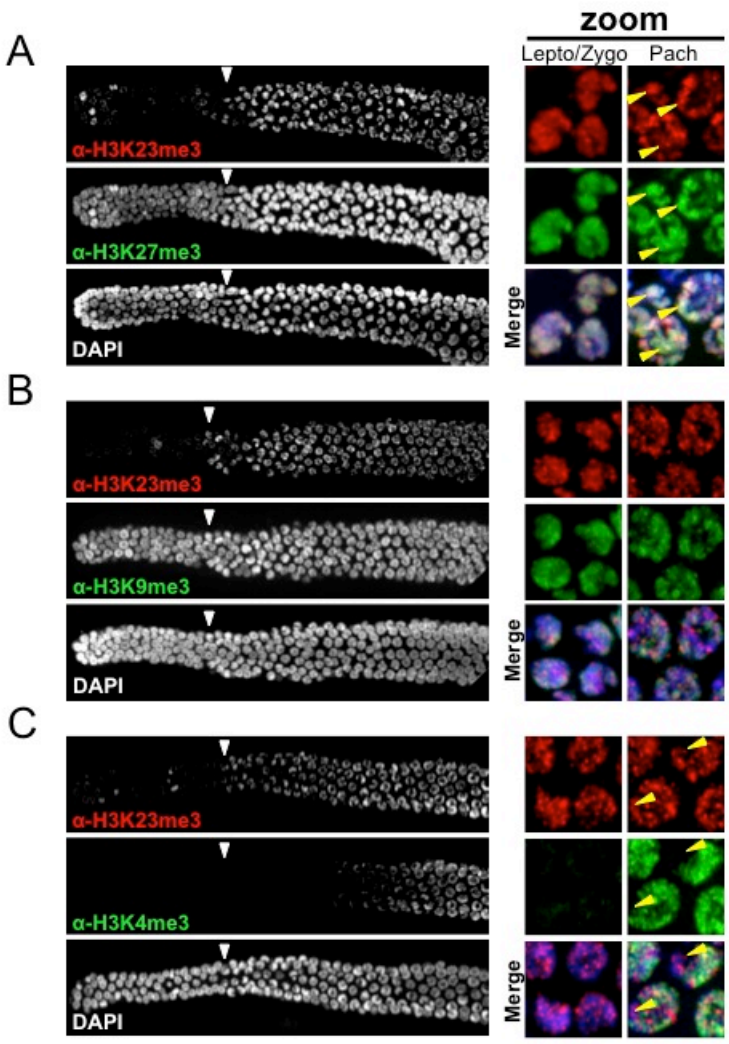


Figure 32. Comparison of heterochromatic histone PTM staining profiles in *C. elegans* gonad. (A-C) Co-immunofluorescence staining between α-H3K23me3 and α-H3K27me3, α-H3K9me3, or α-H3K4me3 in the *C. elegans* gonad. Zoomed images of germline nuclei co-stained with the indicated antibodies. Yellow arrowheads point to the location of the inactive X chromosome.

Figure 33

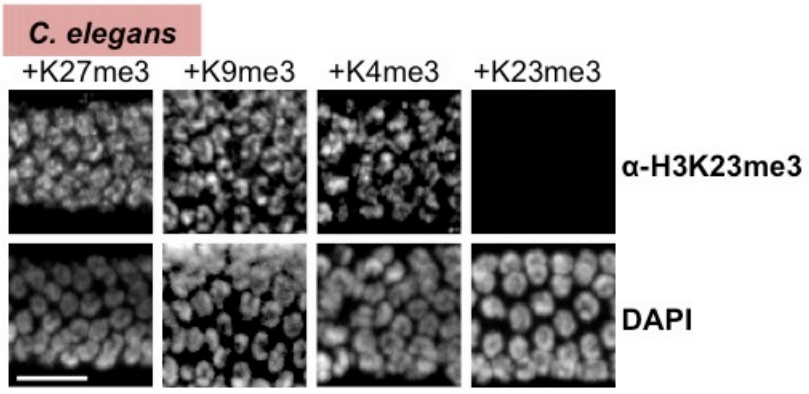


Figure 33. Specificity of α-H3K23me3 in worms. Worm germline immunostained with α-H3K23me3 antibodies preabsorbed with the indicated peptides (on top). Scale bar, 10um.

Figure 34

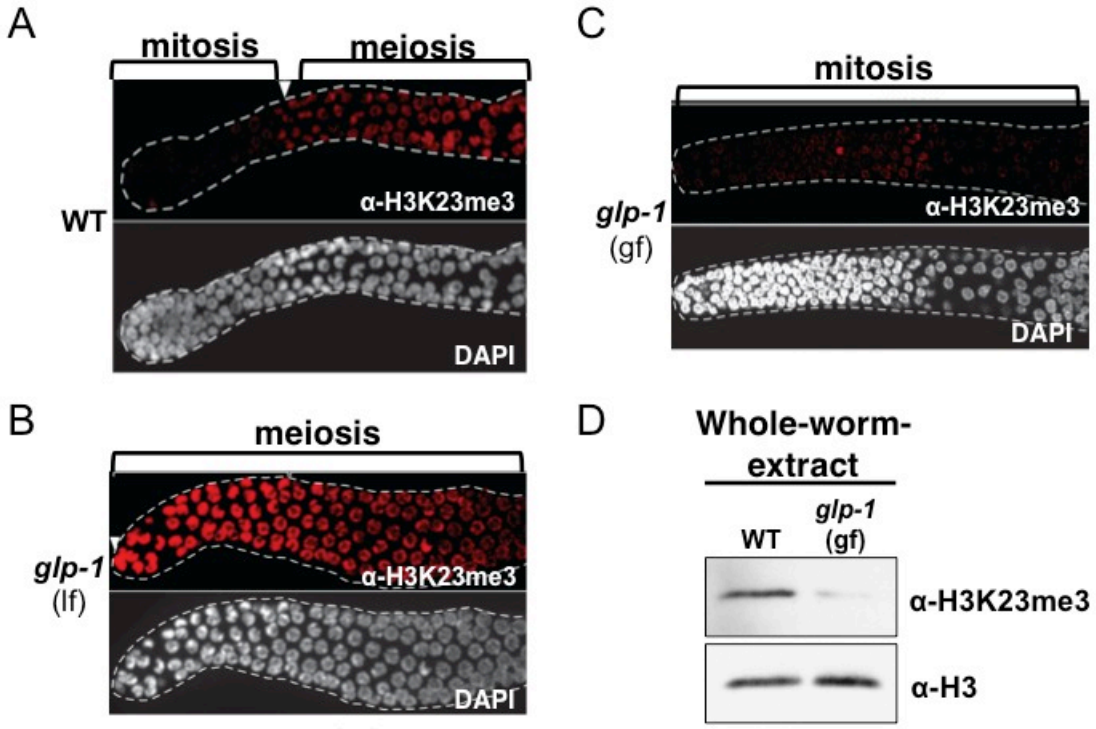


Figure 34. **Localization and levels of H3K23me3 in mutants of Notch signaling**. **(A-C)** H3K23me3 localization in loss-of-function and gain-of-function mutants (*lf* or *gf*, respectively) of the Notch-like receptor GLP-1. **(D)** Whole-worm-extracts as analyzed by Western blotting.

4.3B H3K23ME3 IS STRONGLY LINKED TO MEIOSIS IN NEMATODES

We further examined the correlation between H3K23me3 and meiotic entry in *C. elegans* using mutants of the Notch-like receptor GLP-1, which maintains the proliferative state (mitotic) of germline stem cells.[156, 176] As expected, in the GLP-1 loss-of-function mutant *glp-1(lf)*, which undergoes early differentiation of mitotic cells and early entry into meiosis, we observed a shift of the H3K23me3 signal towards the distal tip of the gonad **(Figure 34A-B).** Remarkably, in the GLP-1 gain-of-function mutant *glp-1(gf)* which prevents entry of mitotic germ cells into meiosis and causes their proliferation, H3K23me3 levels were significantly reduced throughout the gonad **(Figure 34C)**. Moreover, Western blot analysis of *glp-1(gf)* whole-worm-extracts showed that H3K23me3 levels are also dramatically reduced on an organismal level, further indicating the vast majority of H3K23me3 signal in adult *C. elegans* is confined to meiotic germ cells **(Figure 34D)**. By contrast, H3K9me3 and H3K27me3 levels were intact, and perhaps even upregulated, in *glp-1(gf)* whole-worm extracts **(Figure 35)**. Notably, by our immunofluorescence staining we noticed that cells of the gut and head regions were also intensely staining with this antibody (data not shown).

Figure 35

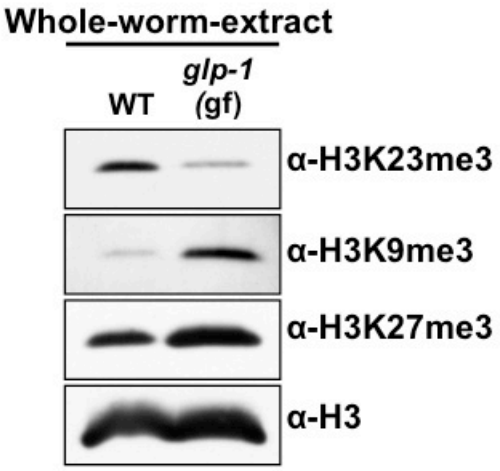


Figure 35. **H3K23me3 is significantly reduced in meiosis deficient worms**. Histone H3 PTM levels in *glp-1* (gain of function) whole-cell-extracts.

4.3C MES-2 AND MES-4 ARE NOT THE WRITERS FOR H3K23ME3 IN NEMATODES

Since the SET domain of *Tetrahymena* Ezl3p is relatively similar to Enhancer of zeste, we examined whether the only Enhancer of zeste homolog present in *C. elegans*, the H3K27 methyltransferase MES-2, could also methylate H3K23. As expected, in the gonad of *mes-2* knockout strains, H3K27me3 was completely undetectable **(Figure 36A)**. In contrast, H3K23me3 remained present in meiotic cells, and surprisingly, H3K23me3 levels within the mitotic germ cell region increased **(Figure 36A)**. Additionally, we investigated H3K23me3 levels following RNAi-mediated knockdown of MES-4, an NSD homolog known to antagonize MES-2 in the germline.[177] Similar to *mes*-2 knockouts, H3K23me3 remained detectable in mitotic and meiotic regions of the *mes-4(RNAi)* germline **(Figure 36B)**. MES-2 and MES-4 regulate gene expression in the germline, and their disruption may have pleotropic effects that include alteration of pathways regulating H3K23me3.[177] Nevertheless, our data suggest neither MES-2 nor MES-4 are exclusively responsible for H3K23me3 in the *C. elegans* gonad. Similarly, H3K23me3 levels were intact in strains deficient in SET-2, SET-12, MET-1, MET-2, and MET-10 methyltransferases, which were tested because of the similarity of their SET domain to that of *Tetrahymena* Ezl3p **(Figure 37)**.

Figure 36

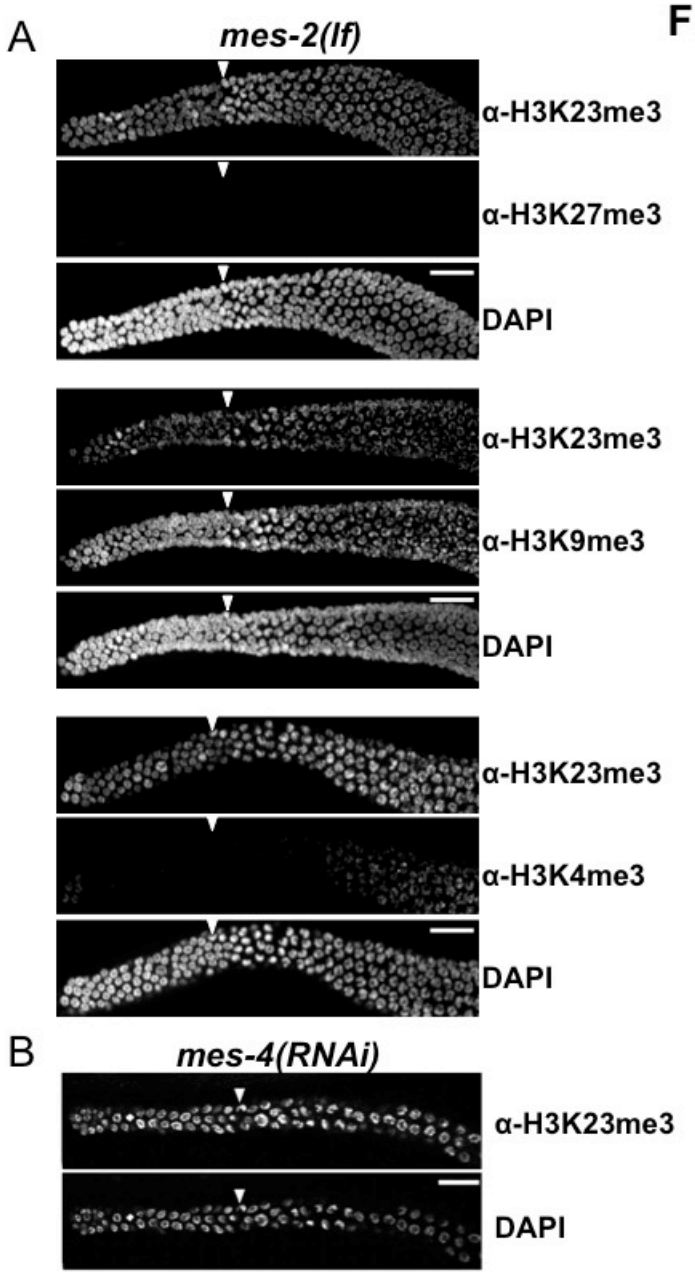


Figure 36. **MES homologs are not the writers for H3K23me3 in worms**. **(A)** Immunofluorescence analysis of H3K23me3 levels in the germline of *C. elegans* lacking MES-2, the only worm homolog of Enhancer of zeste. Scale bar, 10 μm. **(B)** MES-4 is not the writer for H3K23me3 in the worm. Immunofluorescence analysis of H3K23me3 levels in the germline of *C. elegans* lacking MES-4. Scale bar, 10 μm.

