

Study of dynamics of histone H3 variants and H3 proteolysis during mouse spermatogenesis

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch–Naturwissenschaftlichen Fakultät

der Universität Basel

von

Ching-Yeu Liang

aus Taiwan

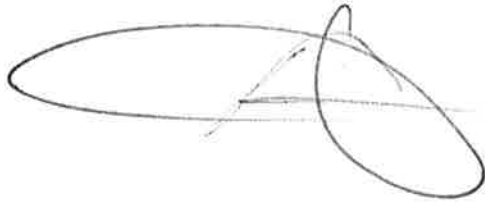
Basel, 2016

Original document stored on the publication server of the University of Basel

edoc.unibas.ch

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät
auf Antrag von

Prof. Dr. Antoine Peters
(Fakultätsverantwortlicher und Referent)

A handwritten signature in black ink, consisting of several overlapping loops and a horizontal line, positioned below the name of Prof. Dr. Antoine Peters.

Prof. Dr. Anja Groth
(Koreferent)

Basel, der 23.Juni.2015

Prof. Jörg Schibler
The Dean of Faculty

Table of Contents

Summary	1
Chapter 1. Introduction and scope of thesis	3
1.1 Epigenetics	3
1.1.1 DNA methylation	4
1.1.2 Histone modification	7
1.1.3 Histone variants	12
1.2 Spermatogenesis	17
1.2.1 Primordial germ cells and spermatogonia	18
1.2.2 Meiotic phase	19
1.2.3 Spermiogenesis	20
1.2.4 Testis-specific histone variants	22
1.2.5 Nucleosome retention	24
1.3 Scope of thesis	25
Chapter 2. Published manuscript	27
PRC1 coordinates timing of sexual differentiation of female primordial germ cells	
Chapter 3. Published manuscript	35
Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa	
Chapter 4. manuscript in preparation	47
Proteolytic cleavage of histone H3 by a serine protease during mouse spermiogenesis	
Chapter 5. manuscript in preparation	68
Characterization of expression of histone H3 variant proteins during spermatogenesis suggests variant specific roles in chromatin reprogramming	
Chapter 6. Discussion and outlook	96
6.1 The dynamics of H3 variants during mouse spermatogenesis	96
6.2 The proteolytic cleavage during spermiogenesis	99
Reference	102
Acknowledgement	114
Curriculum vitae	115

Summary

Spermatogenesis is a complex differentiation process in which male gametes, known as spermatozoa, are produced from spermatogonial stem cells in the seminiferous tubules of the testis. The spermatogenesis process is typically divided into three phases: a mitotic phase, a meiotic phase and post-meiotic spermiogenesis. During mammalian spermiogenesis, haploid round spermatids undergo remarkable morphological changes and an extensive reorganization of chromatin to differentiate into mature spermatozoa. As part of the chromatin reorganization, most histones in round spermatids are replaced by transition proteins and subsequently by protamines. This histone-to-protamine exchange is required for efficient compaction of paternal genome into the sperm head and implicated in male fertility. Nonetheless, previous studies found that 1-10 % histones are still retained at specific loci, particularly at unmethylated CpG-rich promoters, in mouse and human sperm. How spermatid chromatin is reorganized genome-wide during spermiogenesis while some loci are exempted from histone eviction is still elusive.

Our previous study has shown that the residual nucleosomes in mouse sperm largely contain the histone H3 variant, H3.3. The study also revealed differential histone turnover of canonical and variant H3 in round spermatids, which may underlie the final histone composition in mature sperm. In order to determine the dynamics of H3 variants during mouse spermatogenesis, I analyze protein expression of canonical and variant H3 proteins at different stages of male germ cells by triton-acetic acid-urea gel-Western blotting. Surprisingly, I find that mouse testis-specific H3 variant (H3t), not canonical H3, is the most abundant H3 protein from spermatogonia to spermatids and that most canonical H3 is replaced by H3.3 during meiosis. I further observe that a relatively large portion of H3t is removed from chromatin during the process of histone-to-protamine exchange compared with H3.3, which is consistent with that H3.3 is the predominant H3 in residual sperm nucleosomes. Taken together, the first part of my thesis reveals important findings on chromatin composition and dynamics of histone H3 variants during mouse spermatogenesis.

In the second part of my thesis, I describe the discovery that histone H3 is cleaved at its N-terminal tail by a serine protease activity in nuclei of the late-stage mouse spermatids. Arginine 26 and lysine 27 on H3 are important to the H3 protease activity. This proteolytic

cleavage of H3 tail may result in nucleosome destabilization and then contribute to nucleosome eviction during spermiogenesis. Interestingly, I find that the acetylation on H3 can prevent H3 from proteolytic cleavage *in vitro* and that the genome-wide distribution of H3 lysine 27 acetylation (H3K27Ac) is positively correlated to the occupancy of nucleosomes containing transcriptionally active mark in sperm, suggesting that the inhibition of H3 cleavage by acetylated lysine 27 in late-stage spermatids may lead to the nucleosome retention at specific loci in mature sperm. Overall, these findings provide novel insights into the mechanism of nucleosome eviction and retention during spermiogenesis through the regulation of H3 proteolytic cleavage.

Chapter 1: Introduction

1.1 Definition of epigenetics

Deoxyribonucleic acid (DNA) is the macromolecule that consists of a large number of linked nucleotides and is present in nearly all living organisms. DNA stores genetic information that is arranged in hereditary units, called genes, and is required to build and maintain the cells and tissues of an organism. In order to decode genetic information to control cells, DNA is transcribed into messenger ribonucleic acid (mRNA) by RNA polymerases, and then mRNA carries the instructions from DNA to ribosomes to guide protein synthesis, known as translation. Proteins translated from mRNA are regarded as the end products of gene expression and perform most of the functions of cells. Therefore, the two-step process, transcription and translation, is well-documented as the central dogma of molecular biology.

The genetic information encoding DNA can be transmitted from parent cell to daughter cells through semi-conservative replication followed by cell division. Thus, a daughter cell is genetically identical to parent cell. However, while being exposed to environmental changes or stress, cells can modify their gene expression profile that changes cellular and physiological trait to response external environmental effects without alterations in the DNA sequence. Moreover, multicellular organisms consist of many distinct types of cells that are differentiated from a fertilized zygote, the earliest developmental stage of the embryo. Despite containing identical genomic DNA inherited from zygote, each cell type has a unique gene expression profile to maintain its biological function and structure in organisms. To dissect the molecular mechanism by which a cell changes its phenotype without a change in genotype either during development or during environmental stimulation, a relatively new science called epigenetics is quickly growing and aroused wide-spread interest over the last decade.

The term epigenetics was coined by Conrad Waddington in the early 1940s. He proposed the model of “epigenetic landscape” to describe the influence of genetic processes during development and defined epigenetics as “the branch of biology which studies the causal interactions between genes and their products which bring the phenotype into being” (Goldberg et al. 2007). Over the following years, the definition of epigenetics has evolved with the increasing knowledge of genetics. Today the term of epigenetics has been generally

accepted as meaning “the study of changes in gene function that are mitotically and/or meiotically heritable and that do not entail a change in DNA sequence” (Wu and Morris 2001). Nowadays, epigenetics is emerging as an important field associated with the studies of developmental biology, stem cell biology and diseases, such as oncogenesis. The molecular mechanisms of epigenetic regulation described in current literatures include DNA methylation, post-translational histone modification, exchange of histone variants and chromatin architectures and non-coding RNAs. These epigenetic mechanisms construct a regulation network to contribute to proper gene regulation. In sections below, I will introduce DNA methylation, histone modifications and histone variants.

1.1.1 DNA methylation

DNA methylation is the “oldest” epigenetic modification known to regulate gene expression. In prokaryotes, DNA methylation is restricted to adenine and cytosine residues (Marinus 1987). In eukaryotes, DNA methylation occurs at the fifth position of the pyrimidine ring of cytosine bases. 5-methylcytosine (5mC) is primarily restricted to the context of CpG dinucleotides, whereas some non-CpG methylation is also found in mammals (Woodcock et al. 1987; Ramsahoye et al. 2000; Lister et al. 2009). 70 to 80 % of cytosines at CpG sites in mammalian somatic tissues are methylated (Ehrlich et al. 1982; Jabbari and Bernardi 2004). Most of these methylated regions are distributed at repetitive genomic elements, such as satellite sequences, centromeric repeats, transposons, parasitic elements and endogenous retroviruses. Therefore, DNA methylation is considered as a host defense mechanism to suppress the expression of repetitive elements and endogenous retrovirus genes (Slotkin and Martienssen 2007).

Moreover, approximately 10 % CpG dinucleotides in the genome are clustered together in 0.5-5 kb long stretches of DNA, called CpG islands (CGI). CpG islands are often found at the promoter regions or with the first exon of expressed genes and generally lack DNA methylation for the expression of most housekeeping genes and many regulated gene. Approximately 70 % promoters of human genes are associated with CpG islands (Saxonov et al. 2006). It is well-known that hypermethylation of CGI promoters leads to stable transcriptional repression and gene silencing. Because the methyl group on cytosine is situated in the major groove of the DNA helix where many proteins, like transcription

factors, interact with DNA, DNA methylation on CGI can modulate the binding of transcription-relative proteins to promoters (Watt and Molloy 1988; Schubeler 2015). In addition, a family of proteins, known as methyl-CpG-binding proteins, has been identified to bind on methylated CpG sites and then recruit repressor complexes to alter chromatin structure and finally contribute to transcription silencing (Bogdanovic and Veenstra 2009; Baubec and Schubeler 2014). Therefore, DNA methylation and gene expression are inversely correlative.

The pattern of DNA methylation is stable and heritable by daughter cells through mitosis. DNA methyltransferase 1 (Dnmt1) activity is required for maintaining global DNA methylation during DNA replication. Dnmt1 expression is regulated by cell cycle-dependent transcription factors and increases at S phase of cell cycle (Kishikawa et al. 2003). During DNA replication, Dnmt1 is recruited by PCNA and Np95 protein to replication forks and methylates newly synthesized DNA strand based on DNA methylation pattern on parental DNA strand (Chuang et al. 1997; Bostick et al. 2007; Sharif et al. 2007). The deletion of Dnmt1 leads to mouse embryonic lethality at 10.5 days post coitum and significant loss of global DNA methylation (Li et al. 1992), demonstrating that the maintenance of DNA methylation by Dnmt1 is essential for normal mammalian development.

Many studies have shown that the other DNA methylation process, called *de novo* DNA methylation, is present in mammals. During preimplantation development, maternal and paternal genomes undergo a wave of DNA demethylation, at which most of the methylation patterns inherited from the gametes are removed. This loss of DNA methylation is reversed by *de novo* DNA methylation after implantation (Monk et al. 1987; Howlett and Reik 1991; Kafri et al. 1992). In addition, *de novo* DNA methylation is also present in both male and female germ cells during gametogenesis and plays an important role in the establishment of genomic imprinting in germ cells (Smallwood and Kelsey 2012). Two DNA methyltransferases, Dnmt3a and Dnmt3b, are mainly responsible for *de novo* DNA methylation process (Okano et al. 1998) and serve partially redundant function in the establishment of DNA methylation pattern (Okano et al. 1999). But the expression profiles of both *de novo* methyltransferase are quite different during embryonic development (Watanabe et al. 2002), so Dnmt3a mutation and Dnmt3b mutation cause different developmental defects (Okano et al. 1999). Furthermore, the conditional knockout mice of

Dnmt3a and Dnmt3b show that Dnmt3a is required for DNA methylation at most imprinting control regions (ICRs) in the germ cells, while Dnmt3b contributes to methylation of some ICRs in the germs cells (Kaneda et al. 2004). Therefore, Dnmt3a and Dnmt3b have overlapping and distinct functions for DNA methylation depending on the stage of development and the cell type.

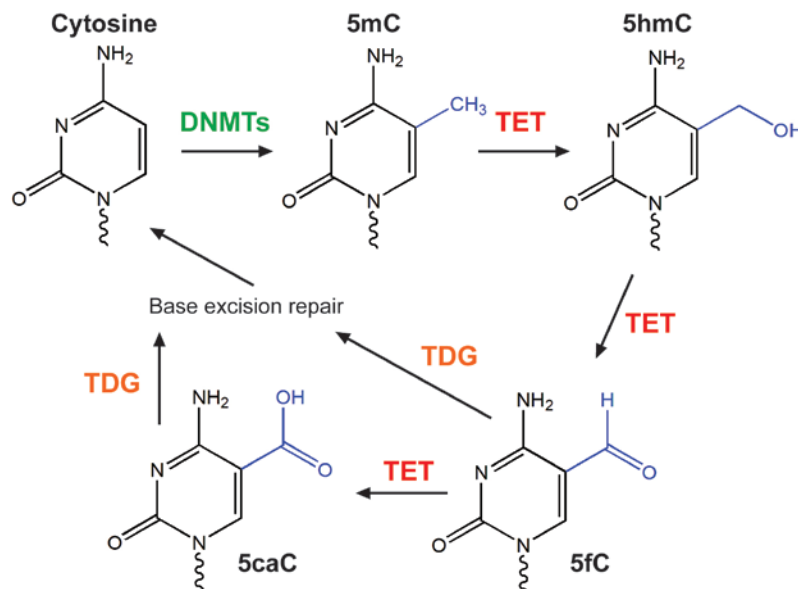


Figure 1. The cycle of DNA methylation and demethylation. DNA methyltransferases (DNMTs) methylate cytosine within the context of a CpG dinucleotide to yield 5-methylcytosine (5mC) through replication-dependent or independent manner. One potential mechanism of DNA demethylation implicate that Ten-eleven translocation (TET) enzymes oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC). 5fC and 5caC can be excised by thymine DNA glycosylase (TDG) and replaced by cytosine through base excision repair (BER).

DNA demethylation in animals can occur through either a passive or an active mechanism. Passive DNA demethylation is replication-coupled and thought to take part in the absence or the prevention of Dnmt1 activity during replication. Without Dnmt1 activity, newly-synthesized DNA strands cannot maintain the DNA methylation pattern from parental strands, and thereby DNA methylation level is reduced after several rounds of replication. By contrast, the active mechanism of DNA demethylation involves enzymes that directly modify 5mC (Figure 1). First TET (ten-eleven translocation) dioxygenase enzymes sequentially oxidize 5mC to 5-hydroxymethylcytosine (5hmC), 5-formylcytosine (5fC) and 5-carboxylcytosine (5caC) (Iyer et al. 2009; Tahiliani et al. 2009; Ito et al. 2011). 5fC and 5caC are removed by thymine DNA glycosylase (TDG) and then replaced by cytosine through BER (base excision repair) pathway (He et al. 2011; Zhang et al. 2012). During the

reprogramming in early embryonic development, mass DNA demethylation is modulated by TET3 and subsequently by replication-dependent passive demethylation (Gu et al. 2011; Guo et al. 2014). Other TET proteins, TET1 and TET2, mediate global DNA demethylation during primordial germ cell development (Hackett et al. 2013; Vincent et al. 2013).

1.1.2 Histone modifications

Genomic DNA in eukaryotic cells is wrapped around histones to form nucleosomes that is used to compact large genomic DNA into the nucleus. Each nucleosome core particle is consisting of approximately 147 base pairs DNA in 1.7 superhelical turns wrapped around a histone octamer containing two copies each of four core histones H2A, H2B, H3, and H4 (Luger et al. 1997; Davey et al. 2002). Histones have a large proportion of positively charged amino acids that neutralize negatively charged DNA backbone to form higher-order chromatin structures. In addition to the core histones, there is a linker histone, called H1, which interacts with linker DNA region (approximately 20-80 nucleotides in length) between nucleosomes to stabilize the chromatin fiber (Thoma et al. 1979). Although DNA is packaged tightly and orderly by nucleosomes, nucleosomes still allowing other proteins to access DNA for biological processes such as transcription, DNA replication and DNA repair. The properties of nucleosomes can be modulated in different ways, including the covalent modification of histones, the replacement of canonical histones to variants and the nucleosome reposition by ATP-dependent chromatin remodeling complexes.

The flexible N-terminal or C-terminal tail of four core histones is extended out from the face of nucleosome (Figure 2). These histone tails are known to contain many positively-charged amino acids that electrostatically interact with negatively-charged phosphate groups along nucleosomal DNA, linker DNA and the acidic patches of the neighboring nucleosomes for the formation of higher-order chromatin structure (Mutskov et al. 1998; Angelov et al. 2001; Davey et al. 2002; Dorigo et al. 2004). Based on *in vitro* studies, the deletion of histone tails alters nucleosome structure, reduces nucleosome stabilization and increases the accessibility of nucleosomal DNA (Ferreira et al. 2007; Biswas et al. 2011; Iwasaki et al. 2013). Therefore, histone tails play a crucial role in maintaining nucleosome structure and dynamics. In addition, histone proteins, especially histone tails, are subject to large numbers and different type of post-translational modifications, such as acetylation, methylation,

phosphorylation, ubiquitination and so on (Figure 2) (Bannister and Kouzarides 2011; Zentner and Henikoff 2013). These histone modifications are involved in various DNA processes including replication, repair and transcription by either directly altering nucleosome structure or providing docking platforms for activators, repressors and chromatin remodeling complexes. With the advent of next-generation sequencing technology, the genome-wide distributions of histone modifications and DNA-binding proteins are more easily characterized by chromatin immunoprecipitation sequencing (ChIP-seq) analysis (Zentner and Henikoff 2014). Also, a growing number of histone modifications such as lysine crotonylation and 2-hydroxyisobutyrylation have been identified through the advanced proteomic technologies (Tan et al. 2011; Arnaudo and Garcia 2013; Dai et al. 2014).

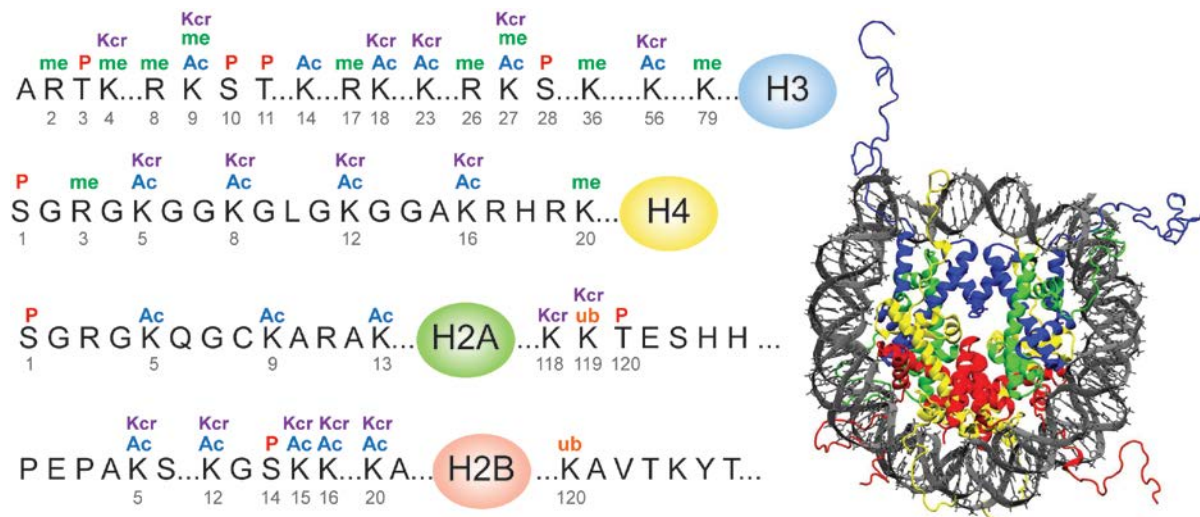


Figure 2. Nucleosome structure and the summary of post-translational modifications identified in core histones. A nucleosome consists of 147 base pairs of DNA wrapping around two copies of each core histones, H3, H4, H2A and H2B (Davey et al. 2002). The tails of core histones protrude from the nucleosome and are covalently modified at several residues by distinct modification, such as acetylation (Ac), methylation (me), phosphorylation (P), ubiquitination (Ub) and crotonylation (Kcr). Particularly, lysine is able to be mono-, di-, or trimethylated, and arginine is able to be mono- or dimethylated *in vivo*.

Histone acetylation

Acetylation is the first described histone modification on lysine residues and linked to transcriptional activation. It is catalyzed by histone acetyltransferases (HATs) with acetyl-CoA as acetyl group donor and is erased by histone deacetylases (HDACs). Acetylation neutralizes the positive charge of lysine to loosen the charge-dependent interaction between histone and DNA and thereby open the chromatin structure to allow transcription

machinery access to DNA. Therefore, histone hyperacetylation is a feature of active promoters and enhancers for transcription initiation. In addition, histone acetylation is also associated with other cellular processes that require DNA access. For example, both histones H4 and H2A show increased levels of acetylation following DNA double-strand break generation to relax histone-DNA contacts for DNA repair factors binding (Bird et al. 2002; Murr et al. 2006; Jiang et al. 2010). And H3 and H4 acetylation are also considered to play an important role in facilitating firing of replication origins before DNA replication (Unnikrishnan et al. 2010). Therefore, the charge neutralization of histone lysine by acetylation is necessary not only for transcription activation but also for efficient DNA replication and repair.

In addition to influence nucleosome structure directly, acetylated lysine residues on histones are recognized by a specific protein domain, the bromodomain, which is found in numerous chromatin-associated proteins including transcription factors, transcription initiation factors, chromatin remodeling factors and acetyltransferases (Zeng and Zhou 2002). Aberrant expression and genetic rearrangements of bromodomain-containing proteins have been implicated in a wide range of human diseases such as cancer, inflammation and neurodegenerative diseases (Muller et al. 2011). Recently, bromodomain-containing proteins have emerged as therapeutic targets in a remarkable range of disease models (Shi and Vakoc 2014).

Histone methylation

Histones can be methylated on the side chains of both lysine and arginine residues with the potential additional of one, two or three methyl groups, and methylation is commonly found on histone H3 and H4. Unlike acetylation and phosphorylation, histone methylation does not alter the charge of lysine and arginine, and so methylation has less effect in directly modulating nucleosome dynamics. But methylation can recruit diverse chromatin effector molecules to regulate chromatin and transcription states (Taverna et al. 2007). So far, there are three families of histone methyltransferases that catalyze the addition of methyl groups donated from S-adenosyl methionine (SAM) to histones. The SET-domain containing proteins and DOT1-like proteins have been shown to methylate lysine residues (Rea et al. 2000; Feng et al. 2002), and the protein arginine N-methyltransferase (PRMT)

family has been shown to methylate arginine residues (Bedford and Clarke 2009). Lysine methylation on histone can be removed by lysine-specific demethylases and Jumonji histone demethylases (Shi et al. 2004; Tsukada et al. 2006). Unlike HAT and HDAC, histone methyltransferases (HMTs) and histone demethylases (HDMs) have very restricted substrate specificities, so generally distinct residues on histones are methylated or demethylated by distinct HMTs or HDMs.

Furthermore, methylations on different histone residues contribute to distinct cellular functions as well as different degrees of residue methylation (mono-, di- and trimethylation). For example, trimethylation at H3 lysine 4 (H3K4me₃) is linked to transcription activation and is enriched at the promoter regions and around transcription start site (TSS) to modulate transcription initiation (Ng et al. 2003; Bernstein et al. 2005; Lauberth et al. 2013), but monomethylated H3 lysine 4 (H3K4me₁) generally marks enhancer regions (Heintzman et al. 2007). Trimethylation at H3 lysine 9 and 27 (H3K9me₃ and H3K27me₃) are associated with transcriptional repression and heterochromatin formation. H3K9me₃ shows a relatively homogenous distribution on inactive regions, whereas H3K27me₃ is enriched around the inactive gene promoters (Kooistra and Helin 2012). Trimethylation at H3 lysine 36 (H3K36me₃) is generally found in downstream of the TSS in the gene body and is involved in suppressing cryptic transcription in gene bodies (Carrozza et al. 2005; Keogh et al. 2005). Therefore, the location of the methyl residue on histone and the degree of methylation cause the different function and distribution of histone methylation marks.

Methylation can be present at multiple lysine and arginine residues on the same histone. However, some histone methylations have mutually antagonistic relationship. For instance, dimethylation on H3R2 prevents the methylation of H3K4 (Guccione et al. 2007; Kirmizis et al. 2007). In addition, H3K4me₃ and H3K4me₂ marks can recruit histone lysine demethylases, PHF8 and KDM7A, to transcription initiation sites to remove repressive marks, such as methylation on H3K9 and H3K27 residues (Horton et al. 2010). Thus H3K9me₃ and H3K27me₃ are excluded from active promoter regions. Interestingly, although H3K4me₃ and H3K27me₃ are usually enriched at active and inactive gene promoters, they are found to colocalize in some genomic regions, termed bivalent domains, in embryonic stem cells. These bivalent domains are thought to play a role in keeping pluripotency by poising developmental regulatory genes for either activation or repression during embryonic

differentiation (Azuara et al. 2006; Bernstein et al. 2006). Mll2, a Set1/Trithorax-type H3K4 methyltransferases, is mainly responsible for H3K4 trimethylation at bivalent domain in embryonic stem cells (Denissov et al. 2014).

Histone phosphorylation

Because phosphate group carries negative charges, histone phosphorylation can modulates the nucleosome dynamics through altering charge-dependent DNA-histone interaction as acetylation. The addition of phosphates on histones creates charge repulsion between the highly negatively charged DNA backbone and histone, potentially weakening the association of DNA and histones (Banerjee and Chakravarti 2011). Therefore, histone phosphorylation functions in various cellular processes such as DNA damage response, transcription regulation and chromatin compaction (Rossetto et al. 2012). Phosphorylation of histone H2A variant H2A.X on serine 139 in mammals, termed to γ H2A.X, has been well-known as a critical marker for DNA damage response (Rogakou et al. 1998). When DNA double-strand break occurs, protein kinases ATM and ATR carry out H2A.X phosphorylation (Burma et al. 2001; Ward and Chen 2001), and γ H2A.X spreads bidirectionally over several kilobases to megabases on each side of the DNA break (Rogakou et al. 1999; Iacovoni et al. 2010). This wide distribution of γ H2A.X increases the DNA accessibility and also provides a binding platform for DNA repair factors and chromatin remodeling complexes to alter chromatin structure and repair DNA breaks. For example, MDC1, a repair mediator, can interact with γ H2A.X through recognition by its BRCT domain and then serves as a scaffold for recruiting more DNA repair-related proteins (Rogakou et al. 1999; Jungmichel and Stucki 2010).

Moreover, it has been reported that histone H3 phosphorylation occurs during mitosis and meiosis. Threonine 3, serine 10, threonine 11 and serine 28 on H3 tails are phosphorylated by distinct protein kinases at prophase of mitosis, and these phosphorylation levels decline at anaphase. A cascade of mitotic histone phosphorylation is associated with the processes of chromatin condensation and kinetochore assembly (Sawicka and Seiser 2012). In particular, H3 serine 10 is adjacent to H3 lysine 9 that was shown to be methylated to recruit HP1 proteins for heterochromatin formation in silence regions. Interestingly, the phosphorylation on H3 serine 10 (H3S10ph) during mitosis was demonstrated to promote the ejection of HP1 proteins bound to the adjacently methylated H3K9 (Fischle et al. 2005;

Hirota et al. 2005). This suggests that H3S10ph influences chromatin structure during mitosis through the regulation of protein binding to chromatin.

Histone ubiquitylation

Except acetylation and methylation, lysine residues can also be modified by ubiquitylation. Unlike other modifications consisting of small chemical group, ubiquitylation is the addition of 76 amino-acid protein ubiquitin to the free amino group of lysine residues to form a branched molecule. Histone ubiquitylation is generally in the form of monoubiquitination that is not relevant to protein degradation as polyubiquitination. Histone H2A and H2B are two of the most abundant ubiquitinated proteins in the nucleus, and approximately 5-15 % of H2A and 1-2% of H2B are modified by ubiquitylation (Cao and Yan 2012). However, both ubiquitinated core histones play distinct functions in transcriptional regulation. Monoubiquitination on H2A lysine 119 (H2AK119ub, uH2A) catalyzed by Ring1A, Ring1B and Bmi1 in polycomb repressive complex 1 is involved in Polycomb-mediated transcriptional repression by restraining RNA pol II from elongation (Wang et al. 2004; Cao et al. 2005; Stock et al. 2007; Zhou et al. 2008). In addition, during DNA damage response, H2A and H2A.X around DNA break sites are ubiquitinated by RNF8 and RNF168 to recruit checkpoint and repair proteins (Mattioli et al. 2012; Panier and Durocher 2013). By contrast, monoubiquitination on H2B lysine 120 in human or lysine 123 in yeast occupies at the gene body of transcriptional active genes to promote transcriptional elongation (Minsky et al. 2008). It is required for establishments of H3K4 methylation by COMPASS and H3K79 methylation by Dot1 during transcription (Sun and Allis 2002; Lee et al. 2007).

1.1.3 Histone variants

In metazoans most nucleosomes are assembled by canonical histone proteins (H3, H4, H2A and H2B). It is well-documented that multiple copies of genes encoding canonical histone proteins are organized as clusters in the genome, and there are approximately 75 distinct canonical histone mRNAs in mammals (Marzluff et al. 2002). Moreover, the transcription of canonical histone genes is tightly coupled to DNA replication. Their mRNAs are highly synthesized at the beginning of S phase of mitosis and then are rapidly degraded at the end of S phase (Marzluff et al. 2008). Thus, canonical histones constitute the main histone supply during DNA replication. The 3' end of canonical histone mRNAs contain a stem-loop

sequence, instead of polyadenylation, that is recognized by a stem-loop binding protein (SLBP) (Battle and Doudna 2001). This interaction between histone mRNA and SLBP is required for histone pre-mRNA processing, histone mRNA stabilization, transport and efficient translation at S phase (Marzluff et al. 2008).

In addition to the canonical histones, replacement variant histones have been described in each histone protein except histone H4 (Maze et al. 2014; Henikoff and Smith 2015). Unlike canonical histone, variant histones are typically encoded by single gene or low copy number of gene in genome. Variant histone mRNAs usually are polyadenylated without stem-loop sequence, and most their synthesis are DNA replication-independent (Wells and Kedes 1985; Ivanova et al. 1994). Moreover, variant histones differ from canonical histone either by the alternation of a few amino acids or by the addition of domain. These differences result in that variant histones have specific incorporation manners, influence post-translational modification and also alter the physical properties of the nucleosome (Maze et al. 2014). Therefore, the substitution of canonical histones by variants has been implicated in many biological processes such as transcription and DNA repair.

Table 1. The summary of histone H3 proteins

Histone H3	name	Expression/ deposition	Chaperones	Functions
Canonical	H3.1	RD	CAF1	Replication and repair
	H3.2			
Variants	H3.3	RI	HIRA/ATRX/DAXX	Transcription activation, heterochromatin maintenance, ERVs silence and MSCI
	CenH3	RI	HJURP/DAXX	Chromosome segregation
Testis-specific variants	H3t	ND	ND	ND
	H3.X			
	H3.Y			
	H3.5			

RD, replication dependent; RI, replication independent; ND, not determined; EVR: endogenous retroviral elements; MSCI: meiotic sex chromosome inactivation.

Canonical H3

In eukaryotes except *S. cerevisiae*, canonical H3 is composed of two H3 proteins, H3.1 and H3.2. The protein sequences of both canonical H3 are almost identical and differ in only one amino acid. They are highly expressed at S phase to provide the main supply for chromatin assembly during DNA replication. Chaperone CAF-1 is well-known to mediate the deposition of canonical H3-H4 into the replication fork (Verreault et al. 1996; Tagami et al. 2004).

During replication, parental nucleosomes ahead of the DNA replication fork are disassembled via Asf1 and FACT chaperon complexes with the MCM2-7 helicase. Asf1 transfers the parental or new synthesized H3-H4 dimer to CAF-1. Subsequently CAF1 is recruited by PCNA and deposits two canonical H3-H4 dimers or a (H3-H4)₂ tetramer into DNA at the replication fork. After (H3-H4)₂ tetramer deposition, two H2A-H2B dimers are added to form a complete nucleosome (Alabert and Groth 2012).

In addition, although both canonical H3 share expression and incorporation manner, the post-translational modifications on H3.1 and H3.2 are slightly different. H3.2 contains more methylated H3K27, a repressive mark, than H3.1 in human cell lines, but H3.1 is enriched for repressive mark (methylated H3K9) as well as active mark (acetylated H3K14) (Hake et al. 2006). Thus, a single amino acid exchange on canonical H3 may influence the genomic localization and then result in different modifications on two canonical H3 proteins.

H3.3

The replacement variants of H3 best characterized in mammals are H3.3 and centromeric H3 (CenH3) in eukaryotes. H3.3 is encoded by two distinct intron-containing genes, *H3f3a* and *H3f3b*, and is synthesized throughout the cell cycle. H3.3 and canonical H3 have only 4-5 amino acid differences that do not affect fundamental nucleosome structure (Tachiwana et al. 2011). But these different amino acids on H3.3 recruit H3.3-specific histone chaperones to direct H3.3 localization to specific loci in the genome. Two H3.3-specific histone chaperones, the HIRA complex and the DAXX/ATRX complex, have been identified to facilitate H3.3 deposition by replication-independent manners (Tagami et al. 2004; Lewis et al. 2010). Based on genome-wide studies, it is well-known that H3.3 is generally enriched at active promoters, gene bodies and enhancers (Mito et al. 2005; Wirbelauer et al. 2005; Jin et al. 2009) and is decorated with active marks of transcription, such as H3K4me₃, H3K36me₃, H3K9ac and H3K27ac (McKittrick et al. 2004; Hake et al. 2006). H3.3-containing nucleosomes with active marks at promoters and enhancers undergo rapid turnover, which may allow for the accessibility of the transcriptional machinery as well as transcription factors (Kraushaar et al. 2013). HIRA chaperone complex is responsible for the H3.3 deposition into genic, euchromatic regions during transcription (Ray-Gallet et al. 2002; Tagami et al. 2004). However, in embryonic stem cells, H3.3 occupancy is also observed at

the bivalent promoters of developmentally regulated genes, which are poised for activation upon differentiation (Goldberg et al. 2010), and H3.3-HIRA can recruit polycomb repressive complex 2 (PRC2) to establish H3K27me3 mark on bivalent promoters of developmentally regulated gene (Banaszynski et al. 2013). In addition, H3.3 deposition is not always linked to transcriptional activation. Recent studies have shown that H3.3 also can be incorporated into silent pericentric heterochromatin, telomeres and endogenous retroviral elements (ERVs) by Daxx/ATRX chaperone complex, instead of HIRA (Goldberg et al. 2010; Lewis et al. 2010; Dhayalan et al. 2011; Elsasser et al. 2015). H3.3 depletion leads to telomere-dysfunction, abnormal karyotype, the reduction of H3K9me3 on ERVs and the upregulation of ERVs expression (Wong et al. 2009; Elsasser et al. 2015). Therefore, H3.3 also plays an important role in the establishment of silenced chromatin states and in maintenance of genome stability.

Although *H3f3a* and *H3f3b* encode the same protein sequence of H3.3, they have different expression patterns and function during mouse development. The *H3f3b* homozygous mutation causes lethality at birth and growth-deficiency in mouse embryos (Bush et al. 2013; Tang et al. 2015), whereas the mice of *H3f3a* homozygous mutation are viable. Furthermore, the male mice of *H3f3b* heterozygous mutation is sterile, but *H3f3a* mutation male mice are fertile (Tang et al. 2015). Therefore, H3.3 from *H3f3b* is relatively more important than H3.3 from *H3f3a* during embryonic development and spermatogenesis.

CenH3

CenH3 (CENP-A in mammals) is other well-characterized replacement H3 variant in eukaryotes and is specifically localizes at centromeres, which link sister chromatids and serve as the attachment sites for the spindle microtubules during mitosis. CenH3 deposition plays an important for maintaining centromere structure, the formation of kinetochores and proper chromosome segregation during mitosis. The deletion of CenH3 results in chromosome missegregation defect and mitotic defect (Howman et al. 2000; Zeitlin et al. 2001; Regnier et al. 2005). In mammals, CenH3 deposition is replication-independent. CenH3 is highly expressed at the late G2 phase of mitosis and is deposited into centromeres during telophase and early G1 phase (Boyarchuk et al. 2011). Holliday Junction-Recognizing protein (HJURP) has been identified to be a CenH3-specific chaperon (Dunleavy et al. 2009;

Foltz et al. 2009; Shuaib et al. 2010). Moreover, CenH3 only has 50 to 60 % identity to canonical H3 at histone fold domain and contains a unique N-terminal tail, so CenH3-containing nucleosome shows the distinct biophysical properties compared with canonical nucleosome. CenH3-histone H4 tetramer are more compact and structurally more rigid than the canonical H3- H4 tetramer (Black et al. 2004). In addition to form octamer nucleosomes, CenH3 particles have been shown to exist as hemisomes, hexasomes and tetramers *in vitro* and *in vivo* in (Dalal et al. 2007; Mizuguchi et al. 2007; Furuyama and Henikoff 2009; Williams et al. 2009; Bui et al. 2012; Shivaraju et al. 2012).

Other H3 variants

Except H3.3 and CenH3, there are four other H3 variants, H3.X, H3.Y, H3t and H3.5, identified in human tissues before. H3.X (also known as H3.Y.2) and H3.Y (also known as H3.Y.1) are primate-specific histone H3 whose expressions are found in human brain, testis, certain tumor tissues and cancer cell lines (Wiedemann et al. 2010). The identity of protein sequence between H3.X and H3.Y is 89.7 %, but H3.X contains a specific long C-terminal tail that does not exist in other H3 proteins. Nutritional- and growth-associated stress stimuli increase the number of H3.Y-expressing cells, and the knockdown of H3.Y influences cell growth and downregulates the expression of genes involved in cell cycle control in U2OS (Wiedemann et al. 2010). But the incorporation mechanism, localization and detail function of H3.X and H3.Y are still unclear.

H3t (also known as H3.1t, TH3 and H3.4) and H3.5 (also known as H3.3C) have been identified in human genome, and both of them are highly expressed in human testis (Witt et al. 1996; Schenk et al. 2011). H3.5 is encoded in *H3f3C* gene whose transcript is highly identical to H3.3/*H3f3B* mRNA. When H3.5 is ectopically expressed in HEK293 cells, it is preferentially localized at euchromatin region (Schenk et al. 2011). So H3.5 may come from *H3f3b* gene duplication and have similar function as H3.3. Furthermore, other testis-specific H3 variant, H3t, have been well characterized in human testis. H3t is specifically transcribed in human testis (Witt et al. 1996). Previous studies have reported that nucleosomes assembled by human H3t are more unstable than canonical nucleosomes *in vitro* and have more rapid turnover *in vivo* because of two human H3t-specific residues, Met71 and Val111 (Tachiwana et al. 2010). The different physical properties of human H3t-containing

nucleosome may function in global chromatin reorganization during meiosis and post-meiotic event in testis. In addition, H3t expression in mouse and rat testes has been described briefly before (Meistrich et al. 1985; Govin et al. 2007; Garcia et al. 2008; Montellier et al. 2013). Although mouse H3t has 98 % identity to human H3t, mouse H3t contains Val71 and Ala111 that are the same to canonical H3, not human H3t. Therefore, mouse H3t-containing nucleosome may have distinct physical properties to human H3t-containing nucleosome. But so far the biological function and incorporation mechanism of human and mouse H3t are unknown.

1.2 Spermatogenesis

Spermatogenesis is a complex differentiation process in which male gametes, commonly called sperm but specifically known as spermatozoa, are derived from primordial germ cells and occurs in the seminiferous tubules of the testis. The spermatogenesis process is typically divided into three phases: mitotic phase, meiotic phase and spermiogenesis (Figure 3). In the proliferative phase, the diploid spermatogonia undergo successive mitotic divisions to form clones of cells that finally form primary spermatocytes. In meiotic phase, the primary spermatocytes undergo two meiotic divisions to form haploid spermatids. In spermiogenesis, the spermatids change their morphology and chromatin structure to differentiate into spermatozoa. Interestingly, because cytokinesis is not complete during mitotic and meiotic division, spermatogenic cells are connected by cytoplasmic bridges throughout spermatogenesis. Thus spermatogenic cells may share essential signals through this intercellular connection to synchronize cell division and differentiation (Greenbaum et al. 2011). In addition to spermatogenic cells, three types of somatic cells, peritubular myoid cells, Sertoli cells and Leydig cells, are present in or adjacent to the seminiferous tubules. Peritubular myoid cells form a single layer at the external side of basement membrane to maintain seminiferous tubule structure. The functions of Sertoli cells include providing structural support and nutrition to spermatogenic cells and regulating the process of spermatogenesis. In addition, Sertoli cells act as phagocytes to degrade degenerating germ cells and residual cytoplasm after sperm release (Breucker et al. 1985; Blanco-Rodriguez and Martinez-Garcia 1999). Leydig cells are found adjacent to the seminiferous tubules and produce androgenic hormones to promote spermatogenesis when

lutetizing hormone (LH) stimulate them. In below sections, I will introduce the epigenetic and chromatin dynamics from primordial germ cells to spermatozoa.

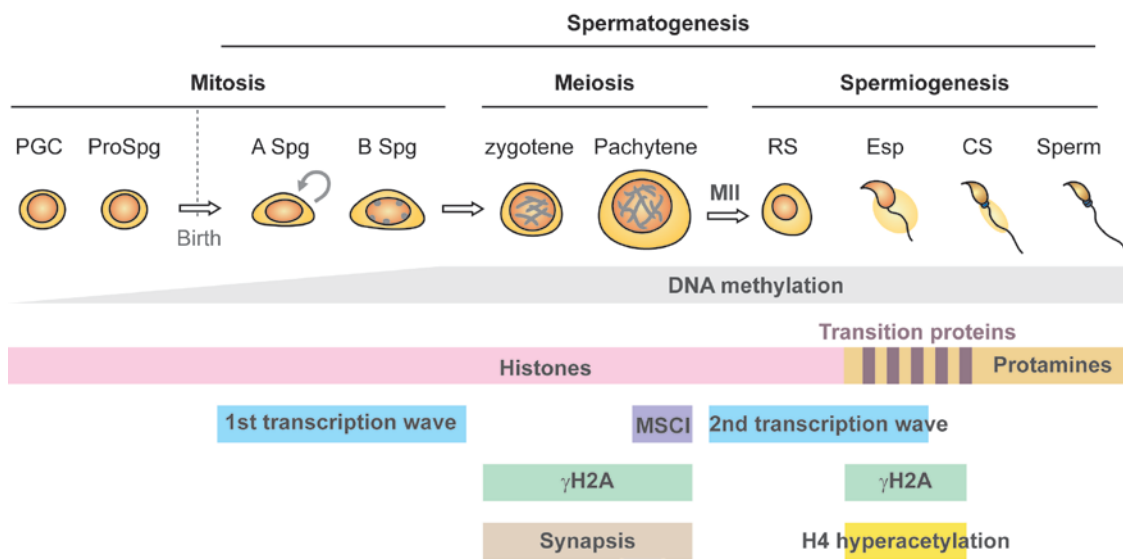


Figure 3. Schematic overview of mouse spermatogenesis. Primordial germ cells (PGC) are derived from epiblast cells of embryo and migrate to the gonadal ridge region. The global DNA demethylation occurs in PGCs. After migration, male PGCs are arrested at G1 phase as prospermatogonia (ProSpg). After birth, prospermatogonia restart transcription and mitosis and become type A spermatogonia (A Spg). Type A spermatogonia undergo multiple mitotic divisions to differentiate into type B spermatogonia (B Spg) and then primary spermatocytes. The formation recombinant synapsis, DNA break response (e.g. γ H2A.X) and meiotic sex chromosome inactivation take part at different stage of primary spermatocytes. After meiosis II, spermatocytes are divided into haploid round spermatids (RS) that undergo dramatic morphological change and chromatin remodeling to differentiate into mature sperm. Most histones are replaced by transition proteins in elongating spermatids (Esp) and then by protamines in condensing spermatids (CS).

1.2.1 Primordial germ cells and spermatogonia

During mammal embryogenesis, primordial germ cells (PGCs) arise from epiblast cells of the postimplantation embryo and subsequently migrate to the gonadal ridge region that is the precursor of the gonads. After PGCs migration and sexual determination, male PGCs enter mitotic arrest at G1 phase of the cell cycle and remain as prospermatogonia (gonocytes) before birth (Western et al. 2008). Interestingly, in order to express germline-specific genes and restore the totipotency to the next generation, two epigenetic reprogramming happens during PGCs migration to gonads. Early PGCs undergo DNA demethylation at imprinted loci, transposons and a subset of germline-specific genes through replication-dependent (passive) and replication-independent (active) mechanism (Lee et al. 2002; Lane et al. 2003; Sato et al. 2003; Hackett et al. 2012; Seisenberger et al. 2012). This reduction of DNA methylation is co-incident with the loss of H3K9me2 and the increase of H3K27me3 (Seki et al. 2005;

Hajkova et al. 2008). When PGCs enter gonads, the second reprogramming occurs at which DNA methylation is completely erased at other regions such as gene bodies and intergenic regions, and the levels of linker histone H1, H3K9me3 and H3K27me3 are also reduced (Seki et al. 2005; Seki et al. 2007; Hajkova et al. 2008). Following complete DNA demethylation in PGCs, the genome must undergo *de novo* methylation by Dnmt3a and Dnmt3b in prospermatogonia (gonocytes) stage, especially imprinted genes and repeat sequences (Davis et al. 1999; Davis et al. 2000; Kato et al. 2007). The acquisition of DNA methylation in male germ cells also continues after birth at mitotic and meiotic phase of spermatogenesis (Oakes et al. 2007).

After birth, prospermatogonia in the fetus resume mitotic proliferation and become post-natal spermatogonial stem cells. Spermatogonial stem cells are localized along the basement membrane of seminiferous tubules and have self-renewal and differentiation capacity to maintain stem cell pool and also sperm production throughout the male life. In rodents, based on their nuclear morphology, spermatogonia have been divided into three subtypes: type A, intermediate and type B spermatogonia. The type A single (A_s) spermatogonia are thought the most undifferentiated spermatogonial stem cells, and heterochromatin is absent from the nucleus of A_s spermatogonia. The A_s spermatogonia can undergo two mitotic divisions to form type A pair (A_{pr}) spermatogonia and subsequently a chain of four type A aligned (A_{al}) spermatogonia that are still in undifferentiated stage. A_{al} spermatogonia continue to go through multiple rounds of division to become differentiating spermatogonia (A1, A2, A3, A4, intermediate and B). Unlike undifferentiated spermatogonial stem cells, type B spermatogonia contain a large amount of heterochromatin. The type B spermatogonia undergo one mitotic division and then give rise to primary spermatocytes that progress into meiosis.

