CYTOSINE METHYLATION AND HYDROXYMETHYLATION AT THE LEPTIN PROMOTER

By Haneen Al-Azzawi, BSc (Hons) MRes

November 2013

UNITED KINGDOM • CHINA • MALAYSIA

School of Biosciences
Sutton Bonington Campus
Loughborough
Leicestershire
LE12 5RD
United Kingdom

Declaration

I hereby declare that this thesis is my own work and that it has not been submitted anywhere for any other degree or award. The work presented herein is my own and where other sources of information have been used, they have been duly acknowledged.

Haneen Al-Azzawi

Abstract

Leptin is an important hormone well known for its role in regulating energy intake and expenditure. DNA methylation levels at the leptin promoter in adult tissues appear to correlate with environmental stresses experienced during early life. This suggests that, once established in early life, DNA methylation is stably transmitted over successive cell generations. The aim of the work presented in this thesis was to determine factors that contribute to the establishment and maintenance of this epigenetic mark at the leptin promoter and to investigate the individual roles of cytosine methylation and cytosine hydroxymethylation at this genomic locus.

No effect of a high fat prenatal diet was observed on leptin promoter DNA methylation levels in the adipose tissue of pigs. However, this genomic region exhibited intermediate levels of DNA methylation, which is usually associated with gene silencing, even though adipose tissue is the primary site of leptin expression.

Double stranded methylation data obtained from DNA methyltransferase (DNMT) mutant mouse embryonic stem cells (mESCs) was used to investigate the contributions of the three catalytically active DNMT enzymes to leptin promoter DNA methylation patterns. Depletion of DNMT3b resulted in increased methylation levels at the leptin promoter, consistent with preliminary data from mutant DNMT3b mouse tissues where similar increases in methylation levels were observed at specific CpG dinucleotides.

Two mESC lines, either hypomethylated or hypermethylated at the leptin promoter, were tested for leptin mRNA expression and neither cell line expressed leptin mRNA, suggesting that some form of methylation may be required for leptin expression. To further investigate the relationship between DNA methylation and leptin expression, *in vitro* differentiated adipocytes were analysed. 3T3-L1 preadipocytes, which do not express leptin, exhibit high levels of DNA methylation and these high methylation levels are maintained after the cells differentiate into leptin-expressing adipocytes. Induction of cytosine hydroxymethylation at the leptin promoter was detected in differentiating and mature adipocytes and evidence is presented to suggest that cytosine hydroxymethylation at the leptin promoter correlates with leptin expression.

Acknowledgments

I would like to express my sincere thanks to the BBSRC for funding this project, my supervisors Reinhard Stöger and Kevin Sinclair for guiding me through my PhD and the many people in the Division of Animal Sciences that helped me along the way.

Extra special thanks to Tom for carrying me over the final hurdle and to Beaver, Pelican and my friends in Stoke for making me realise that in the grand scheme of things, it doesn't really matter if my cells are infected!

And of course a massive thank you to Mum and Dad because without you, I would not exist!

Table of Contents

List of abbreviations

5-caC	5-carboxyl cytosine
5-fC	5-formyl cytosine
5-ghmC	5-glucosyl hydroxymethyl cytosine
5-hmC	5-hydroxymethyl cytosine
5-mC	5-methyl cytosine
AgRP	Agouti related peptide
ANOVA	Analysis of variance
AP	Alkaline phosphatase
ARC	Arcuate nucleus
AZ1	Antizyme 1
BCP	1-bromo 3-chloropropane
BSA	Bovine serum albumin
C/EBPα	Ccaat enhancer binding protein alpha
cDNA	Complementary DNA
CGI	CpG Island
CGI-P	CpG island promoter
CNS	Central nervous system
CpG	Cytosines that are adjacent to guanines on a DNA strand
CREB	cAMP response element binding protein
DEX	Dexamethasone
DKO	DNMT3a$^{(-/-)}$ 3b$^{(-/-)}$ double knockout cell line
DMEM	Dulbecco's modified eagles media
DNA	Deoxyribonucleic acid
DNMT	DNA (cytosine-5)-methyltransferase

DNMT1	DNA (cytosine-5)-methyltransferase 1
DNMT3a	DNA (cytosine-5)-methyltransferase 3a
DNMT3b	DNA (cytosine-5)-methyltransferase 3b
DNMT3L	DNA (cytosine-5)-methyltransferase 3-like protein
dNTP	Deoxynucleotide triphosphates
DTT	Dithiothreitol
ECL	Enhanced chemiluminescence
EDTA	Ethylene diamine tetraacetic acid
EGTA	Ethylene glycol tetraacetic acid
ESC	Embryonic stem cell
FBS	Fetal bovine serum
GAPDH	Glyceraldehydes 3-phosphate dehydrogenase
gDNA	Genomic deoxyribonucleic acid
GMEM	Glasgow minimal essential media
Hp-bss	Hairpin bisulfite sequencing
HPLC	High performance liquid chromatography
HPSF	High purity salt free
HRP	Horse radish peroxidase
IBMX	3-isobutyl-1- methylxanthine
KCl	Potassium chloride
$KRuO_4$	Potassium Perruthenate
LB	Lennox broth
LEP	Leptin gene
LGA	Large for gestational age

LIF	Leukaemia inhibitory factor
lnRNA	Long non-coding RNA
MBD	Methyl CpG binding domain
MeCP2	Methyl CpG binding protein 2
mESC	Mouse embryonic stem cell
miRNA	Micro RNA
MNase	Micrococcal Nuclease
mRNA	Messenger ribonucleic acid
NaCl	Sodium chloride
NPY	Neuropeptide Y
NRC	National Research Council
Nt	Nucleotide
Ob-R	Leptin receptor gene
P	Passage number
pAB	Primary antibody
PBL	Peripheral blood leukocyte
PBS	Phosphate buffered saline
PCNA	Proliferating cell nuclear antigen
PCR	polymerase chain reaction
PMSF	Phenylmethylsulphonyl fluoride
POMC	Proopiomelanocortin
qPCR	Quantitative PCR
RNA	Ribonucleic acid
RPL27	60S ribosomal protein L27

RT	Reverse transcription
Rt-PCR	Reverse transcription polymerase chain reaction
sAB	Secondary antibody
SAT	Subcutaneous adipose tissue
SEM	Standard error of the mean
SGA	Small for gestational age
SNP	Single nucleotide polymorphism
Sp1	Specificity protein 1
TAE	Tris acetate Ethylene diamine tetraacetic acid
TDG	Thymidine-DNA glycosylase
T-DMR	Tissue-specific differentially methylated region
TET	Tet methylcytosine dioxygenase
TET1	Tet methylcytosine dioxygenase 1
TET2	Tet methylcytosine dioxygenase 2
TET3	Tet methylcytosine dioxygenase 3
TF	Transcription factor
TSS	Transcription start site
UHRF1	Ubiquitin-like ring finger domain 1
VAT	Visceral adipose tissue
X-gal	5-bromo 4-chloro-indolyl-ß-D-galactopyranoside
α-MSH	Alpha-melanocyte stimulating hormone

List of figures

List of tables

Introduction

Leptin is an important hormone with a prominent role in the regulation of energy homeostasis. Recently it has emerged that leptin is implicated in other body processes and also in various disease states, indicating that leptin's function may be more complicated than previously thought (Margetic et al, 2002). There are over 20,000 publications involving leptin and although many of these explore leptin expression, information on how leptin gene transcription is regulated is limited and epigenetic mechanisms appear to have been overlooked by the research community.

The leptin promoter is a tissue-specific differentially methylated region (T-DMR) and lies within a CpG rich region, making it a prime target for epigenetic transcriptional regulation through the cytosine modification 5-methylcytosine (5-mC). The main site of leptin expression is the white adipose tissue which exhibits DNA methylation at the leptin promoter in both human and mouse (Stöger, 2006). This is interesting as 5-mC is usually affiliated with transcriptional repression. The recent re-discovery of 5-hydroxymethylcytosine (5-hmC) in mammalian DNA, another cytosine modification that under certain experimental conditions cannot be distinguished from 5-mC (Huang *et al*, 2010; Nestor *et al*, 2010), may provide an explanation for these obscure findings.

Environmental factors can influence cytosine modifications especially during critical windows such as early development. Cytosine modifications at the leptin promoter in adult tissues appear to correlate with stresses experienced during early life suggesting that once established, cytosine modifications are stably transmitted mitotically over successive cell generations and possibly meiotically through transgenerational epigenetic inheritance (Jousse *et al*, 2011; Milagro *et al*, 2009; Stöger, 2006; Tobi *et al*, 2009).

It is, therefore, of importance to determine how this key metabolic gene is transcriptionally regulated through cytosine modifications and the mechanisms by which this locus maintains or changes its state of cytosine modification.

Chapter 1. Literature Review

1.1 Leptin

1.1.1 An introduction to leptin

Leptin is a key metabolic hormone and the journey to its discovery began in 1949 at The Jackson Laboratory. A spontaneous mutation had resulted in mice exhibiting a syndrome of obesity and over ten years later parabiosis experiments were used to determine that these mice lacked a circulating factor. When this factor was reintroduced, the obese phenotype was reversible and, furthermore, the factor was believed to act through receptors in the brain as mutant mice with hypothalamic lesions were unaffected (*Reviewed by* Flier & Maratos-Flier, 2010). Leptin was finally discovered in 1994 and this was a major advance in modern physiology. As little was known about the mechanisms that controlled body weight, the discovery of leptin and its actions changed our understanding of the causes of obesity (Flier & Maratos-Flier, 2010; Zhang *et al*, 1994). Many years after the initial publication that began the journey to leptin's discovery (Hervey, 1959) we have yet to fully elucidate its role in the mammalian body.

Although leptin has a prominent role in relaying information about energy stores to the brain (MacDougald *et al*, 1995; Morris & Rui, 2009), there is increasing evidence indicating that leptin has a more diverse role. Leptin is most abundant in white adipose tissue but it is also expressed in other areas of the body. Although leptin acts on pathways in the brain, leptin receptors exist in other tissues and organs, and leptin plays an important role in many peripheral pathways (Margetic *et al*, 2002). Leptin has been linked to such diverse processes as reproduction (Holness *et al*, 1999; Mantzoros, 2000; Moschos *et al*, 2002), osteogenesis (Zaidi *et al*, 2012) and angiogenesis (Bouloumié *et al*, 1998). Abnormal leptin expression is associated with diseases such as hepatitis (Kukla *et al*, 2011), preeclampsia (Margetic et al, 2002) and reproductive dysfunction (Brzechffa et al, 1996; Chehab et al, 1996; Pandey et al, 2010).

1.1.2 The leptin gene

The leptin gene was initially identified in the mouse (Zhang *et al*, 1994) but orthologous genes have since been identified in humans and many other mammalian species (Doyon *et al*, 2001). For most species only one leptin transcript has been identified to date. This transcript produces a protein consisting of 167 amino acid residues (Margetic et al, 2002). A shorter protein coding transcript has now been identified in the mouse (ensemble annotation GRCM38) which produces a shorter protein product from the same promoter

region. The leptin promoter shows a high level of conservation between many species (Figure 1.1). An enhancer element has also been described for the human leptin gene which is located 1.9 kb upstream of the transcription start site (Bi *et al*, 1997). It has a placental-selective element and only functions in cells of the placenta. No other species are known to possess this enhancer region.

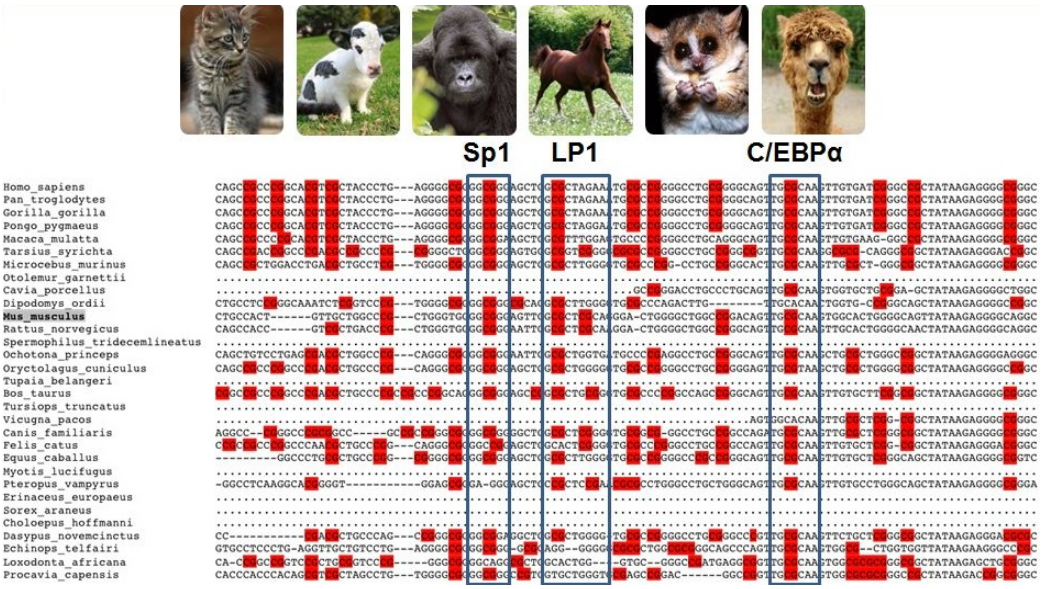

Figure 1.1 The leptin promoter. An alignment of the leptin promoter shows a high level of conservation between many species. Within the core promoter are two transcription factor binding sites, Sp1 and C/EBPα. Site LP1, which does not match the recognition sequence of any known transcription factor, binds a factor present in preadipocytes and adipocytes but not in other cell types (Mason et al, 1998). For an alignment of species studied see Figure 2.14. *Alignment produced using EBI's ClustalW2.*

1.1.3 Leptin synthesis

Leptin mRNA is primarily expressed in adipocytes and to a lesser extent in other tissues like the ovaries, placenta, skeletal muscle and stomach (Table 1.1). White adipose tissue is the predominant site of leptin synthesis although there is some evidence to suggest that brown adipose tissue is also a site for the production of leptin. Studies into levels of leptin mRNA in different sites of adipose tissue have shown that there is site specific variation in both humans and rodents. For example, in humans, leptin mRNA expression is greater in subcutaneous fat depots than in omental adipose tissue (Margetic *et al*, 2002).

Sex-specific differences have also been observed in leptin synthesis and it has been proposed that this is because of the stimulating role of oestrogens and the suppressive effect of androgens. Females appear to have higher plasma leptin levels than males. As leptin production corresponds to the size of body fat stores, this correlation between leptin concentrations and fat mass may explain why females have higher leptin levels as in general they have a higher percentage of body fat than men (Wabitsch *et al*, 1997).

Table 1.1 Sites of leptin mRNA expression. Leptin mRNA is expressed in various tissues in both human and mouse. Corresponding sites of leptin receptor mRNA expression are also indicated (Margetic et al, 2002).

	Leptin	Leptin Receptor
White adipose tissue	✓	
Brown adipose tissue	✓	
Placenta	✓	✓
Ovary	✓	✓
Stomach	✓	
Skeletal muscle	✓	
Liver	✓ (Protein)	✓
Heart	✓(Fetal)	✓

Leptin is translated as a 167 amino acid protein the sequence of which is highly conserved between many species including human, mouse and cow indicating that leptin and its role are important. Leptin is composed of four very similar α-helices and post-translational modifications cleave the protein into a functional molecule that consists of two peptides (Zhang *et al*, 1997).

1.1.4 The leptin receptors

The anorexigenic effects of leptin are mediated through the interaction of leptin with the leptin receptor (*Ob-R*) which is a member of the cytokine receptor family. Leptin protein is secreted into the blood and travels across the blood-brain barrier into the central nervous system (CNS) where it can act on a leptin receptor expressing area of the hypothalamus called the arcuate nucleus (ARC) (Myers, 2004).

The *Ob-R* gene is located on chromosome 1 and 4 in the human and mouse genomes, respectively. At least six splice variants of the leptin receptor exist (named Ob-Ra to Ob-Rf) all of which are the product of the *Ob-R* gene and are the result of alternative mRNA splicing. It is thought that each splice variant has a different function (Margetic *et al*, 2002).

Leptin receptor isoforms can be divided into three groups: long, short and secreted. The long form, Ob-Rb, is the only isoform that can mediate leptin signalling. This is because it is the only isoform that has a full length intracellular signalling domain. The short and long isoforms have identical extracellular and transmembrane domains and the same first 29 amino acids of the intracellular domain but differ in sequence after because of alternative splicing. The secreted form, Ob-Re, contains only the extracellular domain. The functions of the short (Ob-Ra, Ob-Rc, Ob-Rd and Ob-Rf) and secreted (Ob-Re) isoforms is unclear, however, it has been proposed that the short forms facilitate the transport of leptin across the blood-brain barrier and as the secreted form consists of just the extracellular domain that binds circulating leptin, it is thought that Ob-Re regulates the concentrations of free leptin (Munzberg *et al*, 2005).

The human and murine leptin receptors share a high level of amino acid sequence homology within both the intracellular (71%) and extracellular domains (78%) (Tartaglia, 1997). Ob-Ra and Ob-Rb appear to be well conserved among many species although this cannot be said for other isoforms (Munzberg *et al*, 2005)

1.1.5 Leptin signalling

Leptin is the primary adipose hormone that relays information about energy stores to the brain through interaction with the leptin receptor (Figure 1.2). The leptin receptor is expressed in various regions of the brain and within the hypothalamus leptin can act on two sub-populations of Ob-Rb expressing neurons within the ARC of the hypothalamus to mediate its anorexigenic effect (Morris & Rui, 2009).

As energy levels in the body rise and become adequate insulin is released which signals the body to start storing glycogen and also induces the production of leptin in adipocytes which is secreted into the blood (MacDougald *et al*, 1995). For leptin to enter the central nervous system and reach the hypothalamus it must pass through the blood-brain barrier and/or blood-cerebrospinal fluid barrier. It is possible that the short isoforms of the leptin receptor act as transporters and that Ob-Ra may be responsible for transporting leptin from the peripheral blood and into the CNS but it is also argued that transporters come from a different origin (Kawamura *et al*, 2002) (Ziylan *et al*, 2009).

Within the ARC Ob-R is expressed in at least two populations of neurons. These are neurons that co-express Agouti-related protein (AgRP) and Neuropeptide Y (NPY) and neurons that express Proopiomelanocortin (POMC).

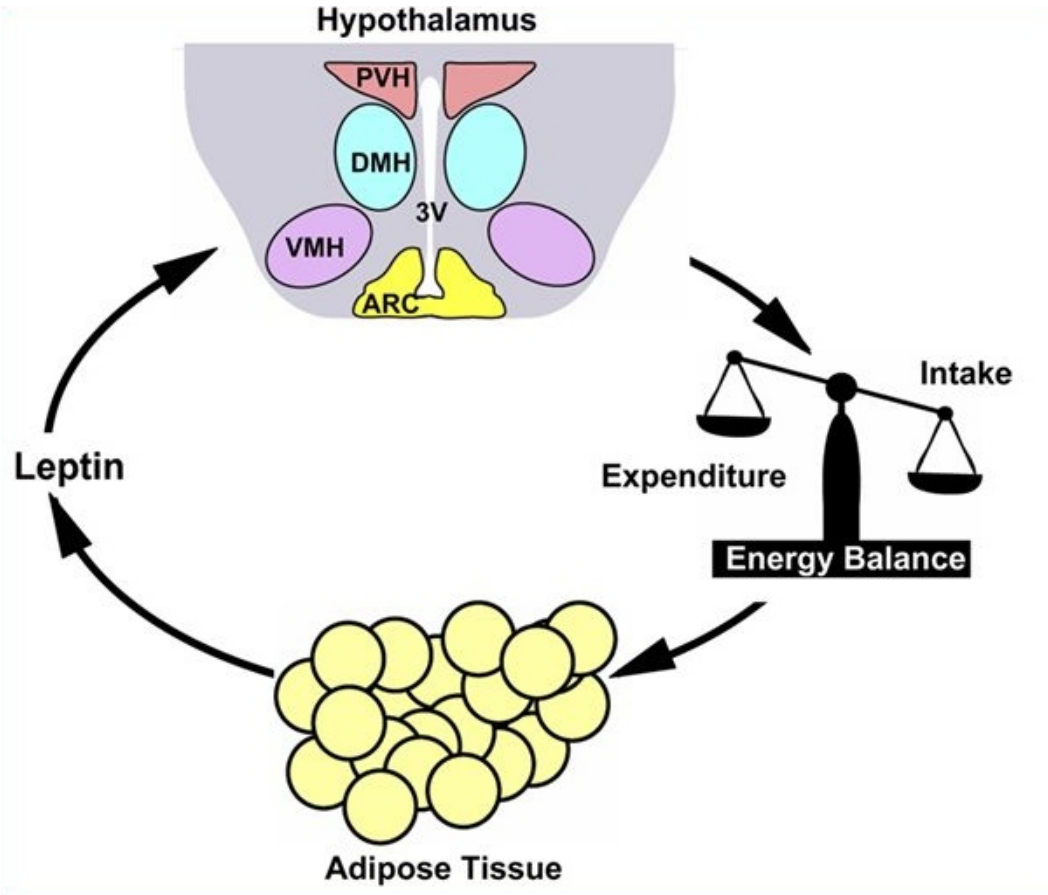

Figure 1.2 Leptin as a satiety factor. Leptin's primary role is as a satiety factor in energy homeostasis. Leptin is produced in and secreted from the white adipose tissue into the blood. Leptin is transported into the CNS and acts on leptin receptors in the hypothalamus to induce leptin signalling and mediate anorexigenic effects. Arcuate nucleus (ARC), ventromedial nucleus (VMH), dorsomedial nucleus (DMH), posteromedial hypothalamus (PMH), third ventricle (3V) *Taken from Morris and Rui, 2009.*

In the Ob-R-POMC neuron, Ob-R activates the neuron by stimulating the expression of POMC which is converted to alpha melanocyte stimulating hormone (α-MSH), a hormone which mediates an appetite suppressing signal. Within the AgRP and NYP co-expressing neurons, AgRP acts as an antagonist of α-MSH signalling and NYP is an appetite stimulating hormone. Leptin interacts with Ob-R to inhibit the NYP-AgRP neurons and activate Ob-R-POMC neurons therefore suppressing the appetite (Myers, 2004).

1.1.6 Multiple roles for leptin

Our current understanding of leptin concludes that its most prominent role is in energy homeostasis as a satiety factor and as an indicator of adiposity.

When leptin was discovered it was thought to be exclusively an adipose tissue-derived, anti-obesity hormone, the role of which was to relay information about energy stores to the brain and regulate metabolism (Zhang *et al*, 1994). There is now a wealth of evidence to suggest that leptin's role is more extensive. It has been proposed to have a role in such processes as inflammation, angiogenesis, immunity, haematopoiesis and reproduction (Margetic *et al*, 2002).

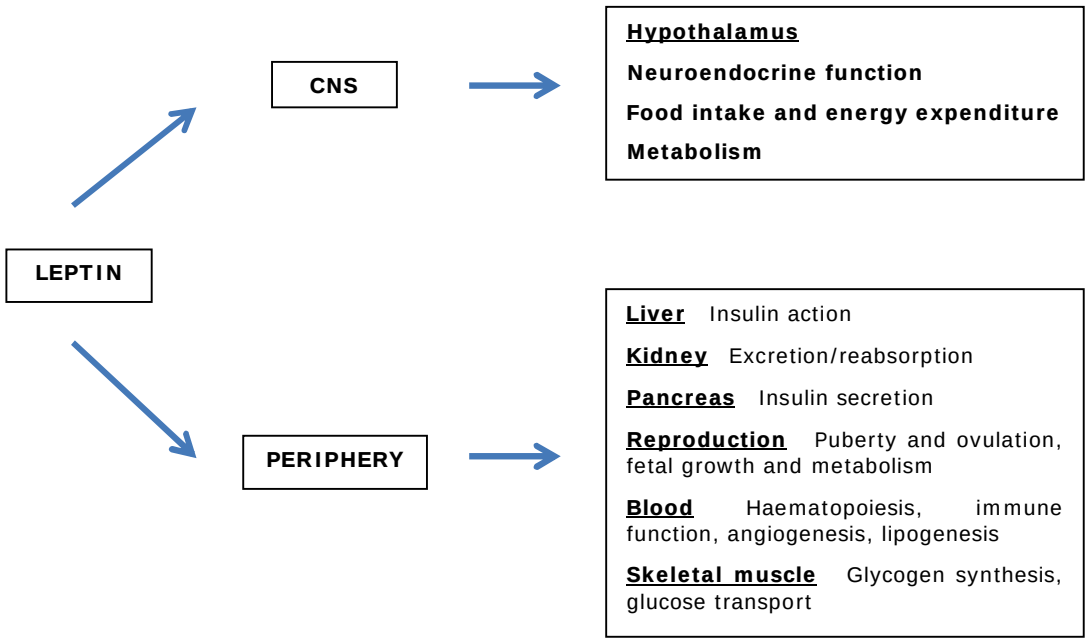

Figure 1.3 The central and peripheral actions of leptin.

1.1.7 Variations in leptin expression and leptin levels

In its primary role as a modulator of energy homeostasis, leptin exhibits circadian oscillations and a strong relationship has been suggested between feeding and the diurnal rhythm of leptin (Bodosi *et al*, 2004). Adipose tissue, the primary site of leptin expression and secretion, exists in various depots around the body and there appears to be a site specific variation of leptin expression levels in rodents (Maffei *et al*, 1995; Trayhurn *et al*, 1995) and humans. For example, in humans, leptin mRNA expression is greater in subcutaneous fat depots than in omental fat depots (Hube *et al*, 1996; Masuzaki *et al*, 1996).

Although variations in leptin protein levels may be normal in the regulation of energy homeostasis, there have been many reports of abnormal leptin expression levels in relation to disease and dysfunction. During gestation, plasma leptin levels in the mother increase. This coincides with an increase in leptin receptor binding protein levels in maternal circulation and could lead to prolonged leptin circulation or to increased synthesis by either maternal fat depots and or the placenta (Margetic *et al*, 2002). Leptin has been found to be differentially expressed in preeclamptic placentas via an increase in mRNA levels (Enquobahrie *et al*, 2008) and maternal plasma levels are also significantly elevated in those with preeclampsia (Mise *et al*, 1998). Low concentrations of leptin in placental and cord blood have been linked to fetal growth retarded pregnancies (Ashworth *et al*, 2000). A pre-pubertal leptin surge has been noted in both boys and girls (Moschos *et al*, 2002) and there is also a leptin surge in neonates which may be involved in nutritional programming (Ahima *et al*, 1998). In patients with conditions associated with food restriction such as anorexia nervosa, serum leptin levels are significantly lower than in control groups although this correlates with body weight and fat (Mantzoros *et al*, 1997).

Variations in leptin expression have been linked to dependence on the number and size of adipocytes. This may be a plausible conclusion in some circumstances such as in anorexia nervosa, where there is less adipose tissue to produce leptin. As individual adipose depots have different characteristics such as differences in the number of large or small adipocytes, the variation in leptin expression differs accordingly. There are circumstances where changes in fat mass and adipocyte cell size do not appear to correlate with increases or decreases in leptin expression such as the neonatal leptin surge (Margetic *et al*, 2002). Under these circumstances the transcriptional regulation of the leptin gene may influence leptin expression levels.

1.1.8 Conclusions

The complexity of the leptin signalling axis requires an improved knowledge to fully understand its roles within the body. Although many studies have provided information concerning variation in leptin expression, it is interesting to note how little information there is regarding the transcriptional regulation of the leptin gene. The leptin gene possesses a CpG rich promoter and DNA methylation detected in this region differs from the bimodal patterns that are usually observed at CpG islands of imprinted loci. Despite this property, leptin expression levels are rarely correlated with epigenetic data. The transcriptional

regulation of the leptin gene appears to have been overlooked by the scientific community.

1.2 Multiple epigenetic mechanisms including DNA methylation

Epigenetics is a relatively new field concerning alterations to the genome that may be heritable through meiosis and mitosis, but that do not involve changes to the DNA sequence itself and its fundamental role is in the regulation of gene expression (Berger *et al*, 2009). DNA is wrapped around histone proteins to form nucleosomes, the basic unit of chromatin. Nucleosomal properties can be changed via several epigenetic mechanisms and the exact way in which this occurs is integral to the cells identity as this allows for genes, promoters and other genetic elements to be accessible when necessary (Kooistra & Helin, 2012). Whilst the focus of this thesis is DNA methylation, a brief overview of other epigenetic modifications will now be given.

The modification state of histone tails are important determinants of chromatin structure (Kooistra & Helin, 2012). The first to be discovered, histone acetylation, is generally associated with active genes suggesting a role in the facilitation of gene expression. Acetylation is thought to neutralise the positive charge produced by lysine residues which influences transcriptional outcomes. In contrast, methylation of a lysine residue does not affect its positive charge and is associated with both gene repression (H3K9 and H3K27) and activation (H3K4 and H3K36) (*Reviewed by* Zentner & Henikoff, 2013).

Additionally, antisense transcripts such as long non-coding RNAs (lnRNAs) and microRNAs (miRNAs) are now known to play a role in gene regulation (Lee, 2012; Nazarov et al, 2013). In the case of lnRNAs, these transcripts may be present when the sense transcript is not transcribed, indicating that they play a role in gene repression (Lee, 2012). There is evidence to suggest that in certain cases they may also promote sense transcript expression (Dietz *et al*, 2012).

Polycomb group (PcG) proteins are transcriptional repressors which function in the modification of histone tails, in order to implement gene silencing. PcGs function in complexes which can be attracted by CpG rich regions and lnRNAs (*Reviewed by* Simon & Kingston, 2013). Consequently it is emerging that epigenetic modifications work in concert with each other, rather than in isolation to define the epigenomic landscape of an individual cell but, for the sake of brevity, they will not be reviewed further. DNA methylation is the most well studied epigenetic marking and involves the covalent modification of nucleotides within the DNA sequence (Feil & Fraga, 2011). In mammals DNA methylation

affects cytosine and predominantly those that are adjacent to guanines on a DNA strand (Deaton & Bird, 2011).

1.2.1 Cytosine modifications in mammalian DNA

In mammalian DNA the cytosine base can be modified to form several cytosine modifications including 5-methyl cytosine (5-mC), the most intensively studied epigenetic mark. Commonly referred to as DNA methylation, we have yet to fully elucidate the role of 5-mC but evidence has shown it to be a key epigenetic modification in mammals that is essential for normal development (Li *et al*, 1992). Recently, three other cytosine modifications, 5-hydroxymethyl cytosine (5-hmC), 5-carboxyl cytosine (5-caC) and 5-formyl cytosine (5-fC) have also been identified (Ito et al, 2011). Although the existence of 5-hmC in mammalian DNA was first reported in the 1970s (Penn *et al*, 1972), its presence was widely forgotten until recently when significant quantities were detected in purkinje neurons (Kriaucionis & Heintz, 2009).

With the re-discovery of 5-hmC came the revelation that many techniques used to explore 5-mC patterns do not distinguish between 5-mC and 5-hmC (Huang *et al*, 2010; Nestor *et al*, 2010). Therefore, much of the scientific literature concerning 5-mC and its effects may be skewed by the presence of 5-hmC, and this necessitates the re-interpretation of past findings which have shaped perceptions regarding the role of DNA methylation. The abrupt influxes of publications concerning 5-hmC have suggested various potential roles for this cytosine modification, adding yet another level of complexity to the phenomenon of epigenetics. It appears that the role of DNA methylation is dependent on the cell type, the gene of interest, and the specific cytosine modification present. This thesis will refer to 5-mC and 5-hmC collectively as DNA methylation and individually as cytosine methylation (5-mC) and cytosine hydroxymethylation (5-hmC).

1.2.2 DNA methylation in mammalian DNA

In mammals DNA methylation occurs mainly in CpG dinucleotides (Ziller et al, 2011). The majority of CpGs in mammalian genomes are subject to DNA methylation. For example, human somatic tissue DNA exhibits cytosine methylation at approximately 70% of all CpG sites (Ehrlich *et al*, 1982). Due to the mutagenic properties of 5-mC, this has resulted in CpG-deficient mammalian genomes (Bird, 1980). Exceptions to this global DNA methylation are CpG islands (CGIs); sections of the genome which are predominantly unmethylated and consequently CpG rich in comparison to the rest of the genome (Suzuki &

Bird, 2008). CGIs often overlap with the promoter regions of individual genes with more than 60% of mammalian genes possessing CpG island promoters (CGI-P) (Illingworth et al, 2010). CGI-Ps can influence local chromatin structure through epigenetic modifications to regulate gene activity. DNA methylation at these regions is commonly thought to be associated with the transcriptional silencing of a gene (*reviewed by* Deaton & Bird, 2011).

The current perceptions of DNA methylation concludes that DNA methylation patterns are established during early embryonic development and are then stably maintained throughout the life of the individual. Unmodified cytosines are firstly methylated to become 5-mC, the pattern of which is established and maintained by the DNA methyltransferases (DNMT) family of enzymes (Cheng & Blumenthal, 2008). 5-mC can then be oxidised by the ten-eleven translocation (TET) family of enzymes to form 5-hmC, 5-fC and 5-caC (Tahiliani *et al*, 2009). We have yet to find an enzyme that can directly hydroxymethylate cytosine to produce 5-hmC, suggesting that the presence of 5-hmC is dependent on the oxidation of pre-existing 5-mC (Figure 1.4). Consequently, this thesis will begin by reviewing the literature on the establishment and maintenance of cytosine methylation patterns by the DNMT family of enzymes.

1.2.3 The establishment and maintenance of cytosine methylation (5-mC) patterns

Cytosine methylation patterns are established and maintained by the DNMT family of enzymes. Three catalytically active members of the mammalian DNMT family, DNMT1, DNMT3a and DNMT3b, are directly responsible for the addition of methyl groups to cytosine to form 5-methyl cytosine (5-mC) (Cheng & Blumenthal, 2008). It is generally accepted that DNMT3a and DNMT3b are responsible for establishing cytosine methylation patterns during embryonic development whilst DNMT1 maintains this established pattern in somatic tissues.

DNMT1 is known as the maintenance methyltransferase and is regarded as being responsible for the transfer of methylation patterns from parent strand to daughter strand during DNA replication (*reviewed by* Qin *et al*, 2011). The CpG dinucleotide is a short, self-complementary sequence which results in palindromic CpG dyads on double stranded DNA. Evidence of symmetrical DNA methylation patterns at CpG dyads (Bird, 1978a; Wigler, 1981) led to the hypothesis that cytosine methylation is maintained in successive cell generations via the transfer of methylation patterns between strands during semi-conservative DNA replication.

The evidence for DNMT1 maintenance activity stems mainly from *in vitro* experiments which found that as well as being specific for CpG, hemi-methylated DNA (where DNA methylation is present on one strand but not on the opposing strand of a DNA molecule) is the preferred substrate for DNMT1 (Bestor & Ingram, 1983). Additionally, DNMT1 has protein domains that can interact with proliferating cell nuclear antigen (PCNA), a DNA polymerase processing protein (Chuang *et al*, 1997), and Ubiquitin-like, containing PHD and RING finger domains 1 (UHRF1) (Sharif *et al*, 2007). These factors may ensure the localisation of DNMT1 to the replication fork, concurrent with the idea that DNMT1 maintains cytosine methylation during DNA replication. DNMT1 is mainly produced during the S phase of the cell cycle (Robertson *et al*, 2000), when it would be required to methylate newly generated hemi-methylated sites. Knockout of DNMT1 in mESCs leads to genome-wide hypomethylation (Li *et al*, 1992) suggesting that DNMT1 is required for maintenance methylation and is responsible for the majority of DNA methylation in these cells.

For cytosine methylation to be maintained throughout successive cell generations it must first be established in the original cell population. At fertilisation, a totipotent embryo that gives rise to all cell types of the offspring is produced from the combination of the maternal and paternal gametes. Epigenetic reprogramming occurs during early embryonic development. Dramatic changes in genome-wide DNA methylation occur during germ cell development and early preimplantation development and this involves the erasure of maternally and paternally inherited DNA methylation patterns and the establishment of an offspring specific DNA methylation pattern (*reviewed by* Mason *et al*, 2012).

Within hours of fertilisation the paternal pronucleus undergoes rapid demethylation (Mayer *et al*, 2000). This occurs at a time when there is no DNA replication suggesting the active removal of 5-mC. The specific mechanism by which this demethylation occurs is a matter of debate, however, hydroxymethylation of the genome could be a potential mechanism and will be discussed later in this review (Section 1.2.4.1).

During the cleavage stages of preimplantation development the maternally inherited genome becomes increasingly demethylated. This is a passive process that occurs during active DNA replication in the absence of maintenance methylation; as DNMT1 is excluded from the nucleus at this point (Cardoso *et al*, 1999). Exceptions to this wave of demethylation are those genes which are imprinted (Nakamura *et al*, 2007); an epigenetic process that results in a

parent-specific expression that is associated with differential DNA methylation of an imprinted control region (Barlow, 2011).

When unmethylated DNA is introduced into somatic cells, it is not subject to DNA methylation but, when introduced into mouse preimplantation embryos, becomes stably methylated (Jähner *et al*, 1982). This has led to the conclusion that the establishment of cytosine methylation patterns is confined to the totipotent stages of embryogenesis. DNMT3a and DNMT3b are both highly expressed during pluripotency and down-regulated post-differentiation. They are also essential for the establishment of methylation patterns (Athanasiadou *et al*, 2010; Okano *et al*, 1999). They are widely referred to as the "*de novo*" DNMTs.

DNMT3a and DNMT3b show equal activity towards unmethylated and hemi-methylated DNA (Hashimoto *et al*, 2012) and their deletion in pluripotent mESCs leads to a loss of *de novo* methylation (Athanasiadou *et al*, 2010; Okano *et al*, 1999). They are highly expressed at totipotent stages of development and down-regulated after differentiation which agrees with the need for them to establish methylation patterns during early development and the assumption that after differentiation they are dispensable (Jones & Liang, 2009). DNMT3a and DNMT3b are therefore thought to be responsible for initiating methylation patterns and for this reason are referred to as the *de novo* DNMTs. In turn they are crucial for normal development as loss of both leads to early embryonic lethality (Okano *et al*, 1999). Although studies suggest that each *de novo* DNMT targets specific sequences, information about the individual contributions of each enzyme is limited and there is evidence to show that their functions overlap and complement each other (Athanasiadou *et al*, 2010). Double knockout of the *de novo* DNMTs leads to genome wide hypomethylation whilst individual knockouts result in different locus-specific losses of DNA methylation (Chen *et al*, 2003; Okano *et al*, 1999). Different phenotypes are observed with inactivation of the individual genes. Loss of DNMT3a leads to a failure in the establishment of methylation patterns at imprinted genes (Hu *et al*, 2008) and severe defects resulting in postnatal lethality (Okano et al, 1999). Inactivation of DNMT3b results in embryonic lethality and hypomethylation of pericentromeric repeats. DNMT3b hypomorphic germline mutations are associated with ICF syndrome and are thought to be responsible for the majority of cases of this disease (Ehrlich, 2003; Jin *et al*, 2008).

1.2.3.1 Establishment and maintenance of DNA methylation re-evaluated

The perception that DNMT3a and DNMT3b establish methylation patterns that DNMT1 subsequently maintains during semi-conservative DNA replication is the generally accepted view. However, there has been an accumulation of data that questions these perceptions and the enzymatic control of DNA methylation appears to be more complicated than is assumed by the scientific community. Much of the evidence that has led to these perceptions stems from publications reporting the *in vitro* activities of DNMT enzymes. *In vivo*, DNA replication has a high processivity (0.035 seconds/nucleotide) (Jackson & Pombo, 1998) whilst, *in vitro*, DNMT1 has a low turnover rate (70-450 sec/methyl group) (Pradhan *et al*, 1999). *In vivo*, DNMT enzymes work in concert with other factors and proteins, the scale of which we are only just beginning to appreciate, culminating in a currently poor understanding of the *in vivo* enzymatic control of cytosine methylation (Arand *et al*, 2012).

DNMT1 is normally regarded as the "maintenance methyltransferase", stably copying methylation patterns from the parent strand to the daughter strand during DNA replication (Gowher & Jeltsch, 2001; Okano *et al*, 1998). This implies that hemi-methylated sites, where one strand of the DNA molecule is methylated whilst the opposing strand is unmethylated, should be found in low abundance in the genome as they immediately become fully methylated after DNA replication. Early landmark experiments (Bird, 1978a) concluded that levels of hemi-methylation are indeed low, but recent more sensitive techniques have found significant levels of hemi-methylation in single-copy sequences and repeat sequences (Liang *et al*, 2002). To address the problems arising from the discrepancies between *in vitro* and *in vivo* experiments, DNMT mutant mouse embryonic stem cells (mESCs) have become a popular model for investigating the actions of DNMT enzymes. mESCs depleted of the *de novo* DNMTs but expressing DNMT1 exhibit hypomethylation of imprinted genes and repeats and gradually lose global DNA methylation with increasing cell divisions (Armstrong et al, 2012; Chen et al, 2003; Liang et al, 2002). This evidence suggests that in the absence of DNMT3a and DNMT3b, DNMT1 is unable to maintain methylation at certain genomic elements. Additionally, although knockout of DNMT1 in mESCs leads to genome-wide hypomethylation (Li *et al*, 1992) analysis of specific loci has determined that in the absence of DNMT1, some sequences maintain a considerable amount of hemi and fully methylated sites (Arand *et al*, 2012). Although DNMT3a and DNMT3b are down-regulated after differentiation, they are still expressed in somatic cells and their simultaneous knockout results

in the hemi-methylation of up to 30% of CpG sites in repetitive elements of the mouse genome (Jones & Liang, 2009). This is concurrent with a gradual loss of DNA methylation with increasing cell divisions of DNMT3a and DNMT3b double knockout mESCs (Jackson *et al*, 2004). These observations suggest that the DNMT3 enzymes may also contribute to maintenance methylation and ongoing methylation is required by DNMT3a and DNMT3b in order for these sequences to retain their methylation. DNMT1 also shows some activity towards unmethylated DNA (Bestor, 2000) and therefore could potentially contribute to the establishment of cytosine methylation patterns.

The action of the DNMTs is generally perceived to be processive. That is, they would either methylate or not methylate every CpG dinucleotide in a given stretch of DNA. This leads to the expectation that methylation levels at every CpG dinucleotide within a given cell type would be 0%, 100% or 50% (if one allele was methylated and the other allele was unmethylated) and also that CpG dinucleotides in the same stretch of DNA would have the same methylation status. However, the quantification of methylation levels rarely gives such clear cut results. This heterogeneity could be due to the faithful maintenance of heterogeneous patterns that were established during early development (*reviewed by* Jones & Liang, 2009). It could equally be due to non-processive actions of the DNMT enzymes. *In vitro*, the maintenance function of DNMT1 appears to occur in a processive manner and a processive maintenance function for DNMT3b has also been suggested (Gowher & Jeltsch, 2002; Hermann *et al*, 2004; Vilkaitis *et al*, 2005). The *de novo* functions of DNMT1 and DNMT3a, however, appear to be distributive (Gowher & Jeltsch, 2001; Gowher *et al*, 2005) although a processive methylation activity for DNMT3a has also been described (Holz-Schietinger & Reich, 2010).

It has also been suggested that the current hypothesis for the maintenance of DNA methylation is flawed as it does not include a mechanism for error correction (*reviewed by* Jones & Liang, 2009). It is unlikely that methylation patterns established during early development could be faithfully maintained by DNMT1 alone without a proofreading mechanism to ensure that the fidelity of the patterns is maintained; and it is doubtful that the genome could be kept methylated by DNMT1 alone (Ooi & Bestor, 2008).

The most popular techniques for DNA methylation analysis require bisulfite conversion of DNA. This renders the two strands of a DNA molecule non-complementary and the opposing strands become denatured. Consequently, it is technically challenging to analyse the DNA methylation pattern on both strands

of an individual DNA molecule. The general consensus regarding the maintenance activity of DNMT1 has led to the majority of scientists disregarding one strand of DNA during methylation analysis with the assumption that because of DNMT1s maintenance activity, it will mimic the methylation pattern of the other strand. With this evidence in mind it is clear that further studies into the specificity and processivity of the DNMTs is required before it can be assumed that both DNA strands of a DNA molecule have the same DNA methylation pattern.

1.2.4 Cytosine hydroxymethylation (5-hmC)

In contrast with 5-mC levels, the 5-hmC content of human tissues varies significantly between tissues (Nestor *et al*, 2012). Current reports have found it to be most abundant in breast tissue (Nestor *et al*, 2012), brain tissue, pre-implantation embryos and embryonic stem cells (Globisch *et al*, 2010; Williams *et al*, 2012). Consequently, many studies investigating the role of 5-hmC use embryonic stem cells (ESCs) as a model and, as evidence suggests this modification to be tissue-specific, these results may not be representative of the role of 5-hmC in all tissue types. Additionally, DNA methylation patterns frequently change in cell culture (Antequera & Bird, 1993), and culture can also result in a dramatic loss of global 5-hmC and TET gene expression levels (Nestor *et al*, 2012).

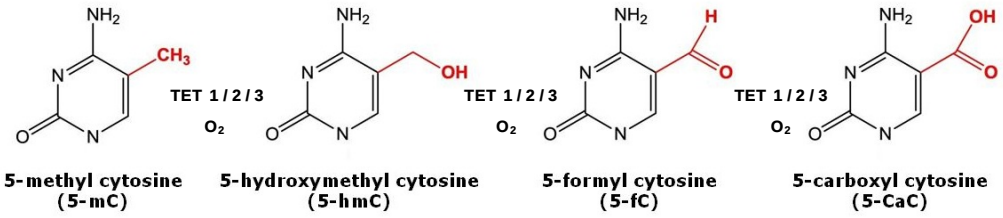

Figure 1.4. The structures of 5-mC, 5-hmC, 5-fC and 5-CaC. 5-mC is formed through the addition of a methyl group to the C-5 position by the DNMT enzymes (highlighted in red). This methyl group can be oxidised by the addition of a hydroxyll group (highlighted in red) to form 5-hmC. 5-hmC can be further oxidised by the TET enzymes to produce 5-fC and 5-CaC. *Figure adapted from phys.org.*

Presently, the oxidation of 5-mC by the TET enzymes is the only known way to acquire 5-hmC in the genome and TET1, TET2 and TET3 are all capable of converting 5-mC to 5-hmC *in vivo* and *in vitro* (Ito *et al*, 2010; Tahiliani *et al*, 2009; Williams *et al*, 2012). It has been reported that the TET enzymes exhibit tissue-specific expression with TET1 mainly being expressed in ESCs whilst TET2

and TET3 are more ubiquitously expressed (Szwagierczak *et al*, 2010; Tahiliani *et al*, 2009; Williams *et al*, 2012).

The majority of studies regarding hydroxymethylation have mainly been carried out in mESCs owing to their global 5-hmC content. Depletion of TET1 in mESCs leads to only a 20-40% reduction in 5-hmC, and to a slight increase in the ratio of 5-mC:5-hmC (Dawlaty *et al*, 2011; Koh *et al*, 2011; Williams *et al*, 2011; Xu *et al*, 2011). The incomplete disappearance of 5-hmC could be attributed to another TET enzyme compensating for the loss of TET1. TET2 is also expressed in mESCs and a functional redundancy between the two enzymes could explain remaining levels of 5-hmC after TET1 depletion whereby TET2 may partly compensate for the loss of TET1 (Williams *et al*, 2012). The limited effect of TET1 knockdown/knockout on 5-mC levels may be expected because there is so much more 5-mC than 5-hmC with an estimated ratio of one 5-hmC nucleotides for every 22 5-mCs nucleotides (Ito et al, 2011). Therefore changes in 5-hmC level may not lead to detectable changes in global 5-mC levels (Williams *et al*, 2012). Gene specific increases in 5-mC after TET1 depletion have been observed in several studies, which were associated with a simultaneous loss of 5-hmC showing that TET1 regulates DNA methylation levels in certain genes (Dawlaty *et al*, 2011; Koh *et al*, 2011; Williams *et al*, 2011; Wu & Zhang, 2011; Xu *et al*, 2011).

There is not a complete overlap with 5-hmC containing regions and TET1 binding. This taken with the limited effect of TET1 depletion on global 5-hmC and 5-mC levels suggests that other pathways and/or proteins are involved in 5-hmC regulation. Although evidence strongly suggests that the existence of 5-hmC is dependent on pre-existing 5-mC (Ficz *et al*, 2011; Williams *et al*, 2011), global 5-hmC levels in human tissues do not correlate with TET1/2/3 gene expression levels (Nestor *et al*, 2012), and peaks in hydroxymethylation content that are independent of changes in 5-mC levels have been observed in the mouse zygote (Salvaing *et al*, 2012) suggesting there may be other factors involved in the conversion of cytosine to 5-hmC.

1.2.4.1 5-hmC as an intermediate in an active or passive demethylation pathway

The DNMT enzymes which catalyse cytosine methylation are well characterised, unlike the enzymes which demethylate DNA. The discovery of 5-hmC may have uncovered a pathway for demethylation, as this cytosine modification has the ability to recruit or displace certain proteins and may act as an intermediate

between 5-mC and cytosine in a demethylation pathway (Globisch *et al*, 2010; Liutkeviciute *et al*, 2009; Tahiliani *et al*, 2009).

5-hmC could act as an intermediate in an active demethylation pathway which ultimately results in the replacement of 5-mC with cytosine. This could serve as a likely demethylation pathway for the paternal genome during epigenetic reprogramming (Williams *et al*, 2012). This demethylation pathway has been suggested to involve DNA repair mechanisms (Guo *et al*, 2011), but the emergence that 5-hmC can further be oxidised by the TET proteins to 5-fC, and 5-caC has uncovered a potential enzymatic pathway for active DNA demethylation. This is supported by studies which demonstrate the presence of the 5-fC and 5-caC in mammalian DNA (Ito et al, 2011; Williams et al, 2012). The TET proteins cannot convert 5-caC to cytosine, therefore, another enzyme such as a decarboxylase or glycosylase must be involved in this step. The enzyme could be thymidine-DNA glycosylase (TDG). Depletion of TDG leads to the accumulation of 5-caC in ESCs (He *et al*, 2011) and another study has shown that TDG is required for active DNA demethylation (Cortellino *et al*, 2011). Enzymes that are capable of dehydroxymethylating 5-hmC and converting it to cytosine have been elusive but, *in vitro,* DNMT3a and DNMT3b have been found to function as DNA hydroxymethyl cytosine dehydroxymethylases (Chen *et al*, 2012).