4.3D H3K23ME3 IN CONSERVED IN GERM CELLS OF MICE

Much like *C. elegans*, cell division programs of mouse germ cells are coupled to differentiation and are highly compartmentalized. Indirect immunofluorescence staining of mouse testes showed that α-H3K23me3 was initially detected in spermatocytes that were also positive for SCP3, a recombination protein expressed during meiotic entry/ leptotene **(Figure 38)**. Peptide preincubations revealed that α-H3K23me3 was specific for H3K23me3 in mouse germ cells **(Figure 39)**. Notably, as germ cells differentiated into spermatids, the H3K23me3 signal became increasingly confined to a tight focal subdomain within highly condensed chromocenters, which contain pericentric heterochromatin **(Figure 38, yellow arrowheads)**.[169] Together with our previous data, these results suggest the upregulation of H3K23me3 during meiosis is conserved among *Tetrahymena, C. elegans*, and mice, and the enrichment of H3K23me3 within pericentric heterochromatin is conserved between mice and *Tetrahymena*.

Figure 37

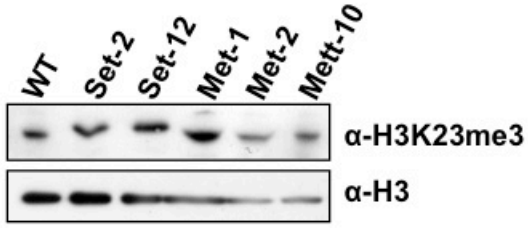


Figure 37. **The search for H3K23me3 writers in the worm**. Western blot analysis of H3K3me3 levels of worms lacking indicated methyltransferases. Knock downs of these methyltransferases were achieved by RNAi.

Figure 38

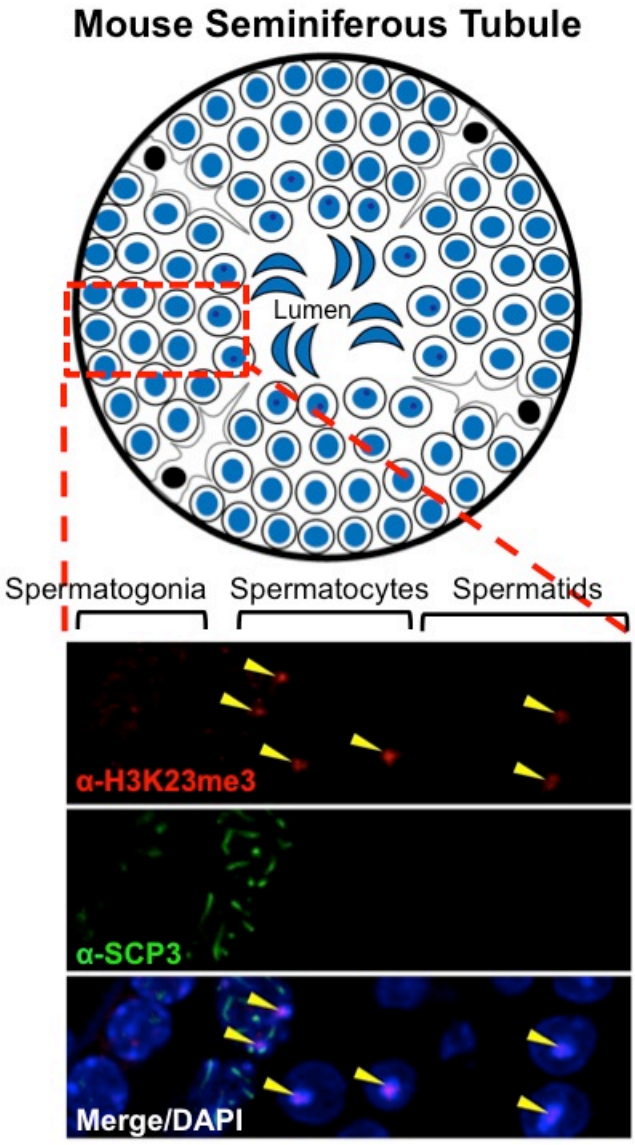


Figure 38. **H3K23me3 in the mouse germline**. Co-staining of the mouse seminiferous tubule with α-H3K23me3 and α-SCP3 (meiotic marker) shows enrichment of H3K23me3 in meiotic spermatocytes. Yellow arrowheads point to H3K23me3-enriched regions within heterochromatic chromocenters.

Figure 39

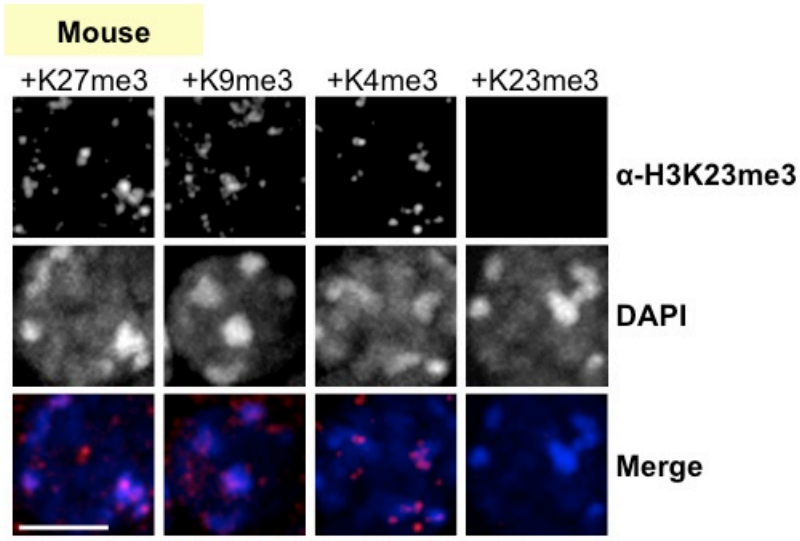


Figure 39. **Specificity of α-H3K23me3 in mouse germ cells**. Mouse germline immunostained with α-H3K23me3 antibodies preabsorbed with the indicated peptides (on top). Scale bar, 5 um.

4.3E H3K23ME3 IS PRESENT IN THE DEVELOPING BRAIN OF MICE

While performing the immunofluorescence assays using α-H3K23me3 in *C. elegans*, we noticed that some somatic cells of the head and gut regions were also strongly stained. The positive stain at the head region suggested that H3K23me3 could be present in neurons. Prompted by these observations we proceeded to test for the presence of H3K23me3 in adult and embryonic mouse brain tissue. As shown in **Figure 40A**, we detected H3K23me3 in the embryonic brain tissues, but not in adult brain tissues by Western blot analysis (in collaboration with Sarven Sabuncyan, JHMI, Baltimore, MD). Surprisingly, the staining profile also indicated that H3K23me3 levels dramatically increase during embryogenesis (from E18.5 to E20.5) and sharply decline during adulthood. By comparison, H3K27me3 levels were relatively stable at the embryonic stages. Control tissue samples of embryonic and adult tissue also highlighted the fact that H3K23me3 is highly enriched in brain tissues. Taken together, these results showed that, in addition to the mouse germline, H3K23me3 is also present and upregulated in the mouse brain during embryogenesis, when neurogenesis at the cortical plate is peaking.

Figure 40

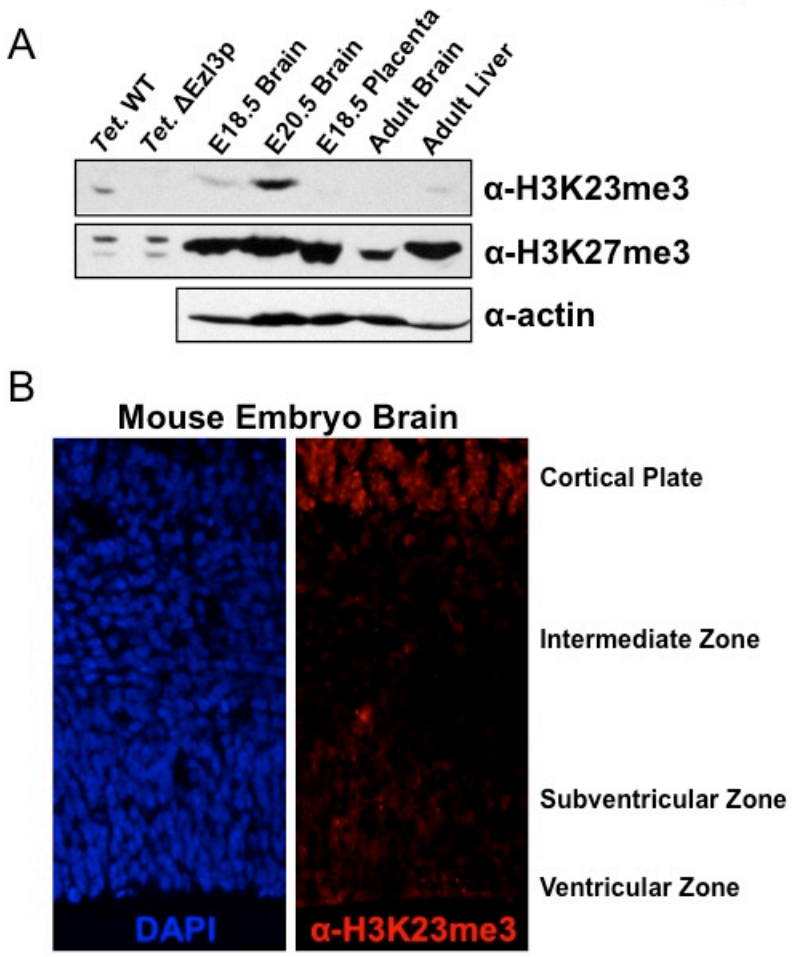


Figure 40. H3K23me3 is present in mouse embryonic brain tissue. (A) Western blot analysis of whole-tissue extracts. **(B)** Immunofluorescence staining of the cerebral cortex with α-H3K23me3 during embyonic development (E14.5).

During mouse embryogenesis, the cortical plate begins to develop at the E12.5 gestational period. At this stage, progenitor cells located near the ventricular zone differentiate as they migrate towards the cortical zone to give rise to the neurons that comprise the new and thickening cortical plate.[178] This process continues up to the point of birth.[178] The increase in H3K23me3 we detected by Western blotting followed the same dynamics as neurogenesis in the cortical plate, suggesting a possible overlap in function. To determine whether H3K23me3 is involved in the formation of the cortical plate, we performed immunofluorescence analysis of brain sections of embryonic mice at E14.5 (in collaboration with Drs. Arnab Kundu and Andre Levchenko, JHMI, Baltimore, MD). Remarkably, we found that H3K23me3 was highly enriched in cortical neurons, and only few neurons near the ventricular zone were positive for this mark **(Figure 40B)**. These results indicate that H3K23me3 levels are upregulated during neuronal differentiation.

Next, to interrogate the dynamics of H3K23 trimethylation during neurogenesis, we employed an *in vitro* model of differentiation using human embryonic stem cells (in collaboration with Valentin Sluch and Dr. Don Zack, JHMI, Baltimore, MD). Using appropriate chemical signals, hESCs can be made to differentiate into neurons. We viewed this as an exciting opportunity to not only ask whether H3K23me3 is present in human neuronal samples, but also to develop an *in vitro* cell line that contains H3K23me3 and would inform and facilitate future biochemical experiments. To date, no other somatic cell culture lines that we tested contained H3K23me3, including Cos7, HEK293, Panc-1, CaCo-2, primary lymphoma, and HeLa cells (data not shown). Thus, we induced neurogenic differentiation in hESCs and collected whole-cell-extracts to

analyze by Western blotting. Our results show that H3K23me3 levels, virtually undetectable in hESCs, were robustly upregulated during neurogenesis **(Figure 41)**. This increase coincided with the appearance of clipped H3 species previously shown to occur during differentiation of mouse embryonic stems cells (ESCs).[179] Given that the major clipping site (alanine 21) of histone H3 in mouse ESCs is proximal to H3K23, these findings suggest a potential overlap in the functions of clipping and H3K23me3.