1.2.2 Meiotic phase of spermatogenesis

During meiosis, a single round of DNA replication is followed by two cycles of cell division, termed meiosis I and meiosis II, to produce four haploid spermatids from one diploid type B spermatogonium. In the first meiotic prophase, replicated homologous chromosomes pair along their lengths and form synapsis from leptotene to zygotene stages and then synapse completely at pachytene stage. Pairing and synapsis of homologous chromosomes are

crucial for correct chromosome segregation during meiosis and ensures that mature gametes contain a full set of chromosomes (Handel and Schimenti 2010). In addition, these chromosomal pairing and synapsis accompanies the recombination, which is thought to increase the genetic diversity to next generations. Recombination is initiated by DNA double-strand breaks that can recruit homologous recombination repair machinery. However, in mammals, males carry X and Y sex chromosomes with homology restricted to only a small portion of their length (Burgoyne 1982). In pachytene stage, the sex chromosomes partially synapse through their pseudo autosomal regions to form the sex- or XY-body and undergo chromatin remodeling to silence X- and Y-linked genes (Handel 2004). This phenomenon of transcription silencing is known as meiotic sex chromosome inactivation (MSCI). It is well-known that many DNA damage response proteins and histone markers, such as γ H2A.X, uH2A, MDC1 and 53BP1, are accumulated in sex body and thought to facilitate the initiation of MSCI (Turner 2007). Interestingly, from early to later pachytene stage, canonical H3.1 and H3.2 are progressively replaced by variant H3.3 only in sex body (van der Heijden et al. 2007). This histone exchange may alter epigenetic modifications in sex chromatin presumably required for MSCI. After lengthy process of the first meiosis, the second meiotic division occurs immediately to generate haploid spermatids.

1.2.3 Spermiogenesis

Spermiogenesis is a post-meiotic process in which the haploid spermatids produced from division of secondary spermatocytes undergo a series of changes to differentiate into mature spermatozoa. Based on the nuclear size and morphology, spermatids can be classified into three subtypes: round, elongating and condensing (elongated) spermatids. The main changes during spermiogenesis include acrosome formation, flagellar tail formation, removal of cytoplasm and nuclear condensation. The acrosome is an organelle derived from Golgi apparatus and forms a cap-structure covering the anterior part of the sperm nucleus. The acrosome contains many hydrolytic enzymes that can break down the zona pellucida around the ovum, allowing the sperm nucleus to enter the ovum during fertilization (Tulsiani et al. 1998). The long flagellum is a specific structure in sperm that connects to posterior to the nucleus and is composed of microtubules. The flagellar movement is driven by dynein motor proteins, which use the energy of ATP hydrolysis to slide the microtubules, to propel sperm through the female reproductive tract (Inaba 2011).

Moreover, when spermiogenesis process is complete, mature sperm is released from Sertoli cells into the lumen of the seminiferous tubule. At this time point, the unnecessary organelle and cytoplasm of spermatids, known as the residual body, are separated from sperm and are phagocytosed by the Sertoli cell (Chemes 1986). Then the released sperm without cytoplasm is transported from testis to the epididymis for storage. However, mature sperm still carries many mitochondria as an energy source for flagellar movement (Piomboni et al. 2012).

During the last stage of spermiogenesis, the nuclear volume of spermatid dramatically reduces. In order to compact male genomic DNA into small sperm nucleus, chromatin in spermatids is subject to dramatic remodeling during which global histones on chromatin are replaced by transition proteins at elongating spermatid stage and finally by protamines. Because protamines are small arginine-rich proteins, protamines can interact with negative charged DNA backbone more strongly than histones. Also, protamines contain many cysteine residues that can form inter- and intra-protamine disulfide bounds essential for the formation of highly compacted chromatin (Carrell et al. 2007). In addition, chromatin compaction by protamines can protect paternal genome in sperm heads from physical and chemical damages. There are two protamines proteins, protamine 1 (*Prm1*) and protamine 2 (*Prm2*), expressed in human and mouse testes. The disruption of *Prm1* or *Prm2* in mice leads to morphologically abnormal sperm (Cho et al. 2001). Although *Prm2*-deficient sperm can activate metaphase II-arrested mouse eggs through intracytoplasmic sperm injection, most of zygotes are unable to develop to blastocyst stage (Cho et al. 2003). Moreover, the absence and mutation of protamine 2 have been reported in infertile male patients (de Yebra et al. 1993; Tanaka et al. 2003). Therefore, the proper protamine expression for histone-to-protamine exchange plays a critical role in sperm maturation, paternal DNA stabilization and male fertility.

During histone-to-protamine exchange processing, the hyperacetylation of histone H4 and DNA breaks have been observed in elongating spermatids (McPherson and Longo 1993a; Sonnack et al. 2002; Laberge and Boissonneault 2005; Govin et al. 2007). Hyperacetylated H4 is thought to facilitate histone eviction by directly reducing DNA-histone interactions and opening the chromatin structure for recruiting chromatin remodeling machinery. In mice and humans, reduced levels of histone H4 hyperacetylation in sperm correlates with

impaired fertility (Sonnack et al. 2002; Fenic et al. 2004). In addition, it has been reported that bromodomain testis-specific protein, Brdt, is able to bind acetylated H4 at the time of histone replacement by transition proteins (Moriniere et al. 2009). Mice expressing a truncated form of BRDT lacking the first bromodomain show defects in fertility caused by abnormal nuclear compaction in post-meiotic spermatids (Shang et al. 2007). Furthermore, SWI/SNF, an ATP-dependent chromatin remodeling complex, can be recruited to hyperacetylation site through interacting with Brdt (Dhar et al. 2012). Therefore, H4 hyperacetylation in elongating spermatid may provide a docking platform for bromodomain protein to regulate histone replacement and chromatin remodeling. Recently, a bromodomain-containing proteasome activator, PA200, has been reported to recognize histone acetylation, especially H4 and H2B acetylation, to degrade histone during spermiogenesis. The deletion of PA200 in mice causes accumulation of core histones in elongated spermatids and reduces male fertility (Qian et al. 2013). Thus H4 hyperacetylation is also relative to histone degradation during spermiogenesis.

Double-strand DNA break is other notable mark that occurs in elongating spermatid during histone-to-protamine exchange processing, so elongating spermatid chromatin contains the high level of γ H2A.X (Leduc et al. 2008). And histone hyperacetylation induced by HDAC Inhibitor treatment can promote DNA breaks in spermatids (Laberge and Boissonneault 2005). It is well-known that the activity of topoisomerase II beta (TOP2B) is responsible for generating DNA breaks in elongating spermatids (McPherson and Longo 1993b; Chen and Longo 1996). Because protamine-bound DNA is less supercoiled than nucleosome-bound DNA, supercoiled DNA is formed after nucleosome removal during spermiogenesis. TOP2B is capable of unwinding and untangling DNA by creating a transient DNA break. Previous studies showed that TOP2B activity in spermatids is regulated by poly(ADP-ribose) polymerase 1 (PARP1) and poly(ADP-ribose) glycohydrolase (PARG1) (Meyer-Ficca et al. 2011). However, in general, the enzyme-bridged DNA breaks transiently introduced by TOP2B do not induce a DNA damage response. The γ H2A.X signal in elongating spermatids may be generated by abortive TOP2B reactions. As post-meiotic spermatid is haploid, non-homologous end-joining (NHEJ) pathway or similar processes is thought to mediate the following DNA repair (Leduc et al. 2008).

1.2.4 Testis-specific histone variants

Various histone variants are specifically expressed in the testis. As described in the above section, four H3 variants, H3t, H3.X, H3.Y and H3.5, have been identified in testis or specifically in human testis. In addition to H3 variants, H1, H2A and H2B also have testis-specific variants in mammals.

There are three linker histone H1 variants, H1t, H1T2 and HILS1, specifically present in the testis. The H1t is detectable in pachytene spermatocytes and round spermatids (Drabent et al. 1996). But H1t-deficient male mice are fertile and have normal spermatogenesis, because canonical H1 expression increases to replace H1t function during spermatogenesis (Drabent et al. 2000). The H1T2, also known as HANP1, is selectively expressed in post-meiotic spermatids. The loss of H1T2 leads to abnormal nuclear condensation during histone-to-protamine exchange and reduces male fertility (Martianov et al. 2005; Tanaka et al. 2005). Moreover, HILS1 is another spermatid-specific linker histone and highly expressed in elongating and condensing spermatids (Yan et al. 2003).

At least four testis-specific H2B variants are identified in the testis so far. TH2B is the well-studied testis-specific H2B variant, which is detectable around meiotic stage and replaces canonical H2B progressively during meiosis (Montellier et al. 2013). Although TH2B deletion has no impact on male fertility, transgenic male mice expressing TH2B with a C-terminal tag are sterile because of aberrant histone-to-protamine exchange (Montellier et al. 2013). Moreover, the spermatid-specific H2B (ssH2B) is found in rat round spermatids and is degraded during nuclear compaction (Unni et al. 1995). And mRNA expression of other two H2B variants, H2BL1 and H2BL2, are specifically detected in the mouse testis. H2BL1 is strongly present in spermatids compared to pachytene spermatids, but H2BL2 expression level is relatively low in meiotic and post-meiotic stages (Govin et al. 2007).

Six non-canonical H2A variants are reported to be highly expressed in the testis. The expression profile of TH2A is similar to TH2B, because their genes are localized adjacently on chromosome 17 and share the same promoter (Huh et al. 1991). TH2A is detectable from meiotic stage during spermatogenesis (Shinagawa et al. 2015). The deletion of both TH2A and TH2B causes incomplete release of cohesin at interkinesis after meiosis I, abnormal histone replacement during spermiogenesis and thereby male infertility. Interestingly, the lack of TH2B is compensated for by overexpression of canonical H2B in spermatocytes and

spermatids, but canonical overexpression of H2A was not found (Shinagawa et al. 2015). Thus TH2A may play a more important role than TH2B during spermatogenesis. Moreover, H2A.Bbd, X-chromosome-encoded histones, is highly expressed in mouse elongating spermatids and is also found in human sperm (Ishibashi et al. 2010). H2A.Bbd-containing nucleosome is unstable *in vitro* (Gautier et al. 2004), so H2A.Bbd may contribute to the displacement of histones by protamines in the later stage of spermatids. The expression of H2A.Lap1 is found in spermatocytes and spermatids. H2A.Lap1 is specifically localized at transcription start sites of active genes in pachytene and round spermatids based on ChIP-seq analysis and is also enriched in sex chromosomes at the late round spermatids (Soboleva et al. 2012). Three novel H2A variants, H2AL1/L2/L3, are identified in mouse genome and is highly expressed in spermatids. H2AL1 and H2AL2 are enriched in pericentric regions in condensing spermatids and may participate in reprogramming of pericentric heterochromatin (Govin et al. 2007).

1.2.5 Nucleosome retention in mature spermatozoa

During human and mouse spermiogenesis, histone-to-protamine exchange is not complete. 10 to 15 % of nucleosomes are retained in human sperm compared with somatic cell, and 1% of nucleosomes are retained in mouse sperm (Gatewood et al. 1987; Brykczynska et al. 2010). The retained nucleosomes are not randomly distributed in sperm genome. Nucleosomes are preferentially retained at promoter regions of genes, especially at unmethylated CpG-rich promoters, and imprinting regions in human and mouse sperm (Hammoud et al. 2009; Brykczynska et al. 2010; Erkek et al. 2013). Retained nucleosomes in sperm are marked by distinct histone modifications depending on the regulatory elements of genes. For example, most promoters of testis-specific and housekeeping genes have strong enrichment of H3K4me2 in sperm, but promoters of development regulatory genes that are repressed in early embryos are marked by H3K27me3 in sperm (Brykczynska et al. 2010). Therefore, the histone marks on the retained nucleosomes in sperm are likely to transmit paternal epigenetic information to next generation after fertilization and then contribute to embryo development.

1.3 Scope of the thesis

Histone H3.3 is a non-canonical H3 variant that is expressed and deposited into chromatin throughout cell cycle. Our previous study has reported that mouse spermatids and sperm have high proportion of histone H3.3 variant compared with embryonic stem cells, and retained nucleosomes at CpG-rich promoter in sperm largely consist of H3.3 (Erkek et al. 2013). Other study also indicated that H3.3 is expressed throughout spermatogenesis, and that its expression level increases at mid-stage of pachytene spermatocytes and persists in post-meiotic phase based on H3.3 immunostaining results (Yuen et al. 2014). When one of two *H3.3* genes, *H3f3b*, is mutated, male mice are sterile because of spermatogenesis arrest at round spermatid stage, abnormal protamine deposition and high rates of apoptosis. (Yuen et al. 2014; Tang et al. 2015). Therefore, H3.3 is required for proper spermatogenesis. Moreover, the existence of a testis-specific H3 variant, called H3t or H3.4, was first identified in human testis and also reported in rat and mouse testis before (Trostle-Weige et al. 1984; Witt et al. 1996; Govin et al. 2007). Compared with canonical H3 nucleosome, human H3t-containing nucleosome is less stable *in vitro*, suggesting that the variant may contribute to histone turnover during meiosis and histone-to-protamine exchange (Tachiwana et al. 2008). However, the detail characterization and biological function of H3t *in vivo* are still unclear. Therefore, the first part of the thesis presented here aim at determining the dynamics of H3 variants compared with canonical H3 during spermatogenesis and also understanding the relationship between H3 modification and H3 variants in each stage of spermatogenetic cells.

During mammalian spermiogenesis, global histones on chromatin are replaced by transition proteins and subsequently by protamines in order to compact paternal genome DNA into sperm head. It is well-known that H4 hyperacetylation and DNA damage response coincide with histone displacement at elongating spermatid stage (Grimes and Henderson 1984; Govin et al. 2007). Hyperacetylated H4 is thought to reduce DNA-histone interaction and may promote histone removal from DNA. A recent study showed that a proteasome activator, PA200, recognizes acetylated H4 and H2B through its bromodomain-like region and is involved in histone degradation in elongating spermatids (Qian et al. 2013). The loss of PA200 causes histone accumulation in the late stage of spermatid and reduces male fertility (Khor et al. 2006; Qian et al. 2013). Moreover, other studies demonstrated that Rnf8,

an E3 ubiquitin-protein ligase of H2A and H2B involving in DNA damage response, is able to regulate MOF-mediated H4K16 acetylation in elongating spermatids and is required for proper histone removal during sperm maturation (Lu et al. 2010). Nevertheless, it is largely unclear what the molecular mechanism of global histone eviction from DNA is during spermiogenesis. Additionally, our previous studies showed that 1 to 10 % of histones are retained at specific genomic regions, particularly at unmethylated CpG-rich promoters, in human and mouse sperm (Gatewood et al. 1987; Hammoud et al. 2009; Brykczynska et al. 2010; Erkek et al. 2013). These residual nucleosomes in sperm genome carry transcriptionally active or repressive modifications that may link to the transmission of paternal epigenetics to next generation. But how specific nucleosomes are kept in specific genomic regions is unresolved. Therefore, in the second part of thesis, the major aim was to investigate the molecular mechanisms of histone replacement and histone retention during mouse spermiogenesis.

Chapter 2

PRC1 coordinates timing of sexual differentiation of female primordial germ cells

Shihori Yokobayashi¹, Ching-Yeu Liang^{1,2}, Hubertus Kohler¹, Peter Nestorov^{1,2}, Zichuan Liu¹, Miguel Vidal³, Maarten van Lohuizen⁴, Tim C. Roloff¹ and Antoine H. F. M. Peters^{1,2}

1. Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, 4058 Basel, Switzerland

2. Faculty of Sciences, University of Basel, 4056 Basel, Switzerland

3. Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas (CSIC), 28040 Madrid, Spain

4. Division of Molecular Genetics and Centre for Biomedical Genetics, the Netherlands Cancer Institute (NKI), 1066 CX Amsterdam, the Netherlands

PRC1 coordinates timing of sexual differentiation of female primordial germ cells

Shihori Yokobayashi^{1,†}, Ching-Yeu Liang^{1,2}, Hubertus Kohler¹, Peter Nestorov^{1,2}, Zichuan Liu¹, Miguel Vidal³, Maarten van Lohuizen⁴, Tim C. Roloff¹ & Antoine H. F. M. Peters^{1,2}

In mammals, sex differentiation of primordial germ cells (PGCs) is determined by extrinsic cues from the environment¹. In mouse female PGCs, expression of stimulated by retinoic acid gene 8 (*Stra8*) and meiosis are induced in response to retinoic acid provided from the mesonephroi^{2–5}. Given the widespread role of retinoic acid signalling during development^{6,7}, the molecular mechanisms that enable PGCs to express *Stra8* and enter meiosis in a timely manner are unknown^{2,8}. Here we identify gene-dosage-dependent roles in PGC development for *Ring1* and *Rnf2*, two central components of the Polycomb repressive complex 1 (PRC1)^{9,10}. Both paralogues are essential for PGC development between days 10.5 and 11.5 of gestation. *Rnf2* is subsequently required in female PGCs to maintain high levels of *Oct4* (also known as *Pou5f1*) and *Nanog* expression¹¹, and to prevent premature induction of meiotic gene expression and entry into meiotic prophase. Chemical inhibition of retinoic acid signalling partially suppresses precocious *Oct4* downregulation and *Stra8*

activation in *Rnf2*-deficient female PGCs. Chromatin immunoprecipitation analyses show that *Stra8* is a direct target of PRC1 and PRC2 in PGCs. These data demonstrate the importance of PRC1 gene dosage in PGC development and in coordinating the timing of sex differentiation of female PGCs by antagonizing extrinsic retinoic acid signalling.

In mammalian somatic cells, PRC1 and PRC2 proteins are transcriptional repressors that function in large multiprotein complexes and that modify chromatin by mono-ubiquitinating histone H2A at lysine 119 (H2AK119u1) and trimethylating histone H3 at lysine 27 (H3K27me3), respectively^{9,12}. At day 12.5 of embryonic development (E12.5), in PGCs marked by *Cdh1* (E-cadherin) staining¹³, we observed nuclear localization of PRC1 components *Rnf2* (also known as *Ring1B*), *Mel18* (also known as *Pcgf2*) and *Rybp* (Fig. 1a and Supplementary Fig. 1) as well as a robust H2AK119u1 signal, suggesting the presence of catalytically active PRC1 complexes (Fig. 1a). To

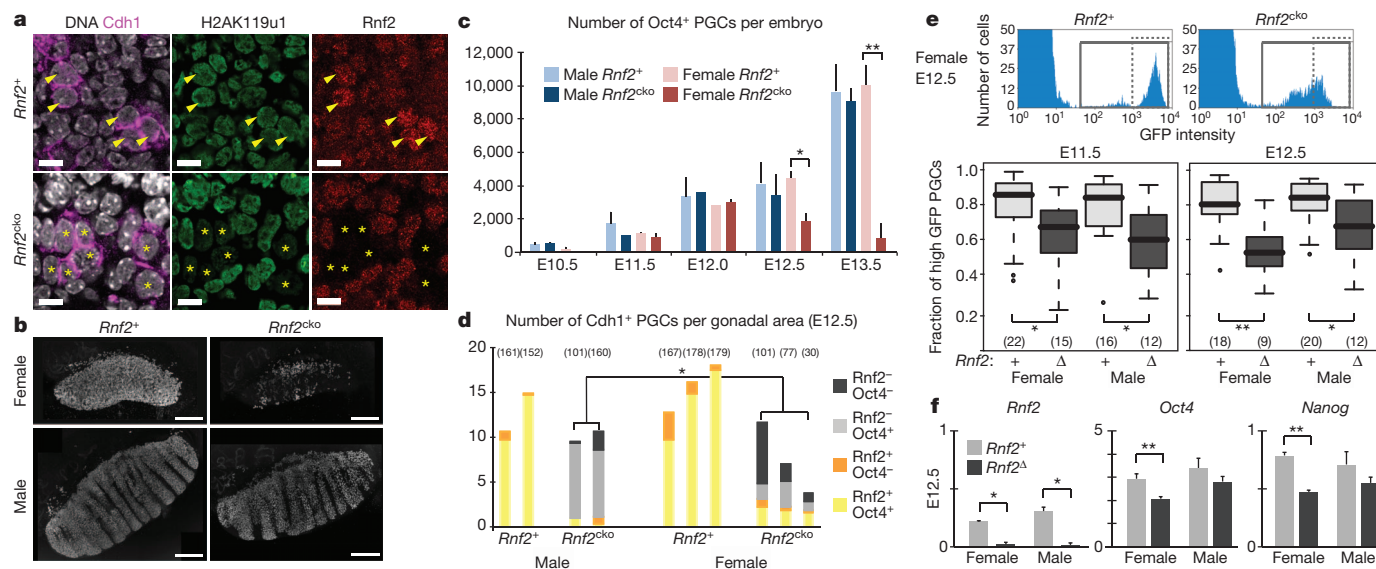


Figure 1 | *Rnf2* regulates PGC development and *Oct4* and *Nanog* expression in *Rnf2*^{cko} female gonads. **a**, Immunofluorescence staining of H2AK119u1, *Rnf2* and *Cdh1* with DAPI (4',6-diamidino-2-phenylindole) in *Rnf2*⁺ and *Rnf2*^{cko} gonadal sections from E12.5 female embryos. Arrowheads denote *Rnf2*⁺ PGCs and asterisks denote *Rnf2*^{cko} PGCs. Scale bars, 10 μm. **b**, Immunofluorescence staining of *Oct4* in E13.5 *Rnf2*⁺ and *Rnf2*^{cko} whole gonads and mesonephroi. Scale bars, 300 μm. **c**, Average number of *Oct4*⁺ cells in whole gonads at E10.5–E13.5. Error bars indicate +s.d. $n = 2–12$. * $P < 0.005$; ** $P < 1.0 \times 10^{-5}$ (Student's *t*-test). **d**, Classification of *Cdh1*-positive PGCs according to *Rnf2* and *Oct4* protein levels in *Rnf2*⁺ and *Rnf2*^{cko} E12.5 gonads. *y* axis represents the number of PGCs that were

normalized to areas analysed (10,000 μm²). Numbers in brackets denote number of PGCs scored per embryo. * $P < 1.0 \times 10^{-8}$ (chi-squared test). **e**, Representative histograms showing *Oct4*(*ΔPE*)–GFP signals in PGCs from female *Rnf2*⁺ and *Rnf2*^{cko} E12.5 gonads. Boxplots showing the ratios of PGCs with high GFP intensity ($>10^3$, enclosed by dashed line in histogram) over all GFP-positive cells (enclosed by solid line) in different embryos. Numbers in brackets denote number of embryos analysed. * $P < 0.05$; ** $P < 0.005$ (Student's *t*-test). **f**, Representative qRT–PCR data of *Rnf2*, *Oct4* and *Nanog* in *Rnf2*⁺ and *Rnf2*^Δ PGCs (normalized to *Tbp*). Error bars indicate +s.d. of 2–3 technical replicates. * $P < 0.05$; ** $P < 0.01$ (Student's *t*-test).

¹Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, CH-4058 Basel, Switzerland. ²Faculty of Sciences, University of Basel, CH-4056 Basel, Switzerland. ³Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas (CSIC), 28040 Madrid, Spain. ⁴Division of Molecular Genetics and Centre for Biomedical Genetics, the Netherlands Cancer Institute (NKI), 1066 CX Amsterdam, the Netherlands. †Present address: Department of Reprogramming Science, Center for iPS Cell Research and Application, Kyoto University, 53 Kawahara-cho, Shogoin Yoshida, Sakyo-ku, Kyoto 606-8507, Japan.

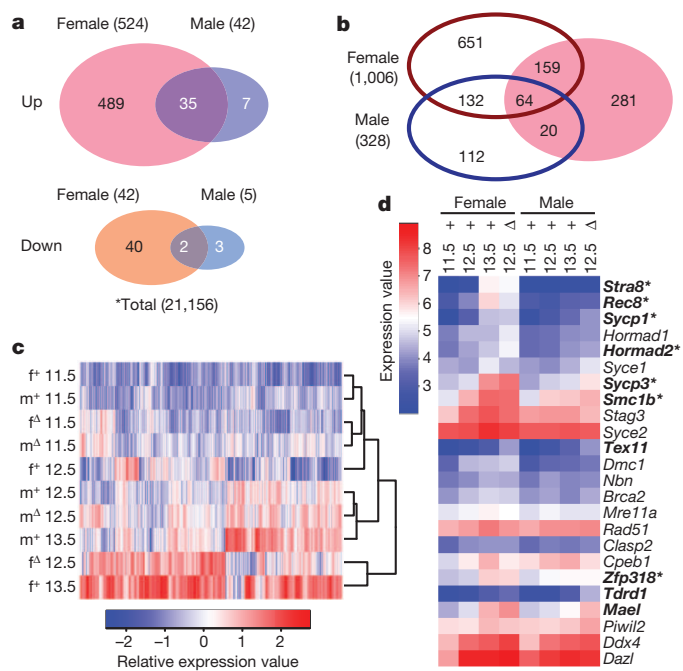


Figure 2 | *Rnf2* deficiency induces extensive transcriptional misregulation in female PGCs. **a**, Venn diagrams showing the numbers of genes misregulated in male and female *Rnf2*^Δ PGCs compared to *Rnf2*⁺ PGCs at E12.5. **b**, Venn diagram showing the numbers of genes upregulated in female (red) and male (blue) control PGCs between E11.5 and E13.5 and in female *Rnf2*^Δ PGCs compared to *Rnf2*⁺ PGCs at E12.5 (pink oval). **c**, Relative expression levels of 1138 probe sets upregulated in male (m) and female (f) *Rnf2*⁺ PGCs between E11.5 and E13.5 (**b**) in various samples indicated. Unsupervised clustering analysis shows clustering of female E12.5 *Rnf2*^Δ PGCs with female E13.5 *Rnf2*⁺ PGCs, whereas E12.5 male *Rnf2*^Δ and *Rnf2*⁺ PGCs clustered together. **d**, Microarray expression values of early meiosis program genes. Genes significantly upregulated in *Rnf2*^Δ PGCs of both sexes are indicated in bold; those in female *Rnf2*^Δ PGCs only in bold with asterisk. *n* = 3 per condition; fold change > 1.5; adjusted *P* value < 0.05.

address the function of PRC1 in PGC development (Supplementary Fig. 2), we conditionally deleted *Rnf2* in PGCs from E9.5 onwards by generating mice carrying a floxed and a mutant allele of *Rnf2* (*Rnf2*^{fl/Δ}) and a Cre recombinase driven by the mouse *Tnap* promoter (*Tnap-cre*)^{14,15}. To concomitantly assess possible functional redundancy with the *Rnf2* paralogue *Ring1* (refs 10, 16), we studied mice that were either heterozygous or homozygous deficient for *Ring1* (ref. 17) (referred to as *Rnf2* conditional knockout (*Rnf2*^{cko}) mice and *Ring1*^Δ *Rnf2*^{cko} mice, respectively) (Supplementary Figs 3 and 6a). This strategy resulted in ~90% deletion efficiency at E11.5 (Supplementary Fig. 4). At E12.5, *Rnf2*, *Mel18*, *Rybp* and *H2AK119u1* were lost in PGCs of *Rnf2*^{cko} embryos but not of *Rnf2*^{fl/Δ} embryos, carrying one functional allele, indicating that complex stability and catalytic activity of PRC1 is regulated by *Rnf2* in PGCs at E12.5 (Fig. 1a and Supplementary Figs 1 and 4). By contrast, *Ezh2* and *H3K27me3* levels were similar in *Rnf2*^{cko} versus control PGCs, suggesting globally unaltered PRC2 function in *Rnf2*^{cko} PGCs (Supplementary Fig. 5).

To study the fate of *Rnf2*-deficient PGCs, we analysed expression of the pluripotency and germ cell marker *Oct4* in whole gonads of E10.5–E13.5 embryos (Fig. 1b and Supplementary Fig. 2). We observed a strong reduction of *Oct4*-positive PGCs, specifically in female *Rnf2*^{cko} embryos but not in male *Rnf2*^{cko} or control embryos, starting around E12.5 of gestation (Fig. 1b, c). By contrast, double deficiency of *Ring1* and *Rnf2* caused a strong reduction of *Oct4*⁺ PGCs by E11.5 in both sexes (Supplementary Fig. 6b, c), indicating an essential role for PRC1 in PGCs after their migration into the embryonic gonad (Supplementary Fig. 2)¹¹.

To further dissect the role of *Rnf2* in regulating *Oct4* expression versus PGC development, we assessed co-expression of *Oct4* and *Rnf2* at E12.5 in *Cdh1*⁺ PGCs in *Rnf2*^{cko} embryos (Fig. 1d). The number of PGCs lacking detectable *Rnf2* protein was strongly reduced in female but not male gonads, despite the fact that gonads of both sexes harboured comparable numbers of *Rnf2*-deficient PGCs at E12.0 (Supplementary Fig. 4a, and data not shown). We further observed a pronounced downregulation of *Oct4* protein in female and some male *Rnf2*-deficient PGCs (Fig. 1d). These data indicate that *Rnf2* contributes to maintaining *Oct4* expression, particularly in female PGCs, beginning between E12.0 and E12.5.

We subsequently investigated the mechanism underlying the reduction in female *Rnf2*-deficient PGCs. We failed to observe increased levels of apoptosis or major changes in cell cycle progression (data not shown). To study changes in gene expression, we introduced a green fluorescent protein (GFP) transgene driven by the promoter of *Oct4* lacking the proximal enhancer (*Oct4*(Δ PE)-GFP)¹⁸ into the *Rnf2*^{cko} strain and isolated pure populations of PGCs by fluorescence-activated cell sorting (FACS) (Fig. 1e). By quantitative reverse transcriptase PCR (qRT-PCR), we barely detected any *Rnf2* transcripts in isolated *Rnf2*^{cko} PGCs, confirming efficient deletion of the *Rnf2*^{fl} allele by *Tnap-cre* (Fig. 1f; see also Supplementary Fig. 4). Hence, we subsequently refer to PGCs from *Rnf2*^{cko} embryos as *Rnf2*^Δ. We also noticed significant reductions in GFP intensities in *Rnf2*^Δ PGCs isolated from male and female E11.5 and E12.5 embryos compared to controls (Fig. 1e), suggesting decreased *Oct4*(Δ PE) promoter activity in these cells. qRT-PCR analysis showed significantly reduced *Oct4* and *Nanog* expression in female GFP-positive *Rnf2*^Δ PGCs at E11.5 and E12.5 (Fig. 1f and data not shown). Thus, *Rnf2* is required for maintaining the expression of pluripotency factors in PGCs.

We next analysed genome-wide expression in purified PGCs. The number of misregulated genes in female compared to male *Rnf2*^Δ PGCs at E12.5 was 12-fold higher (Fig. 2a and Supplementary Table 1). Consistent with the fact that PRC1 is a transcriptional repressor, ~90% of misexpressed genes were upregulated. At E11.5, we only observed a few misregulated genes in *Rnf2*^Δ PGCs (Supplementary Table 1), consistent with the timing of the mutant phenotype. According to Gene Ontology analysis, gene functions related to meiosis (synapsis, sister chromatid cohesion) were highly over-represented among genes upregulated in female *Rnf2*^Δ PGCs. By contrast, nucleosome functions were enriched in downregulated genes, as reported previously for *Ring1* and *Rnf2* double-deficient germinal vesicle oocytes¹⁰ (Supplementary Fig. 7 and Supplementary Table 2). Notably, we barely found any over-representation of developmental gene functions in *Rnf2*^Δ PGCs, a classical feature of PRC1 deficiency in ESCs^{16,19} and germinal vesicle oocytes¹⁰. To test whether this is due to functional redundancy of the *Ring1* paralogue, we profiled expression in *Ring1*^Δ *Rnf2*^Δ PGCs purified from E11.5 embryos. Consistently, many developmental gene functions were over-represented among upregulated genes (Supplementary Fig. 8 and Supplementary Table 2). These data indicate that although *Ring1* expression is sufficient to safeguard global repression of canonical Polycomb target genes in PGCs, *Rnf2* is required in female E12.5 PGCs for repression of genes driving meiosis.

At E11.5, gonads still possess the potential to develop into either ovaries or testes, but are committed to a sex-specific differentiation process 2 days later. To relate aberrant expression in *Rnf2*^Δ PGCs to changes occurring during normal PGC differentiation, we profiled expression in *Rnf2*⁺ PGCs isolated between E11.5 and E13.5. In this developmental period, sixfold more genes were upregulated in female (810) than in male (132) *Rnf2*⁺ PGCs, with an additional 196 genes being upregulated in both sexes, suggesting the activation of female- and male-specific 'PGC-differentiation programs' (PDPs) (Fig. 2b). Among genes upregulated in E12.5 female *Rnf2*^Δ PGCs, 223 (43%) were found to be part of the female PDP (Fig. 2b and Supplementary Fig. 7b) (*P* value 2.67×10^{-159} ; geometric test), being activated about 1 day ahead compared to those in *Rnf2*⁺ PGCs (Fig. 2c).

We next identified that 24 out of 119 genes annotated with the Gene Ontology term ‘meiosis’ (GO:0007126) were part of the female PDP, probably reflecting activation of an ‘early meiosis program’ (Fig. 2d). Among these 24 genes, 10 were precociously upregulated in female *Rnf2^Δ* PGCs, including *Stra8*, required for meiotic initiation⁵, and others such as *Rec8*, *Sycp3*, and *Hormad2*, which have key functions in cohesion, chromosome synapsis and recombination^{20–22}. Using qRT–PCR, we measured high *Stra8*, *Rec8* and *Sycp3* messenger RNA levels at E13.5 in female *Rnf2⁺* PGCs (Fig. 3a). At E11.5 and E12.5, *Stra8* and *Rec8* were precociously activated, with up to 20-fold higher expression levels in female *Rnf2^Δ* versus *Rnf2⁺* PGCs. *Sycp3* was expressed in *Rnf2⁺* PGCs from both sexes with up to fourfold increased levels in female *Rnf2^Δ* PGCs (Fig. 3a). Immunofluorescence analyses revealed strong *Stra8* nuclear localization by E12.5 in PGCs of *Rnf2^{cko}* gonads, whereas the protein only started to accumulate in control PGCs at E13.5 (Fig. 3b and Supplementary Fig. 7c). For *Sycp3*, we observed focal nuclear staining at E13.5 and synaptonemal complex staining at E14.5 in *Rnf2^Δ* germ cells, whereas only axial elements of meiotic chromosomes were visible in *Rnf2⁺* germ cells at E14.5 (Fig. 3c, d, data not shown). These data indicate that precocious transcriptional activation of *Stra8* and other PGC differentiation and meiosis genes in female *Rnf2^Δ* PGCs induces these cells to prematurely stop proliferation and enter into meiotic prophase, hence accounting for the lower number of female *Rnf2^Δ* PGCs at E12.5. We also measured increased expression of genes functioning in retinoic acid metabolism, such as *Aldh1a2*, *Crabp1* and *Crabp2* in female *Rnf2^Δ* PGCs at E11.5 and E12.5 (Supplementary Fig. 7d and data not shown), probably enhancing retinoic acid signalling and meiotic entry in a feed-forward manner²³.

We next wanted to find out why *Stra8* transcription is abnormally activated only in female *Rnf2^Δ* PGCs. In male gonads, retinoic-acid-mediated induction of *Stra8* is counteracted by the somatically expressed retinoid-degrading enzyme *Cyp26b1* and by the *Fgf9* signalling pathway

(Supplementary Fig. 2)^{3,4,24}. PRC1 may therefore not be required to suppress *Stra8* expression in males. To study this possibility, we aimed to overcome the antagonizing activity of *Cyp26b1* and cultured E11.5 genital ridges for 24 h in the presence of all-*trans* retinoic acid (ATRA)^{3,4}. ATRA increased *Stra8* expression in isolated *Rnf2⁺* PGCs of both sexes to levels comparable with what we measured in female *Rnf2^Δ* PGCs treated with vehicle (Fig. 4a). Intriguingly, ATRA additionally enhanced *Stra8* expression in male as in female *Rnf2^Δ* PGCs to >fivefold higher levels compared to *Rnf2⁺* PGCs (Fig. 4a), indicating that PRC1 effectively suppresses retinoic-acid-induced *Stra8* activation in PGCs of both sexes. We observed comparable responses for *Rec8* and *Sycp3* expression (Fig. 4a). By contrast, *Stra8* activation was completely suppressed in all genotypes by treatment with Win-18446 (*N,N'*-octamethylenebis(dichloroacetamide)), an inhibitor of the retinoic acid biogenesis pathway²⁵ (Fig. 4a). Likewise, 2-day *in vivo* exposure of PGCs to Win-18446 in developing embryos suppressed premature *Stra8* protein expression (Supplementary Fig. 9a, b). In addition, numbers of Oct4-expressing PGCs in *Rnf2^{cko}* gonads were significantly increased upon treatment *in vitro* and *in vivo* with Win-18446 (Fig. 4b and Supplementary Fig. 9c, d). These results indicate that sensitization of PGCs to retinoic acid signalling by *Rnf2* deficiency contributes to precocious exit from the PGC state and activation of the meiotic program in female gonads.

Subsequently, most *Rnf2^Δ* PGCs were unable to complete meiosis and to develop into mature oocytes (Supplementary Fig. 10), possibly as a consequence of the precocious entry into meiosis that may eventually impair, perhaps in concert with other aberrant changes in gene expression (Fig. 2b), the execution of the natural female PDP and oogenesis. Alternatively, PRC1 may exert a separate essential function later during meiotic progression.

Finally, we performed a micro-chromatin immunoprecipitation (μ ChIP) assay on isolated *Rnf2⁺* PGCs. At E11.5, the *Stra8* promoter was strongly enriched with PRC2-mediated H3K27me3 and with

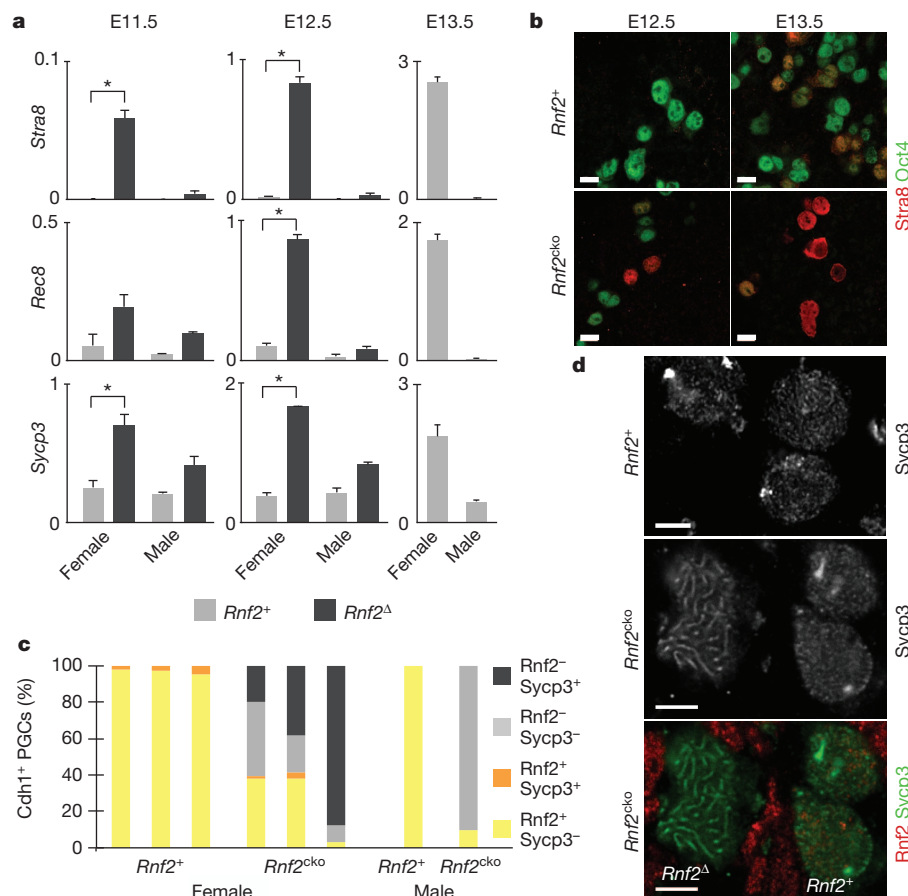


Figure 3 | Female *Rnf2^Δ* PGCs enter precociously into meiotic prophase.

a, Representative qRT–PCR data of *Stra8*, *Rec8* and *Sycp3* transcripts (normalized to *Tbp*) in isolated PGCs. Error bars indicate +s.d. of two technical replicates. * $P < 0.05$ (Student's *t*-test).

b, Immunofluorescence staining of Oct4 and *Stra8* in *Rnf2⁺* and *Rnf2^{cko}* female gonads at E12.5 and E13.5. Scale bars, 10 μ m. **c**, Classification of Cdh1-positive PGCs according to *Rnf2* and *Sycp3* protein levels in *Rnf2⁺* and *Rnf2^{cko}* E13.5 gonads of individual embryos. Numbers in brackets indicate the number of PGCs scored per embryo.

d, Immunofluorescence staining of *Sycp3* and *Rnf2* in *Rnf2⁺* and *Rnf2^{cko}* E14.5 ovaries. Scale bars, 10 μ m.

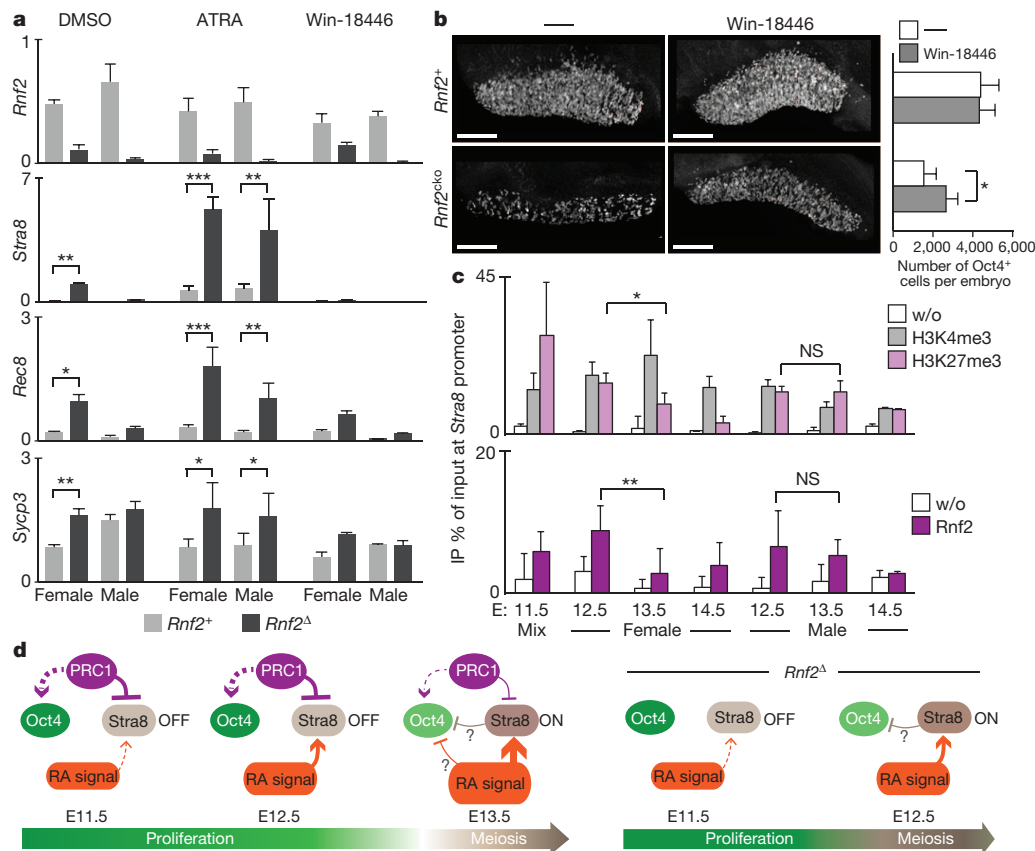


Figure 4 | PRC1 antagonizes retinoic acid signalling and maintains *Stra8* in a repressive chromatin state. **a**, Representative qRT-PCR data of *Rnf2*, *Stra8*, *Rec8* and *Sycp3* transcripts (normalized to *Tbp*) in isolated PGCs. Error bars indicate +s.d. of 2–4 technical replicates. * $P < 0.01$; ** $P < 0.001$; *** $P < 0.0001$ (Student's *t*-test). **b**, Immunofluorescence staining of *Oct4* and number of *Oct4*-positive PGCs in *Rnf2*⁺ and *Rnf2*^{ko} whole gonads and mesonephroi from E12.5 female embryos, treated with Win-18446 or vehicle at E10.5. Graph shows means with +s.d. of 2–4 technical replicates. Scale bars, 300 μ m. **c**, μ ChIP analysis for H3K4me3, H2K27me3 and *Rnf2* at the promoter region of *Stra8* in PGCs isolated from E11.5, 12.5, 13.5 and 14.5 *Rnf2*⁺ gonads. Percentages of immunoprecipitated DNA (IP) compared to input are shown. w/o, without primary antibodies. Error bars indicate +s.d. of 1–3 biological and 2–4 technical replicates. * $P < 0.05$; ** $P < 0.001$ (Student's *t*-test). **d**, Model for transcriptional regulation of meiotic genes (for example, *Stra8*) and pluripotency genes (for example, *Oct4*) during sex differentiation of female PGCs. RA, retinoic acid.

H3K4me3, indicative of a Polycomb repressed, yet potentially transcriptionally primed, chromatin state. During subsequent stages, we noticed a progressive decrease in H3K27me3 levels in female PGCs consistent with *Stra8* transcriptional activation, whereas the promoter remained bivalent in male PGCs (Fig. 4c). By contrast, the *Hoxa9* promoter was bivalent in all conditions (Supplementary Fig. 11). Likewise, we detected *Rnf2* association with the *Stra8* promoter at E11.5 and E12.5 PGCs and a significant decrease in *Rnf2* occupancy in female but not male PGCs at E13.5. Together, these experiments suggest that *Stra8* is directly regulated by PRC2 and PRC1 in PGCs of both sexes.

