THE UNIVERSITY
of EDINBURGH

Regulators of DNA methylation in mammalian cells

Ausma Termanis

Thesis presented for the degree of Doctor of Philosophy

University of Edinburgh

2012

Declaration

I declare that this thesis was composed by myself. The research in this thesis is my own, unless otherwise stated and has not been submitted for any other degree or professional qualification

Ausma Termanis (August 2012)

Table of Contents

Index of Figures

Acknowledgements

First of all, I would like to thank Irina for giving me the opportunity to carry out my PhD research in her lab. I would like to thank her for advice throughout my project, for numerous interesting discussions and for reading the draft of my thesis. Additionally I would like to thank current and past members of the Stancheva lab (Kat, Joe, Chao, Tuo, Sadie, Thomas, Jose, Kevin, Naza, Tristin, Cara, Matt, Andreas) for helpful discussions and encouragement.

Next I would like to thank current and past members of the Bird lab. Especially to Rob Illingworth for help with MBD columns and Sabine for help with southern blots and helpful discussions on experiments.

I would also like to thank Neil Brockdorff' and members of his lab in Oxford for an enjoyable collaboration. Additionally I would like to thank Bernard Ramsahoye at the Edinburgh University Cancer Research UK centre for HPLC analysis of samples. I would like to thank Kathrin Muegge and Yoichi Shinkai for providing me with cell lines. Also I thank the Protein Purification Facility at the University of Edinburgh for help with MBD columns and protein purifications.

Finally I would like to thank Alex for our animated scientific discussions and for much needed time away in between.

I would like to thank Cancer Research UK, who funded my work during the last four year.

Abstract

Although the many cells within a mammal share the same DNA sequence, their gene expression programmes are highly heterogeneous, and their functions correspondingly diverse. This heterogeneity within an isogenic population of cells arises in part from the ability of each cell to respond to its immediate surroundings via a network of signalling pathways. However, this is not sufficient to explain many of the transcriptional and functional differences between cells, particularly those that are more stable, or, indeed, differences in expression between parental alleles within the same cell. This conundrum lead to the emergence of the field of epigenetics - the study of heritable changes in gene expression independent of DNA sequence. Such changes are dependent on "epigenetic modifications", of which DNA methylation is one of the best characterised, and is associated with gene silencing. The establishment of correct DNA methylation patterns is particularly important during early development, leading to cell type specific and parental allele specific gene regulation. Besides DNA methyltransferases, various other proteins have recently been implicated in DNA methylation. The absence of these proteins leads to defects in DNA methylation and development that can be even more severe than those in DNA methyltransferase knockouts themselves. In this study I focus on three such accessory proteins: LSH (a putative chromatin remodelling ATPase), G9a (a histone lysine methyltransferase) and SmcHD1 (a structural maintenance of chromosomes protein). To compare DNA methylation between WT cells and cells knocked out for each of these proteins, I used whole genome methylated DNA affinity purification and subsequent hybridization to promoter microarrays. This enabled me to compare the requirement for each protein in DNA methylation at specific genomic regions.

The absence of LSH in mouse embryonic fibroblasts (MEFs) resulted in the loss of DNA methylation at 20% of usually methylated promoters, and the misregulation of associated protein coding genes. This revealed a requirement for LSH in the establishment of DNA methylation at promoters normally methylated during pre-implantation as well as post-implantation development.

Secondly, I identified hypomethylation at 26% of normally methylated promoters in *G9a-/-* compared to WT ES cells. Strikingly, this revealed that G9a is required for maintenance of DNA methylation at maternal as well as paternal imprinting control regions (ICRs). This is accompanied by expression defects of imprinted genes regulated by these ICRs.

Finally, in collaboration with the Brockdorff lab at the University of Oxford I identified a role for SmcHD1 in establishing DNA methylation at promoters on the X chromosome normally methylated slowly during X chromosome inactivation. Interestingly, SmcHD1 was also required for DNA methylation at autosomal gene promoters, contrary to previous reports that it is mainly involved in X chromosome methylation.

I conclude that different accessory proteins are required to facilitate correct DNA methylation and gene repression at distinct regions of the genome, as well as at different times during development. This function of accessory proteins may be in part dependent on the prior establishment of specific chromatin signatures and developmental signals, together comprising a precisely regulated system to establish and maintain appropriate DNA methylation throughout development.

Chapter 1 - Introduction

1.1 Epigenetics

Since the completion of various genome projects, a wealth of insight has been gained into the evolutionary origins of different organisms as well as their functional divergence. In the case of unicellular organisms this enables extensive genome function characterization throughout their lifecycle. Multicellular organisms provide us with yet widely unanswered questions, despite the availability of genome sequences. Multicellular organisms arise from a series of lineage commitments commencing in the first cell divisions upon fertilization of a single cell. Lineage commitment results in subsets of cells with different functions, despite containing largely the same DNA sequence. Most functional differences between cells occur through the presence of pools of different proteins and non-coding RNAs within the cells. Therefore the question arises: How can cells regulate which products of the genome they contain?

In some early lineage commitment processes it has become apparent, that cells with different cytoplasmic properties can arise through accumulation of different proteins at each cell pole, followed by cell divisions. This leads to temporary differences in protein composition (Plusa et al, 2005; Smith et al, 2006; Xin et al, 2003). However, most differences in protein and RNA composition of cells originate from differential regulation of gene expression. Some differences can arise through topological properties of the early embryo: specific cell-cell contacts have been found to initiate signalling pathways, leading to differential activation or repression of genes. The Hippo signalling pathway in the mouse for example is

initiated in cells in the centre of the developing embryo destined to become part of embryonic tissues, whereas cells on the periphery, destined to become part of the extraembryonic tissues, do not undergo this activation. Activation of the Hippo pathway enables the continued repression of trophectoderm markers such as Cdx2, in cells destined to become part of the embryonic tissues (Avner & Heard, 2001; Heard et al, 2001).

The wealth of cell types with different gene expression profiles suggests the presence of a more complex gene regulatory system. Indeed throughout the years, other gene expression phenomena have become apparent which challenged our understanding of gene regulation. These involve heritable differences in regulation of genomic regions with the same DNA sequence in different cells or even in the same cell. This is now known as epigenetics ('above' genetics). Major phenomena highlighting the importance of epigenetic regulation mechanisms of genome function are genomic imprinting and X chromosome inactivation.

1.1.1 Genomic Imprinting

Genomic imprinting is the parental allele specific differential expression of genes. Both alleles have the same DNA sequence but are differentially expressed within the same cell. Initial evidence for this phenomenon came from studies of embryos derived from only maternal or paternal chromosomes, which failed to develop normally, displaying placental and embryonic defects (McGrath & Solter, 1984; Surani et al, 1984). Several hypotheses have been proposed to explain this phenomenon, but a parental conflict model is the most developed and has received most support (Haig & Graham, 1991; Morton et al, 2011; Myant & Sansom, 2011). This model hypothesizes, that the maternal genome strives to reserve resources for

future offspring, whereas the paternal genome strives to provide his current offspring with as many resources as required. Evidence for the existence of such a conflict between parental genomes comes from studies into functions of imprinted genes.

Firstly, many imprinted gene functions support this model, with paternally expressed genes supporting growth and maternal genes expressing growth inhibitors. An example occurs in the *Igf2* gene cluster, which encodes a paternally expressed growth factor (IGF2) and a maternally expressed growth suppressor (IGF2R), which targets IGF2 for degradation (Dean et al, 1998; Myant et al, 2011; Zvetkova et al, 2005). Interestingly, all imprinted genes analysed are expressed in placental tissue (Reik et al, 2003). The placenta is important in regulation of embryonic growth and development. Imprinted genes in the placenta therefore regulate the supply of nutrients to the developing embryo. In foetal tissues on the other hand they regulate the demand for resources through regulation of growth.

Secondly, a role of imprinted genes in metabolic processes has also been found. The *Gnas* imprinted locus is a good example of this function. Defects in the maternal allele of the *Gnas* locus lead to obesity, whereas paternal defects lead to lean phenotypes, supporting the parent conflict model outlined above (Blewitt et al, 2008; Gibbons et al, 2000; Yu et al, 2001). The *Igf2* locus has also been shown to function in metabolism, with paternal disruption of *Igf2* causing weight reduction (Smith et al, 2006).

Finally, studies have highlighted requirement of imprinted genes for correct development of behavioural traits. Female mice inheriting a paternal mutation of the *Peg3* imprinted gene display altered behaviour towards her litter (making nest,

gathering pups, keeping them warm), frequently resulting in death of offspring (Li et al, 1999). Additionally to these observations, many well studied imprinting disorders are accompanied by neurological and psychiatric impairments (e.g. Prader-Willi Syndrome and Angelman Syndrome) (Ohta et al, 1999).

The above analysis of genomic imprinting highlights the importance of this phenomenon in a plethora of functions during early development as well as in adult life.

1.1.2 X chromosome inactivation

X inactivation is another phenomenon which involves non DNA sequence dependent distinction between chromosomes. This process ensures equal dosage of X linked gene expression between XY males and XX females and is achieved by inactivation of most genes one X chromosome in female mammals, either on the paternal chromosome in extraembryonic tissues or at random in the embryos (Avner & Heard, 2001).

The above examples epitomize epigenetics. Epigenetic mechanisms utilize chemical modifications of DNA, as well as associated proteins and structural changes to DNA/protein complexes to cause alterations in gene expression or other functions. DNA methylation is the most well studied epigenetic mark.

1.2 DNA methylation

1.2.1 Origins of DNA methylation

DNA methylation originates from a genomic immune system to silence invading transposable elements in the bacterial genome (Lee et al, 1999). This use of DNA methylation to protect the genome from genomic instability has passed on to fungi and plants during evolution (Zemach & Zilberman, 2010). Most fungal methylation is to date still located within transposable elements (Zemach & Zilberman, 2010). Transposable element methylation was lost and regained independently during evolution of vertebrates (Zemach & Zilberman, 2010). During the evolution of animals and plants, gene body methylation arose, with vertebrates and seeded plants showing most genome wide DNA methylation. Methylation is seen at promoters of repressed genes and often within exons of transcribed genes. In invertebrates on the other hand there is little evidence for the silencing function of DNA methylation (Zemach & Zilberman, 2010).

During evolution, DNA methylation enzyme gene families also became extinct at various stages, mainly remaining in vertebrates and plants only (Zemach & Zilberman, 2010). Why has this divergence occurred in the use of DNA methylation within eukaryotic genomes? One factor could be that multicellular organisms require more regulatory mechanisms to generate a diverse range of tissues and cell types. Another idea is that DNA methylation may have evolved alongside the evolution of sexual reproduction. Sexual reproduction brings about the continued requirement for DNA methylation to silence transposable elements, as these can be passed on through generations even without selective advantage. Therefore DNA methylation

may have been maintained in these organisms, allowing evolution of new functions (Zemach & Zilberman, 2010).

1.2.2 DNA methylation

In vertebrates the most prevalent type of DNA methylation is CpG methylation, although non-CpG methylation has been demonstrated to occur widely in ES cells and tissue such as human spleen (Ramsahoye et al, 2000). Non-CpG methylation in mammals has not been well characterized yet, but similarly to CpG methylation it has been proposed to function in gene silencing (Franchina & Kay, 2000; Malone et al, 2001). CpG methylation involves the covalent attachment of a methyl group to the carbon five position of cytosine (Bird, 1986). Strand symmetry of CpG methylation allows it to be faithfully passed on during cell division due to the semi-conservative nature of DNA replication (Bird, 2002). Most CpGs in the genome (60-90% of CpGs) are methylated. This has caused the genome to be relatively depleted of CpG dinucleotides due to the frequent spontaneous deamination of 5 methyl cytosines into thymidines (Bird, 1980; Coulondre & Miller, 1978; Coulondre et al, 1978). Exception to the highly methylated regions of the genome are regions called CpG islands (CGIs), which are usually unmethylated (Figure 1.1A). These regions display relatively high GC content as well as less CpG depletion, probably due to the lack of DNA methylation(Bird et al, 1985; Cross et al, 1994). Approximately 60% of annotated gene promoters are CpG islands, which points to their function in gene regulation (Larsen et al, 1992; Saxonov et al, 2006; Zhu et al, 2008). DNA methylation of CGIs leads to stable silencing of associated promoters and is implicated in developmental gene regulation (Figure 1.1A) (Deaton & Bird, 2011).

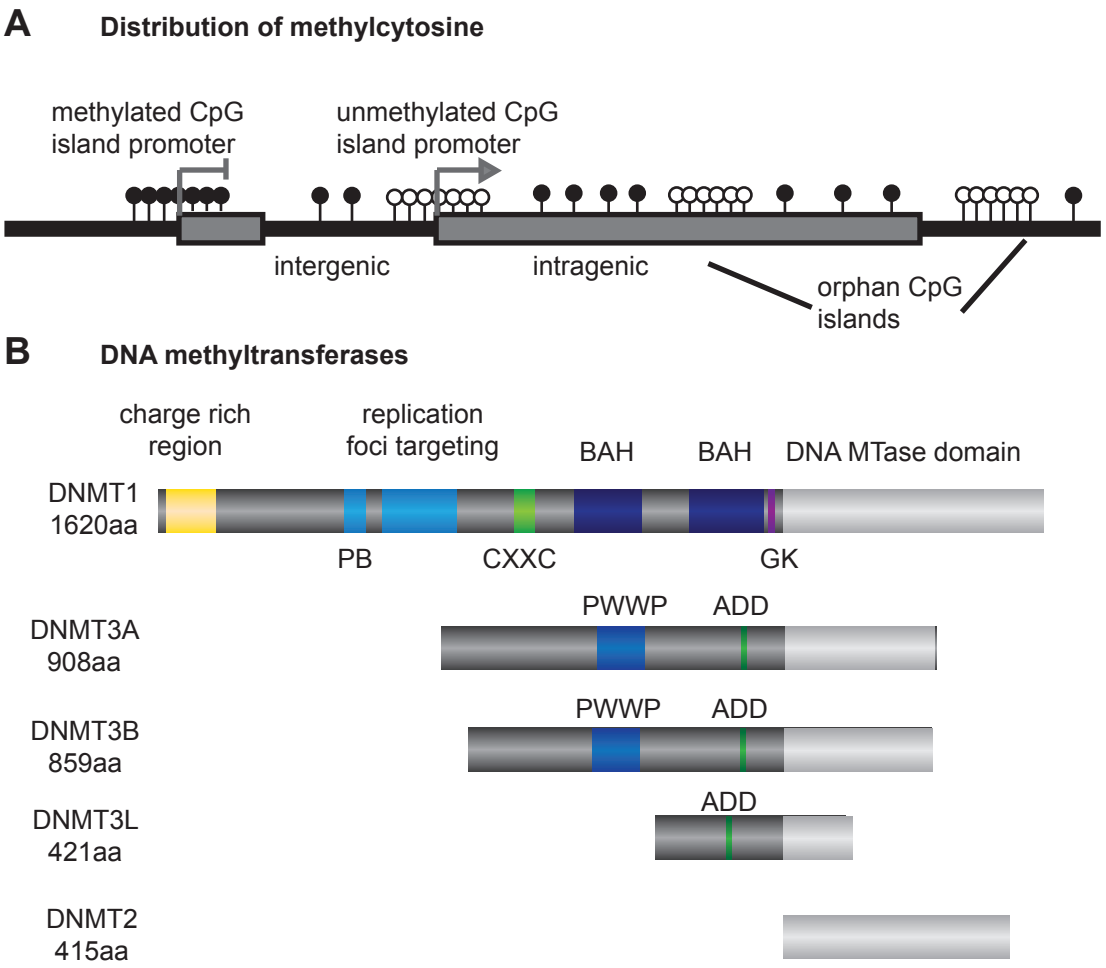


Figure 1.1 DNA methylation

A: Distribution of 5meC. Methylated and unmethylated CpG dinucleotides are indicated by black and white lollipops respectively. Most of the genome is methylated at CG dinucleotides (intergenic and intragenic CG dinucleotides). This has resulted in depleted levels of CG dinucleotides due to frequent spontaneous deamination of 5meC to thymidines. Exception to this are CpG islands (CpG rich regions) which are usually located at promoter regions. CpG islands are usually unmethylated, unless they are involved in developmental regulation of gene activity. CpG islands have been recently found to occur outwith annotated gene promoters (orphan CpG islands). These could be indicators for novel promoters. **B: Murine DNA methyltransferases.** Domain annotation (to scale) of maintenance DNA methyltransferase DNMT1, de novo DNA methyltransferasese DNMT3A and DNMT3B, catalytically inactive family members DNMT3L and DNMT2. PB: PCNA binding domain; CXXC: DNA binding domain; BAH: bromo-adjacent homology domain required for protein-protein interactions; GK: Glycine-Lysine repeats may be required for targeting of DNMT1 to DNA close to replication forks; PWWP: required for chromatin targeting; ADD: ATRX-DNMT3-DNMT3L cysteine rich domain. The ADD domain can bind zinc ions and is required of protein-protein interactions.

CpG islands which do not overlap with annotated gene promoters have been termed orphan CpG islands (Figure 1.1A). A lot of these orphan CpG islands have now been shown (through co-localization with RNA Polymerase II and RNA sequencing data) to be unidentified transcription initiation sites (Illingworth et al, 2010). Substantiating the presence of orphan CGIs at promoters, is the presence of bound RNA polymerase II (RNA PolII) as well as RNAs initiating at these orphan CGIs (Guttman et al, 2009; Illingworth et al, 2010).

In addition to CGI promoters, 40-50% of all promoters do not overlap with CpG islands. This class of promoters has been less studied but recent work has confirmed similar effect of DNA methylation at non-CGI promoters as at CGI promoters (Han et al, 2011).

Gene bodies, unlike promoter regions often display high levels of DNA methylation, especially within highly transcribed regions (Figure 1.1A). This methylation may function to prevent aberrant transcription initiation within the gene body or aid in mRNA splicing (Figure 1.1A) (Laurent et al, 2010).

1.2.3 DNA methyltransferases

Members of the DNA methyltransferase enzyme (DNMT) family establish and maintain DNA methylation. The first member of the family identified was DNMT1, closely followed by DNMT2, DNMT3A, DNMT3B and DNMT3L (Bestor et al, 1988; Gruenbaum et al, 1982; Okano et al, 1998b; Van den Wyngaert et al, 1998; Yoder & Bestor, 1998) (Figure 1.1B). With the exception of DNMT3L, all these enzymes, contain C-terminal conserved catalytic domains which have been shown to be essential for their activity as methyltransferases (Cheng, 1995; Cheng &

Blumenthal, 2008; Jeltsch, 2002; Jurkowska et al, 2011). The catalytic domain contains six conserved motifs with I and X involved in cofactor binding and IV, VI, VIII in catalytic activity. A non-conserved region in the middle of the catalytic motifs is involved in ensuring target specificity (Cheng, 1995; Cheng & Blumenthal, 2008; Jeltsch, 2002).The N terminal regions of DNMTs are responsible for protein-protein interactions, targeting and regulation of these enzymes (Jurkowska et al, 2011).

Maintenance DNA methyltransferase - DNMT1

The first identified DNA methyltransferase preferentially binds hemimethylated DNA and shows reduced specificity to unmethylated DNA (Fatemi et al, 2001). It has been suggested that unmethylated DNA binding to the N terminal portion of DNMT1 may lead to an allosteric inhibitory effect on enzymatic activity (Bacolla et al, 1999). Interaction between the N- and C-termini of the enzyme leads to allosteric activation (Fatemi et al, 2001).

The major function of DNMT1 is in maintaining DNA methylation during cell division, but involvement in *de novo* methylation has also been demonstrated in conjunction with *de novo* methyltransferase DNMT3A (Bestor et al, 1988; Fatemi et al, 2002; Gowher et al, 2005b; Pradhan et al, 1999).Consistent with its role in maintenance methylation DNMT1 localizes diffusely during most of the cell cycle, but to replication foci in S phase and interacts with Proliferating Cell Nuclear Antigen (PCNA) at replication forks (Iida et al, 2002; Leonhardt et al, 1992; Margot et al, 2001). In addition to this, it localizes to centromeric heterochromatin, even after S phase, indicating a replication independent role of DNMT1 (Easwaran et al, 2004).

Expression patterns of DNMT1 reflect its DNA maintenance methylation role. It is highly expressed in proliferating cells and tissues (Robertson et al, 1999) and expression levels of Dnmt1 reach maximum levels during S phase of the cell cycle (Robertson et al, 2000; Szyf, 2001). Additionally, different isoforms of DNMT1 exist and are tissue specifically expressed. DNMT1O is the more stable oocyte specific form of the enzyme (Ding & Chaillet, 2002). Both maternal and zygotic DNMT1 seem to be required for imprinting maintenance in the embryo (Hirasawa et al, 2008).

Disruption of DNMT1 leads to gradual demethylation of the genome, loss of imprinting, defects in X chromosome inactivation and early embryonic lethality (Beard et al, 1995; Howell et al, 2001; Li et al, 1993; Li et al, 1992). Additionally, DNMT1 has been linked to tumorigenesis, playing a role in proliferation and survival of cancer cells (Chen et al, 2007). ES cells on the other hand are viable but die upon differentiation (Li et al, 1992; Tsumura et al, 2006). This may point to a mechanism in ES cells to maintain stable heterochromatin in absence of DNA methylation.

De novo DNA methyltransferases DNMT3A/3B

DNMT3A and DNMT3B are similar in structure to each other, with DNMT3B having a shorter N terminal domain (Figure 1.1B). DNMT3A and DNMT3B are the major *de novo* DNA methyltransferases but have also been implicated in maintenance methylation of repetitive elements in co-operation with DNMT1 (Liang et al, 2002). Consistently, unlike DNMT1, neither of these enzymes show preference for unmethylated or hemimethylated DNA (Okano et al, 1998a).

Initially thought to be redundant in function, DNMT3A and DNMT3B also have non overlapping functions. These are illustrated by differences in phenotypic effects of knockout in embryos and ES cells (Okano et al, 1999). DNMT3A knockout embryos develop to term, but die shortly after birth. DNMT3B knockout embryos on the other hand die at embryonic day 9.5 (E9.5). Knockout ES cells are viable, suggesting DNA methylation is not required for self-renewal properties of ES cells (Tsumura et al, 2006). Additionally, DNMT3A but not DNMT3B knockout shows germ line specific defects in imprinting in the male and female germ line (Kaneda et al, 2004).

Expression patterns of DNMT3A and DNMT3B enzymes reflect their roles in *de novo* methylation. They are both expressed highly in embryos and undifferentiated ES cells, at which time a lot of *de novo* DNA methylation events occur, and less expressed in somatic cells (Okano et al, 1998a). Similarly to DNMT1, different isoforms of DNMT3A and DNMT3B exist. Especially in the case of DNMT3B the full length enzyme (DNMT3B1) is expressed at higher levels in ES cells whereas isoforms lacking several exons are expressed in somatic cells (Okano et al, 1998a). These exons are suggested to be required for correct embryonic development. In addition to their function in cytosine methylation within the CpG context, both enzymes have recently been implicated in non CpG methylation (Gowher & Jeltsch, 2001; Ramsahoye et al, 2000). Consistent with their high expression in ES cells, non-CG methylation is most prevalent in these (Ramsahoye et al, 2000).

Non-canonical DNA methyltransferases

Two structurally and functionally non-canonical members of the DNA methyltransferase family are DNMT2 and DNMT3L. DNMT2 is only comprised of the catalytic domains characteristic of DNMT family members (Figure 1.1B). It is the only mammalian DNA methyltransferase which does not have an N terminal domain involved in targeting and regulation of the enzyme. Transfer RNAs (tRNAs) have been identified as DNMT2 targets, but only weak methylation activity towards DNA has been shown (Hermann et al, 2003; Jurkowski et al, 2008; Tang et al, 2003)

In contrast to DNMT2, DNMT3L has an N-terminal as well as a C-terminal domain (Figure 1.1B). It is catalytically inactive, functioning as a regulator of DNA methylation, predominantly in germ cells (Hata et al, 2002). DNMT3L has been shown to directly interact with DNMT3A and DNMT3B catalytic domains and stimulate their activity (Chedin et al, 2002; Gowher et al, 2005a). Consistently DNMT3L localization is dependent on DNMT3A and DNMT3B (Nimura et al, 2006). DNMT3L functions specifically together with DNMT3A in ensuring correct DNA methylation at imprinted regions (Hata et al, 2002). *Dnmt3l* Knockout mice are viable but males are sterile due to defects in meiosis leading to loss of germ cells and females fail to produce viable offspring (Bourc'his & Bestor, 2004).

1.2.4 Effectors of DNA methylation

DNA methylation can cause gene repression in two ways: Firstly by directly preventing transcription factor binding and secondly through recruitment of effector proteins. Effector binding can lead to recruitment of large co-repressor complexes to aid the establishment of a state repressive to gene transcription and other functions. Three families of methylated DNA binding proteins have been identified. These include MBD proteins (MBD1, MBD2, MBD3, MBD4 and MeCP2), Kaiso family proteins (Kaiso, ZBTB4, ZBTB38) and SRA family proteins (UHRF1 and UHRF2) (Figure 1.2).

MBD family

All MBD family proteins have a conserved methyl CpG binding domain (MBD). MBD3 has a mutation in this domain, rendering it unable to bind methylated DNA (Hendrich & Bird, 1998). MeCP2, MBD1 and MBD2 also have transcription repression domains (TRD). Apart from these conserved domains, each family member has distinguishing additional domains. MBD1 for example contains up to 3 CxxC domains depending on the splice variant. CxxC domains have been previously characterised to bind unmethylated DNA but only the most C terminal domain in MBD1 shows this binding affinity (Jorgensen et al, 2004). MBD2 and MBD3 contain coil-coil motifs involved in protein-protein interactions. MBD4 has a more distinctive glycosylase domain involved in nucleotide excision repair, especially directed at repairing spontaneous deamination products of 5meC (Hendrich et al, 1999). Consistently, MBD4 deficient mice show increased levels of 5meC deamination products throughout the genome (Hendrich et al, 1999; Millar et al, 2002; Wong et al, 2002).

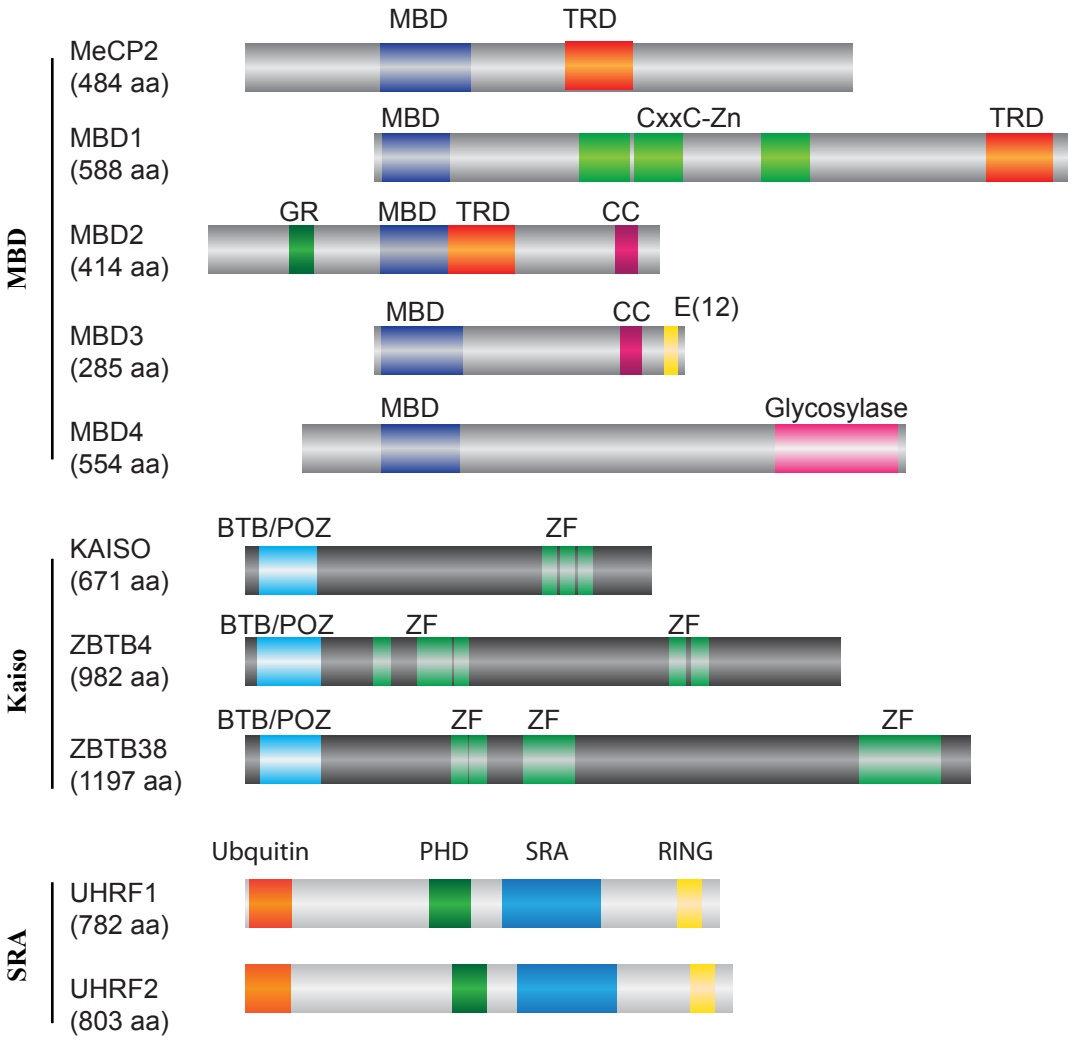


Figure 1.2 Methylated DNA binding proteins

Domain annotation of murine methylated DNA binding proteins (drawn to scale within families). MBD, Kaiso and SRA family proteins are shown. Different domains are indicated. MBD: methylated DNA binding domain; TRD: transcription repression domain; CxxC-Zn: DNA binding domain (third CxxC domain in MBD1 allows binding to unmethylated DNA); GR: Glycine-Lysine repeats; CC: coil-coil domain; E(12): Glutamic acid repeats; Glycosylase: Domain required for base excision DNA repair; BTB/POZ: binds methylated DNA; ZF: zinc finger domains required for methylated DNA and sequence specific DNA binding; PHD: Plant Homeo Domain involved in binding to unmodified and modified histone tails; SRA: Set and Ring associated domain required for binding to 5meC; RING: Really Interesting New Gene domain with ubiquitin ligase activity; Ubiquitin: Ubiquitin-like domain.

MeCP2 is the most studied methyl binding protein due to mutation in its MBD and TRD domains leading to a neurological disorder called Rett syndrome (Amir et al, 1999; Cheadle et al, 2000). Its most studied function is transcriptional repression consistent with binding to methylated DNA and recruitment of co-repressor complexes. Despite this, MeCP2 null mice with Rett syndrome only display small changes in gene expression (Tudor et al, 2002). This gave rise to a series of studies investigating potential other roles. MeCP2 to date has for example been found to associate with splicing factors and also localizes to many actively transcribed genes as well as interacts with transcriptional activator CREB1 (Chahrour et al, 2008; Yasui et al, 2007; Young et al, 2005). Additionally, a chromatin architectural role of MeCP2 has also been proposed, independent of methylated DNA binding (Georgel et al, 2003). MBD1 and MBD4 on the other hand both bind to factors involved in DNA repair (Bellacosa et al, 1999; Watanabe et al, 2003).

Kaiso family

KAISO family proteins all share a conserved triple zinc finger domain involved in methylated DNA and sequence specific DNA binding (Sasai et al, 2010). ZBTB4 and ZBTB38 contain extra zinc finger motifs (Filion et al, 2006). The second characteristic domain present in KAISO family members is the POZ/BTB domain, implicated in protein-protein interaction and transcriptional repression (Perez-Torrado et al, 2006).

ZBTB4 and ZBTB38 are seen to localize to dense heterochromatin in a DNA methylation dependent manner, supporting their transcription repression role in reporter assays (Sasai & Defossez, 2009). Apart from binding co-repressors, KAISO

has also been found to localize to centromeres and spindle microtubules pointing towards a new role in mitosis (Soubry et al, 2010).

SRA family

SRA family members contain SRA domains implicated in hemi-methylated 5meC binding, Tandem Tudor and PHD finger domains which bind modified histones and a RING ubiquitin ligase domain (Bostick et al, 2007; Johnson et al, 2007; Karagianni et al, 2008; Nady et al, 2011; Wang et al, 2011; Xie et al, 2012).

UHRF1 is the most characterized SRA protein family member. UHRF1 accumulates at heterochromatic loci in a hemimethylated DNA dependent manner. Consistent with its binding to hemimethylated DNA, UHRF1 is found to interact with PCNA, DNMT1 and G9A (Bostick et al, 2007; Karagianni et al, 2008; Sharif et al, 2007; Uemura et al, 2000). The suggested mechanism of function is not recruitment of DNMT1, but facilitation of maintenance methylation by DNMT1. It has been proposed that UHRF1 may 'flip' out the unmethylated cytosine during replication making it more accessible to DNMT1 (Arita et al, 2008; Hashimoto et al, 2008).

1.2.5 Developmental dynamics of DNA methylation

The embryonic lethal phenotypes of DNMT knockouts discussed above point to a developmentally important role of DNA methylation. DNA methylation has been found to be implicated in early cell lineage specifications, imprinting and X chromosome inactivation in early embryos. There are two stages during mammalian development at which DNA methylation reprogramming facilitates establishment of these processes (Figure 1.3) (Feng et al, 2010; Saitou et al, 2012).

The first DNA methylation reprogramming event occurs in the pre-implantation embryo (Figure 1.3). Between the one cell and the blastocyst stage the genome undergoes major DNA demethylation. The paternal genome is more highly methylated. A recent study identified the oocyte genome to already be relatively hypomethylated and more similar to early embryonic stages than the sperm genome (Smith et al, 2012). It is thought that the paternal genome may undergo active DNA demethylation before the first S phase followed by passive demethylation of the maternal genome. Consistently, the transition between the sperm genome and the blastocyst shows most differences in DNA methylation whereas the oocyte genome shows fewer changes in methylation during reprogramming.

Some regions have been shown to resist DNA demethylation. These include imprinted differentially methylated regions (DMRs) which have to be stably maintained throughout development after their establishment in the germ line. A recent study expanded on the idea that imprinted DMRs resist DNA demethylation and showed that boundary regions of DMRs do undergo demethylation by the blastocyst stage (Tomizawa et al, 2011).

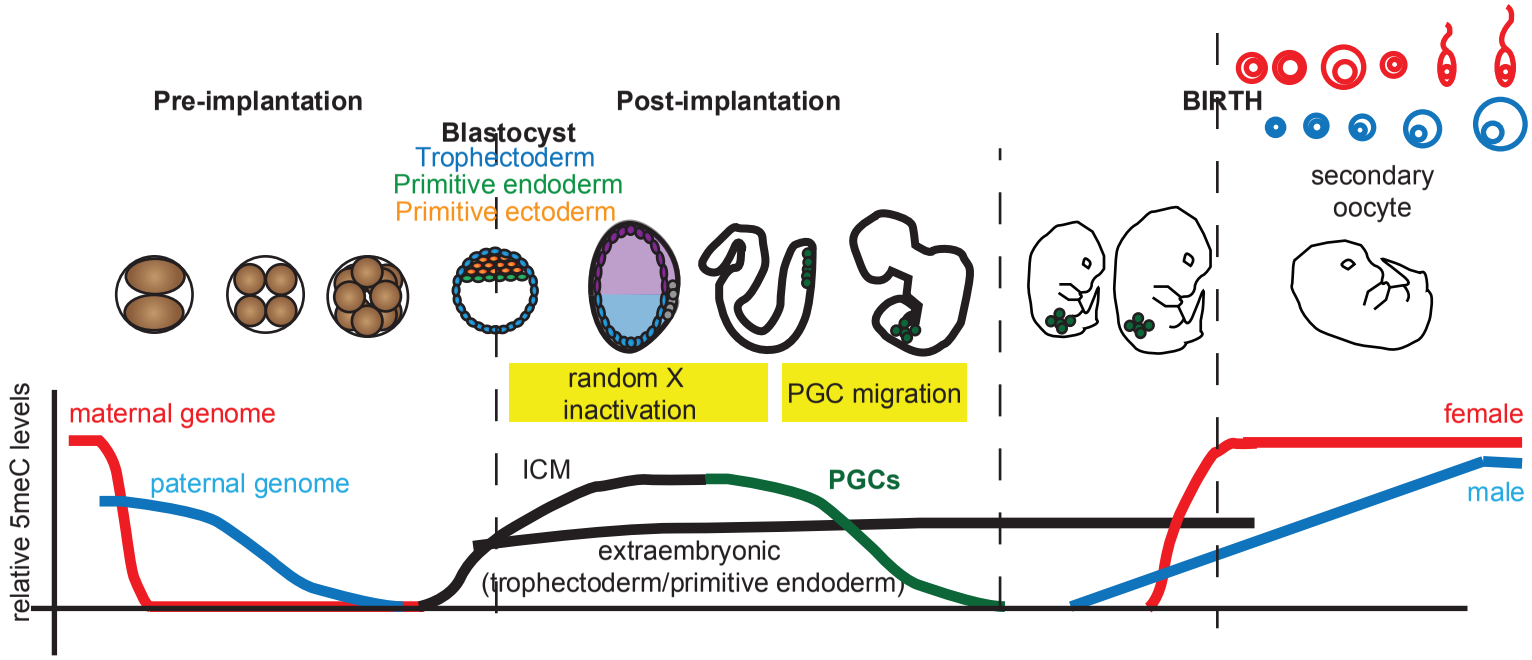


Figure 1.3 DNA methylation dynamics during mammalian development

Top: Displayed are different time points (from fertilization until oocyte maturation after birth) during early mammalian development. **Bottom:** Levels of 5meC during the different stages of development. Male and female DNA methylation are indicated in red and blue respectively. During pre-implantation development DNA methylation is lost from both the maternal and paternal genome. The paternal pronucleus undergoes active DNA demethylation first, followed by passive demethylation of the maternal genome. DNA methylation reaches its low point at the morula stage and is followed by DNA methylation establishment. A second DNA methylation reprograming event occurs in the developing germline. DNA methylation is erased during migration of primordial germ cells to the genital ridge. DNA methylation is subsequently re-established specific to the male and female germline. In males this remethylation occurs before birth, whereas in females remethylation is slower and is completed by the time the secondary oocyte has formed.

Additional germ line specific methylated regions have recently been identified to partially resist DNA methylation (Kobayashi et al, 2012; Smith et al, 2012). Especially oocyte specific methylated CGIs have been identified, which partially resist DNA demethylation (Smallwood et al, 2011). Moreover repetitive elements have also for a long time been thought to remain methylated at this stage, to ensure genome stability, but recent studies point towards repeat elements such as long interspersed repeats (LINEs) and short interspersed repeats (SINEs) losing DNA methylation (Kobayashi et al, 2012; Smith et al, 2012). The above described recent finding give rise to additional complexity as to how DNA methylation and de methylation occur at certain regions.

After genome wide DNA demethylation has reached its low point in the blastocyst, DNA re-methylation events occur to facilitate various processes (Figure 1.3). Random X chromosome inactivation occurs upon differentiation of the inner cell mass (ICM). This is accompanied by loss of pluripotency, methylation and silencing of germ line specific and other tissue specific genes. An indication that this may in turn allow differentiation between different lineages comes from evidence that *de novo* methylation occurs in the ICM, but not cells destined to become part of the extraembryonic tissues (Figure 1.3). Additionally it has been found that the majority of genes methylated are germ line specific genes, leading to their transcriptional silencing (Borgel et al, 2010; Isagawa et al, 2011). This may point to divergence at this stage from a germ line methylation landscape. This is supported by a recent study looking at methylation patterns during differentiation of embryonic stem cells into the three early germ layers (Isagawa et al, 2011). All germ layers showed widespread methylation and silencing of germ line specific genes.

Differences in DNA methylation between germ layers on the other hand did not correspond to expression changes. This is an indication that further lineage differentiation may require other elements apart from DNA methylation.

The second reprograming event occurs during germ line specification (Figure 1.3). The blastocyst in the early embryo consists of the trophectoderm, primitive ectoderm and primitive endoderm (Rossant & Tam, 2009). The primitive ectoderm, gives rise to the epiblast whose cells differentiate into somatic cells as well as primordial germ cells (Rossant & Tam, 2009). Primordial germ cells are specified in the post implantation epiblast. At this stage methylation at imprinted regions and repetitive sequences is present and X inactivation has occurred (Saitou et al, 2012). During migration of germ cells to the genital ridge, reprogramming of germ cells occurs. This involves reactivation of the inactive X chromosome, imprinted region and repeat element DNA methylation erasure, re-establishment of pluripotency and subsequent specification of male and female germ cells (Saitou et al, 2012). DNA methylation after this stage ensures correct parental specific DNA methylation at imprinted regions.

1.3 DNA packaging into chromatin

Up to this point I have discussed DNA methylation in context of naked DNA, rather than regarding DNA as part of a more complex environment. In mammalian cells approximately 3 billion base pairs of DNA have to be packaged into the 5-10 μm nuclear compartment. To achieve this, DNA undergoes several levels of condensation (Figure 1.4).

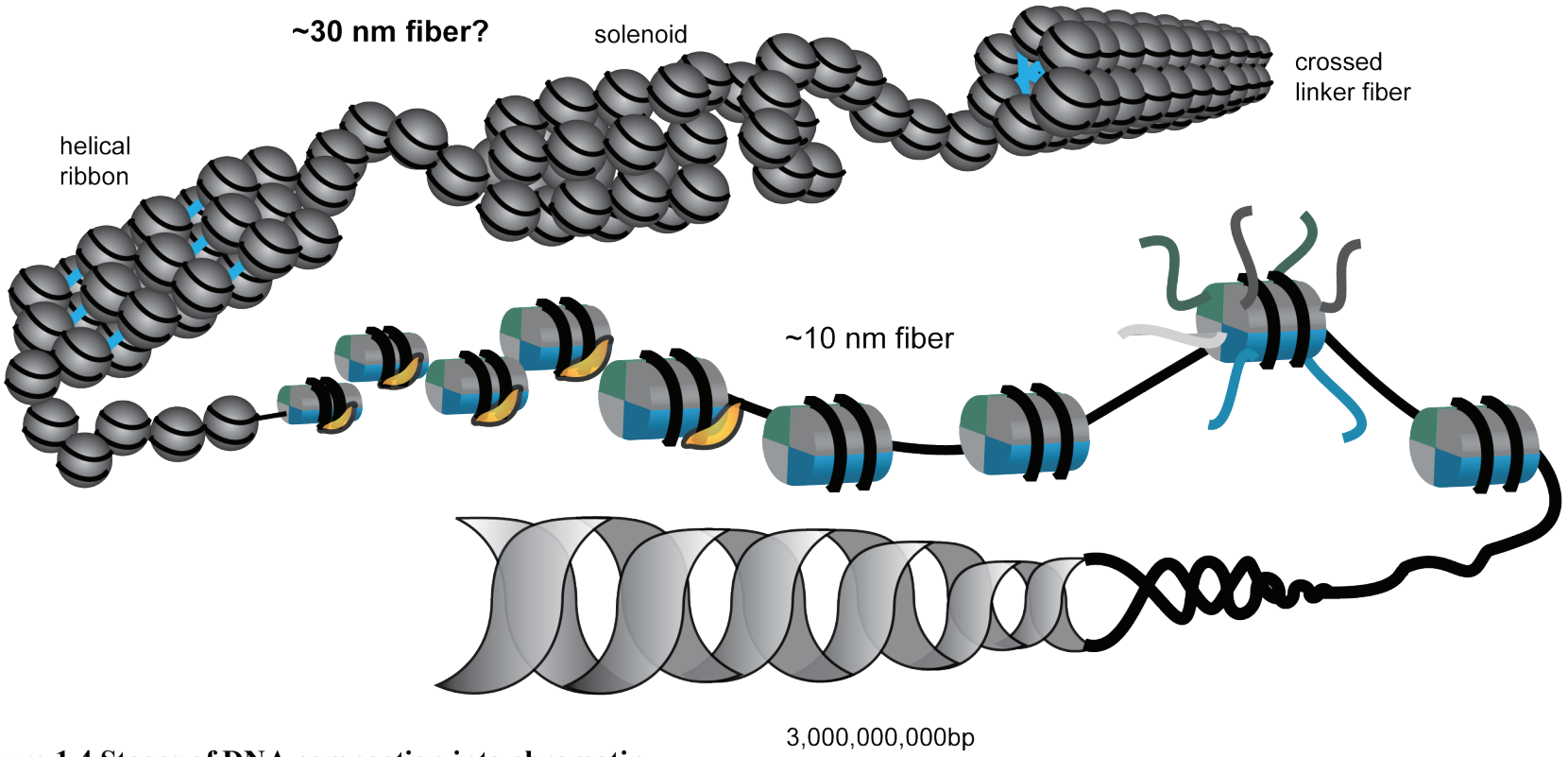


Figure 1.4 Stages of DNA compaction into chromatin

The DNA double helix is displayed , winding 1.75 times around a core of histone octamers to form a nucleosome. Histones are globular apart from their protruding N-terminal tails. Linker histone (yellow) binding to linker DNA and the core histones facilitates further winding of DNA around the histone core. Linker histone is also thought to promote further compaction of nucleosomal DNA. Three different models of further compaction are indicated, resulting in a 30nm chromatin fiber (helical ribbon, solenoid, crossed linker fiber). Recent studies suggest a combination of 30nm and 10nm fibers may comprise the higher order chromatin structure.

The first level involves wrapping of the DNA double helix 1.75 times around a core histone protein octamer to form a nucleosome (Kornberg, 1974; Woodcock et al, 1976). The histone octamer consists of one H3/H4 tetramer and two H2A/H2B dimers (Luger et al, 1997). Core histones are positively charged and globular in structure apart from their protruding N-terminal tails, facilitating the wrapping around of DNA, resulting in up to 10 fold compaction. Like beads on a string nucleosomes are spaced by linker DNA which can be bound by linker histone H1. Unlike the globular core histones, histone H1 has a more extended structure which allows it to bind to both the linker DNA as well as the core histones. Binding of H1 results in final wrapping of DNA twice around the histone octamer (Bednar et al, 1998; Caterino & Hayes, 2011) (Figure 1.4).

Histone H1 is also thought to promote the next level of compaction from a 10nm nucleosomal array into more highly compacted fibres. The exact way DNA is compacted after the 10nm stage is still disputed. For a long time a few models of so called 30nm fibres were pursued. These included the solenoid model, as well as the twisted ribbon and crossed-linker models (Figure 1.4) (Finch & Klug, 1976; Horowitz et al, 1994; Woodcock et al, 1984). Crosslinking experiments between nucleosomes have now suggested the existence of all three above models in one 30nm fibre (Grigoryev et al, 2009).

Throughout recent years more precise studies have indicated the uncertainty of existence of the 30nm fibre, at least on its own. Different types of compaction are seen in the case of nucleosomal arrays with different linker lengths (Perisic et al, 2010; Routh et al, 2008). Additionally chromosome conformation capture (3C) experiments have suggested the existence of a combination of 10nm and 30nm

chromatin segments (at variable proportions) as well as interactions between different chromatin fibres (Lu et al, 2006). Most recently a study demonstrated the irregular folding of nucleosome fibres, rather than a more rigid 30nm structure (Nishino et al, 2012). The varied models outlined above are discussed in a thorough recent review on chromatin structure and demonstrate the continued need for caution when proposing models for higher order structure (Bian & Belmont, 2012).

Compaction of chromatin to the extent outlined above would make it highly inaccessible to regulatory factors and therefore functionally repressed. This gives rise to the question, how different regions of the genome are intricately regulated in different cells and at different stages during development. Analyses of the density of chromatin in the whole cell have shed light on this issue. Not all parts of the genome are condensed to the same extent (Elgin & Grewal, 2003; Trojer & Reinberg, 2007). This is readily visible by cytological staining as very densely compacted heterochromatin and less compacted euchromatin. Euchromatin is gene rich, transcriptionally active and to a large extent decompacted, showing irregular nucleosomal spacing (Elgin & Grewal, 2003). It is useful to consider two types of heterochromatin: constitutive heterochromatin which is permanently condensed, gene poor and repeat rich (e.g. centromeric DNA) and facultative heterochromatin which is less tightly locked into the compact state (e.g. developmentally regulated genomic regions and the inactive X chromosome) (Elgin & Grewal, 2003; Trojer & Reinberg, 2007). Whereas centromeric and telomeric DNA is constitutively compacted, facultative heterochromatin and euchromatin undergo dynamic changes depending on cell type and developmental timing.

This variability and dynamics in chromatin structure demonstrates the need for mechanisms to allow transitions between different chromatin states. DNA methylation, is normally associated with the transition to a repressed chromatin state, though the ability of DNA methylation on its own to generate structural changes in chromatin is unclear. Additionally, there is the question of how DNA methylation is targeted to genomic regions. A recent study demonstrated the inhibitory effect of nucleosomes on methylation of DNA, revealing that linker DNA was readily methylated, whereas nucleosomal DNA remained largely unmethylated (Felle et al, 2011). This suggests, that even relatively 'open' chromatin may still be inhibitory to DNA methyltransferases as well as other regulatory factors.

Several epigenetic mechanisms of regulation of nucleosomal DNA have been identified. These include post-translational modification of histone proteins, as well as moving, removing and exchanging whole or parts of histone octamers by chromatin remodelling ATPases.

1.4 Histone modifications

Histones can be modified on their N-terminal tails, which protrude from the nucleosome. Although histone tail modifications comprise the majority, some modifications have also been found to occur within the globular domain (Williams et al, 2008). Histone N-terminal tails can be variously acetylated, methylated, phosphorylated, ubiquitinated, sumoylated, biotinylated and ADP-ribosylated. The best characterized modifications are lysine acetylation and methylation and serine and threonine phosphorylation. The result of these modifications can be a change in the charge of histone proteins and therefore their interaction with DNA. Lysine

acetylation for example reduces the positive charge of histones and therefore interaction with DNA, resulting in a more open conformation of chromatin (Shogren-Knaak et al, 2006). A more common effect of histone modifications is the generation or inhibition of binding sites for effector proteins.

Histone modifications are involved in many different functional pathways. A 'histone code' has been proposed which links different combinations of histone modifications to specific functional outcomes such as gene activation or repression (Munshi et al, 2009) (Figure 1.5A). Histones can for example be mono-, di-, or tri-methylated and each of these marks can give rise to different functional effects. This is epitomised by methylation of Histone 4 lysien 20 (H4K20). Different histone methyltransferases cause different levels of methylation on this histone residue (mono-, di- and tri methylation), leading to implication in different functions such as cell cycle progression, chromosome condensation, DNA damage and transcription regulation (Wang & Jia, 2009).

1.4.1 Distribution of histone modifications and transcriptional activity

Studies into the distribution of histone modifications across eukaryotic genes have in the past defined distinct profiles at active and inactive genes (Figure 1.5B). H3K4 methylation is usually found at active genes, with H3K4me2 across the gene body, whereas H3K4me3 has a more confined localization at the transcription start sites of genes (Bernstein et al, 2005). Additionally, H3K9 acetylation is also characteristic of active genes (Liang et al, 2004).

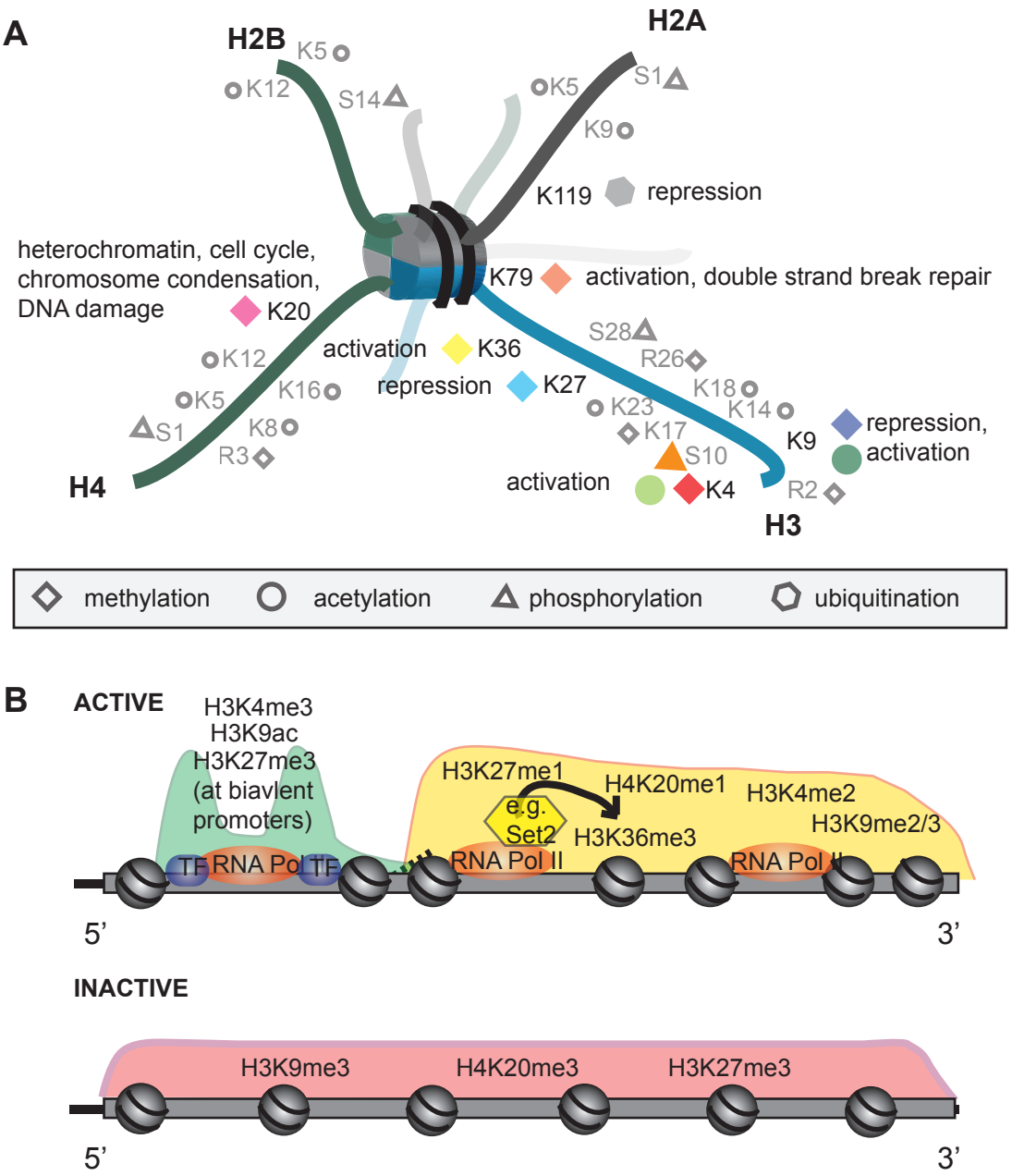


Figure 1.5 Histone modifications

A: Histone modifications and functions. Displayed is a nucleosome with protruding histone protein N-terminal tails. Positions of amino acid residues and their potential chemical modifications are indicated by the one letter amino acid abbreviation as well as a modification symbols. Modification symbols are explained underneath. The most well studied histone modifications are coloured. Functions of these are annotated next to the residues. **B:** Distribution of histone modifications along active and inactive genes. Coloured shading indicates which region of a gene is associated with histone modifications indicated within the shaded region.

Inactive genes on the other hand show characteristic accumulation of H3K27me3, H3K9me2/3 and/or H4K20me3 marks throughout the gene body as well as at the promoter (Cao et al, 2002; Rice et al, 2003). This simple overview of distribution of histone modifications is consistent with functions allocated to different histone modifications in the histone 'code'.

An increasing number of studies demonstrate a more complex nature of interaction between histone modifications at active and inactive genes. One of the most striking discoveries was the presence of H3K9me2/3 (a repressive mark) throughout active genes, apart from transcription start sites (Vakoc et al, 2005; Vakoc et al, 2006). Additional studies identified the presence of repressive marks H3K36me3, H3K27me1 as well as H3K20me1 throughout active genes, excluding transcription start sites (Carrozza et al, 2005; Vakoc et al, 2006). Consistently, SET2 histone methyltransferase responsible for H3K36 methylation has been found to interact with the elongating RNA PolII. Several studies have now pointed towards a function of H3K36 methylation in suppressing aberrant transcription initiation by RNA Pol II throughout transcribed genes (Carrozza et al, 2005; Joshi & Struhl, 2005; Keogh et al, 2005; Kizer et al, 2005). A very recent study demonstrates that H3K36 methylation prevents histone exchange throughout gene bodies. Histone exchange is normally a co-transcriptional event to introduce acetylated histones. This is essential for repressing aberrant transcription initiation within transcribed gene bodies (Venkatesh et al, 2012).

Another interesting discovery was the co-localization of H3K4me3 and H3K9 acetylation at most gene promoters in human embryonic stem cells, independent of their transcription status (Mikkelsen et al, 2007). These so called bivalent chromatin

states have been found to occur widely in mouse ES cells, with enrichment of H3K4me3 and H3K27me3 marks at gene promoters poised for expression(Bernstein et al, 2006; Mikkelsen et al, 2007). Mikkelsen et al. identified that transcription of most of these promoters was initiated, but inactive genes contained paused RNA Polymerase which did not elongate. The above examples highlight a complex system of interactions between histone modification marks, to regulate the gene activity status.

1.4.2 Effectors of histone modifications and crosstalk

The most common mechanism for histone modifications to cause changes in downstream functions is to provide binding platforms for effector proteins (Figure 1.6). This can cause the recruitment of other co-repressor/co-activator complexes (such as Polycomb Repressive Complex 2 by H3K27me3 and Chd1 by H3K4me3) (Torok & Grant, 2006). Proteins that recognize histone modifications are often enzymes themselves, which can catalyse additional histone modifications with similar functional outcome. Often a prerequisite for the functions of these enzymes is their binding to an already established modification. An example occurs during Polycomb mediated silencing of developmental genes discussed in more detail below. In this case Polycomb repressive complex 1 (PRC1) can be recruited to previously established H3K27me3 marks by Polycomb repressive complex 2 (PRC2) (Figure 1.6A). PRC1 in turn catalyses ubiquitination of H2AK119. Both H3K27me3 and H2AK119ub are repressive marks.

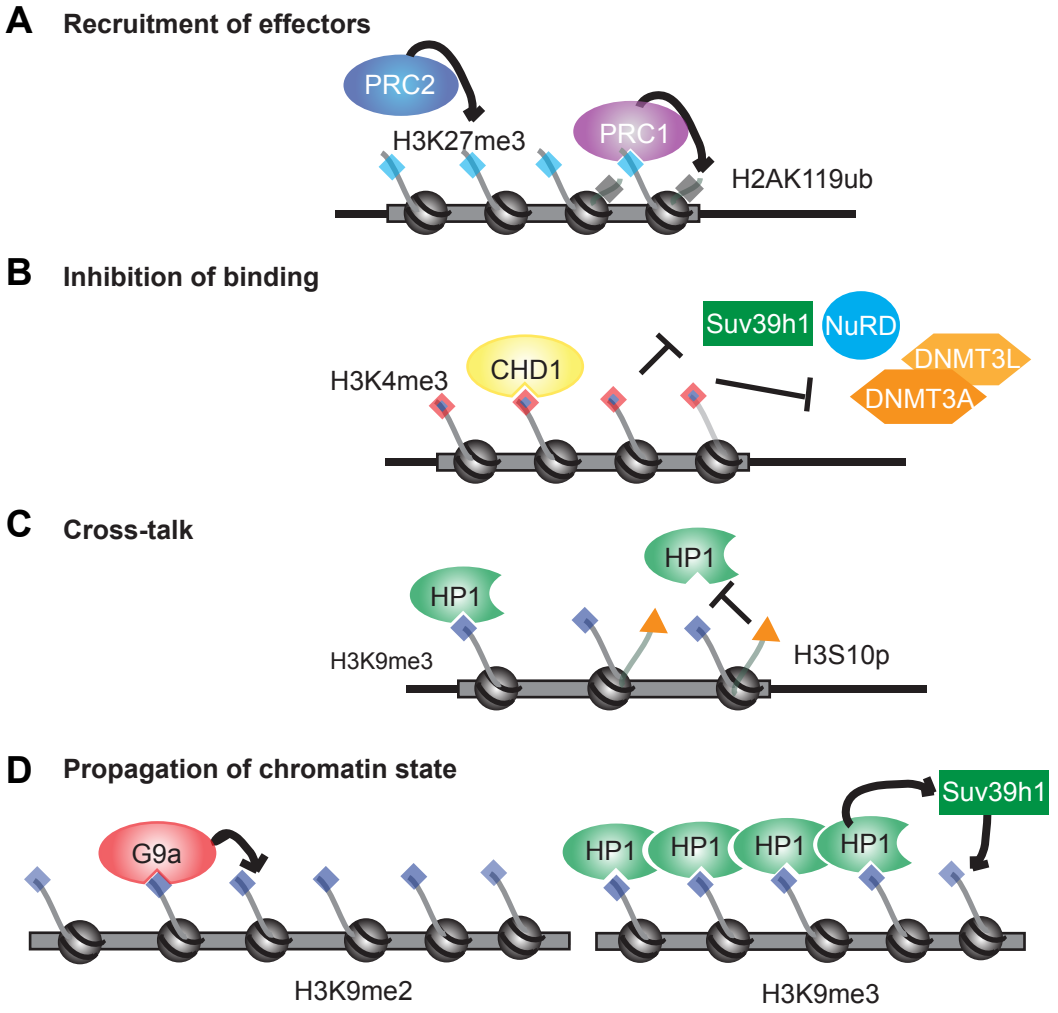


Figure 1.6 Effectors of histone modifications

A: Histone modifications catalysed by one enzyme can act as recruiting platforms for other enzymes, which put in place other histone modifications to support the formation of a specific chromatin state. **B:** Binding of effector proteins can be dependent on certain histone modifications. H3K4me3 can recruit coactivator complexes such as CHD1 and simultaneously inhibit binding of Suv39h1,NuRD and association of DNMT3A/3L complex with chromatin. **C:** Different histone modifications can be involved in crosstalk. H3S10 phosphorulation prevents binding of Heterochromatin protein 1 (HP1) to H3K9me3 marks. **D:** Histone modifications can initiate events to propagate a particular chromatin state. G9a histone methyltransferase catalyses H3K9me2 and also binds to this mark, ensuring continued modification of surrounding histone tails. Suv39h1 catalyses H3K9me3 during formation of heterochromatin. HP1 binds the H3K9me3 mark and can self-dimerize. This may facilitate heterochromatin propagation through repeated recruitment of Suv39h1. Additionally self-dimerization may also facilitate heterochromatin compaction.

In Drosophila additional regulation involves stimulation of H2AK119 ubiquitination by PRC1 through H3K36 demethylation by the dRAF complex (Lagarou et al, 2008).

Initial establishment of a specific chromatin state by binding of co-regulators is enhanced by frequent inhibition of recruitment of factors involved in creating opposite chromatin environments. H3K4 methylation is seen at active genomic regions and recruits CHD1, activator complex. It can simultaneously inhibit H3K9 methylation by SUV39H1 and prevents binding of the NuRD repressive complex and Sir3 silencing factors to H3 tails (Figure 1.6B) (Nishioka et al, 2002; Zegerman et al, 2002).

Interesting, additional regulation by histone modifications occurs through regulation of recognition of relatively stable histone modifications with the aid of more flexible modifications. This may allow stable modifications to be retained on chromatin whilst regulating their recognition at specific times during development or the cell cycle. An example is the phosphor switch which involves serine/threonine phosphorylation inhibiting binding of effector proteins to specific modified histone residues (Figure 1.6C). H3S10 phosphorylation for example is involved in chromosome condensation and segregation. During mitosis it inhibits binding of HP1 to the H3K9me3 mark. This phenomenon can be seen to occur between adjacent modifications on the same N-terminal tail or modifications on different tails or nucleosomes (Fischle et al, 2005; Hirota et al, 2005)(Fischle et al. 2005; Hirota et al. 2005).

Change of a chromatin state is often also accompanied by spreading to neighbouring regions. Firstly, enzymes involved in catalysing a particular histone

modification often possess binding sites for these modifications. Therefore, once the modification is established, these enzymes are retained through binding this modification. An example of this is G9a euchromatin histone H3K9 methyltransferase which contains a catalytic SET domain for H3K9 modification and also contains H3K9me3 binding domains (Figure 1.6D) (Collins et al, 2008). Effector protein self-dimerization can also facilitate the propagation of chromatin states as in the case of HP1. HP1 is recruited to H3K9me3 marks and can, through self-dimerization and recruitment of co-repressor complexes, promote propagation of heterochromatin. Finally, certain effector proteins have more than one chromatin/DNA binding site and can therefore cause bridging between different regions (Brasher et al, 2000; Cowieson et al, 2000; Eissenberg & Elgin, 2000).

1.4.3 Histone modifications in developmental regulation of gene activity – PcG

Histone modifications are proving to be multifaceted regulatory networks of different chromatin states. Drawing together all the function of histone modifications, the Polycomb repressive complexes (PRCs) are an example of how these mechanisms can be applied to developmental regulation of gene activity. Polycomb repression is important in silencing of genes in the *Hox* cluster as well as other lineage specific genes. EZH2 within PRC2 causes H3K27me3 at target genes (Figure 1.7A) (Kirmizis et al, 2004).

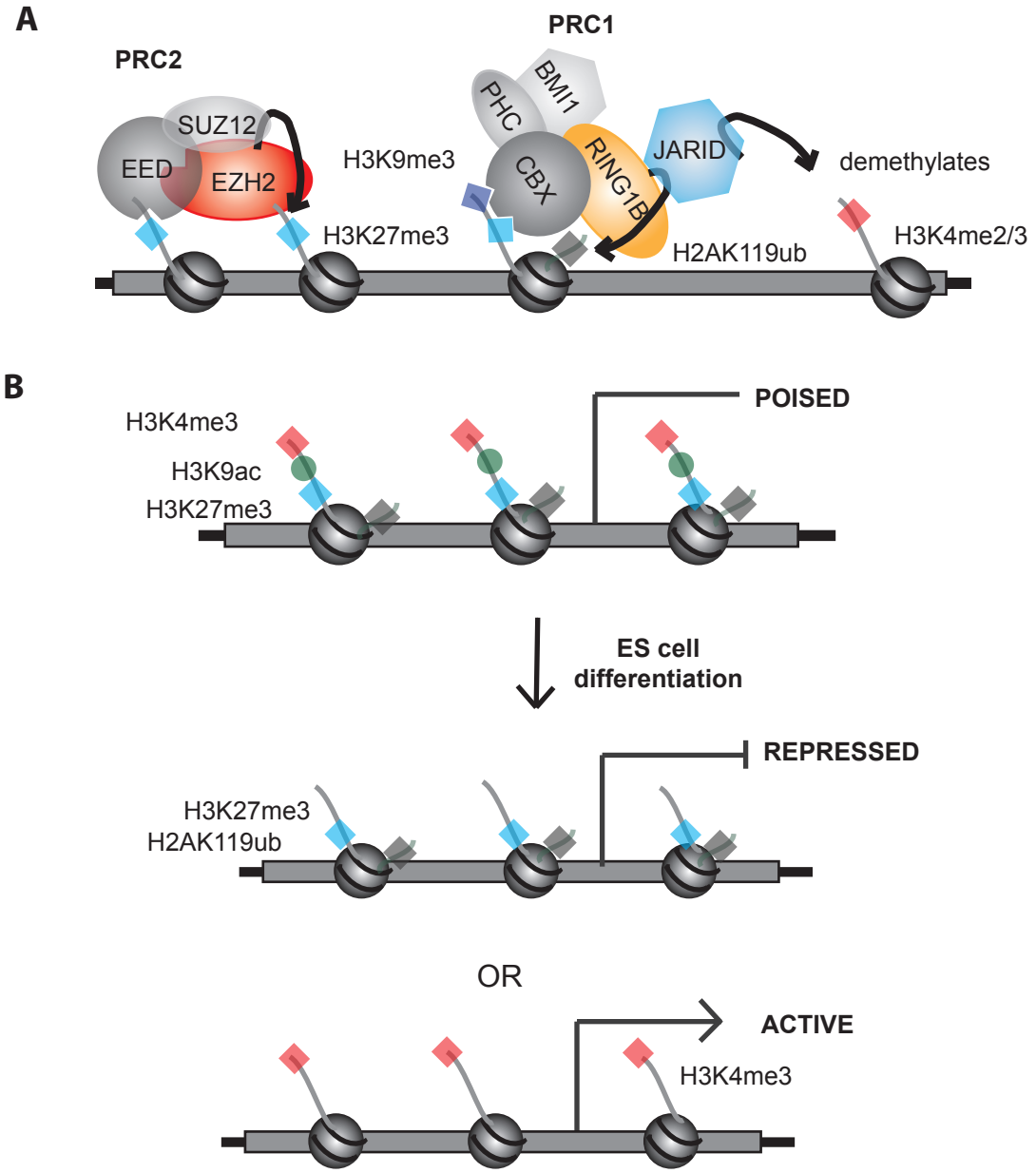


Figure 1.7 Polycomb complexes and regulation of lineage commitment

A: Outline of polycomb mediated repression. PRC2, consisting of SUZ12, EED and EZH2, catalyses the repressive H3K27me3 mark. This histone modification acts as a recruiting platform for PRC1 complex, which in turn catalyses H2AK119 ubiquitination and also recruits accessory proteins such as JARID, to remove the H3K4me2/3 active mark from chromatin. **B:** Many polycomb target gene promoters have recently been found to be marked by a bivalent histone modification state (H3K4me3 as well as H3K27me3). These bivalent state are though to represent a poised chromatin state, allowing rapid activation or continued repression of lineage specific genes during differentiation of stem cells.

PRC1 recognises the H3K27me3 mark through its chromodomain and catalyses the ubiquitination of H2A119 (Wang et al. 2004). In conjunction with this, JARID1D, a histone demethylase is thought to interact with PRC2 which may regulate its activity and causes removal of H3K4me2/3 marks, resulting in a repressive chromatin state (Lee et al, 2007; Morey & Helin, 2010; Pasini et al, 2008)(Figure 1.7A).

Models have been proposed to explain how PRC2 complex is recruited to chromatin. Early models suggested PRC2 recruitment by transcription factors via Polycomb Repressive Elements (DNA sequence elements) (PREs). PREs have been identified in *Drosophila melanogaster.* They are several hundred base pairs long and can be located relatively far away from transcription start sites. They function in recruitment of Polycomb group proteins (PcG) through recruitment of transcription factors (Ringrose & Paro, 2007). A couple of studies have suggested the presence of PREs in vertebrate genomes (Sing et al, 2009; Woo et al, 2010). Still, targeting of PcG proteins to these in vertebrates remains uncertain from this study.

A second model proposes the recruitment of PRC2 through cis- or trans-acting non-coding RNAs (ncRNAs). This method of recruitment has now been suggested in various recent studies. One example is the ncRNA HOTAIR which is involved in transcription repression of the *HoxC* locus. HOTAIR has been shown to recruit PRC2 in trans to the *HoxD* cluster (Rinn et al, 2007). Consistently, HOTAIR binds many PRC2 target genes (Chu et al, 2011). A similar effect is initiated by *Xist* RNA on the X chromosome, during X chromosome inactivation. A non-coding RNA (RepA) within *Xist* was shown to associate with PRC2 components (Zhao et al, 2008). Additional evidence for a Polycomb recruitment mechanism using RNAs comes from studies demonstrating PRC2 association with many other long non-

coding RNAs (lncRNAs) (Khalil et al, 2009; Zhao et al, 2010). Another interesting study revealed that short RNAs transcribed from Polycomb target genes associate with PRC2, suggesting RNA mediated recruitment mechanisms to be widespread (Kanhere et al, 2010).

Polycomb complexes are involved in silencing of cell lineage-regulatory genes in pluripotent cells (e.g. Hox genes). Many Polycomb target genes have recently been found to be occupied by both active and repressive histone marks such as H3K27me3, H3K4me3 and H3K9ac (Figure 1.7B). This is now known as a bivalent chromatin state (Azuara et al, 2006). Interestingly many of these bivalent genes are resolved during differentiation of stem cells into neuronal progenitor cells to leave behind either active or inactive genes (Bernstein et al, 2006). Bivalency in this case is thought to provide a poised gene expression state, allowing quick decision between chromatin states to occur during lineage commitment.

The outline of post-translational modifications to histones above provides a mechanism for initiating changes in chromatin. Since these modifications occur on N-terminal tails of histones, they provide mechanisms to recruit activator and repressor complexes to relatively compacted regions of the genome. These modifications are therefore more versatile than DNA methylation itself. Despite this, only a few histone modifications have been shown to alter chromatin structure directly, leaving us with the still unanswered question of how more large scale chromatin changes can occur. These are required to create nucleosome free regions, to allow access to regulatory proteins. As said above, this is necessary at both ‘open’ and compacted chromatin.

1.5 Chromatin remodelling

1.5.1 Families of chromatin remodelling ATPases

Chromatin remodelling is an ATP dependent process carried out by members of helicase families (Flaus et al, 2006). Three helicase superfamilies (SF1, SF2, SF3) exist, which all contain conserved helicase motifs. SF1 and SF2 are most similar to each other, possessing an ATPase domain comprised of two RecA like lobes with two alpha-helical protrusions in between, separated by a linker region (Figure 1.8A). This ATPase region contains the seven conserved helicase domains.

The SF2 superfamily can be further subdivided into several families, including the Snf2 family (Figure 1.8B). The distinguishing feature of Snf2 family members is the lack of the common DNA strand separation ability. Instead family members show ATP dependent DNA translocation (Flaus et al, 2006). In support of this, the ATPase domain common to all Snf2 family members is very similar to those of known DNA translocation enzymes. An additional distinguishing feature is the longer helicase domain, which contains an insertion between helicase motifs III and IV (Flaus et al, 2006).

Even further subdivision of the Snf2 family gives rise to distinct subfamilies of chromatin remodellers, depending on sequence homology of the helicase domains: SWI/SNF, ISWI, INO80/SWR1, NURD/Mi2/CHD (Cairns, 2007; Flaus et al, 2006) (Figure 1.8C).