Figure 41

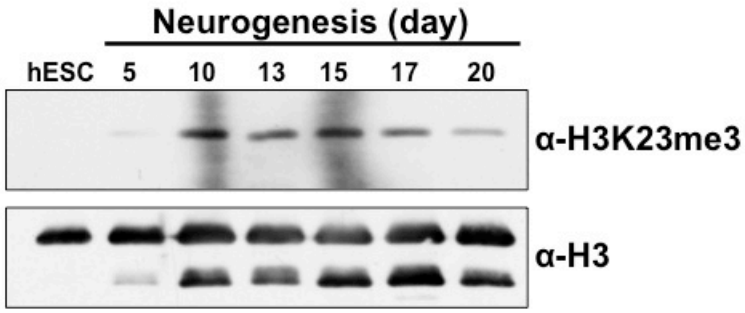


Figure 41. H3K23me3 is upregulated during neurogenesis. Human embryonic stem cells (hESC) were made to differentiate into neurons (retinal ganglion cells). During the differentiation scheme, whole-cell-lysates were collected at the indicated time points and analyzed by western blotting.

4.4 DISCUSSION

4.4A H3K23ME3 IS A CONSERVED PTM ENRICHED IN GERM CELLS

Given that the histone H3 N-terminus is highly conserved across eukaryotic evolution, a histone PTM specific antibody may be used to assess the degree of conservation of that mark. In this work, we generated an antibody specific for H3K23me3 and showed that this mark is conserved in *Tetrahymena*, *C. elegans* and mice. Additionally, we found that H3K23me3 is upregulated in meiotic chromatin of these species. We extend our observations in *Tetrahymena* to propose that H3K23me3 plays an integral and evolutionary conserved role in protecting genomic integrity in germ cells, suggesting a previously unappreciated heterochromatin-forming pathway helps limit DSB formation during meiosis.

Mass spectrometric evidence indicates that trace levels of H3K23me3 exist in mammalian somatic cells such as mouse embryonic fibroblasts, mouse embryonic stem cells, and HeLa cells.[55, 74] Given its low levels, we were unable to detect H3K23me3 in mammalian somatic cells by Western blot or immunofluorescence analyses (data not available). As DSB formation and recombination can occur in mitotic cells, albeit in dramatically reduced rates, the presence of H3K23me3 might serve a similar function in mitotic cells as its proposed function in meiotic cells. Still, we were able to detect high levels of H3K23me3 in mouse germ cells and embryonic brain tissues, indicating H3K23me3-mediated chromatin has specialized properties useful for germ and neuronal biology.

4.4B H3K23ME3 AND H3K9ME3 ARE FUNCTIONALLY DISTINCT

Previous studies concerning H3K23 methylation have hinted at a functional overlap between H3K23me3 and H3K9me3. Specifically, two studies to date have shown that the H3K9me3 reader, HP1, can bind to H3K23me3 *in vitro*.[55, 74] This observation enabled the authors to conclude that H3K23me3 is a heterochromatin-associated modification. Moreover, H3K9me3 plays a crucial role in inhibiting meiotic DSBs and crossover events through compaction of specific genomic regions, like chromocenters, thus limiting Spo11 access.[122, 180] Our findings that H3K23me3 is associated with chromocenters in mouse spermatocytes and with centromere-proximal regions in meiotic micronuclei of *Tetrahymena* are in agreement with a potential overlap in H3K9me3 and H3K23me3 activities.

Despite the striking similarities of H3K9 and H3K23, we have amassed several lines of evidence that suggest H3K9me3 and H3K23me3 are sufficiently distinct in function and localization. Our work in *Tetrahymena* clearly show: 1) H3K9me3 is absent altogether from *Tetrahymena* micronuclei and the loss of the other heterochromatin-associated mark, H3K27me3, does not lead to mislocalized DSBs; 2) there is a pronounced timing differential between the occurrence of H3K23me3 and H3K9me3 during conjugation; and 3) H3K23me3 is not detected in developing new macronuclei, which are enriched for H3K9me3. The difference between the appearance of H3K23me3 and H3K9me3 also suggests that α-H3K23me3 is specific for H3K23me3 rather than the H3K9me3 epitope. Most importantly, our immunofluorescence staining of mouse testis showed that H3K23me3 was restricted to the central region of the chromocenters. This staining profile is clearly different from the H3K9me3 staining profile, which overlaps

with the entire chromocenter.[170, 181] This suggests that H3K23me3 is associated with a subset of DNA elements located within H3K9me3-enriched chromocenters. As several proteins, like HP1 isoforms, have been shown to bind to H3K23me3 in high-throughput and *in vitro* assay, we speculate that these or as yet unknown proteins/complexes specifically interact with H3K23me3 *in vivo*, perhaps in the presence of H3K27me3, to mediate protection of the germline genome from Spo11 access and homologous recombination events during meiosis.[21, 55, 182]

APPENDIX

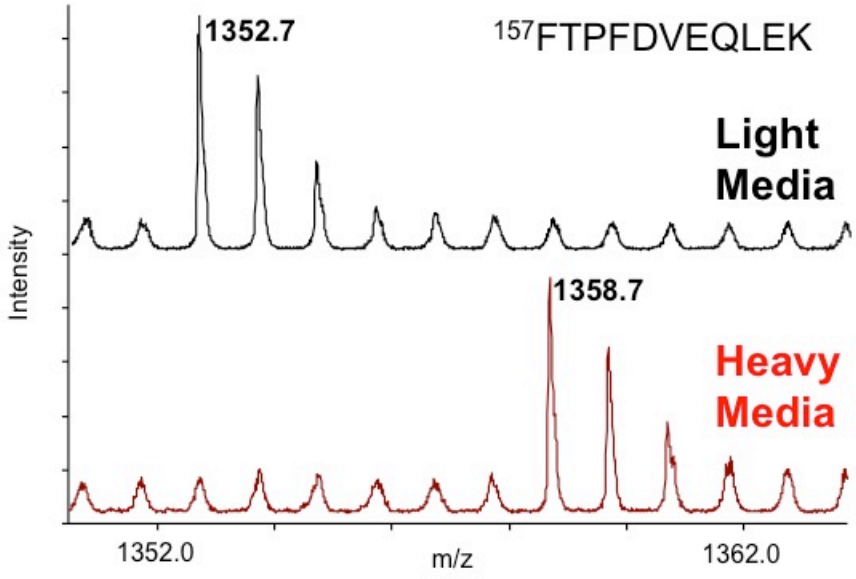
A.1 DEVELOPMENT OF A TECHNIQUE TO IMPROVE THE STUDY OF PROTEIN-PROTEIN INTERACTIONS IN TETRAHYMENA

Stable isotope labeling by amino acids in cell culture (SILAC) culture has proven to be a highly useful technique in circumventing inherent technical difficulties associated with the study of proteins. Among these, the co-purification of non-specific interacting proteins during co-immunoprecipitation assays is one of the most formidable obstacles since it often masks specific and stable protein-protein interactions that occur *in vivo.* I-DIRT (**I**sotopic **D**etermination of **I**nteractions as **R**andom or **T**argeted) was developed to improve weak signal-to-noise ratio typical of antibody-based protein complex purifications. A central feature of this protocol is that parallel cultures are grown in 'heavy' (^{13}C Arg/Lys) or 'light' (^{12}C Arg/Lys) labeled media. The light culture contains cells that express a tagged protein-of-interest. Cells from both cultures are flash frozen to maintain protein-protein interactions and subjected to cryolysis (- 80 °C) using a mixer mill. Equal amounts of light and heavy lysates are then mixed, resuspended in buffer, and immunoprecipitated using a tag- specific antibody. Thus, stable interaction with the tag-protein yields a distinct MS signature designated by the absence of heavy labeled peptides.

I sought to adapt the I-DIRT procedure for *Tetrahymena* to aid in the unbiased and unambiguous identification of protein complexes in this model organism. To determine if I-DIRT is feasible in *Tetrahymena*, I first tested whether these cells are

amenable to metabolic labeling during vegetative growth. A minimal media was generated following the formulation from a previous study[183]; however to make heavy media, light lysine was replaced with ^{13}C labeled lysine providing a 6 Dalton increase in overall mass. As shown in **Appendix Figure 1**, heavy isotope incorporation was highly efficient in *Tetrahymena* indicating that the I-DIRT technique can be developed in *Tetrahymena*.

Appendix Figure 1



Appendix Figure 1. Mass spectrometry analysis of heavy isotope incorporation. *Tetrahymena* were grown in minimal media containing ^{13}C Lys (heavy media) or ^{12}C Lys (light media). Results above indicate that 100% heavy Lys incorporation occurs in *Tetrahymena* growing in heavy media. A shift of 6 Daltons corresponds to the 6 heavy carbon atoms in Lys.

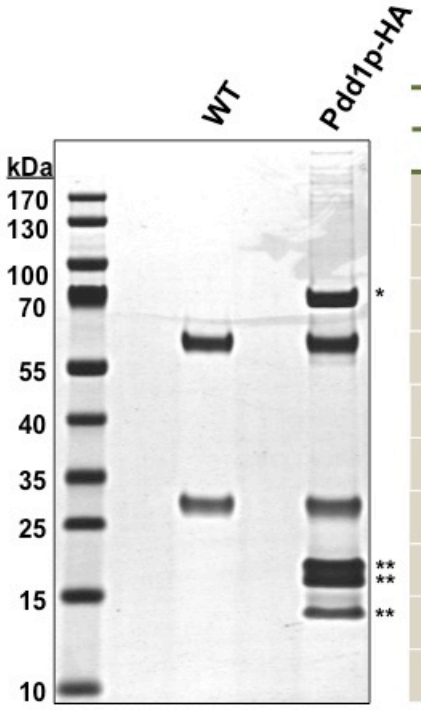
A.2 PDD1P INTERACTOME

Programmed DNA degradation protein (Pdd1p) is a chromodomain-containing protein that binds to H3K27me3 (and H3K9me3) in developing macronuclei during conjugation. This interaction is required for the successful elimination of micronuclear-limited sequences, including transposable elements, and is regulated by the upstream activities of piRNA-like scanRNA and E(z) homolog Ezl1p. Immunofluorescence staining has revealed that Pdd1p is uniformly distributed in new developing macronuclei at hour 9 of conjugation. However, by hour 12, the Pdd1p enriched chromatin forms cytologically distinct, heterochromatic bodies where DNA elimination is thought to occur. This process involves several proteins, but whether they directly bind to Pdd1p is not known, and the exact role of Pdd1p is not entirely clear. To date, no study has successfully identified proteins, including histones, that co-purify with Pdd1p *in vivo*.

Since previous attempts using standard lysing and Co-IP techniques have failed to yield proteins that interact with Pdd1p, we utilized a new technique at that time that is highly stringent in terms of lysing cells but favorable for maintaining protein-protein interactions. This method, which lyses cryogenically frozen cells, was initially developed to lyse yeast. We found that this cryolysis technique completely lysed *Tetrahymena* cells after just six rounds of pulverizing by a mixer mill (30 hertz, 3 min/round). The resulting frozen cell powder was resuspended in binding buffer (containing 300 mM salt), homogenized by drill, and then subjected to standard co-immunoprecipitation procedures. Using an anti-HA tag antibody, we performed co-immunoprecipitation analysis of a control wildtype strain and strain that expresses a HA-tagged Pdd1p protein. Remarkably,

we found that the new lysing conditions yielded several co-purifying proteins. Importantly, we found that Pdd1p pulls down seemingly equimolar levels of histone proteins as determined by Coomassie staining of a SDS-PAGE gel (**Appendix Figure 2**). We cut out several of these bands and subjected them to mass spectrometric identification in collaboration with Dr. Alan Tackett (University of Arkansas, Little Rock, AR). We identified several proteins associated with gene regulation **Appendix Figure 2**. Of note, we identified one peptide corresponding to Twi1p, the argonaute protein that binds to complimentary scanRNA and targets micronuclear-limited sequences for H3K9me3/H3K27me3 in new developing macronuclei. Moreover, we found that altering the salt concentration in the initial binding buffer leads to dramatic changes in the amount and identity of the co-puride proteins. These analyses also reveal that histones are best co-purified with Pdd1p in buffer containing 300 mM salt **(Appendix Figure 3)**.

