creative commons

Attribution-Noncommercial-No Derivative Works 2.5 Switzerland

---

You are free:

to Share — to copy, distribute and transmit the work

Under the following conditions:

Attribution. You must attribute the work in the manner specified by the author or licensor (but not in any way that suggests that they endorse you or your use of the work).

Noncommercial. You may not use this work for commercial purposes.

No Derivative Works. You may not alter, transform, or build upon this work.

- For any reuse or distribution, you must make clear to others the license terms of this work. The best way to do this is with a link to this web page.
- Any of the above conditions can be waived if you get permission from the copyright holder.
- Nothing in this license impairs or restricts the author's moral rights.

Your fair dealing and other rights are in no way affected by the above.

This is a human-readable summary of the Legal Code (the full license) available in German:
http://creativecommons.org/licenses/by-nc-nd/2.5/ch/legalcode.de

Disclaimer:
The Commons Deed is not a license. It is simply a handy reference for understanding the Legal Code (the full license) — it is a human-readable expression of some of its key terms. Think of it as the user-friendly interface to the Legal Code beneath. This Deed itself has no legal value, and its contents do not appear in the actual license. Creative Commons is not a law firm and does not provide legal services. Distributing of, displaying of, or linking to this Commons Deed does not create an attorney-client relationship.

Quelle: http://creativecommons.org/licenses/by-nc-nd/2.5/ch/deed.en Datum: 3.4.2009

Role of histone methylation in paternal transmission of epigenetic information

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch–Naturwissenschaftlichen Fakultät
der Universität Basel

von

Urszula Brykczynska
aus Polen

Basel, 2010

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät der Universität Basel auf Antrag von Dr. Antoine Peters, Professor Dr. Jan Hofsteenge, Professor Dr. Susan Gasser, und Professor Dr. Genevieve Almouzni.

Basel, den 15. September 2009

Professor Dr. E. Parlow
Dekan

Summary

During the development of multi-cellular organisms, one genome gives rise to multiple differentiated cell types. This is achieved by sequence specific transcription factors and different epigenetic mechanisms, which collaborate in reading the genetic information. These epigenetic mechanisms coordinate the establishment and maintenance of transcriptional programs in a lineage specific manner during development. However, very little is known whether such epigenetic information can be also passed to the next generation.

Mammalian gametes may differ in their potential to transmit chromatin encoded epigenetic information. The oocyte genome is organized in a nucleosomal configuration with DNA wrapped around histones that carry various post translational modifications. By contrast, the paternal genome undergoes a major reorganization during the last stages of spermatogenesis. Most of histones are replaced by protamines, which after fertilization, are exchanged by maternally provided histones. Nevertheless, approximately 10% of histones are retained in human spermatozoa, raising a possibility for a paternal, epigenetic contribution to the next generation.

In this thesis, I aimed to determine the genomic localisation of histones retained in sperm and to analyze their potential to influence transcription after fertilization. We show that histones isolated from mouse and human spermatozoa are carrying multiple post translational modifications, many of which have functions in gene regulation. In our genome wide analysis of human promoters, we demonstrate that two of these marks, Trithorax/Set1 mediated dimethylation of lysine 4 of histone H3 (H3K4me2) and Polycomb mediated trimethylation of lysine 27 of histone H3 (H3K27me3), occupy functionally defined groups of genes. H3K4me2-marked promoters control genes with functions in spermatogenesis and cellular homeostasis, suggesting that this mark reflects germline transcription. By contrast, multiple developmental regulators, which are Polycomb targets in pluripotent somatic cells, are marked by H3K27me3 in human sperm. Similarly to somatic cells, the presence of this mark correlates with gene repression during spermatogenesis and in the early embryo. We propose a model in which H3K27me3, transmitted by sperm, assures repression of developmental regulators at the totipotent stage of the preimplantation development. Finally, we demonstrate that a number of these developmental regulators are also marked by H3K27me3 in mouse spermatozoa, implicating an evolutionary conserved role for histone methylation in the paternal transmission of epigenetic information.

Table of contents

1. Introduction

1.1 From Linnaeus to epigenetics

The classical model of Mendelian genetic inheritance has been challenged in recent years. There is growing evidence that phenotypic traits can be inherited across generations without involvement of the DNA encoded genetic information. Schemes of the inheritance of such traits do not follow the rules of genetics and are therefore called non-Mendelian.

One of the first examples of such phenomena dates back to the studies of Linnaeus, who described a naturally occurring mutant of a plant *Linaria Vulgaris* (Linneaus, 1749). Wild type plants have flowers with bilateral symmetry, whereas mutant flowers are radial (Fig. 1a,b). 250 years later Cubas and colleagues showed that the phenotype is caused by the silencing of the *Lcyc* gene (Cubas et al., 1999).The silencing occurs not through mutation of DNA sequence but through methylation of DNA at the promoter of *Lcyc*. The phenotype is stably inherited over many generations. However, spontaneous reversion to the wild-type phenotype, caused by removal of methyl groups from the DNA is also observed (Cubas et al., 1999).

In analogy to the term “mutation”, used in genetics, a stable change in gene expression that does not involve changes in DNA sequence is called “epimutation”. In contrast to mutations, epimutations have different levels of reversibility which affect the schemes of inheritance. Factors that encode epimutated state, such as DNA methylation are called epigenetic.

Paramutations

In some cases an epimutated allele can influence the wild type allele in *trans*, increasing the complexity of non-Mendelian segregation of a phenotype. This phenomenon is called paramutation and was first described in maize (Brink, 1956). When the R^r allele encoding dark purple colour of the maize seeds was crossed with R^{st} allele encoding purple stippled pattern, heterozygotic plants were transmitting a changed R^r allele. The allele was now encoding a much lighter pigmentation of the seeds and was stably inherited. It reverted toward the standard type when made homozygous, but only partially. Currently, multiple examples of paramuation are known, mostly in plants

(reviewed in (Chandler, 2007)). The mechanisms by which the two alleles communicate are not well understood. One well studied example is *b1* locus in maize. The key sequences required for the paramutation at that locus are tandem repeats located upstream of the *b1* transcription start site (Stam et al., 2002). The repeats are transcribed and produce non-coding RNA. This RNA is thought to mediate communication between the two alleles and to modulate the transcription of the *b1* genes by establishing distinct chromatin states (Alleman et al., 2006).

An RNA based mechanism was also described for an epimutation-like phenomenon at the *Kit* locus in mice (Rassoulzadegan et al., 2006). In that study, a dysfunctional *Kit* allele, with an insertion in coding region, was engineered. Both heterozygous animals and their wild type homozygous offspring (coming from heterozygous crosses) had decreased *Kit* expression and showed white feet and white tail tip phenotype. These genetically wild type paramutated animals were transmitting the phenotype to the next generations by both male and female germ line. The proposed mechanism of this phenomenon is linked to the aberrant expression of *Kit* during spermatogenesis and accumulation of Kit RNA in sperm of heterozygous and paramutated males (Rassoulzadegan et al., 2006). Their offspring inherits high levels of Kit RNA, which most probably triggers a silencing response in the early embryo. A recent paper by the same group describes a similar effect, triggered by injection of fragments of Cdk9 RNA into the early embryo (Wagner et al., 2008). Resulting animals showed elevated Cdk9 expression and an associated cardiac disorder. The phenotype was transmitted to at least three generations. Despite the different effect on gene expression triggered in Kit and Cdk9 paramutations (down and up regulation respectively), in both studies aberrant levels of RNA were detected in sperm (Rassoulzadegan et al., 2006; Wagner et al., 2008). These observations point to RNA, as one of epigenetic factors that mediates paternal transgenerational. The means of the maternal inheritance were not analysed in these studies but a similar RNA based mechanism may operate in female germline.

Metastable epialleles

Other examples of non-Mendelian inheritance in mice are coming from analysis of genomic loci which show stochastically variable expression and are called metastable epialleles (Morgan et al., 1999; Rakyan et al., 2003). Metastable epialleles are best analyzed in inbred mice strains, which provide an opportunity to study phenotypic

differences between individuals with identical genotypes. The *Agouti (A)* locus is responsible for the production of yellow hair pigment. In wild type animals, the pigment is produced only during a short period of the hair growth resulting in a light brown (agouti) coat color. A^{vy} is a metastable epiallele of *A*, which is carrying an intracisternal A particle (IAP) retrotransposon, inserted upstream of *A* promoter (Morgan et al., 1999). The retrotransposon is driving the ectopic expression of the *Agouti* locus, causing fully yellow coat color, obesity, diabetes and increased susceptibility to tumors. The A^{vy} animals show mosaic expression of the retrotransposon, resulting in a spectrum of phenotypes with the coat color varying from yellow through variegated yellow/agouti to agouti (Fig.1c). The distribution of the phenotypes among offspring depends on the phenotype of the mother (Morgan et al., 1999). Even though no paternal effect is observed, a phenotype driven by an IAP retrotransposon, inserted into another metastable epiallele $Auxin^{fu}$, is transmitted by both parents (Rakyan et al., 2003). Therefore, there exist means for transmission of the epimutated states by both germ lines. For both A^{vy} and $Auxin^{fu}$ loci different levels of expression correlate with DNA methylation of the inserted retrotransposon sequence (Morgan et al., 1999; Rakyan et al., 2003). However, this modification was shown to be entirely erased from the A^{vy} locus immediately post fertilization (Blewitt et al., 2006). Recently, a number of epigenetic factors have been identified that influence the expression of A^{vy} and show transgenerational effects, supporting other than DNA methylation mechanisms of transmission (Blewitt et al., 2006; Chong et al., 2007a).

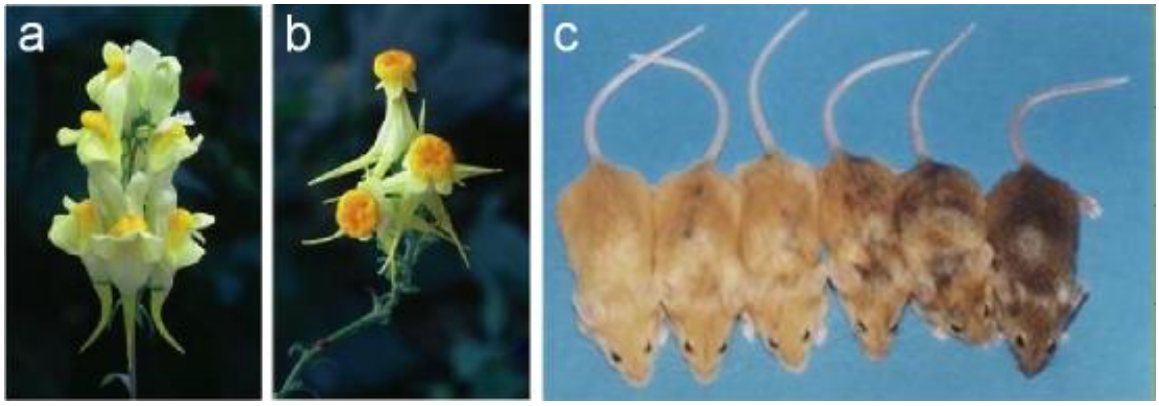


Figure 1. Epimutations in plants and animals (a) wild type *Linaria Vulgaris* **(b)** *Linaria Vulgaris* carrying epimutation of *Lcyc* gene **(c)** Spectrum of coat colours of mice carrying A^{vy} metastable epiallele (adapted from (Morgan et al., 1999))

Imprinted genes

One of the best understood examples of transgenerational inheritance of states of gene expression, so called epigenetic states, is genomic imprinting. Genomic imprinting is a phenomenon in mammals where a gene is expressed only from one allele, either coming from the mother of from the father (reviewed in (Feil, 2009)). This differential expression is dependent on DNA methylation at imprinting control regions (ICRs), which are located within or outside the differentially expressed loci. Depending on their methylation status, ICRs either enhance or repress the expression. Reciprocal DNA methylation patterns are established on ICRs during the male and female germ cell development and are brought to the embryo by spermatozoon and oocyte. There are so far around 80 genes identified that undergo genomic imprinting. Most ICRs are methylated at the maternal allele and only three on the paternal one (reviewed in (Feil, 2009)). Proper DNA methylation of ICRs in sperm and oocyte are necessary for the successful development of the embryo. Oocytes lacking maternal imprints give raise to embryos that die *in utero* (Bourc'his et al., 2001; Kaneda et al., 2004). Males with the impaired DNA methylation on paternal ICRs are infertile, thus its effect on the embryo can not be determined (Bourc'his et al., 2001; Kaneda et al., 2004). Nonetheless, embryos carrying two maternal genomes die *in utero*, showing that the paternal genome is necessary for the development (Surani and Barton, 1983). Further, this phenotype can be rescued by the deletion of two out of three paternal ICRs in one of the genomes, demonstrating that proper regulation of paternally imprinted genes is crucial for the development (Kawahara et al., 2007). Many of imprinted genes are homologues between mouse and human and show the imprinted expression in both species. Reminiscent of the situation in mice, aberrant methylation patterns at paternal ICRs have been found in sperm of infertile men (Marques et al., 2004).

Environmentally induced changes in epigenetic programs

It is a vital question if epigenetic states can be shaped by the environment during the life span of an organism and then be stably transmitted to the next generations.

It has been reported that the methyl donor supply in the maternal diet can influence the coat colour of A^{vy} mice (Cooney et al., 2002; Dolinoy et al., 2006). The establishment of the mosaic DNA methylation pattern at A^{vy} locus is occurring at the post-implantation stages of development. Therefore, high level of methyl donor in the maternal diet could directly increase the number of DNA methylated loci. Consistently,

the range of coat colors of the offspring was shifted towards agouti (Dolinoy et al., 2006). However, the methyl group can be added not only to the DNA, but also to a variety of proteins involved in gene regulation. Therefore, methyl donor supplementation might have primarily affected other than DNA methylation regulatory mechanisms.

The somatic epigenetic changes in response to methyl donor raise the question of whether methyl donor supplementation also affects the germ line, and whether any changes could be maintained to the next generations. Indeed, offspring of the A^{vy} fetus, exposed to methyl donor supplementation showed a shift towards the agouti phenotype as well (Cropley et al., 2006). Yet, it is important to mention that the methyl donor was supplied at the time of primordial germ cell (PGC) formation in the fetus. Hence, all the three generations: the mother, the fetus and its germ cells were directly subjected to the treatment. An analysis that spans over subsequent generations is needed to truly argue for a transgenerational effect. Nevertheless, a transmission of the phenotype by the germ line argues that the epigenetic mark established at A^{vy} locus was maintained through the germ cells development and was not erased after fertilization. Additional criticism of the presented study points out that the first generation of animals was not entirely yellow, but mosaic with a high proportion of agouti. Thus, the supplementation only supported the maintenance of the silenced state, but did not cause *de-novo* establishment of epigenetic information (Waterland et al., 2007).

Several studies have also addressed potential transgenerational effects of endocrine disruptors. A group of widely used pesticides, herbicides and fungicides has characteristics of endocrine disruptors, meaning they can mimic mammalian estrogens and potentially effect development of males and females. Exposure to vinclozolin (fungicide) at the time of gonadal determination has been reported to cause a variety of abnormalities in offspring (Anway et al., 2005). The effects were transmitted down the male germ line for at least three generations (in the absence of vinclozolin) and correlated with an increased DNA methylation in sperm. However, it is not clear whether the phenotype is a direct effect of the DNA methylation levels or if other epigenetic factors are involved. A recent report by another group questions if there is at all a transgenerational effect of endocrine disruptors, suggesting that the results of Anway and colleagues suffered from an inherent artefact of the experimental design (Inawaka et al., 2009).

Evolutionary impact of epigenetic changes

The environmental effects on gene expression described so far can be classified as non-adaptive. Phenotypes inherited by the offspring are either the result of a natural variation of epialleles or the result of the aberrant establishment of epigenetic information. There are no advantages for the offspring in inheriting such phenotypes. The hypothesis that experience of a parent can be transmitted to the offspring and induce some advantageous trait or behaviour has been known as the Lamarckian Hypothesis and was rejected by the Darwinian Theory of Evolution and further by the Modern Synthesis (reviewed in (Rando and Verstrepen, 2007)). It is unreasonable to expect such effects in animals, as germ line and soma are separated very early during development. In contrast, such effects are observed in plants where germ line arises in the adult organism and therefore can be directly affected by the environment (Bastow et al., 2004). Still, a study from *Drosophila melanogaster* implies the existence of an evolutionary mechanism that makes use of random epigenetic variation. The authors propose that in stable environmental conditions specific factors counteract the manifestation of naturally occurring epigenetic variation. In conditions of stress, this suppression is released and an advantageous trait can be selected from a pool of revealed random phenotypes (Ruden et al., 2008; Sollars et al., 2003).

The theory is based on a screen for enhancers of spontaneously occurring morphological phenotypes, which has been performed using flies with a mutation predisposing it for eye malformations (Sollars et al., 2003). Among the identified proteins were Hsp90, a protein chaperon, and several proteins involved in epigenetic gene regulation (from the Trithorax group). Observed phenotypes were heritable through the female germ line and the malformation persisted into the subsequent generations, also in the animals that were not carrying any more mutations for Hsp90 or Trithorax. Several different phenotypes were observed in the same isogenic fly strain used, arguing for the epigenetic and not genetic basis of the malformations. The presence of epigenetic factors like Trithorax among enhancers of the malformations additionally supports this notion. The authors propose that in conditions of stress Hsp90 diverts from its chaperone function, revealing hidden phenotypes and that advantageous ones can be selected and fixed (Ruden et al., 2008; Sollars et al., 2003).

To allow for a selection on epigenetically encoded phenotypes, epigenetic states must be transmitted across multiple generations. Interestingly, in flies there is limited or no DNA methylation reported (Phalke et al., 2009; Tweedie et al., 1997). Inheritance of

expression states must be therefore mediated by other epigenetic factors. Indeed, there are data indicating that the activating and the repressive protein complexes of Trithorax and Polycomb group proteins are involved in transgenerational transmission of epigenetic states in *Drosophila* (Cavalli and Paro, 1998). *Fab-7* is a genetic regulatory element, which can induce silencing of a downstream reporter construct. This silencing is mediated by the association of the Polycomb group proteins with *Fab-7*. Upon activation of the reporter, Polycomb group proteins are displaced from the construct, which remains occupied by Trithorax group proteins. It was observed that even a short pulse of gene activation, at a specific time during embryogenesis, can stably release silencing, and the active sate persists throughout the adult development. Furthermore, the active state is transmitted to the subsequent generations by the female germline. This effect is only observed when *Fab-7* is present, arguing that a specific epigenetic state established on this locus is transmitted (Cavalli and Paro, 1998).

Non-Mendelian inheritance in humans – epigenetic basis of heritable diseases?

The above examples show clearly that information influencing gene expression, so called epigenetic information, can be transmitted across generations in plants and animals. Studies of genomic imprinting provide a proof of principle that epigenetic inheritance exists in humans as well. It is highly probable that some hereditary diseases have an epigenetic basis. Increased knowledge in this field would provide a new spectrum of targets for the drug research. Yet, studies in human are largely limited by the lack of isogenic populations and limited data sets that would span for more than two generations. Several studies indicate that germline transmitted epimutations of DNA mismatch repair genes *MLH1* and *MSH2* are linked to an increased risk of colorectal cancer (Chan et al., 2006; Hitchins et al., 2007; Suter et al., 2004). Two colorectal cancer patients were reported with soma wide abnormal DNA methylation and silencing of the promoter of *MLH1*. No DNA mutations in the region of *MLH1* locus were detected. The same epimutation was also detected in a small fraction of sperm cells of one of the individuals, arguing for germ line transmission (Suter et al., 2004). However, in a subsequent study the authors failed to find any epimutated alleles in sperm of another individual with a somatic epimutation. In contrast, they showed maternal inheritance of an epimutated allele, but without proof for the presence of the epimutation in the female germline (Hitchins et al., 2007). Heritable DNA hypermethylation has also been described for the promoter of *MSH2* (Chan et al., 2006). Three siblings carrying this

epimutation developed colorectal tumors. Further analysis revealed aberrant DNA methylation of *MSH2* in individuals from three generations of the studied family. Interestingly, the epimutation segregated in a Mendelian manner, even though no DNA mutation in the region of the *MSH2* locus was detected. It can not be ruled out that a mutation in other place of the genome was the original cause of aberrant DNA methylation patterns. As was pointed out in a debate, triggered by the publication of (Chan et al., 2006), both *MLH1* and *MSH2* studies are lacking compelling evidence for a germ line epigenetic inheritance (Chong et al., 2007b; Suter and Martin, 2007).

An approach to investigate the epigenetic inheritance in humans is to use statistical data accumulated over generations. Pembery and colleagues analyzed a modern data set on the life style and medical condition of two generations in the Bristol region (Avon Longitudinal Study of Parents and Children). In addition, they analyzed 19th and 20th century parish records from isolated community in Overkalix in Sweden, providing detailed data on births, deaths and diseases of multiple generations (Pembrey et al., 2006). Based on the first data set they found that pre-adolescent paternal smoking was associated with greater body mass index (BMI) in sons, but not daughters. Based on the second data set they found that the paternal grandfather's food supply in pre-adolescence was linked to the mortality risk of grandsons, while the paternal grandmother's food supply was linked to the mortality risk of the granddaughters. Although these studies appear to demonstrate transgenerational effects induced by environmental factors, there is no evidence that transfer of epigenetic information via the germ line is involved. It is very difficult to rule out the involvement of social factors, which can drive such a transgenerational effect. The use of statistical data can become very useful for the research in the future, but only when screening projects involving modern molecular biology methods will be launched on.

Passing the epigenetic information through the germ line

The presented evidence, though some is controversial, suggest a transmission of epigenetic states across generations. Some of the data describe only maternal, some only paternal effects, pointing to differences between the two germ lines, but also showing that both male and female have potential for an epigenetic transmission.

Epigenetic mechanisms are shaping the expression of genes throughout the development of multicellular organisms. To allow for a transition from somatic tissues to specialised germ cells and further to a totipotent embryo, epigenetic patterns of gene

expression have to be erased and re-established. In mammals, this is known to happen in two waves of epigenetic reprogramming – one occurring during the germ cells development and a second one just after fertilization (reviewed in (Reik, 2007)). Nevertheless, there exist sequences naturally resistant to reprogramming, providing a proof of principle for an epigenetic transmission via gametes. Classical examples of such sequences are imprinted genes. The original, parental patterns of DNA methylation at the ICRs are erased during the 1st wave of reprogramming and new sex specific patterns are established during germ cells development. Imprinted sequences escape the global erasure of DNA methylation after fertilization, which is part of the second wave of epigenetic reprogramming (reviewed in (Feil, 2009)). During this process, satellite DNA sequences surrounding centromeres, as well as many transposones, retain their DNA methylation (Lane et al., 2003; Rougier et al., 1998). Nonetheless, a comprehensive list of sequences that escape the reprogramming is not known.

It remains a big question, how complete the reprogramming of epigenetic states encoded by other factors is. Eukaryotic DNA is not naked. It exists in a complex with multiple proteins and RNA molecules, which regulate gene expression, replication and DNA repair. In the coming sections I will review our current knowledge on different components of this complex and their potential to carry epigenetic information across generations.

1.2 Chromatin mediated gene regulation

In eukaryotic cells, DNA is tightly packed in the nucleus. This packing is mediated by proteins that interact with DNA, and constitute together a nucleoprotein complex called chromatin. The basic unit of the chromatin is the nucleosome (Kornberg, 1974) consisting of 146 nucleotides of DNA wrapped around an octamer of four different highly basic proteins called histones.

1.2.1 Posttranslational histone modifications

Nucleosomes, previously considered to function just as structural components of chromatin, are now being recognized as important regulators of chromatin templated processes like transcription, replication and DNA repair. Histones are subjects to a variety of posttranslational modifications such as acetylation of lysines, methylation of lysines and arginines, phosphorylation of serines and threonines, and ubiquitination of

lysines. The majority of these covalently modified residues reside at the flexible N-terminal “tails” of histone H3 and H4 that are localized outside the core of the nucleosome structure.

Histone modifications can affect the strength of histone-DNA interactions and thus directly regulate the accessibility to DNA. On the other hand, such modifications can - either alone or in specific combinations - generate modules that are specifically recognized by certain chromatin associated proteins and that thereby define a specific chromatin state (reviewed in (Kouzarides, 2007)).

The most well understood process regulated by histone modifications is transcription. Generally, active chromatin is characterized by methylation of histone H3 on lysine 4 (H3K4), H3K36 and H3K79 and by acetylation on H3K9, H3K14 and on lysines 5,8, 12 and 16 of histone H4. In contrast, inactive chromatin is enriched in H3K9, H3K27 and H4K20 methylation and in DNA methylation (reviewed in (Kouzarides, 2007)).

Several models have been proposed to explain the function of histone modifications in gene regulation. It is known that histone acetylation or phosphorylation can change the overall charge of the chromatin. The acetylation of histones neutralizes positive charges of histones. Phosphorylation adds a negative charge to chromatin. The charge neutralization model suggests that histone acetylation leads to a decondensation of the chromatin fibre by destabilizing the interaction among nucleosomes and between nucleosomes and DNA. Indeed, there is evidence that histone acetylation can relax chromatin structure *in vivo* and *in vitro* (Shogren-Knaak et al., 2006; Wolffe and Hayes, 1999).

Histone acetylation is generally very dynamic. Acetylation levels are increased by enzymes called histone acetyl transferases (HATs) and removed by histone deacetylases (HDACs). There are many different HATs and HDACs, which target different lysine residues on histones and also other proteins. Most of these enzymes modify more than one lysine residue while some are specific for individual lysines. HATs and HDACs show broad activity, but usually they are part of larger complexes, which show specificity to defined regions of chromatin (reviewed in (Kouzarides, 2007)).

In contrast to histone acetylation, methylation marks are set by histone methyltransferases (HMTs) that recognize their target residues in a sequence specific context. Moreover, each lysine residue has the ability to be modified by mono-, di- or tri-

methylation. The HTMs display specificity towards the level of methylation as well, often modulated by their interaction partners. Except for Dot1, the HMT responsible of methylation at the H3K79 residue localized in the nucleosomal core, all known HMTs belong to a large family of proteins sharing the highly conserved catalytic SET (Suppressor of variegation, Enhancer of Zeste, Trithorax) domain (reviewed in (Kouzarides, 2007)). For many years histone methylation was considered as a permanent mark but this view has been changed drastically with the recent identification of histone demethylases (HDMs) (reviewed in (Swigut and Wysocka, 2007)).

Histone methylation, unlike acetylation, does not change the overall charge of histones. Instead, it functions in recruiting effector proteins to chromatin, which then conduct enzymatic activities such as chromatin remodelling. These proteins can bind to methylated residues via different conserved domains, such as chromodomains (Fischle et al., 2003; Lachner et al., 2001), PHD (Wysocka et al., 2006) and Tudor (Huang et al., 2006) domains .

Depending on the modified residue, histone methylation can have an activating or repressing effect on transcription. In the following section I will concentrate on the role, distribution and possible mechanism of transcriptional regulation of two antagonistic methylation marks: H3K27 methylation established by Polycomb group proteins and H3K4 methylation, mediated by Set1/Trithorax group proteins.

1.2.2 Polycomb and Trithorax group complexes

Polycomb group (PcG) genes were first discovered in *Drosophila melanogaster* as repressors of *Hox* genes, a set of transcription factors that specify cell identity along the anteroposterior axis of segmented animals. PcG proteins form multimeric complexes that are not required to initiate the regulation of *Hox* genes, but rather to maintain their expression state after the initial transcriptional regulators have disappeared from the embryo (Jurgens, 1985; Lewis, 1978). Therefore they provide a mean of a cellular memory that is propagated over the cellular divisions and that maintains specific expression programs. Further experiments have shown that an antagonistic system, involving Trithorax group (TrxG) proteins exists. TrxG proteins are not default activators but function as anti-repressors of PcG target genes. In the absence of TrxG, a homeotic gene can become repressed by the PcG-mediated mechanism even in cells in which it had been active in the early embryo. Therefore, Trx is required continuously throughout development to prevent inappropriate PcG silencing (Klymenko and Muller, 2004; Poux

et al., 2002). These antagonistic mechanisms of gene repression by PcG and gene activation by TrxG are conserved in vertebrates. Members of both groups have been shown to have essential roles in mammalian development (Faust et al., 1995; Glaser et al., 2006; Lee et al., 2006a; O'Carroll et al., 2001; Pasini et al., 2004; Voncken et al., 2003; Yu et al., 1995).

Repressive H3K27 methylation and Polycomb group proteins

Polycomb group proteins act in at least two distinct multi-protein complexes - Polycomb repressive complex 1 and 2 (PRC1 and PRC2). Both PRC1 and PRC2 complexes were primarily described in *Drosophila melanogaster*. In mammals, orthologous complexes have been identified. They were shown to comprise of proteins with highly similar properties and activities as the ones in flies (Cao et al., 2002; Kuzmichev et al., 2002; Levine et al., 2002). In mammals, the PRC2 complex is composed of Enhancer of zeste 2 (Ezh2), Suppressor of zeste 12 (Suz12), Embryonic ectoderm development (Eed) and the histone binding proteins RbAp46/RbAp48. PRC2 mediates histone H3 methylation of lysine 27 through the HMT activity of Ezh2 (Cao et al., 2002; Kuzmichev et al., 2002). This activity is dependent on the presence of the other members of the complex. Suz12 enhances the activity of Ezh2 (Pasini et al., 2004), while Eed, depending on its isoform, can modulate substrate specificity of Ezh2 towards H1K26, H3K27me2/3 or SirT1 (Kuzmichev et al., 2004; Kuzmichev et al., 2005) and is necessary for its activity (Montgomery et al., 2005). Recently, a PRC2-like complex has been described, in which Ezh2 is replaced by its homolog Ezh1. Like PRC2 this complex can mediate H3K27me2/3, though to a lesser extent. By contrast, it is can fully compensate Ezh2 in mono methylation of H3K27 (Margueron et al., 2008; Shen et al., 2008). Interestingly the PRC2/Ezh1 complex can repress transcription in a mode independent from its HMT catalytic activity (Margueron et al., 2008).

One study has revealed an ability of PRC2 complex to interact with DNA methyltransferases (DNMTs). Knock-down of Ezh2 impairs binding and activity of all known DNMTs on PcG target promoters. This observation indicates interdependence between Polycomb and DNA methylation mediated repressive pathways (Vire et al., 2006).

The H3K27me3 mark which is set by PRC2 is recognised by another Polycomb complex PRC1 (Fischle et al., 2003; Min et al., 2003). In mammals, this binding is mediated by the chromodomain of Cbx proteins (Bernstein et al., 2006b). Additionally,

PRC1 comprises of Polyhomeotic 1/2/3 (Phc1/2/3), Ring1a/Rnf2 and Bmi/Mel18. Ring1a and Rnf2 are two homologues proteins that contain a RING domain with an E3 ubiquitin ligase activity. They mediate mono-ubiquitination of Lysine 119 of histone H2A (H2AK119ub) (de Napoles et al., 2004; Wang et al., 2004). Specificity of this reaction toward H2AK119 is directed by Bmi1 and Mel18, two other homologues components of the PRC1 complex (Buchwald et al., 2006; Elderkin et al., 2007).

Targeting of PcG complexes

Binding of PRC1 to the H3K27me3 mark set by PCR2 supports a model where targeting of PRC1 and establishment of H2AK119ub occurs downstream of PRC2. This order of events was demonstrated on the *Hox* gene cluster in mouse embryonic fibroblasts (Cao et al., 2005) and also on several PcG targets in mouse embryonic stem (ES) cells (Boyer et al., 2006). Consistently, in one-cell embryos lacking PRC2 component Ezh2, PRC1 members are depleted from euchromatic (gene rich) regions (Puschendorf et al., 2008).

However, in the same study PRC1 binding to pericentric heterochromatin was not effected. The authors showed that expression of non-coding transcripts from that region is controlled by Rnf2 (Puschendorf et al., 2008). Similarly, in the later stages of pre-implantation development, loss of Ezh2 is not effecting association of PRC1 to chromatin (Terranova et al., 2008). Also, during the initiation of the X inactivation process in female ES cells, PRC1 is recruited independent of H3K27me3 (Schoeftner et al., 2006).

In agreement with the PRC2-independent targeting of PRC1, genome wide analysis of promoter occupancy in mouse ES cells, identified a group of genes associated with Rnf2 and Phc1 (PRC1), but neither Suz12 and Eed (PRC2) nor H3K27 methylation (Boyer et al., 2006). Taken together, these data indicate that PRC1 acts alone at specific developmental stages and may also have other functions that classical Polycomb mediated gene repression.

Sequence components of PcG targeting mechanisms

In *Drosophila melanogaster* specific sequences called Polycomb/Trithorax respond elements (PRE/TRE) have been identified. They are recognized by DNA

binding subunits of Polycomb and Trithorax complexes (reviewed in (Ringrose and Paro, 2007)). However in mammals such sequences have not been identified yet. Genome wide data also indicates much grater spread of Polycomb group proteins on the repressed domains in mammals, comparing to more localized and PRE restricted binding in flies (Boyer et al., 2006; Lee et al., 2006b; Negre et al., 2006).

There is also a positive correlation between high CpG sequence content and Polycomb/H3K27me3 occupancy (Bernstein et al., 2006a; Mikkelsen et al., 2007; Mohn et al., 2008). In a recent genome wide study aimed in identifying mammalian PREs, Ku and colleagues suggested that it is rather the absence of binding motifs for activating transcription factors that define Polycomb repressive domains (Ku et al., 2008).

Mechanisms of repression by PcG and associated histone marks

The mechanism of gene repression by Polycomb group complexes and H3K27 methylation are not well understood. Several *in vitro* studies showed increased compaction and inhibition of chromatin remodelling and transcription upon association of chromatin fiber with PRC1 complex (Francis et al., 2001; King et al., 2002; Shao et al., 1999). However in *Drosophila,* components of the transcription initiation complex (TBP, TFIIB and TFIIF) were found to co-localise with Polycomb group proteins at repressed PcG target promoters (Breiling et al., 2001). Further, recruitment of Polycomb proteins to a PRE inserted at the promoter of a heat shock gene did not interfere with the RNA Polymerase II (RNAPII) binding, but prevented transcription initiation (Dellino et al., 2004). These reports suggest that *in vivo* chromatin structure on promoters of repressed genes may not differ greatly from active ones. Consistently, genome wide ChIP-chip studies in mouse and human ES cells revealed that the majority of Polycomb repressed genes associated with H3K27me3, harbour H3K4 methylation simultaneously, a configuration that was called bivalent (Azuara et al., 2006; Bernstein et al., 2006a). In another genome wide study, H3K4me3 and RNAPII were detected on promoters of the majority of coding human genes, including PcG targets and corresponding short 5' transcripts were detected (Guenther et al., 2007).

These data suggest that Polycomb mediated repression is not inhibiting the transcription initiation but rather the transcription elongation. Indeed, studies of a PcG target gene *Nkx2.2* in ES cells revealed that repression of this gene is dependent on the presence of the PRC1 subunit Ring1a/Rnf2 (Stock et al., 2007). Upon deletion of both homologues, RNAPII changes its conformation from paused to active and produces high

levels of transcripts. Since the levels of H2A119ub are drastically reduced, this mark is proposed to inhibit activation of RNAPII, presumably through blocking the recruitment of other components of the transcriptional machinery.

Activating H3K4 methylation and SET1/Trithorax complexes

In contrast to the two described enzymes mediating H3K27 methylation, there are at least ten known or predicted H3K4 methyltransferases in mammals. The catalytic SET domains of these proteins are either related to yeast Set1 and Drosophila Trx (SET1 family: Mll1, Mll2, Mll3, Mll4, Set1a and Set1b) or unrelated (Ash1L, Set7/9, Smyd1, Smyd3, Meisetz) (reviewed in (Ruthenburg et al., 2007)). Similarly to Ezh1/2, which is active only in the context of the PRC2 complex, SET1 family enzymes exist in multiprotein complexes. SET1-family complexes share, but are not limited, to three common subunits WDR5, RbPB5 and Ash2L. All these three components are required for full HMT activity both *in vitro* and *in vivo* (Dou et al., 2006). RNAi knock down studies showed that WDR5 and RbPB5 are required for di- and tri- methylation whereas Ash2L is required for tri-methylation of H3K4 (Dou et al., 2006; Steward et al., 2006). Interestingly, the WDR5 subunit shows high binding affinity to di-methylated K4 and has been proposed to present this mark for further methylation by the HMT subunit (Wysocka et al., 2005).

Targeting of SET1/Trithorax complexes

Global reduction of H3K4 methylation via RNAi indicates that SET1 family members account for the bulk of H3K4 methylation in the mammalian cells (Dou et al., 2006; Steward et al., 2006). Within the family the six members seem to have non-overlapping functions. Mutant mice lacking either Mll1, Mll2 or Mll3 show all severe but distinguishable phenotypes (Glaser et al., 2006; Lee et al., 2006a; Yu et al., 1995). Genome wide analysis of Mll1 distribution showed that this enzyme is present on the majority of active promoters (Guenther et al., 2005). However another study indicates selective targeting of Mll1 to only a subset of genes (Milne et al., 2005). Two other members of SET1 family Set1A and Set1B were shown to occupy largely non-overlapping nuclear domains, suggesting that Set1A and Set1B each bind to a unique set of target genes (Lee et al., 2007a).

Selective targets of different H3K4 methyltransferases suggest the existence of specific recruitment mechanisms. Set1A and Set1B complexes contain an additional subunit Wdr82, which interacts with the initiating, but not the elongating form of RNA Polymerase II (RNAPII) (Lee and Skalnik, 2008). Importantly, depletion of Wdr38 results in decrease of H3K4me3 near the transcription start sites, but does not affect RNAPII levels. This data suggest that K4 methylation is a downstream consequence of transcription, possibly establishing a memory of an active state. Additionally, Mll1 has been shown to associate with RNAPII at transcription start sites (Milne et al., 2005). In that study, Mll1 was also detected along gene bodies and depletion of Mll1 resulted in a defect in transcription elongation.

Mll proteins contain also multiple chromatin binding domains that can provide additional targeting mechanisms. One of them is the CxxC methyltransferase homology domain, found in Mll1 and Mll2, which specifically recognises DNA with unmethylated CpG sequences (Ayton et al., 2004; Birke et al., 2002). This association suggests recruitment of Mll proteins to CpG rich promoters which are not DNA methylated. Consistently, in Mll1 knock-down cells, promoter of the *Hoxa9* gene gains DNA methylation (Erfurth et al., 2008). Therefore, Mll1 and Mll2 are likely to have a function in protecting CpG rich promoters from DNA methylation. This may be a more general property of the K4 methylating machinery as a subunit of the Set1A/Set1B complexes, CFP1, also contains a CxxC domain (Lee et al., 2007a).

This activity may be independent of the transcriptional machinery as genome wide study of H3K4 methylation and RNAPII occupancy revealed a number of promoters harboring the mark, but not the polymerase (Guenther et al., 2007). Though, it can not be ruled out that different efficiencies of the antibodies influenced this result. Nevertheless, in another genome wide study H3K4 methylation was found to mark majority of the promoters with high CpG content and to be mutually exclusive with DNA methylation (Weber et al., 2007). In mammalian genomes approximately half of promoters harbour starches of sequences with high CpG content, higher than the majority of the surrounding genome. It is thought that these so-called CpG islands arose during evolution by selective protection of certain sequences from cytosine to thymidine conversion. High rate of such mutations can be explained by methylation of cytosines that occurs in the CpG context and makes them prone for a conversion to thymidine in a reaction of de-amination (Shen et al., 1994). Consistently, CpG islands are devoid of DNA methylation, whereas the majority of CpGs in the genome are methylated.

Therefore, reciprocal localisation of DNA methylation and H3K4 methylation supports a model where H3K4 methylation has a function in protecting DNA from cytosine methylation and therefore contributes to conservation of CpG islands (Weber et al., 2007).

Additional level of sequence specificity of SET1 complexes is thought to be achieved by interaction with sequence specific DNA binding factors ((reviewed in (Ruthenburg et al., 2007)).

Reversibility of Trithorax and Polycomb mediated marks

The complexity of gene regulation by activating and repressing histone modifications has largely increased with the discovery of histone de-methylases. The first identified de-methylase was LSD1 which removes methyl groups from di- and mono methylated H3K4 in the reaction of amine oxidation (Shi et al., 2004). Interestingly, specificity of LSD1 can be altered by specific cofactors. When associated with the androgen receptor LSD1 de-methylates di- and monomethylated H3K9 (Metzger et al., 2005). Following discovery and characterisation of the Jumonji family revealed multiple de-methylases encoded in the mammalian genomes (reviewed in (Agger et al., 2008)). In contrast to LSD1, these enzymes catalyse oxidative demethylation and can also remove tri-methylated states of lysines. De-methylation of H3K4me3 is catalysed by enzymes of the JARID1 family. Jarid1d was found in a complex with Polycomb-like protein Ring6a providing a link between de-methylation of H3K4 and gene repression (Lee et al., 2007b). Depletion of Jarid1d led to increased occupancy of the transcriptional machinery and increased transcription of the promoter of the *Engrailed2* gene. Consistently, another member of the family, Jarid1a, was shown to co-localise with PRC2 at Polycomb target genes in ES cells. This localisation was shown to be dependent on PRC2 (Pasini et al., 2008). In another study, Jarid1a was displaced from *Hox* genes promoters in an ES cells differentiation assay correlating with their activation and H3K4 methylation (Christensen et al., 2007). In a similar way as H3K4 de-methylases have roles in gene silencing, the H3K27 de-methylating enzymes UTX and JMJD3 associate with activating Mll/Set1 complexes (De Santa et al., 2007; Lee et al., 2007c). In an ES cells differentiation assay, UTX was shown to directly bind to the *Hoxb1* locus and to be required for its activation (Agger et al., 2007). Moreover, ES cells with RNAi knockdown of JMJD3 failed to differentiate into neurons (Burgold et al., 2008). These cells failed to upregulate neuronal markers *Nestin, Pax6 and Sox1*, although a

direct effect of JMJD3 on the H3K27me3 levels was demonstrated only for *Nestin*. In the light of these studies histone lysine methylation marks are not any more irreversible and can provide means for dynamic gene regulation.

1.2.3 H3K4 and H3K27 methylation in pluripotent embryonic stem cells and during differentiation

In the recent years, development of micro-arrays (alternatively called chips) and further deep-sequencing technologies allowed researchers not only to study selected genes of interest but to analyse entire genomes. These techniques, in combination with chromatin immuno precipitation (ChIP-chip and ChIP-seq), revealed genome wide localisation maps of many chromatin components, including Polycomb and Trithorax proteins and associated histone modifications (Barski et al., 2007; Boyer et al., 2006; Bracken et al., 2006; Ku et al., 2008; Lee et al., 2006b; Mikkelsen et al., 2007; Mohn et al., 2008; Orford et al., 2008; Pan et al., 2007; Weber et al., 2007; Zhao et al., 2007).