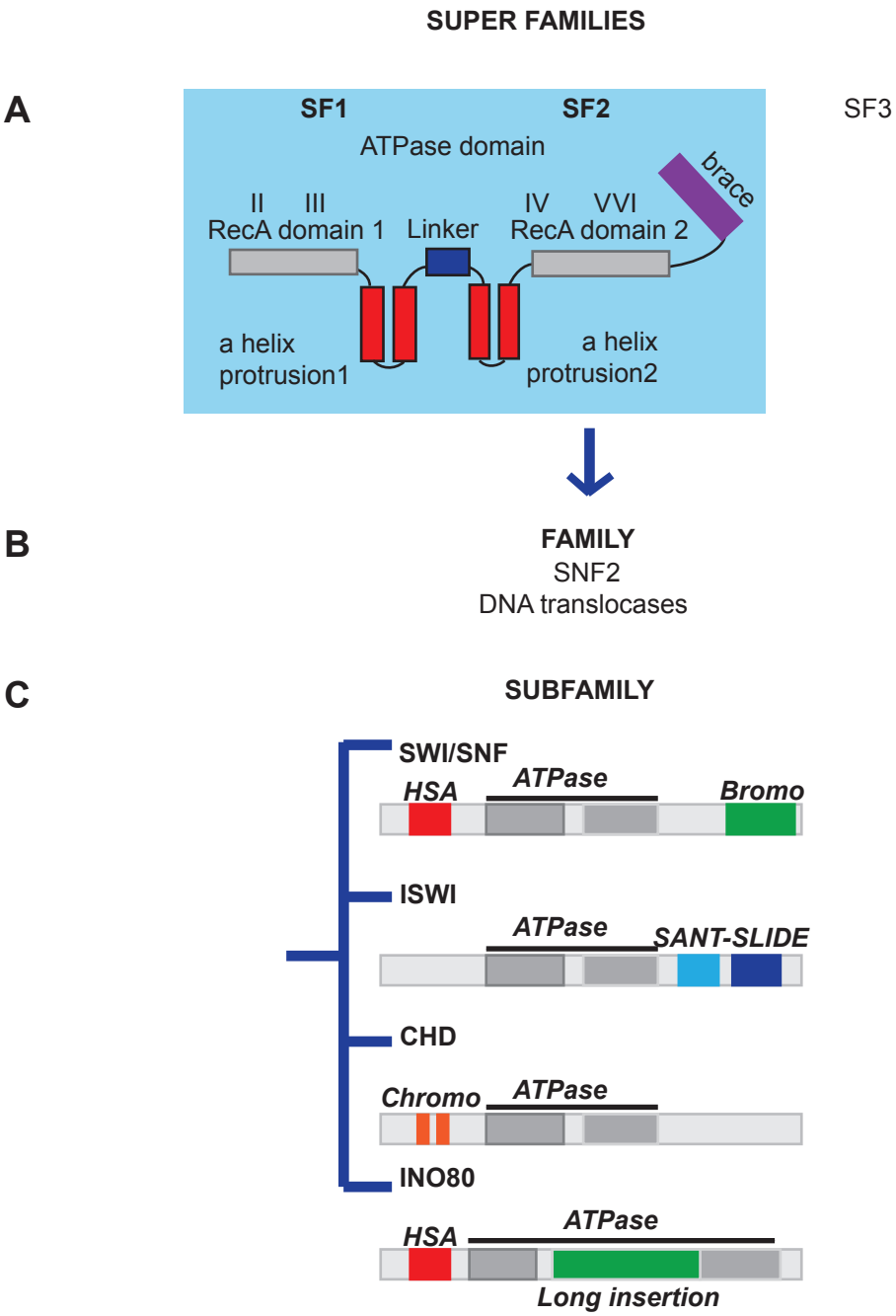


Figure 1.8 SNF2 chromatin remodelling ATPases

A: Helicase superfamilies. SF1, SF2 and SF3 helicase superfamilies exist. SF1 and SF2 superfamilies are most similar to eachother, containing a conserved ATPase domain. **B: Family of DNA translocases.** SF2 superfamily can be subdivided into various families, including the SNF2 family of DNA translocases. **C: Subfamilies.** The SNF2 family can be subdivided further into subfamilies of chromatin remodelling ATPases: SWI/SNF, ISWI, CHD and INO80.

All chromatin remodelling enzymes show affinity for nucleosomes, have domains that recognize histone modifications and associate with other proteins (e.g. transcription factors). Despite this, family members have evolved specialized roles which are in part conferred by various additional domains within and outwith the ATPase region (Clapier & Cairns, 2009)(Figure 1.8C). Many of these domains have been implicated in targeting to chromatin as well as in recruiting factors which stimulate ATPase activity. The HSA domain present in both SWI/SNF and INO80 remodellers binds actin as well as actin related proteins (ARPs), which stimulate ATPase activity (Shen et al, 2003; Szerlong et al, 2008). In addition, INO80 family members have a distinctive extended linker region between the two RecA lobes (split ATPase domain). This allows binding of an additional ARP as well as helicase related protein RVB1/2, both of which are required for remodelling activity (Jonsson et al, 2004). The SANT-SLIDE domain within the ISWI complex contacts unmodified histone tails as well as the DNA itself (Boyer et al, 2004). CHD complexes have two tandem chromodomains, which can interact with H3K4me2/3 as well as DNA. Similarly, SWI/SNF remodellers have a bromo domain which recognizes acetylated histones (Clapier & Cairns, 2009). The above highlights the need for different domains to increase remodeller activity as well as to aid recruitment of remodellers to chromatin.

1.5.2 Mechanism of action

Chromatin remodellers can either reposition or eject nucleosomes or allow loading of histone variants. Studies have identified mechanistic differences between remodellers, which associate them with the above remodelling functions.

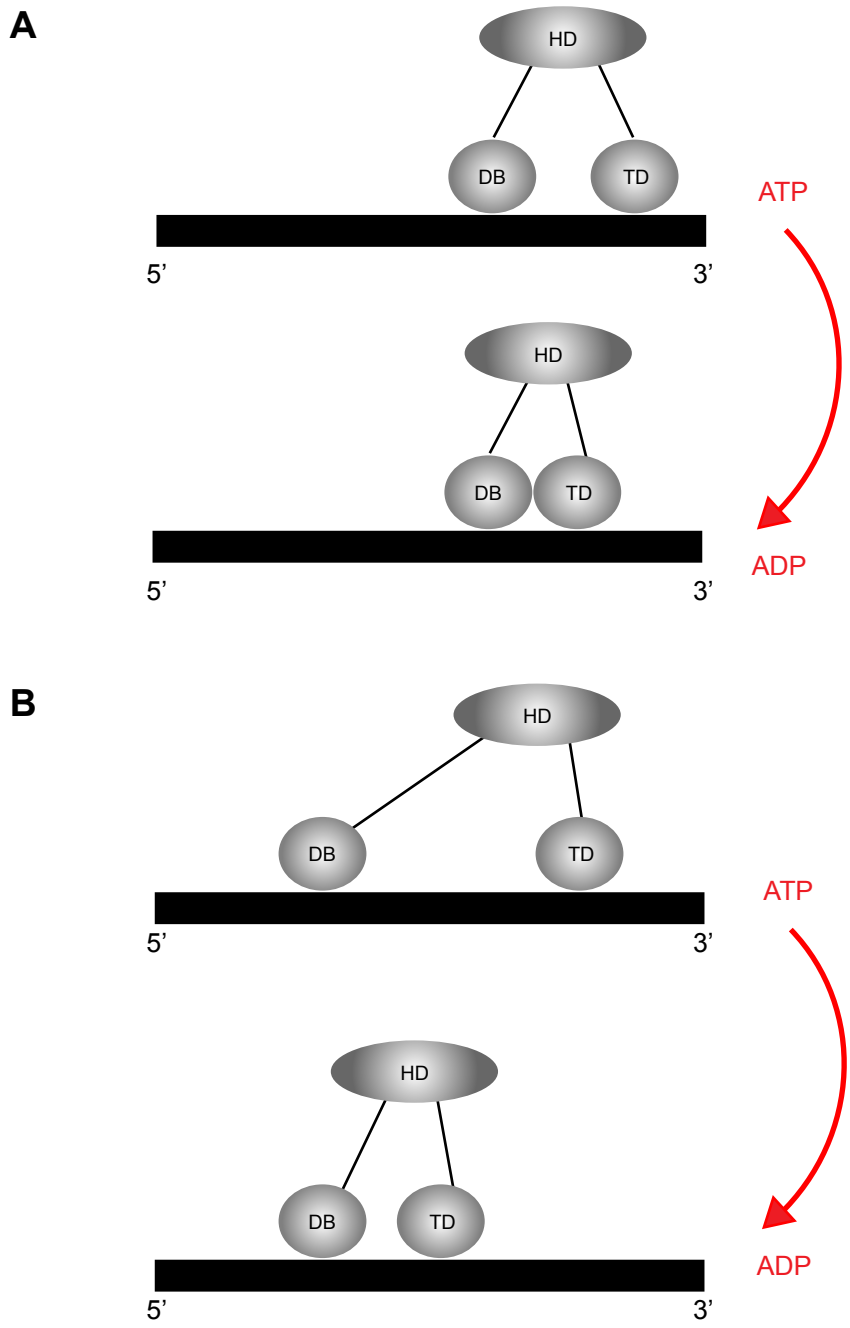


Figure 1.9 DNA translocation by monomeric chromatin remodeller

A: The chromatin remodeller has a hinge domain (HD), DNA binding domain (DB) and a DNA tracking domain (TD). The chromatin remodeller binds to DNA with its DNA binding domain (DB) and DNA tracking domain (TD). ATP hydrolysis induces a conformational change, allowing the TD to translocate in a 3’ to 5’ direction. B: Next, the DB binds further along the DNA to initiate another cycle of ATP hydrolysis and translocation of the TD.

SWI/SNF and ISWI remodellers have been most intensely studied with respect to their remodelling mechanisms, but all remodellers function using DNA translocation mechanisms (Saha et al, 2002; Saha et al, 2005; Whitehouse et al, 2003; Zofall et al, 2006). A translocation mechanism for monomeric remodellers has been proposed which relies on two domains, the DNA tracking domain as well as a DNA binding domain (Figure 1.9). DNA translocation is initiated by ATP binding and hydrolysis, which induces conformational changes between the RecA motifs of the remodeller (Figure 1.9A). This results in the movement of the remodeller tracking domain in a 3' to 5' direction. The DNA binding domain is located in front of the tracking domain.

Once a cycle of ATP binding and hydrolysis has been completed, the DNA binding domain can bind further along the DNA and allow another cycle of translocation to occur (Figure 1.9B)(Cairns, 2007). This model has been in recent years adapted to translocation on the nucleosome surface, minimising steric clashes with histone octames, as well as the twist induced in the DNA (Cairns, 2007). Three models of remodeller translocation and DNA movement around the nucleosome have been proposed. These are the bulge propagation, twist diffusion and histone core swivel models (Figure 1.10) (Bowman, 2010).

1+10 ratchet model (bulge propagation)

In this model, the remodeller binds with its hinge domain to the histone core (Figure 1.10A). The tracking domain and the DNA binding domain bind approximately 1 helical turn apart to the nucleosomal DNA.

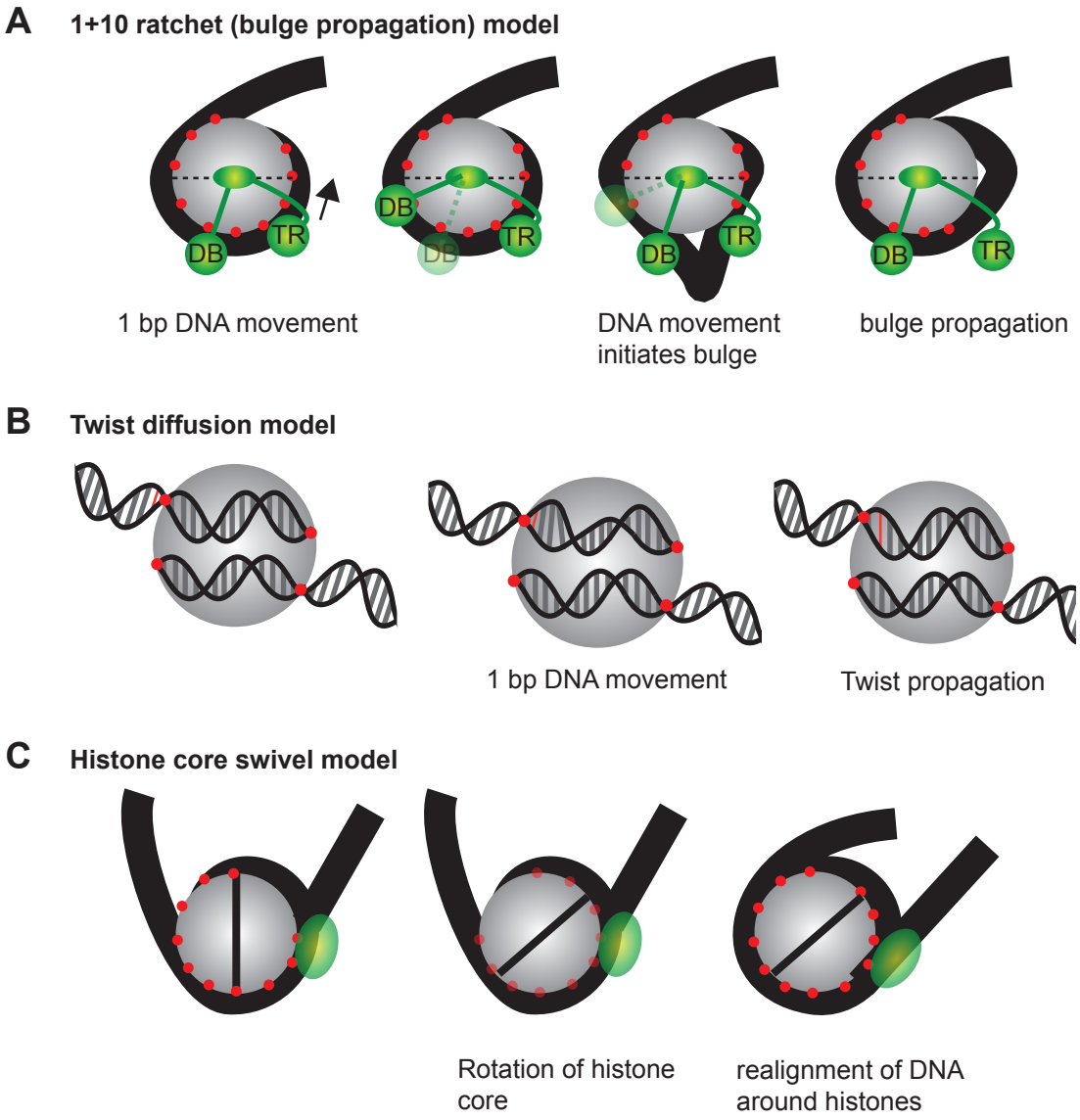


Figure 1.10 Mechanisms of DNA movement around histone octamer

A: 1+10 ratchet/bulge propagation model. Remodeller binds with its hinge domain to the nucleosome core. The DNA tracking (TR) and DNA binding (DB) domains bind two and three turns away from the axis going through the centre of the nucleosome respectively (dashed line). The DNA tracking domain moves 1bp of DNA towards the nucleosomal axis. Next, the DB domain binds 10bp further away from the TR domain and subsequently pulls DNA towards the tracking domain. The final step involves propagation of the DNA loop through the tracking domain. **B: Twist diffusion model.** Initiation of a twist in the DNA helix occurs through movement of 1bp of DNA (blue) from linker DNA to a region in between two histone-DNA contacts (red). The twist is propagated around the nucleosome. The difference to the 1+10 ratchet model is the lack of formation of a big bulge **C: Histone core swivel model:** Remodeller (green) binds two helical turns away from the DNA entry site. ATP hydrolysis induces conformational changes which disrupt histone-DNA contacts and allow the histone core to move one helical turn towards the remodeller. The final stage involves realignment of DNA relative to the histone core.

The first step is thought to involve a 1 base pair movement of DNA by the tracking domain away from the DNA binding domain. In the second step, the DNA binding domain binds one helical turn (10bp) further away from the tracking domain and pulls DNA towards the DNA tracking domain, at the same time returning to its previous position. This produces a DNA loop which is subsequently propagated around the nucleosome through the tracking domain (Bowman, 2010; Cairns, 2007).

Twist diffusion model

The twist diffusion model involves the remodeller binding and pulling 1bp of DNA from the linker DNA into a region between two histone-DNA contacts (Figure 1.10B)(Bowman, 2010). This induces increased twist into the double helix, which is subsequently propagated around the nucleosome. Unlike the bulge diffusion model, this method of DNA movement involves minimal disruption of histone-DNA contact sites.

Histone core swivel model

Unlike the other two models, this model involves rotation of the histone core, followed by a realignment of the DNA around it (Figure 1.10C)(Bowman, 2010). The remodeller binds two helical turns away from the DNA entry site of the nucleosome. ATP hydrolysis by the remodeller is accompanied by a DNA conformational change, which makes it favourable for the histone core to move one helical turn towards the remodeller. The final stage involves the DNA rewrapping with entry and exit sites to the nucleosome being equidistant from the dyad axis.

The three above described mechanisms can cause differences in bulge size as well as amount of force produced. Structural differences in the complex between the

remodeller, DNA and histones may make certain remodellers more amenable to a specific mechanism.

ISWI remodellers, for example, require H4 tails to stabilize the ATPase domain (Dang et al, 2006). In addition they form extensive contacts with the DNA around the nucleosome (Dang et al, 2006). SWI/SNF remodellers in contrast show fewer DNA contacts but on the other hand form more extensive contacts with the histone core (Dang & Bartholomew, 2007; Dechassa et al, 2008). More extensive contacts with histones may allow DNA to be more efficiently pulled from the histone surface, creating the characteristic bigger loops of DNA produced by these remodellers. Consistently, SWI/SNF remodellers are involved in more dramatic remodelling events such as nucleosome ejection or dimer displacement/exchange (Bruno et al, 2003; Kassabov et al, 2003).

An additional contributing factor to the ways of action of remodellers is the requirement of linker DNA for efficient remodelling. Studies have shown that ISWI, Ino80 and CHD remodelling complexes require a certain length of linker DNA to efficiently perform nucleosome remodelling functions (Udugama et al, 2011) (Gangaraju & Bartholomew, 2007; Stockdale et al, 2006). Consistently, ISWI and Ino80 complexes are involved in nucleosome spacing (Udugama et al, 2011). They bind to regions with long linker DNA and dissociate once the nucleosomes have been more equally spaced. SWI/SNF complexes on the other hand have been identified not to be dependent on linker DNA length (Dechassa et al, 2010). This allows SWI/SNF complexes to bind more densely spaced nucleosomes.

Linker DNA requirement may arise from differences in binding of remodellers to the nucleosome. The DNA binding domain of ISWI for example may bind to linker DNA rather than within the nucleosomal DNA. Additionally, some remodellers have been found to bind as dimers, some as monomers (Hota & Bartholomew, 2011). This difference in binding can also be seen within families of remodellers (Racki et al, 2009).

1.5.3 One remodeller, several functions

Consistent with the diversity in chromatin remodelling mechanisms, remodellers are involved in processes such as chromatin assembly, correct spacing of nucleosomes after replication, regulation of accessibility to DNA binding factors, efficient DNA repair and progression of RNA and DNA polymerases (Clapier & Cairns, 2009). Functional diversity of different families of chromatin remodellers arises through differential assembly of subunits around a core chromatin remodeller. Complexes from each family of chromatin remodellers are outlined in Figure 1.11. The ISWI family of remodelling complexes is a good example of this. Seven ISWI complexes have been identified to date, all containing either Snf2H or Snf2L ATPase isoforms. Difference in functionality is conferred by additional subunits. WICH and RSF complexes for example only differ from each other by one subunit. WSTF catalyses the phosphorylation of H2AX and therefore implicates the WICH complex in DNA damage pathways (Xiao et al, 2009). The RSF complex on the other hand is seen to interact with Centromere protein A (CENP-A), regulating centromere structure (Perpelescu et al, 2009).

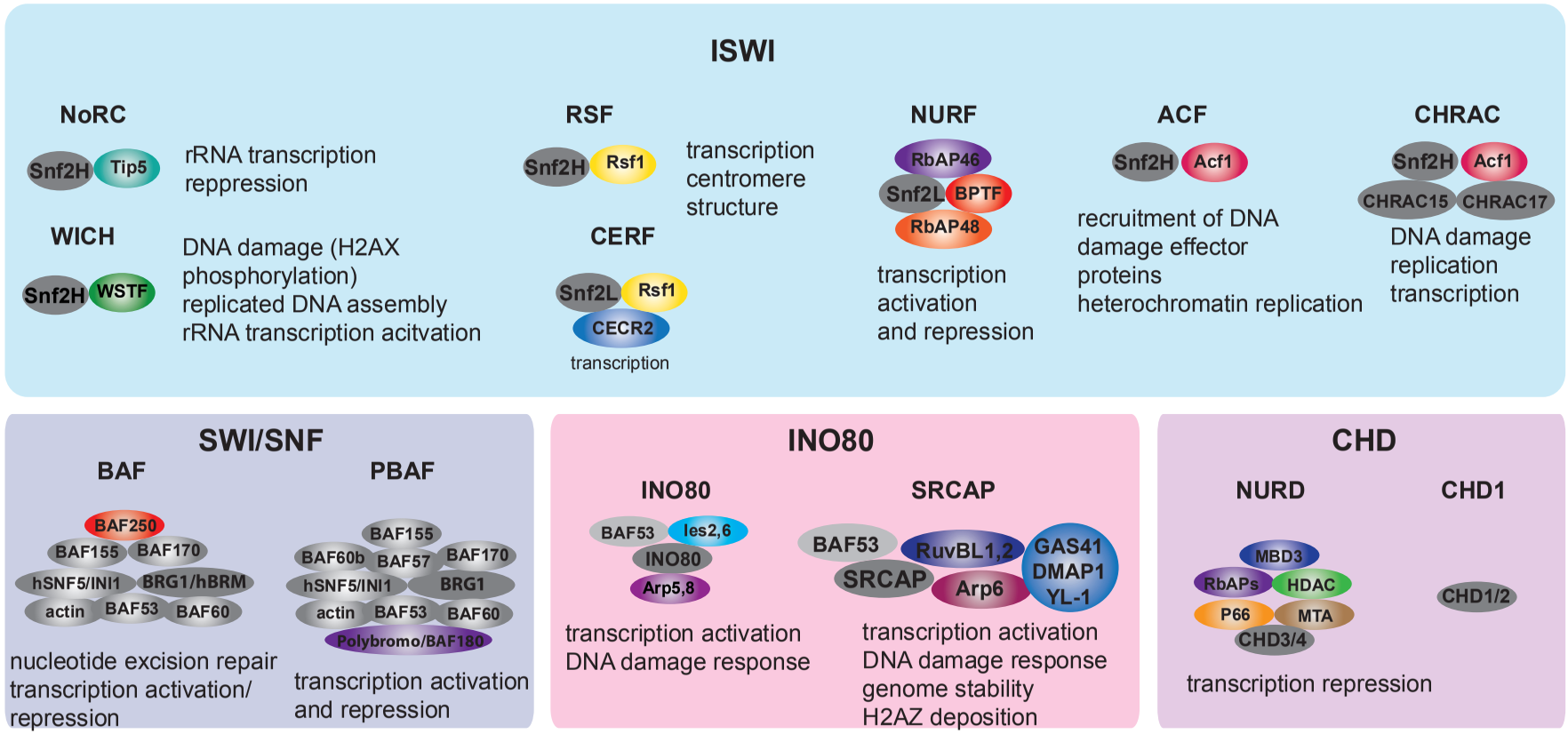


Figure 1.11 Combinatory assembly of mammalian chromatin remodeller complexes for functional diversiy

Displayed is the assembly of ISWI, SWI/SNF, INO80 and CHD remodellers into different complexes. Combinatory assembly can result in complexes with very distinct functions. Functions are indicated below or next to each complex.

Combinatorial assembly of remodeller complex subunits has also been shown to be developmentally dependent (Ho & Crabtree, 2010). Extra dimension is added by families of proteins which can be exchanged during development. This allows assembly of the subunits required for interactions with developmental and cell type specific regulators. To date mammalian BAF chromatin remodelling complex within the SWI/SNF2 family shows the widest variety of combinatorial assemblies of subunits to produce complexes involved at most developmental stages (Ho & Crabtree, 2010). BAF complexes can be assembled around either BRG1 or BRM chromatin remodellers (Figure 1.11). In addition, many BAF subunits are expressed in gene families, whose members can be interchanged to form developmentally distinct complexes. BAF complexes are the complexes most involved in early development. They are required for maintaining the pluripotent state through regulation of the pluripotency transcriptome (Ho et al, 2009a; Ho et al, 2009b). Additionally they are also required for differentiation of stem cells into different cell lineages (Schaniel et al, 2009).

1.6 Crosstalk

1.6.1 Forming DNA methylation and histone modification landscapes

Altered DNA methylation and histone modifications in cancer

Establishment of correct DNA methylation landscapes has proven to be essential for development. Defects in DNA methylation processes have been linked to many human diseases. DNA hypomethylation was the first epigenetic modification defect linked to cancer (Feinberg & Vogelstein, 1983). Hypomethylation often occurs at repetitive elements, retrotransposons and in gene introns, resulting in genome instability through chromosomal rearrangements and

high probability of translocations of retrotransposons (Eden et al, 2003; Weber et al, 2005). Hypomethylation can also affect genic regions, resulting in activation of tumour suppressor genes (Feinberg & Vogelstein, 1983). Consistently, DNA hypomethylation has been found to increase during tumour formation (Fraga et al, 2004). Additionally, DNA hypermethylation has been detected at CpG island promoters as well as CpG island shores, potentially leading to silencing of tumour suppressor genes (Irizarry et al, 2009; Kron et al, 2009; Mori et al, 2006). A recent study carried out a DNA methylation fingerprint of 1628 human samples, revealing the extensive DNA methylation changes that occur in cancer tissues (Fernandez et al, 2012). Apart from a direct effect of DNA methylation defects on gene expression in tumours, DNA hypermethylation has also been demonstrated to create mutational hotspots in somatic cells due to the high mutability of 5meC (Rideout et al, 1990).

In addition to DNA methylation changes that occur during tumour formation, histone modifications also undergo changes. Interestingly, some specific changes in histone modifications have been demonstrated to be markers of high risk of recurrence of tumours and mortality (Barbisan et al, 2008; Barlesi et al, 2007; Seligson et al, 2005). Histone modification changes are often seen on global level, rather than being localized. These include global changes in H4K20me3 and H4K16 acetylation, as well as changes in H3K9 methylation. Histone modification changes have been demonstrated to occur in conjunction with DNA methylation changes (Fraga et al, 2005; Muller-Tidow et al, 2010; Tryndyak et al, 2006). Additionally, tumour types have been found to be distinguishable by expression of histone modifying enzymes and co-factors (HDACs, Polycomb, G9a, LSD1) (Bracken & Helin, 2009; Chase & Cross, 2011; Kondo et al, 2008)

Loss of imprinting disorders

Other diseases involving DNA methylation defects include a wide range of imprinting disorders. Imprinting disorders are accompanied by loss of imprinting (LOI). This can occur through gain or loss of DNA methylation or other changes resulting in defects in allele specific expression. In many cases LOI has been linked to development of cancers. An example is loss of imprinting at the *Igf2/H19* gene cluster. Gain of expression of *Igf2*, encoding a growth promoting factor is a common cause of cancers, such as Wilms Tumour and Beckwith-Widemann syndrome (Riccio et al, 2009). Other imprinting disorders are accompanied by loss of imprinting of large clusters of imprinted genes (e.g. PWS and Angelman syndromes) (Robertson, 2005).

Repeat instability diseases

DNA methylation defects have also been linked to repeat instability disorders such as Fragile X syndrome. This disorder involves the expansion of a CGG repeat at the 5'end of the *Fmr1* gene and subsequent hypermethylation of this region, resulting in transcription repression of this region (Crawford et al, 2001; Oberle et al, 1991).

Defects in DNA methylation machinery

Finally, various well studied diseases arise from mutations in the DNA methylation machinery. Systemic Lupus Erythematosus is accompanied by losses of DNA methylation and reduced levels of DNMT1 in T cells (Richardson et al, 1990). In ICF syndrome individuals, hypomethylation of pericentric heterochromatin is visible on certain chromosomes in lymphoid cells. Most disease cases have been linked to mutations in *Dnmt3b* (Hansen et al, 1999; Tuck-Muller et al, 2000; Xu et al, 1999).

These varied diseases demonstrate the importance of correct establishment of DNA methylation landscapes. A question remaining open for study is, how intricate DNA methylation patterns are achieved during mammalian development as well as in adults. A multitude of studies have highlighted many different DNA sequence requirements in targeting DNA methylation. These include CpG density and transcription factor binding sites for example. This was especially well highlighted in a recent study by Lienert and colleagues, which showed correct establishment of ES cell and differentiated ES cell specific DNA methylation marks at ectopic sites, through introduction of short sequence elements termed methylation determining regions (MDRs) (Lienert et al, 2011) (Meissner, 2011). These regions were sufficient to recapitulate endogenous methylation patterns in ES cells as well as during differentiation. Despite these sequence determinants of DNA methylation, many studies have focused on the co-operation between histone modifications, DNA methylation and chromatin remodelling to generate correct epigenetic landscapes.

One of the earliest examples of histone modification and DNA methylation cooperation came in the study of the mutually exclusive nature of DNA methylation and H3K4me3. Studies have consistently shown the presence of H3K4me3 at CpG island promoters which are devoid of DNA methylation. Recent studies even rely on this mark for identification of novel promoters (Illingworth et al, 2010). Consistently, certain H3K4 methyltransferases have been found to interact with RNA PolII which is mostly present at CpG island promoters in early embryos (Guenther et al, 2007; Terzi et al, 2011). DNMT3L has been shown to bind to unmodified H3K4 residues and recruit DNMT3A and 3B (Ooi et al, 2007). This interaction between DNMT3L and chromatin is disrupted by methylation of H3K4 residues (Meissner et al, 2008;

Mohn et al, 2008; Okitsu & Hsieh, 2007; Weber et al, 2007). This may provide a mechanism of preventing DNA methylation at CpG island promoters in early development.

Additional examples of histone modifications or histone modification enzymes influencing DNA methylation have now been described. G9a H3K9 histone methyltransferase for example interacts with DNA methyltransferases in the process of pluripotency gene silencing as well as in X inactivation (Epsztejn-Litman et al, 2008; Feldman et al, 2006). Suv39h1/2 histone methyltransferase on the other hand recruits DNMT3A/3B at satellite repeats. Consistently, knockout of Suv39h1 in ES cells results in reduced DNA methylation levels at centromeric satellite sequences (Lehnertz et al, 2003). This may be due to H3K9 methylation serving as a binding platform for HP1, which in turn recruits DNA methyltransferases (Fuks et al, 2003; Lehnertz et al, 2003). EZH2, a polycomb component has also been identified to interact with DNMT3A/B (Vire et al, 2006). Finally, genome DNA methylation patterns are also affected by histone acetylation. Inhibition of histone deacetylases by TSA results in reduced DNA methylation levels on an ectopically methylated sequence in HEK293 cells (Cervoni & Szyf, 2001).

Conversely, DNA methylation has also been shown to affect histone modifications. Maintenance methyltransferase DNMT1 for example is found to interact with PCNA and recruit G9a histone methyltransferase to replication foci. This is an example of potential maintenance of histone modifications through DNA methylation during replication, although in many cases it is still unclear whether DNA methylation or histone modifications are the causative marks of a silent chromatin state. Many studies have considered DNA methylation to be a long-term

silencing mark, established towards the end of a silencing process. A good example of this is during X chromosome inactivation. Repression of genes on the inactive X occurs independent of DNA methylation (Keohane et al, 1996; Lock et al, 1987; Sado et al, 2004; Wutz & Jaenisch, 2000). Similarly, DNA methylation is thought to be the secondary event during pluripotency gene silencing, occurring after the establishment of repressive histone marks as well as gene silencing (Feldman et al, 2006).

Additionally, crosstalk between DNA methylation, histone modifications and chromatin remodelling enzymes have also been shown to be essential for correct DNA methylation landscapes. To date, mutations in a few chromatin remodelling ATPases have been linked to diseases displaying DNA methylation defects. Mutations in the chromatin remodelling ATPase ATRX for example is linked to the disease α-thalassemia mental retardation (ATR-X) syndrome, which displays DNA hyper- and hypomethylation (Gibbons et al, 2000; Picketts et al, 1996). The crosstalk between chromatin remodelling, DNA methylation and histone modifications is supported by the interaction of various chromatin remodellers with histone modification enzymes as well as DNA methylation machinery or effector proteins. The mammalian NURD complexes for example contains histone deacetylase subunits, as well as methylated DNA binding proteins (Ho & Crabtree, 2010). Additionally, a novel putative chromatin remodelling ATPase (LSH) also forms a multisubunit complex together with DNMTs as well as HDACs (Myant & Stancheva, 2008).

1.6.2 Characterization of non-DNA methyltransferase proteins which affect DNA methylation

Consistent with the above described ways of crosstalk between histone modifications, DNA methylation and chromatin remodelling, absence of various epigenetic modifiers leads to extensive DNA hypomethylation in murine cell lines as well as embryos, similar or even greater than DNA hypomethylation seen in DNMT knockouts themselves (Table 1.1). These include chromatin remodellers (ATRX and LSH), methylated and unmethylated DNA binding proteins (UHRF1, CGBP), histone modification enzymes (G9a and LSD1) as well non-epigenetic proteins (SMCHD1, PGC7/STELLA, ZFP576). Apart from severe DNA methylation defects, absence of any of these proteins leads to early developmental defects, culminating in embryonic lethality (Table 1.1). Some of these proteins have been extensively studied, revealing their function in facilitating correct establishment and maintenance of DNA methylation. ZFP57 for example is required for correct maintenance of DNA methylation at imprinted regions established in the germ line (Li et al, 2008). UHRF1 has also been implicated in facilitating maintenance of DNA methylation, in association with G9a at replication forks (Zhang et al, 2011). CGBP and LSD1 are both implicated in regulating DNMT1 levels and stability respectively (Carlone et al, 2005; Wang et al, 2009).

Protein	Known function	Knockout phenotype	DNA methylation defect	References
CGBP	Transcriptional activator	Embryonic lethal after implantation	60% DNA hypomethylation (minor satellites /imprinted genes) Maintenance of DNA methylation defects	(Carlone et al, 2005; Carlone & Skalnik, 2001; Fisscher et al, 1996)
ZFP57	KRAB Zinc-finger transcription factor	Null embryos from maternal and zygotic knockouts lethal	Hypomethylation at maternal and paternal imprinted regions. Maintenance of DNA methylation defects	(Li et al, 2008; Quenneville et al, 2011; Zuo et al, 2012)
PGC7/ STELLA	SAP-like DNA binding motif Splicing factor motif	null embryos from Stella-/- females lethal at 4-cell stage	Hypomethylation at maternal and paternal imprinted regions before first cell division Maintenance of DNA methylation defects	(Jonkers et al, 2009; Nakamura et al, 2012; Sato et al, 2002)
UHRF1	Methylated DNA binding protein	Embryonic lethal		(Bostick et al, 2007; Sharif et al, 2007)
G9a	Euchromatic histone lysine methyltransferase	Embryonic lethal	50% DNA hypomethylation in ES cells (repetitive elements and single genes)	(Esteve et al, 2006; Ikegami et al, 2007)
LSD1	Histone lysine demethylase	Embryonic lethal	Progressive loss of DNA methylation in ES cells Maintenance of DNA methylation defects	(Ciccone et al, 2009; Deleris et al, 2010; Foster et al, 2010; Wang et al, 2009)
LSH	Putative SNF2 chromatin remodelling ATPase	Die after birth	60% loss of DNA methylation in embryos and cell lines (repetitive elements and single genes) *De novo* methylation defects	(Donohoe et al, 2009; Fan et al, 2005; Huang et al, 2004)
ATRX	Chromatin remodelling		Hypomethylation of X linked CpG islands	(Gibbons et al, 2000)

	ATPase			
SmcHD1	Structural maintenance of chromosomes	Female nulls embryonic lethal	Hypomethylation of X linked CpG islands	(Blewitt et al, 2008)

Table 1.1 Functions and knockout phenotypes of non-DNA methyltransferase proteins which affect DNA methylation

Studies into some of the above proteins have indicated the possible requirement of different proteins to facilitate DNA methylation at different genomic regions (e.g. ZFP57 at imprinted regions). Regulation of these proteins may therefore highlight a way to provide DNA methyltransferases with increased specificity to different regions at different stages of development. Despite this, few of these proteins have been studied to determine in detail their genome wide DNA methylation targets, especially in protein coding regions. In the long term, genome wide studies such as these will be important in confirming whether different proteins indeed provide a mechanism to target DNA methyltransferases to distinct genomic regions. In the long term, full characterization of protein coding gene target regions of such proteins may allow us to further elucidate how different genomic regions are targeted for DNA methylation or protected from it during various methylation reprograming events during early development, as well as in adult life.

My study therefore focused on further characterizing the role of three such proteins in DNA methylation. These included: Putative chromatin remodelling ATPase LSH whose mechanism of function in DNA methylation has not yet been fully characterized. Secondly, Smchd1, a structural maintenance of chromosomes domain containing protein which has recently been implicated in DNA methylation on the inactive X chromosome. Finally euchromatin histone methyltransferase G9a

which has been more extensively studied, but whose genome wide protein coding targets have not yet been fully characterized either. During my study I aimed to answer three broad questions. More specific questions will be outlined in each chapter.

1.7 Aims

1. At which gene promoter regions do LSH, G9a and SmcHD1 facilitate DNA methylation.

2. Do LSH, G9a and SmcHD1 facilitate DNA methylation at different or overlapping groups of promoters.

3. Are these proteins required for *de novo* or maintenance DNA methylation at target promoters?

Chapter 2 - Materials and Methods

2.1 Materials

2.1.1 Protein analysis Buffers

NE1: 20 mM Hepes pH 7.0, 10 mM KCl, 1 mM $MgCl_2$, 0.1% (v/v) Triton X-100, 20 % (v/v) glycerol. 0.5 mM DTT and complete protease inhibitors (Sigma) were added at time of use.

NE2: 20 mM Hepes pH 7.0, 420 mM NaCl, 10 mM KCl, 1 mM $MgCl_2$, 0.1 % (v/v) Triton X-100, 20 % (v/v) glycerol. 0.5 mM DTT and complete protease inhibitors (Sigma) were added at time of use.

Ponceau S staining solution: 1 % (v/v) glacial acetic acid, 0.5 % (w/v) Ponceau S

Coomassie Brilliant Blue R-250 staining solution: 30 % (v/v) methanol, 10 % (v/v) glacial acetic acid, 0.1 % (w/v) Coomassie Brilliant Blue R-250

SDS PAGE loading buffer (5×): 225 mM Tris-HCl pH 6.8, 50 % glycerol, 5 % SDS, 0.05 % bromophenol blue. 250 mM DTT was added at time of use

SDS Separating gel: 0.1 % SDS, 0.05 % ammonium persulfate, desired concentration of acrylamide and 375 mM Tris pH 8.8 and made up to 10 ml with dH_2O.

SDS Stacking gel: 0.1 % SDS, 0.05 % ammonium persulfate, 4 % Acrylamide, 125 mM Tris pH 6.8 and made up to 10 ml with dH_2O.

Transfer buffer: 25 mM Tris, 250 mM Glycine.

IPTG stock: 1 M IPTG was prepared in dH_2O, filtered through 0.2 um filter and stored at -20 degrees

Sodium Phosphate Buffer for MBD purification: 1 M sodium phosphate buffer (pH 8) was prepared by mixing 26.5 ml 2 M mono-sodium phosphate solution, 473.5 ml di-sodium phosphate solution and 500 µl H_2O.

Lysis/Binding Buffer for MBD purification: 50 mM Sodium Phosphate Buffer (pH8), 300 mM $NaCl_2$, 10 % Glycerol, 10 mM Imidazole, 0.5 mM PMSF

Wash Buffer for MBD purification: 50 mM Sodium Phosphate Buffer (pH 8), 300 mM $NaCl_2$, 10 % Glycerol, 20 mM Imidazole, 0.5 mM PMSF

Elution Buffer for MBD purification: 50 mM Sodium Phosphate Buffer (pH 8), 300 mM $NaCl_2$, 10% Glycerol, 250 mM Imidazole, 0.5 mM PMSF

Wash buffer 1 for MBD column packing: 20 mM HEPES, 0.1 M $NaCl_2$, 0.1 % Triton, 10 % Glycerol, 0.5 mM PMSF

Wash buffer 2 for MBD column packing: 20 mM HEPES, 0.1M $NaCl_2$, 0.1 % Triton, 10 % Glycerol, 0.5 mM PMSF, 10 mM Imidazole

MBD column run buffer A: 20 mM HEPES pH 7.5; 0.1 % Triton; 10 % Glycerol; 0.5 mM PMSF

MBD column run buffer B: 20 mM HEPES pH 7.5; 0.1 % Triton; 10 % Glycerol; 0.5 mM PMSF, 1 M $NaCl_2$

2.1.2 DNA analysis Buffers

Tris-EDTA (TE): 10 mM Tris-HCl pH 7.5, 1 mM EDTA pH 8.0

Church & Gilbert (modified) hybridisation solution: 44.5 g $Na_2HPO_4.H_2O$ and 2 ml H_3PO_4 were dissolved in total volume of 500 ml. 10 g BSA were added slowly to pre-warmed solution on stirrer. Solution was completed with H_2O to 1l. The pH was adjusted to 7.2-7.4.

Southern Blot Wash Buffer I: 80 mM Na_2HPO_4 pH 7.5, 5 % SDS

Southern Blot Wash Buffer II: 80 mM Na_2HPO_4 pH 7.5, 1 % SDS

SSC buffer: 175.3 g $NaCl_2$, 88.2 g $Na_3C_6H_5O_7$ were dissolved in 800 ml water. pH was adjusted to pH 7 with HCl; Volume was added up to 1 l with water.

Bisulfite conversion solution: 3.8 g sodium bisulfite dissolved in 5 ml H_2O and 1.5 ml fresh 2 M NaOH. This solution was resuspended in the dark. 110 mg hydrochinone was dissolved in 1 ml water at 55 degrees for 10 mins and subsequently added to the sodium bisulfite solution.

2.1.3 Primers

All primers were ordered from MWG Eurofins. Lyophilised primers were resuspended in dH_2O to 100 uM stock concentration and stored at -20 degrees.

Gene	Primers
Hoxa5 (RT PCR)	AGC CAC AAA TCA AGC ACA CA
	AGA TCC ATG CCA TTG TAG CC
Hoxa6 (RT PCR)	TAC GGG GCC TCA TGT TTC TA
	CTG CAT CCA GGG GTA AAC AG
Hoxa7 (RT PCR)	CAC ATG GTT CCA GTT TGT GG
	CTG AAC CCC AGA TCC ACT GT
Hoxa10 (RT PCR)	GGA AGC ATG GAC ATT CAG GT
	CCA GGC AAG CAA GAC CTT AG
Hoxa11 (RT PCR)	CCC CTG GTG GTT CAC TCT TA
	CTT GGG GCA CAC AGT TTC TT
Rhox2 (RT PCR)	CAA GCT TGA TGT GGG ACC TGA G
	GAA TCA GGG ACT GGC TCC TCT G
Rhox6 (RT PCR)	AAA CCC ACA TGT TCT GAA TAG G
	GCC ATC ATG GGT GGC TCA GAA G
Rhox9 (RT PCR)	CAG CCA TTC CTG AGG GCA TGA
	GCC TGC TGT GTC TGC GGA AA
Rhox11 (RT PCR)	TGG AGG AGG ACG AGT GTC TGT G
	TGG TGG CCC TCT TTC GGG ATG
C21_Rik (RT PCR)	AGC GAG GAC GAG AAT GAA CC
	ATG CTC AGG AGC TCG GAT AG
Cul4b (RT PCR)	GTG CAG GCA ACA AAG AAG CTA C
	GCA CAT ATG TTG GCC AGT AAC C

Gm9 (RT PCR)	GGC AAA CAC AAG GGC TAT GG
	ACC GGA GGT AAA GGT CGT TG
Actin (RT PCR)	GGC CCA GAG CAA GAG AGG TAT
	ACG CAC GAT TTC CCT CTC AGC

qPCRs (on MBD column fractions)

Igf2/H19	**GGTGGCAGCATACTCCTATAT**
	CTCGGCAACTTCGGTCTTAC
Igf2/H19	CATAGGGGTGGTAAGATGTGTG
	CATAGGTGAACCGCAATTTGG
Igf2R	CGTGATCCTTGGTTGTGCTGAG
	GTTTTAATGCGATTCCGGTTGG
Dlk1	AGCGGCAGTGAGAATGAGAT
	TAAAGGCAAGACGAATCACAAGA
Zdbf2	GAAGTGAACGGAAGTGTCGAA
	AATCACAGCTCCCAATCACC
GnasXL	CACTGAGACCTGCGTCCTCT
	TGGCGGCCAACAACTTTAG
Gnas ex.1	GTCATCGTCGCATAGTTTCCA
Sgce/Peg10	TCCTGACCAACTACGACCTG
	CCATACTCACCACACGAGGA
Snrpn	CAGGACATTCCGGTCAGAG
	TACTAGAATCCACAAGCCCAG
Kcnq1ot1	CAGCACGGATCACTCCAG
	AAAGCTCTCCAAGTAGAATCACA
Zac1	GCATCTGCGATTTGTCACTC
	CTTGCTCTCCAGTCCCGATA
Peg3	CAGAGGACCCTGACAAGGAG
	AGCACAGCACTCTACGCACA
Nnat	GCTGGCATGGGTTTAAAATTCAC
	ATTCTACATGCCGCGTAACCAG
Rasgrf1	CTGCACTTCGCTACCGGTTTC
	AGTAGCAGTCGTGGTAGTTG
Peg13	AGCTGAGCGAACCCCTTTAC
	CGCAGGTCTTCTATCCAACC
Grb10	GTAGCCGTAGGATAGAATTATGTAAC
	CTTACAGCCCAAGCTGGTTC
Ankrd50	TCTCGTCCAAGCCTCTGTC
	GTCGATCACACCGATAACAAC

qPCRs (on cDNA)

Oct4	**CCC ACT TCA CCA CAC TCT AC**
	CCA AAG CTC CAG GTT CTC TT
Dppa2	GCA TTC ATT CAG CGG CTG CCT TT

	TGC GTA GCG TAG TCT GTG TTT GG
Dppa3	GAA CCG CAT TGC AGC CGT ACC T
	GGC TCA CTG TCC CGT TCA AAC TCA
Dppa4	CAA GGG CTT TCC CAG AAC AAA TGC
	GCA GGT ATC TGC TCC TCT GGC AC
GAPDH	GGC TCA TGA CCA CAG TCC ATG CC
	CAC GGA AGG CCA TGC CAG TGA G
Dlk1	TGCGTGATCAATGGTTCTCCCTG
	TCACAGAAGTTGCCTGAGAAGCC
Rtl1	AGCGGTCACCTTCAGACCAC
	GGGCCACCTGATCTCTGCTA
RIAN	CGTGGATTGACACTGTGATGGATTC
	CAAAAGAACGTTGGCTATACTCACCTG
Dio3	GGATGGACTTCGAGACCAAAGG
	CCAAGTTCTTCCCCTTCTCCTTG
Dio3as	CGTAAAGTCAAGGCCTCAACC
	GGAGCTGTAGTATCCGGCTG
Kcnq1	GTGGAAGACAAGGTGACACACTG
	TGGCTACAACTTGGACCTGCTG
Kcnq1ot1	ACACTCAGAACTCGTTGCAATTCAC
	TTGTCTGCCTCTGAAGTGCTAAG
Cdkn1c	CAACGTCTGAGATAGTTAGTTTAGAG
	GTTCCTGCTACATAACGAAAGGTC
Osbpl5	GACTTTGTTCCTGGTGGTGAGAG
	TGTGGAGCTTAACATGGTCGAAGG
Phlda2	CGCTCTGAGTCTGAATGCCTTC
	CTCCTGGGCTCCTGTCTGATG
Gtl2	GCGAGGGAGATAGAGAGTAAC
	CGCGCTCCTAGACTTCATTGT
Snrpn/Snurf	AGGAATTAGAGGCCCACTCC
	AAAGCTTGCAGGTACACAATTTCAC
LSH	TGCTGCACCAGATTTTAACGCCTTT
	CATGTTTGCAATTGTGCGGTTCAC

Bisulfite primers

Rhox2	**TTTTTAGTGAATTGTTATTTTTTTT**
	TTTTTTATTGTTTTGAGTTTTGTAG
	ATCCCACATCAAACTTATAATTAAC
Rhox6/9	AGAGGAAGTATAGGTAGGGGTTTTG
	AAAATACAAATATCCCCCAAACAAC
	CAAACAACACCTTAAAACCAAAAAT
Oct4	TTTGAAGGTTGAAAATGAAGGTTTT
	CATCACCCCCACCTAATAAAAATAA
	CAACCATAAAAAAAATAAACACCCC
Ndn	TTAGTGGTTGGGTTTTG

	ATATAGGAGATTAGGAAAT
	CTATCCTACATCTCACAA T
Peg12	GGTGGTAATGATTAGAATA
	TGTTGTAGTAGGAGGAAGTTGT
	CTT AAC CTC CAC CCA ACC TAA A
IG DMR	TTAAGGTATTTTTTATTGATAAAATAATGTAGTTT
	TTAGGAGTTAAGGAAAAGAAAGAAATAGTATAGT
	TATACACAAAATATATCTATATAACACCATACAA
	CCTACTCTATAATACCCTATATAATTATACCATAA
KvDMR	TTAGGTTTATAGAAGTAGGGGTGGT
	CTACAAAACTCAAAAATCTCCAAAC
Igf2/H19	GGTTTTTTGGTTATTGAATTTTAAAAATTAG
	TTAGTGTGGTTTATTATAGGAAGGTATAGAAGT
	TAAACCTAAAATACTCAAACTTTATCACAAC
	AAAAACCATTCCCTAAAATATCACAAATACC
Uty	TTAGTTGTATTGGTGATG
	TTTAGGGAGAGGTTATGG
	ACAAAATACAATCAAACC
Slitrk1	TGTATTATGTGTTTGGATAGTTTGTG
	TATTGGAGTATTAAATGAATTTGAATTG
	ACCAACTACAACCCCAAAAATAC
Cdkn1c	AGGATTTAGTTGGTAGTAGT (F)
	TTTTCAATTTCAACAACACC (R)
	TATCCTATCCAACTTAAACC (R)

2.1.4 Antibodies

LSH (Mouse monoclonal Santa Cruz sc-46665/H0205): 1:1000 (1:500 for Immunohistochemistry) working concentration in 2 % milk + 0.1 % Tween

HDAC (Rabbit polyclonal Santa Cruz sc-7872 / D171): 1:1000 working concentration in 2-4 % milk + 0.1 % Tween

MRE11: 1:1000 working concentration in 2 % milk + 0.1 % Tween

G9a (Mouse monoclonal Perseus Proteomics): 1:500-1:1000 working concentration in 2-4 % milk + 0.1 % Tween

FLAG M2 (Mouse monoclonal Sigma): 1:2000 working concentration in 2-4 % milk + 0.1 % Tween

Licor secondary antibodies: 1:10,000 working concentration

Alexa Dye secondary antibody for Immunohistochemistry: 1:1000 working concentration.

2.1.5 Cell Lines

Lsh-/- MEFs: Isolated from *Lsh-/-* embryos derived from crosses of heterozygous animals (*Kathrin Muegge*)

G9a-/- ES: Gene conversion of heterozygous G9a+/- ES cells through stringent antibiotic selection. This can cause chromosome loss and duplication *(Yoichi Shinkai)*

Smchd1 mutant MEFs: Cell lines homozygous for non-sense mutation in exon 23 of *Smchd1*, resulting in nonsense mediated RNA decay and a null allele (*Emma Whitelaw).*

2.2 Methods

2.2.1 DNA manipulations

Plasmid Preps: 5 ml culture plasmid preps were carried out using the Quiagen Miniprep Kit. 400 ml culture plasmid preps were carried out using the Quiagen Maxiprep Kit. DNA was eluted in desired volume of Elution Buffer.

Restriction digestion: Restriction digests were carried out in 20 or 50 ul volumes according to manufacturer's recommendations (up to 500 ul volume was used for

gDNA digestions for use in Southern Blotting). Incubation times ranged from 1 h to o/n at 37 °C. Completion of digestion was assessed by agarose gel electrophoresis.

Gel extraction: DNA fragments were excised from agarose gels and extracted using the Quiagen or Fermentas gel extraction kits.

DNA sequencing: Big Dye Reactions were set up in a total volume of 10 ul, including 2 ul BigDye, 1 uM primer, and 500 ng of template DNA or 3.5 ul of PCR product. The following program was used for the sequencing reactions: 96 °C for 1 minute and 25 cycles of cycling between 96 °C for 10 seconds, 50 °C for 5 seconds, and 60 °C for 4 min. The Gene Pool Sanger Sequencing Service at University of Edinburgh was used for Sanger Sequencing.

gDNA extraction: Cell pellets from T175 flasks were resuspended in 4 ml TE buffer. Proteinase K was added to final concentration of 200 ug/ml and SDS added to a final concentration of 1 % o/n at 55 degrees. 100 ug/ml RNAse was added for 1h at 37 °C. The salt concentration was subsequently adjusted to 200 mM with NaCl. Digested peptides were removed through one phenol:chloroform:isoamyl alcohol and one chloroform extraction. DNA was precipitated in 1 volume isopropanol and 1/10-1/5 volume Sodium Acetate (3 M, pH 5.3). DNA was pelleted, washed in Ethanol and resuspended in a suitable volume of TE buffer or water.

Bisulfite conversion of gDNA: 1 ug of gDNA was resuspended in a total volume of 25 ul. To denature, DNA was boiled at 95 °C for 5 mins and chilled immediately on ice. 2.5 ul of fresh NaOH (3 M) was added and samples incubated at 37 °C for 20 minutes. 270 ul sodium bisulfite solution (see buffers section) was added to each DNA sample. Samples were overlaid with mineral oil to prevent evaporation and

incubated o/n for sulfonation at 55 °C. DNA was precipitated in isopropanol and sodium acetate in presence of 2.5 ug glycogen. DNA pellets were resuspended in 25 ul TE buffer. DNA was then denatured by addition of 2.5 ul 3 M NaOH and incubation at 37 °C for 15 mins. Samples were then desulfonated and precipitated by addition of 32.5 ul 5M ammonium acetated (pH 7) and 180ul ethanol. DNA pellet was then resuspended in 35 ul TE buffer.

Methylated DNA affinity purification (MAP): 65 mg of purified His tagged MBD domain (see protein methods) of MeCP2 was bound to 1 ml of nickel coated Chelating Sepharose Fast Flow beads (GE Healthcare). MBD bound beads were packed into a 1 ml Tricorn column (GE healthcare) which was connected to the AKTA liquid chromatography machine. The column was washed using a 0.1 M to 1 M $NaCl_2$ gradient (mixture of column run buffers A and B) three times to eliminate any bacterial DNA carried over from MBD peptide expression and purification from *E.coli*. Genomic DNA was sonicated to an average size of ~300 bp and loaded onto the MBD column . Unmethylated and methylated DNA fragments were eluted at low (up to 0.715 M) and high (1 M) NaCl concentrations. The success of each run was tested by actin and Igf2R-DMR2 PCRs (see primer tables) as unmethylated and differentially methylated controls respectively. Conditions used were: 95 °C– 5 min; 94 °C – 30 s; 60 °C – 1 min; 72 °C – 1min (25 cycles), 72 °C – 10 min). Methylated DNA fractions were pooled, concentrated and de-salted in 30 K filters (Amicon Ultra Centrifugal Filters, Millipore). Samples were then ethanol precipitated and resuspended in 20 ul TE buffer.

Whole genome Amplification: Whole genome amplification of methylated DNA from Methylated DNA Affinity purification was carried out using the Sigma

GenomePlex Genome Amplification Kit. This kit is based on attachment of linker DNA to the ends of the sample DNA and subsequent amplification with linker specific primers. Samples were purified using Quiagen PCR purification kit.

Whole genome methylation analysis: Whole genome methylation levels were determined in different cell lines using the Sigma Imprint Methylated DNA Quantification Kit.

Southern Blotting: 15 ug Genomic DNA was digested with methylation sensitive (HpaII) and insensitive (MspI) restriction enzymes overnight in a total volume of 300-500 ul. Digested DNA was precipitated in isopropanol and resuspended in 25 ul volume and run on a 1 % agarose gel for 8 h. DNA was transferred onto a Nylon membrane overnight in 0.4 M NaOH. The membrane was washed in 2 x SSC buffer. 100 ng of probe was radioactively labelled using the Fermentas DecaLabel DNA Labelling Kit and then purified using GE Healthcare Illustra ProbeQuant G-50 Micro Columns. The membrane was incubated in pre-warmed (65 °C) Church & Gilbert (modified) hybridisation solution for 1h. The probe was added to the membrane in hybridisation solution and incubated at 65 °C o/n on rotor. After hybridisation the membrane was washed for an hour with Wash solution I and half an hour with Wash solution II.

2.2.2 RNA methods

RNA extraction: Cell pellets from confluent T175 flask were resuspended in 1 ml Trizol reagent and 200 ul chloroform added. Samples were centrifuged at 5500 rpm at 4 °C for 15 mins. Top aqueous phase was removed and precipitated with 500 ul isopropanol per ml Trizol used. Precipitated samples were spun at 13000 rpm for 15

min at 4 °C. RNA pellets were washed in cold 75 % ethanol and air dried for 5 mins. RNA was resuspended in RNAse free water and 1 ul RiboLock RNAse inhibitors added.

50 ul RNA was DNAse treated for 1 h at 37 °C. DNAse was inactivated at 90 °C for 5 mins.

cDNA synthesis (for qPCR): 1-5 ug RNA were used in reverse transcription reaction. 1 ul oligo dT primers were added and reaction made up to 11 ul with water. Samples were heated at 65 °C for 5 mins and then cooled on ice. A mastermix was prepared with 1 ul dNTPs (10 mM), 4 ul 5x FS buffer; 2 ul DTT (0.1 M) and 1 ul RNAse OUT (40 U/ul) per sample and 8 ul added to each reaction. Samples were mixed and heated at 42 °C for 2 min and Superscript II added to RT + reactions. The reverse transcription was carried out at 42 °C for 2 h.

Expression Array normalization: See (Myant & Stancheva, 2008) for details

Quantitative PCR reaction mix: 10 ul SYBR qPCR mix, 0.25 uM primers were added to 2 ul diluted cDNA and added up to a total volume of 20 ul.

2.2.3 Protein methods

MBD protein purification: 1 ul of His tagged MBD plasmid was added to 100 ul Bl21 cells and left for 30 min on ice. Cells were heat shocked for 45 s at 42 °C and returned to ice for 2 min. 200 ul LB media was added and cells left to shake for 30 min at 37 °C, before plating on 50 µg/ml Kanamycin plates. Next day, 200 ml kanamycin media was inoculated with a single colony o/n. The o/n culture was diluted 1:50 in kanamycin media and grown to OD 600 of 0.45-0.6 before inducing

MBD expression with 1 mM IPTG for 3-4 h. The induced culture was spun down for 1 h at 4 °C at 4200 rpm. The supernatant was decanted and pellets frozen at – 80 °C o/n to facilitate cell lysis. Cell pellets were resuspended in 30 ml lysis buffer and lysed at 25 KPSI in Constant Systems TS cell Disruptor. Triton was added to 0.1 % and cell lysates spun 30 min at 4 °C at 13000 rpm. Fractogel beads were used for Nickel purification of MBD protein. Beads were washed with 20 ml water per 1 ml bead volume. 10 ml 0.1 M $NiSO_4$ was added per ml of beads and left on rotor for 15-20 min. Beads were washed with 40 ml water and subsequently 20 ml lysis buffer per 1 ml bed volume. The relevant volume of beads were resuspended in cell lysate and left on rotor for 1-2 h. The beads were washed with at least 1 l wash buffer. Then a bed volume of elution buffer was added to the beads and incubated for 10 min. The elution buffer was drained through the beads and collected. 5 more elutions were carried out. The elutions were stored at 4 °C while being assessed by SDS PAGE. Elutions containing most MBD protein and least contaminating proteins were pooled and dialysed for 2 h and then o/n in 2 l lysis buffer without imidazole. Dialysis product was centrifuged at 13000 rpm for 20 min to remove precipitate and Bradford assay was carried out to quantify protein. Dialysed product was stored at 4 °C short term or in 50 % Glycerol at -20 °C long term. 40-60 mg MBD protein was bound to 1 ml of nickel coated beads for 1-2 h and then washed with 20 ml wash buffer 1 , 20 ml wash buffer 2 and 20 ml wash buffer 1 again per ml of beads.

Nuclear extracts: 1 ml of NE1 Buffer was added to cells and homogenized in Dounce Homogenizer. The cell suspension was transferred to 2 ml eppendorf tubes and centrifuged at 3000 rpm for 5 mins at 4 °C to collect nuclei. The nuclear pellet was resuspended in 100 ul NE1 and homogenized with a pipette tip. 250 U of

Benzonase (Merck) was added on ice for 1 h. $NaCl_2$ was added to 420 mM. Tubes were placed on spinning wheel for 20 min at 4 °C. Samples were then centrifuged at 13000 rpm for 15 min at 4 °C. The supernatant contained the nuclear extract.

Protein quantification: Bicinchoninic Acid Protein Assay Kit (Sigma) was used for protein quantification. 0.04 absorbance units are approximately equivalent to 1ug protein. Quantified protein extracts were stored at -80 °C.

SDS gel: Separating gel was prepared and poured first, followed by the stacking gel. 5 ul of sample buffer was added to 44-50 ug of protein, boiled for 5 min and run at 270 V on prepared SDS gel.

Transfer onto Nitrocellulose membrane: Protein from the SDS gel was transferred onto Nitrocellulose membrane for 1 h in Tris/Glycine transfer buffer.

Western Blotting: Membranes were blocked with milk in 1 x TBS with 0.1 % Tween for 1 h at RT. Membranes were incubated o/n at 4 °C in fresh blocking solution with desired concentration of primary antibody. Membranes were then washed three times 20 mins each with 1 x PBS + 0.1 % Tween and then blocked again for 1 h as before. Secondary antibody was added to desired concentration (see materials for concentrations) and left to incubate for 1-3 h depending on the antibody. Membranes were washed as before to remove secondary antibody and scanned on Licor scanner.

Immunohistochemistry: Coverslips were washed in HCL for 15 mins, then dH2O and sterilized in ethanol. Cells were plated at a density of $1\text{-}2x10^5$ per well of a 6-well plate containing one coverslip each. Cells were left to adhere overnight. Media was removed and cells washed with PBS. Washed cells were fixed with 0.04 % w/v

paraformaldehyde for 5 min at RT. Fixing solution was removed and cells washed twice in PBS. 1 ml PBS with 0.1 % Triton was added for 15 min at 37 degrees. Cells were blocked for 1 h in blocking reagent (4 % BSA in PBS) on rocking platform. Blocking solution was changed and correct concentration of primary antibody added and incubated for 1 h on rocking platform. Primary antibody was removed and cells washed three times 5 min with PBS. 1:1000 dilution of secondary antibody (Alexa Dye) in fresh blocking solution was added and incubated for 1h on rocking platform. Secondary antibody was removed and cells washed three times 10 min with PBS. Coverslips were then mounted on microscope slides using Prolong Gold (containing DAPI dye).

2.2.4 Mammalian Cell Culture

MEFs, NIH3T3 and Phoenix cells: Cultured in DMEM + 10 % FBS and PSG

ES cells: Cultured on gelatinized culture flasks in MEM, 10 % FBS, 500 ul 2-mercaptoethanol, 5 ml sodium pyruvate, 5 ml PSG, 5 ml non-essential amino acids and 1:1000 dilution of LIF.

ES cell differentiation: ES cells were cultured for 4 days before starting differentiation. LIF was withdrawn from culture on day 0. Media was changed on Day 2. Retinoic Acid was added to medium on Day 4 of differentiation and media changed with RA on Day 6. Embryoid Bodies were collected on Day 8 and dissociated with trypsin and plated in N2 media at a density of 1-2 x 10^5 cells/cm2 – 1.57 x 10^7 cells/plate. N2 media was changed 2 h after plating and the next day. N2 media was changed to complete neuronal media after 2 days. (Bibel et al, 2004)

Transient transfection: Transient transfections of 10 ug plasmid DNA were carried out using the Polyplus Jetpei or Jetprime reagents on 1 x 10^6 cells in 10 cm dishes. Cells were harvested after 48 h.

Viral packaging and infection: 3.5 x 10^6 PhoenixA cells were seeded for each packaging reaction on 10 cm mammalian culture dishes. 8 ug of DNA was transfected into Phoenix cells using the Polyplus JetPei transfection kit. Transfected cells were left over night with 8 ml media and then media changed to 5 ml. The media containing lentivirus was harvested 48 h and 72 h after transfection, filtered through 0.4 um filter and frozen at -80 °C.

Chapter 3 - Characterization of LSH-dependent DNA methylation at gene promoters

3.1 Introduction

3.1.1 A putative chromatin remodelling ATPase implicated in *Arabidopsis thaliana* DNA methylation

Screening in *Arabidopsis thaliana* for DNA methylation mutants (specifically those showing genome hypomethylation) identified Decrease in DNA methylation 1 (DDM1). DDM1-deficient plants show a 70% reduction in DNA methylation levels with hypomethylation seen at repetitive elements but also at some single copy genes (Navarro et al, 2008; Vongs et al, 1993). Methylation defects are accompanied by gene expression defects with silenced reporter genes for example undergoing reactivation in *Ddm1* mutants. The possibilities of *Ddm1* encoding a DNA methyltransferase or reducing adeno-methionine levels were eliminated early on (Kakutani et al, 1995). Instead, protein sequence similarities suggested DDM1 being a novel SNF2 chromatin remodelling ATPase (Vongs et al, 1993). This is demonstrated in the ClustalW protein alignment of *Arabidopsis thaliana* DDM1 and chromatin remodelling ATPase BRM, displaying highly conserved SNF2 as well as Helicase domains, characteristic of SNF2 chromatin remodelling ATPases (Figure 3.1A).

ATP-dependent nucleosome repositioning activity of DDM1 was later confirmed in an independent study (Brzeski & Jerzmanowski, 2003). DDM1 was therefore the first chromatin remodelling enzyme of the SNF2 family identified to participate in DNA methylation (Jeddeloh et al, 1999).

Furthermore, co-operation between DDM1 and histone methyltransferases leading to repressive H3K9 methylation has been observed (Gendrel et al, 2002).When DDM1 is deleted, this histone mark is replaced by the activating H3K4 methyl mark (Gendrel et al, 2002). Mutant phenotypes of *Ddm1* in *Arabidopsis* include morphological developmental abnormalities, meiotic defects as well as defects in DNA damage response (Melamed-Bessudo & Levy, 2012; Mittelsten Scheid & Paszkowski, 2000; Shaked et al, 2006).

3.1.2 Mammalian homolog of DDM1 is also required for DNA methylation

The mammalian homolog of DDM1, Lymphoid-specific helicase LSH, was identified in early T cells in an attempt to find a helicase involved in V(D)J recombination (Jarvis et al, 1996). *Lsh* was found to be expressed in thymus, early lymphocytes and activated lymphocytes, but is now known to be ubiquitously expressed (Jarvis et al, 1996). The sequence conservation between DDM1, murine and human LSH as well as murine SNF2H chromatin remodelling ATPase (Figure 3.1B) shows high sequence conservation of Coiled-coil, SNF2 as well as Helicase domains. A schematic of LSH protein structure can be seen in Figure 3.1C. Even though sequence conservation with known chromatin remodelling ATPases is evident, significant ATPase and chromatin remodelling activities of LSH are yet to be identified.

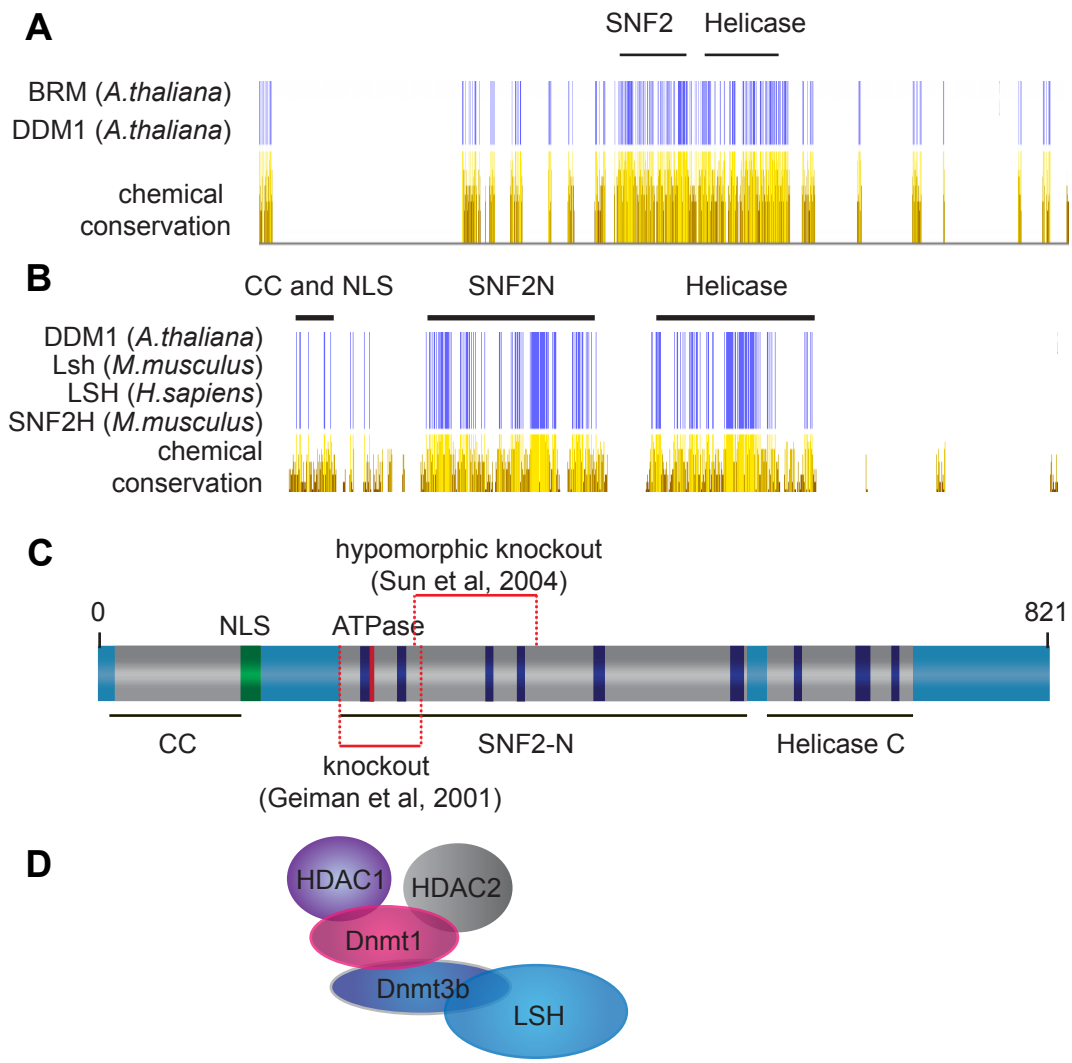


Figure 3.1 Chromatin remodelling ATPases DDM1 (A.thaliana) and LSH (M.musculus)

A: ClustalW alignment of *Arabidopsis thaliana* DDM1 with chromatin remodelling ATPase BRM1. Conservation patterns were visualized using JalView. Top pannel in blue indicates conserved amino acid residues between all proteins. Bottom pannel indicates conservation of chemical properties of amino acid residues with a white – brown – yellow colour gradient of increasing conservation. Despite amino acid differences between proteins, functional conservation may be achieved through the presence of chemically similar amino acids. The location of conserved domains is indicated above the alignment. **B:** ClustalW alignment of Arabidopsis thaliana DDM1 with human and murine LSH as well as murine chromatin remodelling ATPase SNF2H. **C:** Domain structure of murine LSH, drawn to scale. Indicated are the Coiled-coil domain (CC), Nuclear Localization signal (NLS), ATPase site (ATPase), SNF2 N termain domain (SNF2-N) and the Helicase C terminal domain (Helicase C). Indicated on the diagram are the locations of insertions of knockout cassettes leading to full knockout (Geiman et al. 2001) as well as to a hypomorphic knockout (Sun et al. 2004). **D:** Mapped protein interactions by Myant and Stancheva 2008. Note: Only the interaction between LSH and DNMT3B is direct.

Several studies have showed that LSH is required for mammalian cell proliferation and development (Fan et al, 2003; Raabe et al, 2001; Sun et al, 2004). Although *Lsh-/-* mice occur at a normal frequency at birth, they die within a few days (Geiman et al, 2001). These mice display normal breathing and exterior appearance but their body weight was reduced by 22% compared to heterozygous littermates. Newborns also suffer severe renal lesions, but other organs are similar to those of control animals (Geiman et al, 2001). Interestingly, hypomorphic knockout mice in which a truncated protein is detectable survive into adulthood but show signs of premature ageing. Importantly, *Lsh* alterations have been implicated in diseases such as Leukaemia, making the study of LSH an important step towards understanding its function in disease (Fan et al, 2008; Lee et al, 2000).

3.1.3 LSH regulates gene expression through DNA methylation and histone modifications

Similar to the plant homolog DDM1, LSH is involved in DNA methylation. Knockout of *Lsh* causes up to 50% loss of global DNA methylation in cell lines as well as embryos (Donohoe et al, 2009; Tao et al, 2011). More specifically, LSH is now thought to function in DNA *de novo* methylation rather than maintenance of DNA methylation (Zhu et al, 2006). Despite this, one study identified progressive loss of DNA methylation from some repetitive elements upon knockdown of LSH (Suzuki et al, 2008). This may suggest a role of LSH in maintenance of DNA methylation. Similarly to DDM1, LSH deficiency also causes changes in histone modification patterns. H3K4 methylation becomes more abundant at pericentromeric DNA but H3K9 methylation seems to remain unaltered (Tao et al, 2011; Yan et al,

2003b). LSH Immunochemistry after mild extraction with detergents has shown association of LSH with mouse centromeric regions. The same study, as well as co-immunoprecipitation assays, demonstrate colocalization and association of LSH with Heterochromatin Protein 1 (HP1) (Yan et al, 2003a). The above studies suggest that LSH might be essentially for establishment of pericentromeric heterochromatin. Consistent with this inference, LSH has been implicated in meiotic chromosome synapsis, mitosis and silencing of retrotransposons as well as repetitive elements, which all require functional centromeres (De La Fuente et al, 2006; Huang et al, 2004).

There are four previous studies investigating the effects of LSH on DNA methylation and gene expression at single genes (Donohoe et al, 2009; Fan et al, 2005; Huang et al, 2004; Xi et al, 2007). Huang et al investigated LSH targets within the genome and, as previously, identified repetitive elements. RT-PCR showed no difference in single gene expression of *Igf2R*, *Aprt*, *β-actin*, *H19* or *Pgk-1* between WT and *Lsh-/-* embryos (Huang et al, 2004). The same study also carried out an embryonic cDNA microarray experiment. Only ~ 0.5% cDNAs showed more than 2-fold changes in gene expression in liver and brain of *Lsh-/-* mice (Huang et al, 2004). One of the authors' explanations for the few single, non-repeat genes showing expression alterations, was close proximity to repeat elements targeted by LSH, leading to genes in close proximity being subjected to the same repressive or activating state (Huang et al, 2004). The study by Huang et al. also identified that *Lsh* knockout may cause changes in single gene expression, due to changes in cell composition of tissues. *Lsh* knockout leads to impaired lymphoid development, which may cause reduced levels of certain lymphoid cells in tissues analysed, which

may cause the single gene expression changes observed (Huang et al, 2004). The study by Dennis et al, on the other hand, identified various single genes showing hypomethylation. These included *H19* and *Pgk-1*, suggesting that LSH may be required for DNA methylation. This study was closely followed by analysis of the effect of LSH on imprinted genes, identifying increased expression and reduced CpG methylation of the differentially methylated region of *Cdkn1* in *Lsh-/-* cells (Fan et al, 2005). Other imprinted genes tested in this study were shown not to be affected (Fan et al, 2005). The fourth study showed the importance of LSH in developmental regulation of specific genes (Xi et al, 2007). Most homeobox genes (*Hox5*, *Hox6* and *Hox7*) were found to show increases in expression as well as reduced DNA methylation at their promoters upon knockout of *Lsh* (Xi et al, 2007). Together, all these studies highlight the need for LSH in gene expression and DNA methylation of some single genes. Despite this, the extent of requirement for LSH outwith repeat sequences is not yet clear, as all studies focused on a select few genes.

