Investigation into the epigenetic events leading to heritable silencing of Cdkn1c

Michelle Diane Wood

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Summary of thesis

In mammals, there are subsets of genes that are only expressed from one parental allele. This phenomenon, known as genomic imprinting, relies on epigenetic marks to distinguish between the two parental alleles. These epigenetic marks are erased and re-established in the germ line of every generation.

Cdkn1c is an imprinted gene in humans and mice. Maternal specific expression of *Cdkn1c* depends on a germ line DNA methylation mark at a distant cis element (KvDMR1), also known as the imprint control region. When DNA methylation is lost at KvDMR1, *Cdkn1c* expression is suppressed. On the paternal allele, silencing of *Cdkn1c* is associated with the expression of a distantly located non-coding RNA (*Kcnq1ot1/Lit1*) overlapping the KvDMR1, which is presumed to mediate silencing through the formation of heterochromatin. In somatic cells, long term full silencing of paternal *Cdkn1c* is maintained by DNA methylation at a CpG island within *Cdkn1c*.

We have shown that embryonic germ (EG) cells, derived from 12.5 dpc primordial germ cells, are a good model for studying the early epigenetic events that lead to heritable silencing of imprinted genes. *Cdkn1c* expression was suppressed in differentiated EG cells compared to embryonic stem (ES) cells indicating that this domain is silent in EG cells. The *Cdkn1c* CpG island was hypomethylated in both undifferentiated and differentiated EG cells and was also hypomethylated in ES and androgenetic (AG) stem cells indicating that DNA methylation was not required for silencing *Cdkn1c*. Suppression of *Cdkn1c* was likely due to the acquirement of repressive histone marks. H3K27me3 was found enriched at the *Cdkn1c* promoter in differentiated EG cells and in undifferentiated and differentiated AG cells but was absent in undifferentiated EG cells indicating that this mark was specifically recruited during EG cell differentiation.

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ....Meneage..... (candidate) Date ..06/08/08.....

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of .....PhD.....(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed .....Meneage.....(candidate) Date .06/08/08.....

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references.

Signed .....Meneage.....(candidate) Date ....06/08/08.....

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ....Meneage.....(candidate) Date ...06/08/08.....

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Graduate Development Committee.**

Signed ....Meneage.....(candidate) Date ...06/08/08.....

Acknowledgements

The work in this PhD thesis would not be possible without the generous donation of stem cell stocks. I would like to thank M A Surani for supplying the initial stocks of TMAS21G, Sv6.1, CES3 and AKR1 stem cells and Dr. Robert Feil for providing the SF1-1 cell line.

I would also like to thank Pr. Michael Higgins for the ribonuclease protection assay and Dr. Tatyana Nesterova for the southern blot data.

I would like to thank Dr. Rosalind John for being my supervisor on this project, for her encouragement and support, even when things did not go to plan, and for always finding the positive in my results. I would also like to thank Dr. Nick Allen for being my advisor on this project and for all his help and advice on working with stem cells and for providing me with replacement AKR1 cells.

Also to everyone in genetics, thankyou for all your support and many good times.

The work in this PhD thesis was funded by a grant from BBSRC epigenetics initiative.

Abbreviations

PEFs	Primary embryonic fibroblasts
PGCs	Primordial germ cells
EG	Embryonic germ cells
AG	Androgenetic stem cells
PG	Parthenogenetic stem cells
ES	Embryonic stem cells
DMR	Differentially methylated region
ICR	Imprint control region
IC	Imprinting centre
Dpc	Days post coitum
BWS	Beckwith Wiedemann syndrome
5meC	5 methyl-cytosine
Dnmt	DNA methyltransferase
CpG	Cystosine-guanine dinucleotide
GC	Guanine-cytosine
HMT	Histone methyltransferase
HAT	Histone acetyltransferase
HDAC	Histone deacetylase
MBD	Methyl binding domain
PRC	Polycomb repressor complex
PcG	Polycomb group
CTCF	CCCTC binding factor
Lys	Lysine
H3	Histone H3
Me2	Di-methylation
Me3	Tri-methylation
Ac	Acetylation
DNA	Deoxyribonuclease
RNA	Ribonuclease
Kb	Kilobases
Mb	Megabases
bp	Base pairs
Min	Minutes
Sec	Seconds
Hrs	Hours
LINEs	Long interspersed nuclear elements
SINEs	Short interspersed nuclear elements
SP1	Specificity protein 1
HP1	Heterochromatin protein 1
ChIP	Chromatin immunoprecipitation
PCR	Polymerase chain reaction
RT-PCR	Reverse transcribed PCR
qPCR	quantitative PCR
RPA	Ribonuclease protection assay
MSRE	Methylation sensitive restriction endonuclease
IgG	Immunoglobulin G

List of Figures

List of Tables

Abstract

Genomic imprinting is the process of allele-specific gene regulation. Heritable epigenetic modifications such as DNA methylation and chromatin modifications establish and maintain differences between the parental genomes that result in the expression of only one allele. Imprints (including methylation marks) are reset in the germ line during gametogenesis. Mouse distal chromosome 7 contains a 1 Mb imprinted domain that is homologous to human 11p15 chromosome region and contains approximately 16 imprinted genes. Dysregulation of imprinting within this cluster can cause Beckwith-Wiedemann syndrome (BWS) in humans, a developmental disorder characterised by foetal overgrowth, embryonic tumours and developmental abnormalities. The majority of familial cases of BWS, which represent less than 5 % of the total reported, are caused by mutations in the cyclin-dependent kinase inhibitor gene *CDKN1C*. Imprinting of mouse *Cdkn1c* is controlled by the imprinting control region KvDMR1, a differentially methylated CpG island within intron 10 of the imprinted *Kcnq1* gene. Maternal methylation of the KvDMR1 correlates with suppression of the antisense non-coding RNA transcript *Kcnq1ot1/Lit1* and expression of the surrounding genes. On the paternal chromosome, KvDMR1 is unmethylated and *Kcnq1ot1/Lit1* is expressed. The surrounding genes are silent. It has been proposed that *Kcnq1ot1/Lit1* acts in a similar way as *Xist* in X-chromosome inactivation by recruiting histone modifying enzymes such as the polycomb repressor complexes (PRC) and DNA methyltransferases.

Primordial germ cells (PGC) derived from 12.5 dpc embryos can be cultured *'in vivo'* as embryonic germ (EG) cells. They lack DNA methylation at ICRs and are thought to represent an 'imprint erased' cell type. In the absence of an imprint the

maternally expressed genes within the IC2 domain are silent (Lee *et al.* 2002; Tada *et al.* 1998). There appears to be a default sequence of silencing that occurs in the absence of an imprint. Tada *et al.* showed that *Cdkn1c* was not expressed in *in vivo* differentiated EG cells. Moreover, the locus acquired *de novo* methylation when differentiated *in vivo* (Tada *et al.* 1998).

The aim of this study was to determine the default sequence of epigenetic events that lead to the silencing of *Cdkn1c* in EG cells, as reported by Tada *et al.*, by establishing an *in vitro* model for differentiation. Expression of the imprinted genes within the IC2 domain was examined by real time quantitative RT-PCR in order to determine if the genes were silenced during differentiation of the EG cells. Various DNA methylation techniques were used to determine if *Cdkn1c* methylation, which was reported to occur *de novo in vivo*, was established in these differentiated cells. Also the histone modifications associated with 'non-imprinted' and 'imprinted' chromatin were examined in EG and androgenetic (AG) cells in order to determine if there were changes to histone modifications during differentiation that might be involved in the silencing of the IC2 domain genes.

Our data confirmed the absence of DNA methylation at both the germ line and somatic DMRs within the IC2 domain in undifferentiated EG cells. We showed that the maternally expressed genes were not expressed in these cells after differentiation, reflecting the *in vivo* data. In contrast *Kcnq1ot1/Lit1* expression was detected in undifferentiated and differentiated EG cells at levels that appeared biallelic. Our analysis also confirmed the observations of others which suggest the existence of at least two types of imprinted gene, those that require an imprint to be silenced (the normal accepted view of imprinting) and those that require an imprint to be expressed. This is supported by data from cloned embryos made from day 11.5 XX and XY

PGCs where, in the absence of an imprint on KvDMR1, *Cdkn1c* is not expressed (Lee *et al.* 2002).

We showed that methylation at the *Cdkn1c* DMR was not acquired in differentiated EG cells. Surprisingly, we also found that ES cells were hypomethylated at the *Cdkn1c* locus in both undifferentiated and differentiated cells indicating that silencing of paternal expression of *Cdkn1c* occurs independently of DNA methylation. Chromatin immunoprecipitation (ChIP) analysis showed that the repressive H3K27me3 modification was present in both undifferentiated and differentiated AG cells. In contrast, this mark was only present in the differentiated EG cells. This suggested that H3K27me3 was actively recruited to the *Cdkn1c* locus as EG cells differentiated. This supports the hypothesis that EG cells are not equivalent to the silenced paternal allele of the mouse distal chromosome 7 domain but represent an earlier chromatin state. Therefore, we can use these cells to identify the sequence of events that lead to silencing of *Cdkn1c*.

CHAPTER 1:

General Introduction

1. General Introduction

1.1 Epigenetic regulation of gene expression

1.1.1 Genetic and epigenetic information

All the information required for an organism to grow, develop and reproduce is encoded within its DNA. The number of protein coding genes found within the human and mouse genome is approximately 20,000 – 25,000 (Abdellah *et al.* 2004; Lander *et al.* 2001; Waterston *et al.* 2002). Development of an organism is a very orderly process that involves the silencing and expressing of genes at specific time points to give the array of differentiated cell types within the adult organism. This switching on or off of genes is known as 'epigenetics' and was first described by Waddington in the 1940s as a way of modifying the actions of genes during development (Waddington 1942). It is now accepted that, in most cases, chemical modifications to DNA and chromatin are responsible for these complex changes in gene expression and not alterations or deletions to the DNA sequences. This includes the methylation of DNA and the packaging of DNA into chromatin (Hebbes *et al.* 1988; Holliday and Pugh 1975; Monk *et al.* 1987; Razin and Cedar 1977). Epigenetic information is naturally robust and reliable, can be stably inherited through numerous cell divisions and is involved in diverse processes from controlling tissue-specific gene expression, genomic imprinting and X-inactivation (Hebbes *et al.* 1988; Jeppesen and Turner 1993; Monk *et al.* 1987; Roemer *et al.* 1997).

1.1.2 Mono-parental embryos

The non-equivalence of the parental genomes in mammals was first identified in the early 1980s in experiments using pronuclear transplantation to generate 'gynogenetic' (2 maternal genomes) and 'androgenetic' (2 paternal genomes) mouse

embryos (Barton *et al.* 1984; McGrath and Solter 1984; Surani and Barton 1983; Surani *et al.* 1984) (**figure 1.1**). These embryos died at around 10 days of gestation and, when the foetuses were compared, morphological and developmental differences were observed (McGrath and Solter 1984; Surani *et al.* 1984). The gynogenetic embryos developed to the 25-somite stage but were growth retarded and showed poor development of the extraembryonic tissues (Barton *et al.* 1984; Surani and Barton 1983). Androgenetic embryos developed poorly with almost exclusive development of the extraembryonic tissues (Barton *et al.* 1984; Surani *et al.* 1984).