Appendix Figure 2

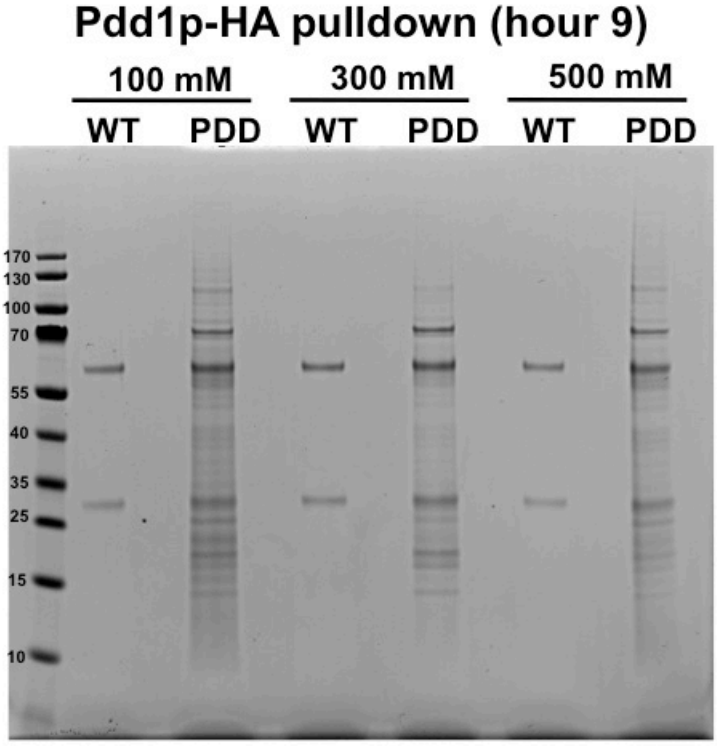


Pdd1p interactome

Protein	Potential function	Accession #
WGR containing protein	ADP ribosylation	TTHERM_00502600
Zinc knuckle family protein	Transcriptional regulator	TTHERM_01091290
Heterochromatin protein 1	Heterochr. formation	TTHERM_01337400
Ubiquitin E1 ligase	Ubiquitination	TTHERM_00997780
SNF2 family protein	Transcriptional regulation	TTHERM_00049310
Twi1p	RNAi binding	TTHERM_01161040
KOW motif family protein	Transcriptional regulator	TTHERM_00028580
JmjC domain protein	Histone Demethylase	TTHERM_00185640
Brg1p	ATPase, remodeling	TTHERM_01245640
SNF2 family protein	nucleosome remodeling	TTHERM_00388250

Appendix Figure 2. **The Pdd1p interactome**. Wildtype or Pdd1p-HA knock-in strains were cryogenically lysed, resuspended in IP buffer, and immunoprecipitated with anti-HA antibody. Differential bands were identified by mass spectrometry. Potential function was determined by BLAST and SMART domain search of identified proteins. Pdd1p is indicated by (*) and core histones are indicated by (**).

Appendix Figure 3



Appendix Figure 3. **Optimizing pulldown conditions for anti-Pdd1p**. HA-Pdd1p pulldowns under different concentrations of salt.

A.3 A PROTOCOL FOR STUDYING MULTIPLY ACETYLATED HISTONES

Citation: Papazyan, R. & Taverna, S. D. Separation and purification of multiply acetylated proteins using cation-exchange chromatography. *Methods Mol. Biol.* **981**, 103-113 (2013).

Title: Separation and purification of multiply acetylated proteins using cation-exchange chromatography

Abstract: High performance liquid chromatography (HPLC) is extremely useful for the study of proteins and the characterization of their post-translational modification. Here we describe a method that utilizes cation-exchange HPLC to separate multiply acetylated histone H3 species on the basis of their charge and hydrophilicity. This high-resolution method allows for the separation of histone H3 species that differ by as few as one acetyl group, and is compatible with subsequent analysis by a variety of techniques, including mass spectrometry and western blotting.

1. Introduction

Dynamic protein acetylation is known to impart key functional changes in biological systems, in particular, acetylation of eukaryotic histone proteins helps govern transcriptional competency of the associated DNA. It is increasingly appreciated that

acetylation can reside at multiple sites on the same histone polypeptide, and can even occur together with other small covalent post-translational modifications (PTMs) like methylation and phosphorylation. Specific combinations of such PTMs likely "code" for a variety of integrative chromatin-templated processes. However, the mechanistic basis of this regulation is poorly understood, in part because detection of multiple PTM states is complicated. For example, antibody-based recognition of PTMs is often dependent on epitope accessibility, which can vary depending on the modification state of surrounding residues. Fortunately, mass spectrometry methodologies have recently emerged that permit relatively unbiased PTM analysis on long stretches of amino acids or even intact proteins, although the complexity of PTM states usually necessitates extensive front-end purification to be suitable for such approaches (***1,2***).

While gel-based systems including acid-urea gel electrophoresis can be used to resolve acetylated (and phosphorylated) protein isoforms, the amount of material purified is relatively small and is in a form that is not compatible with limited proteolytic treatment necessary to utilize the mass spectrometry approaches referred to above. These deficiencies are remedied by cation-exchange HPLC. Under certain conditions, cation-exchange columns, like polyCAT A, permit separation of protein isoforms by exploiting their hydrophilic and electrostatic differences (***3-5***). For example, protein acetylation decreases positive charge and hydrophilicity, which contribute to the superior resolving power of polyCAT A chromatography. Furthermore, this chromatographic approach can also be performed on a preparative scale, yielding amounts of protein that are compatible with subsequent analysis by mass spectrometry and western blotting (***6***).

Below we describe our approach to separate and purify distinctly acetylated isoforms of histone H3 from the ciliated protozoan *Tetrahymena thermophila* by using polyCAT A-based chromatography. We start with highly-purified histone H3 from *Tetrahymena* which contains a complex mixture of acetylation states ranging from no acetylation to five (and possibly more) acetylated lysines. We treat this purified H3 with acetic anhydride (Ac_2O) to generate hyperacetylated isoforms *in vitro*. The hyperacetylated and endogenous H3 samples are independently applied to the polyCAT A column and eluted over a steep salt gradient. The concentrations of salt required to elute hyper- and endogenously acetylated H3 isoforms are used to demarcate elution conditions for all possible acetylated H3 isoforms. Endogenous H3 is then reapplied to the column and eluted using the truncated range of salt concentrations and an extended run time, thus achieving maximal resolution among acetyl isoforms. Lastly, samples are desalted by reversed phase-HPLC (RP-HPLC) and stored for subsequent analysis by mass spectrometry and western blots. Although we resolve histone H3 acetyl states in this protocol, this methodology can be adapted for many other acetylated proteins as well.

2. Materials

All commercially available solutions and reagents should be analytical grade. Water used to prepare solutions should be ultrapure and highly deionized. Additionally, any solution made for HPLC should be purified using at least a 0.45-micron filter and degassed prior to injection into the column.

2.1 In vitro acetylation of proteins

1. Acetic anhydride (Ac_2O)

2. Methanol (MeOH).

3. 2 x Reaction buffer: 100 mM ammonium bicarbonate (NH_4HCO_3, m.w. = 79.05). Dissolve 79 mg in 10 mL of water and do not adjust the pH (*see* **Note 1**).

4. RP-HPLC purified protein dissolved in water (~2 mg/mL) (*see* **Note 2**).

5. Microcentrifuge tubes.

6. Speed-vac centrifuge.

2.2 Desalting by Reversed Phase-HPLC

1. HPLC pump and fraction collector.

2. C8 column (220 x 4.6 mm, RP-300, #0711-0059, Perkin Elmer, Waltham, MA, USA).

3. Acetonitrile (CH_3CN), HPLC grade.

4. Trifluoroacetic acid (TFA).

5. Graduated cylinders, 1 L and 100 mL.

6. Magnetic stirrer, stir bars, and 1 L suction flasks.

7. Kontes vacuum filtration assembly (#XX1504700, Fisher Scientific, Fair Lawn, NJ, USA).

8. Durapore PVDF membrane filters 0.45 microns (#HVLP047000, Millipore, Billerica, MA, USA).

9. 1 mL Hamilton syringe equipped with blunt end needle point.

2.3 PolyCAT A chromatography buffers

1. Phosphate buffer: Stock solution is made by mixing appropriate amounts of NaH_2PO_4 (Sodium phosphate monobasic, solution 1) and Na_2HPO_4 (Sodium

phosphate dibasic, solution 2) to obtain the final buffer at pH 7. Start by making 1 M stocks of each solution. Dissolve 6.9 g of monobasic NaH_2PO_4-H_2O (m.w. = 138) in water to a final volume of 50 mL (Solution 1) and 7.1 g of dibasic Na_2HPO_4 (m.w. = 142) in water to a final volume of 50 mL (Solution 2). Shake both solutions vigorously to help dissolve. While stirring 50 mL of solution 2, adjust to pH 7 by slowly (1 mL at a time) adding solution 1. Store the 1 M Phosphate Buffer pH 7 at room temperature.

2. Urea [$CO(NH_2)_2$].
3. Dithiothreitol (DTT): 1 M solution dissolved in water (*see* **Note 3**).
4. Sodium chloride (NaCl).
5. Bio-Rex MSZ 501 [D] deioninzing resin (#142-7425, BioRAD, Hercules, CA, USA).
6. Gravity column (#732-1010, BioRAD, Hercules, CA, USA).
7. Two 500 mL graduated cylinders.
8. Magnetic stirrer, stir bar, and 1 L flask.

2.4 PolyCAT A column chromatography components

1. HPLC pump and fraction collector.
2. PolyCAT A column (200 X 4.6 mm, 5 μm, 1000Å, PolyLC, Columbia, MD, USA) (*see* **Note 4** and **Note 5**).
3. 1 mL Hamilton syringe equipped with blunt end needle point.
4. Dot blot apparatus (#170-6545, BioRAD, Hercules, CA, USA).
5. PVDF membrane.

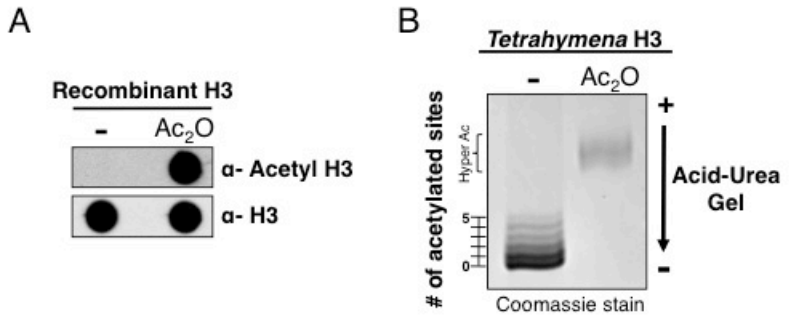
6. Necessary reagents and instructions for immunodetection are described in reference 6.

3. Methods

3.1 In vitro acetylation of proteins

1. Make 50 μL of acetylation buffer by mixing 12.5 μL of acetic anhydride with 37.5 μL of methanol (*see* **Note 6**).
2. In a separate microcentrifuge tube, mix 10 μL 2X reaction buffer and 10 μL of protein (2 mg/mL) (*see* **Note 7**).
3. Combine the two solutions (70 μL final volume) and briefly mix by gentle tapping.
4. Incubate at room temperature for 30 min.
5. Stop the reaction by drying with a speed-vac centrifuge (*see* **Note 8***)*.
6. Store dried samples at -20°C.
7. The extent of acetylation can be tested by mass spectrometry, immunoblot assays, or acid-urea gels after desalting the samples following the procedures in section 3.2 **(Appendix Figure 4)**.

Appendix Figure 4



Appendix Figure 4. ***In vitro*** **acetylation of proteins using acetic anhydride. (A)** Unmodified, recombinant histone H3 was treated with acetic anhydride (Ac_2O) and analyzed by dot blot immunodetection using antibodies specific to acetyled histone H3 or total H3. **(B)** The extent of *in vitro* acetylation of *Tetrahymena* histone H3 was measured by acid-urea gels, which combine size- and charge-based protein separation. The untreated H3 sample contains 6 bands with each band containing the indicated number of acetyl marks. Treatment with acetic anhydride generates hyperacetylated H3, which is represented by a significant upward shift in acid-urea gels.

3.2 Desalting by reversed-phase HPLC

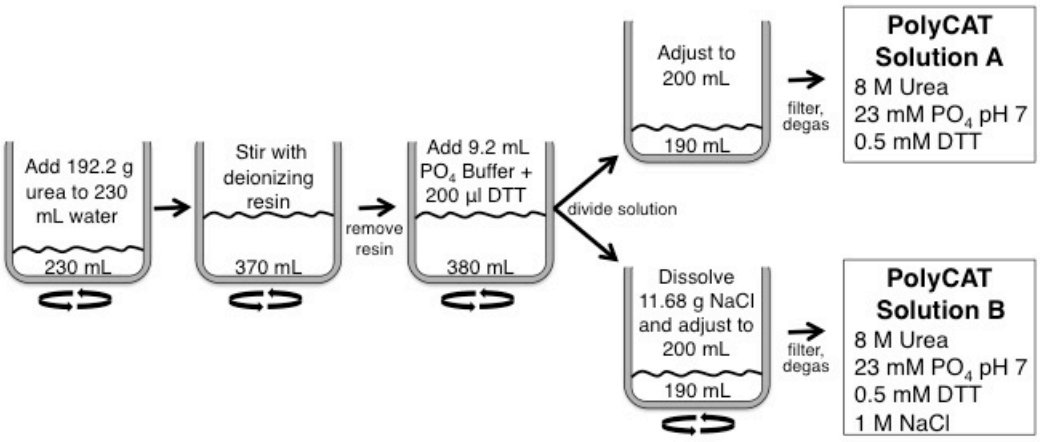
1. Prepare mobile phase for RP-HPLC: 1 L solution A: 5% acetonitrile, 95% water, 0.1% TFA; and 1 L solution B: 90% acetonitrile, 10% water, 0.1% TFA (*see* **Note 9**).
2. Vacuum filter both chromatography solutions using a 0.45 micron PVDF filter (*see* **Note 10**).
3. Transfer solutions into clean 1 L suction flasks with stir bar and stir at medium speed for 10 minutes under vacuum to degas each buffer.
4. Equilibrate the C8 column with 10 mL of solution A at a flow rate of 0.8 mL/min.
5. Take dried samples from section 3.1 and resuspend in 200 ul of solution A by vortexing.
6. Spin down >15,000 x *g* for 5 min to pellet any particulates.
7. Load into Hamilton syringe, inject the sample in the HPLC machine and run the chromatography method at 0.8 mL/min according to **Appendix Table 1** (*see* **Note 11**).
8. Record absorbance at 214 nm and collect relevant fractions.
9. Dry down fractions by speed-vac centrifugation.
10. Once dried, the samples can be stored at -20°C.