These experiments were mainly preformed in *in vitro* cultured cell lines due to the high number of cells required for ChIP and other biochemical assays. Embryonic stem (ES) cells can be derived form the pre-implantation embryos until the blastocyst stage. *In vitro* cultured ES cells can be transplanted into a blastocyst and contribute to all tissues. Therefore ES cells are used for a model of the pluripotent state. Also, several protocols for *in vitro* differentiation of ES cells have been established providing models of early cell fate commitment during development (reviewed in (Niwa, 2007)).

Bivalency

Using ChIP–chip method, mouse and human ES cells were profiled for components of PRC1, PRC2 and H3K27me3 (Boyer et al., 2006; Ku et al., 2008; Lee et al., 2006b; Mikkelsen et al., 2007; Mohn et al., 2008; Pan et al., 2007; Zhao et al., 2007). These studies revealed that apart from classical *Hox* gene targets PcG proteins repress multiple other genes with functions in development, transcriptional regulation and morphogenesis. Upon deletion of PRC1 components Eed or Suz12 majority of Polycomb targets showed increased levels of expression. These targets were preferentially activated when ES cells were induced for differentiation (Boyer et al., 2006; Lee et al., 2006b). The majority of H3K27me3 modified promoters detected in these studies were further shown to harbour also H3K4 methylation, regardless of their repressed state (Azuara et al., 2006; Bernstein et al., 2006a; Mikkelsen et al., 2007; Mohn et al., 2008;

Pan et al., 2007; Zhao et al., 2007). This can be in part explained by default marking of promoters with high CpG content by H3K4 methylation, discussed above (Weber et al., 2007). Indeed, the presence of so called bivalent marking also highly correlates with the CpG density of the underlying DNA sequence (Fig. 2) (Bernstein et al., 2006a; Mikkelsen et al., 2007; Mohn et al., 2008)

In the primary report (Bernstein et al., 2006a), the bivalent state was reported to be a unique feature of ES cells that is resolved upon differentiation either to “H3K4 methylated only” and therefore active or “H3K27 methylated only” and therefore repressed depending on a lineage. However, following studies identified bivalent promoters also in differentiated cells (Barski et al., 2007; Mikkelsen et al., 2007; Mohn et al., 2008). Mohn and colleagues showed that during neuronal differentiation from ES cells to terminal neurons, bivalent domains were formed not only in ES cells but also at the progenitor stage (Mohn et al., 2008). Several neuronal specific genes acquired such a configuration, which was further resolved to “H3K4me2 only” state as they became activated in the terminal neurons. These data suggest that the poising of promoters for activation by the establishment of bivalent domains is a mechanism operating at multiple differentiation stages.

Bivalent promoters are present in both mouse and human ES cells. This argues that they provide a conserved mechanism of gene regulation (Pan et al., 2007; Zhao et al., 2007). Direct comparison of bivalent genes in mouse and human ES cells revealed that 50% of the targets are shared between these two species (Ku et al., 2008). This shared fraction of genes was over-represented for functions in development and transcriptional regulation, which stresses the functional importance for the Polycomb mediated repression at this class of targets. Importantly, occupancy of the PRC1 component Rnf2 was also measured in this study and showed 60% conservation of the targets in the two species.

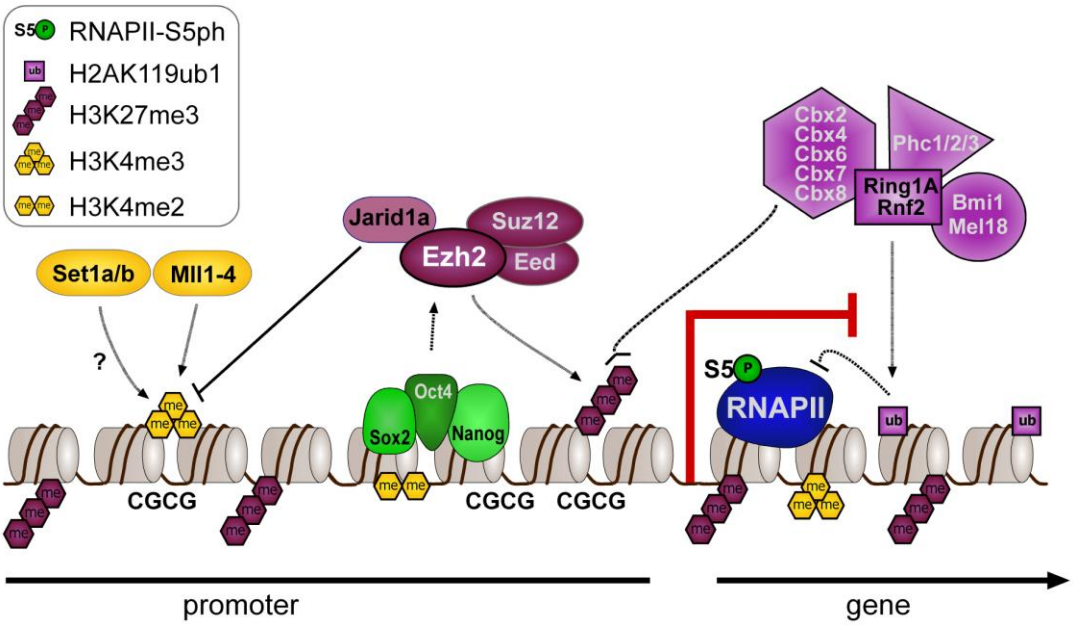


Figure 2. Polycomb mediated repression at bivalent promoters. In mammalian cells, promoters of developmental regulators are marked by both H3K4me2/3 and H3K27me3, and are therefore termed "bivalent". Bivalency correlates strongly with high GC density. Despite the presence of H3K4me2/3, which is likely mediated by the Mll and/or Set1a/b enzymes, bivalent genes are largely repressed by Polycomb mediated mechanisms. The pluripotency transcription factors Sox2, Oct4 and Nanog co-occupy a large fraction of Polycomb-bound genes. PRC2-mediated H3K27me3 provides a binding site for PRC1, which in turn mediates monoubiquitination of H2AK119. Jarid1a targeted by PRC2 downregulates H3K4me2/3 levels. The initiating form of RNA polymerase II (RNAP-S5P) is present at bivalent genes but is arrested before elongation, presumably by H2AK119ub1 inhibiting recruitment of the remodeling complexes.
(adapted from (Hublitz et al., 2009))

Polycomb group proteins and pluripotency

There is no strong data supporting an essential role of PRC2 and H3K27methylation for pluripotency of embryonic stem cells, even though there are curtail developmental regulators of somatic differentiation among their targets. Deletion studies in mouse ES cells showed that upon deletion of Eed, Polycomb target genes are upregulated but nevertheless ES cells can be maintained (Chamberlain et al., 2008; Montgomery et al., 2005). Also Suz12 deficient ES cells could be derived and kept in culture but showed defects in differentiation (Pasini et al., 2007).

In ES cells, Oct4, Nanog and Sox2 form a core regulatory network, required for maintaining pluripotency (reviewed in (Niwa, 2007)) and occupy a large group of important developmental regulators (Boyer et al., 2006). Most of these genes are also targets of PRC2 and harbour H3K27me3 mark, indicating cross talk between PcG and

the pluripotency factors (Lee et al., 2006b). Importantly, in both mentioned above PRC2 mutant cell lines, transcription factors Oct4 and Nanog were still expressed and present at chromatin as detected by stainings (Chamberlain et al., 2008; Pasini et al., 2007). These data indicate that repression of PcG target genes is not crucial for the self renewing capacity of the ES cells and that pluripotency factors can maintain them alone.

Nevertheless, the importance of PRC2 and H3K27 methylation mediated repression is evident based on early lethality of PRC2 mutant embryos (Faust et al., 1995; O'Carroll et al., 2001; Pasini et al., 2004). Interestingly, Eed mutant ES cells can contribute to chimeric embryos, but were not detected in all organs (Morin-Kensicki et al., 2001). Therefore, the key developmental function of PRC2 may be the regulation of the differentiation. Correct establishment of bivalent domains at the pluripotent pre-implantation stage may be crucial for this process. However, subsequent maintenance of repression and establishment of new bivalency is equally important (Ezhkova et al., 2009; Mikkelsen et al., 2007; Mohn et al., 2008). Finally, the removal of H3K27me3, by JMJD3, from the repressed target genes and the resolution of bivalent domains are required for the neuronal differentiation, emphasising the significance of the H3K27me3 mediated repression (Burgold et al., 2008).

By contrast, deletion studies of homologues PRC1 components Ring1A and Rnf2 showed that they are required for maintenance of ES cells (Endoh et al., 2008; van der Stoop et al., 2008). The severity of the phenotype varies between the studies. One mutant cell line shows defect in maintaining pluripotency and differentiates spontaneously, whereas the other one does not survive the deletion (Endoh et al., 2008; van der Stoop et al., 2008, respectively). A link between PRC1 and the ES cells pluripotency circuit was also analysed using cells deficient for Oct4. Oct4 was found to be necessary for Ring1a and Rnf2 recruitment to target genes, confirming that pluripotency factors act upstream of Polycomb complexes (Endoh et al., 2008).

Patterns of H3K4 methylation in pluripotent and differentiated cells

Multiple genome wide studies showed that H3K4 methylation marks the majority of genes both in pluripotent and differentiated cells (Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007; Mohn et al., 2008; Orford et al., 2008; Pan et al., 2007; Weber et al., 2007; Zhao et al., 2007). However, the role of this mark seems to depend on the promoter sequence. As described above, promoters with high CpG content are all

marked by H3K4 methylation. Many of these genes have house keeping function, and for these genes, the level of H3K4 methylation correlates with expression (Weber et al., 2007). A class of developmental regulators that is also marked by H3K27 methylation shows relatively lower levels of H3K4 marking (Zhao et al., 2007). On promoters with low CpG content, the presence of H3K4 methylation correlates with expression and the majority of these genes have tissue specific functions (Weber et al., 2007).

H3K4 methylation localises around transcription start site (TSS), with slight differences between di- (H3K4me2) and tri- (H3K4me3) methylation states. H3K4me2 shows broader distribution, whereas H3K4me3 shows a clear peak downstream of the TSS. Both marks are depleted from the TSS itself, likely to reflect nucleosomal depletion on active genes. On the majority of genes both H3K4me2 and H3K4me3 are present (Barski et al., 2007; Orford et al., 2008). However, the study of Orford and colleagues indicates a specific role for H3K4me2 on promoters with low CpG content. H3K4me2 and H3K4me3 occupancy was analysed in isolated populations of cells at subsequent stages of erythroid differentiation. At a progenitor stage H3K4me2 alone is detected at a subset of inactive promoters with low CpG content. Some of these genes get activated and gain H3K4me3 in differentiated cells, whereas other remain silent and loose the mark, in accordance with the lineage specific expression. Therefore, the presence of H3K4me2 in the absence of H3K4me3 is proposed to represent a poised state (Orford et al., 2008).

Similar to H3K4 methylation H3K27 di and tri methylation states co-localise largely in the mammalian genome and are both associated with repressed genes (Barski et al., 2007). Generally, H3K27 methylation is distributed over a broader region around the TSS than H3K4 methylation. Interestingly, mono methylation of H3K27 has been found on active promoters (Barski et al., 2007).

1.3 Histone methylation marks as a part of the cellular memory

Based on the described H3K4 and H3K27 methylation patterns in ES cells and during differentiation there is a potential role for these marks in the maintenance of specific expression states. Close association with DNA and relative stability of lysine methylation makes it a candidate for a carrier of heritable epigenetic information. To fulfil this function, patterns of histone modifications should be propagated during the cell division.

1.3.1 Propagation of epigenetic marks during replication

During the S phase of the cell cycle, DNA is replicated in a semi-conservative manner (Watson and Crick, 1953). DNA Polymerases synthesize DNA on both strands, in a continuous mode on the leading strand and discontinuous on the lagging strand. They are assisted by the DNA processivity factors PCNA (proliferating cell nuclear antigen) that encircle DNA and greatly increase efficiency of the polymerases. PCNA interacts directly with a number of proteins involved in many different cellular processes, including chromatin remodeling and modifying enzymes (reviewed in (Moldovan et al., 2007)).

Replication of DNA methylation patterns

Similarly to the DNA, DNA methylation marks are replicated in a semi-conservative manner. This is possible due to the symmetrical distribution of methyl groups on two cytosines, in the CpG di-nucleotides. Each strand inherits therefore a template of DNA methylation pattern, which is immediately re-established on both newly replicated strands. This process is mediated by Dnmt1 (DNA methyltransferase1) which interacts with the hemi-methylated DNA (CpG with one of two symmetrical methyl groups) and with PCNA (Chuang et al., 1997). However, the mechanism of inheritance of histones and their marks, during replication, is not well understood.

Supply of histones during the replication

In contrast to multiple histone variants, which are incorporated during specific chromatin processes and are localised on discrete sites in the genome, canonical histones are those which contribute to the vast majority of the nucleosomes and are deposited on the newly replicated DNA. Histone H3 has four major variants: two canonical H3.1 and H3.2, H3.3 deposited on the sites of active transcription (Ahmad and Henikoff, 2002) and CENP-A associated with centromeric sequences (Sullivan et al., 1994).

Just before the entry into the S phase the level of canonical histone transcripts rapidly increases (DeLisle et al., 1983). The onset of expression of histone genes is tightly linked to the cell cycle and depends on the phosphorylation of the nuclear factor NPAT (nuclear protein, ataxia-telangiectasia locus) by Cycline E-CDK2 (cycline-dependent kinase 2) (Ma et al., 2000; Zhao et al., 2000). The levels of histone proteins are further finely regulated during mRNA processing, translation and mRNA degradation.

All these modes of regulation have been shown to depend on a unique stem loop structure at the 3' end of the histone transcripts (Gallie et al., 1996; Pandey and Marzluff, 1987). This 26 nt long motif provides a binding site for SLBP (stem-loop binding protein) that participates directly in translation and degradation of histone transcripts (Cakmakci et al., 2008; Mullen and Marzluff, 2008). The mRNAs of replication-dependent canonical histones are the only nuclear transcripts that lack the poly-A tail. The stem loop structure, which is present instead at the 3' end, might have evolved to allow a coordinated regulation of histone levels, crucial for the successful completion of the S phase (reviewed in (Marzluff et al., 2008)).

New nucleosomes are assembled by association of DNA with a tetramer of histones H3.1 and H4 (two of each), which also exist in an intermediate H3.1–H4 dimeric form, followed by the incorporation of two H2A and H2B dimers. As H2A-H2B dimers are dynamically exchanged also outside of the S phase, the H3 and H4 status is probably more directly dependent on the replication (Kimura and Cook, 2001).

Histones are recruited to the replicating chromatin in a complex with specific binding factors – histone chaperons. Histones H3.1 and H4 are deposited by a coordinated action of two chaperons CAF1 (chromatin assembly factor 1) and ASF1 (antisilencing function 1), which both interact with the replication machinery (Groth et al., 2007; Shibahara and Stillman, 1999). It is not clear if primarily H3.1 and H4 are supplied as dimers or as tetramers. H3 and H4 exist as a stable tetramer in solution, also in the absence of DNA (Baxevanis et al., 1991). However *in vivo,* in a complex with histone chaperons, H3 and H4 are detected as dimers (Tagami et al., 2004). H2A and H2B dimers are deposited by another histone chaperone NAP1 (reviewed in (Zlatanova et al., 2007)).

Replication of histone methylation patterns

There are several models proposed for the distribution of old versus newly assembled nucleosomes during the replication. They can be either randomly distributed on the two daughter strands, or if H3-H4 dimers are released after passing of the replication fork, mixed "new-old" nucleosomes may be assembled. The third possibility would be, that one of the daughter strands inherits all of the old nucleosomes whereas the other one only the new ones (reviewed in (Probst et al., 2009)).

In any of these models, post translational modifications on old nucleosomes can serve as a template for modification of new histones, either in *cis* (inter- or intra-nucleosomal) or in *trans* (between the strands).

Newly synthesized histone H3 is predominantly supplied in a not modified state (Loyola et al., 2006). In contrast, histone H4 carries a conserved set of acetylated lysines 5 and 12. These marks are removed upon incorporation of the H4 into nucleosomes (Loyola et al., 2006; Sobel et al., 1995). The only modification found on H3.1/2, in a non-nucleosomal fraction of histones, was H3K9me1, whereas H3.3 was acetylated at that residue. Importantly, H3K9me3 was not detected on any of H3 variants, suggesting that higher methylation states are acquired in the context of chromatin. Consistently, authors showed preference of Suv39H1 (H3K9 HMT) for a mono methylated substrate. In this model acetylation of K9 at histone H3.3 would prevent it from the acquisition of the repressive mark (Loyola et al., 2006).

At pericentric heterochromatin, Suv39h1/2 mediated H3K9me3 recruits binding factor HP1. HP1 interacts with Suv39h1/2 leading to new methylation events and a self-reinforcing spreading of heterochromatin (Bannister et al., 2001; Lachner et al., 2001). The same mechanism could operate to re-establish H3K9me3 pattern after/during the replication. Both inherited H3K9me3, on old histones, and H3K9me1, on newly incorporated histones, could be used as temples (Loyola et al., 2006). Indeed, HP1 was found to interact with H3.1/2 chaperone CAF1 and to be required for the S phase progression (Murzina et al., 1999; Quivy et al., 2008). Other H3K9 HMTs, SETDB1 (also known as ESET) and G9a are also interacting with the replication machinery (Esteve et al., 2006; Sarraf and Stancheva, 2004). These enzymes are responsible for setting the repressive H3K9 methylation at promoters (Rice et al., 2003; Wang et al., 2003). Their activity along with the replication fork can act both, to propagate H3K9me2/3 mark on repressed genes and also to create an intermediate H3K9me1 substrate for Suv39H1 (Esteve et al., 2006; Loyola et al., 2009; Sarraf and Stancheva, 2004). Importantly, G9a was shown to interact with Dnmt1 in the course of replication, indicating that the deposition of new marks on histones and on DNA is a coordinated process (Esteve et al., 2006).

A similar mechanism of the modification spreading, along with the replication, was recently proposed for H3K27me3. Hansen and colleagues showed that PRC2 complex binds its own product H3K27me3 and co localises with the sites of on-going transcription. Moreover, they showed that PRC2 components are required to transmit

H3K27me3 and the associated repressed state, of a reporter gene, to the next cell generation. This observation leads to a hypothesis, that newly deposited histones are targeted by Ezh2, present at a neighbouring nucleosome (Hansen et al., 2008). The model offers an attractive mechanism for the maintenance of Polycomb mediated repression, which is observed during development and differentiation.

1.3.2 Incorporation of new histones outside of the S phase

There are multiple variants of histones H3, H2A and H2B that are incorporated into the chromatin outside of the S-phase. They are associated either with a specific chromatin processes like transcription and DNA repair, or appear at specific phases of development. Several of these variants are playing roles in the germ line differentiation, which is described in the next chapter. Importantly, as these histones can not be incorporated during the replication, their transmission to the next generation of cells is challenged.

At the sites of ongoing transcription, H3 variant, H3.1/2, is exchange for H3.3 (Ahmad and Henikoff, 2002, Janicki et al., 2004, Wirbelauer et al., 2005). Two canonical H3 variants H3.1 and H3.2 differ only in one amino acid. In contrast, H3.3 differs at 4 and 5 amino acid positions, compared with H3.2 and H3.1 respectively. Three of these substitutions, lying in the core domain of the histone, determine different deposition pathways (Ahmad and Henikoff, 2002). H3.1/2 is deposited by the replication dependent pathway involving CAF1 histone chaperone, whereas H3.3 by the replication independent pathway, involving HIRA chaperone (Tagami et al., 2004).

H3.3 is retained on active genes during the mitotic shutdown of transcription, suggesting that genes marked by H3.3 will be reactivated in the next cell generation (Chow et al., 2005). However, since H3.3 can not be propagated during the replication, inheritance of active state must involve H3.1/2. Studies of isolated oligonucleosomes revealed that H3.3 co-exists with H3.1/2 on the neighbouring nucleosomes and they both carry histone marks associated with gene activation (Loyola et al., 2006). This data suggests that even after “dilution” of H3.3 with H3.1/2, the active state of a gene is preserved. It remains a question whether and when H3.1/2 is exchange back to H3.3 after the replication and if this process is coupled with the transcription. It is also not clear when and how is the spreading of active marks achieved.

In a recent study Ng and colleagues addressed these questions by analysing the mechanism of an uncompleted reprogramming of the genome after the nuclear transfer

(NT) in *Xenopus* (Ng and Gurdon, 2008). They analysed the chromatin status of the *MyoD* locus, which show aberrantly high expression in cloned embryos. The gene is normally highly expressed in endodermal cells, from which the nucleus was taken for the NT. By injecting a tagged histone H3.3, following the NT, they could show that the aberrant expression of *MyoD* correlates with the incorporation of the histone H3.3 at the promoter of *MyoD* at the blastula stage. Furthermore, injection of the increased amounts of H3.3 led to a higher proportion of embryos (53% vs. 73%) with aberrant *MyoD* expression, arguing that H3.3 has a role in the maintenance of the memory of an active state.

To prove a role of H3.3 not only in the maintenance but also in the transmission of this memory, the authors overexpressed a mutated version of H3.3 in the embryos from which the donor nuclei were taken. Substitution of the lysine 4 with the glutamic acid did not affect the transfer efficiency, but led to a reduced number of embryos with the aberrant expression of *MyoD*. These data indicate that lysine 4 of H3.3 may act as an epigenetic mark for the inheritance of an active transcriptional state. Furthermore, the authors propose a model in which the methylation of lysine 4 serves as a mark for the recruitment of new H3.3 and therefore contribute to the spreading of an active state (Ng and Gurdon, 2008). However, as pointed out in the preview to the (Ng and Gurdon, 2008) publication, the observations derived from the experiments with the exogenous tagged histone H3.3 may not illustrate the physiological situation (Lacoste and Almouzni, 2008).

1.4 Epigenetic events in the male germ line

1.4.1 Primordial germ cells – a balance between pluripotency and germ cell commitment

Specification of primordial germ cells (PGCs)

Primordial germ cells (PGC) are the founder cells of the germ cell lineage. Through oogenesis and spermatogenesis they give rise to mature gametes. In *D.melanogaster* and *C.elegans*, PGCs are specified already in the zygote by asymmetrical distribution of maternal factors in the cytoplasm (so-called germ plasm) (reviewed in (Seydoux and Braun, 2006)). In mice PGCs emerge as a group of around 40 cells in the extra-embryonic mesoderm at embryonic day E7.25 (Ginsburg et al.,

1990). Murine PGCs develop from undifferentiated embryonic cells, induced to become germ cells by extra-cellular signals. A coordinated pattern of the signalling molecules BMP4, BMP8b and WNT3, provides a defined zone in the embryo where precursors of PGCs arise at the embryonic day E5.5 (Ohinata et al., 2009). This specific pattern of morphogens is necessary for germ cell specification. Knock-out animals lacking either one of these factors do not develop the germ line (Lawson et al., 1999; Liu et al., 1999; Ying et al., 2000). In response to BMP signalling, precursors of PGCs start to express Blimp1 (Prdm1) and Prdm14, two master regulators of PGC specification (Ohinata et al., 2005; Yamaji et al., 2008).

The major event in the establishment of the germ cell fate is the repression of the mesodermal expression program. Deletion studies showed that Blimp1 plays a major role in repressing the transcription of *Hox1a*, *Hoxb1, Evx1* and multiple other somatic determinants (Kurimoto et al., 2008; Ohinata et al., 2005). The repression of the somatic program is a gradual process and mesodermal markers like *T (brachyury)* and *Fgf8* are initially expressed in PGCs, but subsequently also get down regulated (E8.25) (Saitou et al., 2002; Yabuta et al., 2006). The mechanism by which Blimp1 represses transcription is only partially understood. Although Blimp1 has a histone methyltransferase motif, a SET domain, it does not exhibit any methyltransferase activity. It is proposed that Blimp1 acts on chromatin through its interaction partners. In a recent study, a previously unknown Blimp1-Prmt5 complex has been identified in primordial germ cells (Ancelin et al., 2006). Prmt5 is a histone-arginine methyltransferase that mediates symmetrical di-methylation of arginine 3 of histones H2A and H4. This modification was detected in PGCs between days E8.5 and E11.5. Using a ChIP assay, the complex was detected on one of the repressed genes *Dhx38* at the day E10.5 (Ancelin et al., 2006). This complex was translocated to the cytoplasm at E11.5, concomitant with the activation of *Dhx38*. Even though this study partially reveals the mechanisms of Blimp1 mediated repression, it remains a question what the mode of action of Blimp1 at the time of germ cells specification at day E7.5 is.

Coupled with the repression of mesodermal-specific genes, there is up-regulation of other genes, including *Stella* (*Dppa3*), *Fragilis, Tnap, Kit*. Importantly, the main pluripotency factors *Nanog* and *Sox2* get up-regulated (Yabuta et al., 2006). Expression of *Oct4,* which is progressively repressed in the embryo, by embryonic day E7.5, is exclusively maintained in PGCs (Yeom et al., 1996). Continued *Oct4* expression is

required for the survival of PGCs as *Oct4* null germ cells undergo apoptosis (Kehler et al., 2004). Thus, unlike the surrounding somatic cells, which gradually restrict their developmental potential, PGCs partially re-establish pluripotency. Consistently, pluripotent embryonic germ cells (EG) can be derived *in vitro* from PGC isolated between days E8.5 and E11.5 (Durcova-Hills et al., 2006; Matsui et al., 1992). EG cells, like ES cells can contribute to chimeras when injected into a blastocyst (Matsui et al., 1992). Yet, during their derivation EG cells are reprogrammed by signalling factors like LIF and FGF-2. PGCs themselves are not equally pluripotent and can not contribute to chimeras (Durcova-Hills et al., 2006).

Epigenetic reprogramming of migrating PGCs

Following their specifications around embryonic day 7.25, PGCs proliferate and migrate to the developing gonads between days E7.5 and E10.5. During this migratory phase PGCs undergo extensive epigenetic reprogramming (Hajkova et al., 2008; Seki et al., 2005; Seki et al., 2007).

There is a global loss of the repressive mark H3K9me2 from day E7.5 onwards (Hajkova et al., 2008; Seki et al., 2005). Several JmjC domain histone de-methylates were screened for their expression during this process, but none of them showed PGC exclusive expression. The loss of K9 methylation can be attributed to the PGC specific down-regulation of Glp, a histone methyltransferase which in a complex with G9a mediates H3K9 dimethylation (Seki et al., 2007). Also, an increase of H3K9 acetylation was observed during that time. Therefore, a competition between the acetylation and methylation may enhance the erasure of H3K9me2. Although, it can not be ruled out that the acetylation follows the removal of methylation from H3K9 (Hajkova et al., 2008; Seki et al., 2005). Global DNA methylation is another repressive mark that is disappearing progressively in migrating primordial germ cells. At the same time, the maintenance DNA methyltransferase Dnmt11, is transiently down regulated and the *de-novo* DNA methyltransferases Dnmt3b and Dnmt3a are suppressed and absent, respectively (Seki et al., 2005). By contrast, H3K27me3 is up-regulated from day E8.25 onwards. It possibly compensates the erasure of H3K9me2 and DNA methylation (Hajkova et al., 2008; Seki et al., 2005). H3K27me3 in PGCs is probably mediated by Ezh2, as this enzyme is highly expressed at that time (Yabuta et al., 2006). High levels of H3K27 methylation resemble the situation in ES cells and may contribute to the acquisition of pluripotency. However, a comparison of the modification status at a single gene level is

needed to confirm this hypothesis. The development of sensitive ChIP-seq protocols may soon provide methods to address this question.

At the time of epigenetic reprogramming, a transient repression of RNAPII dependent transcription is observed (Seki et al., 2007). The period of transcriptional quiescence corresponds to the gap between the erasure of H3K9me2 and the establishment of H3K27me3 and may provide a mechanism to prevent miss-regulation of gene expression. Interestingly, the level of H3K4 methylation is continuously high in migrating PGCs. Therefore the down regulation of transcription is neither H3K4 dependent nor causes a decrease of this active chromatin mark (Seki et al., 2007).

The epigenetic reprogramming is not entirely synchronised between the PGCs. The proliferation rates vary between the cells as well. It was observed that at a certain time point during the migration, individual cells enter a prolonged period of the G2 phase. This period may provide a window of opportunity for the reprogramming events to take place (Seki et al., 2007).

X chromosome reactivation

The reprogramming events described so far are taking place in both male and female embryos. However, in the female germ line an additional event of X chromosome re-activation is observed (de Napoles et al., 2007; Monk and McLaren, 1981; Sugimoto and Abe, 2007). In female mammals one of two X chromosomes is inactivated during early embryonic development in order to compensate the gene dosage difference between XY males and XX females. This process is mediated by epigenetic mechanisms and is initiated by the expression of the non-coding RNA *Xist*, followed by the establishment of silent chromatin characterised by high levels of H3K27me3, H2AK119ub, H4K20me1. The silencing is maintained by subsequent changes in chromatin, including the incorporation of the H2A variant macroH2A, hypoacetylation of H4 and DNA methylation of CpG islands (reviewed in (Wutz and Gribnau, 2007)).

At the time of female PGCs specification, inactivation of the X has already been initiated in the embryo. Starting from embryonic day E7 the process of re-activation is observed in PGCs (de Napoles et al., 2007; Sugimoto and Abe, 2007). Gradual re-establishment of equal expression levels at both X chromosomes is observed, starting from embryonic day E7.75 (Sigimoto and Abe, 2007). On other hand, expression of *Xist* RNA, is gradually ceasing from embryonic day E7 onwards (Sugimoto and Abe, 2007). Further, PRC2 complex components and H3K27me3 are selectively disappearing from

the inactive X chromosome, while they remain at high levels on other chromosomes (de Napoles et al., 2007).

Epigenetic events in gonadal PGCs

After their migration through the embryo, PGCs reach the genital ridges (developing gonads) around embryonic day E10.5. Shortly after that, another step of epigenetic reprogramming takes place. It has been proposed that the described steps of PGC re-programming are linked to the partial re-establishment of the pluripotency, while a second step prepares the PGCs to enter the sex specific differentiation programs (Hajkova et al., 2008).

The crucial event occurring at that time is the erasure of the DNA methylation at imprinted loci, between E11.5 and E12.5 (Hajkova et al., 2002). Based on stainings with an antibody against methylated cytosine, the process of global DNA de-methylation initiated during the germ cells migration is continued in post migratory cells (Seki et al., 2005). However, several retrotransposon sequences from the IAP family have been shown to retain high DNA methylation (Hajkova et al., 2002; Lane et al., 2003). Since the imprint erasure at each locus is a rapid process that is completed within one day of development, it is proposed to be realized by an active demethylation (Hajkova et al., 2002). The enzymes that mediate this reaction are currently unknown. The process might be enhanced by the absence of *de-novo* methyltransferases Dnmt3a and Dnmt3b, which are not detectable in the nuclei of PGCs at the time of imprint erasure (Hajkova et al., 2002).

At the time of imprint erasures, several changes in global chromatin organisation and histone modification status have been observed (Hajkova et al., 2008). At embryonic day E11.5, there is a rapid loss of histone H1, followed by the loss of H3K9me3 and downregulation of H3K27me3. Moreover, global re-arrangement of the heterochromatin is observed. Chromocenters, normally visualised by intensively stained DAPI foci, are not easily detectable any more and localise at the nuclear periphery. HP1 and Cbx2, which are normally associated with constitutive and facultative heterochromatin, respectively, are de-localised from the heterochromatin. The restructuring of chromatin is transient, with most PGCs regaining the typical modification status and the heterochromatin arrangement by the embryonic day E12.5 (Hajkova et al., 2008).

The observed loss/downregulation of histone modification is proposed to occur through a histone replacement mechanism, rather then by an active removal of the marks. This model is supported by a concomitant disappearance of the H2A variant H2A.Z at the time of chromatin restructuring. Additionally, high levels of histone chaperons HIRA and NAP-1 are detected, whereas the CAF-1 subunit p150 is depleted from the cell nucleus. Based on its roles in other cell types and its *in-vitro* interactions, NAP-1 may be responsible for the displacement of H2A.Z-H2B dimers and for the removal of H1 (Kepert et al., 2005; Levchenko and Jackson, 2004; Park et al., 2005). Taken together these data imply that pathways of replication independent histone replacement are highly active at the time of chromatin restructuring (Hajkova et al., 2008).

Germ line sex determination

At E12.5, the number of germ cells reaches around 2000. At this time point the gonads of male and female become physically distinguishable. The choice to commit to either germ line is made in PGCs between embryonic day 11.5 and E12.5 in male and 12.5 and 13.5 in female (Adams and McLaren, 2002). Organ culture experiments showed that the commitment of PGCs is dependent on the sex of the gonadal somatic cells, rather than the sex of PGCs: XY PGCs can develop as oocytes when aggregated with a developing ovary and XX PGCs can develop as prospermatogonia when aggregated with a developing testis (Adams and McLaren, 2002).

In most mammals, somatic sex is determined by the presence or absence of *Sry*, a dominant male-determining gene located on the Y chromosome. In mice, *Sry* is expressed between embryonic day E10.5 and E12.0 in the supporting cells of the gonad, before the sexual differentiation of the gonad has occurred. Somatic cells in the gonad are also bi-potential and in response to *Sry* follow rather a male than a female developmental pathway. The supporting cell lineage differentiates into Sertoli rather than granulose cells, and the steroidogenic precursor cells differentiate into Leydig rather than theca cells (reviewed in (Brennan and Capel, 2004)).

Following sex specification, retinoic acid (RA) signaling stimulates female germ cells to enter meiosis. RA activates *Stra8*, which is required for meiosis initiation in both males and females (Anderson et al., 2008; Baltus et al., 2006). In males, however, the entry into meiosis is suppressed until few days after birth. The repression of RA in the

developing testis is achieved through the enzyme CYBP26B1, which degrades RA (Bowles et al., 2006). In the absence of RA and no concomitant *Stra8* activation, male germ cells, which are still highly proliferating at the day E12.5, enter mitotic arrest in G1 cell cycle phase (McLaren, 1984). In human PGCs, a switch from the proliferative phase to the mitotic arrest is observed in an unsynchronized manner between weeks 9 and 18 of the fetal development (Gaskell et al., 2004).

1.4.2 Spermatogenesis

G1-arrested male germ cells are called pro-spermatogonia or gonocytes. Upon birth gonocytes migrate from their original central position in the semniferous tubules towards the periphery and populate an area at the basement membrane around post natal day 3 to 6 (Bellve et al., 1977). These cells form a pool of undifferentiated, self-renewing spermatogonia. Besides self-renewing as a population, undifferentiated spermatogonia generate differentiating spermatogonia, which then differentiate into meiotic spermatocytes, haploid spermatids and spermatozoa. In the semniferous tubules, all types of spermatogonia are localized on the peripheral basement membrane, and the subsequent cell types are arranged in a sequential order towards the lumen (reviewed in (de Rooij, 1998)) (Fig. 3).

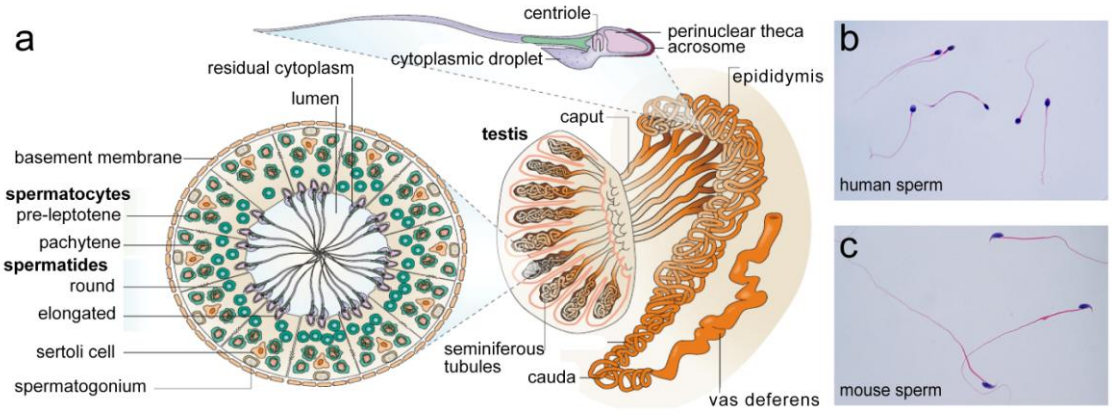


Figure 3. Spermatogenesis. (a) A schematic representation of human testis and a cross section of a semniferous tubule. Cells from the various stages of the spermatogenic pathway are depicted, from the spermatogonial stem cell to the elongating spermatid. In the maturing spermatid, most of the cytoplasm is extruded as a cytoplasmic droplet. The sperm are matured through a caput-to-cauda gradient of RNases, glycosidases and proteases, where approximately 80% of the sperm achieve competence for motility. In humans this is a continual 10-week process. (adapted from (Krawetz, 2005)) **(b)** Mature human ejaculate spermatozoa. **(c)** Mature mouse epididymal spermatozoa**.**

Spermatogonial stem cells

Undifferentiated, self-renewing spermatogonia, so called spermatogonial stem cells, retain expression of stem cell markers like *Oct4, Plzf, Gdnf.* They all play an essential role in the maintenance of spermatogonial stem cells. Their expression persists with various dynamics until the entry into meiosis (Buaas et al., 2004; Costoya et al., 2004; Meng et al., 2000; Pesce et al., 1998). Transition of undifferentiated spermatogonia into differentiating spermatogonia is marked by the expression of *Kit*, also known as *c-Kit*, a receptor tyrosine kinase. As described, *Kit* is a marker of PGCs at the time of their specification and migration to the gonads. Interestingly, from embryonic day E15 to day 3 after birth, *Kit* expression is reduced, coinciding with the period of germ cell quiescence. Expression of *Kit* is re-established in differentiating spermatogonia and continues to be expressed until the onset of meiosis (Prabhu et al., 2006). Multiple studies have shown that *Kit* receptor expression and interaction with Kit ligand (*Kitl*) are essential for both the survival of PGCs and for the progression of spermatogenesis (reviewed in (Mithraprabhu and Loveland, 2009)).

Interestingly, in the first wave of spermatogenesis, gonocytes proceed directly into the differentiating, *Kit* positive spermatogonia. It has been proposed that two specialized niches are established in the neonatal testis – one supporting stem cell renewal and another directing differentiation (Yoshida et al., 2006).

The stem niche established at time of birth is sustaining proliferation and maintenance of spermatogonial stem cells throughout adulthood. Pluripotent germ line stem cells (GS) can be derived from proliferating spermatogonia in mouse and human. These GS cells can be kept in culture and similarly to ES cells contribute to various cell lineages of chimeras after injecting into a blastocyst (Conrad et al., 2008; Guan et al., 2006; Kanatsu-Shinohara et al., 2004).

In the light of the studies on chromatin dynamics in ES cells and during their differentiation, it is a vital question, how similar the epigenetic processes involved in the spermatogonial stem cell maintenance and their differentiation are. Induced deletion of the Mll2 H3K4 methyltransferase in adult mice leads to male and female infertility (Glaser et al., 2009). In males, a developmental block in the differentiation of spermatogonial stem cells was observed, with no effect on the self renewing stem cell pool. A comparable differentiation defect was observed in the ES cells lacking Mll2

(Lubitz et al., 2007). This data suggests that similar epigenetic mechanisms control the differentiation capacity of ES cells and spermatogonia.

During spermatogenesis in *Drosophila,* a progression from spermatogonia to spermatocytes is associated with the binding of testis specific TAFs (TBP associated factors) to genes, involved in spermatid differentiation. It has been proposed that TAFs activate target gene expression in part by counteracting Polycomb repression (Chen et al., 2005). Chromatin immunoprecipitation and deletion studies revealed that testis specific TAFs bind to target promoters, reduce binding of PRC1 components, and promote local accumulation of H3K4me3, a mark of Trithorax action. Testis TAFs also promote relocalization of PRC1 to the nucleoli in spermatocytes (Chen et al., 2005).

Both described reports indicate that Polycomb and Trithorax mediated gene regulation plays a role in the progression from spermatogonial stem cells to the differentiation phases of spermatogenesis.

Re-establishment of paternal imprints

DNA methylation on paternally imprinted loci is progressively established in quiescence germ cells between embryonic day E14.5 and the newborn stage. At that time the *de-novo* methyltransferases Dnmt3a and Dnmt3b are expressed together with an associated protein Dnmt3l (Dnmt3-like). Dnmt3l is related in sequence to Dnmt3a and Dnmt3b but lacks enzymatic activity. It is expressed in germ cells specifically at the time when the *de novo* methylation occurs (Bourc'his and Bestor, 2004). In a complex with either Dnmt3a or Dnmt3b, it stimulates their activity (Suetake et al., 2004). Germline specific knockout studies indicate that Dnmt3a plays a central role in the *de-novo* methylation of all paternally imprinted loci – *Rasgrf1*, *Dlk1-Gtl2, Igf2-H19.* Moreover, Dnmt3b is required for the methylation of *Rasgrf1* (Kaneda et al., 2004; Kato et al., 2007). Dnmt3L contributes to the methylation of all three loci (Bourc'his and Bestor, 2004; Kato et al., 2007; Webster et al., 2005). Interestingly, *de novo* DNA methylation of the maternal and paternal alleles of the paternally imprinted *Igf2-H19* locus occurs at different time points during spermatogenesis. Despite that the inherited DNA methylation marks at the ICRs had been erased in primordial germ cells. These data indicate that the two parental alleles are not equal at the time point of DNA methylation acquisition. They suggest that they possess differential epigenetic marks other than DNA methylation (Davis et al., 2000). However, complete DNA methylation on all paternally imprinted ICRs was observed at birth and no allelic differences have been reported (Kato et al.,

2007). This discrepancy may be caused by a different mouse strains used for the experiments.

It is currently not known how *de-novo* methylation enzymes are targeted to the imprinted loci. Further, the mechanisms promoting the methylation of paternal ICRs and protecting the maternal ones are unknown.

CTCF is a zinc finger protein that binds with high affinity to DNA. It is ubiquitously expressed in somatic cells. Through its insulator function, CTCF mediates intra- and inter-chromosomal contacts. It has been implicated in many cellular processes and is anticipated to have a major role in global organization of chromatin architecture (Reviewed in (Phillips and Corces, 2009)). Among many target sequences, CTCF is binding to the unmethylated ICR of the *Igf2-H19* locus on the female allele. It has been shown to play a major role in protecting this ICR from DNA methylation during oogenesis. Both, introduction of point mutations which prevent CTCF binding, and the RNAi downregulation of CTCF led to aberrant DNA methylation of the *Igf2-H19* ICR in the female germline (Fedoriw et al., 2004; Pant et al., 2003).