5-hmC could also be part of a passive demethylation pathway. DNMT1, the maintenance methyltransferases, has low specificity for 5-hmC (Valinluck & Sowers, 2007). While it is copying methylation patterns from the parent strand to the daughter strand during DNA replication, DNMT1 would treat 5-hmC as if it were an unmethylated cytosine and the cytosine on the daughter strand would remain unmethylated (Tahiliani *et al*, 2009). This may be a possible pathway to passive demethylation in rapidly dividing cells such as ESCs (Williams *et al*, 2012). This theory is supported by the fact that brain cells are not rapidly dividing and still contain high levels of 5-hmC (Williams *et al*, 2012) and data that shows DNMT1 methylates 5-hmC enriched DNA with a lower efficiency to 5-mC hemi-methylated DNA (Valinluck & Sowers, 2007). Significant levels of 5-hmC have been found in various tissues and cell lines suggesting that the presence of this base may be too stable to function solely as an intermediate in a demethylation pathway.

1.2.4.2 5-hmC as a regulator of DNA methylation and a guardian of CpG Islands

The current model for the mechanism of DNA methylation maintenance does not include a proofreading mechanism for error correction (Section 1.2.3.1). In mESCs TET1 binds to a large number of genes and is particularly enriched at transcription start sites (TSSs). Most TET1 bound promoters are CGI-Ps and are mainly unmethylated in mESCs suggesting that TET1 could be a "guardian of CpG Islands" maintaining them free of DNA methylation. The enrichment of 5-hmC at promoter regions and TSSs and the strong evidence suggesting 5-hmC must be derived from pre-existing 5-mC implies that 5-mC is converted to 5-hmC specifically at CpG islands. This suggests a role for TET1 as a regulator of DNA methylation in these regions. Certain events such as increased DNMT activity can lead to random cytosine methylation of CpG islands. TET1 could convert this unwanted cytosine methylation into 5-hmC which could then be followed by active or passive demethylation to return affected cytosines to an unmethylated state. Down-regulation of TET1 during differentiation would then allow for the normal cytosine methylation pattern required to be achieved (*reviewed by* Williams *et al*, 2012).

1.2.4.3 5-hmC as a transcriptional regulator

5-mC at promoter regions has long been associated with gene repression and its oxidation to 5-hmC may extinguish the repressive effects of 5-mC (Branco *et al*, 2012). 5-hmC may alter the local chromatin environment though the recruitment or displacement of proteins (Williams *et al*, 2012). We have yet to discover any 5-hmC binding proteins but the epigenetic regulation of gene expression could be influenced by 5-hmC through the exclusion of methyl-CpG binding domain (MBD) proteins, a hypothesis which has been supported by the inability of methyl-CpG binding protein 2 (MeCP2) to recognise and bind to 5-hmC (Tahiliani *et al*, 2009; Valinluck *et al*, 2004). MBDs recognise methylated cytosines and have been suggested to play an important role in reading 5-mC (Fournier et al, 2012). However, most MBD proteins do not recognise 5-hmC (Jin *et al*, 2010; Valinluck *et al*, 2004) and most likely dissociate from DNA when 5-mC is converted into 5-hmC (Williams *et al*, 2012). The presence of just one 5-hmC base in a DNA fragment significantly inhibits binding of MeCPs (Ruzov et al, 2011; Valinluck et al, 2004).

Predictably it was hypothesised that the TET proteins would regulate transcription by adjusting DNA methylation levels at promoter regions and the

finding that both TET1 and 5-hmC localise to TSSs supported this (Williams *et al*, 2012). Gene expression profiling of TET1 depleted ESCs has found that less than 10% of TET1 target genes change expression after TET1 depletion (Williams *et al*, 2011; Wu & Zhang, 2011). The number of genes down-regulated was similar to the number of genes up-regulated indicating that TET1 and hydroxymethylation can have both repressive and active effects on gene expression (Williams *et al*, 2012).

1.2.5 The regulation of gene expression by DNA methylation

Approximately 70% of human genes, including the leptin gene, possess a CpG rich promoter region. DNA methylation in these regions is frequently related to transcriptional regulation of individual genes and DNA methylation is generally associated with gene silencing. Although evidence points towards the comprehensive role of DNA methylation being in the regulation of gene expression, its function varies with context (Jones & Liang, 2009).

CGI-Ps are usually assumed unmethylated in somatic cells and remain in this state throughout development in all tissues of an animal. However, studies in the mouse have determined that 5-34% of CGIs are tissue-specific differentially methylated regions (T-DMRs) which possess identical genetic information but different DNA methylation patterns between cells and tissues of an individual (Shiota *et al*, 2002; Song *et al*, 2005). DNA methylation at CGI-Ps is associated with gene silencing and there are examples of this demonstrated by X-chromosome inactivation and genomic imprinting where stable methylation is associated with gene repression. There are also examples of genes which exhibit intermediate levels of DNA methylation and yet are still expressed, one example being the leptin promoter (Stöger, 2006). This has questioned the function of DNA methylation in silencing gene expression. Many reports do not take into account the ratios of 5-mC and 5-hmC for DNA methylation. The re-discovery of 5-hmC necessitates the re-evaluation of published data regarding 5-mC. These two cytosine modifications could have different functions which may account for contradictory findings. Genome-wide studies have shown that the type of DNA methylation and its location within a gene is important in determining its function and, although data indicate that 5-mC and 5-hmC often co-exist, differences in their distribution have been observed. 5-hmC is enriched at promoter regions and TSSs (Ficz *et al*, 2011; Pastor *et al*, 2011; Williams *et al*, 2011; Wu & Zhang, 2011; Xu *et al*, 2011) whereas 5-mC is generally depleted from these regulatory regions (Williams *et al*, 2012). 5-hmC is rarely found in heterochromatic regions, which are known to have high 5-mC levels (Ficz *et al*,

2011; Pastor *et al*, 2011; Williams *et al*, 2011). 5-mC at TSSs prevents transcriptional initiation whilst 5-mC in gene bodies may stimulate transcriptional elongation (Bird, 1995). The enrichment of 5-hmC at promoter regions and TSSs, taken with the observation that 5-mC at TSS prevents transcriptional initiation, suggests that 5-hmC may be associated with transcriptional activation and gene expression. There is in fact a report that 5-hmC in the promoter region but not in the gene body strongly represses transcription by either preventing the binding of essential transcription factors or recruiting factors that repress transcription (Robertson *et al*, 2011). This was observed using *in vitro* transcription as a model system. *In vivo* many other factors may influence the regulatory effect of 5-hmC on gene transcription. In mESCs 5-hmC at promoter regions has been associated with gene repression whilst in gene bodies it is associated with transcriptionally active genes (Wu *et al*, 2011).

Although for many years it has been accepted that cytosine methylation is associated with gene repression there exists data which questions this. Predictably, the discovery of 5-hmC generated the hypothesis that 5-hmC was associated with gene activation whilst 5-mC was associated with gene repression. If 5-hmC was solely a transcriptional activator or repressor surely it would be expected that its distribution across tissues would be much more ubiquitous like that of 5-mC? From observations it is clear that the role of DNA methylation is dependent on the tissue, the gene of interest and the location and specific cytosine modification within the gene.

1.3 Current state of knowledge on DNA methylation at the leptin promoter

Leptin is encoded by the obese gene (*LEP*), the promoter region of which lies within a CpG island and exhibits a high level of conservation throughout the vertebrate animal kingdom (Zhang *et al*, 1994). Alignment of the core leptin promoter between many species reveals conserved CpG dinucleotides which could be subject to DNA methylation but much of the literature concerns only the human and mouse leptin promoters (Figure 1.1).

Within the mouse and human leptin promoter regions there are several sites which studies have found to be important for promoter activity and leptin gene transcription and expression. Both species contain a Ccaat enhancer binding protein alpha (C/EBPα) and Specificity protein 1 (Sp1) transcription factor binding site and a TATA box element. The mouse promoter contains an additional site designated LP1 which does not match the recognition sequence of

any known transcription factor but does bind a factor present in preadipocytes and adipocytes, the main site of leptin expression, but not in other cell types (Mason *et al*, 1998). The Sp1 and C/EBPα transcription factor binding sites and LP1 each contain CpG dinucleotides, which could be subject to DNA methylation.

The ability of DNA methylation to silence genes may be achieved by its interference with the binding of transcription factors that activate transcription from a specific gene. There are several transcription factors which are unable to bind to their recognition sequence if it contains a methylated CpG (Watt & Molloy, 1988). DNA methylation of the Sp1, C/EBPα and LP1 sites and CpG sites around them may be important in promoting or repressing leptin transcription. Sp1 can bind to the Sp1 site and activate the leptin promoter and, although it is thought that DNA methylation of this specific site does not directly affect Sp1 binding, DNA methylation around this site has been shown to prevent Sp1 from binding to this site (Zhu *et al*, 2003).

In a human adipose cell line it has been demonstrated that a partially methylated leptin promoter can reduce promoter activity by one third in comparison to an unmethylated promoter when driving the expression of a luciferase reporter gene (Melzner *et al*, 2002). Cytosine methylation of the CpG dinucleotide within the C/EBPα binding site led to a decrease in promoter activity, and cytosine methylation of the two CpG dinucleotides proximal to the TATA box element strongly reduced reporter expression (Melzner *et al*, 2002).

Given the evidence that methylation at the C/EBPα binding site can diminish leptin expression it would be expected that in adipose tissue, where leptin is highly expressed, there would be an absence of DNA methylation. Evidence shows that the leptin promoter is a T-DMR and in human adipose tissue, where leptin is at its most abundant, exhibits a high level of variation in the methylation density of individual epialleles ranging from 2 to 76% methylation (Stöger, 2006).

In imprinted genes and during X-chromosome inactivation, CpG islands are either unmethylated and transcriptionally active or densely methylated and silenced (Deaton & Bird, 2011). The function of intermediate and varying methylation densities is yet to be determined. Adipose tissue is a complex heterogeneous tissue consisting of several cell types (Cinti, 2005). The variation in methylation density between DNA molecules could be attributed to a difference in cell type and leptin expression level, although adipocytes are the predominant cell type. The variation in DNA methylation levels between DNA molecules may have been stably maintained in the adipocytes cell population.

However, adipocytes are formed from the differentiation of precursor cells known as preadipocytes which do not express leptin (Melzner et al, 2002). The variation in DNA methylation between adipose-derived DNA molecules may be attributed to preadipocytes that are undergoing epigenetic transition as they differentiate into mature adipocytes during a pathway to leptin expression (Stöger, 2006).

Methylation of the leptin promoter has also been observed in human and mouse cell lines which can be induced to differentiate into adipocytes. During *in vitro* differentiation of the 3T3-L1 preadipocyte cell line and, therefore, during the transition of leptin from a "switched off" to a "switched on" state, demethylation of the leptin promoter was observed, supporting the theory that partially methylated epialleles may be on a pathway to demethylation for leptin expression (Yokomori *et al*, 2002). Methylation of epialleles from adipose tissue is lower than in peripheral blood leukocytes (PBLs) supporting the idea that demethylation may occur in adipose tissue. The broad range of methylation densities observed in epialleles of human adipose tissue is also observed in human peripheral blood leukocytes but with the majority of epialleles exhibiting an intermediate level of methylation. In general methylation at CpG islands indicates permanent gene silencing but the moderate levels of methylation at the leptin promoter in adipose tissue and PBLs may suggest a less stable state of gene expression, and it has been suggested that this may facilitate the transitioning between active and inactive expression states during the life course of a cell or tissue and the methylation density of the epiallele may modulate the level of leptin expression (Stöger, 2006).

Interestingly, in human adipose tissue the frequency of methylation at the C/EBPα transcription factor binding site is 1.8-fold higher than the average frequency for other sites in the region and even in DNA molecules which are sparsely methylated at the leptin promoter only a few sites remain methylated but often alongside the C/EBPα transcription factor binding site. This is unexpected in adipose tissue as *in vitro* assays have shown that methylation of this site leads to a decrease in leptin promoter activity (Melzner *et al*, 2002). Although the C/EBPα transcription factor binding site exhibits significant levels of methylation in both human and mouse adipose tissue, electromobility shift assays have been used to demonstrate that C/EBPα , which is the predominant C/EBP present in mature adipocytes (Mason *et al*, 1998), can bind to both an unmethylated and a methylated (5-mC) binding site (Melzner *et al*, 2002). This suggests that methylation does not simply prevent transcription factors from

binding but may also be involved in recruiting them. Additionally, the methylation status of a transcription factor binding site may determine which transcription factor binds to the site. cAMP response element-binding protein (CREB) binds the CRE sequence, TGACGTCA, when the internal CpG is unmethylated. However, when the internal CpG is methylated, C/EBPα binds to CRE sequence and the binding of CREB is inhibited (Rishi *et al*, 2010). This information shows that several proteins may be able to bind to a transcription factor binding site and that the methylation status of a CpG within a transcription factor binding site may be important in determining which transcription factor is recruited or repressed.

The specific presence of 5-hmC has not yet been determined at the leptin promoter and it is possible that observed 5-mC patterns actually consisted of 5-mC and/or 5-hmC. As 5-mC is associated with transcriptional silencing, its presence at the leptin promoter in adipose tissue where leptin is highly expressed is somewhat unexpected. Although this may be explained by the hypothesis that DNA molecules exhibiting DNA methylation at the leptin promoter are derived from undifferentiated or differentiating preadipocytes that may be on a demethylation pathway towards a state of leptin expression, it is also plausible that this apparent 5-mC is actually 5-hmC; a mark which has been linked to transcriptional activation (Williams *et al*, 2011). Binding of MeCPs may contribute to the transcriptional silencing of the leptin gene when it is methylated (Pinnick & Karpe, 2011) but the presence of just one 5-hmC base in a DNA sequence can significantly inhibit binding of MeCPs (Ruzov et al, 2011; Valinluck et al, 2004). Although *in vitro* C/EBPα binds to an unmethylated or a methylated binding site, there is no information regarding the binding of C/EBPα to a hemi-methylated or hydroxymethylated binding site. Genome-wide studies have found that 5-hmC appears to be associated with transcription factor binding to distal regulatory sites during adipocytes differentiation of 3T3-L1 cells (Sérandour et al, 2012). Regions gaining 5-hmC are associated with gene expression in adipocytes. Hydroxymethylation of enhancers is an early event of enhancer activation and acquisition of 5-hmC in distal regulatory regions may represent a major event in the progression of an enhancer towards an active state and may participate in the selective activation of tissue-specific genes (Sérandour et al, 2012).

Human and mouse sperm DNA have also been sampled for methylation at the leptin promoter (Stöger, 2006). Predominantly hypomethylated epialleles were recovered from human and mouse sperm although a small number of methylated sequences were observed which may be a result of contamination

with somatic cells; but it cannot be ruled out that they do not originate from the gametes and there is evidence to show that some other genes carry methylated states in the male germ line. As leptin has not been shown to carry a parental specific imprint and the locus is generally unmethylated in sperm, this data suggests that *de novo* methylation of the leptin promoter generally occurs after fertilisation (Stöger, 2006).

1.3.1 Environmental influence on DNA methylation at the leptin promoter

Leptin has been identified as a gene which may be susceptible to environmentally influenced epigenetic variation, and there is evidence to show that DNA methylation at the leptin promoter in adult human tissues correlates with stresses experienced during early life. The environment we experience has an effect on our biology and although the mechanism by which an environment can influence phenotype is not fully understood most dietary, physical, chemical and stochastic environmental influences do not have the ability to promote a change in the genomic DNA sequence or permanently alter genetic processes (Ober & Vercelli, 2011). However, there is evidence to show that some environmental factors have the ability to induce stable alterations in the epigenome which could permanently alter the regulation of genome activity, allowing the genome a level of plasticity in response to the environment (Skinner, 2011). Environmental exposures have been found to influence histone modifications and miRNA expression (Ober & Vercelli, 2011), but most studies have focused on the effect that environmental and lifestyle factors have on DNA methylation (Alegría-Torres *et al*, 2011).

Critical periods during development have been identified where the effect of environmental exposures on DNA methylation have a more dramatic effect. DNA methylation is particularly labile during early development which can result in lifelong effects (*reviewed by* Faulk & Dolinoy, 2011). The more dramatic effect of environment at this stage may be attributed to its effect on epigenetic reprogramming, and to the high DNA synthetic rate, as epigenetic alterations that arise during periconception and embryogenesis would be dramatically amplified by high levels of cell division connected to the rapidly developing organs and systems. The alteration would therefore go on to affect a significantly higher number of cells in the offspring whilst an alteration to the DNA methylation pattern of adult cells would be restricted to those cells or their residing tissues (Feil & Fraga, 2011). This has led to the postulation that early

environmental experiences could be a predictor or indicator of certain phenotypes or disease (Faulk & Dolinoy, 2011).

The majority of studies looking into the effect of environment on DNA methylation have explored dietary factors such as total calorie intake, specific nutrient levels and phytochemicals. Nutrition across the life course is an important influence of adult health and disease risk, an effect which has been linked to the early dietary environment (Thayer & Kuzawa, 2011). The effect of diet on phenotype is demonstrated throughout the animal kingdom. For example, in the honey bees (*Apis mellifera*), early life nutrition is key in determining the type of honey bee that is produced (Drapeau *et al*, 2006). In mice, a prenatal methyl donor rich diet can change the colour coat pattern of mice carrying the agouti viable yellow allele (Waterland & Jirtle, 2003). In both of these classic examples, evidence suggests that DNA methylation influences phenotype (Faulk & Dolinoy, 2011; Gabor Miklos & Maleszka, 2011). Diet could impact DNA methylation by inhibiting the enzymes or restricting the dietary availability of substrates that are required for the process. DNA methylation is related to dietary intake of folate, methionine and S-adenosyl methionine, a derivative of methionine, which serves as a methyl donor and induces methyl group transfer in mammalian DNA methylation reactions. In mice, sustained methyl donor supplementation from periconception results in increased, genome-wide epigenetic variation which becomes progressively amplified in consecutive generations (Li *et al*, 2011). In the sheep, restricting vitamin B12, folate and methionine during periconception leads to adult offspring that are fatter and heavier and also have widespread epigenetic alterations to DNA methylation in the offspring (Sinclair *et al*, 2007).

Leptin has been identified as a gene which appears to be subject to dietary influences. In a study using human subjects that were conceived during the Dutch hunger winter, DNA methylation of the leptin gene was higher in individuals who were periconceptionally exposed to wartime famine when compared to that of their unexposed same sex siblings. In male subjects exposed late in gestation there were also differences in leptin promoter methylation indicating that the effect is timing and sex-specific (Tobi *et al*, 2009). A similar study investigating the effect of a low-protein diet during gestation concluded that mice born to mothers that were fed a low-protein-diet have a lower body weight. They also exhibit a higher food intake in comparison to mice born to mothers which were fed a control diet. These aberrations persisted throughout life and correlated with lower leptin mRNA expression and

protein expression (Jousse *et al*, 2011). Additionally, DNA methylation analysis revealed the demethylation of CpG dinucleotides of the promoter of leptin.

Whilst there appears to be an effect of prenatal diet on leptin promoter methylation, other studies investigating the effect of diet, but not prenatal diet, have found no difference in DNA methylation at the leptin promoter. In a mouse study, a high fat diet correlated with increased leptin expression but no change in leptin promoter methylation was observed (Fan *et al*, 2011).

1.4 Summary

Leptin is a key metabolic hormone and abnormal leptin expression levels have been linked to a variety of conditions and diseases. There is very little information regarding the transcriptional and epigenetic regulation of this gene. As its promoter region lies within a CpG rich region, DNA methylation may play a role in this process. The leptin promoter is a T-DMR and has been described to exhibit cytosine methylation in adipose tissue, the primary site of leptin expression. Cytosine methylation is generally associated with gene silencing but the techniques used to explore DNA methylation at the leptin promoter do not differentiate between 5-mC and 5-hmC, another epigenetic mark which may be involved in transcriptional regulation. Recent evidence shows that the tissue or cell type, the gene of interest and the location and type of DNA methylation present at a promoter are important in determining its function. There is also evidence to show that environmental stresses experienced during early life affect the DNA methylation pattern at the leptin promoter. It is therefore of importance to determine the mechanisms by which the leptin promoter becomes methylated, the specific cytosine modifications present at the leptin promoter and the effect of the environment on these modifications.

1.5 Development of working hypothesis

DNA methylation is the most intensively studied epigenetic mark, particularly in diseased states such as cancer (Sproul & Meehan, 2013). DNA methylation is usually associated with gene silencing and this is demonstrated in the processes of genomic imprinting and X-chromosome inactivation. However, we have yet to determine the comprehensive role of DNA methylation in fundamental processes of the life cycle.

Leptin is a key metabolic hormone and the leptin promoter, which is well conserved between many species, is a CpG rich region. DNA methylation in CpG rich regions is generally thought to regulate gene expression, however there is little information regarding the role of DNA methylation at the leptin promoter. Most studies investigating DNA methylation at the leptin promoter have focused on human, rat and mouse, however, leptin has become the focus of investigations in farm animals with regards to both meat and milk production. Adipose tissue not only affects meat quality and value but also provides energy reserves for pregnancy and lactation. Single nucleotide polymorphisms (SNPs) of leptin and the leptin receptor have been intensively studied and provided information to improve fertility in cattle and pigs and other processes (Wylie, 2011). There is no information regarding DNA methylation of the leptin promoter with regards to improving meat production, lactation and fertility in farm animals. It has also been found that prenatal diet can influence leptin promoter DNA methylation (Tobi *et al*, 2009). Therefore, characterisation of leptin promoter DNA methylation in farm animals may provide basic but valuable information for the improvement of reproduction and meat quality.

Whilst the enzymes which methylate DNA, the DNA methyltransferases (DNMTs) are well characterised, the specific mechanism by which DNA becomes methylated is not. *In vitro* studies have contributed much information to our knowledge of these mechanisms. However, *in vivo*, the DNMTs work in concert with other factors which enhance their activity. Additionally, there is now data derived from *in vivo* studies which questions the general perceptions regarding the individual contributions of the three catalytically active DNMTs, DNMT1, DNMT3a and DNMT3b to DNA methylation patterns. These general perceptions have led to the majority of scientists disregarding one strand of a DNA molecule as it is assumed both strands have the same methylation pattern. To address problems associated with investigating DNMT activity *in vitro*, mESCs with knockouts for the DNMTs were sourced for this study. Hairpin-bisulfite sequencing (Hp-bss), a technique which allows for the analysis of DNA

methylation patterns on both strands of an individual DNA molecule (Laird *et al*, 2004) was used in order to remove the bias that may be caused by analysing only one strand of a DNA molecule.

Until recently, it was thought by many that 5-methyl cytosine (5-mC) was the only cytosine modification present in mammalian DNA and subsequently the term "DNA methylation" predominantly refers to 5-mC. However, in the 1970s a report of another cytosine modification, 5-hydroxymethylcytosine (5-hmC), in mammalian DNA was published (Penn *et al*, 1972). At the time, 5-hmC was determined to be an oxidative damage product and deemed to be of little or no importance and subsequently forgotten. Recently, 5-hmC was rediscovered in the purkinje cells of the brain (Kriaucionis & Heintz, 2009). Additionally, it has been revealed that many techniques used for the analysis of DNA methylation do not distinguish between 5-mC and 5-hmC (Huang *et al*, 2010; Nestor *et al*, 2010). Consequently much of the published data regarding DNA methylation and specifically 5-mC may be skewed by the presence of 5-hmC. These two cytosine modifications may have different roles and whilst 5-mC is generally associated with gene silencing, reports have proposed both a gene silencing and gene expression role for 5-hmC amongst other possibilities. To investigate the relationship between DNA methylation at the leptin promoter and leptin expression the 3T3-L1 preadipocyte cell line, which do not express leptin, were differentiated into adipocytes, which do express leptin.

In summary, the main objectives of this study were to contribute to the gap in knowledge concerning DNA methylation at the leptin promoter, specifically:

i) The effect of a high-fat prenatal diet on leptin expression levels and DNA methylation levels,

ii) The role of the individual DNMT enzymes in methylating this region and

iii) The role of DNA methylation accounting for both 5-mC and 5-hmC in leptin gene expression.

Chapter 2. Characterisation of porcine leptin promoter DNA methylation in visceral and subcutaneous adipose tissue

2.1 Introduction

The leptin promoter displays a high level of conservation between many species (Figure 1.1 and Figure 2.14). DNA methylation at the leptin promoter has been studied mainly in human, mouse and rat; there is currently no information regarding the porcine leptin promoter. Characterising the porcine leptin promoter with regards to DNA methylation may be important to the successful farming of pigs in the context of both meat production and reproduction for the following reasons. Leptin levels have been shown to correlate with birth weight in humans (Marchini *et al*, 1998) and there is a documented relationship between birth weight and the subsequent growth and development of organs and tissues (Morise *et al*, 2008). Leptin resistance appears to contribute to low reproductive efficiency in pigs (Astiz et al, 2013).

Leptin is primarily expressed in adipocytes of white adipose tissue, a heterogeneous endocrine organ which presents itself in the form of two anatomically, structurally and functionally distinct depots. Whereas subcutaneous adipose tissue (SAT) is found in a fat layer directly under the skin, visceral adipose tissue (VAT) is located deeper within the body, surrounding the internal organs (Ibrahim, 2010; Wronska & Kmiec, 2012).

There is evidence to show that VAT and SAT express leptin at different levels, although this is yet to be determined in the pig. In humans several studies have found that SAT has a higher leptin mRNA expression level than VAT (Van Harmelen *et al*, 1998; Zha *et al*, 2009). In rodents differences in leptin expression have also been observed, but the highest expressing depot appears to be dependent on maturity. At 1-2 weeks of age SAT is the main site of leptin expression (Rayner *et al*, 1997) whilst in mature rats leptin expression is greater in the VAT (Maffei *et al*, 1995).

Adipocytes are formed from the differentiation of precursor cells known as preadipocytes. Differences in leptin expression between the two depots may be a consequence of different genetic differentiation programs of adipocyte precursers between the two depots (Tchkonia *et al*, 2007). Adipocytes clearly express leptin and although preadipocytes are considered not to express leptin (Wang *et al*, 2008) there is evidence of leptin expression in both human (Wabitsch *et al*, 1996) and porcine preadipocytes (Barb *et al*, 2001). The different ratio of preadipocytes to adipocytes between VAT and SAT (Wronska &

Kmiec, 2012) could also be a factor influencing leptin expression levels. In humans, the two depots are primarily composed of different types of adipocyte, with visceral adipocytes being smaller than subcutaneous adipocytes (Wronska & Kmiec, 2012). Leptin expression levels correlate with adipocyte size as larger adipocytes give higher leptin expression levels than small adipocytes (Margetic et al, 2002). Therefore, the ratio of small adipocytes to large adipocytes within the depot may also be a factor influencing leptin expression levels between VAT and SAT. If leptin is transcriptionally regulated via DNA methylation, then differences in DNA methylation levels between the two depots may also be apparent.

Sex-specific differences in adipose depots and leptin expression have also been observed. In humans, VAT accounts for more of the male total body fat than in females whilst women have more SAT than males (Karastergiou *et al*, 2012). This coincides with the observation of a sexual dimorphism in leptin mRNA expression where females have higher leptin expression (Montague *et al*, 1997). There is also a correlation between plasma leptin levels and birth weight. Infants born small for gestational age (SGA) have lower plasma leptin levels than infants born large for gestational age (LGA) (Marchini *et al*, 1998). However, there is no information regarding leptin expression levels between SGA infants and LGA infants in VAT and SAT.

In humans and mice the leptin promoter is a tissue-specific differentially methylated region (T-DMR) and evidence suggests its transcription and expression may be regulated epigenetically through DNA methylation (Melzner et al, 2002; Stöger, 2006). Despite reports comparing differences in leptin mRNA expression between VAT and SAT, males and females and SGA and LGA infants, there has been no reported attempt to determine if these variations in leptin expression correlate with differences in DNA methylation at the leptin promoter. Furthermore, leptin has been identified as a gene which is susceptible to epigenetic variation influenced by environmental factors such as prenatal nutrition. This was demonstrated in 2009 by Tobi *et al* who found DNA methylation of the leptin promoter in offspring conceived during the Dutch hunger winter to be higher than in their unaffected, same-sex siblings. The effect of prenatal diet on leptin promoter methylation has been further documented in a more recent study in mouse (Jousse et al, 2011). It was found that plasma leptin levels were lower in adult mice exposed to a low protein perinatal diet in comparison to mice on control diets. This coincided with lower

leptin promoter DNA methylation levels in the adult mice exposed to a low protein perinatal diet.

2.1.1 Aims and objectives

The aim of the work presented in this chapter was to investigate DNA methylation levels at the leptin promoter in porcine VAT and SAT by pyrosequencing, taking into account the effect of sex, size at birth and also maternal diet during pregnancy. Whilst studies have reported interesting observations regarding the influence of these factors on either leptin mRNA expression levels or DNA methylation at the leptin promoter, they rarely correlate both factors. The work presented in this chapter will therefore take both DNA methylation and mRNA expression into account.

2.2 Materials and methods

Pig tissue samples were obtained from an ongoing study by Dr Alison Mostyn (School of Veterinary Medicine and Science, University of Nottingham). All animal procedures described in this thesis were approved by the Ethics Committee for Animal Experiments of the Animal Sciences Group of Wageningen Research Centre, and conducted at Schothorst Feed Research in The Netherlands.

2.2.1 Rearing of animals and diets

Yorkshire x Landrace sows, reared at Schothorst Feed Research in the Netherlands, were artificially inseminated in order to ensure the accurate prediction of their farrowing date as commercially practised in the pig industry. Prior to this study all sows were administered the same diet. The experimental treatments started at the day of breeding when sows (n=13) were randomly assigned to one of two isocaloric gestational diets with varying starch and fat contents (Table 2.1; for full dietary composition see Appendix A).

Sows were fed according to an increasing feeding scheme during gestation (25.1 MJ NE/day from day 0 to 70, and 32.6 MJ NE/day from day 70 to 110) to meet increasing metabolic demand as gestation progressed. The high fat/low starch diet contained higher energy content. To account for this the daily feed allowance for these sows was reduced ensuring equivalent caloric intake between the two groups (Table 2.1). Sows of parities 2 to 6 were distributed equally between the two treatments, with a mean parity of 3.8 (SEM = 0.24). Sows of parity 2 received 95% of the feeding scheme. Diets were provisioned in a factory for research diets and according to the National Research Council

(NRC) nutrient recommendations for gestating and lactating sows. The control group were fed a high-starch diet as is commercial practise in the pig industry. Cereals or tapioca were provided as the source of starch. The high-fat diet contained palm oil, which has an equal ratio of saturated to unsaturated fatty acids, to increase fat content and had a reduced starch content to maintain equal energy intake.

Table 2.1 Feed quantity and composition of gestational diets. Sows were assigned to isocaloric gestational diets with varying starch and fat contents according the stage of gestation. The daily feed allowance of the sows subjected to a high fat diet was decreased to account for higher energy content, ensuring equivalent caloric intake between the two groups.

	Feed quantity during gestation (kg/day)			Diet composition (MJ NE/Kg)		
Gestational diet	**Day 0-40**	**Day 40-70**	**Day 70-110**	**Starch**	**Fat**	**Protein**
Control (n=13)	2.89	2.89	3.75	5.37	0.93	2.09
High-fat (n=13)	2.51	2.51	3.26	2.74	3.29	2.07

Control or high fat diets were fed from the day of mating (day 0) until the day of gestation (day 110). One week prior to the expected farrowing date, sows were transferred to farrowing crates and farrowed naturally. From day 110 and throughout lactation sows were administered the same diet for lactation (Table 2.2; for full dietary composition see Appendix A). Feeding levels was decreased before farrowing, and then steadily increased post-farrowing, as is normal practice in the pig industry (Table 2.2). The feed intake of each sow was recorded throughout pregnancy and lactation. After birth, piglets were allowed to suckle and extra piglets were removed before day three post-farrowing according to the number of functional teats. From week 2 of age all piglets were fed standard creep feed (Table 2.3; for nutrient composition see Appendix B) and at weaning (approximately day 28), piglets were housed in pens with pigs of similar body weights. From 3 weeks post-weaning piglets were administered a commercial post-weaning diet (Table 2.3). At five to six weeks post-weaning, piglets at approximately 25kg were placed in a pen with piglets of mixed sex and equal bodyweights and received a grower diet followed by a commercial diet for finishing pigs (Table 2.3). On reaching six months of age piglets were slaughtered at an approximate commercial slaughter weight of 110kg.

Table 2.2 Feed allowance of sows from day 110 of gestation and throughout lactation

Day	Feed allowance (Kg/day)
111-113	3.2
114-115	2.5
Farrowing	2.0
1	2.5
2	3.0
3	3.5
4	4.0
5	4.5
6	5.0
7	5.5
8	5.5
9-10	6.0
11-28	*Ad libitum*

Table 2.3 Composition of sow and piglet diets

	Diet composition (MJ NE/Kg)					
	Sow diet		**Piglet diet**			
	Prior to study	**During lactation**	**Creep feed**	**Post-weaning**	**Starter diet**	**Finisher diet**
Starch	5.65	5.42	5.65	6.12	6.44	6.39
Fat	2.25	2.33	2.78	2.20	1.82	1.69
Protein	2.65	2.75	2.99	2.98	2.89	2.72

2.2.2 Tissue collection

Tissue collection was performed at Schothorst Feed Research in the Netherlands. In this study, adipose tissues from six month old piglets of small and medium birth weight were sampled. Size category was determined at one week of age when piglets were assigned as being small or medium according to their size. The piglet with the median birth weight in each litter was selected as the medium birth weight sample. Often multiple piglets fell in the median birth weight range and where possible medium birth weight piglets were matched for sex with small birth weight piglet from the same litter. Castration of boars for meat quality is standard commercial practice however males selected for this study were not castrated.

Animals chosen for sampling were sedated by an intramuscular injection of a ketamin-xylazine combination (Ketamin 10%, Alfasan International BV, Woerden, The Netherlands). Piglets were weighed before euthanasia with T-61 dose (Intervet, Boxmeer, The Netherlands). Subcutaneous and visceral adipose tissue were collected, snap-frozen in liquid nitrogen and stored at -80◦C for later analysis.

2.2.3 DNA methylation analysis

2.2.3.1 Genomic DNA (gDNA) extraction

Genomic DNA (gDNA) was extracted using the DNeasy Blood and Tissue kit (Qiagen Ltd., West Sussex, UK). In brief, cells were suspended in 200µl ALT lysis buffer containing 20µl proteinase K solution (both supplied with kit). Lysis was performed by overnight incubation at 56°C on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) and complete disruption was ensured by adding 200µl buffer AL. After the addition of 200µl absolute ethanol, the lysate was passed through a DNeasy spin column to allow gDNA to bind the column matrix. After several wash steps with wash buffers provided to remove residual material, gDNA was eluted from the column in 100µl buffer AE. gDNA was tested for quality and concentration using the Nanodrop ND1000 UV-Vis spectrophotometer (Nanodrop, Wilmington, USA) and by electrophoresing 2µg nucleic acid on an agarose gel (Section 2.2.3.3.3). gDNA was stored at -20°C.

2.2.3.2 Bisulfite Conversion of gDNA

gDNA was converted using a bisulfite conversion protocol as described in (Genereux *et al*, 2008) which results in the conversion of cytosine, 5-carboxylcytosine and 5-formylcytosine to uracil but leave 5-mC and 5-hmC unchanged. DNA was denatured by suspension in 20µl 0.3M sodium hydroxide and incubation at 37°C for 30 minutes on a thermal cycler (Eppendorf, Stevenage, UK). 180µl bisulfite solution consisting of a 45% ammonium bisulfite solution (A1145-51002, Spectrum Laboratories, USA) containing 4M sodium bisulfite and 1M ammonium sulphite monohydrate was pre-incubated at 70°C and added to the denatured DNA. DNA was converted by incubation at 70°C for 80 minutes using a thermal cycler (Eppendorf, Stevenage, UK).

The bisulfite treated DNA was purified using the Qiaquick PCR purification kit (Qiagen Ltd., West Sussex, UK). DNA was mixed with 1ml DNA binding buffer (PB, supplied with kit) and passed through a PCR purification spin column. After a wash step with buffer PE (supplied with kit) converted DNA was eluted from the column with 45µl elution buffer (supplied with kit).

Desulphonation was performed by addition of 5µl 3M sodium hydroxide and incubation at 37°C for 20 minutes on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA). DNA was purified by passing samples through an Illustra S-200 spin column (27-5120-01, GE Healthcare, Buckinghamshire, UK) and stored at -20°C.

2.2.3.3 Pyrosequencing

2.2.3.3.1 Pyrosequencing assay design

The Pig (*Sus scrofa*) Leptin (*LEP*) gene promoter region sequence (accession number GenBank: AF492499.2; 16,810 bp) was obtained from the nucleotide database by the National Center for Biotechnology Information, NCBI (http://www.ncbi.nlm.nih.gov/nuccore/AF492499). The assay was designed to interrogate nine CpGs of the core leptin promoter from 5,602-5,665bp (Appendix E). CpG-5 is located within a C/EBPα transcription factor binding site. The assay was designed to produce one PCR product from which two sequencing primers would be used to pyrosequence the region, one for CpGs 1-3 and one for CpGs 4-9. PCR primers were designed to amplify the sense strand of the porcine leptin promoter and sequence the antisense strand of the PCR amplicon.

2.2.3.3.2 Pyrosequencing PCR

Converted DNA (Section 2.2.3.2) was amplified in a PCR reaction (Table 2.4). Primers were obtained HPLC purified from Eurofins (Eurofins MGM Operon, Ebersberg, Germany). Primer sequences were forward: 5' TTGAGGAATTTTC GGTTGTTATTTTGT 3' and reverse: 5' BIO-TACCCTCCCCTCTTATAAA 3'. The reverse primer was biotinylated at the 5'end in order to sequence the antisense strand of the PCR amplicon. The expected PCR product size was 310bp. A no-template control was included to ensure amplifications were contamination free.

Table 2.4 PCR reaction mixture

Reagent	Volume per sample
HotStar Taq Mastermix	12.5µl
Nuclease free water	4µl
Forward primer (10mM)	1µl
Reverse primer (10mM)	1µl
Bisulfite converted DNA (from 2µg gDNA)	6.5µl
	TOTAL 25µl

2.2.3.3.3 Agarose gel electrophoresis

PCR products were electrophoresed through a 1.5% agarose gel in 1x tris-acetate-EDTA (TAE) electrophoresis buffer. Gels contained 1:10,000 RedGel nucleic acid stain (Biotium Inc., California, USA) to allow the visualisation of the PCR product under U.V. light. A 100bp DNA ladder (Promega, Southhampton, UK) was used to determine PCR product size. Gels were visualised using the Doc-It Imager (UVP, California, USA).

2.2.3.3.4 Pyrosequencing procedure

10-20µl of PCR product was used in pyrosequencing reactions to determine the percentage of methylated/hydroxymethylated cytosines at each of the nine CpG sites analysed (Figure 2.14). The technique was performed as recommended by the manufacturer and all reagents were obtained from Qiagen unless otherwise stated. PCR products were shaken on the Thermomixer Comfort (Eppendorf,

Stevenage, UK) at 1400 rpm for 5 minutes with 40µl binding buffer, 18µl water and 2µl streptavidin sepharose beads (17-5113-01, GE Healthcare, Buckinghamshire, UK) to allow binding of beads to the biotinylated strand of the PCR products.

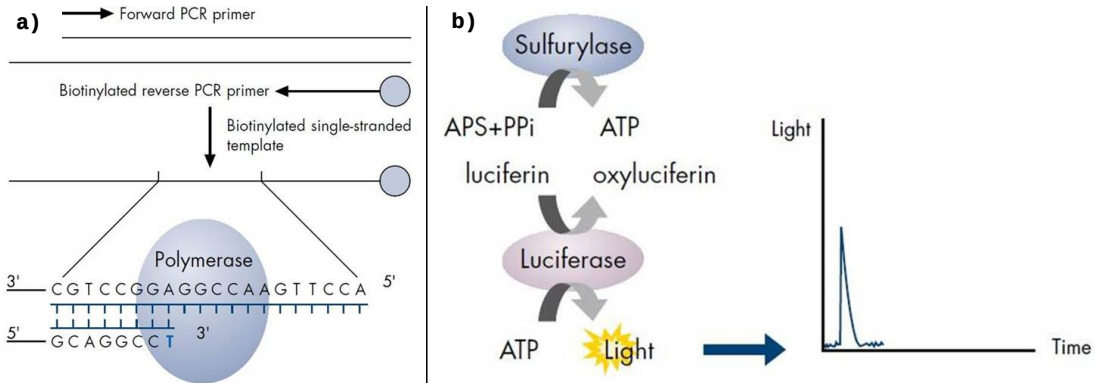

Figure 2.1 An overview of Pyrosequencing for CpG DNA methylation analysis. a) gDNA is bisulfite converted and the region of interest is amplified in a PCR reaction using a biotinylated reverse primer. The biotinylated strand of the PCR product is then purified and annealed to a sequencing primer. b) The pyrosequencer sequentially releases adenine, guanine, cytosine and thymine nucleotides and incorporation of a nucleotide into the sequence results in a chemical reaction from which light is produced. The intensity of the light produced is relative to the number of bases incorporated into the DNA and therefore the ratio of cytosines to thymines can be determined. *Adapted from Qiagen pyrosequencing handbook.*

After cleaning with 70% ethanol, denaturing with denaturing solution and purification with wash buffer to remove the unbiotinylated DNA strand using the Pyromark Q24 Vacuum Workstation (Qiagen Ltd., West Sussex, UK), single stranded biotinylated DNA was eluted into 25µl annealing buffer containing 0.3µM sequencing primer CpGs 1-3: 5'ACGTTGTTTCGTTAGGCGGGG3′ and CpGs 4-9: 5'GGGTGGGTTAGGGTTTGT3′. Biotinylated DNA was incubated at 80°C for 3 minutes on preheated heat blocks supplied by Qiagen to allow annealing of the sequencing primer and then left to cool for 5 minutes.

A pyrosequencing dispensation cartridge was filled with Pyromark Gold Q24 reagents, volumes as suggested by pyrosequencing software, and assays were performed on the Pyromark Q24 Pyrosequencer (Qiagen Ltd, West Sussex, UK). Data was analysed using Qiagen Pyromark Q24 software. Analysis parameters were as recommended by Qiagen.

2.2.4 mRNA expression analysis

2.2.4.1 Total RNA extraction

Total RNA was extracted using Trireagent (Ambion, California, USA). All centrifugation steps were performed at 4°C. Cells were homogenised in 1ml Trireagent solution using the Polytron PT4000 homogeniser (Kinematica AG, Luzern, Switzerland). Lysates were incubated at room temperature for 5 minutes before centrifuging at 10,000 rpm for 10 minutes. 100µl 1-bromo 3-chloropropane (BCP) was added to the supernatant which was shaken vigorously for 15 seconds, incubated at room temperature for 15 minutes and centrifuged at 13,000 rpm for 15 minutes. 500µl isopropanol was added to the RNA containing aqueous phase and the sample was shaken vigorously for 15 seconds, incubated at room temperature for 15 minutes and centrifuged at 10,000 rpm for 10 minutes. The supernatant was discarded and the RNA pellet was purified with 1ml 75% ethanol and centrifuged at 10,000 rpm for 5 minutes. Ethanol was removed and RNA was allowed to air dry before dissolving the pellet in 30µl nuclease free water.

Total RNA were tested for quality and concentration using the Nanodrop and agarose gel electrophoresis (Section 2.2.3.1). RNA was stored at -80°C.

2.2.4.2 cDNA synthesis

Residual gDNA was removed from 1µg RNA by treatment with RQ1 RNase-free DNase (Promega, Southampton, UK). RNA was incubated with DNase at 37°C for 30 minutes on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) after which 1µl of DNase stop solution was added and the mixture was incubated at 65°C for 10 minutes to terminate the reaction.

All reverse transcription (RT) reagents were obtained from Invitrogen Ltd. (Paisley, UK) unless otherwise stated. The DNase treated RNA was primed for first strand cDNA synthesis using 1µl 50µM oligo(Dt)20 and 1µl 10mM dNTPs. Reactions were incubated at 65°C for 5 minutes on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) and snap-cooled on ice for 5 minutes. RT reaction mixtures were performed using a thermal cycler (Eppendorf, Stevenage, UK) at 50°C for 1 hour and 75°C for 15 minutes to inactivate the RT enzyme (Table 3.8). The resulting cDNA was used directly in rt-PCR reactions. cDNA was stored at -20°.

Table 2.5 RT reaction mixture for cDNA synthesis

Reagent	Volume per sample
Primed RNA	13µl
First strand buffer	4µl
0.1M DTT	1µl
Nuclease free water	1µl
Superscript III reverse transcriptase	1µl
	TOTAL 20µl

2.2.4.3 Quantitative PCR (qPCR)

Relative leptin mRNA expression was calculated by quantitative PCR (qPCR) using sybr green and the Roche lightcycler 480 amplification system (Roche Diagnostics Ltd., West Sussex, UK). Amplifications were performed in duplicate (Table 2.6).

Table 2.6 Reaction mixture for qPCR

Reagent	Volume per sample
2x Sybr green	10µl
Forward primer (5µM)	0.25µl
Reverse primer (5µM)	0.25µl
Nuclease free water	8.5µl
cDNA	1µl
	TOTAL 20µl

Primers were obtained high purity salt free (HPSF) purified from Eurofins (Eurofins MGM Operon, Ebersberg, Germany) and were designed to cross intron-exon boundaries to prevent the amplification of gDNA (Table 2.7).

Table 2.7 Primer sequences for qPCR

Gene of interest	Primer sequence	Amplicon size (bp)	Primer efficiency
Leptin	Forward: 5' CATGCAGTCTGTCTCCTCCA 3' Reverse: 5' GAGGTTCTCCAGGTCATTCG 3'	172	1.72
ß-Actin	Forward: 5' TCCCTGGAGAAGAGCTACGA 3' Reverse: 5' CGCACTTCATGATCGAGTTG 3'	151	1.92
RPL27	Forward: 5' CCTCATGCCCACAAGGTACT 3' Reverse: 5' CAAGAAGCAGGCAGACACAG 3'	240	1.64
AZ1	Forward: 5' GAGTTCCAGGGTCTCCATCA 3' Reverse: 5' AGGCAGCGCATACTGCAGGATCCGG 3'	220	1.52

Primers were optimised for annealing temperature, concentration and specificity by performing a melt curve and by agarose gel electrophoresis (Section 2.2.3.3.3). Standard curves were performed for all primer sets using purified PCR products in order to calculate primer efficiencies. A calibrator sample (from the standard curve) was included in every plate in order to account for any variation in PCR efficiency between plates. No-template controls were also included to detect contamination. Relative expression was normalised against three endogenous control genes; ß-Actin, RPL27 and AZ1 which have been deemed suitable for use as housekeeping genes in porcine adipose tissue (Piórkowska et al, 2011).

96 well PCR plates (LightCycler 480 multiwell plate, Roche) were sealed using sealing foil (LightCycler 480 sealing foil, Roche) and centrifuged for 2 minutes prior to amplification. Amplifications were performed using the following thermal profile: 95°C 10 minutes, 95°C 10 seconds, 60°C 50 seconds and 72°C 1 second over 45 cycles and 40°C 10 seconds.

2.2.5 Statistical Analysis

The imbalanced nature of this study, with respect to animal numbers per dietary treatment, piglet sex and size at birth (Appendix C, Appendix D) prohibited conventional ANOVA. Instead data were analysed using a mixed linear restricted maximum likelihood (REML; Genstat, 2011) algorithm in which piglet sex, size at birth, maternal diet, CpG site and maternal diet x CpG site or relative expression formed the fixed effects, and birth mother (sow) and birth mother x piglet formed the random effects. In so doing this model effectively recognised the birth mother as the experimental unit and so correctly apportioned degrees of freedom for the analysis of expression and methylation between experimental factors (i.e. maternal diet, piglet sex and size) and between individual CpGs within piglet. Methylation data were not normally distributed and so were $\log_{10}$ transformed prior to analysis. Methylation data are therefore presented as $\log_{10}$ means (± S.E.M.), and back-transformed geometric means.

2.3 Results

2.3.1 Visceral vs. subcutaneous adipose tissue depots

Differences in cellular composition and leptin expression levels between adipose tissue depots may be reflected by different DNA methylation levels. To assess this, visceral and subcutaneous adipose tissue derived from the same control pigs were pyrosequenced for DNA methylation levels at the leptin promoter (Figure 2.2).

The regional average, defined as the mean methylation level of all nine CpGs between all samples in each group, was compared followed by a comparison of DNA methylation at each of the nine CpGs interrogated. The regional average DNA methylation level was the same between the subcutaneous and visceral tissues at 5% (Figure 2.2a).

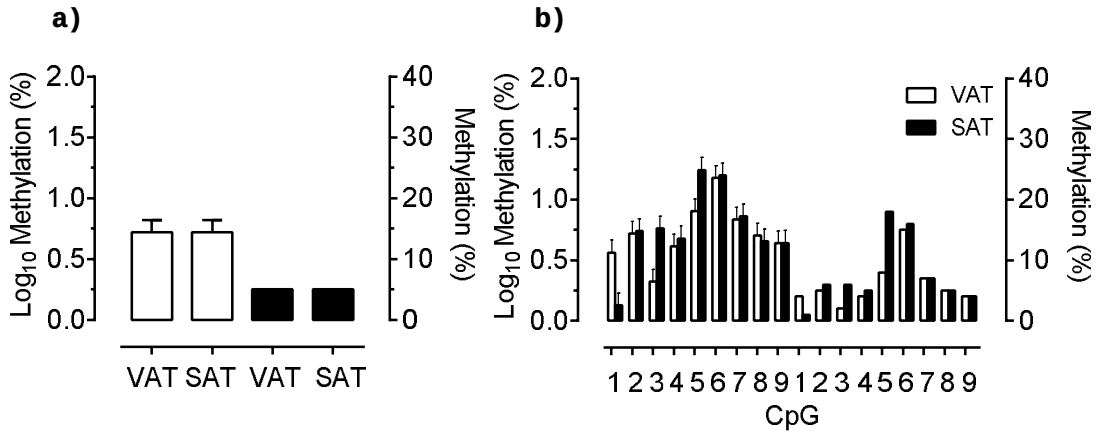

Figure 2.2 Leptin promoter DNA methylation in visceral (VAT) and subcutaneous (SAT) adipose tissue. The percentage DNA methylation determined by pyrosequencing at nine CpGs of the core leptin promoter were plotted as Log_{10} values (left axis) and geometric means (right axis). a) The regional average DNA methylation level across the nine CpGs. White bars indicate Log_{10} values, black bars indicate geometric means b) DNA methylation levels at the nine individual CpGs. White bars indicate VAT, black bars indicate SAT.