3.3 Preparation of the mobile phases for polyCAT A chromatography

These solutions are made fresh before use and discarded at the end of each day. A flow-chart for the procedures in this section and the final concentration of the components for the mobile phases are provided in **Appendix Figure 5**.

1. In a 1 L flask, dissolve 192.2 g of urea (m.w. = 60.06) in 230 mL of water. Stir at room temperature until the urea completely dissolves (*see* **Note 12**).
2. Remove ions from the urea solution by adding 1-5 g of deionizing resin. Deionize for an hour while stirring. Add more resin as needed (*see* **Note 13**).
3. Remove the resin by passing the urea solution through a gravity column. Collect the flow through in a 500 mL graduated cylinder.
4. While stirring on a magnetic plate add 9.2 mL of 1 M Phosphate buffer to the flow through.
5. Add 200 μL of 1 M DTT.
6. Adjust the volume to 380 mL with water.
7. Split the solution evenly into two graduated cylinders (190 mL in each cylinder). In the first cylinder add 10 mL water and label this "solution A" for polyCAT chromatography.
8. While stirring the second cylinder dissolve 11.68 g of NaCl (m.w. = 58.44) to obtain a final concentration of 1 M NaCl and label this as solution "B" for polyCAT chromatography. Adjust the final volume of solution B to 200 mL.
9. Filter and degas both buffers following the instructions above.

Appendix Figure 5



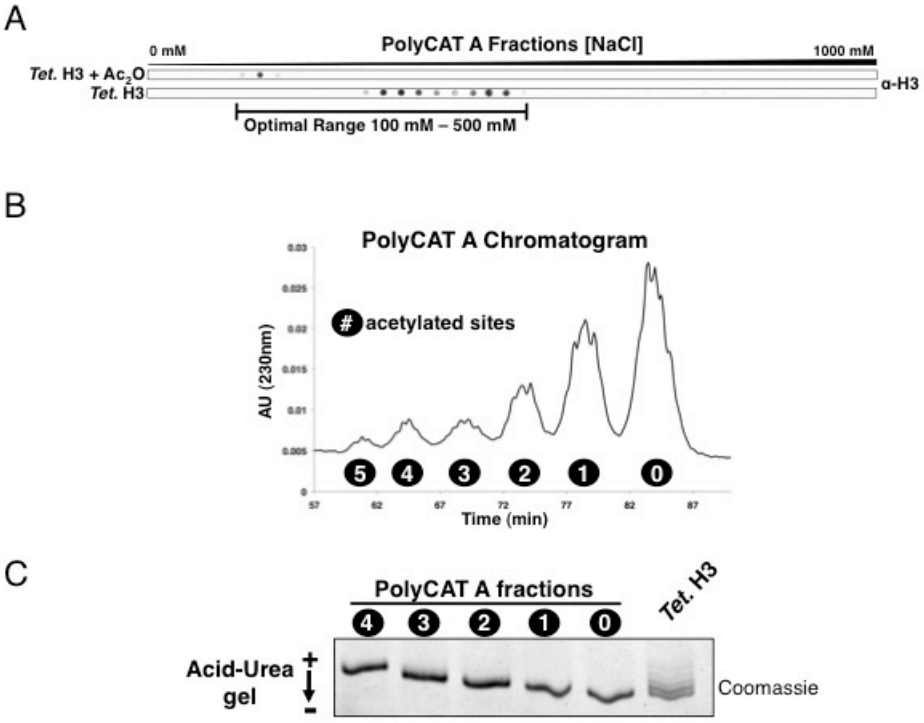
Appendix Figure 5. Flow chart of mobile phase preparation for PolyCAT A chromatography. Circulating arrows indicate that stirring is required at that step.

3.4 PolyCAT A chromatography

1. Equilibrate the column with at least 10 mL polyCAT solution A or until absorbance at 230 nm stabilizes. Use a flow rate of 0.4 mL/min for this step and all steps in the chromatography protocol (*see* **Note 5**).
2. While equilibrating, resuspend dried sample from Section **3.2** in 200 μL polyCAT solution A by vortexing for several seconds (*see* **Note 14**).
3. Pellet any particulates in the tube by centrifuging >15,000 x *g* for 5 min. Immediately after centrifugation, load the sample into a Hamilton syringe fitted with an appropriate blunt needle point designed for injections into HPLC machines.
4. Inject the sample into the HPLC machine and follow the chromatography protocol outlined in **Appendix Table 2**. Detect the presence of protein by constantly measuring absorbance at 230 nm (*see* **Note 15**). Collect fractions every 2 min (0.8 μL fraction volume) during the first 110 min of the run (*see* **Note 16**). Retention time increases as the number of acetylated sites on the protein decrease. For example, hyperacetylated proteins will elute before hypoacetylated species.
5. To obtain an optimal range of elution conditions repeat procedures 1-4 above using RP-HPLC purified protein samples that are not treated with Ac_2O. Compare all fractions from the two polyCAT A chromatography runs by using a dot blot apparatus (load ~10% of each fraction) followed by immunodetection (**Appendix Figure 6A**). The optimal salt concentration range will encompass fractions that include hyperacetylated (Ac_2O treated) and endogenously acetylated species.

6. Inject more sample into the polyCAT A column and adjust the elution gradient by increasing the time of elution within the newly defined elution range, as described in **Fig. 3B** (*see* **Note 17**). Successful optimization of elution conditions will yield individual peaks on the chromatogram that differ by one acetyl group (Appendix **Figure 6B** and *see* **Note 18**).
7. Desalt each relevant fraction by thawing them individually and injecting them onto a RP-HPLC column using the desalting procedures above followed by drying to completion using a speed-vac centrifuge. Dried fractions may be stored in -80°C.
8. Desalted samples may be analyzed by gel electophoresis or mass spectrometry (**Appendix Figure 6C** and reference ***6***).

Appendix Figure 6



Appendix Figure 6. **Separation of acetylated proteins using polyCAT A chromatography**. **(A)** Dot blot analysis comparing *Tetrahymena* H3 species that were either hyperacetylated *in vitro* using Ac_2O or endogenously acetylated. These samples represent the extremes of acetylation that may exist in a biologically relevant sample and serve as ideal starting material to optimize polyCAT A chromatography conditions. **(B)** Chromatogram of an optimized polyCAT A run using 75 μg of endogenously acetylated *Tetrahymena* histone H3. Solutions described in Appendix Figure 5 were used with the following chromatography protocol at a rate of 0.4 mL/min: Inject sample; 0%B for 10 min; 0-10%B for 10 min; 10%B for 15 min; 10-50%B for 65 min; 50-100%B for 10 min; 100%B for 10 min; 100-0%B for 5 min; and 0%B for 10 min. Sub-peaks within each major acetylation peak likely represent differently methylated H3 species. **(C)** Peaks 0-4 from part (B) were desalted and analyzed by acid-urea electrophoresis and Coomassie staining.

4. Notes

1. Due to the volatility of ammonium bicarbonate we recommend making fresh solution each time before use.
2. We obtain purified histone H3 by following the nuclear preparation, acid extraction, and RP-HPLC purification protocols described in Shechter et al. *(7)*. During nuclear preparation, we try to preserve endogenous protein acetylation by including 10 mM butyric acid, an HDAC inhibitor, in buffers.
3. To make 1 M stock solution, dissolve 1.54 g of DTT (m.w.= 154.25) in water to a final volume of 10 mL. Filter, divide to 1 mL aliquots, and store in -20C.
4. The specifications of the column will vary depending on the mass and properties of the protein being studied. Contact the manufacturer for additional help with selecting the proper column.
5. PolyCAT A column performance is significantly enhanced when prewashed with 40 mM EDTA-2Na overnight using a slow flow rate (0.25mL /min). Do not adjust the pH of this solution while dissolving and remember to filter purify the solution prior to passing it through the column. Wash with 10 mL filtered water and proceed to equilibrating the column with "solution A".
6. The *in vitro* acetylation protocol was adapted from reference 8.
7. This procedure is used to acetylate 20 μg of protein. If more acetylated protein is desired adjust the volume of the final reaction as needed and keep the same ratio between the acetylation and reaction buffers.
8. The acetic anhydride reaction is very fast and when using a speed-vac to dry the samples, the reaction yields fully acetylated proteins. In addition to lysines, this

treatment yields acetylated N-termini.

9. Prepare appropriate volumes of acetonitrile and water in separate graduated cylinders prior to mixing. Upon mixing the actual volume (~980 mL) will be less than the expected volume (1000 mL) due to entropy changes. Carefully add TFA in the end and proceed to filtration.
10. This critical step is required for all solutions used in HPLC. Removal of particulates extends the life of the column and minimizes clogging during a run.
11. Prior to this step you can save time by running a pure recombinant sample to determine the proper conditions and time of elution of the protein being studied. In general, acetylation or other small PTMs do not dramatically affect the column retention time of proteins, so hypoacetylated proteins will elute in the same fractions as hyperacetylated proteins using RP-HPLC. If recombinant protein is not available, collect fractions over the entirety of the initial run. After determining the elution time of the sample, the fraction collection window can be narrowed.
12. Do not heat the urea solution to dissolve because heating promotes the breakdown of urea into ions that lead to chemical modification (carbamylation) of the proteins being studied.
13. Deioninzing the urea solution is critical to eliminate carbamylation of proteins and to facilitate the interaction of the protein with the column. This resin contains a colorimetric indicator that turns gold or clear when saturated with ions. Continue adding resin until some of the resin remains blue (1-5 g).
14. Samples injected into the column should be free of salts. When possible, we

recommend drying the protein sample using a speed-vac centrifuge and resuspending it in polyCAT solution A prior to column injection.

15. When possible, absorbance at 230 nm should be recorded for polyCAT A chromatography of histones, because urea saturates the 214 nm signal.
16. This is not necessary for subsequent runs, but for the initial run we recommend collecting all fractions to help determine the extent of binding and retention in the column.
17. While we narrowed the salt concentration to optimize elution conditions, the samples were still resuspended in polyCAT A "solution A" containing no salt. The details of the optimized polyCAT A chromatography protocol is highlighted in the caption for **Appendix Figure 6B**. In general, samples should be injected onto the column using a solution that lacks salt. The presence of salt will disrupt the retention time and resolution of the analyte.
18. Following polyCAT A chromatography, we highly recommend washing the column with filtered water (>20 mL) followed by a total system clean of the HPLC following, especially if the same HPLC machine will be used for subsequent desalting steps by RP-HPLC. Mixing the mobile phases from each type of chromatography may cause precipitation of salts and significant damage to the column or HPLC pump.

Appendix Table 1

RP-HPLC protocol

Gradient	%B	Volume (mL)	Time (min)
Isocratic	0	4	5
Linear	0→100	24	30
Isocratic	100	4	5
Linear	100→0	1.6	2
Isocratic	0	4	5

Appendix Table 2

PolyCAT A protocol

Gradient	%B	Volume (mL)	Time (min)
Isocratic	0	4	10
Linear	0 →100	36	90
Isocratic	100	4	10
Linear	100 → 0	0.8	2
Isocratic	0	4	10

5. Acknowledgements

We thank T. Gilbert and other members of the Taverna laboratory for comments on this protocol. This work was supported by NIH grant RO190035489.

6. References

(1) Garcia, B.A., Shabanowitz, J., and Hunt, D.F. (2007) Characterization of histones and their post-translational modifications by mass spectrometry. *Curr Opin Chem Biol* 11, 66-73.

(2) Thomas, C.E., Kelleher, N.L., and Mizzen, C.A. (2006) Mass spectrometric characterization of human H3: A bird's eye view. *J Prot Res* 5, 240-247.

(3) Lindner, H.H. (2008) Analysis of histones, histone variants, and their post-translationally modified forms. *Electrophoresis* 29, 2516-2532.

(4) Zhang, K., et al. (2004) A mass spectrometric "Western blot" to evaluate the correlations between histone methylation and histone acetylation. *Proteomics* 4, 3765-3775.

(5) Alpert, A.J. (1983) Cation-exchange high-performance liquid chromatography of proteins on poly(aspartic acid)-silica. *J Chromatogr* 266, 23-37.

(6) Taverna, S.D., et al. (2006) Long-distance combinatorial linkage between methylation and acetylation on histone H3 termini. *Proc Natl Acad Sci USA* 104, 2086-2091.