A testis specific paralog of CTCF, CTCFL (named also BORIS) is expressed in developing male germ cells (Loukinov et al., 2002). The expression of CTCFL starts in mitotically arrested gonocytes at embryonic day E14.5 and persists until the spermatogonia stage in new born and adult animals. The expression pattern is mutually exclusive with CTCF, which is expressed only in Sertoli cells and in post meiotic round spermatids (Jelinic et al., 2006). ChIP experiments on embryonic and adult testis revealed that CTCFL occupies both *Igf2-H19* and *Dlk1-Gtl2* ICRs. Since the expression timing of CTCFL correlates with the *de-novo* establishment of paternal imprints, it has been proposed that CTCFL plays a role in recruiting the DNA methylation machinery. Prmt7, an arginine methyltransferase with mediates dimethylation of arginine 3 of histone H4 is proposed to take part in this process. CTCFL interacts with Prmt7 and stimulates its activity *in vitro*. Prmt7 is co-expressed with CTCFL in embryonic and adult testis. H4R3me2 was detected in gonocytes and spermatogonia and association of H4R3me2 with *Igf2-H19* and *Dlk1-Gtl2* ICRs in the adult testis was demonstrated by ChIP. A direct link between CTCFL, Prmt7 and *de-novo* DNA methylation was demonstrated by nuclear co-injection of expression vectors encoding CTCFL, Prmt7, and Dnmt3a, -b and -l, in *Xenopus* oocytes, resulting in the *de-novo* methylation of *Igf2-H19* ICR provided on a plasmid (Jelinic et al., 2006).

CTCFL belongs to a group of cancer- testis genes. These genes are normally present only in the male germ line, but are also expressed in cancer cell lines and in primary tumors. Using such a cancer cell line, an interaction between CTCFL and H3K4me2 histone methyltransferase Set1a was demonstrated (Nguyen et al., 2008). Both CTCFL and Set1a were shown to occupy *Igf2-H19* ICR. It was proposed that CTCFL recruits Set1a to methylate H3K4 on this locus (Nguyen et al., 2008). The observation is not consistent with a hypothesis that CTCFL recruits the DNA methylation machinery to the ICRs in male germ cells. Several evidences, described in detail in the previous chapter, indicate mutually exclusive targeting of DNA methylation and H3K4 methylation. The observed discrepancy may either be caused by a cancer specific interaction of CTCLF or may indicate that the establishment of DNA methylation on the *Igf2-H19* ICR in the male germ line involves a phase of H3K4 methylation. It is currently not known what the other targets of CTCFL are and what effect it has on the chromatin of these loci. Subsequent actions of CTCFL and CTCF may play an important role in the organization of chromatin architecture throughout spermatogenesis.

A detailed ChIP study of the histone modification status of the paternally methylated ICRs during spermatogenesis has been performed (Delaval et al., 2007). As a control two maternally methylated ICRs were included in the analysis. Three developmental stages, meiotic spermatocytes, post meiotic round spermatids and elongating spermatids, were analysed. H3K4me2 was found enriched on maternal but not paternal ICRs in all stages, consistent with the mutually exclusive targeting of DNA methylation and H3K4 methylation. The same pattern was observed for H3K9 acetylation. It has been proposed that the methylation of H3K4 at maternal ICRs protects them from *de novo* DNA methylation. Interestingly, H3K4me2 status in spermatogonia was also analysed. Maternal ICRs showed again higher enrichment of H3K4me2, but also the *Igf2-H19* ICR was relatively more enriched than the other two paternal ICRs (Delaval et al., 2007). This observation together with the CTCFL interaction data (Nguyen et al., 2008) and the described delay in acquisition of DNA methylation on the maternally inherited *Igf2-H19* ICR (Davis et al., 2000), may point to a distinct mechanism of the establishment of the *Igf2-H19* imprint, which involves transient H3K4 methylation.

De-novo DNA methylation of other than imprinted loci

Concomitant with the establishment of paternal imprints, all retrotransposon sequences undergo *de-novo* DNA methylation. Dnmt3a and Dnmt3b have both a role in

this process and show specificity towards different types of retrotransposons (Kato et al., 2007). Dnmt3l is required for the methylation of all retrotransposon sequences (Bourc'his and Bestor, 2004; Kato et al., 2007). Even though Dnmt3l is not expressed beyond the pre-meiotic stage (spermatogonia) of spermatogenesis, knockout studies showed that the inefficient silencing of retrotransposons has severe effect on the meiosis (spermatocytes). In knockout animals, LINE and IAP retrotransposons were highly transcribed in spermatogonia and spermatocytes. The lack of Dnmt3l also caused a meiotic failure with wide spread, non-homologues chromosome synapsis and progressive loss of germ cells by the mid pachytene stage. It resulted in a complete absence of mature sperm in testis of older animals (Bourc'his and Bestor, 2004; Webster et al., 2005). Interestingly pericentric tandem repeats (major and minor satellites) were not affected and displayed a with wild type-like methylation status (Bourc'his and Bestor, 2004). However, in another study on the Dnmt3l knockout mouse, aberrant transcription of major and minor satellites was observed (Kato et al., 2007). The discrepancy may be caused by a different time point of DNA methylation measurement.

Recently, a link between retrotransposon silencing and the small RNA pathway has been discovered. A family of proteins, called Piwi, involved in small RNA biogenesis has been identified in the male germ line, together with a corresponding group of small RNAs - piRNAs. Based on knockout studies, Piwi proteins are implicated in the silencing of retrotransposons, both via a targeted degradation of retrotransposon transcripts by piRNAs and via induced DNA methylation of retrotransposon sequences (Aravin et al., 2006; Aravin et al., 2007; Kuramochi-Miyagawa et al., 2001).

At the prenatal gonocyte stage, not only imprinted loci and retrotransposons are methylated. Based on stainings with an antibody against methylated cytosine, a genome wide acquisition of DNA methylation was reported (Coffigny et al., 1999). Further, analysis using methylation sensitive restriction enzymes revealed that both *de novo* methylation and demethylation occur in spermatogonia and spermatocytes in the early meiotic prophase (Oakes et al., 2007). Alterations include predominantly non-CpG island sequences from both unique loci and repetitive elements. These modifications are progressive and are almost exclusively completed by the end of the pachytene spermatocyte stage (Oakes et al., 2007). Importantly, expression of the Dnmt1, which was silenced in G1 arrested gonocytes, is reestablished after birth. Dnmt1 activity

assures maintenance of the DNA methylation patterns of proliferating spermatogonia (La Salle et al., 2004).

Dynamic changes in histone PTMs and histone variants during meiosis

In the male germ line, the entrance into the meiotic prophase initiates the meiotic cell division which lasts for around 8 days. During the meiotic prophase, homologues chromosomes align in a process called synapsis and fragments of chromosomes are exchanged through homologues recombination. Subsequently, germ cells undergo two successive rounds of cell division, in which homologues chromosomes and their respective sister chromatids separate. Meiosis results in the production of four haploid spermatids.

Several histone methyltransferases have crucial functions in the progression of the meiotic prophase. In mice deficient for both H3K9 tri-methyltransferases genes *Suv39h1* and *Suv39h2,* spermatogenesis is arrested at the meiotic pachytene stage (Peters et al., 2001). Spermatocytes lack pericentric H3K9me3, and chromosomes undergo non-homologous interactions, predominantly at centromeres. These results suggest that Suv39h1/h2 mediated H3K9me3 at pericentric chromatin is necessary for the proper progression of meiosis (Peters et al., 2001).

Abnormalities in meiotic prophase progression have been also demonstrated in mice lacking the H3K9 mono- and dimethylase G9a (Tachibana et al., 2007). Spermatocytes were arrested at the early pachytene stage. The synapsis between homologous chromosomes was not properly formed. Upregulation of several genes has been also observed, suggesting that silencing of these genes via G9a mediated H3K9 mono- and dimethylation may be required for proper synapsis (Tachibana et al., 2007).

Prdm9 (also known as Meisetz), a germ line specific H3K4 tri-methylase, is also crucial for synapsis and recombination of homologous chromosomes during meiotic prophase (Hayashi et al., 2005). In Prdm9-deficient spermatocytes, a number of genes, including those that are specifically expressed in meiotic germ cells, were repressed. These results suggest that Prdm9 mediated H3K4 methylation is involved in the activation of genes important for synapsis and recombination (Hayashi et al., 2005).

In contrast to a complete synapsis between homologues chromosomes, X and Y chromosomes pair only at a short homologues fragment called the pseudo-autosomal region (PAR). The unpaired domains of the sex chromosomes undergo a specific

silencing process, called meiotic silencing of unsynapsed chromosomes (MSUC), which in case of X and Y is called meiotic sex chromosome inactivation (MSCI) (Monesi, 1965; Turner et al., 2000; Turner et al., 2005). At the beginning of the meiotic prophase, both autosomes and sex chromosomes are transcriptionally active. When the synapsis of autosomes is completed, X and Y chromosomes are rapidly silenced and compartmentalized into a peripheral nuclear subdomain called the XY-body (Solari, 1974). MSCI is associated with extensive remodeling of chromatin including changes in histone modifications and the incorporation of several histone variants. One of them is the H2A variant, H2AX (Fernandez-Capetillo et al., 2003; Turner et al., 2004). At the onset of meiotic prophase H2AX localize to sites of DNA double-strand breaks, where it gets phosphorylated at Ser139 (γH2AX) and recruits the DNA repair machinery (Mahadevaiah et al., 2001). γH2AX disappears from autosomes when synapsis is completed, but reappears or remains present on sex chromosomes when the XY body is formed. Analysis of mice lacking H2AX demonstrated that it is essential for XY body formation and MSCI (Fernandez-Capetillo et al., 2003).

At or shortly after the initiation of MSCI, several changes in the histone modifications status are detected. These include ubiquitination of H2A, di-methylation of H3K9, deacetylation of histones H3 and H4 and the disappearance of H3K27 trimethylation (Baarends et al., 1999; Khalil et al., 2004; Takada et al., 2007). Changes also include the incorporation of histone H2A variant macroH2A (Hoyer-Fender et al., 2004). This extensive chromatin remodelling is concomitant with chromosome wide nucleosomal eviction at the XY body (van der Heijden et al., 2007). This was demonstrated by the loss of H3.1/2 from the sex chromatin and the appearance of H3.3, together with transient presence of H3.3 chaperone HirA. Also a temporary absence of several histone modifications was observed, concomitant with the disappearance of H3.1/2. Therefore, part of the rapid changes in the histone modification status at the XY body takes place due to a histone replacement mechanism (van der Heijden et al., 2007).

After completion of meiosis, a partial re-activation of X-chromosome is observed. Based on microarray expression profiling, approximately 13% of the X-linked genes are re-expressed in post-meiotic period (Namekawa et al., 2006). In a recent study, Mueller and colleagues identified a number of X-linked multi-copy gene families, which show post meiotic re-activation. In round spermatids, these genes showed much higher expression level than the single copy genes. The authors suggest that the amplification

of genes on the X chromosome compensates for the post-meiotic repression (Mueller et al., 2008). Despite the continued repressed status of the majority of X-liked genes, there are multiple changes in the chromatin of the XY body occurring after meiosis. γH2AX, macroH2A and H2A ubiquitination are lost, whereas H3.3 and H3K9 methylation persist (Khalil et al., 2004; Namekawa et al., 2006; van der Heijden et al., 2007). Furthermore, another H2A variant H2A.Z is incorporated. H2A.Z is first expressed during the meiotic prophase, but at that time it is excluded from the XY-body. In round spermatids the expression of H2A.Z rapidly increases. At that stage H2A.Z accumulates at the sex chromosomes (Greaves et al., 2006). Therefore, nucleosomes of the post-meiotic X and Y are carrying simultaneously H3.3 and H2A.Z histone variants. This combination has been shown to increase nucleosomal instability (Jin and Felsenfeld, 2007). Such unstable nucleosomes may facilitate the global chromatin remodeling taking place in the following stages of spermatid elongation.

Global chromatin remodelling at last stages of spermatogenesis

After completion of meiosis round spermatids enter a process of spermiogenesis, which involves major structural changes in the nucleus and in the cytoplasm. In the successive stages of spermiogenesis spermatids elongate and shed off their cytoplasm. The nucleus is compacted into a volume of about 5% of that of a somatic cell nucleus. This remarkable condensation is achieved by replacing histones with protamines (Marushige and Marushige, 1975). Protamines are arginine- and cysteine-rich proteins that organize the haploid male genome into a highly specialized, “doughnut-shaped” chromatin structure (Allen et al., 1993; Haaf and Ward, 1995). Protamines are present in the sperm of multiple species from echinoderms to primates and their sequence is highly conserved (reviewed in (Ausio, 1999)).

Unlike many other mammals, which express only protamine 1 (Prm1), mouse and human genomes encode for two different protamine molecules, Prm1 and Prm2. Both are encoded by relatively short genes comprising of two exons. *Prm2* encodes a precursor protein that binds to DNA and during the last stages of spermatid elongation undergoes proteolytic processing. The mature forms of Prm1 and Prm2 bind to 10 and 15 bp of DNA, respectively. This binding neutralizes the negative charge of the DNA backbone and enables the DNA molecules to pack closely together. In a final step of protamine mediated compaction, which is happening after the spermatozoa leave the

testis and proceed through epididymis, a network of bisulfate bonds is formed between the adjacent protamine molecules (Reviewed in (Balhorn, 2007)).

Besides providing a high genome compaction, very little is known about the functions of protamines in the reproductive process. It is not known whether the establishment of protamine domains serves a function in the epigenetic gene regulation, especially in the dramatic transition from the spermatogenic differentiation program to embryonic totipotency. Both protamines are essential for normal sperm development. Haploinsufficiency of either Prm1 or Prm2 causes abnormal sperm morphology, reduced sperm number and infertility (Cho et al., 2001). Furthermore, the proportion between the two protamines has been shown to be critical for fertility. Spermatozoa with reduced levels of Prm2 have increased DNA damage and incomplete chromatin condensation. Embryos fertilized by these spermatozoa through an intra-cytoplasmic sperm injection (ICSI) do not develop beyond the blastocyst stage (Cho et al., 2003). Protamines are also modified posttranslationally. Prm2 undergoes phosphorylation by CamK4 (Ca+/calmodulin –dependent protein kinase 4). Disruption of the *CamK4* gene in mouse results in the failure of the Prm2 incorporation and causes male infertility (Wu et al., 2000). Therefore, phosphorylation of Prm2 plays a major role in the proper chromatin remodeling during spermatogenesis.

In mammals, histones are not replaced directly by protamines. Transition proteins (TP1 and TP2) are intermediates in this process. Targeted mutation of TP1 and TP2 indicate that they have largely redundant functions (Yu et al., 2000; Zhao et al., 2001). Both mutant mice are fertile and display only minor spermiogenesis abnormalities. In mice lacking both TPs, protamine deposition proceeds normally, but many late spermatids showed DNA breaks, and Prm2 was not posttranslationally processed (Zhao et al., 2004). A fraction of spermatozoa could be used for efficient fertilization by ICSI, resulting in birth of healthy offspring. However, the number of spermatozoa was drastically reduced and the majority of them showed highly abnormal morphology, arguing that TPs are required for a normal sperm development and fertility (Zhao et al., 2004).

Expression of both TPs and protamines is tightly regulated during spermatogenesis. Global transcription ceases along with the chromatin condensation. Therefore, transcription and translation of protamines are uncoupled. Protamine transcription starts in fully transcriptionally active round spermatids and the transcripts are stored in the cytoplasm as messenger ribonucleoprotein particles (mRNPs). They

are activated for translation in elongated spermatids only 1 week later (reviewed in (Braun, 2000)). Recently, an H3K9me1/2 de-methylase JHDM2A was shown to be involved in the activation of TP1 and Prm1 genes (Okada et al., 2007). *Jhdm2a*-deficient male mice displayed defects in spermatid elongation, including abnormal nuclear morphology and were infertile. JHDM2A was shown to bind directly to *TP1* and *Prm1* genes, and to be responsible for the reduction of H3K9 methylation at their promoters (Okada et al., 2007).

To achieve a global change in chromatin architecture, a sophisticated machinery that removes nucleosomes from DNA and incorporates TPs and later protamines must exist. As mentioned above, the incorporation of histone variants and changes in histone modification status are thought to contribute to this process. In elongating spermatids, a wave of global histone acetylation on H2A, H2B, H3 and H4 is observed (Hazzouri et al., 2000). This acetylation disappears along with the condensation and the histone to protamine exchange. Additionally, hyperacetylation of histones has been detected in elongating spermatids during spermatogenesis in humans (Faure et al., 2003). Histone acetyltransferases responsible for this process are not known. Cdyl, a histone methyltransferase abundantly expressed in mouse testis, may be a candidate. It is expressed concomitant with histone H4 hyperacetylation and localises to the nuclei of elongating spermatids (Lahn et al., 2002).

One can speculate that histone acetylation, via destabilizing the interaction among nucleosomes and between nucleosomes and DNA, facilitates the removal of histones from chromatin. Furthermore, specific chromatin remodelers that interact with acetylated histones may play a role in this process. Brdt, a testis specific protein that contains two bromodomains, specifically binds the acetylated H4 tail. Interestingly, recombinant Brdt was capable of inducing the reorganisation of chromatin in somatic cells both *in vivo* and in isolated nuclei *in vitro*. This activity was dependent on TSA treatment which induces hyperacetylation (Pivot-Pajot et al., 2003). Targeted deletion of the first bromodomain of *Brdt* in mice resulted in male infertility. Additionally several abnormalities during spermatid elongation were observed (Shang et al., 2007). These data strongly support a model where Brdt has an important function in the acetylation dependent chromatin remodelling in elongating spermatids.

Several testis specific variants are expressed during mammalian spermatogenesis. For the majority of them the function is not well understood (reviewed in (Kimmins and Sassone-Corsi, 2005)). I will shortly describe a few of these variants, which may be implicated in the chromatin remodelling.

H2B variant TSH2B (also known as TH2B) has been detected in human testis throughout spermatogenesis and persists in around 20% of the mature spermatozoa (van Roijen et al., 1998; Zalensky et al., 2002). Negative correlation between sperm chromatin compaction and the TSH2B content has been demonstrated (Singleton et al., 2007). Therefore, it is not clear to what extend TSH2B is a regular component of the mature spermatozoa and to what extend it demonstrates inefficient remodelling of the sperm chromatin. Alternatively, the lower compaction of the TSH2B positive spermatozoa may serve as an advantage in the chromatin decondensation process after fertilisation.

Testis specific histone H1 variants: H1t, H1T2 and HILS were characterised. H1t is expressed exclusively in spermatocytes and spermatids. Targeted disruption of this variant in mice did not result in any effect on spermatogenesis or male fertility. The overall levels of histone H1 were significantly decreased in spermatocytes and spermatids of H1t-null animals, arguing that there was no compensation with other H1 variants (Fantz et al., 2001). Interestingly, H1t was reported to display the lowest condensing effect on chromatin fibres *in vitro* comparing to other variants, suggesting that it may facilitate the histone removal (De Lucia et al., 1994).

H1T2 is expressed exclusively in round and elongating spermatids and localises to chromatin domain at the apical pole. Mice with a disrupted *H1t2* gene have reduced fertility and show delayed nuclear condensation and aberrant elongation (Martianov et al., 2005).

HILS1 is detected in nuclei of elongating and elongated mouse spermatids. A homologus HILS1 gene was also identified in human. Since HILS1 is colocalising with TPs and Prm1, it is proposed to have a function in the nuclear condensation during the elongation stages (Yan et al., 2003).

In a recent study Govin and colleagues performed a detailed analysis of the pericentric heterochromatin (PCH) remodelling at the last stages of murine spermatogenesis (Govin et al., 2007). In round spermatids, PCH is localised into one round chromocenter. During spermatid elongation, chromocenter forms a broader

domain but can be still observed as a distinct compartment in the middle of the cell. During that time PCH becomes enriched with the acetylated histone H4, concomitant with the loss of HP1β. Yet, H3K9me3 persists. Along with the histone to protamine exchange acetylated H4 and H3K9me3 progressively disappear, although data from another group and unpublished data from our laboratory suggest that H4 acetylated on lysines 8 and 12 persist in the mature sperm ((van der Heijden et al., 2006), U.B and A.P unpublished – see Discussion).

Using a biochemical fractionation followed by fluorescent *in situ* hybridisation, Govin and colleagues demonstrated that during the time of histone to protamine transition PCH retains nucleosomal organisation. Consistently, van der Heijden and colleagues used a specific antibody to show that chromocenter of elongating spermatids is associated with nucleosomes (van der Heijden et al., 2006). Further investigation of PCH led to the identification of new histone variants: H2AL1, H2AL2, and H2BL1 (Govin et al., 2007). In elongated spermatids, H2AL1 and H2AL2, together with TH2B (mouse homolog of TSH2B), are present in a nucleosome-like structure, which does not contain histones H3 and H4. This specific structure is observed mainly at the PCH, but is also dispersed in other places in the genome. Histone variants H2AL1 and H2AL2 are retained in mature sperm and are proposed to guide epigenetic reprogramming of the paternal PCH after fertilization (Govin et al., 2007).

Nucleosomal component of the sperm chromatin

As mentioned above, some histones escape the global remodelling and are retained in the mature spermatozoa. Very little is known about their function. They may be a remaining of the imperfect chromatin remodelling. However, sequence specific localisation of the retained histones argues against such scenario.

Histones present in mature human spermatozoa were isolated and for the first time characterised based on their electrophoreyic mobility and amino acid composition (Puwaravutipanich and Panyim, 1975; Tanphaichitr et al., 1978). Further, liquid chromatography was applied to purify sperm histones. Subsequent characterisation revealed presence of all core histones including H2A.X, H2A.Z and H3.3 variants, but excluding histone H1 (Gatewood et al., 1990). Based on these observations histone content of the human sperm was estimated as 15%. It has been also demonstrated that the retained histones are associated with DNA and form the nucleosomes. Following sperm decondensation with reducing agents, beads on a string configuration was

observed by electron microscopy (Gusse and Chevaillier, 1980). Furthermore, sperm treatment with micrococcal nuclease followed by DNA isolation and electrophoresis revealed typical nucleosomal ladder pattern (Zalenskaya et al., 2000). Consistently, the chromatin fraction released by the micrococcal nuclease treatment was exclusively containing histones and not protamines, as revealed by protein analysis using Acetic acid- Urea electrophoresis.

Several further studies tried to address an essential question - where are the retained histones localised? Atomic force microscopy and *in situ* hybridisation were applied to reveal the human sperm chromatin architecture. These studies led to a model where centromeres are organised in a compact chromocenter buried inside the nucleus, whereas telomeres are localised at the periphery (Zalensky et al., 1995; Zalensky et al., 1993). Further, both centromeres and telomeres were shown to be associated with histones. CENP-A a centromeric variant of histone H3 was shown to colocalise with the centromeric DNA in decondensed human sperm nuclei (Zalensky et al., 1993). Biochemical purification and characterisation of a telomere-binding protein complex from the human sperm revealed presence of the histone variant - TSH2B (Gineitis et al., 2000).

Very little is known about the histone component of murine spermatozoa. There is a common knowledge in the field that the amount of the retained histones is approximately 1%, although no solid quantification data has been published so far. *In situ* hybridisation studies revealed similar to human sperm nuclear architecture of the murine sperm, with centromeres located in a chromocenter in the middle and telomeres associated with the periphery (Haaf and Ward, 1995). Furthermore, presence of histones in the chromocenter was demonstrated by immunofluorescence with antibody against acetylated histone H4 (van der Heijden et al., 2006).

This defined architecture of mouse and human sperm may be designed for the coordinated unpacking and activation of the male genome after fertilisation. Furthermore, nucleosomal configuration of centromeres and telomeres, which are two key structural elements of the chromosomes, may assure their proper functionality in the embryo.

Until very recently only sparse data existed about the potential nucleosomal configuration of the gene sequences in mouse and human spermatozoa. Two groups working on human sperm used a low salt extraction to selectively uncover histone associated sequences, which were further released by restriction enzyme digestion

(Gardiner-Garden et al., 1998; Wykes and Krawetz, 2003). Based on these studies γ and ε *globin*, *Prm1, Prm2* and *TP1* loci were reported to be associated with histones. Further another group analysed DNA released by the micrococcal nuclease digestion and reported that the *Igf2* locus is in a nucleosomal configuration (Banerjee and Smallwood, 1998). In a single study on murine sperm, authors analysed DNA released by an endogenous nuclease activity and reported that retrotransposon sequences LINE/L1 are associated with histones (Pittoggi et al., 1999).

Two studies addressing the question of histone localisation in mature human and mouse spermatozoa have been published in the recent two months (Arpanahi et al., 2009; Hammoud et al., 2009). As they were not taken into account in the experimental design of my thesis I will discuss them in the results and discussion sections.

1.4.3 Dynamics of the paternal chromatin after fertilization

After fertilization, paternal protamines are re-exchanged for the maternally provided histones. Protamines are shed from the sperm DNA within 30 min. after the gamete fusion and completely disappear 50 min. later (van der Heijden et al., 2005). It is not known if present in sperm histones are also dissociated from the DNA at that time, but several evidences argue that they are retained (see below). Histone chaperon Hira is detectable at that time at the paternal chromatin. Consistently, histone H3 variant H3.3 and not H3.1/2 is incorporated to the paternal chromatin following sperm decondensation (Torres-Padilla et al., 2006; van der Heijden et al., 2005). At a global level H3.1/2 is detectable on the paternal chromatin only following DNA replication in the later zygote. In contrast, the maternal chromatin is enriched in this variant throughout the zygotic stage (van der Heijden et al., 2005). The same asymmetry of histone H3 variants was observed in human (van der Heijden et al., 2005).

This configuration allowed to demonstrate the transmission of human sperm derived histones to the embryo. After heterologous fertilisation of a mouse oocyte with a human spermatozoon, foci of H3.1/2 were detected on the paternal chromatin, arguing that these histones are of the paternal origin. The signal persisted during and after the incorporation of the maternal histones, suggesting that paternally derived histones are not removed during this process (van der Heijden et al., 2008). Unfortunately, due to the low levels of histones and the limited antibody sensitivity, the same experiment can not be performed using murine spermatozoa.

Decondensing sperm forms the paternal pronucleus, which remains separated from the maternal genome until the first cell division. Upon sperm entry, the oocyte completes the second meiotic division. Half of the genetic material is extruded in a form of a second polar body. The second half, remaining in the oocyte, forms the maternal pronucleus. Maternal and paternal pronuclei largely differ in their histone modification status. Maternal genome is passed on to the embryo in a nucleosomal configuration and carries multiple histone modifications. In contrast, on the paternal genome, majority of histones are acquired *de-novo* and gradually become enriched with different modifications (Adenot et al., 1997; Cowell et al., 2002).

Immediately after incorporation, histones H3 and H4 of the paternal pronucleus acquire acetylation at various residues (Adenot et al., 1997; van der Heijden et al., 2006). Acetylation of lysine 5 and 12 of histone H4 (H4K5ac and H4K12ac) is already present on the maternally provided histones prior to their incorporation (Sobel et al., 1995). Van der Heijden and colleagues observed that at the onset of the maternal histones incorporation, not all the nucleosomes are carrying these marks, representing most probably paternally derived histones (van der Heijden et al., 2006). Further, they showed that H4K8ac is detectable at the decondensing sperm immediately after the fertilization. As already mentioned, H4K8ac and H4K12ac are associated with the PCH in the mature sperm. Together with the fact that H4K8ac is not present on the maternally provided histones, these data are strongly arguing that the observed mark is of the sperm origin (van der Heijden et al., 2006). By contrast, the histones variants H2AL1 and H2AL2, which are also contributing to the PCH of the mature sperm, are rapidly lost from the paternal genome after fertilization (Wu et al., 2008).

The acquisition of histone lysine methylation marks occurs gradually on the paternal pronucleus. The first detectable mark is H4K20me1, followed by mono-methylation of H3K4, H3K9 and H3K27. Before the completion of the first cell division also di- and trimethylation of H3K4 and H3K27 are established, resulting in equal levels of these modifications at both parental genomes. By contrast, di- and trimethylation states of H3K9 are only established around the 4-8 cell stage (Lepikhov and Walter, 2004; Puschendorf et al., 2008; Santos et al., 2005; van der Heijden et al., 2005).

Asymmetrical distribution of histone lysine methylation marks have been also observed in humans (van der Heijden et al., 2009). The analysis was performed on tri-pronuclear zygotes derived by *in vitro* fertilisation. After fertilisation, H3K9me3 and

H3K27me3 were absent from two paternal pronuclei and clearly detectable at the maternal one. Interestingly, the asymmetry was not observed for H3K4me3 which showed high levels at both parental genomes.

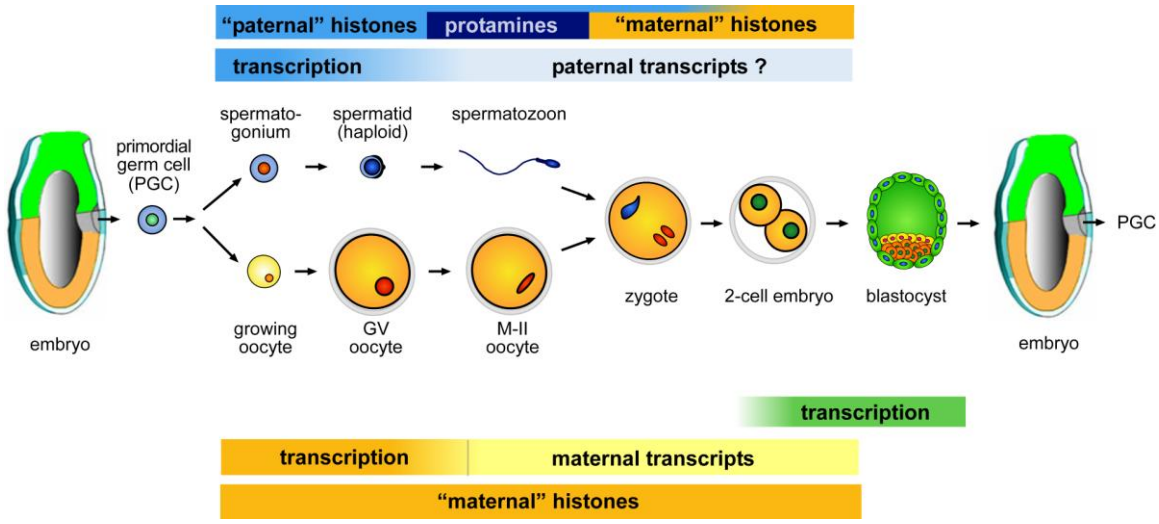


Figure 4. Epigenetic germ line inheritance. Primordial germ cells are specified in the gastrulating embryo and after migration into the gonads, differentiate into either spermatozoa or oocytes. These two gametes differ in their potential to transmit epigenetic information. Like in somatic cells, the oocyte genome is organized in nucleosomal configuration with DNA bound to histones that carry many histone modifications. By contrast, at the last stages of spermatogenesis, most of histones are replaced by protamines, leading to highly compact sperm chromatin structure. After fertilization, protamines are exchanged by maternally provided histones. Nevertheless, about 10% of histones are retained in human spermatozoa and have a potential to influence transcription during development. Concomitant with the histone to protamine exchange, the transcription is ceased in elongating spermatids. Nevertheless, multiple transcripts are retained in mature spermatozoa and may influence the embryonic transcription as well.

1.5 Scope of the thesis

When I started my thesis, it was not known whether the histones, present in mouse and human spermatozoa, are carrying posttranslational modifications. As highlighted in many examples in the first chapter of this introduction, there must exist means for the transmission of epigenetic information across generations, other than DNA methylation. Histone modifications provided a very good candidate to play such a role.

As described in the second chapter, Polycomb and Trithorax complexes play important roles in the gene regulation. Expression states sustained by these complexes are stably maintained in somatic lineages. Therefore, I considered that Polycomb and Trithorax mediated histone modifications could carry epigenetic information not only through somatic cell division, but also through gametes to the next generation.

Germ cells, unlike somatic cells, must maintain the potential to generate a totipotent embryo. Nevertheless, germ cells differentiate into specialised gametes. It was an intriguing question for me, how epigenetic programs of differentiation and totipotency can co-exist in these cells. This seemed particularly challenging for the paternal genome, which undergoes a major re-organisation at the last stages of spermatogenesis.

In the zygote the paternal genome becomes enriched in histone methylation at different lysine residues in a temporal and spatial regulated manner. This led me to hypothesize that the presence of paternally inherited modifications may influence the acquisition of new modifications on the deposited histones *in-cis*. The transmitted histones could influence the *de novo* organization of chromatin and have an effect on the gene expression in the early embryo.

In this PhD project, I initially aimed to analyse the histone modification status in murine spermatozoa. The mouse model would provide me a direct possibility to investigate the functionality of the retained histones. However, very low histone content and high compaction made it difficult to analyse murine spermatozoa. Contrary, human spermatozoa contain 10 times more histones and their chromatin structure is less compact. Thus, I decided to study transmission of modified histones to the next generation, first in human and then in mouse spermatozoa. Five years later, these two data sets allowed me to investigate the conservation of the transmitted histone marks between humans and mice.

2. Results

2.1 Submitted manuscript:

Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa

Urszula Brykczynska, Mizue Hisano, Liliana Ramos[#], Edward J. Oakeley, Tim C. Roloff, Dirk Schübeler, Michael B. Stadler and Antoine H.F.M. Peters*

Friedrich Miescher Institute for Biomedical Research,
Maulbeerstrasse 66,
CH-4058 Basel,
Switzerland

[#]Fertility Laboratory,
Department of Obstetrics and Gynaecology,
Radboud University Nijmegen Medical Center,
P.O. Box 9101, NL-6500 HB Nijmegen,
The Netherlands

*Corresponding author:
Antoine Peters
Phone: +41 61 6978761
Fax: +41 61 6973976
Email: antoine.peters@fmi.ch

Keywords:
chromatin, histone methylation, Polycomb, spermatozoa, transgenerational inheritance, epigenetic memory

In higher eukaryotes, histone methylation is involved in the maintenance of cellular identity during somatic development. During spermatogenesis, most nucleosomes are replaced by protamines. Therefore, it is unclear if histone modifications function in paternal transmission of epigenetic information. Here we show that active H3K4 di-methylation (H3K4me2) and repressive H3K27 tri-methylation (H3K27me3), two modifications important for Trithorax and Polycomb-mediated gene regulation, are present in chromatin of human spermatozoa and show methylation-specific distributions at regulatory regions. H3K4me2-marked promoters control gene functions in spermatogenesis and cellular homeostasis suggesting that this mark reflects germline transcription. In contrast, H3K27me3 marks promoters of key developmental regulators in sperm as in soma. Many H3K27me3-marked genes are never expressed in the male and female germline, and in early "totipotent" embryos, suggesting a function for Polycomb in repressing somatic determinants across generations. Targets of H3K4me2 and H3K27me3 are also modified in mouse spermatozoa, implicating an evolutionary conserved role for histone methylation in chromatin inheritance via the male germline.

By classical Mendelian inheritance, genetic information is transmitted through the generations, underlying phenotypic diversity in sexually reproducing organisms. Nonetheless, non-Mendelian inheritance of traits across generations has been reported in various higher eukaryotes (Chong and Whitelaw, 2004). Furthermore, the low reproductive efficacy of nuclear transfer versus natural reproduction (Hochedlinger and Jaenisch, 2003) suggests that resetting and maturation of chromatin states during gametogenesis is critically important for early embryogenesis, arguing for a transgenerational epigenetic contribution at conception.

In mammals, the dimorphic gametes differ greatly in their potential to transmit epigenetic information encoded in histones and associated posttranslational modifications (Albert and Peters, 2009; Puschendorf et al., 2008; Reik, 2007). Whereas in oocytes chromatin retains a nucleosomal conformation, marked by histone methylations (Puschendorf et al., 2008), the majority of histones are replaced by protamines at the end of spermatogenesis (Balhorn et al., 1977; Gatewood et al., 1987) (Fig. 1a). Following gamete fusion, maternally provided histones replace protamines that subsequently become post-translationally modified by oocyte-derived factors (Albert and

Peters, 2009; Puschendorf et al., 2008) Despite such major remodeling, histones have been reported to reside at specific sequences in human and mouse spermatozoa (Gardiner-Garden et al., 1998; Gatewood et al., 1987; Pittoggi et al., 1999; Wykes and Krawetz, 2003) and to remain associated with the paternal genome during *de novo* nucleosome formation upon fertilization (van der Heijden et al., 2008).

During somatic development, Polycomb (PcG) and Trithorax (TrxG) group proteins serve conserved chromatin-based repressive and anti-repressive roles in epigenetic memory of cell identity, e.g. by controlling expression of developmental regulators that drive differentiation (Hublitz et al., 2009; Sparmann and van Lohuizen, 2006). In mammals, PcG proteins function at least in two distinct Polycomb Repressive Complexes (PRC). PRC2, consisting of Eed, Suz12, and Ezh2 or Ezh1, catalyzes tri-methylation on histone H3 lysine 27 (H3K27me3) (Cao et al., 2002; Kuzmichev et al., 2002; Shen et al., 2008), a modification associated with gene repression in development (Ezhkova et al., 2009; Mohn et al., 2008). Recently, it was shown that PRC2 binds to H3K27me3 and that its catalytic activity is required for long term repression (Hansen et al., 2008). These data suggest that H3K27me3 functions in transcriptional memory of the repressed state. Mammalian TrxG proteins of the Mixed Lineage Leukemia protein family mediate H3K4 methylation, a mark associated with transcriptional activity. Up to date, the roles of H3K27 and H3K4 methylation in transgenerational inheritance have been unknown. Here, we study whether H3K27me3 and H3K4me2 are selectively retained at regulatory sequences in mature human and mouse spermatozoa. Our analyses show that in general promoters of orthologous genes are similarly marked. This finding is compatible with an evolutionary conserved role for histone methylation in chromatin inheritance across generations.

RESULTS

H3K4me2 and H3K27me3 mark functionally distinct gene sets

To evaluate the presence of histones and associated modifications in mature human and mouse spermatozoa, we performed Western blot analyses on highly purified human and mouse spermatozoa (Fig. 1b). Comparable to described previously, we observed approximately 10% of histone H3 in human sperm (Gatewood et al., 1987). In mouse sperm, we only detected about 1% of histone H3 (Fig. 1b). We further detected H3K27me3 and H3K4me2 in spermatozoa of both species (Fig. 1b). Absence of signal for Lamina Associated Polypeptide 2 beta (LAP2β), a marker for somatic and immature

germ cells (Alsheimer et al., 1998), demonstrated purity of sperm samples used. To define the chromosomal localization of modified histones, we developed a chromatin immuno-precipitation (ChIP) approach for H3K4me2 and H3K27me3 that is compatible with the highly condensed chromatin state present in human spermatozoa. Following ChIP on cross-linked chromatin isolated from a pool of human spermatozoa obtained from 9 fertile donors, we amplified and hybridized precipitated genomic DNA to an oligonucleotide array representing over 18'000 human promoters each spanning 2.7 kb around the transcriptional start site. After scanning, we applied a Hidden Markov Model-based peak-finding algorithm (Supplementary Fig. 1, 2 and Methods) and identified over 1'600 and 4'500 promoters that are marked by H3K27me3 and H3K4me2 respectively in three independent ChIP-chip experiments (Fig. 1c, Supplementary Table 1). Thus over 30% of all tested human promoters are positive for these histone modifications. Unfortunately, low immuno-precipitation efficiencies of different anti-H3 antibodies on cross-linked chromatin precluded us from determining the genome-wide nucleosomal occupancy in human spermatozoa. In independent sperm samples, single-gene analyses of 41 selected promoters confirmed that promoters are uniquely modified by either one or both modifications (Fig. 1d, Supplementary Fig. 3). As a separate validation, we performed ChIP experiments under native conditions and obtained similar results arguing for an overall conservation of promoter distributions of H3K27me3 and H3K4me2 in human mature spermatozoa (Fig. 1e, Supplementary Fig. 3).

Finally, we compared our ChIP-chip results to recent ChIP-sequencing data that were obtained for H3K4me3 and H3K27me3 on native chromatin prepared from pools of human spermatozoa of different donors (Hammoud et al., 2009). We observed strong correlations between the average IP enrichment obtained for each promoter region in the current ChIP-chip experiments and the normalized number of ChIP-seq reads aligned to the corresponding promoter region (Hammoud et al., 2009) (Pearson's correlation coefficients of 0.68 for H3K4 methylation and 0.57 for H3K27 methylation) (Supplementary Fig. 4a,b). At the gene level, 79% of promoters enriched in H3K4 methylation in the current study were also enriched in the ChIP-seq experiments, whereas this was 69% for H3K27 methylation (Supplementary Fig. 4e,f). The high reproducibility between the two data sets, despite differences in ChIP and detection methodologies used, argues for a widespread marking of promoters by histone methylation in human spermatozoa. Hammoud and colleagues (Hammoud et al., 2009) also localized micrococcal nuclease resistant nucleosomes in chromatin of human

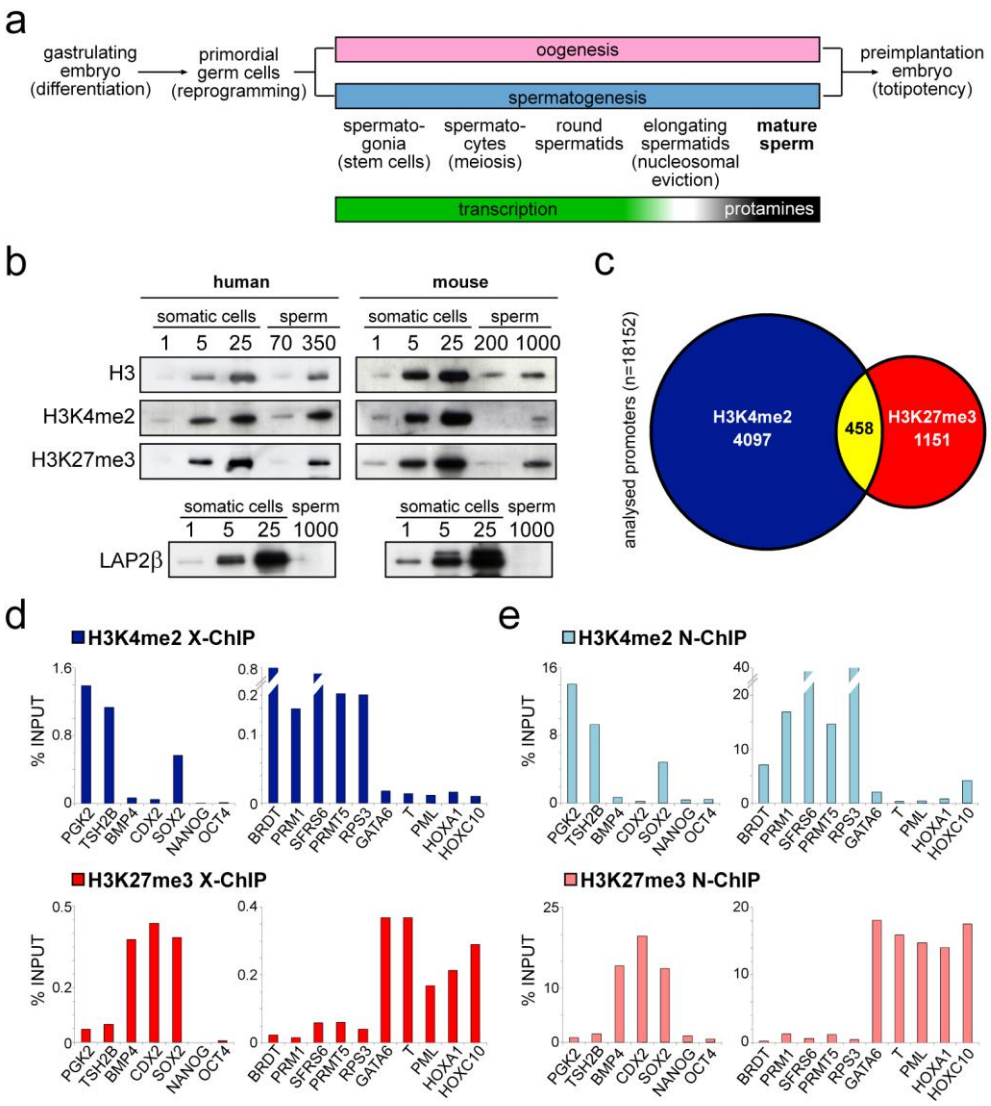


Figure 1. Methylated histones are present in human sperm and localize to distinct promoter sets.
(a) Illustration of mammalian germ cell and embryonic development. Primordial germ cell cells, specified in the proximal epiblast, undergo epigenetic reprogramming including global DNA demethylation. During spermatogenesis, male germ cells first proliferate (spermatogonia), then undergo meiosis (spermatocytes) and convert into spermatozoa after transcriptional arrest and global exchange of histones by protamines (elongating spermatids). Fertilization of the oocyte results in the totipotent early embryo. (**b**) Presence of histone H3, H3K27me3 and H3K4me2 in human and mouse spermatozoa as measured by protein blot analysis. Absence of signal for Lamina associated polypeptide 2 beta (LAP2β), a marker for somatic and immature germ cells (Alsheimer et al., 1998), shows purity of sperm samples used. Cell numbers are given in thousands. (**c**) Venn diagram illustrating the number of promoters that are marked by H3K4me2 and/or H3K27me3 in human sperm in three replicate experiments, as detected by microarray. (**d**, **e**) Validation of promoter array results by real time PCR analyses of immunoprecipitated chromatin after paraformaldehyde fixation (**d**) (X-ChIP) or native conditions (**e**) (N-ChIP) (see also Supplementary Fig. 3 for multiple replicates). Genes were selected on basis of gene function and their modification status at promoters as determined by microarray analyses.

spermatozoa. We observed that 33% and 56% of promoters associated with H3K4me2 and H3K27me3 respectively in the current ChIP-chip experiments carried detectable levels of nucleosomes (Supplementary Fig. 4e,f). Likewise, in the original ChIP-seq study (Hammoud et al., 2009), only 44% and 62% of promoters associated with H3K4me3 and H3K27me3 respectively were associated with nucleosomes (Supplementary Fig. 4e,f). These relatively low but reproducible levels of co-occupancy underscore the difficulty in determining the genome-wide occupancy of nucleosomes in human sperm.