When analysing DNA methylation levels at individual CpGs, no significant difference between the two depots was observed (Figure 2.2b). However, the differences in DNA methylation levels between individual CpGs within each tissue type were significant; visceral ($p=0.002$) and subcutaneous ($p=0.002$). With the exception of CpG-5 and CpG-6, all CpG sites exhibited low DNA methylation levels of between 1-6% in both tissue types. This is comparable to the regional average of 5% (Figure 2.2). At CpG-5, located within a C/EBPα transcription

factor binding site, a 3% increase relative to the regional average was observed in the visceral adipose tissue and in the subcutaneous adipose tissue the increase was more noticeable at 13%. At CpG-6, located ten nucleotides (nt) proximal to a TATA box element, higher levels of DNA methylation relative to the regional average of 10% and 11% were also observed in both the visceral and subcutaneous depot respectively.

Despite there being no significant difference in DNA methylation, leptin mRNA expression levels were significantly higher in the visceral depot ($p=<0.001$) (Figure 2.3).

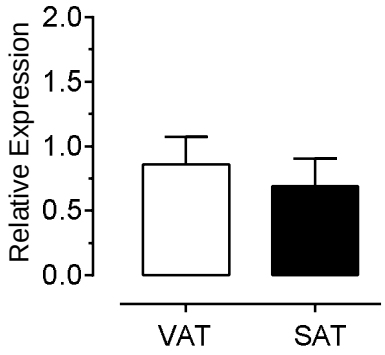

Figure 2.3. Leptin mRNA expression in visceral (VAT) and subcutaneous (SAT) adipose tissue. mRNA expression levels were quantified by qPCR and normalised against three housekeeping genes; β-actin, RPL27 and AZ1.

2.3.2 Male vs. female

Pyrosequencing was used to assess sex-specific differences in DNA methylation at the leptin promoter in both visceral (Figure 2.4) and subcutaneous (Figure 2.6) adipose tissue depots. Uncontrollably the number of male piglets obtained for the study (n=11 for visceral and n=5 for subcutaneous) was higher than the number of female piglets obtained (n=2 for visceral and subcutaneous).

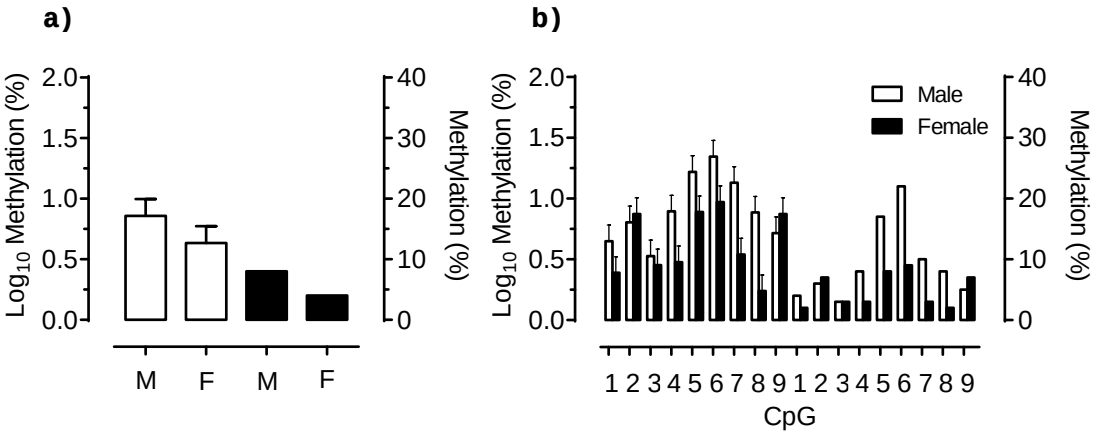

Figure 2.4 Leptin promoter DNA methylation in the visceral adipose tissue of males (M) and females (F). The percentage DNA methylation determined by pyrosequencing at nine CpGs of the core leptin promoter were plotted as Log_{10} values (left axis) and geometric means (right axis). a) The regional average DNA methylation level across the nine CpGs. b) DNA methylation levels at the nine individual CpGs.

In visceral adipose tissue, no significant difference was observed between males and females when analysing both the regional average (Figure 2.4a) and CpG specific (Figure 2.4b) DNA methylation levels. However, differences in DNA methylation levels between individual CpGs within each group were significant; male (p=0.002) female (p=0.002). With the exception of CpG-5 and CpG-6, low levels of DNA methylation between 2 and 8% were observed at all CpGs in both sexes. In males and females, CpG-5 and CpG-6 were more highly methylated than the other CpGs in the region. CpG-5 was 9% and 4% more highly methylated than the regional average in males and females respectively, whilst CpG-6 was 14% and 15% more highly methylated than the regional average in males and females respectively.

There was also no significant difference in leptin mRNA expression levels between the sexes (Figure 2.5).

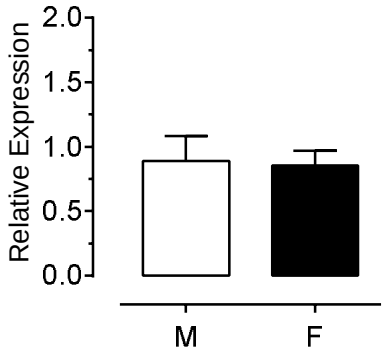

Figure 2.5. Leptin mRNA expression in visceral adipose tissue of males (M) and females (F). mRNA expression levels were quantified by qPCR and normalised against three housekeeping genes; β-actin, RPL27 and AZ1.

In subcutaneous adipose tissue, no significant difference was also observed between males and females when analysing both the regional average (Figure 2.6a) and CpG specific (Figure 2.6b) DNA methylation levels.

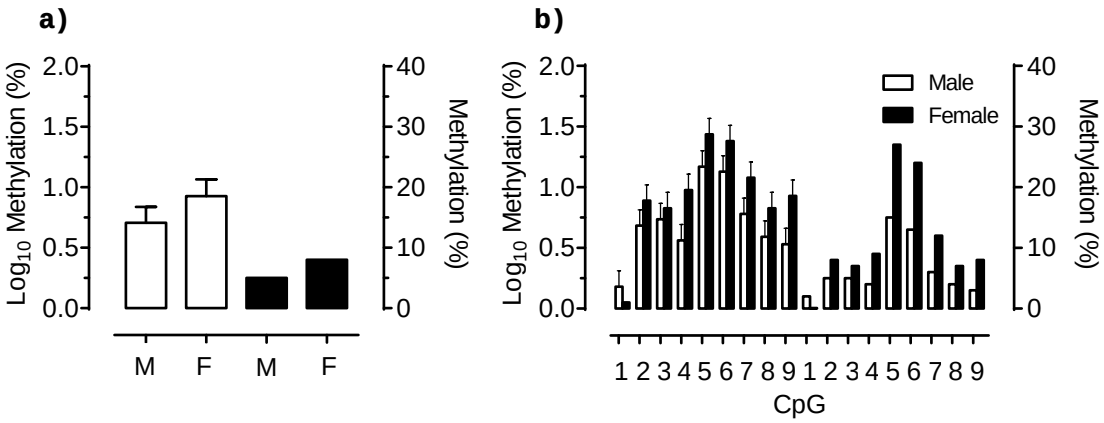

Figure 2.6 Leptin promoter DNA methylation in subcutaneous adipose tissue of males and females. The percentage DNA methylation determined by pyrosequencing at nine CpGs of the core leptin promoter were plotted as Log_{10} values (left axis) and geometric means (right axis). a) The regional average DNA methylation level across the nine CpGs. b) DNA methylation levels at the nine individual CpGs.

However, as observed in the visceral depot, differences in DNA methylation levels between individual CpGs within each group were significant; male ($p=0.005$) female ($p=0.005$). With the exception of CpG-5 and CpG-6, all CpGs exhibited low levels of DNA methylation between 0 and 9% in both sexes. CpG-5 was 10% and 19% more highly methylated than the regional average in the males and females respectively whilst CpG-6 was 8% and 16% more highly methylated than the regional average in the males and females respectively.

In the subcutaneous depot there was also no significant difference in leptin mRNA expression levels between the sexes (Figure 2.7).

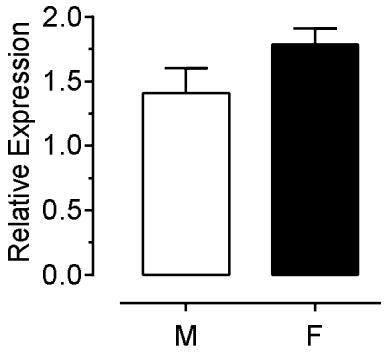

Figure 2.7. Leptin mRNA expression in subcutaneous adipose tissue of males (M) and females (F). mRNA expression levels were quantified by qPCR and normalised against three housekeeping genes; β-actin, RPL27 and AZ1.

2.3.3 Small vs. medium birth weight pigs

It has been reported that the plasma leptin levels of infants born small for gestational age are lower than that observed in infants born large for gestational age (Marchini *et al*, 1998). Pyrosequencing was used to assess the influence of birth weight on DNA methylation at the leptin promoter in adipose tissue (Figure 2.8).

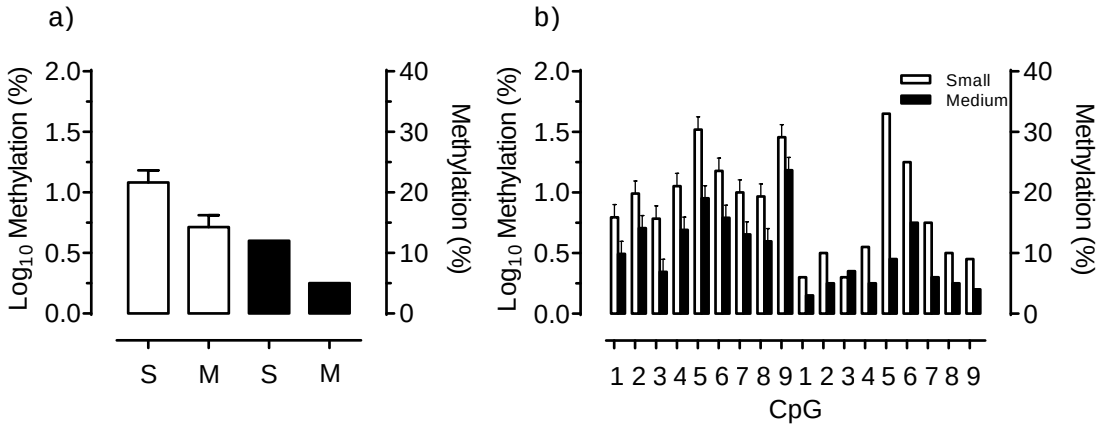

Figure 2.8 Leptin promoter DNA methylation in visceral adipose tissue of small (S) and medium (M) birth weight pigs. The percentage DNA methylation determined by pyrosequencing at nine CpGs of the core leptin promoter were plotted as Log_{10} values (left axis) and anti-logged values (right axis). a) The regional average DNA methylation level across the nine CpGs. b) DNA methylation levels at the nine individual CpGs.

This analysis was performed with only visceral adipose tissue as subcutaneous tissue from low birth weight pigs was not available. As described in Section 2.22, categorisation of small (n=5) and medium (n=8) birth weight piglets was determined at one week of age. Piglets were assigned small or medium according to their size where the piglet with the median birth weight in each litter was selected as the medium birth weight sample.

No significant difference was observed between small and medium birth weight pigs when analysing both the regional average (Figure 2.8a) and CpG specific DNA methylation levels (Figure 2.8b). However, there was a significant difference in DNA methylation levels between CpGs in the same group; small ($p=0.003$) medium ($p=0.003$). With the exception of CpG-5, CpG-6 and CpG-7, each CpG site was characterised by low levels of methylation in both groups.

CpG-5, CpG-6 and CpG-7 exhibited higher levels of DNA methylation relative to the regional average in both groups. CpG-5 was 21% and 4% more highly methylated than the regional average in the small and medium birth weight tissues respectively whilst CpG-6 was 17% and 10% more highly methylated than the regional average in the small and medium birth weight tissues respectively. CpG-7, located six nt proximal to a TATA box element, was also more highly methylated than the regional average being 3% more highly methylated in the small birth weight tissues and 10% more highly methylated in the medium birth weight tissues.

Leptin mRNA expression was significantly higher in small birth weight pigs ($p=0.003$) (Figure 2.9).

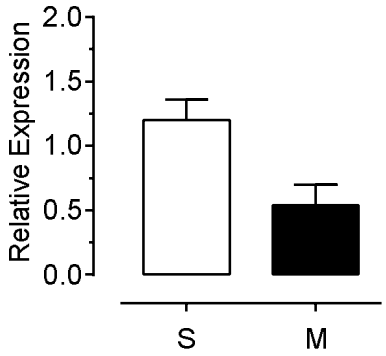

Figure 2.9. Leptin mRNA expression in visceral adipose tissue of small (S) and medium (M) birth weight pigs. mRNA expression levels were quantified by qPCR and normalised against three housekeeping genes; β-actin, RPL27 and AZ1.

2.3.4 Control vs. high-fat diet

To assess the effect of a high fat prenatal diet on leptin promoter DNA methylation, pyrosequencing was used to determine DNA methylation levels at the leptin promoter in the visceral and subcutaneous adipose tissue depots of pigs which experienced a high fat prenatal diet from conception to birth (Figure 2.6). For visceral tissues $n=13$ for both control and high fat tissues and for the subcutaneous depot for the control tissues $n=7$ and for high fat tissues $n=8$ where the experimental unit is the sow. For the visceral depot, tissue from 13 control piglets and 15 high fat piglets was tested and for the subcutaneous depot tissue from 7 control piglets and 11 high fat piglets were tested.

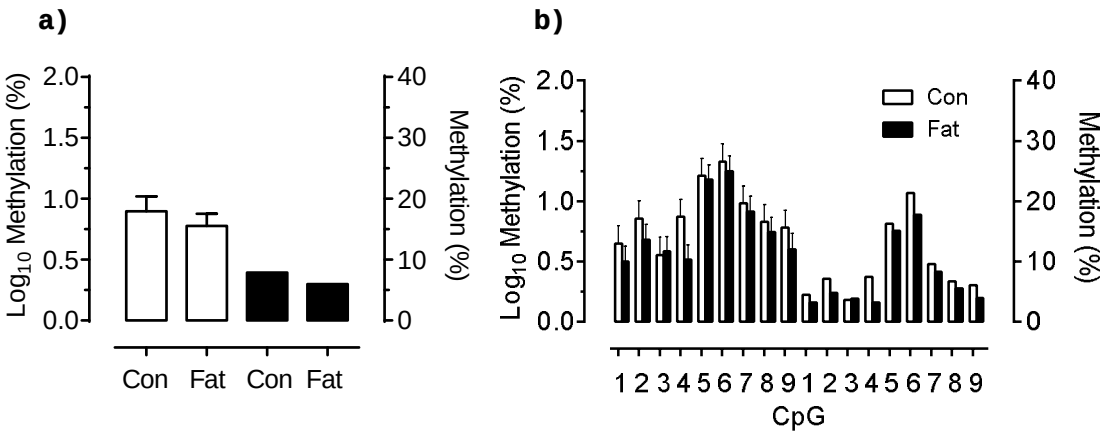

Figure 2.10 Leptin promoter DNA methylation in visceral adipose tissue of pigs subjected to a high fat prenatal diet. The percentage DNA methylation determined by pyrosequencing at nine CpGs of the core leptin promoter was plotted as Log_{10} values (left axis) and geometric means (right axis). a) The regional average DNA methylation level across the nine CpGs. b) DNA methylation levels at the nine individual CpGs. Control=Con, High fat= Fat

In the visceral adipose tissue, the leptin promoter was characterised by low levels of DNA methylation which were comparable between the control and high fat diet groups (Figure 2.10). No interaction between DNA methylation levels and prenatal diet was observed.

There was also no significant difference in leptin mRNA expression between visceral tissues from control and high-fat diet treatments (Figure 2.11).

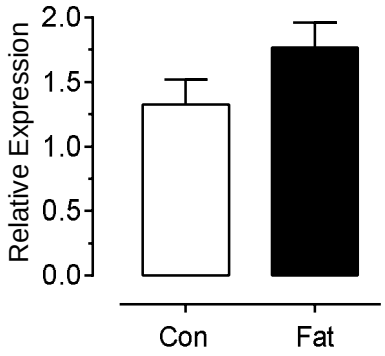

Figure 2.11. Leptin mRNA expression levels in visceral adipose tissue of pigs subjected to a high-fat prenatal diet. mRNA expression levels were quantified by qPCR and normalised against three housekeeping genes; β-actin, RPL27 and AZ1. Control=Con, High fat= Fat

In the subcutaneous adipose tissue, the leptin promoter was also characterised by low levels of DNA methylation which were comparable between the control and high fat diet groups (Figure 2.12). Again, no interaction between DNA methylation levels and prenatal diet was observed.

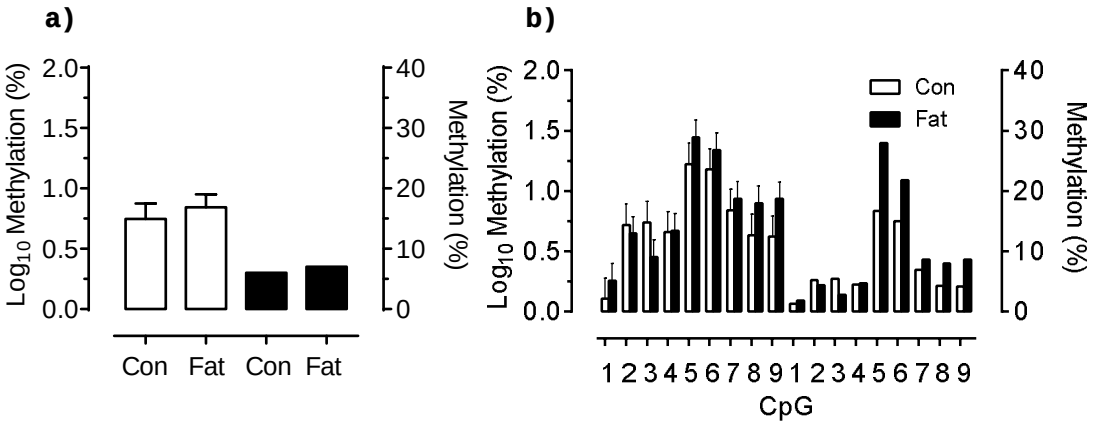

Figure 2.12 Leptin promoter DNA methylation in subcutaneous adipose tissue of pigs subjected to a high fat prenatal diet. The percentage DNA methylation determined by pyrosequencing at nine CpGs of the core leptin promoter was plotted as Log_{10} values (left axis) and geometric means (right axis). a) The regional average DNA methylation level across the nine CpGs. b) DNA methylation levels at the nine individual CpGs. Control=Con, High fat= Fat

As with the visceral tissue, there was also no significant difference in leptin mRNA expression levels between the subcutaneous control and high-fat diet tissues (Figure 2.13).

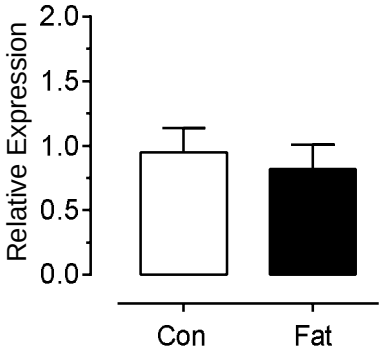

Figure 2.13. Leptin mRNA expression levels in the subcutaneous adipose tissue of pigs subjected to a high-fat prenatal diet. mRNA expression levels were quantified by qPCR and normalised against three housekeeping genes; β-actin, RPL27 and AZ1. Control=Con, High fat= Fat

2.4 Discussion

2.4.1 DNA methylation and leptin expression at the leptin promoter in porcine adipose tissue depots

2.4.1.1 Visceral vs. subcutaneous

Adipose tissue is a heterogeneous tissue consisting of several cell types including adipocytes and non-fat cells such as inflammatory cells, immune cells, preadipocytes and fibroblasts, connective tissues, vascular and neural tissues (Ibrahim, 2010). VAT and SAT show structural, functional and anatomic differences. In comparison to SAT, VAT is described as being more cellular, vascular and innervated and containing a larger number of inflammatory and immune cells. The two depots also have different genetic differentiation programs of adipocyte precursors, a different ratio of preadipocytes to adipocytes and differ in adipocyte morphology. VAT contains a higher proportion of large adipocytes whilst SAT is primarily composed of small adipocytes (Ibrahim, 2010; Wronska & Kmiec, 2012).

There is also evidence to suggest a depot specific variation in leptin expression levels. In a study by Van Harmelen *et al*, leptin mRNA levels and leptin secretion rates were higher in the subcutaneous depot of both non-obese and obese women (Van Harmelen *et al*, 1998). This finding was confirmed by Zha *et al* who found that leptin mRNA levels were higher in the subcutaneous depot of non-obese and obese males and females (Zha *et al*, 2009) but a small study in four obese humans using in situ hybridisation found no difference in leptin expression between the two depots (Lönnqvist *et al*, 1995). In rodents, site specific leptin expression appears to be dependent on age. At 1-2 weeks of age subcutaneous fat is the main site of leptin expression (Rayner *et al*, 1997) whilst in mature rats leptin expression may be lower in subcutaneous fat that in visceral fat (Maffei *et al*, 1995). Here, it was found that porcine VAT expressed leptin more highly that SAT (Figure 2.3). Expression of the leptin gene may be regulated epigenetically through DNA methylation, however, this did not correlate with differences in DNA methylation between depots (Figure 2.2).

2.4.1.2 Male vs. female

In humans VAT accounts for more of the male total body fat than in females whilst women have more SAT than males. Females are known to have higher leptin levels than males, an observation which has been proposed to be attributed to the stimulating effect of estrogens and the repressing effect of testosterone (Kennedy et al, 1997) whilst others have been unable to find a

correlation between sexual dimorphism and sex hormones (Margetic et al, 2002). Here, no statistically significant difference in leptin mRNA expression levels was found between males and females in VAT (Figure 2.5) or SAT (Figure 2.7). DNA methylation levels at the leptin promoter were compared between male and female pigs in both adipose tissues types to determine if there was a gender specific difference (Figure 2.4 and Figure 2.6). No significant differences were observed.

2.4.1.3 Small vs. medium birth weight

Differences in plasma leptin levels between large for gestational age (LGA) and small for gestational age (SGA) infants have been noted where plasma leptin levels were lower in SGA infants than in LGA infants (Marchini et al, 1998). The VAT of pigs assigned small and medium in size at birth was analysed to determine if size had an effect on DNA methylation at the leptin promoter or expression levels of the leptin gene. There was no statistically significant difference in DNA methylation levels at the leptin promoter (Figure 2.8). However, small birth weight pigs had higher leptin mRNA expression ($p=0.003$) (Figure 2.9).

This analysis could not be performed on SAT due to a lack of small control pigs. It has been suggested that SAT is the major source of leptin (Wajchenberg, 2000) and functions exclusively in leptin secretion as it correlates with plasma leptin levels (Ibrahim, 2010). In the pig it appears that the VAT has higher leptin expression levels (Section 2.4.1.1) and therefore this may not apply to pigs. However, as mentioned previously, LGA infants have higher plasma leptin levels than SGA infants and SAT leptin expression may correlate with plasma leptin levels (unfortunately plasma from these pigs was unavailable for analysis). This may be characterised by different DNA methylation levels between SAT and other depots. It would be interesting to repeat this analysis in SAT to confirm this.

2.4.1.4 Control vs. high-fat diet

There is accumulating evidence to suggest that environmental factors such as nutrition influence epigenetic mechanisms (Mazzio & Soliman, 2012; Milagro *et al*, 2009; Tobi *et al*, 2009). Leptin has been identified as a gene which may be susceptible to epigenetic variation influenced by early life nutrition (Jousse et al, 2011; Tobi et al, 2009). A study by Tobi *et al* revealed that offspring conceived at a time of famine during the Dutch hunger winter had higher DNA methylation

levels at the leptin promoter than their unaffected, same sex siblings (Tobi *et al*, 2009). A low protein prenatal diet in mouse appears to result in decreased leptin promoter methylation in white adipose tissue which correlates with lower plasma leptin levels (Jousse et al, 2011).

Here, both the VAT and SAT of piglets which experienced a high fat prenatal diet from conception to birth were tested for leptin promoter DNA methylation and leptin expression levels, however, no significant differences in either factors were observed.

There are several reasons why Tobi *et al* may have seen an effect of diet whilst here, no diet effect was observed. The focus of the Dutch hunger winter study was humans whereas here pigs were used as a model. Pigs may not exhibit the same effects seen in humans although as leptin is a highly conserved protein between the species this is unlikely.

Here, adipose tissue was sampled whereas Tobi *et al* examined DNA methylation levels in whole blood of individuals. It may be common practice to use the whole blood of patients in clinical studies as they are easily accessible but as leptin has been described as a T-DMR, the tissue type examined may be important when investigated DNA methylation. Leptin's primary site of expression is the adipose tissue which is why analysis of DNA methylation in this tissue was chosen for this study. It may be that the diet did not have an effect on DNA methylation on the adipose tissue but did in the whole blood of pigs. However, Jousse *et al* sampled adipose tissue and did observe a difference in leptin promoter methylation between mice subjected to a low protein prenatal diet and control mice. In this case it may be that whilst a low protein perinatal diet influences leptin promoter DNA methylation, a high fat perinatal diet does not.

The study by Tobi *et al* was published before the recent re-discovery of cytosine hydroxymethylation (5-hmC) in mammalian DNA. Mass spectrometry was used to investigate DNA methylation levels, which would have focused solely on cytosine methylation (5-mC). In this study pyrosequencing was used to investigate DNA methylation levels which accounts for total DNA methylation (5-mC and 5-hmC). Although no significant change in total DNA methylation was observed in this study it may be that the ratios of 5-mC to 5-hmC within the total methylation observed did change. These changes would not be observed using pyrosequencing. However, Jousse *et al.*, who did see an effect of diet, also

used pyrosequencing although again this may be a consequence of the different prenatal diets.

Another factor for consideration is the CpGs which were interrogated. Evidence shows that the location of the CpGs within a gene and the specific base modification of the CpG (5-mC or 5-hmC) are important in determining the function of the cytosine modification (Section 1.2.5). It is therefore important to determine that the same CpGs were interrogated in this study and by Tobi *et al* and Jousse *et al* for a comparison of the two studies to be valid. Alignment of the human, mouse and porcine leptin promoters shows that several conserved CpGs were investigated by both studies and this includes the C/EBPα transcription factor binding site.

The source of the dietary fat may also be a factor for consideration as the specific type of fat an individual is exposed to can also modulate adipocyte gene expression. Iyer *et al* found that diets supplemented by coconut oil or sunflower oil led to a gene specific transcript expression decrease in some genes with the pigs exposed to sunflower oil but not in the pigs exposed to coconut oil (Iyer *et al*, 2012). Here, palm oil was used to increase fat content of diets. It may be interesting to repeat this study with other sources of fat as a dietary supplement to determine whether the fat source is influential.

2.4.2 CpG-5 and CpG-6 are more highly methylated than other CpGs in the region

Although significant differences in DNA methylation between the groups tested were not observed, there was a significant difference in DNA methylation between individual CpGs in each sample. Whilst most CpGs exhibited low levels of DNA methylation, CpG-5 and CpG-6 were consistently more highly methylated than other CpGs in the region. In the VAT, CpG-5 was 1.6 fold more highly methylated than the regional average whilst in the SAT, this site was 3.5 fold more highly methylated than the regional average. CpG-6 was also more highly methylated in both depots although the fold increases were more similar at 3.0 fold in the VAT and 3.2 fold in the SAT. A similar observation has been made in human adipose tissue where the CpG located within the C/EBPα transcription factor binding site was 1.5 fold more highly methylated than other CpGs in the region (Stöger, 2006).

CpG-5 is located within a C/EBPα transcription factor binding site which is a highly conserved sequence of the leptin promoter in many species (Figure 1.1).

The methylation status of this specific CpG has proven to be important for leptin promoter activity. As discussed previously, a study by Melzner *et al* found that cytosine methylation (5-mC) of this CpG down-regulates human leptin promoter activity when driving the expression of a luciferase reporter gene. Cytosine methylation (5-mC) of two further CpGs proximal to a TATA box element abrogated leptin promoter activity completely (Melzner *et al*, 2002).

This information suggests that the DNA methylation observed at CpG-5, located within the C/EBPα transcription factor binding site and CpG-6, located six nt proximal to the TATA box element are involved in regulating leptin gene expression in the VAT and SAT. However, when analysis of leptin mRNA expression and DNA methylation was restricted to CpG-5 and/or CpG-6 (data not shown), no correlation was observed.

The technique used to assess DNA methylation levels in this study, pyrosequencing, does not differentiate between 5-mC and 5-hmC and therefore the specific contribution of each cytosine modification to the observed DNA methylation pattern at the leptin promoter is unknown. Although there were no significant differences in total DNA methylation, there may have been a difference in the ratio of 5-mC to 5-hmC within this DNA methylation pattern. This could also explain significant differences in mRNA expression which do not correlate with statistically significant differences in DNA methylayion.

It has been suggested that the binding of MeCP2 to a methylated leptin promoter contributes to silencing of the leptin gene (Melzner *et al*, 2002) and studies have shown that the presence of just one 5-hmC base in a stretch of DNA can significantly inhibit the binding of MeCP2 (Valinluck *et al*, 2004), possibly leading to activation of the leptin promoter and transcription of the leptin gene. With this possibility and the finding of Melzner *et al* that 5-mC at specific sites of the leptin promoter contributes to down-regulation of the leptin promoter it seems more likely that the observed DNA methylation at CpG-5 and CpG-6 could be attributed to 5-hmC. Several studies have shown that demethylation of the leptin promoter modulates leptin expression during adipogenesis (Melzner *et al*, 2002; Yokomori *et al*, 2002) and one potential role for 5-hmC is the suggestion that it is an intermediate in a demethylation pathway. 5-hmC may then be converted to 5-fC and 5-caC and eventually to cytosine which may facilitate the expression of leptin. Although 5-mC is usually associated with gene silencing, Jousse *et al* found that decreases in leptin expression correlated with decreases in leptin promoter DNA methylation (Jousse et al, 2011). This further suggests the possibility that either 5-mC or 5-hmC is associated with leptin expression rather than leptin silencing.

Also important to consider is the heterogeneity of adipose tissue. Although adipose tissue consists primarily of the leptin expressing adipocytes it also contains other cell types which are considered not to express leptin (Wang *et al*, 2008). The DNA methylation observed at CpG-5 and CpG-6 may therefore correspond to the non-leptin expressing portion of cells and may be attributed to 5-mC. VAT and SAT are described to have different cellular compositions with VAT being more cellular, vascular and innervated and containing a larger number of inflammatory and immune cells (Wronska & Kmiec, 2012) and the differences in DNA methylation at CpG-5 may reflect this and the leptin expression levels of those cells. Equally possible, is that the observed DNA methylation is a mixture of 5-mC and 5-hmC relating to different cell types.

2.4.3 DNA methylation may not regulate gene expression at the porcine leptin promoter

Although higher levels of DNA methylation were observed at CpG-5 and CpG-6 in comparison to other CpGs, overall, the porcine leptin promoter exhibited low to intermediate levels of methylation in both adipose tissue depots analysed. This is similar to modification levels detected in the human leptin promoter, whilst the orthologous mouse region is characterised by overall higher levels of methylation (Stöger, 2006).

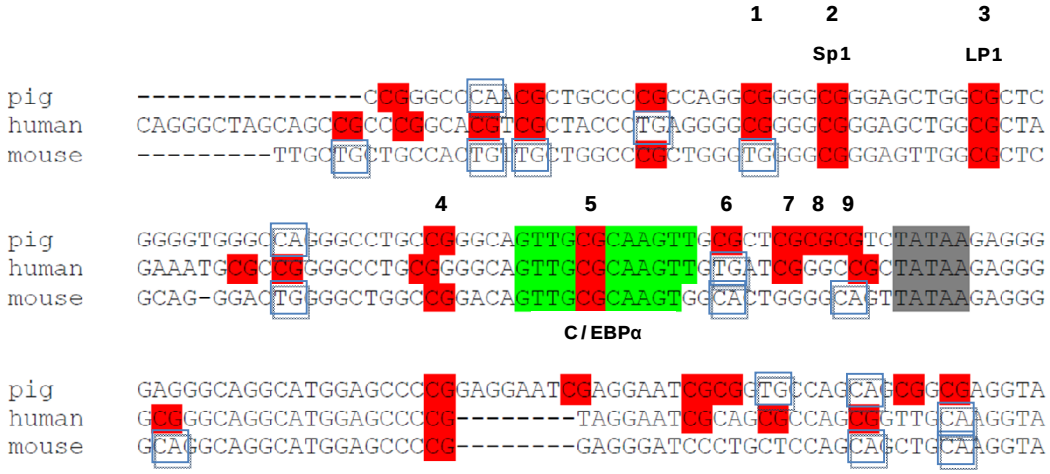

Figure 2.14. The pig, human and mouse core leptin promoters. Alignment of the core leptin promoter sequences from pig human and mouse. CpGs are highlighted in grey, TG and CA nucleotide substitutions are indicated (blue boxes). CpGs 1-9 analysed by pyrosequencing are indicated. The conserved C/EBP-α TF binding site is highlighted in green.

The mouse promoter appears to have become CpG depleted relative to the pig and human promoters (Figure 2.14). This may lead to a higher level of DNA methylation at the mouse leptin promoter because there are fewer CpGs to methylate than in human and pig which may play a role in the regulation of leptin gene expression in this species. In pig and human other epigenetic modifications may be responsible for the regulation of gene expression. Whilst histone modifications and DNA methylation are known to work in concert to regulate gene expression, there is also evidence suggesting histone modifications can work independently of DNA methylation to silence or activate a gene. It has also been suggested that DNA methylation is a secondary event which occurs only once a gene has been silenced by other mechanisms (*Reviewed* by Fuks, 2005).

In addition, polycomb group (PcG) proteins play a major role in the regulation of gene expression. PcG proteins work in complexes to repress genes and in some cases, such as PRC2, their occupancy is associated with unmethylated CpG-rich regions (*Reviewed by* Aloia et al, 2013). It is very unlikely that the low levels of DNA methylation observed are responsible for the regulation of leptin gene expression in this species and other epigenetic modifications, such as histone modifications and PcG protein complexes, are more likely to have an influence.

2.5 Conclusion

The aim of the work presented in this chapter was to investigate DNA methylation levels at the leptin promoter between the visceral and subcutaneous adipose tissues of pigs whilst accounting for gender, size at birth and prenatal diet and also to correlate this data with leptin expression levels. It was found that leptin expression in VAT was higher than in SAT and also that VAT from small birth weight piglets expressed leptin mRNA more highly that that from medium birth weight pigs. However, this did not correlate with differences in DNA methylation. No significant differences in DNA methylation were observed between any groups. Similarly to human, the pig leptin promoter exhibits low to intermediate levels of DNA methylation in adipose tissue and it is more likely that other epigenetic mechanisms such has histone modifications and PcG proteins which can target unmethylated CpG-rich regions regulate gene expression at the human and pig leptin promoter.

Despite this, CpG-5 (located within a C/EBPα transcription factor binding site) and CpG-6 exhibited significantly higher levels of DNA methylation than other CpGs in the region. The action of the DNMTs is generally perceived to be

processive. That is, they would either methylate or not methylate most CpG dinucleotides in a given stretch of DNA. This leads to the expectation that CpG dinucleotides in the same stretch of DNA would have the same methylation status (Section 1.2.3.1). Additionally, DNA methylation is usually associated with gene silencing. Therefore it is interesting that higher levels of DNA methylation were observed in the current study, especially at a transcription factor binding site, in the primary tissue of leptin expression. With this in mind it is of interest to determine the enzymatic mechanisms by which the leptin promoter becomes methylated and the role of individual cytosine modifications at the leptin promoter.

Chapter 3. The influence of the DNA methyltransferases enzymes (DNMTs) on leptin promoter methylation in mouse embryonic stem cells (mESCs)

3.1 Introduction

Cytosines that are adjacent to guanines on a DNA strand (CpGs) are palindromic, producing symmetrical CpG dyads on the opposing strands of a DNA molecule. These cytosines can be modified at the C-5 position to produce 5-methylcytosine (5-mC), a key epigenetic modification in mammals which is essential for normal development (Li *et al*, 1992; Okano *et al*, 1999). It is assumed by many that the symmetrical cytosines of CpG dyads exhibit the same methylation status e.g. they are either both methylated or both unmethylated.

The cytosine methylation (5-mC) pattern of CpG dyads is established and maintained by three catalytically active members of the DNA methyltransferases family (DNMTs). *In vitro* experiments have led to the general agreement that DNMT3a and DNMT3b are responsible for the *de novo* cytosine methylation (5-mC) of unmethylated DNA, whilst DNMT1 maintains symmetrical cytosine methylation (5-mC) patterns by replicating the cytosine methylation (5-mC) pattern of the parent strand to the newly synthesised daughter strand during DNA replication (Gowher & Jeltsch, 2001; Okano *et al*, 1998) (Section 1.2.3).

In section 1.2.3.1, the accumulation of data that questions this theory was considered. *In vitro*, DNMT1 exhibits low activity on unmethylated DNA with its preferred substrate being hemi-methylated DNA (where, within a CpG dyad, one CpG is methylated and one is unmethylated) (Bestor & Ingram, 1983) whereas DNMT3a and DNMT3b perform *de novo* methylation which is independent of the methylation status of the complementary CpG within a CpG dyad (Gowher & Jeltsch, 2001; Okano *et al*, 1998).

In vitro, cytosine methylation (5-mC) by the maintenance function of DNMT1 and by DNMT3b appear to occur in a processive manner, whereas the *de novo* function of DNMT1 and of DNMT3a are distributive (Gowher & Jeltsch, 2001; Gowher & Jeltsch, 2002; Hermann *et al*, 2004; Vilkaitis *et al*, 2005) whilst others have observed a processive methylation activity for DNMT3a (Holz-Schietinger & Reich, 2010).

There are discrepancies in the processivity of the DNMT enzymes between *in vitro* and *in vivo* data. It is assumed that the maintenance function of DNMT1 occurs during DNA replication but, *in vivo*, DNA replication has a high

processivity (0.035sec/nucleotides) (Jackson & Pombo, 1998) whilst, *in vitro*, DNMT1 has a low turnover rate (70-450 sec/methyl group) (Pradhan *et al*, 1999). This coincides with evidence that *in vivo*, DNMT activity is enhanced by other factors such as UHRF1 and DNMT3L which, although unable to directly methylate DNA, work in concert with the DNMTs to establish and maintain cytosine methylation (5-mC) (Bostick *et al*, 2007; Bourc'his *et al*, 2001; Gowher *et al*, 2005; Holz-Schietinger & Reich, 2010; Sharif *et al*, 2007).

The specificity and processivity of the DNMT enzymes appears to be more complicated than is assumed by the scientific community and the discrepancies described between *in vitro* and *in vivo* systems have made it difficult to apply *in vitro* data interpretation in context of data derived from a living cell. To address this issue, several studies have used DNMT mutant mouse embryonic stem cells (mESCs) to investigate the contribution of the individual DNMT enzymes to genome wide and gene specific DNA methylation patterns (Arand *et al*, 2012; Athanasiadou *et al*, 2010; Chen *et al*, 2003; Hu *et al*, 2008; Jackson *et al*, 2004; Li *et al*, 1992; Liang *et al*, 2002; Okano *et al*, 1999; Sharif *et al*, 2007; Tsumura *et al*, 2006).

3.1.1 Aims and Objectives

The primary aim of the work presented in this chapter was to investigate the contributions of DNMT1, DNMT3a and DNMT3b to the DNA methylation pattern at the leptin promoter. DNMT1^{s}, DNMT3a$^{(-/-)}$, DNMT3b$^{(-/-)}$ and DNMT3a$^{(-/-)}$3b$^{(-/-)}$ double knockout (DKO) mESCs, which have been used to investigate the contribution of the DNMT enzymes to genome wide and gene specific DNA methylation, were used as a model *in vivo* system (Li *et al*, 1992; Okano *et al*, 1999). It is generally assumed that CpG dyads have a complementary methylation status e.g. both CpGs which comprise a CpG dyad would either both be methylated or unmethylated (Section 1.2.3). Most techniques used to investigate DNA methylation patterns, such as conventional bisulfite sequencing are limited in that they can only determine the methylation pattern on either the sense strand or antisense strand of an individual DNA molecule, meaning that the methylation status of CpG dyads cannot be determined. Here, hairpin linker technology is used, which permits to monitor the methylation status of complementary CpGs on both strands of an individual DNA molecule (Laird *et al*, 2004).

3.2 Materials and Method

Unless stated otherwise, all materials were sourced from Sigma-Aldrich (Poole, Dorset, UK).

3.2.1 Cell culture

The DNMT3a$^{(-/-)}$, DNMT3b$^{(-/-)}$ and DNMT3a$^{(-/-)}$3b$^{(-/-)}$ (DKO) mESCs (Okano *et al*, 1999) and the DNMT1^{s} knockout mESCs (Li *et al*, 1992) analysed in this study were gifts from Professor Adrian Bird, University of Edinburgh. Gene targeting was used to mutate the DNMT genes in J1 mESCs derived from a male strain 129 embryo. Southern blotting, immunoblotting and enzyme activity assays were performed by the producers of these cell lines to confirm null mutations in the DNMT3a$^{(-/-)}$, DNMT3b$^{(-/-)}$ and DKO cell lines (Okano *et al*, 1999). Targeted mutation of the DNMT1 gene resulted in a hypomorphic mutation, DNMT1^{s} and although a null mutation could not be assured, DNMT1 was nearly undetectable in immunoblot analysis (Li *et al*, 1992).

The J1 wild type cell line could not be sourced. As isogenic cell lines are considered to be genetically similar the CGR8 mouse embryonic stem cell line (WT), which was also derived from a male strain 129 embryo, was obtained from the European Collection of Cell Cultures for use as a control.

All culture reagents were obtained from Invitrogen Ltd. (Paisley, UK) unless otherwise stated. mESCs were maintained in gelatinised flasks at 37°C in the presence of 5% CO2 in growth media consisting of Glasgow Minimal Essential media (GMEM) supplemented with 10% foetal calf serum (FBS) (Biosera, East Sussex, UK), 1mM sodium pyruvate, 1mM non-essential amino acids, 0.2mM 2-mercaptoethanol, 1% penicillin/streptomycin and 1000 units/ml ESGRO LIF (Merck Millipore, Watford, UK) as described in Athanasiadou *et al*., 2010. Tissue culture flasks were obtained from Nunc (Nalge Nunc International, New York, USA) and were coated with 0.1% gelatin (Merck Millipore, Watford, UK) prior to use.

3.2.2 Alkaline Phosphatase staining to confirm pluripotency

The pluripotent state of embryonic stem cells is characterised by a high level of alkaline phosphatase (AP) expression (Pease *et al*, 1990); AP staining was used to confirm that mESCs were in an undifferentiated state. Cells were seeded at low confluency and maintained for five days before fixing with 4% paraformaldehyde in PBS for one minute and staining for AP according to the Millipore AP staining kit protocol (Merck Millipore, Watford, UK). Staining was compared to that in cells grown without LIF, which is required to maintain the

pluripotent state, and colonies were counted to determine that more than 90% remained undifferentiated before downstream applications were performed.

3.2.3 Nucleic acid extraction

Prior to nucleic acid extraction, cells were harvested from flasks by trypsinisation with 0.05% trypsin 0.02% EDTA in Hank's balanced salt solution and washed in PBS to remove any residual growth media.

Genomic DNA (gDNA) and total RNA were extracted from mESCs simultaneously using the Allprep DNA/RNA mini kit (Qiagen Ltd., West Sussex, UK). In brief, cells were disrupted with RTL lysis buffer containing 1% β-mercaptoethanol and homogenised by passing the lysate through a Qiashredder spin column (Qiagen Ltd., West Sussex, UK).

The homogenised lysate was first passed through a DNeasy spin column to allow gDNA to bind the column matrix. An equal volume of 70% ethanol was then added to the eluent before passing the solution through an RNeasy spin column to allow total RNA to bind the column matrix. The DNeasy and RNeasy columns were then washed using buffers supplied with the kit to remove residual material. DNA was eluted in EB buffer and RNA was eluted in RNase-free water (both supplied with the kit).

gDNA and RNA were tested for quality and concentration using the Nanodrop ND1000 UV-Vis spectrophotometer (Nanodrop, Wilmington, USA) and by electrophoresing 2µg nucleic acid on an agarose gel (Section 2.2.3.3.3). RNA integrity was further verified using the Agilent 2100 Bioanalyser (Agilent Technologies UK Ltd., Cheshire, UK). gDNA was stored at -20°C, RNA was stored at -80°C.

3.2.4 Nuclear protein extraction

Prior to nuclear protein extraction, cells were harvested from flasks by trypsinisation with 0.05% Trypsin 0.02% EDTA in Hank's balanced salt solution and washed in ice-cold PBS to remove any residual growth media.

Nuclear protein was extracted from cells using a mini-extract protocol (Schreiber *et al*, 1989). Reagents and samples were kept at 4°C at all times. Cells were suspended in 400µl buffer A (10mM Hepes pH 7.4, 10mM KCL, 0.1mM EDTA, 0.1mM EGTA, 1mM DTT and 0.5mM PMSF), allowed to swell over 15 minutes. 25µl 10% Nonidet NP-40 was added and cells were vortexed vigorously briefly to disrupt the cell membrane. Samples were centrifuged at 3000 rpm for 30 seconds and the nuclear pellet was resuspended in 50µl buffer C (20mM Hepes

pH7.4, 0.4M NaCl, 1mM EDTA, 1mM EGTA, 1mM DTT and 1mM PMSF) by vigorous pipetting. Samples were shaken at 1400 rpm for 15 minutes on the Thermomixer Comfort (Eppendorf, Stevenage, UK) and then centrifuged for 10 minutes at full speed to collect cell debris. Nuclear protein containing supernatant was collected and stored in single use aliquots at -80°C.

Protein concentration was determined using the DC Protein Assay (Biorad, Hertfordshire, UK). 5µl volumes of BSA standards or samples were added to a 96 well plate and 25µl reagent A' followed by 200µl reagent B was added. Colour reactions were allowed to develop for 15 minutes and absorbance was read at 655nm on the Benchmark microplate reader (Biorad, Hertfordshire, UK) at 655nm.

3.2.5 Western Blotting

Although the mutant cells were originally tested for DNMT knockout, western blots were performed to further verify protein inactivation. 20-30µg nuclear protein was mixed with 4x Laemli buffer (Biorad, Hertfordshire, UK) supplemented with 10% ß-mercaptoethanol. After boiling at 95°C for 5 minutes on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) to denature proteins, samples were loaded onto an 8% polyacrylamide gel with a protein marker (Biorad, Hertfordshire, UK) and electrophoresed at 100 volts until sufficient separation was achieved using the Mini Protein II electrophoresis/blotting unit (Biorad Laboratories Ltd., Hertfordshire, UK).

The separated proteins were transferred to an Immuno-blot PVDF membrane (Biorad, Hertfordshire, UK) using the Mini Protean electrophoresis/blotting apparatus (Biorad, Hertfordshire, UK) at 90 volts over 1.5 hours. Success of transfer was determined by staining the gel with coomasie blue dye to detect untransferred protein and staining the membrane with Ponceau S red solution (0.1% Ponceau, 5% acetic acid v/v) to detect transferred protein.

Membranes were blocked in PBS 0.1% v/v Tween 20 containing 5% w/v Marvel skimmed milk powder (Premier International Foods Spalding, Lincolnshire) for 1 hour to prevent non-specific binding of antibodies to the membrane. Blocked membranes were then incubated with primary antibodies (pAB) overnight at 4°C. pABs were diluted in blocking solution (Table 3.1).

Membranes were also probed for histone H3 to validate protein quality. After several washes with PBS 0.1% v/v Tween-20 to remove residual primary antibody, membranes were incubated with secondary antibody (sAB) conjugated

to HRP as for 1 hour (Table 3.1). sAB was also diluted in blocking solution. Membranes were washed further to remove residual secondary antibody and Amersham ECL plus detection reagent was added to the membranes (GE Healthcare, Buckinghamshire, UK). Chemiluminescence was detected using the ChemDoc-It Imager (UVP, California, USA).

Table 3.1 Primary and secondary antibodies used in western blot analysis. Antibodies were sourced from Abcam, Cambridge, UK.

Antibody	Catalogue number	Species of origin	Dilution
Histone H3	Ab1791	Rabbit polyclonal	1:20,000
C/EBPα	Ab40764	Rabbit monoclonal	1:5000
DNMT1	Ab16632	Rabbit polyclonal	1:5000
DNMT3a	Ab23565	Rabbit polyclonal	2:5000
DNMT3b	Ab16049	Rabbit polyclonal	2:5000
Secondary Rabbit IgG HRP	Ab16284	Donkey polyclonal	2:5000

3.2.6 Hairpin Bisulfite Sequencing (Hp-bss)

A hairpin bisulfite sequencing approach was used to analyse DNA methylation patterns at the leptin promoter on both the sense and antisense strand of individual DNA molecules.

DNA sequence information was obtained from the UCSC Genome Browser created by the Genome bioinformatics Group of UC Santa Cruz (http://genome.ucsc.edu/cgi-bin/hgGateway). The position of the mouse leptin CpG island is Chromosome 6: 29009955-29010170 (annotation July 2007 NCBI37/mm9) and the hairpin bisulfite sequencing assay was designed to interrogate nine CpG dyads of the leptin promoter from position 29010127-29009860. Alignment of this region shows a high level of conservation between species (Figure 1.1) and includes an Sp1 transcription factor binding site, a C/EBPα transcription factor binding site and a site designated LP1 by Mason *et al*, mutation of which significantly affects leptin expression (Mason *et al*, 1998) (Figure 3.1).