Both gynogenetic and parthenogenetic embryos contain two maternal genomes. However, gynogenetic embryos are created by pro-nuclear injection whereas parthenogenetic embryos are created by artificially activating an oocyte. Parthenogenetic foetuses demonstrate the same developmental abnormalities as gynogenetic foetuses suggesting that the presence of two female pronuclei is not developmentally viable (Surani and Barton 1983; Surani *et al.* 1984). The pro-nuclear transplantation experiments demonstrated that, in mammals, both parental genomes are required for normal embryonic development (McGrath and Solter 1984; Surani *et al.* 1984).

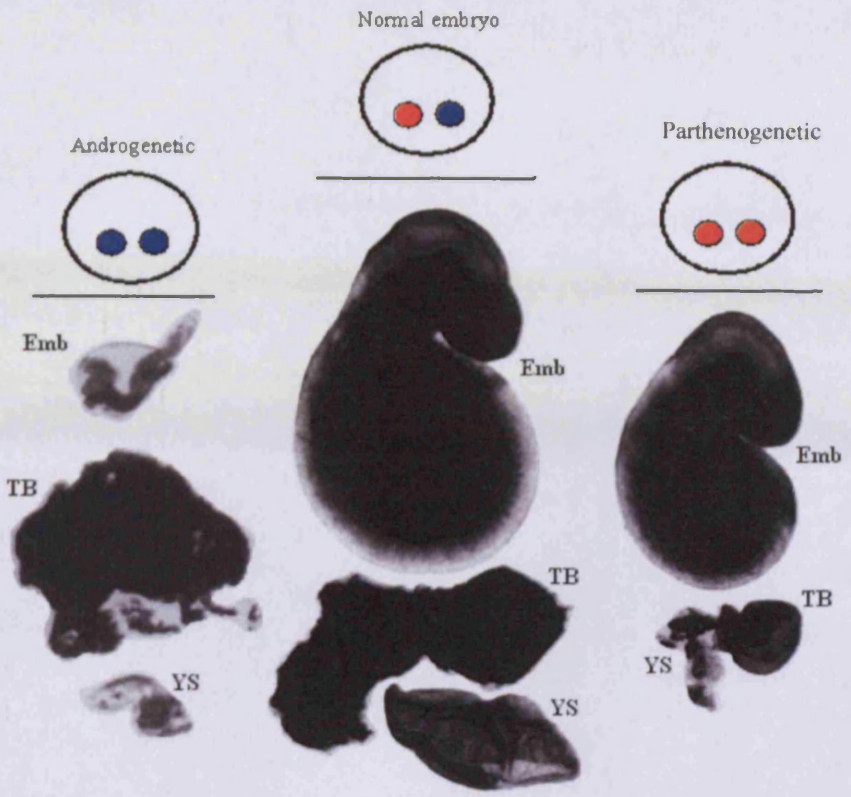


Figure 1.1: Parthenogenetic and androgenetic embryos.

Pro-nuclear transplantation was used to create parthenogenetic (2 maternal genomes) or androgenetic (2 paternal genomes) embryos. Parthenogenetic embryos show relatively normal embryonic development (Emb = Embryo) until 10.5 dpc but poor development of the extraembryonic tissue (YS = yolk sac, TB = trophoblast). Androgenetic embryos have a well defined trophoblast at this time point but poor development of the embryo. Adapted from (Barton *et al.* 1984; Surani *et al.* 1984). Copyright Macmillan Publishers Limited.

1.1.3 Genomic imprinting

Within vertebrate groups, genomic imprinting is a type of gene regulation that is specific to mammals and results in differential expression of some autosomal genes [Reviewed in (Delaval and Feil 2004; Reik and Walter 2001)]. Most autosomal genes are biallelically expressed but there are a subset of genes that are only expressed from either the paternal or the maternal allele (Beechey *et al.* 2008; Cattanach and Kirk 1985). Imprinted genes are expressed (and imprinted) predominantly in the developing embryo, placenta and the perinatal and postnatal brain. They are

commonly involved in embryonic growth, foetal development and behaviour after birth; many also seem to be important for placental development (Beechey *et al.* 2008; Keverne *et al.* 1996; Mann *et al.* 2004; Plagge *et al.* 2005; Salas *et al.* 2004). Thus far, approximately 90 imprinted genes have been identified (Beechey *et al.* 2008). Those that are involved in growth promotion, such as the insulin-like growth factor 2 (*Igf2*), tend to be expressed from the paternal derived chromosome whereas those that are growth suppressing, like the cyclin-dependent kinase inhibitor *Cdkn1c*, tend to be expressed from the maternally derived chromosome (Caspary *et al.* 1999; Morison and Reeve 1998). **Figure 1.2** shows a number of imprinted genes found within the mouse genome.

Imbalanced expression of imprinted genes is probably responsible for the death of gynogenetic and androgenetic mouse embryos. The paternally expressed genes *Peg1/Mest, Igf2, Peg3* and *Snrpn* are not expressed in gynogenetic or parthenogenetic embryos whereas maternally expressed genes *H19* and *Igf2r* are biallelic. Conversely, in the androgenetic embryos *H19* and *Igf2r* are repressed but *Igf2* and *Snrpn* are biallelically expressed (Horii *et al.* 2008; Obata *et al.* 1998; Szabo and Mann 1994).

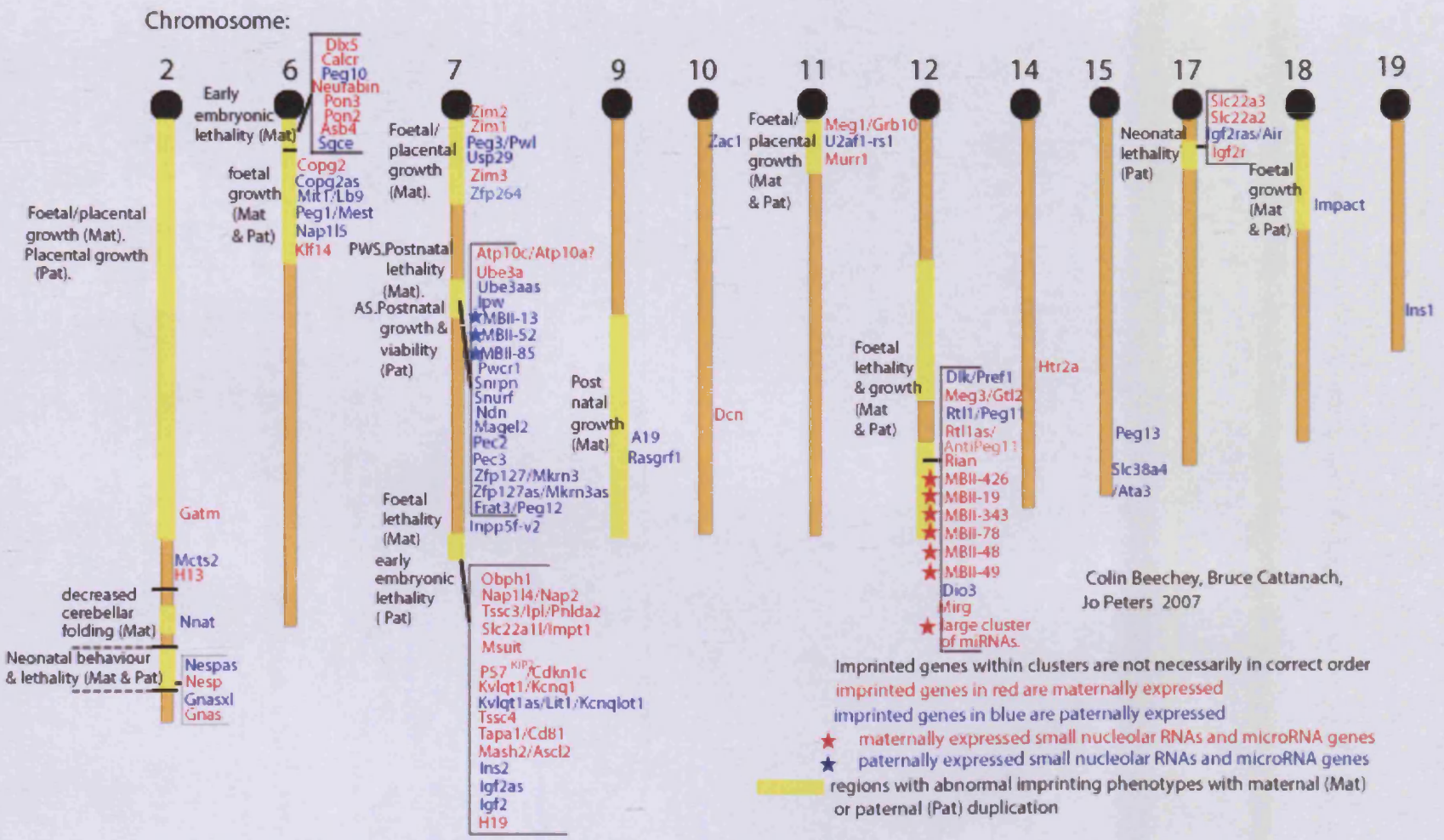


Figure 1.2: Location of imprinted genes in the mouse genome.

The majority of imprinted genes within the mouse genome are found and co-ordinately regulated in clusters and contain both maternally and paternally expressed genes. Around 90 imprinted genes have, so far, been identified. The majority of the imprinted genes have roles in embryonic growth, development and behaviour. From (Beechey *et al.* 2008).

1.1.4 The evolution of imprinting

The evolution of imprinting is a hotly debated topic with many theories surrounding the question of why we have imprinted genes. The majority of the genes in the mammalian genome are diploid (the exceptions are present on the sex chromosomes) which is presumed to reduce the effects of deleterious mutations that can result in a missing, altered or non-functioning protein. Genomic imprinting effectively negates the effect of diploidy so essentially all imprinted genes are functionally haploid. This means that a mutation in the expressed copy of the gene can not be compensated for as the second copy is silent.

The parental conflict theory, developed by David Haig, predicted that paternally expressed genes would be growth promoting whereas maternally expressed genes would be growth inhibiting (Haig 1997; Haig and Westoby 1989) [Reviewed in (Moore and Reik 1996; Wilkins and Haig 2003)]. This applies to mammalian species where there is an unequal provision of resources after conception (Haig 1997; Moore and Haig 1991). In the parental conflict model it is theorised that the paternal genome derived imprinted genes are driven to promote growth by extracting maximal resources for the offspring, while the maternal derived genes aim to limit this demand so as to not affect the mother's own fitness and productivity (Haig and Graham 1991; Moore and Haig 1991). This theory, however, can only be applied to imprinted genes that regulate growth.

Other theories have been proposed to explain why we have imprinted genes [Reviewed in (Wilkins and Haig 2003)]. One theory is that genomic imprinting arose as a mechanism to protect females from ovarian trophoblastic disease (Varmuza and Mann 1994). Another suggests that imprinting may have also developed as a way to protect a subset of alleles from natural selection and so allow mammals to adapt

quickly to environmental changes (Beaudet and Jiang 2002). The fact that many of the imprinted genes are involved in growth and nutrition does seem to suggest that the parental conflict theory is a good explanation for the evolution of imprinting. However, it may be that imprinting evolved as a result of multiple evolutionary pressures.