In summary, we identified an essential role for PRC1 in PGC development between E10.5 and E11.5. Later, between E11.5 and E13.5, *Rnf2* effectively modulates the sensitivity of PGCs to retinoic acid-mediated induction of meiosis by directly controlling the competence of *Stra8* and probably of other meiotic genes for transcriptional activation. Like the proposed role of *Cbx2* in temporal co-linearity of *Hox* gene activation²⁶, we speculate that PRC1 maintains repression of *Stra8* and other genes of the PGC differentiation and early meiosis programs until retinoic acid signalling has reached a certain threshold (Fig. 4d). In addition, PRC1 regulates the expression of pluripotency genes in PGCs, possibly indirectly by suppressing transcription of negative regulators. Impairing PRC1 function in these parallel pathways probably leads to a synergistic effect, thereby promoting premature transition from proliferation into meiosis.

METHODS SUMMARY

Embryos were obtained by timed matings, by scoring noon of the day after mating as E0.5. Genital ridges were cultured in drops of Dulbecco's minimal eagle medium supplemented with 10% fetal calf serum at 37 °C with 5% CO₂ in air. For expression profiling, we collected in triplicate 500 PGCs per embryo by fluorescence-activated cell sorting and processed, hybridized to Affymetrix Mouse Gene 1.0 arrays and analysed as described previously¹⁰. Gene Ontology terms were obtained using Gostat (<http://gostat.welhi.edu.au>). μ ChIP was performed as described previously²⁷.

Full Methods and any associated references are available in the online version of the paper.

Received 27 April 2012; accepted 16 January 2013.

- Brennan, J. & Capel, B. One tissue, two fates: molecular genetic events that underlie testis versus ovary development. *Nature Rev. Genet.* **5**, 509–521 (2004).
- Menke, D. B., Koubova, J. & Page, D. C. Sexual differentiation of germ cells in XX mouse gonads occurs in an anterior-to-posterior wave. *Dev. Biol.* **262**, 303–312 (2003).
- Koubova, J. *et al.* Retinoic acid regulates sex-specific timing of meiotic initiation in mice. *Proc. Natl Acad. Sci. USA* **103**, 2474–2479 (2006).
- Bowles, J. *et al.* Retinoid signaling determines germ cell fate in mice. *Science* **312**, 596–600 (2006).
- Baltus, A. E. *et al.* In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. *Nature Genet.* **38**, 1430–1434 (2006).
- Deschamps, J. Ancestral and recently recruited global control of the *Hox* genes in development. *Curr. Opin. Genet. Dev.* **17**, 422–427 (2007).
- Duester, G. Retinoic acid synthesis and signaling during early organogenesis. *Cell* **134**, 921–931 (2008).
- Oulad-Abdelghani, M. *et al.* Characterization of a premeiotic germ cell-specific cytoplasmic protein encoded by *Stra8*, a novel retinoic acid-responsive gene. *J. Cell Biol.* **135**, 469–477 (1996).
- Sparmann, A. & van Lohuizen, M. Polycomb silencers control cell fate, development and cancer. *Nature Rev. Cancer* **6**, 846–856 (2006).
- Posfai, E. *et al.* Polycomb function during oogenesis is required for mouse embryonic development. *Genes Dev.* **26**, 920–932 (2012).
- Saitou, M., Kagiwada, S. & Kurimoto, K. Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139**, 15–31 (2012).
- Simon, J. A. & Kingston, R. E. Mechanisms of polycomb gene silencing: knowns and unknowns. *Nature Rev. Mol. Cell Biol.* **10**, 697–708 (2009).
- Di Carlo, A. & De Felici, M. A role for E-cadherin in mouse primordial germ cell development. *Dev. Biol.* **226**, 209–219 (2000).
- Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nature Genet.* **40**, 411–420 (2008).
- Lomeli, H., Ramos-Mejia, V., Gertsenstein, M., Lobe, C. G. & Nagy, A. Targeted insertion of Cre recombinase into the TNAP gene: excision in primordial germ cells. *Genesis* **26**, 116–117 (2000).
- Endoh, M. *et al.* Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. *Development* **135**, 1513–1524 (2008).

17. del Mar Lorente, M. *et al.* Loss- and gain-of-function mutations show a polycomb group function for Ring1A in mice. *Development* **127**, 5093–5100 (2000).
18. Yoshimizu, T. *et al.* Germ-line-specific expression of the Oct-4/green fluorescent protein (GFP) transgene in mice. *Dev. Growth Differ.* **41**, 675–684 (1999).
19. Leeb, M. *et al.* Polycomb complexes act redundantly to repress genomic repeats and genes. *Genes Dev.* **24**, 265–276 (2010).
20. Xu, H., Beasley, M. D., Warren, W. D., van der Horst, G. T. & McKay, M. J. Absence of mouse REC8 cohesin promotes synapsis of sister chromatids in meiosis. *Dev. Cell* **8**, 949–961 (2005).
21. Yuan, L. *et al.* Female germ cell aneuploidy and embryo death in mice lacking the meiosis-specific protein SCP3. *Science* **296**, 1115–1118 (2002).
22. Wojtasz, L. *et al.* Meiotic DNA double-strand breaks and chromosome asynapsis in mice are monitored by distinct HORMAD2-independent and -dependent mechanisms. *Genes Dev.* **26**, 958–973 (2012).
23. Maden, M. Retinoic acid in the development, regeneration and maintenance of the nervous system. *Nature Rev. Neurosci.* **8**, 755–765 (2007).
24. Bowles, J. *et al.* FGF9 suppresses meiosis and promotes male germ cell fate in mice. *Dev. Cell* **19**, 440–449 (2010).
25. Hogarth, C. A. *et al.* Suppression of *Stra8* expression in the mouse gonad by WIN 18,446. *Biol. Reprod.* **84**, 957–965 (2011).
26. Bel-Vialar, S. *et al.* Altered retinoic acid sensitivity and temporal expression of *Hox* genes in *polycomb-M33*-deficient mice. *Dev. Biol.* **224**, 238–249 (2000).
27. Dahl, J. A. & Collas, P. A rapid micro chromatin immunoprecipitation assay (ChIP). *Nature Protocols* **3**, 1032–1045 (2008).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank A. Nagy and H. Schöler for the *Tnap-cre* and *Oct4(APE)*-GFP mice, respectively, and M. Griswold for the *Stra8* antibody. We are grateful to L. Gelman (microscopy and imaging), L. Burger, M. Stadler (bioinformatics), E. Cabuy, S. Thiry, K. Jacobeit (functional genomics) and the FMI animal facility for assistance. We thank members of the Peters laboratory, particularly M. Tardat, S. Erkek and M. Gill, for experimental support and discussions. S.Y. is a recipient of a Human Frontier Science Program long-term fellowship and a Japan Society for the Promotion of Science postdoctoral fellowship. Research in the Peters laboratory is supported by the Novartis Research Foundation, the Swiss National Science Foundation (31003A_125386 and NRP 63 - Stem Cells and Regenerative Medicine), SystemsX.ch (Cell plasticity), the Japanese Swiss Science and Technology Cooperation Program, the European Network of Excellence 'The Epigenome' and the EMBO Young Investigator program.

Author Contributions S.Y. and A.H.F.M.P. conceived and designed the experiments. S.Y. performed almost all experiments. C.-Y.L. performed μ ChIP experiments. H.K. performed FACS isolations. Z.L. isolated germinal vesicle oocytes. P.N. isolated RNA for RNA sequencing and provided advice on specific target amplification qRT-PCR. M.V. provided *Ring1*-deficient mice. M.v.L. provided *Rnf2* conditionally deficient mice. T.C.R. assisted in microarray and RNA-sequencing analysis. S.Y., C.-Y.L. and A.H.F.M.P. analysed the data. S.Y. and A.H.F.M.P. wrote the manuscript.

Author Information Microarray and RNA-sequencing data have been deposited in the Gene Expression Omnibus under accession numbers GSE42782 and GSE42852, respectively. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to A.H.F.M.P. (antoine.peters@fmi.ch).

METHODS

Mice and embryo collection. *Rnf2*^{cko} mice with *Rnf2*-deficient PGCs were generated by combining a floxed *Rnf2* (*Rnf2*^{fl}) allele¹⁴ with the *Tnap-cre* transgene allele as illustrated in Supplementary Fig. 3. Introduction of the *Ring1* mutation¹⁷ is illustrated in Supplementary Fig. 6a. For PGC isolation, embryos were sired by fathers that were homozygous for the *Oct4*(*APE*)-GFP transgene. Mice were maintained on a mixed background of 129/Sv and C57BL/6J. Embryos were obtained by timed matings, by scoring noon of the day following mating as 0.5 embryonic day of development (E0.5). The genotype of embryos was determined by PCR as described previously¹⁴. Sex of gonads was determined by PCR for *Ubx1* using the following primers: forward, 5'-TGGTCTGGACCCAAACGCTG TCCACA-3'; reverse, 5'-GGCAGCAGCCATCACATAATCCAGATG-3'. All experiments were performed in accordance with the Swiss Animal Protection laws and institutional guidelines.

Antibodies. For immunofluorescence analyses, the following primary and secondary antibodies were used: polyclonal anti-Oct4 (sc-8628, 1:150), monoclonal anti-Rnf2 (gift from H. Koseki, 1:400)²⁸, monoclonal anti-E-cadherin (Invitrogen, 1:250), polyclonal anti-Stra8 (rabbit, gift from M. Griswold, 1:1,000)²⁹, polyclonal anti-Sycp3 (rabbit, gift from C. Heyting, 1:500)³⁰, polyclonal anti-Mel18 (sc-10774, 1:50), polyclonal anti-Rybp (1:400)³¹, monoclonal anti-H2AK119u1 (Cell signaling, 1:500), polyclonal anti-H3K27me3 (gift from T. Jenuwein, 1:500)³², monoclonal anti-Ezh2 (Novocastra, 1:200), anti-goat IgG-Alexa 488, anti-rabbit IgG-Alexa 488, anti-mouse IgG-Alexa 555 and anti-rat Cy5. For ChIP analysis, anti-H3K4me3 (Millipore, 17-614), anti-H3K27me3 (Millipore, 07-449) and anti-Rnf2 (Active motif, 39663) were used.

Immunofluorescence. For whole-mount stainings, dissected gonads with mesonephroi were fixed for 15 min in 3% paraformaldehyde in PBS (pH 7.4) and permeabilized with 0.5% Triton X-100 in PBS for 20 min on ice. Fixed embryos were blocked overnight at 4 °C in PBS containing 0.1% Triton X-100, 10% BSA and 5% normal donkey serum, and were then incubated with primary antibodies in blocking solution overnight at 4 °C. Gonads were washed three times for 1 h in PBS containing 0.1% Triton X-100 and 2% BSA before application of secondary antibodies. For detection, secondary antibodies were diluted 1:500 in blocking solution and gonads were incubated overnight at 4 °C followed by three washing steps for 1 h in PBS with 0.1% Triton X-100. Gonads were stained briefly with DAPI and mounted in Vectashield (Vector). For gonadal section stainings, the posterior part of embryos or gonads with mesonephroi were frozen in Tissue-Tek Optical Cutting Temperature (OCT) compound (Sakura Finetek) on dry ice. Alternatively, the materials were fixed with 3% paraformaldehyde in PBS (pH 7.4) for 10 min, soaked in 30% sucrose solution overnight and embedded in OCT compound. Twelve-micron-thick cryo-sections were cut from frozen blocks with Microm HM355S. Cryo-sections were fixed with 3% paraformaldehyde for 10 min at room temperature (about 20 °C), permeabilized in 0.5% Triton X-100 in PBS for 4 min at 4 °C and blocked for 30 min in PBS containing 1% BSA at room temperature. Sections were incubated with primary antibodies in blocking solution overnight at 4 °C and subsequently washed three times for 10 min in PBS with 0.05% Tween-20. Incubation of secondary antibodies was done in the blocking solution for 1 h at room temperature.

Microscopy and image analysis. Immunofluorescence stainings of gonads were analysed using the Zeiss LSM 700 confocal microscope. For whole-mount gonads, images were acquired by using a tile function with a z-series of 1- μ m slices in ZEN software and whole image was reconstructed using the XUV-tools software. We counted the number of Oct4-positive cells using a spot function in Imapris (Bitplane) software.

Isolation of PGCs expressing the *Oct4*(*APE*)-GFP transgene by FACS. Dissected gonads were enzymatically disrupted using 0.025% trypsin at 37 °C for 8 min. Trypsin activity was inhibited by adding fetal calf serum in Hank's buffer salt solution without phenol red. Gonads were dispersed by pipetting and subjected to FACS. Embryos were processed individually for expression analysis. PGCs isolated from several embryos were pooled for ChIP analysis.

qRT-PCR. Total RNA was extracted from isolated PGCs or surrounding somatic cells from individual embryos using the PicoPure RNA Isolation Kit (KIT0202) according to the manufacturer's instructions (Stratagene) with the addition of 100 ng *Escherichia coli* ribosomal RNA as carrier. RT-PCR was performed using SuperScript III Reverse Transcriptase (Invitrogen) according to the manufacturer's protocol. qPCR reactions were performed with complementary DNA corresponding to 20 cells using the SYBR Green PCR Master Mix (Applied Biosystem) in an ABI Prism 7000 Real time PCR machine. All qPCR measurements were normalized to the endogenous expression level of *Tbp*. We performed qRT-PCR analyses on multiple pairs of *Rnf2*^{cko} and control littermates. Data in figures present technical replicates for pairs of genotypes. Primers used were as follows: *Rnf2* (forward, 5'-TTAGAAGTGGCAACAAAGAGTG-3'; reverse, 5'-CGCTTCATACTCATCAGAC-3'), *Oct4* (forward, 5'-GATGCTGTGAGCCA

AGGCAAG-3'; reverse, 5'-GGCTCCTGATCAACAGCATCAC-3'), *Nanog* (forward, 5'-CTTTCACCTATTAAGGTGCTTC-3'; reverse, 5'-TGGCATCGG TTCATCATGGTAC-3'), *Stra8* (forward, 5'-CAAAGCCTTGCTGTGTTA-3'; reverse, 5'-AAAGGTCTCCAGGCACTTCA-3'), *Rec8* (forward, 5'-CCAA CAAGGAGCTGGACTTC-3'; reverse, 5'-GGACAGCACAAGAGCAGAT-3'), *Sycp3* (forward, 5'-GTGTTGACAGAGTGGGAAAC-3'; reverse, 5'-GCTTT CATTCTCTGGCTCTGA-3'), *Aldh1a2* (forward, 5'-CCCTGACAGTGGCT TTGAGT-3'; reverse, 5'-CTGTGGGTTGAAGGGAGCTA-3'), *Crabp1* (forward, 5'-GCTTCGAGGAGGAGACAGTG-3'; reverse, 5'-CAGCTCTCGGGTCCAG TAAG-3'), *Crabp2* (forward, 5'-GCCGAGAAGTACCAATGAT-3'; reverse, 5'-GGAAGTCGTCTCAGGCAGTT-3'), *Tbp* (forward, 5'-TGCTGTTGGTGAT TGTTGGT-3'; reverse, 5'-AACTGGCTTGTGTGGGAAAG-3').

Expression profiling of PGCs and data analysis. We performed expression profiling on PGCs isolated from three pairs of *Rnf2*^{cko} and control littermates for each developmental time point. RNA was extracted from 500 PGCs isolated per embryo using the PicoPure RNA Isolation Kit (KIT0202) according to the manufacturer's instructions (Stratagene). The quality of the RNA was assessed using the Agilent 2100 Bioanalyzer and RNA 6000 Pico Chip. The extracted RNA was converted into OmniPlex Whole transcriptome amplification (WTA) cDNA libraries and amplified by WTA PCR using reagents supplied with the TransPlex WTA1 kit (Sigma) following the manufacturer's instructions with minor modifications. The obtained cDNA was purified using the GeneChip cDNA Sample Cleanup Module (Affymetrix). The labelling, fragmentation and hybridization of cDNA was performed according to Affymetrix instructions (GeneChip Whole Transcription Sense Target Labelling technical manual, Rev. 2) with minor modifications. Samples were hybridized to Affymetrix Mouse Gene 1.0 arrays. Microarray quality control and analysis was carried out in R 2.10.0 and Bioconductor 2.5. In brief, array quality was assessed using the 'arrayQualityMetrics' package. Raw data was read into R and normalized with RMA using the 'affy' package and differentially expressed genes were identified using the empirical Bayes method (*F* test) implemented in the LIMMA package. *P* values were adjusted for false discovery rate (FDR) using the Benjamini and Hochberg correction. Probe sets with a log₂ average contrast signal of at least 3, an adjusted *P* value of <0.05, and an absolute linear fold-change of at least 1.5-fold were selected. The *P* values reported for enriched Gene Ontology terms were obtained using Gostat (<http://gostat.wehi.edu.au>) (Supplementary Figs 7 and 8 and Supplementary Table 2). A list of genes belonging to the GO term 'meiosis' (GO:0007126) were obtained using R annotation packages (library(org.Mm.eg.db) and library(GO.db)) (Fig. 2d and Supplementary Table 1). For RNA-sequencing experiments, RNA was extracted from 500 PGCs per embryo using RNeasy Micro Kit (Qiagen). The RNA amplification and cDNA generation were performed using NuGEN Ovation RNA-seq System V2 (Part no. 7102) and sequencing libraries were prepared using TruSeq DNA Sample Preparation Kit (low-throughput protocol) (Part no. 15005180 Rev. C). Barcoded libraries were sequenced in one lane of Illumina HiSeq instrument. The resulting sequencing reads were filtered, aligned to Refseq gene models and weighted as described previously³³.

Treatment of gonads with agonist and antagonist of retinoic acid signalling. Both gonads per embryo were dissected and were cultured together or separately in a drop of Dulbecco's minimal eagle medium supplemented with 10% fetal calf serum at 37 °C in a humidified atmosphere of 5% CO₂ in air. ATRA (Sigma) and Win-18446 (ABCR) were dissolved in dimethylsulphoxide (DMSO). These compounds were added to culture media with concentration of 0.5 μ M for ATRA or 2 μ M for Win-18446. Control cultures were treated with DMSO vehicle as appropriate. For *in vivo* administration of Win-18446, 100 mg per ml stock solution (in DMSO) was diluted with oil and 2.5 mg was injected intraperitoneally into pregnant female mice at E10.5. Embryos were collected at E12.5. DMSO mixed with oil was injected into pregnant female mice for control experiments.

Specific target amplification (STA) qRT-PCR on single germinal vesicle oocytes. Germinal vesicle oocytes were isolated from ovaries and carefully washed with removing cumulus cells. STA was performed using the CellsDirect One-Step qRT-PCR Kit (11753-100) according to the manufacturer's instructions (Invitrogen). In brief, the oocytes were individually added to the reaction mix containing RNase Inhibitor (Ambion) and primers of target genes of interest. After amplification, the samples were treated with Exonuclease I (New England Biolabs) and used as template for qPCR. Primers used were as follows: *Rnf2* (forward, 5'-CAGGCCCATCCAACCTCTTA-3'; reverse, 5'-CAACAGTGGCA TTGCTGAA-3'), *Ssu72* (forward, 5'-GGTGTGCTCGAGTAACCAGAA-3'; reverse, 5'-CAAAGGAGCGGCACTGAAAC-3').

μ ChIP analysis. Small-scale ChIP experiment was performed as previously described²⁷ with some modifications. In brief, 15,000 FACS-sorted PGCs were cross-linked with 0.5% paraformaldehyde in PBS for 10 min at room temperature and quenched with 125 mM glycine. PGCs were then lysed in 50 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% SDS, and protease inhibitors. The cell lysate was sonicated in a

Diagenode Bioruptor to achieve a mean DNA fragment size of around 200–400 base pairs. After centrifugation, the supernatants were diluted with radio immunoprecipitation assay (RIPA) buffer to an equivalent of 2,500 PGCs and incubated with antibody-bound protein G-magnetic beads overnight at 4 °C. The beads were washed four times with the RIPA buffer and one time with 10 mM Tris HCl, pH 8.0, 10 mM EDTA buffer, and bead-bound complexes were incubated with complete elution buffer (20 mM Tris-HCl, pH 7.5, 5 mM EDTA, 50 mM NaCl, 1% SDS, 50 mg per ml proteinase K) at 68 °C for DNA elution, crosslink reversal and protein digestion. Finally, immunoprecipitated DNA was purified by phenol-chloroform extraction and ethanol precipitation and dissolved in MilliQ water for qRT-PCR analysis. Primers used were as follows: *Stra8* (forward, 5'-GTATCGCC GTAACTCCCAGA-3'; reverse, 5'-GCAGATGACCCTCACACAAG-3'), *HoxA9* (forward, 5'-GGAGGGAGGGGAGTAACAAA-3'; reverse, 5'-TCACCTCGCCT AGTTTCTGG-3').

28. Atsuta, T. *et al.* Production of monoclonal antibodies against mammalian Ring1B proteins. *Hybridoma* **20**, 43–46 (2001).
29. Zhou, Q. *et al.* Expression of stimulated by retinoic acid gene 8 (*Stra8*) and maturation of murine gonocytes and spermatogonia induced by retinoic acid in vitro. *Biol. Reprod.* <http://dx.doi.org/10.1095/biolreprod.107.064337> (2008).
30. Lammers, J. H. *et al.* The gene encoding a major component of the lateral elements of synaptonemal complexes of the rat is related to X-linked lymphocyte-regulated genes. *Mol. Cell. Biol.* **14**, 1137–1146 (1994).
31. García, E., Marcos-Gutierrez, C., del Mar Lorente, M., Moreno, J. C. & Vidal, M. RYBP, a new repressor protein that interacts with components of the mammalian Polycomb complex, and with the transcription factor YY1. *EMBO J.* **18**, 3404–3418 (1999).
32. Peters, A. H. *et al.* Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol. Cell* **12**, 1577–1589 (2003).
33. Tippmann, S. C. *et al.* Chromatin measurements reveal contributions of synthesis and decay to steady-state mRNA levels. *Mol. Syst. Biol.* **8**, 593 (2012).

Chapter 3

Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa

Serap Erkek^{1,2,3}, Mizue Hisano¹, Ching-Yeu Liang^{1,2}, Mark Gill¹, Rabih Murr¹, Jürgen Dieker⁴, Dirk Schübeler^{1,2}, Johan van der Vlag⁴, Michael B Stadler^{1,3} and Antoine H F M Peters^{1,2}

1. Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, 4058 Basel, Switzerland

2. Faculty of Sciences, University of Basel, 4056 Basel, Switzerland

3. Swiss Institute of Bioinformatics, Basel, Switzerland

4. Department of Nephrology, Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands

Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa

Serap Erkek¹⁻³, Mizue Hisano¹, Ching-Yeu Liang^{1,2}, Mark Gill¹, Rabih Murr¹, Jürgen Dieker⁴, Dirk Schübeler^{1,2}, Johan van der Vlag⁴, Michael B Stadler^{1,3} & Antoine H F M Peters^{1,2}

In mammalian spermatozoa, most but not all of the genome is densely packaged by protamines. Here we reveal the molecular logic underlying the retention of nucleosomes in mouse spermatozoa, which contain only 1% residual histones. We observe high enrichment throughout the genome of nucleosomes at CpG-rich sequences that lack DNA methylation. Residual nucleosomes are largely composed of the histone H3.3 variant and are trimethylated at Lys4 of histone H3 (H3K4me3). Canonical H3.1 and H3.2 histones are also enriched at CpG-rich promoters marked by Polycomb-mediated H3K27me3, a modification predictive of gene repression in preimplantation embryos. Histone variant-specific nucleosome retention in sperm is strongly associated with nucleosome turnover in round spermatids. Our data show evolutionary conservation of the basic principles of nucleosome retention in mouse and human sperm, supporting a model of epigenetic inheritance by nucleosomes between generations.

In mammals, fusion of two morphologically distinct gametes—oocytes and spermatozoa—leads to the formation of totipotent embryos. Acquisition of totipotency is thought to be mediated by extensive epigenetic reprogramming of parental genomes, and this in turn affects DNA methylation, histone modifications and possibly replication timing and transcriptional activity in a parent-specific manner¹⁻⁴. It is currently unclear to what extent differential reprogramming of maternal and paternal genomes is due to differences in chromatin states inherited from the oocyte and spermatozoon⁴⁻¹¹. Beyond DNA methylation^{1,2,6,12}, it is unknown which types of parental chromatin states are maintained or reprogrammed in early embryos. If certain parental chromatin states were to escape reprogramming in the early embryo, such information could constitute an intrinsic intergenerational epigenetic program directing gene expression in the next generation¹³. If these chromatin states also were to escape reprogramming during gametogenesis, the inheritance program would function transgenerationally¹³. An increasing number of studies point to inter- or transgenerational transmission of acquired phenotypic traits that are related to temporal exposure of grandparents or parents to alternative instructive environmental cues¹⁴⁻¹⁸. Mechanistically, such phenotypic changes may be related to (transient) alterations of an intrinsic inheritance program.

Whether histones and associated post-translational modifications have a role in maternal and paternal transmission of intrinsic or acquired epigenetic information is largely unknown¹³. In many metazoans, male germ cells undergo an extensive chromatin remodeling process in their final differentiation into sperm, during which genomic DNA becomes newly packaged into a highly condensed configuration by sperm-specific proteins. In mammals, removal of

histones is generally not complete^{10,11,19-24}. Furthermore, remaining histones have been reported to stay associated with the paternal genome during *de novo* chromatin formation in the zygote after fertilization⁹.

We and others recently showed that histones retained in human sperm are to some extent enriched at the regulatory sequences of genes^{10,11}. We also demonstrated that H3K4me3- and H3K27me3-marked histones are retained at the promoters of specific sets of genes in mouse spermatozoa¹¹. The extent of evolutionary conservation of nucleosome retention at gene regulatory sequences in spermatozoa and the mechanistic principles of such retention are, however, unknown.

To address conservation and to dissect the molecular logic underlying nucleosome retention, we determined the genome-wide nucleosome occupancy in mouse spermatozoa, which contain only 1% residual histones. We show here that combinatorial effects of sequence composition, histone variants and histone modifications determine the packaging of sperm DNA. Nucleosomes in sperm mainly localize to unmethylated CpG-rich sequences in a histone variant-specific manner and are differentially modified. Comparison of histone-variant profiles between postmeiotic round spermatids and sperm suggests that the retention of variant-specific nucleosomes in sperm is linked to levels of nucleosome turnover in haploid round spermatids.

RESULTS

Nucleosomes localize at GC-rich sequences in mouse sperm

To assess the potential of paternal epigenetic inheritance by nucleosomes in mice, we first aimed to determine the location of

¹Friedrich Miescher Institute for Biomedical Research, Basel, Switzerland. ²Faculty of Sciences, University of Basel, Basel, Switzerland. ³Swiss Institute of Bioinformatics, Basel, Switzerland. ⁴Department of Nephrology, Nijmegen Centre for Molecular Life Sciences, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. Correspondence should be addressed to A.H.F.M.P. (antoine.peters@fmi.ch).

Received 27 February; accepted 29 April; published online 16 June 2013; corrected after print 12 July 2013; doi:10.1038/nsmb.2599

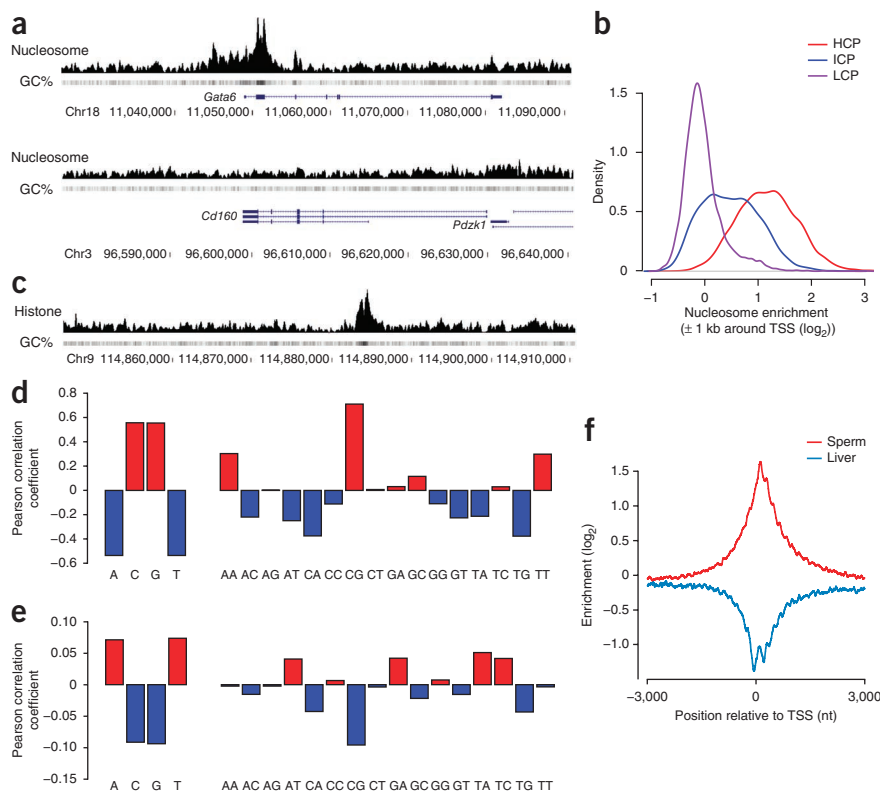
Figure 1 Nucleosome occupancy in sperm is highly dependent on CpG composition.

(a) Nucleosome occupancy and GC percentage at representative CGI and non-CGI loci in mouse sperm. Chr, chromosome. (b) Density plot showing the distribution of nucleosome enrichment ± 1 kb around TSSs of genes classified according to GC content as high (HCP), intermediate (ICP) or low (LCP) CpG promoters. (c) Nucleosome occupancy and GC percentage at an intergenic region in sperm. (d,e) Correlation of single-nucleotide frequencies (left) and dinucleotide frequencies (right) with nucleosome enrichment in sperm (d) and in mouse liver³² (e) in 1-kb regions tiling the mouse genome. (f) Average profiles for nucleosome occupancy in mouse sperm and liver³² ± 3 kb around TSSs of genes.

nucleosomes in spermatozoa. We isolated motile spermatozoa from caudal epididymi and performed deep sequencing of DNA associated with mononucleosomes prepared by micrococcal nuclease (MNase) digestion of sperm chromatin. Genome-wide analyses indicated an approximately ten- and two-fold overrepresentation of nucleosomes at promoter regions and exons, respectively, whereas nucleosomes were underrepresented at introns and repeat regions (Supplementary Fig. 1a,b). We observed promoter enrichment at many but not all genes (Fig. 1a). Classification of promoters according to their GC content, CpG ratio and length of CpG-rich region²⁵ revealed that high-CpG and intermediate-CpG promoters are highly and moderately enriched in nucleosomes, respectively, whereas most promoters with low CpG content lack nucleosomes (Fig. 1b). Nucleosomal enrichment is not restricted to CpG-island (CGI) promoters but is also detected at intra- and intergenic CGIs as well as within GC-rich simple repeat sequences (Fig. 1c and data not shown).

To investigate whether nucleosomal occupancy in sperm correlates with a specific sequence composition, we determined single-nucleotide frequencies in 1-kilobase (kb) windows tiled throughout the genome. Whereas G and C correlate positively with nucleosome occupancy genome wide, A and T do not (Fig. 1d). We next assessed the contribution of different dinucleotides to nucleosome occupancy, independent of single-nucleotide frequencies, by calculating the ratio between observed and expected frequencies for each dinucleotide. Notably, these analyses revealed that, predominantly, the CpG dinucleotide contributes to sequence-related nucleosomal packaging of sperm DNA (Fig. 1d), whereas the GpC dinucleotide has almost no contribution. The ApA and TpT dinucleotides contribute moderately.

To establish whether the observed CpG dinucleotide association reflects an intrinsic DNA sequence preference for nucleosome formation, we reanalyzed *in vitro* nucleosome-reconstitution data of histone octamers assembled onto yeast genomic DNA²⁶. We observed a strong contribution of G and C to *in vitro* nucleosome formation, as reported before (in ref. 27). However, we observed no specific contributions of either CpG or GpC dinucleotides (Supplementary Fig. 2a,b). Thus, the strong association of CpG density with nucleosome retention in mouse sperm does not reflect an intrinsic preference of nucleosomes for CpG-rich DNA. Instead, it represents a new feature



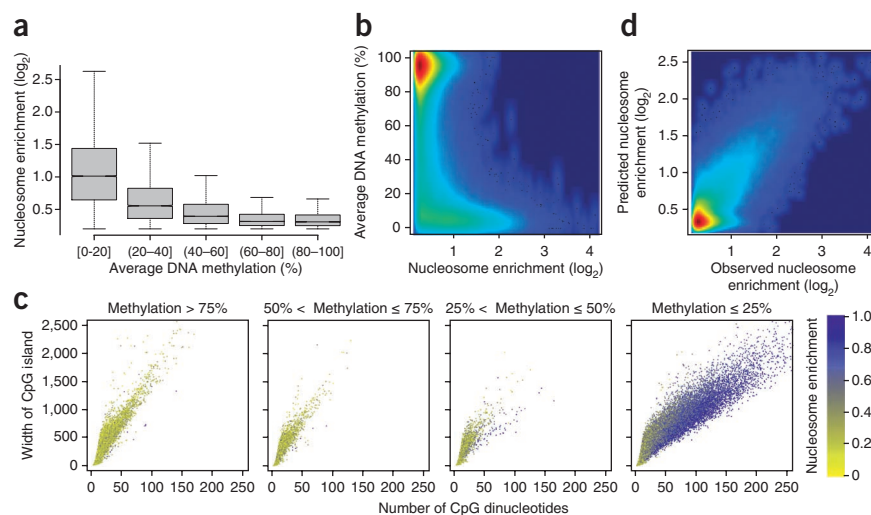
of CGIs in genome function executed during mouse male germ-cell development²⁸. Motif analysis did not reveal any specific sequence compositions other than a strong correlation to GC composition (Supplementary Fig. 1c).

Nucleosomes localize at unmethylated CpG-rich DNA in sperm

The nucleosomal occupancy at CGIs in sperm markedly contrasts with the depletion of nucleosomes at CGI promoters in somatic cells^{29–31}. Indeed, we observed extensive nucleosomal depletion around transcriptional start sites (TSSs) and a clear inverse correlation between nucleosome occupancy and CpG frequency in mouse liver³² (Fig. 1e,f). In somatic cells, however, nucleosomes are not depleted at CGI promoters repressed by Polycomb-group proteins or by DNA methylation³³. Therefore, to investigate whether nucleosomes are preferentially retained at CGIs that are DNA methylated in sperm, we performed bisulfite conversion and high-throughput sequencing of sperm DNA associated with nucleosomes³⁴. In contrast to our expectation, methylated genomic regions were devoid of nucleosomes in sperm (Fig. 2a). We observed a similar inverse relationship by using genome-wide shotgun bisulfite sequencing data from mouse sperm (Fig. 2b)⁶. This exclusive inverse relationship is well illustrated at imprinting control regions (ICRs) in mouse sperm. Whereas paternal ICRs regulating somatic expression of genes such as *H19*, *Dlk1*, *Gtl2* (official symbol *Meg3*) and *Rasgrf1* (ref. 35) are methylated and devoid of nucleosomes, ICRs controlling maternally imprinted gene clusters (for example, *Kcnq1ot1*, *Snrpn* and *Peg10*) are unmethylated and contain nucleosomes (Supplementary Fig. 3). Furthermore, GC-rich retroelements such as LINE1 elements that are methylated in sperm and become demethylated after fertilization¹ lack nucleosomes in sperm (data not shown). These data are compatible with a model in which DNA methylation established early during male germ-cell development³⁶ prevents nucleosome retention during spermiogenesis.

Figure 2 Nucleosome occupancy correlates negatively with DNA methylation in sperm.

(a) Box plot showing distributions of nucleosome enrichments in 1-kb regions of different DNA methylation states genome wide (central bar, median; lower and upper box limits, 25th and 75th percentiles, respectively; whiskers, 1.5 times the interquartile range from the 25th and 75th percentiles). Parentheses and brackets indicate noninclusive and inclusive endpoints in ranges, respectively. (b) Scatter plot showing the correlation of nucleosome occupancy with average DNA methylation according to a previous study⁶, in 1-kb windows genome wide. (c) Relationship between number of CpGs in CGIs and width of CGIs as a function of nucleosome enrichment in sperm. CGIs⁴⁸ were grouped into four classes according to their DNA methylation status in sperm⁶. (d) Correlation of observed to predicted nucleosome enrichment, calculated by a linear model integrating CpG dinucleotide frequency and DNA methylation status in 1-kb windows ($R = 0.789$).



By analyzing sequence characteristics of CGIs and their DNA methylation states, we found strong positive correlations between nucleosomal enrichment and the number and density of CpG dinucleotides within CGIs devoid of DNA methylation (Fig. 2c). Using a linear mathematical model, we can predict nucleosome occupancy in mouse sperm as a function of CpG dinucleotide frequency and DNA methylation level (Fig. 2d).

We and others previously showed that retained histones are not randomly distributed in human sperm but are to some extent enriched at GC-rich regulatory elements of genes^{10,11,37}. As for mice, we observed an inverse relationship between the degree of nucleosomal occupancy and the level of DNA methylation in human sperm³⁸ (Supplementary Fig. 2c–f). Thus, these analyses demonstrate that nucleosome retention at unmethylated CGIs is conserved between mouse and human spermatozoa.

Histone H3 variant-specific occupancy at CGIs in mouse sperm

The unique nucleosomal organization in sperm, highly distinct from that in somatic cells^{29–31}, emphasizes extensive chromatin-remodeling processes occurring during the formation of spermatozoa. Given the important roles of histone variants in transcription, cellular differentiation, reproduction and development^{39–41}, we asked whether canonical H3.1 and H3.2 and variant H3.3 histones may serve specific functions in nucleosome eviction versus retention

during spermiogenesis. We performed western blot analysis with antibodies specific for either histones H3.3 (Supplementary Fig. 4a) or histones H3.1 and H3.2 (referred to as H3.1/H3.2 because the antibody recognizes an epitope shared by histones H3.1 and H3.2)⁴². In comparison to proliferating embryonic stem cells (ESCs) and even to quiescent aging neurons⁴³, in round spermatids and sperm histone H3.3 is incorporated into chromatin to a higher extent than are H3.1/H3.2. This suggests an extensive and rapid replacement of canonical histones by the H3.3 variant, presumably upon entry into meiotic prophase and/or during spermatid differentiation (Fig. 3a). In sperm, histone-H3.3 profiles from chromatin immunoprecipitation and sequencing (ChIP-seq) are highly similar to nucleosomal profiles, whereas H3.1/H3.2 profiles are not (Fig. 3b). Consistently with this, histone H3.3 enrichments are well predicted by the linear model, and this suggests a CpG density-linked retention mechanism for histone H3.3-containing nucleosomes (Fig. 3c). Regions containing H3.1/H3.2 histones are, in contrast, systematically underestimated by the model, a result suggesting that the retention of canonical and H3.3-variant histones may be differentially regulated.

Nucleosome turnover in round spermatids

To understand the timing and mechanisms of chromatin remodeling, we profiled the occupancy of histone H3.3 and H3.1/H3.2 nucleosomes and measured levels of mRNA transcripts by ChIP-seq and RNA sequencing in round spermatids. We observed a widespread reduction in H3.1/H3.2-nucleosome occupancy around TSSs of genes in round spermatids in contrast to sperm (Fig. 4a). We next classified

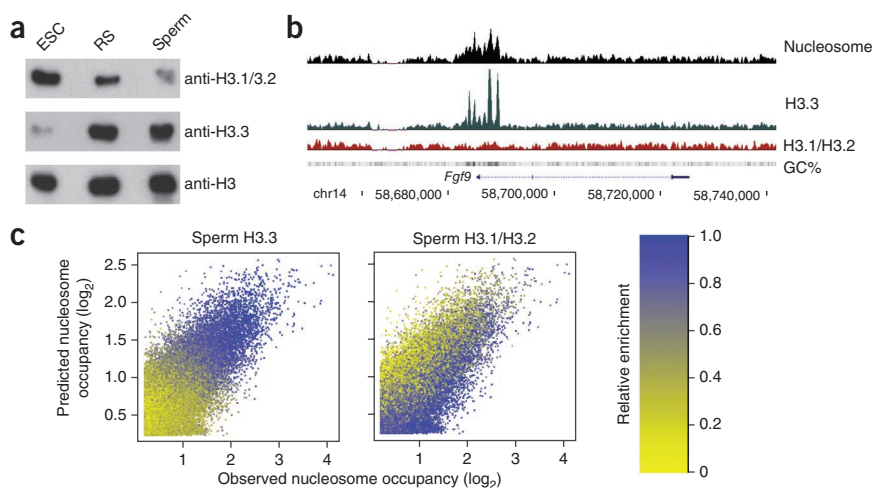


Figure 3 Histone variant-specific packaging of sperm DNA. (a) Western blots showing relative levels of chromatin-bound H3.1/H3.2, H3.3 and total H3 in embryonic stem cells (ESC), round spermatids (RS) and sperm. (b) Occupancy of nucleosomes, H3.3 and H3.1/H3.2 histones and GC percentage at the *Fgf9* locus in sperm. (c) Scatter plots showing the correlation between observed and predicted nucleosome occupancies (1-kb windows) in relation to relative enrichment of H3.3 (left) and H3.1/H3.2 (right) in sperm.

Figure 4 CpG density and gene expression associate with nucleosome eviction in round spermatids. **(a)** Average profiles of H3.3 and H3.1/H3.2 enrichments ± 3 kb around the TSS and transcriptional end site (TES) in sperm and round spermatids. **(b)** Average profiles of H3.1/H3.2 and H3.3 enrichments around TSSs and transcriptional end sites in round spermatids. Genes were classified according to expression status in round spermatids and the percentage of CpGs within ± 1 -kb windows around TSSs of genes (left, CpG% < 3 ; right, CpG% ≥ 3).

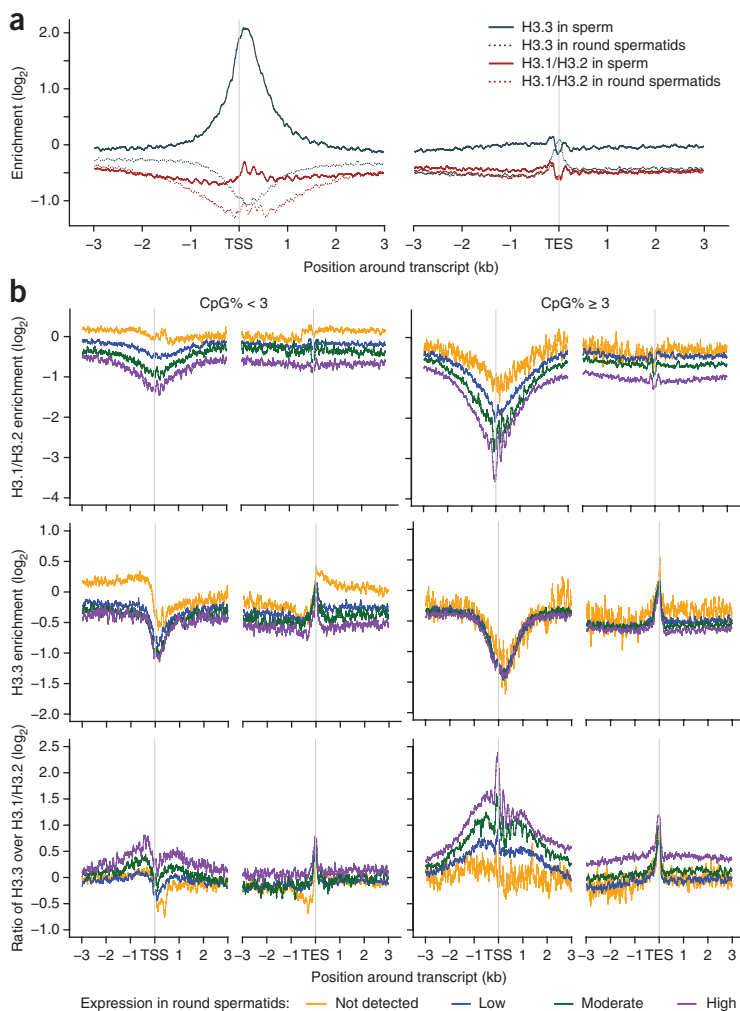
gene promoters according to CpG density and RNA-transcript levels of associated genes (Fig. 4b). For expressed genes, we observed eviction of H3.1/H3.2 nucleosomes around TSSs of CpG-rich ($\geq 3\%$ CpG) and CpG-poor ($< 3\%$ CpG) genes that correlated well with mRNA levels of associated genes. For moderately to highly expressed CpG-rich genes, we also observed clear positioning of remaining nucleosomes around TSSs. These data suggest a transcription-coupled eviction of canonical histones.

For nonexpressed genes, we measured low levels of eviction of canonical histones at CpG-rich TSS regions but not at CpG-poor TSS regions (Fig. 4b). This finding is consistent with studies reporting nucleosome depletion around silent CGI promoters in somatic cells^{30–32}. Depletion of H3.1/H3.2 nucleosomes around TSSs in spermatids is more pronounced than that of histone-H3.2 hemagglutinin-tagged nucleosomes in ESCs (Fig. 4b and Supplementary Fig. 4b)⁴⁴. Possibly, this is due to progressive loss of canonical histones during transcription in post-replicative germ cells.

For histone-H3.3 nucleosomes, we also measured some depletion around TSSs that was more pronounced downstream of TSSs at moderately and highly expressed genes (Fig. 4a,b). Comparison of occupancy levels of histones H3.3 and H3.1/H3.2 suggests extensive transcription-coupled eviction of canonical histones and subsequent replacement by H3.3 nucleosomes in round spermatids. We interpret the ratio of histones H3.3 to H3.1/H3.2 as a proxy measure for nucleosome turnover.

Control of H3 variant-specific occupancy at CGIs in sperm

To understand the relationship between histone variant-specific nucleosome turnover in round spermatids and retention in sperm, we compared the level of occupancy for both variants in both cell types, at regions with nucleosomal enrichments in spermatozoa. We observed that regions that are highly and intermediately enriched for histone H3.3-containing nucleosomes in sperm are actually depleted of such nucleosomes in round spermatids, a result suggesting dynamic redistribution *in cis* or *de novo* incorporation of H3.3 nucleosomes later during spermatid differentiation, for example, in late-round or elongating spermatids (Fig. 5a). In contrast, H3.1/H3.2 nucleosomes are predominantly detected at weak nucleosomal-peak regions in spermatozoa. Furthermore, such local H3.1/H3.2 enrichments in sperm highly resemble those in round spermatids, and this suggests



that H3.1/H3.2 nucleosomes retained in sperm largely reflect reduced turnover of canonical H3.1/H3.2 histones in spermatids (Fig. 5a).