(7) Schechter, D., et al. (2007) Extraction, purification and analysis of histones. *Nat Protoc* 2, 1445-1457.

(8) Acetylation of proteins and peptides. (2011) *http://www.ionsource.com/Card/acetylation/mono0003.htm*. Accessed 7 March 2011.

A.4 CONTRIBUTIONS TO SEQUENCING THE MICRONUCLEAR GENOME

The draft sequence of the macronuclear genome of *Tetrahymena* was published in 2006.[112] This study revealed that the genome size is approximately 104 megabases and is composed of 225 chromosomes. As expected, their findings also confirmed that most of the macronuclear genome is composed of genic DNA, associated regulatory elements, and introns. Interestingly, about 2.3% of the genome is repetitive suggesting that a small subset of repetitive elements can escape DNA elimination during conjugation. However, most of the known micronuclear-limited sequences including many classes of transposable elements and IESs were absent, as predicted by decades of biochemical studies. No transposable elements were detected, underscoring the major role of the DNA elimination pathway in removing these deleterious classes of DNA elements from the transcriptionally permissive macronuclear compartment.

The size of the macronuclear genome revealed that about 15-50% of the micronuclear genome is actively eliminated during conjugation. A recent effort in sequencing the micronuclear genome uncovered a new class of IESs and provided more micronuclear-limited sequences.[184] However, due to the highly repetitive nature and very high AT content of the micronuclear genome, the sequencing reads were not able to support the assembly of entire chromosome, the total genome size was not determined, and to what extent the micronuclear genome was sequenced was not known.

In collaboration with the Broad Institute, Dr. Robert S. Coyne (J. Craig Venter Institute, Rockville, MD) and Dr. Eduardo Orias (University of California, Santa Barbara), we sought to embark on a new micronuclear sequencing project with the goal

of assembling the entire genome. To address the complications above, we placed special emphasis on maximizing size of the sequencing libraries, providing high quality genomic DNA, and minimizing macronuclear DNA contamination. My contribution to this exciting project was to provide highly purified micronuclear DNA and perform some of the initial quality control experiments. The results of the project are summarized in "*Tetrahymena* Comparative Sequencing Project, Broad Institute of Harvard and MIT (http://www.broadinstitute.org/)". Briefly, our findings showed that the micronuclear genome is approximately 158 megabases in length and about 22.08% GC rich. The assembly is currently underway.

REFERENCES

1. Alberts, B. *et al*. in *Molecular Biology of the Cell. 3rd edn. Ch. 22* (Garland Science, New York, 1994).

2. Fujimori, T., Kurotaki, Y., Miyazaki, J. & Nabeshima, Y. Analysis of cell lineage in two- and four-cell mouse embryos. *Development* **130**, 5113-5122 (2003).

3. Piotrowska-Nitsche, K., Perea-Gomez, A., Haraguchi, S. & Zernicka-Goetz, M. Four-cell stage mouse blastomeres have different developmental properties. *Development* **132**, 479-490 (2005).

4. Holliday, R. & Pugh, J. E. DNA modification mechanisms and gene activity during development. *Science* **187**, 226-232 (1975).

5. Riggs, A. D. X inactivation, differentiation, and DNA methylation. *Cytogenet. Cell Genet.* **14**, 9-25 (1975).

6. Rivera, C. M. & Ren, B. Mapping human epigenomes. *Cell* **155**, 39-55 (2013).

7. Heitz, E. Das heterochromatin der moose. *I. Jahrb. Wiss. Botanik* **69**, 762-818 (1928).

8. Passarge, E. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. *Am. J. Hum. Genet.* **31**, 106-115 (1979).

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

10. Oudet, P., Gross-Bellard, M. & Chambon, P. Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. *Cell* **4**, 281-300 (1975).

11. ALLFREY, V. G., FAULKNER, R. & MIRSKY, A. E. Acetylation and Methylation of Histones and their Possible Role in the Regulation of Rna Synthesis. *Proc. Natl. Acad. Sci. U. S. A.* **51**, 786-794 (1964).

12. Olins, D. E. & Olins, A. L. Chromatin history: our view from the bridge. *Nat. Rev. Mol. Cell Biol.* **4**, 809-814 (2003).

13. Finch, J. T., Noll, M. & Kornberg, R. D. Electron microscopy of defined lengths of chromatin. *Proc. Natl. Acad. Sci. U. S. A.* **72**, 3320-3322 (1975).

14. Olins, A. L. & Olins, D. E. Spheroid chromatin units (v bodies). *Science* **183**, 330-332 (1974).

15. Kornberg, R. D. & Lorch, Y. Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. *Cell* **98**, 285-294 (1999).

16. Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F. & Richmond, T. J. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389**, 251-260 (1997).

17. Cao, K. *et al*. High-resolution mapping of h1 linker histone variants in embryonic stem cells. *PLoS Genet*. **9**, e1003417 (2013).

18. Yang, S. M., Kim, B. J., Norwood Toro, L. & Skoultchi, A. I. H1 linker histone promotes epigenetic silencing by regulating both DNA methylation and histone H3 methylation. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 1708-1713 (2013).

19. Oberdoerffer, P. & Sinclair, D. A. The role of nuclear architecture in genomic instability and ageing. *Nat. Rev. Mol. Cell Biol.* **8**, 692-702 (2007).

20. Kouzarides, T. Chromatin modifications and their function. *Cell* **128**, 693-705 (2007).

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

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

23. Jenuwein, T. & Allis, C. D. Translating the histone code. *Science* **293**, 1074-1080 (2001).

24. Turner, B. M. Histone acetylation and an epigenetic code. *Bioessays* **22**, 836-845 (2000).

25. Hong, L., Schroth, G. P., Matthews, H. R., Yau, P. & Bradbury, E. M. Studies of the DNA binding properties of histone H4 amino terminus. Thermal denaturation studies reveal that acetylation markedly reduces the binding constant of the H4 "tail" to DNA. *J. Biol. Chem.* **268**, 305-314 (1993).

26. Li, H. J. & Eckhardt, R. A. in *Chromatin and chromosome structure* 357 (Academic Press, New York, 1977).

27. Turner, B. M. Decoding the nucleosome. *Cell* **75**, 5-8 (1993).

28. Vidali, G., Boffa, L. C., Bradbury, E. M. & Allfrey, V. G. Butyrate suppression of histone deacetylation leads to accumulation of multiacetylated forms of histones H3 and H4 and increased DNase I sensitivity of the associated DNA sequences. *Proc. Natl. Acad. Sci. U. S. A.* **75**, 2239-2243 (1978).

29. Jeppesen, P. & Turner, B. M. The inactive X chromosome in female mammals is distinguished by a lack of histone H4 acetylation, a cytogenetic marker for gene expression. *Cell* **74**, 281-289 (1993).

30. Hebbes, T. R., Thorne, A. W. & Crane-Robinson, C. A direct link between core histone acetylation and transcriptionally active chromatin. *EMBO J.* **7**, 1395-1402 (1988).

31. Grunstein, M. Histone acetylation in chromatin structure and transcription. *Nature* **389**, 349-352 (1997).

32. Sun, Z. *et al*. Deacetylase-Independent Function of HDAC3 in Transcription and Metabolism Requires Nuclear Receptor Corepressor. *Mol. Cell* (2013).

33. Shahbazian, M. D. & Grunstein, M. Functions of site-specific histone acetylation and deacetylation. *Annu. Rev. Biochem.* **76**, 75-100 (2007).

34. Ruiz-Carrillo, A., Wangh, L. J. & Allfrey, V. G. Processing of newly synthesized histone molecules. *Science* **190**, 117-128 (1975).

35. Brownell, J. E. *et al*. Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* **84**, 843-851 (1996).

36. Georgakopoulos, T. & Thireos, G. Two distinct yeast transcriptional activators require the function of the GCN5 protein to promote normal levels of transcription. *EMBO J.* **11**, 4145-4152 (1992).

37. Taunton, J., Hassig, C. A. & Schreiber, S. L. A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* **272**, 408-411 (1996).

38. Struhl, K. Histone acetylation and transcriptional regulatory mechanisms. *Genes Dev.* **12**, 599-606 (1998).

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

40. Zeng, L. & Zhou, M. M. Bromodomain: an acetyl-lysine binding domain. *FEBS Lett.* **513**, 124-128 (2002).

41. Rossetto, D., Avvakumov, N. & Cote, J. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics* **7**, 1098-1108 (2012).

42. Seet, B. T., Dikic, I., Zhou, M. M. & Pawson, T. Reading protein modifications with interaction domains. *Nat. Rev. Mol. Cell Biol.* **7**, 473-483 (2006).

43. Wei, Y., Yu, L., Bowen, J., Gorovsky, M. A. & Allis, C. D. Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell* **97**, 99-109 (1999).

44. Fischle, W. *et al*. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**, 1116-1122 (2005).

45. Prigent, C. & Dimitrov, S. Phosphorylation of serine 10 in histone H3, what for? *J. Cell. Sci.* **116**, 3677-3685 (2003).

46. Fillingham, J., Keogh, M. C. & Krogan, N. J. GammaH2AX and its role in DNA double-strand break repair. *Biochem. Cell Biol.* **84**, 568-577 (2006).

47. Celeste, A. *et al*. Genomic instability in mice lacking histone H2AX. *Science* **296**, 922-927 (2002).

48. Moore, J. D., Yazgan, O., Ataian, Y. & Krebs, J. E. Diverse roles for histone H2A modifications in DNA damage response pathways in yeast. *Genetics* **176**, 15-25 (2007).

49. Redon, C. *et al*. Yeast histone 2A serine 129 is essential for the efficient repair of checkpoint-blind DNA damage. *EMBO Rep*. **4**, 678-684 (2003).

50. Celeste, A. *et al*. Histone H2AX phosphorylation is dispensable for the initial recognition of DNA breaks. *Nat. Cell Biol*. **5**, 675-679 (2003).

51. Paull, T. T. *et al*. A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr. Biol*. **10**, 886-895 (2000).

52. Rogakou, E. P., Pilch, D. R., Orr, A. H., Ivanova, V. S. & Bonner, W. M. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem*. **273**, 5858-5868 (1998).

53. Rogakou, E. P., Boon, C., Redon, C. & Bonner, W. M. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J. Cell Biol*. **146**, 905-916 (1999).

54. Wang, C. *et al*. DNA damage response and cellular senescence in tissues of aging mice. *Aging Cell*. **8**, 311-323 (2009).

55. Liu, H. *et al*. Systematic identification of methyllysine-driven interactions for histone and nonhistone targets. *J. Proteome Res*. **9**, 5827-5836 (2010).

56. Hake, S. B. *et al*. Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *J. Biol. Chem*. **281**, 559-568 (2006).

57. Taverna, S. D. *et al*. Long-distance combinatorial linkage between methylation and acetylation on histone H3 N termini. *Proc. Natl. Acad. Sci. U. S. A*. **104**, 2086-2091 (2007).

58. Eissenberg, J. C. & Shilatifard, A. Histone H3 lysine 4 (H3K4) methylation in development and differentiation. *Dev. Biol*. **339**, 240-249 (2010).

59. Taverna, S. D. *et al*. Yng1 PHD finger binding to H3 trimethylated at K4 promotes NuA3 HAT activity at K14 of H3 and transcription at a subset of targeted ORFs. *Mol. Cell* **24**, 785-796 (2006).

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

61. Bannister, A. J. *et al*. Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* **410**, 120-124 (2001).

62. James, T. C. *et al*. Distribution patterns of HP1, a heterochromatin-associated nonhistone chromosomal protein of Drosophila. *Eur. J. Cell Biol*. **50**, 170-180 (1989).

63. Peters, A. H. *et al*. Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-337 (2001).

64. Jacobs, S. A. *et al*. Specificity of the HP1 chromo domain for the methylated N-terminus of histone H3. *EMBO J*. **20**, 5232-5241 (2001).

65. Wu, R., Terry, A. V., Singh, P. B. & Gilbert, D. M. Differential subnuclear localization and replication timing of histone H3 lysine 9 methylation states. *Mol. Biol. Cell* **16**, 2872-2881 (2005).

66. Krauss, V. Glimpses of evolution: heterochromatic histone H3K9 methyltransferases left its marks behind. *Genetica* **133**, 93-106 (2008).

67. Pinheiro, I. *et al*. Prdm3 and Prdm16 are H3K9me1 methyltransferases required for mammalian heterochromatin integrity. *Cell* **150**, 948-960 (2012).

68. Towbin, B. D. *et al*. Step-wise methylation of histone H3K9 positions heterochromatin at the nuclear periphery. *Cell* **150**, 934-947 (2012).

69. Yokochi, T. *et al*. G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 19363-19368 (2009).

70. Wen, B., Wu, H., Shinkai, Y., Irizarry, R. A. & Feinberg, A. P. Large histone H3 lysine 9 dimethylated chromatin blocks distinguish differentiated from embryonic stem cells. *Nat. Genet.* **41**, 246-250 (2009).

71. Reddy, K. L., Zullo, J. M., Bertolino, E. & Singh, H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**, 243-247 (2008).