Next we addressed whether promoters bound by H3K27me3 and/or H3K4me2 share sequence features or characteristic functions of the associated genes. We grouped promoters according to their CpG density, and observed that H3K27me3 in sperm was restricted to CpG-island containing promoters (Supplementary Fig. 5), as observed in somatic cells (Mikkelsen et al., 2007; Mohn et al., 2008), Gene ontology analysis showed that many developmental regulatory genes (e.g. *SOX2*, *CDX2*, *GATA6*, *BMP4*, *T*) and *HOX* genes are strongly over-represented among H3K27me3 genes, some of which were also marked by H3K4me2 (Fig. 1d, Fig. 2, Supplementary Fig. 3, Supplementary Table 2). H3K4me2 was strongly over-represented among promoters regulating various spermatogenic processes (e.g. *PRM1*, *PGK2*, *BRDT*, and *TSH2B*) (Fig. 1d, 2, Supplementary Fig. 3, Supplementary Table 2). Promoters of genes functioning in cellular homeostasis and gene expression (e.g. *RPS3*, *SFRS6*, *DICER1*, *PRMT5*) were significantly over-represented among H3K4me2- and under-represented among H3K27me3-marked genes (Fig. 2). These data show that functionally distinct gene sets are marked by the two modifications in sperm.

During meiotic prophase, a remarkable chromosome-wide exchange of nucleosomes takes place at both sex chromosomes in the course of meiotic sex chromosome inactivation (van der Heijden et al., 2007). Interestingly, X-linked genes were largely devoid of both H3K4me2 and H3K27me3 in sperm (Supplementary Fig.6), despite partial transcriptional reactivation of X-linked genes in round spermatids (Chalmel et al., 2007; Namekawa et al., 2006). This observation suggests that remodeling events specific for the X chromosome reduce the local retention of modified histones in sperm.

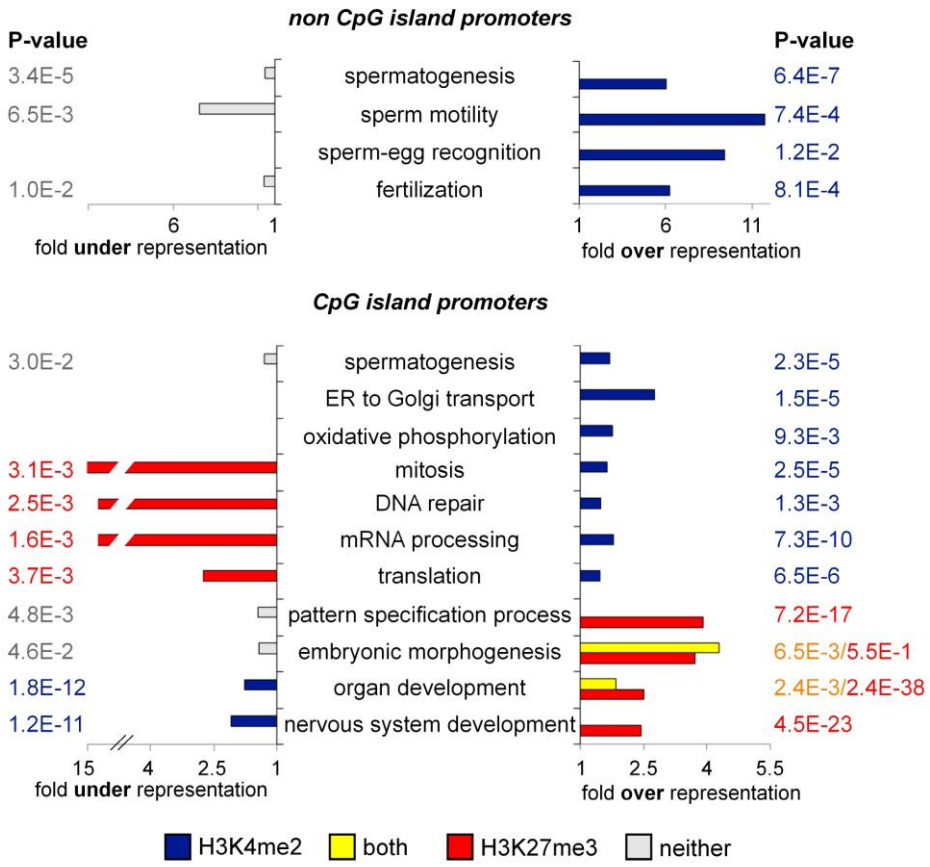


Figure 2. Gene ontology analysis of genes associated with methylated histones in human sperm. Selection of gene ontology (GO) based gene functions significantly over- and under-represented among modified promoters (in comparison to all annotated promoters on the array within a given CpG-density class). H3K4me2 and H3K27me3 occupy sets of genes with mutually exclusive functions, with spermatogenic and house keeping functions for H3K4me2-marked genes and developmental functions for H3K27me3-marked genes. All significantly over- and under-represented GO terms are listed in Supplementary Table 2.

Histone and DNA methylation are largely mutually exclusive at promoters in sperm

In mammals, paternal transmission of DNA methylation is required for imprinted gene regulation in the subsequent generation. To determine a possible interplay between histone and DNA methylation pathways during gametogenesis, we evaluated in human spermatozoa the DNA methylation status at CpG-island promoters, since DNA hypermethylation confers transcriptional repression at such promoters (Weber et al., 2007). When comparing promoters that were previously classified as either DNA methylated or unmethylated (Weber et al., 2007), we observed that both histone modifications were largely mutually exclusive with DNA methylation (Fig. 3a).

Furthermore, since developmental genes were not overrepresented among sperm targets of DNA methylation (data not shown), Polycomb and DNA methylation mark distinct gene targets in the germ line, as in soma. When analyzing DNA methylation levels without applying a defined cut off for the methylated state, we observed that DNA methylation levels were significantly lower at promoters marked by H3K4me2 than those harboring neither mark (Fig. 3b). This suggests that H3K4 and DNA methylation are largely antagonistic during spermatogenesis, a notion consistent with data observed in somatic cells (Imamura et al., 2006; Mohn et al., 2008; Weber et al., 2007).

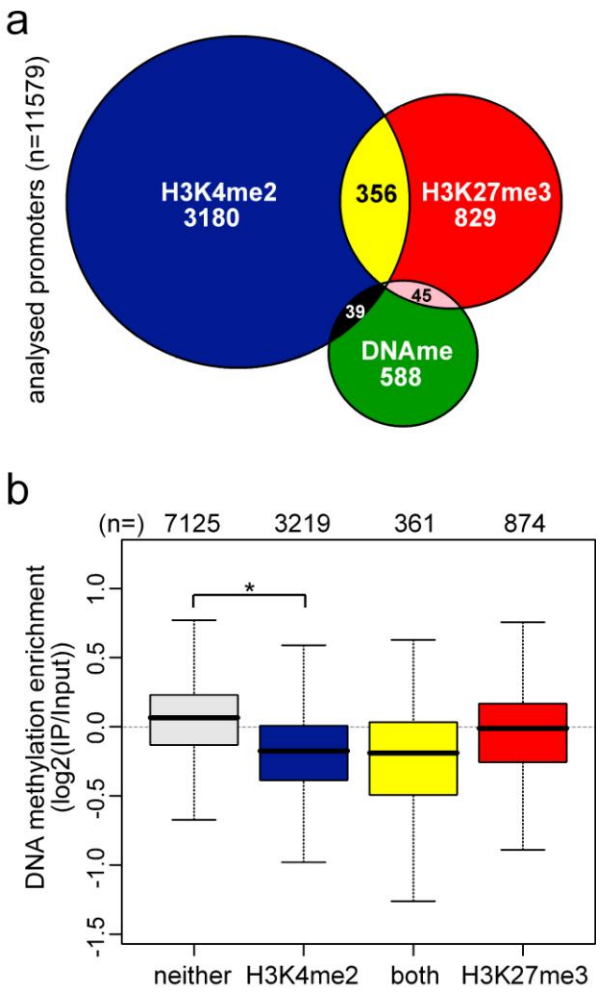


Figure 3. DNA methylation of CpG islands is mutually exclusive with H3K4 methylation in sperm. (**a**) Venn diagram illustrating low frequency of co-occupancy of histone modifications and DNA hypermethylation (Weber et al., 2007) on CpG-island promoters (see also Supplementary Fig. 5). DNA methylated promoters are significantly under-represented among promoters marked by H3K4me2 or H3K27me3 (One-sided hypergeometric test; P-value = 1.85e-58; P-value = 1.24e-3 respectively). (**b**) Boxplot showing relative enrichment for DNA methylation in sperm (Weber et al., 2007) at genes positive or negative for the tested chromatin marks (with the central bar marking the median, lower and upper limits of the box marking 25th and 75th percentiles, and the whiskers extending the 1.5 interquartile range from the box). H3K4me2-marked genes show significantly lower levels of DNA methylation than genes with neither histone mark (*: Wilcoxon test P-value < 2.2e-16).

Histone methylation status in sperm versus somatic cells

To relate genomic localization to chromatin regulation during development, we compared ChIP patterns in sperm to those generated in human embryonic stem cells (hESCs) and primary fibroblasts (Bracken et al., 2006; Pan et al., 2007; Weber et al., 2007; Zhao et al., 2007). The majority of H3K4me2 targets in sperm were equally

marked in somatic cells (Fig. 4a, Supplementary Fig. 7). These targets are associated with gene ontology functions in cellular homeostasis and gene regulation (data not shown). A significant number of H3K4me2 targets in sperm however lack this modification in hESCs and fibroblasts (blue arrow in Fig. 4a and in Supplementary Fig. 7), suggesting testis specific regulation. Consistently, these later genes are highly expressed in human spermatocytes and round spermatids (Chalmel et al., 2007) (Fig. 4b, 4c, Supplementary Fig. 8), two cell populations that represent the meiotic and haploid stages of spermatogenesis preceding the final stage of global transcriptional repression in elongating spermatids (Fig. 1a). Along similar lines, gene ontology analysis revealed a significant over-representation of spermatogenetic functions among this group of H3K4me2-marked genes (data not shown).

Figure 4a further shows that only subsets of genes containing either H3K4me2 or H3K27me3 in hESCs are marked also in sperm. Genes retaining H3K4me2 in sperm were more likely to be expressed and at significantly higher levels during spermatogenesis than genes without H3K4me2 or with H3K27me3 (Fig. 4b, 4c, Supplementary Fig. 8). These data argue that H3K4me2 in sperm largely reflects robust transcription during the final stages of spermatogenesis (Geremia et al., 1977) whereas H3K27me3 likely represents PcG-mediated transcriptional repression at preceding developmental stages.

In sperm, only 28% of H3K27me3 promoters contained also H3K4me2 (Fig. 4a). Compared to hESCs, this represents a 3-fold underrepresentation of doubly marked or “bivalent” promoters, particularly among CpG-island promoters in human sperm (Supplementary Fig. 7). These data point towards a specific regulation of H3K4 methylation at CpG-island promoters during human spermatogenesis, distinct from soma.

To obtain a closer insight into genes marked by H3K27me3 in sperm and/or in ESCs, we performed in-depth gene ontology analyses. Among the targets uniquely marked in sperm, only histone genes were over-represented (Fig. 4d). Closer analyses revealed that over 70% of the 66 canonical histone genes localized in the large histone gene clusters on chromosome 1 and 6 were marked by H3K27me3 (as well as H3K4me2) in sperm whereas in hESCs, these genes were marked by H3K4me2 only (data not shown) (Pan et al., 2007). In contrast, histone variant genes operating beyond DNA replication were not uniformly marked by H3K27me3 in human sperm while harboring H3K4me2 in hESC (data not shown). These data argue for a cluster-wide marking of canonical

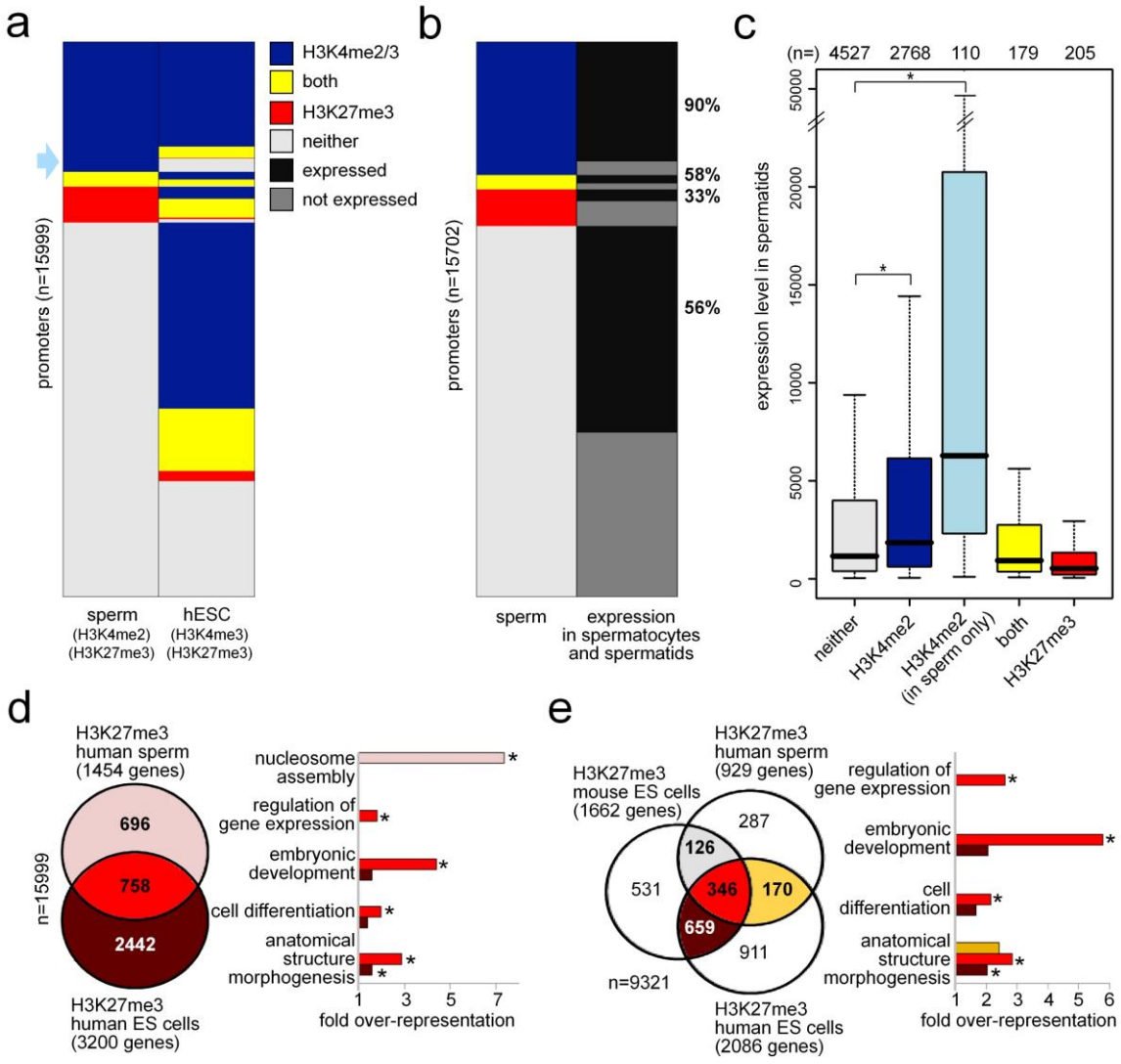


Figure 4. Spermatogenic and highly expressed genes are marked by H3K4me2 in sperm.
(**a**) Statemap showing clustering of 15'999 genes according to the chromatin status of their promoters in human sperm and hESCs (Pan et al., 2007). Arrow indicates genes marked by H3K4me2 in sperm only (see also Supplementary Fig. 7). (**b**) Statemap showing comparison of modification status at 15'702 promoters in sperm with gene expression status in human spermatocytes and spermatids (Chalmel et al., 2007). 90% of H3K4me2 promoters control genes actively transcribed in spermatogenesis. Percentages represent fractions of genes expressed. (**c**) Box plot displaying expression levels in human spermatids (Chalmel et al., 2007) for genes differentially modified in sperm. Genes under the control of promoters that are only H3K4me2 targets in sperm but not in soma (light blue, also indicated by blue arrow in panel **a** are most highly expressed. Other H3K4me2 genes (dark blue) show significantly higher levels of expression compared to genes with neither mark (*: Wilcoxon test P-value < 2.2e-16). (**d, e**) Venn diagram and GO term graphs showing over-representation of developmental gene functions among H3K27me3-marked genes shared between sperm and ESCs (Pan et al., 2007) of human (**d**) and ESCs (Mohn et al., 2008) of mouse (**e**) (* indicates GO terms with P < 1.0e-10). Supplementary Table 3 lists enrichments at all relevant GO terms.

histone genes by repressive H3K27me3 that may result from entry into meiosis and/or cell cycle exit associated with terminal differentiation of male germ cells during spermiogenesis.

Interestingly, developmental GO terms were more strongly over-represented among targets shared by sperm and hESCs compared to targets unique to hESCs (Fig. 4d; Supplementary Table 3). Furthermore, we compared targets in human and mouse ESCs and in human sperm. We observed that genes shared by all three cell types were more significantly over-represented for developmental gene functions than those genes shared by two or one cell types (Fig. 4e; Supplementary Table 3). We conclude that many PRC2 targets are evolutionary conserved between germline and embryonic stem cells of human and mouse.

Transcriptional history and potential of marked genes

To understand the origin and possible future function of modifications present in sperm, we investigated how the observed chromatin patterns relate to expression at multiple developmental time points during gametogenesis and post fertilization. Due to the absence of comprehensive expression datasets for the human germline and embryo we inferred expression states from data in mice (Namekawa et al., 2006; Zeng and Schultz, 2005). The validity of such cross-species approach was supported by direct comparative expression analyses revealing high expression levels for those orthologues expressed in human and mouse spermatocytes and/or spermatids and low expression levels for those expressed only in germ cells of one species (Supplementary Fig. 10). Only genes with one ortholog were considered and classified as inactive or active at each developmental time point. Figure 5 shows the percentage of genes that are active or never expressed at various developmental stages. Similar to the results in human (Figure 4b) the majority of H3K4me2-marked genes were expressed in mouse spermatocytes and/or spermatids (Fig. 5a). Moreover, many H3K4me2 targets were also expressed in oocytes or became activated in 2- or 8-cell embryos (Fig. 5b). Direct comparison of all developmental stages confirmed that over 65% of H3K4me2-marked genes expressed in oocytes and/or embryos were indeed expressed during spermatogenesis (Fig. 5c). These data suggest that H3K4me2 in sperm preferentially marks genes with house keeping functions, commonly expressed in the germ line and during embryogenesis.

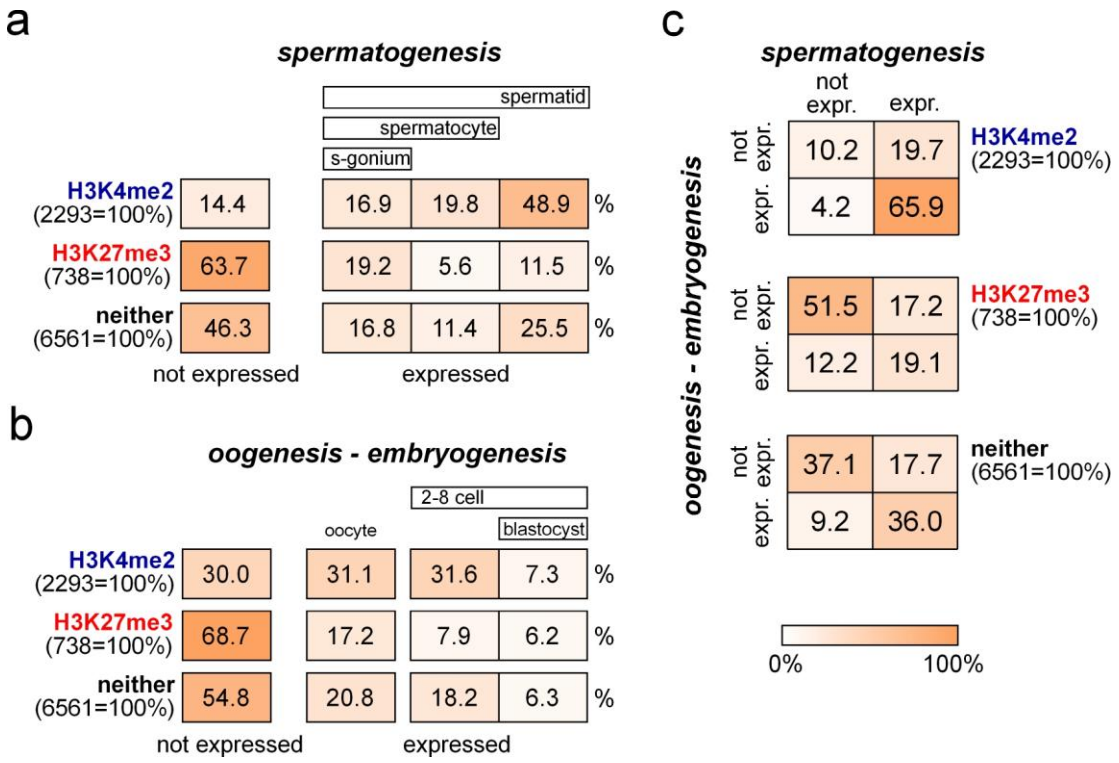


Figure 5. H3K27me3 and H3K4me2 in sperm reflect differential history and potential for expression during development. Classification of mouse genes (n=9859) according to their expression status during (**a**) spermatogenesis (Namekawa et al., 2006), (**b**) oogenesis and embryogenesis (Zeng and Schultz, 2005) (indicated by percentages and colors) in relation to the histone modification status at orthologous genes in human sperm. Genes were classified inactive or active according to (**a**) the last stage of spermatogenesis or (**b**) to the first stage of embryogenesis in which they were expressed. We used the first stage of embryonic expression as criterion for those genes expressed in oocytes and embryos. In (**c**), genes were classified according to their expression state during spermatogenesis versus oogenesis and early embryogenesis. Both H3K27me3 and H3K4me2-marked genes show significantly different distribution from genes with neither mark (Chi-square test P-value < 2.6e-14 for all comparisons). Expression classification of genes marked by both modifications is not shown. Intensity of coloring indicates percentage of genes expressed. S-gonium: spermatogonium.

Targets of H3K27me3 in sperm show opposing behavior as two out of three targets were never expressed during spermatogenesis (Fig. 5a). Almost 20 percent of H3K27me3 targets marked in sperm were expressed in spermatogonia. Expression of several of these genes, like *c-Kit*, *Stra8*, and *Dnmt3a*, has been shown to be required for that developmental stage (Anderson et al., 2008; Kaneda et al., 2004; Ohta et al., 2003). Thus, although not directly investigated in this study, Polycomb-mediated repression may dynamically regulate target genes at specific stages of germ cell development, as observed in other differentiation systems (Ezhkova et al., 2009; Mohn et al., 2008). In oocytes, the majority of H3K27me3 targets were not expressed nor did they become

activated in early embryos, reminiscent of the situation in spermatogenesis (Fig. 5b). Exceptions to this were several key regulatory genes of embryonic and extra-embryonic differentiation that were repressed during spermatogenesis and oogenesis but became activated in the early embryo such as *Cdx2*, *Elf5* and *Bmp4* (Ng et al., 2008; Strumpf et al., 2005). Over 50% of H3K27me3 targets, however, were never expressed during spermatogenesis (Namekawa et al., 2006), oogenesis (Pan et al., 2005) and early embryogenesis (Zeng and Schultz, 2005) (Fig. 5c, Supplementary Fig. 11). Even at earlier stages of gametogenesis such as developing primordial germ cells (PGC), over 90% of this group were not transcribed (Supplementary Fig. 12). Since the H3K27me3 target genes in sperm are highly enriched for key regulators of lineage specification and differentiation in soma (Fig. 2), their repressed state throughout germ cell development and in totipotent early embryos suggests that this modification may serve transgenerational gene regulatory functions.

Evolutionary conservation between human and mouse spermatozoa

If H3K4 and H3K27 methylation would indeed perform transcriptional regulatory functions across generations, we expect them to have evolutionary conserved targets in sperm of human and mouse species. A high level of conservation would imply selection for maintenance of histone modifications at promoters of specific target genes during the extensive chromatin remodeling taking place in elongating spermatids. Since global transcription is shut down in elongating spermatids and mature spermatozoa, the presence of histone methylation at selected loci could only exert its gene regulatory function after fertilization, in line with a role in transgenerational epigenetic inheritance.

To address conservation we profiled H3K4me2 and H3K27me3 at promoters of 39 mouse genes, which are orthologous to the human genes analyzed before (Fig. 1, Supplementary Fig. 3). We developed a ChIP procedure with an increased immunoprecipitation efficiency to accommodate the lower abundance of histones and the higher level of chromatin compaction in mouse spermatozoa compared to human sperm. We performed ChIP under native conditions followed by qPCR detection. For testis specific and house keeping genes, we observed, as in human, strong enrichments for H3K4me2 (Fig. 6a), consistent with their expression in spermatids. Interestingly, all testis specific genes, but none of the house keeping genes, were also marked by H3K27me3 (Fig. 6b). The observed double marking is specific to mouse sperm (Fig. 6c) and suggests that these testis-specific genes acquire H3K27me3 following their

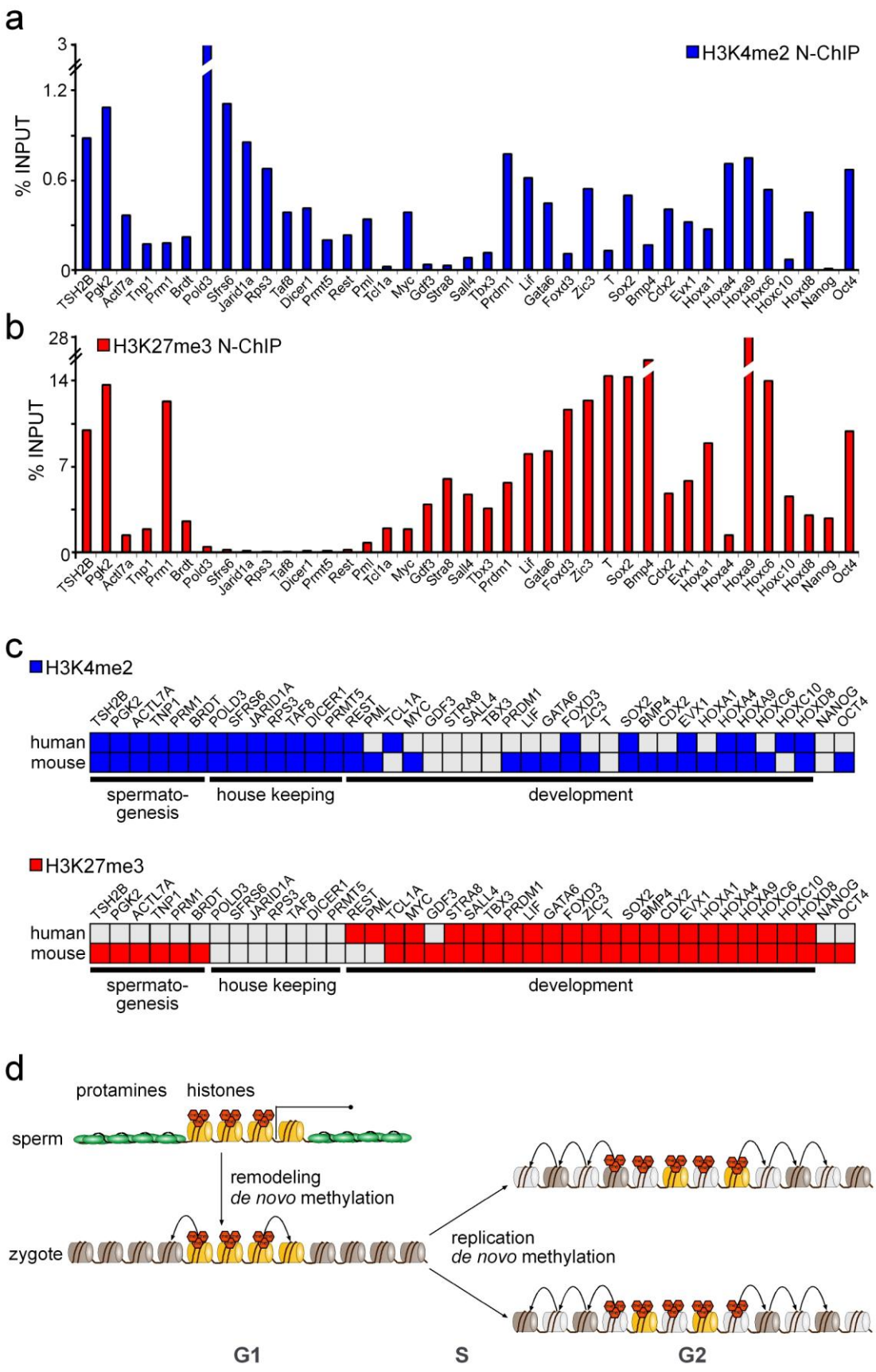

a
H3K4me2 N-ChIP
% INPUT
3
1.2
0.6
0
TSH2B Pgk2 Actl7a Tnp1 Prm1 Brdt Pold3 Sfrs6 Jarid1a Rps3 Taf8 Dicer1 Prmt5 Rest Pml Tcl1a Myc Gdf3 Stra8 Sall4 Tbx3 Prdm1 Lif Gata6 Foxd3 Zic3 T Sox2 Bmp4 Cdx2 Evx1 Hoxa1 Hoxa4 Hoxa9 Hoxc6 Hoxc10 Hoxd8 Nanog Oct4
b
H3K27me3 N-ChIP
% INPUT
28
14
7
0
TSH2B Pgk2 Actl7a Tnp1 Prm1 Brdt Pold3 Sfrs6 Jarid1a Rps3 Taf8 Dicer1 Prmt5 Rest Pml Tcl1a Myc Gdf3 Stra8 Sall4 Tbx3 Prdm1 Lif Gata6 Foxd3 Zic3 T Sox2 Bmp4 Cdx2 Evx1 Hoxa1 Hoxa4 Hoxa9 Hoxc6 Hoxc10 Hoxd8 Nanog Oct4
c
H3K4me2
TSH2B PGK2 ACTL7A TNP1 PRM1 BRDT POLD3 SFRS6 JARID1A RPS3 TAF8 DICER1 PRMT5 REST PML TCL1A MYC GDF3 STRA8 SALL4 TBX3 PRDM1 LIF GATA6 FOXD3 ZIC3 T SOX2 BMP4 CDX2 EVX1 HOXA1 HOXA4 HOXA9 HOXC6 HOXC10 HOXD8 NANOG OCT4
human
mouse
spermato-genesis
house keeping
development
H3K27me3
TSH2B PGK2 ACTL7A TNP1 PRM1 BRDT POLD3 SFRS6 JARID1A RPS3 TAF8 DICER1 PRMT5 REST PML TCL1A MYC GDF3 STRA8 SALL4 TBX3 PRDM1 LIF GATA6 FOXD3 ZIC3 T SOX2 BMP4 CDX2 EVX1 HOXA1 HOXA4 HOXA9 HOXC6 HOXC10 HOXD8 NANOG OCT4
human
mouse
spermato-genesis
house keeping
development
d
protamines
histones
sperm
remodeling
de novo methylation
zygote
replication
de novo methylation
G1
S
G2

Figure 6. Evolutionary conservation of H3K27me3 and H3K4me2-marked promoters in mouse spermatozoa. (**a**) H3K4me2 and (**b)** H3K27me3 status at promoters of 39 mouse genes, orthologous to the human genes analyzed in Fig. 1d, 1e and in Supplementary Fig. 3. Results represent percentage of material immuno-precipitated from input chromatin under native conditions as determined by real time PCR analyses. One representative out of four experiments is presented (see Supplementary Fig. 13 for all replicates). (**c**) Schematic diagram showing evolutionary conservation between human and mouse spermatozoa of H3K4me2 at promoters of spermatogenic and house keeping genes and of H3K27me3 at promoters of developmental genes. In contrast to human, promoters of mouse genes are more frequently marked by both modifications (in particular spermatogenic and some developmental genes). (**d**) Conceptual model for role of H3K27me3 in paternal transmission of epigenetic information. During spermatogenesis, several H3K27me3-modified nucleosomes (yellow) remain associated with promoters of developmental regulators. Upon gamete fusion, protamine-bound sequences undergo chromatin remodelling and incorporate maternally provided H3.3-containing nucleosomes (dark grey). H3K27me3 is subsequently established by maternal PRC2 at neighbouring nucleosomes. During following rounds of replication, H3K27me3 is maintained while H3.1/H3.2-containing nucleosomes (light grey) become incorporated, thereby enabling long-term transcriptional repression of Polycomb target genes.

expression in spermatocytes and round spermatids, possibly to safeguard their repression after fertilization. For developmental regulatory genes, most promoters tested were marked by H3K27me3, as observed in humans. However, several genes also harbored H3K4me2, suggesting that the CpG-island promoters of many developmental genes, marked by PcG-mediated H3K27me3, retain H3K4me2 during spermatogenesis in mouse but not in human. Finally, the promoters of the pluripotency factors *Oct3/4* and *Nanog* harbored both modifications in mouse but not in human sperm. *Nanog* in mouse is also DNA methylated (Farthing et al., 2008; Imamura et al., 2006) whereas both promoters are DNA methylated in human sperm (Weber et al., 2007). In summary, the H3K27me3 status at developmental genes is highly conserved between mouse and human spermatozoa (Fig. 6c). Nonetheless, these promoters in mouse are in general also marked by H3K4me2, suggesting species-specific mechanisms regulating the homeostasis of H3K4 methylation at CpG-island genes during spermatogenesis. Interestingly, H3K27me3 appears more widespread at testis specific and pluripotency genes in mouse compared to human sperm.

DISCUSSION

During sperm development in animals, histones become replaced by sperm-specific histones, protamine-like proteins or protamines. In mammals and other organisms, however, a certain fraction of histones remain present in mature spermatozoa, providing means for epigenetic inheritance (Chu et al., 2006; Dorus et al., 2006; Gardiner-Garden et al., 1998; Gatewood et al., 1987; Ooi and Henikoff, 2007; Pittoggi et al., 1999; Wykes and Krawetz, 2003). Here we demonstrate that promoters with distinct gene functions

are selectively marked by active and/or repressive histone methylation in human and mouse spermatozoa. The high level of evolutionary conservation of modified targets is compatible with a model in which retention of modified nucleosomes is subjected to selection. This model predicts that in embryos sperm-inherited modified nucleosomes remain positioned in the paternal chromatin during its remodeling by maternally provided histones in the course of pronuclear formation. Studies on the pronuclear localization of replication-dependent versus replication-independent H3 variants in early zygotes suggest that sperm-inherited histones are indeed retained within the paternal genome during pronucleus formation (Torres-Padilla et al., 2006; van der Heijden et al., 2005; van der Heijden et al., 2008).

To support transgenerational inheritance of histone methylation, marks need to be maintained during development from fertilization onwards. For H3K27 tri-methylation, the modification becomes microscopically detectable at the paternal genome concurrent with replication in the one-cell embryo (Albert and Peters, 2009; Puschendorf et al., 2008). Furthermore, the persistent presence of H3K27me3 at the originally inactive X chromosome in cloned pre-implantation embryos (Bao et al., 2005) or at the maternal genome in one-cell embryos maternally and zygotically deficient for *Ezh2* (Puschendorf et al., 2008) strongly argues for the lack of substantial zygotic and/or maternal H3K27me3 demethylase activity in early embryos. Together, these studies support a model in which paternally inherited H3K27me3 can be transmitted through subsequent pre-implantation development (Fig. 6d).

Recently, the catalytic activity of Ezh2 was shown to be required for repression over multiple cellular generations, suggesting a role for H3K27me3 in epigenetic memory (Hansen et al., 2008). Presence of H3K27me3 in sperm also correlated positively with gene repression in the male and female germ line as well as during early embryogenesis. Therefore, our study not only suggests transmission of modification-specific patterns of histone methylation via spermatozoa but also opens the possibility that Polycomb represses somatic determinants in the male germ line and in early embryos thereby contributing to the propagation of totipotency across generations (Fig. 6d).

We speculate that Polycomb serves a similar regulatory function at orthologous maternal alleles during oogenesis and early embryogenesis. If maternal alleles of the H3K27me3 targets identified in sperm would be differentially regulated, the situation would resemble classical imprinting. However, the undisturbed embryonic patterning

observed in gynogenones and parthenogenones argues against such a scenario. Furthermore, live-born offspring with two maternal genomes are obtained at a respectable frequency by genetic manipulation of only two imprinted loci (Kawahara et al., 2007). Therefore, there is no strong argument for a restriction of the transgenerational contribution by Polycomb to the regulation of developmental genes on the paternal genome only.

For H3K4 dimethylation, it is unknown whether the mark functions only in the process of transcription or whether it also serves a role in epigenetic memory of the active state in proliferating cells. Nuclear transfer experiments performed in *Xenopus* oocytes provided evidence for a role of lysine 4 if histone H3 in transcriptional memory (Ng and Gurdon, 2008). In *C. elegans*, deficiency for the H3K4me2 demethylase Lsd1/KDM1 caused a progressive sterility over many generations that correlated with transgenerational accumulation of H3K4me2 in the germ line and increased expression of spermatogenic genes in the soma (Katz et al., 2009). These data argue that programmed H3K4 demethylation, possibly of testis-expressed genes, is required for germ line immortality in *C. elegans.* In mouse and human embryos, H3K4 methylation is established along the paternal genome within the first cell cycle (van der Heijden et al., 2005; van der Heijden et al., 2009) providing means for somatic transmission. Nevertheless, the fate of germ-line inherited H3K4me2 at e.g. housekeeping versus testis-specific genes remains to be tested.

Molecular genetic experiments will be required to elucidate the extent and functional significance of methylation at distinct histone residues for transgenerational inheritance. There may exist variability in establishment and retention of modified histones between spermatozoa and/or individuals, possibly in response to environmental influences and/or innate cues, such as incomplete chromatin remodeling during spermatid elongation. Hence, transmission of histone encoded epigenetic information may constitute a novel transgenerational mechanism for phenotypic variation (Chong et al., 2007a).

METHODS

Samples collection and purification

Human sperm samples were obtained from normospermic men visiting the University Medical Center St Radboud, the Netherlands, for routine semen diagnosis. Sperm morphology was assessed using established criteria (Menkveld et al., 2001) and the diagnosis of normospermia was based on criteria of the World Health Organization (*WHO Laboratory Manual* 1999). All donors signed a written informed consent for participation in the study. Sperm samples were collected in sterile containers and purified by three rounds of washing with human tubal fluid medium (HTF; Cambrex, Verviers, Belgium) and density gradient centrifugation (20 min, 500g) using Pure Sperm solution (Nidacom). Purified sperm was then diluted 1:1 with TEST yolk buffer medium (TYB, Irvine Scientific, CA, USA), cooled in vapor phase of liquid nitrogen for 15 min. and subsequently stored in liquid nitrogen.

Motile mature spermatozoa were obtained from CD1 and C57BL/6J (for Western Blot and ChIP respectively) mice by allowing spermatozoa to swim out of caudal epididymal tissue for 1hr at 37°C into sperm motility medium (135 mM NaCl, 5 mM KCl, 1 mM MgSO4, 2 mM CaCl2, 30mM Hepes pH7.4; freshly supplemented with 10 mM lactate acid, 1 mM sodium pyruvate, 20 mg/ml BSA, 25 mM NaHCO3). To avoid contamination of somatic cells, only top fractions containing about $3x10^6$ motile sperm per mouse were used for further assays.

Sample purity was verified by microscopy and by western blot analyses assaying for the absence of presence of Lamina-associated polypeptide 2β (LAP2β) (Fig. 1b), a marker for somatic and immature germ cells (Alsheimer et al., 1998).