We next assessed the connection between CpG density (Fig. 3) and nucleosome turnover in spermatids (Fig. 4) in relation to histone variant-specific nucleosome retention at promoter regions of genes in sperm (Fig. 5b). CGI promoters (with $\geq 3\%$ CpG) that undergo intermediate to high nucleosomal turnover in round spermatids have high levels of histone H3.3 in sperm. In contrast, non-CGI promoters ($< 3\%$ CpG) are subjected to low to intermediate nucleosome turnover and are relatively enriched for H3.1/H3.2. Finally, a group of CGI promoters is enriched for both H3.1/H3.2 and H3.3 (Fig. 5b). These promoters are generally characterized by intermediate turnover in spermatids. Together, these data show that CpG density and the extent of turnover in spermatids relate to the identity of histones retained in sperm.

H3K27me3 associates with H3.1/H3.2 retention in sperm

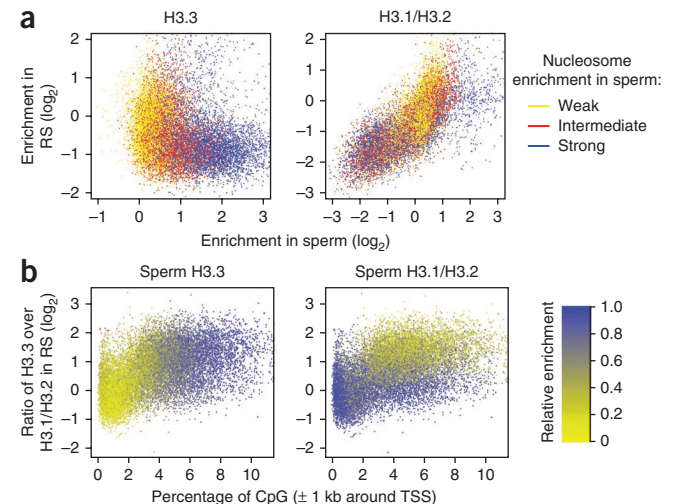
To study whether histone modification states may affect nucleosome dynamics during spermiogenesis, we performed ChIP-seq for H3K4me3 and H3K27me3, two modifications associated with CGIs in somatic cells. We measured comparable enrichments around TSSs for both modifications in round spermatids and sperm (Supplementary Fig. 5a), a result indicating propagation of the modification state during spermiogenesis. CGI promoters (with $\geq 3\%$ CpG) containing histone-H3.3 nucleosomes are generally marked by H3K4me3 in sperm (Fig. 6a). A fraction of CGIs with intermediate H3K4me3 levels are

Figure 5 Extent of nucleosome turnover in round spermatids relates to histone variant–specific retention in sperm. (a) Scatter plots showing the correlation between H3.3 (left) and H3.1/H3.2 (right) enrichments in sperm versus round spermatids (RS) at genomic regions enriched for nucleosomes in sperm. Enriched regions are classified as weak, intermediate and strong according to their relative occupancy by nucleosomes in sperm. (b) Scatter plots showing the correlation between percentage of CpGs at TSSs (± 1 kb) and nucleosome turnover (ratio of H3.3 over H3.1/H3.2) in round spermatids in relation to relative enrichment of H3.3 (left) and H3.1/H3.2 (right) in sperm.

highly enriched in H3K27me3, and this indicates the presence of bivalent promoters in sperm (Fig. 6b; cluster 4 in Supplementary Fig. 5b). Notably, such H3K4me3 and H3K27me3 double-marked CGI promoters also show enrichment for H3.1/H3.2 histones in sperm (Fig. 6a,b). These data suggest the presence of bivalent promoters in sperm that contain a mixture of H3.3 and H3.1/H3.2 nucleosomes. The data further suggest that Polycomb proteins and/or PRC2-mediated H3K27me3 suppress, at least in part, the default eviction of H3.1/H3.2 histones at CGIs in round spermatids and consequently promote the retention of preexisting canonical histones during chromatin remodeling in elongating spermatids. In accordance, CGI promoters with low to intermediate levels of nucleosome turnover in spermatids are H3K27me3 positive in spermatids and in sperm (Supplementary Fig. 5d).

Determinants of nucleosome retention in sperm

On the basis of sequence composition, occupancy levels of nucleosomes and histone variants and histone modifications at gene promoters as well as expression states, we can classify genes into five different clusters (Fig. 6c) that correlate well with different gene functions in cellular homeostasis (clusters 2 and 3), germ-cell and



embryonic development (clusters 1 and 4, respectively) and stimulus perception and host defense (cluster 5) (Supplementary Table 1 and Supplementary Figs. 5b and 6a–j).

To quantify the extent by which expression and the different chromatin characteristics such as histone variants and modifications measured in round spermatids as well as CpG density contribute to nucleosome occupancy in sperm, we performed a variance partitioning analysis for promoter regions (Fig. 6d). Combining all of these variables can explain a total of 79.4% and 70% of the variance in H3.3 and H3.1/H3.2 occupancies in sperm.

For histone-H3.3 occupancy in sperm, CpG density of promoters has, as expected, the highest unique contribution, whereas H3.3

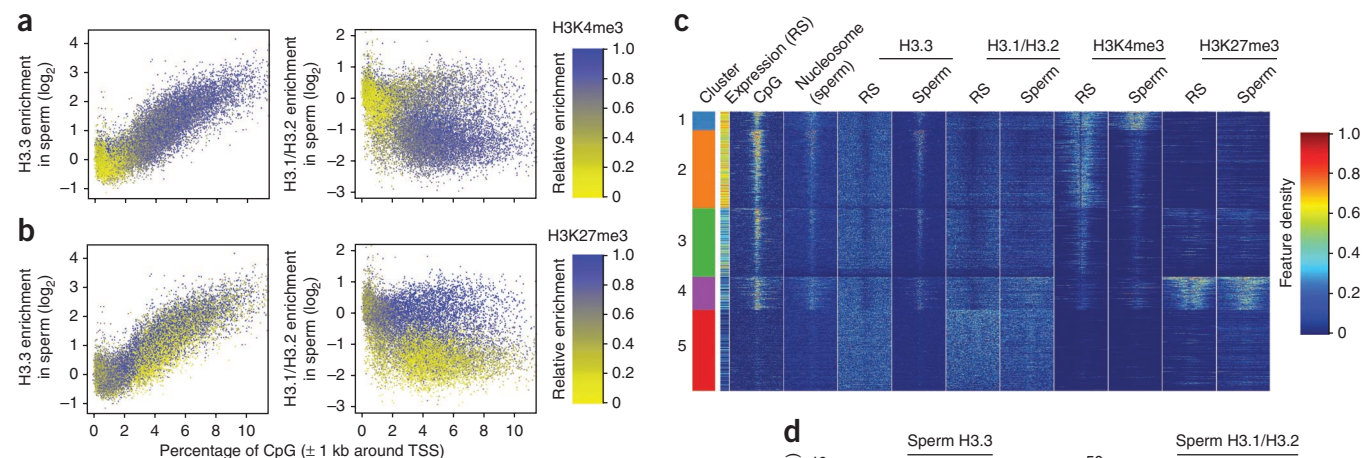


Figure 6 Combinatorial effects of CpG density, histone variants and histone modifications underlie unique packaging of sperm DNA. (a,b) Scatter plots showing the correlation of the percentage of CpGs with enrichment of variant and canonical H3 histones (TSS ± 1 kb) in sperm in comparison to the enrichment of H3K4me3 (a) and H3K27me3 (b). (c) Heat map of genes, illustrating expression status in round spermatids (RS), CpG density and nucleosome coverage in sperm and histone variant and modification coverage around TSSs (± 3 kb) in round spermatids and sperm. Feature density shows the scaled read densities from ChIP-seq experiments. Genes ($n = 19,180$) were grouped by k means into five clusters (1 to 5) containing 1,346, 5,358, 4,468, 2,902 and 5,106 genes, respectively. One thousand genes were randomly selected for visualization. (d) Variance partitioning analysis (details in Online Methods) assessing the unique contribution of different variables (most in round spermatids) to the relative enrichments of H3.3 (left) and H3.1/H3.2 (right) around TSSs (± 1 kb) in sperm. Combinatorial effects refer to variation common to the different combinations of variables included.

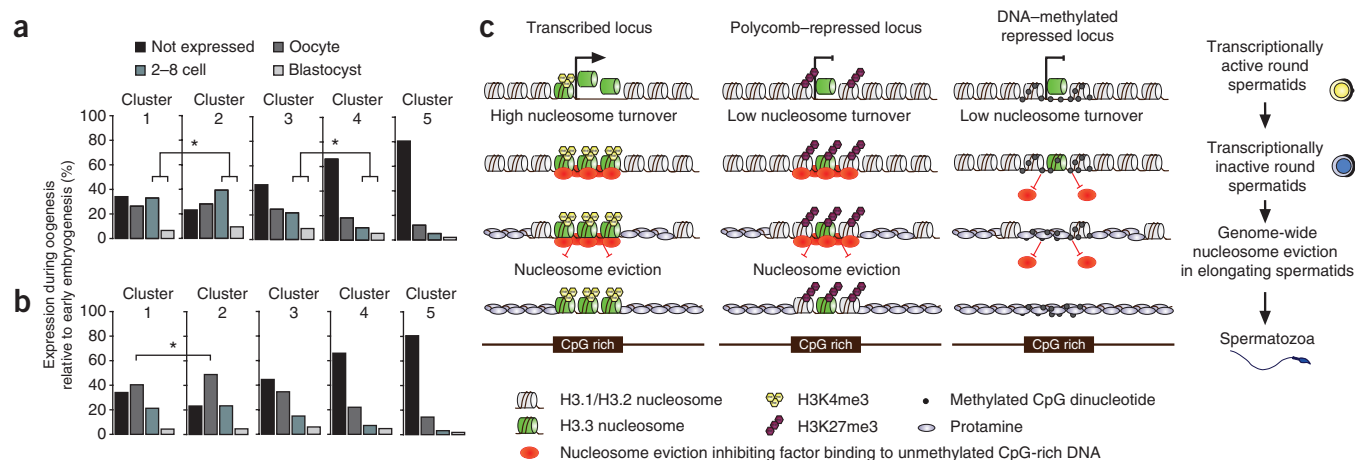


Figure 7 Model of nucleosome retention during spermiogenesis. **(a, b)** Expression states of genes belonging to different clusters (**Fig. 6c**) during oogenesis and early embryogenesis. We classified genes as ‘not expressed’, ‘oocyte’, ‘2–8 cell’ and ‘blastocyst’ as described before¹¹. Embryonic expression was classified according to the first expression stage during development. Genes transcribed in oocytes and two-cell to eight-cell embryos or in oocytes and blastocyst embryos were classified as 2–8 cell or blastocyst **(a)** or as oocyte **(b)**. We matched 14,032 of 19,180 RefSeq genes for expression during oogenesis and embryogenesis⁴⁵. Numbers of genes in each cluster are 1,097, 4,419, 3,431, 2,417 and 2,668, respectively. $*P < 1.0 \times 10^{-6}$ (by Fisher’s exact test). **(c)** Model of nucleosome turnover and retention during spermiogenesis. H3K4me₃, become stably incorporated at unmethylated CGIs in response to cessation of global histone turnover and transcription in late round spermatids. Reduced turnover of H3K27me₃-marked H3.1/H3.2 nucleosomes in round spermatids promotes retention of such nucleosomes in spermatozoa. Nucleosome retention at unmethylated CGIs would be mediated by unknown CGI-binding factors suppressing nucleosome eviction or could alternatively result from a reduced affinity of protamines for CG-rich DNA. In the presence of DNA methylation, protection against eviction is lost, owing to the inability of the CGI-binding factor(s) to bind methylated DNA.

occupancy in spermatids has a small unique contribution (**Fig. 6d**; clusters 1–4 in **Fig. 6c**). Notably, 78% of the variance is explained when CpG density and H3.3 occupancy in spermatids are used together as the only variables in the partitioning analysis. These data suggest that extensive histone H3.1/H3.2 turnover and ensuing H3.3 deposition at CGIs in round spermatids contribute to H3.3 retention at such promoters in sperm.

In contrast, H3.1/H3.2 enrichments in sperm mostly relate to H3.1/H3.2 enrichments in round spermatids (**Fig. 6d**; cluster 5 in **Fig. 6c**). Moreover, CGI promoters marked by H3K27me₃ in round spermatids preferentially retain H3.1/H3.2 in sperm (**Fig. 6d**; cluster 4 in **Fig. 6c**). When H3.1/H3.2 and H3K27me₃ enrichments in round spermatids are taken as the only variables, 68% of the total of 70% variance is explained. These quantifications suggest that low nucleosome turnover at H3K27me₃-marked promoters in spermatids substantially contributes to H3.1/H3.2 retention in sperm.

When performing the variance partitioning analysis genome wide at 1-kb windows that do not intersect with TSS regions and any other CGIs, we observed that only 30% to 42% of variance for histones H3.3 and H3.1/H3.2, respectively, is explained by CpG density and chromatin characteristics measured in round spermatids (**Supplementary Fig. 5e**). Nonetheless, occupancy levels of histone variants in sperm relate well to the occupancy of the corresponding variants in spermatids. This relationship supports a model of nucleosome retention without major remodeling in *cis* during spermatid maturation and protamine incorporation.

H3K27me₃ associates with gene repression in early embryos

To assess the potential of nucleosomes and associated modifications retained in sperm for regulating transcription in the next generation, we analyzed the expression of genes belonging to the five different clusters (shown in **Fig. 6c**) in oocytes and in preimplantation embryos^{11,45}. We observed that housekeeping genes in cluster 2 are

significantly more likely to be *de novo* transcribed (48.7% compared to 39.6%; **Fig. 7a**) in early embryos as well as in oocytes (48.8% compared to 40.4%; **Fig. 7b**) than are genes belonging to cluster 1 enriched for germline functions. Notably, H3K4me₃ is more enriched around TSSs of cluster 1 genes than cluster 2 genes in sperm (**Fig. 6c** and **Supplementary Fig. 5b**). Analogously, genes of clusters 3 and 4 have similar H3K4me₃ enrichments in sperm (**Supplementary Fig. 5b**), yet they display significantly different expression states in early embryos (**Fig. 7a, b**). These data suggest a rather limited potential, if any, of H3K4me₃ nucleosomes in sperm to predetermine transcription in early embryos (**Fig. 7a**). This may relate to the prevalent H3K4 trimethylation at CGIs in spermatids and ESCs and during somatic differentiation that is independent of their transcriptional status²⁸ (**Supplementary Fig. 5c**).

In contrast, only ~16% of CGI promoters marked by H3K27me₃ (and H3K4me₃) in sperm (cluster 4) are expressed in preimplantation embryos (**Fig. 7a**). Moreover, many Polycomb-target genes in sperm are similarly modified by H3K27me₃ in ESCs (**Supplementary Fig. 5c**). These data support a model of H3K27me₃ mediating epigenetic inheritance of transcriptional repression between generations.

DISCUSSION

The role of histones and associated post-translational modifications in maternal and paternal transmission of epigenetic information is currently unknown. Here we describe a systematic genome-wide characterization of chromatin states in mouse spermatids and spermatozoa. We show that the histones retained to a level of 1% in the sperm of mice¹¹ are highly enriched at CGIs that are not methylated at the underlying sequence. Likewise, we demonstrate that in human sperm the CGIs in which the 10–15% of residual histones are somewhat enriched^{10,11,37} are also unmethylated. Because CGIs are frequently associated with gene promoter function, the evolutionarily conserved presence of modified nucleosomes at unmethylated CGIs in the sperm

of mammals suggests a central role for CGIs and retained nucleosomes in paternal intrinsic epigenetic inheritance between generations¹³.

Our data show that CGIs in mouse sperm generally contain the variant histone H3.3 protein (clusters 1–4 in **Fig. 6c**), whereas canonical H3.1/H3.2 proteins are present at only some CGIs (cluster 4 in **Fig. 6c**). Comparative analysis of chromatin states in round spermatids and sperm suggests that the level of nucleosome turnover in round spermatids determines the type of H3 histone retained at CGIs in sperm.

In round spermatids, we measured an extensive eviction of canonical nucleosomes around the TSSs of genes and replacement by histone H3.3-containing nucleosomes. The extent of nucleosome turnover positively correlates with the transcriptional activity of associated gene promoters, as in somatic cells^{32,44}. CGIs in round spermatids also show transcription-independent nucleosome turnover, as observed in somatic cells^{30–32}, and this possibly reflects dynamic competition between nucleosomes and transcription factors for CGI binding. Notably, the overall extent of replacement of H3.1/H3.2 to H3.3 around TSSs is more pronounced in postmitotic round spermatids than, for example, in replicating ESCs. These data suggest that male germ cells undergo extensive remodeling of their chromatin during the approximately 2 weeks after their entry into meiosis and in subsequent differentiation as haploid spermatids.

In sperm, histone H3.3 is enriched at most CGIs, as a reflection of turnover in spermatids. In contrast, H3.1/H3.2 is present at CGIs with low to intermediate nucleosome turnover in spermatids and that are marked by H3K27me3 in round spermatids and sperm. These findings suggest that PRC2 proteins directly or indirectly through H3K27me3 inhibit nucleosome turnover in round spermatids, thereby promoting H3.1/H3.2 retention in sperm. At non-CGI promoter genes (cluster 5 in **Fig. 6c**), we observed only minor enrichment of H3.1/H3.2 nucleosomes around the TSSs of some genes, and this supports the notion of poor nucleosome retention, if any, at non-CGI promoters.

Currently, the mechanisms driving nucleosome retention versus eviction during spermiogenesis are unknown. Our findings support a model in which histone-H3.3 nucleosomes present at CGIs in sperm become stably incorporated into chromatin and marked by H3K4me3 in late round spermatids in response to a global cessation of histone turnover and transcription (**Fig. 7c**). Reduced nucleosome turnover, as observed at H3K27me3-marked CGIs in spermatids, would promote retention of canonical H3.1/H3.2 in sperm. This model entails that CpG-rich DNA would somehow resist loading of transition proteins and protamines in elongating spermatids, thereby enabling nucleosome retention at CGIs as measured in sperm. Resistance to loading could be mediated by CGI-binding proteins binding unmethylated DNA and protecting nucleosomes locally from eviction. Alternatively, it could reflect a reduced intrinsic affinity of protamines for CG-rich DNA. A variation on this model is that transcription factors, chromatin factors and histone-H3.3 nucleosomes would continue to compete for binding to CGIs during the histone-to-protamine exchange process in elongating spermatids. This dynamic process may block protamine incorporation.

Approximately ten-fold more nucleosomes are retained in human sperm in comparison to mouse sperm. While using the same chromatin preparation and high-throughput sequencing procedures, we observed a reduced contribution of CpG dinucleotides to nucleosome occupancy in human versus mouse spermatozoa (0.28 versus 0.71 Pearson correlation coefficient in **Supplementary Fig. 2c**; ref. 11 and **Fig. 1d**). These data may therefore suggest that the eviction of nucleosomes at CpG-poor regions in the genome is less efficient during spermiogenesis in humans than in mice.

Bisulfite sequencing analysis of genomic DNA of mouse and human sperm revealed an inverse correlation between nucleosome occupancy and DNA methylation. These data are compatible with a model in which DNA methylation established early during male germ-cell development³⁶, for example, at paternal ICRs (**Supplementary Fig. 3**), directly or indirectly prevents nucleosome retention during spermiogenesis (**Fig. 7c**). Such a mechanism would preclude transmission of chromatin states associated with methylated DNA in immature male germ cells. The differential reprogramming of DNA methylation in zygotes that were generated by microinsemination of round spermatids versus mature spermatozoa⁷ may thus indicate the presence of specific chromatin states with methylated DNA, for example, at repetitive sequences in round spermatids that are lost in sperm.

A recent study reported enrichment of H3K9 dimethylation at the ICRs of *H19* and *Rasgrf1* (ref. 5). Though we observed minor enrichment for H3K27me3 at these regions, we failed to detect any noteworthy nucleosomal occupancy (comparison of enrichments in **Supplementary Figs. 3 and 6**). In contrast, unmethylated maternal ICRs contain nucleosomes marked by H3K4me3 and/or H3K27me3 (also described in ref. 46). More generally, we observed higher enrichments for modified histones than for core histones as well as extensive enrichments for histone modifications adjacent to relatively narrow peaks of nucleosomes (**Supplementary Figs. 3 and 6**). Technically, differential enrichment is probably due to a higher sensitivity of antibodies for modified histones and an overall lower abundance of modified histones. Biologically, enrichment for histone modifications in the absence of nucleosomes may reflect retention of nucleosomes in only a low percentage of spermatozoa that could possibly lead to variegated paternal transmission. These findings suggest that caution is warranted in interpreting enrichment values for modified histones in sperm in cases where occupancy data of corresponding nucleosomes is absent.

Our study demonstrates that largely the same genes are marked by H3K27me3 in round spermatids and in sperm. In germinal-vesicle oocytes deficient for *Ring1* and *Rnf2*, two key components of the PRC1 complex, 62% of the upregulated genes are marked by H3K27me3 in mouse sperm, whereas only 35% of the unaffected genes are PRC2 targets in sperm⁴⁷. Notably, about 85% of PRC2 targets in sperm remain repressed during preimplantation development. Correspondingly, ESCs contain only a slightly reduced number of PRC2 targets (**Supplementary Fig. 5c**). Promoters of several pluripotency factors such as *Oct3/4* (official symbol *Pou5f1*), *Sox*, *Esrrb* and *Klf5* contain H3K27me3-marked nucleosomes in sperm, whereas *Nanog* is DNA methylated and essentially devoid of nucleosomes. *Klf4* is strongly labeled with H3K4me3-marked nucleosomes and weakly labeled with H3K27me3 (**Supplementary Fig. 6k–p**). Notably, repression of *Sox2* and *Klf4* in germinal-vesicle oocytes is dependent on *Ring1* and *Rnf2* function⁴⁷. Together, these data suggest that Polycomb may mediate gene repression in the male as in the female germ line. Although principally hypothesizing the paternal transmission of modified nucleosomes, epigenetic reprogramming of some H3K27me3-marked genes such as certain pluripotency factors would be required to occur in early preimplantation embryos to support their development. The majority of Polycomb targets, however, remain repressed in early embryos and would not need to be reprogrammed, and this is consistent with a model of intergenerational or possibly transgenerational inheritance of an intrinsic epigenetic memory program.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Accession codes. Data have been deposited with NCBI Gene Expression Omnibus under accession code [GSE42629](https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE42629).

Note: Supplementary information is available in the online version of the paper.

ACKNOWLEDGMENTS

We gratefully thank S. Dessus-Babus and T. Roloff (Friedrich Miescher Institute for Biomedical Research (FMI) functional genomics group), I. Nissen (Laboratory for Quantitative Genomics, D-BSSE, Basel), L. Burger (FMI bioinformatics group), H. Kohler (FMI FACS facility) and the FMI animal facility for excellent assistance. We thank members of the Peters laboratory for fruitful discussions. S.E. is supported as a recipient of a Boehringer Ingelheim Fond fellowship. M.G. and R.M. are supported by European Molecular Biology Organization (EMBO) Long Term fellowships (ALTF 253-2011 and ALTF 600-2008). Research in the Peters and Schübeler labs is supported by the Novartis Research Foundation and the Swiss Initiative in Systems Biology (Cell Plasticity, Systems Biology of Cell Differentiation). The Peters lab further acknowledges support from the Swiss National Science Foundation (31003A_125386 and National Research Programme NRP63, Stem Cells and Regenerative Medicine), the Japanese Swiss Science and Technology Cooperation Program, the FP7 Marie Curie Initial Training Network “Nucleosome4D” and the EMBO Young Investigator Program.

AUTHOR CONTRIBUTIONS

S.E., M.H. and A.H.F.M.P. conceived of and designed the experiments. S.E., M.H., C.-Y.L. and M.G. performed experiments. J.D. and J.v.d.V. provided antibodies. R.M. and D.S. performed and supervised bisulfite sequencing experiments, respectively. M.B.S. provided bioinformatics training and support. S.E., M.H., C.-Y.L., M.G., M.B.S. and A.H.F.M.P. analyzed the data. S.E. and A.H.F.M.P. prepared the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Smith, Z.D. *et al.* A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484**, 339–344 (2012).
- Gu, T.P. *et al.* The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**, 606–610 (2011).
- Aoki, F., Worrall, D.M. & Schultz, R.M. Regulation of transcriptional activity during the first and second cell cycles in the preimplantation mouse embryo. *Dev. Biol.* **181**, 296–307 (1997).
- Puschendorf, M. *et al.* PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. *Nat. Genet.* **40**, 411–420 (2008).
- Nakamura, T. *et al.* PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486**, 415–419 (2012).
- Kobayashi, H. *et al.* Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS Genet.* **8**, e1002440 (2012).
- Kishigami, S. *et al.* Epigenetic abnormalities of the mouse paternal zygotic genome associated with microinsemination of round spermatids. *Dev. Biol.* **289**, 195–205 (2006).
- Bui, H.T. *et al.* Essential role of paternal chromatin in the regulation of transcriptional activity during mouse preimplantation development. *Reproduction* **141**, 67–77 (2011).
- van der Heijden, G.W. *et al.* Sperm-derived histones contribute to zygotic chromatin in humans. *BMC Dev. Biol.* **8**, 34 (2008).
- Hammoud, S.S. *et al.* Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**, 473–478 (2009).
- Brykczynska, U. *et al.* Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nat. Struct. Mol. Biol.* **17**, 679–687 (2010).
- Mayer, W., Niveleau, A., Walter, J., Fundele, R. & Haaf, T. Demethylation of the zygotic paternal genome. *Nature* **403**, 501–502 (2000).
- Gill, M.E., Erkek, S. & Peters, A.H. Parental epigenetic control of embryogenesis: a balance between inheritance and reprogramming? *Curr. Opin. Cell Biol.* **24**, 387–396 (2012).
- Heijmans, B.T. *et al.* Persistent epigenetic differences associated with prenatal exposure to famine in humans. *Proc. Natl. Acad. Sci. USA* **105**, 17046–17049 (2008).
- Kaminen-Ahola, N. *et al.* Maternal ethanol consumption alters the epigenotype and the phenotype of offspring in a mouse model. *PLoS Genet.* **6**, e1000811 (2010).
- Anway, M.D., Memon, M.A., Uzumcu, M. & Skinner, M.K. Transgenerational effect of the endocrine disruptor vinclozolin on male spermatogenesis. *J. Androl.* **27**, 868–879 (2006).
- Carone, B.R. *et al.* Paternally induced transgenerational environmental reprogramming of metabolic gene expression in mammals. *Cell* **143**, 1084–1096 (2010).
- Zeybel, M. *et al.* Multigenerational epigenetic adaptation of the hepatic wound-healing response. *Nat. Med.* **18**, 1369–1377 (2012).
- Balhorn, R., Gledhill, B.L. & Wyrobek, A.J. Mouse sperm chromatin proteins: quantitative isolation and partial characterization. *Biochemistry* **16**, 4074–4080 (1977).
- Gatewood, J.M., Cook, G.R., Balhorn, R., Bradbury, E.M. & Schmid, C.W. Sequence-specific packaging of DNA in human sperm chromatin. *Science* **236**, 962–964 (1987).
- Gardiner-Garden, M., Ballesteros, M., Gordon, M. & Tam, P.P. Histone- and protamine-DNA association: conservation of different patterns within the β -globin domain in human sperm. *Mol. Cell Biol.* **18**, 3350–3356 (1998).
- Wykes, S.M. & Krawetz, S.A. The structural organization of sperm chromatin. *J. Biol. Chem.* **278**, 29471–29477 (2003).
- Pittoggi, C. *et al.* A fraction of mouse sperm chromatin is organized in nucleosomal hypersensitive domains enriched in retroposon DNA. *J. Cell Sci.* **112**, 3537–3548 (1999).
- Arpanahi, A. *et al.* Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. *Genome Res.* **19**, 1338–1349 (2009).
- Mohn, F. *et al.* Lineage-specific polycomb targets and *de novo* DNA methylation define restriction and potential of neuronal progenitors. *Mol. Cell Biol.* **30**, 755–766 (2008).
- Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366 (2009).
- Tillo, D. & Hughes, T.R.G. +C content dominates intrinsic nucleosome occupancy. *BMC Bioinformatics* **10**, 442 (2009).
- Deaton, A.M. & Bird, A. CpG islands and the regulation of transcription. *Genes Dev.* **25**, 1010–1022 (2011).
- Schones, D.E. *et al.* Dynamic regulation of nucleosome positioning in the human genome. *Cell* **132**, 887–898 (2008).
- Ramirez-Carrozzi, V.R. *et al.* A unifying model for the selective regulation of inducible transcription by CpG islands and nucleosome remodeling. *Cell* **138**, 114–128 (2009).
- Fenouil, R. *et al.* CpG islands and GC content dictate nucleosome depletion in a transcription-independent manner at mammalian promoters. *Genome Res.* **22**, 2399–2408 (2012).
- Li, Z., Schug, J., Tuteja, G., White, P. & Kaestner, K.H. The nucleosome map of the mammalian liver. *Nat. Struct. Mol. Biol.* **18**, 742–746 (2011).
- Kelly, T.K. *et al.* Genome-wide mapping of nucleosome positioning and DNA methylation within individual DNA molecules. *Genome Res.* **22**, 2497–2506 (2012).
- Stadler, M.B. *et al.* DNA-binding factors shape the mouse methylome at distal regulatory regions. *Nature* **480**, 490–495 (2011).
- Kacem, S. & Feil, R. Chromatin mechanisms in genomic imprinting. *Mamm. Genome* **20**, 544–556 (2009).
- Smallwood, S.A. & Kelsey, G. *De novo* DNA methylation: a germ cell perspective. *Trends Genet.* **28**, 33–42 (2012).
- Vavouri, T. & Lehner, B. Chromatin organization in sperm may be the major functional consequence of base composition variation in the human genome. *PLoS Genet.* **7**, e1002036 (2011).
- Molaro, A. *et al.* Sperm methylation profiles reveal features of epigenetic inheritance and evolution in primates. *Cell* **146**, 1029–1041 (2011).
- Elsaesser, S.J., Goldberg, A.D. & Allis, C.D. New functions for an old variant: no substitute for histone H3.3. *Curr. Opin. Genet. Dev.* **20**, 110–117 (2010).
- Orsi, G.A., Couble, P. & Loppin, B. Epigenetic and replacement roles of histone variant H3.3 in reproduction and development. *Int. J. Dev. Biol.* **53**, 231–243 (2009).
- Szenker, E., Ray-Gallet, D. & Almouzni, G. The double face of the histone variant H3.3. *Cell Res.* **21**, 421–434 (2011).
- van der Heijden, G.W. *et al.* Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. *Nat. Genet.* **39**, 251–258 (2007).
- Piña, B. & Suau, P. Changes in histones H2a and H3 variant composition in differentiating and mature rat brain cortical neurons. *Dev. Biol.* **123**, 51–58 (1987).
- Goldberg, A.D. *et al.* Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**, 678–691 (2010).
- Zeng, F. & Schultz, R.M. RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. *Dev. Biol.* **283**, 40–57 (2005).
- Delaval, K. *et al.* Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *EMBO J.* **26**, 720–729 (2007).
- Posfai, E. *et al.* Polycomb function during oogenesis is required for mouse embryonic development. *Genes Dev.* **26**, 920–932 (2012).
- Hackenberg, M. *et al.* CpGcluster: a distance-based algorithm for CpG-island detection. *BMC Bioinformatics* **7**, 446 (2006).

ONLINE METHODS

Biological sample collection. Mouse sperm were collected from C57BL/6J mice by using a swim-up procedure as described¹¹. To isolate round spermatids, testicular cells were prepared from 28-d-old C57BL/6J mice. Isolated cells were subjected to Hoechst (Invitrogen, cat. no. 33342) staining for 30 min at 37 °C, and round spermatids were collected by fluorescence-activated cell sorter (FACS) to 90% purity. All experiments were performed in accordance with the Swiss animal protection laws (license 51, Kantonales Veterinäramt, Basel, Switzerland) and institutional guidelines.

Mononucleosomal DNA preparation and native ChIP. We performed chromatin isolation from mature sperm under native conditions as described¹¹. MNase treatment for sperm was performed with 15 U (Roche Nuclease S7, cat. no. 10107921001) at 37 °C for 5 min per 2 million spermatozoa. Round-spermatid chromatin was isolated in a similar way, except for omission of DTT treatment used for sperm. MNase treatment for round spermatids was performed with 5 U at 37 °C for 30 min per 1 million cells. Chromatin immunoprecipitation (ChIP) was carried out with antibodies to H3.3 (Millipore 17-10245 ChIP grade (first replicate) and Millipore 09-838 (second replicate)), H3.1 and H3.2 (H3.1/H3.2)^{42,49}, H3K4me3 (Millipore 17-614) and H3K27me3 (Millipore 07-449) by using approximately 15 million to 20 million sperm or 5 million round spermatids and ~5 µg antibody per ChIP. Both mononucleosomal and immunoprecipitated DNA were resolved by 5% polyacrylamide electrophoresis, and 150-bp DNA was gel-purified. Input genomic DNA control was prepared by treatment of sperm with DTT and detergents as in mononucleosomal preparation, isolation of genomic DNA and subsequent sonication. The reproducibility of nucleosome isolations and ChIP experiments was demonstrated by the use of biological replicates (**Supplementary Fig. 7**).

RNA isolation. RNA from FACS-sorted round spermatids was isolated by the Qiagen RNeasy Mini kit. RNA integrity was confirmed by running RNA samples on Agilent 2100 Bioanalyzer mRNA pico arrays.

Library preparation and sequencing. Library preparation for ChIP-seq was done with the Illumina ChIP-seq DNA Sample Prep Kit (cat. no. IP-102-1001). Before preparation of RNA-seq libraries, rRNA from RNA was depleted by the Ribo-Zero rRNA removal kit (Epicentre Biotechnologies). Strand-specific RNA-seq libraries were prepared by the Illumina directional mRNA-seq library preparation prerelease protocol. The quality of libraries was assessed by Agilent 2100 Bioanalyzer. Libraries were sequenced on Illumina GA II (36-bp reads) and Illumina HiSeq 2000 (51-bp reads).

Chromatin-bound (histone) fractionation and immunoblotting. Round spermatids were isolated from C57BL/6J mouse testes by centrifugal elutriation⁵⁰, and fractionation of chromatin-bound material was performed according to a previous study⁵¹, with some modifications. Briefly, cells were resuspended in buffer A (10 mM HEPES, pH7.5, 10 mM KCl, 1.5 mM MgCl₂, 0.05% Nonidet P-40 and 0.5 mM DTT with protease inhibitors) and incubated for 10 min on ice. After centrifugation, the nuclear pellet was collected and washed twice with buffer A. Nuclei were further lysed in buffer B (3 mM EDTA, 0.2 mM EGTA, 1 mM DTT and protease inhibitors). Then insoluble chromatin was collected by centrifugation, washed twice with buffer B and resuspended in 0.2 M HCl to extract histones. Sperm samples collected by the swim-up procedure were initially treated with 50 mM DTT at room temperature for 2 h. Then the chromatin-bound fraction was isolated as described for round spermatids and was concentrated by trichloroacetic acid precipitation. Chromatin-bound extracts were analyzed by 15% SDS-PAGE gels and transferred onto PVDF membranes that were incubated with antibodies to histones H3 (Abcam ab1791, 1:6,000), H3.3 (Millipore 17-10245, 1:1,500) and H3.1/H3.2 (1:4,000)^{42,49}.

Processing and alignment of the reads. Filtering, alignment and processing of the reads for both ChIP-seq and RNA-seq were done as described³⁴. Reads from native ChIP-seq experiments were shifted by 74 nucleotides, corresponding to half the length of a nucleosome, toward their 3' end to account for the fragment length.

Genomic coordinates. All coordinate regions used in analyzing mouse ChIP-seq and RNA-seq data were based on the mouse mm9 assembly (July 2007 Build

37 assembly by NCBI and Mouse Genome Sequencing Consortium). To obtain 1-kb windows used in genome-wide analysis, the mouse genome was divided into nonoverlapping 1-kb windows. From these, the subset of mappable windows (as defined in ref. 34) was used in the subsequent analysis. RefSeq coordinates were downloaded from UCSC⁵² (<http://hgdownload.cse.ucsc.edu/goldenPath/mm9/database/refGene.txt.gz> from 16 August 2009). For each gene, coordinates corresponding to the longest known transcript were selected.

Genomic regions were classified as promoter, exon, repeat, intron or intergenic as follows. Promoter is defined as the bases covering ± 1 kb surrounding RefSeq transcripts. Exons are exonic sequences of RefSeq transcripts that do not overlap ± 1 kb from the TSS. Repeats are repeat elements of repeat masker (obtained from http://hgdownload.cse.ucsc.edu/goldenPath/mm9/database/chr*_rmsk.txt.gz (where the asterisk is substituted by chromosome number) 30 January 2009) that do not overlap promoter/exon regions. Introns are intronic sequences of RefSeq transcripts that do not overlap promoter/exon/repeats. The remaining part of the genome, which is not promoter/exon/repeat/intron, was classified as intergenic. Genomic regions used in analysis of published human ChIP-seq data were based on the human hg18 assembly (March 2006 Build 36.1 assembly by NCBI and International Human Genome Sequencing Consortium). One-kilobase windows for the human genome were generated in a similar way as for the mouse genome.

Classification of genes according to their promoter GC content. CpG classifications of the genes as high CpG (HCP), intermediate CpG (ICP) and low CpG (LCP) was performed according to criteria defined in ref. 25. For the classifications, coordinates ± 1 kb surrounding TSSs were used (**Fig. 1b**).

Calculation of observed/expected ratios for dinucleotide frequencies. Dinucleotide and single-nucleotide counts per 1-kb window were obtained by the R package Biostrings (Biostrings: string objects representing biological sequences, and matching algorithms; R package version 2.26.2). The observed/expected ratio was calculated as follows: $XY_{cnt}/(X_{cnt} \cdot Y_{cnt}) \cdot (W_{size} - 1)$, where XY_{cnt} is the dinucleotide count of XY in one 1-kb window, X_{cnt} and Y_{cnt} are single nucleotide counts, and W_{size} is the window size (1 kb).

CGI definition and usage. CGI definitions are based on a CpG cluster algorithm⁴⁸. The algorithm was run with default parameters on mm9 to obtain genomic coordinates of CGI.

UCSC tracks. Wiggle files were generated for 100-bp windows and uploaded to the UCSC genome browser⁵². Data were visualized by smoothing over 3 pixels (**Figs. 1a,c and 3b**).

Quantification of enrichment levels genome wide, at promoter regions and nucleosome peaks. Enrichment levels for ChIP-seq experiments were calculated for 1-kb windows, promoter regions of the genes (±1 kb surrounding TSSs) and nucleosome peaks identified by hidden semi-Markov model (**Supplementary Note** describes identification of nucleosome peaks). To calculate enrichment, total read counts mapping to a coordinate region were calculated for ChIP and control (input genomic DNA) samples. Then these counts were normalized to account for different library sizes between ChIP and control samples. Enrichment for each region was calculated as the ratio between library size-normalized read counts for ChIP and control samples according to the following formula: $\log_2(((Cnt_{smp}/LSize_{smp} - \min(LSize_{smp}, LSize_{cnt})) + psnt)/((Cnt_{cnt}/LSize_{cnt} - \min(LSize_{smp}, LSize_{cnt})) + psnt))$, where Cnt_{smp} is the total number of reads mapping to the coordinate in ChIP sample, $LSize_{smp}$ is the total library size for the ChIP sample, Cnt_{cnt} is the total number of reads mapping to the coordinate in the control sample, $LSize_{cnt}$ is the total library size for the control sample and $psnt$ is a constant number ($psnt = 8$) that was used to stabilize enrichments on the basis of low read counts.

Plotting profiles around genomic regions. For each sample, reads mapping to the genomic regions of interest (**Fig. 4** and **Supplementary Fig. 4b**) were summed for every base pair within the genomic region analyzed. Average read counts per bp were calculated by dividing the total number of reads per bp to total number of genomic regions analyzed. To plot average enrichment values for multiple ChIP-seq samples on the same plot, counts were scaled by the library size, and enrichment values were calculated as the ratio between scaled read counts of ChIP

and control samples (sonicated sperm genomic DNA). Profiles were smoothed for plotting by taking the rolling mean over 40 bp.

Heat-map plots. For ChIP-seq experiments, the number of reads covering each base pair in the region ± 3 kb around TSSs of genes was quantified. Read coverage was averaged in 50-bp windows along ± 3 kb TSSs. Within each data set, values were scaled to arrange between 0–1. CpG coverage around ± 3 kb was obtained by Bioconductor-package Biostrings, and coverage intensities were scaled in a similar way as for ChIP-seq features. Expression data for RS was calculated as $\log_2(\text{read count per transcript})$. Clustering was performed by using k means with $k = 5$, empirically selected as the minimal value of k that resulted in distinct clusters consisting of homogenous members.

Variance partitioning analysis. Variance partitioning analysis was performed by R-package yhat (yhat: interpreting regression effects; R package version 1.0-5). Unique and combinatorial effects for each variable were obtained by the function commonalityCoefficients().

GO-term analysis. GO-term analysis was performed by using Bioconductor-package topGO⁵³. Enrichment tests were done by Fisher's exact test (Supplementary Table 1).

49. van der Heijden, G.W. *et al.* Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. *Mech. Dev.* **122**, 1008–1022 (2005).
50. Barchi, M., Geremia, R., Magliozzi, R. & Bianchi, E. Isolation and analyses of enriched populations of male mouse germ cells by sedimentation velocity: the centrifugal elutriation. *Methods Mol. Biol.* **558**, 299–321 (2009).
51. Méndez, J. & Stillman, B. Chromatin association of human origin recognition complex, cdc6, and minichromosome maintenance proteins during the cell cycle: assembly of prereplication complexes in late mitosis. *Mol. Cell Biol.* **20**, 8602–8612 (2000).
52. Kent, W.J. *et al.* The human genome browser at UCSC. *Genome Res.* **12**, 996–1006 (2002).
53. Alexa, A., Rahnenfuhrer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600–1607 (2006).

Erratum: Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa

Serap Erkek, Mizue Hisano, Ching-Yeu Liang, Mark Gill, Rabih Murr, Jürgen Dieker, Dirk Schübeler, Johan van der Vlag,

Michael B Stadler & Antoine H F M Peters

Nat. Struct. Mol. Biol. 20, 868–875 (2013); published online 16 June 2013; corrected after print 12 July 2013

In the version of this article initially published, the parentheses in Figure 2a denoting noninclusive endpoints in ranges had not been indicated. The error has been corrected in the HTML and PDF versions of the article.

Chapter 4

Characterization of expression of histone H3 variant proteins during spermatogenesis suggests variant specific roles in chromatin reprogramming

Ching-Yeu Liang^{1, 2}, Ragna Sack¹, Hubertus Kohler¹ and Antoine H. F. M Peters^{1, 2}

1. Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, 4058 Basel, Switzerland

2. Faculty of Sciences, University of Basel, 4056 Basel, Switzerland

Keywords:

Histone variant, testis-specific histone H3, histone H3.3, mouse spermatogenesis

Abstract

Specialized histone variants incorporated into nucleosomes play important roles in gene regulation and modulation of chromatin structure. Although many testis-specific histones have been reported before, the dynamics of testis-specific and canonical histone H3 expression during male germ cell development in mouse is still unclear. Here, using triton-acetic acid-urea (TAU) gel electrophoresis, we identify that mouse testis-specific histone H3 (H3t) is the most predominant H3 protein from spermatogonia to spermatids. The mouse H3t gene is highly transcribed in mitotic spermatogonia and is downregulated in subsequent germ cell stage. These data suggest that incorporation of H3t into chromatin in spermatogonia is regulated in a DNA replication-coupled manner as described for canonical H3. Interestingly, we find that most canonical H3 is removed and replaced by H3.3 variant during meiosis. Next, during the process of histone-to-protamine exchange, most of H3t is evicted from chromatin in comparison to H3.3, which leads to the predominant presence of H3.3 in the residual nucleosomes of sperm. Moreover, we observe that the transcriptional activity mark, trimethylation on H3 lysine 4 (H3K4me3), is globally erased in pachytene spermatocytes but is restored on H3.3 rather than on H3t in round spermatids, suggesting that H3.3 is involved in the activation of postmeiotic transcription. Collectively, our study demonstrates the characterization of mouse H3t and an extensive dynamics of histone H3 variants in male germ cell development which may prime for histone eviction during spermiogenesis.

Introduction

The fundamental unit of chromatin is the nucleosome that is composed of DNA wrapped around a histone octamer assembled by two each of core histone H3, H4, H2A, and H2B (Luger et al. 1997). This nucleosome structure provides an efficient organization to compact genomic DNA into the nucleus. Histones are subjected to many post-transcriptional modifications, such as methylation, acetylation and phosphorylation, which directly alter the degree of chromatin compaction or recruit different protein complexes to affect chromatin structure and regulate gene expression (Zentner and Henikoff 2013). Based on expression and mode of incorporation into nucleosomes, histone proteins are classified into

canonical histones and replacement variant histones (Skene and Henikoff 2013; Maze et al. 2014). Canonical histones are encoded by multicopy intronless genes that are highly expressed during S-phase of cell cycle, ensuring a major supply for nucleosome incorporation during replication (Marzluff et al. 2008). By contrast, the expression and incorporation of replacement variant histones are generally replication-independent. Except for H4, each of the core histone classes has histone variants. Compared with canonical histones, variant histones have slightly different protein sequences that direct histone incorporation into chromatin outside S-phase and influence the physical properties of nucleosomes. The exchange of histone variants contributes to gene regulation, DNA repair, chromosome segregation and other processes through altering nucleosome stability and chromatin structure (Maze et al. 2014).

In histone H3 family in mammals, H3.1 and H3.2 are defined as canonical H3, which are highly expressed during S phase and are deposited into replication fork by CAF-1 chaperon complex during DNA replication (Tagami et al. 2004). By contrast, the replacement variants of H3 best characterized in mammals are H3.3 and centromeric H3 (CenH3). Two distinct intron-contained genes, *H3f3a* and *H3f3b*, encode H3.3 that are synthesized and deposited throughout the cell cycle. It is well-known that H3.3 deposition at active promoters, gene bodies and transcription factor-binding sites is mediated by HIRA chaperone (Tagami et al. 2004; Goldberg et al. 2010). But recent studies showed that H3.3 also can be incorporated into heterochromatin regions, such as pericentric repeats and telomeres, by Daxx/ATRX complex to maintain genome stabilization (Lewis et al. 2010). CenH3 is a centromere-specific H3 variant that is required for kinetochore assembly and chromosome segregation (Dunleavy et al. 2009; Foltz et al. 2009). CenH3 protein is synthesized in G2 phase and deposited into DNA outside of replication in late mitosis or G1 phase by HJURP chaperon (Boyarchuk et al. 2011).