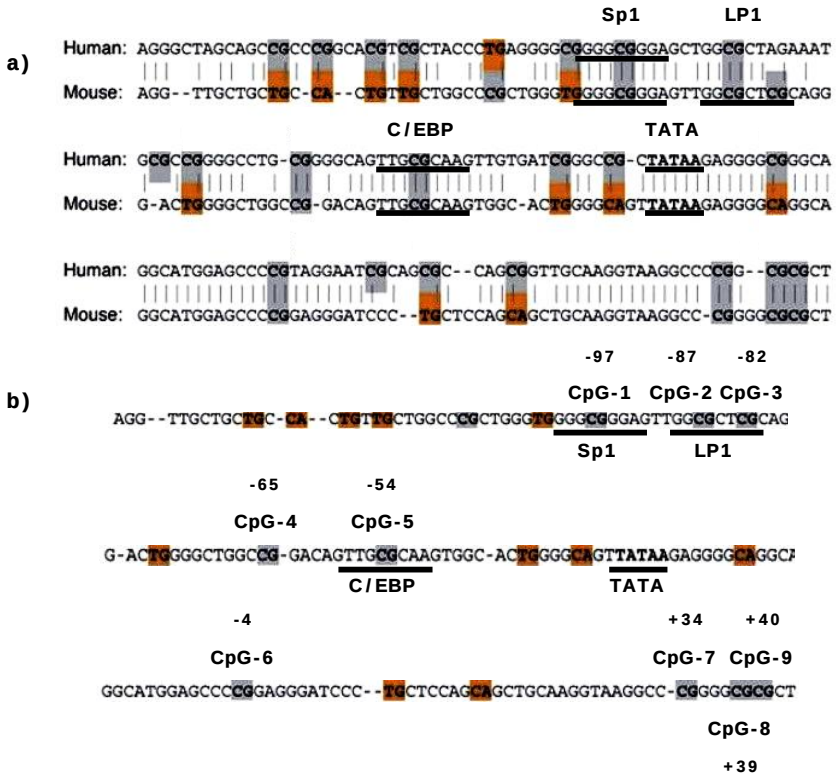

Figure 3.1 The human and mouse leptin promoters. a) Alignment of the core leptin promoter sequence from human and mouse. CpGs are highlighted in grey, TA and CA nucleotide substitutions are highlighted in brown. b) The mouse leptin promoter showing the distribution of CpG dinucleotides. CpGs are highlighted in grey, TA and CA nucleotide substitutions are highlighted in brown and the location of the CpGs in relation to the transcription start site is given. CpGs are numbered 1-9 and relate to the nine CpGs analysed by hp-bss and pyrosequencing. The Sp1, LP1 and C/EBP sites are underlined. *Figure adapted from Stöger, 2006.*

In order to maintain association of complementary DNA strands, individual gDNA molecules were linked together with oligonucleotide hairpin linkers (Figure 3.2) gDNA was isolated (Section 3.2.3) and 5µg was restricted with 16 units FokI restriction endonuclease (New England Biolabs, Hertfordshire, UK) in a 70µl reaction volume at 37°C o/n on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA). The FokI restriction endonuclease cleaved at a recognition site 372bp upstream of the first cytosine of interest on the sense strand and left an overhang of four nucleotides on the sense strand. After buffer change by passing the sample through an Illustra s-200 spin column (27-5120-01, GE Healthcare, Buckinghamshire, UK) the complementary strands in the restriction digest were then joined with 8µl 50µM hairpin oligonucleotide linker (Table 3.2) using 8,000 units high concentration T4 DNA ligase (New England Biolabs, Hertfordshire, UK) at room temperature for 3.5 hrs.

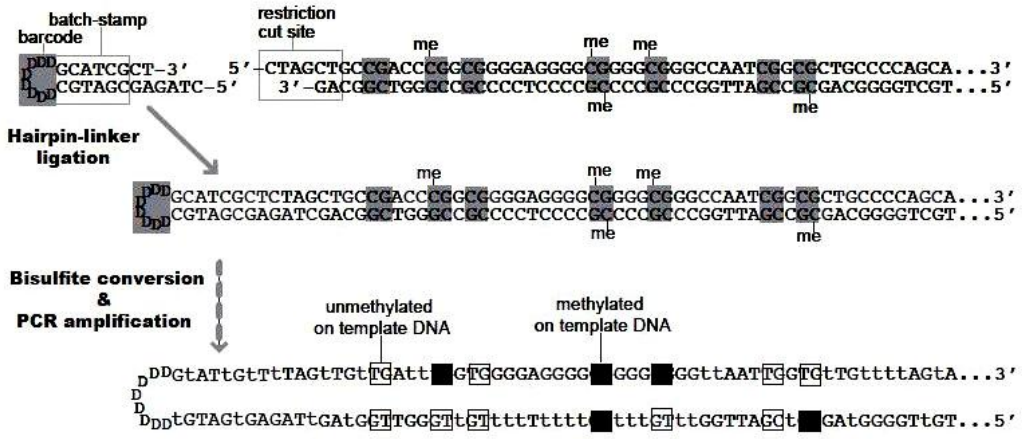

Figure 3.2 An overview of Hp-bss. The sense and bottom strand of individual DNA molecules are linked together via an oligonucleotide hairpin linker containing a batchstamp and barcode which facilitate the removal of contaminant and redundant sequences. DNA then undergoes bisulfite conversion which renders the top and bottom strand non-complementary resulting in denaturation. The top and bottom strand remain attached via the hairpin linker. Hairpin linked, bisulfite converted DNA is then subject to PCR amplification and sub cloning to determine the methylation status of the individual DNA molecules.

Hairpin linkers were 28 nucleotides in length and the first four bases were designed to be complementary to the overhang on the sense strand and were 5' phosphorylated to facilitate ligation to the antisense strand. The next eight bases were designed to be complementary to the last eight bases to allow the hairpin linker to anneal to itself and included a sequence of bases that was unique for each sample (batchstamp) to allow the removal of contaminant sequences from analysis. The eight bases in between the self-annealing sites (barcode) were random nucleotides which allowed the identification and removal of redundant sequences produced in the later PCR amplification from analysis.

Hairpin-linked DNA was cleaned with an equal volume of Microclean (Microzone Ltd., Haywards Heath, UK) and centrifuged at 13,000 rpm for 7 minutes to collect the DNA pellet. The supernatant was removed and the pellet was resuspended in 18µl water.

DNA was then converted using a bisulfite conversion protocol which results in the conversion of cytosine, 5-carboxylcytosine and 5-formylcytosine to uracil but leaves 5-mC and 5-hmC unaffected (Genereux *et al*, 2008). DNA was denatured

by suspension in 20µl 0.3M sodium hydroxide and incubation at 42°C for 2 minutes and 99°C for 2 minutes on a thermal cycler (Eppendorf, Stevenage, UK). 180µl bisulfite solution consisting of a 45% ammonium bisulfite solution (A1145-51002, Spectrum Laboratories, USA) containing 4M sodium bisulfite and 1M ammonium sulphite monohydrate was pre-incubated at 70°C and added to the denatured DNA. DNA was converted by incubation at 99°C for 1 minute and 40 seconds and 70°C for 10 minutes over 5 cycles using a thermal cycler (Eppendorf, Stevenage, UK).

Table 3.2 Oligonucleotide hairpin linkers used in Hp-bss analysis. PHO represents a phosphate group which was attached to the 5' end of the hairpin linker to facilitate ligation.

Cell line	Passage number	Hairpin linker sequence
WT	P6	5'PHO TCTAAGCGATGCDDDDDDDDGCATCGCT 3'
DNMT1s	P15	5'PHO TCTAAGCGTAGCDDDDDDDDGCTACGCT 3'
DNMT3a$^{(-/-)}$	P21	5'PHO TCTAAGGCATGCDDDDDDDDGCATGCCT 3'
DNMT3b$^{(-/-)}$	P21	5' PHO TCTAGACGATGCDDDDDDDDGCATCGTC 3'
DKO	P21	5' PHO TCTATGGCATGCDDDDDDDDGCATGCCA 3'

The bisulfite treated DNA was purified using the Qiaquick PCR purification kit (Qiagen Ltd., West Sussex, UK). DNA was mixed with 1ml DNA binding buffer (PB, supplied with kit) and passed through a PCR purification spin column. After a wash step with buffer PE (supplied with kit) converted DNA was eluted from the column with 45µl elution buffer (supplied with kit).

Desulphonation was performed by addition of 5µl 3M sodium hydroxide and incubation at 37°C for 20 minutes on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA). DNA was purified by passing samples through an Illustra S-200 spin column (27-5120-01, GE Healthcare, Buckinghamshire, UK) and stored at -20°C.

Bisulfite PCR was performed to amplify converted DNA. The sequence for the forward primer was 5' CCAAAACCCTCATCAAGACC 3' and the sequence for the reverse primer was 5' TCATTGGCTATCTGCAGCAC 3'. Primers were obtained high purity salt free (HPSF) purified from Eurofins (Eurofins MGM Operon, Ebersberg, Germany) and designed to amplify a 427bp fragment containing the converted hairpin linked region of interest (Figure 3.1). Amplifications were performed using the HotStar Taq Mastermix kit (Qiagen, Stevenage, UK) (Table 3.3). PCR reactions were performed in a thermal cycler (Eppendorf, Stevenage, UK) using the following program:1 cycle of 95°C for 15 minutes, 41 cycles of 95°C for 30 seconds, 53°C for 30 seconds, 72°C for 1 minute and 1 cycle of 72°C for 5 minutes. PCR products were electrophoresed through a 1.5% agarose gel in TAE buffer containing 1:10,000 RedGel nucleic acid stain (Biotium Inc., California, USA) against a 100bp DNA ladder (Promega, Southhampton, UK) and then visualised on the Doc-It Imager (UVP, California, USA) (Section 2.2.3.3.3). A no-template control was included to ensure amplifications were contamination free.

Table 3.3 Hp-bss PCR reaction mixture

Reagent	Volume per sample
HotStar Taq Mastermix	25µl
Nuclease free water	8µl
Forward primer (50mM)	2µl
Reverse primer (50mM)	2µl
Hairpin-linked converted DNA	13µl
	TOTAL 50µl

PCR products were removed from the gel using the Qiaex II gel extraction kit (Qiagen, Stevenage, UK). After excising the DNA band with a scalpel the gel was dissolved at 50°C on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) in 3 volumes buffer QX1 with 10µl QIAEX II suspension (both

supplied with kit). The QIAEX II suspension, which binds DNA, was then washed using buffers supplied with the kit and the purified PCR products were eluted in nuclease free water.

The purified PCR product was cloned using the PGemT easy vector system (Promega, Southampton, UK). Ligation reactions were assembled and incubated at 4°C overnight to facilitate insertion of the PCR product into the vector (Table 3.4). 50µl DH5α competent cells (New England Biolabs, Hertfordshire, UK) were added to 5µl of the ligation reaction and incubated on ice for 30 minutes to allow the vector to bind the cell wall of the competent cells.

Table 3.4 Ligation reactions. Purified PCR products were cloned using the PGemT easy vector system (Promega, Southampton, UK).

Reagent	Volume per sample
Ligation Buffer	5µl
Nuclease free water	2µl
PGemT vector	0.5µl
Purified PCR product	2 µl (10ng)
T4 DNA ligase	0.5µl
	TOTAL 10µl

The competent cells were then subject to heat shock for 2 minutes at 42°C and returned to ice for 5 minutes to allow the vector to enter the cells. The competent cells were added to 1ml LB and shaken at 150 rpm for 1.5 hours at 37°C. The transformation culture was plated onto LB/Ampicillin/x-gal plates which were then incubated at 37°C overnight. White colonies were picked and transferred to a PCR reaction mixture with primers specific for SP6 (5' ATTTAGGTGACACTATAGAATA 3') and T7 (5' TAATACGACTCACTATAGGG 3') which are sites within the PGemT vector that flank the insertion site (Table 3.5).

PCR reactions were performed on a thermal cycler (Eppendorf, Stevenage, UK) using the following program: 1 cycle of 95°C for 15 minutes, 40 cycles of 95°C

for 30 seconds, 57°C for 30 seconds, 72°C for 1 minute and 1 cycle of 72°C for 5 minutes.

Table 3.5 PCR reaction mixture

Reagent	Volume per sample
HotStar Taq Mastermix	13.5µl
Nuclease free water	10.5µl
T7 primer (5µM)	1µl
SP6 primer (5µM)	1µl
	TOTAL 25µl

PCR products were electrophoresed through a 1.5% agarose gel in TAE buffer containing 1:10,000 RedGel nucleic acid stain (Biotium Inc., California, USA) against a 100bp DNA ladder (Promega, Southhampton, UK) and then visualised on the Doc-It Imager (UVP, California, USA) (Section 2.2.3.3.3). A no-template control was included to ensure amplifications were contamination free. PCR products of 584bp (427bp insert plus 157bp of the PGemT vector) were sequenced (Source Bioscience, Nottingham, UK) with SP6 sequencing primers supplied by Source Bioscience.

3.2.7 Pyrosequencing

Individual pyrosequencing PCR assays were designed to interrogate CpGs on the sense and antisense strands of the mouse leptin promoter using Pyromark Assay Design software 3.0 (Qiagen, West Sussex, UK). The assay was designed to reflect the Hp-bss assay design and amplify the same nine CpGs (Figure 3.1) (Appendix E). Analysis was performed on three consecutive passages of cells (Table 3.6).

Table 3.6 Passage numbers of mESCs used in pyrosequencing analysis

	Passage number		
Cell line	**Replicate 1**	**Replicate 2**	**Replicate 3**
WT	P6	P7	P8
DNMT1^{s}	P15	P16	P17
DNMT3a$^{(-/-)}$	P21	P22	P23
DNMT3b$^{(-/-)}$	P21	P22	P23
DKO	P21	P22	P23

gDNA was bisulfite converted with an incubation step of 70°C for 80 minutes (Section 3.2.6). Converted DNA was then amplified in a PCR reaction (Table 3.7). Amplifications were performed using a thermal cycler (Eppendorf, Stevenage, UK) with the program: 1 cycle of 95°C for 15 minutes, 45 cycles of 95°C for 30 seconds, 58°C for 30 seconds, 72°C for 1 minute and 1 cycle of 72°C for 5 minutes for the sense strand assay and the program: 1 cycle of 95°C for 15 minutes, 45 cycles of 95°C for 30 seconds, 61°C for 30 seconds, 72°C for 1 minute and 1 cycle of 72°C for 5 minutes for the antisense strand assay. Agarose gel electrophoresis was used to visualise PCR products (Section 2.2.3.3.3). A no-template control was included to ensure amplifications were contamination free.

Table 3.7 PCR reaction mixture

Reagent	Volume per sample
HotStar Taq Mastermix	13.5µl
Nuclease free water	4µl
Forward primer (10mM)	1µl
Reverse primer (10mM)	1µl
Bisulfite converted DNA (from 2µg gDNA)	6.5µl
	TOTAL 25µl

10-20µl of PCR product was used in pyrosequencing reactions to determine the percentage of methylated/hydroxymethylated cytosines at each CpG site (Section 2.2.3.3.4).

3.2.8 rt-PCR for leptin, Sp1, C/EBPα, DNMT3L, TET1 and TET2 mRNA expression

cDNA synthesis was performed using 2µg total RNA (Section 2.2.4.2). Rt-PCR amplifications were assembled using the HotStar Taq Mastermix Plus kit (Qiagen, Stevenage, UK) (Table 3.8).

Primers were obtained from Eurofins (Eurofins MGM Operon, Ebersberg, Germany) and were specific to the gene of interest (Table 3.9). PCR reactions were performed in a thermal cycler (Eppendorf, Stevenage, UK) using the following program:1 cycle of 95°C for 15 minutes, 38 cycles of 95°C for 30 seconds, 57°C for 30 seconds, 1 minute and 1 cycle of 72°C for 5 minutes.

PCR products were electrophoresed through a 1.5% agarose gel in TAE buffer containing 1:10,000 RedGel nucleic acid stain (Biotium Inc., California, USA) against a 100bp DNA ladder (Promega, Southhampton, UK) and then visualised on the Doc-It Imager (UVP, California, USA) (Section 2.2.3.3.3). A no-template control was included to ensure amplifications were contamination free and cDNA was also amplified with a housekeeping gene, GAPDH, to ensure cDNA integrity.

Table 3.8 PCR reaction mixture for rt-PCR

Reagent	Volume per sample
HotStar Taq Mastermix	13.5µl
Coraload dye	3.5µl
Nuclease free water	7µl
Forward primer (10mM)	1µl
Reverse primer (10mM)	1µl
Sample cDNA (from RT reaction)	1µl
	TOTAL 25µl

Table 3.9 Primers sequences for rt-PCR

Gene of interest	Primer sequence	Size of PCR product (bp)
Leptin	Forward: 5' CCAAAACCCTCATCAAGACC 3' Reverse: 5' TCATTGGCTATCTGCAGCAC 3'	208
C/EBPα	Forward: 5' TTACAACAGGCCAGGTTTCC 3' Reverse: 5' CTCTGGGATGGATCGATTGT 3'	232
Sp1	Forward: 5' ACAGGAGAGAAGAAATTTGCC 3' Reverse: GCAATACCCTCTGGACAGATG	233
DNMT3L	Forward: 5' CGGCACCAGCTGAAGGCCTTCCATG 3' Reverse: 5' AGGCAGCGCATACTGCAGGATCCGG 3'	342
TET1	Forward: 5' AAGACAGACTTTAACAACAAACC 3' Reverse: 5' CCACTTCTCCACCTCAATTC 3'	168
TET2	Forward: 5' CTGACTCTCAAGTCACAGAAA 3' Reverse: 5' TGTCTGGATTGCATCCTTCAC 3'	193

3.2.9 Restriction enzyme based 5-hmC detection

5-hmC can be glucosylated to form 5-glucosyl hydroxymethylcytosine (5-ghmC). Certain methylation insensitive restriction endonucleases, which cleave DNA when unmethylated, methylated or hydroxymethylated cytosines are present in their recognition site, are blocked by 5-ghmC. A glucosylation step prior to

incubation with the restriction enzyme can therefore be used to differentiate between 5-mC and 5-hmC.

This was utilised in an assay where gDNA was firstly glucosylated in order to convert 5-hmC bases to 5-ghmC, and then cleaved with the restriction enzyme MspI. Glucosylated DNA was also cleaved with HpaII, which is an isoschizomer of MspI that is unable to cleave methylated and hydroxymethylated cytosines. A PCR reaction was then performed with primers that were designed to flank an MspI/HpaII recognition site within the leptin promoter (site four depicted in Figure 3.1). The presence of a PCR product with MspI cleaved glucosylated DNA would indicate the presence of hydroxymethylation at that specific site whilst the presence of a PCR product with HpaII cleaved glucosylated DNA would indicate the presence of methylated and/or hydroxymethylated cytosines at that specific site (Figure 3.3).

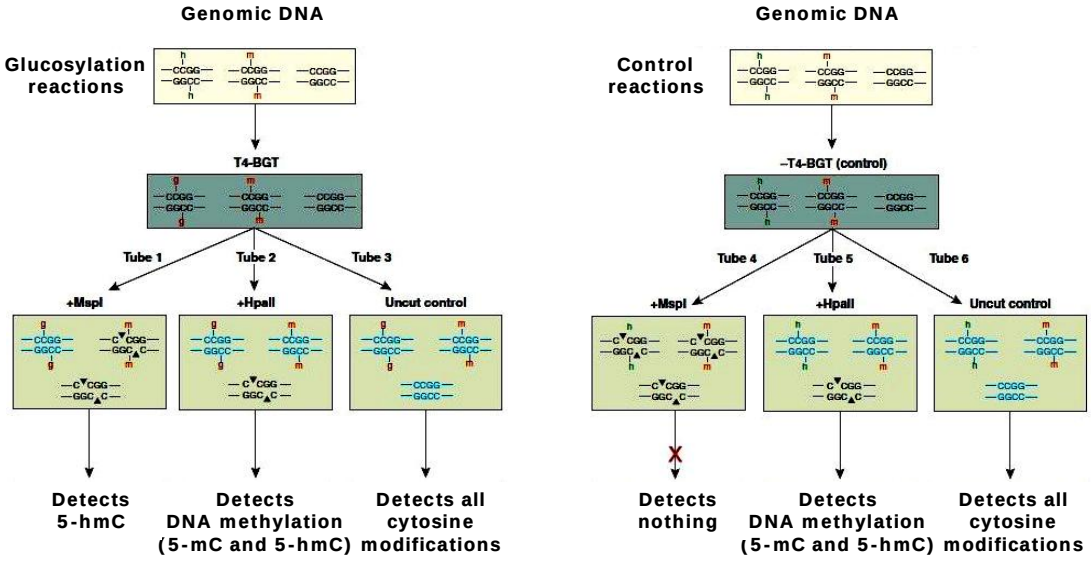

Figure 3.3 Workflow for the restriction enzyme based detection of 5-hydroxymethylcytosine (5-hmC). Glucosylation reactions: gDNA is glucosylated using T4-BGT, converting 5-hmC to 5-ghmC. Glucosylated DNA is digested with MspI or HpaII. MspI cleavage is blocked by 5-ghmC whereas HpaII cleavage is blocked by 5-ghmC, 5-hmC and 5-mC, therefore PCR amplification of glucosylated DNA digested with MspI determines the presence of 5-hmC. Control reactions: gDNA is not glucosylated and therefore should not produce a PCR product after digestion with MspI. *Adapted from the NEB Epimark handbook.*

Reagents were obtained from New England Biolabs, Hertfordshire, UK unless otherwise stated. 200ng gDNA was glucosylated in order to convert any 5-hmC modifications to 5-ghmC by incubation with a ß-glucosyltransferase enzyme overnight at 37°C on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) (Table 3.10).

Table 3.10 Glucosylation reaction mixtures

Reagent	Glucosylation reaction	No glucosylation control
10x reaction buffer	5µl	5µl
UDP-glucose (2mM)	5µl	5µl
ß-glucosyltransferase enzyme (10 units/µl)	2µl	-
gDNA	200ng	200ng
Nuclease-free water	Up to 50µl	Up to 50µl

40 units MspI or HpaII restriction enzyme (or 3µl water to no restriction controls) was added and samples were again incubated overnight at 37°C. Glucosylated, restricted DNA was purified using DNA clean and concentrator columns (Zymo, California, USA). 100µl DNA binding buffer (supplied with kit) was mixed with DNA and passed through the column to allow the binding of DNA to the column matrix. The column was washed several times with wash buffer (supplied with kit) and DNA was eluted with 10µl nuclease free water. The purified DNA was amplified in a PCR reaction using HotStar Taq plus mastermix (Qiagen, West Sussex, UK). PCR reactions were performed using the program: 1 cycle of 95°C for 15 minutes, 30 cycles of 95°C for 30 seconds, 57°C for 30 seconds and 72°C for 1 minute and 1 cycle of 72°C for 5 minutes on a thermal cycler (Eppendorf, Stevenage, UK) (Table 3.11).

Table 3.11 PCR reaction mixture

Reagent	Volume per sample
HotStar Taq Mastermix	13.5µl
Coraload dye	3.5µl
Nuclease free water	3µl
Forward primer (5mM)	1µl
Reverse primer (5mM)	1µl
Template DNA	4µl
	TOTAL 25µl

Primers were obtained from Eurofins (Eurofins MGM Operon, Ebersberg, Germany) and were designed to amplify a 136bp PCR product. The forward primer sequence was 5' CTAGAATGGAGCACTAGGTTGCTG 3' and the reverse primer sequence was 5' CTGCCTGCCCCTCTTATAACTGC 3'. PCR products were separated through a 1.5% agarose gel in TAE buffer containing 1:10,000 RedGel nucleic acid stain (Biotium Inc., California, USA) against a 100bp DNA ladder (Promega, Southhampton, UK) and then visualised on the Doc-It Imager (UVP, California, USA) (Section 2.2.3.3.3). A no-template control was included to ensure amplifications were contamination free.

3.2.10 DNMT3b mutant mouse tissues

DNMT3b mutant adipose and liver mouse tissues were a gift from Claire Francastel at the University of Paris. These tissues were obtained from newborn DNMT3b (mEX3/mEX24) mutant mice which have a missense mutation in the catalytic domain (mutation D823G in exon 24) which results in a partial loss of DNMT3b function and a single base insertion shifting the reading frame and introducing a premature termination signal in the N-terminal region (T insertion in exon 3). This results in hypomorphic mutants with a partial loss of DNMT3b function whilst expression levels of the other DNMTs remains unaffected (Velasco *et al*, 2010).

3.2.11 Statistical analysis

Statistical analysis was performed by conventional ANOVA (Genstat, 2011).

3.3 Results

3.3.1 Validation of mESCs

The primary aim of the work presented in this chapter was to investigate the contributions of DNMT1, DNMT3a and DNMT3b to the DNA methylation pattern at the leptin promoter. Generally these three DNMTs are highly expressed in pluripotent cells with DNMT3a and DNMT3b becoming down-regulated in somatic cells. For this reason pluripotent mouse embryonic stem cells (mESCs) were utilised. Pluripotent cell lines such as mESCs can spontaneously differentiate in culture. The undifferentiated state of each cell line was confirmed prior to performing downstream applications. The undifferentiated state of mESCs can be characterised by high levels of alkaline phosphatase (AP) expression which is then down-regulated in differentiated cells (Pease *et al*, 1990). AP staining is, therefore, routinely used to test for pluripotency. Using the Millipore AP staining kit, cells highly expressing AP produce a deep red colour whilst differentiated cells produce a pink colour or no colour. Staining was performed alongside cells that had lost their pluripotency by depriving them of LIF, a factor which is essential for the maintenance of pluripotency in mESCs. Post-staining, the number of colonies producing a deep red colour was counted and AP staining confirmed that more than 90% of colonies within each mESC stem cell line remained undifferentiated (Figure 3.4).

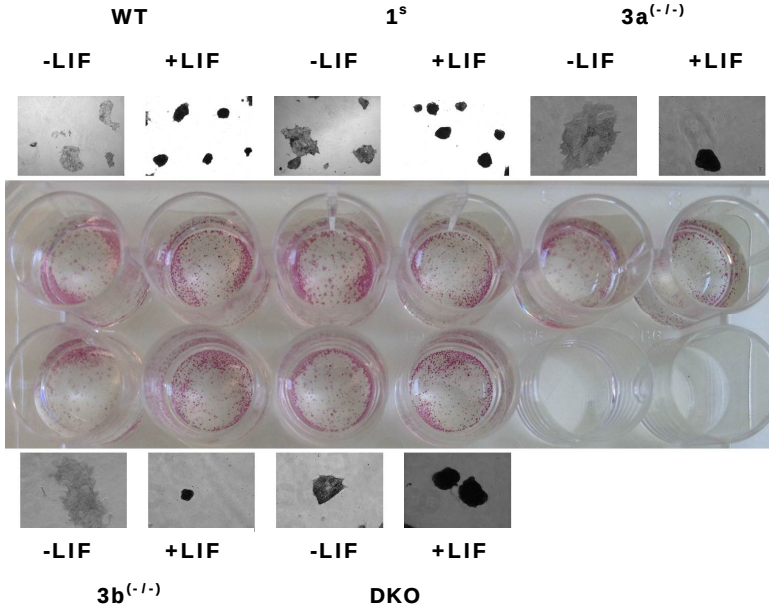

Figure 3.4 Alkaline Phosphatase staining of mESCs. WT, 1^s, $3a^{(-/-)}$, $3b^{(-/-)}$ and DKO mESCs were tested for pluripotency by AP staining. Cells were maintained with (undifferentiated) and without LIF (differentiated). Undifferentiated cells produce a deep red colour whilst differentiated cells produce either no colour or a light pink colour. Colour microscopy equipment was unavailable however black and white images at x40 are shown alongside a colour photograph of the 24 well plate.

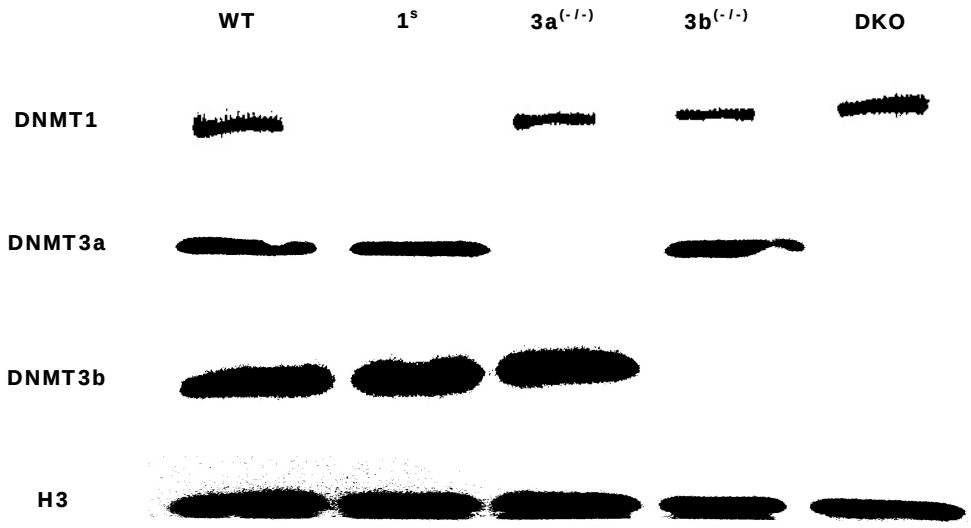

Figure 3.5 Western Blots for DNMT protein expression in mESCs. mESCs were also probed for Histone H3 to verify protein integrity.

To investigate the individual contributions of DNMT1, DNMT3a and DNMT3b to DNA methylation at the leptin promoter, mESCs with knockouts for these enzymes were sourced. In order to confirm that these enzymes had been inactivated, each cell line was subjected to protein analysis by western blotting prior to performing downstream applications. The WT cell line, which expresses DNMT1, DNMT3a and DNMT3b was used as a positive control (Figure 3.5).

DNMT1, DNMT3a and DNMT3b proteins were all detected in the WT mESC line. The DNMT3a$^{(-/-)}$ and DNMT3b$^{(-/-)}$ cell lines are null mutations, and the protein targeted for inactivation was not detected in either cell line. The DNMT1^{s} cell line is a hypomorphic mutation, however DNMT1 protein was not detected by western analysis. Cell lines were probed for histone H3 to validate protein integrity of the samples as this protein should be present in all cells. H3 protein was detected in all cell lines.

3.3.2 DNA methylation (5-mC and 5-hmC) analysis

To assess DNA methylation at the leptin promoter pyrosequencing and Hp-bss were used, each technique having advantages and disadvantages. Pyrosequencing determines the overall percentage of methylated cytosines at each CpG analysed within a given population of DNA strands (Figure 3.6). Although this is a useful technique in determining DNA methylation levels in a total population of DNA molecules, it involves the conventional bisulfite conversion of gDNA, resulting in a loss of complementarity between the sense and antisense strand of DNA molecules. This procedure results in the

denaturation and separation of DNA molecules and therefore the sense and antisense strands must be analysed using two separate assays. Hp-bss also requires the bisulfite conversion of gDNA, however, prior to this DNA molecules are cleaved and the sense and antisense strands are ligated via an oligonucleotide hairpin linker. Upon bisulfite conversion and denaturation of the DNA molecule, the sense and antisense strand remain attached by the hairpin linker. Unlike in pyrosequencing, this allows the analysis of DNA methylation patterns on both the sense and antisense strand of an individual DNA molecule, referred to as an epiallele, and subsequently the analysis of the methylation status of individual CpG dyads. The hairpin linker is also encoded with information that allows the removal of contaminant and redundant epialleles (Miner *et al*, 2004), ensuring a high level of confidence in PCR-derived sequence data. A disadvantage of Hp-bss is the requirement of cloning and sequencing of individual epialleles. The number of analysed epialleles is limited by the cost of sequencing and therefore the data may not represent the whole population of epialleles that may be present within a sample. In this analysis 40 epialleles were obtained from each mESC line. To assess the consistency and congruency of data derived from these techniques, Hp-bss data was compared to that obtained from pyrosequencing.

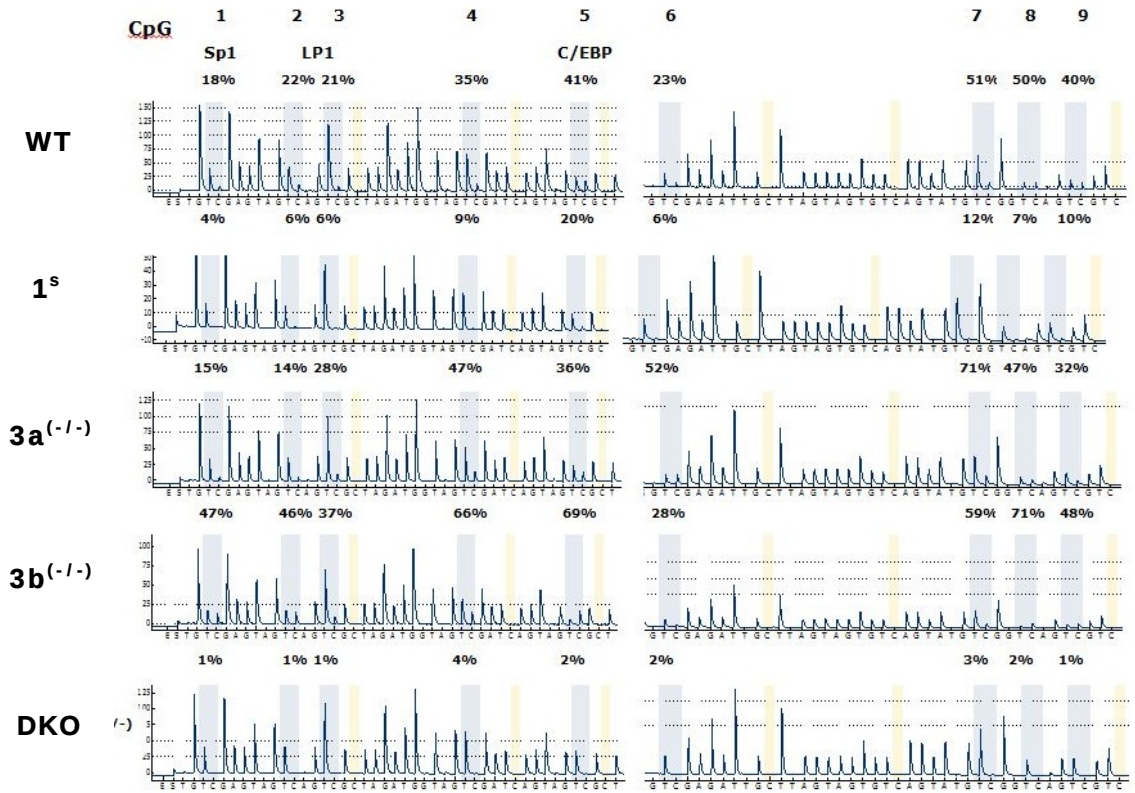

Figure 3.6 DNA methylation levels on the upper strand of the leptin promoter in mESCs using pyrosequencing. Pyrosequencing was used to determine DNA methylation levels at nine CpGs of the leptin promoter in WT, 1^{s}, 3a $^{(-/-)}$, 3b$^{(-/-)}$ and DKO mESCs. CpG one is located within an Sp1 transcription binding site, CpGs two and three are located within a site designated LP1 and CpG five is located within a C/EBPα transcription factor binding site. Internal controls for the detection of an incomplete bisulfite conversion are highlighted in yellow. Numbers given are the percentage of methylated DNA molecules within the sample at each individual CpG site.

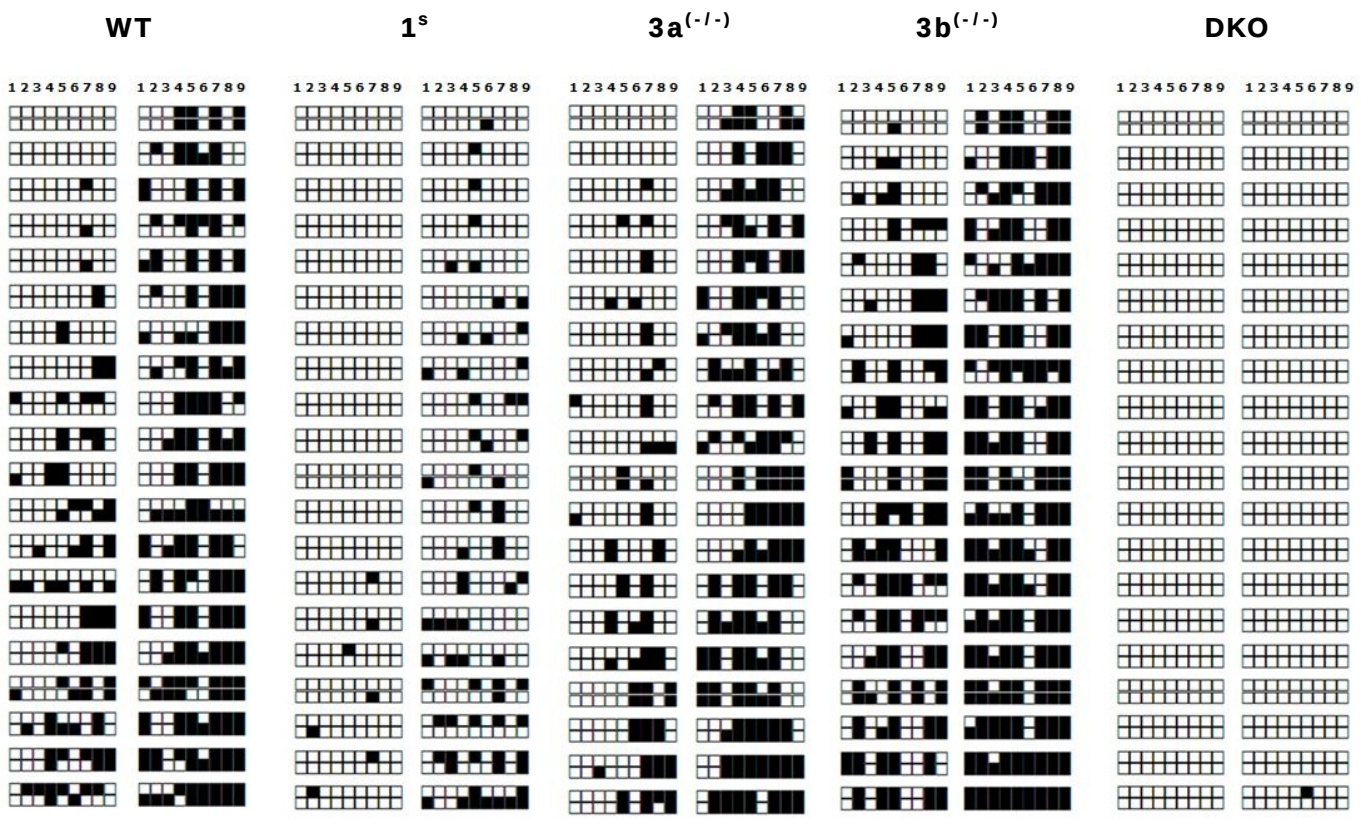

Figure 3.7 DNA methylation at the leptin promoter in mutant mESCs using Hp-bss. 40 epialleles from WT, 1^{s}, $3a^{(-/-)}$, $3b^{(-/-)}$ and DKO mESCs were analysed for DNA methylation at the leptin promoter. Nine complementary CpG sites were interrogated by Hp-bss which allows the analysis of both strands of an individual DNA molecule. The two rows of boxes within each rectangle represent the sense and antisense strand of each epiallele. Unmethylated CpGs are white, methylated CpGs are black. Epialleles within each cell line are sorted by methylation density.

3.3.2.1 Quality monitoring

Several measures were taken to ensure the quality of the data produced. Pyrosequencing analysis was performed on three consecutive passages of cells in order to minimise possible changes in DNA methylation levels between cells of different passages. Standard deviations between the three passages were between 0 and 7% at all sites. Due to the nature of Hp-bss, analysis could only be performed on one passage of cells. However, DNA methylation levels obtained from the two techniques were compared as another quality measure. Individual pyrosequencing assays were designed for the sense and antisense strand of the leptin promoter. It was not possible to rule out that the different assays conferred an amplification bias and for this reason the DNA methylation levels determined by pyrosequencing on the two strands were not compared directly. The data obtained from the two strands was also compared to that obtained from the Hp-bss to assess any bias between the two assays.

3.3.2.1.1 Efficiency of bisulfite conversion

Both pyrosequencing and Hp-bss use a bisulfite conversion step (Section 3.3.6). During this conversion all unmethylated cytosines, including those that are not part of a CpG dinucleotide, are converted to uracil and amplified in downstream PCR reactions as thymine whilst 5-mC and 5-hmC remain unaffected. This allows the differentiation between cytosine and 5-mC or 5-hmC, respectively. An incomplete conversion would result in the inaccurate determination of DNA methylation levels and therefore measures were taken to determine the conversion efficiency. The pyrosequencer was programmed to include a control for an incomplete bisulfite conversion which can be seen in the pyrograms highlighted in yellow (Figure 3.6). At this point in the pyrosequencing step the next base to be sequenced is a thymine which has resulted from a converted cytosine. The machine attempts to incorporate both thymines and cytosines at this step and can therefore determine the ratio of thymines to cytosines and subsequently the conversion efficiency. The required conversion efficiency for a successful run was at the default setting of 95%. There is a report of non-CpG DNA methylation in these mutant mESCs, although not specifically at the leptin promoter (Arand *et al*, 2012). Non-CpG methylation occurs when cytosines that are not part of a CpG dinucleotide are subject to DNA methylation. The presence of non-CpG methylation could skew the efficiency of bisulfite conversion as it is these non-CpG cytosines which are used to test the efficiency. The presence of non-CpG methylation at the leptin promoter in these cell lines was not determined however it did not affect pyrosequencing runs. In the Hp-bss assay,

the commercially synthesised oligonucleotide hairpin linkers were designed to contain unmethylated cytosines which were used to monitor the bisulfite conversion efficiency. This was above 99.5% for all samples.

3.3.2.1.2 Accounting for contamination and redundancy

Another factor which could bias these results is contamination during PCR reactions. All PCR reactions were performed against no-template controls and in the pyrosequencing assay these reactions were also pyrosequenced in order to detect possible contamination. Hairpin linkers for the hp-bss assay were designed to include a batchstamp which was specific to each of the cell lines so that post-sequencing, contaminant sequences could be removed from the analysis. Each individual hairpin linker also contained a randomly generated barcode and any sequence which had the same DNA methylation pattern and the same batchstamp and barcode were assumed to be redundant and removed from the analysis.

3.3.2.2 Overall DNA methylation levels at the leptin promoter in mESCs

DNA methylation levels at nine CpGs of the leptin promoter described in Figure 3.1 were analysed by pyrosequencing and Hp-bss in WT, DNMT1^{s}, DNMT3a$^{(-/-)}$, DNMT3b$^{(-/-)}$ and DKO mESCs. In order to assess the overall effect on DNA methylation the average DNA methylation level over the nine CpGs within each cell line was calculated for both techniques (Figure 3.8). The pyrosequencing data revealed that all cell lines had higher DNA methylation levels on the antisense strand. As explained previously, in pyrosequencing analysis the sense and antisense strands were analysed through two separate assays. It cannot be ruled out that the different PCR reactions between the two assays conferred a bias and therefore no direct comparisons were made between the sense strand and the antisense strand. With the exception of the DNMT3b$^{(-/-)}$ cell line this effect was also observed in the Hp-bss data, where the sense and antisense strands of individual epialleles were analysed simultaneously.

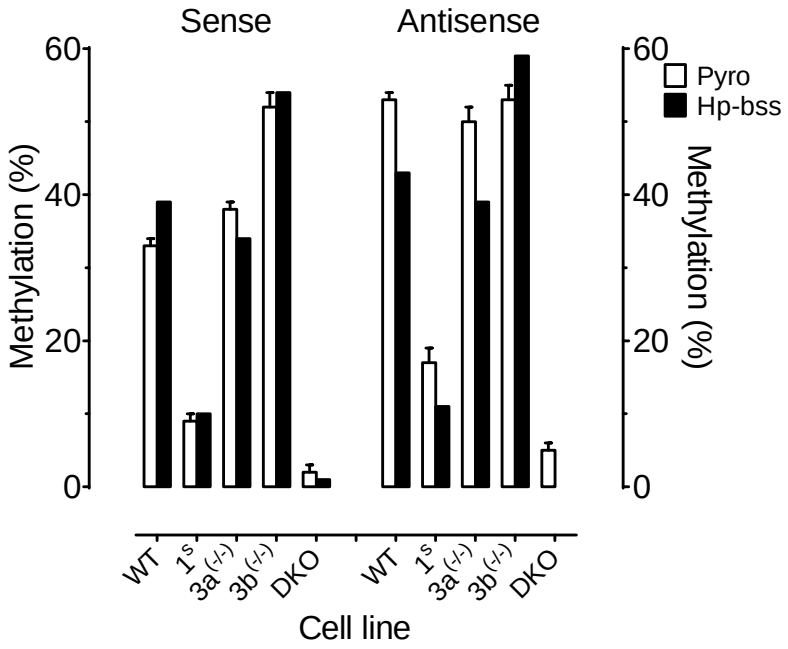

Figure 3.8 Overall DNA methylation levels at the leptin promoter in mESCs. The average DNA methylation level between the nine CpGs in each cell line were calculated by pyrosequencing (white bars) and Hp-bss (black bars).

3.3.2.3.1 The effect of DNMT knockouts on overall DNA methylation levels at the leptin promoter

In both the pyrosequencing and Hp-bss analysis, the WT cell line exhibited intermediate overall DNA methylation levels at the leptin promoter. The effect of DNMT knockouts was analysed relative to the WT cell line (Table 3.12).

The DNMT1^{s} cell line exhibited low levels of DNA methylation on the sense and antisense strand of the leptin promoter (Figure 3.8). Pyrosequencing and Hp-bss analysis gave comparable results between methylation levels on the sense and the antisense strand. Relative to the WT cell line this resulted in a significant decrease in DNA methylation levels on both strands (ANOVA $p=<0.001$) (Table 3.12). The hypomorphic DNMT1^{s} cell line is not guaranteed to be a null mutation and has previously been described as extensively demethylated but not completely demethylated (Lei *et al*, 1996). Further studies which have successfully created a DNMT1$^{(-/-)}$ mouse embryonic stem cell line have confirmed that even with complete knockout of DNMT1, the cell line still contains low and stable levels of DNA methylation (Lei *et al*, 1996). Inactivation of DNMT1, even though the enzyme may not be completely inactivated, caused a substantial reduction in DNA methylation on both the sense

and antisense strand of the leptin promoter. Mentioned previously was the observation that DNA methylation levels determined by pyrosequencing for the antisense strand were higher than the sense strand in every cell line. As separate assays were designed to interrogate the two strands this may have conferred bias which resulted in this observation. In the DNMT1^{s} cell line this effect was also observed in the analysis of Hp-bss data. Although overall DNA methylation levels were higher on the antisense strand, the loss of DNA methylation with the inactivation of DNMT1 was more pronounced on the antisense strand. The dramatic loss of overall DNA methylation levels at the leptin promoter in the absence of DNMT1 is in agreement with the idea that DNMT1 is required in order to maintain the majority of DNA methylation at this genomic locus.

Table 3.12 The effect of DNMT knockouts on overall DNA methylation at the leptin promoter. The percentage of methylation lost or gained was calculated relative to the WT cell line.

Cell line	Sense strand		Antisense strand	
	Pyro	**Hp-bss**	**Pyro**	**Hp-bss**
DNMT1^{s}	-24%	-29%	-36%	-32%
DNMT3a$^{(-/-)}$	+5%	-4%	-3%	-3%
DNMT3b$^{(-/-)}$	+19%	+16%	0%	+16%
DKO	-31%	-39%	-48%	-43%

At the leptin promoter, the inactivation of DNMT3a did not result in a significant change in DNA methylation when compared to the WT cell line. There was a slight discrepancy between the pyrosequencing and Hp-bss results on the sense strand. According to the pyrosequencing data, DNA methylation levels increased by 5% whilst the Hp-bss data displayed a 4% decrease in methylation levels (Table 3.12).

The DNMT3b$^{(-/-)}$ cells exhibited the highest levels of DNA methylation and whilst the inactivation of DNMT3a resulted in no significant change, the inactivation of

DNMT3b produced an significant increase in DNA methylation levels relative to the WT cell line (ANOVA p=<0.001) (Figure 3.8). On the sense strand the increase in DNA methylation was agreeable between the pyrosequencing and Hp-bss results with a 19% and 16% increase in DNA methylation levels respectively, relative to the WT (Table 3.12). On the antisense strand there was a discrepancy in the results obtained from the two techniques. The pyrosequencing data inferred no change whilst the Hp-bss data resulted in a 16% increase in DNA methylation levels relative to the WT. Although the observation of increased DNA methylation levels at the leptin promoter with the inactivation of DNMT3b may be genuine, gene-specific increases in DNA methylation levels in this cell line have not been reported and the limitations of this assay may be responsible for this effect. The original WT for these mutant mESCs could not be obtained and an alternative isogenic cell line, CGR8 (WT), was sourced. It may be possible that the original wild type has higher DNA methylation levels than the CGR8 (WT) cell line which would account for the increased DNA methylation observed in the DNMT3b$^{(-/-)}$ cells. As DNA methylation levels are known to increase with increasing time in culture (Antequera & Bird, 1993), the number of passages each cell line has experienced may also be accountable for this effect. WT cells were used at P6, 7 and 8 whereas DNMT3b$^{(-/-)}$ cells were older at P21, 22 and 23. Although these factors may account for the increased DNA methylation levels relative to the WT, DNMT3b$^{(-/-)}$ mESCs still exhibited higher overall DNA methylation levels by 14-20% than the DNMT3a$^{(-/-)}$ which were at equivalent passages.

The DKO cell line was severely hypomethylated at the leptin promoter in both the pyrosequencing and Hp-bss assays. This cell line has been reported to progressively lose methylation with increasing passages (Armstrong et al, 2012; Jackson et al, 2004). Cells at P21, 22 and 23 were used in these assays and it may be that if earlier passage-cells were used, higher levels of DNA methylation would have been observed at the leptin promoter.