Western Blot

Proteins from murine and human sperm were isolated as described (Lee et al., 1995) with minor modifications. Additional steps of sonication (2 times 30 s 40%, Branson sonicator), and extraction with 1.6 M Urea, 1 M NaCl and 0.28 M β-mercaptoethanol for 30 min. at 37°C were included. After precipitation with 20% trichloroacetic acid protein pellets were boiled for 20 min. in SDS sample buffer and separated by SDS-PAGE. Western blot analyses were performed using following antibodies and dilutions: polyclonal H3K4me2 (Upstate #07030) (1:1000), polyclonal H3K27me3 (Upstate #07449) (1:1000), polyclonal H3 (Abcam #17921) (1:10 000), monoclonal LAP2β

(Dechat et al., 1998) (1:5). Protein extracts from WI38 human primary lung fibroblasts and CCE ES cells were used as controls and were prepared as described above.

Crosslinked chromatin immuno precipitation (X-ChIP)

ChIP-chip experiments were performed on a pool of 9 donor samples to average possible variability between individuals. H3K4me2 and H3K27me3 ChIPs were carried out in parallel on identical sets of samples.

Per ChIP, $2x10^7$ spermatozoa were used. After thawing, pooled samples were washed with PBS to remove cryo-preservation medium (5 min., 800g). ChIP experiments were performed as described before (Weber et al., 2007) with several modifications. Fixation was performed with 0.5% paraformaldehyde for 10 min at room temperature (RT). Lysis was performed in the presence of 0.5% SDS and 10 mM DTT, for 1 h at RT. N-Ethylmaleimide (30 mM) was added to quench DTT and the samples were diluted 2.5 times prior to sonication. Sonication was performed six times for 20 s (Branson sonicator, amplitude 70%) to obtain chromatin with fragment sizes of 300-700 bp.

Sperm chromatin was then used for immuno-precipitation at 4°C overnight with 5 μg of antibody - H3K4me2 (Upstate #07030) or H3K27me3 (Upstate #07449). Following steps included incubation with protein A-Sepharose beads and washing as described in (Weber et al., 2007). Cross-link reversal, DNA isolation and amplification with WGA2 amplification kit (Sigma) was than performed according to (O'Geen et al., 2006). For amplification, 50 ng of input DNA and entire ChIP DNA were used. For each H3K27me3 array experiment 3 simultaneously prepared ChIP samples were pooled, and used for the amplification. For each H3K4me2 array experiment, 1 ChIP sample was used for amplification. A set of four genes was tested for each sample by quantitative PCR and showed similar bound-to-input ratios before and after amplification.

Validation of micro-array results was performed by ChIP-real time PCR analyses using SYBR Green PCR Master Mix (Applied Biosystem) and ABI Prism 7500 Real-time PCR machine and is presented in Supplementary Fig. 3 (for list of primers see Supplementary Table 4). ChIP was performed on pools of sperm obtained from donors different from those used for ChIP-chip analyses.

Native chromatin immuno precipitation (N-ChIP)

Native ChIP on human and mouse sperm was performed according to the protocol by Umlauf and colleagues (http://www.epigenome-noe.net/researchtools/protocol.php?protid=22) with modifications (Umlauf et al., 2004). For one ChIP on human sperm, 3 donor samples were pooled. For both mouse and human $1x10^7$ spermatozoa were used per one ChIP. Prior to ChIP, mouse spermatozoa were treated with 50mM DTT in PBS at room temperature for 2 hours, followed by N-Ethylmaleimide treatment and washing with PBS. Subsequently, human and mouse spermatozoa were lysed in Buffer I (0.3M Sucrose, 15mM Tris (pH 7.5), 60mM KCl, 15mM NaCl, 5mM MgCl2, 0.1mM EGTA, 0.5mM DTT) containing 0.5% deoxycholate and 0.25% NP-40 for 10 min on ice. Chromatin digestion was performed by micrococcal nuclease as described (Umlauf et al., 2004). Immuno-precipitation was than carried out with H3K4me2 (Upstate #07030) or H3K27me3 (Upstate #07449) antibodies following the published protocol (Umlauf et al., 2004). Real-time PCR was performed using SYBR Green PCR Master Mix (Applied Biosystem) and ABI Prism 7500 Real-time PCR machine (for list of primers see Supplementary Table 4).

Microarray design and analysis

ChIP samples were hybridized to human promoter tiling microarray (2006-07-18_HG18_RefSeq_promoter, NimbleGen Systems Inc.), representing 18029 promoter regions spanning on average 2200bp upstream and 500bp downstream of transcription start sites of all RefSeq annotated genes. Probes on the array were isothermal oligos, 50-75 nt long located with 100bp spacing. Sample labeling, hybridization and array scanning were performed by NimbleGen Systems Inc. according to standard procedures. Three independent ChIP-chip experiments were performed for each modification. One array hybridization was performed in a dye swap configuration.

All subsequent analyses were performed using R software (www.r-project.org). Raw fluorescence values were used to calculate log2 of precipitated/input ratios for each probe. Loess normalization (Limma package, R, (Smyth, 2005)) was used to correct for labeling dye artifacts. To remove noise coming from low signal range, probes with average intensity (A= 1/2(log2(IP)+log2(IN))) lower than 10 for H3K27me3 and 9.5 for H3K4me2 were removed from the analysis. After these corrections Pearson correlation coefficients R between the replicates were 0.56, 0.57 and 0.67 for H3K27me3 replicates and 0.72, 0.82 and 0.82 for H3K4me2 replicates.

To identify regions associated with analyzed modifications the genomic intervals represented on the microarray were classified into enriched and not-enriched segments using Hidden Markov Models (HMMs), as described (Birney et al., 2007). The basic premise of HMMs is that observed data are generated stochastically from a pre-determined number of hidden background probability distributions, or states. Here, we used a three-state gaussian emission HMM that was trained on each array using Baum-Welch algorithm to estimate model parameters (the states corresponding to not-enriched, weekly enriched and strongly enriched segments), and all oligos with posterior probabilities larger than 0.8 for the strongly enriched state were classified to be associated to the analyzed modification. Because of the average chromatin fragment size (400-600 bp) and the resulting limited resolution of the ChIP assay, peaks shorter than 300bp were removed from further analysis. The segmentation algorithm was implemented in Python using the GHMM library (Schliep et al., 2004). As expected, the strongly enriched regions identified by this HMM approach typically contained probes with high enrichment values (see Supplementary Fig. 2).

Peak finding algorithm was performed independently for each of 3 replicates. Peaks were associated to the closest Ensembl (release 48, genome build hg18, www.ensembl.org) annotated transcription start site (2.2 kb upstream, 0.5 kb downstream). Any gene associated to at least one peak was called positive and set of targets being positive in all 3 out of 3 replicates was generated for each modification. We performed subsequent bioinformatic analyses only on those genes that were enriched in H3K4me2 and/or H3K27me3 in all three replicates. Lists of genes are provided in Supplementary Table1. Schematic representation of the analysis work flow is provided in Supplementary Fig. 1.

Comparison to the ChIP–Seq data and the nucleosomal sequencing data in human sperm reported by (Hammoud et al., 2009).

We downloaded the sequencing data from (Hammoud et al., 2009), (GEO identifier: GSE15690) and extracted all read sequences from the ELAND alignments. Reads were realigned to the hg18 human genome assembly using Bowtie (Langmead et al., 2009) allowing for up to two mismatches. Similarly, probe sequences of the NimbleGen hg18 RefSeq Promotor array were aligned to the genome. In order to calculate ChIP-seq and ChIP-chip promotor enrichments, we used the genomic intervals corresponding to putative promoter regions that are tiled on the NimbleGen array, and either counted the

number of aligned reads, or determined the average enrichment ratio of oligos mapping to each of these regions. Read counts were normalized for the total number of aligned reads in each sample, and divided by the number of reads found in a control sample as described in (Hammoud et al., 2009). The resulting enrichments were plotted against each other on a log2 scale. Promoter regions that were identified being enriched (with identified peaks) in all three ChIP-chip replicates (this study) are indicated in blue (H3K4me2) or red (H3K27me3) (Supplementary Fig. 4a-d). Venn diagrams (Supplementary Fig. 4e, 4f) are based on published lists of enriched promoters (Hammoud et al., 2009) which were compared to lists of enriched promoters identified in the current study

CpG class annotation

Promoter regions on the array were assigned to three CpG classes (LCP, ICP and HCP: low intermediate and high CpG-content promoters) based on published criteria (Weber et al., 2007). In the text LCP promoters are referred to as non CpG island promoters and ICP and HCP promoters together as CpG island promoters.

GO term analysis

Gene ontology (GO) analyses were performed using GO Stat (Beissbarth and Speed, 2004), (http://gostat.wehi.edu.au). Complete lists of significantly over- and under-represented GO terms for Fig. 2 are provided in Supplementary Table 2 and for Fig. 4d and 4e in Supplementary Table 3.

Matching to published data sets

Schematic representations of all matching steps are provided in Supplementary Fig. 1 for ChIP, MedDIP data and expression in human and in Supplementary Fig. 9 for expression in mouse. For each comparison a master data set of genes common to all array platforms used in different studies was determined and only these genes were used in subsequent analyses.

Comparison to maps of DNA methylation in human sperm and H3K4me2 in human fibroblasts

To match MeDIP measurements of DNA methylation in human sperm and H3K4me2 ChIP measurements in human fibroblasts (Weber et al., 2007) to current data set, we re-

annotated the promoter coordinates of Weber and colleagues (Weber et al., 2007) to the new genome build coordinates (hg 17 to hg18 conversion) and Ensembl transcription start sites within these regions. 5meC and H3K4me2 enrichment measurements were mapped to current study via Ensembl gene ids. Promoters were assigned to CpG-density classes as determined in (Weber et al., 2007). DNA hypermethylation status on ICP and HCP promoters was determined by applying a threshold for 5meC log2 ratio >0.4 as described (Weber et al., 2007). H3K4me2 status was determined by applying a threshold log2 ratio > -1.5, based on the bimodal distribution of the enrichment values allowing to distinguish between positive and negative fractions of promoters.

Comparison to maps of H3K4me3 and H3K27me3 in human ES cells and of H3K27me3 in human fibroblasts and mouse ES cells

Published lists of genes associated with modified histones accompanying (Bracken et al., 2006; Mohn et al., 2008; Pan et al., 2007; Zhao et al., 2007) studies were annotated with Ensembl gene ids and matched to current data set. H3K27me3 mouse ES data set (Mohn et al., 2008) was matched based on mouse orthologs of the genes on the array (see below).

Matching human to mouse genes

For comparison to mouse expression data sets, human genes on the array were annotated with corresponding mouse orthologs, based on Ensembl criteria (www.ensembl.org/info/docs/compara/homology_method.html). Only genes with one-to-one orthology were used in the analysis (13167 of 18152 human Ensembl annotated genes). The approach to compare chromatin modifications at genes in human sperm to expression of orthologous genes in spermatogenic, oogenic and embryonic cells of mouse origin is supported by recent comparative expression studies showing high conservation of expression patterns between mouse and human orthologous genes in many tissues, including testis (Jordan et al., 2005; Xing et al., 2007).

Defining expression states

To analyze expression status of genes during different stages of gametogenesis (primordial germ cell development, oogenesis and spermatogenesis) and embryogenesis, we processed data from publicly available Affymetrix CEL files (Chalmel et al., 2007; Kurimoto et al., 2008; Namekawa et al., 2006; Pan et al., 2005; Zeng and

Schultz, 2005) using Genedata's Expressionist pro 5.0 (Genedata AG). Expression values were estimated using the RMA-Bioconductor function (Irizarry et al., 2003) and their distributions were standardized by quantile normalization and scaled by transforming the median expression value to 20.

Probesets with a detection P-value < 0.04 (Affymetrix default) in both replicates (Chalmel et al., 2007) or in at least three out of four replicates (Kurimoto et al., 2008; Namekawa et al., 2006; Pan et al., 2005; Zeng and Schultz, 2005) were considered to be expressed and annotated with Ensembl gene ids. Lists of genes expressed at a given stage were compared to current data set using the Ensembl gene id for mapping. For analysis of absolute expression levels the mean of replicates was calculated for each probeset and the highest value was taken in case of multiple probesets per gene.

Defining expression profiles

To identify classes of genes with similar expression profiles during spermatogenesis and embryogenesis, we first selected probesets with an expression value >20 (embryogenesis) or >40 (spermatogenesis) in at least one developmental stage. Probes with significantly changing expression levels between developmental states (P-value < 0.05 in N-way ANOVA analysis) were assigned to different expression profiles using the self organizing map (SOM) clustering algorithm. False discovery rate was estimated using a Benjamini-Hochberg test to correct P-values. To distinguish between maternally provided transcripts and *de novo* transcription in 2-cell stage embryos, we compared expression levels in embryos treated and untreated with the transcription inhibitor α-amanitin (Zeng and Schultz, 2005). We combined profiles with similar changes in expression states during development. Probesets in groups of combined profiles were annotated with Ensembl gene ids. Groups were compared to current data set based on mouse Ensembl gene ids.

Statistical analyses

Statistical tests were performed in R, using two-sided Wilcoxon rank sum test (Fig. 3b and 4c) as a non-parametric test of location for non-normal data. In Fig. 3a, the one-sided hypergeometric test (R "Phyper" function) was used to measure the probability of observing an overlap equal or smaller to the one obtained from the real data. Associations in Fig. 5 were tested using Pearson's Chi-squared test on the raw count

data. The P values reported for enriched GO terms (Fig. 2, 4d, 4e) were obtained using GO Stat (http://gostat.wehi.edu.au).

Author contributions

U.B. and A.H.F.M.P. conceived and designed the experiments. U.B. and M.H. performed the experiments. L.R. provided purified samples of human spermatozoa. U.B., M.H., M.B.S. and A.H.F.M.P. analyzed the data. T.C.R and D.S. provided advice on data analyses and the manuscript. U.B. and A.H.F.M.P. wrote the manuscript.

Acknowledgements

We thank Thomas Jenuwein and Roland Foisner for providing antisera. We are grateful to Fred Zilbermann and Carolin Kolb and the FMI animal facility for excellent technical assistance. We thank Peter de Boer and Remi Terranova for critical reading of the manuscript and all members of the Peters and Schübeler laboratories for fruitful discussions. U.B. acknowledges the Boehringer Ingelheim Fonds for her PhD fellowship. T.C.R is supported by the Swiss Cancer Leagues. Research in the Peters lab is supported by the Novartis Research Foundation, the Swiss National Science Foundation, the European Network of Excellence “The Epigenome” and the EMBO YIP program.

2.2 Supplementary data

Supplementary information to the manuscript entitled:

Repressive and active histone methylation mark distinct promoters in human and mouse spermatozoa

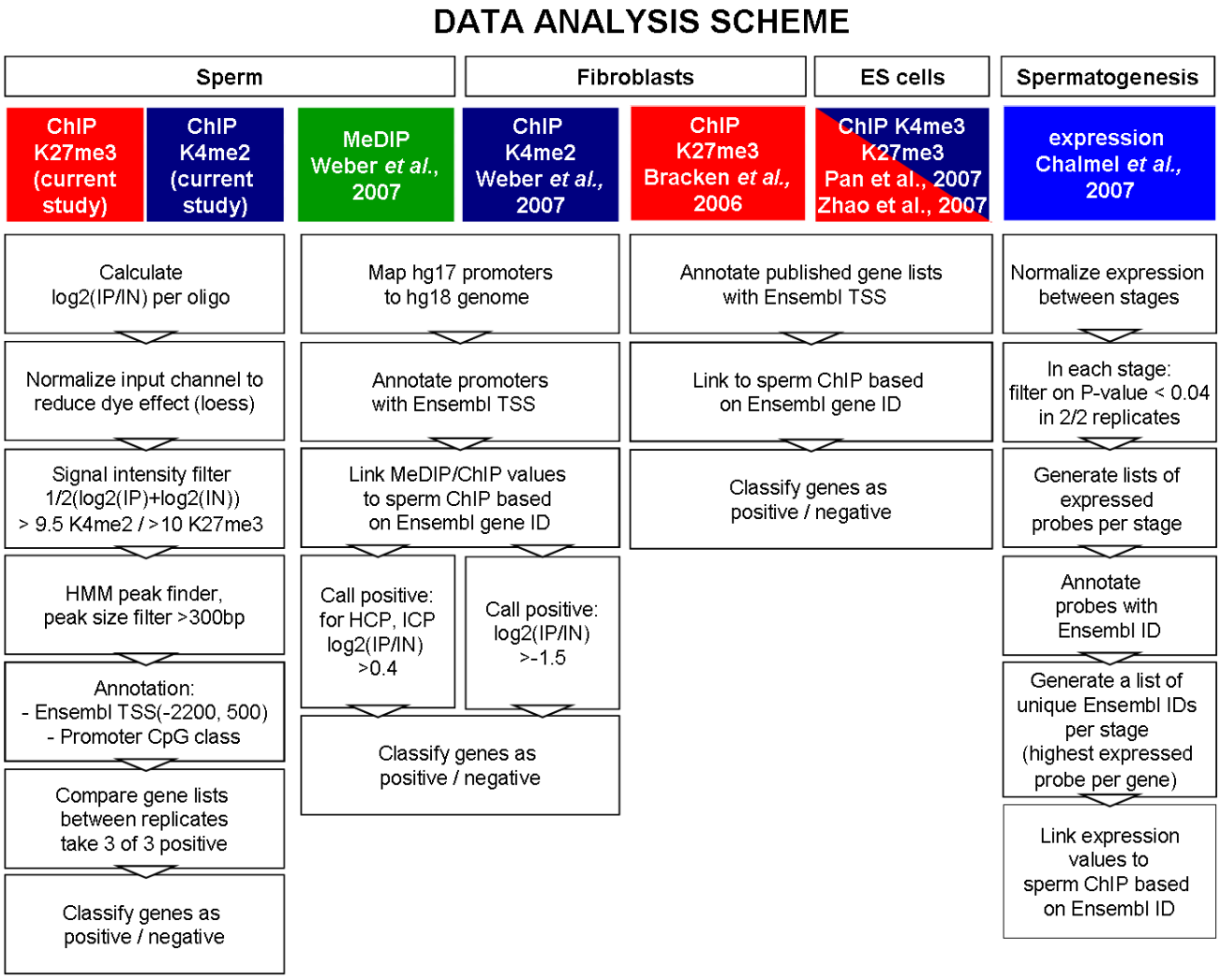
Urszula Brykczynska, Mizue Hisano, Liliana Ramos, Edward J. Oakeley, Tim C. Roloff, Dirk Schübeler, Michael B. Stadler and Antoine H.F.M. Peters

Supplementary Figures

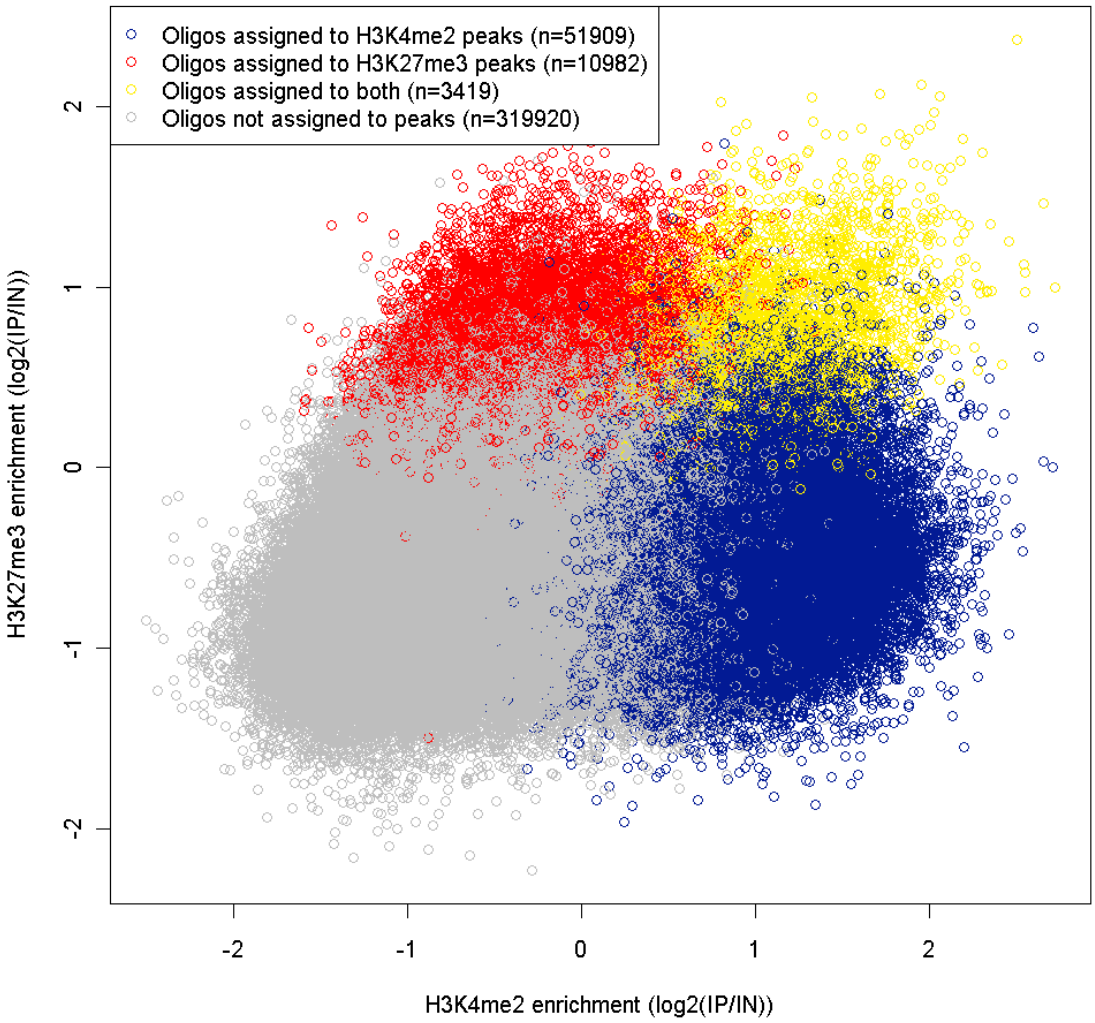
1: Data analysis scheme.
2: Validation of HMM peak prediction.
3: ChIP real-time PCR validation of genomic array data.
4: Comparison to histone methylation states and nucleosomal association reported by Hammoud et al., 2009.
5: Classification of H3K4me2 and H3K27me3 targets according to CpG density of underlying promoter sequence.
6: Number of modified gene promoters per chromosome.
7: Comparison of histone methylation states in human sperm, hESCs and human fibroblasts.
8: Levels of expression in human spermatocytes.
9: Data analysis scheme of comparison to mouse expression data.
10: Validation of comparison between human and mouse expression programs.
11: Majority of genes not expressed during spermatogenesis, in fully grown oocytes and during early embryogenesis are also not expressed during oogenesis
12: Majority of genes not expressed during spermatogenesis, in fully grown oocytes and during early embryogenesis are also not expressed in primordial germ cells
13: ChIP real-time PCR analysis of modified genes in mouse spermatozoa

Supplementary Tables (due to large file sizes these data are not provided in the thesis)

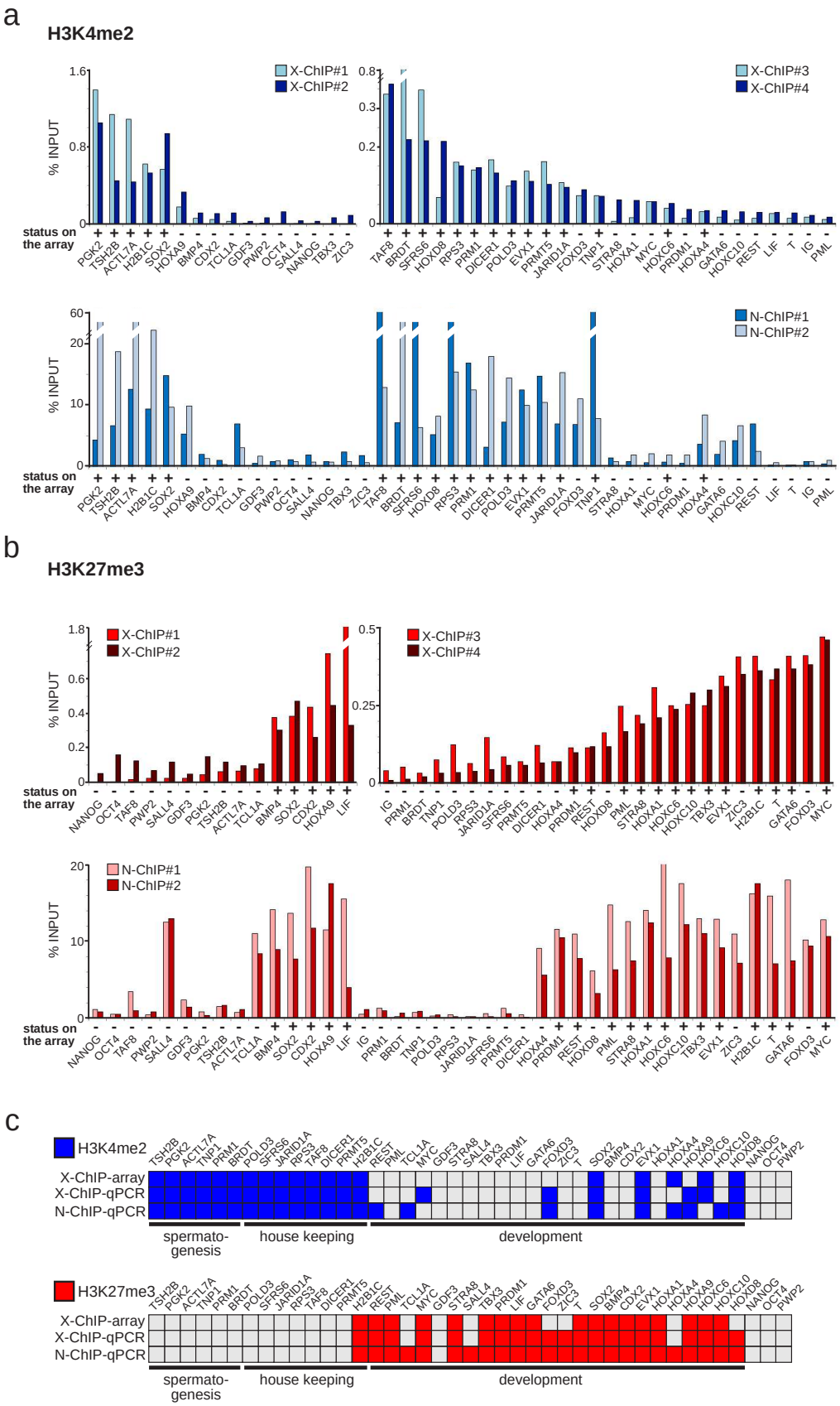
1 List of genes with modification states in human spermatozoa
2: List of significantly over- and under-represented GO terms for Fig. 2
3: List of significantly over- and under-represented GO terms for Fig. 4d, e
4: Real-time PCR primer sequences



Supplementary Fig. 1: Data analysis scheme. Flow chart showing procedures used to identify modified promoters in human sperm and to compare modified promoter sets to published global histone modification and expression data (see Methods for detailed description).

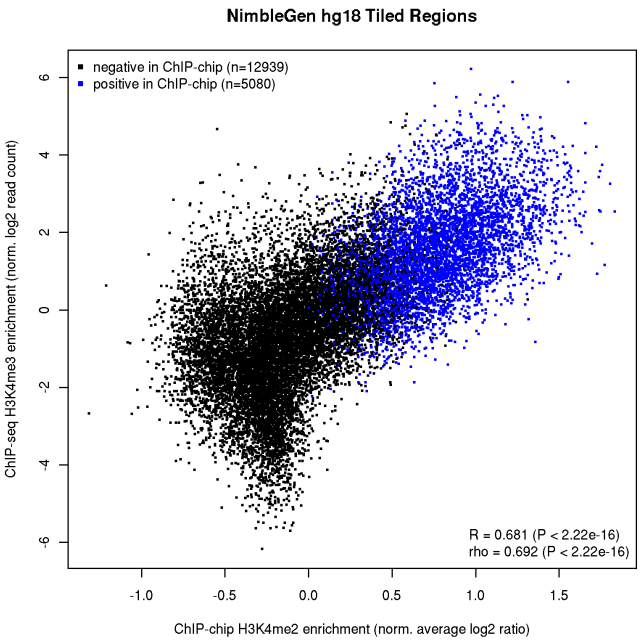


Supplementary Fig. 2: Validation of HMM peak prediction. Scatter plot presenting raw H3K4me2 and H3K27me3 log2 enrichment ratios per oligo (mean average of 3 replicates). Colored circles represent oligos assigned to peaks by HMM peak finding algorithm (see Methods for details).

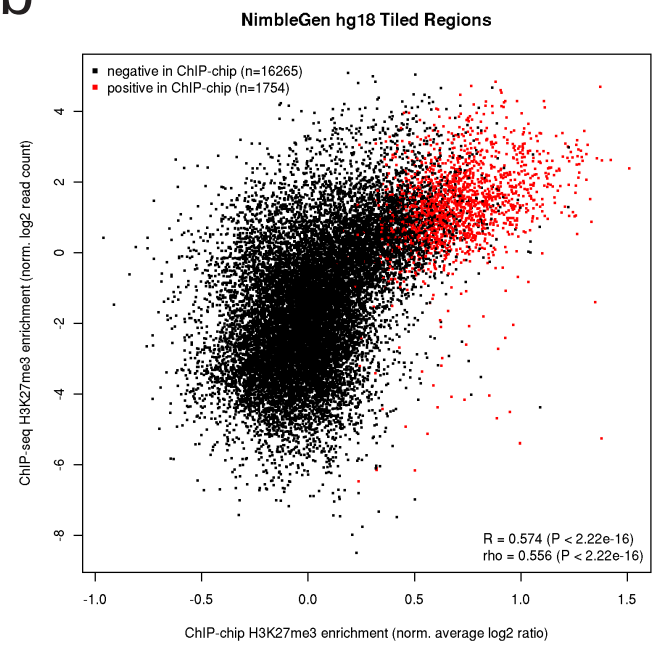


Supplementary Fig. 3: ChIP real time PCR validation of genomic array data. (**a, b**) Promoters of 41 genes and an intergenic negative control (IG) were analyzed for relative enrichments of (**a**) H3K4me2 and (**b**) H2K27me3. For targets, scored positive on the array, real time PCR amplicons overlap with identified peaks. We analyzed each gene in two independent cross-linked ChIP (X-ChIP) experiments and in two independent native ChIP (N-ChIP) experiments. Data are presented as % of immuno-precipitated input chromatin. Linear scaling was applied for the replicas as presented in the same graphs with following multiplication factors: (**a**) X-ChIP#1 times 1, X-ChIP#2 times 3, X-ChIP#3 times 1, X-ChIP#4 times 1, N-ChIP#1 times 1, N-ChIP#2 times 2, (**b**) X-ChIP#1 times 1, X-ChIP#2 times 2, X-ChIP#3 times 1, X-ChIP#4 times 1, N-ChIP#1 times 1, N-ChIP#2 times 2. A selection of results from X-ChIP#1 and #3 and N-ChIP #1 and #2 is shown in Figures 1d and 1e.(**c**) Schematic diagram showing consistency of array results with the real-time PCR validation by two ChIP methods.

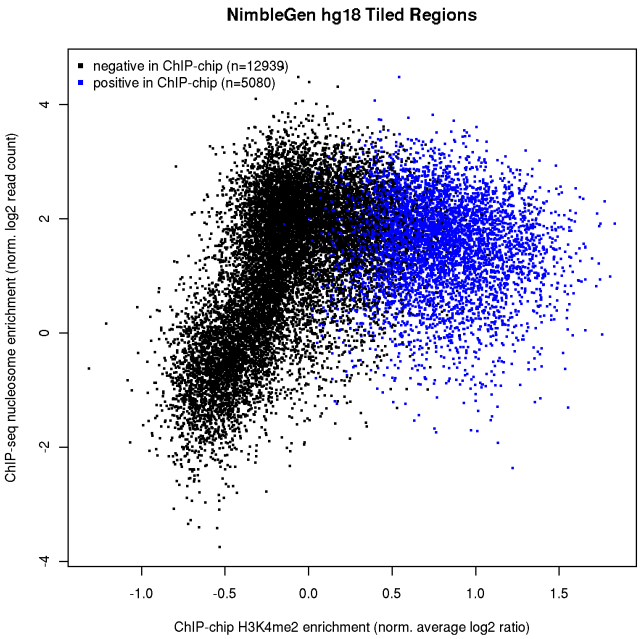
a


NimbleGen hg18 Tiled Regions
negative in ChIP-chip (n=12939)
positive in ChIP-chip (n=5080)
R = 0.681 (P < 2.22e-16)
rho = 0.692 (P < 2.22e-16)
ChIP-seq H3K4me3 enrichment (norm. log2 read count)
ChIP-chip H3K4me2 enrichment (norm. average log2 ratio)

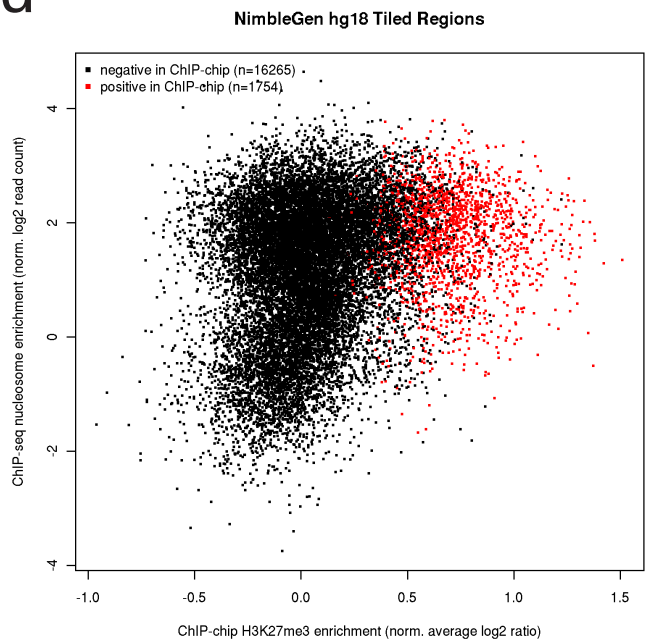
b


NimbleGen hg18 Tiled Regions
negative in ChIP-chip (n=16265)
positive in ChIP-chip (n=1754)
R = 0.574 (P < 2.22e-16)
rho = 0.556 (P < 2.22e-16)
ChIP-seq H3K27me3 enrichment (norm. log2 read count)
ChIP-chip H3K27me3 enrichment (norm. average log2 ratio)

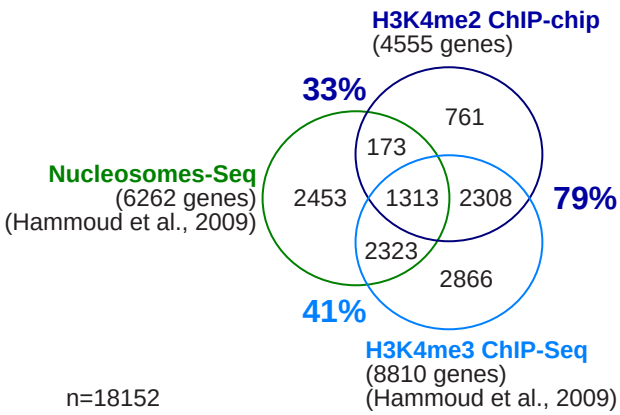
c


NimbleGen hg18 Tiled Regions
negative in ChIP-chip (n=12939)
positive in ChIP-chip (n=5080)
ChIP-seq nucleosome enrichment (norm. log2 read count)
ChIP-chip H3K4me2 enrichment (norm. average log2 ratio)

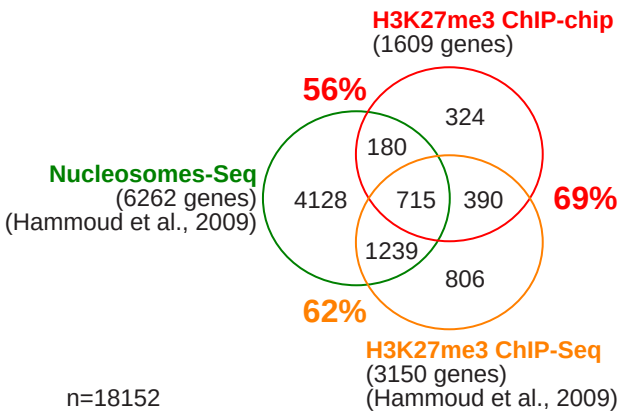
d


NimbleGen hg18 Tiled Regions
negative in ChIP-chip (n=16265)
positive in ChIP-chip (n=1754)
ChIP-seq nucleosome enrichment (norm. log2 read count)
ChIP-chip H3K27me3 enrichment (norm. average log2 ratio)

e


H3K4me2 ChIP-chip
(4555 genes)
33%
761
173
Nucleosomes-Seq
(6262 genes)
(Hammoud et al., 2009)
2453
1313
2308
79%
2323
2866
41%
H3K4me3 ChIP-Seq
(8810 genes)
(Hammoud et al., 2009)
n=18152

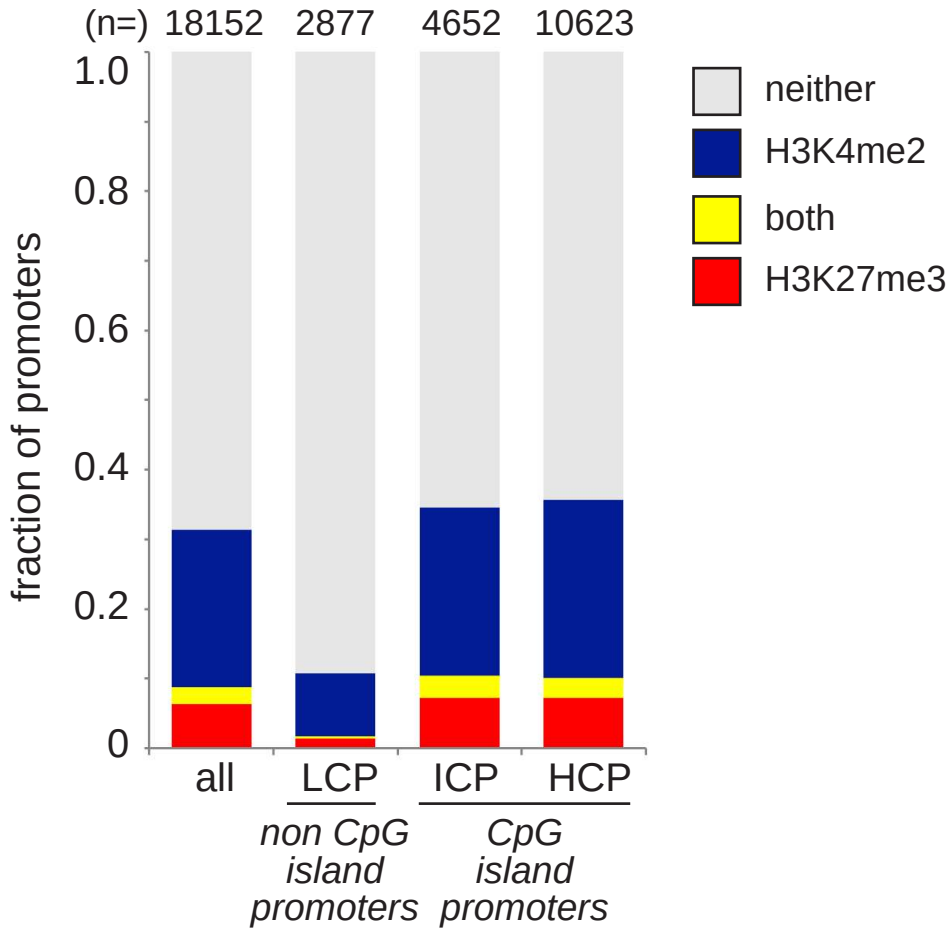
f


H3K27me3 ChIP-chip
(1609 genes)
56%
324
180
Nucleosomes-Seq
(6262 genes)
(Hammoud et al., 2009)
4128
715
390
69%
1239
806
62%
H3K27me3 ChIP-Seq
(3150 genes)
(Hammoud et al., 2009)
n=18152

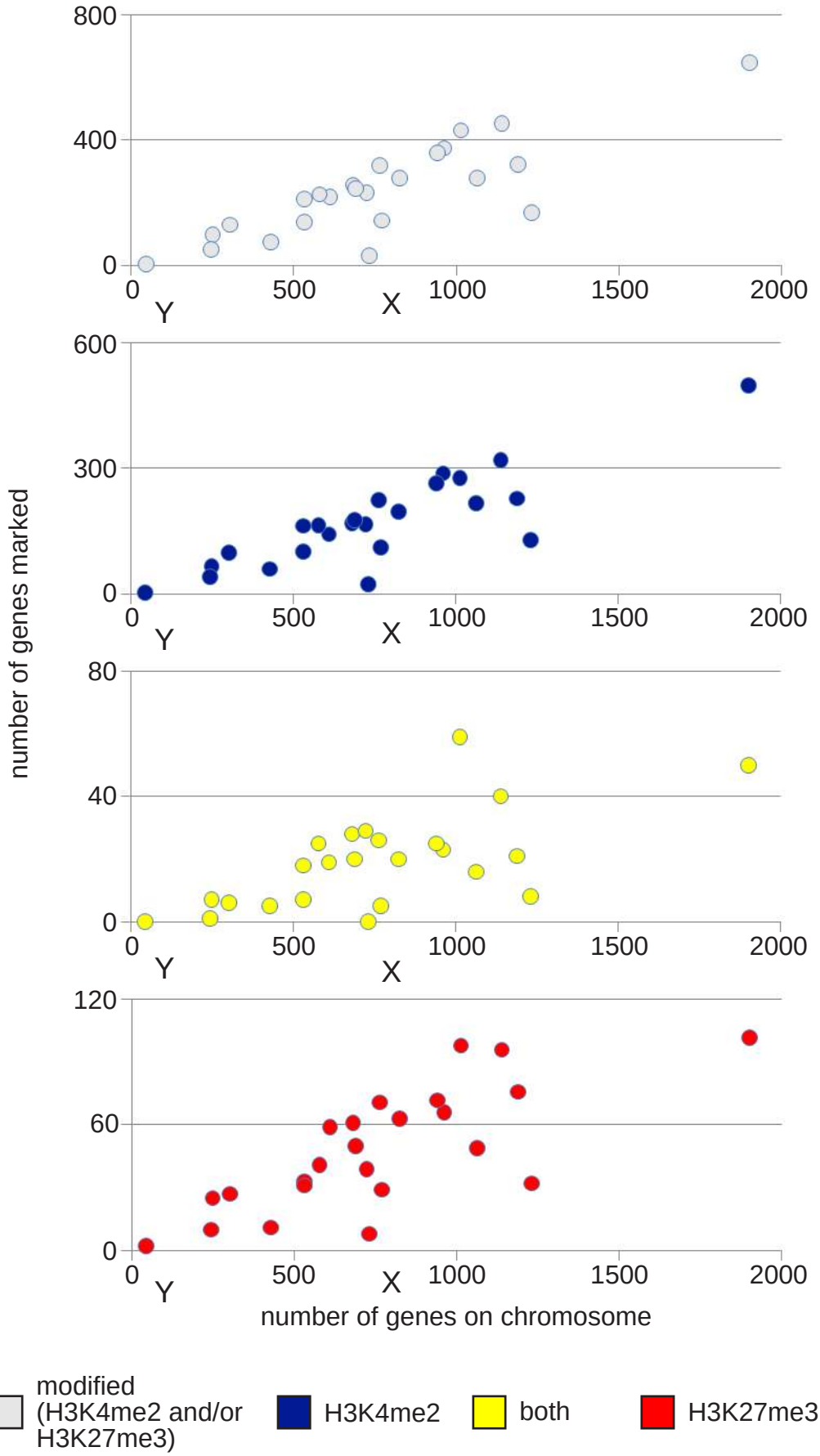
Supplementary Fig. 4: Comparison to histone methylation states and nucleosomal association reported by (Hammoud et al., 2009).