Moreover, four H3 variants are identified in human tissues. The expression of testis-specific H3 (H3t, also known as H3.1t and H3.4) and H3.5 (also known as H3.3C) are specifically detected in human testes (Witt et al. 1996; Schenk et al. 2011). H3.X (also known as H3.Y.2) and H3.Y (also known as H3.Y.1) are primate-specific histone H3 identified and expressed in human brain, testis and certain tumor tissues (Wiedemann et al. 2010). Previous studies have shown that nucleosomes containing human H3t are more unstable than conventional

nucleosomes, suggesting that H3t containing nucleosomes may contribute to global chromatin reorganization during meiosis and post-meiotic event (Tachiwana et al. 2010). In mouse, testicular expression of H3t transcript and protein has been described before (Govin et al. 2007; Montellier et al. 2013), and the *GM12260* pseudogene in mouse genome was predicted to encode for H3t. However, the characterization of mouse H3t function and the dynamics of H3 variants during male germ cell development in mouse testis are still unclear.

Here, by using triton-acetic acid-urea (TAU) gel electrophoresis, we observe a novel H3 protein band specifically expressed in mouse testes compared to somatic adult tissues. Mass spectrometry analysis confirmed its identity as mouse H3t protein. We find that mouse *H3t* transcripts are particularly highly abundant in spermatogonia while being reduced at later stages of spermatogenesis. In agreement, H3t is the predominant H3 protein in spermatogonia over canonical H3.1 and H3.2. During subsequent meiotic and postmeiotic stages, the relative amount of H3.3 increase, and most of H3.1 and H3.2 are replaced by H3t and H3.3. Surprisingly, H3t levels are greatly reduced in mature sperm having undergone the histone-to-protamine replacement. Instead, we measure considerable levels of H3.3 and canonical H3 in the residual nucleosomes, arguing that H3t may function as a replacement variant involved in histone eviction during spermiogenesis while the canonical histones and H3.3 are resistant to eviction and are selectively retained at specific sites in mature sperm, e.g. at CpG rich sequences as reported previously (Erkek et al. 2013).

Results

Identification of mouse testis-specific H3 (H3t) on TAU gel

Our previous study has revealed that H3.3 is more abundant than canonical H3 variants, H3.1 and H3.2, in mouse epididymal sperm using H3.3 and H3.1/3.2-specific antibodies for SDS-PAGE-Western blotting (Erkek et al. 2013). To address whether H3.3 is also enriched in mouse adult testis compared to other mouse tissues, we perform triton-acetic acid-urea (TAU) gel-electrophoresis followed by Western blotting to separate H3 variants by hydrophobicity and charge (Shechter et al. 2007). We subsequently probe it for H3 with an

antibody recognizing the C-terminus of H3 that is highly conserved between all known canonical and variant H3 proteins. As shown in Figure 1A, we observe that histone acid extracts from total testis contained more H3.3 (blue arrowhead) and less H3.1/H3.2 (red/orange arrowheads) as compared to the nuclear extract from mouse embryonic stem (ES) cells and any other mouse tissue tested, including post-replicative brain tissues. Interestingly, we reproducibly note a H3 protein migrating faster than H3.1 (labeled by green arrowhead) present in testis extracts, but not in other tissue extracts. Its expression level is notably higher than that of H3.1, H3.2 and H3.3. Moreover, this faster migrating H3 species is detected by H3.1/H3.2-specific antibodies, and not by a H3.3 antibody, distinguishing between H3 variant specific amino acids at residue 31 of the N-terminal tail of the H3 proteins (Supplementary Figure 1A).

Previously, Montellier and colleagues have identified the presence of H3t peptides corresponding to the *GM12260* pseudogene in mouse spermatocytes and round spermatids by mass spectrometry (Montellier et al. 2013). *GM12260* encodes a predicted protein that is 98% identical similar to human H3t and close to canonical H3.1/3.2, but not H3.3 (Figure 1B). We confirm by RNA-sequencing analysis that *GM12260* is highly transcribed in mouse round spermatids, but not in mouse ES cells (Supplementary Figure 1B) (Erkek et al. 2013). Therefore, we speculate that the novel H3 species present in mouse testis extract and migrating faster than H3.1 in the TAU gel corresponds to mouse H3t.

To confirm this hypothesis, we perform mass spectrometry analyses on the putative H3t (band 1) and control bands including H3.1 (band 3) and H3.3 (band 2 and 4) (Figure 1C). After normalizing for peptides that are identical in all H3 proteins, we measure levels of the H3t specific-peptide (residues 18-26, with Val24) and canonical H3 peptide (residues 18-26, with Ala24) in these four H3 bands. Consistent with our expectation, band 1 isolated from testis extract contains 8 fold more of the H3t peptide than of the canonical H3 peptide. Bands 2-4 barely contain any H3t specific peptide (Figure 1C, 1D). Furthermore, H3t overexpressed in HEK293 cells migrates at a comparable position in the TAU gel as endogenous H3t from mouse testis, just below the H3.1 band (Supplementary Figure 1C). Taken together, these results demonstrate the migration patterns of the various mouse H3 histones in TAU gels and showed that in contrast to embryonic stem cells (Loyola et al.

2006), H3t and H3.3 are the two H3 proteins predominantly present in chromatin of adult mouse testis.

Mouse H3t is initially expressed in mouse testes on postnatal days 8 to 10

We next ask when H3t and H3.3 replace canonical H3 during spermatogenesis. In order to investigate the dynamics in expression of the H3 variants, we first examine *H3t*, *H3f3a* and *H3f3b* transcript levels in testes of juvenile mice undergoing the first wave of spermatogenesis. We use *Stra8*, *Sycp3* and *Tnp2* as pre-meiosis, meiosis and post-meiosis control markers (Figure 2A, bottom panels). After normalization to ribosome S17, we find that *H3t* transcript is expressed in testes at 7 days postpartum (dpp), not in ES cells, and dramatically increases when the first wave of spermatogonia differentiation into meiotic spermatocyte occurs at 8-10 dpp (Figure 2A, upper right panel). Strikingly, two H3.3-encoded genes, *H3f3a* and *H3f3b*, have entirely different transcriptional profiles in the first wave of spermatogenesis. The *H3f3a* transcript is expressed at a low level until the time of formation of haploid spermatids (Figure 2A, upper middle panel). In contrast, *H3f3b* transcript is stably expressed over subsequent days of testis development. (Figure 2A, upper right panel). These results are in line with the observation that mutation of *H3f3a* leads to sperm defects and subfertility, while spermatogenesis is arrested at the round spermatid stage already in *H3f3b* heterozygous mutant animals (Couldrey et al. 1999; Tang et al. 2015).

Next, we evaluate protein levels of each H3 variant in testes of juvenile mice by TAU gelelectrophoresis and Western blotting. Consistent with above transcription results, we observe low H3t protein levels in 8-10 dpp testes. Its expression increases dramatically from 12 dpp until adulthood (Figure 2B and the quantification in Figure 2C). Likewise, H3.3 levels increase upon advancing testicular development. In contrast, H3.1 and H3.2 protein levels get progressively reduced during the first wave of mouse spermatogenesis.

Dynamics of H3 variants during mouse spermatogenesis

Besides male germ cells, the testis contains different somatic cells, such as Sertoli and Leydig cells. To exclude H3 proteins extracted from somatic cells in testis, we purify germ cells from different development stages of adult mouse testes using two different approaches. First, we perform fluorescence-activated cell sorting (FACS) using Hoechst

33342 and c-Kit staining to isolate diploid type-B spermatogonia from adult mouse testes. The fraction contained 75% spermatogonia, mixed mainly with pre-leptotene spermatocytes (Supplementary Figure 2A and B) (Barroca et al. 2009; Hammoud et al. 2014). At the same time, we purify pachytene spermatocytes on the basis of their DNA content using Hoechst staining with 95% cell purity (Lassalle et al. 2004) (Supplementary Figure 2A and B). Moreover, we use the centrifugal elutriation to isolate haploid round and condensing spermatids individually based on cell size (Barchi et al. 2009), and the percent cell purity in both spermatid fractions is over 90% (Supplementary Figure 2B).

Next, we extract mRNA from spermatogonia, spermatocytes and round spermatids to analyze expression of each H3 variant as well as stage-specific germ-cell markers described above. Notably, we find that *H3t* transcript is high in spermatogonia but dramatically decreases in spermatocytes and spermatids (Figure 3A, upper left panel). Since spermatogonia undergo DNA replication, *H3t* expression and incorporation into chromatin may be DNA replication-dependent as canonical H3. Consistent with results in Figure 2A, *H3f3a* transcript level progressively increase from the spermatogonia to round spermatid stage, while *H3f3b* transcript level is generally high and more stably between cell types (Figure 3A, upper middle and left panels).

To relate RNA to protein expression, we examine the relative levels of different H3 proteins from spermatogonia to mature sperm by TAU-Western blotting. We find that all four H3 isoforms H3t, H3.1, H3.2 and H3.3, are expressed together in spermatogonia, but H3t protein is most abundant (Figure 3B and the quantification in Figure 3C). Intriguingly, H3.1 and H3.2 proteins are not detectable in pachytene spermatocytes and spermatids, while H3.3 protein expression is increased in spermatocytes. Thus, H3t and H3.3 likely function as the two major H3 replacement proteins in meiotic and postmeiotic chromatin. Surprisingly, we reproducibly observe increased levels of H3.3 and particularly of H3.1 and H3.2 in mature spermatozoa, while H3t protein levels is drastically reduced (Figure 3B and C). This data argue that H3t-containing nucleosomes are more extensively removed from chromatin than H3.3-containing nucleosomes during spermiogenesis. Truly remarkably, the substantial amount of H3.1 and H3.2 present in mature sperm suggest that these canonical histones, incorporated into chromatin in replicating spermatogonia, resist the histone-to-H3t turnover and histone-to-protamine turnover programs.

H3t has less H3K4me3 modification than H3.3 in spermatogenic cells

It is well-known that trimethylation on H3 lysine 4 (H3K4me3) is a histone mark for transcription activation and is more abundant on histone H3.3 variant than on H3.1/H3.2 in somatic cells (McKittrick et al. 2004; Loyola et al. 2006). Given the extensive replacement of H3.1 and H3.2 nucleosomes by H3t nucleosomes, we wonder about H3K4me3 levels on H3t and H3.3 in male germ cells. By using TAU-Western blotting with anti-H3K4me3 antibody, we find that H3K4me3 is strongly enriched on H3.3 over H3t in spermatids, even more than in ESCs (Figure 4A and the quantification in Figure 4B). Further, we use the reversed phase HPLC to isolate H3.3 and H3t from round spermatid acid extracts and then perform SDS-PAGE-Western blotting with equal loading amounts of H3.3 and H3t. This result is consistent with above finding that H3t contains less H3K4me3 than H3.3 (Figure 4C). In addition, interestingly, TAU gel results in Figure 4A shows that H3K4me3 is dramatically reduced to an undetectable level either on H3.3 or on H3t in pachytene spermatocytes. Although previous studies have reported that pachytene spermatocytes have transcription activity (Turner et al. 2005), the number of transcribed genes may be lower in pachytene spermatocytes than in spermatids because of synapsis formation. Taken together, these results demonstrate that less H3K4me3 is deposited in H3t in male germ cells similar to canonical H3.1/H3.2 in somatic cells, and that global H3K4me3 level changes dramatically from meiotic spermatocyte to post-meiotic spermatids.

Discussion

In the present study, we describe the characterization of mouse testis-specific histone H3 (H3t) and found that H3t and H3.3 are two major H3 proteins during mouse spermatogenesis, not canonical H3.1 and H3.2. In spermatogonia stage, all of four H3 proteins, H3.1, H3.2, H3.3 and H3t, are present, but H3t is the most predominant one. When spermatogonia differentiate into primary spermatocyte, most canonical H3.1 and H3.2 are removed, and H3.3 amount increases at pachytene spermatocyte stage. Interestingly, previous studies showed that testis-specific H2B and H2A variants, TH2B and TH2A, are also highly synthesized and then replace canonical H2B and H2A at meiotic stage in mouse testes (Montellier et al. 2013; Shinagawa et al. 2015). Therefore, these data suggest that meiotic

spermatocytes have remarkably histone variant exchange event, in which canonical histones are replaced by H3.3 and testis-specific histones variants that may destabilize nucleosome structures (Li et al. 2005; Tachiwana et al. 2010) and contribute to global chromatin reprogramming during meiosis.

Based on previous studies, it has been indicated that nucleosomes containing human H3t are less stable than conventional nucleosome *in vitro* and have rapid turnover *in vivo* because of two human H3t-specific residues, Met71 and Val111 (Tachiwana et al. 2010). Furthermore, Val111 on human H3t reduces the efficiency of human Nap1-mediated nucleosome formation *in vitro* (Tachiwana et al. 2008). Mouse H3t encoded by *GM12260* pseudogene has 98 % identify to human H3t. Human and mouse H3t protein share specific residues, Val24, His42, and Ser98, but mouse H3t has not Met71 and Val111 residues as human H3t. Therefore, mouse H3t nucleosome may have different physical properties comparing to human H3t nucleosome. Nonetheless, we observe that most of mouse H3t is removed during histone-to-protamine replacement from condensing spermatid to mature sperm stage, as compared to H3.3. It is possible that mouse H3t may destabilize nucleosome structure as human H3t to promote nucleosome eviction during sperm maturation. Previous studies have showed that meiotic cells have the high level of phosphorylated histones, such as H2A.X phospho-Ser139, H3 phospho-Ser10, H3 phospho-Thr3 and H4 phospho-Ser1 (Hamer et al. 2003; Krishnamoorthy et al. 2006; Nickerson et al. 2007; Nguyen et al. 2014), for DNA repair and chromosome segregation in yeast and mammals. Possibly, H3t Ser98 could be phosphorylated and then involved in chromatin remodeling during meiosis stage. It will be crucial to decipher the importance of two H3t-specific residues during spermatogenesis in future *in vitro* and *in vivo* studies.

Because of the different genomic occupancy of H3 variants, H3 variants are decorated by diverse post-translational modifications that are relative to distinct biological functions. For example, in somatic and ES cells, H3.3 is generally localized at the active promoters, gene-regulation elements and gene bodies of active genes and is enriched in active markers, such as H3K4me3, H3K9Ac and H3K79me3, considered as representative of transcriptionally active state. By contrast, canonical H3.1 and H3.2 are generally distributed in heterochromatin and have more H3K9me3, H3K27me2 and H3K27me3 associated with gene silence and the formation of facultative heterochromatin. And our results show that mouse

H3t considered as H3.1/H3.2 substitute has less H3K4me3 enrichment than H3.3 in mouse spermatids. Therefore, these suggest that mouse H3t in mouse spermatid could play a similar role as H3.1/H3.2 in somatic cells to mediate gene silence and maintain heterochromatin. In addition, we find that total H3K4me3 level is dramatically reduced either on H3t or on H3.3 in pachytene spermatocytes and is restored after meiosis consistent with previous immunostaining findings (Song et al. 2011). Because of DNA repair coupled with chromosomal recombination occurs in early meiosis, the number of transcriptionally active genes may decrease that leads to global H3K4me3 reduction in pachytene spermatocytes. Likewise, erasing most of H3K4me3 in spermatocytes may promote the re-organization of H3K4me3 to quickly switch expression profiles from meiotic genes to post-meiotic genes after meiosis. Together, these results indicate that the transcriptionally active modification on histone H3 is subjected to dynamic change during and after meiosis.

It is well-documented that the transcription and deposition of canonical histones are tightly coincident with DNA replication. Multiple copies of canonical histone genes are rapidly transcribed at the beginning of S-phase to produce a large amount of histones required for nucleosome assembly at replication fork, and canonical histone mRNAs are quickly degraded at the end of S-phase. In the 3' untranslated region (UTR) of canonical histone mRNA, there is a stem loop structure, which can recruit stem-loop binding protein (SLBP) to stabilize and translocate histone mRNA to polyribosomes, instead of a poly(A)-tail (Marzluff et al. 2008). Unlike replication-dependent canonical histones, the expression and deposition of replacement histone variants (non-canonical histones) are not limited at S-phase. Their mRNAs are transcribed independent on replication and polyadenylated at 3'UTR. In this present study, we observe that the transcript level of *H3t* gene encoded by *GM12260* pseudogene is remarkably higher in spermatogonia, in which mitotic division occurs in order to provide a continuous supply of male stem cells and produce primary spermatocytes. Likewise, there is a stem loop-like sequence close to the 3' end of *GM12260* pseudogene, suggesting that H3t expression may be regulated by DNA replication-dependent manner as described for canonical H3. In addition, previous studies used *in vivo* radioactive protein labeling to show that H3t protein in rat is expressed and incorporated into chromatin during mitotic phase of spermatogonia (Trostle-Weige et al. 1984; Meistrich et al. 1985). Taken

together, these findings demonstrate that like canonical H3, H3t expression and incorporation to chromatin could be DNA replication-dependent and occurs at mitotic spermatogonia.

In addition to the dynamics of H3t expression during spermatogenesis in this study, we also find two genes encoding H3.3 are expressed in different pattern during spermatogenesis. *H3f3b* is transcribed stably at high level during the first wave of spermatogenesis and also in different stages of spermatogenic cells, while the level *H3f3a* transcription is relatively low before meiosis and is upregulated in post-meiotic spermatids. In addition, our previous genome-wide study has showed that H3.3 turnover occurs at promoter regions during mouse spermiogenesis. Together, these data suggest that the expressing of *H3f3a* after meiosis may contribute to H3.3 turnover in postmeiotic spermatids, and that *H3f3b* expression may be essential for whole developmental process of male germ cells. Interestingly, a previous study indicated that *H3f3b* mutant mice died at birth, and *H3f3b* heterozygous male mice are sterile with abnormal spermiogenesis at round spermatid stage (Tang et al. 2015). And another study using different knockout strategy indicated that the rate of apoptosis increases at spermatogonia and spermatocyte stage in *H3f3b* mutant mouse testes (Yuen et al. 2014). But *H3f3a* mutant male mice are viable and sub-fertile with sperm deficiency (Tang et al. 2015). Thus, *H3f3b* plays a more important role throughout mouse spermatogenesis than *H3f3a*, which is in line with our *H3f3a* and *H3f3b* transcript profiling results.

Materials and methods

Isolation of mouse spermatogonia, spermatocytes, spermatids and sperm

Male C57BL/6J mice were used in this study. For isolating spermatogonia and pachytene spermatocytes, testicular cells were trypsinized from adult male testes and were stained by Hoechst 33342 (Invitrogen, cat. no. 33342) first and subsequently by c-Kit antibody conjugated with Phycoerythrin (Affymetrix, 12-1171). Pachytene spermatocytes were collected based on DNA content by fluorescence-activated cell sorter (Getun et al. 2010), and spermatogonia were isolated based on c-Kit signal and DNA content according to previous studies (Hammoud et al. 2014). Round, elongating and condensing spermatids

were separated by centrifugal elutriation based on cell size as described (Barchi et al. 2009). Mature sperm was obtained from adult epididymides by performing a swim-up procedure as described (Hisano et al. 2013). The purity of each cell fractions was confirmed by immunostaining with DAPI and specific markers, such as gamma-H2AX, Sycp3, for each stage of spermatogenic cells. All animal experiments were performed in accordance with the Swiss animal protection laws and institutional guidelines.

Histone acid extraction, TAU-gel Western blotting and RP-HPLC

For acid extraction of histones from mouse tissue, tissues from C57BL/6J mice were cut into small pieces, resuspended in tissue lysis buffer (15 mM Tris-HCl pH 7.5, 60 mM KCl, 11 mM CaCl₂, 5 mM NaCl, 5 mM MgCl₂, 250 mM sucrose, 0.5 mM DTT, 5 mM sodium butyrate, 0.3 % NP-40 and protease inhibitor cocktail) and homogenized by dounce homogenizer. After centrifugation, nuclear pellet was resuspended in 0.2 N HCl and incubated at 4°C for at least 2 hours to solubilize histones. For acid extract of histones from purified spermatogenic cells, we lysed cells by Buffer A (10 mM HEPES pH 7.5, 1.5 mM MgCl₂, 10 mM KCl, 0.05 % NP-40, 0.5mM DTT, 5 mM Sodium butyrate and protease inhibitor cocktail) to separate nuclear and cytoplasmic fractions. After centrifugation, histones were extracted from nuclear pellet by 0.2N HCl as above steps. Histones in acid solution were precipitated with trichloroacetic acid (TCA) and finally dissolved into TAU sample buffer or Milli-Q water. Using TAU gel and RP-HPLC for separating H3 variants were done as described (Shechter et al. 2007). Histone samples in TAU sample buffer were loaded into 15 % TAU gel to separate histone proteins and then transferred to PVDF membrane for Western-blotting. Histone samples in water were injected into reverse-phase HPLC with C8 column and fractionated by using an acetonitrile gradient. Fractions containing H3 were verified by SDS-PAGE Western blotting. Primary antibodies used for Western blotting were H3 (Abcam, ab1791), H3K4me3 (Millipore, 07-473), H3.3 (Cosmo, CAC-CE-040B) and H3.1/H3.2 (Cosmo, CAC-CE-039B).

RNA extraction, reverse transcription and Quantitative real-time PCR

Total RNA was extracted from testis tissue and spermatogenic cells by using Qiagen RNeasy mini kit and RNeasy micro kit. RNA was treated with DNase (TURBO DNA-free™ Kit). cDNA was synthesized by using SuperScript® III First-Strand Synthesis System with random hexamers, and then amplification of cDNA was analyzed with Fast SYBR® Green Master Mix

on the 7500 Fast Real-Time PCR System. All transcript levels were normalized to housekeeping gene, ribosomal protein S17. Sequence of primers for real-time PCR were 5'-GGCACGGAAGTCGACGGGA-3' and 5'-CCACCGTGCCAGGGTG-3' for mouse *H3t*, 5'-CGGCGTGTGTAGGGGAA-3' and 5'-CGAAGGCTGCGAACACAA-3' for *H3f3a*, 5'-CGGGGTGAAGAAGCCTCA-3' and 5'-GGTAACGACGGATCTCTCTCA-3' for *H3f3b*, 5'-CAAAGCCTTGGCTGTGTTA-3' and 5'-AAAGGTCTCCAGGCACTTCA-3' for *Stra8*, 5'-GTGTTGCAGCAGTGGGAAC-3' and 5'-GCTTTCATTCTCTGGCTCTGA-3' for *SYCP3* and 5'-AAGTGAGCAAGAGAAAGGCCGTCA-3' and 5'-ACATCCTGGAGTGCCTCACTTGTA-3' for *Tnp2*.

Mass spectrometry analysis

H3 bands were excised from Coomassie-stained TAU gel or SDS-PAGE, processed twice by propionylation of free and monomethylated amino groups prior tryptic digestion and finally treated with propionic anhydride to modify the amino groups of the peptide N-terminus. Peptides were separated and analyzed by nano-HPLC (Easy-nLC 1000) coupled to an LTQ Orbitrap Velos hybrid mass spectrometer (all Thermo Scientific). Mascot 2.3 (Matrix Science) was used for database search. Label-free quantification was performed using Progenesis-LC (nonlinear Dynamics). For normalization, eight H3 peptides (residue 54-63, 64-69, 73-83, 117-128 and 130-134) were selected which are conserved in the H3 family.

References

- Barchi M, Geremia R, Magliozzi R, Bianchi E. 2009. Isolation and analyses of enriched populations of male mouse germ cells by sedimentation velocity: the centrifugal elutriation. *Methods in molecular biology* **558**: 299-321.
- Barroca V, Lassalle B, Coureuil M, Louis JP, Le Page F, Testart J, Allemand I, Riou L, Fouchet P. 2009. Mouse differentiating spermatogonia can generate germinal stem cells in vivo. *Nature cell biology* **11**: 190-196.
- Boyarchuk E, Montes de Oca R, Almouzni G. 2011. Cell cycle dynamics of histone variants at the centromere, a model for chromosomal landmarks. *Current opinion in cell biology* **23**: 266-276.
- Couldrey C, Carlton MB, Nolan PM, Colledge WH, Evans MJ. 1999. A retroviral gene trap insertion into the histone 3.3A gene causes partial neonatal lethality, stunted growth, neuromuscular deficits and male sub-fertility in transgenic mice. *Human molecular genetics* **8**: 2489-2495.
- Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G. 2009. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137**: 485-497.

- Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, Schubeler D, van der Vlag J, Stadler MB, Peters AH. 2013. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nature structural & molecular biology* **20**: 868-875.
- Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, Black BE, Cleveland DW. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137**: 472-484.
- Getun IV, Wu ZK, Khalil AM, Bois PR. 2010. Nucleosome occupancy landscape and dynamics at mouse recombination hotspots. *EMBO reports* **11**: 555-560.
- Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X et al. 2010. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**: 678-691.
- Govin J, Escoffier E, Rousseaux S, Kuhn L, Ferro M, Thevenon J, Catena R, Davidson I, Garin J, Khochbin S et al. 2007. Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *The Journal of cell biology* **176**: 283-294.
- Hamer G, Roepers-Gajadien HL, van Duyn-Goedhart A, Gademan IS, Kal HB, van Buul PP, de Rooij DG. 2003. DNA double-strand breaks and gamma-H2AX signaling in the testis. *Biology of reproduction* **68**: 628-634.
- Hammoud SS, Low DH, Yi C, Carrell DT, Guccione E, Cairns BR. 2014. Chromatin and transcription transitions of mammalian adult germline stem cells and spermatogenesis. *Cell stem cell* **15**: 239-253.
- Hisano M, Erkek S, Dessus-Babus S, Ramos L, Stadler MB, Peters AH. 2013. Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nature protocols* **8**: 2449-2470.
- Krishnamoorthy T, Chen X, Govin J, Cheung WL, Dorsey J, Schindler K, Winter E, Allis CD, Guacci V, Khochbin S et al. 2006. Phosphorylation of histone H4 Ser1 regulates sporulation in yeast and is conserved in fly and mouse spermatogenesis. *Genes & development* **20**: 2580-2592.
- Lassalle B, Bastos H, Louis JP, Riou L, Testart J, Dutrillaux B, Fouchet P, Allemand I. 2004. 'Side Population' cells in adult mouse testis express Bcrp1 gene and are enriched in spermatogonia and germinal stem cells. *Development* **131**: 479-487.
- Lewis PW, Elsaesser SJ, Noh KM, Stadler SC, Allis CD. 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at telomeres. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 14075-14080.
- Li A, Maffey AH, Abbott WD, Conde e Silva N, Prunell A, Siino J, Churikov D, Zalensky AO, Ausio J. 2005. Characterization of nucleosomes consisting of the human testis/sperm-specific histone H2B variant (hTSH2B). *Biochemistry* **44**: 2529-2535.
- Loyola A, Bonaldi T, Roche D, Imhof A, Almouzni G. 2006. PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. *Molecular cell* **24**: 309-316.
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251-260.
- Marzluff WF, Wagner EJ, Duronio RJ. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature reviews Genetics* **9**: 843-854.
- Maze I, Noh KM, Soshnev AA, Allis CD. 2014. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nature reviews Genetics* **15**: 259-271.
- McKittrick E, Gafken PR, Ahmad K, Henikoff S. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 1525-1530.
- Meistrich ML, Bucci LR, Trostle-Weige PK, Brock WA. 1985. Histone variants in rat spermatogonia and primary spermatocytes. *Developmental biology* **112**: 230-240.

- Montellier E, Boussouar F, Rousseaux S, Zhang K, Buchou T, Fenaille F, Shiota H, Debernardi A, Hery P, Curtet S et al. 2013. Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes & development* **27**: 1680-1692.
- Nguyen AL, Gentilello AS, Balboula AZ, Shrivastava V, Ohring J, Schindler K. 2014. Phosphorylation of threonine 3 on histone H3 by haspin kinase is required for meiosis I in mouse oocytes. *Journal of cell science* **127**: 5066-5078.
- Nickerson HD, Joshi A, Wolgemuth DJ. 2007. Cyclin A1-deficient mice lack histone H3 serine 10 phosphorylation and exhibit altered aurora B dynamics in late prophase of male meiosis. *Developmental biology* **306**: 725-735.
- Schenk R, Jenke A, Zilbauer M, Wirth S, Postberg J. 2011. H3.5 is a novel hominid-specific histone H3 variant that is specifically expressed in the seminiferous tubules of human testes. *Chromosoma* **120**: 275-285.
- Shechter D, Dormann HL, Allis CD, Hake SB. 2007. Extraction, purification and analysis of histones. *Nature protocols* **2**: 1445-1457.
- Shinagawa T, Huynh LM, Takagi T, Tsukamoto D, Tomaru C, Kwak HG, Dohmae N, Noguchi J, Ishii S. 2015. Disruption of Th2a and Th2b genes causes defects in spermatogenesis. *Development*.
- Skene PJ, Henikoff S. 2013. Histone variants in pluripotency and disease. *Development* **140**: 2513-2524.
- Song N, Liu J, An S, Nishino T, Hishikawa Y, Koji T. 2011. Immunohistochemical Analysis of Histone H3 Modifications in Germ Cells during Mouse Spermatogenesis. *Acta histochemica et cytochemica* **44**: 183-190.
- Tachiwana H, Kagawa W, Osakabe A, Kawaguchi K, Shiga T, Hayashi-Takanaka Y, Kimura H, Kurumizaka H. 2010. Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 10454-10459.
- Tachiwana H, Osakabe A, Kimura H, Kurumizaka H. 2008. Nucleosome formation with the testis-specific histone H3 variant, H3t, by human nucleosome assembly proteins in vitro. *Nucleic acids research* **36**: 2208-2218.
- Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**: 51-61.
- Tang MC, Jacobs SA, Mattiske DM, Soh YM, Graham AN, Tran A, Lim SL, Hudson DF, Kalitsis P, O'Bryan MK et al. 2015. Contribution of the Two Genes Encoding Histone Variant H3.3 to Viability and Fertility in Mice. *PLoS genetics* **11**: e1004964.
- Trostle-Weige PK, Meistrich ML, Brock WA, Nishioka K. 1984. Isolation and characterization of TH3, a germ cell-specific variant of histone 3 in rat testis. *The Journal of biological chemistry* **259**: 8769-8776.
- Turner JM, Mahadevaiah SK, Fernandez-Capetillo O, Nussenzweig A, Xu X, Deng CX, Burgoyne PS. 2005. Silencing of unsynapsed meiotic chromosomes in the mouse. *Nature genetics* **37**: 41-47.
- Wiedemann SM, Mildner SN, Bonisch C, Israel L, Maiser A, Matheisl S, Straub T, Merkl R, Leonhardt H, Kremmer E et al. 2010. Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *The Journal of cell biology* **190**: 777-791.
- Witt O, Albig W, Doenecke D. 1996. Testis-specific expression of a novel human H3 histone gene. *Experimental cell research* **229**: 301-306.
- Yuen BT, Bush KM, Barrilleaux BL, Cotterman R, Knoepfler PS. 2014. Histone H3.3 regulates dynamic chromatin states during spermatogenesis. *Development* **141**: 3483-3494.
- Zentner GE, Henikoff S. 2013. Regulation of nucleosome dynamics by histone modifications. *Nature structural & molecular biology* **20**: 259-266.

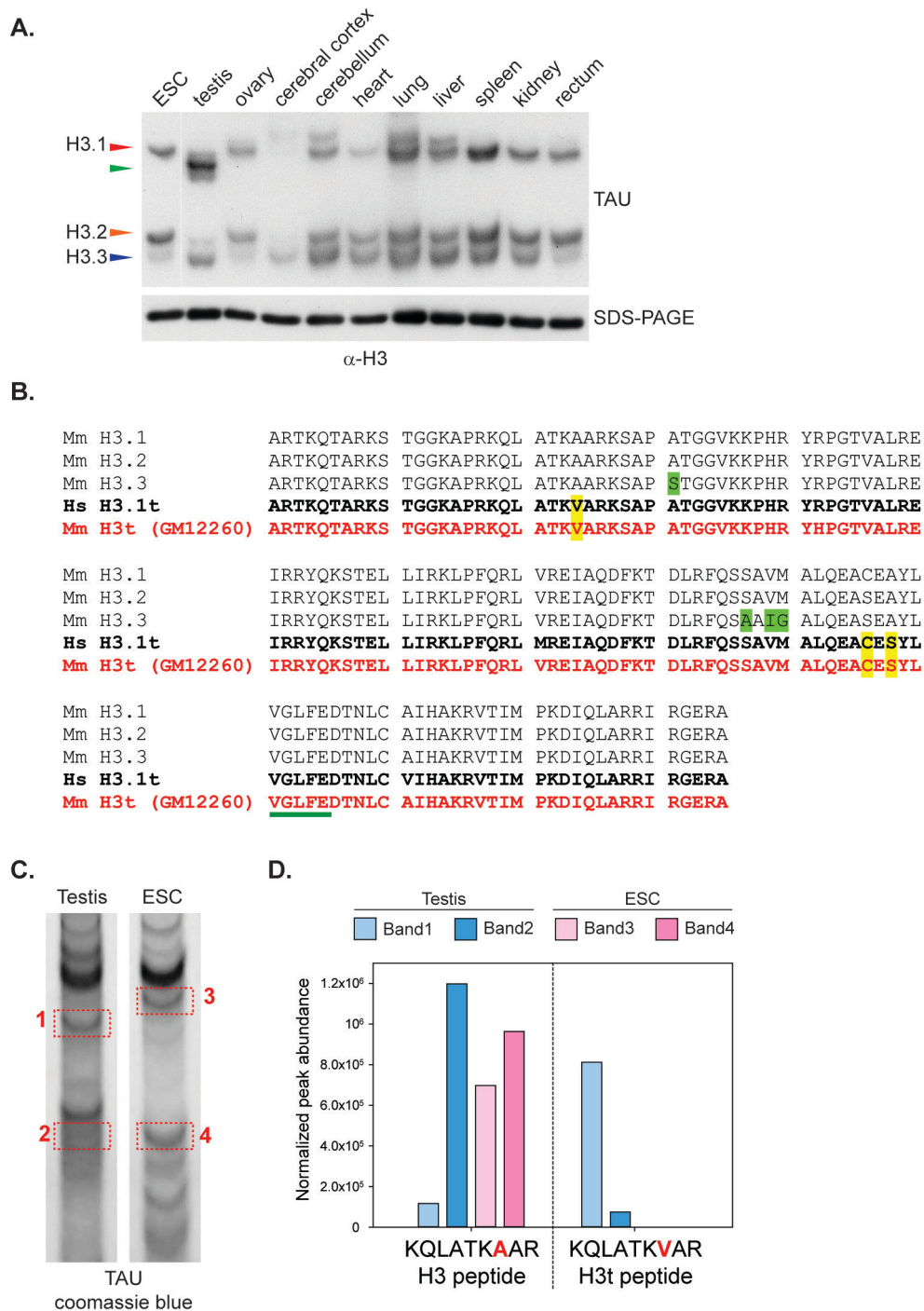


Figure 1. Identification of testis-specific H3 (H3t) in mouse testis. (A) Protein expression of H3 proteins in different mouse tissues. H3.1, H3.2 and H3.3 were separated by TAU gel and analyzed by Western blotting with general H3 antibody. SDS-PAGE-Western blotting was used for verifying loading amount of total H3 extracted from each mouse tissues. (B) Amino acid sequence alignment of mouse H3.1, H3.2, H3.3, human H3t (H3.4), and mouse H3t translated from pseudogene *GM12260*. And amino acids specifically on human and mouse H3t are highlighted in yellow. (C) Coomassie blue staining of histones from testis and mouse embryonic stem cells (ESC) on TAU gel. Red boxes represent H3 bands excited for mass spectrometry analysis. (D) Normalized peak abundances of the canonical H3 (18-26) peptide KQLATKAAR and the human replacement variant H3t(18-26) peptide KQLATKVAR in the TAU gel bands one to four (see Figure 1C). The N-term and the lysines of the peptides are propionylated.

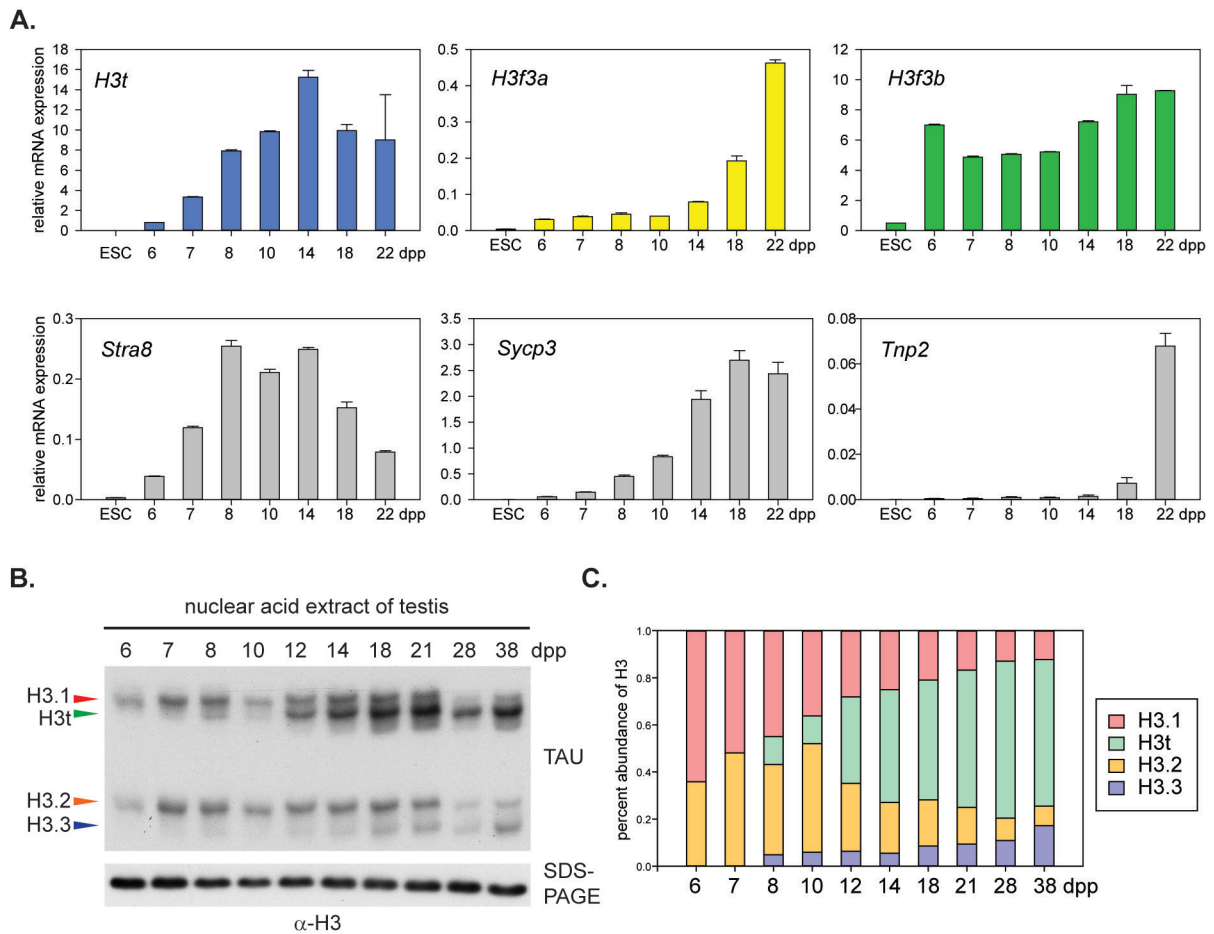


Figure 2. Mouse H3t expression during the first wave of spermatogenesis (A) RNA expression level of *H3t*, *H3f3a* and *H3f3b* in total testes isolated from different days post-partum (dpp) of C57BL/6 mice. *Stra8*, *SycP3* and *Tnp2* RNA expression levels during the first wave of spermatogenesis are as developmental controls. Ct values analyzed by qPCR were normalized to the values of housekeeping gene, ribosomal protein *S17*. (B) Protein expression of four H3 proteins from 4 dpp to adult testes. The equal amount of total H3 from each testis sample was loaded into TAU gel. Arrowheads indicate localizations of H3 variants on TAU gel. (C) Quantification of the proportion of four H3 proteins in testis at different dpp from (B). The intensity of each H3 band was calculated by ImageJ.

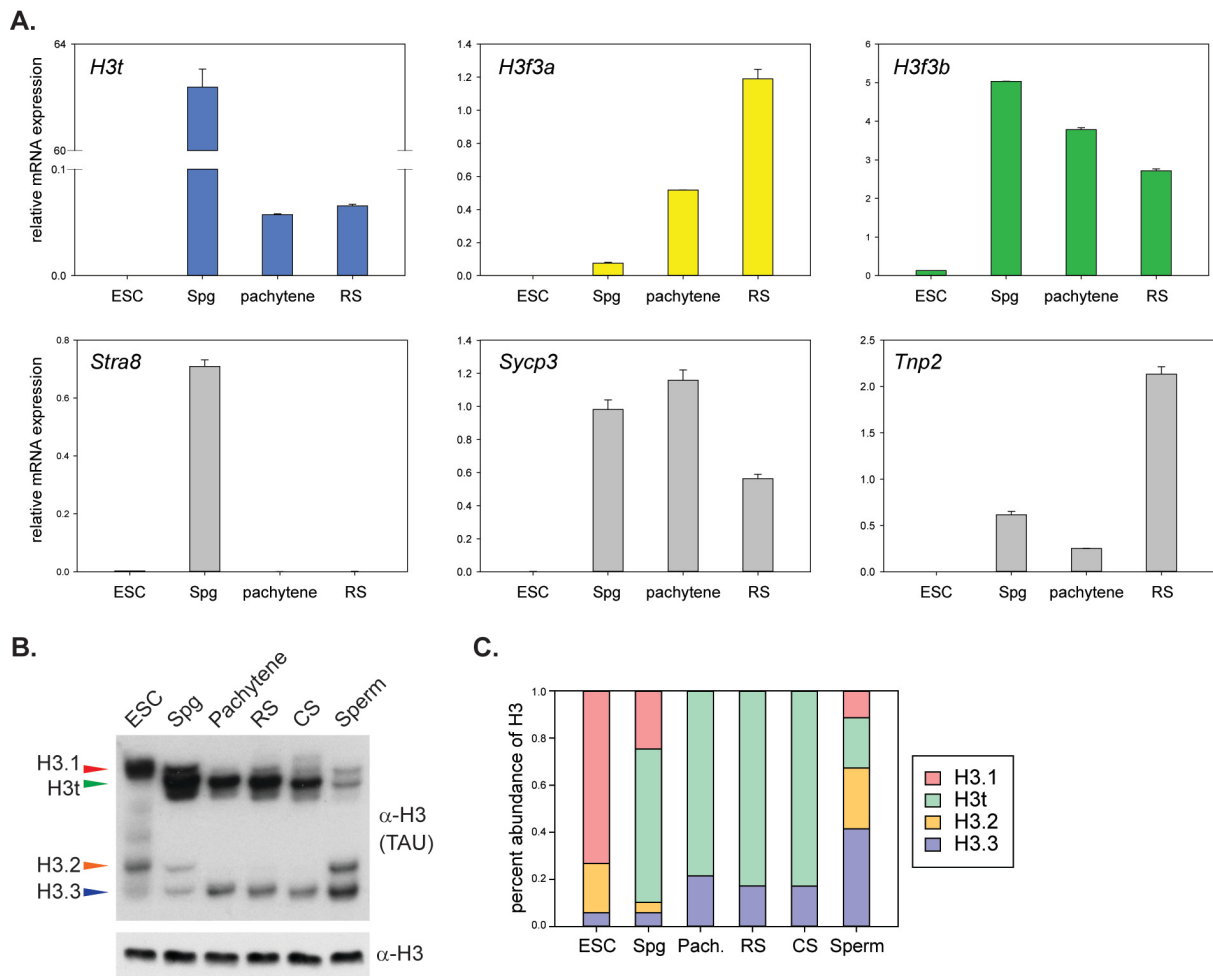


Figure 3. The dynamic expression of H3 variants from spermatogonia to mature sperm. (A) Relative RNA expression level of *H3t*, *H3f3a*, *H3f3b*, and developmental control genes (*Stra8*, *Sycp3* and *Tnp2*) normalized to housekeeper ribosomal protein *S17*. Type-B spermatogonia (Spg) and pachytene spermatocytes (Pachytene) were purified from adult mouse testes by using FACS, and round spermatids (RS) were isolated by centrifugal elutriation. (B) Protein expression of four H3 from spermatogonia to mature sperm compared to ES cells. TAU gel was used to separating H3 proteins, and equivalent of total H3 was loaded. Condensing spermatids (CS) were purified by centrifugal elutriation, and mature sperm was isolated from mouse epididymis by swim-up methods. (C) Quantification of the proportion of each H3 variant in ES cells and spermatogenic cells from (B). The intensity of each H3 band was calculated by ImageJ.