72. Meister, P. & Taddei, A. Building silent compartments at the nuclear periphery: a recurrent theme. *Curr. Opin. Genet. Dev.* **23**, 96-103 (2013).

73. Reddy, K. L. & Feinberg, A. P. Higher order chromatin organization in cancer. *Semin. Cancer Biol.* **23**, 109-115 (2013).

74. Leroy, G. *et al*. Proteogenomic characterization and mapping of nucleosomes decoded by Brd and HP1 proteins. *Genome Biol.* **13**, R68 (2012).

75. Cao, R. *et al*. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. *Science* **298**, 1039-1043 (2002).

76. Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nat. Rev. Mol. Cell Biol.* **10**, 697-708 (2009).

77. Gao, Z. *et al*. PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Mol. Cell* **45**, 344-356 (2012).

78. Di Croce, L. & Helin, K. Transcriptional regulation by Polycomb group proteins. *Nat. Struct. Mol. Biol.* **20**, 1147-1155 (2013).

79. Rastelli, L., Chan, C. S. & Pirrotta, V. Related chromosome binding sites for zeste, suppressors of zeste and Polycomb group proteins in Drosophila and their dependence on Enhancer of zeste function. *EMBO J.* **12**, 1513-1522 (1993).

80. Holdeman, R., Nehrt, S. & Strome, S. MES-2, a maternal protein essential for viability of the germline in Caenorhabditis elegans, is homologous to a Drosophila Polycomb group protein. *Development* **125**, 2457-2467 (1998).

81. O'Carroll, D. *et al*. The polycomb-group gene Ezh2 is required for early mouse development. *Mol. Cell. Biol.* **21**, 4330-4336 (2001).

82. Tie, F., Furuyama, T., Prasad-Sinha, J., Jane, E. & Harte, P. J. The Drosophila Polycomb Group proteins ESC and E(Z) are present in a complex containing the histone-binding protein p55 and the histone deacetylase RPD3. *Development* **128**, 275-286 (2001).

83. Bernstein, B. E. *et al*. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**, 315-326 (2006).

84. Davidovich, C., Zheng, L., Goodrich, K. J. & Cech, T. R. Promiscuous RNA binding by Polycomb repressive complex 2. *Nat. Struct. Mol. Biol.* **20**, 1250-1257 (2013).

85. Zhao, J. *et al*. Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol. Cell* **40**, 939-953 (2010).

86. Endoh, M. *et al*. Histone H2A mono-ubiquitination is a crucial step to mediate PRC1-dependent repression of developmental genes to maintain ES cell identity. *PLoS Genet.* **8**, e1002774 (2012).

87. Wang, H. *et al*. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873-878 (2004).

88. Lee, M. G. *et al*. Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. *Science* **318**, 447-450 (2007).

89. Tavares, L. *et al*. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148**, 664-678 (2012).

90. Margueron, R. *et al*. Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461**, 762-767 (2009).

91. Ruthenburg, A. J., Li, H., Patel, D. J. & Allis, C. D. Multivalent engagement of chromatin modifications by linked binding modules. *Nat. Rev. Mol. Cell Biol.* **8**, 983-994 (2007).

92. Ruthenburg, A. J. *et al*. Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* **145**, 692-706 (2011).

93. Rothbart, S. B. *et al*. Multivalent histone engagement by the linked tandem Tudor and PHD domains of UHRF1 is required for the epigenetic inheritance of DNA methylation. *Genes Dev.* **27**, 1288-1298 (2013).

94. Moriniere, J. *et al*. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461**, 664-668 (2009).

95. Henikoff, S. & Shilatifard, A. Histone modification: cause or cog? *Trends Genet*. **27**, 389-396 (2011).

96. Rando, O. J. Combinatorial complexity in chromatin structure and function: revisiting the histone code. *Curr. Opin. Genet. Dev*. **22**, 148-155 (2012).

97. Wood, V. *et al*. The genome sequence of Schizosaccharomyces pombe. *Nature* **415**, 871-880 (2002).

98. Grewal, S. I. & Jia, S. Heterochromatin revisited. *Nat. Rev. Genet*. **8**, 35-46 (2007).

99. Fuchs, S. M., Krajewski, K., Baker, R. W., Miller, V. L. & Strahl, B. D. Influence of combinatorial histone modifications on antibody and effector protein recognition. *Curr. Biol*. **21**, 53-58 (2011).

100. Gardner, K. E., Allis, C. D. & Strahl, B. D. Operating on chromatin, a colorful language where context matters. *J. Mol. Biol*. **409**, 36-46 (2011).

101. ENCODE Project Consortium *et al*. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57-74 (2012).

102. Djebali, S. *et al*. Landscape of transcription in human cells. *Nature* **489**, 101-108 (2012).

103. ENCODE Project Consortium *et al*. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**, 799-816 (2007).

104. Baker, M. Making sense of chromatin states. *Nat. Methods* **8**, 717-722 (2011).

105. Ernst, J. & Kellis, M. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nat. Biotechnol*. **28**, 817-825 (2010).

106. Filion, G. J. *et al*. Systematic protein location mapping reveals five principal chromatin types in Drosophila cells. *Cell* **143**, 212-224 (2010).

107. Kharchenko, P. V. *et al*. Comprehensive analysis of the chromatin landscape in Drosophila melanogaster. *Nature* **471**, 480-485 (2011).

108. Greider, C. W. & Blackburn, E. H. Identification of a specific telomere terminal transferase activity in Tetrahymena extracts. *Cell* **43**, 405-413 (1985).

109. Zaug, A. J. & Cech, T. R. The intervening sequence RNA of Tetrahymena is an enzyme. *Science* **231**, 470-475 (1986).

110. Yao, M. C. & Yao, C. H. Repeated hexanucleotide C-C-C-C-A-A is present near free ends of macronuclear DNA of Tetrahymena. *Proc. Natl. Acad. Sci. U. S. A*. **78**, 7436-7439 (1981).

111. Chalker, D. L. Dynamic nuclear reorganization during genome remodeling of Tetrahymena. *Biochim. Biophys. Acta* **1783**, 2130-2136 (2008).

112. Eisen, J. A. *et al*. Macronuclear genome sequence of the ciliate Tetrahymena thermophila, a model eukaryote. *PLoS Biol*. **4**, e286 (2006).

113. Gorovsky, M. A., Yao, M. C., Keevert, J. B. & Pleger, G. L. Isolation of micro- and macronuclei of Tetrahymena pyriformis. *Methods Cell Biol*. **9**, 311-327 (1975).

114. Chalker, D. L., Meyer, E. & Mochizuki, K. Epigenetics of ciliates. *Cold Spring Harb Perspect. Biol*. **5**, 10.1101/cshperspect.a017764 (2013).

115. Taverna, S. D., Coyne, R. S. & Allis, C. D. Methylation of histone h3 at lysine 9 targets programmed DNA elimination in tetrahymena. *Cell* **110**, 701-711 (2002).

116. Cervantes, M. D., Xi, X., Vermaak, D., Yao, M. C. & Malik, H. S. The CNA1 histone of the ciliate Tetrahymena thermophila is essential for chromosome segregation in the germline micronucleus. *Mol. Biol. Cell* **17**, 485-497 (2006).

117. Endo, M. & Sugai, T. Amitotic division of the macronucleus in Tetrahymena thermophila: DNA distribution by genomic unit. *Zoolog Sci*. **28**, 482-490 (2011).

118. Gorovsky, M. A. Genome organization and reorganization in Tetrahymena. *Annu. Rev. Genet*. **14**, 203-239 (1980).

119. Bruns, P. J. & Brussard, T. E. Nullisomic tetrahymena: eliminating germinal chromosomes. *Science* **213**, 549-551 (1981).

120. Sugai, T. & Hiwatashi, K. Cytologic and autoradiographic studies of the micronucleus at meiotic prophase in Tetrahymena pyriformis. *J. Protozool*. **21**, 542-548 (1974).

121. Huang, H., Wiley, E. A., Lending, C. R. & Allis, C. D. An HP1-like protein is missing from transcriptionally silent micronuclei of Tetrahymena. *Proc. Natl. Acad. Sci. U. S. A*. **95**, 13624-13629 (1998).

122. Wu, S., Hu, Y. C., Liu, H. & Shi, Y. Loss of YY1 impacts the heterochromatic state and meiotic double-strand breaks during mouse spermatogenesis. *Mol. Cell. Biol*. **29**, 6245-6256 (2009).

123. Allis, C. D., Bowen, J. K., Abraham, G. N., Glover, C. V. & Gorovsky, M. A. Proteolytic processing of histone H3 in chromatin: a physiologically regulated event in Tetrahymena micronuclei. *Cell* **20**, 55-64 (1980).

124. Strahl, B. D., Ohba, R., Cook, R. G. & Allis, C. D. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proc. Natl. Acad. Sci. U. S. A*. **96**, 14967-14972 (1999).

125. Bernstein, E. *et al*. Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. *Mol. Cell. Biol*. **26**, 2560-2569 (2006).

126. Rougeulle, C. *et al*. Differential histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Mol. Cell. Biol.* **24**, 5475-5484 (2004).

127. Huang, H., Smothers, J. F., Wiley, E. A. & Allis, C. D. A nonessential HP1-like protein affects starvation-induced assembly of condensed chromatin and gene expression in macronuclei of Tetrahymena thermophila. *Mol. Cell. Biol.* **19**, 3624-3634 (1999).

128. Wei, Y., Mizzen, C. A., Cook, R. G., Gorovsky, M. A. & Allis, C. D. Phosphorylation of histone H3 at serine 10 is correlated with chromosome condensation during mitosis and meiosis in Tetrahymena. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 7480-7484 (1998).

129. Vavra, K. J., Allis, C. D. & Gorovsky, M. A. Regulation of histone acetylation in Tetrahymena macro- and micronuclei. *J. Biol. Chem.* **257**, 2591-2598 (1982).

130. Coon, J. J. *et al*. Protein identification using sequential ion/ion reactions and tandem mass spectrometry. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 9463-9468 (2005).

131. Garcia, B. A., Pesavento, J. J., Mizzen, C. A. & Kelleher, N. L. Pervasive combinatorial modification of histone H3 in human cells. *Nat. Methods* **4**, 487-489 (2007).

132. Young, N. L. *et al*. High throughput characterization of combinatorial histone codes. *Mol. Cell. Proteomics* **8**, 2266-2284 (2009).

133. Thomas, C. E., Kelleher, N. L. & Mizzen, C. A. Mass spectrometric characterization of human histone H3: a bird's eye view. *J. Proteome Res.* **5**, 240-247 (2006).

134. Allis, C. D. & Dennison, D. K. Identification and purification of young macronuclear anlagen from conjugating cells of Tetrahymena thermophila. *Dev. Biol.* **93**, 519-533 (1982).

135. Sweet, M. T. & Allis, C. D. Isolation and purification of tetrahymena nuclei. *CSH Protoc.* **2006**, 10.1101/pdb.prot4500 (2006).

136. Shechter, D., Dormann, H. L., Allis, C. D. & Hake, S. B. Extraction, purification and analysis of histones. *Nat. Protoc.* **2**, 1445-1457 (2007).

137. Geer, L. Y. *et al*. Open mass spectrometry search algorithm. *J. Proteome Res.* **3**, 958-964 (2004).

138. Harlow, E. in *Antibodies : a laboratory manual* (ed Lane, D.) (Cold Spring Harbor Laboratory, New York, 1988).

139. Laemmli, U. K. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685 (1970).

140. Madireddi, M. T., Davis, M. C. & Allis, C. D. Identification of a novel polypeptide involved in the formation of DNA-containing vesicles during macronuclear development in Tetrahymena. *Dev. Biol.* **165**, 418-431 (1994).

141. Wenkert, D. & Allis, C. D. Timing of the appearance of macronuclear-specific histone variant hv1 and gene expression in developing new macronuclei of Tetrahymena thermophila. *J. Cell Biol.* **98**, 2107-2117 (1984).

142. Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T. & Allis, C. D. Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 1237-1241 (1995).

143. Nightingale, K. P. *et al*. Cross-talk between histone modifications in response to histone deacetylase inhibitors: MLL4 links histone H3 acetylation and histone H3K4 methylation. *J. Biol. Chem.* **282**, 4408-4416 (2007).

144. Smith, J. J., Torigoe, S. E., Maxson, J., Fish, L. C. & Wiley, E. A. A class II histone deacetylase acts on newly synthesized histones in Tetrahymena. *Eukaryot. Cell.* **7**, 471-482 (2008).

145. Garcia, B. A. *et al*. Organismal differences in post-translational modifications in histones H3 and H4. *J. Biol. Chem.* **282**, 7641-7655 (2007).

146. Waterborg, J. H. Sequence analysis of acetylation and methylation in two histone H3 variants of alfalfa. *J. Biol. Chem.* **265**, 17157-17161 (1990).

147. Miao, W. *et al*. Microarray analyses of gene expression during the Tetrahymena thermophila life cycle. *PLoS One* **4**, e4429 (2009).

148. Xiao, B., Wilson, J. R. & Gamblin, S. J. SET domains and histone methylation. *Curr. Opin. Struct. Biol.* **13**, 699-705 (2003).