Among vertebrates, autosomal imprinting has only been found in eutherian mammals and marsupials but not in egg-laying species such as fish, amphibians, reptiles or birds (Killian *et al.* 2001; Nolan *et al.* 2001; Suzuki *et al.* 2005) and not in the oviparous monotremes such as the platypus and the echidna (Killian *et al.* 2000; Killian *et al.* 2001). This suggests that imprinting may have evolved along side the emergence of viviparity and particularly for the development of the extraembryonic tissues including the trophoblast. Many imprinted genes, such as achaete-scute complex homolog-like 2 (*Ascl2)* and tumour-suppressing subchromosomal transferable fragment 4 *(Tssc4),* are expressed solely in the placenta (Cerrato *et al.* 2005; Umlauf *et al.* 2004). However, this would also comply with the parental conflict model as controlling placental development can affect the growth of the developing offspring.

Although imprinting of a considerable number of genes is highly conserved between the mouse and human, there are also several genes where the imprinting status differs between the two species (Kalscheuer *et al.* 1993; Monk *et al.* 2006). The insulin-like growth factor receptor 2 (*IGF2R*) gene is imprinted only during early development then lost in later developmental stages in some humans whereas the imprinting status of the mouse *Igf2r* gene is stable at all developmental stages (Lerchner and Barlow 1997; Monk *et al.* 2006; Riesewijk *et al.* 1998). The *IGF2R* gene is polymorphically imprinted in humans with the majority showing biallelic

expression of this gene (Riesewijk *et al.* 1998; Xu *et al.* 1993). This may be because mice are multiparous while humans are generally uniparous resulting in relaxation of imprinting in humans due to reduced selective pressure on *IGF2R*.

1.1.5 Imprinting disorders and Beckwith-Wiedemann syndrome (BWS)

Genomic imprinting is essential for normal embryonic growth. There are several disorders that result from a disruption in imprinting, often leading to severe growth anomalies or inviability for the developing embryo (Lee 2003). An example is Beckwith-Wiedemann Syndrome (BWS), a congenital disorder characterised by pre- and post-natal macrosomnia, macroglossia, organomegaly, exomphalos and embryonic tumours (Beckwith 1969; Gicquel and Le Bouc 2001) [BWS reviewed in (Maher and Reik 2000)]. The disease was mapped to a 1 Mb imprinting domain on human chromosome 11p15 (Koufos *et al.* 1989; Ping *et al.* 1989) that contains both growth promoting and growth inhibiting genes and is homologous to an imprinting domain on mouse distal chromosome 7 (Onyango *et al.* 2000). BWS can arise as a result of uniparental disomy (UPD), specific mutations or translocations of human chromosome 11p15 (Hatada *et al.* 1996; Henry *et al.* 1991; Hoovers *et al.* 1995; Lam *et al.* 1999). There are also some epigenetic abnormalities that can also lead to disruptions in imprinting such as loss of imprinting of *KCNQ1OT1/LIT1* which is the most common genetic alteration in BWS (Lee *et al.* 1999; Weksberg *et al.* 2001).

Interestingly, the only mutation that results in familial cases of BWS is located within the growth inhibitor gene *CDKN1C* that is normally silent on the paternal chromosome. Maternally inherited mutations have been identified in ~ 40 % of BWS families and ~5 % of sporadic cases (Hatada *et al.* 1996; Lam *et al.* 1999; Lee *et al.* 1997). *CDKN1C* encodes a cyclin dependent kinase inhibitor belonging to the CIP

family of regulators and is involved in regulation of the cell cycle during the transition between the G1 and S phases (Lee *et al.* 1995; Matsuoka *et al.* 1995). From experiments in mice where *Cdkn1c* was deleted, it was identified that loss of *Cdkn1c* expression results in multiple developmental abnormalities, such as exomphalos and renal medullary dysplasia, that are associated with BWS, as well as contributing to the overgrowth phenotype (Zhang *et al.* 1997). This concurs with the observation that exomphalos is more common in patients with mutations in *CDKN1C* but not *IGF2* (Cooper *et al.* 2005; Lam *et al.* 1999). Over-expression of the *IGF2* gene, within the same cluster, contributes not only to the overgrowth phenotype but is involved in Wilm's tumour development (Bliek *et al.* 2001; Cooper *et al.* 2005; Sparago *et al.* 2007). BWS patients with a loss of imprinting of *IGF2* are phenotypically indistinguishable from those with mutations in *CDKN1C* although those with a mutation in *CDKN1C* are unlikely to develop embryonic tumours (Caspary *et al.* 1999; Weksberg *et al.* 2001). The effects of increased *IGF2* in BWS may be mediated through a decrease in *CDKN1C* expression as it has been shown that *Cdkn1c* expression is reduced in mice with high serum levels of *Igf2* (Caspary *et al.* 1999; Grandjean *et al.* 2000).

1.2 Mechanisms of imprinting

1.2.1 ICRs and models of gene silencing

Epigenetic mechanisms such as DNA methylation and chromatin modifications regulate gene expression in several areas of the mammalian genome including X chromosome inactivation, heterochromatin formation, silencing of parasitic elements and genomic imprinting (Hebbes *et al.* 1988; Kohlmaier *et al.* 2004; Lehnertz *et al.* 2003; Li *et al.* 1993; Sado *et al.* 2000).

Imprinted genes generally occur in clusters that extend for distances as large as 4 Mb (Buiting *et al.* 1995; Koufos *et al.* 1989; Lopes *et al.* 2003). They are co-ordinately regulated within these clusters suggesting the presence of shared regulatory elements acting to control imprinting of multiple genes. These regulatory elements are termed imprinting control regions (ICRs) and several have been identified by their differential methylation patterns in imprinted clusters and loss of imprinting when these regions are deleted (Ainscough *et al.* 2000; Buiting *et al.* 1995; Lee *et al.* 1999; Leighton *et al.* 1995; Smilinich *et al.* 1999; Thorvaldsen *et al.* 1998). Heritable modifications of DNA and chromatin may work together to confer imprinting as methylation of an ICR can affect expression of imprinted genes that are several Kb away (Lopes *et al.* 2003; Zwart *et al.* 2001). Not all genes within a domain are imprinted. Therefore, either ICRs have the ability to control which genes within a cluster become imprinted or there are differences in the target sequences between imprinted and non-imprinted genes. Imprinted genes and ICRs are often associated with CpG islands (Smilinich *et al.* 1999; Sutcliffe *et al.* 1994; Wutz *et al.* 1997). These CpG islands often show allele-specific differences in 5-methylcytosine (5meC) patterns and are called differentially methylated regions (DMRs) (Li *et al.* 1993; Sutcliffe *et al.* 1994; Thorvaldsen *et al.* 1998; Tremblay *et al.* 1997). The DMRs lie within the imprinting centre and differential DNA methylation is usually, but not exclusively, established in the germ line and so point to a role in controlling imprinting (Davis *et al.* 2000; Smilinich *et al.* 1999; Tucker *et al.* 1996).

It is not yet fully understood how imprinting occurs over such large domains. Many of the imprinted domains that have been studied so far appear to employ different mechanisms for imprinting ranging from an insulator model, to block access of imprinted genes to enhancers, and to the non-coding RNA model that recruits

chromatin remodelling enzymes. The insulator model has been proposed for the *Igf2-H19* imprinting domain (**figure 1.3 A**) where the ICR functions as a methylation-sensitive binding domain for the zinc-finger transcription factor CTCF (CCCTC-binding factor) which is located between the reciprocally imprinted *Igf2* and *H19* genes (Leighton *et al.* 1995; Schoenherr *et al.* 2003; Thorvaldsen *et al.* 1998; Yang *et al.* 2003). The CTCF protein has been described previously as an activator, repressor and chromatin insulator element and forms a transcriptional switch between *H19* and *Igf2* expression (Bell and Felsenfeld 2000; Ishihara *et al.* 2006; Schoenherr *et al.* 2003; Yang *et al.* 2003). The ICR is unmethylated on the maternal chromosome and can bind CTCF. This insulates *Igf2* from the downstream shared enhancers resulting in expression of *H19*, a non-coding RNA transcript. On the methylated paternal chromosome, CTCF cannot bind the ICR so the enhancers can drive expression of *Igf2* (Schoenherr *et al.* 2003; Thorvaldsen *et al.* 1998; Yang *et al.* 2003). A functional role for the *H19* transcript has not yet been determined but it may serve to recruit repressive histone modifications to the *Igf2* gene to further prevent transcription (Gabory *et al.* 2006; Jones *et al.* 1998a).

The non-coding RNA model was proposed for the *Igf2r* domain where the ICR is found within the *Igf2r* gene (**figure 1.3 B**). The ICR is a differentially methylated CpG island located at the promoter region of an antisense non-coding RNA transcript called *Air* which is expressed on the unmethylated paternal chromosome whereas the surrounding genes are silent (Lyle *et al.* 2000; Sleutels *et al.* 2002; Wutz *et al.* 1997). On the methylated maternal chromosome the antisense transcript is not expressed. Imprinting of the *Igf2r* gene has been shown to occur in opossum and canines but in the absence of both methylation of the DMR and expression of the antisense transcript (O'Sullivan *et al.* 2007; Weidman *et al.* 2006),

suggesting that these mechanisms may have evolved later to regulate silencing within the *Igf2r* imprinting domain.

The non-coding RNA model has also been proposed for the mouse distal chromosome 7 IC2 imprinting domain (**figure 1.3 C**). In this domain maternal methylation of the ICR (known as KvDMR1) results in the expression of the surrounding maternally expressed genes (Fitzpatrick *et al.* 2002; Kanduri *et al.* 2002). On the paternal chromosome, DNA methylation is absent from KvDMR1, the surrounding maternally expressed genes are silent but there is expression of the paternally expressed non-coding RNA, *Kcnq1ot1/Lit1,* from a promoter within KvDMR1 (Diaz-Meyer *et al.* 2003; Fitzpatrick *et al.* 2002; Shin *et al.* 2008).

Non-coding RNAs are found within most imprinting domains and an attractive hypothesis is that they are involved in targeting repressive histone modifications to the flanking genes as has been shown for *Xist* and X chromosome inactivation (Clemson *et al.* 1996). It is also possible that expression of the non-coding RNA in imprint domains simply reflects the activity of other regulatory elements such as boundary and silencer elements. However, recent data for the *Igf2r* cluster (Sleutels *et al.* 2002) and the IC2 domain (Mancini-DiNardo *et al.* 2006) suggests that it is transcription of the non-coding transcript itself that is required to establish silencing of neighbouring genes.

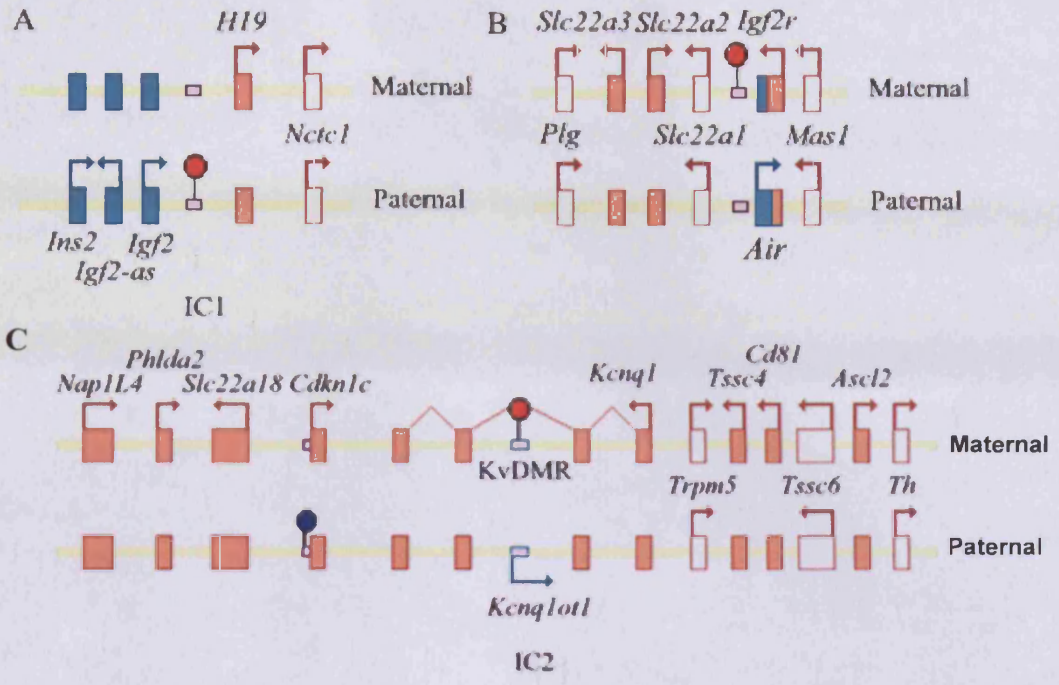


Figure 1.3: Coordinated regulation of imprinted genes within clusters.