3.1.4 Potential mechanisms of LSH-dependent DNA methylation

Previous work in our lab indicated that the native LSH exists mostly as a monomeric protein in nuclear extracts of mammalian cells, unlike other chromatin remodelling ATPases which are associated with large multisubunit complexes (Myant & Stancheva, 2008). Luciferase reporter gene assays combined with histone deacetylase inhibitors (TSA) on the other hand identified the need for histone deacetylases (HDACs) for efficient gene repression. Consistently, depletion of *Hdac1* and *Hdac2* alleviated the ability of LSH to silence the reporter gene. Subsequent co-immunoprecipitation experiments identified the requirement of the N-terminal coiled-coil domain of LSH for interaction with HDAC1 and HDAC2 as well

as DNMT1 and DNMT3B. The interaction with HDACs and DNMT1 was dependent on DNMT3B, suggesting a direct interaction of DNMT3B with LSH (Myant & Stancheva, 2008) (Figure 3.1D). A standing hypothesis is that LSH induces localized accumulation of DNMT3B and other chromatin modifiers upon binding to DNA/chromatin. DNA methylation could result in initiation or further stabilization of a repressive chromatin state established at LSH bound regions (Myant & Stancheva, 2008). Consistent with this model, studies characterizing LSH-mediated formation of pericentromeric heterochromatin identified association of LSH with various repressive epigenetic regulators. LSH for example associates with HP1, the major heterochromatin regulator (Yan et al, 2003b). Another study showed reduced levels of H2A-K119 ubiquitynation at *Hox* genes as well as reduced association of PRC1 components at these genes in *Lsh-/-* MEFs. This suggests that LSH is required for correct assembly of PRC2 complex as well as recruitment of PRC1 complex to chromatin (Xi et al, 2007).

In summary, studies to date have extensively investigated the role of LSH in DNA methylation and silencing of transcription at repetitive elements and have pinpointed that DNA methylation and gene expression of a few single genes are dependent on LSH. Additionally, several hypotheses have been suggested for the function of LSH at repetitive genomic regions as well as single genes. However, it is yet unknown what the extent of LSH function in single gene DNA methylation and gene expression genome wide is. As discussed in the introduction, several non-DNA methyltransferase proteins have been identified to be crucial for correct establishment and maintenance of DNA methylation. Still not well defined is, whether these proteins target different or overlapping regions genome wide. This will

be essential in determining whether these proteins can aid the recruitment of DNMTs to correct regions at different stages of development. I therefore set out to identify genome-wide the promoters of protein coding genes, which require LSH for correct DNA methylation. Secondly I set out to determine whether DNA methylation established by LSH is required for correct expression of associated genes. All results discussed in this chapter have been published in Genome Research (Myant et al, 2011).

3.2 Results

3.2.1 Methylated DNA affinity purification and microarray analysis

Previous work in the lab applied a methylated DNA affinity purification technique (Illingworth et al, 2008), followed by hybridization to mouse promoter microarrays to define genome wide differences in DNA methylation between WT and *Lsh-/-* MEFs (Figure 3.2) (Myant et al, 2011). WT and *Lsh-/-* MEFs were isolated from littermate embryos in Kathrin Muegge lab. Knockout mice were obtained though crossing heterozygous mice for the knockout allele, meaning that WT LSH would have been present during germ cell determination in the parents. Any DNA methylation defects observed in *Lsh-/-* MEFs can therefore only be attributed to defects in maintenance of DNA methylation patterns established in the germline or to de novo methylation defects at later stages of embryo development.

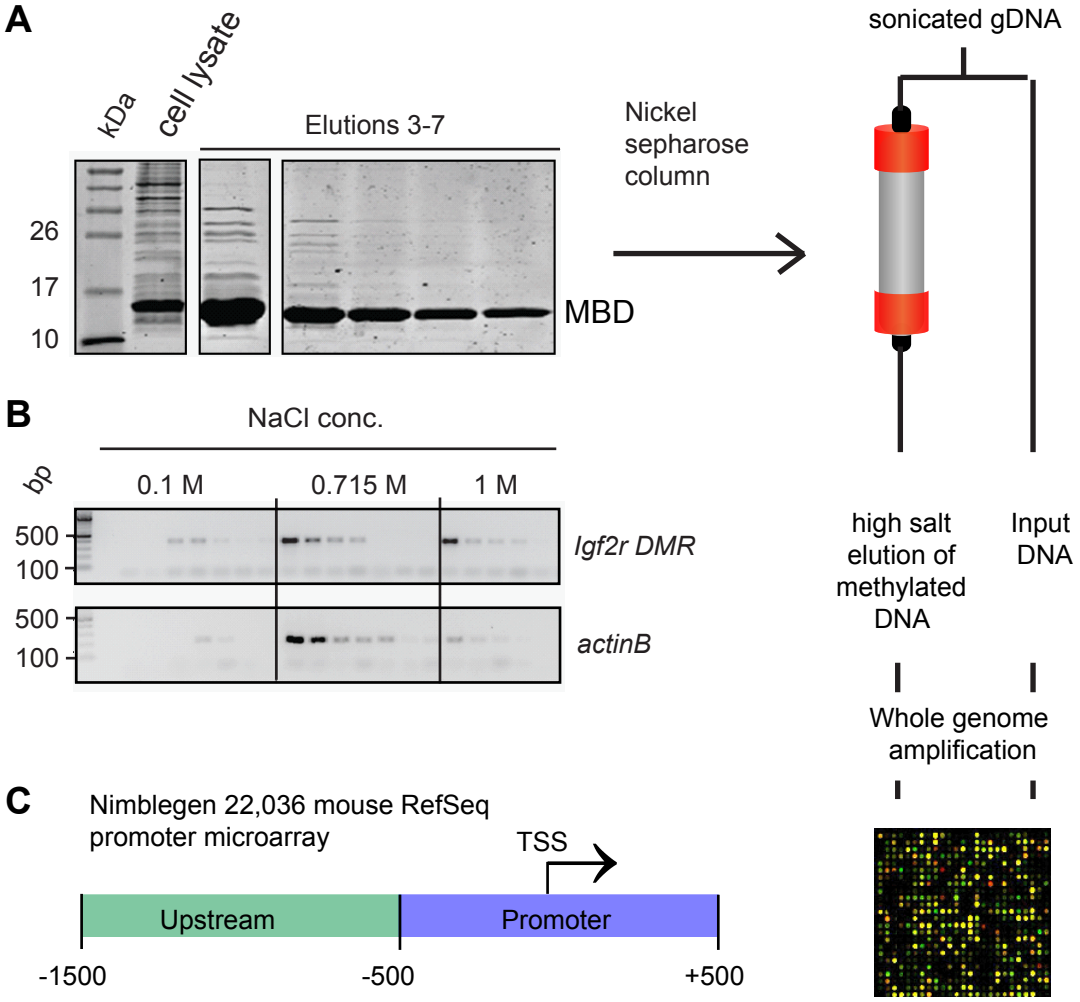


Figure 3.2 Methylated DNA affinity purification and promoter microarray

A: Nickel affinity purification of His tagged MBD domain of rat MeCP2. Shown is the MBD protein within the cell lysate (left pannel) and in subsequent elutions from nickel beads by immidazole. The purified MBD protein is packed onto a nickel sepharose column. Sonicated genomic DNA from cells of interest is passed through the MBD column at low salt to allow binding of methylated DNA. Unmethylated DNA is eluted at low salt and methylated DNA at high salt. **B:** Validation of methylated DNA affinity purifications by PCR on unmethylated *actinB* control and differentially methylated *Igf2r* DMR2. Actin DNA elutes at low salt (0.715M) and *Igf2R DMR2* DNA elutes at low (0.715M) and high (1M) salt. **C:** Hybridization of whole genome amplified methylated DNA and input DNA fractions to Nimblegen microarrays containing 22,036 mouse RefSeq promoters (spanning 1500bp upstream to 500bp downstream of transcription start site).

Methylated DNA affinity purification was carried out on 2 biological replicate WT and KO cell lines. This technique involved the purification of methylated DNA by liquid chromatography on a nickel column containing rat MeCP2 methylated DNA binding domain (MBD) (Figure 3.2A). Genomic DNA isolated from cultured MEFs and sonicated, was passed through the column and fractionated into unmethylated and methylated fragments with the aid of an increasing salt gradient. The purification was assessed by PCR analysis on unmethylated *Actin* and differentially methylated *Igf2r DMR* (Figure 3.2B).

The methylated DNA fractions as well as the input DNA were whole genome amplified and subsequently sent to Roche NimbleGen for hybridization to tiling microarrays representing 24,000 mouse protein-coding gene promoters (MM8 RefSeq) spanning 1500bp upstream to 500bp downstream of transcription start sites (TSSs) (Figure 3.2C). The tiling microarray contained probes mapping to the promoter regions of all protein coding genes. Each promoter region was covered by 25 probes (50-70bp) which are spaced at ~100bp. Each probe was present in multiple copies to allow quantitative analysis.

Microarray data was analysed with the help of Jose de las Heras and Arvind Sundaram (Myant et al, 2011). Raw fluorescence intensities were Loess normalized using the LIMMA package. Subsequently log2 values of MAP/Input were calculated for each probe (probe value). A log2(MAP/Input) value ≤ 0 indicates lack of enrichment in the methylated DNA fraction compared to Input DNA. To allow comparison between microarrays, probe values were adjusted to have the same median absolute deviation. For ease of analysis the region was divided into a 1000bp

upstream region (-500 to -1500bp relative to TSS) and a 1000bp promoter region (-500bp to +500bp relative to TSS) (Figure 3.2C). Aggregate log2 (MAP/Input) values were then generated by determining the median of all probes in the promoter or upstream region (M value). Promoters and upstream regions with less than 5 probes were discarded. Additionally, a low stringency filter was used to eliminate low intensity probe binding. Finally average M values were calculated from replicate experiments.

3.2.2 LSH regulates DNA methylation at protein coding gene promoters

The first stage of my study involved calculation of total numbers of methylated promoters in WT and *Lsh*-/- MEFs (Figure 3.3A). This involved grouping of all promoters into three different classes, according to their CpG density: Low CpG density (LCP), Intermediate CpG density (ICP) and High CpG density (HCP). These categories were determined by using observed/expected CpG density cut-offs described previously (Mohn et al. 2008). Expected CpG densities are calculated by taking into account the base composition of a given promoter region. This allows the calculation of a probability value for a CpG dinucleotide to occur in a specific promoter region. Subsequent multiplication by the total number of base pairs in the given region results in a number of CpGs expected to occur. Different average M value cut-offs were then used to determine the number of methylated promoters in LCP (M value ≥0.8), ICP (M value ≥0.9) and HCP (M value ≥1) groups of promoters to allow for CpG density related differences in binding affinity to the MBD column (Figure 3.3A). The thresholds used were similar to those used in previous studies (Mohn et al, 2008; Weber et al, 2007).

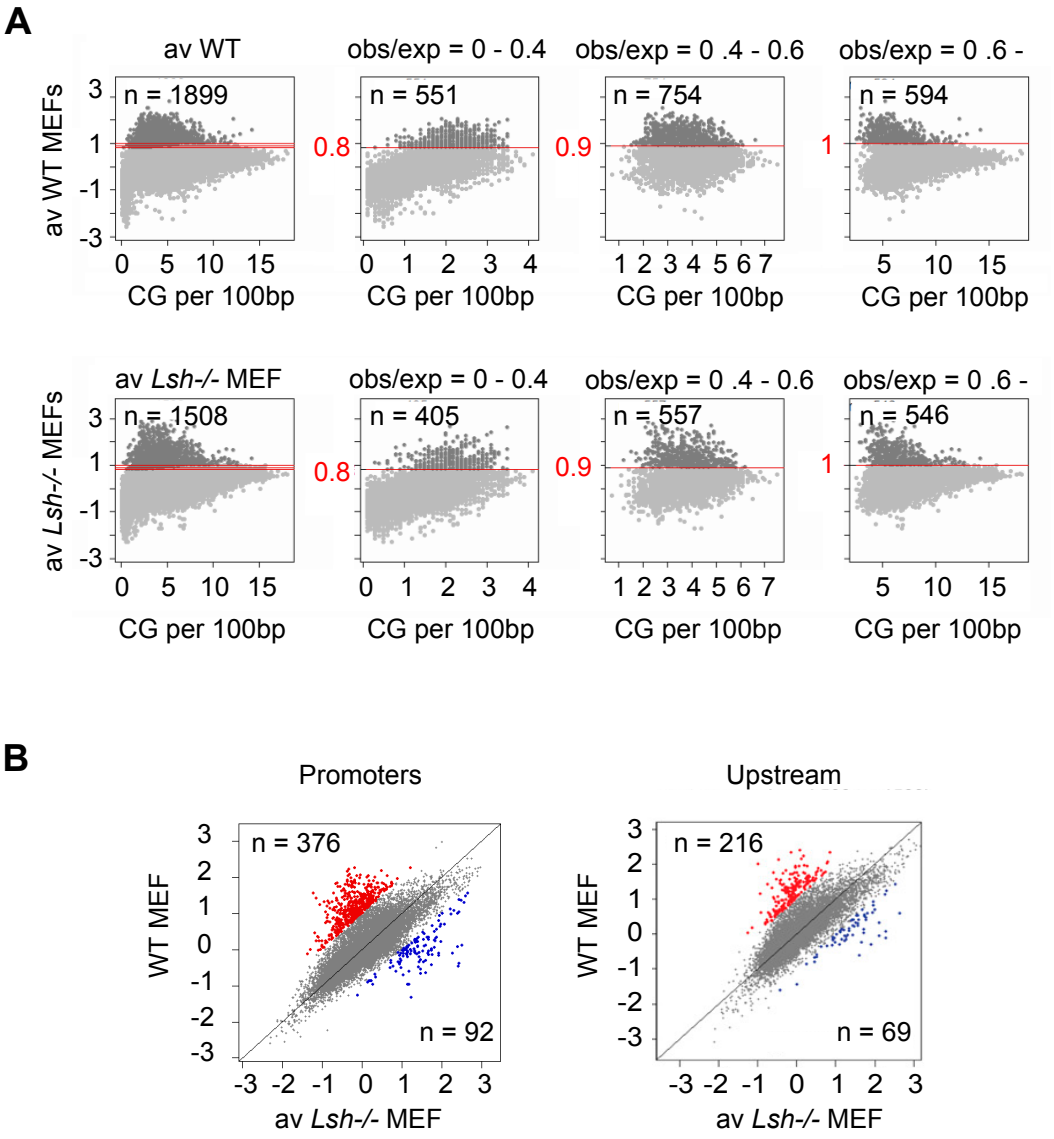


Figure 3.3 Comparison of promoter DNA methylation in WT and *Lsh-/-* MEFs

A: Total number of methylated promoters in WT and *Lsh-/-* MEFs. Median log2(MAP/Input) values are plotted against number of CGs per 100bp. MAP = methylated DNA affinity purified. Low, intermediate and high CpG density promoters were definied according to the observed to expected ratio of CpG dinucleotides in a given promoter region. Different thresholds were used for quantifying total numbers of methylated promoters in low (0.8), intermediate (0.9) and high (1) CpG density classes to account for differences in binding to the MBD column. n indicates number of promoters methylated. Left pannel shows an avarage over all CpG density classes. **B:** Scatter plot of 2-fold changes in DNA methylation in promoter and upstream regions (see Fig 3.2c). All WT log2(MAP/Input) values are plotted agains *Lsh-/-* log2(MAP/Input) values. A threshold was applied to indicate promoters/ upstream regions undergoing 2-fold or more hypomethylation (red) or hypermethylation (blue) in *Lsh-/-* MEFs compared to WT.

In summary, all classes of promoters display loss of DNA methylation in *Lsh-/-* MEFs compared to WT, although a lower percentage of high CpG promoters are hypomethylated. Averages across all CpG density groups highlight that 1899 promoters are methylated in WT MEFs and 1508 are methylated in *Lsh-/-* MEFs. This represents hypomethylation of ~20% of all methylated promoters.

Subsequently, differentially methylated regions were identified by calculating the difference between M values between WT and *Lsh-/-* MEFs. It is unclear, how fold change in enrichment of microarray signal between different cell lines relates to absolute changes in DNA methylation. I therefore used a high threshold (2-fold) to identify significant differences in DNA methylation (Appendix 3). The scatter plot in Figure 3.3B (left) highlights 376 gene promoters (red) being 2-fold or more hypomethylated in *Lsh-/-* MEFs, comprising around 20% of all normally methylated promoters. Interestingly, a smaller group of 92 promoters which are normally unmethylated become methylated in the absence of LSH. A role of LSH in preventing DNA methylation has not been identified before. Whether the hypermethylation we see in *Lsh-/-* MEFs is a direct effect of LSH, or an indirect effect through DNA methylation defects at other gene promoters has yet to be identified. A similar trend was seen in the case of upstream regions (Figure 3.3B right). The rest of the chapter focuses specifically on the 1000bp promoter region.

At this stage, it is important to consider the correlation between replicate experiments carried out, to determine, how significant the determination of differentially methylated regions may be. For this purpose scatter plots comparing replicates with eachother were generated and Spearman's correlation coefficients

calculated (Appendix 1). Both WT and *Lsh-/-* replicates display a high Spearman's correlation coefficient (~0.89). We were therefore satisfied, that replicate experiments were sufficiently similar to carry out further analysis. False discovery rate (FDR) calculations (Benjamini Hochberg) are often used to analyse the significance of microarray data. Due to the nature of the promoter microarray set up, we did not think an FDR analysis would be especially conclusive: 25 probes span the regions represented on the microarray. For a lot of promoter regions, differences in fluorescence intensities are limited to a small part of the entire region. Calculating FDRs for summarised probe values may therefore lead to an overestimation of the number of false discoveries which may be encountered. To overcome this, future work would have to involve re-analysis of the raw data, to calculate FDRs for each probe separately.

In summary, LSH is required for DNA methylation of a large number of intermediate- and low- CpG density class promoters of protein coding genes.

3.2.3 LSH is required for DNA methylation throughout the genome at single gene promoters as well as clusters of promoters

The next objective was to determine the location of LSH dependent promoters across all chromosomes, with the intention to identify whether LSH is equally required for DNA methylation of promoters on all chromosomes. From Figure 3.4A, it can be seen that both hypo- and hypermethylated genes are spread over all chromosomes. Despite the spread across all chromosomes, DNA methylation at single promoters as well as groups of up to 18 adjacent promoters requires LSH. This indicates that lack of LSH affects DNA methylation of large chromosomal regions.

The *Rhox* cluster on the X chromosome (Figure 3.4B) is most striking, as 11 members, including the neighbouring *Gm9* gene, show great reduction in DNA methylation in *Lsh-/-* MEFs. It is therefore the largest contiguous genomic region affected by LSH deficiency. The *Rhox* cluster encodes a group of reproductive homeobox genes on the X chromosome (Maclean et al, 2005). Initially 12 members of the *Rhox* gene family were identified but now several sub members have been shown to exist, suggesting an expansion of the cluster (Wang & Zhang, 2006). These genes are expressed in a cell type specific manner. Nine out of 12 *Rhox* genes are exclusively expressed in reproductive tissues (Maclean et al, 2005). *Rhox4*, *7* and *8* also show expression in thymus, stomach and intestine respectively (Maclean et al, 2005).

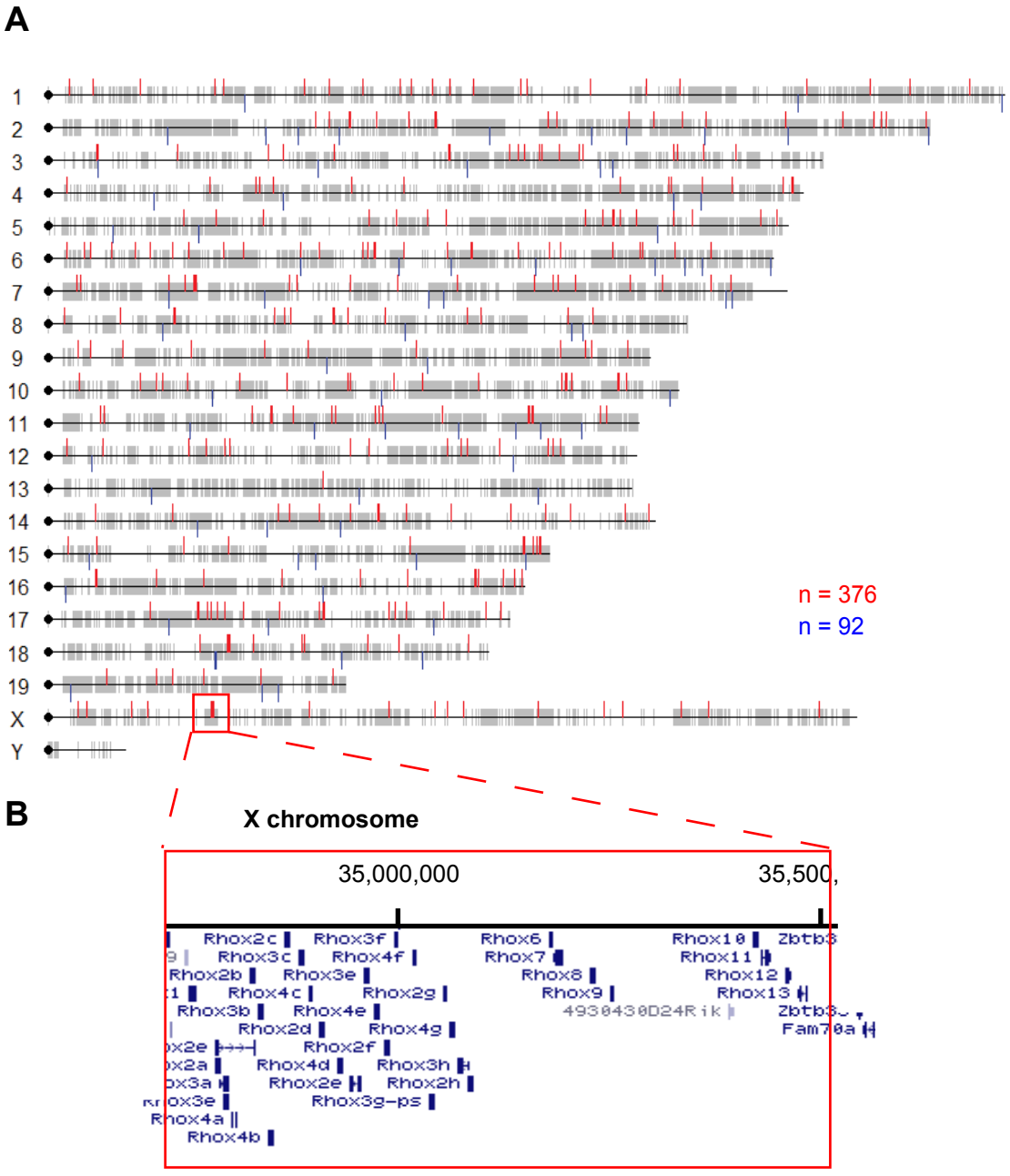


Figure 3.4 Chromosomal distribution of LSH target promoters

A: Chromosomal map of promoters undergoing 2-fold changes in DNA methylation in *Lsh-/-* MEFs (hypomethylated in red and hypermethylated in blue). **B:** Structure of the *Rhox* locus on the X chromosome.

Several of these genes are regulated by hormones and all are colinearly expressed (timing of expression is dependent on chromosomal position). Noticeable is the fact that they are highly expressed in extraembryonic tissues (Maclean et al, 2005).

DNA methylation of *Rhox* genes was previously analysed in *Dnmt3a* and *Dnmt3b* null mouse embryos (Oda et al, 2006). In this study it was found that *Rhox* gene promoters undergo methylation reprogramming during early embryonic development. *Rhox* gene promoters are hypomethylated in the pre-implantation embryo and gain DNA methylation by the E9.5 stage in embryonic, but not extraembryonic tissues (Oda et al, 2006). Importantly, DNA methylation of *Rhox* gene promoters is independent of X inactivation.

Rhox promoter *de novo* methylation in the embryo is dependent on DNMT3B and to a lesser extent on DNMT3A (Oda et al, 2006). Requirement of LSH for DNA methylation of *Rhox* genes highlights the potential need for LSH in early differentiation into the embryonic cell lineages.

3.2.4 LSH may be involved in DNA methylation at different stages of early development

To address the potential role of LSH in early cell lineage commitment I set out to determine the developmental time point at which LSH-dependent promoters are normally methylated. As described in the introduction, LSH is thought to be mainly required for *de novo* methylation. The aim was therefore to identify at which stage of mammalian development LSH may be required for correct DNA methylation establishment. For this purpose I carried out an additional MBD column

and microarray analysis experiment on WT embryonic stem cells (ES cells). This enabled comparison of DNA methylation levels of gene promoters in WT ES cells, WT MEFs and *Lsh-/-* MEFs. The WT ES cells used in this study were not isolated from the same mice as WT and *Lsh-/-* MEFs. Conclusions therefore have to be seen as preliminary and future repeats with cell lines isolated from the same mice will be important. This will be possible once *Lsh-/-* mice have been generated in our lab.

Results of the DNA methylation comparison are presented as a heat map in Figure 3.5. I clustered gene promoters displaying 2-fold or greater differences in DNA methylation between cell lines, giving rise to three groups of LSH dependent gene promoters. Group I is the largest and encompasses 233 gene promoters which are normally methylated in WT ES cells and WT MEFs but are hypomethylated in the absence of LSH. GOFFA gene ontology analysis demonstrates that genes in this cluster show involvement in processes such as signalling pathways, morphogenesis as well as sensory perception. Additionally, genes involved in trophectoderm development were identified. Group II gene promoters (143) are normally unmethylated in ES cells and methylated in MEFs. Ontology analysis reveals presence of genes involved in neural development, cell as well as anatomical morphogenesis. Additional genes involved in embryo development are found, which include pluripotency genes and imprinted genes. The smallest group of promoters (92) shows increases in DNA methylation in the absence of LSH. This cluster of promoters is normally unmethylated in both ES cells and MEFs, which prevents determination of normal timing of DNA methylation in the comparison of these cell lines. Members of this cluster include genes involved in metabolic processes, transport as well as cell differentiation and proliferation.

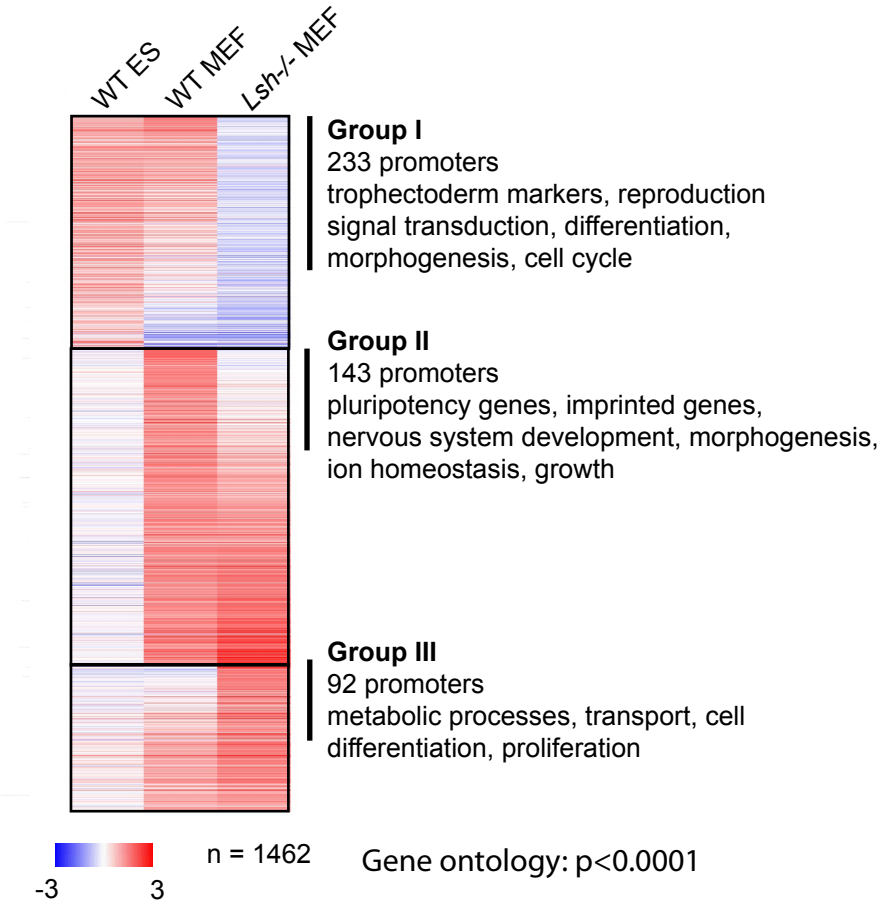


Figure 3.5 Developmental timing of LSH dependent DNA methylation

Comparison of DNA methylation between WT ES, WT MEFs and *Lsh-/-* MEFs. The heat map was generated to show all promoters with 2-fold changes between the three cell lines (total number of promoters = 1462). Each row is a promoter and a blue – white – red colour gradient of increasing methylation (log2(MAP/Input)) was used. -3 is the lowest and 3 the highest M value in the dataset. Top: All promoters with 2-fold less DNA methylation in *Lsh-/-* MEFs compared to WT ES cells, ranked by decreasing methylation in WT MEFs. Middle: All promoters with 2-fold higher DNA methylation in WT MEFs compared to ES cells, ranked by increasing methylation in *Lsh-/-* MEFs. Bottom: All promoters with 2-fold more DNA methylation in *Lsh-/-* MEFs compared to WT ES cells, ranked by increasing methylation in WT MEFs. This enabled assignment of LSH dependent gene promoters into three groups depending on their normal methylation status in the two WT cell lines (indicated on the right of the heat map).

Due to the 2-fold cut-offs (for significant methylation changes) used in this analysis, some genes are allocated to group II instead of group I, when comparing with WT ES cells. These include some members of the *Rhox* gene cluster, which are all known to be methylated in most embryonic tissues. Additionally some genes are omitted from the heat map. These include members of the Prader Willi Syndrome (PWS) imprinted gene cluster. *Ndn* is represented in GroupII, but *Mkrn3* and *Peg12* from the same imprinted gene cluster are omitted due to DNA methylation differences compared to ES cells not being above the 2-fold threshold. This analysis therefore enables rough identification when LSH dependent promoters may normally be methylated during development. It can be used to predict that LSH may be required for DNA methylation during epiblast and extraembryonic tissue specification (group I) and during stem cell differentiation into somatic cells (group II). Group III gene promoters are not methylated in ES cells or MEFs. I could therefore not determine their usual timing of DNA methylation.

Another aspect to consider in this study is that it involved comparison of cell lines not derived from each other. To determine more accurately developmental timing of LSH dependent DNA methylation, embryonic studies or stem cell differentiation experiments would have to be carried out. Our recent study for example investigated the role of LSH in silencing of pluripotency genes during retinoic acid differentiation of *Lsh* knockdown ES cells into neuronal progenitors. Pluripotency genes were silenced less efficiently in the absence of LSH (Myant et al, 2011). This confirms that LSH is required during ES cell differentiation for silencing pluripotency genes. Additionally, inefficient silencing of pluripotency genes during

ES cell differentiation supports the hypothesis that LSH may be required for establishment rather than maintenance of DNA methylation and gene silencing.

3.2.5 Validation of DNA methylation defects in *Lsh-/-* MEFs

The observations described above were based on analyses of high throughput data. Validation of these data was required to exclude technical artefacts of MBD column purification of methylated DNA and analyses of microarray hybridizations. For this purpose, I used sodium bisulfite genomic DNA sequencing to investigate DNA methylation profiles of promoters from each of the three LSH dependent promoter groups outlined in Figure 3.5. These included *Rhox2* (Group I), *Ndn* (Group II) and *Dbt* (Group III). In addition I also carried out bisulfite sequencing of some other gene promoters within the *Rhox* cluster and the imprinted PWS cluster (which contains Ndn). I used *Oct4* as a negative control, which has previously been identified to remain largely methylated in absence of LSH (Athanasiadou et al, 2010).

Bisulfite analysis of several members of the *Rhox* cluster (*Rhox2, 6 and 9*) is shown in Figure 3.6A. It reveals almost complete loss of DNA methylation in *Lsh-/-* MEFs compared to WT MEFs. Additionally, I carried out bisulfite sequencing on genomic DNA from stable *Lsh* knockdown ES cells and non-silencing control ES cells. The bisulfite profile displays no loss of DNA methylation at the *Rhox2* promoter in Lsh knockdown ES cells (Figure 3.6A). This suggests, that LSH is required for *de novo* methylation of gene promoters within the *Rhox* locus.

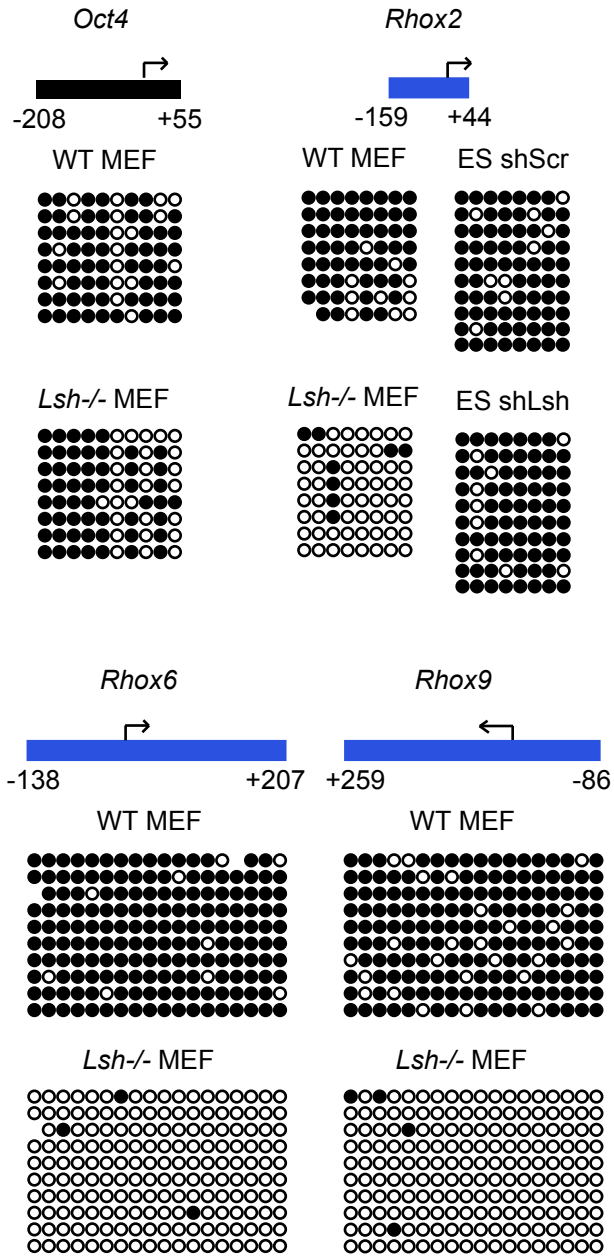


Figure 3.6 A Validation of MBD column and microarray by bisulfite sequencing

Bisulfite genomic DNA sequencing on *Oct4* and *Rhox* genes in WT and *Lsh-/-* MEFs as well a control shRNA and shLSH knockdown ES cells. *Oct4* was used as a control, which had previously been identified to only undergo DNA hypomethylation in some CpG positions. Black circles indicate methylated CpG positions and white circles indicate unmethylated CpG positions. Each row represents a DNA sequence from a different cell in the population (potential clonal sequences were removed in BiQ Analyzer). Diagrams above each bisulfite profile indicate the position of the sequenced region in relation to the transcription start site of the gene associated with the promoter region.

Group II of gene promoters whose methylation was identified to be dependent on LSH in the above high throughput data, includes various imprinted gene promoters. Previous studies, as outlined in the introduction, showed that LSH is required for DNA methylation at the imprinted *Cdkn1c* gene promoter, but not other imprinted differentially methylated regions (DMR). This study argued that *Cdkn1c* was the only DMR analysed which gains differential DNA methylation in the embryo, rather than the germ line. This indicates that LSH may be required for establishing tissue specific imprinted DNA methylation in the embryo, rather than maintaining imprinted DNA methylation which is established in the germ line. I carried out bisulfite sequencing on several imprinted gene promoters which normally gain DNA methylation in the embryo (Figure 3.6B). These included *Ndn*, *Peg12* and *Mkrn3*. These three genes are part of the Prader Willi Syndrome locus on chromosome 7 and are regulated by a differentially methylated region in the *Snrpn* promoter, which is methylated during germ cell specification. My methylated DNA affinity purification and microarray analysis did not indicate hypomethylation of the *Snrpn* promoter. *Ndn*, *Peg12* and *Mkrn3* promoters on the other hand show extensive DNA hypomethylation in bisulfite profiles of *Lsh-/-* MEFs (Figure 3.6B). This supports the idea that LSH is required primarily in establishing DNA methylation patterns at imprinted genes, rather than maintaining them.

Finally, for the validation of Group III promoters, bisulfite sequencing on *Dbt* was carried out, confirming DNA hypermethylation in *Lsh-/-* MEFs. Although significant, the extent of hypermethylation was less than hypomethylation seen in Group I and Group II promoters (Figure 3.6B).

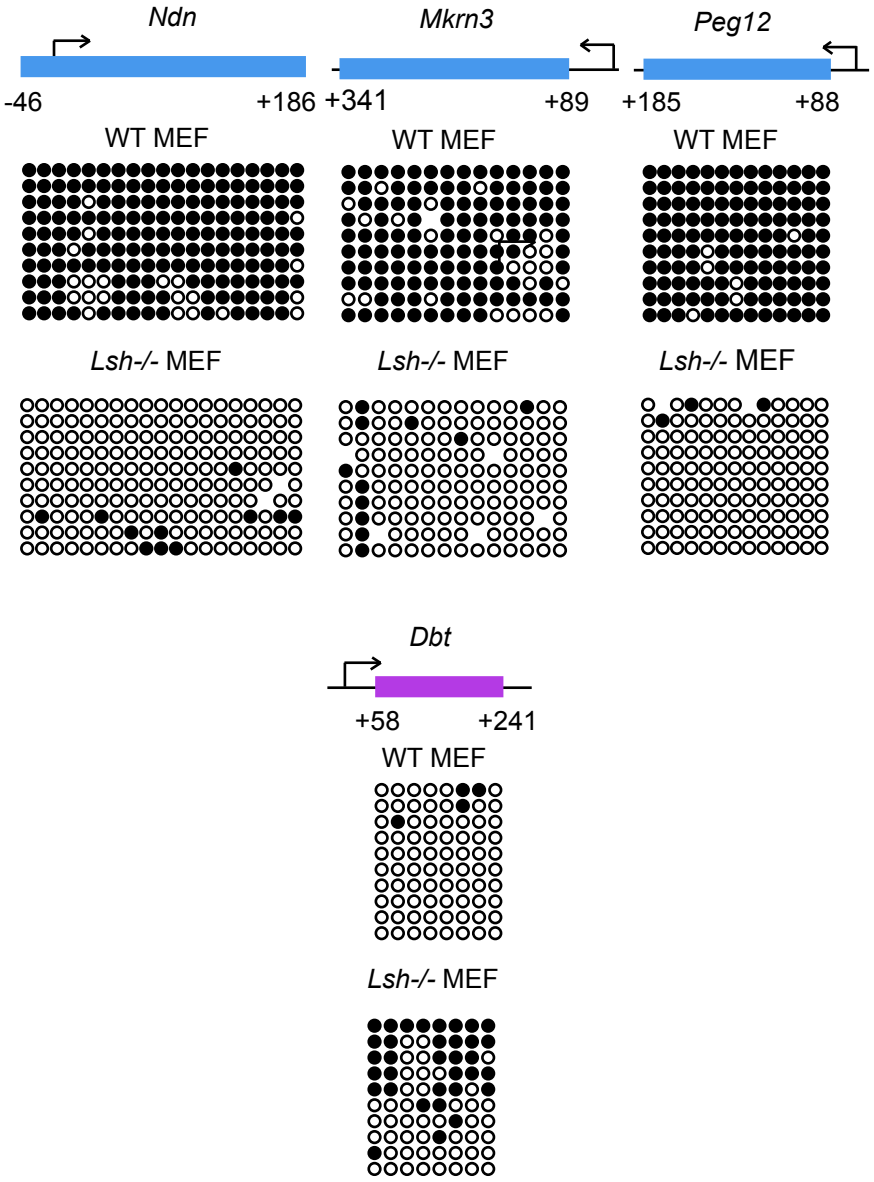


Figure 3.6 B Validation of MBD column and microarray by bisulfite sequencing

Bisulfite genomic DNA sequencing on genes in the PWS imprinted cluster as well as on *Dbt,* a gene displaying increased DNA methylation in *Lsh-/-* MEFs in promoter microarray analysis. Black circles indicate methylated CpG positions and white circles indicate unmethylated CpG positions. Each row represents a DNA sequence from a different cell in the population (potential clonal sequences were removed in BiQ Analyzer). Diagrams above each bisulfite profile indicate the position of the sequenced region in relation to the transcription start site of the gene associated with the promoter region.

In summary, bisulfite sequencing confirms requirement of LSH for DNA methylation at promoters normally methylated during the establishment and differentiation of embryonic stem cells. LSH is especially required for *de novo* rather than maintenance of DNA methylation at the *Rhox2* gene promoter. Additionally, results suggest that imprinted DNA methylation marks are established rather than maintained by LSH. This could be further confirmed by carrying out bisulfite DNA sequencing on imprinted gene promoters in *Lsh* knockdown MEFs.

3.2.6 Gene expression patterns are altered in absence of LSH

From the above analysis the question arises, whether LSH dependent promoter DNA methylation is required for correct gene expression at associated genes or whether other factors are sufficient to maintain gene expression patterns in absence of LSH. For this purpose I isolated cDNA from WT and *Lsh*-/- MEFs for subsequent expression microarray analysis of 25,658 protein coding genes (Nimblegen MM8-60mer-expr-X4) (See materials and methods for normalization methods). 755 genes are overexpressed in *Lsh*-/- MEFs, whereas 516 transcripts show 4-fold or greater decrease in expression in absence of LSH (Figure 3.7A). In summary, 5% of all protein coding genes are miss expressed in *Lsh*-/- MEFs.

Strikingly, *Lsh*-/- MEFs display many more gene expression defects than DNA methylation defects. This could be an indication of either indirect effects of LSH on gene expression or that LSH is required for regulating gene expression independent of DNA methylation. To further investigate this, I sought to determine whether expression and DNA methylation changes are inversely proportional with each other as would normally be predicted (Figure 3.7B).

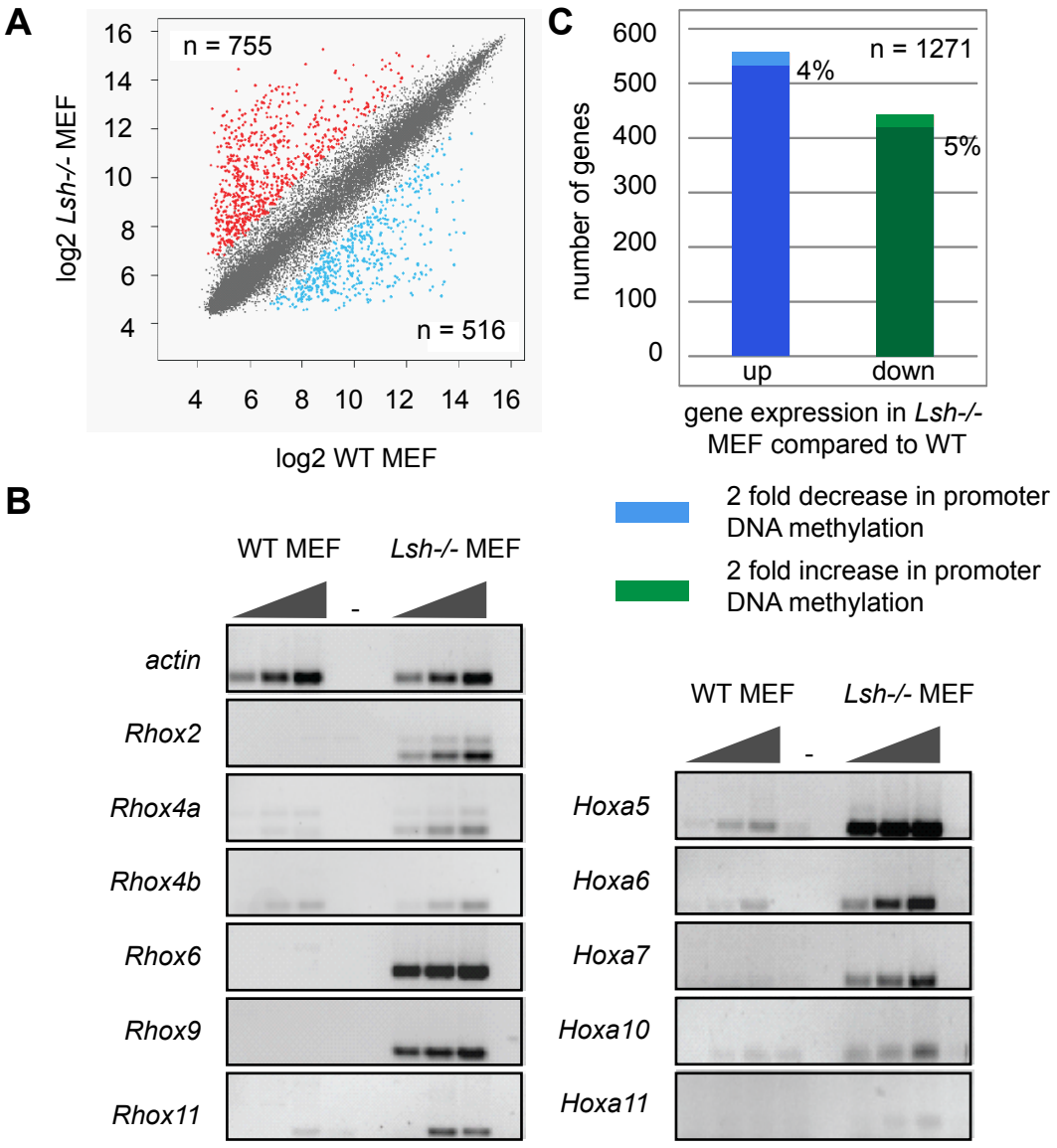


Figure 3.7 Gene expression analysis in WT and *Lsh-/-* MEFs

A: Expression microarray analysis. Plotted are all log2 median fluorescence intensity values for WT and *Lsh-/-* MEFs for each gene analysed on the Nimblegen expression array. In red are genes which show 4-fold or more increase in expression in *Lsh-/-* MEFs and in blue are genes which show 4-fold or more decrease in expression in *Lsh-/-* MEFs. **B:** Validation of expression array by reverse transcription PCR of *Rhox* and *Hox* genes. PCRs were carried out with three different concentrations of cDNA as well as on a negative cDNA synthesis reaction (RT-). **C:** Comparison of expression changes with DNA methylation changes in *Lsh-/-* MEFs. Shown is a histogram of the 4-fold or more up- or down regulated genes which also display 2-fold or higher DNA methylation changes. Highlighted are the genes whose expression changes and methylation changes are inversely proportional. 4% of upregulated genes loose DNA methylation at their gene promoters. 5% of downregulated genes gain DNA methylation at their promoters.

Plotted are the numbers of genes showing 4-fold or greater up or down regulation. Highlighted are those expression changes, which can be directly attributed to 2-fold or higher inversely proportional DNA methylation changes. In summary, ~5% of 4-fold or greater expression changes can be directly attributed to 2-fold or greater promoter or upstream region DNA methylation changes.

To validate expression microarray conclusions that LSH is required for correct expression of specific genes, I carried out reverse transcription PCR analysis of WT and *Lsh-/-* MEF cDNA on *Rhox* and *Hox* genes (Figure 3.7C). Several *Rhox* genes analysed, are activated, consistent with hypomethylation of their promoters in *Lsh-/-* MEF. *Rhox4b* is an exception and remains silent, despite being hypomethylated in *Lsh-/-* MEFs. *Hox* genes are weakly expressed in WT MEFs and are upregulated in absence of LSH, despite not undergoing any changes in promoter DNA methylation. Other work from our lab confirmed the miss expression of other group I and group II genes that show LSH-dependent DNA methylation (Myant et al, 2011).

In summary, absence of LSH causes more extensive expression changes than DNA methylation changes. This could be due to indirect effects of promoter DNA methylation changes on gene expression. If for example the promoter of a repressor of gene expression is hypomethylated in absence of LSH, this could lead to expression changes at genes regulated by this repressors. In some cases DNA methylation defects are not accompanied by expression defects, indicating that DNA methylation is not required for the maintenance of a correct gene expression status.

3.3 Discussion

In summary, the above study allowed more detailed characterization of LSH dependent DNA methylation at gene promoters than ever published before. Before my study, LSH requirement for repetitive element DNA methylation was extensively characterized. Previously identified effects of LSH absence on single gene DNA methylation and gene expression were suggested to be exceptions arising from close proximity to repetitive elements, rather than direct regulation by LSH. Contradicting this notion, I identified extensive DNA methylation and expression defects at protein coding gene promoters genome wide. I identified extensive changes in DNA methylation at single genes as well as gene clusters in *Lsh*-/- MEFs. Both hypo- and hypermethylation was observed, suggesting either a dual role of LSH in ensuring correct DNA methylation patterns or indirect effects of DNA hypomethylation of some promoters on the methylation status of others. This suggests that LSH may indeed be directly involved in regulating DNA methylation and gene expression at non-repeat regions of the genome.

Additionally, I identified LSH to be required for DNA methylation at some gene promoters whose DNA methylation status distinguishes early embryonic lineages (e.g. *Rhox* genes and pluripotency genes). This, suggests, that LSH may be implicated in *de novo* DNA methylation and gene expression patterns during early development. Interestingly *Lsh*-/- embryos survive until after birth, without any major developmental defects, suggesting that *de novo* methylation events which fail to take place in absence of LSH do not result in gross developmental abnormalities during embryonic development. Direct facilitation of DNA methylation by LSH could be tested by Chromatin Immunoprecipitation of LSH at specific gene

promoters. Since my data suggests that LSH is required for DNA methylation of pluripotency gene promoters, ChIP could be carried out on these promoters during differentiation of stem cells into neuronal progenitors. This would enable to determine whether LSH interacts with these promoters to facilitate *de novo* DNA methylation during differentiation.

Since this study focused on *Lsh-/-* or knockdown cell lines, the next step towards confirming a potential involvement of LSH in DNA methylation and gene expression during early development would require studies in embryos as well as more stem cell differentiation experiments (Myant et al, 2011). Use of conditional knockout embryos with the potential to reactivate and inactivate LSH at different developmental stages would allow more precise characterization of the developmental timing of LSH-dependent *de novo* DNA methylation events. However, embryo studies can be difficult and have to be carried out according to guidelines. Therefore gaining as much insight into LSH function in DNA methylation from analysis of cell lines is important. This is the focus of chapter 4.

Chapter 4 - Investigating a cell-autonomous function of LSH in DNA methylation

4.1 Introduction

4.1.1 Potential requirements for LSH dependent DNA methylation

Following the conclusion from chapter 3, that LSH may function in DNA methylation and gene expression regulation as early as in the pre-implantation embryo, some other evidence for an early role of LSH exists. When comparing WT MEFs and WT ES cells, LSH expression is 10-fold reduced in WT MEFs (Figure 4.1B). Accordingly, an ES cell differentiation experiment I carried out demonstrates a 5-fold drop in LSH protein levels at the onset of differentiation. LSH levels peak slightly during differentiation but remain low compared to WT ES cells (Figure 4.1A and B). This substantiates the notion that LSH may function early in development and only carry out limited function in differentiated cells. Additionally, this poses the question, whether LSH can only exert its DNA methylation function during a specific early developmental window. In this chapter I therefore outline my attempt to answer the following questions:

Can LSH function in an autonomous manner in somatic cells, to facilitate DNA methylation establishment at any of the identified LSH dependent loci? Or is LSH function in establishing DNA methylation restricted to a particular developmental window, due to the requirement of other specific co-factors or signals not present in somatic cells? In addition I wanted to determine, whether LSH requires its catalytic activity (ATP hydrolysis) for its function in DNA methylation.

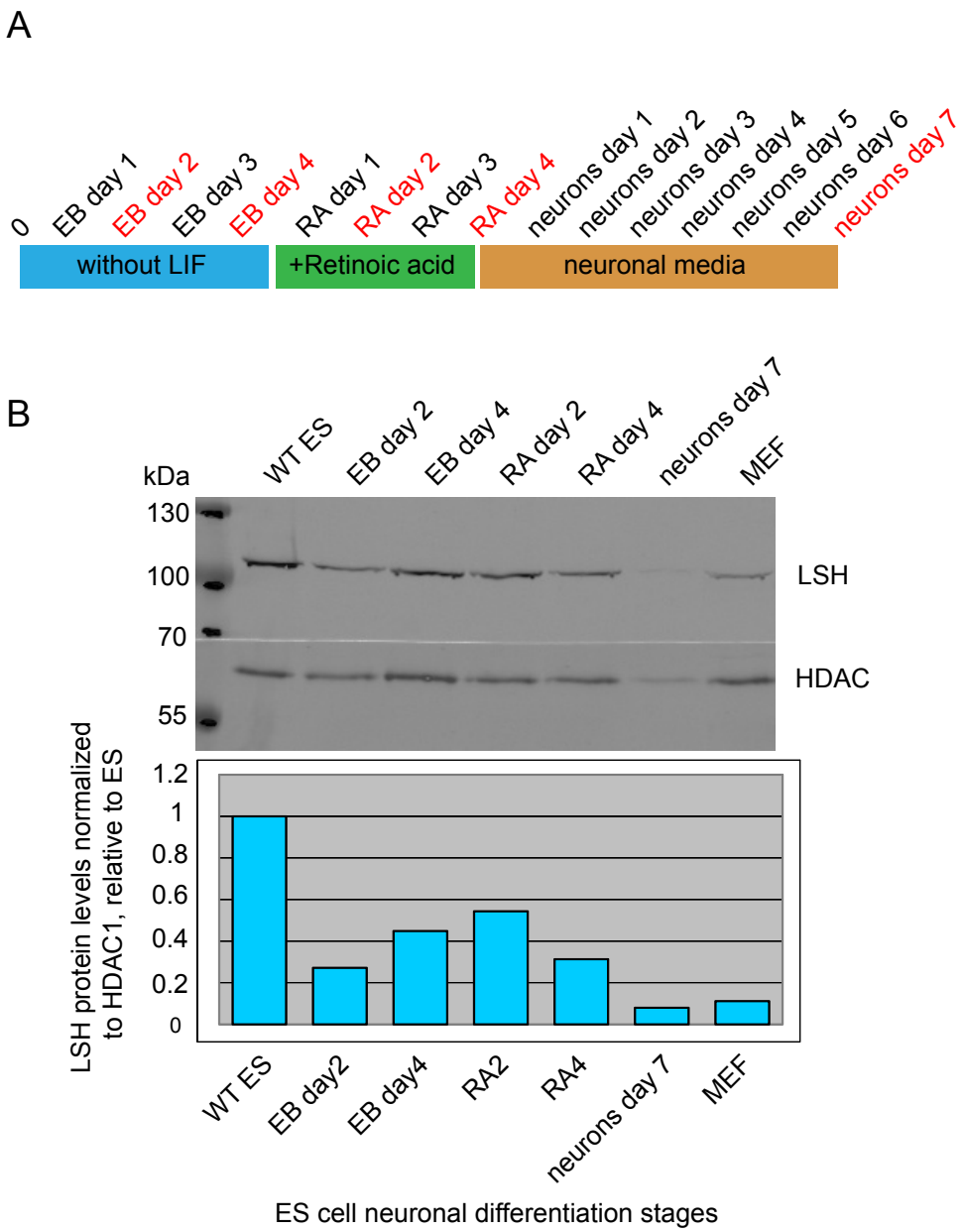


Figure 4.1 LSH protein levels during embryonic stem cell differentiation into neurons

A: ES cell differentiation timeline. Steps of neuronal differentiation are displayed, including removal of Leukemia inhibitory factor, addition of retinoic acid and addition of neuronal media. Time points at which cells were harvested are indicated in red. B: Western blot on LSH and HDAC1 as loading control (top pannel). Quantification of Western blot is shown below. LSH protein levels were normalized to HDAC1 loading control and are represented relative to WT ES cells (bottom pannel).

I sought to answer the above questions through re-expression of WT and catalytically mutant LSH in *Lsh*-/- MEFs. The catalytic mutation used in this study is a conversion of a conserved lysine into glutamine in the catalytic domain of LSH. The inability of this mutant to hydrolyse ATP was confirmed by Kevin Myant, a previous PhD student in our lab.

The primary pre-requisite for establishing correct DNA methylation patterns is the presence of *de novo* methyltransferases. As mentioned in the introduction of chapter 3, LSH has been shown to interact with DNMT3B directly as well as indirectly with DNMT1 and HDACs (Myant & Stancheva, 2008). One immediate restriction to the ability of LSH to rescue DNA methylation in mouse embryonic fibroblasts may therefore be the reduced level of the full length isoform of DNMT3B (DNMT3B1) in somatic cell lines (Okano et al, 1998a). Additionally this DNMT3B isoform is active on its own but is enhanced by presence of DNMT3L. DNMT3L is also expressed at low levels in somatic cells (Ooi et al, 2010).

Apart from this, DNA methylation has for a long time been thought to be a secondary event to gene silencing, perhaps ensuring long term stability of a silent state. This notion is supported by various studies. On the X chromosome for example, DNA methylation is not required for initiation of X chromosome inactivation. Prolonged culture of male ES cells in absence of DNA methylation, on the other hand leads to activation of Xist on the active X chromosome. Additionally DNA methylation also occurs at promoters of genes silenced during X inactivation. This is also thought to be a secondary event after silencing has taken place (Cohen et al, 2012; Jonkers et al, 2008; Sado et al, 2000). During stem cell differentiation DNA

methylation has also been demonstrated to occur at pluripotency gene promoters (especially Oct4) after silencing has occurred. Measurements of levels of Oct3/4 protein, H3K4me3, H3K9ac, H3K9me3 and DNA methylation have been previously profiled at different time points during differentiation of stem cells (Feldman et al, 2006). H3K9 methylation is the first repressive mark, appearing within the first day of differentiation. This initiates silencing of Oct3/4 and is followed by rapid decrease in H3K9ac, H3K4me. However, DNA methylation is only detectable after 2 days of differentiation, by which time Oct3/4 is already silent (Feldman et al, 2006). DNA methylation has been demonstrated to gradually increase throughout differentiation (Feldman et al, 2006). This suggests that DNA methylation may require prior formation of a repressive chromatin state upon initiation of differentiation. Can the establishment of a repressive chromatin state be initiated at early developmental genes in somatic cells, to facilitate subsequent DNA methylation establishment with the help of LSH?

Finally, correct initiation of gene silencing has been demonstrated to require initiation of signalling pathways at specific time points during early development. A recent study for example demonstrated the activation of G9a histone methyltransferase by a signalling cascade resulting in inhibition of APC/C ubiquitin ligase (Yamamizu et al, 2012). This signalling pathway is thought to ensure correct timing of pluripotency gene inactivation during differentiation of stem cells. These type of signals may be absent in somatic cells, potentially preventing re-establishment of pluripotency gene silencing and DNA methylation. Similarly silencing and DNA methylation of genes during divergence between embryonic and extraembryonic lineage cells also requires specific signals generated through

signalling pathways. An example of this is the Hippo signalling pathway which is imitated in ICM cells in the early embryo to allow continued repression of extraembryonic lineage markers (Avner & Heard, 2001; Heard et al, 2001). Re-introduction of LSH into somatic cells may not allow re-establishment of DNA methylation and gene silencing at early developmental loci due to the absence of such initiating signals of gene repression in somatic cells.

The above analysis suggests, there may be a developmental window during which gene silencing and DNA methylation of early developmental gene silencing can be established. Somatic cells may have mechanisms in place to maintain already established silent state, but may not be able to initiate the formation of these. An example of this is in X chromosome inactivation. Studies have demonstrated, that late expression of *Xist* RNA does not allow initiation of X chromosome inactivation. This suggests, that X inactivation occurs during a specific developmental window and cannot be established at later stages (Sado et al, 2004).

4.2 Results

4.2.1 Expression of WT and catalytic mutant LSH in *Lsh-/-* MEFs

Various attempts have been made in the lab to re-express LSH in *Lsh-/-* MEFs. These used expression vectors containing different promoters, different integration methods as well as different tagged LSH sequences. Expression of LSH under a CMV promoter failed completely and expression using a vector containing an MMLV promoter resulted in low transient expression of LSH protein.

Previous studies have identified the efficient silencing of retroviral promoters such as CMV and MMLV in different murine cell lines (Haas et al, 2003; He et al,

2005; Kawasaki et al, 2011; Lorincz et al, 2001; Swindle & Klug, 2002).Therefore I used a lentiviral Murine Stem Cell Virus vector to clone the WT and catalytically mutant *Lsh* (K237Q) open reading frame with a 3' 3 x FLAG tag followed by the 3' untranslated region into the Murine Stem Cell Virus vector (MSCV). The MSCV vector contains long terminal repeats from the murine stem cell PCMV virus which act as a promoter (Figure 4.2). The differences to the MMLV LTR are several point mutations and a deletion which enhance activation and have been shown to prevent endogenous silencing in stem cells. The idea was to try and use this vector to obtain more constitutive expression of LSH in *Lsh-/-* MEFs. The MSCV virus was packaged into lentiviruses using the Phoenix cell method (Figure 4.2). Harvested viruses containing the WT and catalytic mutant LSH constructs were used to infect *Lsh-/-* MEFs. Empty MSCV vector was used in a control infection. Positive pool cell lines were selected with puromycin and individual clonal lines isolated for further analysis.

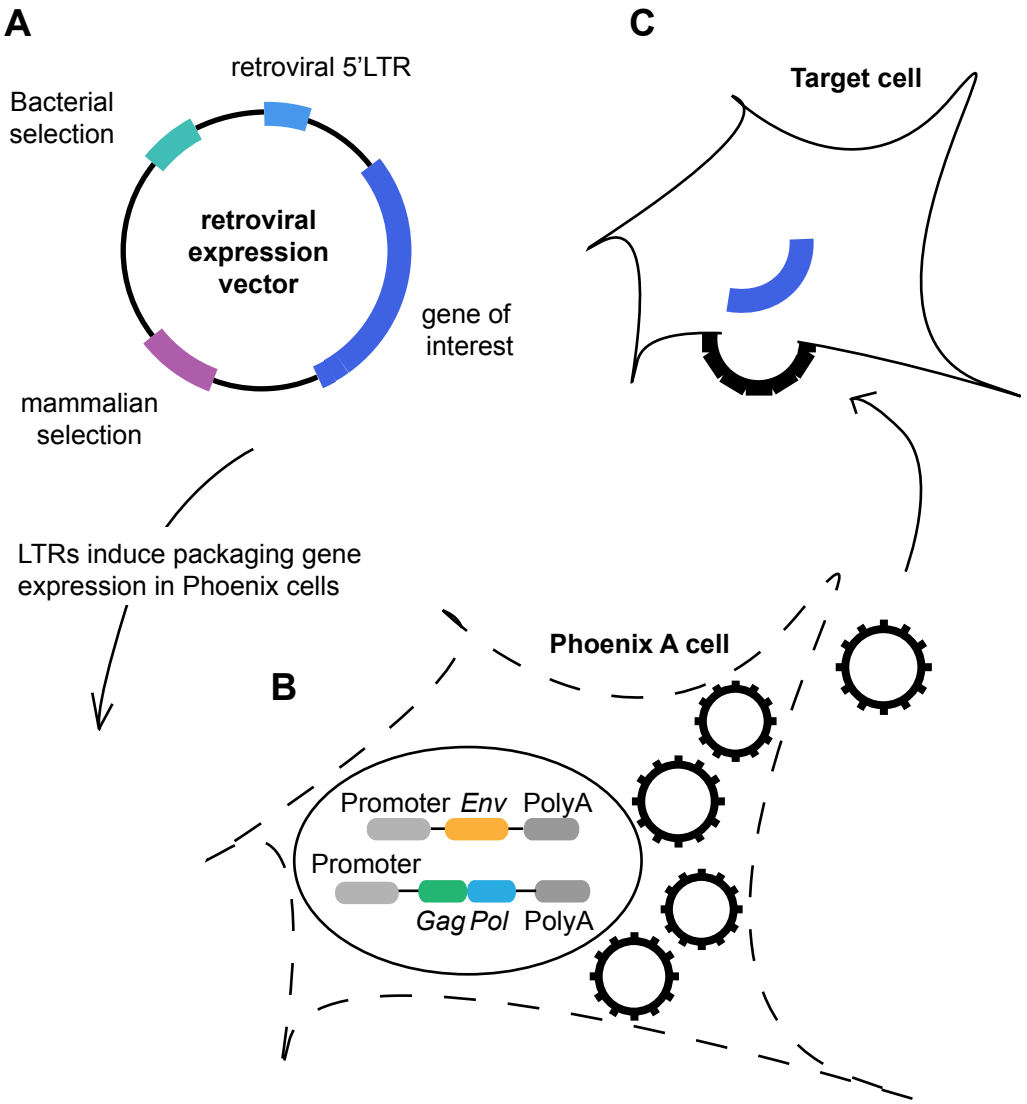


Figure 4.2 Retroviral expression system

A: Retroviral vector containing viral 5' long terminal repeat (LTR) which acts as a promoter for expression of a gene of interest. Vector also contains bacterial and mammalian selection markers. **B:** Transfection of the retroviral vector into PhoenixA cells. Viral LTRs within expression vector induce the expression of viral genes *Env* (envelope protein), *Gag* (retroviral core protein) and *Pol* (reverse transcriptase and integrase) present in Phoenix cells. This results in packaging of the gene of interest into viral particles. **C:** Viruses containing the gene of interest cause Phoenix cell lysis (dashed appearance of phoenix cell) and can be used to introduce the gene of interest into target cells by viral infection.

4.2.2 Characterization of rescue cell lines

I carried out Western blot analysis and quantification of LSH protein levels in the different WT and catalytic mutant clonal cell lines isolated (Figure 4.3A). All clonal lines apart from WT LSH rescue clonal line number 8 express LSH at lower levels compared to WT MEFs. The three mutant LSH rescue cell lines express LSH at higher levels than the WT1 and WT4 rescue cell lines. Any lack of DNA methylation or gene expression changes in mutant compared to WT rescue lines in following experiments can therefore not be attributed to insufficient expression of mutant LSH. I selected three cell lines displaying similar LSH expression levels for further analysis (WT1, WT4 and Mut4). Immunocytochemistry analysis confirms nuclear LSH expression in WT1, WT4 and mut4 cell lines (Figure 4.3B). Additionally, I used the other clonal cell lines (especially WT8) in some experiments to identify, whether different expression levels of LSH had variable effects on DNA methylation and gene expression. Apart from this, using different clonal lines enabled me to exclude the possibility that any DNA methylation or gene expression changes are due to derivation of certain clonal lines, rather than LSH expression itself.

To further characterize LSH rescue cell lines, I carried out growth measurements to determine the population doubling time of each cell line (Figure 4.3C). All rescue cell lines display increased population doubling times compared to WT and *Lsh-/-* MEFs. This indicates that re-expression of wild-type and mutant LSH by retroviral infection causes some growth disadvantage. Re-expression of WT and mutant LSH in these cells may cause aberrant DNA methylation or gene expression, which could lead to growth defects.

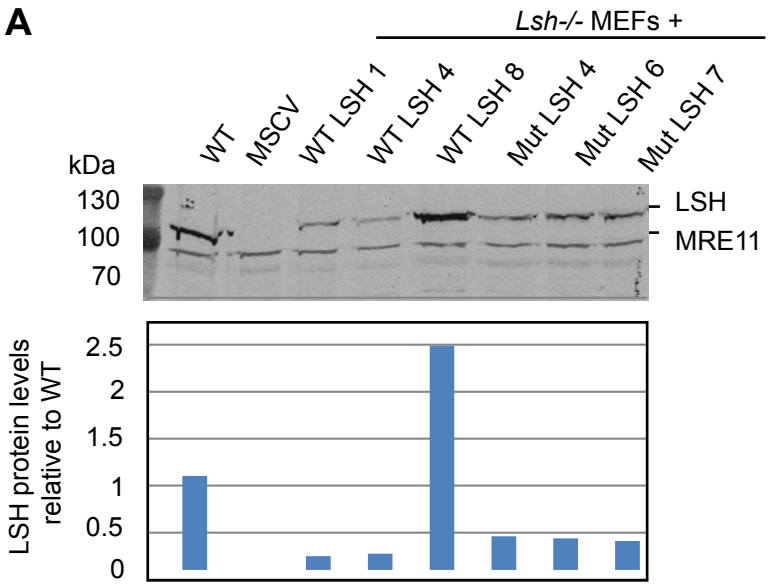


Figure 4.3 A LSH expression levels in rescue cell lines

A: Anti-LSH Western blot. Top pannel: LSH protein detection in three clonal cell lines each from infections of *Lsh-/-* MEFs with WT LSH open reading frame (ORF) or ATPase mutant LSH ORF. An anti-MRE11 antibody was used as a loading control. Bottom pannel: Western Blot quantification using Licor. LSH protein levels are normalized to MRE11 loading control and displayed relative to WT.

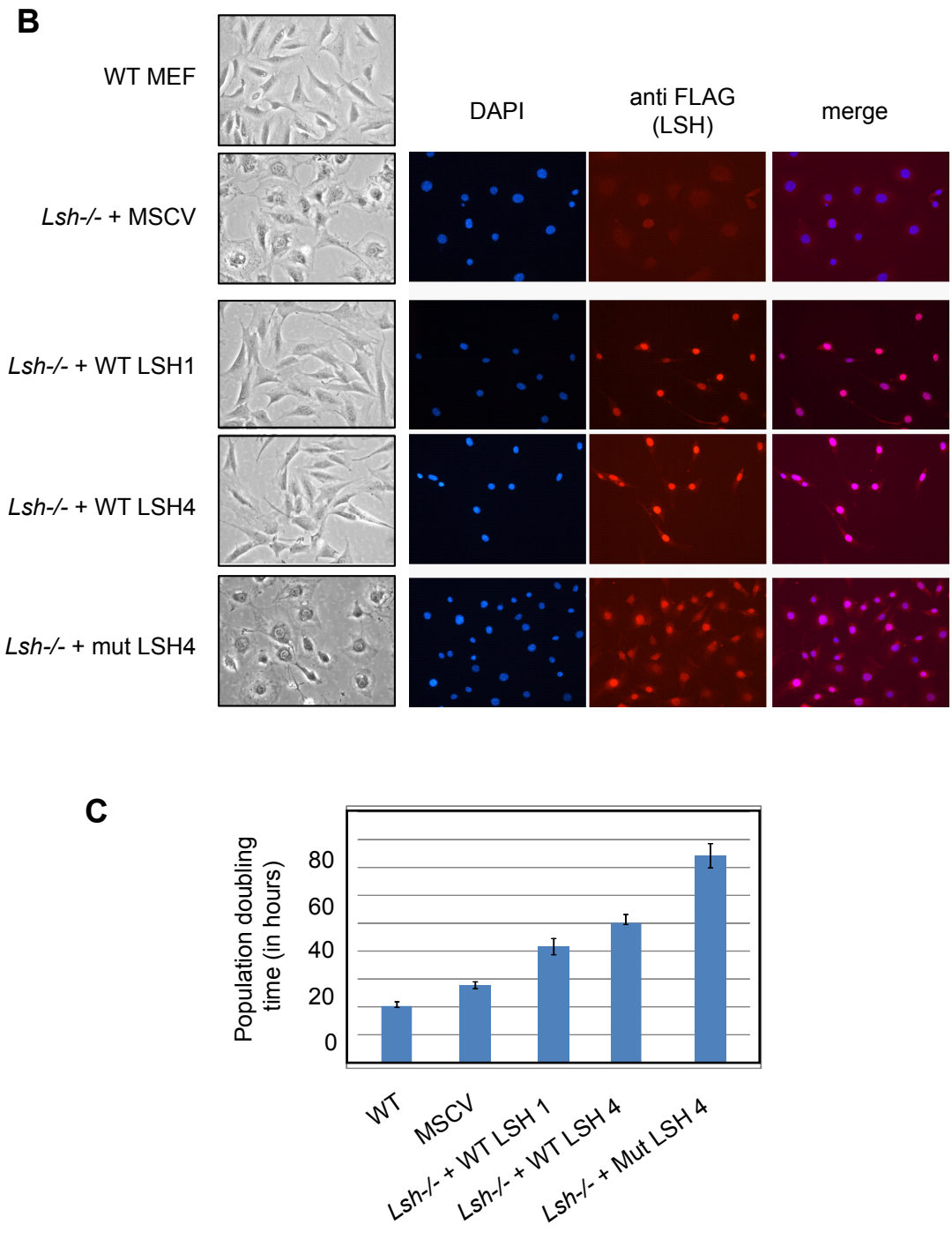


Figure 4.3 B & C LSH rescue cell line characterization

B: Cell morphology and Immunocytochemistry analysis: The left pannel shows control and rescue cell, cell morphology (40x magnification). The right pannel shows anti-FLAG immunofluorescence staining of LSH in MSCV empty vector and rescue cell lines. Merge with the DAPI nuclear signal confirms nuclear LSH expression. **C**: Population doubling time measurements of rescue cell lines. Error bars shown represent 2 x standard deviation of triplicate readings.

The catalytic mutant cell line displays the slowest growth rate, indicating that a mutation in the catalytic domain causes LSH to exert an additional negative effect compared to wild-type LSH. This is the first indication that there is a difference in function of wild-type and catalytically inactive LSH. To expand on this, I compared cell morphology of the clonal lines (Figure 4.3B left column). WT rescue cell lines look morphologically more similar to WT MEFs, whereas the mutant rescue cell lines resemble more the MSCV vector only control cell line. This suggests, that wild-type LSH, but not mutant LSH can cause some reversal of knockout phenotypes in these cells.

4.2.3 Wild-type but not catalytically inactive mutant LSH restores DNA methylation at repetitive sequences in *Lsh-/-* MEFs

After a general characterization of rescue cell lines, the immediate question was, whether LSH re-expression leads to rescue of DNA methylation defects in *Lsh-/-*MEFs. To obtain an indication whether DNA methylation levels had changed in these cell lines I carried out Southern blot analysis on minor satellite repeats, characteristic of centromeric DNA. This involved digestion of genomic DNA with MspI (methylation insensitive) and HpaII (methylation sensitive) restriction enzymes and probing with a 120bp minor satellite repeat fragment. For this analysis, three different clonal lines expressing either wild-type or catalytically inactive LSH were used (Figure 4.4A). All genomic DNAs are equally digested by methylation insensitive enzyme MspI. This was used as a loading control indicating equal DNA amounts as well as equal ability to digest the DNA fully. WT MEF gDNA is only digested to a small extent by methylation sensitive enzyme HpaII, indicating high levels of DNA methylation.