3.3.2.3.2 DNA methylation levels at individual CpGs of the leptin promoter

Next, pyrosequencing was used to determine DNA methylation levels at the nine individual CpGs analysed on both the sense and antisense strand. The results obtained from the pyrosequencer were again compared to that obtained from Hp-bss analysis as a control measure. The percentage DNA methylation at each CpG as determined by pyrosequencing was plotted against that calculated from the Hp-bss data to identify any discrepancies between the two techniques

(Figure 3.9). Overall the two techniques gave a good correlation, however there were some CpGs that exhibited markedly different DNA methylation levels between the two techniques. The CpGs affected were not consistent between the cell lines. This observation may be a consequence of bias in the PCR assays or, more likely, a consequence of analysing only 40 epialleles within the Hp-bss assay. These 40 epialleles may not be representative of the many DNA molecules within each population of cells. For these reasons DNA methylation levels at the individual CpGs was analysed by pyrosequencing only, however, the comparative results from Hp-bss analysis are displayed also (Figure 3.9).

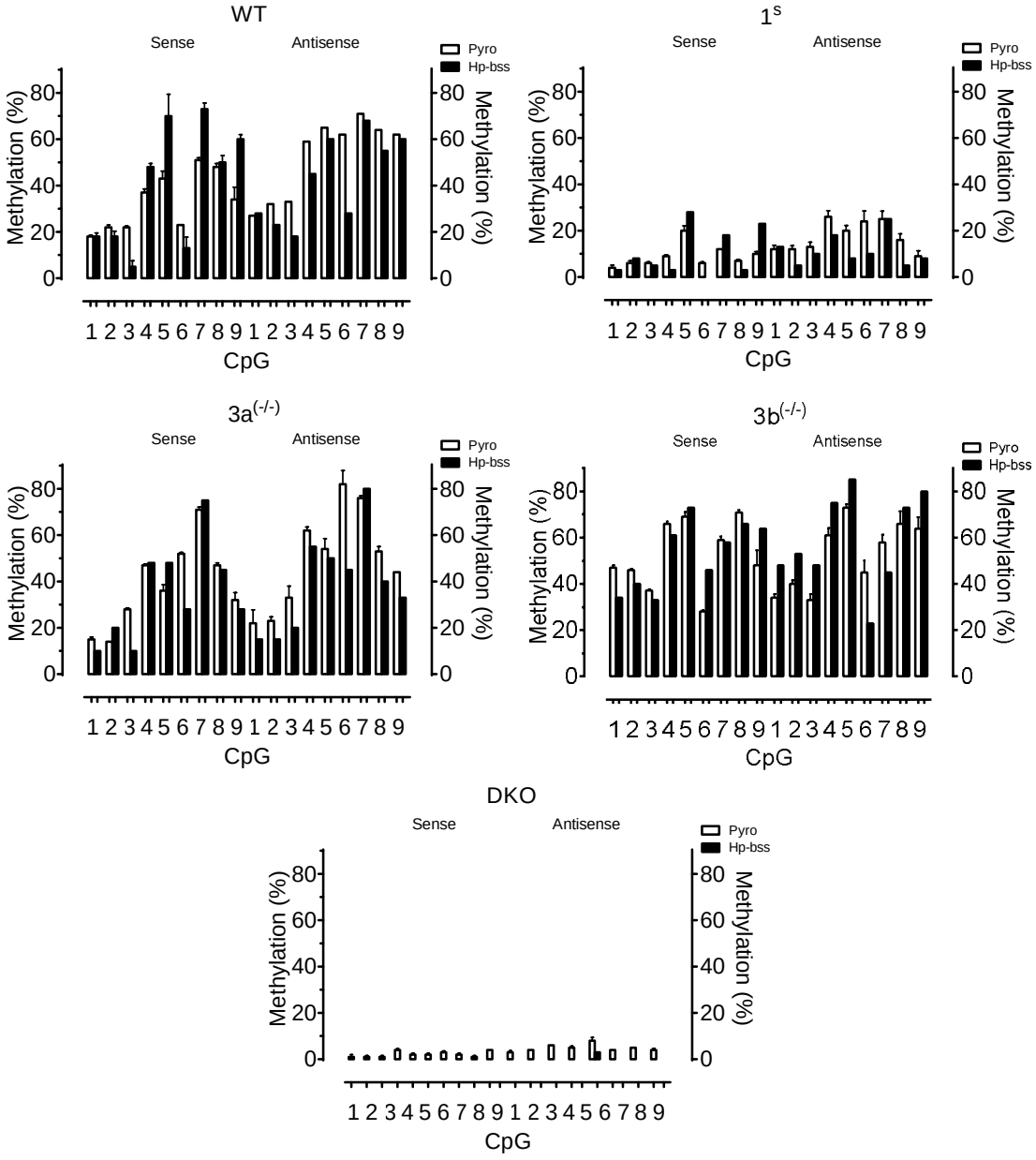

Figure 3.9. DNA methylation levels at individual CpGs of the leptin promoter. The percentage DNA methylation at each CpG in mutant mESCs as determined by pyrosequencing (white bars) and as calculated from Hp-bss (black bars) data.

3.3.2.3.3 The effect of DNMT knockouts on DNA methylation levels at individual CpGs of the leptin promoter

In the WT cell line each of the nine CpGs interrogated exhibited low to intermediate levels of DNA methylation on the sense strand (Figure 3.9). CpG-1, CpG-2, CpG-3 and CpG-6 were the most lowly methylated (18-23%) whilst CpG-4, CpG-5, CpG-7 and CpG-9 were the most highly methylated (34-51%). Higher DNA methylation levels were observed on the antisense strand and although CpG-1, CpG-2 and CpG-3 were still the most lowly methylated CpGs (27-33%), CpG-6 was one of the most highly methylated on this strand. CpG-4 to CpG-9 were the most highly methylated (59-71%).

The inactivation of DNMT1 resulted in a decrease in DNA methylation at each of the nine CpG analysed on both the sense and antisense strand (Table 3.13). The loss in methylation was not uniform between the nine CpGs. On the sense strand, CpG-1, CpG-2, CpG-3, CpG-4, CpG-6 and CpG-8 were now the most lowly methylated (4-9%) whilst CpG-5, CpG-7 and CpG-9 were still the most highly methylated (10-20%). On the antisense strand CpG-1,CpG- 2, CpG-3, CpG-8 and CpG-9 were the most lowly methylated (12-16%) and CpG-4, CpG-5, CpG-6 and CpG-7 were the most highly methylated (20-26%). The double knockout of DNMT3a and DNMT3b led to hypomethylation at each of the nine CpGs on both the sense and antisense strand and at most CpGs in the $DNMT1^{s}$ cell line, the loss in DNA methylation on the sense strand was comparable to that observed in the DKO cell line. CpG-5, however, exhibited 18% higher DNA methylation levels than the DKO cell line, retaining almost half of the DNA methylation observed in the WT cell line. On the antisense strand CpG-5 had only retained under a third of the DNA methylation level observed in the WT cell line.

In the $DNMT3a^{(-/-)}$ cell line, CpG-1, CpG-2 and CpG-3 were again the most lowly methylated (15-28% on the sense strand and 22-33% on the antisense strand). On the sense strand the remaining CpGs, with the exception of CpG-7, gave intermediate levels of DNA methylation (32-52%). CpG-7, however, exhibited higher levels of DNA methylation in comparison to the other sites at 71%. Most sites lost a marginal amount of methylation between 1% and 8%, however, CpG-4, CpG-6 and CpG-7 gave increases of 10%, 29% and 20% respectively in comparison to the WT cell line. On the antisense strand CpG-5, CpG-8 and CpG-9 also gave intermediate levels of DNA methylation (44-54%) but CpG-4, CpG-6 and CpG-7 were more highly methylated (62-82%) and the DNA methylation

levels at these sites had increased relative to the WT. This increase was the most dramatic at CpG-6 (an increase of 19%).

In the DNMT3b$^{(-/-)}$ cell line, on the sense strand, CpG-6 was the most lowly methylated (28%) whilst CpG-1, CpG-2, CpG-3 and CpG-9 gave intermediate levels of methylation (37-48%). CpGs-4, CpG-5, CpG-7 and CpG-8 were the most highly methylated (59-71%). At all nine CpGs there was an increase in DNA methylation levels relative to the WT. DNA methylation levels at most CpGs increased by 14-29% but at CpG-6 and CpG-7 the increase was only 5% and 8% respectively. As discussed previously although the observation of increased DNA methylation levels at the leptin promoter with the inactivation of DNMT3b may be genuine, gene-specific increases in DNA methylation levels in this cell line have not been reported and the limitations of this assay may be responsible for this observation. Interestingly, although DNA methylation levels did also increase on the antisense strand, these increases were marginal (0-8%) in comparison to the sense strand and at CpG-6 and CpG-7, where levels only increased marginally on the sense strand, DNA methylation levels actually decreased by 17% and 13% respectively. On the sense strand CpG-6 was the least methylated whilst on the antisense strand CpG-1, CpG-2, CpG-3 and CpG-6 gave the lowest DNA methylation levels (33-46%). CpG-4, CpG-7, CpG-8 and CpG-9 gave intermediate levels of DNA methylation (58-66%) whilst CpG-5 was the most highly methylated at 73%.

In summary, with the exception of the DKO cell line, the inactivation of the individual DNMTs appeared to have a CpG specific effect. The inactivation of DNMT1 resulted in a loss of DNA methylation at all CpGs, however, this loss was much less substantial on the sense strand at CpG-5 which retained more methylation in comparison to the other CpGs, an effect that was not observed at this CpG on the antisense strand. The inactivation of DNMT3a caused a decrease in methylation at most CpGs, but resulted in increases at CpG-4, CpG-6 and CpG-7. The inactivation of DNMT3b resulted in an increase in DNA methylation at most CpGs but CpG-6 and CpG-7 only increased marginally on the sense strand and on the antisense strand methylation levels actually decreased.

Table 3.13 The effect of DNMT inactivation on DNA methylation at individual CpGs of the leptin promoter. The percentage of methylation lost or gained at each CpG site was calculated relative to the WT cell line.

Sense

CpG	1		2		3		4		5		6		7		8		9	
	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss
1^{s}	-15	-15	-16	-10	-15	0	-27	-45	-7	-43	-17	-13	-39	-55	-41	-48	-24	-38
3a $^{(-/-)}$	-3	-8	-8	+3	-5	+5	+10	0	-7	-23	+29	+15	+20	+3	-1	-5	-2	-33
3b $^{(-/-)}$	+29	+20	+24	+45	+15	+8	+29	+18	+25	+15	+5	+3	+8	-28	+23	+33	+14	+25
DKO	-17	-18	-21	-18	-20	-5	-32	-48	-41	-70	-21	-13	-48	-73	-46	-50	-33	-60

Antisense

CpG	1		2		3		4		5		6		7		8		9	
	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss	Pyro	Hp-bss
1^{s}	-15	-15	-21	-18	-21	-8	-33	-28	-45	-53	-39	-18	-46	-43	-48	-50	-53	-53
3a $^{(-/-)}$	-5	-13	-9	-8	0	+3	+3	+10	-12	-10	+19	+18	+5	+13	-11	-15	-18	-28
3b $^{(-/-)}$	+7	+20	+7	+30	0	+30	+2	+30	+8	-5	-17	-5	-13	-23	+2	+18	+2	+20
DKO	-23	-28	-30	-23	-29	-18	-53	-45	-61	-60	-54	-25	-67	-68	-59	-55	-59	-60

3.3.2.3.4 The effect of DNMT knockout on methylation density in individual epialleles

Pyrosequencing gives valuable information on the methylation of levels of individual CpGs in a given population of cells or DNA molecules but does not allow for the examination of the DNA methylation levels between CpGs in individual DNA molecules, known as epialleles. As DNA methylation is frequently associated with gene silencing and as general ideas of the DNMTs assumes that they are processive, it might be expected that the majority of epialleles would be either hypermethylated or hypomethylated rather than having an intermediate methylation density. Here, Hp-bss was used to investigate the DNA methylation density in 40 epialleles from each cell line, simultaneously accounting for both the sense strand and antisense strand (Figure 3.7).

Each epiallele within the DKO cell line was completely hypomethylated with the exception of one epiallele which was methylated on the antisense strand at CpG-6 (Figure 3.7). In the remaining cell lines, epialleles displayed considerable variation in methylation density. This did not appear to be a consequence of DNMT inactivation as this was also apparent in the WT cell line where, although two epialleles were entirely unmethylated, most epialleles were methylated and showed large variation in methylation densities with the spectrum ranging from 5% (1/18) to 77% (14/18).

Considerably more unmethylated epialleles were present in the DNMT1^{s} mESCs with 13 epialleles being entirely unmethylated. The 27 epialleles that were methylated also showed variation in methylation density, ranging from 5% (1/18) to 50% (9/18), although in comparison to the WT cell line the range of methylation densities was not as wide.

The DNMT3a$^{(-/-)}$ epialleles had the same characteristics as the WT cell line with regards to methylation density. Two epialleles were also entirely unmethylated and the spectrum of methylation densities in the remainder ranged from 5% (1/18) to 77% (14/18).

There were no entirely unmethylated DNMT3b$^{(-/-)}$ epialleles and the spectrum of methylation densities in this cell line was much wider than the others, ranging from 5% (1/18) to 100% (18/18).

3.3.2.3.5 The effect of DNMT knockout on the methylation status of CpG dyads

DNMT1 is generally regarded as the maintenance methyltransferases, faithfully replicating the cytosine methylation (5-mC) pattern of CpG dyads between the parent strand and the newly synthesised daughter strand during DNA replication (Okano *et al*, 1998). This has led to the assumption by many that symmetrical CpGs forming a CpG dyad exhibit a complementary methylation status, e.g. they are either both methylated or both unmethylated, but there has been an accumulation of data that questions this simplistic view (Section 1.2.3.1). The majority of techniques are not able to assess DNA methylation patterns on both the sense and the antisense strand of an individual DNA molecule and therefore many studies have been unable to determine the complementary methylation status of CpG dyads. Here, Hp-bss data was used to investigate the effect of DNMT knockout on DNA methylation patterns at CpG dyads.

Firstly, the percentage of unmethylated, hemi-methylated and methylated CpG dyads within the 40 epialleles obtained for each cell line was calculated (Table 3.14). In the WT cell line CpG dyads were predominantly unmethylated or methylated but, even with the presence of DNMT1, 20% of CpG dyads were hemi-methylated. The inactivation of DNMT1 did lead to an overall increase of unmethylated CpG dyads but the number of hemi-methylated CpG dyads remained similar to that observed in the WT cell line. The inactivation of DNMT3a did not have a drastic effect on the methylation status of CpG dyads within the population but the inactivation of DNMT3b led to an increase in methylated dyads which coincided with a decrease in unmethylated and hemi-methylated dyads. Although the DKO cell line was hypomethylated and the majority of CpG dyads were therefore unmethylated, there was one hemi-methylated CpG dyad within the analysed epialleles.

Table 3.14 The percentage of unmethylated, hemi-methylated and methylated CpG dyads within each cell line

Cell line	%u	%h	%m
WT	49	20	31
1^{s}	82	16	2
$3a^{(-/-)}$	57	12	31
$3b^{(-/-)}$	36	15	49
DKO	99.7	0.3	0

Hemi-methylated epialleles, where methylation is present on one strand but not on the opposing strand, were observed in every cell line even in the presence of DNMT1. The strand which held the methylated CpG within each hemi-methylated epiallele was considered (Table 3.15). The number of hemi-methylated epialleles was indeed highest in the DNMT1^{s} cell line and the location of the methylated CpGs appeared to be equal amongst the sense and antisense strand (Table 3.15). There did not appear to be an obvious strand bias with the hemi-methylated epialleles observed in the other cell lines, although in the DNMT3b$^{(-/-)}$ and DKO mESCs the few hemi-methylated epialleles observed held their methylated CpGs on the antisense strand (Table 3.15).

Table 3.15 The strand location of the methylated CpG in hemi-methylated epialleles

Cell line	Sense strand	Antisense strand
WT	2	3
1^{s}	9	8
3a$^{(-/-)}$	2	2
3b$^{(-/-)}$	0	2
DKO	1	0

Within each cell line, the majority of hemi-methylated epialleles exhibited a low level of methylation density (Figure 3.10). There did not appear to be a obvious bias towards the CpGs that were methylated although within the WT, DNMT1^{s} and DNMT3a$^{(-/-)}$ cell lines several hemi-methylated epialleles were methylated at CpG-7 only.

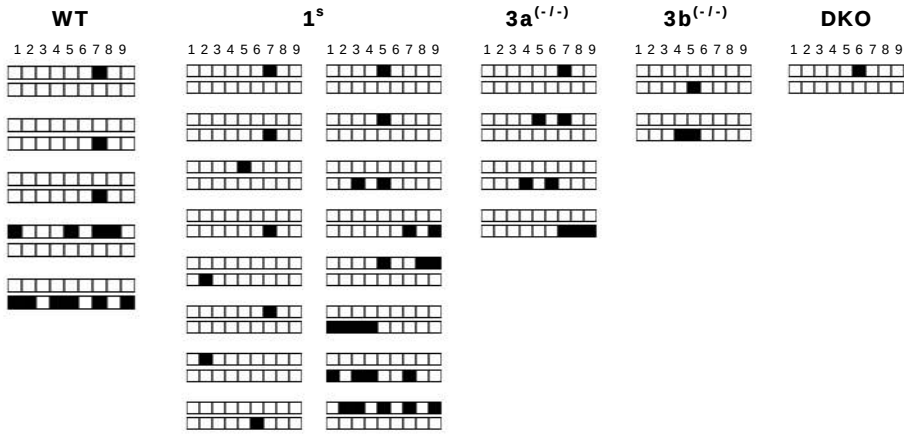

Figure 3.10 The CpG location of the methylated CpG in hemi-methylated epialleles.

The results just described refer to hemi-methylated epialleles where DNA methylation is present on only one strand of an epiallele. There were also epialleles in each cell line which exhibited DNA methylation on both strands but possessed individual CpG dyads that were hemi-methylated.

This was apparent in the WT cell line where 10-27% of epialleles were hemi-methylated at individual CpG sites (Table 3.16). The site exhibiting the highest level of hemi-methylation in this cell line was CpG-5, although there did not appear to be a bias towards which strand possessed the methylated CpG (Table 3.17). At some CpGs there did appear to be a strand bias, for example, at CpG-1 20% of epialleles were hemi-methylated and whilst seven were methylated on the antisense strand, only one epiallele was methylated on the sense strand (Table 3.17).

The inactivation of DNMT1 resulted in decreases in the number of hemi-methylated CpG dyads at each site. This correlated with an increase in the number of unmethylated CpG dyads and is therefore most likely a consequence of the lack of maintenance methylation between CpG dyads and a loss of methylation in the sample. At some sites levels of hemi-methylation were comparable to that observed in the WT but at these sites the WT had a higher proportion of methylated CpG dyads whilst the DNMT1^{s} cells had a higher proportion of unmethylated CpG dyads. This again may be a consequence of the lack of maintenance methylation between CpG dyads and a loss of methylation in the sample. The comparable levels of hemi-methylated CpG dyads at these sites between the WT and DNMT1^{s} cell line may also indicate *de novo* methylation at these CpG sites.

There were also several sites within the DNMT1^{s} cell line which had fully methylated CpG dyads (Table 3.16). This was most pronounced at CpG-7, where four epialleles possessed a methylated CpG dyad. A strand specific bias in hemi-methylation was also observed in this cell line. At CpG-4, whilst in the WT cells a similar level of hemi-methylation was observed between the two strands, four on the sense, three on the antisense, with the inactivation of DNMT1 all six hemi-methylated epialleles possessed the methylated CpG on the antisense strand (Table 3.17). This was also apparent at CpG-9. Whilst in the WT cells the same level of hemi-methylation was observed between the two strands (two on the sense strand, two on the antisense strand) with the inactivation of DNMT1 seven of the methylated CpGs were now located on the sense strand whilst only one was located on the antisense strand (Table 3.17).

With the inactivation of DNMT3a levels of hemi-methylation were slightly lower or remained similar to that observed in the WT at most CpGs and this generally correlated with an increase in unmethylated CpG dyads. At CpG-5, although there was a decrease in hemi-methylated dyads and this correlated with an increase in unmethylated dyads, the number of methylated dyads (43.5%) remained similar to that observed in the WT (50%). At CpG-6 the number of hemi-methylated CpG dyads remained similar to that seen in the WT cell line, however, rather than correlating with an increase in unmethylated dyads the number of these decreased while the proportion of methylated dyads increased from the 7.5% seen in the WT cells to the 25% seen in the DNMT3a$^{(-/-)}$ cells. At CpG-7 whilst the number of hemi-methylated dyads decreased, there was also a decrease in the number of unmethylated dyads but an increase in the number of methylated dyads from 57.5% to 70%.

The inactivation of DNMT3b led to increased DNA methylation levels and this generally correlated with an increased number of methylated CpG dyads at each site analysed. There were exceptions to this however. Most notable was the change at CpG-3. At this CpG site 40% of epialleles were hemi-methylated in the DNMT3b$^{(-/-)}$ mESCs which was over double the amount observed in the WT cell line. Taking into account that the CGR8 (WT) mESCs are not the original wild type, the level of hemi-methylation observed at CpG-3 in the DNMT3b$^{(-/-)}$ mESCs was still double of that observed in the DNMT3a$^{(-/-)}$ mESCs. This correlated with an increase in the number of methylated CpG dyads rather than an increase in unmethylated CpG dyads. Interestingly, when analysing the specific strand on which the methylated CpG was located, one was located on the sense strand whilst 16 were located on the antisense strand. At CpG-7 there was a decrease in the number of methylated dyads which correlated with a decrease in the number of hemi-methylated dyads and an increase in the number of unmethylated dyads. This was not observed at other sites but several sites had dramatic increases in the number of methylated CpG dyads. At most sites there was a 1.4-3.6 fold increase in the number of methylated dyads but at certain CpGs this increase was more pronounced. CpG-2 resulted in a 6.6 fold increase in hemi-methylated dyads and CpG-3 resulted in a four-fold increase.

The DKO cell line exhibited hypomethylation with the exception of one epiallele that was methylated on the antisense strand at CpG-6. This may indicate a *de novo* function for DNMT1 at the leptin promoter whilst it is unable to maintain methylation in the absence of both DNMT3a and DNMT3b.

Table 3.16 The percentage of Unmethylated, hemi-methylated and methylated CpG dyads at each CpG site

CpG	1			2			3			4			5			6			7			8			9		
	u	h	m	u	h	m	u	h	M	u	h	m	u	h	m	u	h	m	u	h	m	u	h	m	u	h	m
WT	67.5	20	13.5	67.5	25	7.5	80	17.5	3.5	45	17.5	37.5	23.5	27.5	50	67.5	25	7.5	17.5	25	57.5	40	15	45	35	10	55
1^{s}	85	15	0	87.5	13.5	0	87.5	10	3.5	83.5	15	3.5	67.5	30	3.5	90	10	0	67.5	23.5	10	93.5	7.5	0	75	20	5
$3a^{(-/-)}$	83.5	10	7.5	80	5	15	75	20	5	43.5	13.5	45	45	13.5	43.5	53.5	23.5	25	15	15	70	53.5	10	37.5	67.5	5	27.5
$3b^{(-/-)}$	47.5	20	33.5	35	15	50	50	40	10	25	10	65	7.5	15	77.5	77.5	7.5	15	53.5	5	43.5	15	15	70	13.5	10	77.5
DKO	100	0	0	100	0	0	100	0	0	100	0	0	100	0	0	97.5	3.5	0	100	0	0	100	0	0	100	0	0

Table 3.17 The location of methylated CpGs within hemi-methylated CpG dyads

CpG	1		2		3		4		5		6		7		8		9	
	Sense	Anti sense	Sense	Anti sense	Sense	Anti sense	Sense	Anti sense	Sense	Anti sense	Sense	Anti sense	Sense	Anti sense	Sense	Anti sense	Sense	Anti sense
WT	1	7	4	6	1	6	4	3	7	4	2	8	7	3	2	4	2	2
1^{s}	1	5	3	2	1	3	0	6	11	1	0	4	3	6	1	2	7	1
$3a^{(-/-)}$	1	3	2	0	2	6	1	4	2	3	1	8	2	4	3	2	0	2
$3b^{(-/-)}$	2	6	5	1	1	16	0	4	3	3	0	3	1	1	5	1	3	1
DKO	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0

3.3.3 DNA methylation in DNMT3b mutant mouse tissues

The DNMT3b$^{(-/-)}$ mESCs exhibited the highest overall levels of DNA methylation between the five cell lines analysed and, with the exception of two CpGs, produced increases in DNA methylation levels at each CpG relative to the WT cell line. It could be argued that these observations might be a consequence of the different wild type cell line used, although they are isogenic. Alternatively, increased methylation could be a result of the different number of passages experienced by the different cell lines (Antequera & Bird, 1993).

To further assess the observation that DNMT3b inactivation leads to hypermethylation of the leptin promoter without the use of a cell line, DNA methylation levels were analysed by pyrosequencing in DNMT3b mutant tissues. The acquirement of DNMT3b$^{(-/-)}$ mouse tissues was not possible because the DNMT3b null allele results in embryonic lethality. Instead adipose and liver tissues from mouse models of human ICF syndrome were obtained (Velasco *et al*, 2010). These mice (mEX3/mEX24) have a missense mutation in the catalytic domain (mutation D823G in exon 24) which results in a partial loss of DNMT3b function and a single base insertion shifting the reading frame and introducing a premature termination signal in the N-terminal region (T insertion in exon 3). This results in hypomorphic mutants with a partial loss of DNMT3b function whilst expression levels of the other DNMTs remains unaffected.

Although perirenal adipose tissue, interscapular brown adipose tissue, flank and perigonadal adipose tissue from these mice were obtained the appropriate controls for these samples could not be sourced and although they were pyrosequenced the data could not be analysed or interpreted accordingly (data not shown). Liver tissues from one wild type newborn mouse and two DNMT3b mutant newborn mice were analysed (Figure 3.11).

On the sense stand great variation was seen between the two mutant embryos at most CpGs, however, CpG-3, CpG-4 and CpG-5 exhibited comparable levels of DNA methylation with differences of no more than 7% between the two samples. At these sites DNA methylation levels were higher in the mutant tissues when compared with the wild type tissues, with increases of between 16% and 39% in both mutant samples (Table 3.18).

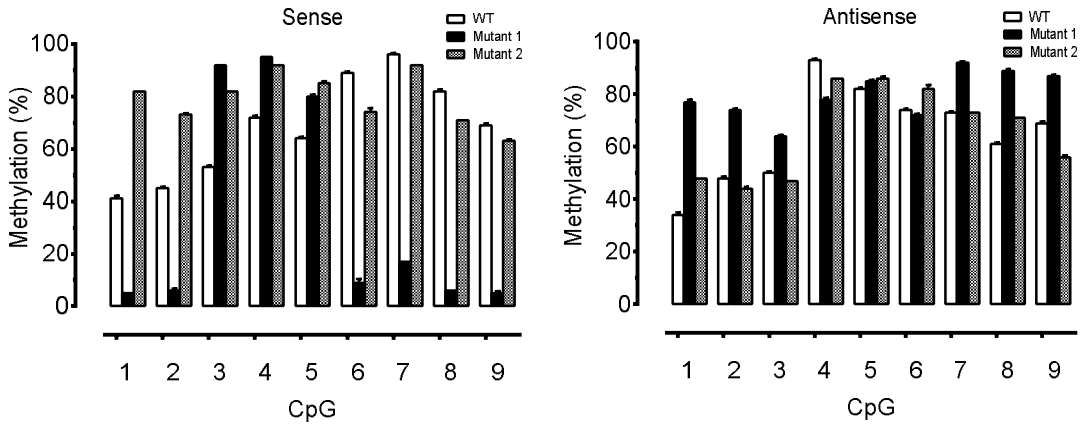

Figure 3.11 DNA methylation levels at the leptin promoter in DNMT3b mutant livers. Pyrosequencing was used to determine DNA methylation levels in DNMT3b mutant liver tissues at nine CpGs on both the sense and antisense strand of the leptin promoter.

Table 3.18 Increases and decreases in DNA methylation levels at the sense strand of the leptin promoter in DNMT3b mutant tissues relative to wild type

Tissue	CpG 1	CpG 2	CpG 3	CpG 4	CpG 5	CpG 6	CpG 7	CpG 8	CpG 9
DNMT3b mutant 1	-36	-39	+39	+23	+16	-80	-79	-76	-64
DNMT3b mutant 2	+41	+28	+29	+20	+21	-15	-4	-11	-6

On the antisense strand these sites gave more modest changes in DNA methylation which varied between the two samples (Table 3.19).

Table 3.19 Increases and decreases in DNA methylation levels at the antisense strand of the leptin promoter in DNMT3b mutant tissues relative to wild type

Tissue	CpG 1	CpG 2	CpG 3	CpG 4	CpG 5	CpG 6	CpG 7	CpG 8	CpG 9
DNMT3b mutant 1	+43	+26	+14	-15	+3	-2	+19	+28	+18
DNMT3b mutant 2	+14	-4	-3	-7	+4	+8	0	+10	-13

Although this is preliminary data and obviously should be assessed in tissues of more animals and different tissue types, this data does bear similarities to the effect observed with the inactivation of DNMT3b in mESCs. In these cells increases in DNA methylation were observed on the sense strand whilst on the

antisense strand more modest changes in DNA methylation occurred. This provides evidence to suggest that the observations made at the leptin promoter with the inactivation of DNMT3b are not likely to be just an effect of using a different wild type cell line or a consequence of using different passages of cells.

3.3.4 Rt-PCRs for Leptin, transcription factors and DNMT3L

DNA methylation at promoters is generally associated with gene silencing. rt-PCRs were performed to assess leptin mRNA expression (Figure 3.12). The WT, DNMT1^{s} and DNMT3a$^{(-/-)}$ expressed leptin mRNA but the DNMT3b$^{(-/-)}$ and DKO cell lines did not. This data was not quantified but the intensity of the bands suggests that the DNMT1^{S} cell line may not express leptin mRNA at the same level as the WT cell line, whilst the intensity of the band produced for the DNMT3a$^{(-/-)}$ cell line is of a similar intensity to that of the WT.

As the inactivation of the DNMT enzymes has genome-wide effects, it may not be the change in DNA methylation levels at the leptin promoter that affects leptin transcript levels. For instance, a change in DNA methylation at a gene encoding transcription factors that regulate leptin gene transcription could be affected. To account for this, rt-PCRs were also performed for mRNAs encoding Sp1 and C/EBPα, two transcription factors which are associated with transcription of the leptin gene. Thus, overall no dramatic change in expression levels of leptin-regulating transcription factors occurred. All cell lines expressed both genes and again, although the data was not quantified, for Sp1 there were no adverse differences in band intensity between the cell lines (Figure 3.12). This was also apparent for C/EBPα, with the exception of the DNMT3b$^{(-/-)}$ cell line, which appeared to produce a band of lower intensity than the other cell lines.

rt-PCRs was also performed for DNMT3l, the non-catalytic member of the DNMT family which has been reported to enhance DNMT function. Inactivation of this enzyme has been reported to have a similar effect of global hypomethylation as observed with the inactivation of DNMT1 (Arand *et al*, 2012). Again, as the inactivation of the DNMTs has a global effect, this was performed to confirm that DNMT3L had not undergone an adverse change in expression levels. All cell lines expressed DNMT3L and, although the data was not quantified, the intensity of the bands was similar between the cell lines. GAPDH housekeeping gene expression analysis also produced similar intensity bands between all cell lines.

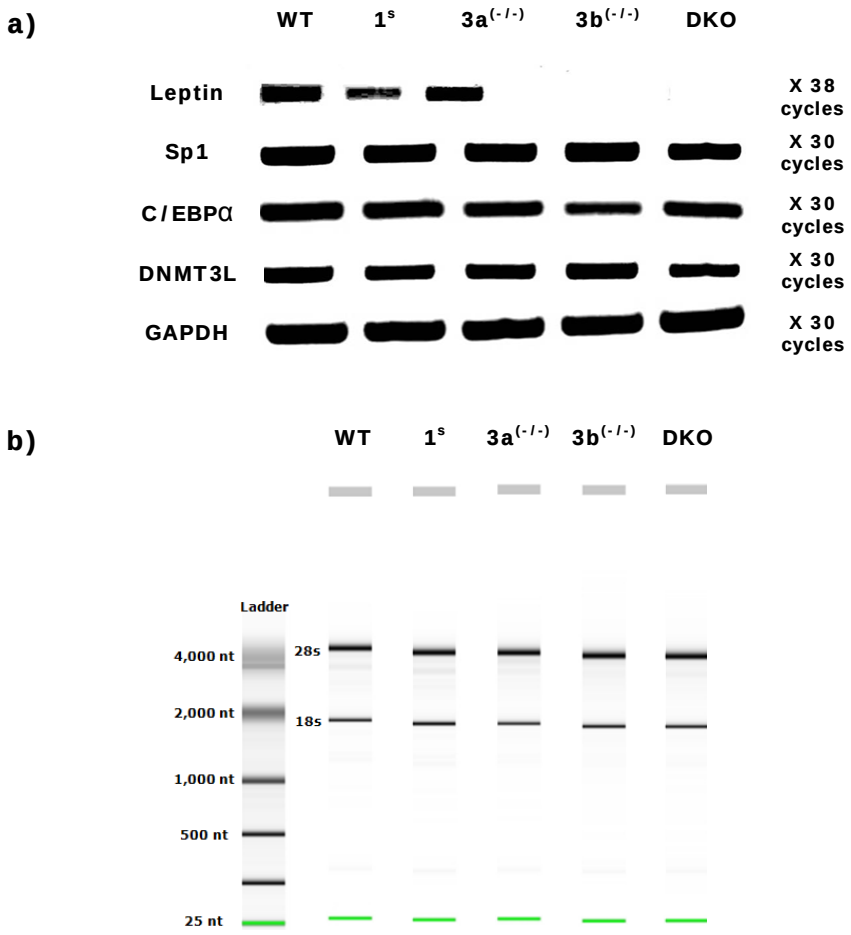

Figure 3.12 mRNA expression analysis of mESCs. a) mESCs were analysed for Leptin, Sp1, C/EBPα and DNMT3L mRNA expression by rt-PCR. b) RNA quality was validated using the Agilent Bioanalyser.

3.3.5 The detection of hydroxymethylation

During the course of the work described in this chapter the re-discovery of cytosine hydroxymethylation (5-hmC) in mammalian DNA (Kriaucionis & Heintz, 2009) and the finding that the TET enzymes can oxidise 5-mC to 5-hmC (Tahiliani *et al*, 2009) were made. Concurrently it was found that 5-hmC is present in mESCs (Xu *et al*, 2011) and for this reason the cell lines were analysed for the presence of 5-hmC at the leptin promoter.

Firstly, rt-PCRs were performed to determine that mRNAs encoding the TET1 and TET2 members of the TET family were present in all cell lines (Figure 3.13). This data was not quantified. However, no obvious differences in band intensity were observed between the samples in the TET1 assay whilst the band intensity for

TET2 appeared to be lower in the DNMT3b$^{(-/-)}$ and DKO cell lines indicating the possibility of a difference in expression levels.

rt-PCRs were also attempted for mRNAs encoding TET3, however, no expression was observed in any of these cell lines and a positive control could not be obtained so it was not clear if the rt-PCR was not fully optimised or if TET3 mRNA was not expressed but it has been reported that TET3 is not expressed in embryonic stem cells (Ito *et al*, 2010).

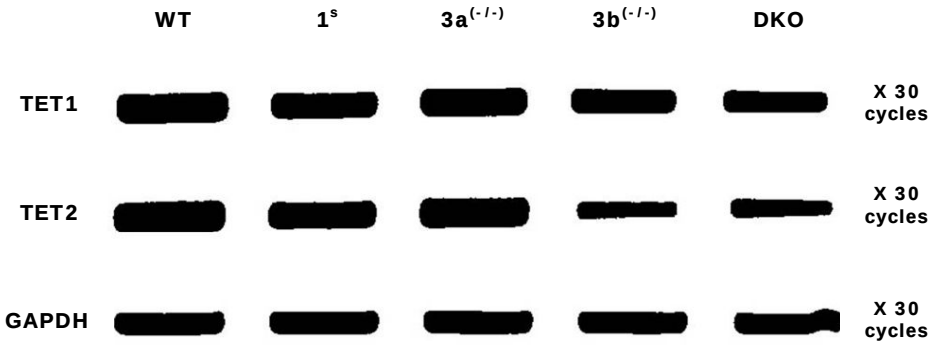

Figure 3.13 rt-PCR for TET1 and TET2 mRNA expression in mESCs

At the time that the work for this chapter was performed, there were few techniques available for gene-specific detection of 5-hmC. The NEB Epimark kit was sourced and used to determine the presence of 5-hmC at CpG four of the leptin promoter (Section 3.3.9). 5-hmC was detected in the WT cell line and the DNMT3a$^{(-/-)}$ mESCs but not in the remaining cell lines (Figure 3.14).

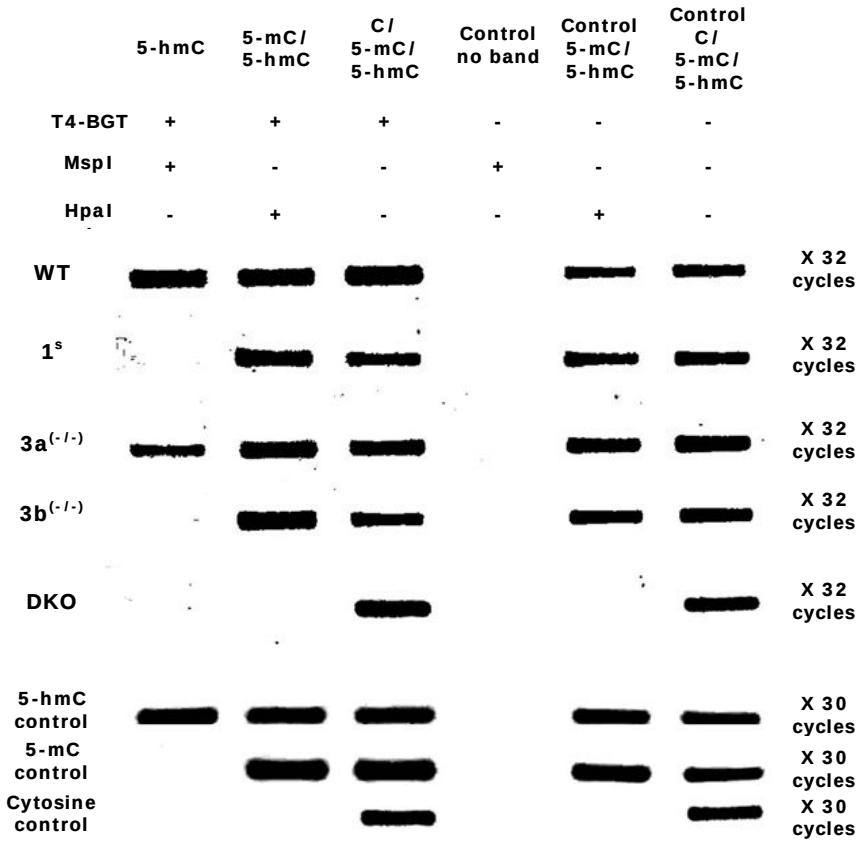

Figure 3.14 5-hmC detection in mESCs. The NEB Epimark assay was used to detect cytosine hydroxymethylation (5-hmC) at CpG four of the leptin promoter.

3.4 Discussion

3.4.1 Limitations of these experiments

Prior to discussing these results the limitations of these experiments must be considered. Firstly, the original wild type cell line that was used to generate the DNMT mutant mESCs could not be obtained and an alternative wild type cell line, CGR8 (WT), was sourced. These cell lines are derived from the same mouse strain and are isogenic. However, there is always the possibility that small genetic differences may have led to any observations when comparing DNA methylation between the wild type and mutant mESCs.

Another limitation is the passage number that the cell lines had experienced. Cell lines have been shown to acquire DNA methylation in culture (Antequera & Bird, 1993) and although the DNMT3a$^{(-/-)}$, DNMT3b$^{(-/-)}$ and DKO cell lines were used at equivalent passages the remaining cell lines were not. The discrepancy in passage numbers may therefore have also led to any differences in methylation observed between these cell lines and the WT and DNMT1^{s} cell lines.

Another factor to consider is the presence of differentiated cells within the sample populations. AP staining was performed on all cell lines to test for pluripotency and cells were only used in downstream reactions if at least 90% of colonies remained undifferentiated. DNMT1^{s} cells have a severely reduced differentiation ability and hypomethylated DKO cells are unable to differentiate (Jackson *et al*, 2004) but any differentiated cells within the other cell lines, even if this was no more than 10%, may have affected downstream analysis.

Results will now be discussed taking into account these limitations.

3.4.2 DNMT1 cannot maintain methylation in the absence of both DNMT3a and DNMT3b

DNMT1 is widely regarded as the maintenance methyltransferase, replicating DNA methylation patterns between CpG dyads semi-conservatively during DNA replication (Gowher & Jeltsch, 2001; Okano *et al*, 1998). DNMT knockout mESCs were used to investigate the role of the individual DNMT enzymes in the observed methylation pattern at the leptin promoter and the work presented in this chapter provides evidence to show that the activity of DNMT1 is not as simple as this.

DNMT3a and DNMT3b are known as the *de novo* DNMTs, establishing methylation patterns during early development which are then maintained by DNMT1 (Gowher & Jeltsch, 2001; Okano *et al*, 1998). Taking this model into account, it would be expected that the simultaneous inactivation of DNMT3a and DNMT3b would not lead to a drastic effect in DNA methylation levels, as DNMT1 would still be present to maintain methylation patterns that had already been established by them. DNMT1 appears in fact unable to maintain DNA methylation at the leptin promoter in the absence of both DNMT3a and DNMT3b as evidenced by the simultaneous knockout of DNMT3a and DNMT3b which led to complete hypomethylation of the leptin promoter, even though DNMT1 protein was still present. This is concurrent with data derived from this DKO cell line in both a gene specific (Arand *et al*, 2012), and genome wide context (Jackson *et al*, 2004). These DKO cells progressively lose DNA methylation with increasing passages (Jackson *et al*, 2004). The inactivation of DNMTs does not specifically target the leptin promoter but has a global effect. It is therefore possible that a change in a factor which enhances DNMT activity has contributed to this effect. DNMT3L is one such cofactor which stimulates DNMT3a and DNMT3b activity and is required for the *de novo* methylation of imprinted patterns (Bourc'his *et al*, 2001; Holz-Schietinger & Reich, 2010). To confirm that DNMT inactivation had

not induced adverse effects on DNMT3L expression levels, which could in turn affect DNA methylation levels, rt-PCRs for mRNAs encoding DNMT3L were performed. DNMT3L was expressed in all cell lines and it is therefore unlikely that down-regulated DNMT3L expression was responsible for this hypomethylation.

The general hypothesis on DNMT1 activity suggests that this enzyme alone maintains DNA methylation. It has been suggested that it is doubtful that a methylation pattern established during development could be maintained by a copying mechanism through DNMT1 alone (Jones & Liang, 2009). The global maintenance methylation efficiency of DNMT1 in the DKO cell line was calculated to be between 97.7% and 98.7% (Jackson *et al*, 2004) whilst a previous calculation at a specific sequence led to a higher maintenance efficiency estimation of 99.9% (Pfeifer *et al*, 1990). A more recent calculation of the maintenance efficiency of DNMT1 in mESCs resulted in the finding that it is dependent on the specific sequence being analysed ranging from 80% to 95% (Arand *et al*, 2012).

Whether the maintenance efficiency of DNMT1 at the leptin promoter is the same as the global level or is specific to the region, it is clear that DNMT1 is unable to maintain DNA methylation in this region in the absence of both DNMT3a and DNMT3b in mESCs.

3.4.3 Hemi-methylation is present at the leptin promoter even in the presence of DNMT1

As DNMT1 is commonly regarded as the maintenance DNMT, stably transferring DNA methylation patterns semi-conservatively, it would be expected that the complementary cytosines of CpG dyads within a stretch of DNA would have the same methylation status and that hemi-methylated CpG dyads should rarely occur. Landmark experiments by Bird have indeed shown that there are few hemi-methylated sites in the genome (Bird, 1978b). More recent analyses, which have allowed for the single base resolution of methylation symmetry at individual CpG dyads, have revealed varying degrees of hemi-methylation between cell types and the element under interrogation. In mESCs, levels of hemi-methylation have been found to be variable between repetitive elements (Arand *et al*, 2012).

Of the 360 CpG dyads analysed, 2% were fully methylated in the DNMT1^{s} cell line. DNMT3a and DNMT3b have been found to exhibit significant maintenance methylation activity as well as *de novo* methylation activity in mESCs and other

cell types (Arand *et al*, 2012). DNMT3a and DNMT3b methylate independently of the methylation status of CpG dyads (Gowher & Jeltsch, 2001; Okano *et al*, 1998). Studies have shown that DNMT3a and DNMT3b can strongly bind to nucleosomes containing methylated DNA (Jeong *et al*, 2009) (Sharma *et al*, 2011) and it has been suggested that this property could allow them to *de novo* methylate hemi-methylated CpG dyads to maintain full methylation status in the absence of DNMT1 (Arand *et al*, 2012). The simultaneous knockout of DNMT3a and DNMT3b results in 30% of CpG sites in mouse repeats being hemi-methylated (Liang *et al*, 2002). This information combined with the observation of decreasing levels of methylation, with increasing passages in DKO cells suggests that ongoing methylation by these "*de novo*" enzymes is required in mESCs to maintain DNA methylation at the leptin promoter.

The DNMT1^{s} cell line is hypomorphic and full DNMT1 inactivation cannot be guaranteed even though DNMT1 protein cannot be detected (Li *et al*, 1992). Therefore, another possibility for the presence of fully methylated CpG dyads in this cell line could be attributed to low levels of DNMT1 activity. DNMT1 is hypothesised to transfer methylation patterns between CpG dyads semi-conservatively during DNA replication. It would therefore be expected that the inactivation of DNMT1 in a somatic state would lead to epialleles that are hemi-methylated, exhibiting DNA methylation on one strand of the DNA molecule but not on the opposing strand with the population of epialleles eventually becoming demethylated. In these mESCs a slightly different outcome might be expected owing to the continual presence of DNMT3a and DNMT3b, which are thought to be responsible for initiating methylation patterns. Under these circumstances, epialleles might be expected to become hemi-methylated and remain hemi-methylated rather than becoming unmethylated. This is because although DNMT1 is not present to maintain methylation, DNMT3a and DNMT3b would still be performing *de novo* methylation. There were in fact 10/40 epialleles with methylation on both strands within the DNMT1^{s} cell line. The presence of methylation on both strands could also be attributed to low levels of DNMT1 activity. If, however, DNMT1 faithfully copies methylation patterns at all CpG dyads, then the presence of DNMT1 cannot account for this observation, because not one of these epialleles had a completely complementary DNA methylation patterns between individual strands. This might be due to selective *de novo* methylation of CpG sites by DNMT3a and DNMT3b. However, although these epialleles did not have a completely complementary methylation status there were some with fully methylated CpG dyads at certain CpG sites. This, as

discussed earlier, could be a consequence of maintenance methylation by DNMT3a and DNMT3b.

A high number of hemi-methylated epialleles (17/40), where DNA methylation was present on one strand but not on the opposing strand, were observed in the DNMT1^{s} cell line and there did not appear to be a bias as to which strand held the methylation. This is concurrent with the idea that DNMT1 was unavailable to transfer methylation patterns semi-conservatively to the opposing strand. There were however also 13/40 unmethylated epialleles within the DNMT1^{s} cell line. As previously discussed, in a somatic tissue DNMT3a and DNMT3b are down-regulated and therefore DNMT1 is the predominant DNMT available to maintain methylation patterns but in mESCs DNMT3a and DNMT3b are still highly expressed. Their presence does not appear to have led to a continual establishment of DNA methylation patterns.

Hemi-methylated epialleles were also observed in all other cell lines, where DNMT1 was expressed. They were not in as high abundance as in the DNMT1^{s} cell line, but they were still present. Although the number of these hemi-methylated epialleles was low in each cell line, there were also a number of hemi-methylated CpG dyads. This is concurrent with observations at the leptin promoter in adipose tissues and peripheral blood leukocytes (Fu *et al*, 2012). Most notable was the effect of DNMT3b inactivation on CpG-3. At CpG-3 not only did this cause the number of hemi-methylated epialleles at this site to increase to double that seen in the DNMT3a$^{(-/-)}$ cell line, but of these 17 epialleles, 16 were methylated on the antisense strand. The site CGCTC has been shown to terminate the processive methylation of hemi-methylated DNA by DNMT1 (Goyal *et al*, 2006). CpG-3 falls just after this sequence. This provides evidence to suggest that *de novo* methylation of this CpG occurs on the antisense strand and that its maintenance might be carried out by DNMT3b rather than DNMT1. With the inactivation of DNMT1 this site was able to retain the same number of fully methylated CpG dyads as observed in the WT cell line. Similarly the inactivation of DNMT1 resulted in 30% of hemi-methylated CpG dyads at CpG-5, located within a C/EBPα transcription factor binding site. Although this was a similar number to that observed in the WT, of these 12 CpG dyads 11 were methylated on the sense strand whereas in the WT cell line there did not appear to be a bias as to which strand was methylated. This observation suggests that DNMT3a or DNMT3b *de novo* methylate this site specifically on the sense strand.

3.4.4 The presence of 5-hmC may account for hemi-methylation

Hemi-methylated CpGs were observed in all cell lines analysed and their presence could be explained by several mechanisms. These CpGs could be selectively *de novo* methylated by the DNMTs or at some point these CpG dyads could have been fully methylated and have become hemi-methylated by selective demethylation. The enzymes which demethylate DNA are not well characterised but the discovery of 5-hmC may have uncovered a pathway for demethylation. This cytosine modification has the ability to recruit or displace certain proteins and may act as an intermediate between 5-mC and cytosine during both active and passive demethylation (Globisch *et al*, 2010; Liutkeviciute *et al*, 2009; Tahiliani *et al*, 2009).

5-hmC could act as an intermediate in an active demethylation pathway which ultimately results in the replacement of 5-mC with cytosine. This had been suggested to involve DNA repair mechanisms (Guo *et al*, 2011) but the emergence that 5-hmC can further be oxidised, by the TET proteins, to 5-fC and 5-caC has uncovered a potential enzymatic pathway for active DNA demethylation. The TET proteins cannot convert 5-caC to cytosine so another enzyme such as a decarboxylase or glycosylase may be involved in this step. The enzyme could be thymidine-DNA glycosylase (TDG) depletion of which leads to the accumulation of 5-caC in ESCs (He *et al*, 2011).