A IC1 domain from mouse distal chromosome 7. When unmethylated the *H19* DMR binds CTCF preventing access of *Igf2* to its downstream enhancers (Cerrato *et al.* 2005). **B** *Igf2r* imprinted domain showing the non-coding RNA model of regulation. The DMR is located on the promoter region of *Air*. When methylated expression of *Air* is prevented resulting in expression of the surrounding imprinted genes (Sleutels *et al.* 2001). **C** IC2 domain from mouse distal chromosome 7. Maternal methylation of the KvDMR1 results in expression of the surrounding maternally expressed genes. On the paternal allele, lack of methylation at the KvDMR1 results in expression of a non-coding RNA *Kcnq1ot1/Lit1* and silencing of the surrounding maternally expressed genes (Cerrato *et al.* 2005).
Blue boxes indicate paternally expressed genes, orange boxes indicate maternally expressed genes, and white boxes are biallelically expressed genes. Red lolly pops are germ line methylation marks, blue lolly pops are somatic methylation marks. Pink boxes are differentially methylated regions (DMRs). Arrows indicate direction of transcription.

1.2.2 DNA methylation and gene silencing

DNA methylation is part of a multi-level control mechanism which regulates the expression of eukaryotic genes. DNA methylation normally has a repressive function when found at the promoters of genes. Genomic methylation can be inhibited

by treating cells with 5-azacytidine resulting in the reactivation of the transcriptionally inactive X-chromosome and abnormal gene expression (Jones *et al.* 1983; Mohandas *et al.* 1981; Vasques *et al.* 2005).

DNA is methylated by the addition of a methyl group at the 5 position of cytosine to form 5-methylcytosine (5meC) (Razin and Cedar 1977). The majority of DNA methylation within the mammalian genome occurs at symmetrical CpG dinucleotides and is mediated by DNA methyltransferases (Bestor *et al.* 1988; Bird 1986; Li *et al.* 1992) [reviewed in (Bird 2002)] (**figure 1.4**). The proportion of 5meC in the vertebrate genome is just 1 - 2 %, potentially as a consequence of the mutagenic nature of 5meC which can undergo deamination (Chan *et al.* 2001; Cooper and Krawczak 1989; Ehrlich *et al.* 1982; Razin and Cedar 1977). CpG dinucleotides are under-represented in the mammalian genome except for discrete regions called CpG islands. CpG islands are defined as regions of DNA that are more than 200 bp in length with a normal CpG (cytosine-guanine) dinucleotide density i.e. CpG dinucleotides are not under represented (Bird *et al.* 1985; Gardiner-Garden and Frommer 1987). Many CpG islands are associated with the promoter regions of genes and the majority are unmethylated (Gardiner-Garden and Frommer 1987; Yamashita *et al.* 2005). During development, some CpG islands become methylated and the associated genes become stably silent (Monk *et al.* 1987). CpG islands can also be found in many parasitic elements such as retrotransposons, LINEs and SINEs and these are usually methylated (Walsh *et al.* 1998; Yoder *et al.* 1997).

Methylation of CpG islands is normally linked with gene silencing. Methylation of cytosine is associated with alterations to the chromatin structure, affecting the interaction of transcription factor complexes with DNA [reviewed in (Bird and Wolffe 1999; Jones and Takai 2001; Santos *et al.* 2005)]. Specificity protein

1 (SP1) is a transcription factor found in eukaryotes that binds to a conserved 'GC box' motif containing a potentially methylated CpG site. Methylation of this site, however, does not prevent SP1 from binding (Harrington *et al.* 1988). SP1 binding can form a physical barrier that prevents nearby CpG islands from becoming methylated and may be required to keep frequently active genes switched on (Brandeis *et al.* 1994; Siegfried *et al.* 1999). There is evidence to suggest that removal of the SP1 binding sites flanking a CpG island can lead to its *de novo* methylation during development (Brandeis *et al.* 1994; Macleod *et al.* 1994; Siegfried *et al.* 1999). There are a large number of parasitic DNAs within the mammalian genome and these elements must be constantly silenced suggesting that it may be necessary for the genome to stay in a default silent, and so methylated state (Walsh *et al.* 1998; Yoder *et al.* 1997).

DNA methylation can recruit non-histone proteins, such as the methylated CpG-binding domain (MBD) proteins. These are transcriptional repressors that interact with, and establish complexes with, histone deacetylating complexes (HDACs) and nucleosome remodelling proteins (Jones *et al.* 1998b; Nan *et al.* 1998; Reese *et al.* 2007). MECP2 is a MBD protein involved in gene silencing and mutations in *MECP2* have been found in Rett syndrome patients (Amir *et al.* 1999; Makedonski *et al.* 2005). This protein is thought to promote silencing by binding to methylated DNA and recruiting HDACs to the surrounding chromatin, blocking transcription in mammals (Jones *et al.* 1998b; Nan *et al.* 1998).

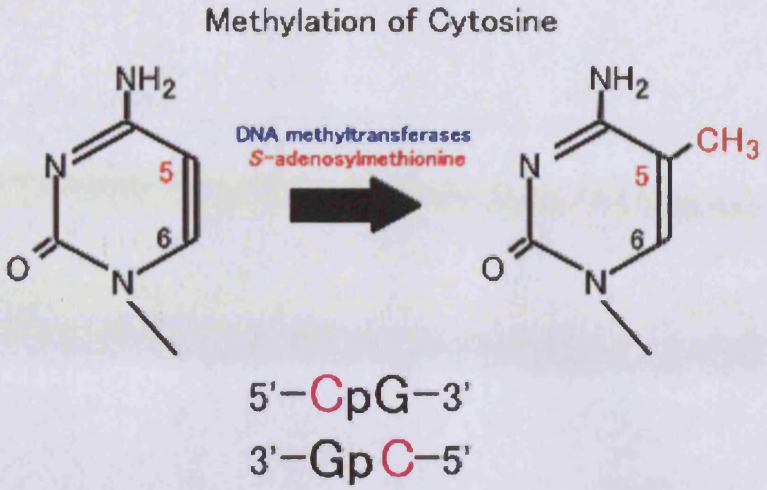


Figure 1.4: Methylation of cytosine at CpG dinucleotides.

Cytosines are methylated at the 5 position of cytosine at the symmetrical CpG dinucleotide by the action of DNA methyltransferases (Dnmts). From **http://www.sci.osaka-u.ac.jp/introduction/eng/biology.html**

1.2.3 DNA methyltransferases (Dnmts)

DNA methyltransferases (Dnmts) are enzymes that are responsible for establishing and maintaining methylation at CpG dinucleotides. In mammals there are three distinct families of DNA methyltransferases: Dnmt1, Dnmt2 and Dnmt3. Dnmt3 enzymes are thought to control *de novo* methylation in the germ line and in the pre-implantation embryo (Hsieh 1999; Okano *et al.* 1998a; Xie *et al.* 1999) (**Figure 1.5**).

Dnmt1 is a maintenance methylation enzyme as it has a high affinity for hemi-methylated DNA *in vivo*. Targeted deletion of the *Dnmt1* gene in mice results in global de-methylation causing mid-gestational lethality, growth retardation and dysregulation of some, but not all, imprinted genes (Caspary *et al.* 1998; Lei *et al.* 1996; Li *et al.* 1993; Li *et al.* 1992). *Xist* repression in visceral endoderm is unaffected in *Dnmt1-/-* embryos, suggesting that Dnmt1 is not required for imprinted X

chromosome inactivation. Random X inactivation, however, is unstable suggesting that *Dnmt1* expression may be necessary to maintain random X inactivation in the embryo (Sado *et al.* 2000; Vasques *et al.* 2005). *Dnmt1* null ES cells are viable although they contain low levels of DNA methylation probably because there are other Dnmts in ES cells that are responsible for *de novo* methylation (Lei *et al.* 1996; Li *et al.* 1992).

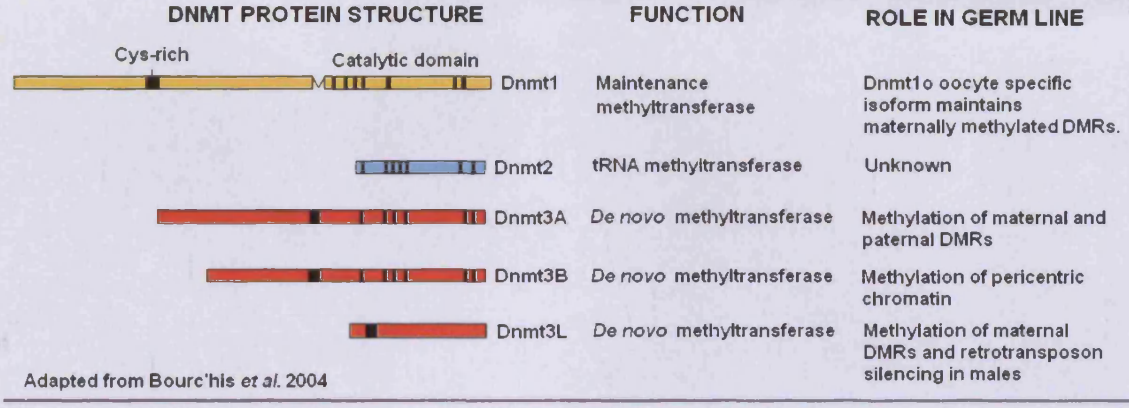


Figure 1.5: Function of DNA methyltransferases (Dnmts) in establishing methylation marks at DMRs in germ cells.

Dnmt1 contains both the catalytic domain, responsible for enzymatic activity of the enzyme, and a cysteine rich domain of no known function. Dnmt2, Dnmt3a and Dnmt3b all share the catalytic domain however, Dnmt2 lacks the cysteine rich domain. Dnmt3L has no catalytic activity and is thought to direct methylation of other Dnmts to specific targets. Dnmts are expressed during gametogenesis at stages where imprints are established and have different functions depending on whether they are expressed in males or females. Vertical bars inside the catalytic domain correspond to highly conserved motifs found in most DNA methyltransferases.