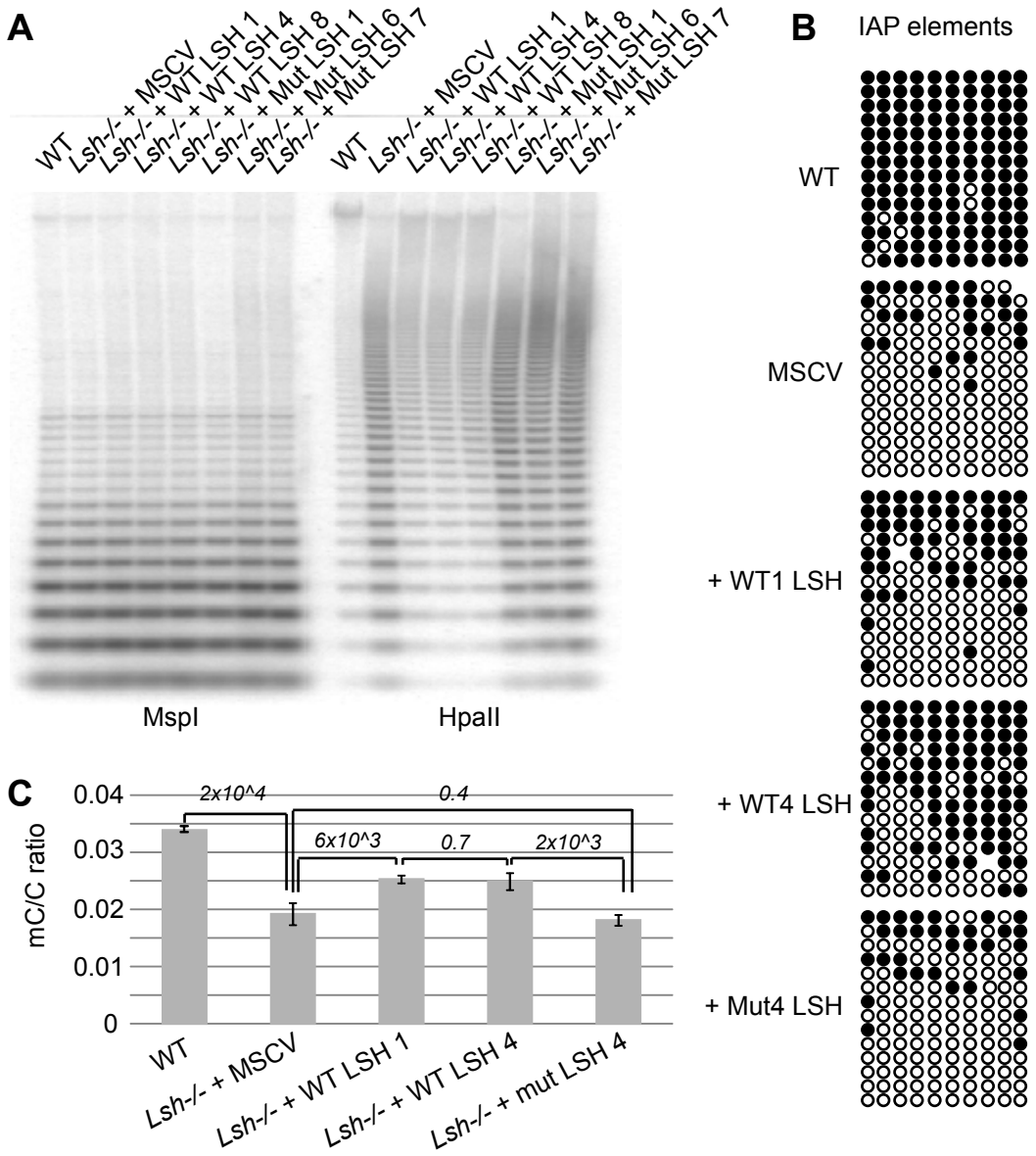


Figure 4.4 Repetitive element and whole genome DNA methylation analysis

A: Minor satellite DNA methylation by southern blotting of MspI (methylation insensitive) and HpaII (methylation sensitive) digested rescue cell line genomice DNA. The southern blot was probed with a minor satellite repeat probe. **B:** IAP element bisulfite sequencing on rescue cell line gDNA. **C:** Left: Whole genome DNA methylation levels in rescue cell lines were determined by HPLC analysis of genomic DNA. Values displayed are ratios of methyl cytosine to cytosine. Error bars show standard deviation of triplicates. Significance of DNA methylation differences between rescue cell lines were determined by two-tailed t- tests. P values are indicated on the graph.

Lsh-/- MSCV control cell gDNA display higher levels of digestion, confirming a loss of DNA methylation in these cells. This confirmed previous studies which showed minor satellite hypomethylation in *Lsh*-/- cell lines (Dennis et al, 2001). Southern blot analysis on wild-type LSH rescue cell lines displays intermediate levels of digestion by HpaII. This indicates increased levels of methylation at minor satellite repeats upon re-expression of wild-type LSH. Mutant cell lines on the other hand display similar HpaII digestion patterns to MSCV control cell lines. This suggests that catalytically mutated LSH is unable to cause increases in DNA methylation at minor satellite repeats.

Additionally, I determined DNA methylation patterns at Intracisternal A particle (IAP) repeat elements by sodium bisulfite DNA sequencing in three different rescue cell lines (Figure 4.4B). IAP elements are similar to retroviral repeats and are present only in rodent genomes (Qin et al, 2010). Consistent with minor satellite repeats, IAP elements display increased DNA methylation upon re-expression of WT *Lsh* (especially WT4), but not in cells expressing the mutant form of LSH (Figure 4.4B). Interestingly WT1 rescue cells display completely methylated IAP repeats as well as some completely unmethylated IAP repeats. Unlike bisulfite sequencing at single genes, repetitive element bisulfite sequencing amplifies repeats located in different genomic regions. Differences in DNA methylation of non-clonal sequences from the same rescue cell line could therefore be representative of DNA methylation at different IAP repeats within the genome being rescued to different extents.

To extend this study and determine global levels of 5meC, I prepared genomic DNA samples for HPLC analysis by Bernard Ramsahoye's group in the University of Edinburgh Cancer Research UK centre. HPLC analysis involves hydrolysis of DNA into single nucleotides and subsequent separation by High-performance reverse phase liquid chromatography (Ramsahoye, 2002). Presence of different bases can then be identified by detecting UV absorbances. Consistently with southern blot analysis of repeat elements, both WT rescue cell lines display genome methylation levels higher than those in the control *Lsh-/-* MSCV cells (Figure 4.4C). However, the cell lines expressing wild-type LSH showed only partial (~50%) rescue of DNA methylation, when comparing with WT MEFs. The mutant cell line showed no significant difference in genome-wide 5meC levels compared to *Lsh-/-* MEFs. Two-tailed t-tests between all samples confirmed these differences in DNA methylation (Figure 4.4C).

In summary, whole genome and repetitive sequence analysis reveals a partial rescue of DNA methylation in WT but not catalytic mutant rescue cell lines. The question arising is, why full DNA methylation could not be restored in these cell lines. I investigated several possible reasons for this partial rescue. Firstly, most rescue cell lines used showed LSH expression levels lower than those in WT MEFs. Because of this I included WT rescue cell line number 8 in my analysis of minor satellite DNA methylation as this cell line showed overexpression of LSH compared to WT. Despite LSH being overexpressed in this cell line, minor satellite DNA methylation levels were similar to other WT rescue cell lines investigated by Southern blot analysis. This suggests, that partial rescue of DNA methylation observed in the cell lines analysed is not due to low expression of LSH. Secondly, all

analysis was carried out on clonal cell lines, which were derived by seeding cells at low density and subsequently isolating colonies. Despite this, immunofluorescence analysis revealed non-homogeneous populations of cells, with some expressing LSH whereas others not to detectable levels by eye (Figure 4.5A). We therefore counted the number of cells (Rebekah Tillotson, honours project student) expressing LSH in the two WT rescue and one mutant rescue cell population (Figure 4.5B). All rescue cell populations express LSH in over 50% of cells. This suggests that 50% rescue of repetitive element and whole genome DNA methylation is not directly proportional to the number of cells expressing detectable levels of LSH. Additionally, a similar number of cells in the LSH mutant rescue population express LSH as within the wild-type LSH4 rescue cell population. Despite this, mutant rescue cells do not show increase in DNA methylation at repetitive elements and in the genome as a whole. This substantiates the conclusion that the ability of LSH to hydrolyse ATP is required for DNA methylation at repeat elements.

In summary, my analysis shows, that wild-type but not catalytic mutant LSH can partially restore DNA methylation at repetitive elements when introduced into *Lsh-/-* MEFs. This partial rescue of DNA methylation is not directly proportional to the number of cells expressing LSH in the population. Additionally, rescue of DNA methylation is not enhanced by increased levels of LSH expression. This suggests, that partial rescue of DNA methylation could have occurred due to the absence of other developmental factors or signals required.

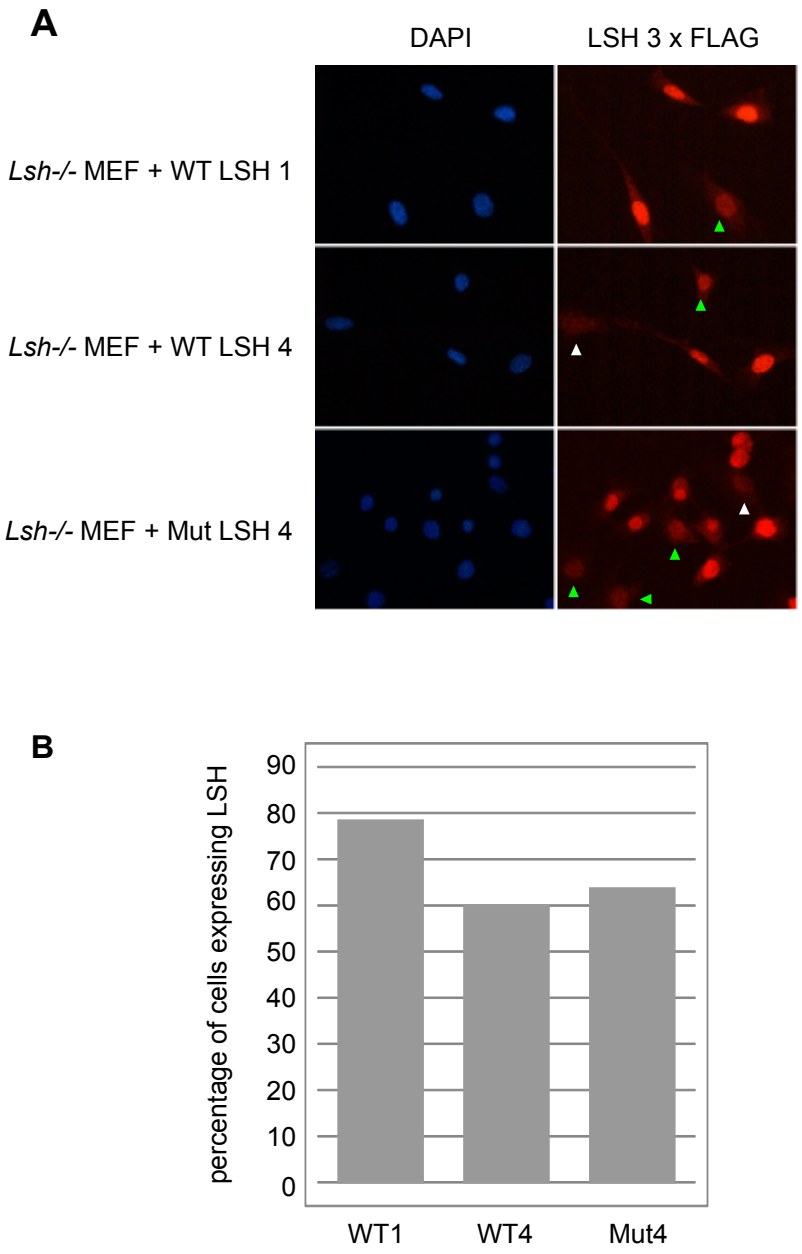


Figure 4.5 Homogeneity of LSH rescue populations.

A: Immunofluorescence images taken at 1s exposure of *Lsh-/-* MEFs expressing WT or mutant LSH. DAPI staining of the nuclei is shown on the left and anti-FLAG stained cells showing LSH expression on the right. White arrows indicate cells with nuclear FLAG staining indistinguishable from background. Green arrows indicate cells expressing LSH at very low levels, highlighting non-homogenous cell population. B: Histogram showing percentage of cells in each rescue cell line population expressing LSH to detectable levels by eye.

4.2.4 Wild-type, but not mutant LSH, restores DNA methylation at promoters of protein-coding genes in *Lsh-/-* MEFs

The next question was, whether LSH re-expression can also restore correct DNA methylation patterns at protein-coding gene promoters. In chapter 3 I described and discussed protein coding gene promoter DNA methylation defects in *Lsh-/-* MEFs. The objective was to investigate promoters classified as group I (early methylation events) and group II (late methylation events), which underwent hypomethylation in *Lsh-/-* MEFs. Together with honours student, Rebekah Tillotson, I carried out bisulfite genomic DNA sequencing on *Rhox2, Rhox6,* and *Rhox9* (Figure 4.6). These results show that *Lsh-/-* MEFs expressing wild-type LSH display elevated DNA methylation at *Rhox2, Rhox6 and Rhox9* gene promoters. Mutant LSH had no effect on DNA methylation, confirming the data obtained for repetitive elements and whole genome DNA methylation. Bisulfite sequencing on single genes was also a way to determine whether different cells in the population undergo different levels of *de novo* DNA methylation at gene promoters. Each row in a bisulfite diagram reflects a non-clonal sequence from a single cell. This generates an overview of DNA methylation in a cell population. As can be seen from bisulfite sequencing of *Rhox2* and *Rhox6*, both WT rescue cell lines show even DNA methylation in all unique sequences (apart from *Rhox2* in WT1 cell line). This suggests that even though the population of cells might not be uniform in LSH expression levels, most cells have sufficient levels of LSH to allow DNA methylation at these genes.

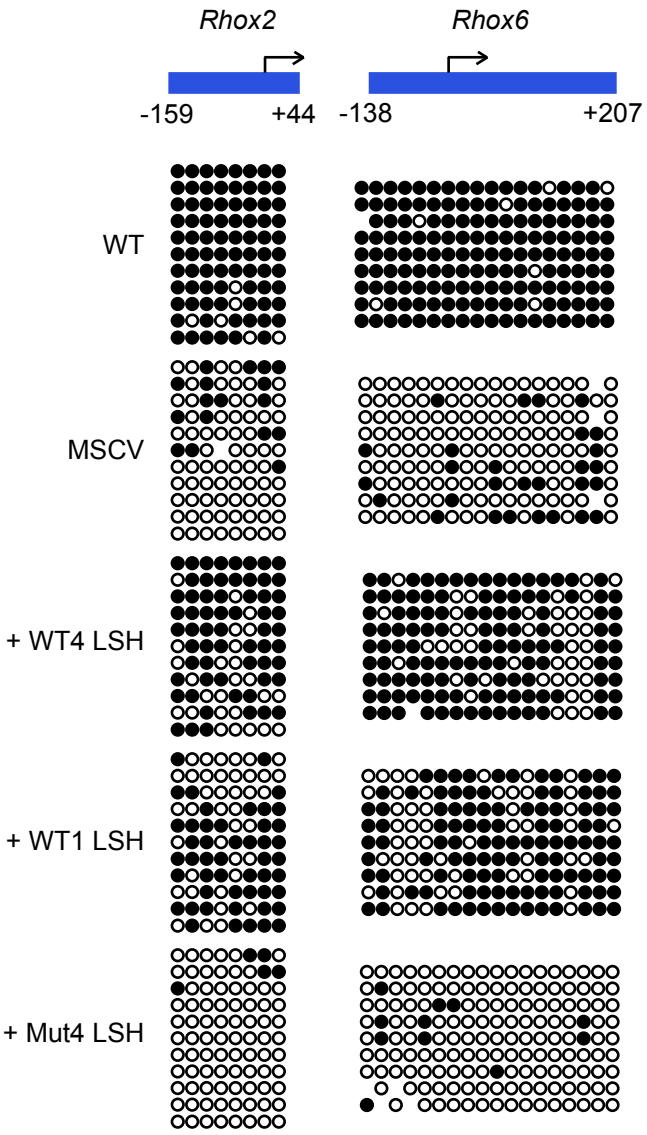


Figure 4.6 Bisulfite DNA sequencing of *Rhox2* and *Rhox6* in rescue cell lines

Black circles indicate methylated CpGs and unmethylated CpGs are indicated by white circles. Location of bisulfite sequencing product relative to transcription start site of gene is indicated above the diagram.

4.2.5 Restoration of DNA methylation at protein coding gene promoters is accompanied by rescue of gene expression defects at some genes

To investigate whether LSH dependent *de novo* DNA methylation at selected promoters is accompanied by reversal of gene expression defects, we carried out qPCR analysis of cDNA from WT4 and Mut4 cell lines (Rebekah Tillotson, honours student). This was also an easier way to assay more different genes in the rescue cell lines, than by bisulfite sequencing. Quantitative PCR was carried out on transcripts in the *Rhox* and *Hox* clusters, pluripotency genes (*Dppa4)* and imprinted genes (*Ndn, Peg12, Peg3*). These represent genes from group I as well as group II of LSH dependent methylated loci.

Expression of wild-type LSH in *Lsh-/-* MEFs leads to most genes analysed (apart from *Peg3*) being expressed to more similar levels to WT MEFs (Figure 4.7). The *Rhox* cluster is hypomethylated as well as overexpressed in absence of LSH. Gm9 is the neighbouring gene of the *Rhox* cluster shown to be hypomethylated and overexpressed as well (Myant et al. 2011). *Gm9, Rhox2a and Rhox6* show up to 10-fold repression in rescue cell lines containing WT LSH. Other transcripts analysed only display up to 2-fold repression. Re-expression of WT LSH causes reduced expression of *Hoxa4* as well as *Hoxc6*. *Dppa4* is a pluripotency gene and also shows reduction of expression in presence of WT LSH. Out of the three imprinted genes investigated, both *Ndn* and *Peg* show reduced expression in presence of WT LSH.

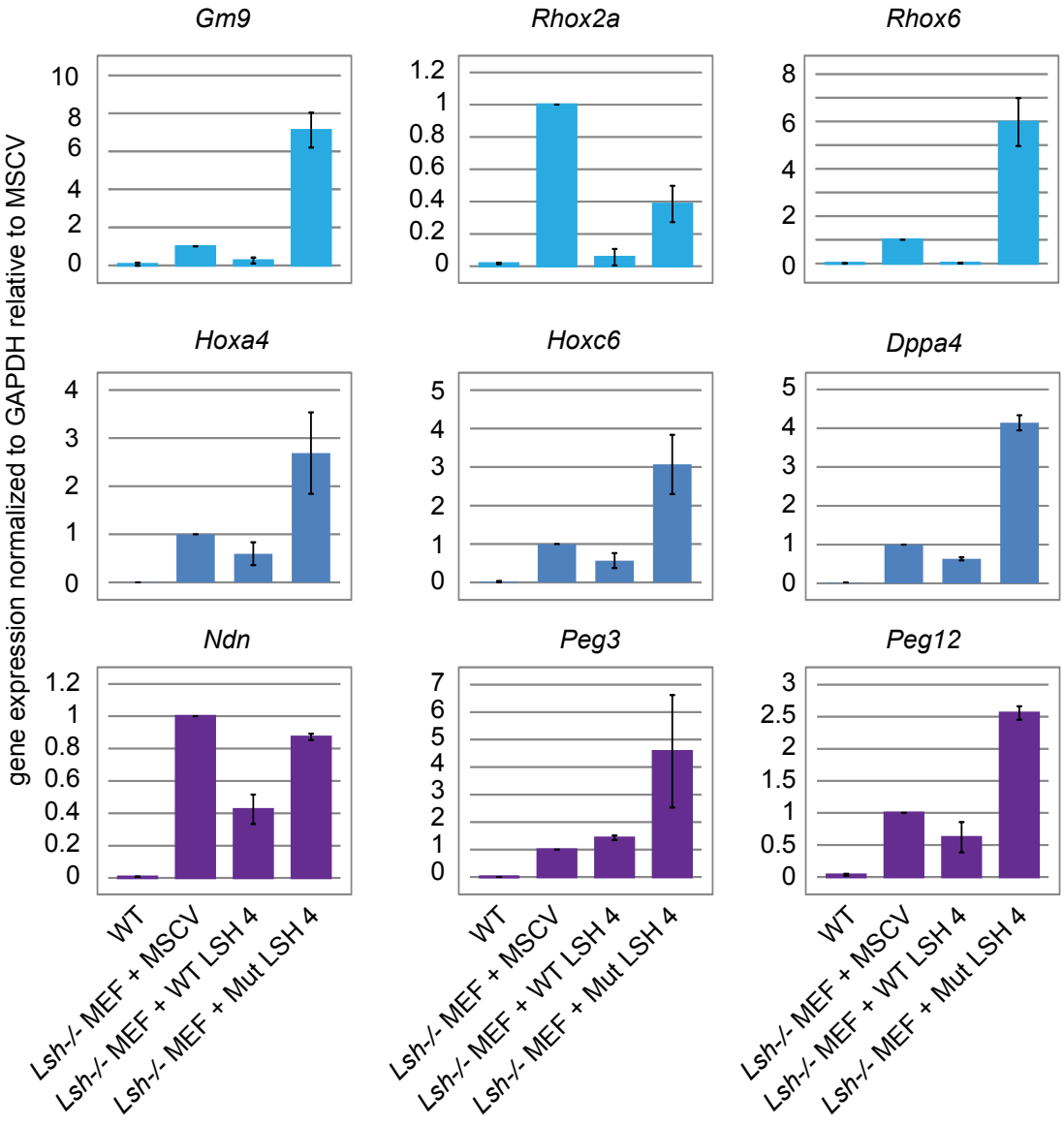


Figure 4.7 Gene expression in LSH rescue cell lines

Gene expression levels in rescue cell lines by quantitative qPCR, normalized to GAPDH. All expression values are displayed relative to MSCV control cell levels. Different groups of genes are indicated in different colours. Light blue: *Rhox* cluster. Dark blue: Differentiation associated genes. Purple: imprinted genes. Error bars represent 2 x standard deviation of triplicate experiments.

Cells expressing catalytically inactive LSH seem to be more varied in expression of the analysed transcripts (Figure 4.7). Only *Rhox2a* and *Ndn* show mildly reduced expression in the presence of mutant LSH. Other transcripts analysed show increase in expression compared to MSCV control cells. This is interesting, as it suggests that the mutant LSH has a negative effect, enhancing the defects seen in absence of LSH.

4.3 Discussion

In summary, I found that expression of WT LSH, but not catalytically inactive LSH in *Lsh-/-* MEFs results in 50% rescue of DNA methylation at repetitive sequences such as minor satellites and IAP elements. Additionally, re-expression of WT, but not mutant LSH, caused gain of DNA methylation at promoters which are hypomethylated in the absence of LSH. This was accompanied by reduced expression of associated genes. Strikingly, the Rhox cluster is a specific example where genes with increased methylation showed up to 10-fold reduced transcription upon re-expression of LSH. This resembled more their expression patterns in WT MEFs. However, re-expression of WT LSH only caused up to 2-fold reduction in Hox gene, pluripotency gene and imprinted gene expression. The methylation and expression patterns of pluripotency and *Hox* genes undergo changes during ES cell differentiation. *Hox* genes become expressed in a lineage specific manner, whereas Pluripotency genes become silenced. Imprinted gene methylation and silencing is also regulated during ES cell differentiation into different cell types. All these genes therefore have a common developmental time when they are regulated, with the help of LSH. An explanation why expression of imprinted genes may not be restored efficiently in rescue cell lines is the presence of DNMT3L at very low levels. DNMT3L has been shown to function in DNA methylation at imprinted genes and

may be required for DNA methylation of these by LSH. Pluripotency genes are silenced upon differentiation of stem cells. Since LSH is expressed at high levels in ES cells, before silencing of pluripotency genes, initiation of their inactivation must require other factors or chromatin states in addition to LSH which may not be present in MEFs. *Hox* genes are silent in stem cells and usually marked by bivalent chromatin signatures (H3K4 methylation and H3K27 methylation) which are then resolved during differentiation. In LSH knockout cells this bivalent chromatin state may not exist, which may be a pre-requisite for efficient gene silencing and subsequent DNA methylation by LSH. This suggests, that expression of LSH on its own may not be sufficient to allow their silencing. Additionally, in the presence of all required factors for DNA methylation to occur, there may only be a particular time frame (directed by signals in vivo) during which silencing and DNA methylation of early developmental genes can occur.

Immediate future work should focus on characterizing promoters of protein coding genes genome wide, to determine at which promoters LSH can facilitate DNA methylation in a cell autonomous manner. Additionally, expression array analysis would be useful to correlate re-establishment of DNA methylation with gene expression in the rescue cell lines. From the experiments in this study, I attempted to provide explanations, why gene silencing was re-established at some loci, but not others in Lsh-/- cells re-expressing WT LSH. High throughput studies may help us to more closely define these loci and therefore infer more accurately, which developmental factors/signals may be missing, preventing re-establishment of DNA methylation and gene expression at some loci.

4.3.1 Functional significance of LSH catalytic mutation – towards a functional mechanism

This study demonstrates that ATPase hydrolysis is required for LSH dependent DNA methylation at repetitive sequences, gene promoters as well as on the whole genome level. The only exception to this was *Rhox2*, which displayed reduced expression levels in presence of WT as well as mutant LSH. Intriguingly, mutant LSH did have an additional effect on the expression of genes analysed. Most genes analysed (*Rhox, Hox, Dppa* and some imprinted genes) showed increases in expression in presence of mutant LSH. This enhancement of gene expression defects in cell lines expressing mutant LSH compared to cells with completely absent LSH, suggests that the mutant LSH has a dominant effect on gene expression.

Previous studies have documented negative effects of chromatin remodeller ATPase site mutations. *Drosophila melanogaster* chromatin remodeller Brahma carrying a mutation in its ATPase site renders the protein inactive but does not prevent assembly of the BRM complex (Elfrig et al. 1998). This poses the possibility that the ATPase mutant LSH may have similar effects. Experiments in our lab have shown that catalytically inactive LSH still binds DNA as efficiently as wild-type LSH. It is on the other hand yet unknown, whether mutant LSH still binds DNMTs as well as HDACs. It is most likely that mutant LSH can still bind these factors, as the catalytic mutation only affect one amino acid. Despite this, a previous study from our lab identified the N terminal portion of LSH sufficient for silencing a reporter gene (Myant & Stancheva, 2008). This suggests that the N-terminal (amino acids 1-226) region may be required for interactions with DNMTs and HDACs. The catalytic

mutation is located between amino acids 202 and 394. Therefore interactions between mutant LSH and HDACs and DNMTs would still have to be tested.

To come to a potential model for the mechanism of how mutant LSH may cause a dominant effect on gene expression it is helpful to consider the effect of the mutant LSH on DNA methylation and gene expression again. DNA methylation is almost completely lost in *Lsh-/-* MEFs from loci requiring LSH for DNA methylation. Consistently, catalytically inactive LSH is unable to cause even more reduced levels of DNA methylation. Despite this, mutant LSH causes a dominant effect on gene expression, with LSH target genes being even more upregulated in presence of mutant LSH than in complete absence of LSH. This suggests, that mutant LSH may negatively affect silencing mechanisms such as histone deacetylation, instead of DNA methylation. One idea is that in absence of LSH, gene expression increases at analysed loci, but is still to some extent repressed by other remodellers or co-repressor complexes being able to recruit HDACs to the promoter regions of genes. Catalytic mutant LSH can bind DNA (unpublished work, Irina Stancheva) and may still be able to bind HDACs and DNMTs but the mutation in the ATPase domain would prevent LSH from translocating along chromatin. HDACs and DNMTs may therefore be permanently sequestered to a localized region, unable to be targeted throughout the locus by other repressor complexes. This may result in increased gene expression. Additionally, 'static' LSH (due to lack of ATPase hydrolysis activity) may inhibit binding of other repressive complexes to target regions. This may lead to inhibition of turnover between LSH and other repressive complexes.

These effects of the mutant LSH may result in elevated levels of acetylation throughout promoter regions compared to those in *Lsh*-/- MEFs. Chromatin immunoprecipitation of acetylated histones throughout regions spanning the promoters of LSH target regions in *Lsh*-/- MEFs and *Lsh*-/- MEFs re-expressing mutant LSH could be carried out. In *Lsh*-/- MEFs, histone deacetylation may still occur, facilitated by co-repressors other than LSH. This may preserve some extent of gene repression despite DNA hypomethylation in absence of WT LSH. The catalytic mutant LSH on the other hand may bind to chromatin together with HDACs and DNMTs. HDACs may therefore be sequestered, unable to be recruited by other co-repressors. Additionally, the catalytic mutation would prevent LSH to translocate along the promoter region of genes, to facilitate histone deacetylation. The hypothesis therefore is, that acetylation in *Lsh*-/- MEFs expressing the catalytic mutant LSH may be more widespread (with only localized deacetylation) than in *Lsh*-/- MEFs.

4.3.2 Applications of the LSH expression rescue system

I have identified that LSH can act in a cell autonomous manner at some genomic regions but not others. The regions identified to be rescued by WT LSH re-expression could therefore be used in further studies of LSH function. Different truncated or mutated forms of LSH could be re-expressed in *Lsh*-/- MEFs, using the retroviral system described above. I have proposed different models for a dominant effect of catalytically inactive LSH on gene expression. One of these argues, that mutant LSH could sequester HDACs, which may then be unavailable to be recruited to LSH target loci by other repressor complexes. This could lead to even more increased gene expression compared to *Lsh*-/- MEFs. A way to test this hypothesis

would be to make an N-terminal truncation of catalytically inactive LSH. The N-terminal domain has been proposed to be required for the interactions between LSH, DNMTs and HDACs. If LSH exerts a dominant effect by sequestering HDACs, truncation of the domain required for these interactions should lead to release of HDACs and ability for them to be recruited to LSH target genes by other repressor complexes. If on the other hand static binding of LSH prevents binding of other repressor complexes, an N-terminal mutation would release HDACs but still prevent recruitment of these by other repressors. In this case no change in gene expression would be observed.

Additionally, an N-terminal truncation on its own would also enable us to determine, whether interaction between LSH, HDACs and DNMTs is required for correct DNA methylation and gene silencing. This mutant may have a lesser effect on DNA methylation and gene expression than the *Lsh* knockout, as it may still be able to remodel chromatin, facilitating access to other repressor complexes.

Finally this system could also be used to re-introduce several factors into Lsh-/- MEFs at the same time. DNMT3B1 and DNMT3L would be candidates for this. Since these DNMTs are expressed at low levels in somatic cells, the ability for LSH to rescue DNA methylation and gene expression may be reduced.

Chapter 5 - G9a function in maintenance of imprinted DNA methylation and gene expression

5.1 Introduction

5.1.1 G9a, a histone lysine methyltransferase, with a role in DNA methylation

Increasingly more non-DNA methyltransferase proteins are found to play a role in facilitating DNA methylation. Knockout of major euchromatic histone lysine methyltransferase G9a in ES cells leads to 50% loss of DNA methylation (Dong et al, 2008). A role for G9a in *de novo* DNA methylation has been demonstrated in silencing of pluripotency genes, to prevent their reactivation after silencing (Athanasiadou et al, 2010; Epsztejn-Litman et al, 2008; Feldman et al, 2006). Additionally, a study has described G9a function in DNA methylation at the imprinted *Snrpn* gene (Xin et al, 2003). Both of these studies support a developmental role for G9a. Apart from genic regions, G9a is required for correct DNA methylation at retrotransposons and major satellite repeats as well as in proviral gene silencing (Dong et al, 2008; Leung et al, 2011). A role for G9a in maintenance of DNA methylation has been proposed based on interactions with DNMT1 as well as PCNA at replication forks (Esteve et al, 2006). This suggests a mechanism for maintaining correct DNA methylation as well as histone modifications during replication. A recent study on the other hand dismissed a role for G9a in maintenance of DNA methylation in somatic cells, suggesting that G9a may function more during early development (Sharma et al, 2012).Consistent with these roles in *de novo* and maintenance DNA methylation, G9a interacts with DNA

methyltransferases DNMT1 (at replication forks), 3A and 3B (through its ankyrin domain) (Epsztejn-Litman et al, 2008; Esteve et al, 2006).

Interestingly DNA methylation and gene repression of G9a target genes and repeat elements can be re-established in*G9a-/-* ES cells through reintroduction of a catalytically inactive G9a transgene, whilst H3K9me2 levels remain low (Figure 5.1B) (Dong et al, 2008; Zhao et al, 2008). This points towards the function of G9a in DNA methylation being independent of its histone methyltransferase activity. The role of G9a in DNA methylation in comparison with the role of LSH was interesting to me, as it may elucidate whether G9a has non-overlapping or overlapping DNA methylation targets to LSH. This may provide insight into different mechanisms to ensure correct establishment and maintenance of DNA methylation at various genomic regions.

5.1.2 Structure of G9a histone methyltransferase complex

H3K9 methylation is one of the most well studied repressive histone modification. Several different enzymes have been identified to catalyse this modification. Distinguishing features of enzymes include the type and position of histone residues they target, as well as whether they function in euchromatic or heterochromatic regions of the genome. The first identified H3K9 methyltransferase was SUV39H1/2, now widely recognised as the major heterochromatin histone methyltransferase (Peters et al, 2003; Rea et al, 2000; Rice et al, 2003). It catalyses mainly H3K9me3 with the help of a SET domain and consistent with its heterochromatin function it has been shown to interact with HP1 (Peters et al, 2003; Rice et al, 2003; Schotta et al, 2004). Several other SET domain containing histone

H3K9 methyltransferases have now been studied. These include G9a, GLP and SETDB1.

G9a is the second most well studied histone methyltransferase. Similarly to SUV39H1, it contains a catalytic SET domain and has been implicated in H3K9 methylation (Tachibana et al, 2001). The protein structure of G9a can be seen in Figure 5.1A. In contrast to Suv39H, G9a catalyses H3K9me1/2 and displays 10-20 fold higher H3K9 methylation activity (Tachibana et al, 2002).

Additionally, weak H3K27 methylation by G9a has also been demonstrated in vitro, recently argued to occur in vivo as well (Tachibana et al, 2001; Trojer et al, 2009; Weiss et al, 2010; Wu et al, 2011). The most striking difference between SUV39H1/2 and G9a came from studies of nuclear localization of these enzymes. Whereas SUV39H1/2 colocalizes with heterochromatin, G9a associates with euchromatin, making it to date the major euchromatin histone methyltransferase (Tachibana et al, 2002)

Consistently, *G9a-/-* embryos undergo dramatic reduction in H3K9 mono and di methylation. They develop severe growth impairments, leading to early lethality by E8.5 (Tachibana et al, 2002). *G9a-/-* ES cells have also been analysed and similarly display 80% loss of H3K9 di methylation compared to WT levels (Tachibana et al, 2002). Additionally H3K9 acetylation and H3K4 methylation levels, both signs of actively transcribed regions, increase in knockout cells. In contrast to embryos, *G9a-/-* ES cells remain viable but display growth defects upon differentiation. These effects can be reversed by expression of exogenous *G9a* (Tachibana et al, 2002).

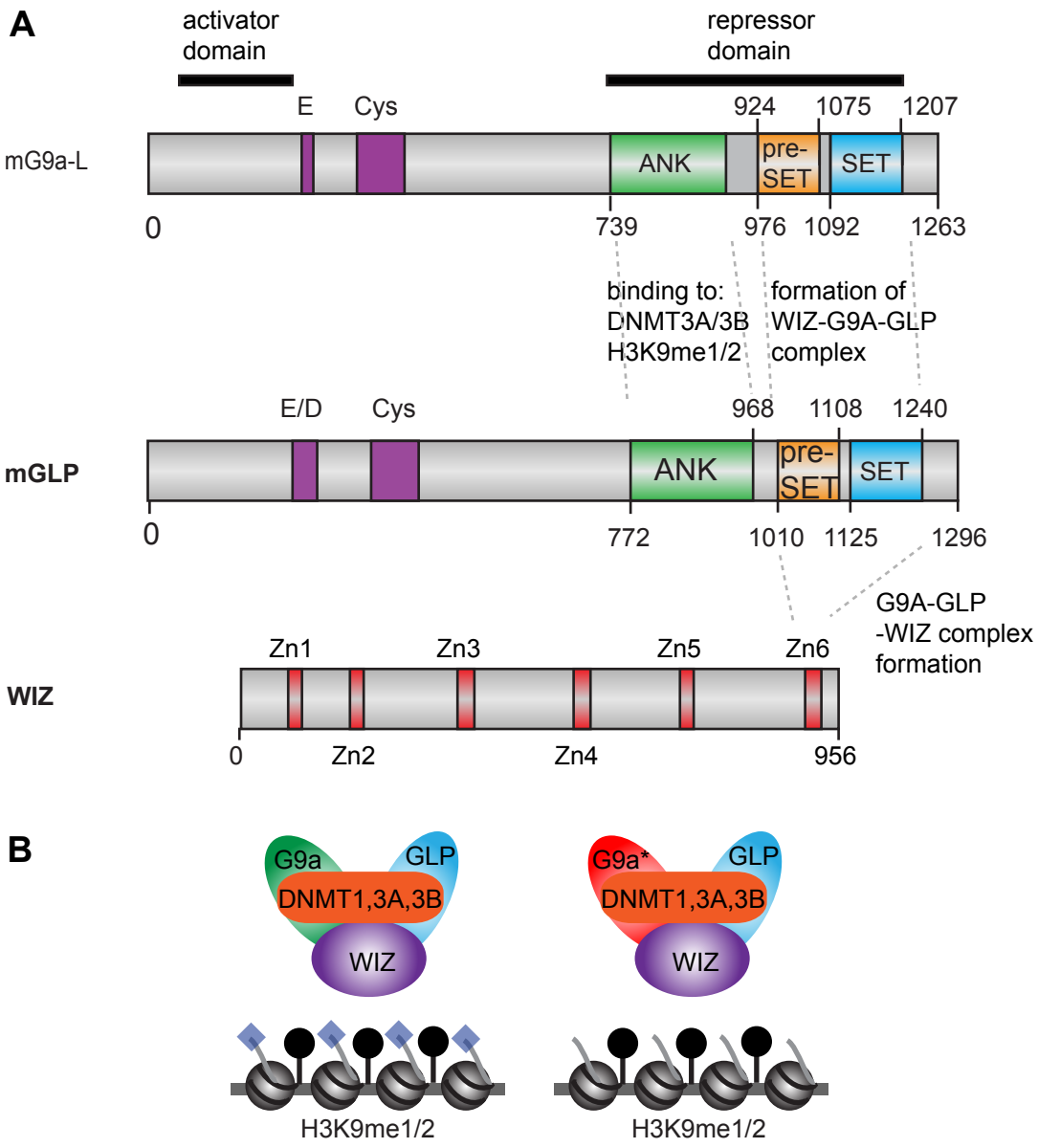


Figure 5.1 Domain structures and function of G9a and interacting partners GLP and WIZ

A: Protein domain structure (drawn to scale) of murine G9a-L, GLP and WIZ. Both G9a and GLP have N-terminal Cysteine rich regions (Cys). GLP also has a glutamate/aspartate rich region (E/D) and G9a has an aspartate rich region (E). In addition both G9a and GLP contain Ankyrine repeats (ANK) for interaction with DNMT3A/B and with H3K9me1/2 and Pre-SET and SET domains for interactions between G9a, GLP and WIZ. WIZ contains six Zinc finger motifs (Zn), the sixth of which interacts with G9a and GLP pre-SET and SET domains. All domain interactions are highlighted (Adapted from Shinkai and Tachibana 2011). **B (left)**: DNA methylation (black 'lollipops') and H3K9me1/2 (purple diamonds) at target genomic regions. **B (right):** A point mutation in the G9a catalytic SET domain prevents H3K9me1/2 by the complex but does not affect facilitation of correct DNA methylation by the complex.

Another SET domain containing histone methyltransferase, GLP, has been identified, with very similar target specificities and functions to G9a (Tachibana et al, 2005). G9a and GLP are structurally almost identical (Figure 5.1A). *Glp* knockout also leads to early embryonic lethality and mimics the phenotype of *G9a-/-* embryos (Tachibana et al, 2005). Studies have now identified the preferential formation of a heteromeric complex between G9a and GLP, dependent on their pre-SET and SET domains (Figure 5.1B). This complex results in increased stability of G9a in ES cells. Accordingly, absence of GLP is accompanied by reduced levels of G9a protein (Tachibana et al, 2005).

Another interacting partner of G9a and GLP has also been implicated in stability of the methyltransferase complex (Ueda et al, 2006). This is the zinc finger protein WIZ (Figure 5.1A). WIZ can interact with G9a or GLP on their own but is more stable in the trimolecular complex containing all three proteins (Figure 5.1B). Similar to GLP, WIZ increases the stability of G9a. Consistently knockdown of WIZ also results in reduced levels of G9a (Ueda et al, 2006).

5.1.3 G9a is required for gene repression

G9a/GLP histone methyltransferase complex has been extensively studied with regards to its effect on gene activity. HDAC independent repression of a reporter gene by G9a was demonstrated. Accordingly, several co-repressors interact with or recruit G9a to chromatin and several studies have characterized its repressive functions (Table 5.1) (Shinkai & Tachibana, 2011). Repressive functions include silencing of the pluripotency genes during stem cell differentiation and imprinted gene regulation in the placenta and ES cells, which may explain early embryo and

germ line defects in *G9a* knockout mice (Chang et al, 2006; Feldman et al, 2006; Redrup et al, 2009; Xin et al, 2003). Non-embryonic functions of G9a have also been identified. These include establishment of proviral gene silencing, silencing of neuronal genes in non-neural cells and expression of lineage specific cytokines in the immune system (Lehnertz et al, 2010; Leung et al, 2011; Roopra et al, 2004).

G9a interacting/ recruiting factors	Repressive function
SMARCAD1	Required for establishment of repressive chromatin
Human Blimp-1 homolog PRD1-BF1	Differentiation of B lymphocytes
UHRF1	Recruits transcriptional repressors (DNMTs and HDACs)
CDP/Cut	Co-repressor complex
REST/NRSF	Repression of transcription of neural genes outside nervous system
HP1	Heterochromatin formation
DNMT1,3a,3b	DNA methylation
ZNF200	Transcription regulation

Table 5.1 G9a interacting partners and repressive functions

Despite this well-known repressive function of G9a, several studies have identified a role of G9a in gene activation. Functions of G9a dependent gene activation include activation of beta globin gene transcription in adult erythroid cells, co-activator functions on nuclear receptors and gene activation during T helper cell differentiation and function (Chaturvedi et al, 2009; Lee et al, 2006; Lehnertz et al, 2010). A recent study demonstrated that gene activation is a direct function of G9a,

identifying different G9a protein domains being required for gene repression (SET domain) and activation (N terminal domain) (Figure 5.1A) (Purcell et al, 2011). This may indicate that G9a functions in gene silencing and activation at different loci as well as at different stages of development.

5.2 Aims

Only two larger scale studies have characterized genomic regions whose DNA methylation and gene expression are dependent on G9a. One study used restriction landmark genomic scanning (RLGS) analysis in *G9a-/-* ES cells. This method uses two-dimensional gel electrophoresis with genomic DNA digested with a methylation sensitive enzyme (Ikegami et al, 2007; Smiraglia et al, 2007). The other study carried out microarray analysis on *G9a* conditional knockout ES cells to detect aberrant gene expression (Ikegami et al, 2007; Yokochi et al, 2009). Despite this, both of these studies only focused on a subset of genomic regions (late replicating regions at nuclear periphery and CpG rich loci). I aimed to generate a genome wide overview of G9a dependent DNA methylation at promoters of protein coding genes. The second aim was to compare identified promoters whose DNA methylation is dependent on G9a with those whose DNA methylation is dependent on LSH.

5.3 Results

5.3.1 DNA methylation defects at promoters of protein coding genes in *G9a-/-* ES cells

I used the methylated DNA affinity purification technique, as described in Chapter 3, and promoter microarray hybridization analysis, to compare the genome wide promoter DNA methylation between WT and *G9a-/-* ES cells. Cell lines for this experiment were obtained from Yoichi Shinkai's lab by gene conversion. This involved treatment of heterozygous cells for the knockout with increased selection drug concentrations, leading to forced loss of heterozygosity through chromosomal loss. Three technical replicates were carried out in this study. Biological replicates could not be carried out at this time, due to the unavailability of *G9a-/-* mice in our lab.

Using the microarray data for purified methylated DNA, I first calculated the total numbers of methylated promoters in WT and *G9a-/-* ES cells across promoters with different CpG density, to allow for CpG density related differences in binding affinity to the MBD column (Figure 5.2A). All classes of promoters display DNA hypomethylation in *G9a-/-* compared to WT ES cells, although a slightly lower percentage of high density CpG promoters are hypomethylated in *G9a-/-* ES cells. Averages across all CpG density groups indicate that 620 promoters are methylated in WT ES cells and 436 are methylated in *G9a-/-* ES cells. This represents DNA hypomethylation of ~29% of all normally methylated promoters.

It is unclear, how fold change in enrichment of microarray signal between different cell lines relates to absolute changes in DNA methylation. Therefore I used

a high threshold to identify significant differences in DNA methylation. I highlighted all 2-fold differences in DNA methylation between WT and *G9a-/-* ES cells (Figure 5.2B). Consistent with other studies there are promoters which lose DNA methylation in absence of G9a (9 promoters) as well as promoters which gain DNA methylation (6 promoters). To determine, whether variation occurred between replicate experiments, scatter plots were plotted and Spearman correlation coefficients calcluated (Appendix 2). Average Spearman correlation coefficient for WT replicates is 0.8 and for *G9a-/-* ES cells is 0.85. These corraltion coefficients demonstrate significant correlation between replicates.

For initial validation of the MBD column and microarray experiment, I searched for previously identified promoters whose DNA methylation is dependent on G9a. These include some *Mage* genes such as *Magea-2* and *Magea-8*, which did not undergo 2-fold or greater changes in DNA methylation in *G9a-/-* compared to WT ES cells, according to my experiment. The question arose whether a 2-fold threshold was too high to detect significant differences in DNA methylation between WT and *G9a-/-* ES cells. I therefore reduced the threshold to highlight all 1.4 fold changes in DNA methylation, which also included previously identified G9a target genes (Figure 5.2C). 163 gene promoters display hypomethylation in *G9a-/-* ES cells compared to WT ES cells (Appendix 3). These promoters represent 26% of all normally methylated promoters in WT ES cells. Additionally 76 gene promoters display hypermethylation in *G9a-/-* compared to WT ES cells (Appendix 3).

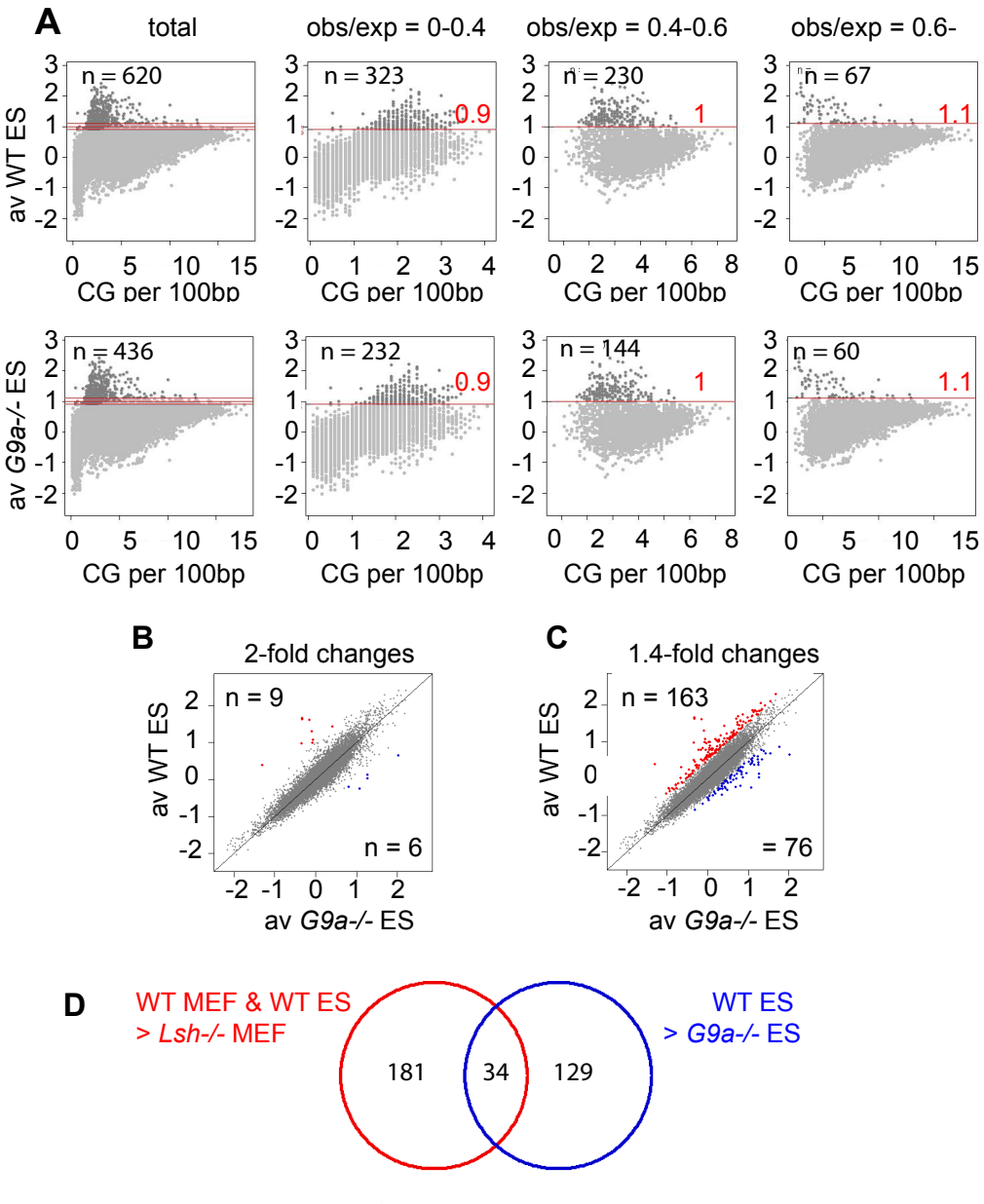


Figure 5.2 DNA methylation analysis in WT and *G9a-/-* ES cells

A: Total number of methylated promoters in WT and *G9a-/-* ES cells. Median log2(MAP/Input) values are plotted against number of CGs per 100bp. Different thresholds were used for quantifying total numbers of methylated promoters in low (0.9), intermediate (1) and high (1.1) CpG denstiy classes to account for differences in binding to the MBD column. n indicates number of promoters methylated. Left pannel shows an average over all CpG density classes. **B:** Scatter plot comparing DNA methylation levels in WT and *G9a-/-* ES cells. All WT log2(MAP/Input) values are plotted agains *G9a-/-* log2(MAP/Input) values. A threshold was applied to indicate promoters undergoing 2-fold hypomethylation (red) or hypermethylation (blue) in *G9a-/-* compared to WT ES cells. C: Scatter plot highlighting all 1.4 fold changes in DNA methylation. **D:** Venn Diagram of overlapping G9a and LSH target promoters.

5.3.2 There is some overlap between G9a and LSH target promoters

Comparing the number of promoters whose DNA methylation depends on G9a with those whose DNA methylation depends on LSH, there are fewer G9a target promoters than LSH target promoters. This may suggest the requirement for G9a in maintenance of DNA methylation at a smaller and more defined group of promoters. Following this, I decided to define the overlap between promoters requiring G9a or LSH for their DNA methylation. The Venn diagram in Figure 5.2D outlines the extent of overlap between G9a and LSH dependent promoters. 34 promoters require both LSH and G9a to ensure correct DNA methylation. In summary, 21% of promoters which require G9a for their correct DNA methylation also require LSH. This suggests that G9a and LSH facilitate DNA methylation at different promoters, but that some promoters require both proteins for full methylation.

Bisulfite sequencing analysis carried out by others in the lab showed that reproductive homeobox genes (*Rhox*) are hypomethylated in absence of G9a. *Rhox2* and *Rhox9* are 70-80% hypomethylated in *G9a-/-* ES cells. In *Lsh-/-* MEFs, these genes are hypomethylated almost completely. Further analysis suggested that LSH is required for correct targeting of G9a to the *Rhox* promoters. In *Lsh-/-* MEFs G9a was no longer bound to these loci. This demonstrates that LSH facilitates the recruitment of G9a to *Rhox* gene promoters (Myant et al, 2011). Both LSH and G9a allow correct establishment of DNA methylation at these promoters. In absence of G9a, LSH can still facilitate DNA methylation, although only partially. In absence of LSH, G9a fails to be recruited and neither factor is present to allow DNA methylation to be established. This would explain the more severe effect of the LSH knockout on DNA methylation at these loci.

My microarray data on DNA methylation in WT and *G9a-/-* ES cells did not highlight the need for G9a for DNA methylation at *Rhox* promoters above the threshold of significance (1.4 fold differences in log2MAP/input). This suggests that DNA methylation differences between WT and *G9a-/-* ES cells may have been underestimated during methylated DNA affinity purification and microarray hybridization. However, my data provide a more comprehensive list than before, of gene promoters whose DNA methylation depends on G9a. Further experiments may be needed to generate a more complete list of promoters whose DNA methylation is dependent on G9a.

5.3.3 Imprinted gene promoters are hypomethylated in absence of G9a

Strikingly, within the 163 hypomethylated gene promoters in the absence of G9a, I found several promoters of imprinted genes (Figure 5.3A).

Imprinting is an epigenetic phenomenon, which leads to parent-of-origin allele specific expression of a gene. Imprinting was first identified when embryos containing only paternal or maternal chromosomes failed to develop to term (Barton et al, 1984; McGrath & Solter, 1984). An imprinted status is achieved through differential epigenetic modifications (DNA methylation and histone modifications) on the two parental alleles. As described in the introduction, differentially methylated imprinting marks are established during DNA methylation reprogramming in developing germ cells. These regions are the so called Imprinting Control Centres (ICRs) whose DNA methylation is stable in the second round of DNA methylation reprogramming in the pre-implantation embryo (Arnaud, 2010). Imprinted genes are located in clusters, each controlled by one, or in some cases two ICRs (Bartolomei & Ferguson-Smith, 2011).

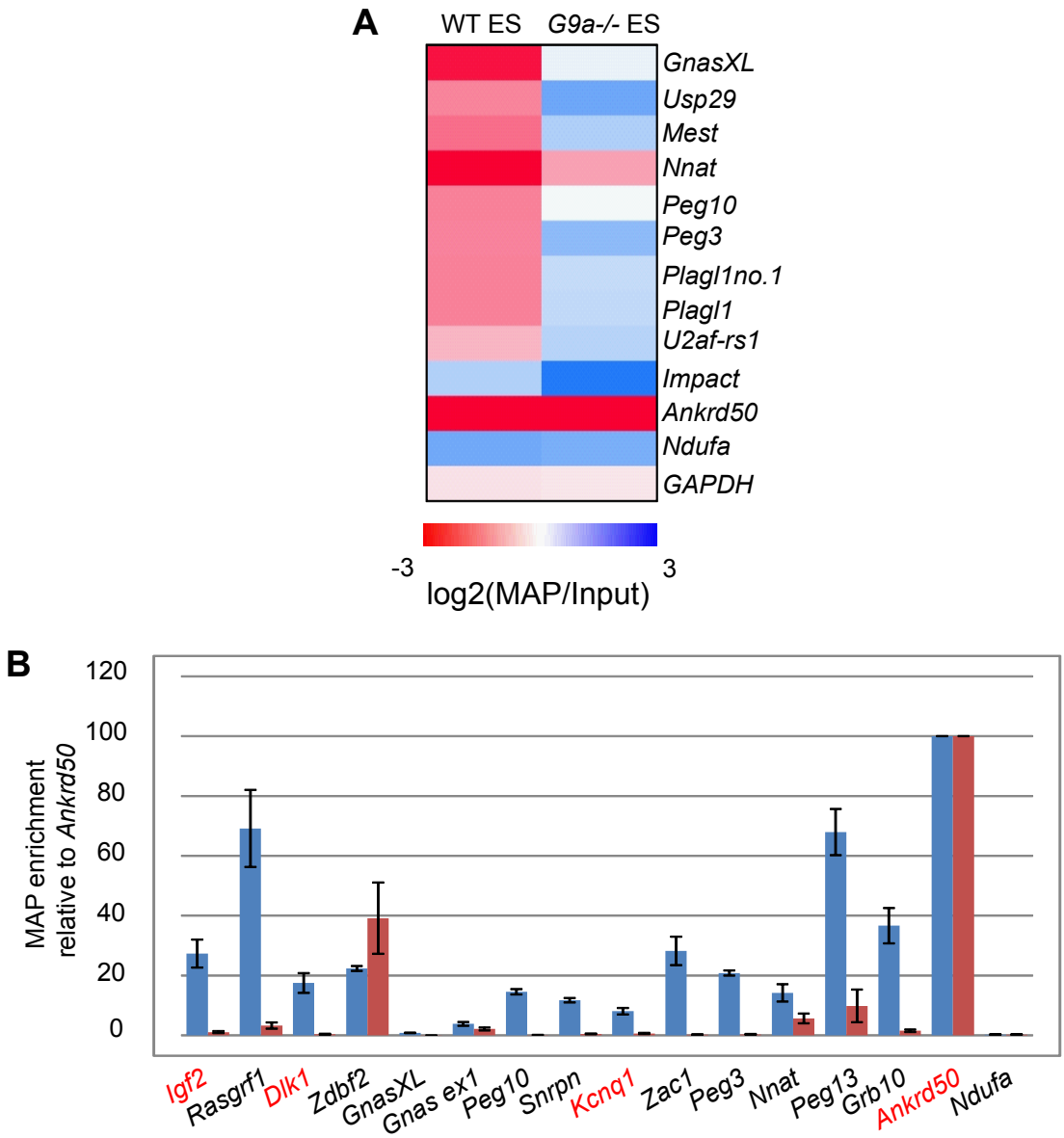


Figure 5.3 Imprinting control region (ICR) DNA methylation analysis in WT and *G9a-/-* ES cells

A: Heatmap displaying DNA methylation differences between WT and *G9a-/-* ES cells at imprinted gene promoters. Differences are diplayed with a blue-white-red colour gradient of increasing DNA methylation. *Ankrd50* is used as positive control, methylated in both cell lines. *Ndufa* and *GAPDH* are negative controls, unmethylated in both cell lines. These controls were identified from the MBD microarray experiment **B:** Quantitative PCR of Imprinting Control Region (ICR) enrichment in the methylated DNA fraction compared to input from MBD purification of WT (blue) and *G9a-/-* ES (red) cell gDNA. Error bars highlight 2 x standard deviation of technical replicates. All values represent methylated DNA fraction enrichment compared to Input, normalized to *Ankrd50*. ICRs indicated in red were subsequently validated by bisulfite sequencing

Many imprinted genes have secondary differentially methylated regions (DMRs) associated with them (Bartolomei & Ferguson-Smith, 2011). DNA methylation at secondary DMRs is often established in a cell and developmental timing specific manner.

22 germ line DMRs have been found to date (Table 5.2) (Arnaud, 2010). 12 germ line DMRs are maternally imprinted and are located within protein coding gene promoters and are therefore represented on the promoter methylation microarray (Table 5.2). Paternally imprinted ICRs are not represented on the microarray, as they are intergenic (Table 5.2). Additionally, ICRs located in promoters of non-coding RNAs (such as *Kcnq1ot1*) are also absent from the microarray.

Eight germ line DMRs represented on the microarray are hypomethylated in absence of G9a (Figure 5.3A). This is intriguing, as imprinted DNA methylation in the germ line is thought to be established and stably maintained throughout development. Hypomethylation of germ line imprints in absence of G9a is an important discovery, as recent studies identified other factors (such as ZFP57/STELLA) which may explain how imprinted regions escape various epigenetic reprograming events during development (Messerschmidt, 2012; Nakamura et al, 2007).

	germ line DMR	imprint	on microarray	1.4 fold change
Dlk1	between Dlk1 and Gtl2	paternal		
Rasgrf1	upstream of Rasgrf1	paternal		
Igf2	upstream of H19	paternal		
Zdbf2	upstream of Zdbf2	paternal		
Mcts2	promoter	maternal		
Gnas	GnasXL/Nespas promoter	maternal	x	x
	Gnas 1A promoter	maternal	x	
Peg10	promoter	maternal	x	x
Mest	promoter	maternal	x	x
Nap1L5	promoter	maternal	x	
Peg3	promoter	maternal	x	x
Snrpn	promoter	maternal	x	
Inpp5f	promoter	maternal	x	
Kcnq1	Kcnq1ot1 promoter	maternal		
Plagl1	promoter	maternal	x	x
Grb10	promoter	maternal	x	
U2af1-rs1	promoter	maternal	x	x
KcnK9	promoter	maternal	x	
Slc38aA	promoter	maternal		
Igf2r	Airn promoter	maternal		
Impact	promoter	maternal	x	x
Nnat	promoter	maternal	x	x

Table 5.2 All known Imprinting Control Regions (and whether they are maternal or paternally imprinted are shown). Indicated are the ones represented on the microarray and those which display 1.4-fold hypomethylation.

5.3.4 Most Imprinting Control Regions are hypomethylated in *G9a-/-* ES cells

To expand these observations, I decided to carry out quantitative PCR on the enrichment of imprinted regions in the input DNA and purified methylated DNA fraction used for the microarray analysis. I included most of the maternal germ line DMRs appearing in the microarray analysis (Table 5.2). Additionally I quantified the enrichment of some other maternal germ line DMRs located at non-coding RNA promoters, such as *Kcnq1ot1,* as well as the four paternal germ line DMRs (*Dlk1, Rasgrf1,Igf2/H19 and Zdbf2*). *Ankrd50* and *Ndufa* were used as positive and negative controls respectively as they did not show changes in DNA methylation between WT and *G9a-/-* ES cells in the microarray data (Figure 5.3A).

Quantitative PCR allowed me to determine the enrichment of germ line imprinted regions in the methylated DNA fraction compared to input DNA in WT and *G9a-/-* ES cells (Figure 5.3B). All germ line imprinted regions analysed, apart from *Zdbf2*, were hypomethylated in *G9a-/-* ES cells compared to WT ES cells. The reason why I did not identify some of these ICRs to be significantly hypomethylated in promoter microarray analysis, may come from imprinted regions only being 50% methylated in WT ES cells. Only one allele of these regions will bind to the MBD column, reducing the enrichment in the methylated fraction (compare imprinted gene MAP enrichment to Ankrd50 in WT ES). Any loss of DNA methylation may therefore be less detectable than loss at non imprinted regions. Quantitative PCR may be a more sensitive method of quantifying differences in enrichment in the methylated DNA fraction.

Validation of Imprinting Control Region hypomethylation

To validate the observation that DNA methylation at imprinting control regions is dependent on G9a, I carried out bisulfite genomic DNA sequencing on ICRs located in the *Igf2-H19*, *Dlk1* and *Kcnq1* imprinted clusters. Additionally I included the *Uty* gene promoter (a non-imprinted region control affected by G9a absence), as well as *Ankrd50* (a highly methylated region showing no change in DNA methylation in G9a-/- ES cells) in the analysis (Figure 5.4). Although MBD column and microarray analysis did not indicate a significant change in DNA methylation at *Ankrd50* gene promoter in *G9a-/-* ES cells compared to WT ES cells, bisulfite DNA sequencing displays a mild reduction in DNA methylation (74% in WT ES cells and 63% in *G9a-/-* ES cells). As Ankrd50 was used for normalization in the above qPCRs, the changes in DNA methylation at imprinted regions may have been underestimated.

Bisulfite sequencing of *IG DMR*, *KvDMR* and *Igf2/H19* DMR indicate almost complete lack of DNA methylation at these genomic regions in *G9a-/-* ES cells (Figure 5.4). Bisulfite sequencing and MeDIP analysis on other imprinting control regions have also confirmed my observations (Tuo Zhang, unpublished). Not all ICRs are 50% methylated (indicating DNA methylation on one parental allele only) in WT ES cells, especially IG DMR (Figure 5.4). A recent study identified dynamic changes in the size of germ line DMRs during early development (variable outer boundaries of DMRs) (Tomizawa et al, 2011). The IG-DMR region I sequenced falls slightly out with the blastocyst DMR, but within the embryonic DMR. As ES cells originate from the blastocyst, the region I investigated by bisulfite sequencing may be on the boundary of the differentially methylated region.

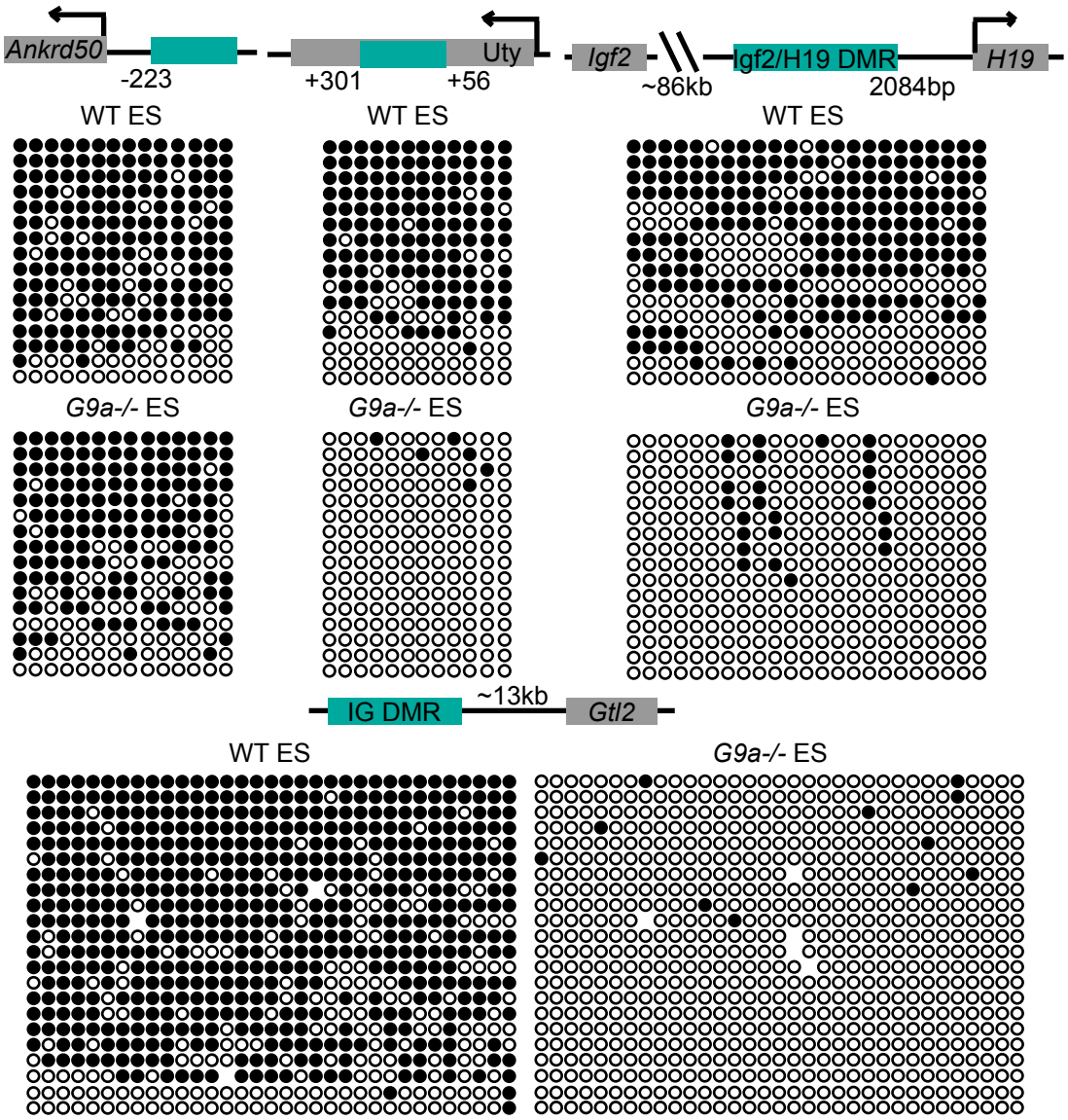
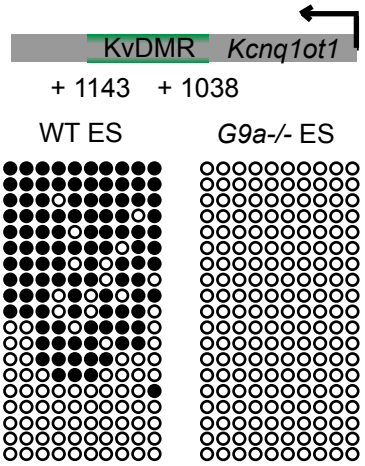


Figure 5.4 Validation of DNA hypomethylation



Bisulfite DNA sequencing validation of germline imprinted region DNA hypomethylation in *G9a-/-* ES cells compared to WT ES cells. *Ankrd50* was used as a positive control which displayed high levels of DNA methylation in both cell lines in the MBD column microarray experiment. *Uty* is a non-imprinted region whose DNA methylation is dependent of G9a. *KvDMR*, *Igf2/H19* and *IG DMR* are germline imprinted regions. Regions to which the bisulfite sequenced data maps to (green) in relation to adjacent genes (grey) are shown above bisulfite profiles. Transcription start sites are indicated by directional arrows.

The region I sequenced may therefore not display a characteristic differential methylation pattern between the two parental alleles. Additionally, WT ES cells often display unstable DNA methylation, especially at imprinted regions. Other divergence from 50% methylation in WT ES cells in my data may therefore be due to this instability. Finally, bisulfite sequencing of *Uty* (a non-imprinted control identified to be hypomethylated in *G9a-/-* ES cells on the promoter microarray) also displays extensive hypomethylation in absence of G9a.

From the above analysis, absence of G9a in murine ES cells results in loss of DNA methylation at the majority of Imprinting Control Regions. The question was, whether these DNA methylation defects are accompanied by changes in gene expression of imprinted genes regulated by these ICRs. As discussed in the introduction, in many cases DNA methylation is considered to be the secondary epigenetic mark, lack of which does not necessarily lead to expression changes at genes. Two clusters previously shown to encompass imprinted genes expressed in ES cells are the *Dlk1* and the *Kcnq1* clusters. The *Dlk1* cluster is controlled by a paternally imprinted ICR, whereas the *Kcnq* ICR is maternally imprinted. My attempts at qPCR analysis of imprinted gene expression within the *Dlk1* and *Kcnq1* clusters in WT and *G9a-/-* ES cells were unsuccessful. Many imprinted genes, although imprinted, are lowly expressed in ES cells, making qPCR analysis of these difficult (Li et al, 2010). Therefore I decided to determine their expression levels in embryoid bodies (EBs), derived from WT and *G9a-/-* ES cells after differentiation induced by withdrawal of leukaemia inhibitory factor (LIF) and addition of retinoic acid (Bibel et al, 2004).

5.3.5 Expression defects of genes at the *Dlk1/Gtl2* imprinted locus in *G9a-/-* EBs

The *Dlk1* locus is located on chromosome 12 and consists of three paternally expressed protein coding genes *Dlk1*, *Rtl1* and *Dio3* as well as an array of maternally expressed non-coding RNAs (Figure 5.5A) (da Rocha et al, 2008). A germ line imprinting control regions is located between the *Dlk1* and *Rtl1* genes (da Rocha et al, 2008). This region is unmethylated on the maternal allele and methylated on the paternal allele. Deletion studies of the imprinting control region from both parental chromosomes point towards a role of the maternal unmethylated ICR in imprinting control of all genes within the locus (Lin et al, 2003). Deletion of the maternal unmethylated ICR leads to 'epigeneotype' switching of the maternal allele to resemble the paternal allele gene expression patterns. One hypothesis is that the unmethylated ICR promotes the expression of non-coding RNAs within this cluster, which in turn regulate the silencing of protein coding genes. A second hypothesis is, that the ICR directly regulates protein coding gene expression and repression on the maternal chromosome (Lin et al, 2003). Deletion of the ICR from the paternal chromosome on the other hand had no effect on expression of genes within the imprinted cluster. This suggests that there may be another mechanism to activate protein coding gene promoters on the paternal chromosome, which is independent of the ICR (Lin et al, 2003).

DNA methylation is important in regulation of protein coding gene expression on the paternal chromosome as seen by silencing of *Dlk1* in *Dnmt1-/-* embryos (Hirasawa et al, 2008). Therefore the question I posed is, whether DNA hypomethylation of IG-DMR, in absence of G9a, is required for correct expression

of imprinted genes within the *Dlk1* locus. I hypothesized, that loss of DNA methylation at the paternal ICR may cause the paternal allele to resemble the maternal allele in terms of gene expression status. Quantitative PCR analysis showed that in the absence of G9a and DNA methylation at the ICR, *Dlk1* was repressed, consistent with the idea that DNA methylation is required for expression of *Dlk1* (Figure 5.5C). *Rtl1* and *Dio3* on the other hand were not repressed (Figure 5.5C). *Rtl1* was upregulated and *Dio3* is only marginally repressed. The qPCR detection of *Rtl1* may have been biased due to the expression of microRNAs which overlap with the *Rtl1* transcript (Figure 5.5A) and may therefore have been detected in the qPCR analysis. In the absence of DNA methylation at the paternal ICR, non-coding RNAs would be expected to become biallelically expressed. This may dilute down the effect of repression of *Rtl1* seen in this analysis. *Dio3,* on the other hand, has been found to be partially imprinted, with the paternal allele accounting for only 84% of the expression and leaky transcription from the maternal allele in embryos and the placenta (Tsai et al, 2002). The absence of ICR DNA methylation on the paternal chromosome in *G9a-/-* ES cells may generate two chromosomes with leaky transcription, reducing the extent of loss of gene expression seen. The reduction in expression of *Dio3as* cannot be explained by the hypothesis, that loss of DNA methylation at the ICR results in epigenotype reversion.

Snrpn was used as a positive control in this experiment, as it has previously been identified to be biallelically expressed in the absence of G9a (Xin et al, 2003). Consistently Snrpn expression doubles in *G9a-/-* embryoid bodies (Figure 5.5C).