In rapidly dividing cells such as ESCs, 5-hmC may also function as part of a passive demethylation pathway (Williams *et al*, 2012). DNMT1 has low specificity for 5-hmC (Valinluck & Sowers, 2007). While it is copying methylation patterns from the parent strand to the daughter strand during DNA replication, it would treat 5-hmC as if it were an unmethylated cytosine; consequently the cytosine would remain unmethylated on the daughter strand (Tahiliani *et al*, 2009). Whether 5-hmC does block DNMT1 mediated maintenance methylation is yet to be confirmed as UHRF1, a factor which enhances DNMT activity, has been shown to bind 5-hmC enriched DNA possibly moderating the effect of 5-hmC recognition by DNMT1 (Frauer *et al*, 2011).

Due to technical constraints in the gene specific detection of 5-hmC, only CpG-4 was interrogated for 5-hmC content and the limitations of the technique must be taken into account with the interpretation of the data. Firstly the sensitivity of the assay is unknown, for example, how many 5-hmc bases are required to produce a band for hydroxymethylation? There may be a very small number of epialleles that are hydroxymethylated at CpG-4 but the assay may not be

sensitive enough to detect them. Another limitation is the interrogation of only one CpG of the leptin promoter. Although 5-hmC might not be detected at CpG-4, it does not mean that other CpGs in the region are not hydroxymethylated.

5-hmC was not detected in the DKO cell line and this was expected as the cell line was hypomethylated. 5-hmC was detected at CpG-4 in the WT and DNMT3a cells only. In these cells 17.5% and 13.5% of CpG dyads at this site were hemi-methylated. In the WT cell line the remainder were predominantly unmethylated, correlating with demethylation whilst in the DNMT3a$^{(-/-)}$ cell line the remainder were approximately equal between unmethylated and methylated. 5-hmC was not detected in the DNMT1^{s} or DNMT3b$^{(-/-)}$ cell line. With regards to the DNMT1^{s} cell line the inactivation of DNMT1 may have caused a lack of 5-hmC or, as the cell line lost a significant amount of methylation, this may be due to detection sensitivity issues. In the DNMT3b$^{(-/-)}$ cell line 5-hmC was not detected at CpG-4 but 10% of CpGs at this site were still hemi-methylated. This was similar to the WT cell line, however, in the WT cell line this correlated with predominantly unmethylated CpGs dyads whilst in the DNMT3b$^{(-/-)}$ cell line CpG dyads were predominantly fully methylated. It appears that in the absence of DNMT3b CpG-4 has a tendency to become fully methylated. DNMT3b may be important for demethylation of this site or the protection of this site from methylation.

This overall increase in DNA methylation observed in the DNMT3b$^{(-/-)}$ cell line may be attributed to a disruption in this 5-hmC demethylation pathway as caused by the inactivation of DNMT3b. rt-PCRs were also performed for the detection of TET1 and TET2 mRNAs, enzymes which work together to sustain hydroxymethylation in embryonic stem cells (Koh *et al*, 2011). TET1 and TET2 were both expressed in all cell lines. This data was not quantified but it was noted that the DNMT3b$^{(-/-)}$ and DKO cell lines produced bands of lower intensity for TET2 than the other cell lines. This suggests the possibility that the single inactivation of DNMT3b and the simultaneous inactivation of DNMT3a and DNMT3b, which would have a global effect on DNA methylation, have affected expression levels of TET2 in these cell lines. Consequently, this may have resulted in the lack of hydroxymethylation.

It has been reported that sequences in ESCs which are enriched for 5-hmC do not express *de novo* methylation activity of DNMT1 (Arand *et al*, 2012). The absence of 5-hmC may have activated the *de novo* methylation activity of DNMT1 and therefore caused the observed increase in DNA methylation levels. DNMT3a and DNMT3b have been reported to act as dehydroxymethylases *in*

vitro (Chen *et al*, 2012). Thus, these enzymes may have a function in regulating hydroxymethylation and in the context of a demethylation pathway, thereby influence the levels of DNA methylation at the leptin promoter. The lack of 5-hmC may be a consequence of an inability to hydroxymethylate the promoter and allow demethylation. There may be 5-hmC at the other CpG sites that could not be analysed. They may require DNMT3b to dehydroxymethylate and in its absence there has been a build up of methylation at these sites. The functioning of 5-hmC would have resulted in the presence of 5-caC. During bisulfite conversion this cytosine modification acts as an unmethylated cytosine. A technique for investigating 5-caC at single base resolution would help to support the theory that 5-hmC mediated demethylation is responsible for hemi-methylation in these cell lines.

It would also be interesting to determine the presence of 5-hmc at CpG- 5 and CpG-6 as in the DNMT3b$^{(-/-)}$ cells these adjacent sites experienced opposite effects with regards to the methylation status of CpG dyads. Whilst CpG-5 had 15% hemi-methylated CpG dyads and 77.5% fully methylated dyads, CpG-6 had 7.5% hemi-methylated dyads and 77.5% unmethylated dyads. The presence or absence of 5-hmC at these sites would give further clues about the effects of DNMT inactivation of hydroxymethylation and its role in a demethylation pathway.

3.4.5 Inactivation of DNMT3b leads to hypermethylation of the leptin promoter

Unexpectedly, the loss of DNMT3b led to an increase in DNA methylation levels on the sense strand although no change was observed on the antisense strand. Increases in DNA methylation have not been reported before in this DNMT3b$^{(-/-)}$ cell line. As discussed previously, elevated methylation could be a consequence of the different wild type used or the higher number of passages that the cells had gone through. That said, DNA methylation levels were still higher relative to the DNMT3a$^{(-/-)}$ cells which were at the same passage number as the DNMT3b$^{(-/-)}$ mESCs. From this several scenarios must be taken into account when discussing these results. The first scenario is that the original wild type bears the same methylation pattern as observed in the CGR8 wild type (WT) used here. In this case the results would indicate that overall the inactivation of DNMT3b results in increased levels of DNA methylation at the leptin promoter whilst the inactivation of DNMT3a results in a decrease in methylation levels. This suggests that the interaction between DNMT3a and DNMT3b is important for regulating the levels of DNA methylation at the leptin promoter. Looking in a CpG specific context the

inactivation of DNMT3a resulted in increases at CpG-6 and CpG-7. The inactivation of DNMT3b resulted in a modest increase in methylation at these CpGs. These observations further promoter the idea that DNMT3a and DNMT3b are both required to preserve DNA methylation levels at the leptin promoter.

The second scenario would be that the DNA methylation levels in the original wild type do not reflect that seen in the CGR8 wild type (WT) used here. The levels in the original wild type might be higher and in this scenario the inactivation of DNMT3a and DNMT3b has resulted in a decrease in methylation levels, although more profound with the inactivation of DNMT3a. This may infer that DNMT3a is predominantly required for the overall preservation of the DNA methylation pattern although it cannot be determined how this would affect individual CpGs and this again may lead to the conclusion that both DNMT3s are required to preserve the methylation pattern.

The third scenario assumes that again the DNA methylation levels in the original wild type do not reflect that seen in the CGR8 wild type (WT) used here and are actually similar to that seen with the inactivation of DNMT3b. This would infer that DNMT3a is specifically required for the preservation of methylation levels at the leptin promoter whilst DNMT3b may be dispensable.

There is evidence to show that at some genomic loci both DNMT3s are required to preserve methylation patterns whilst at other only one is required (Arand *et al*, 2012). Further work is needed to determine which scenario is true. As a preliminary test to account for the possibility that the observed hypermethylation was in fact a flaw in the experimental design, DNMT3b mutant tissues were sourced and analysed for leptin promoter DNA methylation levels by pyrosequencing. These tissues were obtained from murine models of human ICF syndrome which feature mutations in the DNMT3b gene resulting in hypomorphic mutants. It has been reported that although human ICF patients exhibit an overall decrease in genomic DNA methylation, some loci become hypermethylated (Heyn et al, 2012). These mutants exhibit a partial loss of DNMT3b function which does not affect the expression levels of other DNMTs. This partial loss of function also results in the hypomethylation of centromeric minor repeats known to be a target of DNMT3b (Okano *et al*, 1999), whilst the methylation status of major satellite repeats remains unaffected (Velasco *et al*, 2010). The number of tissues that could be obtained were low (one wild type liver and two DNMT3b mutant livers) but within the liver tissues of these mice, there were three CpG sites that showed the same hypermethylation as with the DNMT3b$^{(-/-)}$ mESCs. Due to the low numbers and lack of experimental repeats,

statistical analysis was not performed. Nevertheless, this provides preliminary data to support evidence of increased DNA methylation levels at the leptin promoter with the inactivation of DNMT3b.

Although leptin has not been linked with ICF syndrome, leptin has been found to be important in immunity (Carbone *et al*, 2012) and ICF is an immune disorder. Although this is a weak link, it would be interesting to determine if ICF patients suffer from hypermethylation of the leptin promoter.

3.4.6 An unmethylated promoter does not correlate with leptin expression

Leptin mRNA was expressed in the WT, DNMT1^{s} and DNMT3a$^{(-/-)}$ cell lines. Before discussing this it is important to point out that a high number of cycles were required to perform this PCR (38) in the mESCs and it is possible that the observed leptin mRNA expression may be due to the small number of differentiated cells within the samples or may be an artefact. RNA-seq data from other mESCs available from the ENCODE project reports no RNA alignments for leptin. Complementary ChIP-seq data suggests that these mESCs are enriched for the histone modifications H3K9me3 and H3K27me3 at the leptin promoter (Appendix H). These modifications are associated with gene repression (Zentner & Henikoff, 2013) and taken together with the RNA-seq data provide evidence to suggest that the expression observed here is an artefact. However, Leptin mRNA is expressed in the blastocyst from which mESCs are derived so it is also a possibility that these pluripotent cells do express leptin (Kawamura *et al*, 2002).

The WT, DNMT1^{s} and DNMT3a$^{(-/-)}$ cell lines expressed leptin mRNA whilst the DNMT3b$^{(-/-)}$ and DKO cell lines did not. Gene expression is usually associated with an unmethylated promoter however the total population of cells from the three leptin expressing cell lines exhibit DNA methylation. Analysis of individual DNA molecules gave epialleles within each cell line that were entirely unmethylated. In the analysis of leptin promoter DNA methylation patterns in adipose tissue it has been suggested that these unmethylated epialleles may be responsible for the observed leptin expression (Stöger, 2006). The DKO cell line however was hypomethylated and did not express leptin mRNA. C/EBPα, the predominant transcription factor associated with the leptin gene possesses a CpG within its binding site (CpG-5). C/EBPα can bind to both an unmethylated and a methylated (5-mC) binding site (Melzner *et al*, 2002). These findings provide evidence to suggest that DNA methylation at the leptin promoter is actually

required for leptin expression but, in contrast the DNMT3b$^{(-/-)}$ cell line was hypermethylated relative to the other cell lines and still did not express leptin mRNA. As the inactivation of DNMTs has a global effect it is possible that it is a change in DNA methylation at another gene that has affected leptin mRNA expression levels. Rt-PCRs were also performed for mRNAs encoding C/EBPα and Sp1, another transcription factor associated with leptin transcription to confirm that with DNMT inactivation they had not suffered any adverse effects. mRNAs encoding both transcription factors were expressed in all cell lines. Although this data was not quantified, the intensity of the bands produced for Sp1 was similar between the cell lines whilst that for C/EBPα appeared to be less intense in the DNMT3b$^{(-/-)}$ cell line. It may be possible that the down-regulation of C/EBPα is responsible for the lack of leptin expression in this cell line.

The specific cytosine modifications present within the observed DNA methylation pattern may also play a role in regulating leptin expression. Indeed the WT and DNMT3a$^{(-/-)}$ cells show evidence of 5-hmC enrichment at CpG-4, whilst the DNMT3b$^{(-/-)}$ and DKO cell line did not. The DNMT1^{s} cell line was not enriched for 5-hmC but the sensitivity of the assay could be responsible for this result, as it would be expected that the dramatic decrease in DNA methylation would also result in a decrease in 5-hmC levels. A frequently mentioned possible role of hydroxymethylation is that it acts as an intermediate in a demethylation pathway. However, the presence of this cytosine modification in non-dividing cells such as within brain tissue, suggests that this may not be its only function (Kriaucionis & Heintz, 2009). Hydroxymethylation has been associated with both gene repression and gene activation (Williams *et al*, 2011). *In vitro*, the presence of just one 5-hmC in a sequence of DNA significantly inhibits the binding of MeCP2, a factor associated with gene silencing (Ruzov et al, 2011; Valinluck et al, 2004). It has been suggested that MeCP2 contributes to the transcriptional silencing of the leptin gene when it exhibits cytosine methylation (5-mC) (Melzner *et al*, 2002). The ability of C/EBPα to bind to its leptin promoter binding site when the internal CpG is hydroxymethylated is unknown.

3.5 Conclusion

The data presented in this chapter shows that DNMT1, DNMT3a and DNMT3b are all required for the preservation of overall DNA methylation levels at the leptin promoter in mESCs. Whilst DNMT1 is regarded as the maintenance methyltransferase, in the absence of DNMT3a and DNMT3b, DNMT1 is unable to maintain DNA methylation at the leptin promoter and hemi-methylation was

observed in all cell lines. With the individual inactivation of DNMT3a and DNMT3b certain CpG sites exhibited strand specific hemi-methylation. The acquirement of double-stranded DNA methylation data allows for the analysis of DNA methyltransferase activities by a hidden markov model, a statistical analysis which can estimate CpG-specific rates of methylation events and DNMT processivity (Fu *et al*, 2012). Work is ongoing to analyse this data for probabilities of *de novo* methylation and maintenance activity at the leptin promoter via this model, however, this analysis requires a number for epialleles that are hemi-methylated for each individual CpG site. From this it can be determined whether the hemi-methylation is due to a *de novo* methylation event or a loss of methylation (Fu *et al*, 2012). Unfortunately it is unsure whether enough of these epialleles have been obtained for each CpG site.

Whilst DNA methylation is frequently associated with gene silencing, in these cell lines examples of both an unmethylated and a methylated leptin promoter which did not correlate with leptin mRNA expression were found here. The DNA methylation levels at individual CpGs of the leptin promoter and the specific cytosine modifications present within the observed DNA methylation pattern may also influence on leptin expression status. However, mESCs are not a suitable model for exploring the role of specific cytosine modifications in leptin gene expression as it is a possibility that they do not express leptin mRNA. Further work is required to elucidate the individual roles of 5-methyl cytosine (5-mC) and 5-hydroxymethyl cytosine (5-hmC) at the leptin promoter with regards to leptin expression.

Chapter 4. Investigating the role of hydroxymethylation in leptin expression

4.1. Introduction

Leptin is an important hormone with an infamous role in energy metabolism and an emerging role in other fundamental life processes such as reproduction, inflammation and angiogenesis (Fantuzzi & Faggioni, 2000; Hausman *et al*, 2012) (Section 1.1.6). Many publications have reported divergent leptin levels between individuals, one example being the higher plasma leptin levels observed in infants born large for gestational age in comparison to infants born small for gestational age (Marchini *et al*, 1998). The leptin promoter is a tissue specific differentially methylated region (T-DMR) (Stöger, 2006) and cytosine methylation (5-mC) of specific CpGs in this region correlates with the down-regulation of leptin expression (Melzner *et al*, 2002). Despite this information, data regarding divergent leptin expression levels between individuals is rarely correlated with DNA methylation data.

In addition to this the leptin promoter exhibits intermediate levels of DNA methylation in adipose tissue, the primary site of leptin expression (Stöger, 2006). This is an intriguing observation as promoter DNA methylation is frequently associated with gene silencing. Although the primary cells are leptin expressing adipocytes, adipose tissue is highly heterogeneous (Ibrahim, 2010). Therefore, the observed DNA methylation levels could, in part be attributed to the non-leptin expressing portion of cells. Described in the previous chapter was the analysis for mRNAs encoding leptin in various mESC lines and hypermethylation of the leptin promoter did indeed correlate with leptin gene silencing, however also presented were mESCs with intermediate levels of DNA methylation that did express leptin. In these leptin expressing mESCs, cytosine hydroxymethylation (5-hmC) was also detected within this DNA methylation pattern.

Published research which investigates DNA methylation at the leptin promoter does not distinguish between 5-methyl cytosine (5-mC) and 5-hydroxymethyl cytosine (5-hmC) and at present there have been no reported attempts to determine if 5-hmC is present at the leptin promoter. The exact roles of 5-hmC in gene expression are yet to be elucidated, but since 2010 there has been a rapid accumulation of publications providing information on possible roles of 5-hmC. A frequently cited role of 5-hmC is as an intermediate in a DNA

demethylation pathway. 5-hmC is formed from the oxidation of pre-existing 5-mC which can then be further oxidised to form 5-fC and 5-caC and eventually leading to conversion back to cytosine (Ito et al, 2011). The 5-hmC detected in leptin expressing mESCs may pertain to a demethylation pathway which allows conversion of 5-mC to cytosine thereby regulating levels of 5-mC and facilitating leptin expression levels. However, an mESC line that was hypomethylated at the leptin promoter also did not express leptin mRNA and therefore, an unmethylated leptin promoter does not necessarily correlate with leptin expression.

5-hmC has also been proposed to play a role as both a transcription activator and a transcriptional repressor. In mESCs it has been reported that 5-hmC positivity does not correlate with transcriptional activation and most 5-hmC positive genes are not expressed in mESCs (Williams *et al*, 2011). However, 5-hmC is predominantly associated with euchromatin suggesting that it is associated with gene activity rather than repression (Ficz *et al*, 2011; Szulwach *et al*, 2011).

Furthermore, it appears that the specific cell type, the location of the observed DNA methylation within this cell type and the specific cytosine modification, 5-mC or 5-hmC, are all important factors in determining the transcriptional effect of DNA methylation. In neurons, gene expression correlates with the enrichment of 5-hmC over the entire transcriptional unit of the gene whilst it is depleted from the transcription start site (TSS) and intergenic regions. The relationship between gene body 5-hmC enrichment and gene expression levels appears to be dependent on neuronal cell type (Mellén *et al*, 2012). In mESCs the presence of 5-mC in the promoter region is associated with low levels of gene transcription whilst promoters enriched for 5-hmC or both 5-mC and 5-hmC are more highly transcribed (Ficz *et al*, 2011). 5-hmC in exons also correlates with increased levels of transcription and genes specifically enriched for this cytosine modification are actually more highly transcribed than those that were unmethylated (Ficz *et al*, 2011). These observations in neurons and mESCs contradict *in vitro* data where 5-hmC in the promoter region strongly represses gene transcription whilst 5-hmC in the gene body has minimal effects on transcription (Robertson *et al*, 2011).

4.1.1. Aims and Objectives

The primary aim of the work presented in this chapter was to investigate DNA methylation levels at the leptin promoter in the leptin expressing adipocytes whilst accounting for the specific composition of 5-mC and 5-hmC within the observed DNA methylation pattern. Adipocytes are formed from the differentiation of preadipocytes during adipogenesis, a process which is accompanied by a transition in leptin expression from a "switched off" to a "switched on" state (MacDougald *et al*, 1995). Using the 3T3-L1 mouse preadipocyte cell line as a model of adipogenesis, pyrosequencing is used to determine DNA methylation levels at the leptin promoter in 3T3-L1 preadipocytes and adipocytes to determine if 5-mC and/or 5-hmC are associated with leptin repression or leptin expression in these cell types.

4.2 Materials and methods

Unless otherwise specified, all materials used for these experiments were sourced from Sigma-Aldrich (Poole, Dorset, UK).

4.2.1. Cell culture

The 3T3-L1 mouse preadipocyte cell line was obtained from and maintained as recommended by the European Collection of Cell Cultures. All culture reagents were obtained from Invitrogen Ltd. (Paisley, UK) unless otherwise stated. Approximately $4.7x10^5$ 3T3-L1 preadipocytes per well were seeded in 12 well plates (Nalge Nunc International, New York, USA) and cultured in growth media consisting of Dulbecco's Modified Eagles media (DMEM) supplemented with 10% foetal calf serum (FBS) (Biosera, East Sussex, UK), 1mM sodium pyruvate, 1mM non-essential amino acids and 1% penicillin/streptomycin and maintained at 37°C in the presence of 5% CO2.

Preadipocytes were differentiated into adipocytes by replacing growth media with differentiation induction media consisting of growth media supplemented with 1:100 0.0115g/ml isobutylmethylxanthine (IBMX), 1µM dexamethasone (DEX) and 167nM bovine insulin to 2 day post confluent cells. After 48 hours, differentiation induction media was replaced with growth media supplemented with 167nM bovine insulin. Media was refreshed with growth media every 48 hours until day eight when cells were fully differentiated and harvested by trypsinisation (Section 3.2.3). To confirm cells had differentiated, an oil red o staining was performed. Oil red o binds to the lipid droplets which accumulate within differentiated adipocytes (Lillie & Ashburn, 1943). In brief, culture medium was removed from wells and cells were rinsed and fixed with 60%

isopropanol for one minute with gentle rotation. Cells were then incubated for 15 minutes with gentle rotation in an oil red o solution consisting of two parts deionised H_20 and three parts 0.5% oil red o in isopropanol. Cells were counterstained with haematoxylin in order to identify individual cells by their nuclei. Cells stained red were then counted and after confirming the full differentiation of cells (<90%) at eight days the process was repeated for a time course experiment and cells were harvested at six hourly intervals between 20 and 96 hours post induction of differentiation and at eight days (192 hours) full differentiation. Experiments were repeated on three passages (P) of cells; P12, P13 and P14.

4.2.2 Nucleic acid extraction

Prior to nucleic acid extraction, cells were harvested from flasks by trypsinisation with 0.05% Trypsin 0.02% EDTA in Hank's balanced salt solution and washed in PBS to remove any residual growth media. Cells were then separated into two aliquots, one for total RNA and one for gDNA extraction.

gDNA was extracted using the DNeasy Blood and Tissue kit (Qiagen Ltd., West Sussex, UK) (Section 2.2.3.1)

Total RNA was extracted using Trireagent (Ambion, California, USA) (Section 2.2.4.1)

4.2.3 Rt-PCR

Rt-PCRs were performed for leptin, Sp1, C/EBPα, TET1 and TET2 mRNA expression (Section 3.2.8). All PCRs were performed using the same thermal profile described in previous sections and against GAPDH as a housekeeping gene.

4.2.4 Pyrosequencing

DNA methylation analysis of the leptin promoter was performed using pyrosequencing technology (Section 3.2.7). Nine CpGs of the core leptin promoter were analysed including an Sp1 binding site, a C/EBPα binding site and LP1 (Figure 4.1).

4.2.5 5-hmC analysis

5-hmC was detected by a restriction enzyme based technique (Section 3.2.9) and a magnetic bead capture technique. Detection of 5-hmc at single base resolution was also attempted however, due to time restraints, the results could not be fully completed or verified with repeats.

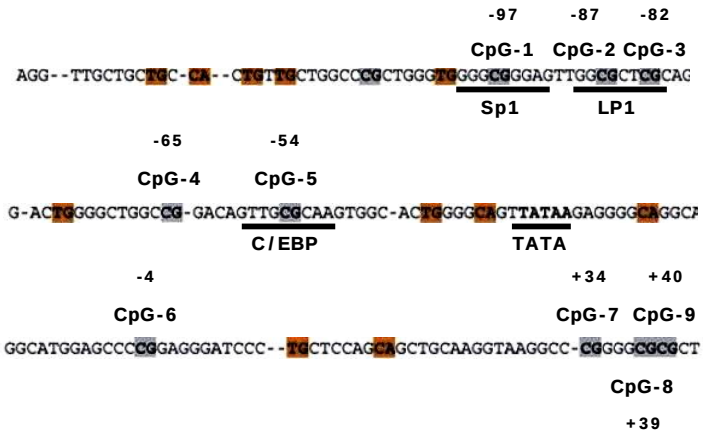

Figure 4.1 The mouse leptin promoter. The mouse leptin promoter showing the distribution of CpG dinucleotides. CpGs are highlighted in grey, TA and CA nucleotide substitutions are highlighted in brown and the location of the CpGs in relation to the transcription start site is given. CpGs are numbered 1-9 and relate to the nine CpGs analysed by pyrosequencing. The Sp1, LP1 and C/EBP sites are underlined. *Figure adapted from Stöger, 2006.*

4.2.5.1 5-hmC detection by magnetic bead capture

The 5-hmC detection assay described in Section 3.2.9 was limited by the interrogation of only CpG-4 (Figure 3.1). The presence of 5-hmC over the leptin promoter region, rather than at a specific CpG, was investigated using the Hydroxymethyl Collector kit (Active Motif, California, USA). Using this kit gDNA is first fragmented and then glucosylated in order to convert any 5-hmC bases to 5-ghmC. A biotin conjugate is then chemically attached to the modified glucose and magnetic streptavidin beads are used to capture the biotinylated 5-ghmC fragments. Enriched DNA can then be amplified in a PCR reaction with primers specific to the region of interest in order to determine the presence of 5-hmC.

gDNA was sheared to an average fragment size of 600-100bp by restriction with Micrococcal Nuclease (MNase) (New England Biolabs, Hertfordshire, UK) (Table 4.1). Fragmentation reactions were incubated at 37°C for 15 minutes on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) (Table 4.1).

Table 4.1 DNA fragmentation reaction mixture

Reagent	Volume per sample
gDNA	4µg
10x Reaction buffer	5µl
100x BSA	1µl
Micrococcal Nuclease (2 units/µl)	2µl
Nuclease free water	Up to 50µl

MNase was inactivated by adding EDTA pH 8.0 to a final concentration of 50mM. Fragmented DNA was purified using DNA Clean and Concentrator columns (Zymo, California, USA) and to ensure fragments ranged from 100-600bp DNA was visualised on an agarose gel (Section 2.2.3.3.3).

200ng fragmented DNA was glucosylated by incubation with a ß-glucosyltransferase enzyme at 37°C for 1 hour on the PCH-1 dry block heater (Grant Bio Instruments, New Jersey, USA) (Table 4.2).

Table 4.2 DNA glucosylation reactions

Reagent	Glucosylation reaction	No glucosylation control
Fragmented DNA	200ng	200ng
10mM DTT	5µl	5µl
3mM UDP-azide-glucose	2.5µl	-
ß-glucosyltransferase (10 units/µl)	2µl	2µl
Nuclease free water	Up to 50µl	Up to 50µl

20µl biotin conjugation solution was added and samples were incubated at 37°C for a further hour to allow the biotin conjugate to bind 5-ghmC modifications. Glucosylated, biotinylated DNA was purified using DNA purification columns supplied with the kit. 370µl DNA purification binding buffer and 3M sodium hydroxide was added to DNA until the mixture turned yellow and was transferred to a DNA purification column (supplied with kit). DNA was passed through the column to allow binding of DNA to the column matrix. After washing with DNA purification wash buffer (supplied with kit) DNA was eluted with 50µl DNA purification elution buffer.

25ul magnetic streptavidin beads and 25µl binding buffer AM13 (both supplied with the kit) were added to the purified biotinylation reaction and samples were incubated for 30 minutes at room temperature with end to end rotation on the Stuart Scientific tube rotator (Stuart, Staffordshire, UK) to allow the streptavidin beads to bind biotinylated 5-ghmC. Using a magnet (supplied with kit), beads were washed 5 times with binding buffer AM13 and enriched DNA was eluted from the beads with 100µl 1x elution buffer AM2, by pipetting several times and incubation for 30 minutes at room temperature with end to end rotation on the Stewart Scientific tube rotator (Stewart, Staffordshire, UK). Enriched DNA was purified using the DNA clean and concentrator (Zymo, California, USA) and amplified in a PCR reaction alongside input fragmented DNA to detect 5-hmC (Section 3.2.9).

4.2.5.2 Detection of 5-hmC at single base resolution

Potassium perruthenate ($KRuO_4$) is an oxidising agent with the ability to convert 5-hmC to 5-fC which behaves as an unmethylated cytosine during bisulfite conversion (Booth *et al*, 2012). Thus, by combining pyrosequencing with a $KRuO_4$ pre-oxidation, the levels of 5-hmC can be determined when compared to a non-oxidised control. 3µg gDNA was suspended in 24µl 0.05M NaOH on ice. 1µl 15mM $KRuO_4$ (Alpha Aeser, Karlsruhe, Germany) in 0.5M NaOH was added and samples were incubated on ice for one hour with occasional vortexing. Samples were purified by passage through a minispin oligo column (Roche, Welwyn, U.K.) and then subjected to DNA methylation analysis by pyrosequencing (Section 4.2.4).

4.2.6 Statistical Analysis

Statistical analysis was performed by conventional ANOVA (Genstat, 2011).

4.3 Results

4.3.1 Differentiation of 3T3-L1 preadipocytes into adipocytes

3T3-L1 preadipocytes were differentiated into adipocytes over eight days (Section 4.2.1). Differentiated cells displayed a distinct change in morphology which was characterised by the accumulation of lipid droplets within the adipocytes (Figure 4.1). Oil red o staining, which binds to lipid droplets, was used to confirm that a high proportion of cells (>90%) had differentiated into adipocytes (Figure 4.2).

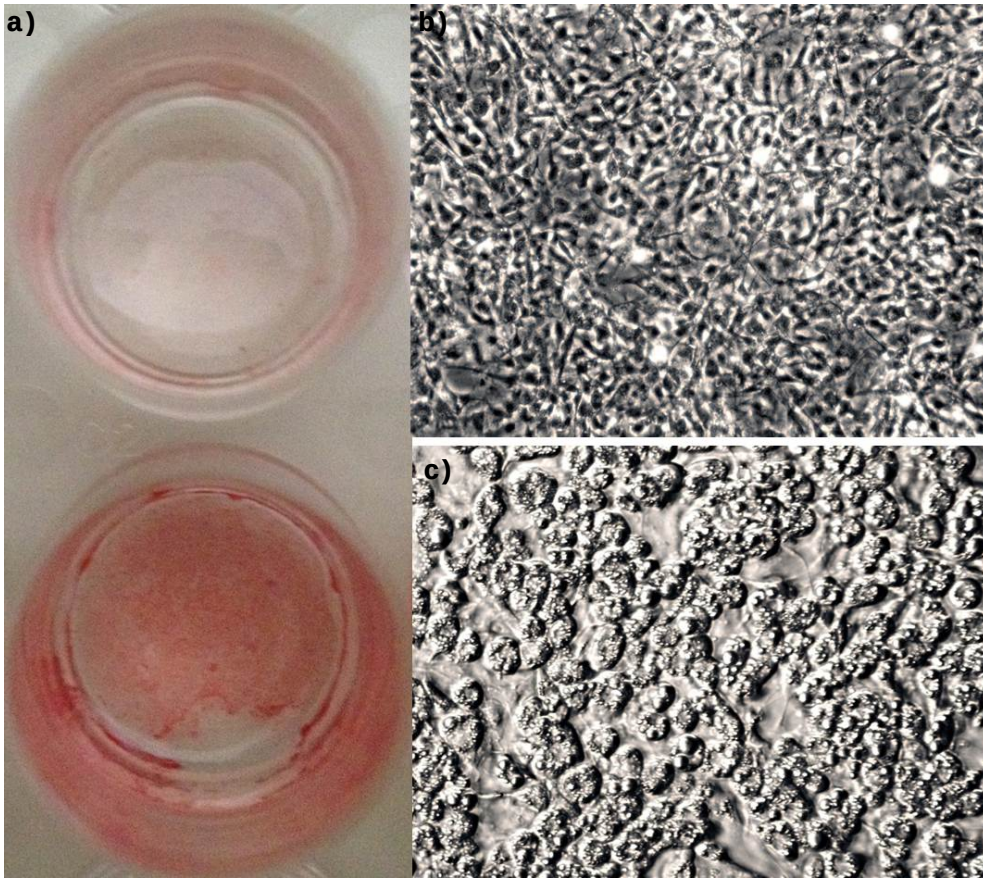

Figure 4.2 Differentiation of 3T3-L1 preadipocytes into adipocytes. a) Oil red o staining. Undifferentiated control cells (top well) and differentiated cells (bottom well) were subjected to oil red o staining which binds to lipid droplets present in adipocytes. b) Undifferentiated control cells eight days post induction of differentiation x20 magnification. c) Differentiated control cells eight days post induction of differentiation x20 magnification.

4.3.2 rt-PCRs for mRNA expression

3T3-L1 mouse preadipocytes do not express leptin but after differentiation, the adipocytes generated do express leptin (MacDougald *et al*, 1995). Rt-PCRs were performed to determine the presence of leptin mRNA in 2µg total RNA, isolated

from approximately $5x10^6$ undifferentiated control cells and adipocytes. Undifferentiated control cells did not express leptin whilst differentiated adipocytes did (Figure 4.3).

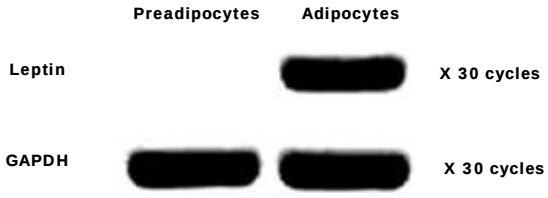

Figure 4.3 Leptin mRNA expression in preadipocytes and adipocytes

Rt-PCRs were also performed for mRNAs encoding Sp1 and C/EBPα, transcription factors which are associated with expression of the leptin gene, to determine if any adverse changes in their expression had occurred. These mRNAs were expressed in both undifferentiated and differentiated cells (Figure 4.4).

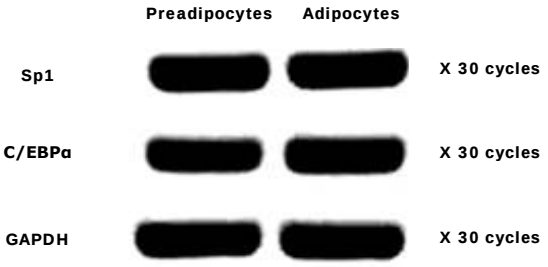

Figure 4.4 Sp1 and C/EBPα mRNA expression in preadipocytes and adipocytes

Presented in the previous chapter were Rt-PCRs for mRNAs encoding DNMT3L, the catalytically inactive member of the DNA methyltransferase family which is known to enhance DNMT3a and DNMT3b function (Gowher *et al*, 2005). In order to determine if changes in DNMT3L expression correlated with the adipogenic transition, rt-PCRs were repeated on preadipocytes and adipocytes. mRNAs encoding DNMT3L were not detected in preadipocytes but differentiation into adipocytes correlated with DNMT3L mRNA expression (Figure 4.5).

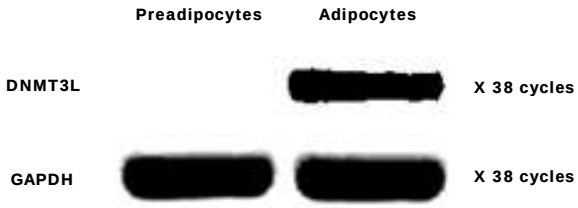

Figure 4.5 DNMT3L mRNA expression in preadipocytes and adipocytes

Rt-PCRs were also performed for TET1 and TET2, members of the family of enzymes which facilitate the oxidation of 5-mC to 5-hmC. mRNAs encoding TET1 and TET2 were detected in both preadipocytes and adipocytes (Figure 4.6). Rt-PCRs were also attempted for TET3 mRNA however no band was observed and a positive control could not be obtained so it is uncertain as to whether these cells did not express TET3 mRNA or if they assay was not fully optimised (data not shown).

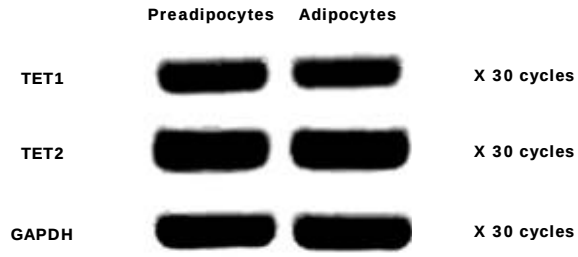

Figure 4.6 TET1 and TET2 mRNA expression in preadipocytes and adipocytes

4.3.3 DNA methylation at the leptin promoter in preadipocytes and adipocytes

Pyrosequencing for both the sense and antisense strand of the leptin promoter was used to determine DNA methylation levels at the leptin promoter in preadipocytes and adipocytes eight days post differentiation. Pyrosequencing portrays the percentage of DNA molecules exhibiting DNA methylation at individual CpG sites. This analysis was performed on DNA samples from three individual differentiation experiments and the mean DNA methylation level from all nine CpG sites combined was used for data analysis. Standard deviations between the three differentiations were no more than 7% at each CpG site.

4.3.3.1 Overall DNA methylation levels at the leptin promoter

To assess differences in overall DNA methylation levels at the core leptin promoter, the mean DNA methylation level between the nine CpGs was calculated and compared between preadipocytes and adipocytes (Figure 4.7).

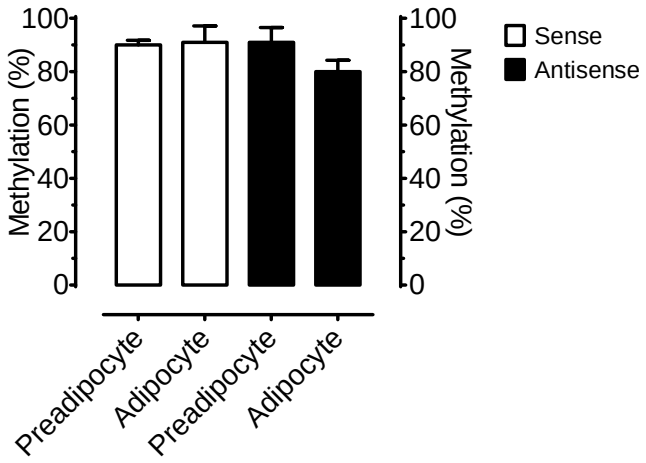

Figure 4.7 Overall DNA methylation levels on both strands of the leptin promoter in preadipocytes and adipocytes

On the sense strand both the preadipocytes and adipocytes exhibited a high level of DNA methylation at the leptin promoter at 90% and 91% methylation respectively. On the antisense strand high levels of DNA methylation were again observed in the preadipocytes (91% methylation), however, the adipocytes exhibited an 11% loss in methylation in comparison to the preadipocytes (ANOVA p=0.003).

4.3.3.2 DNA methylation levels at individual CpGs of the leptin promoter

Next, DNA methylation levels at individual CpGs of the leptin promoter were compared between preadipocytes and adipocytes. On the sense strand of the leptin promoter, preadipocytes exhibited high levels of DNA methylation between 76% and 100% (Figure 4.8). The adipocytes also exhibited high levels of DNA methylation which was either the same as, or slightly lower/higher at each of the nine CpGs in comparison to the preadipocytes (Table 4.3).

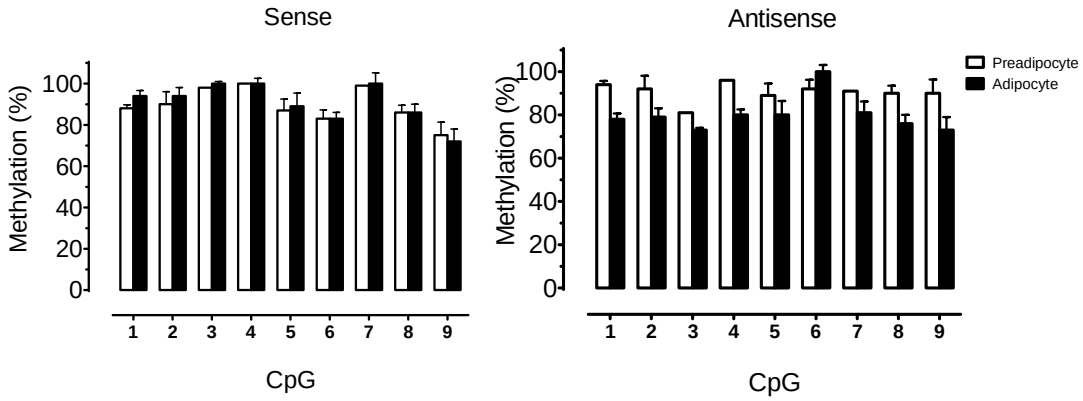

Figure 4.8 DNA methylation levels at individual CpGs of the leptin promoter in preadipocytes and adipocytes

On the antisense strand of the leptin promoter preadipocytes also exhibited high levels of DNA methylation between 81% and 96% at all CpGs analysed. High levels of DNA methylation were also observed in the adipocytes however at most CpGs, methylation levels had decreased in comparison to the preadipocytes. At CpG-6, levels had increased marginally (Table 4.3).

Table 4.3 Percentage changes in DNA methylation levels at individual CpGs of the leptin promoter between preadipocytes and adipocytes

CpG	Sense	Antisense
1	+6	-16
2	+4	-13
3	+2	-8
4	0	-16
5	+2	-9
6	0	+8
7	+1	-10
8	0	-14
9	-3	-17

4.3.4 Detection of 5-hmC

Although the decrease in DNA methylation levels in the adipocytes on the sense strand of the leptin promoter was statistically significant, there did not appear to be a dramatic change in DNA methylation levels between the non-leptin expressing preadipocytes and the leptin expressing adipocytes. Pyrosequencing does not differentiate between 5-mC and 5-hmC, therefore, this analysis did not take into account the specific cytosine modifications that were present within the observed DNA methylation and although total DNA methylation did not change noticeably the actual ratio of 5-mC to 5-hmC may have. Two separate assays were used to detect 5-hmC, the NEB Epimark kit (Figure 4.9) and Active Motifs Hydroxymethyl collector (Figure 4.9).

The NEB Epimark kit was used to detect 5-hmC at CpG-4, located adjacent to a C/EBPα transcription factor binding site. 5-hmC was not detected in the preadipocytes and it appeared that all or the majority of DNA methylation at this CpG consisted of 5-mC (Figure 4.10).

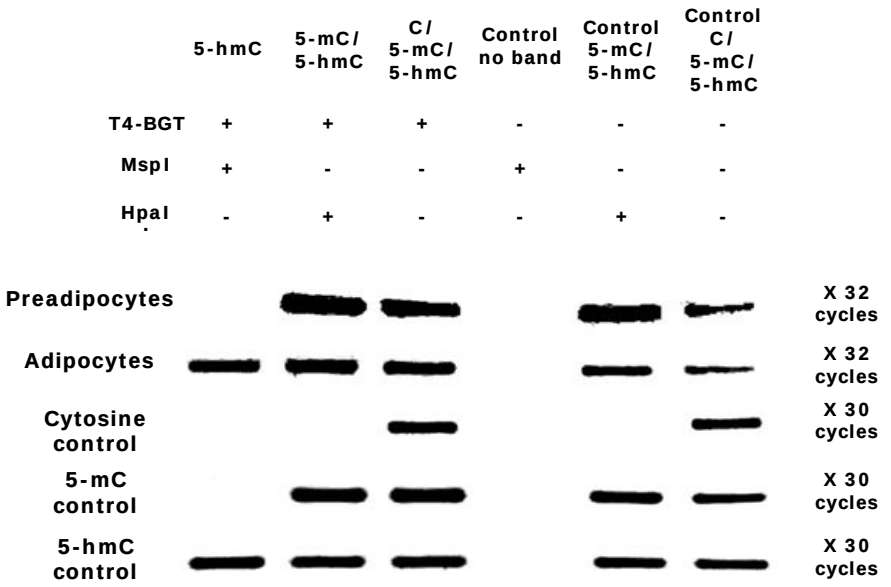

Figure 4.9 Detection of 5-hmC using the NEB Epimark kit

In the adipocytes, however, 5-hmC was detected at CpG-4 and, although this assay was not quantified, the band produced for total DNA methylation (5-mC and 5-hmC) was of a similar intensity to the band produced for 5-hmC. This suggests that, in the adipocytes, the DNA methylation at CpG-4 may consist entirely of 5-hmC. Active Motif's Hydroxymethyl Collector was used to detect the presence of 5-hmC over the whole of the region analysed rather than at a

specific CpG (Figure 4.10). This confirmed the previous findings from the NEB Epimark kit. In the preadipocytes no 5-hmC was detected whilst in the adipocytes 5-hmC was detected.

Figure 4.10 Detection of 5-hmC using the Active Motif's Hydroxymethylcollector

4.3.5 Time course experiment

3T3-L1 preadipocytes are fully differentiated into adipocytes by day eight post induction of differentiation. In the previous experiments differences in mRNA expression and DNA methylation between adipocytes and preadipocytes were determined at day eight post induction of differentiation. It was found that the differentiation process correlated with the induction of leptin and DNMT3L mRNA expression. Although no dramatic changes in overall DNA methylation levels were observed, the presence of 5-hmC at the leptin promoter correlated with leptin expression. The differentiation experiment was repeated, harvesting cells at 20, 26, 32, 38, 44, 50, 96 hours and eight days post induction of differentiation to determine the time point at which these changes occurred during the adipogenic process.

Leptin mRNA was detected in the adipocytes from 32 hours post induction of differentiation and was not detected at any time points within the preadipocytes. Although the rt-PCR data were not quantified, the intensity of the PCR band produced increased at every time point after 32 hours (Figure 4.10). mRNAs encoding the transcription factors Sp1 and C/EBPα were detected in both the preadipocytes and adipocytes at every time point and there did not appear to be any obvious variation in band intensity between the cell types or between the time points, suggesting that they are stably expressed throughout the adipogenic process (Figure 4.10). DNMT3L was not expressed in the preadipocytes and mRNAs encoding DNMT3L were detected in the adipocytes only at day eight, the end point of differentiation (Figure 4.10). This rt-PCR was performed over 38 cycles in comparison to other rt-PCRs which were performed

over 30 cycles. TET1 and TET2, the enzymes which facilitate the oxidation of 5-mC to 5-hmC, were also expressed in both the preadipocytes and adipocytes at every time point analysed (Figure 4.11).

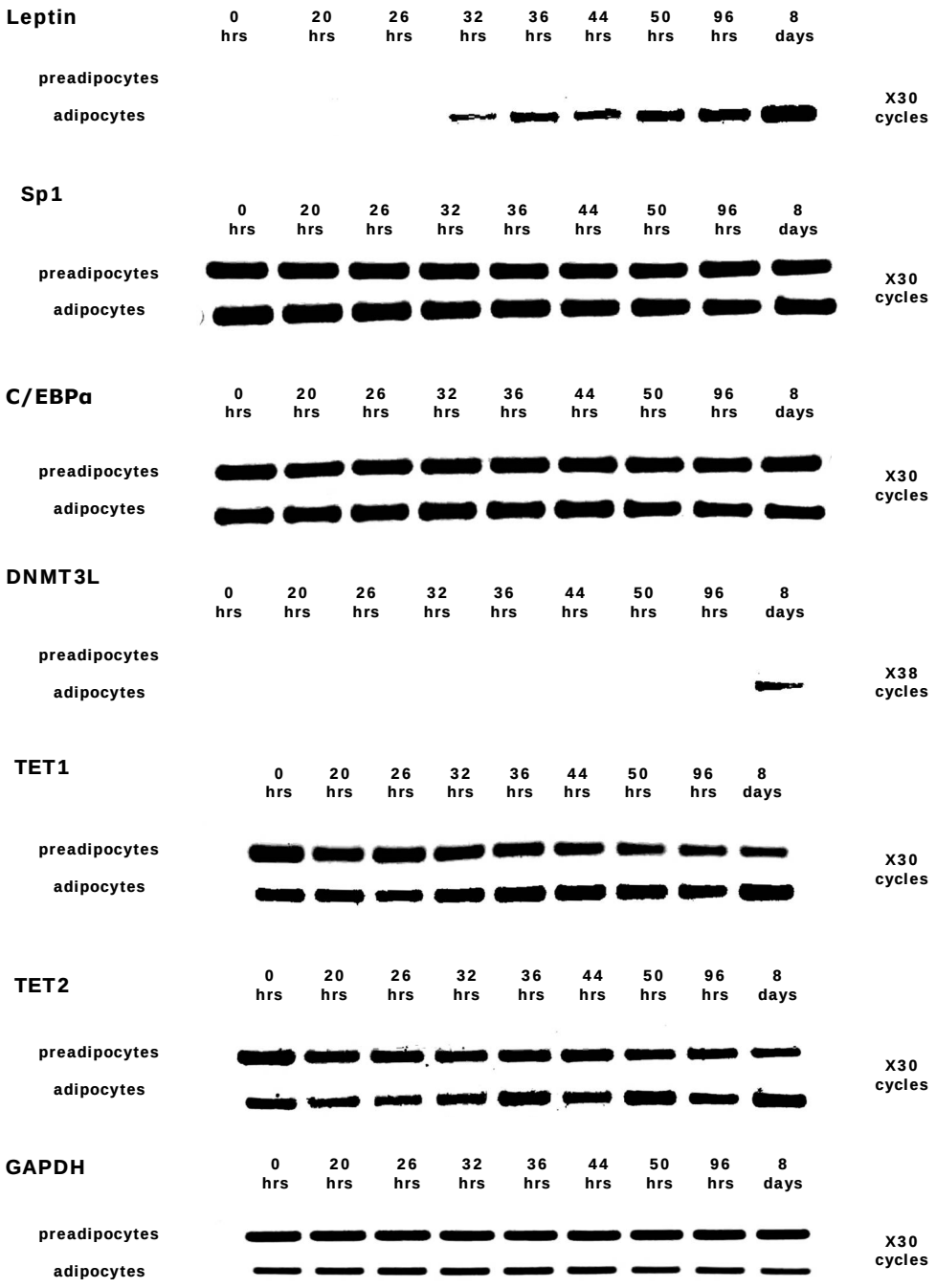

Figure 4.11 mRNA expression at various time point during the differentiation of 3T3-L1 preadipocytes into adipocytes

Pyrosequencing analysis for DNA methylation levels at the leptin promoter was repeated for each time point analysed. Although the time course differentiation was repeated three times, due to technical difficulties with the pyrosequencer, pyrosequencing analysis could only be performed once and only on the sense strand of the leptin promoter.

Overall DNA methylation levels at the leptin promoter remained high throughout the differentiation process in both the preadipocytes and the adipocytes which exhibited between 88% and 93% DNA methylation at every time point analysed (Figure 4.12).