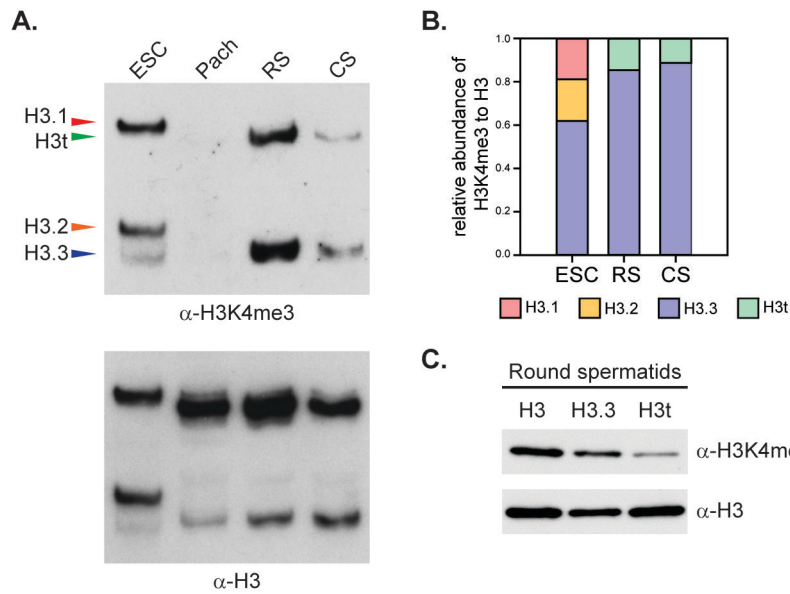
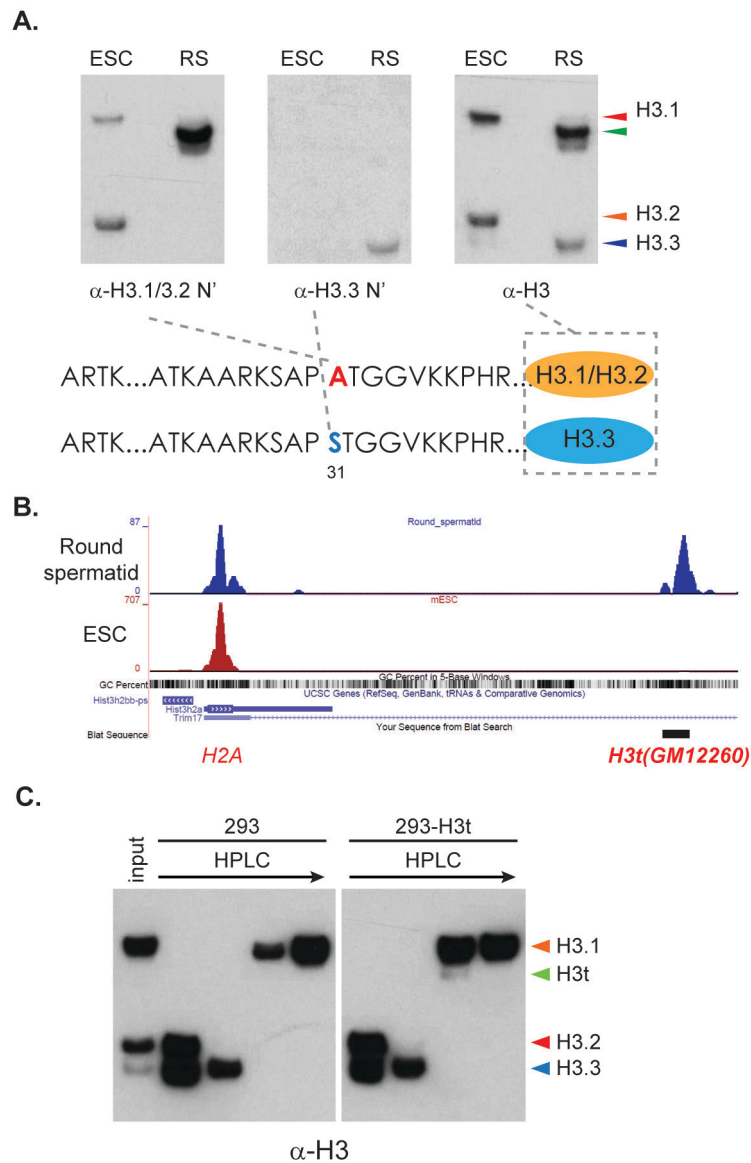
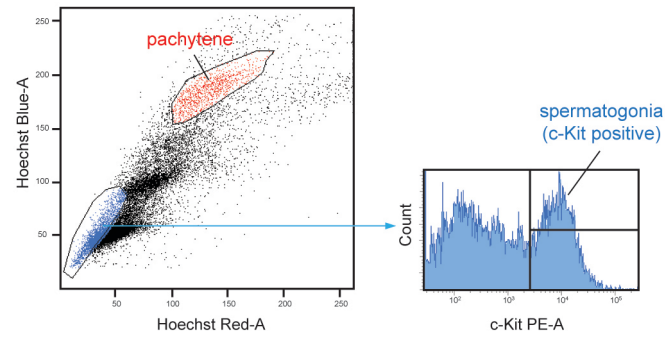


Figure 4. H3t contains less H3K4me3 modification in spermatogenic cells. (A) The results of TAU gel-Western blotting probed with H3K4me3 and H3 antibodies. Histone H3 was extracted from mouse embryonic stem cell (ESC), pachytene spermatocyte (Pach), round spermatid (RS) and condensing spermatid (CS). (B) The percent abundance of H3K4me3 on H3 variants. The intensity of each H3K4me3 band on TAU gel-Western blotting from figure A was measured by ImageJ and normalized by the intensity of corresponding H3 bands. (C) The levels of H3K4me3 on H3.3 and H3t from round spermatid extract were analyzed by SDS-PAGE-Western blotting. RP-HPLC was used for separating H3.3 and H3t from round spermatid acid extract.

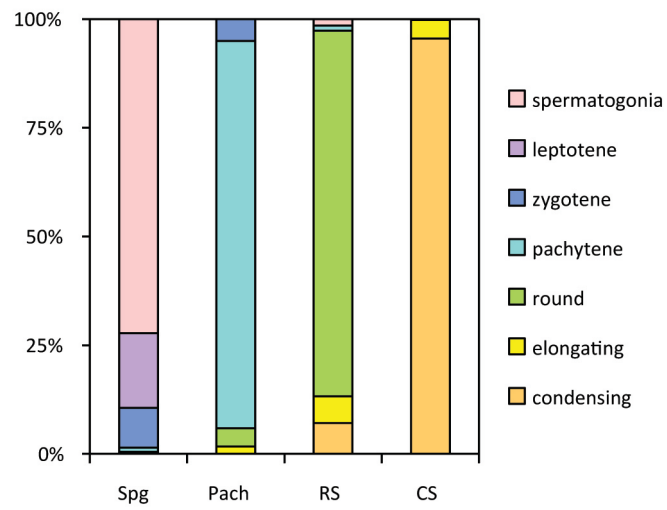


Supplementary figure 1. The characterization of mouse H3t. (A) Histones were extracted from in mouse embryonic stem cells (ESC) and round spermatids (RS) for TAU gel-Western blotting probed with general H3, H3.3-specific and H3.1/H3.2-specific antibodies. (B) *GM12260* pseudogene is highly transcribed in mouse spermatids. UCSC genome browser snapshot of *GM12260* (*H3t*) locus shows RNA-seq results from round spermatid and ESC. (C) The confirmation of H3t localization on TAU gel. H3t was overexpressed in 293 cells, and separated by RP-HPLC and subsequently by TAU gel. Green arrowhead indicated over-expressed H3t.

A.



B.



Supplementary figure 2. The purification of different stage of spermatogenic cells from mouse testes. (A) The Hoechst FACS profile of testicular cells. The population of pachytene spermatocyte labeled by red was isolated by FACS. Spermatogonia were purified based on DNA content and c-Kit level. (B) The cell purity of each spermatogenic cells fraction. Spermatogonia (Spg) and pachytene spermatocyte (Pach) were isolated by FACS. Round and condensing spermatid (RS and CS) were isolated by centrifugal elutriation.

Chapter 5

Proteolytic cleavage of histone H3 by a serine protease during mouse spermiogenesis

Ching-Yeu Liang^{1, 2}, Helene Royo¹, Ragna Sack¹, Frédéric Zilbermann¹, Hubertus Kohler¹, Jeremy Keusch¹ and Antoine H. F. M. Peters^{1,2}

1. Friedrich Miescher Institute for Biomedical Research (FMI), Maulbeerstrasse 66, 4058 Basel, Switzerland

2. Faculty of Sciences, University of Basel, 4056 Basel, Switzerland

Keywords:

Histone H3, proteolysis, H3 cleavage, mouse spermiogenesis and serine protease

Abstract

Mammalian spermiogenesis is a post-meiotic differentiation process during which haploid male germ cells dramatically remodel their chromatin structure and undergo a remarkable cellular morphogenesis to become mature sperm. As part of the chromatin remodeling events, most of histones are replaced by transition proteins and subsequently by basic protamines. This histone-to-protamine exchange results in the efficient chromatin condensation to package paternal DNA into the sperm head and is essential for male fertility. Nonetheless, previous studies have shown that around 1% to 10% nucleosomes are retained in mouse and human mature sperm. The residual nucleosomes may provide means for the transmission of paternal epigenetic information between generations. The mechanism of global chromatin remodeling during spermiogenesis remains largely unclear. Here, we observe that histone H3 is cleaved at its N-terminal tail by a serine protease activity in nuclei of late stage spermatids and that cleaved H3 histones are subject to a rapid degradation by the proteasome. This H3 proteolytic cleavage is conserved from mouse to human. We propose that removal of the H3 N-terminal tail destabilizes nucleosome structure and may thereby contribute to nucleosome disassembly and global histone removal during spermiogenesis. We map the cleavage sites by mass spectrometry analysis and identify that arginine 26 and lysine 27 on the H3 N-terminal tail are crucial residues for H3 protease activity. Strikingly, we find that acetylation on H3 inhibits H3 cleavage process *in vitro*. Likewise, the genome-wide distribution of acetylation on H3 Lysine 27 (H3K27Ac) is positively correlated with the occupancy of nucleosomes containing transcriptionally active mark in mature sperm, indicating that preventing H3 from proteolytic cleavage by H3K27Ac may contribute to specific nucleosome retention in sperm.

Introduction

Nucleosome is a basic unit of chromatin organization in eukaryotes, which packages genomic DNA efficiently into the cell nucleus. It consists of 146 bp DNA wrapped around a histone octamer assembled by two copies of core histone variants, H3, H4, H2A and H2B (Luger et al. 1997). Besides being a packaging unit, nucleosomes play important regulatory roles in many nuclear processes such as transcription, DNA replication and DNA repair by

modulating DNA accessibility. It is well-known that histones are subject to post-translational modifications by histone-modifying enzymes that alter the histone-DNA interaction and provide docking sites for chromatin-associated proteins to regulate chromatin structure (Musselman et al. 2012; Zentner and Henikoff 2013). Moreover, nucleosome structure can be perturbed by ATP-dependent chromatin remodeling enzymes that use the energy of ATP hydrolysis to evict nucleosomes, mobilize nucleosomes along DNA and exchange histones (Becker and Workman 2013). Thus, histone modifying and chromatin remodeling enzymes are crucial for the regulation of nucleosome dynamics and DNA accessibility. Recently, the proteolytic cleavage of histone has been reported to facilitate the nucleosome turnover. For example, histone H3 is proteolytically cleaved by cathepsin L protease on its N-terminal tail during mouse embryonic stem cell differentiation (Duncan et al. 2008). And similar H3 tail cleavage is also reported in other cellular processes, such as cell senescence, viral infection and sporulation (Falk et al. 1990; Tesar and Marquardt 1990; Santos-Rosa et al. 2009; Duarte et al. 2014). Because histone tails can stabilize nucleosome structure and contain many post-translational modification sites for recruiting chromatin regulator, histone tail clipping may influence chromatin structure and then promote histone eviction.

During mammalian spermiogenesis, chromatin in spermatids is subject to dramatic remodeling. Most of histones are replaced in elongating spermatids by transition proteins and subsequently by protamines (Rousseaux et al. 2005). Protamines are arginine-rich proteins strongly interacting with DNA and to condense genomic DNA into the small sperm head. The histone-to-protamine exchange plays a critical role in sperm maturation and male fertility (Oliva 2006). Based on previous studies, H4 hyperacetylation observed at elongating spermatid stage is thought to reduce nucleosome stability and promote histone removal from DNA (Grimes and Henderson 1984; Govin et al. 2007). Also Rnf8-dependent histone ubiquitination facilitates histone removal during histone-to-protamine exchange in elongating spermatids through regulating H4K16 acetylation (Lu et al. 2010). In addition, a recent study showed that PA200, a proteasome activator highly expressed testes, recognizes acetylated H4 and H2B through its Bromodomain-like region and involves in the histone degradation during spermatogenesis (Qian et al. 2013). Nonetheless, the molecular mechanism of global histone eviction from chromatin during spermiogenesis is largely unclear. We and others recently showed that 1-10 % of nucleosomes are retained at specific

genomic regions, particularly at unmethylated CpG-rich promoters, in human and mouse sperm (Gatewood et al. 1987; Hammoud et al. 2009; Brykczynska et al. 2010; Erkek et al. 2013). These residual nucleosomes carry active or repressed modifications that may involve in the paternal transmission of epigenetic information to next generation. But the mechanisms of specific nucleosome retention in sperm chromatin are unresolved.

In this study, we identify a proteolytic cleavage of the N-terminal tail of histone H3 during at the late stage of spermiogenesis, which may destabilize nucleosome structure and thereby promote nucleosome eviction during the histone-to-protamine exchange. This H3 cleavage processing is catalyzed by a putative serine protease present in spermatid nuclei. We show that residue Arg26 and Lys27 on H3 are important for H3 protease activity and that H3 acetylation inhibits H3 protease activity. Further, chromatin immunoprecipitation analysis indicates that acetylation on H3 Lys27 is correlated with nucleosome retention at promoter regions containing transcriptionally active mark in mature sperm. Taken together, our studies demonstrate a H3 tail cleavage process during spermiogenesis which may contribute to global nucleosome eviction.

Results

Identification of short form of H3 found in the later stage of spermiogenesis

It is well-known that histones are evicted from the genome and replaced by transition proteins and then by protamines during mammal spermiogenesis. In order to investigate the dynamics of histone turnover during mouse spermiogenesis, we use fluorescence-activated cell sorting (FACS) to isolated pachytene spermatocytes from adult mouse testes by DNA content and performed centrifugal elutriator to isolate different stages of spermatids by cell size. While performing histone acid extraction and then immunoblotting with antibody against C-terminal part of general histone H3, we observe a faster running species of H3 specifically in elongating/condensing spermatids (Esp/CS) and sperm compared with spermatocytes (Spc), round spermatids (RS) and embryonic stem cells (ESC) (Figure 1A and 1B). But we do not detect any fast migrating bands of other core histones, H4, H2A and H2B in mouse spermatids (Figure 1C). These results indicate that the proteolytic cleavage of H3

N-terminal tail occurs during spermiogenesis. Previous *in vitro* studies have showed that the removal of H3 N-terminal tail decreases nucleosome stability and affects histone-DNA interaction (Biswas et al. 2011; Iwasaki et al. 2013). Thus, H3 cleavage occurring at the late stage of spermatid may promote global nucleosome eviction by nucleosome destabilization. Moreover, we detect the cleaved form of H3 also in human mature sperm (Figure 1B), suggesting that N-terminal H3 cleavage is conserved in mammalian species.

Furthermore, a previous study demonstrated that proteasome containing PA200 activator is required for histone degradation in mouse spermatids (Qian et al. 2013). To confirm whether cleaved H3 can be degraded by proteasomes, we isolate elongating/condensing spermatids from mouse testes, and cultured them with proteasome inhibitor, MG132, for 24 hours. As showed in Figure 1D and 1E, inhibiting the proteasome by MG132 causes the accumulation of cleaved form of H3 in spermatids. Thus, this finding supports that cleaved H3 could be subsequently degraded by proteasomes.

H3 protease in spermatids is from serine protease family

To determine which mouse spermatogenic cells contain H3 protease activity, we incubated nuclei isolated from ES cells, pachytene spermatocytes, round spermatids and elongating/condensing spermatids in presence or absence of protease inhibitors at 32 °C for half hour. Strikingly, we find that H3 in spermatid nuclei is more extensively cleaved in the absence of the protease inhibitors *in vitro*, not in spermatocyte and ES cell nuclei. Thus H3 protease is present in the nuclei of round and elongating/condensing spermatids (Figure 2A). Likewise, treating spermatids with 0.5 % paraformaldehyde before *in vitro* incubation also inhibits further H3 cleavage *in vitro* (Figure 2B). However we do not detect any cleaved H3 in round spermatids *in vivo* (Figure 1A). The possible explanation is that round spermatids have an inhibition mechanism of H3 protease activity, but it may be disturbed during nucleus isolation. Furthermore, in order to identify which protease family H3 protease in spermatid belongs, we use the same *in vitro* assay but replaced general protease inhibitors to individual protease inhibitors that inactivate specific protease families. The experiments showed that the serine protease inhibitors, AEBSF and Aprotinin, prevent H3 cleavage *in vitro* (Figure 2C), but not cysteine protease inhibitor, E64. Thus, we propose that H3 protease activity in mouse spermatids is from serine protease family.

The characterization of H3 protease from spermatids *in vitro*

To characterize the putative H3 protease in mouse spermatids, we establish an *in vitro* H3 cleavage assay. Because H3.3, a replication-independent H3 variant, is abundant in mouse spermatid and sperm (Erkek et al. 2013), and we have known that the cleavage site is localized on H3 N-terminal tail, we purify recombinant histone H3.3 with C-terminal 6xHis tag from *E. coli* or mononucleosomes containing overexpressed H3.3 with C-terminal 3xFlag tag from HEK293 cells as substrates. We then incubate them with nuclear extracts of elongating/condensing spermatids at 32°C and detected the cleaved H3 products by immunoblotting with anti-His or anti-Flag antibody (Figure 3A). Interestingly, although both H3 substrates are cleaved at N-terminus by the H3 protease present in nuclear extracts of spermatids in this assay, the H3 in nucleosomes isolated from HEK293 cells is a better substrate than native H3 purified from *E. coli* (Figure 3B and C). It is possible that nucleosome structure and post-translational modifications on H3 affect the enzyme activity and the binding affinity of H3 protease. Moreover, this H3 cleavage *in vitro* reaction is also inhibited by the serine protease inhibitor AEBSEF, not by proteasome inhibitor MG132 (Figure 3D and Supplementary figure 1A). Therefore, H3 cleavage *in vitro* results from a serine protease activity, not from proteasome degradation.

To identify the subcellular localization of the H3 protease activity in spermatids, we fractionate spermatid extract into cytoplasmic, nuclear-soluble and chromatin-bound fractions for the *in vitro* H3 cleavage assay. Additionally, it is well-known that spermatid has an acrosomal compartment that is derived from the Golgi apparatus and contains abundant proteases for fertilization. In order to exclude a possible contamination by proteases from acrosome in each fraction, we use an antibody recognizing Golgi protein, RCAS1 (Engelsberg et al. 2003), to verify where Golgi-derived acrosomal components are. The result show that H3 protease activity is present in nuclear-soluble and chromatin-bound fraction, but not in cytoplasmic fraction containing Golgi proteins (Figure 3E). Therefore H3 protease activity measured in spermatid nuclei is not due to contaminating acrosomal proteases during cell extraction.

Arg26 and Lys27 on H3 is important for H3 protease activity

We next ask which amino acid residue on H3 tail is important for H3 protease activity. Since the molecular weight of the cleaved H3 is around 2 to 3 kD less than a full-length H3, we hypothesize that a major cleavage site would localize between Ala21 and Pro30 residues on H3 N-terminal tail and that some residues in this region may be necessary for H3 protease recognition, binding or activity. We, therefore, express a series of mutant H3-3xFlag individually in HEK293 cells and purified mononucleosomes containing mutant H3 by using anti-Flag resins as substrates for *in vitro* cleavage assay. We find that only Arg26 and Lys27 double mutations (R26A/K27A) on H3 completely inhibit cleavage process *in vitro*, but not other substitutions (Figure 3F and Supplementary figure 1B). Therefore Arg26 and Lys 27 residues are crucial for H3 protease activity or binding. In addition, this substrate preference is similar to trypsin-like proteases, which cleave peptide bonds following a positively charged amino acid, such as lysines and arginines.

Determination of proteolytic cleavage sites of H3 in spermatids

To study in more detail the proteolytic cleavages sites on the H3 N-terminal tail, we first use antibodies against different H3 modifications and variants for immunoblotting to check whether they can recognize cleaved H3. Remarkably, besides the antibody against C-terminal part of H3 that we had used above, we also observe that the cleaved H3 produce is detected by antibodies recognizing trimethylation on H3 Lys27 (H3K27me3), phosphorylation on H3.3 Ser31 (H3.3S31ph) and C-terminal part of H3.3. By contrast, antibodies again trimethylation on H3 Lys4 (H3K4me3) and acetylation on H3 Lys27 (H3K27Ac) did not recognize the cleaved H3 (Figure 4A). These results reveal that proteolytic cleavage sites are localized around Lys27 residue on H3 tail and that acetylation and methylation on Lys27 may influence H3 protease activity or target recognition.

Next, we examine precious cleavage sites by mass spectrometry analysis. Because cleaved H3 may be degraded by proteasomes quickly in elongating spermatids after protease proteolysis, cleaved H3 existed in a small portion of total H3. Therefore, we use cleaved H3 3xFlag produced abundantly from *in vitro* H3 cleavage assay to verify cleavage sites by mass spectrometry analysis. Moreover, H3 contains several lysine and arginine residues on its tail that leads to too short trypsin-digested peptides to be analyzed efficiently by MS. We therefore replace the wild-type H3 subtract to H3 R26A mutation that did not affect

cleavage processing *in vitro* (Figure 3F) and can generate a longer peptide (from Lys18 to R40) covering putative cleavage sites after lysine propionylation and subsequently trypsin digestion. Strikingly, compared to full-length H3 band, mass spectrometry analysis shows that the cleaved H3 band has remarkably more four N-terminal truncated peptides beginning at Ala24, Lys27, Ser28 and Thr32 (Figure 4B). And the truncated peptide from Ser28 to Arg40 was relatively more abundant than other truncated peptides in cleaved H3 band (Figure 4C). Therefore we propose that a major cleavage site is between amino acids 27 and 28 on H3 tail, and that a minor cleavage site is between amino acids 23 and 24.

Identification of H3 protease in the mouse testis

In order to identify H3 protease in mouse spermatids, we first use CaptoQ sepharose to fractionate the nuclear extract of spermatids by the gradient of salt and then examine H3 protease activity from each fraction by the *in vitro* H3 cleavage assay (Figure 5A). In addition, based on above findings, the putative H3 protease in mouse spermatids has a substrate preference that is similar to that of trypsin-like proteases. Therefore, we use benzamidine sepharose, which binds strongly with the catalytic domain of trypsin or trypsin-like proteases, to fractionate testis extract. Proteins binding on benzamidine sepharose are eluted by acid glycine solution, and then we confirm H3 protease activity in each fraction by the *in vitro* H3 cleavage assay (Figure 5B). Through protein identification by mass spectrometry analysis, we find that the peptides of one serine protease, tripeptidyl peptidase II (Tpp2), is specifically detected in fractions exhibiting H3 protease activity from both sepharose columns (Supplementary figure 2). In addition, *Tpp2* is highly expressed in mouse spermatids compared with *cathepsin L*, a well-known H3 protease in mouse ES cells (Duncan et al. 2008) (Figure 5C). Therefore we propose that Tpp2 is a putative H3 protease in mouse spermatids.

H3 acetylation prevent H3 from cleavage

It is well-known that histone H4 becomes generally hyperacetylated in elongating spermatids during the process of histone-to-protamine exchange. Likewise we observed acetylation on Lys27 of H3 in elongating spermatids by immunostaining (Supplementary figure 3). So we next ask whether H3 acetylation regulates H3 cleavage processing. To address this question, we incubate native H3 with 6xHis tag purified from *E. coli* with p300

catalytic domain that can acetylate lysine residues, including Lys27, on H3 *in vitro* (Figure 6A). Surprisingly, when using acetylated H3 as substrate for *in vitro* cleavage assay, all cleavage is inhibited (Figure 6B). And based on above findings, we had already known a major cleavage site is at Lys27 that also can be acetylated by p300 *in vitro*. Together, we propose that acetylation on Lys27 may be an important factor preventing cleavage processing.

H3K27Ac is relative to nucleosome retention with active marker in mature sperm

Our previous study has found that there are around 1% nucleosomes retained in mouse sperm (Brykczynska et al. 2010), and that these retained nucleosomes are enriched at CpG-rich promoters marked by either H3K4me3 or H3K27me3 (Erkek et al. 2013). Therefore, we next ask whether H3K27Ac preventing H3 from cleavage is relative to nucleosome retention during sperm maturation. To verify the relationship between H3K27Ac and nucleosome retention, we perform chromatin immunoprecipitation coupled to deep sequencing (ChIP-seq) with anti-H3K27Ac antibody in round, elongating and condensing spermatids. H3K27Ac distribution is consistent from round to condensing spermatids (Supplementary figure 4A) and also is enriched more at CpG-rich promoters than CpG-low promoters (Supplementary figure 4B and C). Compared to nucleosome occupancy in mouse sperm from our published data, we observe that H3K27Ac enrichment at transcriptional start sites (TSS) in round spermatids correlates modestly with nucleosome-retention region in sperm (Figure 6C and 6D left panel). But when we divide TSS into CpG island (CGI)-contained TSS and non-CGI TSS and then overlay scatter plots with H3K27me3 enrichment in round spermatids, a repressive histone marker, the results indicate that nucleosome-retained regions in sperm without H3K27Ac accumulation are marked by H3K27me3, specifically at CGI-contained TSS (Figure 6D middle and right panel). Moreover, we observe that H3K27Ac distribution in spermatids is positively correlated with the occupancy of the residual nucleosomes containing transcriptionally active histone mark, H3K4me3, at TSS in mature sperm (Figure 6E). We therefore speculate that H3K27Ac may protect H3 containing active markers at TSS from proteolytic cleavage and then modulate nucleosome retention in mature sperm.

Discussion

During mammalian spermiogenesis, spermatids undergo dramatic morphological and chromatin-structural changes to differentiate into spermatozoa. In this differentiation processing, most nucleosomes are evicted from chromatin and are replaced by transition proteins and then by protamines to condense paternal DNA into the small sperm head. Compared to somatic cells, only 1 to 10 % nucleosomes are retained in mouse and human sperm chromatin (Brykczynska et al. 2010). Despite remarkable progress in the elucidation of histone-to-protamine exchange during spermiogenesis, a comprehensive explanation of the mechanisms of nucleosome eviction and retention remains elusive. We observe a proteolytic cleavage of histone H3 occurring at the late stage of spermatids and sperm. Based on the crystal structure of the nucleosome core particle, N-terminal tail of H3 exposed on the surface of the nucleosome contains an α -helices (amino acid 14-20) that can interact with DNA at the edge of nucleosome (Biswas et al. 2011). The truncation of H3 tail has been shown to result in the alternation of histone-DNA contact, nucleosome destabilization, and variation in the rate of nucleosome sliding on DNA *in vitro* (Ferreira et al. 2007). In addition, expressing tail-truncated H3 in cells influences the replication-coupled nucleosome assembly and disrupts the formation of a silent higher order chromatin structure (Ahmad and Henikoff 2002; Sperling and Grunstein 2009). Together, H3 tail is required for assembling nucleosome, maintaining nucleosome structure, regulating DNA accessibility. Therefore, H3 proteolytic cleavage occurring in spermatid nuclei may destabilize nucleosome structure and then contribute to global nucleosome eviction during histone-to-protamine exchange processing.

H3 tail cleavage has been reported in different biological systems, and several proteases belonging to either cysteine or serine protease families are involved in H3 cleavage processing. For example, cathepsin L, a cysteine protease, was found to cleave H3 tail around Ala21 residue during mouse embryonic stem cell differentiation, and acetylation on Lys23 can reduce cathepsin L-mediated cleavage of H3 (Duncan et al. 2008). Moreover, cathepsin L also manipulates H3 cleavage during oncogene-induced senescence in fibroblasts and melanocytes (Duarte et al. 2014). But the H3 protease activity identified in human embryonic stem cells is exerted by a serine protease activity, not by cathepsin L (Vossaert et al. 2014). And recently two serine proteases, glutamate dehydrogenase (GDH) in chicken livers (Mandal et al. 2013; Purohit et al. 2013) and PRB1 in yeast (Xue et al. 2014) have been

identified as histone H3 specific proteases. These suggest that H3 can be clipped by distant proteases depend on cell type and species. In our study, the H3 protease activity in spermatid nuclei is from a trypsin-like serine protease. After fractionating spermatid nuclear extract and analyzing protein content by mass spectrometry, we identify a trypsin-like protease, tripeptidyl peptidase II (Tpp2), enriched in fractions with H3 protease activity. In future, we will continue to investigate whether Tpp2 mediates H3 cleavage processing during spermiogenesis.

Acrosome formation during spermiogenesis has been reported to influence nuclear compaction and chromatin reorganization during spermiogenesis (Kang-Decker et al. 2001; Yao et al. 2002; Lin et al. 2007; Xiao et al. 2009; Fujihara et al. 2012). A DAPI-intense doughnut-like structure is observed in round spermatid nuclei (De Vries et al. 2012). This structure is co-localized with acrosomal sac, a Golgi-derived sac-like structure attached to spermatid nucleus. And less nucleosomes and histones are localized at the doughnut-like structure, but transition proteins are enriched in it in spermatids. Moreover, H4 acetylation is also found to be enriched in doughnut-like structure. Therefore, this DNA-intense doughnut-like structure is thought to be an initiation site of chromatin remodeling in histone-to-protamine exchange processing (De Vries et al. 2012). Because a putative H3 protease, Tpp2, have been found to be accumulated in acrosome in mature sperm (Zhou et al. 2013), it is possible that Tpp2 could be shuttle between spermatid nuclei and acrosome through this doughnut-like structure and subsequently cleave H3 to promote nucleosome removal at this structure.

Proteasome has been reported to play an important role in histone degradation during spermiogenesis. PA200, an activator of proteasome, is highly expressed in testes, and its deficiency leads to defective spermatogenesis with abnormal spermatocytes and spermatids and the reduction of male fertility (Khor et al. 2006). Recently, Qian et al. showed that core histones are accumulated at elongating spermatids in PA200-deficient testis. PA200 contains bromodomain-like regions that recognize acetylated histones, especially acetyl-H2B and acetyl-H4, to degrade acetylated histones (Qian et al. 2013). Thus PA200-mediated histone degradation is acetylation-dependent, not polyubiquitin-dependent. Furthermore, in this present study, we find that inhibiting proteasomes by MG132 inhibitor causes the accumulation of cleaved form of H3 in elongating/condensing spermatids. Therefore, as in

Figure 7, we propose that histone H3 is cleaved on its tail by a putative serine protease that might destabilize nucleosome structure, and subsequently, destabilized nucleosomes may be evicted by uncharacterized chromatin remodelers. It is possible that the H4 hyperacetylation and H3 proteolytic cleavage are present together in the same nucleosomes at elongating spermatid stage, so evicted nucleosomes containing H4 acetylation and cleaved H3 are subsequently degraded by PA200 proteasome or other proteasomes. In future research, it is interesting to dissect the correlation between hyperacetylated H4 and H3 proteolytic cleavage during histone-to-protamine exchange in elongating spermatids. In addition, we find H3K27Ac prevents H3 from cleavage proteolysis and is correlated to nucleosome retention at active promoters in mature sperm. However, whether H3K27me3 also has the same role as H3K27Ac in protecting H3 from cleavage remains an important issue that requires further investigations. In summary, our findings reveal a histone H3 proteolysis processing during mouse spermiogenesis that may contribute to global nucleosome removal.

Materials and methods

Antibodies

Primary antibodies used for Western blotting and chromatin immunoprecipitation were H3 (Abcam, ab1791), gamma-H2A.X (Millipore, 05-636), H4Ac (Millipore, 06-866), beta-tubulin (SIGMA, T4026), H4 (Abcam, ab10158), Tnp2 (Santa Cruz, sc-21106), His-tag (Invitrogen, 37-2900), Flag-tag (SIGMA, F7425), H3K4me3 (Millipore, 07-473), H3K27Ac (Abcam, ab 4729), H3Ac (Millipore, 06-599), H3K27me3 (Millipore, 07-449), H3.3-phospho S31 (Abcam, ab92628), H3.3 (Millipore, 09-838).

Isolation of mouse spermatocyte, spermatid and sperm

Male C57BL/6J mice were used in this study. For isolating pachytene spermatocytes, testicular cells were trypsinized from adult male testes and then were stained by Hoechst 33342 (Invitrogen, cat. no. 33342) for 30 min at 37 °C. Pachytene spermatocytes were collected by fluorescence-activated cell sorter (FACS) according to previous study (Getun et al. 2010). And round, elongating and condensing spermatids were separated individually

from testicular cells by centrifugal elutriation based on cell size as described (Barchi et al. 2009). Mature sperm was obtained from adult epididyma by performing a swim-up procedure as described (Hisano et al. 2013). All experiments were performed in accordance with the Swiss animal protection laws and institutional guidelines.

Histone extraction

Cells were incubated with Buffer A (10 mM HEPES pH 7.5, 1.5 mM MgCl₂, 10 mM KCl, 0.05 % Nonidet P-40, 0.5mM DTT, 5 mM Sodium butyrate and protease inhibitor cocktail) to separate nucleus and cytoplasm. After centrifugation, nuclear pellet was re-suspended into 0.2 N HCl and incubated on ice for at least 1 hour to solubilize histones. Then histones were precipitated by trichloroacetic acid (TCA) and finally were dissolved into Milli-Q water.

Protease activity in mouse spermatid nuclei

In Figure 2, cells were incubated with ice-cold Buffer A (10 mM HEPES pH 7.5, 1.5 mM MgCl₂, 10 mM KCl, 0.05 % Nonidet P-40, 0.5mM DTT, 5 mM Sodium butyrate) without any protease inhibitors. After centrifugation, nuclear pellet was re-suspended into Buffer A and incubated at 32 °C for 30 min without or with protease inhibitors, such as AEBSF, Aprotinin and E64. Nuclear pellet was collected again by centrifugation at 4 °C and lysed by 1X SDS sample buffer. Western blotting analysis with general H3 antibody was used to confirm cleaved H3 amount compared with full-length H3.

Purification of recombinant H3 protein and its mono-nucleosome

For native H3 purification, histone H3.3 with C-terminal 6x-His tag was overexpressed in BL21 *E.coli* after IPTG induction and purified as substrate by using TALON metal affinity resin. For H3 nucleosome purification, plasmid containing H3.3 ORF with C-terminal 3x-Flag tag was transfected into HEK293 cells to overexpress ectopic H3.3 for 2 days. These cells were lysed by Buffer A to obtain nuclei that were continually treated with micrococcal nuclease (MNase). After MNase treatment, nucleosomes were extracted by high salt buffer (10 mM Tris-HCl pH 7.4, 2 mM MgCl₂, 2 mM EDTA, 600 mM NaCl, 10 mM sodium butyrate and protease inhibitor cocktail), and mono-nucleosomes were collected by using sucrose gradient. Mono-nucleosomes assembled by ectopic H3.3-3xFlag were isolated by anti-FLAG[®] M2 magnetic beads (Sigma, M8823).

***In vitro* H3 cleavage assay**

Nuclear extract from elongating/condensing spermatids was prepared as described (Duncan et al. 2008). H3.3-6xHis purified from *E.coli* and H3.3-3xFlag mono-nucleosome isolated

from HEK293 cells were incubated with spermatid nuclear extract at 32 °C for 15 to 120 min, and then reactions stopped by adding SDS sample buffer. Protein level of cleaved H3 was analyzed by Western blotting with His-tag and Flag-tag antibodies.

***In vitro* mouse spermatid culture**

Elongating/condensing spermatids were isolated by centrifugal elutriation, incubated into Dulbecco's Modified Eagle's Medium (DMEM) including 10% fetal bovine serum, L-glutamine, and penicillin-streptomycin, and treated with 10 μM proteasome inhibitor, MG132 at 32 °C for 24 hours. Then spermatids were lysed by 1x SDS sample buffer and boiled at 95 °C. Proteins were analyzed by Western blotting.

***In vitro* H3 acetylation assay**

H3.3-6xHis purified from *E.coli* was incubated with recombinant p300 catalytic domain protein (Active motif, 31205) and 50 μM Acetyl-CoA in protease buffer (10 mM HEPES pH7.5, 10 mM KCl, 1.5 mM MgCl₂, 0.34M sucrose, 10% glycerol, 0.1 mM DTT) for 37 °C for 1 hour. Acetylation level on H3 was verified by Western blotting with H3Ac and H3K27Ac antibodies.

Mass spectrometry for mapping cleavage sites on H3

Full-length H3 and cleaved H3 bands from *in vitro* H3 cleavage assay were excised from coomassie-stained SDS-PAGE, propionylated by using propionic anhydride and subsequently digested by trypsin for overnight. Peptides were separated and analyzed by nano-HPLC (Agilent 1100 nanoLC system, Agilent Technologies) coupled to an LTQ Orbitrap Velos hybrid mass spectrometer (Thermo Scientific).

Protease enrichment and identification

Nuclear extract of mouse spermatids was prepared as described above, injected into CaptoQ column or benzamidine column and fractionated with 40 mM to 1 M KCl gradient or 50 mM glycine pH3. Each fraction was then analyzed by *in vitro* H3 cleavage assay as described above and subjected to mass spectrometry analysis

Native chromatin precipitation (N-ChIP)

Native ChIP was performed according to a previous study (Hisano et al. 2013) with some modifications. Briefly, condensing spermatid and elongating spermatids were treated 50 mM DTT in PBS with 5 mM sodium butyrate for 1 and 0.5 hours before cell lysis, except round spermatid. Cells were lysed into Buffer 1 with detergents (0.3 M Sucrose, 15 mM Tris-

HCl pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM MgCl₂, 0.1 mM EGTA, 0.5 mM DTT and 0.5 % NP-40 and 1 % DOC) and then treated with proper MNase to digest chromatin to mononucleosomes. This chromatin solution from spermatids or sperm mixed with H3K27Ac antibody at 4 °C for overnight incubation, and Dynabeads® Protein G was added in to pull down H3K27Ac nucleosomes. After elution and protease K digestion, immunoprecipitated DNA were resolved by 5% polyacrylamide electrophoresis, and 150-bp DNA was purified from gel for further library preparation.

Sequence data processing and analysis

The UCSC genome assembly mm10 was used as a reference genome. Transcript annotations are based on the UCSC knownGene database and were obtained from the Bioconductor package TxDb.Mmusculus.UCSC.mm10.knownGene. Genomic coordinates of CpG islands were obtained from the UCSC genome annotation database (<http://hgdownload.soe.ucsc.edu/goldenPath/mm10/database/>; file name: cpGIslandExt.txt.gz). Datasets of sperm MNase-seq, H3K27me3 ChIP-seq in round spermatids and H3K4me3 ChIP-seq in round spermatids and sperm were obtained from (Erkek et al. 2013). These datasets are available under GEO accession number GSE42629. Sequencing reads were aligned using Rbowtie (parameters '-m 1 --best --strata') against the mouse genome (UCSC version mm10), by using QuasR (Gaidatzis et al. 2015). Read counts were calculated in R, for transcription start site regions, i.e. 1kb or 2kb regions centered on transcription start sites (for genes with multiple transcription start sites annotated, the most 5' transcription start site was considered). In order to regularize small count values, a pseudo-count of 8 was added to read counts. In addition and for display in the scatter plots only, a small amount of noise was added to read counts to reduce overplotting. In supplementary figure 4 B-C, read counts were normalized by library size. CpG density was calculated as the ratio of observed over expected CpG dinucleotides in transcription start site regions.

References

- Ahmad K, Henikoff S. 2002. The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. *Molecular cell* **9**: 1191-1200.
- Barchi M, Geremia R, Magliozzi R, Bianchi E. 2009. Isolation and analyses of enriched populations of male mouse germ cells by sedimentation velocity: the centrifugal elutriation. *Methods in molecular biology* **558**: 299-321.
- Becker PB, Workman JL. 2013. Nucleosome remodeling and epigenetics. *Cold Spring Harbor perspectives in biology* **5**.

- Biswas M, Voltz K, Smith JC, Langowski J. 2011. Role of histone tails in structural stability of the nucleosome. *PLoS computational biology* **7**: e1002279.
- Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB, Peters AH. 2010. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nature structural & molecular biology* **17**: 679-687.
- De Vries M, Ramos L, Housein Z, De Boer P. 2012. Chromatin remodelling initiation during human spermiogenesis. *Biology open* **1**: 446-457.
- Duarte LF, Young AR, Wang Z, Wu HA, Panda T, Kou Y, Kapoor A, Hasson D, Mills NR, Ma'ayan A et al. 2014. Histone H3.3 and its proteolytically processed form drive a cellular senescence programme. *Nature communications* **5**: 5210.
- Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD. 2008. Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. *Cell* **135**: 284-294.
- Engelsberg A, Hermosilla R, Karsten U, Schulein R, Dorken B, Rehm A. 2003. The Golgi protein RCAS1 controls cell surface expression of tumor-associated O-linked glycan antigens. *The Journal of biological chemistry* **278**: 22998-23007.
- Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, Schubeler D, van der Vlag J, Stadler MB, Peters AH. 2013. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nature structural & molecular biology* **20**: 868-875.
- Falk MM, Grigera PR, Bergmann IE, Zibert A, Multhaup G, Beck E. 1990. Foot-and-mouth disease virus protease 3C induces specific proteolytic cleavage of host cell histone H3. *Journal of virology* **64**: 748-756.
- Ferreira H, Somers J, Webster R, Flaus A, Owen-Hughes T. 2007. Histone tails and the H3 alphaN helix regulate nucleosome mobility and stability. *Molecular and cellular biology* **27**: 4037-4048.
- Fujihara Y, Satouh Y, Inoue N, Isotani A, Ikawa M, Okabe M. 2012. SPACA1-deficient male mice are infertile with abnormally shaped sperm heads reminiscent of globozoospermia. *Development* **139**: 3583-3589.
- Gaidatzis D, Lerch A, Hahne F, Stadler MB. 2015. QuasR: quantification and annotation of short reads in R. *Bioinformatics* **31**: 1130-1132.
- Gan H, Wen L, Liao S, Lin X, Ma T, Liu J, Song CX, Wang M, He C, Han C et al. 2013. Dynamics of 5-hydroxymethylcytosine during mouse spermatogenesis. *Nature communications* **4**: 1995.
- Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. 1987. Sequence-specific packaging of DNA in human sperm chromatin. *Science* **236**: 962-964.
- Getun IV, Wu ZK, Khalil AM, Bois PR. 2010. Nucleosome occupancy landscape and dynamics at mouse recombination hotspots. *EMBO reports* **11**: 555-560.
- Govin J, Escoffier E, Rousseaux S, Kuhn L, Ferro M, Thevenon J, Catena R, Davidson I, Garin J, Khochbin S et al. 2007. Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *The Journal of cell biology* **176**: 283-294.
- Grimes SR, Jr., Henderson N. 1984. Hyperacetylation of histone H4 in rat testis spermatids. *Experimental cell research* **152**: 91-97.
- Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. 2009. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**: 473-478.
- Hisano M, Erkek S, Dessus-Babus S, Ramos L, Stadler MB, Peters AH. 2013. Genome-wide chromatin analysis in mature mouse and human spermatozoa. *Nature protocols* **8**: 2449-2470.
- Iwasaki W, Miya Y, Horikoshi N, Osakabe A, Taguchi H, Tachiwana H, Shibata T, Kagawa W, Kurumizaka H. 2013. Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS open bio* **3**: 363-369.
- Kang-Decker N, Mantchev GT, Juneja SC, McNiven MA, van Deursen JM. 2001. Lack of acrosome formation in Hrb-deficient mice. *Science* **294**: 1531-1533.

- Khor B, Bredemeyer AL, Huang CY, Turnbull IR, Evans R, Maggi LB, Jr., White JM, Walker LM, Carnes K, Hess RA et al. 2006. Proteasome activator PA200 is required for normal spermatogenesis. *Molecular and cellular biology* **26**: 2999-3007.
- Lin YN, Roy A, Yan W, Burns KH, Matzuk MM. 2007. Loss of zona pellucida binding proteins in the acrosomal matrix disrupts acrosome biogenesis and sperm morphogenesis. *Molecular and cellular biology* **27**: 6794-6805.
- Lu LY, Wu J, Ye L, Gavrilina GB, Saunders TL, Yu X. 2010. RNF8-dependent histone modifications regulate nucleosome removal during spermatogenesis. *Developmental cell* **18**: 371-384.
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251-260.
- Mandal P, Verma N, Chauhan S, Tomar RS. 2013. Unexpected histone H3 tail-clipping activity of glutamate dehydrogenase. *The Journal of biological chemistry* **288**: 18743-18757.
- Musselman CA, Lalonde ME, Cote J, Kutateladze TG. 2012. Perceiving the epigenetic landscape through histone readers. *Nature structural & molecular biology* **19**: 1218-1227.
- Oliva R. 2006. Protamines and male infertility. *Human reproduction update* **12**: 417-435.
- Purohit JS, Tomar RS, Panigrahi AK, Pandey SM, Singh D, Chaturvedi MM. 2013. Chicken liver glutamate dehydrogenase (GDH) demonstrates a histone H3 specific protease (H3ase) activity in vitro. *Biochimie* **95**: 1999-2009.
- Qian MX, Pang Y, Liu CH, Haratake K, Du BY, Ji DY, Wang GF, Zhu QQ, Song W, Yu Y et al. 2013. Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* **153**: 1012-1024.
- Rousseaux S, Caron C, Govin J, Lestrat C, Faure AK, Khochbin S. 2005. Establishment of male-specific epigenetic information. *Gene* **345**: 139-153.
- Santos-Rosa H, Kirmizis A, Nelson C, Bartke T, Saksouk N, Cote J, Kouzarides T. 2009. Histone H3 tail clipping regulates gene expression. *Nature structural & molecular biology* **16**: 17-22.
- Sperling AS, Grunstein M. 2009. Histone H3 N-terminus regulates higher order structure of yeast heterochromatin. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 13153-13159.
- Tesar M, Marquardt O. 1990. Foot-and-mouth disease virus protease 3C inhibits cellular transcription and mediates cleavage of histone H3. *Virology* **174**: 364-374.
- Vossaert L, Meert P, Scheerlinck E, Glibert P, Van Roy N, Heindryckx B, De Sutter P, Dhaenens M, Deforce D. 2014. Identification of histone H3 clipping activity in human embryonic stem cells. *Stem cell research* **13**: 123-134.
- Xiao N, Kam C, Shen C, Jin W, Wang J, Lee KM, Jiang L, Xia J. 2009. PICK1 deficiency causes male infertility in mice by disrupting acrosome formation. *The Journal of clinical investigation* **119**: 802-812.
- Xue Y, Vashisht AA, Tan Y, Su T, Wohlschlegel JA. 2014. PRB1 is required for clipping of the histone H3 N terminal tail in *Saccharomyces cerevisiae*. *PLoS one* **9**: e90496.
- Yao R, Ito C, Natsume Y, Sugitani Y, Yamanaka H, Kuretake S, Yanagida K, Sato A, Toshimori K, Noda T. 2002. Lack of acrosome formation in mice lacking a Golgi protein, GOPC. *Proceedings of the National Academy of Sciences of the United States of America* **99**: 11211-11216.
- Zentner GE, Henikoff S. 2013. Regulation of nucleosome dynamics by histone modifications. *Nature structural & molecular biology* **20**: 259-266.
- Zhou Y, Ru Y, Wang C, Wang S, Zhou Z, Zhang Y. 2013. Tripeptidyl peptidase II regulates sperm function by modulating intracellular Ca²⁺ stores via the ryanodine receptor. *PLoS one* **8**: e66634.