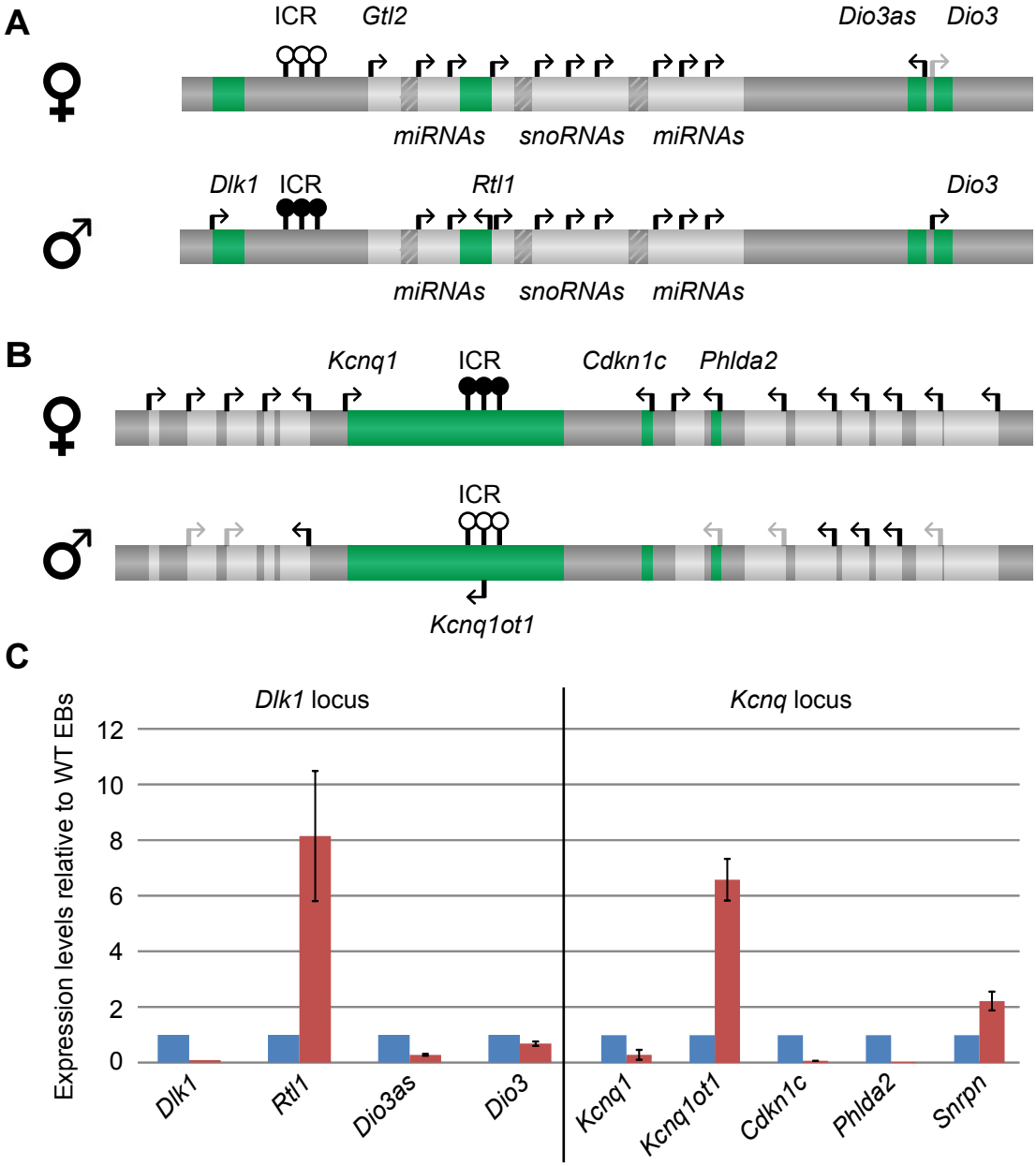


Figure 5.5 Gene expression at the *Kcnq1* and *Dlk1/Gtl2* imprinted loci

A and B: Structure and transcription status of maternal and paternal chromosomes at two imprinted loci (*Kcnq1* and *Dlk1-Gtl2* loci). Transcription start sites are depicted by directional arrows. Light grey arrows indicate leaky transcriptional activity. Differentially methylated ICRs are depicted by black (methylated) and white (unmethylated) 'lollipops'. Some genes are completely silent on one parental chromosome (no arrow). **C:** Quantitative PCR on imprinted genes within the *Kcnq1* and *Dlk1-Gtl2* loci in WT (blue) and *G9a-/-* differentiated ES cells (red) (embryoid bodies day 8). qPCR results were normalized to *Gapdh*. Error bars represent 2 x standard deviations of technical triplicates.

5.3.6 Expression defects of genes at the *Kcnq* imprinted locus in *G9a-/-* EBs

The *Kcnq1* cluster is located on chromosome 7 and consists of 12 differentially expressed genes (Figure 5.5B) (Ideraabdullah et al, 2008). Amongst these is the antisense non-coding RNA (ncRNA) *Kcnq1ot1* whose promoter is located within intron 10 of the *Kcnq1* gene (Ideraabdullah et al, 2008). This cluster is regulated by one imprinting control element (KvDMR1) which overlaps with the *Kcnq1ot1* promoter (Ideraabdullah et al, 2008). This ICR is methylated on the maternal chromosome and unmethylated on the paternal chromosome (Engemann et al, 2000; Ideraabdullah et al, 2008). The unmethylated ICR on the paternal chromosome allows transcription of the *Kcnq1ot1* ncRNA (Ideraabdullah et al, 2008)(Smilinich et al. 1999). All other imprinted genes are expressed primarily from the maternal chromosome, although some leaky expression is seen on the paternal chromosome.

Deletion studies have also been undertaken in this locus. Paternal deletion of KvDMR1 resulted in disruption of silencing of six genes on both sides of the deletion (Fitzpatrick et al, 2002; Mancini-Dinardo et al, 2006). Maternal deletion of the ICR, on the other hand did not lead to imprinted expression defects. Further insight into the regulation of imprinting at this cluster came from deletion studies of the *Kcnq1ot1* promoter. On the paternal chromosome this resulted in activation of normally silent genes. Equally, premature termination of *Kcnq1ot1* transcript on the paternal chromosome led to imprinted expression defects (Fitzpatrick et al, 2002). These observations confirmed that the *Kcnq1ot1* transcript or the complete transcription process of *Kcnq1ot1* is required for correct imprinting in this cluster

(Mancini-Dinardo et al, 2006). Deletion on the maternal chromosome again had no effect.

I carried out qPCR analysis of expression of several genes within the *Kcnq1ot1* cluster in WT and *G9a-/-* EBs (Figure 5.5C). I detected an increase in *Kcnq1ot1* gene expression and reduction in expression of Kcnq1, *Cdkn1c* and *Phlda2*. Unlike the *Dlk1* cluster, all genes follow my hypothesis that loss of DNA methylation at the ICR would lead to reversal of the previously methylated maternal allele to a state similar to the paternal allele. This hypothesis would have to be confirmed by either Fluorescence In Situ Hybridization (FISH) to detect expression from both parental chromosomes. Alternatively analysis of gene expression could be carried out in cells derived from strains containing parental allele specific polymorphisms, which can be used to distinguish expression from both alleles on mRNA level.

The question arising from my work is, whether it is the lack of DNA methylation or histone modification function of G9a, which causes the defects in imprinted gene expression in *G9a-/-* ES cells. Other work in our lab demonstrated that DNA methylation at ICRs could not be re-established in *G9a-/-* ES cells through introduction of a G9a transgene (Tuo Zhang, unpublished). This demonstrates, that G9a is unable to establish imprinting marks in ES cells, without earlier developmental signals. Additionally, this observation means, that a catalytically inactive G9a cannot be used to test whether DNA methylation at ICRs is able, on its own, to ensure correct expression of imprinted genes. Alternatively, an inhibitor of G9a catalytic activity could be used to address this question. This would help to identify, whether the absence of the histone methyltransferase function of G9a is

responsible for the gene expression defects at imprinted genes rather than absence of the DNA methylation function in *G9a-/-* ES cells.

5.4 Discussion

In summary, I set out to characterize dependence of DNA methylation at promoters of protein coding genes on G9a, in mouse ES cells. Methylated DNA affinity purification and subsequent hybridization to mouse promoter microarrays allowed genome-wide characterization of DNA methylation patterns at promoters of protein-coding genes in *G9a-/-* compared to WT ES cells. Strikingly, amongst these were several maternally imprinted regions, which displayed decrease in DNA methylation. Further, analysis showed that most maternal and paternal germ line Imprinting Control Regions (ICRs) were hypomethylated in absence of G9a. Additionally, this was accompanied by expression defects of several genes regulated by these ICRs. These observations are intriguing, since stable imprinting is essential in the developing embryo and adult mammals and imprinted regions normally escape epigenetic reprogramming during development. Such dramatic defects in imprinted DNA methylation or gene expression have only been seen upon deletion of maternal and zygotic *Dnmt1*, zinc finger protein Zfp57(Quenneville et al, 2011), Kap1/Trim28 and PGC7/STELLA (Hirasawa et al, 2008; Messerschmidt, 2012; Nakamura et al, 2007). ZFP57 and PGC7/STELLA have recently been characterized in detail for their role in establishing and maintaining imprinting marks. Hypomethylation of ICRs in G9a-/- ES cells may therefore point towards a role of G9a in maintenance of DNA methylation at maternal and paternal ICRs in conjunction with these other proteins.

5.4.1 Mechanism of G9a dependent maintenance of DNA methylation at ICRs

G9a function in *de novo* or maintenance DNA methylation?

The *G9a-/-* ES cells used in this study were obtained through gene conversion of heterozygous cells by exposure to high concentrations of selection antibiotics (Tachibana et al, 2002). The ES cells used for this were initially derived from wild-type blastocysts. As discussed in the introduction, ICR DNA methylation is set up in the parental male and female germ line at regions called Imprinting Control Regions. ES cells derived from WT blastocysts would therefore have correctly established ICR DNA methylation. Any defects in ICR DNA methylation and imprinted gene expression seen in this study can therefore be attributed to a defect in the maintenance, rather than establishment of imprinting. This is consistent with the knowledge that G9a interacts with maintenance DNA methyltransferase DNMT1 at replication forks. Whether G9a is also implicated in *de novo* DNA methylation at imprinted regions is yet unknown. Previous studies have identified interactions of G9a with DNMT3A and DNMT3B as well as roles in *de novo* DNA methylation at non-imprinted genes (discussed in the introduction). Suggesting an additional function of G9a in *de novo* methylation at imprinted regions is the dramatic increase in *G9a* expression in postnatal day15 mouse oocytes (Kageyama et al, 2007). This is the stage at which maternal ICR *de novo* methylation occurs.

At which developmental stage does G9a maintain DNA methylation?

After establishment of imprinting marks in the germ line, they are stably maintained throughout mammalian development, even during genome-wide DNA methylation reprogramming in the pre-implantation embryo. At this stage, several

questions arise, regarding the mechanisms which ensure stability of DNA methylation at imprinting control regions. The first wave of DNA demethylation in the pre-implantation embryo occurs in the paternal pronucleus through an active pathway prior to the first cell division (Saitou et al, 2012; Santos et al, 2002). How are paternal imprinted DNA methylation marks maintained at this stage? At the onset of the first cell divisions, both maternal and paternal genomes undergo passive DNA demethylation (Saitou et al, 2012; Santos et al, 2002). Again, at this stage methylation at imprinted regions has to be stably maintained in both parental genomes. The differences in dynamics of DNA methylation of the parental genomes were suggested to potentially point to the need for different mechanisms to maintain maternal and paternal imprints. The loss of DNA methylation at most maternal and paternal ICRs in *G9a-/-* ES cells is therefore surprising.

Upon completion of DNA demethylation of the genome, a wave of DNA *de novo* methylation leads to re-establishment of DNA methylation specific to cells destined to become embryonic as well as extraembryonic tissues (Saitou et al, 2012; Santos et al, 2002). This poses the opposite question, of how monoallelic DNA methylation at imprinted regions is maintained, whilst other regions undergo *de novo* DNA methylation. Finally, long term stability of imprinted regions in somatic cells also has to be ensured.

Defects in DNA methylation of imprinting control regions in *G9a-/-* ES cells in my study suggest the requirement of G9a in long term maintenance of DNA methylation. This function of G9a was confirmed through knockdown of G9a in ES cells, which also resulted in hypomethylation of ICRs (Tuo Zhang, unpublished). A yet unanswered question is, whether G9a is also required for maintenance of DNA

methylation during genome reprograming. An answer to this question could be obtained from analysis of ICR DNA methylation at the stages of active and passive genome DNA demethylation during early development. If ICRs lose DNA methylation, G9a may be required for protecting ICRs from demethylation during genome reprograming. Previous studies have argued that G9a is not required for DNA methylation at imprinted loci in embryos, since imprinting defects were only demonstrated in *G9a-/-* ES cells, but not embryos (Xin et al, 2003). This gives rise to difficulties in studying G9a function in embryonic development. Potential explanations for these observations will be discussed later.

How does G9a prevent demethylation?

An additional question is, whether G9a prevents passive DNA demethylation through co-operation with DNMT1, or does G9a inhibit active DNA demethylation pathways. To test whether G9a protects ICRs from active DNA demethylation, components of active DNA demethylation pathways (such as TET proteins) could be inhibited at the same time as knocking down *G9a* in ES cells. If inhibition of active DNA demethylation pathways prevents ICR hypomethylation in absence of G9a, G9a may function to prevent active demethylation.

How does G9a fit into current models of DNA methylation at imprinted regions?

The combination of several other studies into proteins required for correct establishment and maintenance of DNA methylation at ICRs have led to a proposed model (Figure 5.6) (Strogantsev & Ferguson-Smith, 2012). Establishment of DNA methylation at imprinted regions requires mainly DNMT3A and DNMT3L (Figure 5.6A)(Kaneda et al, 2010; Kaneda et al, 2004). DNMT3B has been implicated in *de*

novo methylation of some regions such as *Rasgrf1*, a paternal ICR (Watanabe et al, 2011). Another protein, ZFP57, has a minor role in *de novo* DNA methylation at ICRs (at the *Snrpn* locus) (Li et al, 2008). ZFP57 recruits a co-repressor KAP1, which in turn recruits SETDB1 histone lysine methyltransferase and NuRD co-repressor complex (Lechner et al, 2005; Schultz et al, 2002; Underhill et al, 2000). Interestingly ZFP57 binds specifically to a motif containing methylated CG dinucleotides (Quenneville et al, 2011). This is more consistent with a role in maintenance of DNA methylation.

Maintenance of DNA methylation at ICRs requires ZFP57, complexed with co-repressor molecules (Figure 5.6B). Additionally, ZFP57 interacts with maintenance methyltransferase DNMT1 as well as *de novo* methyltransferases DNMT3A and 3B (Quenneville et al, 2011; Zuo et al, 2012). DNMT1 has been shown to prevent passive DNA demethylation of ICRs in male and female pronuclei in the pre-implantation embryo. Unlike the oocyte form of DNMT1 (DNMT1-o) (which becomes excluded from the nucleus to facilitate passive DNA demethylation), the somatic DNMT1-s is still present in the nucleus during this period (Kurihara et al, 2008). SETDB1 which interacts with ZFP57 causes H3K9 trimethylation for stabilization of the repressive state (Schultz et al, 2002). Additionally, PGC7/STELLA inhibits Tet-protein mediated active DNA demethylation at imprinted loci in the maternal genome during genome reprograming upon fertilization (Jonkers et al, 2009).

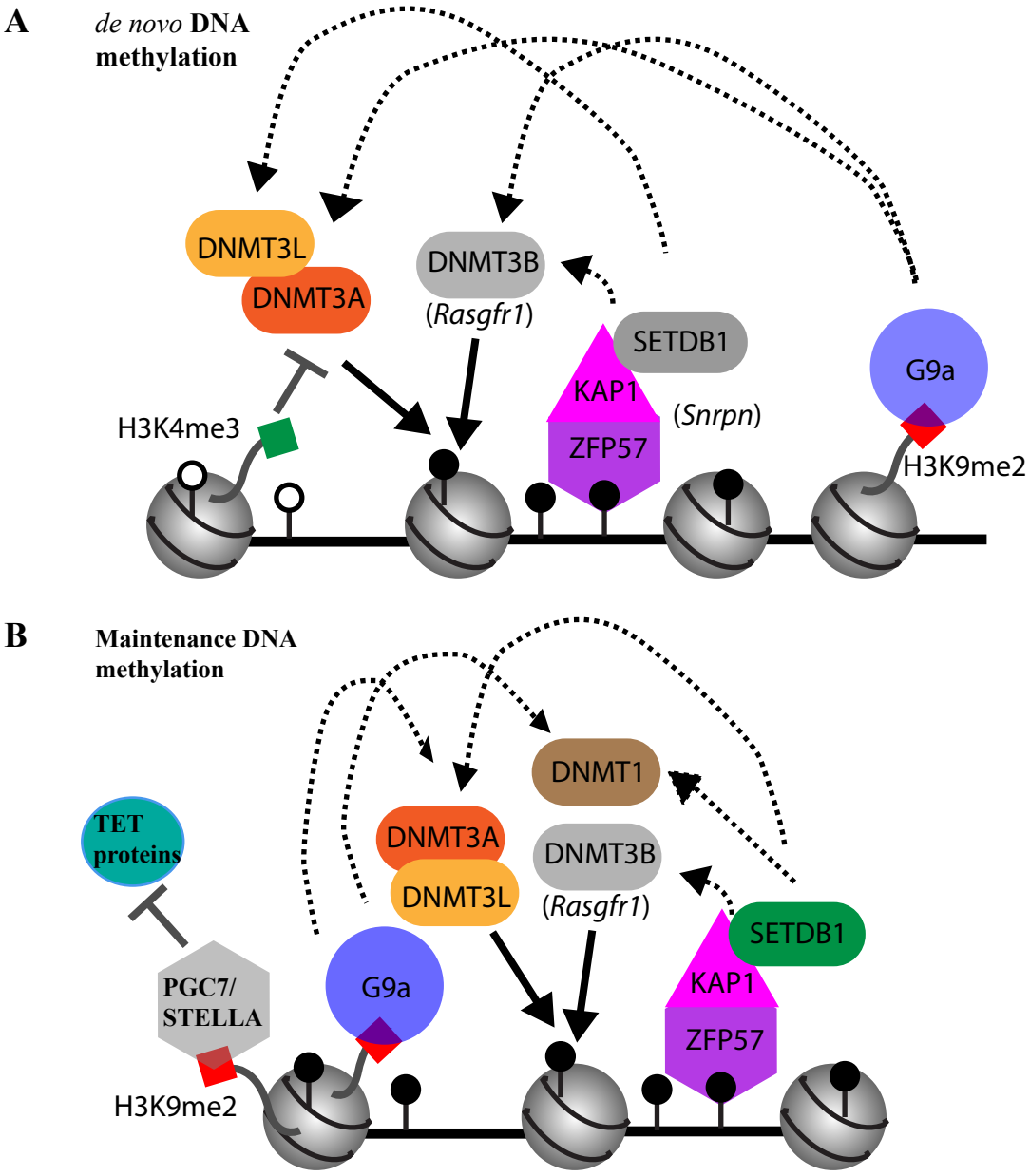


Figure 5.6 Current model of de novo and maintenance DNA methylation at ICRs (adapted from Strogantsev and Ferguson-Smith, 2012)

A: De novo DNA methylation of ICRs. DNMT3A/3L complex is recruited to nucleosomes depleted of H3K4me3. DNMT3B and ZFP57 complex are only required at a few ICRs. Interactions between proteins are highlighted by dashed arrows. DNA methylation is indicated by black 'lollipops' and unmethylated CpGs by white lollipops. The potential interaction of G9a with DNMT3A and DNMT3B in this pathway is indicated. **B:** ZFP57 is the main factor to date, required for maintaining DNA methylation at ICRs. It interacts with corepressor KAP1, Histone lysine methyltransferase SETDB1 as well as DNMT1. The potential interaction of G9a with DNMT1, PGC7/STELLA and DNMT3A/3L or other components of the maintenance methylation pathway is indicated.

The above model highlights the requirement for different factors in *de novo* DNA methylation at ICRs, as well as in preventing demethylation at different stages of development. This may shed light on ways G9a may function. From my study as well as others, evidence points towards a role in maintenance of DNA methylation. The question is, why knockout of *G9a* leads to complete loss of DNA methylation at ICRs, despite the continued presence of factors such as ZFP57 and PGC7/STELLA in *G9a-/-* ES cells (at known ZFP57 target ICRs). This either demonstrates non-redundant functions of these three proteins, or suggests regulation of ZFP57 and PGC7/STELLA by G9a. Absence of G9a may prevent ZFP57 and PGC/STELLA from exerting their function in maintaining DNA methylation.

First of all, a co-operative function of G9a, PGC7/STELLA and ZFP57 could be investigated by determining any potential interactions between these proteins. Mass spectrometry analysis in our lab of interacting partners of G9a only identified GLP, WIZ and DNMT3A (Tuo Zhang, unpublished). An interaction with ZFP57 and PGC7/STELLA was not detected. Interestingly, interactions between G9a and DNMT1 could not be identified in mass spectrometry analysis either. This suggests, that this interaction may be more transient, especially compared to interactions of G9a with GLP and WIZ. Any interaction between G9a and ZFP57 may also be transient and therefore more difficult to detect.

A recent study into the function of PGC7/STELLA in maintaining DNA methylation identified the interaction of PGC7/STELLA with H3K9me2 (Figure 5.6B)(Nakamura et al, 2012). The interaction of PGC7/STELLA with chromatin was reduced in absence of G9a. The same study also identified the need for PGC7/STELLA in maintaining DNA methylation at maternal and paternal imprinted

regions which have H3K9me2 (Nakamura et al, 2012). Finally the study demonstrates that Tet3 binding is inhibited by PGC7/STELLA binding to H3K9me2 (Nakamura et al, 2012). This suggests, that G9a mediated H3K9me2 may be necessary for maintenance of imprinted DNA methylation by PGC7/STELLA. However, experiments in our lab demonstrate that G9a catalytic activity is not required for maintenance of DNA methylation at ICRs. Additionally, rescue of G9a expression in *G9a-/-* ES cells through reintroduction of a transgene does not allow re-establishment of DNA methylation at ICRs (Tuo Zhang, unpublished). However, PGC7/STELLA interaction with chromatin is re-established. These observations suggest, that G9a may facilitate DNA methylation at ICRs in co-operation with PGC7/STELLA, but that G9a also functions in maintaining DNA methylation independent of H3K9 methylation and PGC7/STELLA, perhaps at a different time point.

5.4.2 G9a dependent maintenance of DNA methylation in embryos

A difficulty in studying G9a dependent maintenance of DNA methylation at ICRs in vivo becomes apparent from previous studies. Several studies reported that G9a is required for imprinted gene regulation in the placenta (Redrup et al, 2009; Wagschal et al, 2008). These studies did not detect imprinting defects in the embryo proper. One study to date has identified the requirement of G9a in DNA methylation and correct monoallelic expression of *Snrpn* ICR in ES cells (Xin et al, 2003). The same study showed absence of any defects at this locus in *G9a-/-* embryos. One of the reasons argued are, that DNA methylation in ES cells is often unstable. Female ES cells in particular frequently undergo DNA hypomethylation at imprinted genes during their in vitro culture (Dean et al, 1998; Zvetkova et al, 2005). The WT and

G9a-/- ES cells used in this study are male ES cells which do not normally undergo large scale DNA hypomethylation events during culture, but sometimes display localized DNA methylation defects at imprinted genes. Variations in how prone imprinted regions are to these defects are seen.

My bisulfite sequencing of the KvDMR in WT ES cells indicated that DNA methylation is not quite limited to 50% of sequences, as would normally be expected (Figure 4). Still, only marginal DNA hypermethylation is seen. The *Igf2/H19* region has been previously described to have very unstable DNA methylation in ES cells. In my data 50% DNA methylation exists at Igf2/H19 ICR, but is not evenly distributed over 50% of the sequences.

Despite this, I detect DNA hypomethylation in *G9a-/-* ES cells. Additionally knockdown of *G9a* in ES cells lead to hypomethylation of most ICRs compared to control cells (Tuo Zhang, unpublished). This confirms that DNA hypomethylation seen is due to absence of G9a, rather than due to non-related effects of in vitro ES cell culture. Hypomethylation may be enhanced by the reduced stability of imprinting in ES cells as well as differences in mechanism of silencing in embryos compared to ES cells. This could occur through presence of another factor in embryos, which ensures correct imprinted DNA methylation, independent of G9a.

Chapter 6 - DNA methylation, facilitated by a chromosome structural protein

6.1 Introduction

6.1.1 A Structural Maintenance of Chromosome Protein is required for correct X chromosome inactivation

SmcHD1 was initially identified as a modifier of position effect variegation in screens for modifiers of transgene variegation (Rius & Lyko, 2012). It is a well conserved protein which contains a structural maintenance of chromosomes (SMC) hinge domain normally found in proteins involved in chromosome condensation, cohesion, segregation as well as DNA repair (Figure 6.1A). Interestingly, homozygous mutant female embryos die before E13.5 whilst male mutants remain healthy (Sorge et al, 2013). This was the first indication that SmcHD1 may have an X chromosome linked role.

X chromosome inactivation is a means to achieve dosage compensation of gene expression on the X chromosomes between males (which only have one X chromosome) and females (which have two X chromosomes). A role of an SMC protein in dosage compensation has only been described once before, in *C.elegans.* The process of X chromosome dosage compensation in these organisms is achieved through transcription repression on both X chromosomes in female nematodes by half. SMC proteins DPY-27 and MIX-1 form a complex with non-SMC proteins to function in this gene repression (Gendrel, 2012; Prinapori et al, 2012). SmcHD1 is therefore the first mammalian SMC protein to be implicated in X chromosome inactivation.

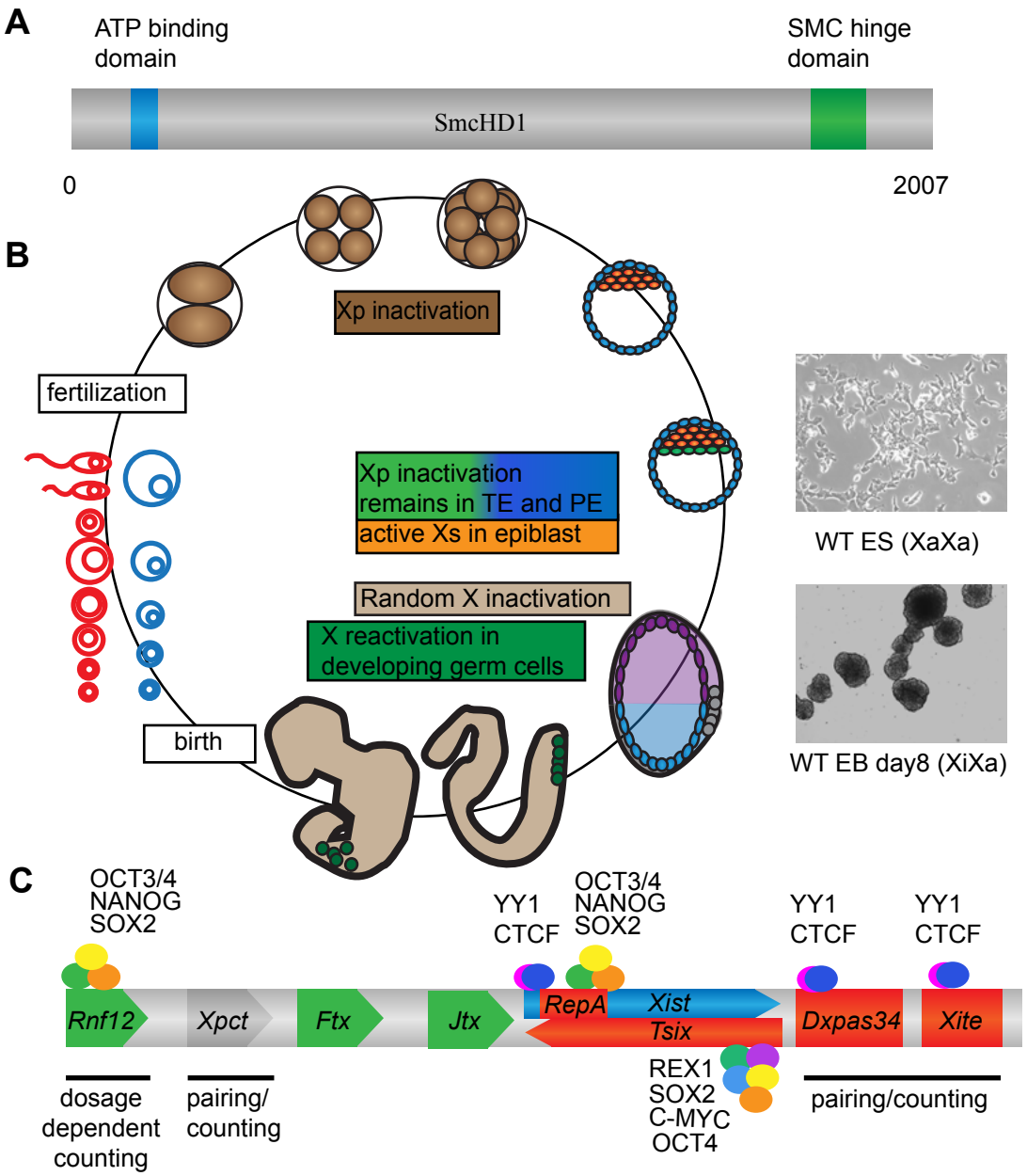


Figure 6.1 X inactivation mechanism and dynamics

A: SmcHD1 protein domain structure (to scale) showing the ATP binding and SMC hinge domains. **B: Left:** X inactivation dynamics during early mammalian development. Different stages of development are indicated as a circular timeline. Changes in X inactivation are indicated in boxes. **Right:** ES cells are derived from the ICM/epiblast and their differentiation into embryoid bodies is used as a model system for random X inactivation. ES cells have two active X chromosomes (XaXa) and inactivate one X chromosome during differentiation (XiXa). **C:** Genomic structure of the X inactivation centre on the X chromosome. Genes are indicated by directional arrows and enhancers by rectangles. Green and red indicate elements which enhance or inhibit *Xist* expression respectively. Binding of transcription factors is indicated above and below. Functions of different domains are indicated below the diagram by black lines.

6.1.2 Mammalian X inactivation

X inactivation is a dynamic process in early mammalian development (Figure 6.1B) (Gendrel & Cohen, 2012). At fertilization the two X chromosomes are active. The first wave of X inactivation occurs between the 4-cell and blastocyst stage at which imprinted X inactivation of the paternal X chromosome occurs. Upon expansion of the blastocyst, the paternal X chromosome is reactivated in embryonic lineage cells (epiblast) but remains inactive in extraembryonic lineage cells (trophectoderm and primitive endoderm) (Gendrel & Cohen, 2012). This is closely followed by the onset of random X inactivation in differentiating embryonic cells. From this stage onwards, random X inactivation is stably transmitted through cell divisions. The only exception occurs during the establishment of the germ line. At this stage random X chromosome inactivation ceases during formation of primordial germ cells (Gendrel & Cohen, 2012). When considering the dynamics of X inactivation during early development it becomes apparent that X inactivation is linked to loss of pluripotency (Gendrel & Cohen, 2012). This has led to the use of embryonic stem cell differentiation as a model for X inactivation (Figure 6.1B right). Undifferentiated ES cells, derived from the ICM/epiblast have two active X chromosomes whereas one X is inactivated during differentiation.

X inactivation has been widely studied in mammals since the discovery by Mary Lyon. It has now been established that X inactivation is orchestrated by various non-coding RNAs as well as protein-coding genes on the X chromosome (Figure 6.1C) (Gendrel et al, 2012; Lee, 2011). These regulators of X inactivation are located in a ~1MB region termed the X inactivation centre (XIC). This region encodes *Xist*, a

non-coding RNA which is seen to coat the inactive X chromosome in females (Brockdorff et al, 1991; Brockdorff et al, 1992; Hong et al, 1999; Jonkers et al, 2008). Regulation of *Xist* is of great importance in ensuring correct X inactivation. *Xist* has to be expressed on the X chromosome to be inactivated in females and repressed on the other female and the male X chromosome. Several elements of the XIC are implicated in positive or negative regulation of *Xist* (Figure 6.1C). *Tsix*, an antisense non-coding RNA overlapping with *Xist* is a negative regulator of *Xist (Lee et al, 1999)*. Consistently in ES cells, where both X chromosomes are active, the *Tsix* promoter is bound by pluripotency transcription factors, promoting its expression (Navarro et al, 2010). *Xist* promoter on the other hand is bound by other pluripotency transcription factors, causing its repression (Donohoe et al, 2009; Navarro et al, 2008). Several other genes encoded by the XCI have now been found to regulate *Xist* and *Tsix* expression. FTX, *Jpx* and RNF12 for example are positive regulators of *Xist* (Barakat et al, 2011; Jonkers et al, 2009; Tian et al, 2010). Additionally enhancer regions Xite and Dxpas34 regulate correct Tsix expression (Navarro et al, 2010; Ogawa & Lee, 2003).

The XIC region is thought to be involved in various stages of X inactivation. The first of these is chromosome pairing and counting (Bacher et al, 2006; Jeon et al, 2012). As the differentiation process of ICM stem cells initiates, pluripotency genes are downregulated, leading to biallelic expression of Xist and loss of *Tsix* expression. It is thought, that at this stage, the two active X chromosomes associate with eachother through several regions (Figure 6.1C) (Xu et al, 2007). This is thought to occur via CTCF interactions with pluripotency factors at the *Tsix* promoter at the onset of differentiation, before pluripotency genes are downregulated (Xu et al,

2007). This initiates subsequent activation of *Tsix* on one allele only, followed rapidly by coating of the *Xist* expressing X chromosome with *Xist* (Jeon et al, 2012). At this stage *Xist* expression on one allele is promoted by RNF12 and *Jpx* in cis and trans. *Xist* coating of the X chromosome is followed by heterochromatinization. Recent studies have identified possible mechanisms of spreading of a repressive state along the inactive X chromosome (Jeon et al, 2012). Binding of PRC2 components to a region called the A-repeats which are part of Xist but can also be separately transcribed is one example (Zhao et al, 2008). This promotes the establishment and spreading of repressive H3K27me3 marks along the inactive X chromosome. Additional silencing mechanisms include CpG methylation as well as incorporation of histone variants (macro H2A) (Chang et al, 2006).

6.1.3 Function of DNA methylation in mammalian X inactivation

DNA methylation plays an important role in X inactivation. Firstly, DNA methylation is seen at the *Xist* promoter on the active X chromosome. DNA methylation at the *Xist* promoter has been shown to be dispensable for the initiation of X inactivation (Cohen et al, 2012; Sado et al, 2004). Prolonged culture of male ES cells in the absence of DNA methylation on the other hand leads to activation of *Xist* on the active X chromosome. This suggests that DNA methylation is required for stable long term repression of *Xist* on the active X chromosome (Sado et al, 2004). Secondly, DNA methylation is put in place at CpG island promoters on the inactive X, to allow long-term silencing of X linked genes. Accordingly, studies have revealed DNA methylation to be a late step in the X inactivation process (Jonkers et al, 2008).

A recent study from the Brockdorff lab we collaborated with has addressed the dynamics of DNA methylation on the X chromosome (Gendrel et al. 2012). In summary, sequencing of purified methylated DNA from day7 and day10 embryoid bodies demonstrated three different classes of CGIs on the X chromosome, which gain DNA methylation at different stages during X chromosome inactivation (DNA methylation gained by day 7, by day 10 and outwith this timeframe) (Figure 6.2). Fast methylating CGIs displayed high CpG content, which is surprising, as elsewhere in the genome sequences with lower CpG content are more readily methylated. The relationship between expression levels in ES cells and DNA methylation kinetics during X inactivation was more conclusive. Fast methylating CGIs tend to fall into the class of repressed Polycomb target genes in ES cells. The reason for the ease of methylation of these CGIs may come from the absence of H3K4 methylation, since previous studies have described inhibition of DNMT3A, 3B and 3L recruitment by H3K4 methylation. The class of slow methylating CGIs on the other hand was enriched in genes highly expressed in ES cells. This may be the reason for their slow methylation, as active chromatin marks inhibit DNA methyltransferases.

6.1.4 Smchd1 is required for correct DNA methylation and gene silencing on the inactive X

Studies so far on SmcHD1 have focused on characterizing the role of SmcHD1 on the X chromosome, due to the limited phenotypes observed in *Smchd1*-/- male embryos. Forced X inactivation by mutating *Xist* on one X chromosome and insertion of a GFP transgene on the other X chromosome was used to show that SmcHD1 is required for silencing of one of the X chromosomes. GFP expression was only visible in SmcHD1 mutant female embryos as well as extraembryonic

tissues, but not in WT tissues (Sorge et al, 2013). Further investigation showed that X inactivation initiates normally, indicating that SmcHD1 may be involved in maintenance of correct X chromosome inactivation in females (Sorge et al, 2013). A further observation supporting SmcHD1 role in X inactivation was the colocalization of SmcHD1 with *Xist* on the inactive X chromosome (Sorge et al, 2013).

A link between DNA methylation on the inactive X and Smchd1 was identified in the same study (Sorge et al, 2013). The study investigated 16 CpG islands subject to X inactivation as well as CpG islands which are methylated independently of X inactivation (Sorge et al, 2013). This showed dependence of CpG island methylation on SmcHD1 in response to X inactivation. X inactivation independent CpG islands such as those present within reproductive homeobox loci remained unaltered in female *Smchd1* mutant cell lines (Sorge et al, 2013). Consistently, many but not all X linked genes are upregulated in female *Smchd1* mutant embryos.

The above observations have therefore defined SmcHD1 as a novel epigenetic regulator of X inactivation. This was an interesting observation as, unlike LSH and G9a, this factor is not a known histone, DNA or chromatin modifying enzyme, demonstrating that not only epigenetic modifiers can play a role in facilitating correct DNA methylation establishment and maintenance. Previous chapters defined a wide role of LSH in facilitating DNA methylation genome wide. The most striking targets of G9a dependent DNA methylation on the other hand are ICR. This suggests, that different proteins may be required for targeting DNA methylation to different regions. Studying SmcHD1 was of interest to me, as it represents a protein which may have a more confined role in facilitating DNA methylation on the X

chromosome, since male embryos mutated in *Smchd1* display limited phenotypes. The combined study of all three factors may provide an overview of how DNA methylation is correctly targeted to a variety of genomic regions.

Although, it has been demonstrated that SmcHD1 is required for correct DNA methylation on the inactive X chromosome, several questions still remain unanswered:

- Is SmcHD1 required for DNA methylation of all X inactivation dependent CGIs?
- Is SmcHD1 dispensable for autosomal CGI methylation?

6.2 Characterization of SmcHD1 dependent DNA methylation genome wide

To answer the above questions I collaborated with the Brockdorff lab to determine genome wide the SmcHD1 dependent CGIs. For this purpose I carried out methylated DNA affinity purification on female WT and *Smchd1*-/- MEFs. *Smchd1* knockout arise from a non-sense mutation in exon 23, leading to non-sense mediated RNA decay. *Smchd1-/-* MEFs were isolated from knockout embryos. *Smchd1-/-* Male WT MEFs were used as controls for genes which become methylated independent of X chromosome inactivation. Male *Smchd-/-* MEFs were used as a control to confirm previous data demonstrating that Smchd1 is not required for DNA methylation on the active X chromosome. I validated the methylated DNA purifications by PCR amplification of different elution fractions with actin and *Igf2r DMR2* primers (Figure 6.2). Actin is an unmethylated control gene and *Igf2r* DMR2

is a differentially methylated region. PCR analysis confirms efficient enrichment of methylated DNA in the high salt elution fractions. These samples could therefore be used for DNA methylation analysis.

All affinity purified methylated DNA samples were processed for Solexa sequencing by the Brockdorff lab (Gendrel et al, 2012). Purified methylated DNA was end-repaired, A-tailed and adaptors were ligated using an Illumina DNA sample prep Kit. The DNA with attached adapters was amplified by 18 cycles of PCR and size selected for 200-500bp fragments. The resultant library was sequenced with an Illumina Genome Analyser II using 51bp reads. 25 million reads were obtained, which were subsequently mapped to the mouse mm9 genome (Gendrel et al, 2012). Numbers of mapped reads for different experiments were: 13903829 (XX WT MEF), 1937780 (7 day differentiated XX ES cell), 2663037 (10 day differentiated XX ES cells), 29104137 (XX Smchd1 null MEF), 8903268 (XY WT MEF) and 7150486 (XY Smchd1 mutant MEF). Data visualization was carried out in Gbrowse (Stein et al., 2002). CGI were annotated, using those detected in cerebellum DNA using CAP-seq (Illingworth et al., 2010). Only those CGIs were annotated which displayed MBD-seq peaks in XX but not XY MEFs. For analysis of DNA methylation between samples, reads were counted logged to base 2, then normalized to the largest (XX Smchd1 null MEF) dataset using Seqmonk (http://www.bioinformatics.bbsrc.ac.uk/projects/seqmonk/) probes designed around Ensembl gene features (Gendrel et al, 2012).

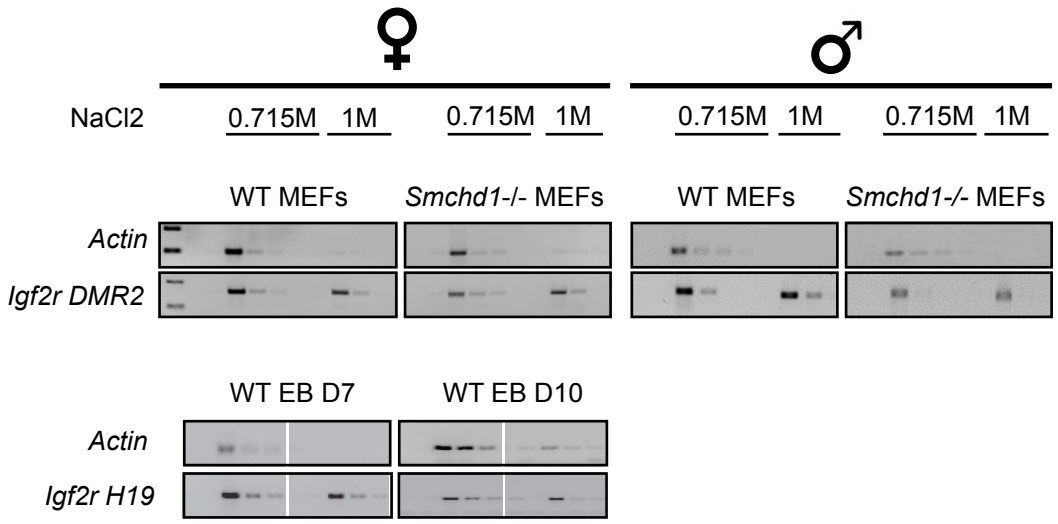


Figure 6.2 X chromosome DNA methylation dynamics

MBD column validation using *Actin* (unmethylated control) or *Igf2r* (differentially methylated) PCRs. PCRs are shown for the low salt and high salt elutions of MBD column runs.

Firstly sequencing demonstrated that SmcHD1 is required for DNA methylation at 1299 CGIs genome wide (>2-fold hypomethylation in absence of SmcHD1) (Appendix). Of these, 177 CGIs are X linked, which confirms previous results that SmcHD1 is required for DNA methylation on the X chromosome. Analysis of this data by the Brockdorff lab identified a significant overlap between SmcHD1 dependent CGIs and the class of CGIs methylated slowly during X chromosome inactivation, as described in the introduction (Gendrel et al, 2012). This suggests that SmcHD1 may be required for facilitating DNA methylation at CGIs, which have less favourable characteristics for DNA methylation (high expression levels, H3K4me3). Additionally SmcHD1 is recruited to the X chromosome when DNMT3B levels are low. This suggests requirement for SmcHD1 in facilitating DNA methylation in presence of low levels of DNMT3B (Gendrel et al, 2012). This is supported by the observation that fast methylating CGIs do not require Smchd1 for their methylation, but are not methylated to the same levels in mutant cells as WT cell (Gendrel et al, 2012). The above observations are summarised in Figure 6.3.

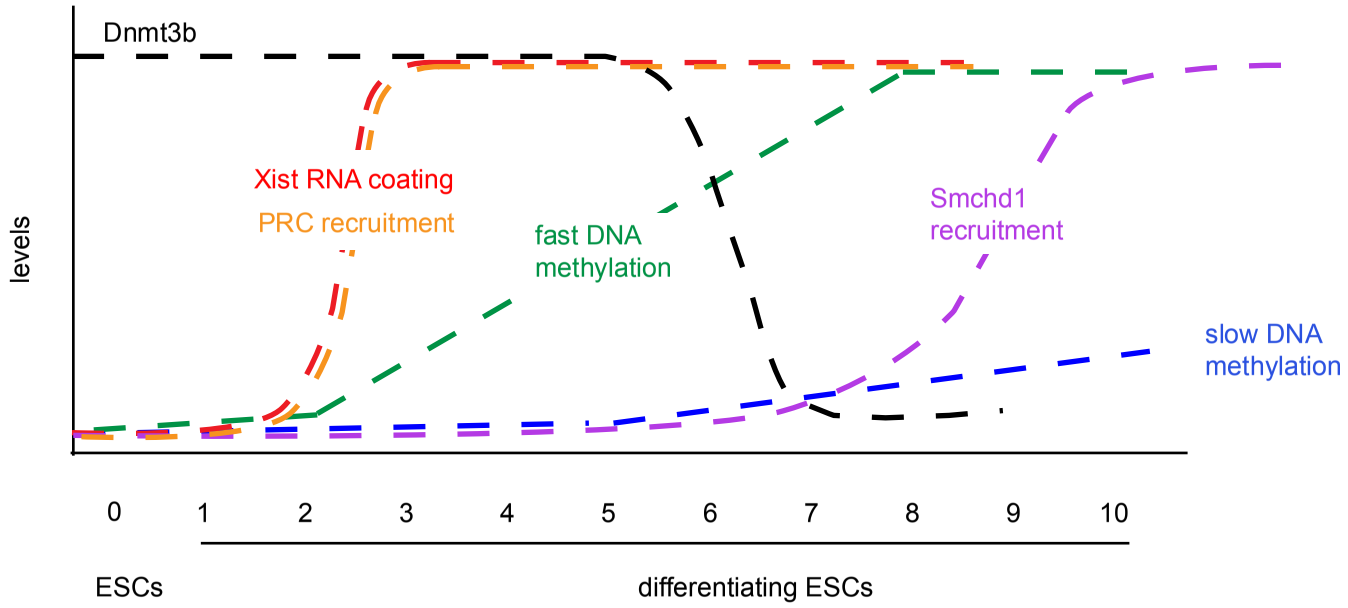


Figure 6.3 X chromosome DNA methylation dynamics and SmcHD1

Timings at which DNA methylation is acquired at X-linked CGIs during ES cell differentiation is indicated (fast and slow DNA methylation). This is displayed in relation to Xist RNA coating of the inactive X chromosome, PRC recruitment. SmcHD1 recuritment is also indicated, coinciding with the onset of 'slow' DNA methylation as well as reduced levels of DNMT3B (Gendrel et al. 2012).

During my PhD studies I was interested in the requirement of different non-DNA methyltransferase proteins in DNA methylation genome wide, not just on the X chromosome. In the case of SmcHD1, studies up to now have focused on the involvement of SmcHD1 in DNA methylation on the X chromosome. The minimal effect of *Smchd1* mutation in males suggests function primarily in female X chromosome inactivation. It is unclear whether SmcHD1 is dispensable for DNA methylation elsewhere in the genome. Indications of a role wider than its function in X inactivation potential come from studies of SmcHD1 homologs in *Arabidopsis thaliana*, which have been implicated in RNA dependent DNA methylation, as well as in humans, in DNA repair (Lee & Lu, 1999; Lorkovic et al, 2012). Therefore I was interested to look at the autosomal sequencing data for methylated DNA I purified from WT and Smchd1-/- MEFs.

1122 autosomal CGIs were identified to be hypomethylated in *Smchd1* mutant MEFs (Appendix 3). This is a greater number than expected from the lack of phenotype in male mutant embryos, suggesting, that SmcHD1 may have a more widespread role than previously outlined. Figure 6.4 shows some examples of autosomal CGIs, which are hypomethylated in Smchd1-/- MEFs compared to WT. Interestingly, Peg12 in the PWS imprinted gene cluster is more than 2-fold hypomethylated in absence of SmcHD1. Visualization of methylation peaks in the entire PWS gene cluster demonstrates loss of DNA methylation at Peg12 as well as at the Ndn promoter CGIs. When examining the imprinting control region (Snrpn ICR) for this imprinted gene cluster, no change in DNA methylation can be observed. Strikingly the Snrpn ICR does not appear as a methylated peak in WT MEFs.

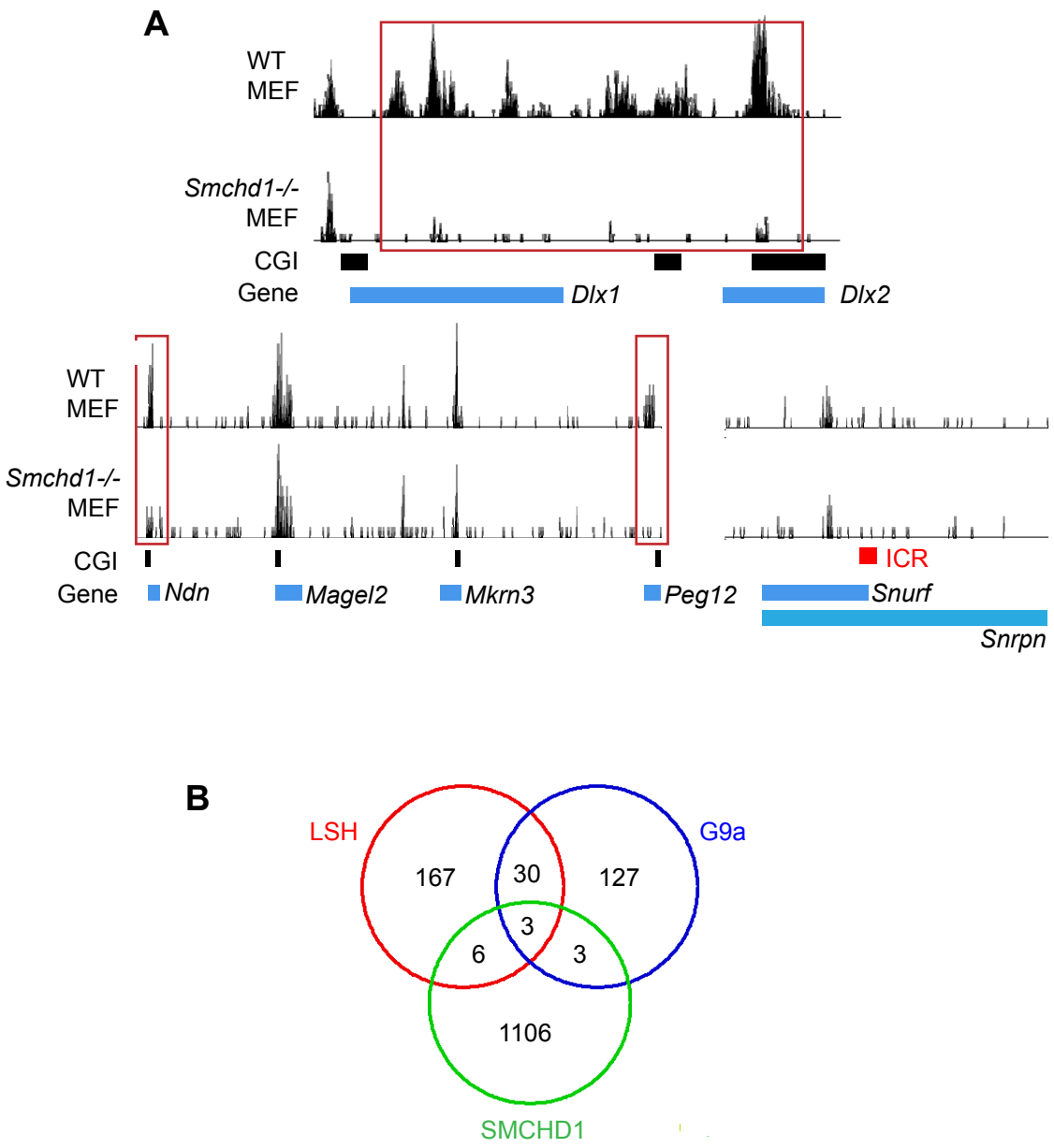


Figure 6.4 SmcHD1 requirement for autosomal DNA methylation and gene expression

A: Sequencing results of affinity purified methylated DNA from WT and *Smchd1* mutant MEFs. Peak hights indicate amount of methylation. CpG islands (CGIs) and genes analysed are indicated below. **B:** Venn Diagram outlining the overlap between LSH, G9a and SMCHD1 target genes (Gendrel et al. 2012)

This suggests that the sequencing depth may not have been sufficient to identify all differences in DNA methylation between cell lines, especially at imprinted regions, which have one methylated and one unmethylated allele. Future work may therefore require resequecing with an increased number of reads per sample. Additionally validation of the sequencing results by bisulfite sequencing would be necessary to confirm the differences.

Despite the drawbacks of the sequencing data, I decided to compare the identified SmcHD1 target genes with those of G9a and LSH determined in previous chapters. The Venn diagram in Figure 6.3 B displays minimal overlap between all autosomal SmcHD1 target genes and those of LSH and G9a (Appendix 4). This suggests, that SmcHD1 may be involved in regulating DNA methylation at mostly distinct genomic regions to those regulated by LSH and G9a.

Chapter 7 Discussion

7.1.1 Achieving diverse DNA methylation patterns during mammalian development

The initial question posed, when DNA methylation was identified to be widespread in vertebrates and plants, was what its functions are. As mentioned in the introduction, DNA methylation is thought to have originated from a genomic immune system to silence transposable elements, which has evolved to function in gene regulation, imprinting, X chromosome inactivation and regulation of genome stability through silencing of transposable elements in mammals (Lee et al, 1999). Throughout the years it has also become apparent, that DNA methylation forms distinct landscapes in different cell types as well as at different stages of mammalian development. As discussed in the introduction, correct generation of DNA methylation landscapes is important to prevent diseases such as cancer. In somatic cells, DNA methylation is relatively stable, but during early development many dynamic changes occur to DNA methylation landscapes. Somatic methylation patterns have to be erased in germ cells, to be substituted with gender-specific DNA methylation at imprinted regions as well as novel differentially methylated regions. Differential DNA methylation at imprinted regions is thereafter maintained throughout early development and adult life. Erasure of DNA methylation and re-establishment is also important in the divergence between extraembryonic and embryonic tissues as well as during differentiation of embryonic stem cells.

The above outlines the importance of correct targeting and maintenance of DNA methylation. It is known, that DNA methyltransferases establish and maintain

DNA methylation, but how can DNA methyltransferases be directed to only the regions requiring methylation at different stages of development? This is the focus of many recent studies. One possibility, as to how DNA methylation landscapes are formed, is that the default state of the genome may be a methylated state. Factors may exist to prevent DNA methylation at some regions but not at others. One example is the inhibitory effect of histone modifications, for example H3K4me3, on DNA methylation (Meissner et al, 2008; Mohn et al, 2008; Okitsu & Hsieh, 2007; Weber et al, 2007). More recent studies are in support of the existence of an opposite model whereby DNA methylation is targeted to some regions, but not others. Sequence elements have been identified, which preferentially recruit DNA methylation (Lienert et al, 2011; Meissner, 2011). Additionally, different isoforms of DNMTs seem to target common as well as different genomic regions (Navarro et al, 2010). These mechanisms could explain some of the complexity of ensuring correct DNA methylation landscapes.

Additional, several non-DNMT proteins have been shown to facilitate DNA methylation. Absence of these can lead to severe loss of DNA methylation, sometimes similar to the hypomethylation seen in *Dnmt* knockouts themselves (Donohoe et al, 2009; Jonkers et al, 2009; Navarro et al, 2008; Tian et al, 2010; Xin et al, 2003).

7.1.2 Accessory proteins target DNA methylation to distinct genomic regions

In this study I sought to characterize three such ‘accessory’ proteins. LSH (putative chromatin remodelling ATPase), G9a (histone methyltransferase) and

SmcHD1 (protein related to structural maintenance of chromosomes proteins) have all been implicated in DNA methylation, due to DNA methylation defects seen in knockout cells and embryos.

Both LSH and G9a have previously been reported to be required for DNA methylation at repetitive sequences and few single genes. In my study, I successfully expanded these findings, characterizing promoters of protein-coding genes genome-wide which require either LSH or G9a for DNA methylation. This enabled me to define the role of LSH in *de novo* methylation of genes that normally undergo methylation in the pre- and post- implantation embryo. Additionally, it was striking to find G9a to be necessary for maintenance of DNA methylation at imprinting control regions. DNA methylation at ICRs is normally established in the germ line and maintained throughout development.

SmcHD1 is a novel factor required for correct DNA methylation on the inactive X chromosome. In this study we found SmcHD1 to be essential for DNA methylation of a subset of CGIs normally methylated slowly during X chromosome inactivation. Analysis of targets of SmcHD1 identified extensive hypomethylation of autosomal CGIs in absence of SmcHD1. This finding suggests a wider role for SmcHD1 apart from on the X chromosome. Most striking is the need for SmcHD1 in DNA methylation at the autosomal PWS imprinted locus. Differentially methylated regions in the Peg12 as well as Ndn promoters were found to be hypomethylated in absence of SmcHD1. Other imprinted regions did not display hypomethylation in absence of SmcHD1 in this experiment. Due to low sequencing depth, these are only preliminary findings, which would have to be confirmed through bisulfite sequencing as well as through additional sequencing at higher depth.

In summary, the methylated DNA affinity purification method enabled me to characterize the requiremenet for different proteins in DNA methylation genome wide. This method can be adapted to analyse the requirement of any protein of interest in DNA methylation at different stages of mammalian development. My findings suggest LSH, SmcHD1 and G9a facilitate DNA methylation at partially overlapping time points during development, but at large groups of unique target genes (Figure 7.1). This supports the hypothesis, that different accessory proteins may facilitate DNA methylation at different genomic regions.

There are several considerations which have to be made in drawing the above conclusions about the different 'accessory proteins'. First of all, different cell types were used in this study to investigate LSH, G9a and SmcHD1 roles in DNA methylation. *Lsh-/-* and *SmcHD1-/-* MEFs were used, whereas the *G9a-/-* was in ES cells. Especially in the case of G9a, some later developmental target regions may have been missed, due to the use of ES cells, rather than MEFs. Additionally, there are also a few drawbacks to the methylated DNA affinity purification method used. This method has been previously used to identify highly methylated CpG rich regions genome wide. In my study we have demonstrated that we could enrich for methylated low, intermediate and high CpG density promoters. Despite this there may still be a skew towards detection of more CpG rich promoters. Additionally, differences in methylation at differentially methylation regions as well as at regions with smaller changes (e.g. ICRs) are harder to detect using the MBD-microarray method. This was demonstrated during the analysis of MBD-microarray data on WT and *G9a-/-* ES cells.

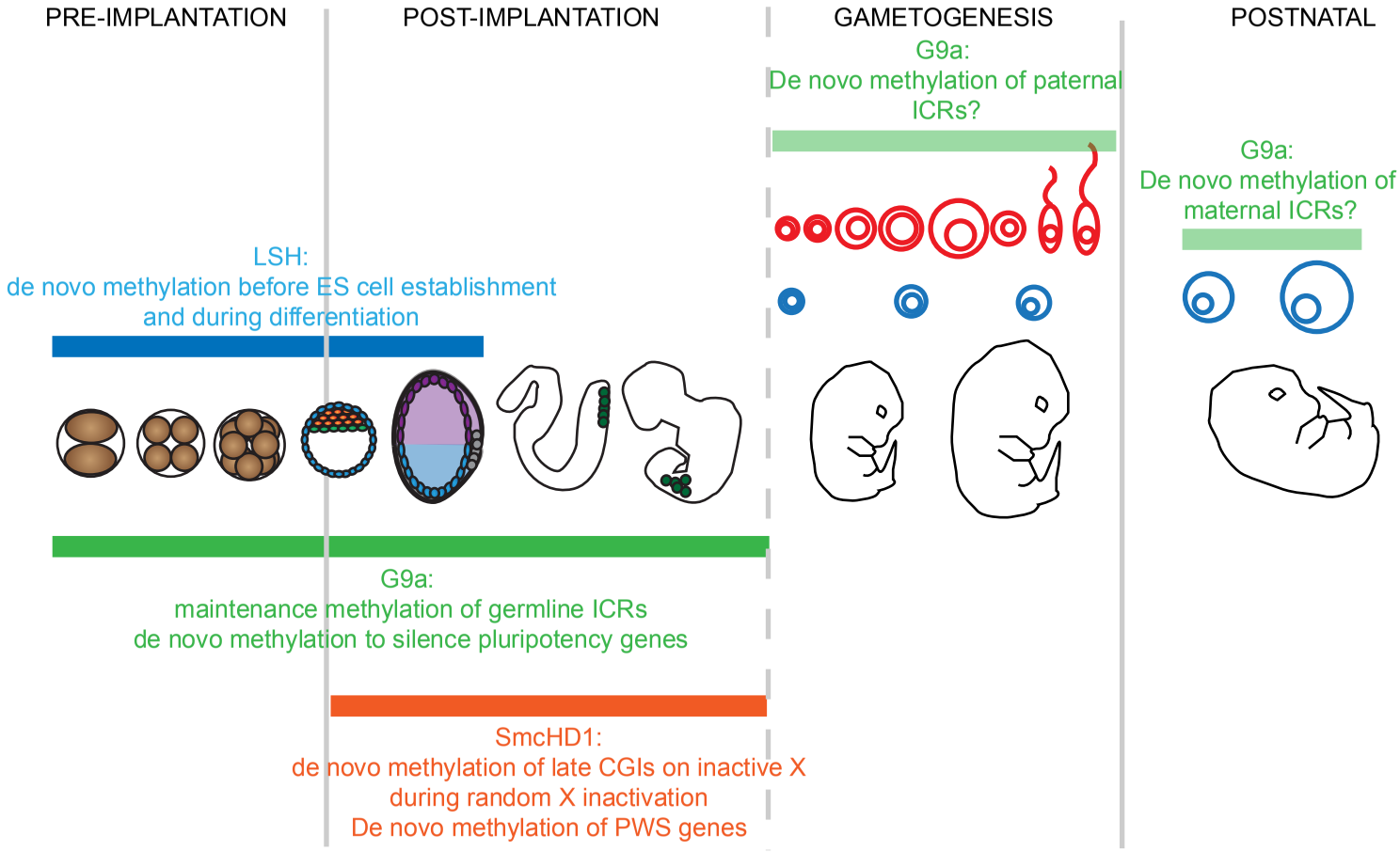


Figure 7.1 Function of LSH, G9a and SmcHD1 at diferent stages of development

Displayed are different time points of early development from pre-implantation to the postnatal stage. Functions of LSH, G9a and SmcHD1 at different stages and at different genomic regions are indicated by coloured bars.

Hypomethylation of some imprinted regions could be detected, but the extent of DNA hypomethylation at ICRs only became apparent during qPCR analysis of purified methylated DNA. This suggests, that microarray analysis of purified methylated DNA may not always be sensitive enough to detect smaller changes in DNA methylation.

7.1.3 Suitability of different accessory proteins for targeting DNA methylation to specific genomic regions

The first question arising from my study is why different accessory proteins may be suited to facilitate DNA methylation at specific regions. Comparing target regions between the three proteins I studied, identified some overlap, but the majority were unique regions. Can other functions of these proteins aid our understanding of how they facilitate DNA methylation?

LSH being a putative chromatin remodeller gives rise to the hypothesis, that it may facilitate targeting of DNMTs and other factors to difficult to access regions, through chromatin remodelling. Similarly, previous studies have found other chromatin remodelling ATPases to facilitate DNA methylation. These include the LSH homolog DDM1 as well as another SNF2 family ATPase, ATRX (Navarro et al, 2008; Tian et al, 2010). Attempts to demonstrate LSH chromatin remodelling activity *in vitro* have been unsuccessful, with only a weak ATPase activity detectable. The question arising was, whether chromatin remodelling is required for LSH to exert its function in promoting DNA methylation. Instead, LSH may cause accumulation of DNA methyltransferases and histone deacetylases at target regions

without the need for remodelling. Therefore, the question I posed was, whether ATPase activity is required for LSH dependent DNA methylation.

I carried out rescue experiments with WT and catalytically inactive LSH in *Lsh*-/- MEFs. The catalytically inactive LSH was unable to re-establish DNA methylation at repetitive elements as well as single genes, compared to wild-type LSH. This suggests that LSH may use its ATPase activity to remodel nucleosomes, making chromatin more accessible to DNA methyltransferases and other chromatin modifiers. In vivo, since DNA methylation is often seen as the secondary effect of gene silencing, which promotes a long term silent state, DNA methyltransferases may have to access an already compacted chromatin environment. This could explain the need for a chromatin remodeller. However, reintroduction of catalytically inactive LSH prevented establishment of DNA methylation and gene silencing at the expressed *Rhox* loci in Lsh-/- MEFs. This suggests, that even in relatively 'open' chromatin, remodelling activity of LSH is required. In support, a recent study demonstrated that nucleosomes inhibit DNA methylation (Felle et al, 2011). Therefore, even in active regions, nucleosomes may have to be remodelled. Additionally, this poses the question whether regions, whose DNA methylation is dependent on LSH are different from others in terms of nucleosome density or stability, which may necessitate the presence of a chromatin remodeller.

SmcHD1 on the other hand is similar to many structural maintenance of chromosomes proteins. SMC proteins have been found to link strands of nucleotides together (Losada & Hirano, 2005). Well known roles include chromosome cohesion, condensation, segregation and DNA repair (Losada & Hirano, 2005). A couple of recent studies identified a plant SMC homolog to be involved in RNA-dependent

DNA methylation (Barakat et al, 2011; Ogawa & Lee, 2003). RNA-dependent DNA methylation usually utilises short siRNAs, made by the enzyme, DICER. These siRNAs can in turn direct DNA methylation to particular sequences (Xu et al, 2007). Based on their known function, SMC proteins have therefore been proposed to 'clamp' short RNA molecules to DNA regions to facilitate DNA methylation (Xu et al, 2007).

The function of SmcHD1 in X inactivation could potentially utilise a similar mechanism, the difference being, that X inactivation is achieved through a long non-coding RNA, rather than small RNAs. SmcHD1 may interact with Xist and facilitate the spreading of DNA methylation. Consistently previous studies have indicated that DNA methylation on the inactive X chromosome may commence close to the Xist gene. In our study, the class of CGIs methylated at intermediate kinetics were located closer to the Xist locus than other classes, which partially supports this notion (Gendrel et al, 2012). Requirement of SmcHD1 in DNA methylation at the PWS locus is more unclear, as this locus does not utilise RNAs in the silencing mechanism. Additionally, other imprinted regions did not seem to require SmcHD1, despite being regulated by non-coding RNAs.

Why G9a is required for DNA methylation at imprinting control regions is yet unclear. My study, combined with others highlights that G9a is required for DNA methylation in ES cells. On the other hand, in embryos, DNA methylation was unaffected at G9a target genes and especially at several imprinted regions (Chang et al, 2006; Xin et al, 2003; Zhao et al, 2008). This suggests a mechanism of maintaining DNA methylation in embryos, independent of G9a.

7.1.4 Regulation of DNA methylation through regulation of accessory proteins

In summary the above discussion, combined with previous studies highlights the requirement of specialised proteins to facilitate DNA methylation at different genomic loci. This provides a mechanism of controlling tissue- and developmental stage-specific DNA methylation through regulation of accessory proteins. Consistently, LSH for example is highly expressed in stem cells but is downregulated during differentiation. This highlights its function in DNA methylation during early development. A recent study into G9a requirement for pluripotency gene silencing demonstrated stimulation of G9a expression in presence of Protein Kinase A (Yamamizu et al, 2012). This was required for correct regulation of timing of differentiation of stem cells. Additionally, G9a is upregulated in oocytes, at the time of establishment of DNA methylation marks at Imprinting Control Regions (Kageyama et al, 2007). This suggests that G9a could also be involved in the establishment of imprinted methylation patterns in the germ line. Expression patterns of DNA methylation accessory proteins can therefore also be used to infer possible developmental and cell type specific function. SmcHD1 for example is highly expressed in cells of the immune system, which may point to additional functions of SmcHD1 apart form in CpG island methylation establishment during X inactivation.

7.1.5 Therapeutic uses

Understanding the mechanisms of how proteins function in the mammalian cell is crucial for studies into diseases. As discussed in detail in the introduction, DNA methylation is increasingly found to be defective in diseases, especially cancers. Absence of all three proteins I have studied (LSH, G9a and SmcHD1)

results in extensive DNA hypomethylation. DNA hypomethylation can be linked to genome instability and loss of imprinting, which are frequently characteristic of cancers. LSH for example has been demonstrated to play a role in ensuring stability of repetitive elements, regulating cell proliferation and cell divisions (Fan et al, 2003). Additionally, LSH expression is characteristic of highly proliferating tissues such as thymus, testis and bone marrow (Raabe et al, 2001). These are all important functions, defects of which are frequent hallmarks of cancers (Hanahan & Weinberg, 2011). Consistently, LSH deficiency has been demonstrated to play a role in development of leukaemia. One study identified DNA hypomethylation in absence of LSH to cause activation of oncogenes, which may promote development of erythroleukemia (Fan et al, 2008). Additionally another study identified frequent occurrence of a deletion in LSH protein in various leukaemias (Lee et al, 2000). The above are observations made in the case of murine LSH. The human homolog of LSH, HELLS, has also been widely characterized regarding its function in the development of cancers. Several studies have highlighted overexpression of HELLS in many cancers (Ryu et al, 2007).

Histone methyltransferase G9a has similarly been found to be implicated in cancers. G9a is overexpressed in many human cancers. Consistently, knockdown of G9a inhibits growth of cancer cells. G9a has also been found to methylate p53, resulting in inactivation of p53. This may be of importance as over 50% of cancers display inactive p53.

The importance of continued development of more specific DNA demethylation therapeutics has been highlighted in various studies and reviews (Esteller, 2005; Rius & Lyko, 2012). Inhibiting specific DNA methylation accessory

proteins rather than using DNA demethylation drugs may provide a treatment with less widespread effects on gene regulation.

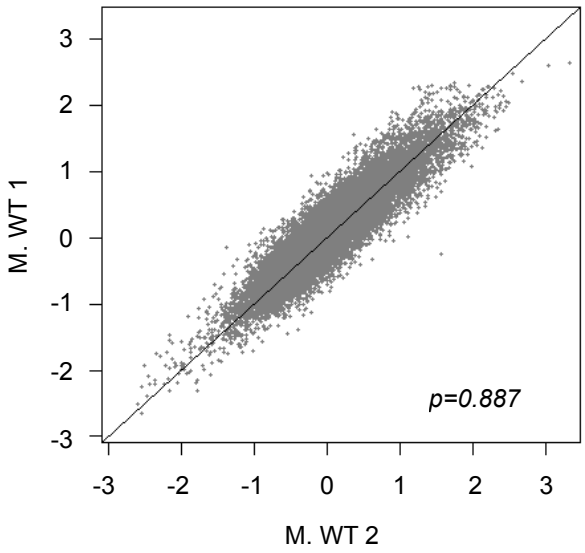
Apart from their function in DNA methylation, LSH, G9a and SmcHD1 are also implicated in other functions. In addition to its function in DNA methylation several studies have for example implicated LSH and its plant homolog DDM1 in DNA damage repair pathways (Geiman et al, 1998; Shaked et al, 2006; Yao et al, 2012). A recent study in our lab, run in parallel to my study focused on characterizing the role of LSH in DNA damage repair (Burrage et al., accepted for publication). This study demonstrated increased sensitivity of *Lsh-/-* mouse and human fibroblasts to ionizing radiation. More detailed analysis identified defects in the events downstream of a double strand DNA damage events. Use of LSH rescue cell lines (same as the ones used in my study) in the study demonstrated the requirement of catalytic activity of LSH in its role in DNA damage repair. Additionally, the DNA damage repair function of LSH was suggested to be independent of its DNA methylation function.

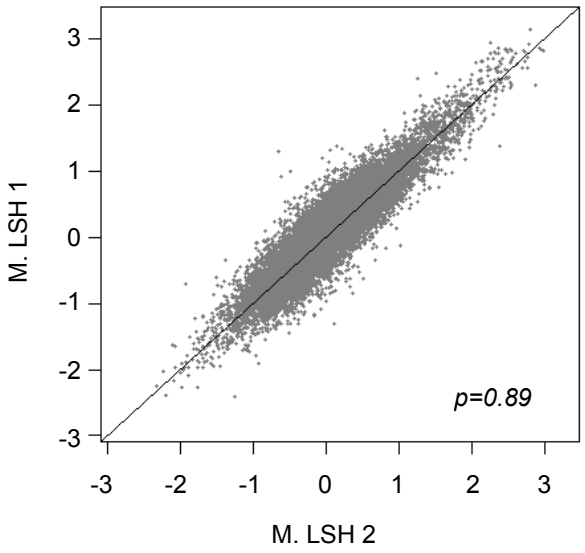
One of the most common hallmarks of cancer is genomic instability, which occur through for example random mutations and defects in repetitive element silencing which can lead to chromosomal rearrangements (Hanahan & Weinberg, 2011). The discovery of proteins such as LSH which are implicated in various distinct functions related to genomic instability makes it a sought after target for cancer therapy. My study and the study of Burrage et al. (accepted) highlight the independence of DNA methylation and DNA damage repair functions of LSH, which may make LSH even more amenable to new cancer therapeutics, as different

inhibitors may be able to target the different functions of LSH, so as not to disrupt the other function.

Appendix 1

Correlation of LSH MBD column microarray replicates

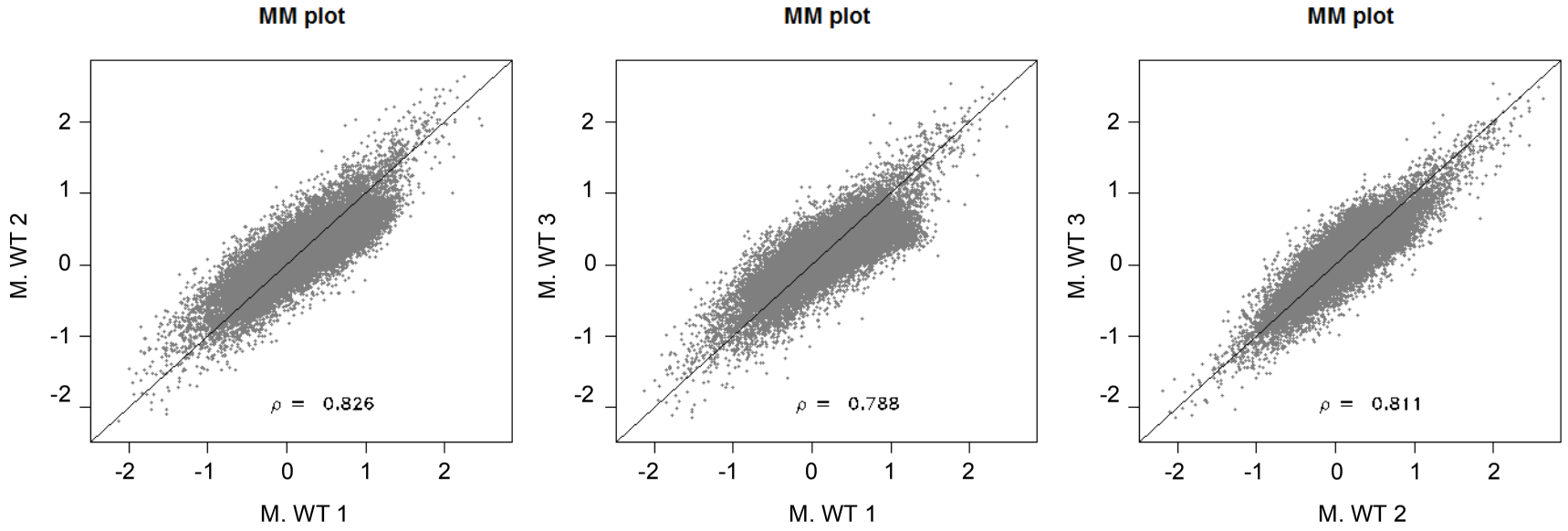




Appendix 1: Correlation between LSH MBD microarray replicate samples
Axes represent M values of two different biological replicates. Indicated on the bottom right are the Spearman's correlation coefficients.