(**a, b**) Scatter plots showing correlation of normalized log2 number of aligned reads obtained by ChIP-Seq (Hammoud et al., 2009) and average log2 enrichment ratios of oligos obtained by ChIP-chip (current study), mapping to tiled promoter regions on the NimbleGen array (each point represents one promoter region), for H3K4me2/3 (**a**) and H3K27me3 (**b**). Correlation measures R (Pearson's correlation coefficient), and rho (Spearman's rank correlation coefficient), indicate high correlation between the two studies performed using different ChIP and detection techniques. Promoter regions that were identified being enriched in all three ChIP-chip replicates (this study) are indicated in blue (H3K4me2) or red (H3K27me3). (**c, d**) Scatter plots showing correlation of normalized log2 number of aligned reads obtained by sequencing of isolated nucleosomes (Hammoud et al., 2009) and average log2 enrichment ratios of oligos obtained by ChIP-chip for H3K4me2 (**c**) and H3K27me3 (**d**) (current study), mapping to tiled promoter regions on the NimbleGen array (each point represents one promoter region). Plots show that majority of promoters detected to be associated with H3K4me2 (**c**) and H3K27me3 (**d**) by ChIP-chip also carry detectable levels of nucleosomes. Promoters with low modification enrichments have variable levels of nucleosomes. Promoter regions that were identified being enriched in all three ChIP-chip replicates (this study) are indicated in blue (H3K4me2) or red (H3K27me3).

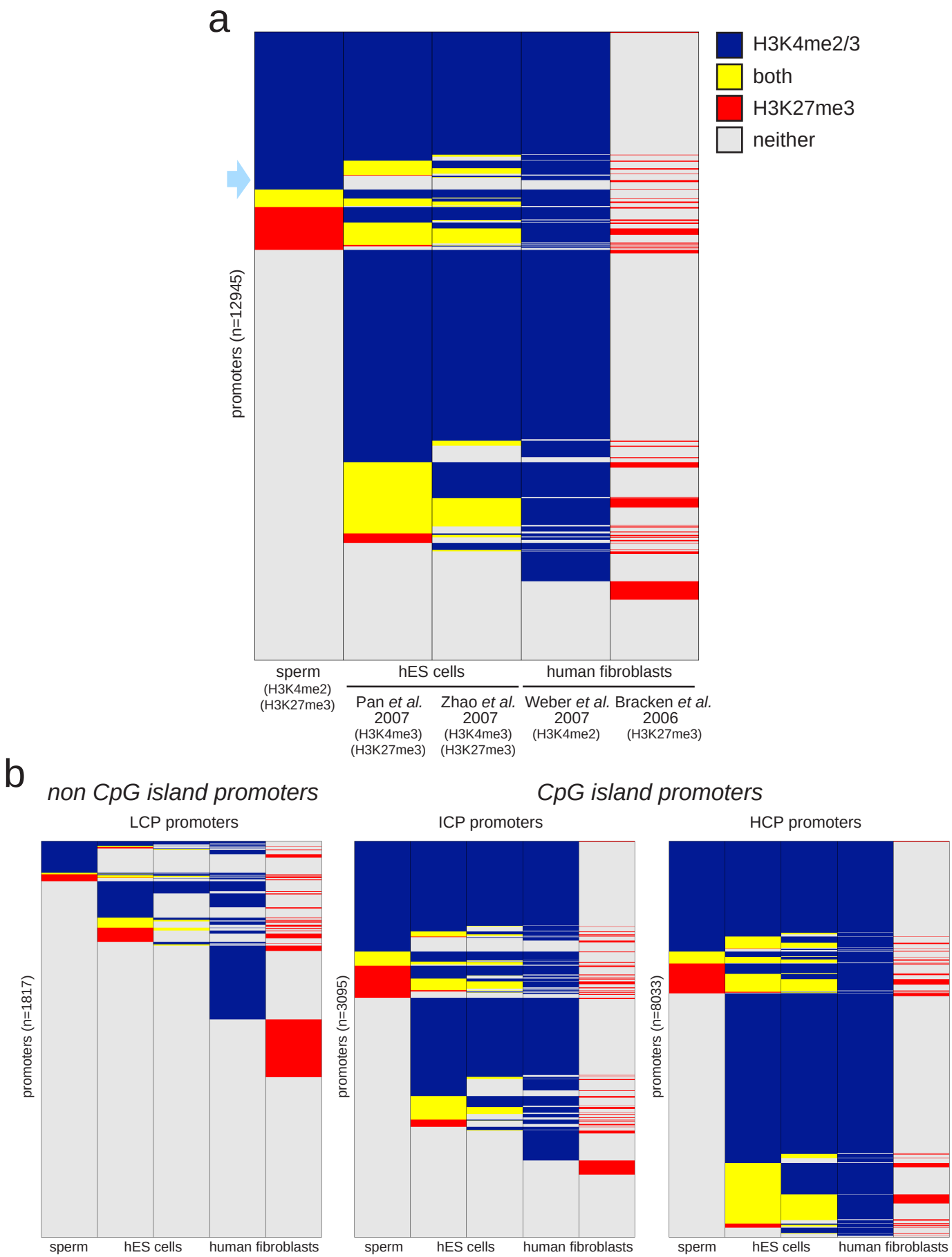
(**e, f**) Venn diagrams showing overlap between identified groups of modified promoters (this study) and published lists of genes associated with modified histones and nucleosomes (Hammoud et al., 2009) that are present on the NimbleGen promoter array, for H3K4me2 (**e**), for H3K27me3 (**f**). Percentages indicate overlap between pairs of analyzed groups.



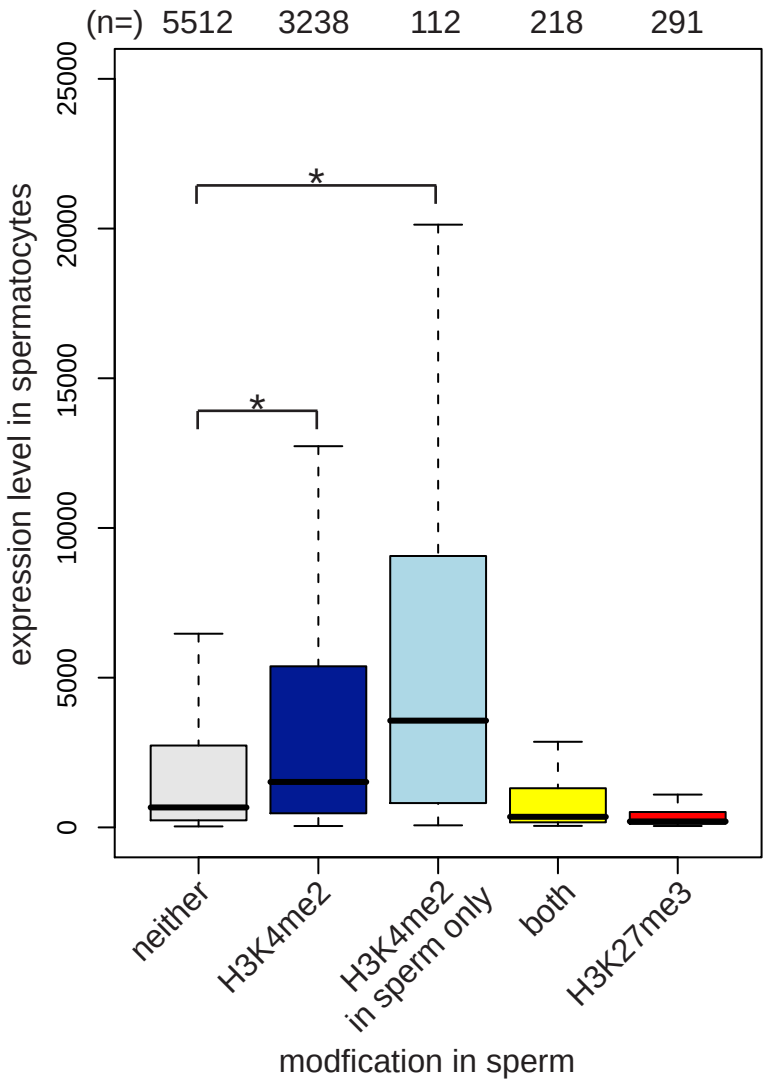
Supplementary Fig. 5: Classification of H3K4me2 and H3K27me3 promoters according to CpG density of underlying promoter sequence. Bar diagrams showing preferential marking of CpG island promoters, also called intermediate (ICP) and high CpG-content (HCP) promoters, by H3K4me2 and H3K27me3. Promoters with low CpG-content are labeled as non CpG island or low CpG-content promoters (LCP). Classification on CpG density is based on published criteria (Weber et al., 2007).



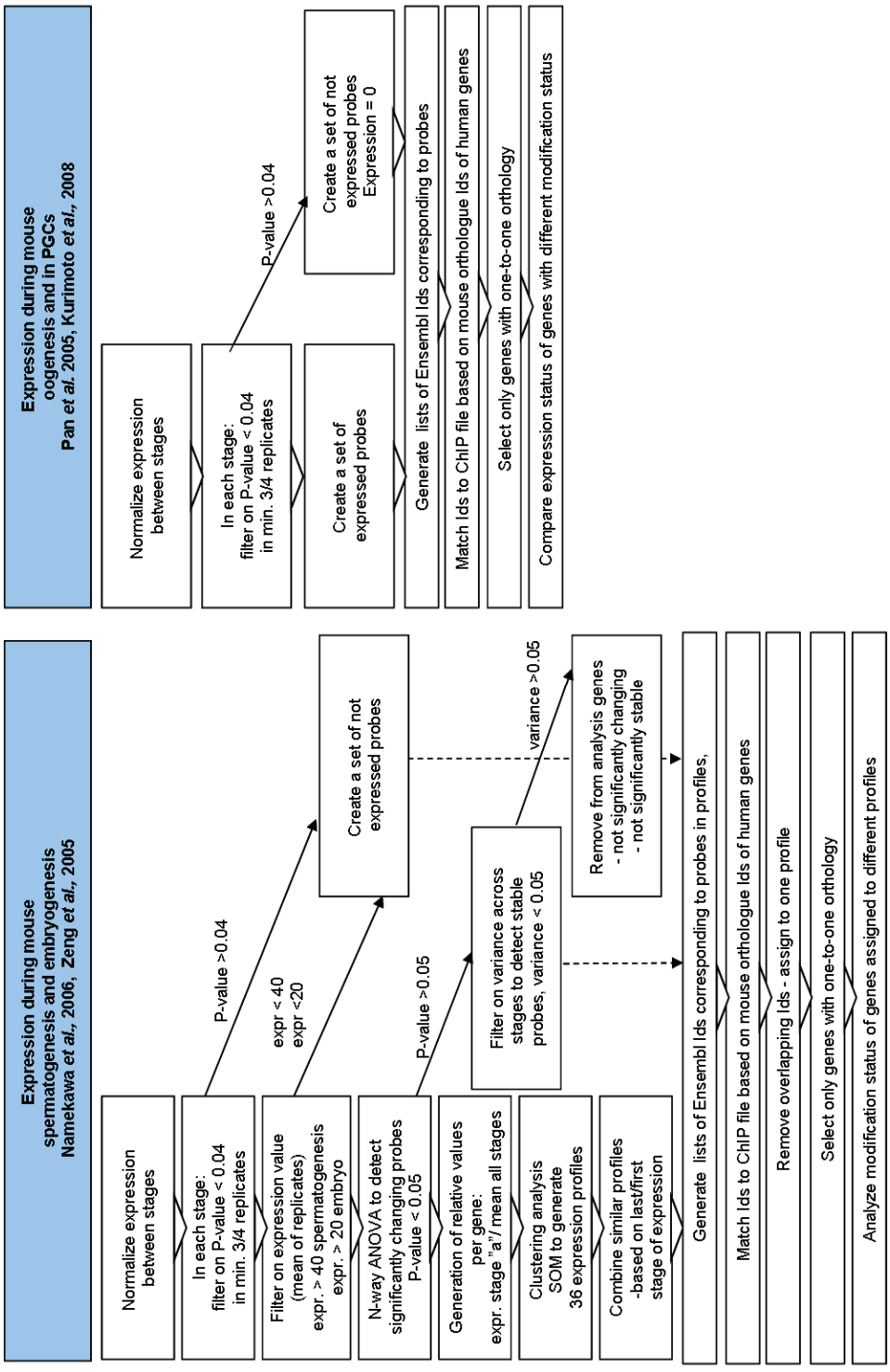
Supplementary Fig. 6: Number of modified gene promoters per chromosome. Scatter plots revealing linear relationships between the number of gene promoters modified by H3K4me2 and/or H3K2me3 and the total number of genes per chromosome (present on the array). The X-chromosome is largely devoid of genes marked by either modification.



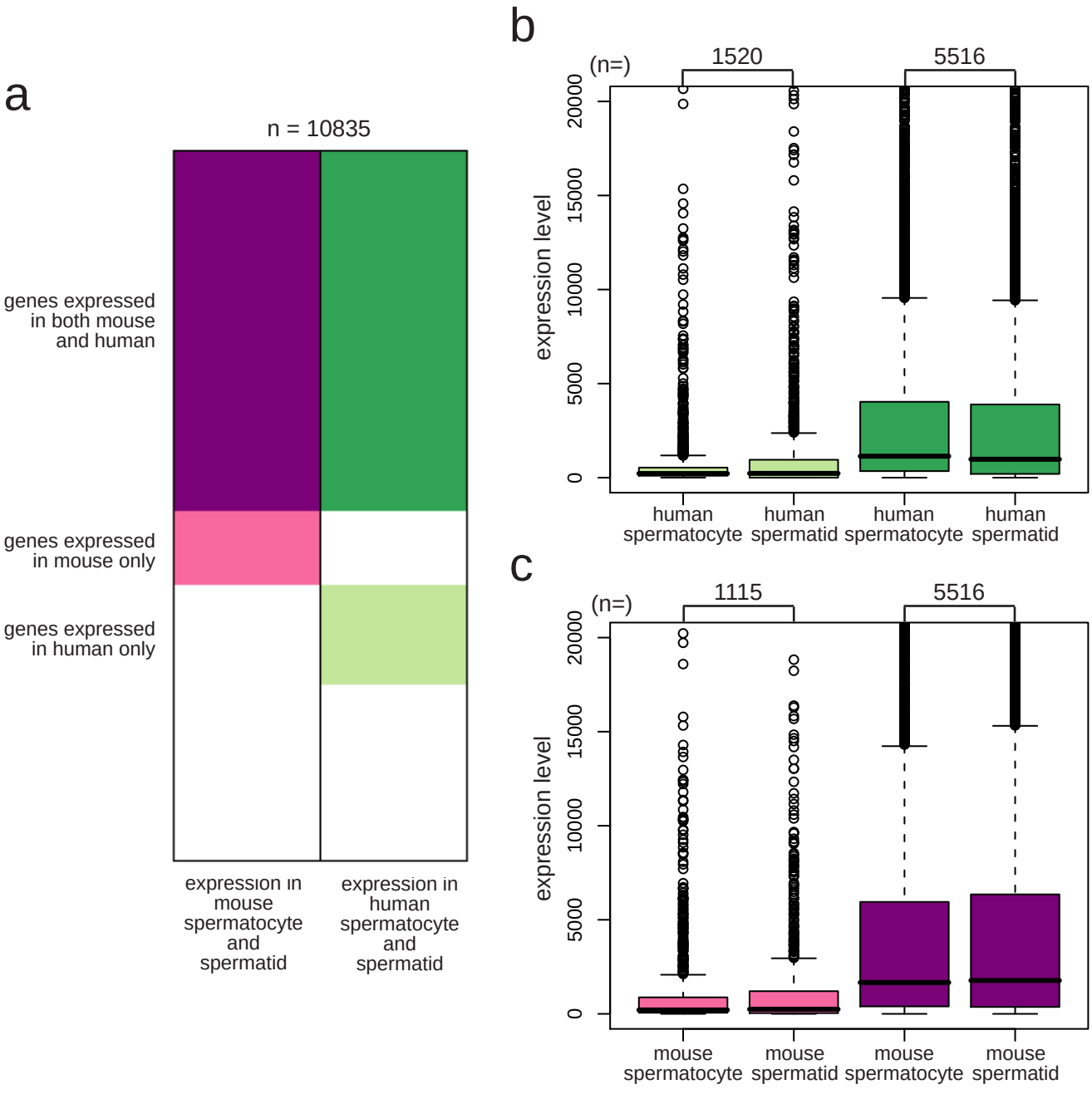
Supplementary Fig. 7: Comparison of histone methylation states in human sperm, hESCs and human fibroblasts. (**a**) Clustering of all promoters according to their methylation status in human sperm, ESCs (Pan et al., 2007; Zhao et al., 2007) and fibroblasts (Bracken et al., 2006; Weber et al., 2007) (see Supplementary Fig. 1 and Methods for details of comparison). Arrow indicates genes marked by H3K4me2 in sperm only (see also Fig. 4a). (**b**) Clustering of promoters according to their CpG-density and histone methylation status in human sperm, ESCs and fibroblasts.



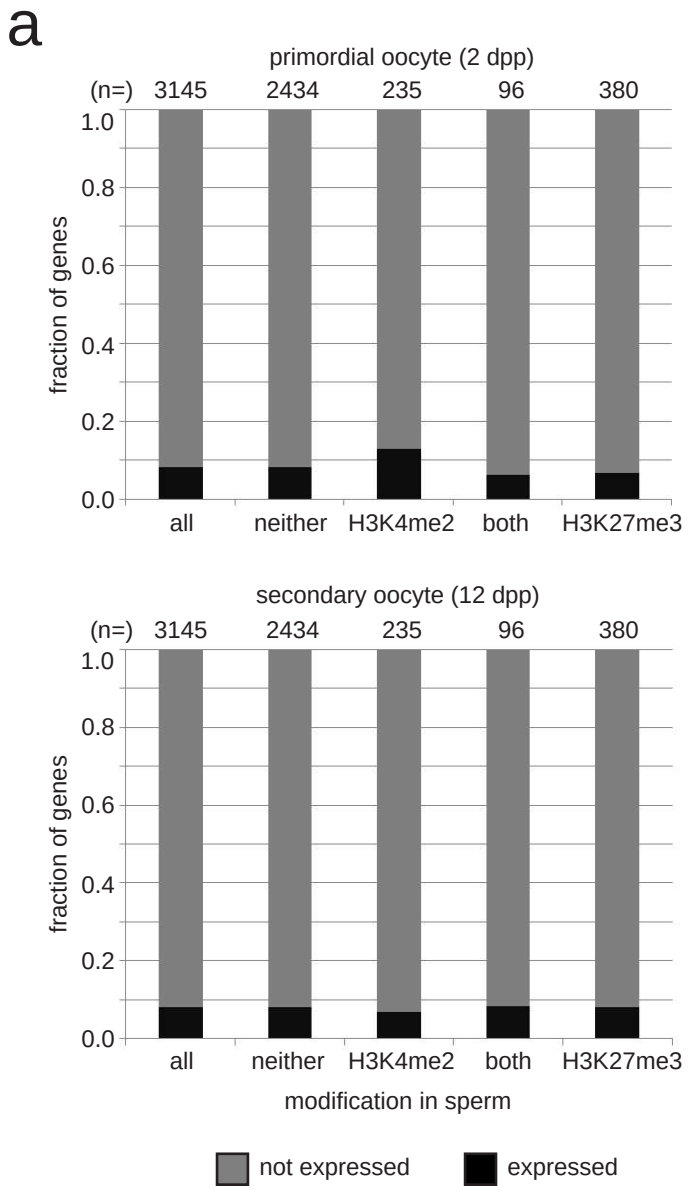
Supplementary Fig. 8: Levels of expression in human spermatocytes. Box plot displaying expression levels in human spermatocytes for genes differentially modified in sperm (Chalmel et al., 2007). Genes under control of promoters that are H3K4me2 targets only in sperm but not in soma (light blue, also indicated by blue arrow in Fig. 4a and Supplementary Fig 7) are most highly expressed. Other H3K4me2 genes (dark blue) show significantly higher levels of expression compared to genes with neither mark (*: Wilcoxon test P-value < 2.2e-16). In the plot, the central bar marks the median, lower and upper limits of the box indicate the 25th and 75th percentiles, and the whiskers extend 1.5 interquartile range from the box.



Supplementary Fig. 9: Data analysis scheme of comparison to mouse expression data. Flow chart showing procedures used to compare modification status of genes in human sperm to expression status during murine spermatogenesis, embryogenesis, oogenesis and primordial germ cell (PGCs) development (see Methods for detailed description).



Supplementary Fig. 10: Validation of comparison between human and mouse expression programs. (**a**) Representation of 10835 one-to-one orthologous genes expressed in mouse (purple) and human (green) spermatocytes and spermatids (Chalmel et al., 2007; Namekawa et al., 2006). Genes commonly expressed in both species are marked in dark color and genes expressed in one species only in light color. (**b**) Box plot representing expression levels in human spermatocytes and spermatids for groups identified in panel (**a**). (**c**) Box plot representing expression levels in mouse spermatocytes and spermatids for groups identified in panel (**a**). Genes expressed in one species only show significantly lower expression levels than genes expressed in both cells from both species.

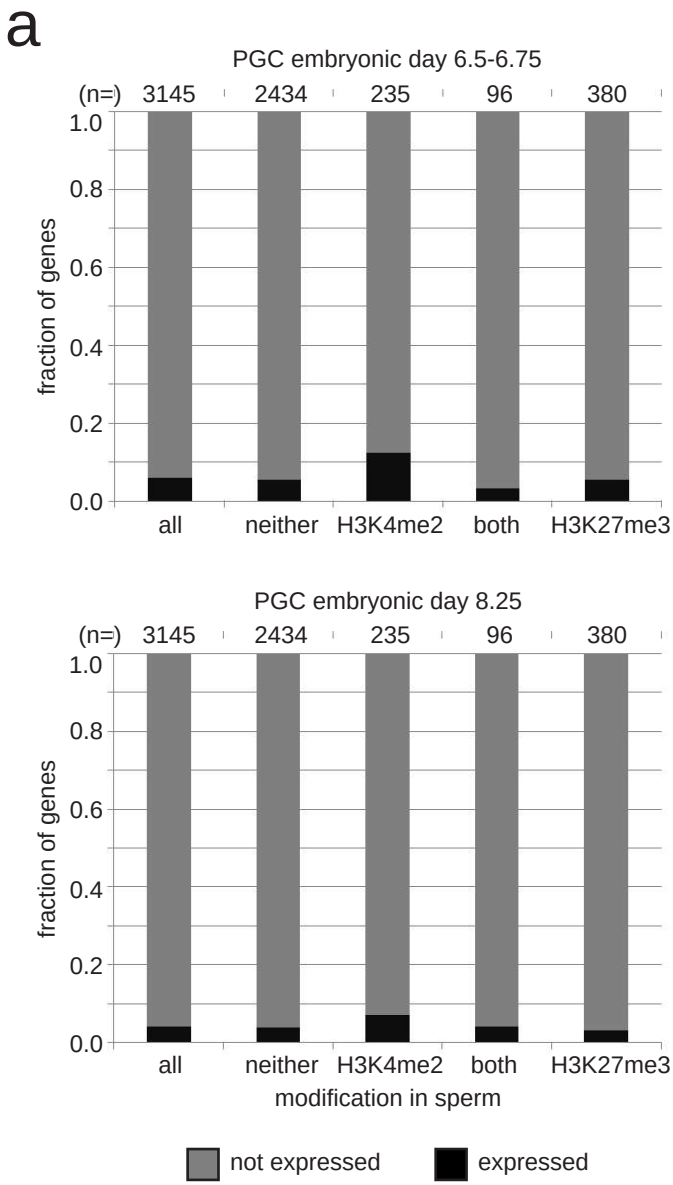


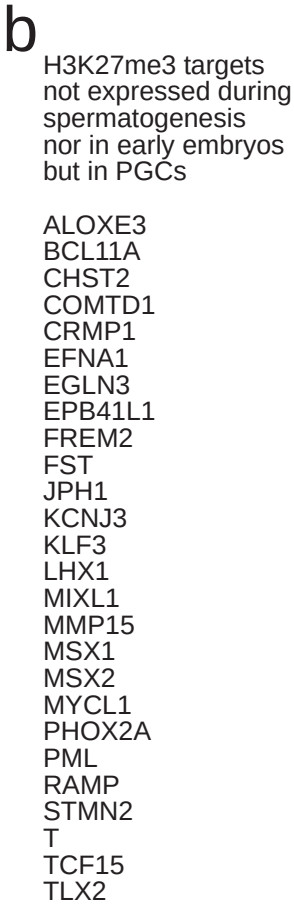
b

H3K27me3 targets not expressed during spermatogenesis nor in early embryo but during oogenesis

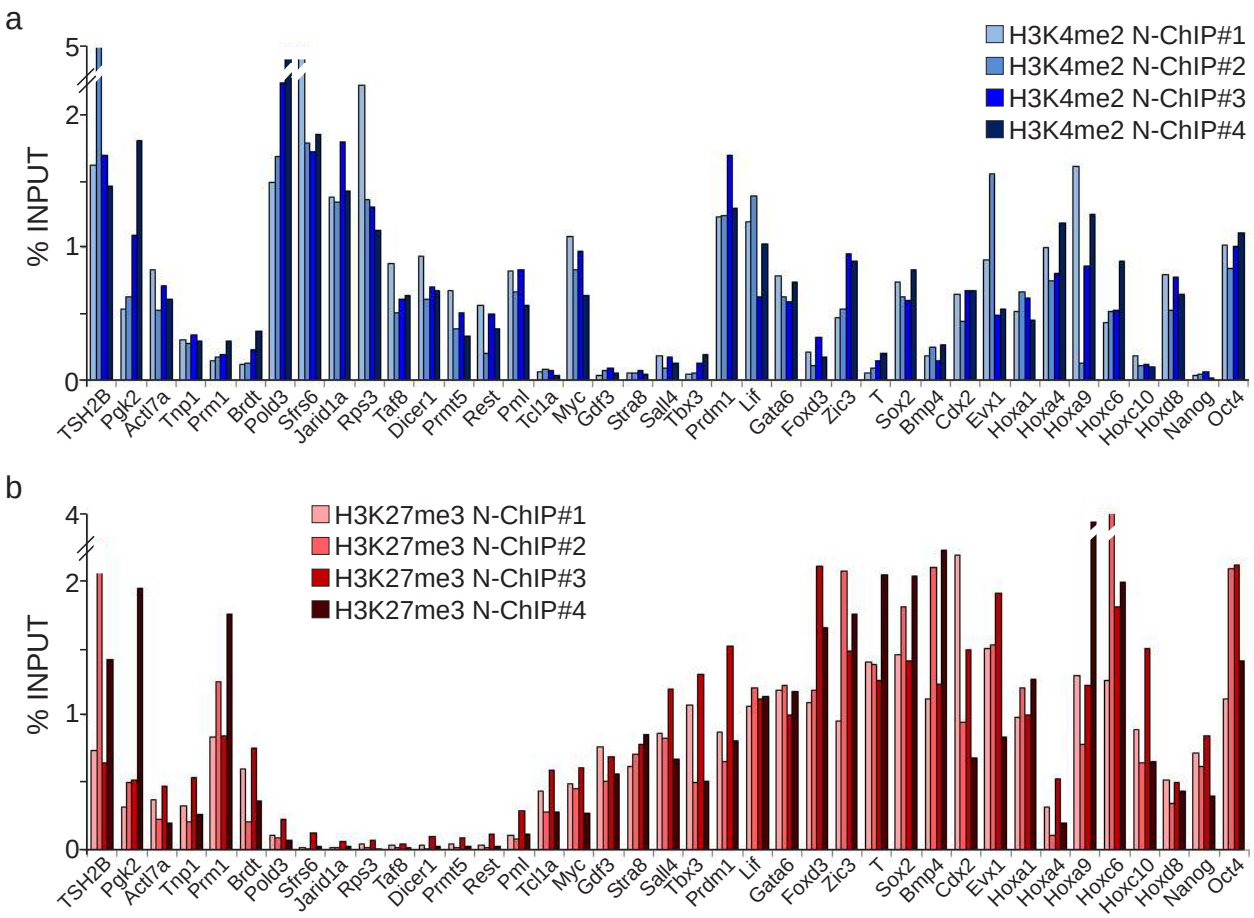
ABCA4	HTR1B
ADCY1	JPH1
ALDH1L1	KCNJ3
ALOXE3	KLF3
C9orf91	LRRC3B
CAMK1G	LYPD3
CBR3	LYPD6
CD8A	MC3R
CHRDL2	NMUR2
CHRNB2	PARD3
COL11A1	PCDH1
COL8A1	PHOX2A
CRHBP	PTPN5
CYP2E1	SCN11A
DAPK2	SEMA6C
EGLN3	SH3RF2
EPB41L1	SLC16A13
GABRB2	SLC17A3
GABRB3	SLC44A1
GPNMB	ST8SIA2
GSC	TBX20
HOXA13	TLL1
HRK	TRPA1
HTR1A	TRPC5
	VSX2

Supplementary Fig. 11: Majority of genes not expressed during spermatogenesis, in fully grown oocytes and during early embryogenesis are also not expressed during oogenesis (**a**) Bar graphs showing expression status in primordial oocytes and growing oocytes (Pan et al., 2005). Analysis was done for genes not expressed during spermatogenesis, in fully grown oocytes and during early embryogenesis (as identified in Fig. 5c). (**b**) List of H3K27me3-targets that are expressed in primordial and/or growing oocytes (representing expressed fraction from panel (**a**)





Supplementary Fig. 12: Majority of genes not expressed during spermatogenesis, in fully grown oocytes and during early embryogenesis are also not expressed in primordial germ cells (**a**) Bar graphs showing expression status in mouse primordial germ cells at E6.5-6.75 and E8.25 of their development (Kurimoto et al., 2008). Analysis was done for genes not expressed during spermatogenesis, in fully grown oocytes and during early embryogenesis (as identified in Fig. 5c). (**b**) List of H3K27me3-targets that are expressed in PGCs at one or both developmental stages (representing expressed fraction from panel (**a**). Consistent with the origin of PGCs, some of these H3K27me3-marked genes have been implicated in mesoderm differentiation (e.g. *T* (Saitou et al., 2002)).



Supplementary Fig. 13: Real time PCR analysis of modified genes in mouse spermatozoa 39 mouse genes were analyzed by native ChIP (N-ChIP) with (**a**) H3K4me2, (**b**) H2K27me3 antibodies, followed by real time PCR. Genes were selected based on orthology with human genes analyzed in Supplementary Fig. 3. Data are presented as % of immuno-precipitated input chromatin. Linear scaling was applied for the replicas presented in the same graphs with following multiplication factors: For H3K4me2: N-ChIP#1 times 1, N-ChIP#2 times 1/2, N-ChIP#3 times 3.5, N-ChIP#4 times 1.6; For H3K27me3: N-ChIP#1 times 1, N-ChIP#2 times 1/4, N-ChIP#3 times 1, N-ChIP#4 times 1/7. For both modifications, N-ChIP #4 is shown in Fig. 6 a,b.

3. Discussion

Functionality of the sperm derived methylated histones

In this thesis we show that methylated histones are retained in mature human and mouse spermatozoa. These histones are residing on functionally defined groups of promoters. Therefore, it is improbable that such retained histones represent a random leftover of inefficient chromatin remodeling during spermatogenesis. Defined, common functional features of genes associated with histones in sperm strongly argue that the retention is purposeful. Furthermore, a high conservation between the modification status in mice and humans supports the model of transgenerational transmission of epigenetic information through the male germ line.

Among mammals, the presence of histones in sperm is not unique to mouse and human. A histone component was also reported for rat, bull and boar sperm (Banerjee and Smallwood, 1998; Codrington et al., 2007; Gatewood et al., 1990; Palmer et al., 1990; van der Heijden et al., 2006). Interestingly, histones only represent a minor fraction of DNA bound proteins in sperm of species other than human. Thus, in human histone levels are exceptionally high. Our demonstration that H3K27me3 is associated with developmental genes in both mouse and human, indicates that there may exist a group of regulatory sequences with conserved nucleosomal configuration. As protein coding sequences occupy only around 1% in mammalian genomes, 1% of retained histone may be enough to well serve the regulatory function.

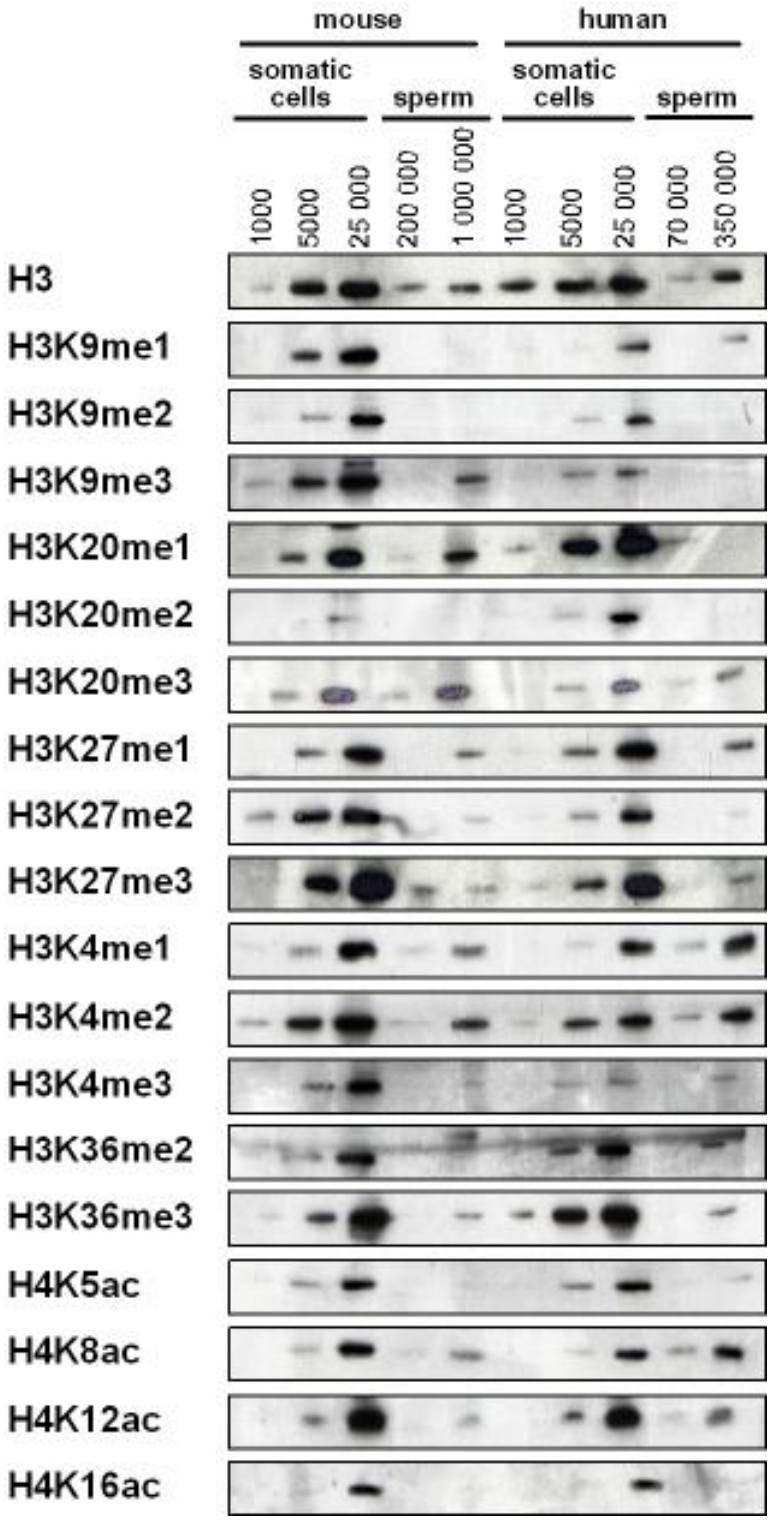
Nevertheless, we have not provided direct evidences for a function of sperm derived histones for the subsequent embryonic development. Conservation of the methylation marks from mouse to human, will allow us to use the powerful mouse model to answer these questions. Conditional transgenic mice, overexpressing either a H3K4 or H3K27 histone de-methylase at the last stages of spermatogenesis, will be generated for that purpose. We hope that using this strategy, histone modifications in spermatozoa will be specifically removed, without affecting the spermatogenic process. Developmental potential of the embryos generated with sperm lacking a given mark will reveal whether sperm derived modified histones play an essential role in the embryo.

Paternal versus maternal epigenetic contributions

For many years, sperm has been regarded as an “inert” container for delivering of the paternal DNA. The fundamentally different chromatin configuration of sperm comparing

to that of somatic cells, made it very difficult to investigate its nature with established biochemical methods. In contrast, oocytes retain the nucleosomal configuration of the genome. This, along with a high volume of cytoplasm, makes oocytes potential carriers of proteins and RNAs necessary for pre-implantation development. Furthermore, analysis of the developmental potential of gynogenic embryos demonstrates that the sole manipulation of two out of the three paternally imprinted loci solves the need for a paternally derived genome. Therefore, genomic imprinting is the only barrier for the parthenogenic development in mammals. (Kawahara et al., 2007). One could argue that this is a proof for a very marginal paternal contribution to the embryo. However, in a natural situation, the two parental genomes are joined to assure phenotypic variation and survival of the species. Besides the differences coming from the physiological roles of male and female gametes during reproduction and differences in genomic imprinting, one can propose that the two gametes should not fundamentally differ in their epigenetic contributions. Consistently, when we analyzed the presence of histone modifications in mouse and human spermatozoa, we detected multiple marks, many of which have recognized functions in gene regulation (Figure, page 95). We decided to concentrate on two of them, H3K4me2 and H3K27me3, which in somatic cells are associated with active and repressed genes, respectively. In the presented genome wide study, we show that these marks are associated with a substantial number of promoters in mature spermatozoa. Furthermore, multiple developmental regulators, which are Polycomb targets in pluripotent somatic cells, are marked by the Polycomb deposited H3K27me3 in sperm. Similarly, to somatic cells, the presence of this mark correlates with gene repression at the preceding and following stages of development. We propose that H3K27me3 transmitted by sperm assures repression of developmental regulators at the totipotent stage of the preimplantation development. We further hypothesize that the same sequences carry H3K27me3 in the oocyte genome, so that both inherited alleles are protected from the aberrant expression in the early embryo.

Injection of round spermatids into oocytes results in the development of healthy and fertile offspring (Kimura and Yanagimachi, 1995; Tamashiro et al., 1999). Even though the success rates of such experiments are higher when mature spermatozoa are injected, these data argue that the epigenetic state of the paternal genome, at the round spermatid stage, is compatible with embryonic development. As we show in the presented results, the majority of genes carrying repressive H3K27me3 in sperm was



Histone post translational modifications present in mouse and human spermatozoa. Western blot analysis was performed according to the Methods in the Results section.

already repressed either in PGCs or at early spermatogenic stages. Data from the round spermatid injection experiments thus support a model in which epigenetic marks with major importance for embryonic development, are established early or continuously exist in the germ cell lineage.

In *C.elegans* worms lacking the H3K4 de-methylase SPR-5 (a worm homolog of LSD1), H3K4me2 accumulates over generations in the male germline and causes aberrant upregulation of spermatogenic genes. The authors suggest that epigenetic determinants of the spermatogenic program are selectively erased from the genome after fertilization, in order to maintain totipotency (Katz et al., 2009). In our study, H3K4me2 is retained on spermatogenic genes as well. We hypothesize that similarly to the situation in *C.elegans* these marks are selectively removed after fertilization, whereas marks on other loci may be retained.

Defined genomic localization of histones in mouse and human spermatozoa

In the last two months, two papers describing genomic localization of nucleosomes and modified histones in mouse and human spermatozoa have been published (Arpanahi et al., 2009; Hammoud et al., 2009). We performed a detailed comparison of our data with the (Hammoud et al., 2009) study. We compared the data sets both based on the lists of genes associated with methylated H3K4 and H3K27 and based on the raw ChIP-seq ((Hammoud et al., 2009)) and ChIP-chip (our study) enrichment scores. In both comparisons we found a high positive correlation. The very different format of the data by Arpanahi and colleagues precluded us from making a direct comparison to their data set. The arrays used in their study allow to localize nucleosomal bound sequences with a 35 kb resolution and therefore do not provide information about the status of single promoters. Nevertheless, the gene ontology analysis of sequences associated with histones in both published studies and in our study, revealed over-representation of developmental processes. Furthermore, consistently with our results on selected loci, Arpanahi and colleagues found an over-representation of developmental processes among histone bound sequences in mouse spermatozoa.

Similar to our study, Hammoud and colleagues reported that genes with developmental functions are associated with H3K27me3 or are in the bivalent state. Contrary to our data, they also observed a group of developmental genes associated with H3K4me3 only. This discrepancy may be caused by the fact that in our study, H3K4

dimethylation and in their study trimethylation was analyzed. Nevertheless, high correlation of our H3K4me2 data with their H3K4me3 data suggests that in spermatozoa the two modifications occupy largely overlapping sets of promoters, consistent with multiple studies in different cell types in human and mouse (Barski et al., 2007; Guenther et al., 2007; Mikkelsen et al., 2007; Mohn et al., 2008; Orford et al., 2008; Weber et al., 2007). Furthermore, consistent with our results, Hammoud and colleagues find genes expressed during spermatogenesis and in the early embryo, among H3K4 methylated targets.

In contrast to our study, which was limited to promoters, the two described publications provide genome-wide data. Interestingly, Arpanahi and colleagues report that in both human and mouse spermatozoa, promoter sequences are significantly higher enriched among nucleosomal bound regions than intergenic sequences. Furthermore, they find a high positive correlation between nucleosomal bound regions and CTCF binding sites in human spermatozoa. Hammound and colleagues report that several miRNA clusters and imprinted loci are associated with histones in human spermatozoa. Taken together these observations support a hypothesis, in which histones are selectively retained on regulatory sequences and may serve evolutionary conserved functions.