149. Hennig, L. & Derkacheva, M. Diversity of Polycomb group complexes in plants: same rules, different players? *Trends Genet.* **25**, 414-423 (2009).

150. Liu, Y. *et al*. RNAi-dependent H3K27 methylation is required for heterochromatin formation and DNA elimination in Tetrahymena. *Genes Dev.* **21**, 1530-1545 (2007).

151. Schoeberl, U. E., Kurth, H. M., Noto, T. & Mochizuki, K. Biased transcription and selective degradation of small RNAs shape the pattern of DNA elimination in Tetrahymena. *Genes Dev.* **26**, 1729-1742 (2012).

152. Mochizuki, K., Fine, N. A., Fujisawa, T. & Gorovsky, M. A. Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in tetrahymena. *Cell* **110**, 689-699 (2002).

153. Aronica, L. *et al*. Study of an RNA helicase implicates small RNA-noncoding RNA interactions in programmed DNA elimination in Tetrahymena. *Genes Dev.* **22**, 2228-2241 (2008).

154. Baudat, F., Imai, Y. & de Massy, B. Meiotic recombination in mammals: localization and regulation. *Nat. Rev. Genet.* **14**, 794-806 (2013).

155. Lichten, M. & de Massy, B. The impressionistic landscape of meiotic recombination. *Cell* **147**, 267-270 (2011).

156. Kimble, J. & Crittenden, S. L. Controls of germline stem cells, entry into meiosis, and the sperm/oocyte decision in Caenorhabditis elegans. *Annu. Rev. Cell Dev. Biol.* **23**, 405-433 (2007).

157. Bruns, P. J. & Cassidy-Hanley, D. Biolistic transformation of macro- and micronuclei. *Methods Cell Biol.* **62**, 501-512 (2000).

158. Hamilton, E. P. & Orias, E. Genetic crosses: setting up crosses, testing progeny, and isolating phenotypic assortants. *Methods Cell Biol.* **62**, 219-228 (2000).

159. Chung, P. H. & Yao, M. C. Tetrahymena thermophila JMJD3 homolog regulates H3K27 methylation and nuclear differentiation. *Eukaryot. Cell.* **11**, 601-614 (2012).

160. Loidl, J. & Scherthan, H. Organization and pairing of meiotic chromosomes in the ciliate Tetrahymena thermophila. *J. Cell. Sci.* **117**, 5791-5801 (2004).

161. Mochizuki, K., Novatchkova, M. & Loidl, J. DNA double-strand breaks, but not crossovers, are required for the reorganization of meiotic nuclei in Tetrahymena. *J. Cell. Sci.* **121**, 2148-2158 (2008).

162. Brachet, E., Sommermeyer, V. & Borde, V. Interplay between modifications of chromatin and meiotic recombination hotspots. *Biol. Cell.* **104**, 51-69 (2012).

163. Tsai, J. H. & McKee, B. D. Homologous pairing and the role of pairing centers in meiosis. *J. Cell. Sci.* **124**, 1955-1963 (2011).

164. Dobson, M. J., Pearlman, R. E., Karaiskakis, A., Spyropoulos, B. & Moens, P. B. Synaptonemal complex proteins: occurrence, epitope mapping and chromosome disjunction. *J. Cell. Sci.* **107 (Pt 10)**, 2749-2760 (1994).

165. Boateng, K. A., Bellani, M. A., Gregoretti, I. V., Pratto, F. & Camerini-Otero, R. D. Homologous pairing preceding SPO11-mediated double-strand breaks in mice. *Dev. Cell.* **24**, 196-205 (2013).

166. Lee, J. M. *et al.* EZH2 generates a methyl degron that is recognized by the DCAF1/DDB1/CUL4 E3 ubiquitin ligase complex. *Mol. Cell* **48**, 572-586 (2012).

167. Colaiacovo, M. P. *et al.* Synaptonemal complex assembly in C. elegans is dispensable for loading strand-exchange proteins but critical for proper completion of recombination. *Dev. Cell.* **5**, 463-474 (2003).

168. Cooke, H. J. & Saunders, P. T. Mouse models of male infertility. *Nat. Rev. Genet.* **3**, 790-801 (2002).

169. Govin, J. *et al.* Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *J. Cell Biol.* **176**, 283-294 (2007).

170. Martianov, I. *et al.* Distinct functions of TBP and TLF/TRF2 during spermatogenesis: requirement of TLF for heterochromatic chromocenter formation in haploid round spermatids. *Development* **129**, 945-955 (2002).

171. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77**, 71-94 (1974).

172. Timmons, L. & Fire, A. Specific interference by ingested dsRNA. *Nature* **395**, 854 (1998).

173. Maddox, P. S., Hyndman, F., Monen, J., Oegema, K. & Desai, A. Functional genomics identifies a Myb domain-containing protein family required for assembly of CENP-A chromatin. *J. Cell Biol.* **176**, 757-763 (2007).

174. Bender, L. B., Cao, R., Zhang, Y. & Strome, S. The MES-2/MES-3/MES-6 complex and regulation of histone H3 methylation in C. elegans. *Curr. Biol.* **14**, 1639-1643 (2004).

175. Kelly, W. G. *et al.* X-chromosome silencing in the germline of C. elegans. *Development* **129**, 479-492 (2002).

176. Crittenden, S. L., Troemel, E. R., Evans, T. C. & Kimble, J. GLP-1 is localized to the mitotic region of the C. elegans germ line. *Development* **120**, 2901-2911 (1994).

177. Gaydos, L. J., Rechtsteiner, A., Egelhofer, T. A., Carroll, C. R. & Strome, S. Antagonism between MES-4 and Polycomb repressive complex 2 promotes appropriate gene expression in C. elegans germ cells. *Cell. Rep.* **2**, 1169-1177 (2012).

178. Kwan, K. Y., Sestan, N. & Anton, E. S. Transcriptional co-regulation of neuronal migration and laminar identity in the neocortex. *Development* **139**, 1535-1546 (2012).

179. Duncan, E. M. *et al.* Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. *Cell* **135**, 284-294 (2008).

180. Gerton, J. L. *et al.* Global mapping of meiotic recombination hotspots and coldspots in the yeast Saccharomyces cerevisiae. *Proc. Natl. Acad. Sci. U. S. A.* **97**, 11383-11390 (2000).

181. Cocquet, J. *et al.* The multicopy gene Sly represses the sex chromosomes in the male mouse germline after meiosis. *PLoS Biol.* **7**, e1000244 (2009).

182. Moore, K. E. *et al.* A general molecular affinity strategy for global detection and proteomic analysis of lysine methylation. *Mol. Cell* **50**, 444-456 (2013).

183. Roberts, C. T.,Jr & Morse, D. E. Galactokinase-deficient mutants of Tetrahymena thermophila: selection and characterization. *Mol. Gen. Genet.* **180**, 129-134 (1980).

184. Fass, J. N. *et al.* Genome-Scale Analysis of Programmed DNA Elimination Sites in Tetrahymena thermophila. *G3 (Bethesda)* **1**, 515-522 (2011).

COPYRIGHT LICENSES

SPRINGER LICENSE TERMS AND CONDITIONS

Feb 09, 2014

This is a License Agreement between Romeo Papazyan ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3314280636034
License date	Jan 22, 2014
Licensed content publisher	Springer
Licensed content publication	Springer eBook
Licensed content title	Separation and Purification of Multiply Acetylated Proteins Using Cation-Exchange Chromatography
Licensed content author	Romeo Papazyan
Licensed content date	Jan 1, 2013
Type of Use	Thesis/Dissertation
Portion	Full text
Number of copies	1
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	
Title of your thesis / dissertation	IDENTIFICATION AND CHARACTERIZATION OF HISTONE H3 LYSINE 23 METHYLATION IN GERMLINE CHROMATIN
Expected completion date	Feb 2014
Estimated size(pages)	200

SPRINGER LICENSE TERMS AND CONDITIONS

Feb 09, 2014

This is a License Agreement between Romeo Papazyan ("You") and Springer ("Springer") provided by Copyright Clearance Center ("CCC"). The license consists of your order details, the terms and conditions provided by Springer, and the payment terms and conditions.

All payments must be made in full to CCC. For payment instructions, please see information listed at the bottom of this form.

License Number	3314280679199
License date	Jan 22, 2014
Licensed content publisher	Springer
Licensed content publication	Springer eBook
Licensed content title	Separation and Purification of Multiply Acetylated Proteins Using Cation-Exchange Chromatography
Licensed content author	Romeo Papazyan
Licensed content date	Jan 1, 2013
Type of Use	Thesis/Dissertation
Portion	Figures
Author of this Springer article	Yes and you are a contributor of the new work
Order reference number	
Title of your thesis / dissertation	IDENTIFICATION AND CHARACTERIZATION OF HISTONE H3 LYSINE 23 METHYLATION IN GERMLINE CHROMATIN
Expected completion date	Feb 2014
Estimated size(pages)	200

CURRICULUM VITAE

Romeo Papazyan
rpapazy1@jhmi.edu

Educational History:

Ph.D. Expected	2014	*Department of Pharmacology and Molecular Sciences* Mentor: Sean D. Taverna Ph.D.	Johns Hopkins School of Medicine
B.S.	2008	Biology	University of California, Los Angeles

Other Professional Experience:

Undergradaute Research Assistant	02/2003-06/2008	Laboratory of Osvaldo Rey Ph.D., University of California, Los Angeles

Honors:

2013 Judge, Poster Contest, Biomed Central: Epigenetics and Chromatin Conference (Boston, MA)
2012 Scheinberg Travel Award
2010 First place poster, department retreat (Johns Hopkins University)
2005 Undergraduate Research Fellowship (University of California, Los Angeles)

Publications:

Papazyan R, Voronina E, Chapman JR, Gilbert TM, Meier E, Mackintosh SG, Shabanowitz J, Tackett AJ, Coyne RS, Hunt DF, Liu Y, and Taverna SD. (2013). "Loss of H3K23 methylation during meiosis promotes programmed DNA double-strand breaks in heterochromatin and reduces progeny viability". *Submitted*.

Papazyan R, Taverna SD. **Separation and purification of multiply acetylated proteins using cation exchange chromatography**. (2013) Methods Mol Biol. 2013 981; 103-113. PMID:23381856

Dancy BC, Ming SA, **Papazyan R,** Jelinek CA, Majumdar A, Sun Y, Dancy BM, Drury WJ, Cotter RJ, Taverna SD, Cole PA. **Azalysine analogues as probes for protein lysine deacetylation and demethylation**. (2012) J Am Chem Soc. 2012 Mar 21; 134(11):5138-5148. Epub 2012 Mar 12. PMID:22352831

Coyne RS, Orias E, Taverna SD, **Papazyan R. Tetrahymena Comparative Sequencing Project, Broad Institute of Harvard and MIT**. (2011) http://www.broadinstitute.org/.

Papazyan R, Doche M, Waldron RT, Rozengurt ER, Rey O. **Protein Kinase D isozymes activation and localization during mitosis**". (2008) Exp Cell Res. 2008 Oct 1;314(16):3057-68. Epub 2008 jul 26. PMID:18692497

Papazyan R, Rozengurt E, Rey O. **The C-terminal tail of protein kinase D2 and protein kinase D3 regulates their intracellular distribution**. (2006) Biochem Biophys Res Commun. 2006 Apr 14;342(3):685-9. Epub 2006 Feb 13. PMID:16494840

Rey O, **Papazyan R,** Waldron RT, Young SH, Lippincott-Schwartz J, Jacamo R, Rozengurt E. **The nuclear import of protein kinase D3 requires its catalytic activity**. (2006) J Biol Chem. 2006 Feb 24;281(8):5149-57. Epub 2005 Dec 27. PMID:16380377

Rey O, Young SH, **Papazyan R,** Shapiro MS, Rozengurt E. **Requirement of the TRPC1 cation channel in the generation of transient Ca2+ oscillations by the calcium- sensing receptor**. (2006) J Biol Chem. 2006 Dec 15;281(50):38730-7. Epub 2006 Oct 17. PMID:17046820

Posters:

Papazyan R, Voronina E, Chapman JR, Gilbert TM, Meier E, Shabanowitz J, Hunt DF, Liu Y, Taverna SD. (2012) **Methylation of histone H3 at lysine 23 insulates meiotic heterochromatin from programmed DNA double-strand breaks**. TCNP symposium, JHU, Baltimore, Maryland

Papazyan R, Sinnett-Smith J, Waldron RT, Rozengurt E, Rey O. **Protein kinase D3 localizes to mitotic spindles and centrosomes**. (2007) CURE Digestive Disease meeting, UCLA, Los Angeles, California

Papazyan R, Rozengurt E, Rey O. **The C-terminal tail of protein kinase D2 and protein kinase D3 regulates their intracellular distribution**. (2006) CURE Digestive Disease meeting, UCLA, Los Angeles, California

Talks:

Papazyan R, Voronina E, Chapman JR, Gilbert TM, Meier E, Shabanowitz J, Hunt DF, Liu Y, Taverna SD. **Methylation of histone H3 at lysine 23 in meiotic heterochromatin**. (2013) Epigenetics and Chromatin: Interactions and processes, Harvard School of Medicine, Boston, Massachusetts