Appendix 2

Correlation of G9a MBD column microarray replicates



Appendix 2: Correlation between G9a MBD microarray replicate samples
Axes represent M values for two different technical replicate experiments. Indicated on the bottom right are the Spearman correlation coefficients

Appendix 3

LSH, G9a and SmcHD1 target promoters

Loss of DNA methylation in *Lsh-/-* MEFs (2-fold)

Chr	Start	Stop	Gene	av_LSH KO	av_WT	DM(WT-KO)
1	4485994	4486994	Sox17	-1.0	0.6	1.6
1	9283578	9284578	Sntg1	0.1	1.1	1.0
1	9283952	9284952	Sntg1	0.1	1.3	1.2
1	9284223	9285223	Sntg1	0.2	1.3	1.1
1	19193627	19194627	Tcfap2b	-0.2	1.1	1.2
1	34504966	34505966	Tesp2	-0.1	1.3	1.4
1	36217462	36218462	Lincr	0.3	1.5	1.1
1	53005707	53006707	Gdf8	-0.5	0.9	1.4
1	58154442	58155442	Aox4	0.2	1.5	1.3
1	64997185	64998185	Crygd	-0.3	1.0	1.3
1	72755889	72756889	Igfbp2	-0.5	1.0	1.5
1	74868299	74869299	B230363K08Rik	-0.2	0.8	1.0
1	79318514	79319514	Scg2	0.0	1.0	1.0
1	79318614	79319614	Scg2	0.0	1.0	1.0
1	82916955	82917955	Slc19a3	0.0	1.3	1.2
1	87852616	87853616	Spata3	-0.2	0.9	1.1
1	97497471	97498471	N/A St8sia4	0.2	1.3	1.0
1	98702144	98703144	9530060I07	0.3	1.5	1.2
1	111810783	111811783	Cdh7	-0.2	1.3	1.5
1	123355409	123356409	Htr5b	0.5	1.7	1.2
1	123355511	123356511	Htr5b	0.3	1.4	1.1
1	130419255	130420255	Cxcr4	0.4	1.9	1.4
1	156488074	156489074	Cacna1e	0.2	1.7	1.5
1	169525732	169526732	Lmx1a	0.2	1.9	1.7
1	177815289	177816289	B020018G12Rik	-0.3	0.8	1.1
1	189963321	189964321	Ush2a	-0.3	1.4	1.7
2	55252041	55253041	Kcnj3	0.3	1.5	1.2
2	57974948	57975948	Pscdbp	-0.3	1.3	1.6
2	62212421	62213421	Dpp4	0.2	1.4	1.1
2	62212570	62213570	Dpp4	0.0	1.1	1.1
2	62374253	62375253	Fap	-0.3	1.1	1.4
2	62374360	62375360	Fap	-0.3	0.9	1.2
2	67918066	67919066	B3galt1	-0.7	0.3	1.0
2	74475413	74476413	Hoxd12	0.9	1.9	1.0
2	79929278	79930278	Pde1a	-0.4	0.7	1.1
2	80092813	80093813	Prdx6-rs1	0.0	1.4	1.4
2	103224415	103225415	Elf5	0.3	1.5	1.2
2	105024682	105025682	0610012H03Rik	0.4	2.2	1.8
2	119936023	119937023	Pla2g4e	0.1	1.4	1.4
2	119936276	119937276	Pla2g4e	0.0	1.3	1.3
2	124844670	124845670	Slc12a1	-0.1	1.3	1.4
2	130850961	130851961	1700037H04Rik	0.2	1.3	1.1
2	135748867	135749867	6330527O06Rik	-0.2	0.8	1.0
2	135749170	135750170	6330527O06Rik	-0.2	1.2	1.4
2	152067507	152068507	Gm123	0.0	1.3	1.2
2	152271253	152272253	Defb19	-0.3	1.2	1.5
2	163912549	163913549	Wfdc15	-0.6	1.4	2.1
2	170292722	170293722	4930470P17Rik	-0.5	0.8	1.3
2	171939196	171940196	Mc3r	-0.2	0.8	1.0
2	172795511	172796511	Pck1	-0.1	0.9	1.1
2	181142078	181143078	Srms	0.2	1.2	1.0
3	10179254	10180254	Fabp9	0.2	1.2	1.0
3	10283215	10284215	1700008G05Rik	-0.2	1.6	1.8
3	26828203	26829203	Spata16	0.0	1.4	1.4
3	45474123	45475123	Pcdh10	0.0	1.7	1.7
3	45474305	45475305	Pcdh10	0.2	1.6	1.4
3	48704780	48705780	1700018B24Rik	0.1	1.6	1.5
3	58972641	58973641	Ush3a	0.0	1.2	1.2
3	82633513	82634513	Npy2r	0.0	1.1	1.0
3	82989401	82990401	Lrat	0.4	2.0	1.7
3	95219288	95220288	Lysmd1	-0.1	1.3	1.5
3	97010835	97011835	Gja8	-0.1	1.1	1.1
3	98307057	98308057	Reg4	0.2	1.3	1.1
3	101416471	101417471	Cd2	0.0	1.1	1.1
3	102025431	102026431	N/A	-0.2	1.1	1.3
3	105586371	105587371	Kcnd3	0.1	1.1	1.0
3	109701562	109702562	Vav3	-0.5	0.6	1.1
3	110270994	110271994	Ntng1	0.5	1.5	1.0
3	129191464	129192464	Pitx2	0.1	1.6	1.5
3	129746409	129747409	Egf	-0.1	1.1	1.1
3	135284923	135285923	4933425K02Rik	-0.5	1.2	1.7
3	141919138	141920138	Bmpr1b	-0.3	1.5	1.7
4	3946229	3947229	Rdhe2	-0.7	0.5	1.2
4	33323944	33324944	Rragd	0.3	1.7	1.4

4	42894795	42895795	BC049635	-0.7	0.5	1.2
4	43727166	43728166	Olfr71	0.3	1.5	1.3
4	46595743	46596743	LOC634450	-0.3	0.7	1.0
4	62657907	62658907	Kif12	-0.6	0.8	1.4
4	73484172	73485172	Frmd3	0.2	1.3	1.1
4	118066879	118067879	1110020C03Rik	0.1	1.6	1.6
4	128120293	128121293	Tlr12	0.0	1.0	1.0
4	128620562	128621562	Hpca	-0.1	1.7	1.8
4	134898393	134899393	1700029M20Rik	-0.3	2.0	2.3
4	134958099	134959099	Il28ra	0.3	1.4	1.0
4	141194002	141195002	Fhad1	-0.2	1.2	1.4
4	151540839	151541839	Gm833	-0.1	1.2	1.3
4	153511033	153512033	Actrt2	-0.2	1.2	1.4
4	153742327	153743327	2810405K02Rik	0.2	1.4	1.1
5	28172233	28173233	Htr5a	-0.1	1.0	1.1
5	34842202	34843202	Sh3bp2	0.0	2.0	2.0
5	44388189	44389189	LOC624120	0.1	1.6	1.5
5	66282920	66283920	9130230L23Rik	0.1	1.4	1.3
5	71978650	71979650	Gabrb1	-0.1	1.7	1.7
5	78189398	78190398	Hod	0.2	1.7	1.5
5	78285709	78286709	Spink2	-0.1	1.3	1.4
5	82095653	82096653	Lphn3	0.1	1.4	1.3
5	110855286	110856286	Galnt9	-0.6	0.4	1.0
5	114426016	114427016	Acacb	0.0	1.3	1.3
5	116549462	116550462	Ccdc60	-0.2	1.2	1.3
5	116852306	116853306	1500001A10Rik	0.6	1.8	1.1
5	118127648	118128648	Nos1	-0.1	2.2	2.3
5	121269635	121270635	Rph3a	-0.1	1.0	1.1
5	129050453	129051453	Piwil1	0.3	1.7	1.4
5	132826607	132827607	N/A	1.0	2.0	1.0
5	146894253	146895253	Gpr12	0.0	1.6	1.6
5	150286553	150287553	Lgr8	-0.3	1.8	2.1
6	3917854	3918854	Tfpi2	-0.4	1.6	2.0
6	7504570	7505570	Tac1	0.4	1.8	1.4
6	8899314	8900314	Nxph1	0.2	1.3	1.1
6	13174465	13175465	Gm467	0.0	1.3	1.3
6	18058555	18059555	Asz1	-0.2	1.7	1.9
6	21165609	21166609	Kcnd2	0.4	1.6	1.2
6	25640239	25641239	Gpr37	-0.3	1.1	1.5
6	29009730	29010730	Lep	-0.2	0.9	1.1
6	36317503	36318503	9330158H04Rik	-0.5	1.9	2.4
6	40372056	40373056	BC052883	-0.5	0.6	1.1
6	52242575	52243575	Evx1	0.3	1.5	1.1
6	55946592	55947592	Gsbs	-0.4	0.6	1.1
6	65073808	65074808	Ptgds2	-0.2	1.6	1.9
6	65881686	65882686	4930544G11Rik	-0.2	0.9	1.1
6	67305215	67306215	Il12rb2	0.0	1.6	1.6
6	67464874	67465874	Tacstd2	0.2	2.0	1.9
6	73398628	73399628	4931417E11Rik	-0.3	1.7	2.0
6	82367632	82368632	Tacr1	0.1	1.2	1.1
6	82367944	82368944	Tacr1	0.0	1.9	1.9
6	87338498	87339498	1810036H07Rik	-0.5	0.6	1.1
6	87394173	87395173	Bmp10	-0.2	1.2	1.4
6	87498131	87499131	Arhgap25	0.0	1.5	1.5
6	87498376	87499376	Arhgap25	-0.1	1.0	1.1
6	97025334	97026334	C130034I18Rik	-0.2	0.8	1.0
6	103476148	103477148	Chl1	0.0	1.1	1.0
6	103476275	103477275	Chl1	0.0	1.5	1.5
6	105643048	105644048	Cntn4	0.1	1.3	1.1
6	116471634	116472634	Olfr212	-0.1	1.5	1.6
6	116521545	116522545	Olfr214	-0.6	0.6	1.3
6	122247951	122248951	Klrg1	-0.4	1.0	1.4
6	122591542	122592542	Dppa3	-0.1	1.2	1.4
6	122707622	122708622	Slc2a3	0.0	1.7	1.7
6	129296946	129297946	4922502D21Rik	-0.4	0.8	1.2
6	136745438	136746438	Gucy2c	0.1	1.4	1.4
7	6119012	6120012	Olfr1349	-0.1	1.1	1.3
7	6813475	6814475	BC023179	-0.8	1.4	2.2
7	25085082	25086082	4732471J01Rik	0.0	1.7	1.7
7	28249144	28250144	Sycn	0.2	1.6	1.4
7	30131563	30132563	AY078069	0.0	1.4	1.4
7	30223381	30224381	Gm1082	0.2	1.9	1.7
7	30379844	30380844	4930479M11Rik	-0.3	0.8	1.1
7	30532035	30533035	Ffar2	-0.5	1.3	1.8
7	49777668	49778668	Slc6a5	0.5	1.8	1.3
7	51378547	51379547	Tmem16e	0.1	1.4	1.3
7	62298589	62299589	Mkrn3	0.1	1.2	1.1
7	72179608	72180608	Gm489	0.2	1.6	1.4

7	84481412	84482412	Fah	0.8	1.9	1.1
7	100456665	100457665	Mrpl48	0.1	1.5	1.4
7	104092136	104093136	Trim6	0.2	1.2	1.0
7	105211035	105212035	Olfr691	-0.8	0.4	1.1
7	108873999	108874999	Ric3	0.1	1.2	1.1
7	120119318	120120318	Abca15	-0.4	1.6	2.1
7	126842625	126843625	1810010M01Rik	-0.5	0.9	1.4
7	136789110	136790110	Mgmt	-0.3	1.0	1.3
7	140959494	140960494	Trp53i5	-0.5	0.8	1.3
8	3352936	3353936	A430078G23Rik	0.0	1.2	1.2
8	15045546	15046546	LOC434285	-0.2	1.6	1.9
8	26139798	26140798	Adam18	0.1	1.6	1.5
8	26190833	26191833	Adam3	-0.1	1.2	1.3
8	46806783	46807783	Klkb1	-0.5	1.6	2.0
8	48850665	48851665	Stox2	-0.4	1.1	1.5
8	50173007	50174007	Odz3	-0.3	2.3	2.6
8	58832528	58833528	Glra3	-0.5	0.8	1.4
8	59186234	59187234	Hpgd	0.0	1.4	1.4
8	61803396	61804396	BC030500	-0.6	1.4	1.9
8	69632269	69633269	Npy1r	-0.3	0.7	1.0
8	86430750	86431750	Olfr370	-1.0	1.7	2.6
8	89456195	89457195	Abcc12	-0.1	1.2	1.2
8	107181803	107182803	Cmtm2a	-0.4	1.0	1.4
8	107210865	107211865	Cmtm2b	0.0	1.4	1.4
8	112503145	112504145	Pkd1l3	0.0	1.3	1.3
9	6168083	6169083	Pdgfd	0.2	1.3	1.1
9	8861760	8862760	Pgr	0.1	1.4	1.3
9	15624353	15625353	Mtnr1b	0.1	1.2	1.1
9	29712075	29713075	Hnt	-0.3	0.9	1.2
9	44805682	44806682	A530065I17Rik	0.2	1.2	1.0
9	53772938	53773938	Elmod1	0.4	1.4	1.0
9	76108955	76109955	Hcrtr2	0.0	1.0	1.1
9	105667248	105668248	E330026B02Rik	-0.2	1.2	1.4
9	110795860	110796860	Lrrc2	-0.2	1.0	1.3
9	112032115	112033115	Arpp21	-1.0	1.3	2.2
9	112032306	112033306	Arpp21	-0.8	1.2	2.0
9	119672058	119673058	Scn11a	-0.3	1.1	1.4
10	6473374	6474374	Plekhg1	-0.2	1.4	1.6
10	19045452	19046452	Olig3	-0.2	1.4	1.6
10	22509044	22510044	Tcf21	-0.5	1.3	1.8
10	23638473	23639473	Taar3	-0.6	1.3	2.0
10	23649409	23650409	Taar4	-0.6	0.4	1.0
10	23798450	23799450	Taar9	-0.9	0.5	1.4
10	28832754	28833754	6330407J23Rik	-0.3	1.7	2.0
10	28833110	28834110	N/A	-0.4	1.8	2.2
10	39675559	39676559	BC021785	-0.2	1.3	1.6
10	49471324	49472324	Grik2	-0.3	1.9	2.3
10	61817314	61818314	Hkdc1	0.3	1.4	1.1
10	62456473	62457473	0610038K03Rik	0.5	1.6	1.1
10	71375490	71376490	1700049L16Rik	-0.6	0.7	1.3
10	77296048	77297048	1700009J07Rik	-0.7	0.9	1.7
10	88935926	88936926	Nr1h4	-0.5	0.8	1.4
10	105873948	105874948	Ppfia2	0.0	1.2	1.2
10	106889744	106890744	Myf5	-0.2	1.7	1.9
10	106898339	106899339	Myf6	-0.3	0.8	1.1
10	108103702	108104702	C6.1AL	-0.2	1.2	1.3
10	117607434	117608434	Il22 Iltifb	-0.2	1.1	1.3
10	117697509	117698509	Iltifb	-0.6	1.2	1.8
10	117843539	117844539	Ifng	-0.6	0.6	1.1
10	119221948	119222948	Grip1	-0.7	1.0	1.7
10	119222071	119223071	Grip1	-0.8	0.8	1.6
11	11014246	11015246	A930041G11Rik	0.0	1.6	1.6
11	11780157	11781157	Ddc	-0.5	0.9	1.4
11	42262679	42263679	Gabrb2	-0.1	1.1	1.2
11	46078479	46079479	C030019I05Rik	-0.1	1.1	1.2
11	46247652	46248652	RP23-273O7.4	-0.4	1.2	1.7
11	50730366	50731366	Zfp454	0.2	1.6	1.4
11	58593823	58594823	Olfr315	-0.4	1.1	1.6
11	59428660	59429660	Olfr225	0.0	1.0	1.0
11	67586848	67587848	Glp2r	0.0	1.0	1.1
11	68409749	68410749	Ccdc42	0.2	1.4	1.2
11	69211041	69212041	Gm740	-0.4	1.0	1.4
11	81389234	81390234	1700071K01Rik	0.5	1.6	1.1
11	99059570	99060570	6330509G02Rik	-0.4	0.8	1.2
11	99138531	99139531	4631426H08Rik	-0.1	1.4	1.6
11	99153557	99154557	4732407F15Rik	0.0	1.4	1.4
11	99166684	99167684	Krt1-c29	0.0	1.5	1.5
11	99253726	99254726	9030623C06Rik	-0.2	0.9	1.2

11	99308694	99309694	Krt1-23	-0.5	1.9	2.4
11	99336848	99337848	4732494G06Rik	-0.3	1.6	1.9
11	99911814	99912814	Krt1-24	0.3	1.9	1.6
11	100023100	100024100	Krt1-14	-0.9	0.4	1.3
11	113881541	113882541	Sdk2	0.5	1.5	1.0
11	115150368	115151368	Otop3	0.6	1.6	1.0
12	3954466	3955466	Pomc1	0.0	1.2	1.1
12	11462397	11463397	Vsnl1	-0.1	1.0	1.0
12	29168251	29169251	Allc	-0.7	0.4	1.1
12	32792957	32793957	Pik3cg	-0.5	1.1	1.5
12	32793217	32794217	Pik3cg	-0.4	1.1	1.5
12	36502683	36503683	Agr2	-0.5	0.5	1.0
12	37618311	37619311	Meox2	-0.2	2.1	2.3
12	37618412	37619412	Meox2	-0.2	2.2	2.4
12	62442144	62443144	Lrfn5	-0.2	1.3	1.5
12	62442314	62443314	Lrfn5	0.0	1.2	1.2
12	66143621	66144621	Wdr20	-0.2	1.6	1.8
12	66143834	66144834	Wdr20	0.0	1.8	1.8
12	82403389	82404389	1810020G14Rik	-0.4	1.1	1.4
12	85306447	85307447	Rnf113a2	0.2	1.3	1.1
12	86575322	86576322	Batf	-0.3	0.8	1.1
12	93289101	93290101	LOC625794	-0.6	0.7	1.3
12	103299855	103300855	Cox8c	0.1	1.4	1.2
12	104096079	104097079	Serpina1f	-0.7	0.9	1.5
12	105623256	105624256	Tcl1	-0.4	1.5	1.9
13	56904710	56905710	Trpc7	-0.1	1.1	1.2
14	9954786	9955786	Fhit	0.1	1.7	1.6
14	25729626	25730626	Asb14	-0.5	1.0	1.5
14	47586557	47587557	Otx2	0.5	1.7	1.2
14	49791993	49792993	Tep1	-0.3	1.4	1.6
14	56078831	56079831	Gjb6	-0.2	0.8	1.0
14	62462702	62463702	Tdh	-0.1	1.2	1.3
14	68124782	68125782	Nkx2-6	0.1	1.1	1.0
14	68343647	68344647	Loxl2	0.4	1.4	1.0
14	68344055	68345055	Loxl2	0.2	1.5	1.3
14	73870079	73871079	Lcp1	-0.2	0.8	1.0
14	83253089	83254089	Pcdh17	0.2	1.7	1.6
14	83253197	83254197	Pcdh17	0.3	1.6	1.3
14	95401668	95402668	Klhl1	-0.2	1.1	1.3
14	102728875	102729875	Ednrb	0.2	2.2	1.9
14	102729581	102730581	Ednrb	-0.2	1.6	1.8
14	107795590	107796590	Slitrk1	-0.8	1.1	1.9
14	119177117	119178117	Oxgr1	-0.3	1.5	1.9
14	123859749	123860749	Fgf14	-0.6	0.7	1.3
14	123859855	123860855	Fgf14	-0.6	0.6	1.3
15	4325144	4326144	Plcxd3	-0.2	0.9	1.2
15	10121870	10122870	Prlr	-0.1	1.0	1.1
15	10121988	10122988	Prlr	0.0	1.0	1.0
15	28147854	28148854	Dnahc5	-0.2	0.9	1.0
15	37177124	37178124	Grhl2	0.7	1.9	1.3
15	37177335	37178335	Grhl2	0.8	2.0	1.2
15	74645830	74646830	Hemt1	0.0	1.4	1.4
15	98168473	98169473	Olfr284	-0.5	1.4	1.9
15	98206642	98207642	Olfr283	-0.2	1.8	2.0
15	98283845	98284845	Olfr281	-0.2	1.6	1.9
15	100192118	100193118	Yghl1-4	-0.3	0.9	1.2
15	100763858	100764858	Scn8a	-0.4	1.8	2.2
15	101301011	101302011	Krt2-10	-0.8	0.4	1.1
15	101378193	101379193	Krt2-20	0.2	1.4	1.2
15	101389933	101390933	Krt2-18	-0.2	0.8	1.0
15	101401438	101402438	Krtcap1	-0.3	1.0	1.3
15	101645669	101646669	Krt2-17	-0.1	0.9	1.1
15	101752245	101753245	Krt2-4	0.2	1.8	1.6
16	9905611	9906611	Grin2a	0.6	1.8	1.1
16	10083305	10084305	4930517K11Rik	0.0	2.4	2.4
16	22537048	22538048	Dgkg	-0.2	1.3	1.5
16	32084723	32085723	Lrrc33	0.5	1.7	1.1
16	48350362	48351362	Morc1	0.0	1.0	1.0
16	64420836	64421836	Ckt2	-0.3	2.0	2.3
16	74293400	74294400	Robo2	-0.1	1.2	1.3
16	88178587	88179587	Grik1	0.2	1.8	1.6
16	88395337	88396337	Cldn17	-0.3	0.8	1.1
16	88784808	88785808	Krtap8-2	-0.4	0.7	1.2
16	93896877	93897877	Cldn14	-0.1	1.0	1.1
16	96305113	96306113	Sh3bgr	0.2	1.5	1.3
16	97765450	97766450	Tmprss2	0.4	1.5	1.0
17	21087942	21088942	A830058L05Rik	-0.1	1.0	1.1
17	30924258	30925258	Tmprss3	-0.3	0.9	1.2

17	30925487	30926487	Tmprss3	-0.1	0.9	1.0
17	31005554	31006554	Tsga2	0.0	1.2	1.2
17	31023692	31024692	Slc37a1	0.3	1.4	1.1
17	32876175	32877175	Morc2b	-0.5	1.3	1.7
17	33808082	33809082	Psmb8	0.5	1.6	1.2
17	35113603	35114603	Pou5f1	-0.1	0.9	1.0
17	36588790	36589790	2410137M14Rik	-0.3	0.7	1.1
17	40257060	40258060	Crisp2	-0.7	1.3	1.9
17	47916568	47917568	Bzrpl1	0.2	1.4	1.1
17	55923716	55924716	M6prbp1	-0.4	1.0	1.4
17	56790228	56791228	Tnfsf9	0.2	2.2	2.0
17	56963934	56964934	Vav1	0.2	1.7	1.5
17	70420423	70421423	Dlgap1	0.0	1.1	1.1
17	71657242	71658242	BC027072	-0.1	1.3	1.4
17	73854537	73855537	Xdh	-0.2	1.0	1.1
17	81646195	81647195	Slc8a1	-0.5	1.4	1.9
17	90362539	90363539	LOC623954	0.0	1.1	1.0
17	93503212	93504212	Adcyap1	-0.3	1.0	1.3
18	31459638	31460638	Rit2	-0.1	1.2	1.3
18	37078148	37079148	Pcdha4	0.2	1.4	1.2
18	37117940	37118940	Pcdha8	0.2	1.5	1.3
18	37390471	37391471	Pcdhb1	-0.1	1.1	1.2
18	37568044	37569044	Pcdhb13	-0.3	1.8	2.1
18	42520115	42521115	Pou4f3	0.4	1.5	1.1
18	52456404	52457404	Ftmt	0.1	1.5	1.5
18	52818558	52819558	Gykl1	-0.5	1.0	1.5
18	66094970	66095970	Cplx4	-0.2	0.8	1.0
18	72475447	72476447	Dcc	0.1	1.4	1.3
18	86846740	86847740	Cbln2	0.2	1.2	1.0
19	12103285	12104285	Olfr1423	-0.3	0.8	1.1
19	22514804	22515804	Trpm3	-0.3	0.8	1.1
19	25739007	25740007	Dmrt2	0.2	1.9	1.7
19	32168637	32169637	Asah2	0.8	1.9	1.0
19	44039299	44040299	Cpn1	0.3	1.8	1.5
19	58723114	58724114	Pnlip	-0.3	1.3	1.5
X	6198063	6199063	Ccnb3	0.2	1.2	1.0
X	8189943	8190943	Fthl17	0.2	1.4	1.2
X	17302328	17303328	1700094E07Rik	-0.1	1.5	1.6
X	20517323	20518323	LOC331392	0.1	1.8	1.7
X	33641986	33642986	Gm9	0.8	1.8	1.1
X	33675936	33676936	Rhox2	0.3	1.8	1.5
X	33696261	33697261	LOC665169 Rhox4a	-0.2	1.9	2.0
X	33696404	33697404	Rhox4a	-0.1	2.0	2.2
X	33737970	33738970	Rhox4b	-0.2	2.1	2.3
X	33785722	33786722	Rhox4c	-0.4	1.8	2.1
X	33819884	33820884	Rhox4b	-0.1	1.9	2.1
X	33819991	33820991	Rhox4d	-0.1	1.9	2.0
X	33862906	33863906	Rhox4e	-0.3	2.3	2.6
X	33913143	33914143	Rhox4h	-0.2	1.8	2.0
X	33956784	33957784	Rhox4g	-0.1	1.7	1.8
X	33967141	33968141	LOC621852	-0.2	1.3	1.5
X	34071181	34072181	Rhox6	0.1	1.4	1.3
X	34145836	34146836	Rhox9	0.1	1.5	1.5
X	34310601	34311601	Rhox10	0.2	1.6	1.4
X	34329266	34330266	Rhox11	0.0	1.6	1.6
X	53849985	53850985	Gpr101	0.2	1.3	1.0
X	70429503	70430503	Tktl1	0.3	1.8	1.5
X	79740143	79741143	Tsga8	0.1	1.4	1.3
X	82508438	82509438	CN716893	0.3	1.7	1.3
X	85661908	85662908	Pet2	0.4	1.8	1.4
X	101059375	101060375	Magee2	0.1	1.3	1.1
X	114597090	114598090	Tgifx1	0.0	1.3	1.3
X	118513726	118514726	Nap1l3	0.2	1.2	1.1
X	118513829	118514829	Nap1l3	0.1	1.1	1.0
X	130648179	130649179	Pramel3 N/A	0.0	1.9	1.9
X	130719147	130720147	4930481M05 N/A	-0.1	1.9	2.1
X	136159297	136160297	Tex13	0.2	1.8	1.6
X	159104934	159105934	1700045I19Rik	0.2	1.5	1.3

Gain of DNA methylation in *Lsh-/-* MEFs (2-fold)

Chr	Start	Stop	Gene	av_LSH KO	av_WT	DM(WT-KO)
1	40734703	40735703	4931419K03Rik	1.0	-0.1	-1.1
1	154716818	154717818	Nmnat2	1.0	-0.5	-1.5
1	196639109	196640109	Cd34	1.9	0.0	-1.9
2	24783930	24784930	2810443J12Rik	1.7	-0.6	-2.2
2	44934525	44935525	Zfhx1b	1.6	0.4	-1.3
2	51756396	51757396	0610033I05Rik	1.7	0.3	-1.4
2	51756672	51757672	0610033I05Rik	1.6	0.1	-1.5
2	60183603	60184603	Ly75	1.2	0.1	-1.2
2	91244961	91245961	1110051M20Rik	1.7	-0.3	-2.0
2	112085026	112086026	Slc12a6	1.2	-0.2	-1.4
2	119353739	119354739	Ndufaf1	1.1	0.0	-1.1
2	135432770	135433770	Plcb4	1.3	0.2	-1.0
2	152653446	152654446	Ttll9	1.7	-0.6	-2.3
2	181528054	181529054	Samd10	2.2	1.0	-1.2
3	10333333	10334333	Zfand1	1.1	0.0	-1.0
3	55869953	55870953	Mab21l1	2.0	-0.1	-2.1
3	86633642	86634642	Lrba Mab21l2	2.6	-0.3	-2.9
3	114022094	114023094	Col11a1	1.1	0.1	-1.1
3	114022228	114023228	Col11a1	1.1	0.1	-1.1
3	116504626	116505626	Dbt	2.4	-0.4	-2.8
4	21970007	21971007	Coq3	1.5	0.0	-1.5
4	48684113	48685113	2310039E09Rik	1.4	-0.5	-1.9
4	129125682	129126682	Ccdc28b	1.4	-0.7	-2.1
4	134766020	134767020	Npal3	1.0	0.0	-1.0
5	13404814	13405814	Sema3a	1.3	0.2	-1.1
5	31197975	31198975	Khk	1.1	0.0	-1.0
5	125629553	125630553	Scarb1	1.6	0.2	-1.5
6	52093914	52094914	Hoxa2	2.6	1.4	-1.2
6	52147656	52148656	Hoxa7	1.7	0.4	-1.3
6	72309401	72310401	Vamp5	1.5	0.4	-1.1
6	83121511	83122511	Wdr54 1700003E16Rik	1.9	0.4	-1.5
6	83121622	83122622	1700003E16Rik	2.0	0.3	-1.7
6	100670331	100671331	BC049816	1.0	0.0	-1.0
6	125180699	125181699	Vamp1	1.0	0.0	-1.0
6	131258410	131259410	2010012C16Rik	1.6	0.2	-1.4
6	135030878	135031878	Gprc5a	1.3	0.0	-1.4
6	149098185	149099185	4833442J19Rik	2.5	1.1	-1.4
6	149098290	149099290	4833442J19Rik	2.6	1.1	-1.4
7	24928107	24929107	Dedd2	0.5	-0.5	-1.0
7	44842316	44843316	Cpt1c	1.4	0.0	-1.5
7	78450796	78451796	Ntrk3	2.1	0.9	-1.1
7	78450988	78451988	Ntrk3	2.1	1.1	-1.0
7	81662444	81663444	3110040N11Rik	1.1	-0.7	-1.8
7	139829543	139830543	6430531B16Rik	2.3	1.1	-1.1
7	141117221	141118221	Irf7	1.7	0.0	-1.6
8	23631447	23632447	Nek3	1.1	-0.5	-1.6
8	73695138	73696138	Ifi30	1.1	-0.3	-1.3
8	108143732	108144732	C76566	1.4	0.2	-1.2
8	110291812	110292812	Nqo1	1.1	-0.3	-1.4
9	57634121	57635121	Arid3b	1.6	-0.4	-2.0
9	78400960	78401960	Cd109	1.3	0.2	-1.1
10	33986387	33987387	Tspyl4	1.8	-0.1	-1.9
10	68928921	68929921	Ank3	0.4	-0.7	-1.1
10	68929022	68930022	Ank3	0.5	-0.5	-1.0
10	128336266	128337266	Itga7	2.5	1.1	-1.4
11	29414528	29415528	2010316F05Rik	1.5	0.0	-1.4
11	54941039	54942039	Gm2a	1.7	0.5	-1.1
11	69619363	69620363	Tmem102	1.8	0.0	-1.8
11	84644659	84645659	BC022224	1.6	0.3	-1.4
11	96643305	96644305	Nfe2l1	2.3	0.8	-1.5
11	96643545	96644545	Nfe2l1	2.2	0.8	-1.5
11	101709944	101710944	Meox1	1.2	0.1	-1.1
11	110153265	110154265	Abca5	1.2	0.2	-1.0
12	8999399	9000399	Wdr35	1.2	0.0	-1.3
12	8999518	9000518	Wdr35	1.2	0.0	-1.3
12	96087584	96088584	Flrt2	2.7	0.2	-2.5
13	21447866	21448866	4921509E05Rik	2.8	1.7	-1.1
13	21448009	21449009	4921509E05Rik	2.7	1.4	-1.3
13	64138030	64139030	Slc35d2	2.5	0.1	-2.4
13	101115331	101116331	Mccc2	1.0	0.0	-1.0
14	30994300	30995300	Ncoa4	1.0	-0.1	-1.0
14	45308565	45309565	Bmp4	1.1	0.0	-1.1
14	60314033	60315033	Ebpl	1.4	-0.1	-1.5
15	8657125	8658125	Slc1a3	1.9	-0.4	-2.3

15	8657445	8658445	Slc1a3	1.8	-0.3	-2.1
15	51707069	51708069	D530033C11Rik	1.3	0.0	-1.3
15	51707318	51708318	D530033C11Rik	1.1	0.1	-1.0
15	55137333	55138333	Col14a1	2.1	1.1	-1.0
15	76134000	76135000	Oplah	2.5	1.5	-1.0
15	98555664	98556664	Fkbp11	1.9	0.3	-1.5
16	3787102	3788102	Zfp597	2.5	1.5	-1.0
16	56810614	56811614	Tmem45a	0.8	-0.4	-1.3
17	24949681	24950681	Unkl	2.0	0.7	-1.3
17	45616590	45617590	Mad2l1bp	1.6	-0.2	-1.8
17	79622867	79623867	Cyp1b1	1.4	0.1	-1.3
18	34533723	34534723	Pkd2l2	1.3	0.1	-1.2
18	34703401	34704401	Nme5	2.4	1.3	-1.1
18	60626360	60627360	2010002N04Rik	1.7	0.1	-1.5
18	77251237	77252237	3110023G01Rik	1.4	-0.5	-1.9
19	4854628	4855628	Ctsf	1.3	0.3	-1.0
19	44346898	44347898	Scd2	1.4	0.0	-1.3
19	47589831	47590831	Obfc1	2.1	-0.1	-2.2

Loss of DNA methylation in *G9a-/-* ES (1.4-fold)

Chr	Start	Stop	Gene	av_G9a KO	av_WT	M(WT-KO)
1	98702144	98703144	9530060I07	0.22	1.04	0.82
1	95655179	95656179	D2hgdh	-0.04	0.57	0.61
1	87431373	87432373	LOC434484	-0.87	-0.27	0.6
1	1.36E+08	1.36E+08	LOC667414	1.24	1.74	0.5
1	98702041	98703041	Slco4c1	0.29	1.06	0.77
1	90042629	90043629	Ugt1a1	-0.21	0.35	0.56
2	1.52E+08	1.52E+08	2400002F11Rik	-0.15	1.61	1.76
2	1.79E+08	1.79E+08	4930591A17Rik	-0.01	0.63	0.64
2	1.30E+08	1.30E+08	AU015228	0.59	1.25	0.66
2	1.74E+08	1.74E+08	Gnas	0.42	1.43	1.01
2	55252041	55253041	Kcnj3	0.32	0.84	0.52
2	1.57E+08	1.57E+08	Nnat	0.87	1.74	0.87
2	24916640	24917640	Noxa1	1.11	1.85	0.74
2	85521745	85522745	Olfr1009	0.24	0.87	0.63
2	61604791	61605791	Tbr1	0.42	0.93	0.51
3	1.03E+08	1.03E+08	5730470L24Rik	0.19	0.72	0.53
3	96014182	96015182	Anp32e	-0.93	-0.35	0.58
3	36244596	36245596	BC050811	0.38	0.91	0.53
3	1.01E+08	1.01E+08	Cd2	0.82	1.35	0.53
3	1.30E+08	1.30E+08	Egf	-0.01	0.51	0.52
3	90084072	90085072	Hax1	-1.02	-0.42	0.6
3	49852976	49853976	Pcdh18	0.04	0.63	0.59
3	26828203	26829203	Spata16	0.48	1.22	0.74
3	96266131	96267131	Sv2a	-0.53	0	0.53
3	57382320	57383320	Tm4sf1	1.02	1.55	0.53
3	1.46E+08	1.46E+08	Wdr63	0.03	0.99	0.96
4	33323944	33324944	Rragd	0.28	0.79	0.51
5	1.46E+08	1.46E+08	1700001J03Rik	0.2	0.76	0.56
5	93423554	93424554	Cxcl10	0.03	0.59	0.56
5	71978650	71979650	Gabrb1	-0.3	0.35	0.65
5	1.15E+08	1.15E+08	Oasl2	-0.12	0.82	0.94
5	98420779	98421779	Prdm8	0.11	0.64	0.53
5	3241895	3242895	Rpl26	0.89	1.54	0.65
5	1.38E+08	1.38E+08	Trfr2	0.01	0.71	0.7
6	1.31E+08	1.31E+08	2010012C16Rik	-1.3	0.39	1.69
6	72291390	72292390	2500002L14Rik	-0.94	-0.29	0.65
6	1.29E+08	1.29E+08	4922502D21Rik	-0.16	0.78	0.94
6	89061255	89062255	4933427D06Rik	0.64	1.25	0.61
6	1.37E+08	1.37E+08	Art4	0.84	1.36	0.52
6	18058555	18059555	Asz1	0.74	1.52	0.78
6	1.29E+08	1.29E+08	BC048546	-0.03	0.6	0.63
6	86403921	86404921	C87436	-0.72	-0.13	0.59
6	1.29E+08	1.29E+08	Cd69	-0.62	-0.12	0.5
6	1.25E+08	1.25E+08	Emg1	-0.53	0.12	0.65
6	25640239	25641239	Gpr37	-0.78	-0.19	0.59
6	1.35E+08	1.35E+08	Gprc5d	0.06	0.65	0.59
6	1.37E+08	1.37E+08	Gucy2c	1.29	1.83	0.54
6	82739531	82740531	Hk2	-0.36	0.14	0.5
6	48484037	48485037	Lrrc61	-0.96	-0.27	0.69
6	30687572	30688572	Mest	0.19	1.07	0.88
6	17440740	17441740	Met	0.6	1.1	0.5
6	1.35E+08	1.35E+08	Pbp2	0.65	1.21	0.56
6	4696805	4697805	Peg10	0.46	1	0.54
6	57464504	57465504	Ppm1k	-0.27	0.33	0.6
7	30379844	30380844	4930479M11Rik	-0.22	0.77	0.99

7	4135029	4136029	6030429G01Rik	0.05	0.78	0.73
7	12177159	12178159	6330408A02Rik	0.38	0.91	0.53
7	1.20E+08	1.20E+08	Abca15	1.32	2.05	0.73
7	1.25E+08	1.25E+08	Jmjd5	0.42	1.09	0.67
7	12176961	12177961	Lig1	0.48	0.99	0.51
7	97274777	97275777	Ndufc2	0.06	0.66	0.6
7	6119012	6120012	Olfr1349	0.66	1.29	0.63
7	6333993	6334993	Peg3	0.06	0.99	0.93
7	6334168	6335168	Usp29	-0.07	0.98	1.05
7	5824864	5825864	Zfp371	0.59	1.24	0.65
8	38461598	38462598	AI429214	-0.34	0.97	1.31
8	70423363	70424363	Nat2	-0.16	0.52	0.68
8	50173007	50174007	Odz3	1.42	1.98	0.56
8	86430750	86431750	Olfr370	1.14	1.8	0.66
9	78153316	78154316	Dppa5	0.86	1.4	0.54
9	43863546	43864546	Rnf26 no.1	0.86	1.45	0.59
9	96540527	96541527	Zbtb38	0.33	1.17	0.84
10	1.08E+08	1.08E+08	C6.1AL	0.71	1.62	0.91
10	99188644	99189644	Csl no.1	1.1	1.69	0.59
10	99188846	99189846	Csl	1.06	1.58	0.52
10	57631517	57632517	LOC245263	1.21	1.82	0.61
10	21379760	21380760	LOC432436	0.82	1.35	0.53
10	1.20E+08	1.20E+08	Msrb3	1.22	1.73	0.51
10	1.07E+08	1.07E+08	Myf5	0.63	1.19	0.56
10	12780485	12781485	Plagl1 no.1	0.28	1	0.72
10	12780606	12781606	Plagl1	0.26	1	0.74
10	50583565	50584565	Sim1 no.1	-0.46	0.08	0.54
10	50583772	50584772	Sim1	-0.51	0.01	0.52
10	22509044	22510044	Tcf21	-0.01	0.58	0.59
11	16951436	16952436	1500041B16Rik	0.03	0.61	0.58
11	81389234	81390234	1700071K01Rik	1.42	1.98	0.56
11	84773534	84774534	Car4	0.27	0.78	0.51
11	68409749	68410749	Ccdc42	0.81	1.55	0.74
11	41843242	41844242	Gabrg2 no.1	-0.45	0.24	0.69
11	41843455	41844455	Gabrg2	-0.46	0.05	0.51
11	69211041	69212041	Gm740	0.49	1.09	0.6
11	36787166	36788166	Odz2	0.21	0.71	0.5
11	58441739	58442739	Olfr323	0.95	1.51	0.56
11	16951884	16952884	Plek	0.12	0.65	0.53
11	80626109	80627109	Tmem98	-0.64	-0.1	0.54
11	22871536	22872536	U2af1-rs1	0.22	0.79	0.57
11	11361911	11362911	Zpbp	-0.22	0.29	0.51
12	1.03E+08	1.03E+08	Cox8c	0.82	1.57	0.75
12	34513171	34514171	Ferd3l	-0.03	0.95	0.98
12	40709238	40710238	Gm889	0.55	1.05	0.5
12	37618311	37619311	Meox2 no.1	-0.12	0.48	0.6
12	37618412	37619412	Meox2	-0.06	0.54	0.6
12	66143621	66144621	Wdr20 no.1	0.74	1.46	0.72
12	66143834	66144834	Wdr20	0.76	1.68	0.92
13	99736990	99737990	0610009B10Rik	1.69	2.29	0.6
13	67999105	68000105	A530094I17Rik	-0.46	0.09	0.55
13	1.01E+08	1.01E+08	Birc1f	0.21	0.76	0.55
13	84211581	84212581	C130071C03Rik	-0.26	0.29	0.55
13	3803048	3804048	Calml3	0.78	1.45	0.67
13	70276753	70277753	D13Wsu50e	-0.06	0.44	0.5
13	40870605	40871605	Gcnt2	0.77	1.34	0.57
13	45517940	45518940	Gmpr	0.08	0.59	0.51
13	26777080	26778080	Hdgfl1	0.7	1.2	0.5
13	21723376	21724376	Hist1h2ac Hist1h2ai	0.14	0.81	0.67
13	21840440	21841440	Hist1h2ao N/A	-0.08	0.49	0.57
13	23941121	23942121	Hist1h2ba	0.99	1.65	0.66
13	21729062	21730062	Hist1h2bm	0.17	0.68	0.51
14	64220310	64221310	Fbxo16	-0.27	0.37	0.64
14	9954786	9955786	Fhit	-0.31	0.45	0.76
14	73870079	73871079	Lcp1	0.65	1.16	0.51
14	27844920	27845920	Lrtm1	-0.34	0.21	0.55
14	1.08E+08	1.08E+08	Slitrk1	-0.09	1.3	1.39
15	77945715	77946715	Cacng2	0.33	0.87	0.54
15	54749582	54750582	Enpp2 no.1	0.1	0.77	0.67
15	54749692	54750692	Enpp2	0.21	0.78	0.57
15	74645830	74646830	Hemt1	1.34	1.88	0.54
15	98283845	98284845	Olfr281	1.28	1.84	0.56
15	78000863	78001863	Rabl4	-0.78	-0.17	0.61
16	4633787	4634787	1110025F24Rik	0.54	1.08	0.54
16	52372922	52373922	Alcam	0.11	0.81	0.7
16	22537048	22538048	Dgkg	0.62	1.13	0.51
17	93503212	93504212	Adcyap1	0.57	1.08	0.51
17	40257060	40258060	Crisp2	0.38	1.01	0.63

17	6417939	6418939	Dynlt1	-0.41	0.1	0.51
17	89108501	89109501	Fshr	0.63	1.15	0.52
17	58698095	58699095	Nudt12	0.04	0.82	0.78
17	35140753	35141753	Psors1c2	0.07	0.62	0.55
17	6522768	6523768	Sytl3	1.17	1.85	0.68
17	30925487	30926487	Tmprss3	0.57	1.14	0.57
18	9618966	9619966	Cetn1	1.33	1.85	0.52
18	52818558	52819558	Gykl1	-0.05	1.07	1.12
18	13114920	13115920	Impact	-0.36	0.2	0.56
18	84220255	84221255	Tshz1	-0.76	-0.26	0.5
18	13987078	13988078	Zfp521	1.55	2.1	0.55
19	8955447	8956447	1810009A15Rik	0.12	0.67	0.55
19	46760283	46761283	As3mt	0.06	0.89	0.83
19	34592912	34593912	N/A Lip1	-0.04	0.78	0.82
X	5258382	5259382	AU022751	0.73	1.24	0.51
X	1.60E+08	1.60E+08	Bmx	0.66	1.17	0.51
X	49864486	49865486	Cxx1a Cxx1b	0.12	0.77	0.65
X	1.02E+08	1.02E+08	Cypt2	0.51	1.08	0.57
X	33967141	33968141	LOC621852	1.11	1.7	0.59
X	1.50E+08	1.50E+08	Magea2	0.32	0.91	0.59
X	1.50E+08	1.50E+08	Magea8	0.29	1.02	0.73
X	1.31E+08	1.31E+08	Nxf3	0.03	0.58	0.55
X	1.30E+08	1.30E+08	Rpl36a	-0.65	0.31	0.96
X	38615	39615	Sybl1	-0.34	0.49	0.83
X	1.40E+08	1.40E+08	Trpc5 no.1	-0.67	0.01	0.68
X	1.40E+08	1.40E+08	Trpc5	-0.67	0.01	0.68
X	27106937	27107937	Xmr	-0.83	-0.3	0.53
X	95139999	95140999	Yipf6	-0.47	0.04	0.51
Y	581563	582563	Uty no.1	-0.32	1.66	1.98
Y	581667	582667	Uty	-0.33	1.62	1.95

Gain of DNA methylation in *G9a-/-* ES

Chr	Start	Stop	GeneName	av_G9a KO	av_WT	M(WT-KO)
1	58657339	58658339	Cflar	1.28	0.12	-1.16
1	58739477	58740477	Casp8	0.81	-0.2	-1.01
1	58739585	58740585	Casp8	0.81	-0.2	-1.01
1	74594460	74595460	Ttll4	1.09	0.42	-0.67
1	1.30E+08	1.30E+08	Rab3gap1	0.49	-0.31	-0.8
1	1.64E+08	1.64E+08	Myoc	0.56	-0.01	-0.57
1	1.72E+08	1.72E+08	1700084C01Rik	0.51	-0.12	-0.63
2	37180741	37181741	Pdcl	1.08	-0.25	-1.33
2	1.51E+08	1.51E+08	Nsfl1c	0.85	0.17	-0.68
2	1.53E+08	1.53E+08	Ttll9	1.34	0.51	-0.83
2	1.55E+08	1.55E+08	Acss2	1.44	0.67	-0.77
2	1.69E+08	1.69E+08	1700101G07Rik	0.46	-0.07	-0.53
2	1.74E+08	1.74E+08	Gnas	1.07	0.39	-0.68
2	1.81E+08	1.81E+08	BC050777	0.27	-0.31	-0.58
3	27373576	27374576	Aadacl1	-0.31	-0.86	-0.55
3	51748665	51749665	Mgst2	0.76	0.22	-0.54
3	97973612	97974612	Pde4dip	0.44	-0.09	-0.53
4	1.36E+08	1.36E+08	6030445D17Rik	0.43	-0.08	-0.51
4	1.47E+08	1.47E+08	Fv1	2.02	0.64	-1.38
5	78090297	78091297	1700023E05Rik	1.77	0.86	-0.91
5	1.15E+08	1.15E+08	N/A	0.71	-0.01	-0.72
5	1.24E+08	1.24E+08	Ogfod2	1.26	0.36	-0.9
6	83740914	83741914	Tex261	0.4	-0.14	-0.54
7	28888661	28889661	Psmd8	0.47	-0.07	-0.54
7	28888774	28889774	Psmd8	0.47	-0.07	-0.54
7	29940862	29941862	Polr2i	0.13	-0.46	-0.59
7	1.28E+08	1.28E+08	Inpp5f	0.88	0.3	-0.58
8	4779034	4780034	Shcbp1	0.82	0.21	-0.61
8	11839992	11840992	1700018L24Rik	0.59	0.09	-0.5
8	20438598	20439598	A430108E01Rik 2610005L07Rik	1.27	0.02	-1.25
8	95743977	95744977	Aytl1	0.27	-0.27	-0.54
9	13269840	13270840	LOC434635	0.17	-0.37	-0.54
9	19372003	19373003	Zfp75	0.37	-0.29	-0.66
9	20135060	20136060	Zfp560	1.15	0.36	-0.79
9	21688175	21689175	BC018242	0.85	0.11	-0.74
9	21938067	21939067	9530015I07Rik	1.11	0.47	-0.64
9	44276916	44277916	Blr1	0.35	-0.2	-0.55
9	1.07E+08	1.07E+08	Cish	0.35	-0.22	-0.57
10	81359779	81360779	BC024063	0.35	-0.44	-0.79
11	29030243	29031243	N/A Pnpt1	0.45	-0.26	-0.71
11	69478266	69479266	Mpdu1	0.03	-0.59	-0.62
11	97591467	97592467	1700001P01Rik	0.77	0.25	-0.52

12	18157029	18158029	LOC245297	0.57	0.05	-0.52
12	85532143	85533143	7420416P09Rik	0.01	-0.55	-0.56
12	1.09E+08	1.09E+08	Wdr25	0.82	0.19	-0.63
13	1.21E+08	1.21E+08	LOC633640	0.76	0.2	-0.56
13	1.21E+08	1.21E+08	3110070M22Rik	1.15	0.62	-0.53
14	19462394	19463394	Sec24c	0.88	0.02	-0.86
14	19840208	19841208	Ap3m1	0.44	-0.07	-0.51
14	51941865	51942865	C920008G01Rik	1.33	0.77	-0.56
14	53385316	53386316	Mmp14	-0.11	-0.69	-0.58
14	53559613	53560613	4931414P19Rik	0.64	-0.04	-0.68
14	62941390	62942390	4930578I06Rik	0.45	-0.05	-0.5
14	66660255	66661255	Cdca2	0.57	0.02	-0.55
14	1.23E+08	1.23E+08	Vgcnl1	0.19	-0.41	-0.6
15	99981929	99982929	AI317237	0.9	0.4	-0.5
16	4656340	4657340	Hmox2	0	-0.53	-0.53
16	13022813	13023813	Ercc4	0.05	-0.48	-0.53
16	92189677	92190677	1190017O12Rik	1.26	0.44	-0.82
17	20439296	20440296	V1re6	0.58	-0.1	-0.68
17	25093432	25094432	Mcpt6	0.67	0.08	-0.59
17	25093649	25094649	Mcpt6	0.96	0.14	-0.82
17	28417076	28418076	Slc26a8	0.69	-0.24	-0.93
17	30353446	30354446	Dnahc8	0.39	-0.13	-0.52
17	42351655	42352655	Cd2ap	0.73	0.12	-0.61
17	56559071	56560071	Acsbg2	0.81	0.3	-0.51
18	37055657	37056657	Pcdha1	1.24	0.52	-0.72
18	37055758	37056758	Pcdha1	1.08	0.52	-0.56
18	37071680	37072680	Pcdha3	1.39	0.75	-0.64
18	37123353	37124353	Pcdha9	1.55	0.68	-0.87
18	37145703	37146703	Pcdha12	1.02	0.45	-0.57
18	37146153	37147153	Pcdha11	1.11	0.55	-0.56
18	37862409	37863409	Pcdhga9	0.43	-0.24	-0.67
18	37881246	37882246	Pcdhga11	0.31	-0.29	-0.6
18	74393140	74394140	Mbd1	1.24	0.59	-0.65
19	42308593	42309593	4933417O08Rik	0.35	-0.15	-0.5

Loss of DNA methylation in *Smchd1-/-*

Chr	Start	Stop	Gene Name	av_Smchd1 KO	av_WT	M(WT-KO)
1	137250294	137250399	U6.608	-2.5	0.4	2.8
1	169153958	169154050	Uck2	-2.5	0.1	2.6
1	188949017	188949862	1700007P06Rik	-1.3	0.9	2.2
1	139863216	139863295	Mir181b-1	-2.3	-0.3	2.0
1	134710805	134710925	SNORA17.507	-2.5	-0.5	2.0
1	171899540	171902526	3110045C21Rik	-2.3	-0.5	1.8
1	172207507	172216900	Sh2d1b1	0.0	1.8	1.8
1	185505657	185506073	AC107787.1	-2.5	-0.8	1.7
1	59538991	59543791	Fzd7	1.6	3.3	1.7
1	188944748	188948244	Gm15509	-0.7	0.8	1.6
1	9999895	9999998	U6.223	-2.3	-0.8	1.6
1	55294021	55296902	Mars2	-2.3	-0.8	1.6
1	174886899	174887948	Olfr16	-2.3	-0.8	1.6
1	175151402	175157360	Fcer1a	-2.3	-0.8	1.6
1	132697479	132711317	Fcamr	2.3	3.9	1.5
1	65126804	65128862	Crygb	-0.7	0.7	1.5
1	175942129	175961886	Ifi205	-1.3	0.1	1.4
1	177907908	177909245	B020018G12Rik	-0.3	1.1	1.4
1	87407708	87423736	Gm16092	-2.5	-1.1	1.4
1	139863032	139863118	Mir181a-1	-2.5	-1.1	1.4
1	136538831	136543097	AC122771.1	1.3	2.6	1.3
1	193916253	193923511	Gm16621	1.4	2.6	1.2
1	152006173	152012078	AC114655.1	-0.7	0.5	1.2
1	173155185	173156510	Apoa2	0.3	1.5	1.2
1	136401915	136402040	SNORA61.5	-1.3	-0.1	1.2
1	136705852	136705945	AC131591.1	-1.3	-0.1	1.2
1	148267564	148268004	AL592462.1	-1.3	-0.1	1.2
1	176066877	176067824	Olfr424	-1.3	-0.1	1.2
1	180251239	180254302	Gm16586	-1.3	-0.1	1.2
1	23298540	23298623	Mir30c-2	-2.3	-1.1	1.2
1	133997216	133997331	U6.779	-2.3	-1.1	1.2
1	175145738	175146771	Olfr1404	-2.3	-1.1	1.2
1	175999431	176000384	Olfr430	-2.3	-1.1	1.2
1	94407625	94415092	Olfr1414	1.4	2.6	1.2
1	73061657	73062473	Tnp1	1.8	2.9	1.2
1	155251957	155255726	E330020D12Rik	2.0	3.1	1.1
1	68051660	68072470	Gm15671	-0.3	0.8	1.1

1	132916422	132921551	Il10	1.0	2.1	1.1
1	74928083	74931993	Fev	0.3	1.4	1.1
1	195333432	195335997	AL365334.1	-0.7	0.4	1.1
1	174405489	174408706	Slamf9	1.1	2.2	1.1
1	44260915	44275806	4832428D23Rik	1.3	2.3	1.0
1	15258880	15259878	Gm10566	-1.3	-0.3	1.0
1	94376258	94377196	Olfr1416	-1.3	-0.3	1.0
1	171862829	171864379	Gm15852	0.7	1.7	1.0
1	176019169	176020111	Olfr429	-1.3	-0.3	1.0
1	180567036	180567202	U1.47	-1.3	-0.3	1.0
1	132829235	132836650	Il19	2.0	3.0	1.0
1	135971747	135975697	Btg2	1.3	2.3	1.0
2	89814866	89818869	Olfr1260	-2.3	0.7	3.0
2	89136508	89137425	Olfr1230	-2.5	0.4	2.8
2	24146716	24149340	Il1f10	-2.3	0.5	2.8
2	71381466	71384749	Dlx2	1.4	4.2	2.8
2	125108424	125109714	Gm14003	-1.3	1.4	2.8
2	168886620	168886740	AL772258.1	-2.5	0.2	2.7
2	129390783	129391804	Gm14041	-2.3	0.4	2.7
2	166015197	166015301	5S_rRNA.83	-2.3	0.2	2.6
2	171131668	171131931	SCARNA13.2	-2.3	0.2	2.6
2	60249582	60250636	Gm13580	-1.3	1.2	2.5
2	71366740	71372038	Dlx1	1.8	4.3	2.4
2	24070938	24081222	Il1f6	-2.3	0.1	2.4
2	160751277	160752384	Gm14228	-2.3	0.1	2.4
2	152395074	152395274	AL845162.1	-2.5	-0.1	2.4
2	89861295	89862221	Olfr1264	-0.7	1.6	2.3
2	130230995	130231810	1700020A23Rik	-2.3	-0.1	2.2
2	147149298	147150823	Gm14110	-2.3	-0.1	2.2
2	10433594	10433689	Mir466d	-2.5	-0.3	2.2
2	89122296	89128111	Olfr1229	-2.5	-0.3	2.2
2	117074007	117074170	U1.99	-2.5	-0.3	2.2
2	19579689	19582217	Otud1	0.5	2.7	2.2
2	86989877	86990816	Olfr1111	-2.3	-0.3	2.0
2	9804668	9811167	9230102O04Rik	2.2	4.2	2.0
2	151296278	151301889	4921509C19Rik	-2.5	-0.5	2.0
2	127343028	127343545	AL731831.1	-1.3	0.6	1.9
2	158182160	158182290	SNORA71.8	-1.3	0.6	1.9
2	58759370	58760029	Gm11099	-0.7	1.2	1.9
2	151691051	151692696	Gm14157	0.7	2.5	1.8
2	86037960	86038907	Olfr1045	-2.3	-0.5	1.8
2	87622693	87623643	Olfr152	-2.3	-0.5	1.8
2	88326891	88327826	Olfr1184	-2.3	-0.5	1.8
2	130101973	130102105	SNORA51.4	-2.3	-0.5	1.8
2	151654803	151657844	Gm14155	-0.7	1.1	1.8
2	32077490	32077574	SNORD62.2	-2.5	-0.8	1.7
2	85744226	85745209	Olfr1024	-2.5	-0.8	1.7
2	129190306	129196875	Il1b	0.8	2.6	1.7
2	169407581	169408733	1700101G07Rik	-2.5	-0.8	1.7
2	179710698	179713718	4921531C22Rik	-2.5	-0.8	1.7
2	130389478	130394431	Mrps26	1.4	3.1	1.7
2	25562399	25563237	Bmyc	0.3	2.0	1.7
2	51219092	51221019	Gm13499	-1.3	0.4	1.7
2	103683575	103683670	AL928544.1	-1.3	0.4	1.7
2	144068977	144070585	4930444E06Rik	-0.3	1.4	1.7
2	167363695	167368314	Snai1	2.6	4.3	1.7
2	29617048	29621994	Gm13547	1.8	3.4	1.7
2	174661465	174666675	Gm14617	0.0	1.6	1.6
2	173220067	173221221	Gm14642	0.8	2.4	1.6
2	167684085	167688042	Gm14319	2.6	4.2	1.6
2	71585564	71588257	Gm13663	-0.3	1.2	1.6
2	87802109	87808414	Olfr1157	-1.3	0.2	1.6
2	128249390	128250301	Gm14010	-0.3	1.2	1.6
2	131124281	131125571	Gm14233	-1.3	0.2	1.6
2	147187277	147188145	Al646519	-1.3	0.2	1.6
2	11306630	11308691	Gm13291	-2.3	-0.8	1.6
2	36667669	36668610	Olfr3	-2.3	-0.8	1.6
2	49762291	49762400	Y_RNA.20	-2.3	-0.8	1.6
2	51483031	51483927	Tas2r134	-2.3	-0.8	1.6
2	130103752	130103823	Snord57	-2.3	-0.8	1.6
2	135495189	135496161	Gm14209	-2.3	-0.8	1.6
2	158185534	158185665	SNORA71.1	-2.3	-0.8	1.6
2	177282920	177297014	Gm14403	-2.3	-0.8	1.6
2	71586132	71588219	AL928963.1	-0.3	1.2	1.5
2	172113633	172116020	Gm14275	-0.3	1.2	1.5
2	3670258	3672449	Gm13185	0.8	2.3	1.4
2	75567649	75568229	Gm13656	-1.3	0.1	1.4
2	103905983	103907643	Gm10912	-1.3	0.1	1.4

2	118571494	118574546	5430417L22Rik	-1.3	0.1	1.4
2	158205270	158205402	SNORA71.3	-1.3	0.1	1.4
2	163983140	163983257	n-R5s207	-1.3	0.1	1.4
2	27395946	27398882	Gm13421	2.1	3.5	1.4
2	25508640	25511737	Lcn8	1.1	2.5	1.4
2	36334707	36335648	Olfr341	-2.5	-1.1	1.4
2	86932614	86933552	Olfr1109	-2.5	-1.1	1.4
2	103880826	103880964	SCARNA21.1	-2.5	-1.1	1.4
2	106649492	106649615	SCARNA4.3	-2.5	-1.1	1.4
2	164080099	164082376	Svs3b	-2.5	-1.1	1.4
2	131679414	131685483	Erv3	1.1	2.5	1.4
2	158201374	158211881	Snhg11	3.0	4.4	1.4
2	75615245	75620363	Gm13657	0.8	2.2	1.4
2	111086288	111096320	Gm13962	-0.7	0.6	1.4
2	79292678	79296908	Neurod1	1.0	2.3	1.3
2	106896297	106899813	Fshb	0.5	1.8	1.3
2	33571881	33574419	9430024E24Rik	1.0	2.3	1.3
2	156203661	156204704	Gm14170	0.0	1.3	1.3
2	167027076	167028299	Gm14291	1.1	2.4	1.3
2	74529948	74533160	Hoxd10	0.3	1.6	1.3
2	152418927	152425135	Defb45	1.9	3.2	1.3
2	154923525	154925817	2310005A03Rik	1.4	2.7	1.3
2	94092036	94104915	E530001K10Rik	2.0	3.3	1.3
2	31872164	31872864	Gm13609	0.5	1.8	1.3
2	179628614	179629398	Gm10711	0.5	1.8	1.3
2	153606615	153620918	Efcab8	3.2	4.5	1.3
2	164481440	164482868	Wfdc10	1.8	3.0	1.3
2	10636794	10639296	Gm10852	-0.7	0.5	1.2
2	144649402	144650151	1700010M22Rik	-0.7	0.5	1.2
2	167073817	167074078	AL591854.1	-0.7	0.5	1.2
2	18594011	18597858	Commd3	-0.3	0.9	1.2
2	89684223	89685128	Olfr48	0.0	1.2	1.2
2	150396151	150404680	Cst7	1.5	2.7	1.2
2	163340252	163340354	U6.703	-1.3	-0.1	1.2
2	171813695	171814695	Gm14272	0.8	2.1	1.2
2	10429257	10429341	Mir467d	-2.3	-1.1	1.2
2	23678573	23678747	7SK.177	-2.3	-1.1	1.2
2	25126559	25127958	Ssna1	-2.3	-1.1	1.2
2	26312664	26313836	Gm13568	-2.3	-1.1	1.2
2	37075011	37075940	Olfr366	-2.3	-1.1	1.2
2	83540814	83541273	Gm13686	-2.3	-1.1	1.2
2	111219124	111220187	Olfr1284	-2.3	-1.1	1.2
2	150574017	150576498	Gm16962	0.5	1.7	1.2
2	156211843	156212058	7SK.215	-2.3	-1.1	1.2
2	162505191	162506265	Gm14246	-2.3	-1.1	1.2
2	174093005	174093086	Mir298	-2.3	-1.1	1.2
2	180133084	180133187	Mir133a-2	-2.3	-1.1	1.2
2	169408912	169413245	Gm11011	1.8	3.0	1.2
2	43914758	43920844	A430092G05Rik	0.7	1.9	1.2
2	25532306	25537128	Lcn6	1.6	2.8	1.2
2	4208641	4222213	Gm13187	2.7	3.9	1.2
2	93820191	93823732	4921507L20Rik	1.6	2.7	1.2
2	84581272	84581335	Mir130a	0.0	1.2	1.2
2	156688864	156699299	1110008F13Rik	3.1	4.3	1.2
2	112319229	112320974	Chrm5	2.7	3.9	1.1
2	152398480	152400680	Defb21	1.3	2.4	1.1
2	164014967	164016344	Wfdc12	1.7	2.8	1.1
2	164101688	164104195	Svs4	0.5	1.6	1.1
2	26301216	26314910	0610009E02Rik	1.8	2.9	1.1
2	31064990	31066867	Gm13405	1.3	2.4	1.1
2	163961439	163966155	Gm11455	2.2	3.3	1.1
2	25480186	25483550	Fcna	2.3	3.4	1.1
2	25632537	25635799	Lcn11	0.7	1.8	1.1
2	157421319	157423056	2410019O14Rik	0.5	1.6	1.1
2	129024500	129037352	Slc20a1	3.6	4.7	1.1
2	31001977	31007811	D330023K18Rik	1.1	2.2	1.1
2	32429704	32431204	9430097D07Rik	1.6	2.6	1.1
2	31817210	31818067	Gm13442	0.3	1.3	1.0
2	36762431	36763514	Olfr354	-1.3	-0.3	1.0
2	89142806	89143747	Olfr1231	-1.3	-0.3	1.0
2	94569171	94570895	Gm13793	-1.3	-0.3	1.0
2	104932175	104933917	Gm11060	-1.3	-0.3	1.0
2	164115004	164117236	Svs3a	-1.3	-0.3	1.0
2	164439022	164443712	Wfdc6b	2.6	3.7	1.0
2	166099408	166120331	Gm11468	5.1	6.1	1.0
2	164395195	164398959	Spint3	1.4	2.4	1.0
2	25182970	25188091	AL732557.1	1.1	2.2	1.0
2	25475587	25477525	Tmem141	1.1	2.2	1.0

2	132520168	132524259	Gm14097	2.6	3.7	1.0
2	152284791	152290356	Defb23	2.6	3.6	1.0
2	180940267	180947890	Srms	3.0	4.0	1.0
3	86349503	86352205	Mab21l2	0.0	4.4	4.4
3	5212838	5215209	Gm10748	-2.3	1.1	3.4
3	55586432	55589207	Mab21l1	1.3	4.5	3.3
3	108301036	108301158	U6atac.18	-2.3	0.2	2.6
3	68112747	68113103	Gm5848	-2.5	-0.3	2.2
3	90407692	90409967	S100a4	-1.3	0.7	2.0
3	24231969	24232474	Gm7536	-2.3	-0.3	2.0
3	88187423	88188386	Bglap	-2.5	-0.5	2.0
3	4338998	4339116	n-R5s194	-2.3	-0.5	1.8
3	92592972	92593391	Lce1j	-2.3	-0.5	1.8
3	139080721	139080827	U6.346	-2.3	-0.5	1.8
3	153452269	153455051	5730460C07Rik	-0.3	1.4	1.8
3	108164469	108165174	Gm12525	0.3	2.0	1.7
3	90330778	90332755	S100a14	0.0	1.4	1.4
3	100542375	100542636	7SK.81	-1.3	0.1	1.4
3	89031038	89031120	Mir92b	-2.5	-1.1	1.4
3	89771921	89772085	U1.55	-2.5	-1.1	1.4
3	102755872	102756036	U1.86	-2.5	-1.1	1.4
3	116169537	116169633	SNORA17.486	-2.5	-1.1	1.4
3	89049888	89053644	Krtcap2	1.7	3.1	1.4
3	93246252	93252999	Tchh	0.7	2.0	1.4
3	14766217	14778459	Car1	-0.3	1.0	1.3
3	38773492	38787181	C230034O21Rik	3.1	4.4	1.3
3	90412445	90415702	S100a5	-0.7	0.5	1.2
3	9402175	9402485	7SK.156	-1.3	-0.1	1.2
3	27052948	27055941	1700125G22Rik	-1.3	-0.1	1.2
3	126135801	126142187	RP23-60D3.2	1.5	2.7	1.2
3	30166747	30168869	Gm10258	-2.3	-1.1	1.2
3	148691896	148692002	U6.284	-2.3	-1.1	1.2
3	13471906	13474596	Gm10958	0.0	1.2	1.2
3	107356616	107357991	Ubl4b	0.0	1.1	1.1
3	36058444	36065142	Gm9760	1.3	2.3	1.1
3	89799378	89802608	Hax1	-1.3	-0.3	1.0
3	121040610	121040877	7SK.62	-1.3	-0.3	1.0
3	94762956	94765970	B230398E01Rik	0.0	1.0	1.0
4	43833207	43834305	Olfr156	-2.3	1.0	3.3
4	129985406	129985918	Gm10570	-2.3	0.9	3.2
4	62917885	62917966	Mir455	-1.3	1.7	3.0
4	120778256	120781831	Tmco2	-2.3	0.5	2.8
4	152592285	152592704	Gm13174	-2.3	0.4	2.7
4	34631128	34631434	7SK.169	-2.5	0.1	2.6
4	149940978	149942653	Gm13091	-1.3	1.1	2.4
4	155135301	155137669	2610204G22Rik	-1.3	1.0	2.3
4	45652808	45655176	Gm12408	0.0	2.3	2.3
4	94755774	94758110	4930551L18Rik	-1.3	0.9	2.2
4	9194470	9197062	AL844180.1	-2.3	-0.1	2.2
4	41586411	41588449	Gm12406	-2.3	-0.1	2.2
4	41763904	41812834	Gm13307	-2.3	-0.1	2.2
4	115452439	115453623	4930544O15Rik	-2.3	-0.1	2.2
4	123536922	123540408	4933427I04Rik	-2.3	-0.1	2.2
4	118626878	118627524	Gm12863	-2.5	-0.3	2.2
4	43864437	43868335	Olfr155	0.3	2.4	2.1
4	63017196	63020545	Orm3	-0.7	1.3	2.0
4	155573533	155574927	Isg15	-1.3	0.7	2.0
4	56908101	56908170	Mir32	-2.5	-0.5	2.0
4	83151051	83151458	Gm10154	-2.5	-0.5	2.0
4	133520653	133524565	Hmgn2	-2.5	-0.5	2.0
4	86204350	86205708	4930457A20Rik	-1.3	0.5	1.8
4	88400772	88401320	Gm13271	-2.3	-0.5	1.8
4	120445211	120445302	Mir30e	-2.3	-0.5	1.8
4	120756591	120756754	U1.157	-2.5	-0.8	1.7
4	114353409	114360200	Gm12829	1.6	3.3	1.7
4	118105506	118109957	Cdc20	1.0	2.6	1.6
4	20270853	20274965	Gm11872	-1.3	0.2	1.6
4	43782922	43783881	Olfr159	-1.3	0.2	1.6
4	101562625	101566394	C130073F10Rik	-0.3	1.2	1.6
4	71829802	71830483	Gm11250	-2.3	-0.8	1.6
4	34853373	34858730	2410114N07Rik	1.5	3.0	1.5
4	116309115	116324285	Akr1a4	1.4	2.8	1.4
4	149154133	149154823	Gm13073	0.5	1.9	1.4
4	116796582	116797771	Gm13014	-1.3	0.1	1.4
4	108690260	108690335	Mir761	-2.5	-1.1	1.4
4	130305549	130305621	Snord85	-2.5	-1.1	1.4
4	137771629	137772104	2310026L22Rik	-2.5	-1.1	1.4
4	44956625	44961202	Gm12678	0.5	1.9	1.4

4	120807262	120807580	7SK.2	0.0	1.4	1.4
4	116791824	116798271	Btbd19	0.3	1.6	1.4
4	141538907	141542712	Gm13037	2.5	3.9	1.3
4	153281495	153282775	Gm13115	0.0	1.3	1.3
4	155177645	155183885	Mrpl20	0.0	1.3	1.3
4	63257936	63266295	Gm11213	1.1	2.4	1.3
4	118656663	118669803	Gm12865	1.5	2.8	1.3
4	63023483	63026912	Orm2	0.7	2.0	1.3
4	141736517	141740054	Gm13062	1.9	3.2	1.3
4	122638552	122640718	4933421A08Rik	-0.7	0.5	1.2
4	147359897	147361314	Nppb	0.7	1.9	1.2
4	86500564	86503316	Rps6	-1.3	-0.1	1.2
4	151784909	151785371	Gm16333	-1.3	-0.1	1.2
4	88451435	88454284	Gm13288	-2.3	-1.1	1.2
4	136714997	136715786	Gm13003	0.5	1.7	1.2
4	141573813	141584911	4930455G09Rik	2.7	3.8	1.2
4	57184943	57189426	Gm12530	1.5	2.6	1.2
4	82527655	82531052	Cer1	1.4	2.5	1.2
4	123970936	123977091	Gm12916	0.7	1.8	1.1
4	25726376	25727293	Fut9	-0.7	0.4	1.1
4	33114117	33118298	4933421O10Rik	0.3	1.4	1.1
4	124760101	124762191	Gm12932	-0.7	0.4	1.1
4	88401099	88401494	Gm16686	0.7	1.8	1.1
4	98590501	98602234	Usp1	0.0	1.1	1.1
4	133683321	133684704	Sh3bgrl3	1.3	2.3	1.1
4	43672494	43682017	4930412F15Rik	2.2	3.3	1.1
4	39397328	39398809	1700009N14Rik	1.3	2.3	1.0
4	47021188	47023653	Gm568	1.3	2.3	1.0
4	114806437	114810469	Gm12835	-1.3	-0.3	1.0
4	135904465	135906249	9130020K20Rik	-1.3	-0.3	1.0
4	53575855	53577569	1700060J05Rik	1.8	2.8	1.0
4	133559245	133574756	Lin28a	2.1	3.1	1.0
4	32683935	32689735	Gja10	0.0	1.0	1.0
5	99938963	99939701	Gm16227	1.0	5.2	4.2
5	99646445	99681946	Rasgef1b	6.1	9.4	3.2
5	99726263	100158084	A930011G23Rik	8.6	11.4	2.8
5	115266797	115267341	Gm13820	-1.3	1.2	2.6
5	138644837	138648536	Gm454	-0.7	1.8	2.6
5	96290662	96291912	Gm5559	-2.3	0.1	2.4
5	31208258	31210161	Ost4	-1.3	0.9	2.2
5	139003068	139005922	1700123K08Rik	-1.3	0.7	2.0
5	27879091	27880415	Gm16058	-2.3	-0.3	2.0
5	125197448	125197545	U6.752	-2.3	-0.3	2.0
5	143916709	143916806	AC123686.1	-2.3	-0.3	2.0
5	113473164	113473272	U6.319	-2.5	-0.5	2.0
5	116265841	116265919	AC162461.1	-2.5	-0.5	2.0
5	139919553	139919659	U6.290	-2.5	-0.5	2.0
5	119747254	119747370	n-R5s175	-1.3	0.6	1.9
5	77085537	77087380	Gm10430	-0.3	1.4	1.8
5	115547698	115548304	Gm10401	0.0	1.7	1.7
5	123930162	123932594	Il31	1.0	2.7	1.7
5	92775664	92777915	Cxcl10	1.7	3.4	1.7
5	25004885	25007359	AC134910.1	-2.5	-0.8	1.7
5	138293375	138299277	Pilrb1	-2.5	-0.8	1.7
5	134680339	134680626	7SK.71	-0.3	1.4	1.7
5	115795651	115798990	Cox6a1	0.0	1.7	1.7
5	112171869	112177196	Gm16019	-0.3	1.3	1.6
5	92416501	92419330	Gm1045	-1.3	0.2	1.6
5	24694179	24694461	7SK.236	-2.3	-0.8	1.6
5	60892072	60892179	U6.637	-2.3	-0.8	1.6
5	120514344	120514452	n-R5s176	-2.3	-0.8	1.6
5	135390119	135391396	Cldn13	1.3	2.7	1.4
5	52765892	52766785	Gm447	-1.3	0.1	1.4
5	67376330	67376880	Gm15949	-1.3	0.1	1.4
5	122167675	122177412	Gm16552	1.0	2.4	1.4
5	86508791	86508960	U1.39	-2.5	-1.1	1.4
5	91322406	91322522	5S_rRNA.26	-2.5	-1.1	1.4
5	117082917	117083471	AC156984.1	-2.5	-1.1	1.4
5	138307057	138312986	Pilrb2	-2.5	-1.1	1.4
5	31252439	31257464	Abhd1	1.7	3.1	1.4
5	116762369	116768290	B230112J18Rik	1.7	3.0	1.4
5	73548253	73549119	1700071G01Rik	-0.7	0.6	1.4
5	116755691	116759369	Gm13838	1.7	3.0	1.3
5	31594628	31598250	Fndc4	1.8	3.2	1.3
5	23839741	23843034	Gm15589	0.7	1.9	1.2
5	110424530	110425417	AC123679.1	-1.3	-0.1	1.2
5	10236829	10237461	Gm10482	-2.3	-1.1	1.2
5	92731200	92731318	U6atac.21	-2.3	-1.1	1.2