Dnmt1 is known to interact with histone modifying enzymes and methyl CpG binding proteins (Fuks *et al.* 2000; Kimura and Shiota 2003; Robertson *et al.* 2000). It also interacts with the *de novo* DNA methyltransferases Dnmt3a and Dnmt3b suggesting that Dnmt1 may play a role in establishing as well as maintaining methylation (Margot *et al.* 2003). *Dnmt3b* and *Dnmt1* double knockout cells have an approximately 95 % reduction in the 5meC content compared with only 20 % reduction seen in the single *Dnmt1* knockout and less than 3 % seen in the *Dnmt3b*

knockout (Rhee *et al.* 2002). This leads to reactivation of normally silent genes such as $p16^{INK4a}$, *TIMP-3* and *TWIST* suggesting that Dnmt3b and Dnmt1 co-operate to maintain global patterns of DNA methylation (Leu *et al.* 2003; Rhee *et al.* 2002).

An oocyte-specific isoform of Dnmt1 (Dnmt1o) has been shown to accumulate in the nucleus at the beginning of oocyte growth (Howell *et al.* 2001; Mertineit *et al.* 1998). In the pre-implantation embryo, Dnmt1o enters the nucleus transiently at the eight-cell stage and maintains the maternal methylation pattern (Mertineit *et al.* 1998; Ratnam *et al.* 2002). Deletion of the oocyte-specific isoform of *Dnmt1* does not affect global methylation patterns yet results in reactivation of some of the normally silent imprinted genes probably due to the loss of maternal methylation marks (Bourc'his *et al.* 2001; Howell *et al.* 2001; Li *et al.* 1992). This indicates that the oocyte specific isoform of Dnmt1 is a maintenance methyltransferase required to specifically maintain imprinted loci in oocytes and early embryonic development (Cirio *et al.* 2008; Howell *et al.* 2001).

Dnmt2 is a conserved gene found in eukaryotes (Jeltsch *et al.* 2006). For a long time the biological function of Dnmt2 was unknown. It has DNA methyltransferase activity but ES cells null for *Dnmt2* are viable and show normal patterns of methylation (Goll *et al.* 2006; Okano *et al.* 1998b). Recent evidence now suggests that Dnmt2 has a role as a novel tRNA methyltransferase. Dnmt2 methylates $tRNA^{asp}$ with high affinity, however the biological function of this is still unclear (Goll *et al.* 2006). It is possible that Dnmt2 plays a role in RNA interference and remodelling of chromatin as Dnmt2 activity has been associated with these two processes in *Dictyostelium* (Jeltsch *et al.* 2006).

There are two '*de novo*' DNA methyltransferases, Dnmt3a and 3b, that are expressed in the germ line and are just two of the factors known to establish

methylation patterns in germ cells (Kaneda *et al.* 2004; Kato *et al.* 2007; Okano *et al.* 1999). Dnmt3a and Dnmt3b are *de novo* methyltransferases as they are responsible for the addition of methyl groups on unmethylated DNA. They are essential for *de novo* DNA methyltransferase activity in murine ES cells and in early embryos (Chen *et al.* 2003; Okano *et al.* 1999).

Mice with a targeted deletion of *Dnmt3a* appear normal at birth but become growth restricted and die at about 4 weeks of age (Hiura *et al.* 2006; Kaneda *et al.* 2004). Conditional knockouts of *Dnmt3a* in germ cells show that it plays a role in DNA methylation during gametogenesis (Kaneda *et al.* 2004). Offspring of female conditional mutants die *in utero* with an absence of DNA methylation and show loss of maternal allele-specific expression of *Cdkn1c* and *Igf2r* and biallelic expression of paternally expressed genes *Peg1, Peg3* and *Snrpn* (Hiura *et al.* 2006; Kaneda *et al.* 2004). *Dnmt3a* conditional mutant males show impaired spermatogenesis and loss of methylation at the *Igf2-H19* DMR, IG-DMR of the *Dlk1-Gtl2* locus and at a short interspersed repeat SINE B1 suggesting that Dnmt3a is required for methylation of these loci (Kaneda *et al.* 2004). The results of these conditional knockouts are similar to the phenotype seen in the Dnmt3L knockout mice suggesting that both Dnmt3L and Dnmt3a are essential for spermatogenesis (Bourc'his *et al.* 2001; Kaneda *et al.* 2004; Nimura *et al.* 2006; Webster *et al.* 2005).

Dnmt3b knockout mice are embryonic lethal. However, the offspring of mice with a conditional knockout of *Dnmt3b* in their germ cells showed no apparent phenotype (Kaneda *et al.* 2004) except for a reduction in methylation at satellite repeats (Kato *et al.* 2007). Dnmt3a and Dnmt3b are structurally very similar and may show some redundancy in their function indicated by the double knockout ES cells which show a more severe global demethylation phenotype (Bourc'his *et al.* 2001;

Okano *et al.* 1999). ICF is a rare recessive human disorder resulting from a mutation in *DNMT3B* and causes immunodeficiency, centromeric instability and facial abnormalities (Ehrlich *et al.* 2006; Hansen *et al.* 1999; Vasques *et al.* 2005). The disorder is characterised by hypomethylation of satellite regions and subsequent chromosome instability (Ehrlich *et al.* 2006; Hansen *et al.* 1999; Xu *et al.* 1999). Mutation in the orthologous gene in mice is lethal but not in humans which is probably due to a reduced effectiveness of the protein in humans as a result of mis-sense or in-frame mutations rather than from a complete loss of the functional protein (Hansen *et al.* 1999). Patients with ICF have normal expression of DNMT3A suggesting that some functions are unique to DNMT3B. These two proteins also exist as multiple isoforms in humans and mice so may have a greater diversity of function (Chen *et al.* 2005; Xie *et al.* 2006).

Dnmt3L is another member of the Dnmt3 family but lacks the catalytic domain found in Dnmt3a and Dnmt3b. The catalytic domain is required for activation of the target cytosine, binding of the methyl donor s-adenosyl L-methionine and for sequence recognition (Aapola *et al.* 2000; Bourc'his *et al.* 2001). Dnmt3L shows sequence homology in the cysteine rich region of the protein and is thought to function as a co-adaptor protein to recruit other Dnmts to locally induce DNA methylation (Aapola *et al.* 2000; Bourc'his *et al.* 2001) (**figure 1.5**). In the mouse, Dnmt3L is expressed at high levels in both the testis and ovary and correlates with stages of gametogenesis where methylation at DMRs of imprinted genes are established, indicating that Dnmt3L is involved in establishing DNA methylation during gametogenesis (Aapola *et al.* 2004; Aapola *et al.* 2000; Bourc'his *et al.* 2001).

Recent studies have shown that Dnmt3L is essential for the establishment of correct DNA methylation at maternal DMRs in oocytes (Bourc'his *et al.* 2001;

Bourc'his and Bestor 2004; Hata *et al.* 2002; Lucifero *et al.* 2007). Dnmt3L is only expressed in females in growing oocytes arrested at the diplotene stage of meiosis I (Bourc'his *et al.* 2001; Bourc'his and Bestor 2004). In female mice with a targeted deletion of *Dnmt3L (Dnmt3L-/-),* oogenesis proceeds as normal, however, the maternal methylation patterns are not correctly established in the oocytes and, as a result, the heterozygous progeny of these mice die at mid-gestation (~9.5 dpc) (Arima *et al.* 2006; Bourc'his *et al.* 2001; Bourc'his and Bestor 2004). It was suggested that death was a result of having a non-functional placenta caused by an imprinting defect (Arima *et al.* 2006). Dnmt3L binds and co-localises with Dnmt3a and Dnmt3b in the nuclei of mammalian cells (Hata *et al.* 2002; Margot *et al.* 2003). It is possible that Dnmt3L could direct methylation to ICRs by interacting with Dnmt3a which is known to put methyl groups on ICRs in germ cells (Margot *et al.* 2003). *Snrpn* and *Igf2r* ICRs are both maternally methylated DMRs that require Dnmt3a activity (Chedin *et al.* 2002). The *Snrpn* ICR is a good target for *de novo* methylation by a Dnmt3a and Dnmt3L complex as there is no evidence of methylation in the absence of Dnmt3a and very little methylation in the presence of Dnmt3a alone. When Dnmt3a is co-expressed with Dnmt3L there is an increase in *de novo* methylation at the ICRs indicating that Dnmt3L has a role in stimulating *de novo* methylation by Dnmt3a (Chedin *et al.* 2002; Chen *et al.* 2005; Hata *et al.* 2002). However, it has also been shown that, in the absence of Dnmt3L, Dnmt3a can also act alone to mark other DMRs (Arnaud *et al.* 2006). It has been shown that Dnmt3L physically interacts with a variant of Dnmt3a (Dnmt3a2) in ES cells and embryonic testes, localising Dnmt3L to the nucleus (Jia *et al.* 2007; Nimura *et al.* 2006). ES cells derived from *Dnmt3L-/-* mutant embryos show loss of maternal germ line methylation patterns which leads to biallelic expression of some maternally silenced genes, such as *Kcnq1ot1/Lit1* and

Snrpn, and repression of several maternally expressed genes including *Cdkn1c, Mash2* and *Igf2r* (Arima *et al.* 2006; Bourc'his *et al.* 2001; Hata *et al.* 2002). Those genes regulated by paternal methylation, such as *Igf2* and *H19*, show normal monoallelic expression (Arima *et al.* 2006; Bourc'his *et al.* 2001).

During spermatogenesis, Dnmt3L is required for the acquisition of methylation at transposable elements, some paternally methylated DMRs, unique non-pericentric heterchromatic sequences and interspersed repeats (Bourc'his and Bestor 2004; Kato *et al.* 2007; Webster *et al.* 2005). Dnmt3L is expressed in non-dividing pro-spermatogonia between 12.5 dpc and birth and declines when prospermatogonia have differentiated into dividing prospermatogonial stem cells (Bourc'his and Bestor 2004). Male mice that lack Dnmt3L are sterile because of a complete absence of germ cells in the adults although there is a normal germ cell complement at birth. This suggests that there is progressive loss of prospermatogonia during development into adulthood (Bourc'his and Bestor 2004; Webster *et al.* 2005). The DNA from mutant germ cells is less methylated than that from wild type germ cells. Methylation is lost specifically from the LINE-1 and IAP retrotransposons whereas the major and minor satellite repeats remain unaffected. This indicates that in male germ cells, Dnmt3L is responsible for methylation of dispersed repeats within euchromatin, possibly by recruiting Dnmt3a, whereas pericentric tandem repeats are methylated by a different DNA methyltransferase (Bourc'his and Bestor 2004; Kato *et al.* 2007). This probably accounts for the genomic instability seen in mutant male germ cells that results in sterility (Webster *et al.* 2005). The DMR of the *Dlk1-Gtl2* imprinting cluster did not show loss of methylation in male mutant germ cells and only partial demethylation was seen at the *H19* DMR indicating that, in males, Dnmt3L is not involved in establishing these paternal germ line DMRs but is probably involved in establishing

life-long silencing of retrotransposons in pro-spermatogonia (Bourc'his and Bestor 2004; Kato *et al.* 2007). In female germ cells, meiosis and the methylation of repeat sequences is normal indicating that Dnmt3L plays different roles in male and female germ cell development (Bourc'his *et al.* 2001).

It is interesting to note that *Dnmt3L* is not found in chickens and fish but is found in species in which imprinting occurs suggesting that this gene may have evolved alongside genomic imprinting (Yokomine *et al.* 2006). As well as having a role in *de novo* methylation of DMRs, Dnmt3L may be an important factor for recruiting histone and non-histone proteins as studies have shown that Dnmt3L can repress transcription by binding directly to HDAC1 through its ATRX domain (Aapola *et al.* 2002; Aapola *et al.* 2000; Deplus *et al.* 2002).