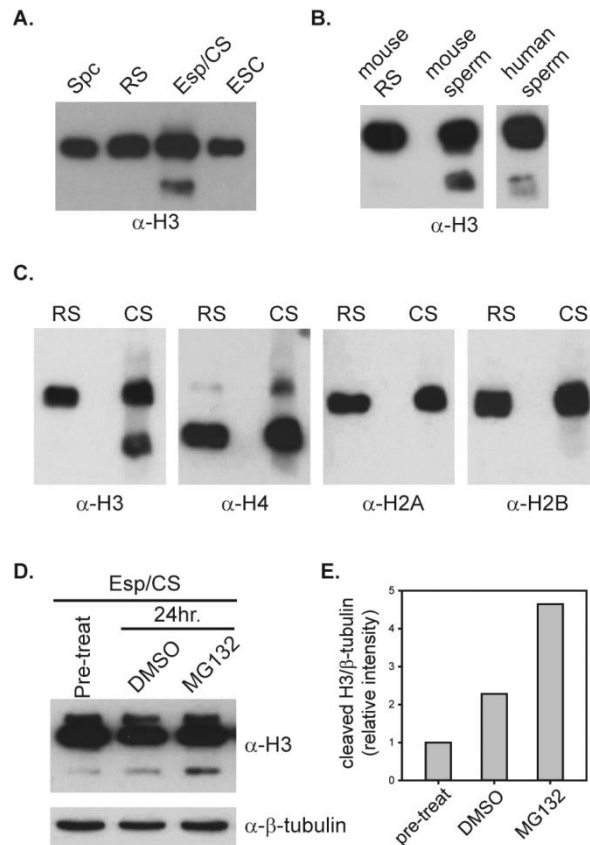


Figure 1. Histone H3 proteolysis during spermiogenesis. (A) Western blot analysis probed with an antibody against the C terminus of H3. Pachytene spermatocytes (Spc) were purified from mouse testes by FACS, and round spermatids (RS) and elongating/condensing spermatids (Esp/Cs) were isolated by centrifugal elutriator. Mouse embryonic stem cell line (ESC) was a control. (B) Western blot analysis of cleaved H3 in mouse and human sperm (C) The nuclear extract from round and condensing spermatids were analyzed by Western blotting with antibodies against C-terminal domain of histone H4, H2A and H2B. (D) The dynamics of cleaved H3 in elongating/condensing spermatids after proteasome inhibitor treatment. Elongating/condensing spermatids isolated from mouse testis were treated with 10 μ M MG132 at 32°C for 24 hours *in vitro*. (E) Analyses of band intensity from (D) are presented as the relative ratio of cleaved H3 to β -tubulin.

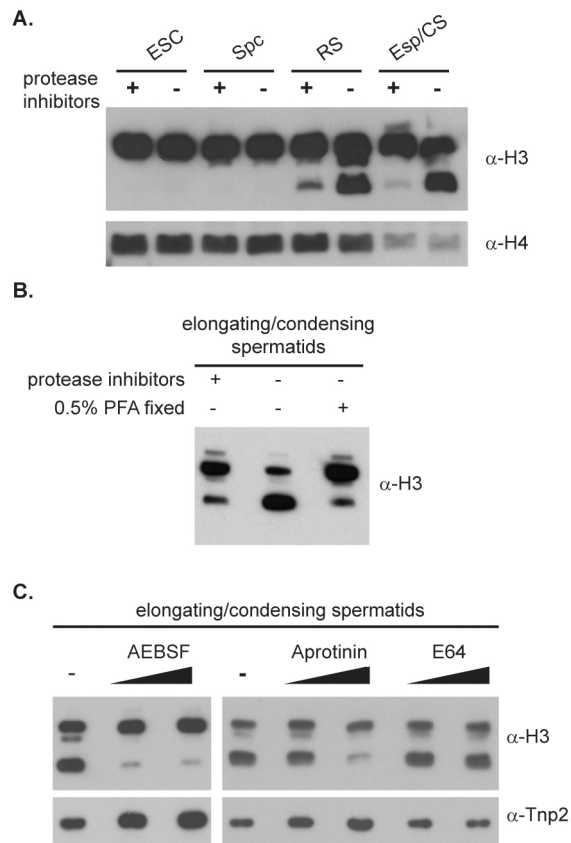


Figure 2. H3 protease in mouse spermatids is a serine protease. (A) H3 protease activity in the nuclei of mouse spermatids. The nuclei were isolated from ESC, pachytene spermatocytes (Spc), round spermatids (RS) and elongating/condensing spermatids (Esp/CS) and incubated at 32°C for 30 min with or without protease inhibitor cocktail. (B) H3 protease activity was inhibited by the treatment of 0.5 % paraformaldehyde (PFA) before nucleus isolation as described in (A). (C) H3 protease activity is inhibited by serine protease inhibitors. The nuclei of elongating/condensing spermatids were incubated with serine protease inhibitors, AEBSF and Aprotinin, and cysteine protease inhibitor, E64.

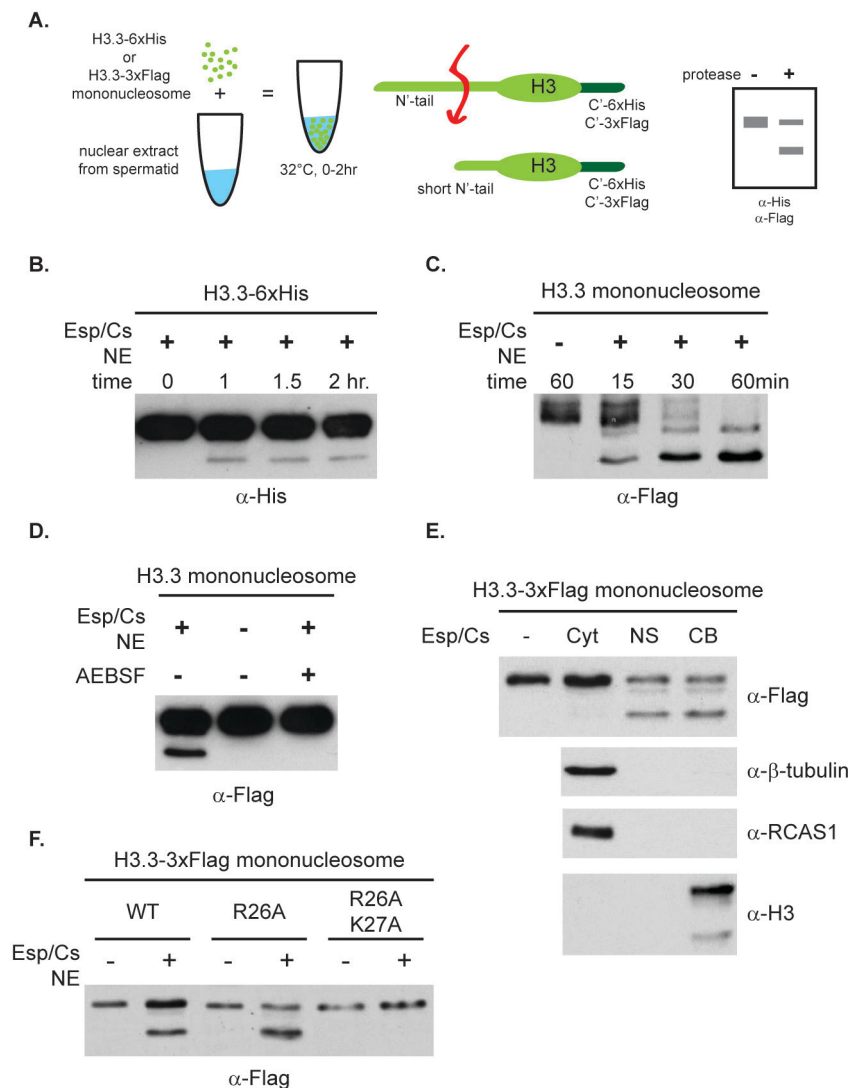


Figure 3. *In vitro* characterization of H3 protease activity from mouse spermatid. (A) Schematic of *in vitro* H3 cleavage assay. (B) *In vitro* H3 cleavage assay using native H3 as substrate. H3 with 6x-His tag was overexpressed and purified from *E.coli*. H3-6x-His was incubated with spermatid nuclear extract (NE) at 32°C for different incubation times. (C) *In vitro* H3 cleavage assay using nucleosome as substrate. The mononucleosomes containing H3 with C-terminal Flag tag were isolated by Flag resins from 293 cells and then incubated with spermatid nuclear extract. (D) The serine protease inhibitor AEBSF blocks H3 protease activity in *in vitro* H3 cleavage assay. (E) The nuclear-soluble fraction (NE) and chromatin-bound fraction (CS) contain H3 protease activity, not cytoplasmic fraction (Cyt). (F) The double mutation of H3 on Arg26 and Lys27 residues prevents H3 from cleavage in *in vitro* H3 cleavage assay.

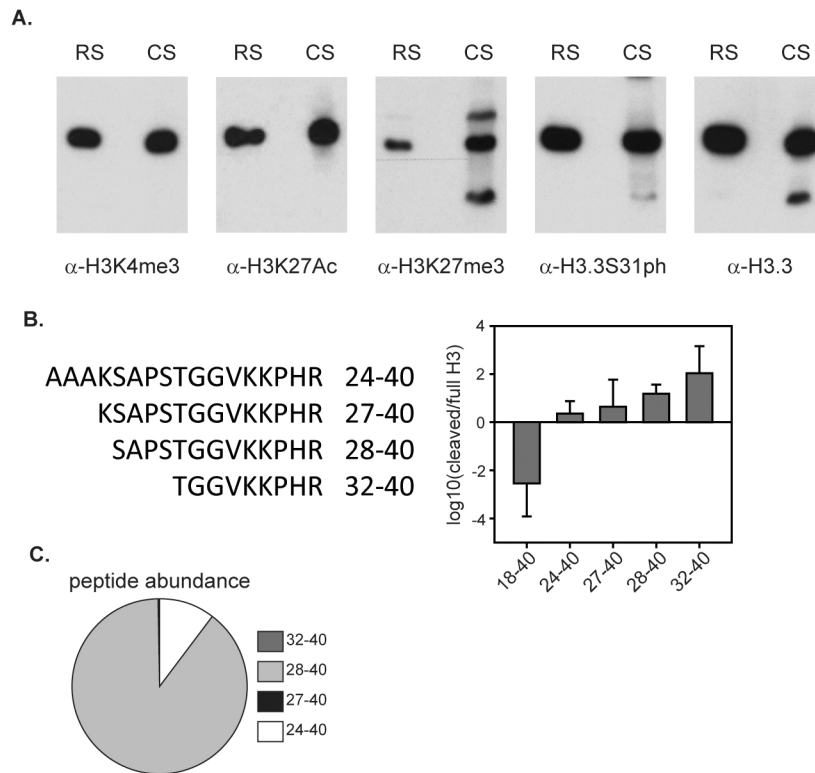


Figure 4. Identification of the cleavage site on H3 (A) Western blot analysis with antibodies against distinct modifications on H3 and specific H3 variants in round spermatid (RS) and condensing spermatid (CS) samples. (B) The comparison of truncated peptide abundance between cleaved H3 and full-length H3 by mass spectrometry analysis. Mononucleosomes assembled by H3 R26A mutation was used as substrate in *in vitro* H3 cleavage assay. Cleaved H3 and full-length H3 bands were excised from SDS-PAGE and digested by trypsin after lysine propionylation. The protein sequences of four truncated peptides identified were shown at left panel. (C) Pie chart demonstrating the percent abundance of four truncated peptides identified in cleaved H3 band.

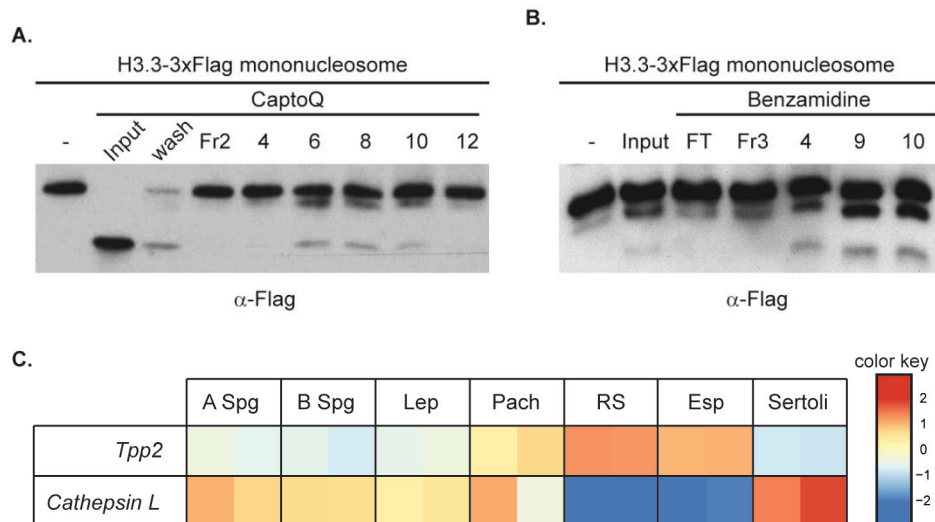


Figure 5. Fractionation of nuclear extract by CaptoQ and benzamidine columns to enrich H3 protease. (A) CaptoQ fractions from spermatid nuclear extract were assayed by *in vitro* H3 cleavage assay. (B) Benzamidine fractions from testis nuclear extract were assayed in *in vitro* H3 cleavage assay. (C) Heat map depicting the gene expression of *Tpp2* and *cathepsin L* in type A spermatogonia (A Spg), type B spermatogonia (B Spg), leptotene spermatocytes (Lep), pachytene spermatocytes (Pach), round spermatids (RS), elongating spermatids (Esp) and Sertoli cells based on published RNA-seq data (Gan et al. 2013). *Cathepsin L* is identified as H3 protease in mouse embryonic stem cells (Duncan et al. 2008), but its expression is relatively low in mouse spermatids. By contrast, *Tpp2* is highly transcribed in mouse spermatids.

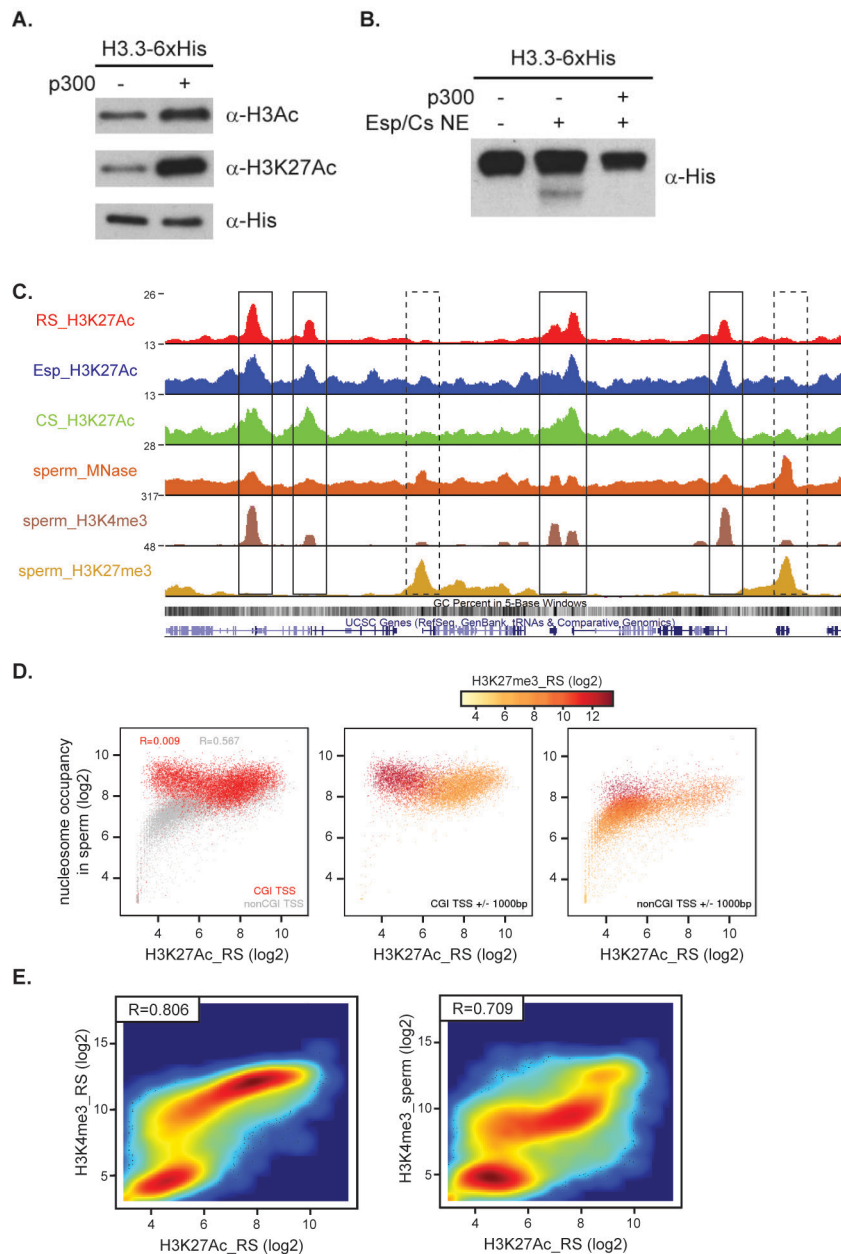


Figure 6. Acetylation on H3 inhibits H3 cleavage *in vitro* and relates to nucleosome retention in mature sperm (A) Native H3-6xHis purified from *E. coli* was acetylated by p300 catalytic domain *in vitro*. H3K27Ac and pan H3Ac antibodies were used for Western blotting to confirm the acetylation level on H3. (B) *In vitro* H3 cleavage assay using acetyl-H3-6xHis as substrate. (C) The snapshot from genome browser showing the genomic distribution of H3K27Ac in round, elongating and condensing spermatids compared to the nucleosome occupancy and the distribution of H3K4me3 and H3K27me3 in mature sperm. (D) Scatter plots showing the correlation between H3K27Ac enrichment in round spermatids and nucleosome occupancy in sperm at transcription start site (TSS) \pm 1 kb. Gene promoters (TSS \pm 1 kb) were classified into CpG Islands (CGI)-containing promoters (red dots in left panel) and non-CGI-containing promoters (grey dots in left panel). The scatter plot of CGI promoters and non-CGI promoter were separated and overlaid by H3K27me3 enrichment in round spermatids (middle and right panels). (E) Scatter plots showing the correlation between H3K27Ac enrichment in round spermatids and H3K4me3 enrichment in round spermatids (left panel) and sperm (right panel) at transcription start site (TSS) \pm 1 kb.

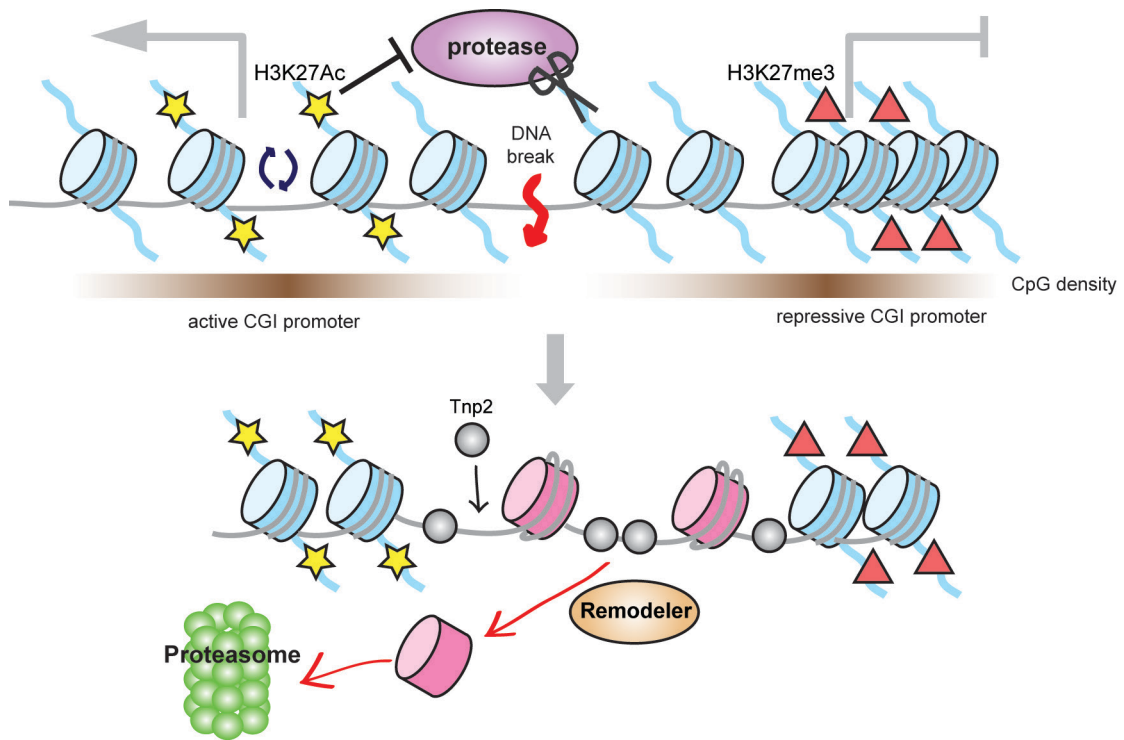
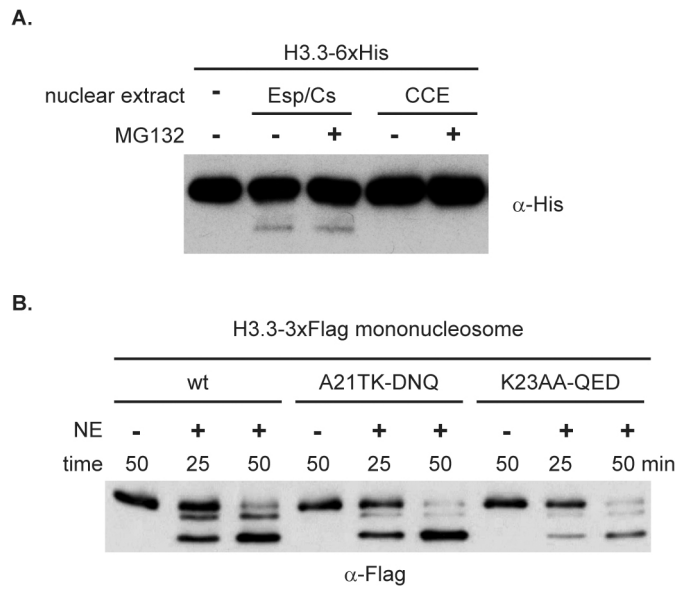


Figure 7. The model of H3 proteolytic cleavage in chromatin remodeling processing during spermatogenesis. N-terminal tail of H3 is cleaved at Arg26 and Lys27 by a putative serine protease activity, present in spermatid nuclei during spermiogenesis. Acetylation on H3 Lys27 (H3K27Ac) prevents H3 from proteolytic cleavage and is relative to nucleosome retention in sperm. Methylation on Lys27 of H3 may also have the same function as H3K27Ac to prevent cleavage. After cleavage, nucleosomes might be destabilized and easily evicted by unknown chromatin remodellers or chaperones. Finally cleaved H3 and evicted nucleosomes are degraded by proteasomes.

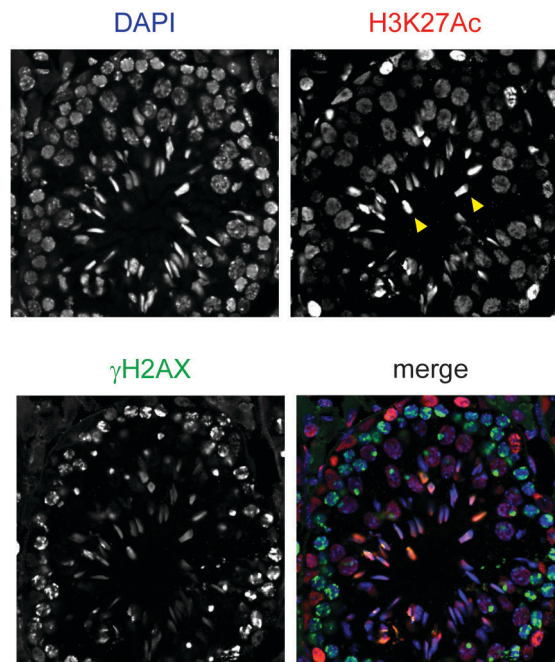


Supplementary figure 1. (A) MG132, a proteasome inhibitor, treatment in *in vitro* H3 cleavage assay. H3 protease activity in spermatids is not from proteasomes. (B) *In vitro* H3 cleavage assay using H3 with various mutation sites (mutations on residue 21-23 and residue 23-26) as substrates.

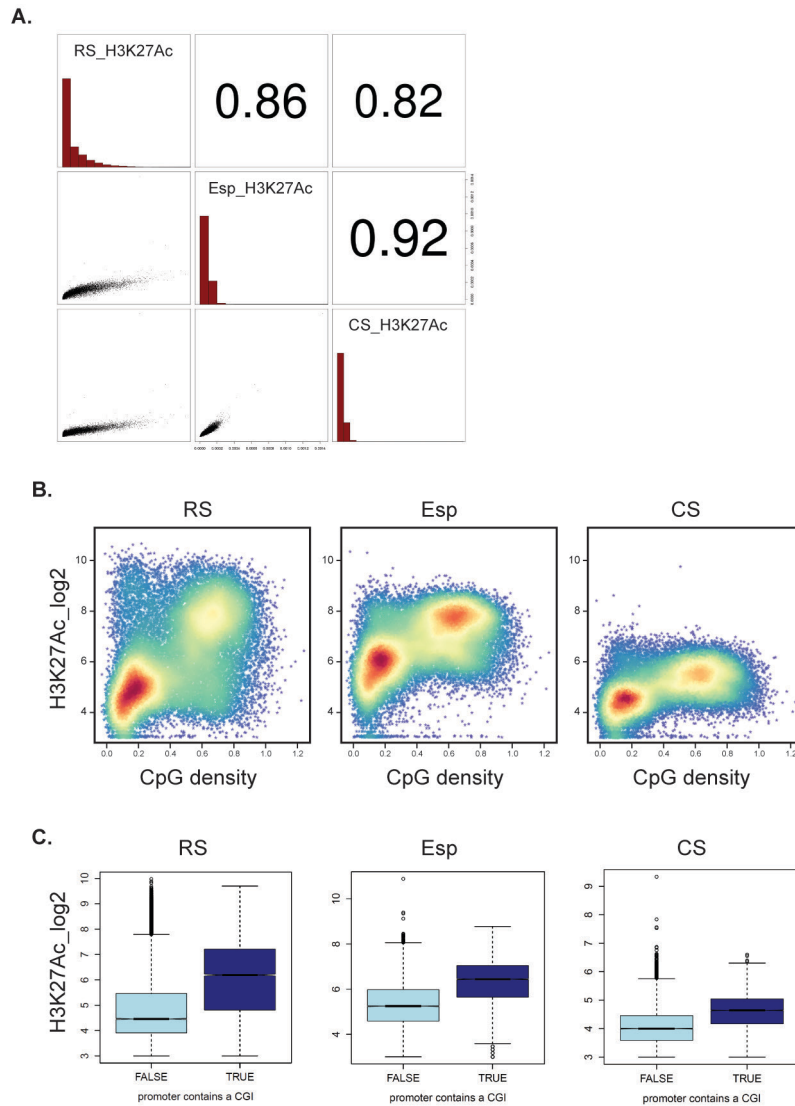
>sp|Q64514|TPP2_MOUSE Tripeptidyl-peptidase 2 OS=Mus musculus GN=Thpp2 PE=1 SV=3

MATAATEEPPFFHGLLPKKEGTGASSFLCRYPEYDGRGVLIAVLDTGVDPGAPGMQVTTDGGPK**IIDIIDTTGSGDVNTAT**
EVEPKDGEIIGLSGRVLIKIPANWTNPLGKYHIGIKNGYDFYPKALKERIQKERKEKIWDPIDHRVALAEACRKQEEFDIAN
NGSSQANKLIKEELQSOVELLNSFEKKYSDPGPVYDCLVWHDGETWRACVDSNENGDLKCAVLRNYKEAQEYSSFGTAE
MLNYSVNIYDDGNLLSIVTSGGAHGTHVASIAAGHFPEEPERNGVAPGAQILSIKIGDTRLSTMETGTGLIRAMIEVINH
KCDLVNYSYGEATHWPNSGRICEVINEAVWKHNTIYVSSAGNNGPCLSTVGCPCGGTTSSVIGVGAYVSPDMMVAEYSLRE
KLPANQYTWSSRGPSADGALGVSISAPGGAIASVPNWTLRGTQLMNGTSMSSPNACGGIALVLSGLKANNVDYTVHSVRR
ALENTAIKADNIEVFAQGHGIIQVDKAYDYLIQNTSFANRLGFTVTVGNNRGIYLRDPVQVAAPSDHGVGIEPVFPENTE
NSEKISFQLHLALTSNSSWQCPSHLELMNQCRHINIRVDPRGLREGLHYTEVCGYDIASPNAGPLFRVPITAVIAAKVN
ESSHYDLAFTDVHFKPGQIRRHFEVPEGATWAEVTVCSSESSEVSAKFVLHAVQLVKQRAYSHEFYKFCSLPEKGTLIE
AFPVLGGKAIIEFCIARWWASLSDVNIDYTI SFHGIVCTAPQLNIHASEGINRFDVQSSLYEDLAPCITLKSQWQTLRPV
NAKTRPLGSRDVLNNRQLYEMVLTYSFHQPKSGEVTSPCLLCELLYESEFDSQLWIIFDQNKRMGSGDAYPHQYSLK
LEKGDYTI RQLIRHEQISDLDRDKDLPFIVSHRLSNTLSLDIHENHSLALLGKKKSSSLTLPPKYNQPFVTSLPDDKIP
KGAGPGCYLAGSLTLSKTELGKKAGQSAKRQGFKKDVIPVHYLLIPPPTKIKNGSKDKEKDSEKEKDLKEEFTEALRD
LKIQWMTKLDSTDIYNELKETYPAYLPLYVARLHQLDAEKERMKRLNEIVDAANAVISHIDQALAVYIAMKTDPRPDAA
TIKNDMDKQKSTLIDALCRKGCALADHLLHTQPHDGAAGDAEAKEEGESTMESLSEYWETTKWTDLFDTKVLIFAYK
HALVNKMYGRGLKFATKLVEEKPTKENWKNCIQLMKLLGWTHCASFTENWLPIMYPPDYCVF

Supplementary figure 2. Mass spectrometry analysis of CaptoQ and benzamidine fractions identified the presence of tripeptidyl-peptidase II (Tpp2) specifically in the fractions exhibiting H3 cleavage activity. Tpp2 peptides identified in the CaptoQ fractions with protease activity are marked by underline. Tpp2 peptides identified in benzamidine fractions with protease activity are marked by red bold characters.



Supplementary figure 3. Immunofluorescence analyses of H3K27 acetylation in wildtype seminiferous tubules at spermatogenic stage VIII. H3K27Ac is detected at elongating spermatids (yellow arrows).



Supplementary figure 4. (A) The correlation of H3K27Ac enrichment at +/- 1 kb TSS between three stages of spermatids (round spermatid, RS; elongating spermatid, Esp; condensing spermatid, CS). (B) Scatter plots showing the correlation between H3K27Ac and CpG density at +/- 1 kb TSS from round to condensing spermatids. (C) The box plots showing the distribution of H3K27Ac enrichment in CpG island-contained TSS and non-CpG TSS at +/- 1 kb window.

Chapter 6: Discussion and Outlook

6.1 The dynamics of H3 variants during mouse spermatogenesis

In the first project, I first performed triton-acetic acid-urea gel to separate histone H3 proteins and thereby found that testis-specific H3 (H3t) is abundant in mouse testis and is highly expressed from spermatogonia stage. Strikingly, the protein expression of canonical H3.1 and H3.2 are lower in spermatogonia compared with H3t, and dramatically decreased to undetectable level during meiosis. By contrast, from spermatogonia to spermatid stage, H3t protein is still maintained at high level, and H3.3 protein level increases during meiosis. Therefore, H3t and H3.3 become two predominant H3 proteins in spermatocytes and spermatids. Previous studies have shown that H3.3 protein is required for correct spermatogenesis. The disruption of one of H3.3 genes, *H3f3b*, causes male infertility because of arrest of round spermatids or abnormal protamine incorporation (Yuen et al. 2014; Tang et al. 2015). However, whether H3t also plays a crucial role during spermatogenesis is still unclear. In order to dissect the biological role of H3t, we are generating H3t-knockout mice. In the future, it will be interesting to know the importance of H3t during male germ cell development and to investigate whether the lack of H3t is compensated by enhanced or prolonged of canonical H3 proteins.

There are two waves of transcription found during spermatogenesis. The first wave of transcription happens at spermatogonia stage before meiosis to express mitotic- and meiotic genes, and the second wave one happens at the late spermatocytes and round spermatids to express genes involved in spermiogenesis (Sassone-Corsi 2002; White-Cooper and Davidson 2011). I found that mouse H3t encoded by *GM12260* pseudogene is highly transcribed during the first wave of transcription in spermatogonia, and its transcript decreases dramatically during meiotic and post-meiotic stage. In addition, 3' UTR of *H3t* gene contains a stem loop sequence, which is generally present at 3' end of canonical histone mRNA and is required for replication-dependent histone expression. Therefore, these results suggest that H3t incorporation into nucleosomes might be regulated by a DNA replication-coupled process as described for canonical H3.

In somatic cells, H3.3 is localized at promoter regions and gene bodies of active genes and carries histone modifications associated with transcription activation, like H3K4me3 and

H3K27Ac. In contrast, canonical H3 is enriched in histone modifications known to correlate with transcription repression (McKittrick et al. 2004; Hake et al. 2006). In mouse spermatids, I found that H3K4me3 is also enriched on H3.3 than H3t, suggesting that H3t may have a similar “decoration” of histone marks to somatic canonical H3 and replace the function of canonical H3 in male germ cells. It will be of great interest in future studies to further investigate the detail of histone marks on H3t compared with H3.3 by mass spectrometry analysis and immunoblotting with specific antibodies against different histone modifications.

In addition, surprisingly, I observed that H3K4me3 reduces to an undetectable level either on H3.3 or on H3t in pachytene spermatocytes, but after meiosis round spermatids is able to restore H3K4me3 level. Previous studies showed that pachytene spermatocytes still have transcriptional activity occurring in synapsed meiotic chromosomes except unsynapsed sex chromosomes (Turner et al. 2005). It is possible that the number of expressed genes is lower in pachytene spermatocytes than in other cell stages, so the global H3K4me3 level decreases a lot in pachytene stages. In addition, this low level of H3K4me3 at the late stage of spermatocytes may promote a quick reconstitution of H3K4me3 on the promoters of spermiogenesis-relative genes to switch gene expression profile after meiosis. Interestingly, two H3K4me3 demethylase, KDM5A and KDM5D, have been reported to be highly expressed in spermatogonia and spermatocytes (Akimoto et al. 2008; Nishio et al. 2014). Therefore, deciphering potential developmental roles of the H3K4me3 demethylases in meiotic stage will be interesting to explore in future studies.

Furthermore, when the spermatid differentiates into mature sperm, global nucleosomes are removed and replaced by transition proteins and then by protamines. Only 1 % of nucleosomes are specifically retained at unmethylated CpG-rich promoters in mature mouse sperm according to our previous studies (Brykczynska et al. 2010; Erkek et al. 2013). In this project, I observed that most H3t is removed during histone-to-protamine exchange process, and 1 % residual nucleosomes in sperm contain a large amount of H3.3. This result represents that more H3t is evicted from chromatin than H3.3 during spermiogenesis. I propose two possible causes for the significant reduction of H3t. First, previous study indicated that human H3t can destabilize nucleosome structure *in vitro* (Tachiwana et al. 2010). Although mouse and human H3t have 3 amino acids difference, mouse H3t-containing nucleosome may also have less stability than canonical nucleosome to enhance

H3t eviction from chromatin. Further *in vitro* biochemical analyses are necessary to examine the influence of mouse H3t on nucleosome structure. Second, like canonical H3 in somatic cells, H3t may be widely distributed in chromatin of round spermatids, whereas H3.3-nucleosomes carrying histone marks of active transcription may be incorporated at promoter of active genes and undergo rapid turnover in round spermatids (Erkek et al. 2013; Kraushaar et al. 2013). During histone-to-protamine exchange process, transcriptional nucleosome turnover is abolished, which causes H3.3 accumulation around promoter regions, especially at unmethylated CpG-island promoters (Erkek et al. 2013). At the same time, unknown complexes may bind to unmethylated CpG-islands to suppress the eviction process of H3.3-nucleosomes, or histone marks specifically on H3.3 could prevent H3.3-containing nucleosomes from proteolytic cleavage and eviction. By contrast, H3t does not have specific localization and histone marks to prevent eviction (Figure 1). Thus most H3t is removed from chromatin than H3.3 during spermiogenesis. In addition, I detected canonical H3.1 and H3.2 in mature sperm. The possible explanation is that a very small portion of canonical H3 under detectable level is not replaced by H3t during meiosis and is stably kept in heterochromatin with repressive histone marks from spermatogonia to sperm. After H3t removal, the proportion of canonical H3 in residual nucleosomes in mature sperm is increased to detectable level.

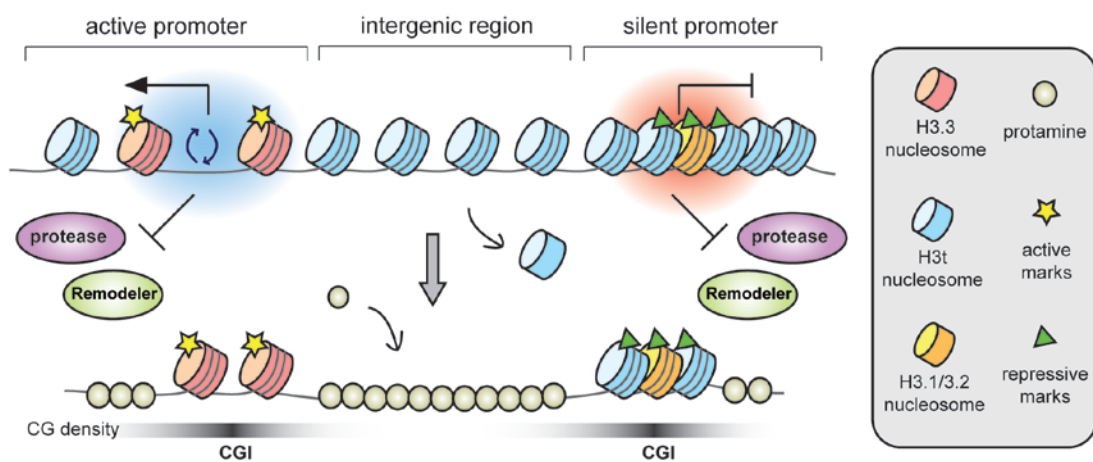


Figure 1. Model of histone H3t eviction and H3.3 retention during histone-to-protamine exchange process. H3t-nucleosomes are widely distributed in the chromatin in round spermatids, but H3.3-nucleosome is specific localized around active CpG island (CGI) promoters with active histone marks. A very small portion of canonical H3.1/H3.2 with repressive marks is kept in heterochromatin regions throughout spermatogenesis. The genome localization and histone marks could influence the process of nucleosome eviction.

Taken together, the findings in this part of study describe the characterization of mouse H3t and also reveal an unexpected dynamic of H3 variant during mouse spermatogenesis (Figure 2).

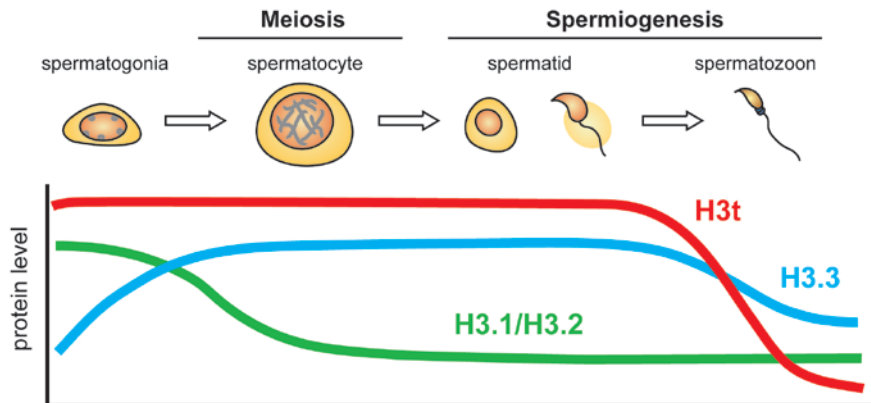


Figure 2. Summary of the dynamics of H3 variants during spermatogenesis

6.2 The proteolytic cleavage of histone H3 during spermiogenesis

Histone H3 is well-known to have a long N-terminal tail protruding from the nucleosome core particle. Many residues on H3 tail are added by various post-translational modifications to regulate nucleosome structure directly and also to recruit other proteins to change chromatin structure. Previous studies showed that H3 tail interacts with DNA at the edge of nucleosome to form higher order structure of chromatin (Biswas et al. 2011). The lack of H3 tail leads to destabilize nucleosome *in vitro* and increase accessibility within silent heterochromatin *in vivo* (Ferreira et al. 2007; Sperling and Grunstein 2009). Interestingly, the endogenous proteolytic cleavage of N-terminal tail of histone H3 has been reported in different cell types. During the differentiation of mouse embryonic stem cells (ESCs), H3 is cleaved by cathepsin L, a cysteine protease, and the primary H3 cleavage site is between amino acids 21 and 22 of the amino terminus (Duncan et al. 2008). However, it remains to elucidated the role of H3 cleavage by cathepsin L during in ESCs differentiation. Likewise, a recent study also showed that cathepsin L mediates H3 proteolytic cleavage during cellular senescence in human fibroblasts, and ectopic expression of N-terminus truncated H3 induces fibroblast senescence (Duarte et al. 2014). Thus, H3 proteolytic cleavage functions as a key regulator of cellular senescence. Moreover, the H3 N-terminus cleavage by other proteases, especially serine protease family, also occurs in yeast sporulation, chicken liver and human ESCs. The vacuolar proteinase B (PRB1) is identified as H3 protease in yeast

under condition of nutrient deprivation (Santos-Rosa et al. 2009; Xue et al. 2014). In chicken liver, glutamate dehydrogenase (GDH) is found to have the serine protease activity to clip H3 tail (Mandal et al. 2013; Purohit et al. 2013). In contrast to mouse ESCs, H3 in human ESCs is cleaved by unidentified serine protease, not by cathepsin L, during differentiation (Vossaert et al. 2014). Therefore, the proteolytic cleavage of H3 tail is universal phenomenon and believed to play a role in histone turnover.

In the second project, I found that N-terminal tail of histone H3 is subject to the proteolytic cleavage by a serine protease activity in the late-stage spermatids, and H3 proteolytic cleavage during spermiogenesis is conserved in human and mouse. We propose that the H3 cleavage process in spermatids is a potential mechanism to alter nucleosome structure and to contribute to global nucleosome removal during spermiogenesis. In order to further investigate the function of H3 cleavage during spermiogenesis, we are generating a mouse expressing “cleavage-resistant” H3 protein. Based on my finding, the major cleavage sites on H3 tail are localized around Arg26 and Lys27 residues. Mutations on both residues protect H3 from cleavage. Thus Arg26 and Lys27 are important for H3 protease recognition and/or activity in mouse spermatids. In addition, I identified that H3t is a predominant H3 protein in mouse spermatids and need to be largely removed during histone-to-protamine exchange. Therefore, we directly mutated both residues 26 and 27 to alanines on H3t in mouse zygotes by the CRISPR/Cas-mediated genome engineering. In the future, through the characterization of this mutant mouse line, we hope to identify whether the proteolytic cleavage of H3 contributes to histone removal during sperm maturation.

DNA damage response and histone H4 hyperacetylation are well-known marks coincident with histone displacement in elongating spermatids. The high level of H2A ubiquitination induced by DNA damage is observed in elongating spermatids (Chen et al. 1998; Baarends et al. 1999). These histone ubiquitination in elongating spermatids is believed to mediate chromatin reorganization. Based on previous studies, the lack of RNF8, an ubiquitin E3 ligase of H2A and H2B, causes the reduction of acetylated H4 on Lys16, abnormal histone accumulation in sperm and finally male infertility (Lu et al. 2010). Therefore RNF8-dependent histone ubiquitination is required for the replacement of histones by protamines during spermiogenesis. Furthermore, a recent study found that a testis-specific proteasome activator, PA200, recognizes acetyl-H4 and acetyl-H2B via its bromodomain-like region and

then mediates the acetylation-dependent degradation of histones during spermiogenesis (Qian et al. 2013). Therefore, H4 acetylation not only directly affects nucleosome structure but acts as a signal for histone degradation. Likewise, I found that treating elongating spermatids with proteasome inhibitor leads to the accumulation of cleaved H3. This finding suggests that cleaved H3 is degraded by proteasomes after proteolytic cleavage. But it is unclear whether H4 hyperacetylation and RNF8-mediated histone ubiquitination are associated with H3 proteolysis. In the future, it is interesting to examine the genome-side distribution of cleaved H3 in elongating spermatids and clarify whether cleaved H3 is enriched in nucleosomes containing hyperacetylated H4 and ubiquitinated H2A/H2B.

In addition to H3 cleavage that we identified in the late stage spermatids, the proteolytic cleavage is found to involve in protamine maturation during spermiogenesis. In human and mouse, two types of protamines are identified: the protamine 1 and the family of protamine 2 proteins. There is a difference between the syntheses of the two protamines. Protamine 1 (P1) is synthesized directly as mature protein, whereas protamine 2 (P2) is synthesized as a long precursor (pre-P2). After binding to DNA, pre-P2 is proteolytically cleaved at its N-terminus to form mature P2 in the late stage spermatids (Carre-Eusebe et al. 1991; Chauviere et al. 1992). The normal P1/P2 ratio is roughly 1 in fertile human sperm (Belokopytova et al. 1993). But, sperm from many infertile patients has P1/P2 ratios higher than 1, and pre-P2 is abnormally accumulated in them (Belokopytova et al. 1993; de Yebra et al. 1993; de Yebra et al. 1998). Therefore the proteolytic processing of protamine 2 precursor is thought to influence male fertility (de Mateo et al. 2011). Although the specific proteases and mechanisms mediating the proteolysis of pre-P2 have not yet been identified, these findings combining my finding suggest that chromatin of the late stage spermatid is exposed in protease-rich environment, and proteases may function in the regulation of chromatin reorganization during spermiogenesis through cleaving histones and protamines.