5	139461462	139461950	AC117698.1	-2.3	-1.1	1.2
5	145803859	145806547	1700018F24Rik	-2.3	-1.1	1.2
5	106243792	106248007	Gm15817	2.3	3.5	1.2
5	31788232	31790849	4930548H24Rik	0.0	1.2	1.2
5	144780673	144787539	Ocm	2.1	3.2	1.1
5	73601348	73603873	RP23-318I19.1	2.7	3.8	1.1
5	130384410	130387150	4930579G22Rik	0.7	1.8	1.1
5	143664794	143668433	Actb	0.3	1.3	1.0
5	33773665	33773844	Nkx1-1	-1.3	-0.3	1.0
5	65527541	65527659	5S_rRNA.6	-1.3	-0.3	1.0
5	149818538	149824352	Gm15408	2.4	3.5	1.0
5	149632343	149640984	Gm15407	3.5	4.5	1.0
6	144618947	144621240	Gm10396	-1.3	1.7	3.0
6	27886655	27886750	Mir592	-2.5	0.2	2.7
6	42861524	42862456	Olfr447	-2.3	0.4	2.7
6	36371742	36371825	Mir490	-2.3	-0.1	2.2
6	42694038	42694991	Olfr453	-1.3	0.8	2.1
6	127268359	127270593	Gm10414	-1.3	0.7	2.0
6	116490473	116491456	Olfr213	-2.3	-0.3	2.0
6	42877232	42878158	Olfr446	-2.5	-0.5	2.0
6	125208067	125208195	SNORA70.40	-2.5	-0.5	2.0
6	89090999	89097347	Gm1965	0.3	2.2	1.9
6	89090999	89097379	Gm1965	0.3	2.2	1.9
6	136391504	136391634	AC124779.1	-1.3	0.6	1.9
6	85815416	85819152	Cml2	-2.3	-0.5	1.8
6	125071742	125071877	SCARNA11.2	-2.3	-0.5	1.8
6	125141819	125145601	Mrpl51	-2.3	-0.5	1.8
6	52127497	52130850	5730596B20Rik	1.0	2.7	1.7
6	99525885	99526131	7SK.32	-2.5	-0.8	1.7
6	121275598	121275694	AC153982.1	-2.5	-0.8	1.7
6	43117009	43118026	Olfr437	-1.3	0.4	1.7
6	89199462	89199596	5S_rRNA.9	-1.3	0.4	1.7
6	116531956	116533049	Olfr215	0.8	2.5	1.7
6	144672085	144673427	Gm7518	0.5	2.1	1.6
6	42846452	42847384	Olfr448	-0.7	0.8	1.6
6	70676429	70676751	Igk-C	-0.7	0.8	1.6
6	116506423	116507430	Olfr214	0.8	2.3	1.5
6	91962388	91966111	4930466I24Rik	1.3	2.7	1.5
6	119429686	119433904	Fbxl14	1.8	3.2	1.5
6	117163674	117164066	AC155646.1	1.0	2.4	1.4
6	30093801	30094127	7SK.7	0.8	2.3	1.4
6	116715547	116715670	SNORA17.412	-1.3	0.1	1.4
6	42436162	42437213	Olfr456	-2.5	-1.1	1.4
6	69942459	69942971	Gm16689	-2.5	-1.1	1.4
6	85710643	85715738	Cml3	-2.5	-1.1	1.4
6	98508733	98508825	AC158665.1	-2.5	-1.1	1.4
6	83028380	83030841	Pcgf1	1.0	2.4	1.4
6	86315677	86320051	Fam136a	0.8	2.2	1.4
6	51233531	51237854	Gm9953	1.4	2.7	1.3
6	52212769	52217597	Gm15053	0.8	2.2	1.3
6	40494035	40494976	Olfr461	-2.3	-1.1	1.2
6	68784840	68785301	Gm16905	-2.3	-1.1	1.2
6	71832551	71832687	SNORD94.1	-2.3	-1.1	1.2
6	86474493	86476165	Pcbp1	-2.3	-1.1	1.2
6	129588711	129593355	Klrc3	-2.3	-1.1	1.2
6	112646771	112654662	Gm16598	1.6	2.7	1.2
6	89161159	89166231	Gm839	0.8	2.0	1.2
6	42767472	42768404	Olfr450	-0.7	0.4	1.1
6	83621201	83627851	Cd207	2.0	3.1	1.1
6	48879895	48883685	1600015I10Rik	2.5	3.6	1.1
6	42350237	42351118	Tas2r143	0.0	1.1	1.1
6	83391749	83398316	B230319C09Rik	1.0	2.1	1.1
6	122576428	122580290	Dppa3	2.1	3.2	1.1
6	24509557	24512070	AC153635.1	-0.3	0.7	1.0
6	32461073	32465576	B230378P21Rik	0.7	1.7	1.0
6	124665250	124665520	snoU89.1	-0.3	0.7	1.0
7	111063576	111078425	Olfr631	-2.3	1.1	3.4
7	69606757	69609396	Peg12	-0.7	2.6	3.4
7	109818415	109819437	Olfr556	-2.5	0.2	2.7
7	111232557	111233689	Olfr645	-1.3	1.4	2.7
7	106883655	106884605	Olfr520	-2.3	0.2	2.6
7	110640429	110641393	Olfr609	-2.3	0.2	2.6
7	31484802	31485399	2200002J24Rik	-2.3	0.1	2.4
7	106628073	106628217	Snord15b	-2.3	0.1	2.4
7	111254795	111255733	Olfr646	-1.3	1.1	2.4
7	112204762	112205700	Olfr677	-2.5	-0.1	2.4
7	3218920	3219001	Mir291a	-2.3	-0.1	2.2
7	6432927	6433883	Olfr5	-2.3	-0.1	2.2

7	6296993	6300606	CAAA01104267.1.18771.1	-2.5	-0.3	2.2
7	25507751	25511082	Gm4763	-2.5	-0.3	2.2
7	15974297	16142195	Obox1	-2.3	-0.3	2.0
7	51480807	51484987	Klk1	-2.3	-0.3	2.0
7	112328827	112329914	Olfr685	-2.3	-0.3	2.0
7	112351873	112352890	Olfr686	-2.3	-0.3	2.0
7	110698882	110704018	Olfr613	-2.5	-0.5	2.0
7	130153626	130154138	AC125221.1	-2.5	-0.5	2.0
7	53967368	53971046	Saa3	-0.3	1.6	1.9
7	110654511	110655458	Olfr610	-0.7	1.2	1.9
7	109752870	109753823	Olfr552	-0.3	1.6	1.9
7	16418595	16425028	Obox6	-0.3	1.5	1.8
7	109788762	109789715	Olfr554	-0.3	1.5	1.8
7	52436552	52437208	Pth2	-0.7	1.1	1.8
7	111530579	111539100	Trim30c	-2.3	-0.5	1.8
7	111881956	111882876	Olfr664	-2.3	-0.5	1.8
7	147347838	147348913	Olfr522	-1.3	0.5	1.8
7	117771175	117773331	Adm	0.7	2.5	1.8
7	4985720	4986848	4632433K11Rik	-2.5	-0.8	1.7
7	29868655	29868713	Mir1963	-2.5	-0.8	1.7
7	51495118	51495405	7SK.186	-2.5	-0.8	1.7
7	80790871	80790975	U6.362	-2.5	-0.8	1.7
7	128850703	128851525	1700069B07Rik	-2.5	-0.8	1.7
7	51307660	51312075	Klk1b27	-1.3	0.4	1.7
7	52308983	52318047	Prrg2	1.1	2.8	1.7
7	110333317	110336055	Olfr592	-0.7	0.9	1.7
7	110599853	110600812	Olfr606	-0.3	1.3	1.6
7	147142904	147143938	Gm2044	-0.3	1.3	1.6
7	103106725	103110123	Gm9966	-0.7	0.8	1.6
7	112064828	112065808	Olfr667	-1.3	0.2	1.6
7	19996973	19997621	AC118017.2	-2.3	-0.8	1.6
7	35036057	35036512	Gm6096	-2.3	-0.8	1.6
7	137178608	137178717	U6.369	-2.3	-0.8	1.6
7	13533941	13547117	AC107704.2	2.6	4.2	1.5
7	7154833	7162506	Zfp772	0.0	1.5	1.5
7	20355960	20360752	Gm16733	1.6	3.0	1.5
7	112885106	112891439	Ilk	0.3	1.7	1.5
7	4916862	4922770	Gm1078	1.1	2.6	1.4
7	31455741	31456610	4930479M11Rik	-1.3	0.1	1.4
7	54007203	54009684	Saa2	-1.3	0.1	1.4
7	112955172	112959613	Mrpl17	-0.3	1.1	1.4
7	80953108	80961389	Gm10619	1.0	2.4	1.4
7	112516744	112517869	Olfr692	0.0	1.4	1.4
7	109690452	109691384	Olfr548-ps1	-0.7	0.6	1.4
7	147604741	147605670	Olfr532	-0.7	0.6	1.4
7	49700832	49703450	4933421I07Rik	-0.3	1.0	1.3
7	133551964	133558384	Cd19	1.4	2.6	1.2
7	148031464	148032667	Scgb1c1	1.4	2.6	1.2
7	106700779	106700873	Mir326	-1.3	-0.1	1.2
7	106915674	106916640	Olfr521	-0.3	0.9	1.2
7	109846730	109847786	Olfr557	-1.3	-0.1	1.2
7	5929764	5929844	AC026949.1	-2.3	-1.1	1.2
7	53995871	53998350	Saa1	-2.3	-1.1	1.2
7	93511886	93512845	Olfr305	-2.3	-1.1	1.2
7	111160159	111161214	Olfr639	-2.3	-1.1	1.2
7	112087042	112087995	Olfr669	-2.3	-1.1	1.2
7	112217983	112218924	Olfr678	-2.3	-1.1	1.2
7	114767485	114768675	5330417H12Rik	-2.3	-1.1	1.2
7	115209282	115210310	Olfr480	-2.3	-1.1	1.2
7	125628628	125629997	Itpripl2	-2.3	-1.1	1.2
7	19399286	19406956	Psg17	0.7	1.9	1.2
7	50906440	50915531	Siglece	3.1	4.3	1.2
7	110486526	110487575	Olfr599	0.0	1.2	1.2
7	130605808	130611517	Aqp8	1.4	2.5	1.2
7	30345534	30360497	4930432E11Rik	2.1	3.2	1.1
7	103960943	103965995	Gm15416	-0.3	0.8	1.1
7	119670486	119679197	A230071N21Rik	1.3	2.4	1.1
7	109736255	109737333	Olfr551	0.5	1.6	1.1
7	27716650	27724405	AC167659.1	-0.7	0.4	1.1
7	111095082	111096020	Olfr633	0.3	1.4	1.1
7	123834893	123834999	U6.670	-0.7	0.4	1.1
7	148153120	148155724	Ifitm1	0.3	1.4	1.1
7	134115234	134119612	D830044I16Rik	0.0	1.1	1.1
7	148657759	148661031	Polr2l	0.8	1.9	1.1
7	106626406	106632248	Rps3	1.6	2.6	1.1
7	3218627	3218709	Mir290	-1.3	-0.3	1.0
7	87065641	87067407	AC109221.2	-1.3	-0.3	1.0
7	111791394	111795819	Olfr658	-0.3	0.7	1.0

7	112108306	112109244	Olfr670	-1.3	-0.3	1.0
7	117201501	117204839	AA474408	-0.3	0.7	1.0
7	3583716	3596489	Gm15927	2.9	4.0	1.0
7	3581587	3594088	Prpf31	2.8	3.9	1.0
7	31853191	31861723	Fxyd3	2.9	3.9	1.0
7	134355122	134357812	Mylpf	1.6	2.6	1.0
8	110075144	110075213	Mir140	-2.5	0.8	3.3
8	123640071	123642792	Foxc2	0.3	3.0	2.8
8	106875211	106876739	Gm9853	-0.3	2.1	2.4
8	48804600	48804723	n-R5s96	-2.5	-0.1	2.4
8	74198856	74200573	Gm9933	-0.7	1.3	2.0
8	12356153	12356469	7SK.277	-2.3	-0.3	2.0
8	26789422	26789743	7SK.147	0.0	2.0	2.0
8	96702989	96704225	Mt1	0.0	2.0	2.0
8	99036240	99037431	1700047G07Rik	-2.5	-0.5	2.0
8	124354336	124354442	U6.507	-2.5	-0.5	2.0
8	47292151	47296638	Slc25a4	1.7	3.6	1.9
8	37088807	37089190	Gm10063	-1.3	0.5	1.8
8	3621575	3624227	2900053A13Rik	0.0	1.8	1.8
8	97523005	97523118	SNORA7.1	-2.5	-0.8	1.7
8	41908363	41912215	Adam39	-1.3	0.4	1.7
8	47377404	47380025	Helt	-1.3	0.2	1.6
8	106845467	106845687	Gm11020	-1.3	0.2	1.6
8	108139366	108139473	Mir1966	-1.3	0.2	1.6
8	44285176	44285744	AC142104.1	-2.3	-0.8	1.6
8	4253102	4256219	Snapc2	0.5	1.9	1.4
8	74411163	74414811	Zfp709	-1.3	0.1	1.4
8	4115513	4115765	7SK.54	-2.5	-1.1	1.4
8	33587355	33587801	Gm5908	-2.5	-1.1	1.4
8	74623661	74624605	Olfr373	-2.5	-1.1	1.4
8	96903137	96910872	Gm15890	2.8	4.1	1.3
8	109553575	109555088	Gm16208	0.0	1.3	1.3
8	111772617	111774008	4922502B01Rik	0.0	1.3	1.3
8	96661104	96662931	Mt4	1.6	2.9	1.3
8	11477929	11480241	AC124475.1	1.7	3.0	1.3
8	12672186	12679229	4931415C17Rik	1.6	2.8	1.2
8	26742422	26742551	SNORA70.12	-0.7	0.5	1.2
8	73988825	73989651	Mrpl34	-0.7	0.5	1.2
8	13896311	13896631	Gm7676	-1.3	-0.1	1.2
8	87574106	87574175	SNORD41.1	-1.3	-0.1	1.2
8	22812991	22814117	AY761184	-2.3	-1.1	1.2
8	90698459	90698927	9430002A10Rik	-2.3	-1.1	1.2
8	109963051	109963463	Rps26-ps1	-2.3	-1.1	1.2
8	13405081	13421945	Grk1	4.8	6.0	1.2
8	35112316	35120100	Gm9951	2.8	4.0	1.2
8	48794698	48799285	Cdkn2aip	1.8	2.9	1.2
8	63462234	63464705	B230317F23Rik	-0.3	0.8	1.1
8	12706944	12719127	Gm15348	3.6	4.7	1.1
8	67505550	67511545	Trim75	2.5	3.6	1.1
8	35178084	35188239	Mboat4	3.6	4.7	1.1
8	97846914	97849163	Zfp319	0.5	1.6	1.1
8	42125110	42127407	Gm16192	1.5	2.6	1.1
8	86064992	86066087	Olfr370	2.2	3.3	1.1
8	112514337	112523703	Tat	4.0	5.1	1.1
8	10924427	10928457	3930402G23Rik	2.0	3.0	1.0
8	11838978	11840400	Gm15875	0.3	1.3	1.0
8	3655770	3660110	Retn	1.5	2.5	1.0
8	22083372	22084427	AY761185	-1.3	-0.3	1.0
8	25548088	25549418	1810011O10Rik	-1.3	-0.3	1.0
8	3665762	3668905	1810033B17Rik	2.0	3.0	1.0
8	12278819	12280657	A230072I06Rik	0.0	1.0	1.0
9	118833991	118834895	Gm16295	-2.3	0.8	3.1
9	39757422	39758354	Olfr975	-1.3	1.1	2.4
9	57965622	57966796	Gm16131	-2.3	0.1	2.4
9	88979462	88987673	Gm11058	-2.3	0.1	2.4
9	58993481	58994017	Gm7589	-2.3	-0.1	2.2
9	120038469	120038671	SNORA73.1	-2.5	-0.3	2.2
9	44981080	44983244	RP23-284E19.7	-0.3	1.8	2.1
9	110907315	110909334	Gm10030	-0.7	1.2	2.0
9	20611478	20611798	7SK.80	-2.5	-0.5	2.0
9	39972215	39972544	Gm4893	-2.5	-0.5	2.0
9	107576808	107581923	Gnat1	2.7	4.6	1.9
9	28005472	28007106	Gm15606	-1.3	0.5	1.8
9	48855737	48856390	AC160137.1	-0.7	1.1	1.8
9	86597857	86597984	SNORA17.577	-2.3	-0.5	1.8
9	108888208	108889209	Ucn2	-2.3	-0.5	1.8
9	109173350	109190190	Fbxw14	-2.3	-0.5	1.8
9	111516402	111516519	n-R5s90	-2.3	-0.5	1.8

9	19122980	19123909	Olfr844	-2.5	-0.8	1.7
9	38635200	38636267	Olfr923	-2.5	-0.8	1.7
9	39526814	39527752	Olfr965	-2.5	-0.8	1.7
9	78169106	78170164	AC138587.2	-2.5	-0.8	1.7
9	81524950	81526222	Htr1b	1.9	3.6	1.7
9	69324836	69327405	Gm16144	1.3	2.9	1.7
9	123891782	123893241	Ccr1l1	-0.7	0.9	1.7
9	106783260	106784620	Rbm15b	-2.3	-0.8	1.6
9	30888202	30890060	CT025653.1	0.5	2.0	1.6
9	107886557	107888247	6230427J02Rik	1.7	3.2	1.5
9	107465586	107468439	Tusc2	1.7	3.1	1.5
9	39908129	39909073	Olfr984	0.0	1.4	1.4
9	119068279	119070787	Gm10608	0.0	1.4	1.4
9	78159779	78162604	Omt2a	-1.3	0.1	1.4
9	89989016	89990740	AC140392.1	-1.3	0.1	1.4
9	53683583	53683959	Gm11147	-2.5	-1.1	1.4
9	110887449	110892217	Rtp3	0.7	2.0	1.4
9	81525158	81538763	D430036J16Rik	2.4	3.7	1.3
9	109730457	109735220	3000002C10Rik	-0.7	0.5	1.2
9	111147096	111147371	7SK.260	-0.7	0.5	1.2
9	21760197	21762714	2310047B19Rik	-1.3	-0.1	1.2
9	106056455	106056544	Mir135a-1	-1.3	-0.1	1.2
9	39807791	39808835	Olfr979	-2.3	-1.1	1.2
9	53448082	53448210	SNORA61.1	-2.3	-1.1	1.2
9	60267670	60267792	SNORA17.443	-2.3	-1.1	1.2
9	64023230	64023328	Snord16a	-2.3	-1.1	1.2
9	116134326	116134401	AC131777.1	-2.3	-1.1	1.2
9	123544506	123546443	RP24-298N7.5	-2.3	-1.1	1.2
9	124016689	124028296	Ccr2	-2.3	-1.1	1.2
9	56454500	56462278	Gm16960	3.3	4.5	1.2
9	69879303	69885895	Gcnt3	1.9	3.1	1.2
9	24429629	24430271	Cypt4	-0.3	0.8	1.1
9	20695201	20695280	snoU105B.1	-0.7	0.4	1.1
9	112085316	112088238	AC100550.1	-0.7	0.4	1.1
9	110842114	110848662	Tdgf1	2.4	3.5	1.1
9	109334833	109351654	Fbxw16	1.5	2.6	1.1
9	44777655	44788514	Cd3g	1.8	2.8	1.1
9	58336770	58337735	CT030640.1	-1.3	-0.3	1.0
9	71908720	71909406	Rpl15-ps2	-1.3	-0.3	1.0
9	36339988	36349548	Gm5615	0.0	1.0	1.0
9	56995987	56999744	Gm16493	0.0	1.0	1.0
9	32122368	32151822	Kcnj5	4.1	5.1	1.0
10	117099338	117099444	U6.213	-2.3	0.9	3.2
10	117641998	117647098	Il22	0.5	3.7	3.2
10	116724852	116729924	Lyz1	1.4	4.4	3.0
10	117170735	117183414	Slc35e3	5.2	8.2	3.0
10	117726685	117732094	AC158600.1	0.3	3.2	3.0
10	117578843	117606053	Mdm1	6.3	9.2	2.9
10	117066556	117124408	Cpm	7.2	10.1	2.9
10	117878103	117882948	Ifng	2.6	5.4	2.8
10	117187680	117229761	Nup107	7.4	10.0	2.6
10	116669293	116675759	9530003J23Rik	4.1	6.6	2.6
10	22057446	22057528	AC159330.1	-2.3	0.2	2.6
10	127100895	127102540	Gm16229	0.0	2.5	2.5
10	116683788	116685202	Gm10936	-0.3	2.1	2.4
10	117251527	117282991	Rap1b	5.7	8.1	2.4
10	93378914	93379081	SNORA73.10	-2.5	-0.1	2.4
10	117125944	117147814	Mdm2	5.3	7.7	2.4
10	116652198	116661563	Yeats4	5.9	8.2	2.4
10	116714390	116719377	Lyz2	1.6	3.9	2.3
10	116781724	116814029	Cpsf6	5.1	7.0	1.9
10	77832758	77833065	7SK.19	-2.5	-0.8	1.7
10	85270560	85271990	AC122408.1	-1.3	0.2	1.6
10	30662185	30662291	U6.221	-2.3	-0.8	1.6
10	95135625	95136085	Gm6843	-2.3	-0.8	1.6
10	77888415	77890364	Cstb	-0.7	0.5	1.2
10	116407374	116409665	Gm10271	0.0	1.2	1.2
10	69746983	69747701	Gm7097	1.7	2.7	1.1
11	102977639	102981039	Hexim1	-1.3	1.8	3.1
11	87702461	87703396	Olfr462	-2.3	0.6	2.9
11	94638531	94644017	Tmem92-ps	-2.3	0.6	2.9
11	97022304	97022989	Gm11583	-2.5	0.4	2.8
11	20731484	20731922	Gm12037	-2.3	0.5	2.8
11	72153872	72154142	7SK.224	-2.3	0.5	2.8
11	12548080	12548279	AL645934.1	-2.5	0.2	2.7
11	83476086	83478185	Ccl4	-2.3	0.4	2.7
11	107855336	107855945	Gm11657	-2.3	0.4	2.7
11	87911560	87913610	2210416O15Rik	-2.3	0.2	2.6

11	75275041	75280178	2210403K04Rik	0.0	2.3	2.3
11	98485665	98486975	Gm12355	-2.3	-0.1	2.2
11	58141223	58143841	Gm9900	-2.3	-0.3	2.0
11	73201583	73205716	Olfr377-ps1	-2.3	-0.3	2.0
11	120326180	120328921	Arl16	-0.7	1.2	2.0
11	101139282	101140428	Ccdc56	-1.3	0.6	1.9
11	115277003	115281276	Atp5h	-0.3	1.5	1.8
11	53514681	53514990	7SK.11	-2.3	-0.5	1.8
11	69208989	69210182	Lsmd1	-2.3	-0.5	1.8
11	82577847	82581124	Zfp830	-2.3	-0.5	1.8
11	96038865	96040114	Gm11535	-1.3	0.5	1.8
11	109298486	109299459	Gm15642	-2.3	-0.5	1.8
11	35163571	35164287	Gm12122	-2.5	-0.8	1.7
11	49106799	49107816	Olfr1392	-2.5	-0.8	1.7
11	99475331	99476160	Krtap2-4	-2.5	-0.8	1.7
11	97702094	97705413	B230217C12Rik	0.0	1.6	1.6
11	4185463	4186942	Gm11958	-1.3	0.2	1.6
11	69476545	69476627	Mir1934	-1.3	0.2	1.6
11	97141442	97144211	Gm11592	-1.3	0.2	1.6
11	106047758	106047901	U1.180	-2.3	-0.8	1.6
11	116535107	116535178	snoR38.2	-2.3	-0.8	1.6
11	70327018	70328118	Vmo1	1.8	3.3	1.5
11	58259651	58266425	4930504O13Rik	1.7	3.2	1.5
11	48973518	48974456	Olfr1394	0.3	1.8	1.5
11	97702607	97705413	B230217C12Rik	-0.3	1.2	1.5
11	49154036	49154971	Olfr1390	-0.7	0.7	1.5
11	70465352	70468146	Pfn1	-0.7	0.7	1.5
11	119081474	119083787	Gm11752	1.3	2.7	1.5
11	16851378	16854775	Fbxo48	-1.3	0.1	1.4
11	53235683	53236672	Leap2	-1.3	0.1	1.4
11	73934793	73935740	Olfr401	-1.3	0.1	1.4
11	113064629	113070983	4732490B19Rik	2.2	3.6	1.4
11	83658167	83658474	7SK.60	-2.5	-1.1	1.4
11	87507984	87508083	U6.596	-2.5	-1.1	1.4
11	96711707	96711779	Mir152	-2.5	-1.1	1.4
11	120207004	120209856	Actg1	-2.5	-1.1	1.4
11	54060773	54063169	Csf2	0.5	1.8	1.3
11	101471928	101477270	Gm11635	1.9	3.3	1.3
11	58234613	58239526	1810065E05Rik	1.6	2.9	1.3
11	44399819	44404731	4930597A21Rik	0.7	1.9	1.2
11	49257355	49258290	Olfr1388	0.3	1.5	1.2
11	96064768	96066699	Gm11538	-0.7	0.5	1.2
11	99248654	99249211	Gm11940	-0.7	0.5	1.2
11	115114592	115114699	AL603828.1	-0.7	0.5	1.2
11	67594445	67598816	Gm12302	-1.3	-0.1	1.2
11	3434241	3434910	Gm11947	-2.3	-1.1	1.2
11	33595894	33597908	Gm12119	-2.3	-1.1	1.2
11	50324404	50324679	7SK.300	-2.3	-1.1	1.2
11	51938032	51941281	AL935177.1	0.5	1.7	1.2
11	65439896	65440005	U6.227	-2.3	-1.1	1.2
11	71836257	71836359	U6atac.16	-2.3	-1.1	1.2
11	75236233	75236842	Gm12336	-2.3	-1.1	1.2
11	98807758	98807864	U6.37	-2.3	-1.1	1.2
11	117275703	117276682	Gm11729	-2.3	-1.1	1.2
11	120957955	120958521	Gm11791	1.0	2.2	1.2
11	58246061	58255124	Gm12253	0.8	2.0	1.2
11	110927478	110938135	Kcnj2	2.2	3.4	1.2
11	97945771	97946738	Gm11632	0.5	1.7	1.2
11	69413238	69419444	Atp1b2	2.8	4.0	1.2
11	84642680	84644370	4930502E09Rik	0.0	1.2	1.2
11	58303221	58304255	Olfr332	1.0	2.2	1.2
11	73797164	73803561	Olfr398	1.8	2.9	1.2
11	100107405	100110276	Krt16	-0.3	0.8	1.1
11	120459735	120469512	Anapc11	2.5	3.7	1.1
11	69943831	69951052	Mgl2	1.8	2.9	1.1
11	84331479	84338923	Lhx1	3.7	4.8	1.1
11	121282715	121292602	Fn3krp	1.9	3.0	1.1
11	58599641	58600771	Olfr314	2.1	3.2	1.1
11	96055630	96058761	Hoxb13	1.6	2.7	1.1
11	85660292	85663835	Gm11444	2.5	3.6	1.1
11	6531490	6538652	Wap	2.4	3.5	1.1
11	101902823	101916934	Cd300lg	4.3	5.3	1.1
11	67611689	67629507	Dhrs7c	3.3	4.4	1.1
11	51498888	51502376	0610009B22Rik	1.6	2.6	1.1
11	87514898	87518833	1110028F11Rik	1.6	2.6	1.1
11	107271704	107274863	Gm11713	1.6	2.6	1.1
11	101848370	101853570	1700006E09Rik	1.3	2.3	1.0
11	23458058	23458842	Gm12060	-1.3	-0.3	1.0

11	49243929	49245001	Olfr1389	-0.3	0.7	1.0
11	49477112	49478620	Scgb3a1	-0.3	0.7	1.0
11	81915347	81916902	Ccl12	-1.3	-0.3	1.0
11	83560442	83566144	1100001G20Rik	-0.3	0.7	1.0
11	87570366	87570429	Mir142	-0.3	0.7	1.0
11	69496348	69499286	Tnfsf13	1.4	2.4	1.0
11	68735337	68740583	Odf4	1.1	2.2	1.0
11	83251188	83254734	1700020L24Rik	1.1	2.2	1.0
11	33863013	33873641	Kcnmb1	2.4	3.5	1.0
11	5815318	5816734	Gm11967	0.5	1.5	1.0
11	35894166	35897356	Gm12126	1.0	2.0	1.0
11	69435526	69437372	Sat2	1.0	2.0	1.0
11	97884269	97888059	Gm11629	0.0	1.0	1.0
12	110948385	110948480	Mir411	-2.3	1.1	3.4
12	34613290	34614175	Ferd3l	-1.3	1.7	3.0
12	87059427	87059545	n-R5s64	-2.5	-0.1	2.4
12	110950527	110950646	Mir1197	-1.3	0.9	2.2
12	110965032	110965106	Mir381	-2.3	-0.1	2.2
12	28025562	28027574	Sox11	0.8	2.7	1.9
12	70258709	70260173	Rps29	-2.3	-0.5	1.8
12	104966366	104976368	Serpina1b	-1.3	0.5	1.8
12	110956957	110957036	Mir495	-2.3	-0.5	1.8
12	9581248	9588305	Osr1	2.9	4.7	1.8
12	110844667	110844743	SNORD112.1	-2.5	-0.8	1.7
12	33039750	33040067	7SK.8	-0.3	1.4	1.7
12	89591799	89592682	CT485612.1	-0.7	0.9	1.7
12	77893322	77896541	Gpx2	-0.3	1.3	1.6
12	33063569	33067808	RP24-143P3.1	-2.3	-0.8	1.6
12	110977329	110977407	Mir496	-2.3	-0.8	1.6
12	110978720	110978787	Mir377	-2.3	-0.8	1.6
12	114131774	114133577	AC160929.2	1.0	2.5	1.5
12	102383824	102387492	AC122507.1	0.3	1.8	1.5
12	110981925	110982005	Mir410	-0.7	0.7	1.5
12	75416096	75417381	Gm11042	-1.3	0.1	1.4
12	87818328	87818415	AC132189.1	-2.5	-1.1	1.4
12	114812453	114812892	AC090887.2	-2.5	-1.1	1.4
12	11228267	11229016	1700034J04Rik	0.5	1.8	1.3
12	70285145	70287759	Mgat2	-0.3	0.9	1.2
12	105366206	105377755	Serpina3b	0.0	1.2	1.2
12	105501332	105507582	Serpina3i	-1.3	-0.1	1.2
12	110832716	110832809	Mir434	-1.3	-0.1	1.2
12	58733565	58734002	Rpl26-ps2	-2.3	-1.1	1.2
12	105091799	105104137	Serpina1a	-2.3	-1.1	1.2
12	106414568	106419353	Tcl1b5	-2.3	-1.1	1.2
12	114184574	114188723	AC160929.1	2.3	3.4	1.1
12	82046491	82051494	Srsf5	1.9	3.0	1.1
12	108192022	108194138	Gm16087	1.3	2.3	1.0
12	88221653	88225764	6430527G18Rik	-1.3	-0.3	1.0
12	106269285	106269559	Scarna13	-0.3	0.7	1.0
12	110947270	110947335	Mir379	-1.3	-0.3	1.0
12	81787357	81793453	0610009B14Rik	2.3	3.4	1.0
12	104645032	104646030	9330161L09Rik	1.1	2.2	1.0
12	114509391	114511457	Gm900	1.0	2.0	1.0
12	114509484	114511358	Gm900	1.0	2.0	1.0
13	63478069	63482353	Gm16132	-2.5	0.1	2.6
13	99754045	99754520	1700024P04Rik	-2.3	-0.3	2.0
13	96100555	96100692	Snora47	-1.3	0.6	1.9
13	78328234	78339237	Nr2f1	3.0	4.9	1.9
13	34218877	34222223	Tubb2b	-0.7	1.1	1.8
13	60630657	60630770	AC123619.1	-2.3	-0.5	1.8
13	112383634	112383833	Gm10736	-2.3	-0.5	1.8
13	93344586	93346187	Spz1	1.5	3.3	1.8
13	14179706	14179812	U6.86	-2.5	-0.8	1.7
13	49809488	49809609	SNORA24.6	-2.5	-0.8	1.7
13	25029270	25033286	9330162012Rik	1.3	2.9	1.7
13	23519844	23522891	C230035I16Rik	-0.7	0.8	1.6
13	74145600	74146802	Gm15912	-1.3	0.2	1.6
13	84544444	84545077	AC155294.1	-1.3	0.2	1.6
13	4613426	4613908	Rpl29-ps2	-2.3	-0.8	1.6
13	24578531	24578633	U6.899	-2.3	-0.8	1.6
13	53621072	53621428	Gm5449	-2.3	-0.8	1.6
13	9622626	9622873	7SK.175	-1.3	0.1	1.4
13	27404496	27411121	Prl6a1	-1.3	0.1	1.4
13	49588687	49591945	Tes3-ps	-1.3	0.1	1.4
13	22106337	22107239	Gm11290	0.3	1.6	1.3
13	41896152	41897164	1700061E18Rik	-1.3	-0.1	1.2
13	54500453	54500695	Gm16248	0.0	1.2	1.2
13	67001557	67006286	Uqcrb	-1.3	-0.1	1.2

13	67461607	67462325	RP24-242F4.3	-1.3	-0.1	1.2
13	12681832	12682058	SNORA73.5	-2.3	-1.1	1.2
13	24428139	24429212	Gm11342	-2.3	-1.1	1.2
13	31066073	31067845	Gm5447	-2.3	-1.1	1.2
13	62441548	62441628	CT573034.2	-2.3	-1.1	1.2
13	30032152	30034130	Gm11365	2.1	3.3	1.2
13	23898862	23917049	Slc17a2	2.6	3.8	1.2
13	23370219	23374136	4933404K08Rik	0.5	1.7	1.2
13	55028124	55031566	4930526F13Rik	0.5	1.7	1.2
13	112045221	112047957	Actbl2	2.3	3.5	1.2
13	94395337	94398588	Gm15622	0.3	1.4	1.2
13	48367785	48369253	A330048O09Rik	-0.3	0.8	1.1
13	33341481	33352715	Serpinb9e	-0.7	0.4	1.1
13	62318904	62319329	Gm10260	-0.7	0.4	1.1
13	67005850	67010102	Gm10767	-1.3	-0.3	1.0
13	22118568	22124065	Gm11292	1.4	2.4	1.0
13	101185514	101185929	Gm9858	0.5	1.5	1.0
13	29034917	29043540	AL606528.1	1.6	2.6	1.0
14	54464909	54465501	Gm13893	-2.5	0.7	3.2
14	96280485	96281994	4921530L21Rik	-1.3	1.8	3.1
14	56636801	56639228	Mcpt1	-2.5	0.5	3.0
14	55567897	55567979	Mir208a	-2.3	0.4	2.7
14	23377304	23377675	Gm5670	-2.3	0.2	2.6
14	62226319	62230753	Kcnrg	-2.3	0.2	2.6
14	78710956	78711107	Mir2145-2	-2.5	0.1	2.6
14	121148574	121149193	AC125405.1	-2.5	-0.1	2.4
14	32350143	32350444	7SK.278	-2.3	-0.1	2.2
14	42660348	42904201	Gm7995	-1.3	0.8	2.1
14	27813548	27815492	Hesx1	-0.7	1.3	2.0
14	44534734	44540745	Ang5	-2.5	-0.5	2.0
14	53029564	53030602	Olfr1510	-2.5	-0.5	2.0
14	54450780	54451501	Gm13894	-2.5	-0.5	2.0
14	54837511	54837576	Tcra-J	-2.5	-0.5	2.0
14	54741901	54741916	Gm16622	-2.3	-0.5	1.8
14	54781665	54781718	CT030634.5	-2.3	-0.5	1.8
14	56775419	56778416	Gzmg	-0.3	1.4	1.8
14	54768337	54768876	Gm6683	0.3	2.0	1.7
14	67556202	67557573	4930578I07Rik	0.3	2.0	1.7
14	49272174	49272277	U6.587	-2.5	-0.8	1.7
14	51669126	51669838	Rnase11	-2.5	-0.8	1.7
14	51734240	51737466	Eddm3b	-2.5	-0.8	1.7
14	54799637	54799699	Gm16669	-2.5	-0.8	1.7
14	54811336	54811400	Gm16820	-2.5	-0.8	1.7
14	70828303	70828404	U11.5	-2.5	-0.8	1.7
14	105676096	105676203	U6.111	-2.5	-0.8	1.7
14	56269019	56271376	Tssk4	1.6	3.2	1.6
14	52991785	52992726	Olfr1512	0.3	1.9	1.6
14	54495659	54496426	Gm13892	-0.7	0.8	1.6
14	53009490	53010446	Olfr1511	-2.3	-0.8	1.6
14	54794417	54794473	Gm16925	-2.3	-0.8	1.6
14	51874941	51875787	Ear11	1.0	2.5	1.5
14	12060066	12060506	Rpl21-ps4	-1.3	0.1	1.4
14	51815161	51815598	Ang2	-1.3	0.1	1.4
14	63616939	63619903	Defb47	1.0	2.4	1.4
14	44579329	44584083	Ang6	-2.5	-1.1	1.4
14	54781114	54781176	CT030634.4	-2.5	-1.1	1.4
14	27208685	27215007	Gm16241	1.5	2.8	1.4
14	55840844	55860272	Dhrs2	3.1	4.5	1.4
14	41945071	41949741	Sftpa1	1.0	2.3	1.3
14	29321260	29324578	CT025649.1	2.8	4.1	1.3
14	53070068	53071077	Olfr1509	0.0	1.3	1.3
14	56206278	56209950	Psme2	0.7	2.0	1.3
14	47380216	47391200	Cdkn3	2.1	3.3	1.3
14	21522249	21523639	Chchd1	-0.3	0.9	1.2
14	52968778	52969719	Olfr1513	-0.3	0.9	1.2
14	56039106	56039561	Gm8894	-1.3	-0.1	1.2
14	56678581	56681147	Mcpt4	-1.3	-0.1	1.2
14	109091411	109091860	Rpl27a-ps2	-1.3	-0.1	1.2
14	54832363	54832424	Gm16920	-2.3	-1.1	1.2
14	115443527	115443613	Mir19b-1	-2.3	-1.1	1.2
14	35183460	35187975	Sncg	2.9	4.1	1.2
14	9498904	9505892	4930455B14Rik	1.5	2.7	1.2
14	60705908	60714930	4930563I02Rik	2.1	3.3	1.2
14	45607870	45623495	Ptger2	2.8	4.0	1.2
14	20894868	20902394	1810063B07Rik	2.3	3.4	1.1
14	49277354	49287319	Otx2	2.3	3.4	1.1
14	37891036	37911497	Cdhr1	3.3	4.4	1.1
14	25728900	25732926	Gm10398	-0.3	0.7	1.0

14	33106110	33108160	RP24-84D8.2	-1.3	-0.3	1.0
14	115442893	115442976	Mir17	-0.3	0.7	1.0
14	75536077	75536719	Gm15628	1.0	2.0	1.0
14	9046662	9056081	Kctd6	3.3	4.3	1.0
14	49922981	49933884	Slc35f4	1.5	2.5	1.0
14	56295455	56300654	Tinf2	2.5	3.5	1.0
14	18717322	18717647	7SK.70	1.0	2.0	1.0
14	51764677	51766442	Rnase1	0.0	1.0	1.0
14	55657299	55661816	Gm10364	0.0	1.0	1.0
14	109309208	109313456	Slitrk1	2.0	3.0	1.0
14	115443073	115443168	Mir18	0.0	1.0	1.0
15	98143020	98143979	Olfr285	-2.3	1.4	3.7
15	79501291	79503292	Tomm22	-2.3	0.5	2.8
15	58678133	58678876	Gm5959	-0.3	2.2	2.5
15	84947259	84947720	Gm10923	-2.5	-0.1	2.4
15	97900183	97900311	SNORA17.406	-2.3	-0.1	2.2
15	38172915	38174731	AC122459.2	0.5	2.6	2.1
15	74643884	74649245	Gml	-1.3	0.7	2.0
15	100446069	100451184	Dazap2	-0.3	1.7	2.0
15	60650551	60656635	Fam84b	1.8	3.8	2.0
15	60654520	60663348	9930014A18Rik	1.7	3.6	1.9
15	98057106	98058074	Olfr286	-0.7	1.1	1.8
15	75866889	75867046	U1.201	-2.5	-0.8	1.7
15	98267901	98268828	Olfr282	2.1	3.8	1.7
15	33334727	33335731	1700084J12Rik	0.0	1.7	1.7
15	31123960	31124093	SNORA17.166	-2.3	-0.8	1.6
15	74875447	74879260	Ly6c1	-2.3	-0.8	1.6
15	77452631	77456620	AL592187.1	-2.3	-0.8	1.6
15	97765716	97766026	7SK.276	-2.3	-0.8	1.6
15	37172158	37172917	Gm16137	1.4	2.9	1.5
15	99556049	99558567	2310016M24Rik	1.0	2.5	1.5
15	98708209	98711835	Rhebl1	2.4	3.8	1.5
15	78742396	78744090	1700027A07Rik	1.4	2.8	1.4
15	79908211	79913836	Rpl3	0.3	1.7	1.4
15	74987123	74989169	Ly6g	-2.5	-1.1	1.4
15	79522932	79523055	n-R5s40	-2.5	-1.1	1.4
15	74573269	74577117	2300005B03Rik	0.7	2.0	1.4
15	54143516	54143920	Gm5215	-0.7	0.6	1.4
15	58765365	58771044	Ndufb9	2.5	3.8	1.3
15	95654890	95660503	Gm10970	2.8	4.0	1.2
15	85980014	85980281	7SK.172	-0.7	0.5	1.2
15	59042458	59043338	Gm5045	-1.3	-0.1	1.2
15	74825308	74828064	Ly6a	-1.3	-0.1	1.2
15	76551896	76554275	C030006K11Rik	0.8	2.1	1.2
15	36135015	36135077	U6.662	-2.3	-1.1	1.2
15	73559594	73559720	U4.20	-2.3	-1.1	1.2
15	76778969	76781159	1110038F14Rik	1.0	2.2	1.2
15	86016158	86017582	Gm15569	-2.3	-1.1	1.2
15	94817438	94817569	SNORA17.154	-2.3	-1.1	1.2
15	98761856	98764996	Tuba1b	-2.3	-1.1	1.2
15	99198460	99198566	U6.114	-2.3	-1.1	1.2
15	85645106	85652163	Pkdrej	4.2	5.5	1.2
15	100504856	100518351	Cela1	4.6	5.7	1.2
15	89350527	89354108	Gm15609	2.4	3.6	1.2
15	63689826	63700852	Gsdmc3	-0.3	0.8	1.1
15	80091044	80094736	Rps19bp1	1.7	2.8	1.1
15	82385538	82389843	Cyp2d12	-0.3	0.8	1.1
15	98816396	98819712	4930578M01Rik	1.3	2.4	1.1
15	75447109	75455730	Zfp41	3.5	4.6	1.1
15	31520555	31531602	Cct5	3.4	4.5	1.1
15	102493299	102501478	Atp5g2	1.8	2.8	1.1
15	31370965	31383444	Ropn1l	3.7	4.7	1.1
15	95882047	95909984	D030018L15Rik	5.3	6.3	1.1
15	98310831	98313114	Lalba	1.1	2.2	1.1
15	102953501	102977259	D930007P13Rik	3.3	4.4	1.0
15	76168038	76168835	Rps6-ps1	-1.3	-0.3	1.0
15	82620537	82624675	Cyp2d26	3.1	4.2	1.0
15	82219584	82224452	Cyp2d11	0.5	1.5	1.0
15	99005281	99005368	7SK.118	0.0	1.0	1.0
16	96367083	96369191	Gm15317	-2.5	0.4	2.8
16	10783901	10785629	Socs1	-0.7	2.0	2.7
16	31427962	31428377	Gm15743	-2.3	0.4	2.7
16	17139768	17140576	Gm15646	-1.3	1.1	2.4
16	17711761	17712198	Rpl26-ps4	-2.5	-0.1	2.4
16	17980514	17981366	Vpreb2	-2.3	-0.1	2.2
16	38442943	38443349	CT571259.2	-2.3	-0.1	2.2
16	77599181	77599245	Mir99a	-2.3	-0.1	2.2
16	88877758	88878283	Krtap16-5	-2.3	-0.1	2.2

16	23423495	23423808	7SK.13	-2.5	-0.3	2.2
16	59035606	59039575	Olfr187	-2.5	-0.3	2.2
16	4229230	4231206	Gm5766	-1.3	0.8	2.1
16	13453933	13454019	Mir365-1	-2.3	-0.3	2.0
16	8449591	8449879	7SK.16	-2.5	-0.5	2.0
16	81261948	81262260	7SK.61	-1.3	0.6	1.9
16	35234882	35235009	U4.25	-2.3	-0.5	1.8
16	5324129	5324250	n-R5s30	-2.5	-0.8	1.7
16	8996178	8996459	AC165274.1	-1.3	0.4	1.7
16	3838975	3839913	Olfr15	1.1	2.7	1.6
16	18355709	18355856	U1.190	-1.3	0.2	1.6
16	30064470	30067882	Hes1	-1.3	0.2	1.6
16	13531521	13531644	SNORA17.179	-2.3	-0.8	1.6
16	14316167	14317378	Gm15869	-2.3	-0.8	1.6
16	32108951	32113829	Gm15729	1.1	2.7	1.6
16	32186278	32186976	Bex6	-2.3	-0.8	1.6
16	16868495	16870936	Vpreb1	-0.7	0.7	1.5
16	36359468	36367834	BC100530	-1.3	0.1	1.4
16	36367156	36387492	Gm15845	-1.3	0.1	1.4
16	90554409	90555129	1110008E08Rik	-1.3	0.1	1.4
16	18240810	18240935	SNORA77.1	-2.5	-1.1	1.4
16	33001989	33002107	n-R5s33	-2.5	-1.1	1.4
16	88924295	88924462	Gm9829	-2.5	-1.1	1.4
16	11313905	11320167	Tnfrsf17	2.4	3.8	1.4
16	33056539	33060275	Rpl35a	0.7	2.0	1.3
16	50589972	50591267	Ccdc54	-0.7	0.5	1.2
16	16048958	16049433	Gm10121	-1.3	-0.1	1.2
16	20604525	20604609	Mir1224	-1.3	-0.1	1.2
16	58404991	58408389	4930461C15Rik	0.0	1.2	1.2
16	48872721	48874611	Retnlg	-2.3	-1.1	1.2
16	48283848	48294405	Dppa4	1.8	3.1	1.2
16	5211921	5222392	AU021092	2.7	3.9	1.2
16	3707215	3718124	Mefv	1.6	2.8	1.2
16	96122005	96123335	Gm15340	0.5	1.7	1.2
16	48816969	48819004	Retnlb	0.0	1.2	1.2
16	8766262	8766369	U6.574	-0.3	0.8	1.1
16	91677513	91689010	Donson	2.7	3.8	1.1
16	23889667	23890930	Sst	-0.7	0.4	1.1
16	88896213	88896695	Krtap8-2	-0.7	0.4	1.1
16	96027636	96029924	2810404F17Rik	1.1	2.2	1.1
16	89570421	89571428	AC140433.1	1.7	2.7	1.1
16	92319008	92321686	4930563D23Rik	2.8	3.8	1.1
16	89403272	89404019	Krtap16-7	-0.3	0.7	1.0
16	24391698	24393674	AC116484.1	1.0	2.0	1.0
16	4941420	4964330	Anks3	4.3	5.3	1.0
16	22918455	22939839	Fetub	3.7	4.7	1.0
16	32400592	32408313	Tm4sf19	3.1	4.1	1.0
16	88506052	88507223	Cldn17	1.5	2.5	1.0
16	91320697	91329240	Gm15966	3.2	4.2	1.0
16	14311934	14314107	Gm15868	0.0	1.0	1.0
16	89507949	89508568	Krtap7-1	0.0	1.0	1.0
17	35302064	35304124	CR974444.1	-2.5	1.2	3.6
17	15080189	15082944	1600012H06Rik	-1.3	1.2	2.6
17	7047683	7047764	AC125143.1	-2.3	0.2	2.6
17	76683755	76684201	Gm16391	-2.3	-0.1	2.2
17	23658600	23658678	AC154753.1	-2.3	-0.5	1.8
17	35836148	35836254	U6.372	-2.3	-0.5	1.8
17	13714322	13721299	Smok4a	-0.7	1.0	1.7
17	37862944	37863936	Olfr120	-2.5	-0.8	1.7
17	36097675	36097759	Mir877	-0.7	0.8	1.6
17	14846073	14847062	Gm3222	-2.3	-0.8	1.6
17	57977850	57978656	Rpl7a-ps5	-2.3	-0.8	1.6
17	87815325	87815381	AC084383.1	-0.7	0.7	1.5
17	29158238	29159347	Gm16195	-1.3	0.1	1.4
17	35250889	35252345	Ly6g5b	1.3	2.6	1.4
17	27700520	27702694	Al413582	0.3	1.6	1.4
17	21622141	21625839	Gm10150	-1.3	-0.1	1.2
17	35118805	35122834	Lsm2	-0.3	0.9	1.2
17	50378729	50380849	AY702103	-1.3	-0.1	1.2
17	33352340	33353287	Gm4461	-2.3	-1.1	1.2
17	36245531	36248633	Gm8815	-2.3	-1.1	1.2
17	36979171	36979297	SNORA51.5	-2.3	-1.1	1.2
17	37287952	37288996	Olfr93	-2.3	-1.1	1.2
17	47427905	47427977	5S_rRNA.105	-2.3	-1.1	1.2
17	56382411	56382521	Mir7b	-2.3	-1.1	1.2
17	24830136	24833101	Gfer	1.8	3.0	1.2
17	33896937	33898807	CT030732.1	-1.3	-0.3	1.0
17	46141895	46141949	AC127690.1	-1.3	-0.3	1.0

17	24300833	24306647	Atp6v0c	1.0	2.0	1.0
18	37815561	37818391	3222401L13Rik	-0.7	1.4	2.1
18	84888396	84890167	D030046N08Rik	-0.7	1.2	2.0
18	35929972	35930325	AC132837.1	-2.3	-0.5	1.8
18	78546493	78547440	AC151578.1	-2.3	-0.5	1.8
18	48204989	48208032	Gm5506	-1.3	0.2	1.6
18	12657194	12657637	AC102131.1	-2.3	-0.8	1.6
18	54143975	54144982	Gm5507	-2.3	-0.8	1.6
18	61807470	61807557	Mir145	-1.3	0.1	1.4
18	31979516	31980067	AC161511.1	-2.5	-1.1	1.4
18	61557489	61557554	Mir378	-2.5	-1.1	1.4
18	61298163	61298267	U6.855	-1.3	-0.1	1.2
18	24919664	24919940	7SK.245	-2.3	-1.1	1.2
18	39153289	39153846	Gm15334	-2.3	-1.1	1.2
18	42412040	42412489	Gm6636	-2.3	-1.1	1.2
18	74224756	74225238	Gm9925	-2.3	-1.1	1.2
18	75069806	75069887	AC155261.1	-2.3	-1.1	1.2
18	50727310	50728353	Hdhd1a	-0.3	0.8	1.1
18	35696237	35696605	Gm5239	0.0	1.0	1.0
19	8847012	8848957	Tmem179b	-2.3	0.9	3.2
19	29209769	29209865	Mir101b	-2.3	0.6	2.9
19	5848508	5851372	Gm9783	-2.3	0.2	2.6
19	45823998	45826800	Gm15491	-1.3	0.8	2.1
19	11944636	11945795	Olfr1419	-0.7	1.4	2.1
19	11122103	11125056	AW112010	-1.3	0.5	1.8
19	12183090	12184034	Olfr1428	-2.3	-0.5	1.8
19	29156223	29156303	AC157914.1	-2.3	-0.5	1.8
19	51217118	51217225	AC101812.1	-2.3	-0.5	1.8
19	11769531	11769669	SCARNA17.2	-2.5	-0.8	1.7
19	34895177	34895263	Mir107	-2.5	-0.8	1.7
19	41986361	41993154	Pgam1	0.5	2.2	1.7
19	41667853	41670648	AI606181	1.5	3.1	1.6
19	5728087	5729653	Mtvr2	-1.3	0.2	1.6
19	4078299	4078512	U3.1	-2.3	-0.8	1.6
19	29077895	29078332	Gm10136	-2.3	-0.8	1.6
19	12921263	12921763	AC135633.1	-0.3	1.2	1.5
19	41920468	41922622	Frat2	0.7	2.2	1.5
19	6094419	6098233	Gm550	1.6	3.0	1.5
19	11844881	11849434	Mrpl16	1.1	2.6	1.4
19	12838150	12840122	Cntf	-1.3	0.1	1.4
19	28837651	28837920	AC163993.1	-0.3	1.1	1.4
19	18807231	18807394	U1.45	-2.5	-1.1	1.4
19	47132652	47132758	U6.16	-2.5	-1.1	1.4
19	12110298	12111230	Olfr1423	0.8	2.2	1.4
19	36806227	36811143	Ppp1r3c	2.2	3.5	1.3
19	47806246	47810078	6330577E15Rik	1.5	2.7	1.3
19	29038410	29041288	Ppapdc2	0.3	1.5	1.2
19	33897398	33899561	Gm5519	-1.3	-0.1	1.2
19	23717899	23718000	SNORA17.363	-2.3	-1.1	1.2
19	25746907	25753500	Dmrt2	2.9	4.1	1.2
19	30620353	30624155	Dkk1	2.4	3.6	1.2
19	30078280	30080510	Trpd52l3	2.1	3.3	1.2
19	4839366	4842528	Ccdc87	1.1	2.3	1.2
19	50104614	50105225	Rpl13a-ps1	0.0	1.2	1.2
19	56438752	56440731	RP23-116D4.2	1.9	3.1	1.2
19	53893844	53899618	Gm16299	3.3	4.4	1.1
19	47244959	47245624	AC132288.1	0.5	1.6	1.1
19	10743447	10752561	Pga5	3.9	5.1	1.1
19	31157331	31161080	Cstf2t	2.0	3.1	1.1
19	44213756	44220936	Bloc1s2	2.2	3.3	1.1
19	41136964	41151603	Opalin	3.3	4.4	1.1
19	5366813	5371511	Eif1ad	1.6	2.6	1.1
19	11970489	11971540	Olfr1420	-0.3	0.7	1.0
19	38126374	38129018	I830134H01Rik	-1.3	-0.3	1.0
19	41819294	41819543	7SK.153	-1.3	-0.3	1.0
19	52338813	52339946	Ins1	-0.3	0.7	1.0
19	53446056	53446256	Gm10197	-0.3	0.7	1.0
19	6268795	6276605	Gm14963	3.4	4.4	1.0
19	60831890	60835816	Nanos1	0.5	1.5	1.0
19	29772411	29782938	Mlana	3.8	4.8	1.0
X	50408223	50410367	C430049B03Rik	-2.3	2.8	5.1
X	20192452	20194745	Ndufb11	-2.5	2.0	4.4
X	163645025	163647247	Tmsb4x	-2.5	2.0	4.4
X	132703523	132706881	Tceal8	-2.5	1.5	4.0
X	71611057	71614012	Ubl4	-1.3	2.3	3.6
X	131291134	131295994	Armcx3	-1.3	2.3	3.6
X	71032367	71036169	Ssr4	0.3	3.8	3.5
X	58146828	58151242	Gm14662	-0.3	3.2	3.5

X	130222319	130241801	Gm16795	0.0	3.5	3.5
X	50911099	50912309	Cxx1c	-1.3	2.0	3.3
X	131135718	131141599	Hnrnph2	-0.7	2.5	3.2
X	132779618	132781678	Wbp5	-2.3	0.8	3.1
X	119508174	119510994	Nap1l3	-1.3	1.8	3.1
X	70473645	70476134	F8a	0.0	3.1	3.1
X	10904948	10905277	Gm4906	-2.3	0.7	3.0
X	132277272	132281861	Armcx5	-1.3	1.6	2.9
X	137176499	137176800	7SK.64	-2.3	0.6	2.9
X	71516151	71518474	Rpl10	-0.7	2.0	2.8
X	131120193	131122601	Rpl36a	0.0	2.6	2.6
X	78316033	78318441	Gm11014	0.3	2.9	2.6
X	151647675	151650992	Sat1	-0.7	1.8	2.6
X	34335646	34338802	Slc25a5	0.7	3.2	2.5
X	92217846	92222021	Spin4	0.7	3.1	2.4
X	70528235	70538315	Xlr5c	-2.3	0.1	2.4
X	97944760	97945087	7SK.252	-2.3	0.1	2.4
X	131280034	131281452	Gm10344	-2.3	0.1	2.4
X	149470709	149472106	Mageh1	-0.3	2.1	2.4
X	133276397	133277682	BC065397	0.0	2.3	2.3
X	96131655	96144356	Yipf6	2.5	4.9	2.3
X	70158871	70163750	Cetn2	0.0	2.3	2.3
X	71542436	71550060	Atp6ap1	1.0	3.3	2.3
X	49321872	49325949	1700080O16Rik	0.7	3.0	2.3
X	71527251	71535490	Taz	1.3	3.5	2.3
X	149441012	149443959	Usp51	0.7	2.9	2.3
X	55275877	55294913	Zic3	3.2	5.4	2.2
X	70459672	70467795	Xlr4b	-2.3	-0.1	2.2
X	159186488	159186606	SNORA76.2	-2.5	-0.3	2.2
X	35239616	35250409	Gm14543	-1.3	0.8	2.1
X	133461864	133462912	1700014N06Rik	1.0	3.1	2.1
X	117040805	117043774	Pabpc5	-0.7	1.4	2.1
X	97818222	97821907	Pdzd11	1.3	3.3	2.1
X	34873092	34888471	Rhox3c	-1.3	0.7	2.0
X	7904550	7913363	Ssxb10	-2.3	-0.3	2.0
X	34894844	34899631	Rhox4c	-2.3	-0.3	2.0
X	152057332	152057904	Gm15156	0.3	2.3	2.0
X	34450829	34451980	Gm15008	0.0	2.0	2.0
X	131338684	131343760	Armcx2	0.8	2.8	2.0
X	13788274	13788782	Gm5382	-2.5	-0.5	2.0
X	134742704	134744119	Tex13a	-2.5	-0.5	2.0
X	131282994	131285956	Armcx6	0.0	2.0	2.0
X	14687712	14688588	4930403L05Rik	-0.3	1.6	1.9
X	35180224	35183034	Rhox6	0.3	2.2	1.9
X	131071795	131076404	Timm8a1	0.3	2.2	1.9
X	7990431	7999284	Ssxb1	-2.3	-0.5	1.8
X	9424073	9424626	1700054O13Rik	-0.7	1.1	1.8
X	33735899	33739153	Zcchc12	0.3	2.1	1.8
X	34805370	34810131	Rhox4a	-1.3	0.5	1.8
X	35987726	35988259	AL590633.1	-2.3	-0.5	1.8
X	90531846	90543696	Arx	3.2	5.0	1.8
X	6953051	6955168	Gm14379	0.3	2.0	1.8
X	34731240	34732457	Rnf113a1	1.7	3.5	1.8
X	166412975	166416849	G530011O06Rik	0.5	2.2	1.8
X	7291459	7299226	Gm14720	0.7	2.4	1.7
X	149932775	149936080	Ubqln2	2.3	4.0	1.7
X	102315735	102319248	Magee1	3.1	4.8	1.7
X	5745814	5750133	Nudt10	-1.3	0.4	1.7
X	71162212	71167283	Naa10	1.8	3.4	1.7
X	99442720	99447520	Cited1	3.0	4.7	1.7
X	46048181	46063900	Rbmx2	1.9	3.6	1.6
X	108009729	108010910	Pou3f4	2.1	3.7	1.6
X	139534137	139538878	Rgag1	1.0	2.6	1.6
X	133267481	133278229	Morf4l2	2.1	3.7	1.6
X	123346994	123350616	4932411N23Rik	-0.7	0.8	1.6
X	131219101	131221001	B230119M05Rik	-1.3	0.2	1.6
X	25903110	25972691	Six	-2.3	-0.8	1.6
X	35083504	35087784	Rhox2h	-2.3	-0.8	1.6
X	35457254	35463964	Rhox12	1.1	2.7	1.6
X	99903204	99910548	Dmrtc1b	-2.3	-0.8	1.6
X	124909930	124911507	Gm4916	-2.3	-0.8	1.6
X	71500026	71506887	Emd	2.0	3.5	1.5
X	148436388	148438985	Hsd17b10	1.0	2.5	1.5
X	51272614	51278177	Gm15482	0.8	2.4	1.5
X	101678157	101702829	Uprt	1.1	2.6	1.5
X	133649936	133656622	Esx1	2.5	4.0	1.5
X	100379515	100381979	Nap1l2	0.5	2.0	1.5
X	8380817	8386512	Ssxb5	1.0	2.5	1.5

X	138159381	138160706	Gm15295	1.6	3.0	1.5
X	72650185	72661923	Mtcp1	1.1	2.6	1.5
X	72650185	72661927	Mtcp1	1.1	2.6	1.5
X	151758461	151775297	Prdx4	1.8	3.3	1.4
X	8469479	8472103	B630019K06Rik	1.5	2.9	1.4
X	71654822	71674533	G6pdx	2.6	4.0	1.4
X	20281032	20297665	Usp11	2.8	4.2	1.4
X	20518848	20539143	Uxt	1.8	3.2	1.4
X	40944337	40946151	Gm362	-1.3	0.1	1.4
X	150266768	150268786	Spin2	-0.3	1.1	1.4
X	23585725	23585819	AL773580.1	-2.5	-1.1	1.4
X	35017402	35022193	Rhox4f	-2.5	-1.1	1.4
X	50945665	50946940	Cxx1a	-2.5	-1.1	1.4
X	70319684	70327817	Xlr4a	-2.5	-1.1	1.4
X	71024302	71032236	Idh3g	2.2	3.6	1.4
X	17722996	17723816	Dusp21	1.1	2.5	1.4
X	46817122	46853421	Arhgap36	3.6	4.9	1.4
X	149993136	149993972	Cypt3	-0.7	0.6	1.4
X	71550337	71557201	Gdi1	2.2	3.5	1.3
X	7245067	7248516	Plp2	1.5	2.8	1.3
X	50880264	50912156	Gm14597	1.1	2.5	1.3
X	78316037	78343211	Tmem47	1.8	3.1	1.3
X	71518556	71527676	Dnase1l1	1.3	2.6	1.3
X	67639052	67642592	1110012L19Rik	0.7	2.0	1.3
X	49929020	49972224	A630012P03Rik	3.3	4.6	1.3
X	7501402	7505733	Eras	2.0	3.3	1.3
X	162895185	162917799	Rab9	2.8	4.1	1.3
X	120217104	120302404	3110007F17Rik	1.8	3.0	1.3
X	160346949	160371598	Ap1s2	2.4	3.6	1.3
X	5624571	5632228	Nudt11	-0.7	0.5	1.2
X	72818791	72820607	Gm15063	0.3	1.5	1.2
X	99380280	99384733	Rps4x	0.7	1.9	1.2
X	70310126	70313529	Pnma3	0.8	2.1	1.2
X	131252502	131256456	Armcx1	-0.3	0.9	1.2
X	8325558	8331713	Ssx9	-2.3	-1.1	1.2
X	25627804	25658481	Gm2012	-2.3	-1.1	1.2
X	26099903	26129764	Gm14525	-2.3	-1.1	1.2
X	34985928	34990666	Rhox2f	-2.3	-1.1	1.2
X	35061008	35065834	Rhox4g	-2.3	-1.1	1.2
X	35367085	35371992	4930525M21Rik	-2.3	-1.1	1.2
X	70479027	70488469	Xlr4c	-2.3	-1.1	1.2
X	132567906	132569485	6530401D17Rik	-2.3	-1.1	1.2
X	146219016	146261237	Gm15097	-2.3	-1.1	1.2
X	149357996	149358303	7SK.4	-2.3	-1.1	1.2
X	99261883	99266643	Rgag4	3.6	4.8	1.2
X	91780813	91787482	Maged1	2.8	4.1	1.2
X	53443004	53466058	Gm773	1.8	3.1	1.2
X	7471645	7476395	Pqbp1	3.2	4.4	1.2
X	8828054	8834025	Gm14862	1.4	2.6	1.2
X	35542970	35550223	Zbtb33	1.3	2.5	1.2
X	98624657	98643932	Nono	2.3	3.5	1.2
X	160508637	160514425	Siah1b	2.4	3.6	1.2
X	66609771	66614444	Gm14705	1.8	3.0	1.2
X	34622515	34625397	Rpl39	1.3	2.4	1.2
X	96654364	96656092	Gm14808	0.3	1.4	1.2
X	7215597	7230382	Syp	3.7	4.8	1.2
X	71259253	71269257	Irak1	2.1	3.2	1.2
X	131743226	131745226	Tceal6	0.0	1.2	1.2
X	50144377	50152645	Ccdc160	1.5	2.6	1.2
X	34138201	34146074	Pgrmc1	2.7	3.8	1.2
X	7581326	7595618	Glod5	2.6	3.7	1.2
X	20889336	20911853	Gm6938	0.7	1.8	1.1
X	159198334	159217024	Rbbp7	1.8	3.0	1.1
X	8681688	8683468	4930402K13Rik	0.5	1.6	1.1
X	135868222	135872708	Trap1a	2.1	3.2	1.1
X	136991152	137010679	Prps1	2.1	3.2	1.1
X	6948402	6952956	Usp27x	2.8	4.0	1.1
X	71558372	71565488	Fam50a	1.9	3.0	1.1
X	51080295	51100883	Xlr	-0.7	0.4	1.1
X	133527694	133533011	Zcchc18	0.8	2.0	1.1
X	138661313	138674235	Nxt2	1.1	2.2	1.1
X	138145541	138159806	Irs4	3.5	4.6	1.1
X	144200534	144267999	Gm15128	0.5	1.6	1.1
X	7762455	7770638	Ebp	2.8	3.8	1.1
X	34666790	34690741	Nkap	2.7	3.7	1.1
X	7638297	7651886	Suv39h1	3.8	4.8	1.1
X	71167460	71176189	Renbp	2.8	3.8	1.1
X	102695111	102695978	Cypt2-ps	-1.3	-0.3	1.0

X	139173543	139173640	Mir652	-1.3	-0.3	1.0
X	54636707	54646213	Rbmx	2.5	3.6	1.0
X	7305565	7308190	Praf2	1.6	2.6	1.0
X	70918489	70927841	Slc6a8	2.9	3.9	1.0
X	45694535	45761226	Bcorl1	5.6	6.7	1.0
X	35252274	35254947	Rhox9	0.8	1.9	1.0
X	156560141	156576345	Pdha1	3.1	4.1	1.0
X	58294916	58329243	Gm14664	1.5	2.5	1.0
X	72817385	72823570	Rab39b	1.5	2.5	1.0
X	35184863	35194450	Rhox7	0.5	1.5	1.0
X	56162537	56173481	4930550L24Rik	1.3	2.3	1.0
X	71630058	71638611	Fam3a	2.0	3.0	1.0
X	58144545	58146605	Sox3	3.1	4.1	1.0
X	7848259	7857305	Slc38a5	2.7	3.7	1.0

Appendix 4

Overlap between LSH, G9a and SmcHD1 target promoters

LSH & G9a	LSH & SmcHD1	G9a & SmcHD1	LSH & G9a & SmcHD1
1700071K01Rik	B020018G12Rik	Cxcl10	4930479M11Rik
4922502D21Rik	Cldn17	Ferd3l	Olfr370
Abca15	Ifng	Hax1	Slitrk1
Adcyap1	Krtap8-2		
Asz1	Olfr1423		
AU015228	Olfr214		
BC050811			
C6.1AL			
Ccdc42			
Cd2			
Cetn1			
Cox8c			
Crisp2			
Dgkg			
Gm740			
Gucy2c			
Gykl1			
Hemt1			
Lcp1			
LOC621852			
Myf5			
Oasl2			
Odz3			
Olfr1349			
Olfr281			
Psors1c2			
Spata16			
Tcf21			
Tmprss3			
Wdr20			

References

Amir RE, Van den Veyver IB, Wan M, Tran CQ, Francke U, Zoghbi HY (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nature genetics* **23:** 185-188

Arita K, Ariyoshi M, Tochio H, Nakamura Y, Shirakawa M (2008) Recognition of hemi-methylated DNA by the SRA protein UHRF1 by a base-flipping mechanism. *Nature* **455:** 818-821

Arnaud P (2010) Genomic imprinting in germ cells: imprints are under control. *Reproduction* **140:** 411-423

Athanasiadou R, de Sousa D, Myant K, Merusi C, Stancheva I, Bird A (2010) Targeting of de novo DNA methylation throughout the Oct-4 gene regulatory region in differentiating embryonic stem cells. *PloS one* **5:** e9937

Avner P, Heard E (2001) X-chromosome inactivation: counting, choice and initiation. *Nat Rev Genet* **2:** 59-67

Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, Gouti M, Casanova M, Warnes G, Merkenschlager M, Fisher AG (2006) Chromatin signatures of pluripotent cell lines. *Nat Cell Biol* **8:** 532-538

Bacher CP, Guggiari M, Brors B, Augui S, Clerc P, Avner P, Eils R, Heard E (2006) Transient colocalization of X-inactivation centres accompanies the initiation of X inactivation. *Nat Cell Biol* **8:** 293-299

Bacolla A, Pradhan S, Roberts RJ, Wells RD (1999) Recombinant human DNA (cytosine-5) methyltransferase. II. Steady-state kinetics reveal allosteric activation by methylated dna. *J Biol Chem* **274:** 33011-33019

Barakat TS, Gunhanlar N, Pardo CG, Achame EM, Ghazvini M, Boers R, Kenter A, Rentmeester E, Grootegoed JA, Gribnau J (2011) RNF12 activates Xist and is essential for X chromosome inactivation. *PLoS genetics* **7:** e1002001

Barbisan F, Mazzucchelli R, Santinelli A, Stramazzotti D, Scarpelli M, Lopez-Beltran A, Cheng L, Montironi R (2008) Immunohistochemical evaluation of global DNA methylation and histone acetylation in papillary urothelial neoplasm of low malignant potential. *International journal of immunopathology and pharmacology* **21:** 615-623

Barlesi F, Giaccone G, Gallegos-Ruiz MI, Loundou A, Span SW, Lefesvre P, Kruyt FA, Rodriguez JA (2007) Global histone modifications predict prognosis of resected non small-cell lung cancer. *Journal of clinical oncology : official journal of the American Society of Clinical Oncology* **25:** 4358-4364

Bartolomei MS, Ferguson-Smith AC (2011) Mammalian genomic imprinting. *Cold Spring Harbor perspectives in biology* **3**

Barton SC, Surani MA, Norris ML (1984) Role of paternal and maternal genomes in mouse development. *Nature* **311:** 374-376

Beard C, Li E, Jaenisch R (1995) Loss of methylation activates Xist in somatic but not in embryonic cells. *Genes Dev* **9:** 2325-2334

Bednar J, Horowitz RA, Grigoryev SA, Carruthers LM, Hansen JC, Koster AJ, Woodcock CL (1998) Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. *Proc Natl Acad Sci U S A* **95:** 14173-14178

Bellacosa A, Cicchillitti L, Schepis F, Riccio A, Yeung AT, Matsumoto Y, Golemis EA, Genuardi M, Neri G (1999) MED1, a novel human methyl-CpG-binding endonuclease, interacts with DNA mismatch repair protein MLH1. *Proc Natl Acad Sci U S A* **96:** 3969-3974

Bernstein BE, Kamal M, Lindblad-Toh K, Bekiranov S, Bailey DK, Huebert DJ, McMahon S, Karlsson EK, Kulbokas EJ, 3rd, Gingeras TR, Schreiber SL, Lander ES (2005) Genomic maps and comparative analysis of histone modifications in human and mouse. *Cell* **120:** 169-181

Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, Fry B, Meissner A, Wernig M, Plath K, Jaenisch R, Wagschal A, Feil R, Schreiber SL, Lander ES (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **125:** 315-326

Bestor T, Laudano A, Mattaliano R, Ingram V (1988) Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of molecular biology* **203:** 971-983

Bian Q, Belmont AS (2012) Revisiting higher-order and large-scale chromatin organization. *Current opinion in cell biology* **24:** 359-366

Bibel M, Richter J, Schrenk K, Tucker KL, Staiger V, Korte M, Goetz M, Barde YA (2004) Differentiation of mouse embryonic stem cells into a defined neuronal lineage. *Nature neuroscience* **7:** 1003-1009

Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev* **16:** 6-21

Bird A, Taggart M, Frommer M, Miller OJ, Macleod D (1985) A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* **40:** 91-99

Bird AP (1980) DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res* **8:** 1499-1504

Bird AP (1986) CpG-rich islands and the function of DNA methylation. *Nature* **321:** 209-213

Blewitt ME, Gendrel AV, Pang Z, Sparrow DB, Whitelaw N, Craig JM, Apedaile A, Hilton DJ, Dunwoodie SL, Brockdorff N, Kay GF, Whitelaw E (2008) SmcHD1, containing a structural-maintenance-of-chromosomes hinge domain, has a critical role in X inactivation. *Nature genetics* **40:** 663-669

Borgel J, Guibert S, Li Y, Chiba H, Schubeler D, Sasaki H, Forne T, Weber M (2010) Targets and dynamics of promoter DNA methylation during early mouse development. *Nature genetics* **42:** 1093-1100

Bostick M, Kim JK, Esteve PO, Clark A, Pradhan S, Jacobsen SE (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* **317:** 1760-1764

Bourc'his D, Bestor TH (2004) Meiotic catastrophe and retrotransposon reactivation in male germ cells lacking Dnmt3L. *Nature* **431:** 96-99

Bowman GD (2010) Mechanisms of ATP-dependent nucleosome sliding. *Current opinion in structural biology* **20:** 73-81

Boyer LA, Latek RR, Peterson CL (2004) The SANT domain: a unique histone-tail-binding module? *Nature reviews Molecular cell biology* **5:** 158-163

Bracken AP, Helin K (2009) Polycomb group proteins: navigators of lineage pathways led astray in cancer. *Nature reviews Cancer* **9:** 773-784

Brasher SV, Smith BO, Fogh RH, Nietlispach D, Thiru A, Nielsen PR, Broadhurst RW, Ball LJ, Murzina NV, Laue ED (2000) The structure of mouse HP1 suggests a unique mode of single peptide recognition by the shadow chromo domain dimer. *EMBO J* **19:** 1587-1597

Brockdorff N, Ashworth A, Kay GF, Cooper P, Smith S, McCabe VM, Norris DP, Penny GD, Patel D, Rastan S (1991) Conservation of position and exclusive expression of mouse Xist from the inactive X chromosome. *Nature* **351:** 329-331