1.2.4 Chromatin modifications

DNA exists in a highorder structure within each cell, tightly packaged around nucleoproteins (Kornberg 1974; Olins and Olins 1974). This sophisticated structure allows for storage, usage and replication of DNA. To allow transcription of DNA, chromatin must be dynamic to allow access for the transcriptional machinery. The basic repeating unit is the nucleosome consisting of an octamer of histone proteins – a central tetramer of H3 and H4 and two dimers of H2A and H2B (Arents *et al.* 1991; Finch *et al.* 1977; Stein *et al.* 1977) (**figure 1.6**). Approximately 147 bp of super helical DNA wrap around each nucleosome twice and each nucleosome is separated from the next by a 10-60 bp 'linker' region of DNA to form the 11 nm bead like structure (Luger *et al.* 1997; McGhee *et al.* 1983). The DNA is packaged further by the addition of linker histone H1 (Misteli *et al.* 2000; Pennings *et al.* 1994).

The bulk chromatin structure can be altered by the covalent modification of the N-terminal tails of core histones. Modification of the histone tails is important for

regulating a variety of processes including transcription, mitosis and the formation of heterochromatin (Hebbes *et al.* 1988; Jeppesen and Turner 1993; Lehnertz *et al.* 2003; Soejima *et al.* 2004; Wei *et al.* 1999). Modifications include acetylation, methylation, phosphorylation and ubiquitination; a combination of which makes up the histone code (Chen *et al.* 1998; Hebbes *et al.* 1988; Strahl *et al.* 1999; Wei *et al.* 1999) [reviewed in (Jenuwein and Allis 2001; Kouzarides 2007)].

Acetylation of specific lysine residues on the N-terminal tails of core histones can activate gene expression by promoting a more open chromatin conformation (Hebbes *et al.* 1988; Soejima *et al.* 2004), yet the methylation of specific lysine residues can promote either active or inactive regions of chromatin. For example, methylation of lysine 4 (K4) on H3 has been associated with gene activation, while methylation of lysine 9 (K9) on H3 is associated with heterochromatin (Lehnertz *et al.* 2003; Strahl *et al.* 1999). In imprinted regions, the active parental allele has an open chromatin structure while the inactive allele has a more heterochromatin-like conformation that is transcriptionally repressive and corresponds to a difference in histone acetylation status (Fournier *et al.* 2002; Fulmer-Smentek and Uta 2001; Gregory *et al.* 2001; Jeppesen and Turner 1993). It is likely that modifications to the chromatin surrounding imprinted genes are responsible for maintaining the imprints and also maybe involved in establishing them. Tri-methylation at Lysine-27 of histone H3 (H3K27Me3) and di-methyl Lys-9 of histone H3 (H3K9Me2) are markers of repressed chromatin whereas acetylation Lys-9 of histone H3 (H3K9Ac) and di-methyl Lys-4 of histone H3 (H3K4Me2) are markers of active chromatin (Cao *et al.* 2002; Fournier *et al.* 2002; Higashimoto *et al.* 2003; Strahl *et al.* 1999; Umlauf *et al.* 2004). These histone modifications have been identified around several imprinted genes including the paternally repressed *Kcnq1* and *Cdkn1c* (Fournier *et al.* 2002;

Higashimoto *et al.* 2003; Soejima *et al.* 2004; Umlauf *et al.* 2004). It may be that repressive histone modifications propagate along the imprinted cluster from the primary imprinting mark to silence the surrounding genes.

DNA methylation and methyl binding domain (MBD) proteins can recruit histone modifying enzymes and is evidence for a link between histone modifications and DNA methylation (Estève *et al.* 2006; Jones *et al.* 1998b; Nan *et al.* 1998). Chromatin is remodelled by a variety of remodelling enzymes such as histone methyltransferases (HMTs), histone acetylases (HAT) and deacetylases (HDACs) (Deplus *et al.* 2002; Fuks *et al.* 2000; Jones *et al.* 1998b; Wagschal *et al.* 2008; Zhang *et al.* 1999). Dnmts and MBD proteins may also play a role in RNA-directed silencing as these proteins have been shown to form complexes with RNA (Jeffery and Nakielny 2004).

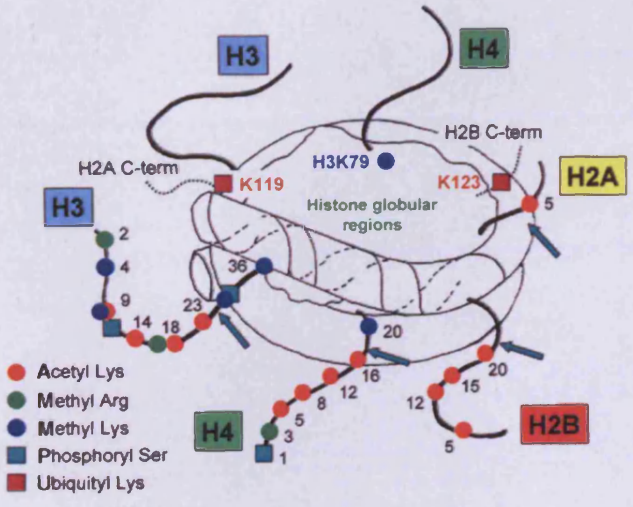


Figure 1.6: Structure of the nucleosome.

DNA is packaged around an octamer of histone proteins. The N-terminal tails of the histones can be modified to allow or prevent access to the DNA by transcriptional machinery. From (Turner 2005). Copyright Macmillan Publishers Limited.

1.2.5 Chromatin remodelling factors

Establishing the differential chromatin states on imprinted genes probably involves the recruitment of non-histone proteins. Polycomb group protein (PcG) complexes are thought to be required for maintaining long-term gene silencing during development by altering the local chromatin environment (Cao *et al.* 2002; Mager *et al.* 2003). PcG proteins play specialised roles in silencing genes in ES cells (Boyer *et al.* 2006; Mager *et al.* 2003) and are thought to modify histones by forming a more condensed chromatin structure, which is not easily accessible to the transcriptional machinery and so maintains long-term gene silencing during development (Boyer *et al.* 2006).

There are two distinct complexes of PcGs (PRC1 and PRC2) that are conserved from fruit flies to humans and occupy a different set of target genes (Boyer *et al.* 2006; Mager *et al.* 2003). Polycomb repressor complex 1 (PRC1) contains a minimal core complex of polycomb (Pc), polyhomeotic (Ph), posterior sex combs (Psc) and Ring1 that has the ability to block SWI/SNF remodelling and transcription *in vitro* (Breiling *et al.* 2004). Polycomb repressor complex 2 (PRC2) contains the H3K27/K9 methyltransferase Enhancer of Zeste Homologue 2 (Ezh2 or Enx1 in the mouse) (Breiling *et al.* 2004; Cao *et al.* 2002), the histone-lysine methyltransferase zinc finger protein Suz12 (Pasini *et al.* 2004) and the Embryonic ectoderm development (Eed) protein (Mager *et al.* 2003). These proteins were identified in a complex that co-localised with histone methyltransferase activity in humans and is specific for H3K27Me3 (Cao *et al.* 2002). PcG targets include a number of transcription factors that have important roles in development and lineage commitment (Boyer *et al.* 2006; Lan *et al.* 2007). During ES cell differentiation, the majority of the PcG targets are upregulated indicating that the PcG complexes are

required to repress genes that are poised for activation on differentiation (Boyer *et al.* 2006).

Eed has been shown to be required for the H3K27 di- and tri-methylation directed by the PRC2 complex (Boyer *et al.* 2006; Cao *et al.* 2002; Montgomery *et al.* 2005). The PRC2 complex also has the ability to interact with histone deacetylases and to co-localise with *Xist* on the imprinted X chromosome (Kalantry and Magnuson 2006; Kalantry *et al.* 2006). It has been proposed that antisense non-coding RNA molecules, such as *Air* and *Kcnq1ot1/Lit1,* have the same ability to recruit histone modifying enzymes and PRCs to silence imprinted genes. For example, the paternally expressed non-coding RNA *Kcnq1ot1/Lit1* may have the ability to recruit H3-K9 and K27 methylation along the adjacent repressed paternal allele of mouse distal chromosome 7. There is also evidence that the PRC2 complex can interact with and recruit the PRC1 complex. PRC1 binding can lead to the formation of a repressive chromatin state by preventing access of the nucleosome remodelling factors SWI/SNF (Cao *et al.* 2002).

Montgomery *et al.* (Montgomery *et al.* 2005) showed that Eed expression is also responsible for establishing global H3-K27 mono-methylation independently of the PRC2 complex suggesting that Eed may play other roles in gene silencing. During stem cell differentiation, Ezh2 and Eed levels decrease whereas Suz12 levels remain constant. Suz12 is also enriched on the imprinted X chromosome in the early stages of chromosome inactivation independent of *Xist* expression suggesting that Suz12, like Eed, may play other roles in gene silencing independent of the PRC2 complex (de la Cruz *et al.* 2005).

It is likely that differential DNA methylation and histone modifications work together to establish imprinting across large domains. The link between DNA

methylation and histone modifications has been shown in a number of studies (Fuks *et al.* 2000; Kimura and Shiota 2003; Robertson *et al.* 2000). Heterochromatin protein 1 (HP1) family members have been shown to mediate communication between histone modifications and DNA methyltransferases. H3K9me3 creates a binding platform for HP1 which then interacts with Dnmt1 to locally increase DNA methylation (Smallwood *et al.* 2007).

1.2.6 Mouse distal chromosome 7 IC2 domain

The mouse distal chromosome 7 imprinting cluster consists of two independently regulated imprinting domains (**figure 1.7**) known as IC1 and IC2 (Cerrato *et al.* 2005). The centromeric cluster (IC2) of mouse distal chromosome 7 contains a number of genes imprinted in both the embryo and placenta (Caspary *et al.* 1998), including *Kcnq1* (*Kvlqt1*), a primary channel-forming α-subunit of voltage gated potassium channels (Gould and Pfeifer 1998); *Slc22a18* (*Impt1*) a poly-specific organic cation transporter (Dao *et al.* 1998); *Phlda2* (*Impl*), a growth gene involved in placental development (Frank *et al.* 2002); and *Cdkn1c* ($p57^{KIP2}$), a cyclin dependent kinase inhibitor (Hatada and Mukai 1995). There is also an unspliced and untranslated RNA, known as the *Kcnq1* overlapping transcript 1 (*Kcnq1ot1*) or the long intronic transcript 1 (*Lit1*) that originates from a DMR within intron 10 of the *Kcnq1* gene (Smilinich *et al.* 1999). The DMR (KvDMR1) acts as the ICR for the cluster which was shown by the deletion of this region resulting in loss of imprinting of at least six imprinted genes within the IC2 domain, including *Kcnq1* and *Cdkn1c,* in both mice and humans (Fitzpatrick *et al.* 2002; Niemitz *et al.* 2004).

On the maternal allele, the KvDMR1 is methylated which is presumed to prevent transcription of *Kcnq1ot1/Lit1*. The surrounding genes are expressed. On the paternal allele the KvDMR1 is not methylated, *Kcnq1ot1/Lit1* is expressed and the

surrounding genes are silent. How the KvDMR1 affects silencing of the surrounding genes is unknown but may be similar to the non-coding RNA model seen for *Igf2r* (Wutz *et al.* 1997). The *Kcnq1ot1/Lit1* RNA transcript may localise histone modifying enzymes to the surrounding genes as seen with *Xist* and imprinted X-chromosome inactivation. The results of truncating the *Kcnq1ot1/Lit1* transcript indicates that it is not the germ line methylation mark at KvDMR1 but the RNA transcript that is the key to silencing within this cluster (Shin *et al.* 2008).