References

- Akimoto C, Kitagawa H, Matsumoto T, Kato S. 2008. Spermatogenesis-specific association of SMCY and MSH5. *Genes to cells : devoted to molecular & cellular mechanisms* **13**: 623-633.
- Alabert C, Groth A. 2012. Chromatin replication and epigenome maintenance. *Nature reviews Molecular cell biology* **13**: 153-167.
- Angelov D, Vitolo JM, Mutskov V, Dimitrov S, Hayes JJ. 2001. Preferential interaction of the core histone tail domains with linker DNA. *Proceedings of the National Academy of Sciences of the United States of America* **98**: 6599-6604.
- Arnaudo AM, Garcia BA. 2013. Proteomic characterization of novel histone post-translational modifications. *Epigenetics & chromatin* **6**: 24.
- Azura V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merckenschlager M et al. 2006. Chromatin signatures of pluripotent cell lines. *Nature cell biology* **8**: 532-538.
- Baarends WM, Hoogerbrugge JW, Roest HP, Ooms M, Vreeburg J, Hoeijmakers JH, Grootegoed JA. 1999. Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. *Developmental biology* **207**: 322-333.
- Banaszynski LA, Wen D, Dewell S, Whitcomb SJ, Lin M, Diaz N, Elsasser SJ, Chapgier A, Goldberg AD, Canaani E et al. 2013. Hira-dependent histone H3.3 deposition facilitates PRC2 recruitment at developmental loci in ES cells. *Cell* **155**: 107-120.
- Banerjee T, Chakravarti D. 2011. A peek into the complex realm of histone phosphorylation. *Molecular and cellular biology* **31**: 4858-4873.
- Bannister AJ, Kouzarides T. 2011. Regulation of chromatin by histone modifications. *Cell research* **21**: 381-395.
- Battle DJ, Doudna JA. 2001. The stem-loop binding protein forms a highly stable and specific complex with the 3' stem-loop of histone mRNAs. *Rna* **7**: 123-132.
- Baubec T, Schubeler D. 2014. Genomic patterns and context specific interpretation of DNA methylation. *Current opinion in genetics & development* **25**: 85-92.
- Bedford MT, Clarke SG. 2009. Protein arginine methylation in mammals: who, what, and why. *Molecular cell* **33**: 1-13.
- Belokopytova IA, Kostyleva EI, Tomilin AN, Vorob'ev VI. 1993. Human male infertility may be due to a decrease of the protamine P2 content in sperm chromatin. *Molecular reproduction and development* **34**: 53-57.
- Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR et al. 2005. Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120**: 169-181.
- Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K et al. 2006. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125**: 315-326.
- Bird AW, Yu DY, Pray-Grant MG, Qiu Q, Harmon KE, Megee PC, Grant PA, Smith MM, Christman MF. 2002. Acetylation of histone H4 by Esa1 is required for DNA double-strand break repair. *Nature* **419**: 411-415.
- Biswas M, Voltz K, Smith JC, Langowski J. 2011. Role of histone tails in structural stability of the nucleosome. *PLoS computational biology* **7**: e1002279.
- Black BE, Foltz DR, Chakravarthy S, Luger K, Woods VL, Jr., Cleveland DW. 2004. Structural determinants for generating centromeric chromatin. *Nature* **430**: 578-582.
- Blanco-Rodriguez J, Martinez-Garcia C. 1999. Apoptosis is physiologically restricted to a specialized cytoplasmic compartment in rat spermatids. *Biology of reproduction* **61**: 1541-1547.
- Bogdanovic O, Veenstra GJ. 2009. DNA methylation and methyl-CpG binding proteins: developmental requirements and function. *Chromosoma* **118**: 549-565.

- Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE. 2007. UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317**: 1760-1764.
- Boyarchuk E, Montes de Oca R, Almouzni G. 2011. Cell cycle dynamics of histone variants at the centromere, a model for chromosomal landmarks. *Current opinion in cell biology* **23**: 266-276.
- Breucker H, Schafer E, Holstein AF. 1985. Morphogenesis and fate of the residual body in human spermiogenesis. *Cell and tissue research* **240**: 303-309.
- Brykczynska U, Hisano M, Erkek S, Ramos L, Oakeley EJ, Roloff TC, Beisel C, Schubeler D, Stadler MB, Peters AH. 2010. Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa. *Nature structural & molecular biology* **17**: 679-687.
- Bui M, Dimitriadis EK, Hoischen C, An E, Quenet D, Giebe S, Nita-Lazar A, Diekmann S, Dalal Y. 2012. Cell-cycle-dependent structural transitions in the human CENP-A nucleosome in vivo. *Cell* **150**: 317-326.
- Burgoyne PS. 1982. Genetic homology and crossing over in the X and Y chromosomes of Mammals. *Human genetics* **61**: 85-90.
- Burma S, Chen BP, Murphy M, Kurimasa A, Chen DJ. 2001. ATM phosphorylates histone H2AX in response to DNA double-strand breaks. *The Journal of biological chemistry* **276**: 42462-42467.
- Bush KM, Yuen BT, Barrilleaux BL, Riggs JW, O'Geen H, Cotterman RF, Knoepfler PS. 2013. Endogenous mammalian histone H3.3 exhibits chromatin-related functions during development. *Epigenetics & chromatin* **6**: 7.
- Cao J, Yan Q. 2012. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Frontiers in oncology* **2**: 26.
- Cao R, Tsukada Y, Zhang Y. 2005. Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. *Molecular cell* **20**: 845-854.
- Carre-Eusebe D, Lederer F, Le KH, Elsevier SM. 1991. Processing of the precursor of protamine P2 in mouse. Peptide mapping and N-terminal sequence analysis of intermediates. *The Biochemical journal* **277 (Pt 1)**: 39-45.
- Carrell DT, Emery BR, Hammoud S. 2007. Altered protamine expression and diminished spermatogenesis: what is the link? *Human reproduction update* **13**: 313-327.
- Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP et al. 2005. Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123**: 581-592.
- Chauviere M, Martinage A, Debarle M, Sautiere P, Chevaillier P. 1992. Molecular characterization of six intermediate proteins in the processing of mouse protamine P2 precursor. *European journal of biochemistry / FEBS* **204**: 759-765.
- Chemes H. 1986. The phagocytic function of Sertoli cells: a morphological, biochemical, and endocrinological study of lysosomes and acid phosphatase localization in the rat testis. *Endocrinology* **119**: 1673-1681.
- Chen HY, Sun JM, Zhang Y, Davie JR, Meistrich ML. 1998. Ubiquitination of histone H3 in elongating spermatids of rat testes. *The Journal of biological chemistry* **273**: 13165-13169.
- Chen JL, Longo FJ. 1996. Expression and localization of DNA topoisomerase II during rat spermatogenesis. *Molecular reproduction and development* **45**: 61-71.
- Cho C, Jung-Ha H, Willis WD, Goulding EH, Stein P, Xu Z, Schultz RM, Hecht NB, Eddy EM. 2003. Protamine 2 deficiency leads to sperm DNA damage and embryo death in mice. *Biology of reproduction* **69**: 211-217.
- Cho C, Willis WD, Goulding EH, Jung-Ha H, Choi YC, Hecht NB, Eddy EM. 2001. Haploinsufficiency of protamine-1 or -2 causes infertility in mice. *Nature genetics* **28**: 82-86.
- Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF. 1997. Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* **277**: 1996-2000.

- Dai L, Peng C, Montellier E, Lu Z, Chen Y, Ishii H, Debernardi A, Buchou T, Rousseaux S, Jin F et al. 2014. Lysine 2-hydroxyisobutyrylation is a widely distributed active histone mark. *Nature chemical biology* **10**: 365-370.
- Dalal Y, Wang H, Lindsay S, Henikoff S. 2007. Tetrameric structure of centromeric nucleosomes in interphase *Drosophila* cells. *PLoS biology* **5**: e218.
- Davey CA, Sargent DF, Luger K, Maeder AW, Richmond TJ. 2002. Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *Journal of molecular biology* **319**: 1097-1113.
- Davis TL, Trasler JM, Moss SB, Yang GJ, Bartolomei MS. 1999. Acquisition of the H19 methylation imprint occurs differentially on the parental alleles during spermatogenesis. *Genomics* **58**: 18-28.
- Davis TL, Yang GJ, McCarrey JR, Bartolomei MS. 2000. The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Human molecular genetics* **9**: 2885-2894.
- de Mateo S, Ramos L, de Boer P, Meistrich M, Oliva R. 2011. Protamine 2 precursors and processing. *Protein and peptide letters* **18**: 778-785.
- de Yebra L, Balleca JL, Vanrell JA, Bassas L, Oliva R. 1993. Complete selective absence of protamine P2 in humans. *The Journal of biological chemistry* **268**: 10553-10557.
- de Yebra L, Balleca JL, Vanrell JA, Corzett M, Balhorn R, Oliva R. 1998. Detection of P2 precursors in the sperm cells of infertile patients who have reduced protamine P2 levels. *Fertility and sterility* **69**: 755-759.
- Denissov S, Hofemeister H, Marks H, Kranz A, Ciotta G, Singh S, Anastassiadis K, Stunnenberg HG, Stewart AF. 2014. Mll2 is required for H3K4 trimethylation on bivalent promoters in embryonic stem cells, whereas Mll1 is redundant. *Development* **141**: 526-537.
- Dhar S, Thota A, Rao MR. 2012. Insights into role of bromodomain, testis-specific (Brdt) in acetylated histone H4-dependent chromatin remodeling in mammalian spermiogenesis. *The Journal of biological chemistry* **287**: 6387-6405.
- Dhayalan A, Tamas R, Bock I, Tattermusch A, Dimitrova E, Kudithipudi S, Ragozin S, Jeltsch A. 2011. The ATRX-ADD domain binds to H3 tail peptides and reads the combined methylation state of K4 and K9. *Human molecular genetics* **20**: 2195-2203.
- Dorigo B, Schalch T, Kulangara A, Duda S, Schroeder RR, Richmond TJ. 2004. Nucleosome arrays reveal the two-start organization of the chromatin fiber. *Science* **306**: 1571-1573.
- Drabent B, Bode C, Bramlage B, Doenecke D. 1996. Expression of the mouse testicular histone gene H1t during spermatogenesis. *Histochemistry and cell biology* **106**: 247-251.
- Drabent B, Saftig P, Bode C, Doenecke D. 2000. Spermatogenesis proceeds normally in mice without linker histone H1t. *Histochemistry and cell biology* **113**: 433-442.
- Duarte LF, Young AR, Wang Z, Wu HA, Panda T, Kou Y, Kapoor A, Hasson D, Mills NR, Ma'ayan A et al. 2014. Histone H3.3 and its proteolytically processed form drive a cellular senescence programme. *Nature communications* **5**: 5210.
- Duncan EM, Muratore-Schroeder TL, Cook RG, Garcia BA, Shabanowitz J, Hunt DF, Allis CD. 2008. Cathepsin L proteolytically processes histone H3 during mouse embryonic stem cell differentiation. *Cell* **135**: 284-294.
- Dunleavy EM, Roche D, Tagami H, Lacoste N, Ray-Gallet D, Nakamura Y, Daigo Y, Nakatani Y, Almouzni-Pettinotti G. 2009. HJURP is a cell-cycle-dependent maintenance and deposition factor of CENP-A at centromeres. *Cell* **137**: 485-497.
- Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C. 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic acids research* **10**: 2709-2721.
- Elsasser SJ, Noh KM, Diaz N, Allis CD, Banaszynski LA. 2015. Histone H3.3 is required for endogenous retroviral element silencing in embryonic stem cells. *Nature*.

- Erkek S, Hisano M, Liang CY, Gill M, Murr R, Dieker J, Schubeler D, van der Vlag J, Stadler MB, Peters AH. 2013. Molecular determinants of nucleosome retention at CpG-rich sequences in mouse spermatozoa. *Nature structural & molecular biology* **20**: 868-875.
- Feng Q, Wang H, Ng HH, Erdjument-Bromage H, Tempst P, Struhl K, Zhang Y. 2002. Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain. *Current biology* : **CB 12**: 1052-1058.
- Fenic I, Sonnack V, Failing K, Bergmann M, Steger K. 2004. In vivo effects of histone-deacetylase inhibitor trichostatin-A on murine spermatogenesis. *Journal of andrology* **25**: 811-818.
- Ferreira H, Somers J, Webster R, Flaus A, Owen-Hughes T. 2007. Histone tails and the H3 alphaN helix regulate nucleosome mobility and stability. *Molecular and cellular biology* **27**: 4037-4048.
- Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD. 2005. Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438**: 1116-1122.
- Foltz DR, Jansen LE, Bailey AO, Yates JR, 3rd, Bassett EA, Wood S, Black BE, Cleveland DW. 2009. Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* **137**: 472-484.
- Furuyama T, Henikoff S. 2009. Centromeric nucleosomes induce positive DNA supercoils. *Cell* **138**: 104-113.
- Garcia BA, Thomas CE, Kelleher NL, Mizzen CA. 2008. Tissue-specific expression and post-translational modification of histone H3 variants. *Journal of proteome research* **7**: 4225-4236.
- Gatewood JM, Cook GR, Balhorn R, Bradbury EM, Schmid CW. 1987. Sequence-specific packaging of DNA in human sperm chromatin. *Science* **236**: 962-964.
- Gautier T, Abbott DW, Molla A, Verdel A, Ausio J, Dimitrov S. 2004. Histone variant H2ABbd confers lower stability to the nucleosome. *EMBO reports* **5**: 715-720.
- Goldberg AD, Allis CD, Bernstein E. 2007. Epigenetics: a landscape takes shape. *Cell* **128**: 635-638.
- Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, Dewell S, Law M, Guo X, Li X et al. 2010. Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* **140**: 678-691.
- Govin J, Escoffier E, Rousseaux S, Kuhn L, Ferro M, Thevenon J, Catena R, Davidson I, Garin J, Khochbin S et al. 2007. Pericentric heterochromatin reprogramming by new histone variants during mouse spermiogenesis. *The Journal of cell biology* **176**: 283-294.
- Greenbaum MP, Iwamori T, Buchold GM, Matzuk MM. 2011. Germ cell intercellular bridges. *Cold Spring Harbor perspectives in biology* **3**: a005850.
- Grimes SR, Jr., Henderson N. 1984. Hyperacetylation of histone H4 in rat testis spermatids. *Experimental cell research* **152**: 91-97.
- Gu TP, Guo F, Yang H, Wu HP, Xu GF, Liu W, Xie ZG, Shi L, He X, Jin SG et al. 2011. The role of Tet3 DNA dioxygenase in epigenetic reprogramming by oocytes. *Nature* **477**: 606-610.
- Guccione E, Bassi C, Casadio F, Martinato F, Cesaroni M, Schuchlantz H, Luscher B, Amati B. 2007. Methylation of histone H3R2 by PRMT6 and H3K4 by an MLL complex are mutually exclusive. *Nature* **449**: 933-937.
- Guo F, Li X, Liang D, Li T, Zhu P, Guo H, Wu X, Wen L, Gu TP, Hu B et al. 2014. Active and passive demethylation of male and female pronuclear DNA in the Mammalian zygote. *Cell stem cell* **15**: 447-458.
- Hackett JA, Reddington JP, Nestor CE, Dunican DS, Branco MR, Reichmann J, Reik W, Surani MA, Adams IR, Meehan RR. 2012. Promoter DNA methylation couples genome-defence mechanisms to epigenetic reprogramming in the mouse germline. *Development* **139**: 3623-3632.
- Hackett JA, Sengupta R, Zyllicz JJ, Murakami K, Lee C, Down TA, Surani MA. 2013. Germline DNA demethylation dynamics and imprint erasure through 5-hydroxymethylcytosine. *Science* **339**: 448-452.

- Hajkova P, Ancelin K, Waldmann T, Lacoste N, Lange UC, Cesari F, Lee C, Almouzni G, Schneider R, Surani MA. 2008. Chromatin dynamics during epigenetic reprogramming in the mouse germ line. *Nature* **452**: 877-881.
- Hake SB, Garcia BA, Duncan EM, Kauer M, Dellaire G, Shabanowitz J, Bazett-Jones DP, Allis CD, Hunt DF. 2006. Expression patterns and post-translational modifications associated with mammalian histone H3 variants. *The Journal of biological chemistry* **281**: 559-568.
- Hammoud SS, Nix DA, Zhang H, Purwar J, Carrell DT, Cairns BR. 2009. Distinctive chromatin in human sperm packages genes for embryo development. *Nature* **460**: 473-478.
- Handel MA. 2004. The XY body: a specialized meiotic chromatin domain. *Experimental cell research* **296**: 57-63.
- Handel MA, Schimenti JC. 2010. Genetics of mammalian meiosis: regulation, dynamics and impact on fertility. *Nature reviews Genetics* **11**: 124-136.
- He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L et al. 2011. Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333**: 1303-1307.
- Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA et al. 2007. Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nature genetics* **39**: 311-318.
- Henikoff S, Smith MM. 2015. Histone variants and epigenetics. *Cold Spring Harbor perspectives in biology* **7**: a019364.
- Hirota T, Lipp JJ, Toh BH, Peters JM. 2005. Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438**: 1176-1180.
- Horton JR, Upadhyay AK, Qi HH, Zhang X, Shi Y, Cheng X. 2010. Enzymatic and structural insights for substrate specificity of a family of jumonji histone lysine demethylases. *Nature structural & molecular biology* **17**: 38-43.
- Howlett SK, Reik W. 1991. Methylation levels of maternal and paternal genomes during preimplantation development. *Development* **113**: 119-127.
- Howman EV, Fowler KJ, Newson AJ, Redward S, MacDonald AC, Kalitsis P, Choo KH. 2000. Early disruption of centromeric chromatin organization in centromere protein A (Cenpa) null mice. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 1148-1153.
- Huh NE, Hwang IW, Lim K, You KH, Chae CB. 1991. Presence of a bi-directional S phase-specific transcription regulatory element in the promoter shared by testis-specific TH2A and TH2B histone genes. *Nucleic acids research* **19**: 93-98.
- Iacovoni JS, Caron P, Lassadi I, Nicolas E, Massip L, Trouche D, Legube G. 2010. High-resolution profiling of gammaH2AX around DNA double strand breaks in the mammalian genome. *The EMBO journal* **29**: 1446-1457.
- Inaba K. 2011. Sperm flagella: comparative and phylogenetic perspectives of protein components. *Molecular human reproduction* **17**: 524-538.
- Ishibashi T, Li A, Eirin-Lopez JM, Zhao M, Missiaen K, Abbott DW, Meistrich M, Hendzel MJ, Ausio J. 2010. H2A.Bbd: an X-chromosome-encoded histone involved in mammalian spermiogenesis. *Nucleic acids research* **38**: 1780-1789.
- Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y. 2011. Tet proteins can convert 5-methylcytosine to 5-formylcytosine and 5-carboxylcytosine. *Science* **333**: 1300-1303.
- Ivanova VS, Hatch CL, Bonner WM. 1994. Characterization of the human histone H2A.X gene. Comparison of its promoter with other H2A gene promoters. *The Journal of biological chemistry* **269**: 24189-24194.
- Iwasaki W, Miya Y, Horikoshi N, Osakabe A, Taguchi H, Tachiwana H, Shibata T, Kagawa W, Kurumizaka H. 2013. Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS open bio* **3**: 363-369.
- Iyer LM, Tahiliani M, Rao A, Aravind L. 2009. Prediction of novel families of enzymes involved in oxidative and other complex modifications of bases in nucleic acids. *Cell cycle* **8**: 1698-1710.

- Jabbari K, Bernardi G. 2004. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene* **333**: 143-149.
- Jiang X, Xu Y, Price BD. 2010. Acetylation of H2AX on lysine 36 plays a key role in the DNA double-strand break repair pathway. *FEBS letters* **584**: 2926-2930.
- Jin C, Zang C, Wei G, Cui K, Peng W, Zhao K, Felsenfeld G. 2009. H3.3/H2A.Z double variant-containing nucleosomes mark 'nucleosome-free regions' of active promoters and other regulatory regions. *Nature genetics* **41**: 941-945.
- Jungmichel S, Stucki M. 2010. MDC1: The art of keeping things in focus. *Chromosoma* **119**: 337-349.
- Kafri T, Ariel M, Brandeis M, Shemer R, Urven L, McCarrey J, Cedar H, Razin A. 1992. Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. *Genes & development* **6**: 705-714.
- Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H. 2004. Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429**: 900-903.
- Kato Y, Kaneda M, Hata K, Kumaki K, Hisano M, Kohara Y, Okano M, Li E, Nozaki M, Sasaki H. 2007. Role of the Dnmt3 family in de novo methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Human molecular genetics* **16**: 2272-2280.
- Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ et al. 2005. Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123**: 593-605.
- Khor B, Bredemeyer AL, Huang CY, Turnbull IR, Evans R, Maggi LB, Jr., White JM, Walker LM, Carnes K, Hess RA et al. 2006. Proteasome activator PA200 is required for normal spermatogenesis. *Molecular and cellular biology* **26**: 2999-3007.
- Kirmizis A, Santos-Rosa H, Penkett CJ, Singer MA, Vermeulen M, Mann M, Bahler J, Green RD, Kouzarides T. 2007. Arginine methylation at histone H3R2 controls deposition of H3K4 trimethylation. *Nature* **449**: 928-932.
- Kishikawa S, Murata T, Ugai H, Yamazaki T, Yokoyama KK. 2003. Control elements of Dnmt1 gene are regulated in cell-cycle dependent manner. *Nucleic acids research Supplement*: 307-308.
- Kooistra SM, Helin K. 2012. Molecular mechanisms and potential functions of histone demethylases. *Nature reviews Molecular cell biology* **13**: 297-311.
- Kraushaar DC, Jin W, Maunakea A, Abraham B, Ha M, Zhao K. 2013. Genome-wide incorporation dynamics reveal distinct categories of turnover for the histone variant H3.3. *Genome biology* **14**: R121.
- Laberge RM, Boissonneault G. 2005. On the nature and origin of DNA strand breaks in elongating spermatids. *Biology of reproduction* **73**: 289-296.
- Lane N, Dean W, Erhardt S, Hajkova P, Surani A, Walter J, Reik W. 2003. Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. *Genesis* **35**: 88-93.
- Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG. 2013. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**: 1021-1036.
- Leduc F, Maquennehan V, Nkoma GB, Boissonneault G. 2008. DNA damage response during chromatin remodeling in elongating spermatids of mice. *Biology of reproduction* **78**: 324-332.
- Lee J, Inoue K, Ono R, Ogonuki N, Kohda T, Kaneko-Ishino T, Ogura A, Ishino F. 2002. Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* **129**: 1807-1817.
- Lee JS, Shukla A, Schneider J, Swanson SK, Washburn MP, Florens L, Bhaumik SR, Shilatifard A. 2007. Histone crosstalk between H2B monoubiquitination and H3 methylation mediated by COMPASS. *Cell* **131**: 1084-1096.
- Lewis PW, Elsaesser SJ, Noh KM, Stadler SC, Allis CD. 2010. Daxx is an H3.3-specific histone chaperone and cooperates with ATRX in replication-independent chromatin assembly at

- telomeres. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 14075-14080.
- Li E, Bestor TH, Jaenisch R. 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69**: 915-926.
- Lister R, Pelizzola M, Dowen RH, Hawkins RD, Hon G, Tonti-Filippini J, Nery JR, Lee L, Ye Z, Ngo QM et al. 2009. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**: 315-322.
- Lu LY, Wu J, Ye L, Gavrilina GB, Saunders TL, Yu X. 2010. RNF8-dependent histone modifications regulate nucleosome removal during spermatogenesis. *Developmental cell* **18**: 371-384.
- Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* **389**: 251-260.
- Mandal P, Verma N, Chauhan S, Tomar RS. 2013. Unexpected histone H3 tail-clipping activity of glutamate dehydrogenase. *The Journal of biological chemistry* **288**: 18743-18757.
- Marinus MG. 1987. DNA methylation in Escherichia coli. *Annual review of genetics* **21**: 113-131.
- Martianov I, Brancorsini S, Catena R, Gansmuller A, Kotaja N, Parvinen M, Sassone-Corsi P, Davidson I. 2005. Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 2808-2813.
- Marzluff WF, Gongidi P, Woods KR, Jin J, Maltais LJ. 2002. The human and mouse replication-dependent histone genes. *Genomics* **80**: 487-498.
- Marzluff WF, Wagner EJ, Duronio RJ. 2008. Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nature reviews Genetics* **9**: 843-854.
- Mattioli F, Vissers JH, van Dijk WJ, Ikpa P, Citterio E, Vermeulen W, Marteijs JA, Sixma TK. 2012. RNF168 ubiquitinates K13-15 on H2A/H2AX to drive DNA damage signaling. *Cell* **150**: 1182-1195.
- Maze I, Noh KM, Soshnev AA, Allis CD. 2014. Every amino acid matters: essential contributions of histone variants to mammalian development and disease. *Nature reviews Genetics* **15**: 259-271.
- McKittrick E, Gafken PR, Ahmad K, Henikoff S. 2004. Histone H3.3 is enriched in covalent modifications associated with active chromatin. *Proceedings of the National Academy of Sciences of the United States of America* **101**: 1525-1530.
- McPherson S, Longo FJ. 1993a. Chromatin structure-function alterations during mammalian spermatogenesis: DNA nicking and repair in elongating spermatids. *European journal of histochemistry : EJH* **37**: 109-128.
- McPherson SM, Longo FJ. 1993b. Nicking of rat spermatid and spermatozoa DNA: possible involvement of DNA topoisomerase II. *Developmental biology* **158**: 122-130.
- Meistrich ML, Bucci LR, Trostle-Weige PK, Brock WA. 1985. Histone variants in rat spermatogonia and primary spermatocytes. *Developmental biology* **112**: 230-240.
- Meyer-Ficca ML, Lonchar JD, Ihara M, Meistrich ML, Austin CA, Meyer RG. 2011. Poly(ADP-ribose) polymerases PARP1 and PARP2 modulate topoisomerase II beta (TOP2B) function during chromatin condensation in mouse spermiogenesis. *Biology of reproduction* **84**: 900-909.
- Minsky N, Shema E, Field Y, Schuster M, Segal E, Oren M. 2008. Monoubiquitinated H2B is associated with the transcribed region of highly expressed genes in human cells. *Nature cell biology* **10**: 483-488.
- Mito Y, Henikoff JG, Henikoff S. 2005. Genome-scale profiling of histone H3.3 replacement patterns. *Nature genetics* **37**: 1090-1097.
- Mizuguchi G, Xiao H, Wisniewski J, Smith MM, Wu C. 2007. Nonhistone Scm3 and histones CenH3-H4 assemble the core of centromere-specific nucleosomes. *Cell* **129**: 1153-1164.
- Monk M, Boubelik M, Lehnert S. 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* **99**: 371-382.

- Montellier E, Boussouar F, Rousseaux S, Zhang K, Buchou T, Fenaille F, Shiota H, Debernardi A, Hery P, Curtet S et al. 2013. Chromatin-to-nucleoprotamine transition is controlled by the histone H2B variant TH2B. *Genes & development* **27**: 1680-1692.
- Moriniere J, Rousseaux S, Steuerwald U, Soler-Lopez M, Curtet S, Vitte AL, Govin J, Gaucher J, Sadoul K, Hart DJ et al. 2009. Cooperative binding of two acetylation marks on a histone tail by a single bromodomain. *Nature* **461**: 664-668.
- Muller S, Filippakopoulos P, Knapp S. 2011. Bromodomains as therapeutic targets. *Expert reviews in molecular medicine* **13**: e29.
- Murr R, Loizou JI, Yang YG, Cuenin C, Li H, Wang ZQ, Herceg Z. 2006. Histone acetylation by Trrap-Tip60 modulates loading of repair proteins and repair of DNA double-strand breaks. *Nature cell biology* **8**: 91-99.
- Mutskov V, Gerber D, Angelov D, Ausio J, Workman J, Dimitrov S. 1998. Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding. *Molecular and cellular biology* **18**: 6293-6304.
- Ng HH, Robert F, Young RA, Struhl K. 2003. Targeted recruitment of Set1 histone methylase by elongating Pol II provides a localized mark and memory of recent transcriptional activity. *Molecular cell* **11**: 709-719.
- Nishio H, Hayashi Y, Moritoki Y, Kamisawa H, Mizuno K, Kojima Y, Kohri K. 2014. Distinctive changes in histone H3K4 modification mediated via Kdm5a expression in spermatogonial stem cells of cryptorchid testes. *The Journal of urology* **191**: 1564-1572.
- Oakes CC, La Salle S, Smiraglia DJ, Robaire B, Trasler JM. 2007. Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. *Developmental biology* **307**: 368-379.
- Okano M, Bell DW, Haber DA, Li E. 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99**: 247-257.
- Okano M, Xie S, Li E. 1998. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature genetics* **19**: 219-220.
- Panier S, Durocher D. 2013. Push back to respond better: regulatory inhibition of the DNA double-strand break response. *Nature reviews Molecular cell biology* **14**: 661-672.
- Piomboni P, Focarelli R, Stendardi A, Ferramosca A, Zara V. 2012. The role of mitochondria in energy production for human sperm motility. *International journal of andrology* **35**: 109-124.
- Purohit JS, Tomar RS, Panigrahi AK, Pandey SM, Singh D, Chaturvedi MM. 2013. Chicken liver glutamate dehydrogenase (GDH) demonstrates a histone H3 specific protease (H3ase) activity in vitro. *Biochimie* **95**: 1999-2009.
- Qian MX, Pang Y, Liu CH, Haratake K, Du BY, Ji DY, Wang GF, Zhu QQ, Song W, Yu Y et al. 2013. Acetylation-mediated proteasomal degradation of core histones during DNA repair and spermatogenesis. *Cell* **153**: 1012-1024.
- Ramsahoye BH, Binizskiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R. 2000. Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proceedings of the National Academy of Sciences of the United States of America* **97**: 5237-5242.
- Ray-Gallet D, Quivy JP, Scamps C, Martini EM, Lipinski M, Almouzni G. 2002. HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Molecular cell* **9**: 1091-1100.
- Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD et al. 2000. Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406**: 593-599.
- Regnier V, Vagnarelli P, Fukagawa T, Zerjal T, Burns E, Trouche D, Earnshaw W, Brown W. 2005. CENP-A is required for accurate chromosome segregation and sustained kinetochore association of BubR1. *Molecular and cellular biology* **25**: 3967-3981.
- Rogakou EP, Boon C, Redon C, Bonner WM. 1999. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *The Journal of cell biology* **146**: 905-916.

- Rogakou EP, Pilch DR, Orr AH, Ivanova VS, Bonner WM. 1998. DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *The Journal of biological chemistry* **273**: 5858-5868.
- Rossetto D, Avvakumov N, Cote J. 2012. Histone phosphorylation: a chromatin modification involved in diverse nuclear events. *Epigenetics : official journal of the DNA Methylation Society* **7**: 1098-1108.
- Santos-Rosa H, Kirmizis A, Nelson C, Bartke T, Saksouk N, Cote J, Kouzarides T. 2009. Histone H3 tail clipping regulates gene expression. *Nature structural & molecular biology* **16**: 17-22.
- Sassone-Corsi P. 2002. Unique chromatin remodeling and transcriptional regulation in spermatogenesis. *Science* **296**: 2176-2178.
- Sato S, Yoshimizu T, Sato E, Matsui Y. 2003. Erasure of methylation imprinting of Igf2r during mouse primordial germ-cell development. *Molecular reproduction and development* **65**: 41-50.
- Sawicka A, Seiser C. 2012. Histone H3 phosphorylation - a versatile chromatin modification for different occasions. *Biochimie* **94**: 2193-2201.
- Saxonov S, Berg P, Brutlag DL. 2006. A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proceedings of the National Academy of Sciences of the United States of America* **103**: 1412-1417.
- Schenk R, Jenke A, Zilbauer M, Wirth S, Postberg J. 2011. H3.5 is a novel hominid-specific histone H3 variant that is specifically expressed in the seminiferous tubules of human testes. *Chromosoma* **120**: 275-285.
- Schubeler D. 2015. Function and information content of DNA methylation. *Nature* **517**: 321-326.
- Seisenberger S, Andrews S, Krueger F, Arand J, Walter J, Santos F, Popp C, Thienpont B, Dean W, Reik W. 2012. The dynamics of genome-wide DNA methylation reprogramming in mouse primordial germ cells. *Molecular cell* **48**: 849-862.
- Seki Y, Hayashi K, Itoh K, Mizugaki M, Saitou M, Matsui Y. 2005. Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. *Developmental biology* **278**: 440-458.
- Seki Y, Yamaji M, Yabuta Y, Sano M, Shigeta M, Matsui Y, Saga Y, Tachibana M, Shinkai Y, Saitou M. 2007. Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. *Development* **134**: 2627-2638.
- Shang E, Nickerson HD, Wen D, Wang X, Wolgemuth DJ. 2007. The first bromodomain of Brdt, a testis-specific member of the BET sub-family of double-bromodomain-containing proteins, is essential for male germ cell differentiation. *Development* **134**: 3507-3515.
- Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K et al. 2007. The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450**: 908-912.
- Shi J, Vakoc CR. 2014. The mechanisms behind the therapeutic activity of BET bromodomain inhibition. *Molecular cell* **54**: 728-736.
- Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA, Shi Y. 2004. Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. *Cell* **119**: 941-953.
- Shinagawa T, Huynh LM, Takagi T, Tsukamoto D, Tomaru C, Kwak HG, Dohmae N, Noguchi J, Ishii S. 2015. Disruption of Th2a and Th2b genes causes defects in spermatogenesis. *Development*.
- Shivaraju M, Unruh JR, Slaughter BD, Mattingly M, Berman J, Gerton JL. 2012. Cell-cycle-coupled structural oscillation of centromeric nucleosomes in yeast. *Cell* **150**: 304-316.
- Shuaib M, Ouararhni K, Dimitrov S, Hamiche A. 2010. HJURP binds CENP-A via a highly conserved N-terminal domain and mediates its deposition at centromeres. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 1349-1354.
- Slotkin RK, Martienssen R. 2007. Transposable elements and the epigenetic regulation of the genome. *Nature reviews Genetics* **8**: 272-285.
- Smallwood SA, Kelsey G. 2012. De novo DNA methylation: a germ cell perspective. *Trends in genetics : TIG* **28**: 33-42.

- Soboleva TA, Nekrasov M, Pahwa A, Williams R, Huttley GA, Tremethick DJ. 2012. A unique H2A histone variant occupies the transcriptional start site of active genes. *Nature structural & molecular biology* **19**: 25-30.
- Sonnack V, Failing K, Bergmann M, Steger K. 2002. Expression of hyperacetylated histone H4 during normal and impaired human spermatogenesis. *Andrologia* **34**: 384-390.
- Sperling AS, Grunstein M. 2009. Histone H3 N-terminus regulates higher order structure of yeast heterochromatin. *Proceedings of the National Academy of Sciences of the United States of America* **106**: 13153-13159.
- Stock JK, Giadrossi S, Casanova M, Brookes E, Vidal M, Koseki H, Brockdorff N, Fisher AG, Pombo A. 2007. Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. *Nature cell biology* **9**: 1428-1435.
- Sun ZW, Allis CD. 2002. Ubiquitination of histone H2B regulates H3 methylation and gene silencing in yeast. *Nature* **418**: 104-108.
- Tachiwana H, Kagawa W, Osakabe A, Kawaguchi K, Shiga T, Hayashi-Takanaka Y, Kimura H, Kurumizaka H. 2010. Structural basis of instability of the nucleosome containing a testis-specific histone variant, human H3T. *Proceedings of the National Academy of Sciences of the United States of America* **107**: 10454-10459.
- Tachiwana H, Osakabe A, Kimura H, Kurumizaka H. 2008. Nucleosome formation with the testis-specific histone H3 variant, H3t, by human nucleosome assembly proteins in vitro. *Nucleic acids research* **36**: 2208-2218.
- Tachiwana H, Osakabe A, Shiga T, Miya Y, Kimura H, Kagawa W, Kurumizaka H. 2011. Structures of human nucleosomes containing major histone H3 variants. *Acta crystallographica Section D, Biological crystallography* **67**: 578-583.
- Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. 2004. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* **116**: 51-61.
- Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L et al. 2009. Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324**: 930-935.
- Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N et al. 2011. Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. *Cell* **146**: 1016-1028.
- Tanaka H, Iguchi N, Isotani A, Kitamura K, Toyama Y, Matsuoka Y, Onishi M, Masai K, Maekawa M, Toshimori K et al. 2005. HANP1/H1T2, a novel histone H1-like protein involved in nuclear formation and sperm fertility. *Molecular and cellular biology* **25**: 7107-7119.
- Tanaka H, Miyagawa Y, Tsujimura A, Matsumiya K, Okuyama A, Nishimune Y. 2003. Single nucleotide polymorphisms in the protamine-1 and -2 genes of fertile and infertile human male populations. *Molecular human reproduction* **9**: 69-73.
- Tang MC, Jacobs SA, Mattiske DM, Soh YM, Graham AN, Tran A, Lim SL, Hudson DF, Kalitsis P, O'Bryan MK et al. 2015. Contribution of the Two Genes Encoding Histone Variant H3.3 to Viability and Fertility in Mice. *PLoS genetics* **11**: e1004964.
- Taverna SD, Li H, Ruthenburg AJ, Allis CD, Patel DJ. 2007. How chromatin-binding modules interpret histone modifications: lessons from professional pocket pickers. *Nature structural & molecular biology* **14**: 1025-1040.
- Thoma F, Koller T, Klug A. 1979. Involvement of histone H1 in the organization of the nucleosome and of the salt-dependent superstructures of chromatin. *The Journal of cell biology* **83**: 403-427.
- Trostle-Weige PK, Meistrich ML, Brock WA, Nishioka K. 1984. Isolation and characterization of TH3, a germ cell-specific variant of histone 3 in rat testis. *The Journal of biological chemistry* **259**: 8769-8776.

- Tsukada Y, Fang J, Erdjument-Bromage H, Warren ME, Borchers CH, Tempst P, Zhang Y. 2006. Histone demethylation by a family of JmjC domain-containing proteins. *Nature* **439**: 811-816.
- Tulsiani DR, Abou-Haila A, Loeser CR, Pereira BM. 1998. The biological and functional significance of the sperm acrosome and acrosomal enzymes in mammalian fertilization. *Experimental cell research* **240**: 151-164.
- Turner JM. 2007. Meiotic sex chromosome inactivation. *Development* **134**: 1823-1831.
- Turner JM, Mahadevaiah SK, Fernandez-Capetillo O, Nussenzweig A, Xu X, Deng CX, Burgoyne PS. 2005. Silencing of unsynapsed meiotic chromosomes in the mouse. *Nature genetics* **37**: 41-47.
- Unni E, Zhang Y, Kangasniemi M, Saperstein W, Moss SB, Meistrich ML. 1995. Stage-specific distribution of the spermatid-specific histone 2B in the rat testis. *Biology of reproduction* **53**: 820-826.
- Unnikrishnan A, Gafken PR, Tsukiyama T. 2010. Dynamic changes in histone acetylation regulate origins of DNA replication. *Nature structural & molecular biology* **17**: 430-437.
- van der Heijden GW, Derijck AA, Posfai E, Giele M, Pelczar P, Ramos L, Wansink DG, van der Vlag J, Peters AH, de Boer P. 2007. Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. *Nature genetics* **39**: 251-258.
- Verreault A, Kaufman PD, Kobayashi R, Stillman B. 1996. Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. *Cell* **87**: 95-104.
- Vincent JJ, Huang Y, Chen PY, Feng S, Calvopina JH, Nee K, Lee SA, Le T, Yoon AJ, Faull K et al. 2013. Stage-specific roles for tet1 and tet2 in DNA demethylation in primordial germ cells. *Cell stem cell* **12**: 470-478.
- Vossaert L, Meert P, Scheerlinck E, Glibert P, Van Roy N, Heindryckx B, De Sutter P, Dhaenens M, Deforce D. 2014. Identification of histone H3 clipping activity in human embryonic stem cells. *Stem cell research* **13**: 123-134.
- Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. 2004. Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**: 873-878.
- Ward IM, Chen J. 2001. Histone H2AX is phosphorylated in an ATR-dependent manner in response to replicational stress. *The Journal of biological chemistry* **276**: 47759-47762.
- Watanabe D, Suetake I, Tada T, Tajima S. 2002. Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. *Mechanisms of development* **118**: 187-190.
- Watt F, Molloy PL. 1988. Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes & development* **2**: 1136-1143.
- Wells D, Kedes L. 1985. Structure of a human histone cDNA: evidence that basally expressed histone genes have intervening sequences and encode polyadenylated mRNAs. *Proceedings of the National Academy of Sciences of the United States of America* **82**: 2834-2838.
- Western PS, Miles DC, van den Bergen JA, Burton M, Sinclair AH. 2008. Dynamic regulation of mitotic arrest in fetal male germ cells. *Stem cells* **26**: 339-347.
- White-Cooper H, Davidson I. 2011. Unique aspects of transcription regulation in male germ cells. *Cold Spring Harbor perspectives in biology* **3**.
- Wiedemann SM, Mildner SN, Bonisch C, Israel L, Maiser A, Matheisl S, Straub T, Merkl R, Leonhardt H, Kremmer E et al. 2010. Identification and characterization of two novel primate-specific histone H3 variants, H3.X and H3.Y. *The Journal of cell biology* **190**: 777-791.
- Williams JS, Hayashi T, Yanagida M, Russell P. 2009. Fission yeast Scm3 mediates stable assembly of Cnp1/CENP-A into centromeric chromatin. *Molecular cell* **33**: 287-298.
- Wirbelauer C, Bell O, Schubeler D. 2005. Variant histone H3.3 is deposited at sites of nucleosomal displacement throughout transcribed genes while active histone modifications show a promoter-proximal bias. *Genes & development* **19**: 1761-1766.

- Witt O, Albig W, Doenecke D. 1996. Testis-specific expression of a novel human H3 histone gene. *Experimental cell research* **229**: 301-306.
- Wong LH, Ren H, Williams E, McGhie J, Ahn S, Sim M, Tam A, Earle E, Anderson MA, Mann J et al. 2009. Histone H3.3 incorporation provides a unique and functionally essential telomeric chromatin in embryonic stem cells. *Genome research* **19**: 404-414.
- Woodcock DM, Crowther PJ, Diver WP. 1987. The majority of methylated deoxycytidines in human DNA are not in the CpG dinucleotide. *Biochemical and biophysical research communications* **145**: 888-894.
- Wu C, Morris JR. 2001. Genes, genetics, and epigenetics: a correspondence. *Science* **293**: 1103-1105.
- Xue Y, Vashisht AA, Tan Y, Su T, Wohlschlegel JA. 2014. PRB1 is required for clipping of the histone H3 N terminal tail in *Saccharomyces cerevisiae*. *PLoS one* **9**: e90496.
- Yan W, Ma L, Burns KH, Matzuk MM. 2003. HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. *Proceedings of the National Academy of Sciences of the United States of America* **100**: 10546-10551.
- Yuen BT, Bush KM, Barrilleaux BL, Cotterman R, Knoepfler PS. 2014. Histone H3.3 regulates dynamic chromatin states during spermatogenesis. *Development* **141**: 3483-3494.
- Zeitlin SG, Shelby RD, Sullivan KF. 2001. CENP-A is phosphorylated by Aurora B kinase and plays an unexpected role in completion of cytokinesis. *The Journal of cell biology* **155**: 1147-1157.
- Zeng L, Zhou MM. 2002. Bromodomain: an acetyl-lysine binding domain. *FEBS letters* **513**: 124-128.
- Zentner GE, Henikoff S. 2013. Regulation of nucleosome dynamics by histone modifications. *Nature structural & molecular biology* **20**: 259-266.
- . 2014. High-resolution digital profiling of the epigenome. *Nature reviews Genetics* **15**: 814-827.
- Zhang L, Lu X, Lu J, Liang H, Dai Q, Xu GL, Luo C, Jiang H, He C. 2012. Thymine DNA glycosylase specifically recognizes 5-carboxylcytosine-modified DNA. *Nature chemical biology* **8**: 328-330.
- Zhou W, Zhu P, Wang J, Pascual G, Ohgi KA, Lozach J, Glass CK, Rosenfeld MG. 2008. Histone H2A monoubiquitination represses transcription by inhibiting RNA polymerase II transcriptional elongation. *Molecular cell* **29**: 69-80.

Acknowledgement

First of all, I would like to express my deepest thanks to my advisor, Antoine Peters, for giving me an opportunity to work in his lab and in FMI. I am very grateful to have an advisor who provided me the freedom to explore the projects that I chose as well as the helpful guidance when my step faltered.

I would also like to thank my committee members, Anja Groth, Marc Bühler and Ulrich Rass for serving as my committee members and for providing me brilliant comments and suggestions during each committee.

I would like to thank all of present and past members of the Antoine's lab. I especially want to thank Shihori Yokobayashi for her guidance and discussion in a wonderful cooperation that I learned a lot from her. Also thanks to Mizue Hisano for her patience to introduce me when I joined the lab and for her encouragement. Thanks to Helene Royo and Serap Erkek for their support in the bioinformatic analysis. Thanks to Mathieu Tardat for his great discussion in experiments and for providing the enjoyment every day. Thanks to Mark Gill for sharing his knowledge. Thanks to Zichuan Liu and Xiaojun Hou for their help in mouse experiments.

Next, I would like to express my sincere gratitude to FMI facilities. Especially, thanks to Ragna Sack from Protein Analysis Facility for her indispensable help in mass spectrometry analysis. Thanks to Hubertus Kohler from the Flow Cytometry Facility for his patience to assist me in each cell sorting experiment. Thanks to Sophie Dessus-Babus for performing ChIP-seq experiments. Thanks to Jeremy from Protein Structure Facility for giving assistance in FPLC experiments. Thanks to Animal Facility for animal housing and care. Without their technical contribution, my research would not have even been possible.

The support and encouragement of my family and friends were unforgettable. Thanks to all of my friends in Switzerland and Taiwan who always made my student life full of happiness. Finally I want to extend my sincerest thanks to my parents, brother and boyfriend, Yen-Lin, for their lovely support to let me focus on my study. Thank you.

Thanks to mice.