Brockdorff N, Ashworth A, Kay GF, McCabe VM, Norris DP, Cooper PJ, Swift S, Rastan S (1992) The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell* **71:** 515-526

Bruno M, Flaus A, Stockdale C, Rencurel C, Ferreira H, Owen-Hughes T (2003) Histone H2A/H2B dimer exchange by ATP-dependent chromatin remodeling activities. *Mol Cell* **12:** 1599-1606

Brzeski J, Jerzmanowski A (2003) Deficient in DNA methylation 1 (DDM1) defines a novel family of chromatin-remodeling factors. *J Biol Chem* **278:** 823-828

Cairns BR (2007) Chromatin remodeling: insights and intrigue from single-molecule studies. *Nat Struct Mol Biol* **14:** 989-996

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

Carlone DL, Lee JH, Young SR, Dobrota E, Butler JS, Ruiz J, Skalnik DG (2005) Reduced genomic cytosine methylation and defective cellular differentiation in embryonic stem cells lacking CpG binding protein. *Molecular and cellular biology* **25:** 4881-4891

Carlone DL, Skalnik DG (2001) CpG binding protein is crucial for early embryonic development. *Molecular and cellular biology* **21:** 7601-7606

Carrozza MJ, Li B, Florens L, Suganuma T, Swanson SK, Lee KK, Shia WJ, Anderson S, Yates J, Washburn MP, Workman JL (2005) Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* **123:** 581-592

Caterino TL, Hayes JJ (2011) Structure of the H1 C-terminal domain and function in chromatin condensation. *Biochem Cell Biol* **89:** 35-44

Cervoni N, Szyf M (2001) Demethylase activity is directed by histone acetylation. *J Biol Chem* **276:** 40778-40787

Chahrour M, Jung SY, Shaw C, Zhou X, Wong ST, Qin J, Zoghbi HY (2008) MeCP2, a key contributor to neurological disease, activates and represses transcription. *Science* **320:** 1224-1229

Chang SC, Tucker T, Thorogood NP, Brown CJ (2006) Mechanisms of X-chromosome inactivation. *Frontiers in bioscience : a journal and virtual library* **11:** 852-866

Chase A, Cross NC (2011) Aberrations of EZH2 in cancer. *Clinical cancer research : an official journal of the American Association for Cancer Research* **17:** 2613-2618

Chaturvedi CP, Hosey AM, Palii C, Perez-Iratxeta C, Nakatani Y, Ranish JA, Dilworth FJ, Brand M (2009) Dual role for the methyltransferase G9a in the maintenance of beta-globin gene transcription in adult erythroid cells. *Proc Natl Acad Sci U S A* **106:** 18303-18308

Cheadle JP, Gill H, Fleming N, Maynard J, Kerr A, Leonard H, Krawczak M, Cooper DN, Lynch S, Thomas N, Hughes H, Hulten M, Ravine D, Sampson JR, Clarke A (2000) Long-read sequence analysis of the MECP2 gene in Rett syndrome patients: correlation of disease severity with mutation type and location. *Human molecular genetics* **9:** 1119-1129

Chedin F, Lieber MR, Hsieh CL (2002) The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proc Natl Acad Sci U S A* **99:** 16916-16921

Chen T, Hevi S, Gay F, Tsujimoto N, He T, Zhang B, Ueda Y, Li E (2007) Complete inactivation of DNMT1 leads to mitotic catastrophe in human cancer cells. *Nature genetics* **39:** 391-396

Cheng X (1995) DNA modification by methyltransferases. *Current opinion in structural biology* **5:** 4-10

Cheng X, Blumenthal RM (2008) Mammalian DNA methyltransferases: a structural perspective. *Structure* **16:** 341-350

Chu C, Qu K, Zhong FL, Artandi SE, Chang HY (2011) Genomic maps of long noncoding RNA occupancy reveal principles of RNA-chromatin interactions. *Mol Cell* **44:** 667-678

Ciccone DN, Su H, Hevi S, Gay F, Lei H, Bajko J, Xu G, Li E, Chen T (2009) KDM1B is a histone H3K4 demethylase required to establish maternal genomic imprints. *Nature* **461:** 415-418

Clapier CR, Cairns BR (2009) The biology of chromatin remodeling complexes. *Annual review of biochemistry* **78:** 273-304

Cohen JF, Leis A, Lecarpentier T, Raymond J, Gendrel D, Chalumeau M (2012) Procalcitonin predicts response to beta-lactam treatment in hospitalized children with community-acquired pneumonia. *PloS one* **7:** e36927

Collins RE, Northrop JP, Horton JR, Lee DY, Zhang X, Stallcup MR, Cheng X (2008) The ankyrin repeats of G9a and GLP histone methyltransferases are mono- and dimethyllysine binding modules. *Nat Struct Mol Biol* **15:** 245-250

Coulondre C, Miller JH (1978) Analysis of base substitutions induced by ultraviolet radiation in Escherichia coli. *National Cancer Institute monograph***:** 115-119

Coulondre C, Miller JH, Farabaugh PJ, Gilbert W (1978) Molecular basis of base substitution hotspots in Escherichia coli. *Nature* **274:** 775-780

Cowieson NP, Partridge JF, Allshire RC, McLaughlin PJ (2000) Dimerisation of a chromo shadow domain and distinctions from the chromodomain as revealed by structural analysis. *Curr Biol* **10:** 517-525

Crawford DC, Acuna JM, Sherman SL (2001) FMR1 and the fragile X syndrome: human genome epidemiology review. *Genetics in medicine : official journal of the American College of Medical Genetics* **3:** 359-371

Cross SH, Charlton JA, Nan X, Bird AP (1994) Purification of CpG islands using a methylated DNA binding column. *Nature genetics* **6:** 236-244

da Rocha ST, Edwards CA, Ito M, Ogata T, Ferguson-Smith AC (2008) Genomic imprinting at the mammalian Dlk1-Dio3 domain. *Trends Genet* **24:** 306-316

Dang W, Bartholomew B (2007) Domain architecture of the catalytic subunit in the ISW2-nucleosome complex. *Molecular and cellular biology* **27:** 8306-8317

Dang W, Kagalwala MN, Bartholomew B (2006) Regulation of ISW2 by concerted action of histone H4 tail and extranucleosomal DNA. *Molecular and cellular biology* **26:** 7388-7396

De La Fuente R, Baumann C, Fan T, Schmidtmann A, Dobrinski I, Muegge K (2006) Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells. *Nat Cell Biol* **8:** 1448-1454

Dean W, Bowden L, Aitchison A, Klose J, Moore T, Meneses JJ, Reik W, Feil R (1998) Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. *Development* **125:** 2273-2282

Deaton AM, Bird A (2011) CpG islands and the regulation of transcription. *Genes Dev* **25:** 1010-1022

Dechassa ML, Sabri A, Pondugula S, Kassabov SR, Chatterjee N, Kladde MP, Bartholomew B (2010) SWI/SNF has intrinsic nucleosome disassembly activity that is dependent on adjacent nucleosomes. *Mol Cell* **38:** 590-602

Dechassa ML, Zhang B, Horowitz-Scherer R, Persinger J, Woodcock CL, Peterson CL, Bartholomew B (2008) Architecture of the SWI/SNF-nucleosome complex. *Molecular and cellular biology* **28:** 6010-6021

Deleris A, Greenberg MV, Ausin I, Law RW, Moissiard G, Schubert D, Jacobsen SE (2010) Involvement of a Jumonji-C domain-containing histone demethylase in DRM2-mediated maintenance of DNA methylation. *EMBO reports* **11:** 950-955

Dennis K, Fan T, Geiman T, Yan Q, Muegge K (2001) Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes Dev* **15:** 2940-2944

Ding F, Chaillet JR (2002) In vivo stabilization of the Dnmt1 (cytosine-5)- methyltransferase protein. *Proc Natl Acad Sci U S A* **99:** 14861-14866

Dong KB, Maksakova IA, Mohn F, Leung D, Appanah R, Lee S, Yang HW, Lam LL, Mager DL, Schubeler D, Tachibana M, Shinkai Y, Lorincz MC (2008) DNA methylation in ES cells requires the lysine methyltransferase G9a but not its catalytic activity. *EMBO J* **27:** 2691-2701

Donohoe ME, Silva SS, Pinter SF, Xu N, Lee JT (2009) The pluripotency factor Oct4 interacts with Ctcf and also controls X-chromosome pairing and counting. *Nature* **460:** 128-132

Easwaran HP, Schermelleh L, Leonhardt H, Cardoso MC (2004) Replication-independent chromatin loading of Dnmt1 during G2 and M phases. *EMBO reports* **5:** 1181-1186

Eden A, Gaudet F, Waghmare A, Jaenisch R (2003) Chromosomal instability and tumors promoted by DNA hypomethylation. *Science* **300:** 455

Eissenberg JC, Elgin SC (2000) The HP1 protein family: getting a grip on chromatin. *Curr Opin Genet Dev* **10:** 204-210

Elgin SC, Grewal SI (2003) Heterochromatin: silence is golden. *Curr Biol* **13:** R895-898

Engemann S, Strodicke M, Paulsen M, Franck O, Reinhardt R, Lane N, Reik W, Walter J (2000) Sequence and functional comparison in the Beckwith-Wiedemann region: implications for a novel imprinting centre and extended imprinting. *Human molecular genetics* **9:** 2691-2706

Epsztejn-Litman S, Feldman N, Abu-Remaileh M, Shufaro Y, Gerson A, Ueda J, Deplus R, Fuks F, Shinkai Y, Cedar H, Bergman Y (2008) De novo DNA methylation promoted by G9a prevents reprogramming of embryonically silenced genes. *Nat Struct Mol Biol* **15:** 1176-1183

Esteller M (2005) DNA methylation and cancer therapy: new developments and expectations. *Current opinion in oncology* **17:** 55-60

Esteve PO, Chin HG, Smallwood A, Feehery GR, Gangisetty O, Karpf AR, Carey MF, Pradhan S (2006) Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes Dev* **20:** 3089-3103

Fan T, Hagan JP, Kozlov SV, Stewart CL, Muegge K (2005) Lsh controls silencing of the imprinted Cdkn1c gene. *Development* **132:** 635-644

Fan T, Schmidtmann A, Xi S, Briones V, Zhu H, Suh HC, Gooya J, Keller JR, Xu H, Roayaei J, Anver M, Ruscetti S, Muegge K (2008) DNA hypomethylation caused by Lsh deletion promotes erythroleukemia development. *Epigenetics* **3:** 134-142

Fan T, Yan Q, Huang J, Austin S, Cho E, Ferris D, Muegge K (2003) Lsh-deficient murine embryonal fibroblasts show reduced proliferation with signs of abnormal mitosis. *Cancer research* **63:** 4677-4683

Fatemi M, Hermann A, Gowher H, Jeltsch A (2002) Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. *European journal of biochemistry / FEBS* **269:** 4981-4984

Fatemi M, Hermann A, Pradhan S, Jeltsch A (2001) The activity of the murine DNA methyltransferase Dnmt1 is controlled by interaction of the catalytic domain with the N-terminal part of the enzyme leading to an allosteric activation of the enzyme after binding to methylated DNA. *Journal of molecular biology* **309:** 1189-1199

Feinberg AP, Vogelstein B (1983) Hypomethylation distinguishes genes of some human cancers from their normal counterparts. *Nature* **301:** 89-92

Feldman N, Gerson A, Fang J, Li E, Zhang Y, Shinkai Y, Cedar H, Bergman Y (2006) G9a-mediated irreversible epigenetic inactivation of Oct-3/4 during early embryogenesis. *Nat Cell Biol* **8:** 188-194

Felle M, Hoffmeister H, Rothammer J, Fuchs A, Exler JH, Langst G (2011) Nucleosomes protect DNA from DNA methylation in vivo and in vitro. *Nucleic Acids Res* **39:** 6956-6969

Feng S, Jacobsen SE, Reik W (2010) Epigenetic reprogramming in plant and animal development. *Science* **330:** 622-627

Fernandez AF, Assenov Y, Martin-Subero JI, Balint B, Siebert R, Taniguchi H, Yamamoto H, Hidalgo M, Tan AC, Galm O, Ferrer I, Sanchez-Cespedes M, Villanueva A, Carmona J, Sanchez-Mut JV, Berdasco M, Moreno V, Capella G, Monk D, Ballestar E, Ropero S, Martinez R, Sanchez-Carbayo M, Prosper F, Agirre X, Fraga MF, Grana O, Perez-Jurado L, Mora J, Puig

S, Prat J, Badimon L, Puca AA, Meltzer SJ, Lengauer T, Bridgewater J, Bock C, Esteller M (2012) A DNA methylation fingerprint of 1628 human samples. *Genome research* **22:** 407-419

Filion GJ, Zhenilo S, Salozhin S, Yamada D, Prokhortchouk E, Defossez PA (2006) A family of human zinc finger proteins that bind methylated DNA and repress transcription. *Molecular and cellular biology* **26:** 169-181

Finch JT, Klug A (1976) Solenoidal model for superstructure in chromatin. *Proc Natl Acad Sci U S A* **73:** 1897-1901

Fischle W, Tseng BS, Dormann HL, Ueberheide BM, Garcia BA, Shabanowitz J, Hunt DF, Funabiki H, Allis CD (2005) Regulation of HP1-chromatin binding by histone H3 methylation and phosphorylation. *Nature* **438:** 1116-1122

Fisscher U, Weisbeek P, Smeekens S (1996) A tobacco nuclear protein that preferentially binds to unmethylated CpG-rich DNA. *European journal of biochemistry / FEBS* **235:** 585-592

Fitzpatrick GV, Soloway PD, Higgins MJ (2002) Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nature genetics* **32:** 426-431

Flaus A, Martin DM, Barton GJ, Owen-Hughes T (2006) Identification of multiple distinct Snf2 subfamilies with conserved structural motifs. *Nucleic Acids Res* **34:** 2887-2905

Foster CT, Dovey OM, Lezina L, Luo JL, Gant TW, Barlev N, Bradley A, Cowley SM (2010) Lysine-specific demethylase 1 regulates the embryonic transcriptome and CoREST stability. *Molecular and cellular biology* **30:** 4851-4863

Fraga MF, Ballestar E, Villar-Garea A, Boix-Chornet M, Espada J, Schotta G, Bonaldi T, Haydon C, Ropero S, Petrie K, Iyer NG, Perez-Rosado A, Calvo E, Lopez JA, Cano A, Calasanz MJ, Colomer D, Piris MA, Ahn N, Imhof A, Caldas C, Jenuwein T, Esteller M (2005) Loss of acetylation at Lys16 and trimethylation at Lys20 of histone H4 is a common hallmark of human cancer. *Nature genetics* **37:** 391-400

Fraga MF, Herranz M, Espada J, Ballestar E, Paz MF, Ropero S, Erkek E, Bozdogan O, Peinado H, Niveleau A, Mao JH, Balmain A, Cano A, Esteller M (2004) A mouse skin multistage carcinogenesis model reflects the aberrant DNA methylation patterns of human tumors. *Cancer research* **64:** 5527-5534

Franchina M, Kay PH (2000) Evidence that cytosine residues within 5'-CCTGG-3' pentanucleotides can be methylated in human DNA independently of the methylating system that modifies 5'-CG-3' dinucleotides. *DNA and cell biology* **19:** 521-526

Fuks F, Hurd PJ, Deplus R, Kouzarides T (2003) The DNA methyltransferases associate with HP1 and the SUV39H1 histone methyltransferase. *Nucleic Acids Res* **31:** 2305-2312

Gangaraju VK, Bartholomew B (2007) Dependency of ISW1a chromatin remodeling on extranucleosomal DNA. *Molecular and cellular biology* **27:** 3217-3225

Geiman TM, Durum SK, Muegge K (1998) Characterization of gene expression, genomic structure, and chromosomal localization of Hells (Lsh). *Genomics* **54:** 477-483

Geiman TM, Tessarollo L, Anver MR, Kopp JB, Ward JM, Muegge K (2001) Lsh, a SNF2 family member, is required for normal murine development. *Biochimica et biophysica acta* **1526:** 211-220

Gendrel AV, Apedaile A, Coker H, Termanis A, Zvetkova I, Godwin J, Tang YA, Huntley D, Montana G, Taylor S, Giannoulatou E, Heard E, Stancheva I, Brockdorff N (2012) Smchd1-dependent and -independent pathways determine developmental dynamics of CpG island methylation on the inactive X chromosome. *Developmental cell* **23:** 265-279

Gendrel AV, Lippman Z, Yordan C, Colot V, Martienssen RA (2002) Dependence of heterochromatic histone H3 methylation patterns on the Arabidopsis gene DDM1. *Science* **297:** 1871-1873

Gendrel D (2012) [Anti-meningococcal vaccines: diversity of vaccination policies and recommendations]. *Archives de pediatrie : organe officiel de la Societe francaise de pediatrie* **19 Suppl 2:** S70-76

Gendrel D, Cohen R (2012) [Meningococcal vaccine. Editorial]. *Archives de pediatrie : organe officiel de la Societe francaise de pediatrie* **19 Suppl 2:** S47-48

Georgel PT, Horowitz-Scherer RA, Adkins N, Woodcock CL, Wade PA, Hansen JC (2003) Chromatin compaction by human MeCP2. Assembly of novel secondary chromatin structures in the absence of DNA methylation. *J Biol Chem* **278:** 32181-32188

Gibbons RJ, McDowell TL, Raman S, O'Rourke DM, Garrick D, Ayyub H, Higgs DR (2000) Mutations in ATRX, encoding a SWI/SNF-like protein, cause diverse changes in the pattern of DNA methylation. *Nature genetics* **24:** 368-371

Gowher H, Jeltsch A (2001) Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG [correction of non-CpA] sites. *Journal of molecular biology* **309:** 1201-1208

Gowher H, Liebert K, Hermann A, Xu G, Jeltsch A (2005a) Mechanism of stimulation of catalytic activity of Dnmt3A and Dnmt3B DNA-(cytosine-C5)-methyltransferases by Dnmt3L. *J Biol Chem* **280:** 13341-13348

Gowher H, Stockdale CJ, Goyal R, Ferreira H, Owen-Hughes T, Jeltsch A (2005b) De novo methylation of nucleosomal DNA by the mammalian Dnmt1 and Dnmt3A DNA methyltransferases. *Biochemistry* **44:** 9899-9904

Grigoryev SA, Arya G, Correll S, Woodcock CL, Schlick T (2009) Evidence for heteromorphic chromatin fibers from analysis of nucleosome interactions. *Proc Natl Acad Sci U S A* **106:** 13317-13322

Gruenbaum Y, Cedar H, Razin A (1982) Substrate and sequence specificity of a eukaryotic DNA methylase. *Nature* **295:** 620-622

Guenther MG, Levine SS, Boyer LA, Jaenisch R, Young RA (2007) A chromatin landmark and transcription initiation at most promoters in human cells. *Cell* **130:** 77-88

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458:** 223-227

Haas DL, Lutzko C, Logan AC, Cho GJ, Skelton D, Jin Yu X, Pepper KA, Kohn DB (2003) The Moloney murine leukemia virus repressor binding site represses expression in murine and human hematopoietic stem cells. *Journal of virology* **77:** 9439-9450

Haig D, Graham C (1991) Genomic imprinting and the strange case of the insulin-like growth factor II receptor. *Cell* **64:** 1045-1046

Han H, Cortez CC, Yang X, Nichols PW, Jones PA, Liang G (2011) DNA methylation directly silences genes with non-CpG island promoters and establishes a nucleosome occupied promoter. *Human molecular genetics* **20:** 4299-4310

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. *Cell* **144:** 646-674

Hansen RS, Wijmenga C, Luo P, Stanek AM, Canfield TK, Weemaes CM, Gartler SM (1999) The DNMT3B DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proc Natl Acad Sci U S A* **96:** 14412-14417

Hashimoto H, Horton JR, Zhang X, Bostick M, Jacobsen SE, Cheng X (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature* **455:** 826-829

Hata K, Okano M, Lei H, Li E (2002) Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* **129:** 1983-1993

He J, Yang Q, Chang LJ (2005) Dynamic DNA methylation and histone modifications contribute to lentiviral transgene silencing in murine embryonic carcinoma cells. *Journal of virology* **79:** 13497-13508

Heard E, Rougeulle C, Arnaud D, Avner P, Allis CD, Spector DL (2001) Methylation of histone H3 at Lys-9 is an early mark on the X chromosome during X inactivation. *Cell* **107:** 727-738

Hendrich B, Bird A (1998) Identification and characterization of a family of mammalian methyl-CpG binding proteins. *Molecular and cellular biology* **18:** 6538-6547

Hendrich B, Hardeland U, Ng HH, Jiricny J, Bird A (1999) The thymine glycosylase MBD4 can bind to the product of deamination at methylated CpG sites. *Nature* **401:** 301-304

Hermann A, Schmitt S, Jeltsch A (2003) The human Dnmt2 has residual DNA-(cytosine-C5) methyltransferase activity. *J Biol Chem* **278:** 31717-31721

Hirasawa R, Chiba H, Kaneda M, Tajima S, Li E, Jaenisch R, Sasaki H (2008) Maternal and zygotic Dnmt1 are necessary and sufficient for the maintenance of DNA methylation imprints during preimplantation development. *Genes Dev* **22:** 1607-1616

Hirota T, Lipp JJ, Toh BH, Peters JM (2005) Histone H3 serine 10 phosphorylation by Aurora B causes HP1 dissociation from heterochromatin. *Nature* **438:** 1176-1180

Ho L, Crabtree GR (2010) Chromatin remodelling during development. *Nature* **463:** 474-484

Ho L, Jothi R, Ronan JL, Cui K, Zhao K, Crabtree GR (2009a) An embryonic stem cell chromatin remodeling complex, esBAF, is an essential component of the core pluripotency transcriptional network. *Proc Natl Acad Sci U S A* **106:** 5187-5191

Ho L, Ronan JL, Wu J, Staahl BT, Chen L, Kuo A, Lessard J, Nesvizhskii AI, Ranish J, Crabtree GR (2009b) An embryonic stem cell chromatin remodeling complex, esBAF, is essential for embryonic stem cell self-renewal and pluripotency. *Proc Natl Acad Sci U S A* **106:** 5181-5186

Hong YK, Ontiveros SD, Chen C, Strauss WM (1999) A new structure for the murine Xist gene and its relationship to chromosome choice/counting during X-chromosome inactivation. *Proc Natl Acad Sci U S A* **96:** 6829-6834

Horowitz RA, Agard DA, Sedat JW, Woodcock CL (1994) The three-dimensional architecture of chromatin in situ: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. *J Cell Biol* **125:** 1-10

Hota SK, Bartholomew B (2011) Diversity of operation in ATP-dependent chromatin remodelers. *Biochimica et biophysica acta* **1809:** 476-487

Howell CY, Bestor TH, Ding F, Latham KE, Mertineit C, Trasler JM, Chaillet JR (2001) Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* **104:** 829-838

Huang J, Fan T, Yan Q, Zhu H, Fox S, Issaq HJ, Best L, Gangi L, Munroe D, Muegge K (2004) Lsh, an epigenetic guardian of repetitive elements. *Nucleic Acids Res* **32:** 5019-5028

Ideraabdullah FY, Vigneau S, Bartolomei MS (2008) Genomic imprinting mechanisms in mammals. *Mutation research* **647:** 77-85

Iida T, Suetake I, Tajima S, Morioka H, Ohta S, Obuse C, Tsurimoto T (2002) PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. *Genes to cells : devoted to molecular & cellular mechanisms* **7:** 997-1007

Ikegami K, Iwatani M, Suzuki M, Tachibana M, Shinkai Y, Tanaka S, Greally JM, Yagi S, Hattori N, Shiota K (2007) Genome-wide and locus-specific DNA hypomethylation in G9a deficient mouse embryonic stem cells. *Genes to cells : devoted to molecular & cellular mechanisms* **12:** 1-11

Illingworth R, Kerr A, Desousa D, Jorgensen H, Ellis P, Stalker J, Jackson D, Clee C, Plumb R, Rogers J, Humphray S, Cox T, Langford C, Bird A (2008) A novel CpG island set identifies tissue-specific methylation at developmental gene loci. *PLoS biology* **6:** e22

Illingworth RS, Gruenewald-Schneider U, Webb S, Kerr AR, James KD, Turner DJ, Smith C, Harrison DJ, Andrews R, Bird AP (2010) Orphan CpG islands identify numerous conserved promoters in the mammalian genome. *PLoS genetics* **6**

Irizarry RA, Ladd-Acosta C, Wen B, Wu Z, Montano C, Onyango P, Cui H, Gabo K, Rongione M, Webster M, Ji H, Potash JB, Sabunciyan S, Feinberg AP (2009) The human colon cancer methylome shows similar hypo- and hypermethylation at conserved tissue-specific CpG island shores. *Nature genetics* **41:** 178-186

Isagawa T, Nagae G, Shiraki N, Fujita T, Sato N, Ishikawa S, Kume S, Aburatani H (2011) DNA methylation profiling of embryonic stem cell differentiation into the three germ layers. *PloS one* **6:** e26052

Jarvis CD, Geiman T, Vila-Storm MP, Osipovich O, Akella U, Candeias S, Nathan I, Durum SK, Muegge K (1996) A novel putative helicase produced in early murine lymphocytes. *Gene* **169:** 203-207

Jeddeloh JA, Stokes TL, Richards EJ (1999) Maintenance of genomic methylation requires a SWI2/SNF2-like protein. *Nature genetics* **22:** 94-97

Jeltsch A (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. *Chembiochem : a European journal of chemical biology* **3:** 274-293

Jeon Y, Sarma K, Lee JT (2012) New and Xisting regulatory mechanisms of X chromosome inactivation. *Curr Opin Genet Dev* **22:** 62-71

Johnson LM, Bostick M, Zhang X, Kraft E, Henderson I, Callis J, Jacobsen SE (2007) The SRA methyl-cytosine-binding domain links DNA and histone methylation. *Curr Biol* **17:** 379-384

Jonkers I, Barakat TS, Achame EM, Monkhorst K, Kenter A, Rentmeester E, Grosveld F, Grootegoed JA, Gribnau J (2009) RNF12 is an X-Encoded dose-dependent activator of X chromosome inactivation. *Cell* **139:** 999-1011

Jonkers I, Monkhorst K, Rentmeester E, Grootegoed JA, Grosveld F, Gribnau J (2008) Xist RNA is confined to the nuclear territory of the silenced X chromosome throughout the cell cycle. *Molecular and cellular biology* **28:** 5583-5594

Jonsson ZO, Jha S, Wohlschlegel JA, Dutta A (2004) Rvb1p/Rvb2p recruit Arp5p and assemble a functional Ino80 chromatin remodeling complex. *Mol Cell* **16:** 465-477

Jorgensen HF, Ben-Porath I, Bird AP (2004) Mbd1 is recruited to both methylated and nonmethylated CpGs via distinct DNA binding domains. *Molecular and cellular biology* **24:** 3387-3395

Joshi AA, Struhl K (2005) Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* **20:** 971-978

Jurkowska RZ, Jurkowski TP, Jeltsch A (2011) Structure and function of mammalian DNA methyltransferases. *Chembiochem : a European journal of chemical biology* **12:** 206-222

Jurkowski TP, Meusburger M, Phalke S, Helm M, Nellen W, Reuter G, Jeltsch A (2008) Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism. *Rna* **14:** 1663-1670

Kageyama S, Liu H, Kaneko N, Ooga M, Nagata M, Aoki F (2007) Alterations in epigenetic modifications during oocyte growth in mice. *Reproduction* **133:** 85-94

Kaneda M, Hirasawa R, Chiba H, Okano M, Li E, Sasaki H (2010) Genetic evidence for Dnmt3a-dependent imprinting during oocyte growth obtained by conditional knockout with Zp3-Cre and complete exclusion of Dnmt3b by chimera formation. *Genes to cells : devoted to molecular & cellular mechanisms*

Kaneda M, Okano M, Hata K, Sado T, Tsujimoto N, Li E, Sasaki H (2004) Essential role for de novo DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* **429:** 900-903

Kanhere A, Viiri K, Araujo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, Pombo A, Fisher AG, Young RA, Jenner RG (2010) Short RNAs are transcribed from repressed polycomb target genes and interact with polycomb repressive complex-2. *Mol Cell* **38:** 675-688

Karagianni P, Amazit L, Qin J, Wong J (2008) ICBP90, a novel methyl K9 H3 binding protein linking protein ubiquitination with heterochromatin formation. *Molecular and cellular biology* **28:** 705-717

Kassabov SR, Zhang B, Persinger J, Bartholomew B (2003) SWI/SNF unwraps, slides, and rewraps the nucleosome. *Mol Cell* **11:** 391-403

Kawasaki H, Kosugi I, Arai Y, Iwashita T, Tsutsui Y (2011) Mouse embryonic stem cells inhibit murine cytomegalovirus infection through a multi-step process. *PloS one* **6:** e17492

Keogh MC, Kurdistani SK, Morris SA, Ahn SH, Podolny V, Collins SR, Schuldiner M, Chin K, Punna T, Thompson NJ, Boone C, Emili A, Weissman JS, Hughes TR, Strahl BD, Grunstein M, Greenblatt JF, Buratowski S, Krogan NJ (2005) Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* **123:** 593-605

Keohane AM, O'Neill L P, Belyaev ND, Lavender JS, Turner BM (1996) X-Inactivation and histone H4 acetylation in embryonic stem cells. *Developmental biology* **180:** 618-630

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* **106:** 11667-11672

Kirmizis A, Bartley SM, Kuzmichev A, Margueron R, Reinberg D, Green R, Farnham PJ (2004) Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes Dev* **18:** 1592-1605

Kizer KO, Phatnani HP, Shibata Y, Hall H, Greenleaf AL, Strahl BD (2005) A novel domain in Set2 mediates RNA polymerase II interaction and couples histone H3 K36 methylation with transcript elongation. *Molecular and cellular biology* **25:** 3305-3316

Kobayashi H, Sakurai T, Imai M, Takahashi N, Fukuda A, Yayoi O, Sato S, Nakabayashi K, Hata K, Sotomaru Y, Suzuki Y, Kono T (2012) Contribution of intragenic DNA methylation in mouse gametic DNA methylomes to establish oocyte-specific heritable marks. *PLoS genetics* **8:** e1002440

Kondo Y, Shen L, Ahmed S, Boumber Y, Sekido Y, Haddad BR, Issa JP (2008) Downregulation of histone H3 lysine 9 methyltransferase G9a induces centrosome disruption and chromosome instability in cancer cells. *PloS one* **3:** e2037

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

Kron K, Pethe V, Briollais L, Sadikovic B, Ozcelik H, Sunderji A, Venkateswaran V, Pinthus J, Fleshner N, van der Kwast T, Bapat B (2009) Discovery of novel hypermethylated genes in prostate cancer using genomic CpG island microarrays. *PloS one* **4:** e4830

Kurihara Y, Kawamura Y, Uchijima Y, Amamo T, Kobayashi H, Asano T, Kurihara H (2008) Maintenance of genomic methylation patterns during preimplantation development requires the somatic form of DNA methyltransferase 1. *Developmental biology* **313:** 335-346

Lagarou A, Mohd-Sarip A, Moshkin YM, Chalkley GE, Bezstarosti K, Demmers JA, Verrijzer CP (2008) dKDM2 couples histone H2A ubiquitylation to histone H3 demethylation during Polycomb group silencing. *Genes Dev* **22:** 2799-2810

Larsen F, Gundersen G, Lopez R, Prydz H (1992) CpG islands as gene markers in the human genome. *Genomics* **13:** 1095-1107

Laurent L, Wong E, Li G, Huynh T, Tsirigos A, Ong CT, Low HM, Kin Sung KW, Rigoutsos I, Loring J, Wei CL (2010) Dynamic changes in the human methylome during differentiation. *Genome research* **20:** 320-331

Lechner MS, Schultz DC, Negorev D, Maul GG, Rauscher FJ, 3rd (2005) The mammalian heterochromatin protein 1 binds diverse nuclear proteins through a common motif that

targets the chromoshadow domain. *Biochemical and biophysical research communications* **331:** 929-937

Lee DW, Zhang K, Ning ZQ, Raabe EH, Tintner S, Wieland R, Wilkins BJ, Kim JM, Blough RI, Arceci RJ (2000) Proliferation-associated SNF2-like gene (PASG): a SNF2 family member altered in leukemia. *Cancer research* **60:** 3612-3622

Lee DY, Northrop JP, Kuo MH, Stallcup MR (2006) Histone H3 lysine 9 methyltransferase G9a is a transcriptional coactivator for nuclear receptors. *J Biol Chem* **281:** 8476-8485

Lee JT (2011) Gracefully ageing at 50, X-chromosome inactivation becomes a paradigm for RNA and chromatin control. *Nature reviews Molecular cell biology* **12:** 815-826

Lee JT, Davidow LS, Warshawsky D (1999) Tsix, a gene antisense to Xist at the X-inactivation centre. *Nature genetics* **21:** 400-404

Lee JT, Lu N (1999) Targeted mutagenesis of Tsix leads to nonrandom X inactivation. *Cell* **99:** 47-57

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

Lehnertz B, Northrop JP, Antignano F, Burrows K, Hadidi S, Mullaly SC, Rossi FM, Zaph C (2010) Activating and inhibitory functions for the histone lysine methyltransferase G9a in T helper cell differentiation and function. *The Journal of experimental medicine* **207:** 915-922

Lehnertz B, Ueda Y, Derijck AA, Braunschweig U, Perez-Burgos L, Kubicek S, Chen T, Li E, Jenuwein T, Peters AH (2003) Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Curr Biol* **13:** 1192-1200

Leonhardt H, Page AW, Weier HU, Bestor TH (1992) A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. *Cell* **71:** 865-873

Leung DC, Dong KB, Maksakova IA, Goyal P, Appanah R, Lee S, Tachibana M, Shinkai Y, Lehnertz B, Mager DL, Rossi F, Lorincz MC (2011) Lysine methyltransferase G9a is required for de novo DNA methylation and the establishment, but not the maintenance, of proviral silencing. *Proc Natl Acad Sci U S A* **108:** 5718-5723

Li E, Beard C, Jaenisch R (1993) Role for DNA methylation in genomic imprinting. *Nature* **366:** 362-365

Li E, Bestor TH, Jaenisch R (1992) Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* **69:** 915-926

Li L, Keverne EB, Aparicio SA, Ishino F, Barton SC, Surani MA (1999) Regulation of maternal behavior and offspring growth by paternally expressed Peg3. *Science* **284:** 330-333

Li SS, Yu SL, Singh S (2010) Epigenetic states and expression of imprinted genes in human embryonic stem cells. *World journal of stem cells* **2:** 97-102

Li X, Ito M, Zhou F, Youngson N, Zuo X, Leder P, Ferguson-Smith AC (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. *Developmental cell* **15:** 547-557

Liang G, Chan MF, Tomigahara Y, Tsai YC, Gonzales FA, Li E, Laird PW, Jones PA (2002) Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. *Molecular and cellular biology* **22:** 480-491

Liang G, Lin JC, Wei V, Yoo C, Cheng JC, Nguyen CT, Weisenberger DJ, Egger G, Takai D, Gonzales FA, Jones PA (2004) Distinct localization of histone H3 acetylation and H3-K4 methylation to the transcription start sites in the human genome. *Proc Natl Acad Sci U S A* **101:** 7357-7362

Lienert F, Wirbelauer C, Som I, Dean A, Mohn F, Schubeler D (2011) Identification of genetic elements that autonomously determine DNA methylation states. *Nature genetics* **43:** 1091-1097

Lin SP, Youngson N, Takada S, Seitz H, Reik W, Paulsen M, Cavaille J, Ferguson-Smith AC (2003) Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the Dlk1-Gtl2 imprinted cluster on mouse chromosome 12. *Nature genetics* **35:** 97-102

Lock LF, Takagi N, Martin GR (1987) Methylation of the Hprt gene on the inactive X occurs after chromosome inactivation. *Cell* **48:** 39-46

Lorincz MC, Schubeler D, Groudine M (2001) Methylation-mediated proviral silencing is associated with MeCP2 recruitment and localized histone H3 deacetylation. *Molecular and cellular biology* **21:** 7913-7922

Lorkovic ZJ, Naumann U, Matzke AJ, Matzke M (2012) Involvement of a GHKL ATPase in RNA-directed DNA methylation in Arabidopsis thaliana. *Curr Biol* **22:** 933-938

Losada A, Hirano T (2005) Dynamic molecular linkers of the genome: the first decade of SMC proteins. *Genes Dev* **19:** 1269-1287

Lu X, Klonoski JM, Resch MG, Hansen JC (2006) In vitro chromatin self-association and its relevance to genome architecture. *Biochem Cell Biol* **84:** 411-417

Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* **389:** 251-260

Maclean JA, 2nd, Chen MA, Wayne CM, Bruce SR, Rao M, Meistrich ML, Macleod C, Wilkinson MF (2005) Rhox: a new homeobox gene cluster. *Cell* **120:** 369-382

Malone CS, Miner MD, Doerr JR, Jackson JP, Jacobsen SE, Wall R, Teitell M (2001) CmC(A/T)GG DNA methylation in mature B cell lymphoma gene silencing. *Proc Natl Acad Sci U S A* **98:** 10404-10409

Mancini-Dinardo D, Steele SJ, Levorse JM, Ingram RS, Tilghman SM (2006) Elongation of the Kcnq1ot1 transcript is required for genomic imprinting of neighboring genes. *Genes Dev* **20:** 1268-1282

Margot JB, Cardoso MC, Leonhardt H (2001) Mammalian DNA methyltransferases show different subnuclear distributions. *Journal of cellular biochemistry* **83:** 373-379

McGrath J, Solter D (1984) Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell* **37:** 179-183

Meissner A (2011) Guiding DNA methylation. *Cell stem cell* **9:** 388-390

Meissner A, Mikkelsen TS, Gu H, Wernig M, Hanna J, Sivachenko A, Zhang X, Bernstein BE, Nusbaum C, Jaffe DB, Gnirke A, Jaenisch R, Lander ES (2008) Genome-scale DNA methylation maps of pluripotent and differentiated cells. *Nature* **454:** 766-770

Melamed-Bessudo C, Levy AA (2012) Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in Arabidopsis. *Proc Natl Acad Sci U S A* **109:** E981-988

Messerschmidt DM (2012) Should I stay or should I go: Protection and maintenance of DNA methylation at imprinted genes. *Epigenetics* **7**

Mikkelsen TS, Ku M, Jaffe DB, Issac B, Lieberman E, Giannoukos G, Alvarez P, Brockman W, Kim TK, Koche RP, Lee W, Mendenhall E, O'Donovan A, Presser A, Russ C, Xie X, Meissner A, Wernig M, Jaenisch R, Nusbaum C, Lander ES, Bernstein BE (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **448:** 553-560

Millar CB, Guy J, Sansom OJ, Selfridge J, MacDougall E, Hendrich B, Keightley PD, Bishop SM, Clarke AR, Bird A (2002) Enhanced CpG mutability and tumorigenesis in MBD4-deficient mice. *Science* **297:** 403-405

Mittelsten Scheid O, Paszkowski J (2000) Transcriptional gene silencing mutants. *Plant molecular biology* **43:** 235-241

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

Morey L, Helin K (2010) Polycomb group protein-mediated repression of transcription. *Trends Biochem Sci* **35:** 323-332

Mori Y, Cai K, Cheng Y, Wang S, Paun B, Hamilton JP, Jin Z, Sato F, Berki AT, Kan T, Ito T, Mantzur C, Abraham JM, Meltzer SJ (2006) A genome-wide search identifies epigenetic silencing of somatostatin, tachykinin-1, and 5 other genes in colon cancer. *Gastroenterology* **131:** 797-808

Morton JP, Myant KB, Sansom OJ (2011) A FAK-PI-3K-mTOR axis is required for Wnt-Myc driven intestinal regeneration and tumorigenesis. *Cell Cycle* **10:** 173-175

Muller-Tidow C, Klein HU, Hascher A, Isken F, Tickenbrock L, Thoennissen N, Agrawal-Singh S, Tschanter P, Disselhoff C, Wang Y, Becker A, Thiede C, Ehninger G, zur Stadt U, Koschmieder S, Seidl M, Muller FU, Schmitz W, Schlenke P, McClelland M, Berdel WE, Dugas M, Serve H, Study Alliance L (2010) Profiling of histone H3 lysine 9 trimethylation levels predicts transcription factor activity and survival in acute myeloid leukemia. *Blood* **116:** 3564-3571

Munshi A, Shafi G, Aliya N, Jyothy A (2009) Histone modifications dictate specific biological readouts. *Journal of genetics and genomics = Yi chuan xue bao* **36:** 75-88

Myant K, Sansom O (2011) Efficient Wnt mediated intestinal hyperproliferation requires the cyclin D2-CDK4/6 complex. *Cell division* **6:** 3

Myant K, Stancheva I (2008) LSH cooperates with DNA methyltransferases to repress transcription. *Molecular and cellular biology* **28:** 215-226

Myant K, Termanis A, Sundaram AY, Boe T, Li C, Merusi C, Burrage J, de Las Heras JI, Stancheva I (2011) LSH and G9a/GLP complex are required for developmentally programmed DNA methylation. *Genome research* **21:** 83-94

Nady N, Lemak A, Walker JR, Avvakumov GV, Kareta MS, Achour M, Xue S, Duan S, Allali-Hassani A, Zuo X, Wang YX, Bronner C, Chedin F, Arrowsmith CH, Dhe-Paganon S (2011) Recognition of multivalent histone states associated with heterochromatin by UHRF1 protein. *J Biol Chem* **286:** 24300-24311

Nakamura T, Arai Y, Umehara H, Masuhara M, Kimura T, Taniguchi H, Sekimoto T, Ikawa M, Yoneda Y, Okabe M, Tanaka S, Shiota K, Nakano T (2007) PGC7/Stella protects against DNA demethylation in early embryogenesis. *Nat Cell Biol* **9:** 64-71

Nakamura T, Liu YJ, Nakashima H, Umehara H, Inoue K, Matoba S, Tachibana M, Ogura A, Shinkai Y, Nakano T (2012) PGC7 binds histone H3K9me2 to protect against conversion of 5mC to 5hmC in early embryos. *Nature* **486:** 415-419

Navarro P, Chambers I, Karwacki-Neisius V, Chureau C, Morey C, Rougeulle C, Avner P (2008) Molecular coupling of Xist regulation and pluripotency. *Science* **321:** 1693-1695

Navarro P, Oldfield A, Legoupi J, Festuccia N, Dubois A, Attia M, Schoorlemmer J, Rougeulle C, Chambers I, Avner P (2010) Molecular coupling of Tsix regulation and pluripotency. *Nature* **468:** 457-460

Nimura K, Ishida C, Koriyama H, Hata K, Yamanaka S, Li E, Ura K, Kaneda Y (2006) Dnmt3a2 targets endogenous Dnmt3L to ES cell chromatin and induces regional DNA methylation. *Genes to cells : devoted to molecular & cellular mechanisms* **11:** 1225-1237

Nishino Y, Eltsov M, Joti Y, Ito K, Takata H, Takahashi Y, Hihara S, Frangakis AS, Imamoto N, Ishikawa T, Maeshima K (2012) Human mitotic chromosomes consist predominantly of irregularly folded nucleosome fibres without a 30-nm chromatin structure. *EMBO J* **31:** 1644-1653

Nishioka K, Chuikov S, Sarma K, Erdjument-Bromage H, Allis CD, Tempst P, Reinberg D (2002) Set9, a novel histone H3 methyltransferase that facilitates transcription by precluding histone tail modifications required for heterochromatin formation. *Genes Dev* **16:** 479-489

Oberle I, Rousseau F, Heitz D, Kretz C, Devys D, Hanauer A, Boue J, Bertheas MF, Mandel JL (1991) Instability of a 550-base pair DNA segment and abnormal methylation in fragile X syndrome. *Science* **252:** 1097-1102

Oda M, Yamagiwa A, Yamamoto S, Nakayama T, Tsumura A, Sasaki H, Nakao K, Li E, Okano M (2006) DNA methylation regulates long-range gene silencing of an X-linked homeobox gene cluster in a lineage-specific manner. *Genes Dev* **20:** 3382-3394

Ogawa Y, Lee JT (2003) Xite, X-inactivation intergenic transcription elements that regulate the probability of choice. *Mol Cell* **11:** 731-743

Ohta T, Gray TA, Rogan PK, Buiting K, Gabriel JM, Saitoh S, Muralidhar B, Bilienska B, Krajewska-Walasek M, Driscoll DJ, Horsthemke B, Butler MG, Nicholls RD (1999) Imprinting-mutation mechanisms in Prader-Willi syndrome. *American journal of human genetics* **64:** 397-413

Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* **99:** 247-257

Okano M, Xie S, Li E (1998a) Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature genetics* **19:** 219-220

Okano M, Xie S, Li E (1998b) Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Res* **26:** 2536-2540

Okitsu CY, Hsieh CL (2007) DNA methylation dictates histone H3K4 methylation. *Molecular and cellular biology* **27:** 2746-2757

Ooi SK, Qiu C, Bernstein E, Li K, Jia D, Yang Z, Erdjument-Bromage H, Tempst P, Lin SP, Allis CD, Cheng X, Bestor TH (2007) DNMT3L connects unmethylated lysine 4 of histone H3 to de novo methylation of DNA. *Nature* **448:** 714-717

Ooi SK, Wolf D, Hartung O, Agarwal S, Daley GQ, Goff SP, Bestor TH (2010) Dynamic instability of genomic methylation patterns in pluripotent stem cells. *Epigenetics & chromatin* **3:** 17

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

Perez-Torrado R, Yamada D, Defossez PA (2006) Born to bind: the BTB protein-protein interaction domain. *BioEssays : news and reviews in molecular, cellular and developmental biology* **28:** 1194-1202

Perisic O, Collepardo-Guevara R, Schlick T (2010) Modeling studies of chromatin fiber structure as a function of DNA linker length. *Journal of molecular biology* **403:** 777-802

Perpelescu M, Nozaki N, Obuse C, Yang H, Yoda K (2009) Active establishment of centromeric CENP-A chromatin by RSF complex. *J Cell Biol* **185:** 397-407

Peters AH, Kubicek S, Mechtler K, O'Sullivan RJ, Derijck AA, Perez-Burgos L, Kohlmaier A, Opravil S, Tachibana M, Shinkai Y, Martens JH, Jenuwein T (2003) Partitioning and plasticity of repressive histone methylation states in mammalian chromatin. *Mol Cell* **12:** 1577-1589

Picketts DJ, Higgs DR, Bachoo S, Blake DJ, Quarrell OW, Gibbons RJ (1996) ATRX encodes a novel member of the SNF2 family of proteins: mutations point to a common mechanism underlying the ATR-X syndrome. *Human molecular genetics* **5:** 1899-1907

Plusa B, Frankenberg S, Chalmers A, Hadjantonakis AK, Moore CA, Papalopulu N, Papaioannou VE, Glover DM, Zernicka-Goetz M (2005) Downregulation of Par3 and aPKC function directs cells towards the ICM in the preimplantation mouse embryo. *Journal of cell science* **118:** 505-515

Pradhan S, Bacolla A, Wells RD, Roberts RJ (1999) Recombinant human DNA (cytosine-5) methyltransferase. I. Expression, purification, and comparison of de novo and maintenance methylation. *J Biol Chem* **274:** 33002-33010

Prinapori R, Guinaud J, Khalil A, Lecuyer H, Gendrel D, Lortholary O, Nassif X, Viscoli C, Zahar JR (2012) Risk associated with a systematic search of extended-spectrum beta-lactamase-producing Enterobacteriaceae. *American journal of infection control*

Purcell DJ, Jeong KW, Bittencourt D, Gerke DS, Stallcup MR (2011) A distinct mechanism for coactivator versus corepressor function by histone methyltransferase G9a in transcriptional regulation. *J Biol Chem* **286:** 41963-41971

Qin C, Wang Z, Shang J, Bekkari K, Liu R, Pacchione S, McNulty KA, Ng A, Barnum JE, Storer RD (2010) Intracisternal A particle genes: Distribution in the mouse genome, active subtypes, and potential roles as species-specific mediators of susceptibility to cancer. *Molecular carcinogenesis* **49:** 54-67

Quenneville S, Verde G, Corsinotti A, Kapopoulou A, Jakobsson J, Offner S, Baglivo I, Pedone PV, Grimaldi G, Riccio A, Trono D (2011) In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. *Mol Cell* **44:** 361-372

Raabe EH, Abdurrahman L, Behbehani G, Arceci RJ (2001) An SNF2 factor involved in mammalian development and cellular proliferation. *Dev Dyn* **221:** 92-105

Racki LR, Yang JG, Naber N, Partensky PD, Acevedo A, Purcell TJ, Cooke R, Cheng Y, Narlikar GJ (2009) The chromatin remodeller ACF acts as a dimeric motor to space nucleosomes. *Nature* **462:** 1016-1021

Ramsahoye BH (2002) Measurement of genome-wide DNA cytosine-5 methylation by reversed-phase high-pressure liquid chromatography. *Methods in molecular biology* **200:** 17-27

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP, Jaenisch R (2000) Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. *Proc Natl Acad Sci U S A* **97:** 5237-5242

Rea S, Eisenhaber F, O'Carroll D, Strahl BD, Sun ZW, Schmid M, Opravil S, Mechtler K, Ponting CP, Allis CD, Jenuwein T (2000) Regulation of chromatin structure by site-specific histone H3 methyltransferases. *Nature* **406:** 593-599

Redrup L, Branco MR, Perdeaux ER, Krueger C, Lewis A, Santos F, Nagano T, Cobb BS, Fraser P, Reik W (2009) The long noncoding RNA Kcnq1ot1 organises a lineage-specific nuclear domain for epigenetic gene silencing. *Development* **136:** 525-530

Reik W, Constancia M, Fowden A, Anderson N, Dean W, Ferguson-Smith A, Tycko B, Sibley C (2003) Regulation of supply and demand for maternal nutrients in mammals by imprinted genes. *The Journal of physiology* **547:** 35-44

Riccio A, Sparago A, Verde G, De Crescenzo A, Citro V, Cubellis MV, Ferrero GB, Silengo MC, Russo S, Larizza L, Cerrato F (2009) Inherited and Sporadic Epimutations at the IGF2-H19 locus in Beckwith-Wiedemann syndrome and Wilms' tumor. *Endocrine development* **14:** 1-9

Rice JC, Briggs SD, Ueberheide B, Barber CM, Shabanowitz J, Hunt DF, Shinkai Y, Allis CD (2003) Histone methyltransferases direct different degrees of methylation to define distinct chromatin domains. *Mol Cell* **12:** 1591-1598

Richardson B, Scheinbart L, Strahler J, Gross L, Hanash S, Johnson M (1990) Evidence for impaired T cell DNA methylation in systemic lupus erythematosus and rheumatoid arthritis. *Arthritis and rheumatism* **33:** 1665-1673

Rideout WM, 3rd, Coetzee GA, Olumi AF, Jones PA (1990) 5-Methylcytosine as an endogenous mutagen in the human LDL receptor and p53 genes. *Science* **249:** 1288-1290

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

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, Goodnough LH, Helms JA, Farnham PJ, Segal E, Chang HY (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* **129:** 1311-1323

Rius M, Lyko F (2012) Epigenetic cancer therapy: rationales, targets and drugs. *Oncogene* **31:** 4257-4265

Robertson KD (2005) DNA methylation and human disease. *Nat Rev Genet* **6:** 597-610

Robertson KD, Keyomarsi K, Gonzales FA, Velicescu M, Jones PA (2000) Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G(0)/G(1) to S phase transition in normal and tumor cells. *Nucleic Acids Res* **28:** 2108-2113

Robertson KD, Uzvolgyi E, Liang G, Talmadge C, Sumegi J, Gonzales FA, Jones PA (1999) The human DNA methyltransferases (DNMTs) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumors. *Nucleic Acids Res* **27:** 2291-2298

Roopra A, Qazi R, Schoenike B, Daley TJ, Morrison JF (2004) Localized domains of G9a-mediated histone methylation are required for silencing of neuronal genes. *Mol Cell* **14:** 727-738

Rossant J, Tam PP (2009) Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* **136:** 701-713

Routh A, Sandin S, Rhodes D (2008) Nucleosome repeat length and linker histone stoichiometry determine chromatin fiber structure. *Proc Natl Acad Sci U S A* **105:** 8872-8877

Ryu B, Kim DS, Deluca AM, Alani RM (2007) Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. *PloS one* **2:** e594

Sado T, Fenner MH, Tan SS, Tam P, Shioda T, Li E (2000) X inactivation in the mouse embryo deficient for Dnmt1: distinct effect of hypomethylation on imprinted and random X inactivation. *Developmental biology* **225:** 294-303

Sado T, Okano M, Li E, Sasaki H (2004) De novo DNA methylation is dispensable for the initiation and propagation of X chromosome inactivation. *Development* **131:** 975-982

Saha A, Wittmeyer J, Cairns BR (2002) Chromatin remodeling by RSC involves ATP-dependent DNA translocation. *Genes Dev* **16:** 2120-2134

Saha A, Wittmeyer J, Cairns BR (2005) Chromatin remodeling through directional DNA translocation from an internal nucleosomal site. *Nat Struct Mol Biol* **12:** 747-755

Saitou M, Kagiwada S, Kurimoto K (2012) Epigenetic reprogramming in mouse pre-implantation development and primordial germ cells. *Development* **139:** 15-31

Santos F, Hendrich B, Reik W, Dean W (2002) Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental biology* **241:** 172-182

Sasai N, Defossez PA (2009) Many paths to one goal? The proteins that recognize methylated DNA in eukaryotes. *The International journal of developmental biology* **53:** 323-334

Sasai N, Nakao M, Defossez PA (2010) Sequence-specific recognition of methylated DNA by human zinc-finger proteins. *Nucleic Acids Res* **38:** 5015-5022

Sato M, Kimura T, Kurokawa K, Fujita Y, Abe K, Masuhara M, Yasunaga T, Ryo A, Yamamoto M, Nakano T (2002) Identification of PGC7, a new gene expressed specifically in preimplantation embryos and germ cells. *Mechanisms of development* **113:** 91-94

Saxonov S, Berg P, Brutlag DL (2006) A genome-wide analysis of CpG dinucleotides in the human genome distinguishes two distinct classes of promoters. *Proc Natl Acad Sci U S A* **103:** 1412-1417

Schaniel C, Ang YS, Ratnakumar K, Cormier C, James T, Bernstein E, Lemischka IR, Paddison PJ (2009) Smarcc1/Baf155 couples self-renewal gene repression with changes in chromatin structure in mouse embryonic stem cells. *Stem cells* **27:** 2979-2991

Schotta G, Lachner M, Sarma K, Ebert A, Sengupta R, Reuter G, Reinberg D, Jenuwein T (2004) A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* **18:** 1251-1262

Schultz DC, Ayyanathan K, Negorev D, Maul GG, Rauscher FJ, 3rd (2002) SETDB1: a novel KAP-1-associated histone H3, lysine 9-specific methyltransferase that contributes to HP1-mediated silencing of euchromatic genes by KRAB zinc-finger proteins. *Genes Dev* **16:** 919-932

Seligson DB, Horvath S, Shi T, Yu H, Tze S, Grunstein M, Kurdistani SK (2005) Global histone modification patterns predict risk of prostate cancer recurrence. *Nature* **435:** 1262-1266

Shaked H, Avivi-Ragolsky N, Levy AA (2006) Involvement of the Arabidopsis SWI2/SNF2 chromatin remodeling gene family in DNA damage response and recombination. *Genetics* **173:** 985-994

Sharif J, Muto M, Takebayashi S, Suetake I, Iwamatsu A, Endo TA, Shinga J, Mizutani-Koseki Y, Toyoda T, Okamura K, Tajima S, Mitsuya K, Okano M, Koseki H (2007) The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* **450:** 908-912

Sharma S, Gerke DS, Han HF, Jeong S, Stallcup MR, Jones PA, Liang G (2012) Lysine methyltransferase G9a is not required for DNMT3A/3B anchoring to methylated nucleosomes and maintenance of DNA methylation in somatic cells. *Epigenetics & chromatin* **5:** 3

Shen X, Ranallo R, Choi E, Wu C (2003) Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol Cell* **12:** 147-155

Shinkai Y, Tachibana M (2011) H3K9 methyltransferase G9a and the related molecule GLP. *Genes Dev* **25:** 781-788

Shogren-Knaak M, Ishii H, Sun JM, Pazin MJ, Davie JR, Peterson CL (2006) Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311:** 844-847

Sing A, Pannell D, Karaiskakis A, Sturgeon K, Djabali M, Ellis J, Lipshitz HD, Cordes SP (2009) A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* **138:** 885-897

Smallwood SA, Tomizawa S, Krueger F, Ruf N, Carli N, Segonds-Pichon A, Sato S, Hata K, Andrews SR, Kelsey G (2011) Dynamic CpG island methylation landscape in oocytes and preimplantation embryos. *Nature genetics* **43:** 811-814

Smiraglia DJ, Kazhiyur-Mannar R, Oakes CC, Wu YZ, Liang P, Ansari T, Su J, Rush LJ, Smith LT, Yu L, Liu C, Dai Z, Chen SS, Wang SH, Costello J, Ioshikhes I, Dawson DW, Hong JS, Teitell MA, Szafranek A, Camoriano M, Song F, Elliott R, Held W, Trasler JM, Plass C, Wenger R (2007) Restriction landmark genomic scanning (RLGS) spot identification by second generation virtual RLGS in multiple genomes with multiple enzyme combinations. *BMC genomics* **8:** 446

Smith FM, Garfield AS, Ward A (2006) Regulation of growth and metabolism by imprinted genes. *Cytogenetic and genome research* **113:** 279-291

Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, Meissner A (2012) A unique regulatory phase of DNA methylation in the early mammalian embryo. *Nature* **484:** 339-344

Sorge F, Gendrel D, le Groupe de pediatrie t (2013) [Counselling child traveller]. *Archives de pediatrie : organe officiel de la Societe francaise de pediatrie* **20:** 95-99

Soubry A, Staes K, Parthoens E, Noppen S, Stove C, Bogaert P, van Hengel J, van Roy F (2010) The transcriptional repressor Kaiso localizes at the mitotic spindle and is a constituent of the pericentriolar material. *PloS one* **5:** e9203

Stockdale C, Flaus A, Ferreira H, Owen-Hughes T (2006) Analysis of nucleosome repositioning by yeast ISWI and Chd1 chromatin remodeling complexes. *J Biol Chem* **281:** 16279-16288

Strogantsev R, Ferguson-Smith AC (2012) Proteins involved in establishment and maintenance of imprinted methylation marks. *Brief Funct Genomics* **11:** 227-239

Sun LQ, Lee DW, Zhang Q, Xiao W, Raabe EH, Meeker A, Miao D, Huso DL, Arceci RJ (2004) Growth retardation and premature aging phenotypes in mice with disruption of the SNF2-like gene, PASG. *Genes Dev* **18:** 1035-1046

Surani MA, Barton SC, Norris ML (1984) Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* **308:** 548-550

Suzuki T, Farrar JE, Yegnasubramanian S, Zahed M, Suzuki N, Arceci RJ (2008) Stable knockdown of PASG enhances DNA demethylation but does not accelerate cellular senescence in TIG-7 human fibroblasts. *Epigenetics* **3:** 281-291

Swindle CS, Klug CA (2002) Mechanisms that regulate silencing of gene expression from retroviral vectors. *Journal of hematotherapy & stem cell research* **11:** 449-456

Szerlong H, Hinata K, Viswanathan R, Erdjument-Bromage H, Tempst P, Cairns BR (2008) The HSA domain binds nuclear actin-related proteins to regulate chromatin-remodeling ATPases. *Nat Struct Mol Biol* **15:** 469-476

Szyf M (2001) The role of DNA methyltransferase 1 in growth control. *Frontiers in bioscience : a journal and virtual library* **6:** D599-609

Tachibana M, Sugimoto K, Fukushima T, Shinkai Y (2001) Set domain-containing protein, G9a, is a novel lysine-preferring mammalian histone methyltransferase with hyperactivity and specific selectivity to lysines 9 and 27 of histone H3. *J Biol Chem* **276:** 25309-25317

Tachibana M, Sugimoto K, Nozaki M, Ueda J, Ohta T, Ohki M, Fukuda M, Takeda N, Niida H, Kato H, Shinkai Y (2002) G9a histone methyltransferase plays a dominant role in euchromatic histone H3 lysine 9 methylation and is essential for early embryogenesis. *Genes Dev* **16:** 1779-1791

Tachibana M, Ueda J, Fukuda M, Takeda N, Ohta T, Iwanari H, Sakihama T, Kodama T, Hamakubo T, Shinkai Y (2005) Histone methyltransferases G9a and GLP form heteromeric complexes and are both crucial for methylation of euchromatin at H3-K9. *Genes Dev* **19:** 815-826

Tang LY, Reddy MN, Rasheva V, Lee TL, Lin MJ, Hung MS, Shen CK (2003) The eukaryotic DNMT2 genes encode a new class of cytosine-5 DNA methyltransferases. *J Biol Chem* **278:** 33613-33616

Tao Y, Xi S, Shan J, Maunakea A, Che A, Briones V, Lee EY, Geiman T, Huang J, Stephens R, Leighty RM, Zhao K, Muegge K (2011) Lsh, chromatin remodeling family member, modulates genome-wide cytosine methylation patterns at nonrepeat sequences. *Proc Natl Acad Sci U S A* **108:** 5626-5631

Terzi N, Churchman LS, Vasiljeva L, Weissman J, Buratowski S (2011) H3K4 trimethylation by Set1 promotes efficient termination by the Nrd1-Nab3-Sen1 pathway. *Molecular and cellular biology* **31:** 3569-3583

Tian D, Sun S, Lee JT (2010) The long noncoding RNA, Jpx, is a molecular switch for X chromosome inactivation. *Cell* **143:** 390-403

Tomizawa S, Kobayashi H, Watanabe T, Andrews S, Hata K, Kelsey G, Sasaki H (2011) Dynamic stage-specific changes in imprinted differentially methylated regions during early mammalian development and prevalence of non-CpG methylation in oocytes. *Development* **138:** 811-820

Torok MS, Grant PA (2006) The generation and recognition of histone methylation. *Results and problems in cell differentiation* **41:** 25-46

Trojer P, Reinberg D (2007) Facultative heterochromatin: is there a distinctive molecular signature? *Mol Cell* **28:** 1-13

Trojer P, Zhang J, Yonezawa M, Schmidt A, Zheng H, Jenuwein T, Reinberg D (2009) Dynamic Histone H1 Isotype 4 Methylation and Demethylation by Histone Lysine Methyltransferase G9a/KMT1C and the Jumonji Domain-containing JMJD2/KDM4 Proteins. *J Biol Chem* **284:** 8395-8405

Tryndyak VP, Kovalchuk O, Pogribny IP (2006) Loss of DNA methylation and histone H4 lysine 20 trimethylation in human breast cancer cells is associated with aberrant expression of DNA methyltransferase 1, Suv4-20h2 histone methyltransferase and methyl-binding proteins. *Cancer biology & therapy* **5:** 65-70

Tsai CE, Lin SP, Ito M, Takagi N, Takada S, Ferguson-Smith AC (2002) Genomic imprinting contributes to thyroid hormone metabolism in the mouse embryo. *Curr Biol* **12:** 1221-1226

Tsumura A, Hayakawa T, Kumaki Y, Takebayashi S, Sakaue M, Matsuoka C, Shimotohno K, Ishikawa F, Li E, Ueda HR, Nakayama J, Okano M (2006) Maintenance of self-renewal ability of mouse embryonic stem cells in the absence of DNA methyltransferases Dnmt1, Dnmt3a and Dnmt3b. *Genes to cells : devoted to molecular & cellular mechanisms* **11:** 805-814

Tuck-Muller CM, Narayan A, Tsien F, Smeets DF, Sawyer J, Fiala ES, Sohn OS, Ehrlich M (2000) DNA hypomethylation and unusual chromosome instability in cell lines from ICF syndrome patients. *Cytogenetics and cell genetics* **89:** 121-128

Tudor M, Akbarian S, Chen RZ, Jaenisch R (2002) Transcriptional profiling of a mouse model for Rett syndrome reveals subtle transcriptional changes in the brain. *Proc Natl Acad Sci U S A* **99:** 15536-15541

Udugama M, Sabri A, Bartholomew B (2011) The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Molecular and cellular biology* **31:** 662-673

Ueda J, Tachibana M, Ikura T, Shinkai Y (2006) Zinc finger protein Wiz links G9a/GLP histone methyltransferases to the co-repressor molecule CtBP. *J Biol Chem* **281:** 20120-20128

Uemura T, Kubo E, Kanari Y, Ikemura T, Tatsumi K, Muto M (2000) Temporal and spatial localization of novel nuclear protein NP95 in mitotic and meiotic cells. *Cell structure and function* **25:** 149-159

Underhill C, Qutob MS, Yee SP, Torchia J (2000) A novel nuclear receptor corepressor complex, N-CoR, contains components of the mammalian SWI/SNF complex and the corepressor KAP-1. *J Biol Chem* **275:** 40463-40470

Vakoc CR, Mandat SA, Olenchock BA, Blobel GA (2005) Histone H3 lysine 9 methylation and HP1gamma are associated with transcription elongation through mammalian chromatin. *Mol Cell* **19:** 381-391

Vakoc CR, Sachdeva MM, Wang H, Blobel GA (2006) Profile of histone lysine methylation across transcribed mammalian chromatin. *Molecular and cellular biology* **26:** 9185-9195

Van den Wyngaert I, Sprengel J, Kass SU, Luyten WH (1998) Cloning and analysis of a novel human putative DNA methyltransferase. *FEBS letters* **426:** 283-289

Venkatesh S, Smolle M, Li H, Gogol MM, Saint M, Kumar S, Natarajan K, Workman JL (2012) Set2 methylation of histone H3 lysine 36 suppresses histone exchange on transcribed genes. *Nature*

Vire E, Brenner C, Deplus R, Blanchon L, Fraga M, Didelot C, Morey L, Van Eynde A, Bernard D, Vanderwinden JM, Bollen M, Esteller M, Di Croce L, de Launoit Y, Fuks F (2006) The Polycomb group protein EZH2 directly controls DNA methylation. *Nature* **439:** 871-874

Vongs A, Kakutani T, Martienssen RA, Richards EJ (1993) Arabidopsis thaliana DNA methylation mutants. *Science* **260:** 1926-1928

Wagschal A, Sutherland HG, Woodfine K, Henckel A, Chebli K, Schulz R, Oakey RJ, Bickmore WA, Feil R (2008) G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Molecular and cellular biology* **28:** 1104-1113

Wang C, Shen J, Yang Z, Chen P, Zhao B, Hu W, Lan W, Tong X, Wu H, Li G, Cao C (2011) Structural basis for site-specific reading of unmodified R2 of histone H3 tail by UHRF1 PHD finger. *Cell research* **21:** 1379-1382

Wang J, Hevi S, Kurash JK, Lei H, Gay F, Bajko J, Su H, Sun W, Chang H, Xu G, Gaudet F, Li E, Chen T (2009) The lysine demethylase LSD1 (KDM1) is required for maintenance of global DNA methylation. *Nature genetics* **41:** 125-129

Wang X, Zhang J (2006) Remarkable expansions of an X-linked reproductive homeobox gene cluster in rodent evolution. *Genomics* **88:** 34-43

Wang Y, Jia S (2009) Degrees make all the difference: the multifunctionality of histone H4 lysine 20 methylation. *Epigenetics* **4:** 273-276

Watanabe S, Ichimura T, Fujita N, Tsuruzoe S, Ohki I, Shirakawa M, Kawasuji M, Nakao M (2003) Methylated DNA-binding domain 1 and methylpurine-DNA glycosylase link transcriptional repression and DNA repair in chromatin. *Proc Natl Acad Sci U S A* **100:** 12859-12864

Watanabe T, Tomizawa S, Mitsuya K, Totoki Y, Yamamoto Y, Kuramochi-Miyagawa S, Iida N, Hoki Y, Murphy PJ, Toyoda A, Gotoh K, Hiura H, Arima T, Fujiyama A, Sado T, Shibata T, Nakano T, Lin H, Ichiyanagi K, Soloway PD, Sasaki H (2011) Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. *Science* **332:** 848-852

Weber M, Davies JJ, Wittig D, Oakeley EJ, Haase M, Lam WL, Schubeler D (2005) Chromosome-wide and promoter-specific analyses identify sites of differential DNA methylation in normal and transformed human cells. *Nature genetics* **37:** 853-862

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

Weiss T, Hergeth S, Zeissler U, Izzo A, Tropberger P, Zee BM, Dundr M, Garcia BA, Daujat S, Schneider R (2010) Histone H1 variant-specific lysine methylation by G9a/KMT1C and Glp1/KMT1D. *Epigenetics & chromatin* **3:** 7

Whitehouse I, Stockdale C, Flaus A, Szczelkun MD, Owen-Hughes T (2003) Evidence for DNA translocation by the ISWI chromatin-remodeling enzyme. *Molecular and cellular biology* **23:** 1935-1945

Williams SK, Truong D, Tyler JK (2008) Acetylation in the globular core of histone H3 on lysine-56 promotes chromatin disassembly during transcriptional activation. *Proc Natl Acad Sci U S A* **105:** 9000-9005

Wong E, Yang K, Kuraguchi M, Werling U, Avdievich E, Fan K, Fazzari M, Jin B, Brown AM, Lipkin M, Edelmann W (2002) Mbd4 inactivation increases Cright-arrowT transition mutations and promotes gastrointestinal tumor formation. *Proc Natl Acad Sci U S A* **99:** 14937-14942

Woo CJ, Kharchenko PV, Daheron L, Park PJ, Kingston RE (2010) A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* **140:** 99-110

Woodcock CL, Frado LL, Rattner JB (1984) The higher-order structure of chromatin: evidence for a helical ribbon arrangement. *J Cell Biol* **99:** 42-52

Woodcock CL, Sweetman HE, Frado LL (1976) Structural repeating units in chromatin. II. Their isolation and partial characterization. *Exp Cell Res* **97:** 111-119

Wu H, Chen X, Xiong J, Li Y, Li H, Ding X, Liu S, Chen S, Gao S, Zhu B (2011) Histone methyltransferase G9a contributes to H3K27 methylation in vivo. *Cell research* **21:** 365-367

Wutz A, Jaenisch R (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell* **5:** 695-705

Xi S, Zhu H, Xu H, Schmidtmann A, Geiman TM, Muegge K (2007) Lsh controls Hox gene silencing during development. *Proc Natl Acad Sci U S A* **104:** 14366-14371

Xiao A, Li H, Shechter D, Ahn SH, Fabrizio LA, Erdjument-Bromage H, Ishibe-Murakami S, Wang B, Tempst P, Hofmann K, Patel DJ, Elledge SJ, Allis CD (2009) WSTF regulates the H2A.X DNA damage response via a novel tyrosine kinase activity. *Nature* **457:** 57-62

Xie S, Jakoncic J, Qian C (2012) UHRF1 double tudor domain and the adjacent PHD finger act together to recognize K9me3-containing histone H3 tail. *Journal of molecular biology* **415:** 318-328

Xin Z, Tachibana M, Guggiari M, Heard E, Shinkai Y, Wagstaff J (2003) Role of histone methyltransferase G9a in CpG methylation of the Prader-Willi syndrome imprinting center. *J Biol Chem* **278:** 14996-15000

Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ, Viegas-Pequignot E (1999) Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* **402:** 187-191

Xu N, Donohoe ME, Silva SS, Lee JT (2007) Evidence that homologous X-chromosome pairing requires transcription and Ctcf protein. *Nature genetics* **39:** 1390-1396

Yamamizu K, Fujihara M, Tachibana M, Katayama S, Takahashi A, Hara E, Imai H, Shinkai Y, Yamashita JK (2012) Protein kinase A determines timing of early differentiation through epigenetic regulation with G9a. *Cell stem cell* **10:** 759-770

Yan Q, Cho E, Lockett S, Muegge K (2003a) Association of Lsh, a regulator of DNA methylation, with pericentromeric heterochromatin is dependent on intact heterochromatin. *Molecular and cellular biology* **23:** 8416-8428

Yan Q, Huang J, Fan T, Zhu H, Muegge K (2003b) Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. *EMBO J* **22:** 5154-5162

Yao Y, Bilichak A, Golubov A, Kovalchuk I (2012) ddm1 plants are sensitive to methyl methane sulfonate and NaCl stresses and are deficient in DNA repair. *Plant cell reports* **31:** 1549-1561

Yasui DH, Peddada S, Bieda MC, Vallero RO, Hogart A, Nagarajan RP, Thatcher KN, Farnham PJ, Lasalle JM (2007) Integrated epigenomic analyses of neuronal MeCP2 reveal a role for long-range interaction with active genes. *Proc Natl Acad Sci U S A* **104:** 19416-19421

Yoder JA, Bestor TH (1998) A candidate mammalian DNA methyltransferase related to pmt1p of fission yeast. *Human molecular genetics* **7:** 279-284

Yokochi T, Poduch K, Ryba T, Lu J, Hiratani I, Tachibana M, Shinkai Y, Gilbert DM (2009) G9a selectively represses a class of late-replicating genes at the nuclear periphery. *Proc Natl Acad Sci U S A* **106:** 19363-19368

Young JI, Hong EP, Castle JC, Crespo-Barreto J, Bowman AB, Rose MF, Kang D, Richman R, Johnson JM, Berget S, Zoghbi HY (2005) Regulation of RNA splicing by the methylation-dependent transcriptional repressor methyl-CpG binding protein 2. *Proc Natl Acad Sci U S A* **102:** 17551-17558

Yu S, Castle A, Chen M, Lee R, Takeda K, Weinstein LS (2001) Increased insulin sensitivity in Gsalpha knockout mice. *J Biol Chem* **276:** 19994-19998

Zegerman P, Canas B, Pappin D, Kouzarides T (2002) Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. *J Biol Chem* **277:** 11621-11624

Zemach A, Zilberman D (2010) Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Curr Biol* **20:** R780-785

Zhang J, Gao Q, Li P, Liu X, Jia Y, Wu W, Li J, Dong S, Koseki H, Wong J (2011) S phase-dependent interaction with DNMT1 dictates the role of UHRF1 but not UHRF2 in DNA methylation maintenance. *Cell research* **21:** 1723-1739

Zhao J, Ohsumi TK, Kung JT, Ogawa Y, Grau DJ, Sarma K, Song JJ, Kingston RE, Borowsky M, Lee JT (2010) Genome-wide identification of polycomb-associated RNAs by RIP-seq. *Mol Cell* **40:** 939-953

Zhao J, Sun BK, Erwin JA, Song JJ, Lee JT (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science* **322:** 750-756

Zhu H, Geiman TM, Xi S, Jiang Q, Schmidtmann A, Chen T, Li E, Muegge K (2006) Lsh is involved in de novo methylation of DNA. *EMBO J* **25:** 335-345

Zhu J, He F, Hu S, Yu J (2008) On the nature of human housekeeping genes. *Trends Genet* **24:** 481-484

Zofall M, Persinger J, Kassabov SR, Bartholomew B (2006) Chromatin remodeling by ISW2 and SWI/SNF requires DNA translocation inside the nucleosome. *Nat Struct Mol Biol* **13:** 339-346

Zuo X, Sheng J, Lau HT, McDonald CM, Andrade M, Cullen DE, Bell FT, Iacovino M, Kyba M, Xu G, Li X (2012) Zinc finger protein ZFP57 requires its co-factor to recruit DNA methyltransferases and maintains DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. *J Biol Chem* **287:** 2107-2118

Zvetkova I, Apedaile A, Ramsahoye B, Mermoud JE, Crompton LA, John R, Feil R, Brockdorff N (2005) Global hypomethylation of the genome in XX embryonic stem cells. *Nature genetics* **37:** 1274-1279