The *Cdkn1c* gene contains a large CpG island that is differentially methylated in the mouse but not in the human (Onyango *et al.* 2000). This may account for the low levels of expression (~10 %) from the imprinted paternal allele observed in humans but not in mice (Chung *et al.* 1996; Matsuoka *et al.* 1996). *Cdkn1c* promoter methylation is a somatic methylation mark acquired in embryos during post-implantation development and is not present in the germ line (Bhogal *et al.* 2004; Tada *et al.* 1998). It may be that the methylation mark on the mouse *Cdkn1c* CpG island serves to lock down silencing in mice whereas imprinting of *CDKN1C* in humans is more reliant on other epigenetic marks.

Mouse distal chromosome 7 is homologous to the human 11p15 imprinting domain (Engemann *et al.* 2000; Onyango *et al.* 2000). This imprinting domain shows a high degree of conservation between the two species including sequence, gene order, gene orientation and imprinting (Engemann *et al.* 2000; Paulsen *et al.* 2000). However, around and between the two domains lie genes, that are also regulated by KvDMR1, but where imprinting is less conserved and show differences between mice and humans (Engemann *et al.* 2000; Paulsen *et al.* 2000). *Osbpl5 (Opbh1), Nap1L4 (Nap2), Cd81, Tssc4* and *Ascl2* (*Mash2*) are imprinted in mice but expressed biallelically in human foetal tissues and placenta (Beatty *et al.* 2006; Monk *et al.*

2006). *Trpm5,* a gene encoding for a taste receptor, is imprinted in humans but biallelically expressed in the mouse (Beatty *et al.* 2006). Conserved imprinted expression is seen for *Phlda2*, *Slc22a18* and *Cdkn1c* and these genes remain imprinted throughout development (Monk *et al.* 2006). Imprinting of *Kcnq1* and *Kcnq1ot/Lit1* is also conserved and are imprinted in early foetal tissues and first trimester placenta but imprinting is relaxed in term placenta and late juvenile and adult stages (Caspary *et al.* 1998; Gould and Pfeifer 1998; Monk *et al.* 2006).

Imprinting in the IC2 domain is unaffected by loss of imprinting in the IC1 domain. Mice lacking *H19* show normal imprinting within the IC2 domain indicating that these two imprinting domains act independently.

Cdkn1c is maternally expressed in mice and *Cdkn1c* null mice show a variety of phenotypes similar to BWS including renal dysplasia, lens and gastrointestinal defects and skeletal abnormalities (Caspary *et al.* 1999; Zhang *et al.* 1997). Contrary to what is seen in humans, loss of *Cdkn1c* is embryonic lethal with approximately 10 % of offspring dying *in utero* and the remainder dying within the first two days of birth (Takahashi *et al.* 2000; Yan *et al.* 1997; Zhang *et al.* 1997). This phenotype can be rescued by deleting the KvDMR1 ICR from the paternal chromosome resulting in expression of paternal *Cdkn1c* (Horike *et al.* 2000; Mancini-DiNardo *et al.* 2006). Loss of imprinting of *IGF2* in humans is also associated with Beckwith Wiedemann syndrome. Mice with loss of imprinting of *Igf2* are born approximately 30 % bigger than their wild type litter mates whereas the *Cdkn1c* mutants do not show any overgrowth at birth (Caspary *et al.* 1999; Leighton *et al.* 1995; Zhang *et al.* 1997). However, we have recently shown that loss of *Cdkn1c* is associated with overgrowth at an earlier time point, 13.5 dpc, so this growth advantage may be lost during embryonic development (Andrews *et al.* 2007). *Igf2* is paternally expressed and it

may be that *Cdkn1c* and *Igf2* interact and have an antagonistic effect in the tissues in which they are co-expressed (Caspary *et al.* 1999).

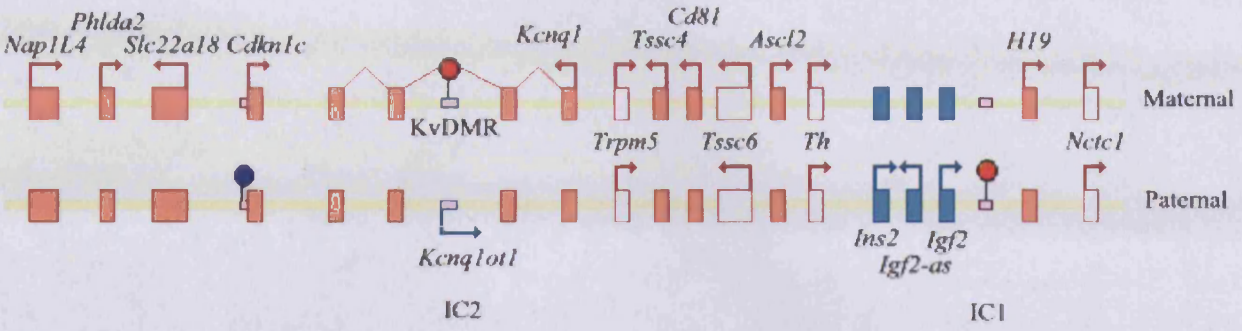


Figure 1.7: Mouse distal chromosome 7 imprinting domain.

Map of the imprinted genes within mouse distal chromosome 7 showing their relative position and allelic origin of expression. This imprinting domain contains two imprinted clusters IC1 and IC2. Within these clusters the imprinted genes are co-ordinately regulated by differentially methylated regions (DMRs). Germ line methylation marks are acquired during gametogenesis (red lolly pop). Within the IC2 domain there is a somatic differentially methylated mark (blue lolly pop) and is located in the promoter region of *Cdkn1c*. Orange boxes represent maternally expressed genes, turquoise boxes represent paternally expressed genes, white boxes are non-imprinted genes. CpG are shown in pink. Arrows indicate gene expression.

1.2 Epigenetic reprogramming in the germ line

1.3.1 Gametogenesis and imprinting

Mouse primordial germ cells (PGCs) develop from the pluripotent epiblast and can be detected as alkaline phosphatase positive staining cells at around 7.5 dpc (Ginsburg *et al.* 1990; Matsui *et al.* 1992; Yamazaki *et al.* 2005). These cells rapidly proliferate and migrate via the hindgut and dorsal mesentery at around 8.5-9.5 dpc. They reach the genital ridge at around 10.5-11.5 dpc where they continue to proliferate until 13.5 dpc (Durcova-Hills *et al.* 2001; Matsui *et al.* 1992; Tam and Snow 1981) (**figure 1.8**). The parental imprints are erased in early PGCs during migration and colonisation into the early embryonic gonad [reviewed in (Reik *et al.* 2001)] where there is also reactivation of the silent X-chromosome (Kato *et al.* 1999;

Monk *et al.* 1987). Loss of methylation at the DMRs of the imprinted genes correlates with the loss of monoallelic gene expression in these cells (Szabo and Mann 1995).

On colonizing the genital ridge, female PGCs enter meiotic prophase and meiotic arrest at around 13.5 dpc, whereas the male PGCs remain mitotic until approximately 15.5 dpc when they enter a state of mitotic arrest (Hajkova *et al.* 2002; Hiura *et al.* 2006). The new imprints on the ICRs are then established in a sex-specific manner. In oocytes, methylation occurs postnatally during the growth phase of the oocytes during follicular development (Hiura *et al.* 2006), whereas in sperm, re-methylation takes place earlier before meiosis in late foetal stages (Kerjean *et al.* 2000; Ueda *et al.* 2000).

It is not yet understood how the germ line specific imprints are established and what the signals are that determine which parental imprint is to be set. *Igf2*, *H19*, *Rasgrf1*, *Dlk1* and *Gtl2* have been identified as genes which are epigenetically modified during spermatogenesis to give allele-specific gene expression in the embryo (Hiura *et al.* 2006; Kerjean *et al.* 2000). However, the majority of the germ line methylation is maternal and is established during oocyte growth (Obata *et al.* 1998). This corresponds with the observation that *de novo* Dnmts are expressed throughout the oocyte growth phase (Lucifero *et al.* 2007; Lucifero *et al.* 2004). In the early embryo there is genome wide de-methylation and re-methylation (Monk *et al.* 1987; Oswald *et al.* 2000; Santos *et al.* 2002), yet the germ line methylation marks of the imprinted genes are maintained (Kato *et al.* 1999; Lei *et al.* 1996). Imprinted genes may be protected from de-methylation by the nucleosomal structure in these regions (Kafri *et al.* 1993; Oswald *et al.* 2000). For example, the female pronucleus becomes de-methylated more slowly which may protect the oocyte imprints (Bouniol *et al.* 1995; Ram and Schultz 1993; Santos *et al.* 2002). Undifferentiated ES cells have

a relatively high *de novo* DNA methylation activity but this activity is low in the post-gastrulation embryo and adult somatic tissues (Monk *et al.* 1987).

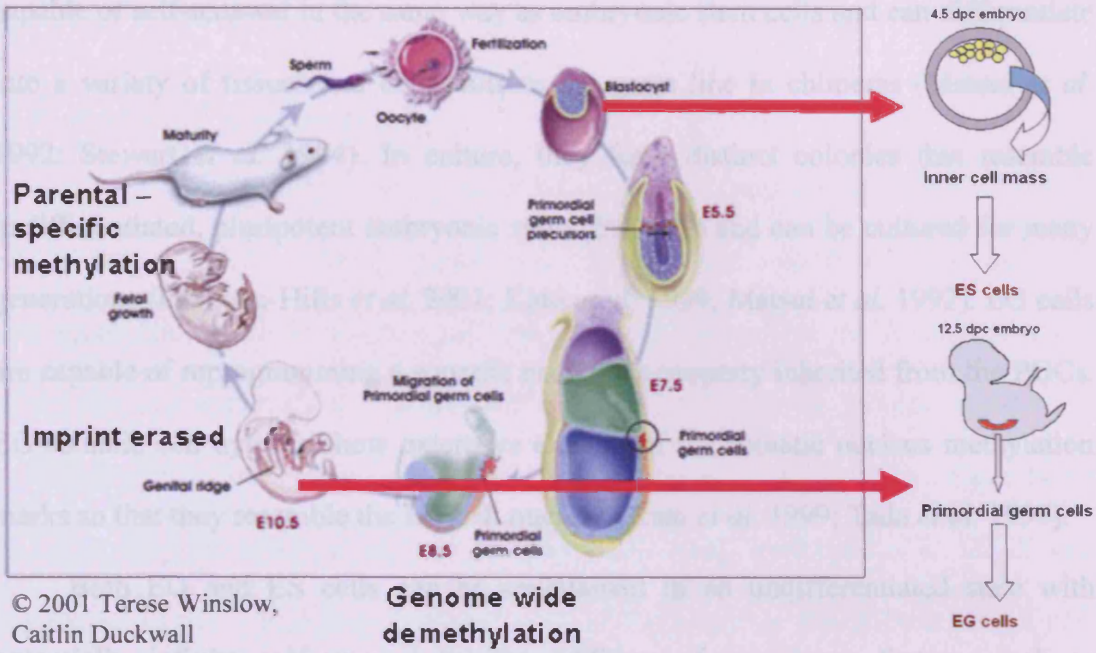


Figure 1.8: Gametogenesis in the mouse and establishing parental specific methylation patterns in the germ line.