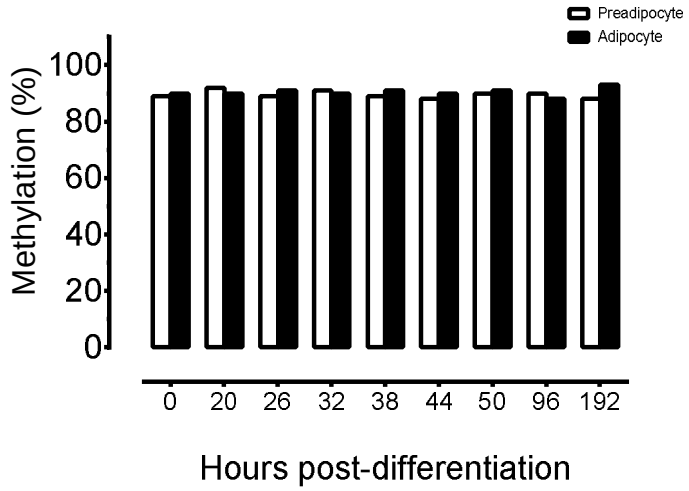

Figure 4.12 Overall DNA methylation levels at the leptin promoter during adipogenic differentiation.

At each time point only minor increases/decreases in DNA methylation of between -2 to +5% were observed between the preadipocytes and the adipocytes (Table 4.3).

Table 4.4 Percentage changes in overall DNA methylation levels at the leptin promoter between preadipocytes and adipocytes at various time points

Time	Sense strand
0 hrs	+1
20 hrs	-2
26 hrs	+2
32 hrs	-1
38 hrs	+2
44 hrs	+2
50 hrs	+1
96 hrs	-2
8 days	+5

Next DNA methylation levels at individual CpGs of the leptin promoter during the differentiation process were analysed (Figure 4.13). DNA methylation levels remained comparable between the preadipocytes and adipocytes throughout the differentiation process, although there were some increases/decreases at certain CpGs (Table 4.5). At 20 hours post-induction of differentiation DNA methylation levels at CpG-8 decreased by 11% in the adipocytes relative to the preadipocytes and at 26 hours post-induction of differentiation DNA methylation levels at CpG-1 increased by 15% in the adipocytes relative to the preadipocytes. At 96 hours a decrease of 11% was also observed at CpG-9.

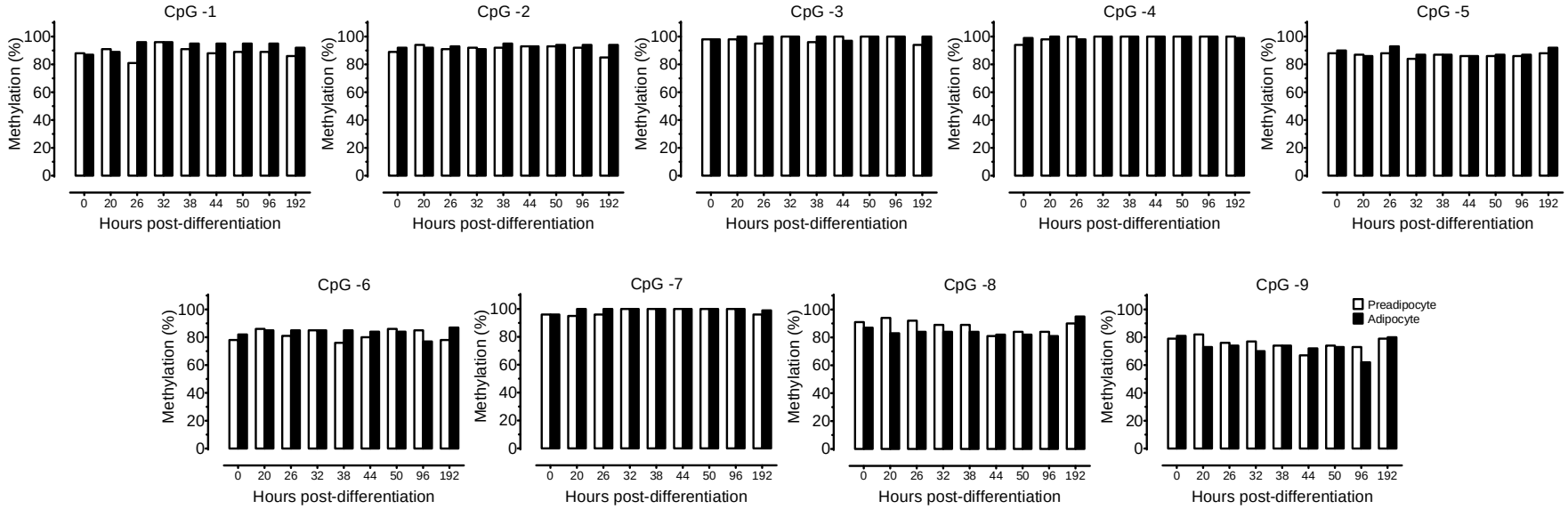

Figure 4.13 DNA methylation levels at individual CpGs of the leptin promoter during the differentiation of preadipocytes into adipocytes

Table 4.5 Percentage changes in DNA methylation levels at individual CpGs of the leptin promoter between preadipocytes and adipocytes at various time points

	0 hrs	20 hrs	26 hrs	32 hrs	38 hrs	44 hrs	50 hrs	96 hrs	192 hrs
CpG-1	-1	-2	+15	0	+4	+7	+6	+6	+6
CpG-2	+3	-2	+2	-1	+3	0	+1	+2	+9
CpG-3	0	+2	+5	0	+4	-3	0	0	+6
CpG-4	+5	+2	-2	0	0	0	0	0	-1
CpG-5	+2	-1	+5	+3	0	0	+1	+1	+4
CpG-6	+4	-1	+4	0	+9	+4	-2	-8	+9
CpG-7	0	+5	+4	0	0	0	0	0	+3
CpG-8	-4	-11	-8	-5	-5	+1	-2	-3	+5
CpG-9	+2	-9	-2	-7	0	+5	-1	-11	+1

As the pyrosequencing analysis does not differentiate between 5-mC and 5-hmC, the potential presence of 5-hmC was again analysed by the two techniques described previously. The NEB Epimark kit was used to determine the presence of hydroxymethylation at CpG-4 at each of the nine time points analysed (Figure 4.13). 5-hmC was not detected in the preadipocytes at any of the time points analysed. In the differentiating adipocytes, 5-hmC was initially detected at 26 hours post-induction of differentiation. However, at the next two time points of 32 and 38 hours 5-hmC was not detectable. Although 5-hmC was not detected at these two time points, a PCR band was still produced for total DNA methylation (5-mC and 5-hmC). This suggests that the DNA methylation present at these time points is purely or predominantly 5-mC. 5-hmC was re-detected at 44 hours and remained detected at every subsequent time point. Although this data was not quantified the intensity of the bands was comparable at 26, 44 and 50 hours suggesting that levels of 5-hmC may be similar at these time points. The intensity of the PCR bands produced for 5-hmC increased at 96 hours and remained relatively similar at 192 hours, again suggesting that at these time points 5-hmC levels may be similar. At these late stages of

differentiation (time points 96 hours and 192 hours), the intensity of the band produced for 5-hmC was comparable to that produced for total DNA methylation suggesting that at these time points the DNA methylation observed at CpG-4 in the adipocytes may consist purely or predominantly of 5-hmC.

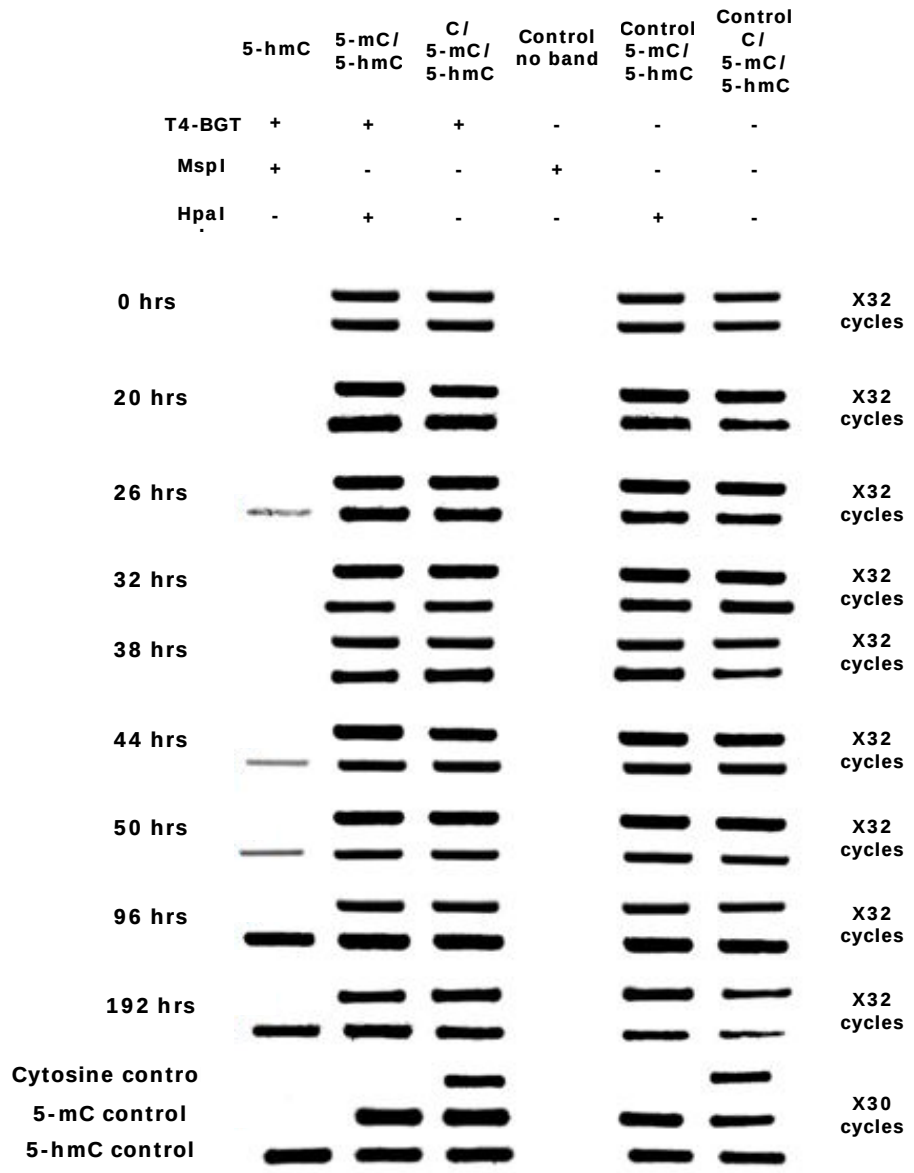

Figure 4.14 Detection of 5-hmC using the NEB Epimark kit at various time points during differentiation. At each time point the top row of bands represents preadipocytes and the bottom row of bands represents differentiating adipocytes.

Active Motifs Hydroxymethyl Collector was used to detect 5-hmC within the region of interest rather than just one specific CpG site (Figure 4.14). Whereas using the CpG site specific technique (CpG-4), 5-hmC became detectable at 26 hours, using this technique, 5-hmc was first detected at 20 hours. This may

stem from differences in the sensitivities between the two different techniques or may suggest that CpG-4 becomes hydroxymethylated at a later time point than other CpGs in the region. 5-hmC was detected at every time point subsequent to 20 hours and although this data was not quantified the intensities of the bands produced suggests that there may be differences in the level of hydroxymethylation present at the different time points.

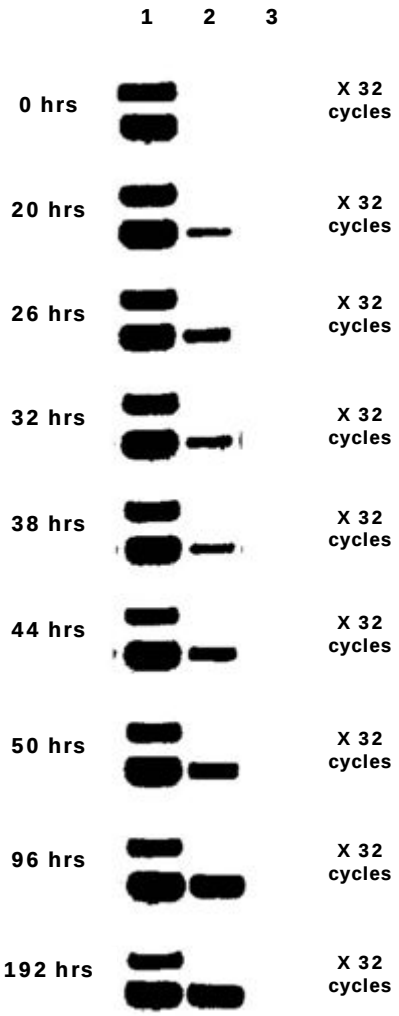

Figure 4.15 Detection of 5-hmC using Active Motif's Hydroxymethyl Collector at various time points during differentiation. At each time point the top row of bands represents preadipocytes and the bottom row of bands represents differentiating adipocytes.

At 26 hours the band produced was of greater intensity that the previous time point of 20 hours and the next time point of 32 hours. This is concurrent with

the data produced from the NEB Epimark assay where 5-hmC was first detected at 26 hours whilst previous to this and at the next two time points 5-hmc was not detected. From 38 hours the bands gradually increase in intensity and at 96 and 192 hours are comparable in intensity. This was again also observed in the data from the NEB Epimark kit where at 96 and 192 hours bands of similar intensity were produced suggesting comparable levels of 5-hmC.

4.3.6 5-hmC at single base resolution

Pyrosequencing was repeated in fully differentiated 3T3-L1 cells (8 day) with the prior treatment of DNA with $KRuO_4$, an oxidising agent that converts 5-hmC to 5-fC, thereby removing 5-hmC from analysis (Section 4.2.5.2). The 5-hmC oxidising ability of KRuO4 was reported near to the deadline for completion of laboratory work for this thesis (Booth *et al*, 2012). Additionally, technical difficulties were experienced and the data set could not be completed (Appendix I). Preliminary data for CpG-6 to CpG-9 at the antisense strand of the leptin promoter was obtained and is presented here (Figure 4.16). This data is described as preliminary due to a lack of experimental repeats and the pyrosequencer analysis software flagging the run as "uncertain due to low peak height". From previous experience, data with this error message can still give accurate estimates of DNA methylation levels, however, the experiment would need to be repeated and optimised to draw solid conclusions.

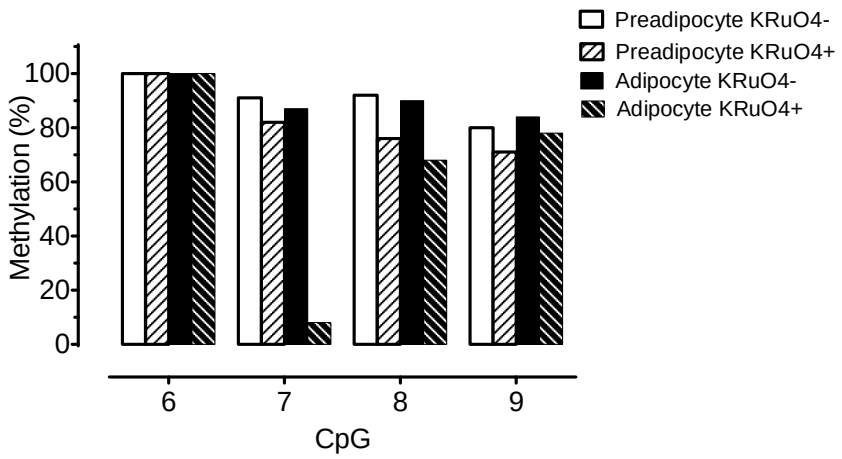

Figure 4.16 5-hmC at single base resolution in 3T3-L1 preadipocytes and adipocytes.

At CpG-6 to CpG-9, on the lower strand of the leptin promoter, DNA methylation in the preadipocytes appeared to consist predominantly of 5-mC with some evidence of marginal levels of 5-hmC (Table 4.6). DNA methylation at CpG-6 in

the adipocytes appeared to consist of only 5-mC, whilst at CpG-8 and CpG-9 there was evidence for the presence of 5-hmC (Table 4.6). Post-KRuO4 oxidation, the most dramatic change in DNA methylation levels was at CpG-7 in the adipocytes. DNA methylation level dropped by 79%, indicating that hydroxymethylation is the predominant cytosine modification at this CpG site in fully differentiated 3T3-L1 adipocytes.

Table 4.6. The percentage contributions of 5-mC and 5-hmC to DNA methylation levels at the leptin promoter in preadipocytes and adipocytes

CpG	Preadipocytes		Adipocytes	
	5-mC	5-hmC	5-mC	5-hmC
6	100	0	100	0
7	82	9	8	79
8	76	16	68	22
9	71	9	78	6

4.4 Discussion

4.4.1 Dramatic changes in overall DNA methylation levels do not correlate with leptin expression

Previous to the recent re-discovery of 5-hmC in mammalian DNA, DNA methylation at CpG rich promoters was considered to be an epigenetic modification associated with gene silencing. The leptin promoter, which is a CpG rich region, exhibits a broad range of DNA methylation levels in mouse adipose tissue, the primary site of leptin expression (Stöger, 2006). This suggests that DNA methylation at the leptin promoter could be compatible with or even facilitate leptin expression rather than leading to silencing. Although the predominant component are the leptin expressing adipocytes, adipose tissue is a heterogeneous organ also containing nerves, immune cells, endothelial cells, fibroblasts and fibroblastic preadipocyte cells (Cristancho & Lazar, 2011). As a consequence, DNA methylation pattern in adipose tissue may also be attributed to cell types that may not express leptin such as connective tissue or the adipocyte precursor cells, preadipocytes. Using 3T3-L1 cells is a great

advantage as analysing methylation levels purely in leptin expressing adipocytes reduces the complexity.

In the results described in this chapter, a high level of DNA methylation was observed at the leptin promoter in 3T3-L1 preadipocytes which, consistent with a large body of literature, do not express leptin (MacDougald *et al*, 1995; Yokomori *et al*, 2002). After differentiation into adipocytes, which do express leptin, no significant change in DNA methylation was observed on the sense strand of the leptin promoter. On the antisense strand there was a drop in methylation levels, although not a dramatic drop. This was statistically significant when analysing the whole region but not under CpG specific analysis. This raises the question what is the role of DNA methylation here and could 5-hmC be part of the puzzle?

4.4.2 Induction of leptin mRNA expression correlates with hydroxymethylation of the leptin promoter

Although there was no change in total DNA methylation levels (5-mC and 5-hmC) between the preadipocytes and the leptin expressing adipocytes, hydroxymethylation was detected by two separate methods post differentiation, again supporting the evidence presented in chapter three, for the relationship between leptin expression and the presence of hydroxymethylation at the leptin promoter.

In vitro transcription assays have been used to demonstrate that 5-mC at specific sites of the human leptin promoter diminishes leptin expression (Melzner *et al*, 2002). This may be attributed to the binding of MeCP2, which has been suggested to contribute to the silencing of the leptin gene when it exhibits cytosine methylation (Melzner *et al*, 2002). MeCP2 is part of the MBD family of proteins, most of which do not recognise 5-hmC (Jin *et al*, 2010; Valinluck *et al*, 2004) and most likely dissociate from DNA when 5-mC is converted into 5-hmC (Williams *et al*, 2012). The presence of just one 5-hmC base in a DNA fragment significantly inhibits binding of MeCPs (Ruzov et al, 2011; Valinluck et al, 2004). The epigenetic regulation of gene expression could be influenced by 5-hmC through the exclusion of methyl CpG binding proteins, a theory which has been supported by the inability of MeCP2 to recognise and bind to 5-hmC. The high levels of DNA methylation observed in the 3T3-L1 preadipocytes may be attributed to cytosine methylation (5-mC) to which MeCP2 would be bound, preventing leptin transcription. Upon differentiation this 5-mC may be oxidised by the TET enzymes, which are expressed both pre and post differentiation,

forming 5-hmC and, as evidence shows hydroxymethylation diminishes MeCP2 binding, would allow the release of MeCP2 and transcription and expression of the leptin gene.

The observed lack of change in DNA methylation levels contradicts a previous report by Yokomori *et al* who observed demethylation of the leptin promoter after the differentiation of 3T3-L1 preadipocytes into adipocytes (Yokomori *et al*, 2002). The group used conventional bisulfite sequencing to investigate DNA methylation levels and the observed demethylation may have been a consequence of analysing too few epialleles which may have biased the result. A repeat of the 3T3-L1 differentiation experiment by Stöger using the Hp-bss technique (data unpublished) also revealed no distinct change in DNA methylation levels post-differentiation. More recently, slight decreases in DNA methylation post-differentiation of these cells has been reported (Bahar *et al*, 2013). In the work presented in this chapter, pyrosequencing was used to determine DNA methylation levels. This technique determines the total DNA methylation levels in a population of cells rather than in individual DNA molecules therefore eliminating this possible bias. Melzner *et al* also reported demethylation of the leptin promoter in a human adipocyte cell line (Melzner *et al*, 2002). 5-hmC may play a role as an intermediate in a demethylation pathway. It is also possible that the observed hydroxymethylation facilitates demethylation of specific CpGs within the leptin promoter and that the demethylation observed by Yokomori *et al* and Melzner *et al* was a consequence of this.

4.4.3 Leptin expression correlates with CpG specific hydroxymethylation

The restriction enzyme and pulldown techniques used here to detect hydroxymethylation do not allow for single base resolution and only inform of the presence of 5-hmC at CpG-4 or at any CpG within the leptin promoter. In 2012, two novel techniques were reported which allowed single base resolution of 5-hmC (Booth *et al*, 2012; Yu *et al*, 2012). One involved the oxidation of DNA with potassium perruthenate pre-bisulfite conversion which would allow the oxidation of 5-hmC to 5-fC and therefore removing it from the pyrosequencing analysis. Although this technique was attempted, the results are preliminary and due to time constraints could not be repeated.

Preliminary results revealed that on the antisense strand, CpG-7 lost 79% of the DNA methylation observed in the non-potassium perruthenate treated cells indicating that, post-differentiation, CpG-7 becomes highly hydroxymethylated

whilst CpG-6, CpG-8 and CpG-9 exhibit little or no hydroxymethylation. Melzner *et al* found that methylation (5-mC) of the C/EBPα transcription factor binding site down-regulated leptin expression and methylation of CpGs proximal to the TATA box element completely diminished promoter activity. CpG-7 is located 62bp upstream of a TATA box element and hydroxymethylation of this site may facilitate transcription of the leptin gene. The C/EBPα transcription factor binding site exhibited high levels of DNA methylation both pre and post differentiation and has been reported to be 1.8 fold as methylated as other sites in the same region even though this is the binding site for the predominant transcription factor of the leptin gene (Stöger, 2006). It is likely that the observed DNA methylation at this site in leptin expressing cells is actually hydroxymethylation and further experimentation is required to determine this.

4.4.4 Limitations of this experiment

Hydroxymethylation is reported to be a tissue specific modification (Nestor *et al*, 2012) and may be important for the process of adipogenesis as a genome wide study has found that 5-hmC is associated with transcription factor binding to regulatory sites during 3T3-L1 differentiation and may be involved in the activation of tissue specific genes (Sérandour et al, 2012). It could also be possible that the 5-hmC detected at the leptin promoter may be a result of the adipogenic transition rather than specifically to facilitate leptin expression. It would be interesting to assess other tissues which express leptin to further investigate the correlation between leptin expression and 5-hmC. The level of leptin expression in other tissues is marginal in comparison to adipocytes and more sensitive gene specific techniques may be required for the detection of 5-hmC. The further optimisation of pre-incubation with $KRuO_4$ coupled to pyrosequencing will allow for this. DNA methylation levels frequently change in cultured cells (Antequera & Bird, 1993) and it has been reported that 5-hmC and TET levels are also affected in culture (Nestor *et al*, 2012). It is therefore important to continue with an *in vivo* approach, possibly examining adipose tissue for hydroxymethylation, although the heterogeneity of this tissue may make this technically challenging.

4.5 Conclusion

Although a previous report suggests that demethylation of the leptin promoter modulates leptin expression of 3T3-L1 adipocytes, in this chapter evidence was provided to suggest that the leptin promoter retains overall DNA methylation

levels during adipogenesis in the 3T3-L1 cell line. Although total DNA methylation levels did not change dramatically, the specific ratio of 5-mC to 5-hmC within the observed DNA methylation did change. The leptin promoter appears to become hydroxymethylated during the transition to a leptin-expressing state and this hydroxymethylation is CpG-site specific rather than occurring over the whole of the leptin promoter. Further work is required to validate the necessity of hydroxymethylation in the expression of leptin and also to confirm the CpG specific hydroxymethylation levels in 3T3-L1 adipocytes at single base resolution.

Chapter 5. General Discussion

Leptin is an important hormone, the full role of which we have yet to elucidate. The leptin promoter lies within a CpG rich region and predominantly cytosines of CpG dinucleotides can be subject to DNA methylation. The role of this modification appears to lie in the regulation of gene expression. Despite this knowledge there is little information regarding DNA methylation at the leptin promoter and there have been no reported attempts to differentiate between the types of cytosine modifications that may be present. The aim of this thesis was to contribute information regarding DNA methylation at the leptin promoter.

Investigation of the roles of the individual DNMT enzymes, which are responsible for establishing and maintaining cytosine methylation patterns revealed that whilst DNMT1 is regarded as the maintenance methyltransferase, at the leptin promoter it appears that this DNMT alone is unable to maintain DNA methylation in the absence of DNMT3a and DNMT3b. Additionally, DNMT3a and DNMT3b may also contribute to maintenance methylation at the leptin promoter. Whilst it is assumed that DNA methylation patterns between cytosines of opposing CpG dinucleotides have the same methylation status, evidence was presented here and also in other publications to show that this is not always the case.

Past reports have suggested that early life nutrition correlates with DNA methylation levels at the leptin promoter (Jousse *et al*, 2011; Milagro *et al*, 2009; Tobi *et al*, 2009). A high fat prenatal diet appeared to have no effect on DNA methylation levels in both visceral and subcutaneous adipose tissue of pigs. However, regardless of the tissue type or group, the pig leptin promoter exhibited higher levels of DNA methylation at specific CpGs. This included a CpG which located within a C/EBPα transcription factor binding site. This has also been demonstrated in human and mouse adipose tissue (Stöger, 2006). DNA methylation is usually associated with gene silencing and it is interesting that in primary tissues of leptin expression, higher levels of DNA methylation would be observed at a CpG which is imperative to the transcription of the leptin gene. Indeed, *in vitro*, DNA methylation at this particular CpG site leads to down-regulation of leptin promoter activity and these findings may be explained by the heterogeneous nature of adipose tissue. The DNA methylation may actually correspond to a non-leptin expressing cell type rather than the leptin expressing adipocytes. However, *in vitro* differentiated murine adipocytes, with robust transcription of leptin, also revealed high levels of DNA methylation. Furthermore, there was no noticeable difference between DNA methylation levels

at the leptin promoter in preadipocytes, which do not express leptin, and adipocytes, which do express leptin.

In both mESCs and adipocytes, evidence is presented that leptin expression correlates with cytosine hydroxymethylation. Many techniques used to investigate DNA methylation do not distinguish between 5-mC and 5-hmC (Huang *et al*, 2010; Nestor *et al*, 2010). It could be the case that failure to distinguish between these two cytosine modifications has led to an inaccurate interpretation of data. These findings may explain previous observations of DNA methylation at the leptin promoter in leptin expressing tissues as several roles have been proposed for 5-hmC, including one as a transcriptional activator (*reviewed by* Williams *et al*, 2011). Although *in vitro* 5-mC at specific CpGs of the leptin promoter down-regulates leptin promoter activity (Melzner *et al*, 2002), the effect of 5-hmC is yet to be determined. Preliminary data presented in this thesis suggests that cytosine hydroxymethylation of the leptin promoter is compatible with expression. However, before this is discussed it is important to address the possibility that another epigenetic modification is responsible for regulation of leptin expression. DNA methylation is one of many markings which constitute the epigenetic landscape and work in concert to regulate gene expression. The low levels of DNA methylation at the leptin promoter in human (Stöger, 2006) and porcine adipose tissue suggest that DNA methylation does not regulate expression of this gene in these species. Whilst here it is suggested that 5-mC and 5-hmC play a role in the regulation of leptin gene expression in mouse, it is also a possibility that 5-hmC is merely an artefact left by transcriptional machinery and therefore is not specifically required for gene expression. Data from the ENCODE project shows that mESCs do not express leptin and are enriched for the repressive histone marks H3K9me3 and H3K27me3 (Appendix H). Histone modifications, either in concert with or independently of DNA methylation, may play a role in regulating transcription and expression of leptin. Trimethylated histone marks which are associated with repression of the leptin gene in mESCs are known to stimulate the polycomb group protein complex, PCR2 (Margueron *et al,* 2009). These proteins have an important role in the regulation of repressive chromatin states and the activity of PCR2 can then stimulate trimethylation of other unmodified residues, thereby mediating the silencing process (Margueron *et al*, 2009). This is one alternative pathway to DNA methylation however, specifically in adipocytes, CpG specific hydroxymethylation correlates with leptin expression. *In vitro*, it has been shown that C/EBPα, which exhibits higher DNA methylation levels than other CpGs in the region can bind to both an unmethylated and a methylated (5-mC)

binding site (Melzner *et al*, 2002; Rishi *et al*, 2010) but the specific effect of a hydroxymethylated binding site is yet to be determined.

Conclusion

Although the work presented in the thesis focuses only on the leptin promoter, there are several general findings which are important for the investigation of DNA methylation at any gene. When performing experiments for DNA methylation analysis, many studies focus only on one of the strands of a DNA molecule. Whilst the general consensus on the maintenance of DNA methylation patterns leads many to believe that the opposing CpG dinucleotides of CpG dyads will have the same methylation status, there is evidence presented in this thesis and in the work of others to show that this is not necessarily the case (Arand *et al*, 2012). Furthermore, the preliminary finding that hydroxymethylation of just one CpG in a stretch of DNA may lead to gene expression exemplifies the importance of *i*) analysing individual CpGs and *ii*) the necessity to differentiate between 5-mC and 5-hmC and possibly other cytosine modifications. Even with the development of techniques which can differentiate between these two cytosine modifications, many studies currently choose to ignore the ability to differentiate between the two. Demonstrated here is the possibility of opposing roles of these two cytosine modifications. In order to fully advance our knowledge concerning the role of DNA methylation in the epigenetic landscape it will be prudent to investigate DNA methylation patterns on complementary strands, to carefully consider the specific CpGs under analysis and to differentiate between 5-mC and 5-hmC or any other, additional types of cytosine modifications that are discovered in the future.

Appendices

Appendix A. Full composition of experimental diets

	Control	High fat	Lactation
Feed stuff		**% / kg of diet**	
Tapioca	28.1	3.1	-
Rapeseed meal	10.0	10.0	4.0
Sunflower seed meal	4.0	4.0	2.0
Soybean hulls	13.0	17.0	2.0
Sugar beet pulp	10.0	10.0	2.0
Palm oil	-	6.6	3.1
Soybean oil	0.5	0.50	0.97
Maize	-	-	10.0
Soybean meal	-	-	11.7
Wheat	10.0	10.0	26.4
Barley	10.0	10.0	15.0
Wheat middling's	15.0	15.0	15.0
Molasses	4.0	4.0	4.0
Monocalcium phosphate	0.22	0.22	0.48
Salt	-	-	0.37
Limestone	0.50	0.60	1.56
Premixed vitamin and minerals	0.50	0.50	0.50
Lysine (25%)	0.17	0.17	-
Lysine-HCl (L, 79%)	-	-	0.17
Threonine (L, 98%)	-	-	0.02
Phytase	0.50	0.50	0.50
Threonine 15% 2390	0.01	0.01	-
Sodium bicarbonate	0.60	0.60	0.16
Diet composition		**g / kg**	
Nutrients			
Ash	64	52	61
Crude protein	123	122	161
Crude fat	25	89	63
Crude fibre	120	121	52
ID starch and sugar	368	212	367
Sugar	61	60	59
Starch	316	161	319
AID LYS	4.03	4.12	7.10
AID MET	1.57	1.58	2.14
AID M+C	3.20	3.24	4.47
AID THR	2.62	2.68	4.29
AID TRP	0.98	0.99	1.53
LYS	5.92	5.98	8.58
MET	2.07	2.04	2.55
CYS	2.46	2.46	3.05
M&C	4.53	4.50	5.60
THR	4.54	4.50	5.87
TRP	1.48	1.46	1.96
ILE	4.37	4.35	6.13
ARG	7.10	7.04	9.83
FEN	5.05	5.02	7.35
HIS	3.16	3.13	4.09
LEU	7.81	7.76	11.46
TYR	3.81	3.86	5.15
VAL	5.84	5.79	7.49
ALA	5.39	5.27	7.07
ASP	9.57	9.53	13.53
GLU	22.04	21.74	32.66
GLY	6.17	6.26	7.09
PRO	7.66	7.64	11.08
SER	5.44	5.46	7.49

Ca	6.78	6.70	9.31
Na	2.00	1.98	2.17
Ferm CHO	235	234	145
Ind. CHO	122	110	80
Dig. P	2.47	2.41	3.27
NSP	343	330	209
Dig. NSP	224	223	132
dEB	301	280	220
C18:2	7.7	14.6	15.7
C18:3	1.04	1.22	1.46
C18:2+3	8.8	15.8	17.1
C16+18>0	12.59	44.20	33.83
C16+18.0	2.42	30.26	16.92

Appendix B. Nutrient composition of piglet diets (g/kg)

	Creep	Post-weaning	Grower	Finisher
Net energy (MJ)	10.42	10.04	9.70	9.50
Ash	57.25	55.54	55.56	48.79
Crude protein	175.61	175.15	139.98	160.03
Crude fat	75.08	59.45	49.12	45.74
Crude fibre	29.11	39.75	44.26	58.09
ID starch + sugar	402.20	409.71	416.18	409.96
Starch	332.54	359.84	378.95	376.09
AID LYS	11.33	10.26	8.80	7.16
AID MET	4.32	3.54	3.03	2.38
AID M+C	6.67	5.595	5.40	4.68
AID THR	6.57	5.86	5.23	4.28
AID TRP	2.20	1.94	1.64	1.34
CA	7.96	7.01	7.01	4.84
Dig. P	3.87	3.57	2.85	1.88
NSP	152.33	172.10	179.20	213.70

Appendix C. Piglets from which subcutaneous and visceral adipose tissue was collected

Piglet ID	Sow ID	Diet	Sex	Size
426	997	control	male	medium
523	199	control	male	medium
992	125	control	male	medium
189	177	control	male	medium
324	42	control	male	medium
605	545	control	female	medium
969	14	control	female	medium
141	92	high fat	male	medium
560	612	high fat	male	medium
449	52	high fat	female	medium
655	557	high fat	male	medium
149	231	high fat	female	medium
445	3	high fat	male	medium
878	89	high fat	male	medium
254	482	high fat	female	medium
130	92	high fat	female	small
146	231	high fat	female	small
555	612	high fat	male	small

Appendix D. Piglets from which only visceral adipose tissue was collected

Piglet ID	Sow ID	Diet	Sex	Size
151	12	control	male	medium
516	199	control	male	small
188	177	control	male	small
150	12	control	male	small
318	42	control	male	small
993	125	control	male	small
654	557	high fat	female	small
448	52	high fat	female	small
444	3	high fat	male	small
880	89	high fat	male	small

Appendix D. Pig pyrosequencing assay design

Pig (*Sus scrofa*) leptin (Ob) gene

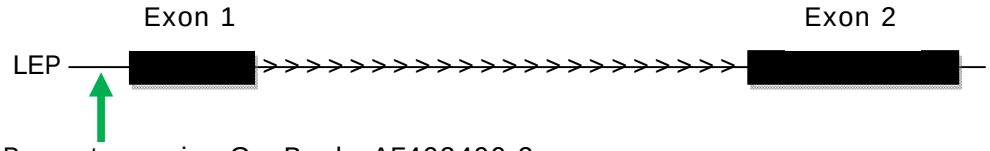

Figure D.1 Structure of the pig leptin gene. The pig leptin gene consists of two exons (black rectangles). The green arrow indicates the location of the core leptin promoter for which the pyrosequencing assay was designed.

PCR and Sequencing primer location

Bisulfite converted sense strand DNA starting from 5383bp (GenBank: AF492499.2). C/EBPα TF binding site is highlighted in green (CpG-5)

ttgaggaattttCGgttgttattttgtttCGgCGgtCGttttCGaggCGCGagaggtCGattCGgg
CGttatttCGaggatttttatttttgttagtCGttCGtagtatttttggCGgttattCGagggtgCG
CGCGgagtttttggCGtagCGgagtttttttttCGaggtttCGattagttttgttCGggtttaaCGtt
gtttCGttaggCGgggCGggagttggCGttCGgggtgggttagggtttgtCGggtagttgCGtaagtt
gCGttCGCGCGtttataagaggggagggta

Location of forward PCR primer

Location of reverse PCR primer

Location of sequencing primer 1 CpG 1-3: AAGTTGTTTAGTTAGGAGGGG

Location of sequencing primer 2 CpG 4-9: GGGTGGGTTAGGGTTTGT

Appendix E. Mouse pyrosequencing assay leptin (Ob) gene

Mus musculus leptin gene

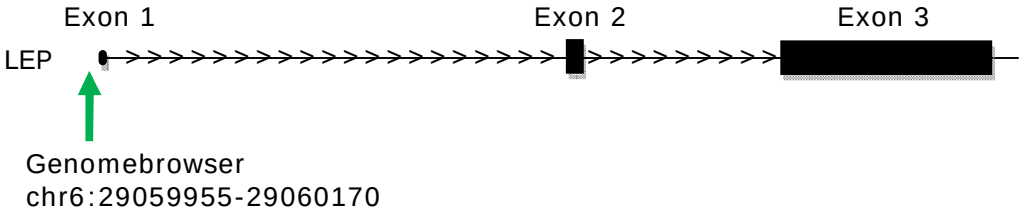

Figure E.1 Structure of the mouse leptin gene. The mouse leptin gene consists of three exons (black rectangles). The green arrow indicates the location of the core leptin promoter for which both the sense and antisense pyrosequencing assays were designed.

Sense strand: PCR and Sequencing primer location

Bisulfite converted sense strand DNA starting from chr6:29059955bp. C/EBPα TF binding site is highlighted in green (CpG-5).

TATTTGGATTTTTGGTTATTTTTAATTTTTAATAATTAGTTGAGGTTTTA
AGATAGAAAGTTGAATTGTTTTGTATTTGTTTTTTTTTATTAGGAATTTA
ATGATATTTTTGTTTTAATGAAAAAAAAAAAAAAAAAAAATAATAATCGA
AAATAATAATAATAATAAATTAAATAAATATTGTTTGATTAGGGATAGGT
TTTTGTTATTTGTTTTTGGGATTGGTTTAGAGTATTTTTAAATATATTTG
CGATTTTCGAAGTAGGTGTATTTTGTGATGTTATAGGATGAAATGAGACG
ATTGTTTTTCGGGTATTAAAGGAAGATAAGTTGTTTTGAGTTTGGGATTA
GTTTTTTTTTTGAGTAGTTAGGTTAGGTATGTAAAGAGTTGTCGGAAAAA
GTAGTTGGTAGAGTTTTGGTTTATTGGTTTTTTTGTTTTTAAGTTAGTTT
TTTGTAGTTTTTTGTTTTTTGCGGTGTTGGAAGTATTATTTTAAGGGATT
CGTTTTTAAATTATCGTTGTTTAGTAGTTGTTGGTCGGATTTCGAGGATT
ATCGGTTTATATTAAGCGTTTTTAAATTTGTATTCGAGGGCGCGGTTGAA
GTTTTTTTTCGAGGCGTTTAGAATGGAGTATTAGGTTGTTGTTGTTATTG
TTGTTGGTTCGTTGGGTGGGGCGGGAGTTGGCGTTCGTAGGGATTGGGGT
TGGTCGGATAGTTGCGTAAGTGGTATTGGGGTAGTTATAAGAGGGGTAGG
TAGGTATGGAGTTTCGGAGGGATTTTTGTTTTAGTAGTTGTAAGGTAAGG
TTCGGGGCGCGTTATTTTTTTTTTTATTAGTTTTTTTAATAGTATTTTAT
TTAGTTTTGGAAATTAGAGAAATTGAGGTAAGAAGGAGGTTATGTGGATA
GTTTGGTGTTGAATTTAGTAGTTTTGTAGCGAGGGATTTTGTAGATAGAG
GGAAGGTTGTTGGGAGTTGAAAGGTATTAAATTTTTGGTTGGGTCGGTTT
GTGTTTATAGGGTATTAATTTGTTTTTGTAAAGTTAGTTAGGGTTTTAAG
TTTGTTTTTGGTTTGTAGATTTGAGATGGTTAGGTTTTTTATAGCGTTTA
GATGGAAAAAGTAAGGTTGGACGTGTAGTGAGGTAGTTATAGTTTTTCGT
TTTTTTGAAGGTTAGATATAATTTTGTTTTGTTGTTGAGAGTTTATTTTT
TAGTGAGGGTTTATTT

Forward PCR primer: AGAATGGAGTATTAGGTTGTTGTTGTTA

Sequencing primer 1 (CpG 1-5): GTTGGTTCGTTGGGTG
Sequencing primer 2 (CpG 6-9): GGTAGGTATGGAGTTT

Reverse PCR primer: CATAACCTCCTTCTTACCTCAATTTCTC

Antisense strand: PCR and Sequencing primer location

Bisulfite converted DNA antisense strand DNA from chr6: 29060170bp. C/EBPα TF binding site is highlighted in green (CpG-5).

AGGTAGGTTTTTATTGGAAAATGAATTTTTAATAGTAAAGTAAGGTTGTATTTGGTTTTTAGGAAGGC
GGAAAGTTGTGGTTGTTTTATTATACGTTTAGTTTTGTTTTTTTTATTTGAACGTTGTGGGAAATTTA
ATTATTTTAAGTTTGTAGATTAAAGATAGGTTTGGAGTTTTGGTTGGTTTTGTAAGAATAGGTTGGTG
TTTTGTGGGTATAGGTCGGTTTAGTTAGGAATTTAATGTTTTTTAGTTTTTAATAGTTTTTTTTTTGT
TTGTAGAGTTTTTCGTTGTAAAATTATTGAATTTAATATTAAGTTGTTTATATGATTTTTTTTTTGTT
TTAGTTTTTTTAATTTTTAGAGTTGGATGGGGTGTTATTAGAAAGATTGGTGGAGGAGAAAGTAGCGC
GTTTCGGGTTTTATTTTGTAGTTGTTGGAGTAGGGATTTTTTCGGGGTTTTATGTTTGTTTGTTTTTT
TTATAATTGTTTTAGTGTTATTTGCGTAATTGTTCGGTTAGTTTTAGTTTTTGCGAGCGTTAATTTTC
GTTTTATTTAGCGGGTTAGTAATAGTGGTAGTAGTAATTTAGTGTTTTATTTTAGGCGTTTCGAGGGA
GAATTTTAGTCGCGTTTTCGAGTGTAAGTTTGGGGGCGTTTGGTATGAGTCGGTAATTTTCGAGGTTC
GGTTAGTAGTTATTGAGTAGCGGTAGTTTAAGGACGGGTTTTTTGGGATGGTGTTTTTAGTATCGTAG
GGAGTAAGAGGTTATAGAAGGTTGGTTTGGGGATAGGGAGATTAGTGAGTTAGGATTTTGTTAGTTGT
TTTTTTCGATAGTTTTTTGTATATTTAATTTGGTTGTTTAGAGGAGAAATTGGTTTTAAGTTTAGGGT
AATTTGTTTTTTTTTGGTATTCGAAGAATAGTCGTTTTATTTTATTTTGTGATATTATAGAATGTATT
TGTTTCGAAGGTCGTAAGTGTGTTTAGGGATGTTTTGGATTAGTTTTAGGGGTAGGTGGTAAGGGTTT
GTTTTTGATTAAATAATGTTTGTTTGGTTTGTTGTTGTTGTTTTCGGTTGTTGTTTTTTTTTTTTT
TTTTTTTTTATTAAGGTAAGAATATTATTGGGTTTTTAATGGGAGGAAATAGATATAGAATAGTTTAG
TTTTTTGTTTTAAAATTTTAGTTAGTTGTTAAGGATTAGAGATAGTTAGGGATTTAGGTG

Forward PCR primer: TGGATGGGGTGTTATTAGAAAGA

Reverse PCR primer: AGTGTAAGTTTGGGGGGGTTTGGTAT

Sequencing primer 1 (CpG 9-6): ATTGGTGGAGGAGAAA

Sequencing primer 2 (CpG 5-1): TTTATAATTGTTTTAGTGTTATTTG

Appendix F Individual western blots for validation of mESCs

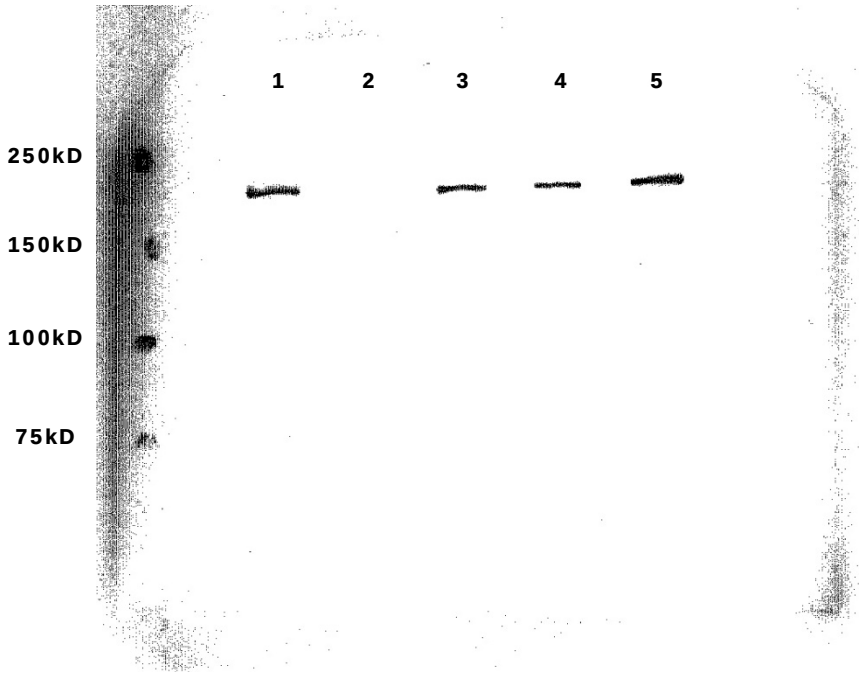

Figure F.1 DNMT1 protein expression in mutant mESCs. Western blots were performed to validate the presence of DNMT1 protein in WT (1), DNMT3a$^{(-/-)}$ (3), DNMT3b$^{(-/-)}$ (4) and DKO (5) cells. DNMT1 protein was not detected in DNMT1^{s} (2) cells.

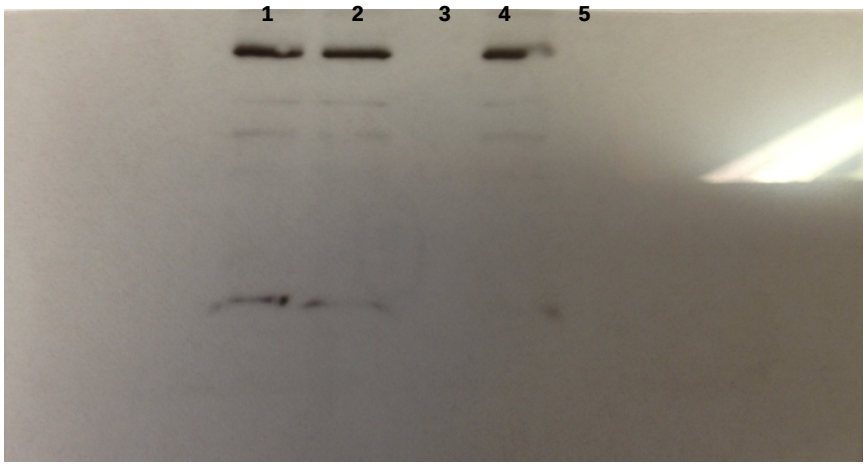

Figure F.2 DNMT3a protein expression in mutant mESCs. Western blots were performed to validate the presence of DNMT3a protein in WT (1), DNMT1^{s} (2) and DNMT3b$^{(-/-)}$ (4) cells. DNMT3a protein was not detected in DNMT3a$^{(-/-)}$ (3) and DKO (5) cells.

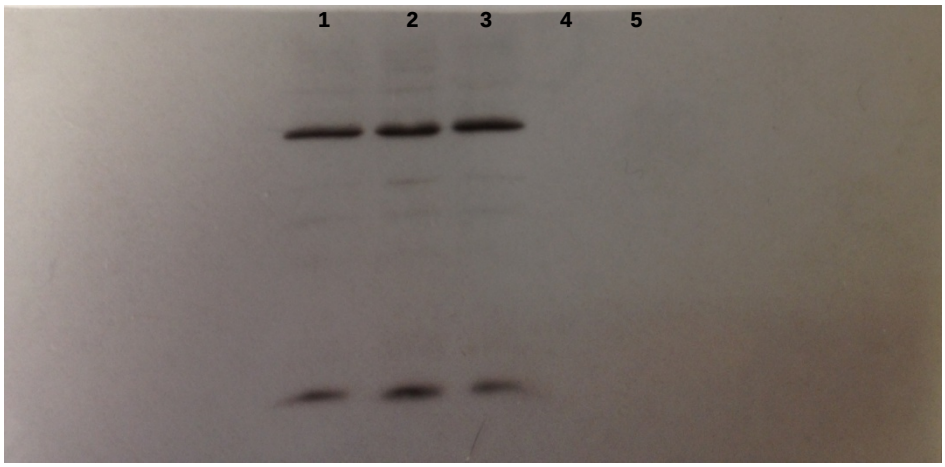

Figure F.3 DNMT3b protein expression in mutant mESCs. Western blots were performed to validate the presence of DNMT3b protein in WT (1), DNMT1^{s} (2) and DNMT3a$^{(-/-)}$ (3) cells. DNMT3b protein was not detected in DNMT3b$^{(-/-)}$ (4) and DKO (5) cells.