Fate of the sperm derived epigenetic information in the early embryo

Besides modified histones, other epigenetic factors may be involved in the paternal transmission of epigenetic information. In both mouse and human spermatozoa, DNA methylation is not only present at imprinted genes, but also marks multiple promoters (Farthing et al., 2008; Weber et al., 2007). Furthermore, various protein coding RNAs and micro RNAs have been identified in human spermatozoa (Ostermeier et al., 2002; Ostermeier et al., 2005).

Taking global changes in the paternal genome just after the fertilization into account, one must keep in mind that the epigenetic information encoded by histone modifications, DNA methylation and RNAs could be immediately erased in the embryo. However, during the pronuclear stages preceding the 1st genome replication, paternal histones can be distinguished from the maternally provided ones based on the differences in H3 variants and on the acetylation patterns on histone H4. Sperm derived histones are observed to retain on the paternal genome, arguing that they are not

removed along with the protamines (Adenot et al., 1997; van der Heijden et al., 2006; van der Heijden et al., 2008).

The erasure of DNA methylation from the paternal genome that takes place after fertilization, is not complete. Both imprinted genes and certain retrotransposons are resistant to this process (Olek and Walter, 1997; Rougier et al., 1998). A comprehensive list of sequences that escape de-methylation, however is not known. Furthermore, the heritable phenotype of the *Kit* epimutation correlates with the aberrant levels of RNA in sperm, arguing that this RNA is not degraded after fertilization (Rassoulzadegan et al., 2006; Wagner et al., 2008).

Heterogeneity of epigenetic traits

The described work was performed on pools of spermatozoa coming from multiple individuals. It can not be ruled out that the patterns of histone modifications observed in human sperm are a sum of variable individual patterns. However, the conservation of histone marks between sperm coming from humans and from an inbred mouse strain argues that the patterns maybe similar between individuals. Furthermore, in the recent study of Hammoud and colleagues, the analysis of nucleosomal positioning in sperm coming from a single donor and from an independent pool of donors, resulted in a high positive correlation between these two samples (r=0.7) (Hammoud et al., 2009). On the other hand, these data suggest that there is a certain level of variability between the individuals. Similarly, a DNA methylation profiling of spermatozoa from 21 healthy individuals revealed overall similarity, but also significant differences at specific loci (Flanagan et al., 2006). It is an attractive hypothesis that epigenetic features can also be subjected to environmental selection (Ruden et al., 2008; Sollars et al., 2003). Differences between individuals may therefore lead to a phenotypic variation and selection that is advantageous for the species. Furthermore, recent data on the effect of maternal diet or the exposure to chemicals like endocrine disruptors, suggest that environmental factors can alter epigenetic information that is transmitted by the germ line (Anway et al., 2005; Cropley et al., 2006). Thus, transgenerational effects are emerging as important issues for medicine and environmental protection. The essential contribution of the histone methylation to the sperm epigenome provides a new insight into the potential role of the father in transmitting epigenetic information to the next generation.

References

Adams, I.R., and McLaren, A. (2002). Sexually dimorphic development of mouse primordial germ cells: switching from oogenesis to spermatogenesis. Development *129*, 1155-1164.

Adenot, P.G., Mercier, Y., Renard, J.P., and Thompson, E.M. (1997). Differential H4 acetylation of paternal and maternal chromatin precedes DNA replication and differential transcriptional activity in pronuclei of 1-cell mouse embryos. Development *124*, 4615-4625.

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

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

Ahmad, K., and Henikoff, S. (2002). The histone variant H3.3 marks active chromatin by replication-independent nucleosome assembly. Mol Cell *9*, 1191-1200.

Albert, M., and Peters, A.H. (2009). Genetic and epigenetic control of early mouse development. Curr Opin Genet Dev *19*, 113-121.

Alleman, M., Sidorenko, L., McGinnis, K., Seshadri, V., Dorweiler, J.E., White, J., Sikkink, K., and Chandler, V.L. (2006). An RNA-dependent RNA polymerase is required for paramutation in maize. Nature *442*, 295-298.

Allen, M.J., Lee, C., Lee, J.D.t., Pogany, G.C., Balooch, M., Siekhaus, W.J., and Balhorn, R. (1993). Atomic force microscopy of mammalian sperm chromatin. Chromosoma *102*, 623-630.

Alsheimer, M., Fecher, E., and Benavente, R. (1998). Nuclear envelope remodelling during rat spermiogenesis: distribution and expression pattern of LAP2/thymopoietins. J Cell Sci *111 (Pt 15)*, 2227-2234.

Ancelin, K., Lange, U.C., Hajkova, P., Schneider, R., Bannister, A.J., Kouzarides, T., and Surani, M.A. (2006). Blimp1 associates with Prmt5 and directs histone arginine methylation in mouse germ cells. Nat Cell Biol *8*, 623-630.

Anderson, E.L., Baltus, A.E., Roepers-Gajadien, H.L., Hassold, T.J., de Rooij, D.G., van Pelt, A.M., and Page, D.C. (2008). Stra8 and its inducer, retinoic acid, regulate meiotic initiation in both spermatogenesis and oogenesis in mice. Proc Natl Acad Sci U S A *105*, 14976-14980.

Anway, M.D., Cupp, A.S., Uzumcu, M., and Skinner, M.K. (2005). Epigenetic transgenerational actions of endocrine disruptors and male fertility. Science *308*, 1466-1469.

Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., Morris, P., Brownstein, M.J., Kuramochi-Miyagawa, S., Nakano, T.*, et al.* (2006). A novel class of small RNAs bind to MILI protein in mouse testes. Nature *442*, 203-207.

Aravin, A.A., Sachidanandam, R., Girard, A., Fejes-Toth, K., and Hannon, G.J. (2007). Developmentally regulated piRNA clusters implicate MILI in transposon control. Science *316*, 744-747.

Arpanahi, A., Brinkworth, M., Iles, D., Krawetz, S.A., Paradowska, A., Platts, A.E., Saida, M., Steger, K., Tedder, P., and Miller, D. (2009). Endonuclease-sensitive regions of human spermatozoal chromatin are highly enriched in promoter and CTCF binding sequences. Genome Res *19*, 1338-1349.

Ausio, J. (1999). Histone H1 and evolution of sperm nuclear basic proteins. J Biol Chem *274*, 31115-31118.

Ayton, P.M., Chen, E.H., and Cleary, M.L. (2004). Binding to nonmethylated CpG DNA is essential for target recognition, transactivation, and myeloid transformation by an MLL oncoprotein. Mol Cell Biol *24*, 10470-10478.

Azuara, V., Perry, P., Sauer, S., Spivakov, M., Jorgensen, H.F., John, R.M., Gouti, M., Casanova, M., Warnes, G., Merkenschlager, M.*, et al.* (2006). Chromatin signatures of pluripotent cell lines. Nat Cell Biol *8*, 532-538.

Baarends, W.M., Hoogerbrugge, J.W., Roest, H.P., Ooms, M., Vreeburg, J., Hoeijmakers, J.H., and Grootegoed, J.A. (1999). Histone ubiquitination and chromatin remodeling in mouse spermatogenesis. Dev Biol *207*, 322-333.

Balhorn, R. (2007). The protamine family of sperm nuclear proteins. Genome Biol *8*, 227.

Balhorn, R., Gledhill, B.L., and Wyrobek, A.J. (1977). Mouse sperm chromatin proteins: quantitative isolation and partial characterization. Biochemistry *16*, 4074-4080.

Baltus, A.E., Menke, D.B., Hu, Y.C., Goodheart, M.L., Carpenter, A.E., de Rooij, D.G., and Page, D.C. (2006). In germ cells of mouse embryonic ovaries, the decision to enter meiosis precedes premeiotic DNA replication. Nat Genet *38*, 1430-1434.

Banerjee, S., and Smallwood, A. (1998). Chromatin modification of imprinted H19 gene in mammalian spermatozoa. Mol Reprod Dev *50*, 474-484.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature *410*, 120-124.

Bao, S., Miyoshi, N., Okamoto, I., Jenuwein, T., Heard, E., and Azim Surani, M. (2005). Initiation of epigenetic reprogramming of the X chromosome in somatic nuclei transplanted to a mouse oocyte. EMBO Rep *6*, 748-754.

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

Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. Nature *427*, 164-167.

Baxevanis, A.D., Godfrey, J.E., and Moudrianakis, E.N. (1991). Associative behavior of the histone (H3-H4)2 tetramer: dependence on ionic environment. Biochemistry *30*, 8817-8823.

Beissbarth, T., and Speed, T.P. (2004). GOstat: find statistically overrepresented Gene Ontologies within a group of genes. Bioinformatics *20*, 1464-1465.

Bellve, A.R., Cavicchia, J.C., Millette, C.F., O'Brien, D.A., Bhatnagar, Y.M., and Dym, M. (1977). Spermatogenic cells of the prepuberal mouse. Isolation and morphological characterization. J Cell Biol *74*, 68-85.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K.*, et al.* (2006a). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell *125*, 315-326.

Bernstein, E., Duncan, E.M., Masui, O., Gil, J., Heard, E., and Allis, C.D. (2006b). Mouse polycomb proteins bind differentially to methylated histone H3 and RNA and are enriched in facultative heterochromatin. Mol Cell Biol *26*, 2560-2569.

Birke, M., Schreiner, S., Garcia-Cuellar, M.P., Mahr, K., Titgemeyer, F., and Slany, R.K. (2002). The MT domain of the proto-oncoprotein MLL binds to CpG-containing DNA and discriminates against methylation. Nucleic Acids Res *30*, 958-965.

Birney, E., Stamatoyannopoulos, J.A., Dutta, A., Guigo, R., Gingeras, T.R., Margulies, E.H., Weng, Z., Snyder, M., Dermitzakis, E.T., Thurman, R.E.*, et al.* (2007). Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. Nature *447*, 799-816.

Blewitt, M.E., Vickaryous, N.K., Paldi, A., Koseki, H., and Whitelaw, E. (2006). Dynamic reprogramming of DNA methylation at an epigenetically sensitive allele in mice. PLoS Genet *2*, e49.

Bourc'his, D., and Bestor, T.H. (2004). Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. Nature *431*, 96-99.

Bourc'his, D., Xu, G.L., Lin, C.S., Bollman, B., and Bestor, T.H. (2001). Dnmt3L and the establishment of maternal genomic imprints. Science *294*, 2536-2539.

Bowles, J., Knight, D., Smith, C., Wilhelm, D., Richman, J., Mamiya, S., Yashiro, K., Chawengsaksophak, K., Wilson, M.J., Rossant, J.*, et al.* (2006). Retinoid signaling determines germ cell fate in mice. Science *312*, 596-600.

Boyer, L.A., Plath, K., Zeitlinger, J., Brambrink, T., Medeiros, L.A., Lee, T.I., Levine, S.S., Wernig, M., Tajonar, A., Ray, M.K.*, et al.* (2006). Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature *441*, 349-353.

Bracken, A.P., Dietrich, N., Pasini, D., Hansen, K.H., and Helin, K. (2006). Genome-wide mapping of Polycomb target genes unravels their roles in cell fate transitions. Genes Dev *20*, 1123-1136.

Braun, R.E. (2000). Temporal control of protein synthesis during spermatogenesis. Int J Androl *23 Suppl 2*, 92-94.

Breiling, A., Turner, B.M., Bianchi, M.E., and Orlando, V. (2001). General transcription factors bind promoters repressed by Polycomb group proteins. Nature *412*, 651-655.

Brennan, J., and Capel, B. (2004). One tissue, two fates: molecular genetic events that underlie testis versus ovary development. Nat Rev Genet *5*, 509-521.

Brink, R.A. (1956). A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. Genetics *41*, 872-889.

Buaas, F.W., Kirsh, A.L., Sharma, M., McLean, D.J., Morris, J.L., Griswold, M.D., de Rooij, D.G., and Braun, R.E. (2004). Plzf is required in adult male germ cells for stem cell self-renewal. Nat Genet *36*, 647-652.

Buchwald, G., van der Stoop, P., Weichenrieder, O., Perrakis, A., van Lohuizen, M., and Sixma, T.K. (2006). Structure and E3-ligase activity of the Ring-Ring complex of polycomb proteins Bmi1 and Ring1b. EMBO J *25*, 2465-2474.

Burgold, T., Spreafico, F., De Santa, F., Totaro, M.G., Prosperini, E., Natoli, G., and Testa, G. (2008). The histone H3 lysine 27-specific demethylase Jmjd3 is required for neural commitment. PLoS One *3*, e3034.

Cakmakci, N.G., Lerner, R.S., Wagner, E.J., Zheng, L., and Marzluff, W.F. (2008). SLIP1, a factor required for activation of histone mRNA translation by the stem-loop binding protein. Mol Cell Biol *28*, 1182-1194.

Cao, R., Tsukada, Y., and Zhang, Y. (2005). Role of Bmi-1 and Ring1A in H2A ubiquitylation and Hox gene silencing. Mol Cell *20*, 845-854.

Cao, R., Wang, L., Wang, H., Xia, L., Erdjument-Bromage, H., Tempst, P., Jones, R.S., and Zhang, Y. (2002). Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science *298*, 1039-1043.

Cavalli, G., and Paro, R. (1998). The Drosophila Fab-7 chromosomal element conveys epigenetic inheritance during mitosis and meiosis. Cell *93*, 505-518.

Chalmel, F., Rolland, A.D., Niederhauser-Wiederkehr, C., Chung, S.S., Demougin, P., Gattiker, A., Moore, J., Patard, J.J., Wolgemuth, D.J., Jegou, B.*, et al.* (2007). The conserved transcriptome in human and rodent male gametogenesis. Proc Natl Acad Sci U S A *104*, 8346-8351.

Chamberlain, S.J., Yee, D., and Magnuson, T. (2008). Polycomb repressive complex 2 is dispensable for maintenance of embryonic stem cell pluripotency. Stem Cells *26*, 1496-1505.

Chan, T.L., Yuen, S.T., Kong, C.K., Chan, Y.W., Chan, A.S., Ng, W.F., Tsui, W.Y., Lo, M.W., Tam, W.Y., Li, V.S.*, et al.* (2006). Heritable germline epimutation of MSH2 in a family with hereditary nonpolyposis colorectal cancer. Nat Genet *38*, 1178-1183.

Chandler, V.L. (2007). Paramutation: from maize to mice. Cell *128*, 641-645.

Chen, X., Hiller, M., Sancak, Y., and Fuller, M.T. (2005). Tissue-specific TAFs counteract Polycomb to turn on terminal differentiation. Science *310*, 869-872.

Cho, C., Jung-Ha, H., Willis, W.D., Goulding, E.H., Stein, P., Xu, Z., Schultz, R.M., Hecht, N.B., and Eddy, E.M. (2003). Protamine 2 deficiency leads to sperm DNA damage and embryo death in mice. Biol Reprod *69*, 211-217.

Cho, C., Willis, W.D., Goulding, E.H., Jung-Ha, H., Choi, Y.C., Hecht, N.B., and Eddy, E.M. (2001). Haploinsufficiency of protamine-1 or -2 causes infertility in mice. Nat Genet *28*, 82-86.

Chong, S., Vickaryous, N., Ashe, A., Zamudio, N., Youngson, N., Hemley, S., Stopka, T., Skoultchi, A., Matthews, J., Scott, H.S.*, et al.* (2007a). Modifiers of epigenetic reprogramming show paternal effects in the mouse. Nat Genet *39*, 614-622.

Chong, S., and Whitelaw, E. (2004). Epigenetic germline inheritance. Curr Opin Genet Dev *14*, 692-696.

Chong, S., Youngson, N.A., and Whitelaw, E. (2007b). Heritable germline epimutation is not the same as transgenerational epigenetic inheritance. Nat Genet *39*, 574-575; author reply 575-576.

Chow, C.M., Georgiou, A., Szutorisz, H., Maia e Silva, A., Pombo, A., Barahona, I., Dargelos, E., Canzonetta, C., and Dillon, N. (2005). Variant histone H3.3 marks promoters of transcriptionally active genes during mammalian cell division. EMBO Rep *6*, 354-360.

Christensen, J., Agger, K., Cloos, P.A., Pasini, D., Rose, S., Sennels, L., Rappsilber, J., Hansen, K.H., Salcini, A.E., and Helin, K. (2007). RBP2 belongs to a family of demethylases, specific for tri-and dimethylated lysine 4 on histone 3. Cell *128*, 1063-1076.

Chu, D.S., Liu, H., Nix, P., Wu, T.F., Ralston, E.J., Yates, J.R., 3rd, and Meyer, B.J. (2006). Sperm chromatin proteomics identifies evolutionarily conserved fertility factors. Nature *443*, 101-105.

Chuang, L.S., Ian, H.I., Koh, T.W., Ng, H.H., Xu, G., and Li, B.F. (1997). Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. Science *277*, 1996-2000.

Codrington, A.M., Hales, B.F., and Robaire, B. (2007). Chronic cyclophosphamide exposure alters the profile of rat sperm nuclear matrix proteins. Biol Reprod *77*, 303-311.

Coffigny, H., Bourgeois, C., Ricoul, M., Bernardino, J., Vilain, A., Niveleau, A., Malfoy, B., and Dutrillaux, B. (1999). Alterations of DNA methylation patterns in germ cells and Sertoli cells from developing mouse testis. Cytogenet Cell Genet *87*, 175-181.

Conrad, S., Renninger, M., Hennenlotter, J., Wiesner, T., Just, L., Bonin, M., Aicher, W., Buhring, H.J., Mattheus, U., Mack, A.*, et al.* (2008). Generation of pluripotent stem cells from adult human testis. Nature *456*, 344-349.

Cooney, C.A., Dave, A.A., and Wolff, G.L. (2002). Maternal methyl supplements in mice affect epigenetic variation and DNA methylation of offspring. J Nutr *132*, 2393S-2400S.

Costoya, J.A., Hobbs, R.M., Barna, M., Cattoretti, G., Manova, K., Sukhwani, M., Orwig, K.E., Wolgemuth, D.J., and Pandolfi, P.P. (2004). Essential role of Plzf in maintenance of spermatogonial stem cells. Nat Genet *36*, 653-659.

Cowell, I.G., Aucott, R., Mahadevaiah, S.K., Burgoyne, P.S., Huskisson, N., Bongiorni, S., Prantera, G., Fanti, L., Pimpinelli, S., Wu, R.*, et al.* (2002). Heterochromatin, HP1 and methylation at lysine 9 of histone H3 in animals. Chromosoma *111*, 22-36.

Cropley, J.E., Suter, C.M., Beckman, K.B., and Martin, D.I. (2006). Germ-line epigenetic modification of the murine A vy allele by nutritional supplementation. Proc Natl Acad Sci U S A *103*, 17308-17312.

Cubas, P., Vincent, C., and Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. Nature *401*, 157-161.

Davis, T.L., Yang, G.J., McCarrey, J.R., and Bartolomei, M.S. (2000). The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. Hum Mol Genet *9*, 2885-2894.

De Lucia, F., Faraone-Mennella, M.R., D'Erme, M., Quesada, P., Caiafa, P., and Farina, B. (1994). Histone-induced condensation of rat testis chromatin: testis-specific H1t versus somatic H1 variants. Biochem Biophys Res Commun *198*, 32-39.

de Napoles, M., Mermoud, J.E., Wakao, R., Tang, Y.A., Endoh, M., Appanah, R., Nesterova, T.B., Silva, J., Otte, A.P., Vidal, M*., et al.* (2004). Polycomb group proteins Ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. Dev Cell *7*, 663-676.

de Napoles, M., Nesterova, T., and Brockdorff, N. (2007). Early loss of Xist RNA expression and inactive X chromosome associated chromatin modification in developing primordial germ cells. PLoS One *2*, e860.
de Rooij, D.G. (1998). Stem cells in the testis. Int J Exp Pathol *79*, 67-80.

De Santa, F., Totaro, M.G., Prosperini, E., Notarbartolo, S., Testa, G., and Natoli, G. (2007). The histone H3 lysine-27 demethylase Jmjd3 links inflammation to inhibition of polycomb-mediated gene silencing. Cell *130*, 1083-1094.

Dechat, T., Gotzmann, J., Stockinger, A., Harris, C.A., Talle, M.A., Siekierka, J.J., and Foisner, R. (1998). Detergent-salt resistance of LAP2alpha in interphase nuclei and phosphorylation-dependent association with chromosomes early in nuclear assembly implies functions in nuclear structure dynamics. Embo J *17*, 4887-4902.

Delaval, K., Govin, J., Cerqueira, F., Rousseaux, S., Khochbin, S., and Feil, R. (2007). Differential histone modifications mark mouse imprinting control regions during spermatogenesis. EMBO J *26*, 720-729.

DeLisle, A.J., Graves, R.A., Marzluff, W.F., and Johnson, L.F. (1983). Regulation of histone mRNA production and stability in serum-stimulated mouse 3T6 fibroblasts. Mol Cell Biol *3*, 1920-1929.

Dellino, G.I., Schwartz, Y.B., Farkas, G., McCabe, D., Elgin, S.C., and Pirrotta, V. (2004). Polycomb silencing blocks transcription initiation. Mol Cell *13*, 887-893.

Dolinoy, D.C., Weidman, J.R., Waterland, R.A., and Jirtle, R.L. (2006). Maternal genistein alters coat color and protects Avy mouse offspring from obesity by modifying the fetal epigenome. Environ Health Perspect *114*, 567-572.

Dorus, S., Busby, S.A., Gerike, U., Shabanowitz, J., Hunt, D.F., and Karr, T.L. (2006). Genomic and functional evolution of the Drosophila melanogaster sperm proteome. Nat Genet *38*, 1440-1445.

Dou, Y., Milne, T.A., Ruthenburg, A.J., Lee, S., Lee, J.W., Verdine, G.L., Allis, C.D., and Roeder, R.G. (2006). Regulation of MLL1 H3K4 methyltransferase activity by its core components. Nat Struct Mol Biol *13*, 713-719.

Durcova-Hills, G., Adams, I.R., Barton, S.C., Surani, M.A., and McLaren, A. (2006). The role of exogenous fibroblast growth factor-2 on the reprogramming of primordial germ cells into pluripotent stem cells. Stem Cells *24*, 1441-1449.

Elderkin, S., Maertens, G.N., Endoh, M., Mallery, D.L., Morrice, N., Koseki, H., Peters, G., Brockdorff, N., and Hiom, K. (2007). A phosphorylated form of Mel-18 targets the Ring1B histone H2A ubiquitin ligase to chromatin. Mol Cell *28*, 107-120.

Endoh, M., Endo, T.A., Endoh, T., Fujimura, Y., Ohara, O., Toyoda, T., Otte, A.P., Okano, M., Brockdorff, N., Vidal, M*., et al.* (2008). Polycomb group proteins Ring1A/B are functionally linked to the core transcriptional regulatory circuitry to maintain ES cell identity. Development *135*, 1513-1524.

Erfurth, F.E., Popovic, R., Grembecka, J., Cierpicki, T., Theisler, C., Xia, Z.B., Stuart, T., Diaz, M.O., Bushweller, J.H., and Zeleznik-Le, N.J. (2008). MLL protects CpG clusters from methylation within the Hoxa9 gene, maintaining transcript expression. Proc Natl Acad Sci U S A *105*, 7517-7522.

Esteve, P.O., Chin, H.G., Smallwood, A., Feehery, G.R., Gangisetty, O., Karpf, A.R., Carey, M.F., and Pradhan, S. (2006). Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. Genes Dev *20*, 3089-3103.

Ezhkova, E., Pasolli, H.A., Parker, J.S., Stokes, N., Su, I.H., Hannon, G., Tarakhovsky, A., and Fuchs, E. (2009). Ezh2 orchestrates gene expression for the stepwise differentiation of tissue-specific stem cells. Cell *136*, 1122-1135.

Fantz, D.A., Hatfield, W.R., Horvath, G., Kistler, M.K., and Kistler, W.S. (2001). Mice with a targeted disruption of the H1t gene are fertile and undergo normal changes in structural chromosomal proteins during spermiogenesis. Biol Reprod *64*, 425-431.

Farthing, C.R., Ficz, G., Ng, R.K., Chan, C.F., Andrews, S., Dean, W., Hemberger, M., and Reik, W. (2008). Global mapping of DNA methylation in mouse promoters reveals epigenetic reprogramming of pluripotency genes. PLoS Genet *4*, e1000116.

Faure, A.K., Pivot-Pajot, C., Kerjean, A., Hazzouri, M., Pelletier, R., Peoc'h, M., Sele, B., Khochbin, S., and Rousseaux, S. (2003). Misregulation of histone acetylation in Sertoli cell-only syndrome and testicular cancer. Mol Hum Reprod *9*, 757-763.

Faust, C., Schumacher, A., Holdener, B., and Magnuson, T. (1995). The eed mutation disrupts anterior mesoderm production in mice. Development *121*, 273-285.

Fedoriw, A.M., Stein, P., Svoboda, P., Schultz, R.M., and Bartolomei, M.S. (2004). Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science *303*, 238-240.

Feil, R. (2009). Epigenetic asymmetry in the zygote and mammalian development. Int J Dev Biol *53*, 191-201.

Fernandez-Capetillo, O., Mahadevaiah, S.K., Celeste, A., Romanienko, P.J., Camerini-Otero, R.D., Bonner, W.M., Manova, K., Burgoyne, P., and Nussenzweig, A. (2003). H2AX is required for chromatin remodeling and inactivation of sex chromosomes in male mouse meiosis. Dev Cell *4*, 497-508.

Fischle, W., Wang, Y., Jacobs, S.A., Kim, Y., Allis, C.D., and Khorasanizadeh, S. (2003). Molecular basis for the discrimination of repressive methyl-lysine marks in histone H3 by Polycomb and HP1 chromodomains. Genes Dev *17*, 1870-1881.

Flanagan, J.M., Popendikyte, V., Pozdniakovaite, N., Sobolev, M., Assadzadeh, A., Schumacher, A., Zangeneh, M., Lau, L., Virtanen, C., Wang, S.C*., et al.* (2006). Intra- and interindividual epigenetic variation in human germ cells. Am J Hum Genet *79*, 67-84.

Francis, N.J., Saurin, A.J., Shao, Z., and Kingston, R.E. (2001). Reconstitution of a functional core polycomb repressive complex. Mol Cell *8*, 545-556.

Gallie, D.R., Lewis, N.J., and Marzluff, W.F. (1996). The histone 3'-terminal stem-loop is necessary for translation in Chinese hamster ovary cells. Nucleic Acids Res *24*, 1954-1962.

Gardiner-Garden, M., Ballesteros, M., Gordon, M., and Tam, P.P. (1998). Histone- and protamine-DNA association: conservation of different patterns within the beta-globin domain in human sperm. Mol Cell Biol *18*, 3350-3356.

Gaskell, T.L., Esnal, A., Robinson, L.L., Anderson, R.A., and Saunders, P.T. (2004). Immunohistochemical profiling of germ cells within the human fetal testis: identification of three subpopulations. Biol Reprod *71*, 2012-2021.

Gatewood, J.M., Cook, G.R., Balhorn, R., Bradbury, E.M., and Schmid, C.W. (1987). Sequence-specific packaging of DNA in human sperm chromatin. Science *236*, 962-964.

Gatewood, J.M., Cook, G.R., Balhorn, R., Schmid, C.W., and Bradbury, E.M. (1990). Isolation of four core histones from human sperm chromatin representing a minor subset of somatic histones. J Biol Chem *265*, 20662-20666.

Geremia, R., Boitani, C., Conti, M., and Monesi, V. (1977). RNA synthesis in spermatocytes and spermatids and preservation of meiotic RNA during spermiogenesis in the mouse. Cell Differ *5*, 343-355.

Gineitis, A.A., Zalenskaya, I.A., Yau, P.M., Bradbury, E.M., and Zalensky, A.O. (2000). Human sperm telomere-binding complex involves histone H2B and secures telomere membrane attachment. J Cell Biol *151*, 1591-1598.

Ginsburg, M., Snow, M.H., and McLaren, A. (1990). Primordial germ cells in the mouse embryo during gastrulation. Development *110*, 521-528.

Glaser, S., Lubitz, S., Loveland, K.L., Ohbo, K., Robb, L., Schwenk, F., Seibler, J., Roellig, D., Kranz, A., Anastassiadis, K*., et al.* (2009). The histone 3 lysine 4 methyltransferase, Mll2, is only required briefly in development and spermatogenesis. Epigenetics Chromatin *2*, 5.

Glaser, S., Schaft, J., Lubitz, S., Vintersten, K., van der Hoeven, F., Tufteland, K.R., Aasland, R., Anastassiadis, K., Ang, S.L., and Stewart, A.F. (2006). Multiple epigenetic maintenance factors implicated by the loss of Mll2 in mouse development. Development *133*, 1423-1432.

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. J Cell Biol *176*, 283-294.

Greaves, I.K., Rangasamy, D., Devoy, M., Marshall Graves, J.A., and Tremethick, D.J. (2006). The X and Y chromosomes assemble into H2A.Z-containing [corrected] facultative heterochromatin [corrected] following meiosis. Mol Cell Biol *26*, 5394-5405.

Groth, A., Corpet, A., Cook, A.J., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. (2007). Regulation of replication fork progression through histone supply and demand. Science *318*, 1928-1931.

Guan, K., Nayernia, K., Maier, L.S., Wagner, S., Dressel, R., Lee, J.H., Nolte, J., Wolf, F., Li, M., Engel, W*., et al.* (2006). Pluripotency of spermatogonial stem cells from adult mouse testis. Nature *440*, 1199-1203.

Guenther, M.G., Jenner, R.G., Chevalier, B., Nakamura, T., Croce, C.M., Canaani, E., and Young, R.A. (2005). Global and Hox-specific roles for the MLL1 methyltransferase. Proc Natl Acad Sci U S A *102*, 8603-8608.

Guenther, M.G., Levine, S.S., Boyer, L.A., Jaenisch, R., and Young, R.A. (2007). A chromatin landmark and transcription initiation at most promoters in human cells. Cell *130*, 77-88.

Gusse, M., and Chevaillier, P. (1980). Electron microscope evidence for the presence of globular structures in different sperm chromatins. J Cell Biol *87*, 280-284.

Haaf, T., and Ward, D.C. (1995). Higher order nuclear structure in mammalian sperm revealed by in situ hybridization and extended chromatin fibers. Exp Cell Res *219*, 604-611.

Hajkova, P., Ancelin, K., Waldmann, T., Lacoste, N., Lange, U.C., Cesari, F., Lee, C., Almouzni, G., Schneider, R., and Surani, M.A. (2008). Chromatin dynamics during epigenetic reprogramming in the mouse germ line. Nature *452*, 877-881.

Hajkova, P., Erhardt, S., Lane, N., Haaf, T., El-Maarri, O., Reik, W., Walter, J., and Surani, M.A. (2002). Epigenetic reprogramming in mouse primordial germ cells. Mech Dev *117*, 15-23.

Hammoud, S.S., Nix, D.A., Zhang, H., Purwar, J., Carrell, D.T., and Cairns, B.R. (2009). Distinctive chromatin in human sperm packages genes for embryo development. Nature *460*, 473-478.

Hansen, K.H., Bracken, A.P., Pasini, D., Dietrich, N., Gehani, S.S., Monrad, A., Rappsilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol *10*, 1291-1300.

Hayashi, K., Yoshida, K., and Matsui, Y. (2005). A histone H3 methyltransferase controls epigenetic events required for meiotic prophase. Nature *438*, 374-378.

Hazzouri, M., Pivot-Pajot, C., Faure, A.K., Usson, Y., Pelletier, R., Sele, B., Khochbin, S., and Rousseaux, S. (2000). Regulated hyperacetylation of core histones during mouse spermatogenesis: involvement of histone deacetylases. Eur J Cell Biol *79*, 950-960.

Hitchins, M.P., Wong, J.J., Suthers, G., Suter, C.M., Martin, D.I., Hawkins, N.J., and Ward, R.L. (2007). Inheritance of a cancer-associated MLH1 germ-line epimutation. N Engl J Med *356*, 697-705.

Hochedlinger, K., and Jaenisch, R. (2003). Nuclear transplantation, embryonic stem cells, and the potential for cell therapy. N Engl J Med *349*, 275-286.

Hoyer-Fender, S., Czirr, E., Radde, R., Turner, J.M., Mahadevaiah, S.K., Pehrson, J.R., and Burgoyne, P.S. (2004). Localisation of histone macroH2A1.2 to the XY-body is not a response to the presence of asynapsed chromosome axes. J Cell Sci *117*, 189-198.

Huang, Y., Fang, J., Bedford, M.T., Zhang, Y., and Xu, R.M. (2006). Recognition of histone H3 lysine-4 methylation by the double tudor domain of JMJD2A. Science *312*, 748-751.

Hublitz, P., Albert, M., and Peters, A.H. (2009). Mechanisms of transcriptional repression by histone lysine methylation. Int J Dev Biol *53*, 335-354.

Imamura, M., Miura, K., Iwabuchi, K., Ichisaka, T., Nakagawa, M., Lee, J., Kanatsu-Shinohara, M., Shinohara, T., and Yamanaka, S. (2006). Transcriptional repression and DNA hypermethylation of a small set of ES cell marker genes in male germline stem cells. BMC Dev Biol *6*, 34.

Inawaka, K., Kawabe, M., Takahashi, S., Doi, Y., Tomigahara, Y., Tarui, H., Abe, J., Kawamura, S., and Shirai, T. (2009). Maternal exposure to anti-androgenic compounds, vinclozolin, flutamide and procymidone, has no effects on spermatogenesis and DNA methylation in male rats of subsequent generations. Toxicol Appl Pharmacol *237*, 178-187.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003). Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res *31*, e15.

Jelinic, P., Stehle, J.C., and Shaw, P. (2006). The testis-specific factor CTCFL cooperates with the protein methyltransferase PRMT7 in H19 imprinting control region methylation. PLoS Biol *4*, e355.

Jin, C., and Felsenfeld, G. (2007). Nucleosome stability mediated by histone variants H3.3 and H2A.Z. Genes Dev *21*, 1519-1529.

Jordan, I.K., Marino-Ramirez, L., and Koonin, E.V. (2005). Evolutionary significance of gene expression divergence. Gene *345*, 119-126.

Jurgens, G. (1985). A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. Nature *316*, 153–155.

Kanatsu-Shinohara, M., Inoue, K., Lee, J., Yoshimoto, M., Ogonuki, N., Miki, H., Baba, S., Kato, T., Kazuki, Y., Toyokuni, S.*, et al.* (2004). Generation of pluripotent stem cells from neonatal mouse testis. Cell *119*, 1001-1012.

Kaneda, M., Okano, M., Hata, K., Sado, T., Tsujimoto, N., Li, E., and 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., and 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. Hum Mol Genet *16*, 2272-2280.

Katz, D.J., Edwards, T.M., Reinke, V., and Kelly, W.G. (2009). A C. elegans LSD1 demethylase contributes to germline immortality by reprogramming epigenetic memory. Cell *137*, 308-320.

Kawahara, M., Wu, Q., Takahashi, N., Morita, S., Yamada, K., Ito, M., Ferguson-Smith, A.C., and Kono, T. (2007). High-frequency generation of viable mice from engineered bi-maternal embryos. Nat Biotechnol *25*, 1045-1050.

Kehler, J., Tolkunova, E., Koschorz, B., Pesce, M., Gentile, L., Boiani, M., Lomeli, H., Nagy, A., McLaughlin, K.J., Scholer, H.R.*, et al.* (2004). Oct4 is required for primordial germ cell survival. EMBO Rep *5*, 1078-1083.

Kepert, J.F., Mazurkiewicz, J., Heuvelman, G.L., Toth, K.F., and Rippe, K. (2005). NAP1 modulates binding of linker histone H1 to chromatin and induces an extended chromatin fiber conformation. J Biol Chem *280*, 34063-34072.

Khalil, A.M., Boyar, F.Z., and Driscoll, D.J. (2004). Dynamic histone modifications mark sex chromosome inactivation and reactivation during mammalian spermatogenesis. Proc Natl Acad Sci U S A *101*, 16583-16587.

Kimmins, S., and Sassone-Corsi, P. (2005). Chromatin remodelling and epigenetic features of germ cells. Nature *434*, 583-589.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. J Cell Biol *153*, 1341-1353.

Kimura, Y., and Yanagimachi, R. (1995). Mouse oocytes injected with testicular spermatozoa or round spermatids can develop into normal offspring. Development *121*, 2397-2405.

King, I.F., Francis, N.J., and Kingston, R.E. (2002). Native and recombinant polycomb group complexes establish a selective block to template accessibility to repress transcription in vitro. Mol Cell Biol *22*, 7919-7928.

Klymenko, T., and Muller, J. (2004). The histone methyltransferases Trithorax and Ash1 prevent transcriptional silencing by Polycomb group proteins. EMBO Rep *5*, 373-377.

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

Kouzarides, T. (2007). Chromatin modifications and their function. Cell *128*, 693-705.
Krawetz, S.A. (2005). Paternal contribution: new insights and future challenges. Nat Rev Genet *6*, 633-642.

Ku, M., Koche, R.P., Rheinbay, E., Mendenhall, E.M., Endoh, M., Mikkelsen, T.S., Presser, A., Nusbaum, C., Xie, X., Chi, A.S.*, et al.* (2008). Genomewide analysis of PRC1 and PRC2 occupancy identifies two classes of bivalent domains. PLoS Genet *4*, e1000242.

Kuramochi-Miyagawa, S., Kimura, T., Yomogida, K., Kuroiwa, A., Tadokoro, Y., Fujita, Y., Sato, M., Matsuda, Y., and Nakano, T. (2001). Two mouse piwi-related genes: miwi and mili. Mech Dev *108*, 121-133.

Kurimoto, K., Yabuta, Y., Ohinata, Y., Shigeta, M., Yamanaka, K., and Saitou, M. (2008). Complex genome-wide transcription dynamics orchestrated by Blimp1 for the specification of the germ cell lineage in mice. Genes Dev *22*, 1617-1635.

Kuzmichev, A., Jenuwein, T., Tempst, P., and Reinberg, D. (2004). Different EZH2-containing complexes target methylation of histone H1 or nucleosomal histone H3. Mol Cell *14*, 183-193.

Kuzmichev, A., Margueron, R., Vaquero, A., Preissner, T.S., Scher, M., Kirmizis, A., Ouyang, X., Brockdorff, N., Abate-Shen, C., Farnham, P.*, et al.* (2005). Composition and histone substrates of polycomb repressive group complexes change during cellular differentiation. Proc Natl Acad Sci U S A *102*, 1859-1864.

Kuzmichev, A., Nishioka, K., Erdjument-Bromage, H., Tempst, P., and Reinberg, D. (2002). Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev *16*, 2893-2905.

La Salle, S., Mertineit, C., Taketo, T., Moens, P.B., Bestor, T.H., and Trasler, J.M. (2004). Windows for sex-specific methylation marked by DNA methyltransferase expression profiles in mouse germ cells. Dev Biol *268*, 403-415.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. Nature *410*, 116-120.

Lacoste, N., and Almouzni, G. (2008). Epigenetic memory: H3.3 steps in the groove. Nat Cell Biol *10*, 7-9.

Lahn, B.T., Tang, Z.L., Zhou, J., Barndt, R.J., Parvinen, M., Allis, C.D., and Page, D.C. (2002). Previously uncharacterized histone acetyltransferases implicated in mammalian spermatogenesis. Proc Natl Acad Sci U S A *99*, 8707-8712.

Lane, N., Dean, W., Erhardt, S., Hajkova, P., Surani, A., Walter, J., and Reik, W. (2003). Resistance of IAPs to methylation reprogramming may provide a mechanism for epigenetic inheritance in the mouse. Genesis *35*, 88-93.

Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol *10*, R25.

Lawson, K.A., Dunn, N.R., Roelen, B.A., Zeinstra, L.M., Davis, A.M., Wright, C.V., Korving, J.P., and Hogan, B.L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. Genes Dev *13*, 424-436.

Lee, J.H., and Skalnik, D.G. (2008). Wdr82 is a C-terminal domain-binding protein that recruits the Setd1A Histone H3-Lys4 methyltransferase complex to transcription start sites of transcribed human genes. Mol Cell Biol *28*, 609-618.

Lee, J.H., Tate, C.M., You, J.S., and Skalnik, D.G. (2007a). Identification and characterization of the human Set1B histone H3-Lys4 methyltransferase complex. J Biol Chem *282*, 13419-13428.

Lee, K., Haugen, H.S., Clegg, C.H., and Braun, R.E. (1995). Premature translation of protamine 1 mRNA causes precocious nuclear condensation and arrests spermatid differentiation in mice. Proc Natl Acad Sci U S A *92*, 12451-12455.

Lee, M.G., Norman, J., Shilatifard, A., and Shiekhattar, R. (2007b). Physical and functional association of a trimethyl H3K4 demethylase and Ring6a/MBLR, a polycomb-like protein. Cell *128*, 877-887.

Lee, M.G., Villa, R., Trojer, P., Norman, J., Yan, K.P., Reinberg, D., Di Croce, L., and Shiekhattar, R. (2007c). Demethylation of H3K27 regulates polycomb recruitment and H2A ubiquitination. Science *318*, 447-450.

Lee, S., Lee, D.K., Dou, Y., Lee, J., Lee, B., Kwak, E., Kong, Y.Y., Lee, S.K., Roeder, R.G., and Lee, J.W. (2006a). Coactivator as a target gene specificity determinant for histone H3 lysine 4 methyltransferases. Proc Natl Acad Sci U S A *103*, 15392-15397.

Lee, T.I., Jenner, R.G., Boyer, L.A., Guenther, M.G., Levine, S.S., Kumar, R.M., Chevalier, B., Johnstone, S.E., Cole, M.F., Isono, K.*, et al.* (2006b). Control of developmental regulators by Polycomb in human embryonic stem cells. Cell *125*, 301-313.

Lepikhov, K., and Walter, J. (2004). Differential dynamics of histone H3 methylation at positions K4 and K9 in the mouse zygote. BMC Dev Biol *4*, 12.

Levchenko, V., and Jackson, V. (2004). Histone release during transcription: NAP1 forms a complex with H2A and H2B and facilitates a topologically dependent release of H3 and H4 from the nucleosome. Biochemistry *43*, 2359-2372.

Levine, S.S., Weiss, A., Erdjument-Bromage, H., Shao, Z., Tempst, P., and Kingston, R.E. (2002). The core of the polycomb repressive complex is compositionally and functionally conserved in flies and humans. Mol Cell Biol *22*, 6070-6078.

Lewis, E.B. (1978). A gene complex controlling segmentation in Drosophila. Nature *276*, 565-570.

Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., and Bradley, A. (1999). Requirement for Wnt3 in vertebrate axis formation. Nat Genet *22*, 361-365.

Loukinov, D.I., Pugacheva, E., Vatolin, S., Pack, S.D., Moon, H., Chernukhin, I., Mannan, P., Larsson, E., Kanduri, C., Vostrov, A.A.*, et al.* (2002). BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. Proc Natl Acad Sci U S A *99*, 6806-6811.