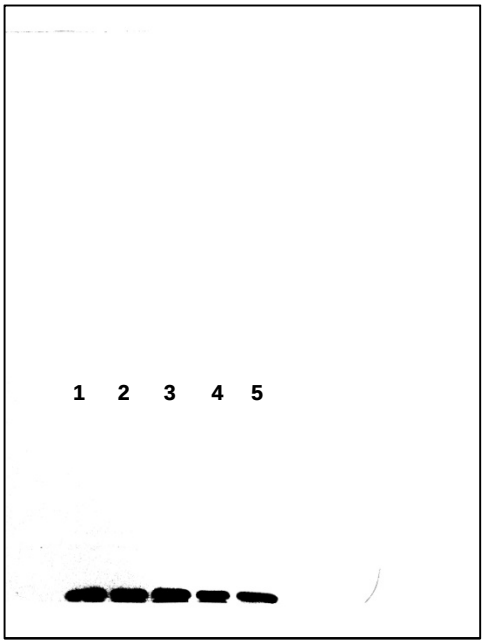

Figure F.4 Histone H3 protein expression in mutant mESCs. Western blots were performed to validate the presence of H3 protein in WT (1), DNMT1^{s} (2) and DNMT3a$^{(-/-)}$ (3) DNMT3b$^{(-/-)}$ (4) and DKO (5) cells.

Appendix G. Amplification of cDNA templates with primers specific to gDNA.

With the exception of C/EBPα all rt-PCR primers were designed to span intron exon boundaries and therefore would be unable to amplify gDNA. In the unlikely scenario that they could amplify gDNA, this would produce an amplicon of a different size however, DNAse treatment was also performed to remove residual gDNA from samples. The only exception is the primer set designed to amplify C/EBPα. This gene consists of one exon and therefore intron-exon boundaries could not be spanned and any residual gDNA would produce an amplicon of the same size as cDNA. To determine that all gDNA was removed from templates, templates were amplified with primers specific to gDNA (Figure G.1).

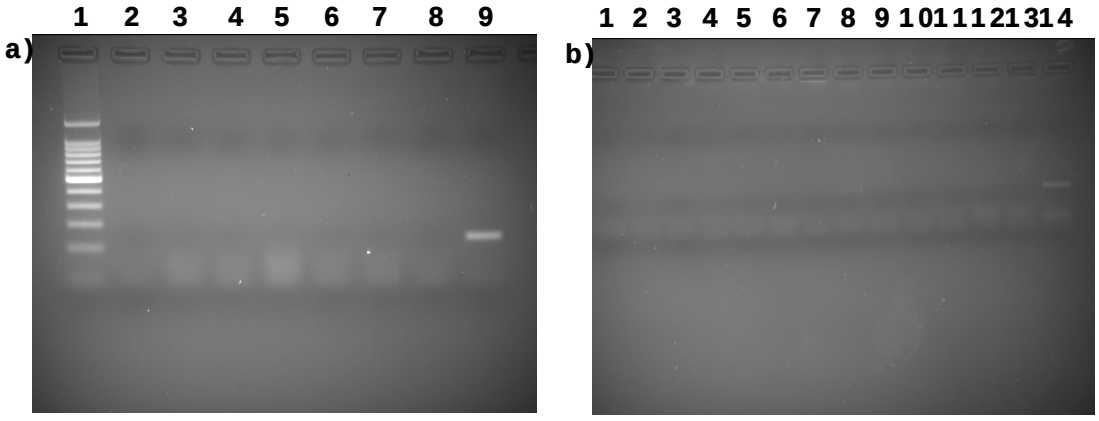

Figure G.1 Amplification of cDNA templates with primers specific to gDNA. Templates were amplified with 5-hmC detection primers described in section 3.2.9 to determine that DNAse treatment was effective in removing residual gDNA. a) Mutant mESC templates Lane 1: 100bp ladder, 2: No template control, 3: WT, 4: DNMT1^{s}, 5: DNMT3a$^{(-/-)}$, 6: DNMT3b$^{(-/-)}$, 7: DKO, 8: Control sample which did not undergo DNAse treatment. b) 3T3-L1 templates Lane 1: No template control, 2: Undiff 0hrs, 3: diff 0hrs, 4: undiff 20hrs, 5: diff 20hrs, 6: undiff 26hrs, 7: diff 26hrs, 8: undiff 32hrs, 8: diff 32 hrs, 9: undiff 38hrs, 10: diff 38hrs, 11: undiff 96 hrs, 12: diff 96 hrs, 13: undiff 192 hrs, 14: diff 192 hrs.

Appendix H. ChIP-seq and RNA-seq data from ENCODE project

Figure H1. mESC ChIP-seq and RNA-seq data. Data was mined from the ENCODE project for various mESC lines and analysed for RNA transcripts and histone modifications specifically at the leptin promoter.

Appendix I. Optimisation of oxidative pyrosequencing protocol

In the first year of the production of this thesis, 5-hmC was re-discovered in mammalian DNA which changed the course of the project. Unfortunately there were few techniques for the locus specific detection of 5-hmC, none of which allowed base specific resolution of this cytosine modification and so the depth of the information that could be gathered was restricted. Close to the deadline for the end of laboratory work for this thesis, two papers were published which described techniques for the base specific resolution of 5-hmC (Booth *et al*, 2012; Yu *et al*, 2012). Booth *et al* described the selective oxidation of 5-hmC to 5-fC using the chemical potassium perruthenate which, coupled to bisulfite sequencing, was used to quantify levels of 5-hmC and also determine the exact CpGs in the sequence of interest that were hydroxymethylated. The work presented in chapter four provided evidence for the association of leptin promoter hydroxymethylation and leptin expression in differentiated of 3T3-L1 cells. In an attempt to determine which specific CpGs of the core leptin promoter became hydroxymethylated and the extent to which they became hydroxymethylated the potassium perruthenate oxidation was attempted and coupled to pyrosequencing to quantify 5-hmC levels. At first glance, it appeared that this would be a simple incubation of DNA with potassium perruthenate followed by pyrosequencing (Section 4.2.5.2), however, potassium perruthenate is a difficult chemical. It is only slightly soluble in water and degrades above 40°C (information obtained from Sigma).

In the first attempt to optimise the protocol 2µg gDNA was incubated with 15mM potassium perruthenate in 0.05M NaOH on ice over one hour. This was followed by column cleaning the oxidised DNA using the Zymo clean and concentrator columns. However, undissolved potassium perruthenate blocked the columns that were used to clean up the DNA in preparation for bisulfite conversion.

In the second attempt the amount of undissolved potassium perruthenate was reduced by centrifugation prior to column purification and bigger Qiagen PCR purification columns were used. The oxidised, purified DNA was subject to PCR with the primers designed to detect 5-hmC (Section 3.2.9) to determine that DNA had survived the purification process. Pyrosequencing was then attempted (Section 4.2.5.2) however a PCR band was not produced for the oxidised DNA. Therefore the quantification of DNA levels could not be attempted.

The protocol was repeated, this time eluting template DNA for pyrosequencing into 10µl to allow more DNA into the PCR reaction. A band was produced and

pyrosequenced however the run was failed because due to low peak height (Figure 5.1).

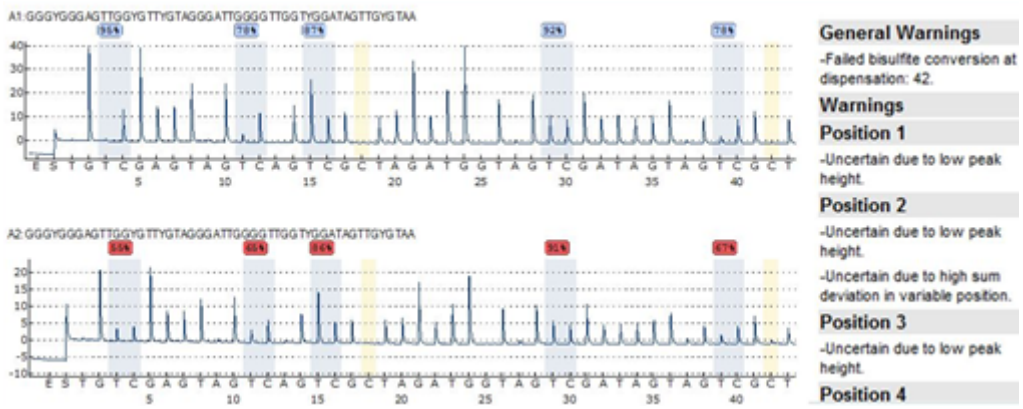

Figure I.1 Pyrosequencing reactions for the detection of 5-hmC at single base resolution. Whilst the unoxidised DNA passed the pyrosequencing run (above), KRuO4 oxidised DNA failed (below).

A high cycle PCR was already in use so rather than increasing cycle number to increase the concentration of the PCR product, 10µg of DNA rather than 2µg was oxidised and bisulfite converted so that more DNA could be inserted into the PCR reaction. The PCR reactions were not consistent, sometimes they were successful and sometimes they were not. The PCR reactions also produced primer-dimer bands which affected pyrosequencing reactions and so all products were purified from the gel prior to pyrosequencing. This led to an improvement in pyrosequencing reactions, however runs were still not completely successful as the software flagged them as uncertain (Figure 5.2)

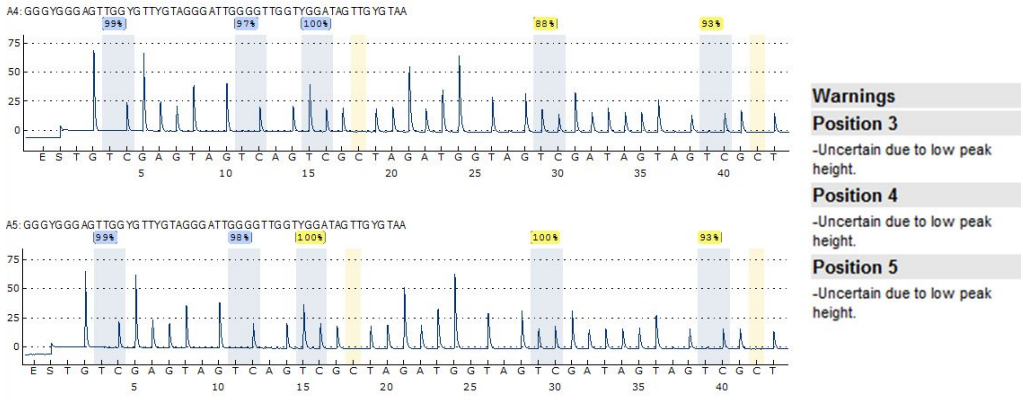

Figure I.2 Pyrosequencing reactions for the detection of 5-hmC at single base resolution. Whilst the unoxidised DNA passed the pyrosequencing run at most CpGs (above), KRuO4 oxidised DNA runs were uncertain at some CpG sites.

Due to the technical difficulties and time restraints, it was decided to cease attempting this experiment for this thesis and present the results obtained as preliminary data.

References

Ahima RS, Prabakaran D, Flier JS (1998) Postnatal leptin surge and regulation of circadian rhythm of leptin by feeding. Implications for energy homeostasis and neuroendocrine function. *J Clin Invest* **101:** 1020-1027

Alegría-Torres JA, Baccarelli A, Bollati V (2011) Epigenetics and lifestyle. *Epigenomics* **3:** 267-277

Aloia L, Di Stefano B, Di Croce L (2013) Polycomb complexes in stem cells and embryonic development. *Development* **140:** 2525-2534

Antequera F, Bird A (1993) CpG islands. *EXS* **64:** 169-185

Arand J, Spieler D, Karius T, Branco MR, Meilinger D, Meissner A, Jenuwein T, Xu G, Leonhardt H, Wolf V, Walter J (2012) In vivo control of CpG and non-CpG DNA methylation by DNA methyltransferases. *PLoS Genet* **8:** e1002750

Armstrong CA, Jones GD, Anderson R, Iyer P, Narayanan D, Sandhu J, Singh R, Talbot CJ, Tufarelli C (2012) DNMTs are required for delayed genome instability caused by radiation. *Epigenetics* **7:** 892-902

Ashworth C, Hoggard N, Thomas L, Mercer J, Wallace J, Lea R (2000) Placental leptin. *Rev Reprod* **5:** 18-24

Astiz S, Gonzalez-Bulnes A, Perez-Solana M, Sanchez-Sanchez R, Torres-Rovira L (2013) In vitro Release of Ovarian Progesterone is Decreased During the Oestrous Cycle and Pregnancy of Swine with Obesity/Leptin Resistance. *Reprod Domest Anim* **48:** 44-48

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

Bahar B, O'Doherty JV, O'Doherty AM, Sweeney T (2013) Chito-Oligosaccharide Inhibits the De-Methylation of a 'CpG' Island within the Leptin (LEP) Promoter during Adipogenesis of 3T3-L1 Cells. *PLoS One* **8:** e60011

Barb CR, Hausman GJ, Houseknecht KL (2001) Biology of leptin in the pig. *Domest Anim Endocrinol* **21:** 297-317

Barlow DP (2011) Genomic imprinting: a mammalian epigenetic discovery model. *Annu Rev Genet* **45:** 379-403

Berger SL, Kouzarides T, Shiekhattar R, Shilatifard A (2009) An operational definition of epigenetics. *Genes Dev* **23:** 781-783

Bestor TH (2000) The DNA methyltransferases of mammals. *Hum Mol Genet* **9:** 2395-2402

Bestor TH, Ingram VM (1983) Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. *Proc Natl Acad Sci U S A* **80:** 5559-5563

Bi S, Gavrilova O, Gong D, Mason M, Reitman M (1997) Identification of a placenta enhancer for the human leptin gene. *Journal of Biological Chemistry* **272:** 30583-30588

Bird AP (1978a) The occurrence and transmission of a pattern of DNA methylation in Xenopus laevis ribosomal DNA. *Philos Trans R Soc Lond B Biol Sci* **283:** 325-327

Bird AP (1978b) Use of restriction enzymes to study eukaryotic DNA methylation: II. The symmetry of methylated sites supports semi-conservative copying of the methylation pattern. *J Mol Biol* **118:** 49-60

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

Bird AP (1995) Gene number, noise reduction and biological complexity. *Trends Genet* **11:** 94-100

Bodosi B, Gardi J, Hajdu I, Szentirmai E, Obal F, Krueger JM (2004) Rhythms of ghrelin, leptin, and sleep in rats: effects of the normal diurnal cycle, restricted feeding, and sleep deprivation. *Am J Physiol Regul Integr Comp Physiol* **287:** R1071-1079

Booth MJ, Branco MR, Ficz G, Oxley D, Krueger F, Reik W, Balasubramanian S (2012) Quantitative sequencing of 5-methylcytosine and 5-hydroxymethylcytosine at single-base resolution. *Science* **336:** 934-937

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

Bouloumié A, Drexler H, Lafontan M, Busse R (1998) Leptin, the product of Ob gene, promotes angiogenesis. *Circ Res* **83:** 1059-1066

Bourc'his D, Xu GL, Lin CS, Bollman B, Bestor TH (2001) Dnmt3L and the establishment of maternal genomic imprints. *Science* **294:** 2536-2539

Branco MR, Ficz G, Reik W (2012) Uncovering the role of 5-hydroxymethylcytosine in the epigenome. *Nat Rev Genet* **13:** 7-13

Brzechffa P, Jakimiuk A, Agarwal S, Weitsman S, Buyalos R, Magoffin D (1996) Serum immunoreactive leptin concentrations in women with polycystic ovary syndrome. *J Clin Endocrinol Metab* **81:** 4166-4169

Carbone F, La Rocca C, Matarese G (2012) Immunological functions of leptin and adiponectin. *Biochimie* **94:** 2082-2088

Cardoso MC, Sporbert A, Leonhardt H (1999) Structure and function in the nucleus: subnuclear trafficking of DNA replication factors. *J Cell Biochem* **Suppl 32-33:** 15-23

Chehab F, Lim M, Lu R (1996) Correction of the sterility defect in homozygous obese female mice by treatment with the human recombinant leptin. *Nat Genet* **12:** 318-320

Chen CC, Wang KY, Shen CK (2012) The Mammalian de Novo DNA Methyltransferases DNMT3A and DNMT3B Are Also DNA 5-Hydroxymethylcytosine Dehydroxymethylases. *J Biol Chem* **287:** 33116-33121

Chen T, Ueda Y, Dodge JE, Wang Z, Li E (2003) Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Mol Cell Biol* **23:** 5594-5605

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

Chuang LS, Ian HI, Koh TW, Ng HH, Xu G, Li BF (1997) Human DNA-(cytosine-5) methyltransferase-PCNA complex as a target for p21WAF1. *Science* **277:** 1996-2000

Cinti S (2005) The adipose organ. *Prostaglandins Leukot Essent Fatty Acids* **73:** 9-15

Cortellino S, Xu J, Sannai M, Moore R, Caretti E, Cigliano A, Le Coz M, Devarajan K, Wessels A, Soprano D, Abramowitz LK, Bartolomei MS, Rambow F, Bassi MR, Bruno T, Fanciulli M, Renner C, Klein-Szanto AJ, Matsumoto Y, Kobi D, Davidson I, Alberti C, Larue L, Bellacosa A (2011) Thymine DNA glycosylase is essential for active DNA demethylation by linked deamination-base excision repair. *Cell* **146:** 67-79

Cristancho AG, Lazar MA (2011) Forming functional fat: a growing understanding of adipocyte differentiation. *Nat Rev Mol Cell Biol* **12:** 722-734

Dawlaty MM, Ganz K, Powell BE, Hu YC, Markoulaki S, Cheng AW, Gao Q, Kim J, Choi SW, Page DC, Jaenisch R (2011) Tet1 is dispensable for maintaining pluripotency and its loss is compatible with embryonic and postnatal development. *Cell Stem Cell* **9:** 166-175

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

Dietz WH, Masterson K, Sittig LJ, Redei EE, Herzing LB (2012) Imprinting and expression of Dio3os mirrors Dio3 in rat. *Front Genet* **3:** 279

Doyon C, Drouin G, Trudeau VL, Moon TW (2001) Molecular evolution of leptin. *Gen Comp Endocrinol* **124:** 188-198

Drapeau MD, Albert S, Kucharski R, Prusko C, Maleszka R (2006) Evolution of the Yellow/Major Royal Jelly Protein family and the emergence of social behavior in honey bees. *Genome Res* **16:** 1385-1394

Ehrlich M (2003) The ICF syndrome, a DNA methyltransferase 3B deficiency and immunodeficiency disease. *Clin Immunol* **109:** 17-28

Ehrlich M, Gama-Sosa MA, Huang LH, Midgett RM, Kuo KC, McCune RA, Gehrke C (1982) Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res* **10:** 2709-2721

Enquobahrie D, Meller M, Rice K, Psaty B, Siscovick D, Williams M (2008) Differential placental gene expression in preeclampsia. *Am J Obstet Gynecol* **199:** 566.e561-511

Fan C, Liu X, Shen W, Deckelbaum RJ, Qi K (2011) The Regulation of Leptin, Leptin Receptor and Pro-opiomelanocortin Expression by N-3 PUFAs in Diet-Induced Obese Mice Is Not Related to the Methylation of Their Promoters. *Nutr Metab (Lond)* **8:** 31

Fantuzzi G, Faggioni R (2000) Leptin in the regulation of immunity, inflammation, and hematopoiesis. *J Leukoc Biol* **68:** 437-446

Faulk C, Dolinoy DC (2011) Timing is everything: the when and how of environmentally induced changes in the epigenome of animals. *Epigenetics* **6:** 791-797

Feil R, Fraga MF (2011) Epigenetics and the environment: emerging patterns and implications. *Nat Rev Genet* **13:** 97-109

Ficz G, Branco MR, Seisenberger S, Santos F, Krueger F, Hore TA, Marques CJ, Andrews S, Reik W (2011) Dynamic regulation of 5-hydroxymethylcytosine in mouse ES cells and during differentiation. *Nature* **473:** 398-402

Flier JS, Maratos-Flier E (2010) Lasker lauds leptin. *Cell* **143:** 9-12

Fournier A, Sasai N, Nakao M, Defossez PA (2012) The role of methyl-binding proteins in chromatin organization and epigenome maintenance. *Brief Funct Genomics* **11:** 251-264

Frauer C, Hoffmann T, Bultmann S, Casa V, Cardoso MC, Antes I, Leonhardt H (2011) Recognition of 5-hydroxymethylcytosine by the Uhrf1 SRA domain. *PLoS One* **6:** e21306

Fu AQ, Genereux DP, Stöger R, Burden AF, Laird CD, Stephens M (2012) Statistical inference of in vivo properties of human DNA methyltransferases from double-stranded methylation patterns. *PLoS One* **7:** e32225

Fuks F (2005) DNA methylation and histone modifications: teaming up to silence genes. *Curr Opin Genet Dev* **15:** 490-495

Gabor Miklos GL, Maleszka R (2011) Epigenomic communication systems in humans and honey bees: from molecules to behavior. *Horm Behav* **59:** 399-406

Genereux DP, Johnson WC, Burden AF, Stöger R, Laird CD (2008) Errors in the bisulfite conversion of DNA: modulating inappropriate- and failed-conversion frequencies. *Nucleic Acids Res* **36:** e150

Genstat (2011) *Genstat Version 14.1.0.5943 Reference Manual*, Oxford, UK: Clarendon Press, Oxford Science Publications.

Globisch D, Münzel M, Müller M, Michalakis S, Wagner M, Koch S, Brückl T, Biel M, Carell T (2010) Tissue distribution of 5-hydroxymethylcytosine and search for active demethylation intermediates. *PLoS One* **5:** e15367

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. *J Mol Biol* **309:** 1201-1208

Gowher H, Jeltsch A (2002) Molecular enzymology of the catalytic domains of the Dnmt3a and Dnmt3b DNA methyltransferases. *J Biol Chem* **277:** 20409-20414

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

Goyal R, Reinhardt R, Jeltsch A (2006) Accuracy of DNA methylation pattern preservation by the Dnmt1 methyltransferase. *Nucleic Acids Res* **34:** 1182-1188

Guo JU, Su Y, Zhong C, Ming GL, Song H (2011) Emerging roles of TET proteins and 5-hydroxymethylcytosines in active DNA demethylation and beyond. *Cell Cycle* **10:** 2662-2668

Hashimoto H, Liu Y, Upadhyay AK, Chang Y, Howerton SB, Vertino PM, Zhang X, Cheng X (2012) Recognition and potential mechanisms for replication and erasure of cytosine hydroxymethylation. *Nucleic Acids Res* **40:** 4841-4849

Hausman GJ, Barb CR, Lents CA (2012) Leptin and reproductive function. *Biochimie* **94:** 2075-2081

He YF, Li BZ, Li Z, Liu P, Wang Y, Tang Q, Ding J, Jia Y, Chen Z, Li L, Sun Y, Li X, Dai Q, Song CX, Zhang K, He C, Xu GL (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science* **333:** 1303-1307

Hermann A, Goyal R, Jeltsch A (2004) The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* **279:** 48350-48359

Hervey GR (1959) The effects of lesions in the hypothalamus in parabiotic rats. *J Physiol* **145:** 336-352

Heyn H, Vidal E, Sayols S, Sanchez-Mut JV, Moran S, Medina I, Sandoval J, Simó-Riudalbas L, Szczesna K, Huertas D, Gatto S, Matarazzo MR, Dopazo J, Esteller M (2012) Whole-genome bisulfite DNA sequencing of a DNMT3B mutant patient. *Epigenetics* **7:** 542-550

Holness M, Munns M, Sugden M (1999) Current concepts concerning the role of leptin in reproductive function. *Mol Cell Endocrinol* **157:** 11-20

Holz-Schietinger C, Reich NO (2010) The inherent processivity of the human de novo methyltransferase 3A (DNMT3A) is enhanced by DNMT3L. *J Biol Chem* **285:** 29091-29100

Hu YG, Hirasawa R, Hu JL, Hata K, Li CL, Jin Y, Chen T, Li E, Rigolet M, Viegas-Péquignot E, Sasaki H, Xu GL (2008) Regulation of DNA methylation activity through Dnmt3L promoter methylation by Dnmt3 enzymes in embryonic development. *Hum Mol Genet* **17:** 2654-2664

Huang Y, Pastor WA, Shen Y, Tahiliani M, Liu DR, Rao A (2010) The behaviour of 5-hydroxymethylcytosine in bisulfite sequencing. *PLoS One* **5:** e8888

Hube F, Lietz U, Igel M, Jensen PB, Tornqvist H, Joost HG, Hauner H (1996) Difference in leptin mRNA levels between omental and subcutaneous abdominal adipose tissue from obese humans. *Horm Metab Res* **28:** 690-693

Ibrahim MM (2010) Subcutaneous and visceral adipose tissue: structural and functional differences. *Obes Rev* **11:** 11-18

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 Genet* **6:** e1001134

Ito S, D'Alessio AC, Taranova OV, Hong K, Sowers LC, Zhang Y (2010) Role of Tet proteins in 5mC to 5hmC conversion, ES-cell self-renewal and inner cell mass specification. *Nature* **466:** 1129-1133

Ito S, Shen L, Dai Q, Wu SC, Collins LB, Swenberg JA, He C, Zhang Y (2011) Tet Proteins Can Convert 5-Methylcytosine to 5-Formylcytosine and 5-Carboxylcytosine. *Science* **6047**

Iyer MN, Sarmah BC, Tamuli MK, Das A, Kalita D (2012) Effect of dietary sunflower oil and coconut oil on adipose tissue gene expression, fatty acid composition and serum lipid profile of grower pigs. *Arch Anim Nutr* **66:** 271-282

Jackson DA, Pombo A (1998) Replicon clusters are stable units of chromosome structure: evidence that nuclear organization contributes to the efficient activation and propagation of S phase in human cells. *J Cell Biol* **140:** 1285-1295

Jackson M, Krassowska A, Gilbert N, Chevassut T, Forrester L, Ansell J, Ramsahoye B (2004) Severe global DNA hypomethylation blocks differentiation and induces histone hyperacetylation in embryonic stem cells. *Mol Cell Biol* **24:** 8862-8871

Jeong S, Liang G, Sharma S, Lin JC, Choi SH, Han H, Yoo CB, Egger G, Yang AS, Jones PA (2009) Selective anchoring of DNA methyltransferases 3A and 3B to nucleosomes containing methylated DNA. *Mol Cell Biol* **29:** 5366-5376

Jin B, Tao Q, Peng J, Soo HM, Wu W, Ying J, Fields CR, Delmas AL, Liu X, Qiu J, Robertson KD (2008) DNA methyltransferase 3B (DNMT3B) mutations in ICF syndrome lead to altered epigenetic modifications and aberrant expression of genes regulating development, neurogenesis and immune function. *Hum Mol Genet* **17:** 690-709

Jin SG, Kadam S, Pfeifer GP (2010) Examination of the specificity of DNA methylation profiling techniques towards 5-methylcytosine and 5-hydroxymethylcytosine. *Nucleic Acids Res* **38:** e125

Jones PA, Liang G (2009) Rethinking how DNA methylation patterns are maintained. *Nat Rev Genet* **10:** 805-811

Jousse C, Parry L, Lambert-Langlais S, Maurin AC, Averous J, Bruhat A, Carraro V, Tost J, Letteron P, Chen P, Jockers R, Launay JM, Mallet J, Fafournoux P (2011) Perinatal undernutrition affects the methylation and expression of the leptin gene in adults: implication for the understanding of metabolic syndrome. *FASEB J* **25:** 3271-3278

Jähner D, Stuhlmann H, Stewart CL, Harbers K, Löhler J, Simon I, Jaenisch R (1982) De novo methylation and expression of retroviral genomes during mouse embryogenesis. *Nature* **298:** 623-628

Karastergiou K, Smith SR, Greenberg AS, Fried SK (2012) Sex differences in human adipose tissues - the biology of pear shape. *Biol Sex Differ* **3:** 13

Kawamura K, Sato N, Fukuda J, Kodama H, Kumagai J, Tanikawa H, Nakamura A, Tanaka T (2002) Leptin promotes the development of mouse preimplantation embryos in vitro. *Endocrinology* **143:** 1922-1931

Kennedy A, Gettys TW, Watson P, Wallace P, Ganaway E, Pan Q, Garvey WT (1997) The metabolic significance of leptin in humans: gender-based differences in relationship to adiposity, insulin sensitivity, and energy expenditure. *J Clin Endocrinol Metab* **82:** 1293-1300

Koh KP, Yabuuchi A, Rao S, Huang Y, Cunniff K, Nardone J, Laiho A, Tahiliani M, Sommer CA, Mostoslavsky G, Lahesmaa R, Orkin SH, Rodig SJ, Daley GQ, Rao A (2011) Tet1 and Tet2 regulate 5-hydroxymethylcytosine production and cell lineage specification in mouse embryonic stem cells. *Cell Stem Cell* **8:** 200-213

Kooistra SM, Helin K (2012) Molecular mechanisms and potential functions of histone demethylases. *Nat Rev Mol Cell Biol* **13:** 297-311

Kriaucionis S, Heintz N (2009) The nuclear DNA base 5-hydroxymethylcytosine is present in Purkinje neurons and the brain. *Science* **324:** 929-930

Kukla M, Mazur W, Bułdak RJ, Zwirska-Korczala K (2011) Potential role of leptin, adiponectin and three novel adipokines--visfatin, chemerin and vaspin--in chronic hepatitis. *Mol Med* **17:** 1397-1410

Laird CD, Pleasant ND, Clark AD, Sneeden JL, Hassan KM, Manley NC, Vary JC, Morgan T, Hansen RS, Stöger R (2004) Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. *Proc Natl Acad Sci U S A* **101:** 204-209

Lee JT (2012) Epigenetic regulation by long noncoding RNAs. *Science* **338:** 1435-1439

Lei H, Oh SP, Okano M, Jüttermann R, Goss KA, Jaenisch R, Li E (1996) De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* **122:** 3195-3205

Li CC, Cropley JE, Cowley MJ, Preiss T, Martin DI, Suter CM (2011) A sustained dietary change increases epigenetic variation in isogenic mice. *PLoS Genet* **7:** e1001380

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

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. *Mol Cell Biol* **22:** 480-491

Lillie R, Ashburn L. (1943) Supersaturated solutions of fat stains in dilute isopropanol for demonstration of acute fatty degeneration not shown by Herxheimer's technique., Arch Pathol, Vol. 36, pp. 432-440.

Liutkeviciute Z, Lukinavicius G, Masevicius V, Daujotyte D, Klimasauskas S (2009) Cytosine-5-methyltransferases add aldehydes to DNA. *Nat Chem Biol* **5:** 400-402

Lönnqvist F, Arner P, Nordfors L, Schalling M (1995) Overexpression of the obese (ob) gene in adipose tissue of human obese subjects. *Nat Med* **1:** 950-953

MacDougald O, Hwang C, Fan H (1995) Regulated expression of the obese gene product (leptin) in white adipose tissue and 3T3-L1 adipocytes. *Proceedings of the National Academy of Sciences of the United States of America* **92:** 9034-9037

Maffei M, Halaas J, Ravussin E, Pratley RE, Lee GH, Zhang Y, Fei H, Kim S, Lallone R, Ranganathan S (1995) Leptin levels in human and rodent: measurement of plasma leptin and ob RNA in obese and weight-reduced subjects. *Nat Med* **1:** 1155-1161

Mantzoros C (2000) Role of leptin in reproduction. *Ann N Y Acad Sci* **900:** 174-183

Mantzoros C, Flier J, Lesem M, Brewerton T, Jimerson D (1997) Cerebrospinal fluid leptin in anorexia nervosa: correlation with nutritional status and potential role in resistance to weight gain. *J Clin Endocrinol Metab* **82:** 1845-1851

Marchini G, Fried G, Ostlund E, Hagenäs L (1998) Plasma leptin in infants: relations to birth weight and weight loss. *Pediatrics* **101:** 429-432

Margetic S, Gazzola C, Pegg G, Hill R (2002) Leptin: A Review of its peripheral actions and interactions. *International Journal of Obesity* **26:** 1407-1433

Margueron R, Justin N, Ohno K, Sharpe ML, Son J, Drury WJ, Voigt P, Martin SR, Taylor WR, De Marco V, Pirrotta V, Reinberg D, Gamblin SJ (2009) Role of the polycomb protein EED in the propagation of repressive histone marks. *Nature* **461:** 762-767

Mason K, Liu Z, Aguirre-Lavin T, Beaujean N (2012) Chromatin and epigenetic modifications during early mammalian development. *Anim Reprod Sci* **134:** 45-55

Mason MM, He Y, Chen H, Quon MJ, Reitman M (1998) Regulation of leptin promoter function by Sp1, C/EBP, and a novel factor. *Endocrinology* **139:** 1013-1022

Masuzaki H, Hosoda K, Ogawa Y, Shigemoto M, Satoh N, Mori K, Tamura N, Nishi S, Yoshimasa Y, Yamori Y, Nakao K (1996) Augmented expression of obese (ob) gene during the process of obesity in genetically obese-hyperglycemic Wistar fatty (fa/fa) rats. *FEBS Lett* **378:** 267-271

Mayer W, Niveleau A, Walter J, Fundele R, Haaf T (2000) Demethylation of the zygotic paternal genome. *Nature* **403:** 501-502

Mazzio EA, Soliman KF (2012) Basic concepts of epigenetics: impact of environmental signals on gene expression. *Epigenetics* **7:** 119-130

Mellén M, Ayata P, Dewell S, Kriaucionis S, Heintz N (2012) MeCP2 Binds to 5hmC Enriched within Active Genes and Accessible Chromatin in the Nervous System. *Cell* **151:** 1417-1430

Melzner I, Scott V, Dorsch K, Fischer P, Wabitsch M, Brüderlein S, Hasel C, Möller P (2002) Leptin gene expression in human preadipocytes is switched on by maturation-induced demethylation of distinct CpGs in its proximal promoter. *J Biol Chem* **277:** 45420-45427

Milagro FI, Campión J, García-Díaz DF, Goyenechea E, Paternain L, Martínez JA (2009) High fat diet-induced obesity modifies the methylation pattern of leptin promoter in rats. *J Physiol Biochem* **65:** 1-9

Miner BE, Stöger RJ, Burden AF, Laird CD, Hansen RS (2004) Molecular barcodes detect redundancy and contamination in hairpin-bisulfite PCR. *Nucleic Acids Res* **32:** e135

Mise H, Sagawa N, Matsumoto T, Yura S, Nanno H, Itoh H, Mori T, Masuzaki H, Hosoda K, Ogawa Y, Nakao K (1998) Augmented placental production of leptin in preeclampsia: possible involvement of placental hypoxia. *J Clin Endocrinol Metab* **83:** 3225-3229

Montague CT, Prins JB, Sanders L, Digby JE, O'Rahilly S (1997) Depot- and sex-specific differences in human leptin mRNA expression: implications for the control of regional fat distribution. *Diabetes* **46:** 342-347

Morise A, Louveau I, Le Huërou-Luron I (2008) Growth and development of adipose tissue and gut and related endocrine status during early growth in the pig: impact of low birth weight. *Animal* **2:** 73-83

Morris DL, Rui L (2009) Recent advances in understanding leptin signaling and leptin resistance. *Am J Physiol Endocrinol Metab* **297:** E1247-1259

Moschos S, Chan J, Mantzoros C (2002) Leptin and reproduction: a review. *Fertil Steril* **77:** 433-444

Munzberg H, Bjornholm M, Bates SH, Jr MGM (2005) Leptin receptor action and mechanisms of leptin resistance. *Cellular and Molecular Life Sciences* **62:** 642-652

Myers M (2004) Leptin receptor signalling and the regulation of mammalian physiology. *Recent Progress in Hormone Research* **59:** 287-304

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

Nazarov PV, Reinsbach SE, Muller A, Nicot N, Philippidou D, Vallar L, Kreis S (2013) Interplay of microRNAs, transcription factors and target genes: linking dynamic expression changes to function. *Nucleic Acids Res* **41:** 2817-2831

Nestor C, Ruzov A, Meehan R, Dunican D (2010) Enzymatic approaches and bisulfite sequencing cannot distinguish between 5-methylcytosine and 5-hydroxymethylcytosine in DNA. *Biotechniques* **48:** 317-319

Nestor CE, Ottaviano R, Reddington J, Sproul D, Reinhardt D, Dunican D, Katz E, Dixon JM, Harrison DJ, Meehan RR (2012) Tissue type is a major modifier of the 5-hydroxymethylcytosine content of human genes. *Genome Res* **22:** 467-477

Ober C, Vercelli D (2011) Gene-environment interactions in human disease: nuisance or opportunity? *Trends Genet* **27:** 107-115

Okano M, Bell DW, Haber DA, Li E (1999) DNA methyltransferases DNMT3a and DNMT3b are essential for de novo methylation and mammalian development. *Cell* **99:** 247-257

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

Ooi SK, Bestor TH (2008) Cytosine methylation: remaining faithful. *Curr Biol* **18:** R174-176

Pandey N, Kriplani A, Yadav R, Lyngdoh B, Mahapatra S (2010) Peritoneal fluid leptin levels are increased but adiponectin levels are not changed in infertile patients with pelvic endometriosis. *Gynecol Endocrinol* **26:** 843-849

Pastor WA, Pape UJ, Huang Y, Henderson HR, Lister R, Ko M, McLoughlin EM, Brudno Y, Mahapatra S, Kapranov P, Tahiliani M, Daley GQ, Liu XS, Ecker JR, Milos PM, Agarwal S, Rao A (2011) Genome-wide mapping of 5-hydroxymethylcytosine in embryonic stem cells. *Nature* **473:** 394-397

Pease S, Braghetta P, Gearing D, Grail D, Williams RL (1990) Isolation of embryonic stem (ES) cells in media supplemented with recombinant leukemia inhibitory factor (LIF). *Dev Biol* **141:** 344-352

Penn NW, Suwalski R, O'Riley C, Bojanowski K, Yura R (1972) The presence of 5-hydroxymethylcytosine in animal deoxyribonucleic acid. *Biochem J* **126:** 781-790

Pfeifer GP, Steigerwald SD, Hansen RS, Gartler SM, Riggs AD (1990) Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability. *Proc Natl Acad Sci U S A* **87:** 8252-8256

Pinnick KE, Karpe F (2011) DNA methylation of genes in adipose tissue. *Proc Nutr Soc* **70:** 57-63

Piórkowska K, Oczkowicz M, Różycki M, Ropka-Molik K, Piestrzyńska-Kajtoch A (2011) Novel porcine housekeeping genes for real-time RT-PCR experiments normalization in adipose tissue: assessment of leptin mRNA quantity in different pig breeds. *Meat Sci* **87:** 191-195

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

Qin W, Leonhardt H, Pichler G (2011) Regulation of DNA methyltransferase 1 by interactions and modifications. *Nucleus* **2:** 392-402

Rayner DV, Dalgliesh GD, Duncan JS, Hardie LJ, Hoggard N, Trayhurn P (1997) Postnatal development of the ob gene system: elevated leptin levels in suckling fa/fa rats. *Am J Physiol* **273:** R446-450

Rishi V, Bhattacharya P, Chatterjee R, Rozenberg J, Zhao J, Glass K, Fitzgerald P, Vinson C (2010) CpG methylation of half-CRE sequences creates C/EBPalpha binding sites that activate some tissue-specific genes. *Proc Natl Acad Sci U S A* **107:** 20311-20316

Robertson J, Robertson AB, Klungland A (2011) The presence of 5-hydroxymethylcytosine at the gene promoter and not in the gene body negatively regulates gene expression. *Biochem Biophys Res Commun* **411:** 40-43

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

Ruzov A, Tsenkina Y, Serio A, Dudnakova T, Fletcher J, Bai Y, Chebotareva T, Pells S, Hannoun Z, Sullivan G, Chandran S, Hay DC, Bradley M, Wilmut I, De Sousa P (2011) Lineage-specific distribution of high levels of genomic 5-hydroxymethylcytosine in mammalian development. *Cell Res* **21:** 1332-1342

Salvaing J, Aguirre-Lavin T, Boulesteix C, Lehmann G, Debey P, Beaujean N (2012) 5-Methylcytosine and 5-hydroxymethylcytosine spatiotemporal profiles in the mouse zygote. *PLoS One* **7:** e38156

Schreiber E, Matthias P, Müller MM, Schaffner W (1989) Rapid detection of octamer binding proteins with 'mini-extracts', prepared from a small number of cells. *Nucleic Acids Res* **17:** 6419

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, De Carvalho DD, Jeong S, Jones PA, Liang G (2011) Nucleosomes containing methylated DNA stabilize DNA methyltransferases 3A/3B and ensure faithful epigenetic inheritance. *PLoS Genet* **7:** e1001286

Shiota K, Kogo Y, Ohgane J, Imamura T, Urano A, Nishino K, Tanaka S, Hattori N (2002) Epigenetic marks by DNA methylation specific to stem, germ and somatic cells in mice. *Genes Cells* **7:** 961-969

Simon JA, Kingston RE (2013) Occupying chromatin: Polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. *Mol Cell* **49:** 808-824

Sinclair KD, Allegrucci C, Singh R, Gardner DS, Sebastian S, Bispham J, Thurston A, Huntley JF, Rees WD, Maloney CA, Lea RG, Craigon J, McEvoy TG, Young LE (2007) DNA methylation, insulin resistance, and blood pressure in offspring determined by maternal periconceptional B vitamin and methionine status. *Proc Natl Acad Sci U S A* **104:** 19351-19356

Skinner MK (2011) Role of epigenetics in developmental biology and transgenerational inheritance. *Birth Defects Res C Embryo Today* **93:** 51-55

Song F, Smith JF, Kimura MT, Morrow AD, Matsuyama T, Nagase H, Held WA (2005) Association of tissue-specific differentially methylated regions (TDMs) with differential gene expression. *Proc Natl Acad Sci U S A* **102:** 3336-3341

Sproul D, Meehan RR (2013) Genomic insights into cancer-associated aberrant CpG island hypermethylation. *Brief Funct Genomics* **12:** 174-190

Stöger R (2006) In vivo methylation patterns of the leptin promoter in human and mouse. *Epigenetics* **1:** 155-162

Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* **9:** 465-476

Szulwach KE, Li X, Li Y, Song CX, Han JW, Kim S, Namburi S, Hermetz K, Kim JJ, Rudd MK, Yoon YS, Ren B, He C, Jin P (2011) Integrating 5-hydroxymethylcytosine into the epigenomic landscape of human embryonic stem cells. *PLoS Genet* **7:** e1002154

Szwagierczak A, Bultmann S, Schmidt CS, Spada F, Leonhardt H (2010) Sensitive enzymatic quantification of 5-hydroxymethylcytosine in genomic DNA. *Nucleic Acids Res* **38:** e181

Sérandour AA, Avner S, Oger F, Bizot M, Percevault F, Lucchetti-Miganeh C, Palierne G, Gheeraert C, Barloy-Hubler F, Péron CL, Madigou T, Durand E, Froguel P, Staels B, Lefebvre P, Métivier R, Eeckhoute J, Salbert G (2012) Dynamic hydroxymethylation of deoxyribonucleic acid marks differentiation-associated enhancers. *Nucleic Acids Res* **40:** 8255-8265

Tahiliani M, Koh KP, Shen Y, Pastor WA, Bandukwala H, Brudno Y, Agarwal S, Iyer LM, Liu DR, Aravind L, Rao A (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science* **324:** 930-935

Tartaglia LA (1997) The Leptin Receptor. *The Journal of Biological Chemistry* **272:** 6093-6096

Tchkonia T, Lenburg M, Thomou T, Giorgadze N, Frampton G, Pirtskhalava T, Cartwright A, Cartwright M, Flanagan J, Karagiannides I, Gerry N, Forse RA, Tchoukalova Y, Jensen MD, Pothoulakis C, Kirkland JL (2007) Identification of depot-specific human fat cell progenitors through distinct expression profiles and developmental gene patterns. *Am J Physiol Endocrinol Metab* **292:** E298-307

Thayer ZM, Kuzawa CW (2011) Biological memories of past environments: epigenetic pathways to health disparities. *Epigenetics* **6:** 798-803

Tobi EW, Lumey LH, Talens RP, Kremer D, Putter H, Stein AD, Slagboom PE, Heijmans BT (2009) DNA methylation differences after exposure to prenatal famine are common and timing- and sex-specific. *Hum Mol Genet* **18:** 4046-4053

Trayhurn P, Thomas ME, Duncan JS, Rayner DV (1995) Effects of fasting and refeeding on ob gene expression in white adipose tissue of lean and obese (oblob) mice. *FEBS Lett* **368:** 488-490

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 Cells* **11:** 805-814

Valinluck V, Sowers LC (2007) Endogenous cytosine damage products alter the site selectivity of human DNA maintenance methyltransferase DNMT1. *Cancer Res* **67:** 946-950

Valinluck V, Tsai HH, Rogstad DK, Burdzy A, Bird A, Sowers LC (2004) Oxidative damage to methyl-CpG sequences inhibits the binding of the methyl-CpG binding domain (MBD) of methyl-CpG binding protein 2 (MeCP2). *Nucleic Acids Res* **32:** 4100-4108

Van Harmelen V, Reynisdottir S, Eriksson P, Thörne A, Hoffstedt J, Lönnqvist F, Arner P (1998) Leptin secretion from subcutaneous and visceral adipose tissue in women. *Diabetes* **47:** 913-917

Velasco G, Hubé F, Rollin J, Neuillet D, Philippe C, Bouzinba-Segard H, Galvani A, Viegas-Péquignot E, Francastel C (2010) Dnmt3b recruitment through E2F6 transcriptional repressor mediates germ-line gene silencing in murine somatic tissues. *Proc Natl Acad Sci U S A* **107:** 9281-9286

Vilkaitis G, Suetake I, Klimasauskas S, Tajima S (2005) Processive methylation of hemimethylated CpG sites by mouse Dnmt1 DNA methyltransferase. *J Biol Chem* **280:** 64-72

Wabitsch M, Blum WF, Muche R, Braun M, Hube F, Rascher W, Heinze E, Teller W, Hauner H (1997) Contibution of androgens to the gender difference in Leptin Production in obese children and adolescents. *Journal of Clinical Investigation* **100:** 808-813

Wabitsch M, Jensen PB, Blum WF, Christoffersen CT, Englaro P, Heinze E, Rascher W, Teller W, Tornqvist H, Hauner H (1996) Insulin and cortisol promote leptin production in cultured human fat cells. *Diabetes* **45:** 1435-1438

Wajchenberg BL (2000) Subcutaneous and visceral adipose tissue: their relation to the metabolic syndrome. *Endocr Rev* **21:** 697-738

Wang B, Wood IS, Trayhurn P (2008) Hypoxia induces leptin gene expression and secretion in human preadipocytes: differential effects of hypoxia on adipokine expression by preadipocytes. *J Endocrinol* **198:** 127-134

Waterland RA, Jirtle RL (2003) Transposable elements: targets for early nutritional effects on epigenetic gene regulation. *Mol Cell Biol* **23:** 5293-5300

Watt F, Molloy PL (1988) Cytosine methylation prevents binding to DNA of a HeLa cell transcription factor required for optimal expression of the adenovirus major late promoter. *Genes Dev* **2:** 1136-1143

Wigler MH (1981) The inheritance of methylation patterns in vertebrates. *Cell* **24:** 285-286

Williams K, Christensen J, Helin K (2012) DNA methylation: TET proteins-guardians of CpG islands? *EMBO Rep* **13:** 28-35

Williams K, Christensen J, Pedersen MT, Johansen JV, Cloos PA, Rappsilber J, Helin K (2011) TET1 and hydroxymethylcytosine in transcription and DNA methylation fidelity. *Nature* **473:** 343-348

Wronska A, Kmiec Z (2012) Structural and biochemical characteristics of various white adipose tissue depots. *Acta Physiol (Oxf)* **205:** 194-208

Wu H, D'Alessio AC, Ito S, Wang Z, Cui K, Zhao K, Sun YE, Zhang Y (2011) Genome-wide analysis of 5-hydroxymethylcytosine distribution reveals its dual function in transcriptional regulation in mouse embryonic stem cells. *Genes Dev* **25:** 679-684

Wu H, Zhang Y (2011) Tet1 and 5-hydroxymethylation: A genome-wide view in mouse embryonic stem cells. *Cell Cycle* **10**

Wylie AR (2011) Leptin in farm animals: where are we and where can we go? *Animal* **5:** 246-267

Xu Y, Wu F, Tan L, Kong L, Xiong L, Deng J, Barbera AJ, Zheng L, Zhang H, Huang S, Min J, Nicholson T, Chen T, Xu G, Shi Y, Zhang K, Shi YG (2011) Genome-wide regulation of 5hmC, 5mC, and gene expression by Tet1 hydroxylase in mouse embryonic stem cells. *Mol Cell* **42:** 451-464

Yokomori N, Tawata M, Onaya T (2002) DNA demethylation modulates mouse leptin promoter activity during the differentiation of 3T3-L1 cells. *Diabetologia* **45:** 140-148

Yu M, Hon GC, Szulwach KE, Song CX, Zhang L, Kim A, Li X, Dai Q, Shen Y, Park B, Min JH, Jin P, Ren B, He C (2012) Base-resolution analysis of 5-hydroxymethylcytosine in the mammalian genome. *Cell* **149:** 1368-1380

Zaidi M, Buettner C, Sun L, Iqbal J (2012) Minireview: The link between fat and bone: does mass beget mass? *Endocrinology* **153:** 2070-2075

Zentner GE, Henikoff S (2013) Regulation of nucleosome dynamics by histone modifications. *Nat Struct Mol Biol* **20:** 259-266

Zha JM, Di WJ, Zhu T, Xie Y, Yu J, Liu J, Chen P, Ding G (2009) Comparison of gene transcription between subcutaneous and visceral adipose tissue in Chinese adults. *Endocr J* **56:** 935-944

Zhang F, Basinski M, Beals J, Briggs S, Churgay L, Clawson D, DiMarchi R, Furman T, Hale J, Hsiung H, Schoner B, Smith D, Zhang X, Wery J, Schevitz R (1997) Crystal structure of the obese protein leptin-E100. *Nature* **387:** 206-209

Zhang Y, Proenca R, Maffei M, Barone M, Leopold L, Friedman JM (1994) Positional cloning of the mouse obese gene and its human homologue. *Nature* **372:** 425-432

Zhu WG, Srinivasan K, Dai Z, Duan W, Druhan LJ, Ding H, Yee L, Villalona-Calero MA, Plass C, Otterson GA (2003) Methylation of adjacent CpG sites affects Sp1/Sp3 binding and activity in the p21(Cip1) promoter. *Mol Cell Biol* **23:** 4056-4065

Ziller MJ, Müller F, Liao J, Zhang Y, Gu H, Bock C, Boyle P, Epstein CB, Bernstein BE, Lengauer T, Gnirke A, Meissner A (2011) Genomic distribution and inter-sample variation of non-CpG methylation across human cell types. *PLoS Genet* **7:** e1002389

Ziylan YZ, Baltaci AK, Mogulkoc R (2009) Leptin transport in the central nervous system. *Cell Biochemistry and Function* **27:** 63-70