Loyola, A., Bonaldi, T., Roche, D., Imhof, A., and Almouzni, G. (2006). PTMs on H3 variants before chromatin assembly potentiate their final epigenetic state. Mol Cell *24*, 309-316.

Loyola, A., Tagami, H., Bonaldi, T., Roche, D., Quivy, J.P., Imhof, A., Nakatani, Y., Dent, S.Y., and Almouzni, G. (2009). The HP1alpha-CAF1-SetDB1-containing complex provides H3K9me1 for Suv39-mediated K9me3 in pericentric heterochromatin. EMBO Rep *10*, 769-775.

Lubitz, S., Glaser, S., Schaft, J., Stewart, A.F., and Anastassiadis, K. (2007). Increased apoptosis and skewed differentiation in mouse embryonic stem cells lacking the histone methyltransferase Mll2. Mol Biol Cell *18*, 2356-2366.

Ma, T., Van Tine, B.A., Wei, Y., Garrett, M.D., Nelson, D., Adams, P.D., Wang, J., Qin, J., Chow, L.T., and Harper, J.W. (2000). Cell cycle-regulated phosphorylation of p220(NPAT) by cyclin E/Cdk2 in Cajal bodies promotes histone gene transcription. Genes Dev *14*, 2298-2313.

Mahadevaiah, S.K., Turner, J.M., Baudat, F., Rogakou, E.P., de Boer, P., Blanco-Rodriguez, J., Jasin, M., Keeney, S., Bonner, W.M., and Burgoyne, P.S. (2001). Recombinational DNA double-strand breaks in mice precede synapsis. Nat Genet *27*, 271-276.

Margueron, R., Li, G., Sarma, K., Blais, A., Zavadil, J., Woodcock, C.L., Dynlacht, B.D., and Reinberg, D. (2008). Ezh1 and Ezh2 maintain repressive chromatin through different mechanisms. Mol Cell *32*, 503-518.

Marques, C.J., Carvalho, F., Sousa, M., and Barros, A. (2004). Genomic imprinting in disruptive spermatogenesis. Lancet *363*, 1700-1702.

Martianov, I., Brancorsini, S., Catena, R., Gansmuller, A., Kotaja, N., Parvinen, M., Sassone-Corsi, P., and Davidson, I. (2005). Polar nuclear localization of H1T2, a histone H1 variant, required for spermatid elongation and DNA condensation during spermiogenesis. Proc Natl Acad Sci U S A *102*, 2808-2813.

Marushige, Y., and Marushige, K. (1975). Transformation of sperm histone during formation and maturation of rat spermatozoa. J Biol Chem *250*, 39-45.

Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. Nat Rev Genet *9*, 843-854.

Matsui, Y., Zsebo, K., and Hogan, B.L. (1992). Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. Cell *70*, 841-847.

McLaren, A. (1984). Meiosis and differentiation of mouse germ cells. Symp Soc Exp Biol *38*, 7-23.

Meng, X., Lindahl, M., Hyvonen, M.E., Parvinen, M., de Rooij, D.G., Hess, M.W., Raatikainen-Ahokas, A., Sainio, K., Rauvala, H., Lakso, M.*, et al.* (2000). Regulation of cell fate decision of undifferentiated spermatogonia by GDNF. Science *287*, 1489-1493.

Menkveld, R., Wong, W.Y., Lombard, C.J., Wetzels, A.M., Thomas, C.M., Merkus, H.M., and Steegers-Theunissen, R.P. (2001). Semen parameters, including WHO and strict criteria morphology, in a fertile and subfertile population: an effort towards standardization of in-vivo thresholds. Hum Reprod *16*, 1165-1171.

Metzger, E., Wissmann, M., Yin, N., Muller, J.M., Schneider, R., Peters, A.H., Gunther, T., Buettner, R., and Schule, R. (2005). LSD1 demethylates repressive histone marks to promote androgen-receptor-dependent transcription. Nature *437*, 436-439.

Mikkelsen, T.S., Ku, M., Jaffe, D.B., Issac, B., Lieberman, E., Giannoukos, G., Alvarez, P., Brockman, W., Kim, T.K., Koche, R.P.*, et al.* (2007). Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. Nature *448*, 553-560.

Milne, T.A., Dou, Y., Martin, M.E., Brock, H.W., Roeder, R.G., and Hess, J.L. (2005). MLL associates specifically with a subset of transcriptionally active target genes. Proc Natl Acad Sci U S A *102*, 14765-14770.

Min, J., Zhang, Y., and Xu, R.M. (2003). Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. Genes Dev *17*, 1823-1828.

Mithraprabhu, S., and Loveland, K. (2009). Control of KIT signalling in male germ cells: What can we learn from other systems? Reproduction.

Mohn, F., Weber, M., Rebhan, M., Roloff, T.C., Richter, J., Stadler, M.B., Bibel, M., and Schubeler, D. (2008). Lineage-specific polycomb targets and de novo DNA methylation define restriction and potential of neuronal progenitors. Mol Cell *30*, 755-766.

Moldovan, G.L., Pfander, B., and Jentsch, S. (2007). PCNA, the maestro of the replication fork. Cell *129*, 665-679.

Monesi, V. (1965). Differential rate of ribonucleic acid synthesis in the autosomes and sex chromosomes during male meiosis in the mouse. Chromosoma *17*, 11-21.

Monk, M., and McLaren, A. (1981). X-chromosome activity in foetal germ cells of the mouse. J Embryol Exp Morphol *63*, 75-84.

Montgomery, N.D., Yee, D., Chen, A., Kalantry, S., Chamberlain, S.J., Otte, A.P., and Magnuson, T. (2005). The murine polycomb group protein Eed is required for global histone H3 lysine-27 methylation. Curr Biol *15*, 942-947.

Morgan, H.D., Sutherland, H.G., Martin, D.I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. Nat Genet *23*, 314-318.

Morin-Kensicki, E.M., Faust, C., LaMantia, C., and Magnuson, T. (2001). Cell and tissue requirements for the gene eed during mouse gastrulation and organogenesis. Genesis *31*, 142-146.

Mueller, J.L., Mahadevaiah, S.K., Park, P.J., Warburton, P.E., Page, D.C., and Turner, J.M. (2008). The mouse X chromosome is enriched for multicopy testis genes showing postmeiotic expression. Nat Genet *40*, 794-799.

Mullen, T.E., and Marzluff, W.F. (2008). Degradation of histone mRNA requires oligouridylation followed by decapping and simultaneous degradation of the mRNA both 5' to 3' and 3' to 5'. Genes Dev *22*, 50-65.

Murzina, N., Verreault, A., Laue, E., and Stillman, B. (1999). Heterochromatin dynamics in mouse cells: interaction between chromatin assembly factor 1 and HP1 proteins. Mol Cell *4*, 529-540.

Namekawa, S.H., Park, P.J., Zhang, L.F., Shima, J.E., McCarrey, J.R., Griswold, M.D., and Lee, J.T. (2006). Postmeiotic sex chromatin in the male germline of mice. Curr Biol *16*, 660-667.

Negre, N., Hennetin, J., Sun, L.V., Lavrov, S., Bellis, M., White, K.P., and Cavalli, G. (2006). Chromosomal distribution of PcG proteins during Drosophila development. PLoS Biol *4*, e170.

Ng, R.K., Dean, W., Dawson, C., Lucifero, D., Madeja, Z., Reik, W., and Hemberger, M. (2008). Epigenetic restriction of embryonic cell lineage fate by methylation of Elf5. Nat Cell Biol *10*, 1280-1290.

Ng, R.K., and Gurdon, J.B. (2008). Epigenetic memory of an active gene state depends on histone H3.3 incorporation into chromatin in the absence of transcription. Nat Cell Biol *10*, 102-109.

Nguyen, P., Bar-Sela, G., Sun, L., Bisht, K.S., Cui, H., Kohn, E., Feinberg, A.P., and Gius, D. (2008). BAT3 and SET1A form a complex with CTCFL/BORIS to modulate H3K4 histone dimethylation and gene expression. Mol Cell Biol *28*, 6720-6729.

Niwa, H. (2007). How is pluripotency determined and maintained? Development *134*, 635-646.

O'Carroll, D., Erhardt, S., Pagani, M., Barton, S.C., Surani, M.A., and Jenuwein, T. (2001). The polycomb-group gene Ezh2 is required for early mouse development. Mol Cell Biol *21*, 4330-4336.

O'Geen, H., Nicolet, C.M., Blahnik, K., Green, R., and Farnham, P.J. (2006). Comparison of sample preparation methods for ChIP-chip assays. Biotechniques *41*, 577-580.

Oakes, C.C., La Salle, S., Smiraglia, D.J., Robaire, B., and Trasler, J.M. (2007). Developmental acquisition of genome-wide DNA methylation occurs prior to meiosis in male germ cells. Dev Biol *307*, 368-379.

Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., and Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. Cell *137*, 571-584.

Ohinata, Y., Payer, B., O'Carroll, D., Ancelin, K., Ono, Y., Sano, M., Barton, S.C., Obukhanych, T., Nussenzweig, M., Tarakhovsky, A.*, et al.* (2005). Blimp1 is a critical determinant of the germ cell lineage in mice. Nature *436*, 207-213.

Ohta, H., Tohda, A., and Nishimune, Y. (2003). Proliferation and differentiation of spermatogonial stem cells in the w/wv mutant mouse testis. Biol Reprod *69*, 1815-1821.

Okada, Y., Scott, G., Ray, M.K., Mishina, Y., and Zhang, Y. (2007). Histone demethylase JHDM2A is critical for Tnp1 and Prm1 transcription and spermatogenesis. Nature *450*, 119-123.

Olek, A., and Walter, J. (1997). The pre-implantation ontogeny of the H19 methylation imprint. Nat Genet *17*, 275-276.

Ooi, S.L., and Henikoff, S. (2007). Germline histone dynamics and epigenetics. Curr Opin Cell Biol *19*, 257-265.

Orford, K., Kharchenko, P., Lai, W., Dao, M.C., Worhunsky, D.J., Ferro, A., Janzen, V., Park, P.J., and Scadden, D.T. (2008). Differential H3K4 methylation identifies developmentally poised hematopoietic genes. Dev Cell *14*, 798-809.

Ostermeier, G.C., Dix, D.J., Miller, D., Khatri, P., and Krawetz, S.A. (2002). Spermatozoal RNA profiles of normal fertile men. Lancet *360*, 772-777.

Ostermeier, G.C., Goodrich, R.J., Moldenhauer, J.S., Diamond, M.P., and Krawetz, S.A. (2005). A suite of novel human spermatozoal RNAs. J Androl *26*, 70-74.

Palmer, D.K., O'Day, K., and Margolis, R.L. (1990). The centromere specific histone CENP-A is selectively retained in discrete foci in mammalian sperm nuclei. Chromosoma *100*, 32-36.

Pan, G., Tian, S., Nie, J., Yang, C., Ruotti, V., Wei, H., Jonsdottir, G.A., Stewart, R., and Thomson, J.A. (2007). Whole-genome analysis of histone H3 lysine 4 and lysine 27 methylation in human embryonic stem cells. Cell Stem Cell *1*, 299-312.

Pan, H., O'Brien M, J., Wigglesworth, K., Eppig, J.J., and Schultz, R.M. (2005). Transcript profiling during mouse oocyte development and the effect of gonadotropin priming and development in vitro. Dev Biol *286*, 493-506.

Pandey, N.B., and Marzluff, W.F. (1987). The stem-loop structure at the 3' end of histone mRNA is necessary and sufficient for regulation of histone mRNA stability. Mol Cell Biol *7*, 4557-4559.

Pant, V., Mariano, P., Kanduri, C., Mattsson, A., Lobanenkov, V., Heuchel, R., and Ohlsson, R. (2003). The nucleotides responsible for the direct physical contact between the chromatin insulator protein CTCF and the H19 imprinting control region manifest parent of origin-specific long-distance insulation and methylation-free domains. Genes Dev *17*, 586-590.

Park, Y.J., Chodaparambil, J.V., Bao, Y., McBryant, S.J., and Luger, K. (2005). Nucleosome assembly protein 1 exchanges histone H2A-H2B dimers and assists nucleosome sliding. J Biol Chem *280*, 1817-1825.

Pasini, D., Bracken, A.P., Hansen, J.B., Capillo, M., and Helin, K. (2007). The polycomb group protein Suz12 is required for embryonic stem cell differentiation. Mol Cell Biol *27*, 3769-3779.

Pasini, D., Bracken, A.P., Jensen, M.R., Lazzerini Denchi, E., and Helin, K. (2004). Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. Embo J *23*, 4061-4071.

Pasini, D., Hansen, K.H., Christensen, J., Agger, K., Cloos, P.A., and Helin, K. (2008). Coordinated regulation of transcriptional repression by the RBP2 H3K4 demethylase and Polycomb-Repressive Complex 2. Genes Dev *22*, 1345-1355.

Pembrey, M.E., Bygren, L.O., Kaati, G., Edvinsson, S., Northstone, K., Sjostrom, M., and Golding, J. (2006). Sex-specific, male-line transgenerational responses in humans. Eur J Hum Genet *14*, 159-166.

Pesce, M., Wang, X., Wolgemuth, D.J., and Scholer, H. (1998). Differential expression of the Oct-4 transcription factor during mouse germ cell differentiation. Mech Dev *71*, 89-98.

Peters, A.H., O'Carroll, D., Scherthan, H., Mechtler, K., Sauer, S., Schofer, C., Weipoltshammer, K., Pagani, M., Lachner, M., Kohlmaier, A.*, et al.* (2001). Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell *107*, 323-337.

Phalke, S., Nickel, O., Walluscheck, D., Hortig, F., Onorati, M.C., and Reuter, G. (2009). Retrotransposon silencing and telomere integrity in somatic cells of Drosophila depends on the cytosine-5 methyltransferase DNMT2. Nat Genet *41*, 696-702.

Phillips, J.E., and Corces, V.G. (2009). CTCF: master weaver of the genome. Cell *137*, 1194-1211.

Pittoggi, C., Renzi, L., Zaccagnini, G., Cimini, D., Degrassi, F., Giordano, R., Magnano, A.R., Lorenzini, R., Lavia, P., and Spadafora, C. (1999). A fraction of mouse sperm chromatin is organized in nucleosomal hypersensitive domains enriched in retroposon DNA. J Cell Sci *112 (Pt 20)*, 3537-3548.

Pivot-Pajot, C., Caron, C., Govin, J., Vion, A., Rousseaux, S., and Khochbin, S. (2003). Acetylation-dependent chromatin reorganization by BRDT, a testis-specific bromodomain-containing protein. Mol Cell Biol *23*, 5354-5365.

Poux, S., Horard, B., Sigrist, C.J., and Pirrotta, V. (2002). The Drosophila trithorax protein is a coactivator required to prevent re-establishment of polycomb silencing. Development *129*, 2483-2493.

Prabhu, S.M., Meistrich, M.L., McLaughlin, E.A., Roman, S.D., Warne, S., Mendis, S., Itman, C., and Loveland, K.L. (2006). Expression of c-Kit receptor mRNA and protein in the developing, adult and irradiated rodent testis. Reproduction *131*, 489-499.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. Nat Rev Mol Cell Biol *10*, 192-206.

Puschendorf, M., Terranova, R., Boutsma, E., Mao, X., Isono, K., Brykczynska, U., Kolb, C., Otte, A.P., Koseki, H., Orkin, S.H.*, et al.* (2008). PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos. Nat Genet *40*, 411-420.

Puwaravutipanich, T., and Panyim, S. (1975). The nuclear basic proteins of human testes and ejaculated spermatozoa. Exp Cell Res *90*, 153-158.

Quivy, J.P., Gerard, A., Cook, A.J., Roche, D., and Almouzni, G. (2008). The HP1-p150/CAF-1 interaction is required for pericentric heterochromatin replication and S-phase progression in mouse cells. Nat Struct Mol Biol *15*, 972-979.

Rakyan, V.K., Chong, S., Champ, M.E., Cuthbert, P.C., Morgan, H.D., Luu, K.V., and Whitelaw, E. (2003). Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. Proc Natl Acad Sci U S A *100*, 2538-2543.

Rando, O.J., and Verstrepen, K.J. (2007). Timescales of genetic and epigenetic inheritance. Cell *128*, 655-668.

Rassoulzadegan, M., Grandjean, V., Gounon, P., Vincent, S., Gillot, I., and Cuzin, F. (2006). RNA-mediated non-mendelian inheritance of an epigenetic change in the mouse. Nature *441*, 469-474.

Reik, W. (2007). Stability and flexibility of epigenetic gene regulation in mammalian development. Nature *447*, 425-432.

Rice, J.C., Briggs, S.D., Ueberheide, B., Barber, C.M., Shabanowitz, J., Hunt, D.F., Shinkai, Y., and Allis, C.D. (2003). Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. Mol Cell *12*, 1591-1598.

Ringrose, L., and Paro, R. (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. Development *134*, 223-232.

Rougier, N., Bourc'his, D., Gomes, D.M., Niveleau, A., Plachot, M., Paldi, A., and Viegas-Pequignot, E. (1998). Chromosome methylation patterns during mammalian preimplantation development. Genes Dev *12*, 2108-2113.

Ruden, D.M., Jamison, D.C., Zeeberg, B.R., Garfinkel, M.D., Weinstein, J.N., Rasouli, P., and Lu, X. (2008). The EDGE hypothesis: epigenetically directed genetic errors in repeat-containing proteins (RCPs) involved in evolution, neuroendocrine signaling, and cancer. Front Neuroendocrinol *29*, 428-444.

Ruthenburg, A.J., Allis, C.D., and Wysocka, J. (2007). Methylation of lysine 4 on histone H3: intricacy of writing and reading a single epigenetic mark. Mol Cell *25*, 15-30.

Saitou, M., Barton, S.C., and Surani, M.A. (2002). A molecular programme for the specification of germ cell fate in mice. Nature *418*, 293-300.

Santos, F., Peters, A.H., Otte, A.P., Reik, W., and Dean, W. (2005). Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. Dev Biol *280*, 225-236.

Sarraf, S.A., and Stancheva, I. (2004). Methyl-CpG binding protein MBD1 couples histone H3 methylation at lysine 9 by SETDB1 to DNA replication and chromatin assembly. Mol Cell *15*, 595-605.

Schliep, A., Steinhoff, C., and Schonhuth, A. (2004). Robust inference of groups in gene expression time-courses using mixtures of HMMs. Bioinformatics *20 Suppl 1*, i283-289.

Schoeftner, S., Sengupta, A.K., Kubicek, S., Mechtler, K., Spahn, L., Koseki, H., Jenuwein, T., and Wutz, A. (2006). Recruitment of PRC1 function at the initiation of X inactivation independent of PRC2 and silencing. EMBO J *25*, 3110-3122.

Seki, Y., Hayashi, K., Itoh, K., Mizugaki, M., Saitou, M., and Matsui, Y. (2005). Extensive and orderly reprogramming of genome-wide chromatin modifications associated with specification and early development of germ cells in mice. Dev Biol *278*, 440-458.

Seki, Y., Yamaji, M., Yabuta, Y., Sano, M., Shigeta, M., Matsui, Y., Saga, Y., Tachibana, M., Shinkai, Y., and Saitou, M. (2007). Cellular dynamics associated with the genome-wide epigenetic reprogramming in migrating primordial germ cells in mice. Development *134*, 2627-2638.

Seydoux, G., and Braun, R.E. (2006). Pathway to totipotency: lessons from germ cells. Cell *127*, 891-904.

Shang, E., Nickerson, H.D., Wen, D., Wang, X., and Wolgemuth, D.J. (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.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J.R., Wu, C.T., Bender, W., and Kingston, R.E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell *98*, 37-46.

Shen, J.C., Rideout, W.M., 3rd, and Jones, P.A. (1994). The rate of hydrolytic deamination of 5-methylcytosine in double-stranded DNA. Nucleic Acids Res *22*, 972-976.

Shen, X., Liu, Y., Hsu, Y.J., Fujiwara, Y., Kim, J., Mao, X., Yuan, G.C., and Orkin, S.H. (2008). EZH1 mediates methylation on histone H3 lysine 27 and complements EZH2 in maintaining stem cell identity and executing pluripotency. Mol Cell *32*, 491-502.

Shi, Y., Lan, F., Matson, C., Mulligan, P., Whetstine, J.R., Cole, P.A., and Casero, R.A. (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell *119*, 941-953.

Shibahara, K., and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell *96*, 575-585.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. Science *311*, 844-847.

Singleton, S., Zalensky, A., Doncel, G.F., Morshedi, M., and Zalenskaya, I.A. (2007). Testis/sperm-specific histone 2B in the sperm of donors and subfertile patients: variability and relation to chromatin packaging. Hum Reprod *22*, 743-750.

Smyth, G.K. (2005). Limma: linear models for microarray data. In *Bioinformatics and Computational Biology Solutions using R and Bioconductor*
V.C. R. Gentleman, S. Dudoit, R. Irizarry, W.Huber, ed. (New York, Springer).

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

Solari, A.J. (1974). The behavior of the XY pair in mammals. Int Rev Cytol *38*, 273-317.

Sollars, V., Lu, X., Xiao, L., Wang, X., Garfinkel, M.D., and Ruden, D.M. (2003). Evidence for an epigenetic mechanism by which Hsp90 acts as a capacitor for morphological evolution. Nat Genet *33*, 70-74.

Sparmann, A., and van Lohuizen, M. (2006). Polycomb silencers control cell fate, development and cancer. Nat Rev Cancer *6*, 846-856.

Stam, M., Belele, C., Dorweiler, J.E., and Chandler, V.L. (2002). Differential chromatin structure within a tandem array 100 kb upstream of the maize b1 locus is associated with paramutation. Genes Dev *16*, 1906-1918.

Steward, M.M., Lee, J.S., O'Donovan, A., Wyatt, M., Bernstein, B.E., and Shilatifard, A. (2006). Molecular regulation of H3K4 trimethylation by ASH2L, a shared subunit of MLL complexes. Nat Struct Mol Biol *13*, 852-854.

Stock, J.K., Giadrossi, S., Casanova, M., Brookes, E., Vidal, M., Koseki, H., Brockdorff, N., Fisher, A.G., and Pombo, A. (2007). Ring1-mediated ubiquitination of H2A restrains poised RNA polymerase II at bivalent genes in mouse ES cells. Nat Cell Biol *9*, 1428-1435.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Suetake, I., Shinozaki, F., Miyagawa, J., Takeshima, H., and Tajima, S. (2004). DNMT3L stimulates the DNA methylation activity of Dnmt3a and Dnmt3b through a direct interaction. J Biol Chem *279*, 27816-27823.

Sugimoto, M., and Abe, K. (2007). X chromosome reactivation initiates in nascent primordial germ cells in mice. PLoS Genet *3*, e116.

Sullivan, K.F., Hechenberger, M., and Masri, K. (1994). Human CENP-A contains a histone H3 related histone fold domain that is required for targeting to the centromere. J Cell Biol *127*, 581-592.

Surani, M.A., and Barton, S.C. (1983). Development of gynogenetic eggs in the mouse: implications for parthenogenetic embryos. Science *222*, 1034-1036.

Suter, C.M., and Martin, D.I. (2007). Inherited epimutation or a haplotypic basis for the propensity to silence? Nat Genet *39*, 573; author reply 576.

Suter, C.M., Martin, D.I., and Ward, R.L. (2004). Germline epimutation of MLH1 in individuals with multiple cancers. Nat Genet *36*, 497-501.

Swigut, T., and Wysocka, J. (2007). H3K27 demethylases, at long last. Cell *131*, 29-32.

Tachibana, M., Nozaki, M., Takeda, N., and Shinkai, Y. (2007). Functional dynamics of H3K9 methylation during meiotic prophase progression. EMBO J *26*, 3346-3359.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell *116*, 51-61.

Takada, Y., Isono, K., Shinga, J., Turner, J.M., Kitamura, H., Ohara, O., Watanabe, G., Singh, P.B., Kamijo, T., Jenuwein, T.*, et al.* (2007). Mammalian Polycomb Scmh1 mediates exclusion of Polycomb complexes from the XY body in the pachytene spermatocytes. Development *134*, 579-590.

Tamashiro, K.L., Kimura, Y., Blanchard, R.J., Blanchard, D.C., and Yanagimachi, R. (1999). Bypassing spermiogenesis for several generations does not have detrimental consequences on the fertility and neurobehavior of offspring: a study using the mouse. J Assist Reprod Genet *16*, 315-324.

Tanphaichitr, N., Sobhon, P., Taluppeth, N., and Chalermisarachai, P. (1978). Basic nuclear proteins in testicular cells and ejaculated spermatozoa in man. Exp Cell Res *117*, 347-356.

Terranova, R., Yokobayashi, S., Stadler, M.B., Otte, A.P., van Lohuizen, M., Orkin, S.H., and Peters, A.H. (2008). Polycomb group proteins Ezh2 and Rnf2 direct genomic contraction and imprinted repression in early mouse embryos. Dev Cell *15*, 668-679.

Torres-Padilla, M.E., Bannister, A.J., Hurd, P.J., Kouzarides, T., and Zernicka-Goetz, M. (2006). Dynamic distribution of the replacement histone variant H3.3 in the mouse oocyte and preimplantation embryos. Int J Dev Biol *50*, 455-461.

Turner, J.M., Aprelikova, O., Xu, X., Wang, R., Kim, S., Chandramouli, G.V., Barrett, J.C., Burgoyne, P.S., and Deng, C.X. (2004). BRCA1, histone H2AX phosphorylation, and male meiotic sex chromosome inactivation. Curr Biol *14*, 2135-2142.

Turner, J.M., Mahadevaiah, S.K., Benavente, R., Offenberg, H.H., Heyting, C., and Burgoyne, P.S. (2000). Analysis of male meiotic "sex body" proteins during XY female meiosis provides new insights into their functions. Chromosoma *109*, 426-432.

Turner, J.M., Mahadevaiah, S.K., Fernandez-Capetillo, O., Nussenzweig, A., Xu, X., Deng, C.X., and Burgoyne, P.S. (2005). Silencing of unsynapsed meiotic chromosomes in the mouse. Nat Genet *37*, 41-47.

Tweedie, S., Charlton, J., Clark, V., and Bird, A. (1997). Methylation of genomes and genes at the invertebrate-vertebrate boundary. Mol Cell Biol *17*, 1469-1475.

Umlauf, D., Goto, Y., and Feil, R. (2004). Site-specific analysis of histone methylation and acetylation. Methods Mol Biol *287*, 99-120.

van der Heijden, G.W., Derijck, A.A., Posfai, E., Giele, M., Pelczar, P., Ramos, L., Wansink, D.G., van der Vlag, J., Peters, A.H., and de Boer, P. (2007). Chromosome-wide nucleosome replacement and H3.3 incorporation during mammalian meiotic sex chromosome inactivation. Nat Genet *39*, 251-258.

van der Heijden, G.W., Derijck, A.A., Ramos, L., Giele, M., van der Vlag, J., and de Boer, P. (2006). Transmission of modified nucleosomes from the mouse male germline to the zygote and subsequent remodeling of paternal chromatin. Dev Biol *298*, 458-469.

van der Heijden, G.W., Dieker, J.W., Derijck, A.A., Muller, S., Berden, J.H., Braat, D.D., van der Vlag, J., and de Boer, P. (2005). Asymmetry in histone H3 variants and lysine methylation between paternal and maternal chromatin of the early mouse zygote. Mech Dev *122*, 1008-1022.

van der Heijden, G.W., Ramos, L., Baart, E.B., van den Berg, I.M., Derijck, A.A., van der Vlag, J., Martini, E., and de Boer, P. (2008). Sperm-derived histones contribute to zygotic chromatin in humans. BMC Dev Biol *8*, 34.

van der Heijden, G.W., van den Berg, I.M., Baart, E.B., Derijck, A.A., Martini, E., and de Boer, P. (2009). Parental origin of chromatin in human monopronuclear zygotes revealed by asymmetric histone methylation patterns, differs between IVF and ICSI. Mol Reprod Dev *76*, 101-108.

van der Stoop, P., Boutsma, E.A., Hulsman, D., Noback, S., Heimerikx, M., Kerkhoven, R.M., Voncken, J.W., Wessels, L.F., and van Lohuizen, M. (2008). Ubiquitin E3 ligase Ring1b/Rnf2 of polycomb repressive complex 1 contributes to stable maintenance of mouse embryonic stem cells. PLoS One *3*, e2235.

van Roijen, H.J., Ooms, M.P., Spaargaren, M.C., Baarends, W.M., Weber, R.F., Grootegoed, J.A., and Vreeburg, J.T. (1998). Immunoexpression of testis-specific histone 2B in human spermatozoa and testis tissue. Hum Reprod *13*, 1559-1566.

Vire, E., Brenner, C., Deplus, R., Blanchon, L., Fraga, M., Didelot, C., Morey, L., Van Eynde, A., Bernard, D., Vanderwinden, J.M*., et al.* (2006). The Polycomb group protein EZH2 directly controls DNA methylation. Nature *439*, 871-874.

Voncken, J.W., Roelen, B.A., Roefs, M., de Vries, S., Verhoeven, E., Marino, S., Deschamps, J., and van Lohuizen, M. (2003). Rnf2 (Ring1b) deficiency causes gastrulation arrest and cell cycle inhibition. Proc Natl Acad Sci U S A *100*, 2468-2473.

Wagner, K.D., Wagner, N., Ghanbarian, H., Grandjean, V., Gounon, P., Cuzin, F., and Rassoulzadegan, M. (2008). RNA induction and inheritance of epigenetic cardiac hypertrophy in the mouse. Dev Cell *14*, 962-969.

Wang, H., An, W., Cao, R., Xia, L., Erdjument-Bromage, H., Chatton, B., Tempst, P., Roeder, R.G., and Zhang, Y. (2003). mAM facilitates conversion by ESET of dimethyl to trimethyl lysine 9 of histone H3 to cause transcriptional repression. Mol Cell *12*, 475-487.

Wang, H., Wang, L., Erdjument-Bromage, H., Vidal, M., Tempst, P., Jones, R.S., and Zhang, Y. (2004). Role of histone H2A ubiquitination in Polycomb silencing. Nature *431*, 873-878.

Waterland, R.A., Travisano, M., and Tahiliani, K.G. (2007). Diet-induced hypermethylation at agouti viable yellow is not inherited transgenerationally through the female. FASEB J *21*, 3380-3385.

Watson, J.D., and Crick, F.H. (1953). Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid. Nature *171*, 737-738.

Weber, M., Hellmann, I., Stadler, M.B., Ramos, L., Paabo, S., Rebhan, M., and Schubeler, D. (2007). Distribution, silencing potential and evolutionary impact of promoter DNA methylation in the human genome. Nat Genet *39*, 457-466.

Webster, K.E., O'Bryan, M.K., Fletcher, S., Crewther, P.E., Aapola, U., Craig, J., Harrison, D.K., Aung, H., Phutikanit, N., Lyle, R*., et al.* (2005). Meiotic and epigenetic defects in Dnmt3L-knockout mouse spermatogenesis. Proc Natl Acad Sci U S A *102*, 4068-4073.

WHO Laboratory Manual for the Examination of Human Spermatozoa and Semen-Cervical Mucus Interaction, (1999) 4th edn (Cambridge University Press).

Wolffe, A.P., and Hayes, J.J. (1999). Chromatin disruption and modification. Nucleic Acids Res *27*, 711-720.

Wu, F., Caron, C., De Robertis, C., Khochbin, S., and Rousseaux, S. (2008). Testis-specific histone variants H2AL1/2 rapidly disappear from paternal heterochromatin after fertilization. J Reprod Dev *54*, 413-417.

Wu, J.Y., Ribar, T.J., Cummings, D.E., Burton, K.A., McKnight, G.S., and Means, A.R. (2000). Spermiogenesis and exchange of basic nuclear proteins are impaired in male germ cells lacking Camk4. Nat Genet *25*, 448-452.

Wutz, A., and Gribnau, J. (2007). X inactivation Xplained. Curr Opin Genet Dev *17*, 387-393.

Wykes, S.M., and Krawetz, S.A. (2003). The structural organization of sperm chromatin. J Biol Chem *278*, 29471-29477.

Wysocka, J., Swigut, T., Milne, T.A., Dou, Y., Zhang, X., Burlingame, A.L., Roeder, R.G., Brivanlou, A.H., and Allis, C.D. (2005). WDR5 associates with histone H3 methylated at K4 and is essential for H3 K4 methylation and vertebrate development. Cell *121*, 859-872.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P.*, et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature *442*, 86-90.

Xing, Y., Ouyang, Z., Kapur, K., Scott, M.P., and Wong, W.H. (2007). Assessing the conservation of mammalian gene expression using high-density exon arrays. Mol Biol Evol *24*, 1283-1285.

Yabuta, Y., Kurimoto, K., Ohinata, Y., Seki, Y., and Saitou, M. (2006). Gene expression dynamics during germline specification in mice identified by quantitative single-cell gene expression profiling. Biol Reprod *75*, 705-716.

Yamaji, M., Seki, Y., Kurimoto, K., Yabuta, Y., Yuasa, M., Shigeta, M., Yamanaka, K., Ohinata, Y., and Saitou, M. (2008). Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat Genet *40*, 1016-1022.

Yan, W., Ma, L., Burns, K.H., and Matzuk, M.M. (2003). HILS1 is a spermatid-specific linker histone H1-like protein implicated in chromatin remodeling during mammalian spermiogenesis. Proc Natl Acad Sci U S A *100*, 10546-10551.

Yeom, Y.I., Fuhrmann, G., Ovitt, C.E., Brehm, A., Ohbo, K., Gross, M., Hubner, K., and Scholer, H.R. (1996). Germline regulatory element of Oct-4 specific for the totipotent cycle of embryonal cells. Development *122*, 881-894.

Ying, Y., Liu, X.M., Marble, A., Lawson, K.A., and Zhao, G.Q. (2000). Requirement of Bmp8b for the generation of primordial germ cells in the mouse. Mol Endocrinol *14*, 1053-1063.

Yoshida, S., Sukeno, M., Nakagawa, T., Ohbo, K., Nagamatsu, G., Suda, T., and Nabeshima, Y. (2006). The first round of mouse spermatogenesis is a distinctive program that lacks the self-renewing spermatogonia stage. Development *133*, 1495-1505.

Yu, B.D., Hess, J.L., Horning, S.E., Brown, G.A., and Korsmeyer, S.J. (1995). Altered Hox expression and segmental identity in Mll-mutant mice. Nature *378*, 505-508.

Yu, Y.E., Zhang, Y., Unni, E., Shirley, C.R., Deng, J.M., Russell, L.D., Weil, M.M., Behringer, R.R., and Meistrich, M.L. (2000). Abnormal spermatogenesis and reduced fertility in transition nuclear protein 1-deficient mice. Proc Natl Acad Sci U S A *97*, 4683-4688.

Zalenskaya, I.A., Bradbury, E.M., and Zalensky, A.O. (2000). Chromatin structure of telomere domain in human sperm. Biochem Biophys Res Commun *279*, 213-218.

Zalensky, A.O., Allen, M.J., Kobayashi, A., Zalenskaya, I.A., Balhorn, R., and Bradbury, E.M. (1995). Well-defined genome architecture in the human sperm nucleus. Chromosoma *103*, 577-590.

Zalensky, A.O., Breneman, J.W., Zalenskaya, I.A., Brinkley, B.R., and Bradbury, E.M. (1993). Organization of centromeres in the decondensed nuclei of mature human sperm. Chromosoma *102*, 509-518.

Zalensky, A.O., Siino, J.S., Gineitis, A.A., Zalenskaya, I.A., Tomilin, N.V., Yau, P., and Bradbury, E.M. (2002). Human testis/sperm-specific histone H2B (hTSH2B). Molecular cloning and characterization. J Biol Chem *277*, 43474-43480.

Zeng, F., and Schultz, R.M. (2005). RNA transcript profiling during zygotic gene activation in the preimplantation mouse embryo. Dev Biol *283*, 40-57.

Zhao, J., Kennedy, B.K., Lawrence, B.D., Barbie, D.A., Matera, A.G., Fletcher, J.A., and Harlow, E. (2000). NPAT links cyclin E-Cdk2 to the regulation of replication-dependent histone gene transcription. Genes Dev *14*, 2283-2297.

Zhao, M., Shirley, C.R., Hayashi, S., Marcon, L., Mohapatra, B., Suganuma, R., Behringer, R.R., Boissonneault, G., Yanagimachi, R., and Meistrich, M.L. (2004). Transition nuclear proteins are required for normal chromatin condensation and functional sperm development. Genesis *38*, 200-213.

Zhao, M., Shirley, C.R., Yu, Y.E., Mohapatra, B., Zhang, Y., Unni, E., Deng, J.M., Arango, N.A., Terry, N.H., Weil, M.M*., et al.* (2001). Targeted disruption of the transition protein 2 gene affects sperm chromatin structure and reduces fertility in mice. Mol Cell Biol *21*, 7243-7255.

Zhao, X.D., Han, X., Chew, J.L., Liu, J., Chiu, K.P., Choo, A., Orlov, Y.L., Sung, W.K., Shahab, A., Kuznetsov, V.A*., et al.* (2007). Whole-genome mapping of histone H3 Lys4 and 27 trimethylations reveals distinct genomic compartments in human embryonic stem cells. Cell Stem Cell *1*, 286-298.

Zlatanova, J., Seebart, C., and Tomschik, M. (2007). Nap1: taking a closer look at a juggler protein of extraordinary skills. FASEB J *21*, 1294-1310.

Acknowledgments

I would like to thank Antoine Peters for being a great supervisor. By his guidance and his enthusiasm for science he taught me a lot. My PhD in Antoine's lab was a great, unforgettable time.

I would like to thank the members of my PhD committee, Genevieve Almouzni, Susan Gasser and Jan Hofsteenge. My yearly meetings with them have been very helpful and encouraging. I would like also to thank Fred Mains who agreed to chair my PhD defence.

I would like to thank all the present and past members of the Peter's lab for creating a great, stimulating and joyful working atmosphere. Especially, I thank Mizue Hisano with whom I collaborated on my project and who greatly contributed to the results presented in this thesis.

I would like to thank Liliana Ramos from the Radboud University Medical Center in The Netherlands, for her great help and for a very effective collaboration on the human sperm project.

I would like to thank Dirk Schübeler and his lab members for many discussions and helpful suggestions.

I acknowledge all the excellent facilities at the FMI. Especially, I thank Michael Stadler from the bioinformatic support, for his great help and expertise, which greatly contributed to my thesis.

I acknowledge The Boehringer Ingelheim Foundation (BIF) for providing me with funding and also for organizing great meetings and seminars.

I thank Rico, my family and friends for supporting me during these five years.

CURRICULUM VITAE

URSZULA BRYKCZYNSKA

PERSONAL DETAILS

Current position: PhD student
Friedrich Miescher Institute for Biomedical Research (FMI)
Maulbeerstrasse 66, CH-4058 Basel, Switzerland

Phone: +41 61 6976626
E-mail: urszula.brykczynska@fmi.ch

Personal: Date of birth: 11th May 1981
Nationality: Polish

EDUCATION

2004 to date **PhD student in Biochemistry**
Friedrich Miescher Institute for Biomedical Research (FMI), Basel, Switzerland
- Supervisor: Dr. Antoine H. F. M. Peters
- Funding: Boehringer Ingelheim Fonds PhD Fellowship

Thesis: Role of histone methylation in paternal transmission of epigenetic information via mouse and human spermatozoa

2002 - 2004 **Masters of Science in Biotechnology** (with distinction)
Faculty of Biology, Warsaw University, Poland
Mass Spectrometry Group, University of Amsterdam, The Netherlands
- Supervisors: Prof. Andrzej Jerzmanowski, Dr. Luitzen de Jong
- Funding: University of Warsaw Scholarship, Erasmus programme

Project: Identification of histone posttranslational modifications by mass spectrometry
Thesis: Reconstitution of plant chromatin *in vitro*

1999 - 2002 **Bachelors of Science in Biotechnology** (with distinction)
Faculty of Biology, Warsaw University, Poland
- Supervisor: Prof. Andrzej Jerzmanowski
- Funding: University of Warsaw Scholarship

Thesis: Application of microarray technology in plant molecular biology

RESEARCH TRAINING AND EXPERIENCE

2008 **Workshop** on Computational and Statistical Aspects of Microarray Analysis, Bressanone, Italy

2007 **Workshop** on Scientific Communication, Cold Spring Harbor, USA

2003 **Summer Internship** University of Toronto, Canada, Laboratory of Andrew Spence, Student project on the genetics of *C.elegans*

2000 **Summer Internship** University of Geneva, Switzerland, Laboratory of Jan Pawlowski, Phylogenetic research on Antarctic Foraminifers

PUBLICATIONS

Brykczynska U, Hisano M, Ramos L, Oakeley EJ, Roloff TC, Schübeler D, Stadler MB and AHFM Peters; Transmission of methylated histones with modification-specific genomic localization by human spermatozoa. *Submitted*

Haueter S, Kawasumi M, Asner I, **Brykczynska U**, Cinelli P, Moisyadi S, Bürki K, Peters AHFM, and P Pelczar. Genetic vasectomy - overexpression of Prm1-EGFP fusion protein in elongating spermatids causes dominant male sterility in mice. *Submitted*

Madan V, Madan B, **Brykczynska U**, Zilbermann F, Hogeveen K, Dohner K, Dohner H, Weber O, Blum C, Rodewald HR, Sassone-Corsi P, Peters AHFM, and HJ Fehling; Impaired function of primitive hematopoietic cells in mice lacking the Mixed-Lineage-Leukemia homolog Mll5. *Blood*, 113(7):1444-54 (2009).

Puschendorf M, Terranova R, Bautsma E, Mao X, Isono K, **Brykczynska U**, Kolb C, Otte AP, Koseki H, Orkin SH, van Lohuisen M, and AHFM Peters; PRC1 and Suv39h specify parental asymmetry at constitutive heterochromatin in early mouse embryos; *Nat Genet* 40, 411-20 (2008).

Pawlowski J, Fahrni JF, **Brykczynska U**, Habura A, Bowser SS; Molecular data reveal high taxonomic diversity of allogromiid foraminifers in Explorers Cove (McMurdo Sound, Antarctica); *Polar Biology* 25: 96-105 (2002).

TALKS AND POSTER PRESENTATIONS

2009 TriRhena Chromatin & Transcription Club Meeting, Freiburg, Germany, **talk**

2009 Epigenetics, Development and Human Disease, Breckenridge, USA, **talk**

2008 FMI annual retreat and scientific evaluation, Grindelwald, Switzerland, **talk**

2007 Chromatin and Epigenetics, EMBO conference, Heidelberg, Germany, poster

2006 Novartis Biomedical Research Conference, Venice, Italy, poster

2006 Epigenetics and Chromatin Remodeling in Development, Keystone, USA, poster

TEACHING EXPERIENCE

2003 - 2004 **Teaching assistant**, Workshops in Molecular Biology for high school students, Science Festival School, Warsaw, Poland

MISCELLANEOUS

- Languages: Polish: native, English: fluent, French: fluent, German: basic
- Student representative at the FMI
- Sports and activities: mountaineering, skiing, running