Imprints are established in the germ line during gametogenesis. As the PGCs migrate to the genital ridge they undergo genome wide de-methylation. At 10.5 – 11.5 dpc imprints are erased as PGCs colonise the genital ridge, they then proceed down either the male or female germ line to acquire parental-specific methylation patterns. Primordial germ cells can be cultured as embryonic germ line stem cells (EG) in the same way as Embryonic stem cells (ES).
Adapted from (In Stem cell information [world wide web site]. Bethesda, MD: National Institutes of Health, U.S. Department of Health and Human Services, 2006 [cited 24/07/2008] Available at http://stemcells.nih.gov/info/scireport/appendixa).

1.3.2 Culturing EG cells

Cells isolated from the inner cell mass (ICM) of blastocyst stage embryos can be cultured *in vitro* as embryonic stem (ES) cells (Evans and Kaufman 1981; Martin 1981). ES cells are similar to ICM cells, but do not undergo the rapid differentiation they would *in vivo,* and contain the imprints of both parental genomes (Dean *et al.* 1998; Evans and Kaufman 1981; Rugg-Gunn *et al.* 2005). PGCs can be isolated from

the developing gonadal ridges of 11.5 dpc and 12.5 dpc embryos and cultured *in vitro* as embryonic germ (EG) cells (Lee *et al.* 2002; Tada *et al.* 1998). These EG cells are capable of self-renewal in the same way as embryonic stem cells and can differentiate into a variety of tissues and contribute to the germ line in chimeras (Matsui *et al.* 1992; Stewart *et al.* 1994). In culture, they form distinct colonies that resemble undifferentiated, pluripotent embryonic stem (ES) cells and can be cultured for many generations (Durcova-Hills *et al.* 2001; Kato *et al.* 1999; Matsui *et al.* 1992). EG cells are capable of reprogramming a somatic nucleus, a property inherited from the PGCs. EG-somatic cell hybrids show extensive erasure of the somatic nucleus methylation marks so that they resemble the EG cell nucleus (Kato *et al.* 1999; Tada *et al.* 1997).

Both EG and ES cells can be maintained in an undifferentiated state with potentially infinite self-renewal by the addition of exogenous factors, such as leukaemia inhibitory factor (LIF), to the culture media (Labosky *et al.* 1994; Matsui *et al.* 1992; Resnick *et al.* 1992) and express stem cell markers such as Oct-4, SSEA-1 and endogenous alkaline phosphatase (Durcova-Hills *et al.* 2001; Rohwedel *et al.* 1996). XY EG cell lines can be established much more frequently than XX cell lines. XX cell lines are unstable, probably due to the loss of imprinting of one of the X chromosomes (Durcova-Hills *et al.* 2001; Zvetkova *et al.* 2005). The inactive X chromosome is reactivated in PGCs at around 12.5 – 13.5 dpc and both chromosomes continue to be active throughout oogenesis. In the male germ line the single X chromosome is transcriptionally inactivated and remains silent until after fertilisation in female pre-implantation embryos (Salido *et al.* 1992). EG and ES cells can be differentiated, as embryoid bodies, into diverse cell types and can be used to investigate how, when and where the early imprints in the two parental genomes and other epigenetic modifications are established (Doetschman *et al.* 1985; Labosky *et al.*

1994; Rohwedel *et al.* 1996). Studies have shown that germ line imprints are epigenetically stable in ES cells even after prolonged culture (Rugg-Gunn *et al.* 2005; Szabo and Mann 1994).

1.4 Aims of the project

The precise sequence of epigenetic events that result in allele-specific gene expression is unclear. Several studies have looked at the establishment of the germ line epigenetic marks in oocytes and spermatocytes as well as the final imprints in stem cells and somatic cells. Much is also known about these epigenetic marks in migrating primordial germ cells at specific time points such as prior to imprint erasure (9.5 dpc) (Durcova-Hills *et al.* 2001; Szabo *et al.* 2002; Yamazaki *et al.* 2005) and post-erasure (11.5 – 12.5 dpc) (Hajkova *et al.* 2002; Lee *et al.* 2002; Tada *et al.* 1998; Ueda *et al.* 2000). Very little is known about the sequence of events that lead to silencing of imprinted genes within the germ line at particular imprinted domains. The mouse distal chromosome 7 domain relies on several key elements to establish differential gene expression including DNA methylation, histone modifications and the transcription of the non-coding RNA, *Kcnq1ot1/Lit1*, all of which are required for imprinting of this domain. One particular gene within this domain is of particular interest. *Cdkn1c* requires a germ line DNA methylation mark at the imprint control region, KvDMR1, to be expressed and a somatic DNA methylation mark within the *Cdkn1c* locus for its silencing.

The aim of this study was to investigate the epigenetic mechanisms that lead to heritable silencing of *Cdkn1c* using an *in vitro* system. The work in this study was based on the observation that *Cdkn1c* was not expressed in EG cells derived from imprint-erased PGCs (Tada *et al.* 1998). These cells lack DNA methylation at all the germ line and somatic DMRs. In the study by Tada *et al.*, the authors derived EG cell

lines from 11.5 – 12.5 dpc PGCs (Tada *et al.* 1998). The DMRs of several imprinted genes within these cells were shown by the authors to be unmethylated. The EG cells were then aggregated with wild type embryos to create chimeras in order to determine how these cells contribute to the germ layers and to determine whether germ line transmission was possible. Primary embryonic fibroblasts (PEFs) were derived from 13.5 dpc chimeric embryos. Fibroblasts derived from the EG cells were selected for by G418 resistance. The authors examined the germ line DMRs of several genes including *Igf2r* and *Nnat* and determined that these regions are unmethylated even after differentiation *in vitro*. The exceptions were *H19*, which showed partial methylation, and *Cdkn1c*, which showed *de novo* methylation of its somatic DMR. The expression levels of the genes were also examined. No expression of *Igf2r, H19* and *Cdkn1c* was observed in both the EG cells and the differentiated PEFs. The results of this study appeared to indicate that *Cdkn1c* acquires *de novo* DNA methylation in the absence of germ line methylation of KvDMR1 suggesting that perhaps certain imprinted genes undergo dynamic and active silencing steps in the absence of a germ line imprint so providing an *in vitro* system in which these events could be studied.

The aims of this project were:

- to examine the expression status of *Cdkn1c* and other genes within the distal 7 imprinting domain and to determine if these genes were actively silenced in EG cell lines after *in vitro* differentiation.
- to determine whether *Cdkn1c* underwent *de novo* DNA methylation after differentiation, similar to the *in vivo* situation.
- to accurately map the spread of methylation across the *Cdkn1c* DMR over time.

- to examine histone modifications in undifferentiated and differentiated EG cells and compare these to androgenetic ES cells (containing two silent paternal alleles) in order to determine if alterations to the chromatin structure were associated with silencing.
- to identify other domains that establish silencing in a similar way to the genes on mouse distal chromosome 7.

To achieve these goals, real time quantitative RT-PCR assays were designed to amplify the mouse distal chromosome 7 genes including *Cdkn1c.* These assays were applied to cDNA prepared from somatic tissues and undifferentiated and differentiated ES and EG cells. DNA methylation assays were also developed for both the KvDMR1 region and the *Cdkn1c* DMR and were applied to genomic DNA isolated from somatic tissues and undifferentiated and differentiated ES and EG cells. This was to confirm that these regions were unmethylated in the undifferentiated EG cells and then to determine whether any changes occurred at different time points during the differentiation of these cells. Chromatin immunoprecipitation (ChIP) assays were designed to amplify the promoter regions of *Phlda2, Cdkn1c, Kcnq1* and *Kcnq1ot1/Lit1* in order to examine the epigenetic status of these regions in undifferentiated and differentiated EG and AG cells. Published ChIP primers were available, however these were designed to span polymorphisms between *M. Spretus* and 129/Sv, in order to apply allele-specific ChIP on *M.Spretus* x 129/Sv crosses, so were not necessarily within the promoter regions of imprinted genes. Designing primers to amplify the promoter regions of imprinted genes allowed us to specifically map the changes to the chromatin structure that possibly affect expression of the imprinted genes. The real-time quantitative RT-PCR assays were also used to identify

other imprinted genes that might be predicted to behave as *Cdkn1c* and these primers were applied to undifferentiated and differentiated stem cell material.

Through these assays, the project aimed to determine whether *Cdkn1c* is actively silenced by a series of epigenetic events that occur by default, in the absence of DNA methylation at KvDMR1, or whether the gene was already silent i.e. germ line DNA methylation and other active steps are required to switch the gene on. The project would also identify other genes that behave in a similar fashion.

Chapter 2:

Materials and Methods

2.1 Cell Culture

ES cell lines:

- CES3, XY diploid biparental ES cell lines derived from 129/sv ICM (gift from M A Surani).
- SF1-1, XY diploid biparental ES cell line derived from hybrid embryo [CBA/C57BL/6 F1 x *M. spretus*] ICM (gift from R Feil) (Allen *et al.* 1994; Feil *et al.* 1997).
- AKR1, XY diploid monoparental ES cell line derived from ICM of androgenetic embryo (gift from N Allen & M A Surani) (Allen *et al.* 1994).

EG cell lines:

- TMAS21G, XY diploid EG cell line derived from mixed background Rosa26 E12.5 genital ridge (Tada *et al.* 1998).
- Sv6.1, XY diploid EG cell line derived from genital ridge of a pure 129/Sv E12.5 embryo (Durcova-Hills *et al.* 2002).

Stem cells were cultured on inactivated SNLs (SNL 76/7, mouse SIM strain embryonic fibroblast), a mouse primary embryonic fibroblast STO cell line containing *Neomycin* and *Lif* transgenes, for several passages to maintain ES cells in their undifferentiated state before harvesting for DNA and RNA analysis. SNLs were inactivated by treating with 10 μg/ml Mitomycin C (Sigma) for 1 – 2 hrs. Mouse ES cells and SNLs were cultured on 0.1 % gelatin (Sigma) coated 10 cm tissue culture plates in Knockout DMEM (Invitrogen) supplemented with 10 % foetal calf serum (Globepharm), 2 mM L-glutamine (Invitrogen), 50 μg/ml Penicillin/Streptomycin (Invitrogen), 0.1 mM non-essential amino acids (Invitrogen), 1000 U/ml Leukaemia Inhibitory Factor (Chemicon Int) and 1 μM β-mercaptoethanol (Invitrogen). The

addition of Leukaemia Inhibitory Factor (LIF) to the culture media and the production of LIF from the SNL feeder layer is required for long term proliferation potential of mouse primordial germ cells in culture and works in a dose dependent manner with a peak response at 1000 U/ml (Resnick *et al.* 1992).

Media was changed approximately every 24 hours and cells were passaged. Stem cells were split on to new gelatin coated 10 cm plates or equivalent in to flasks. For DNA and RNA analysis, feeder cells were removed by panning, allowing the trypsinised feeder cells to settle down for 20 min then removing the medium containing the remaining stem cells. Stem cells were differentiated by the removal of LIF from the media and cultured without feeders in suspension on non-tissue culture petri dishes. Cells formed aggregates called embryoid bodies (**figure 2.1**) and were passaged by harvesting the suspension culture into 15 ml falcon tubes and allowing the embryoid bodies to settle out for 15 min. The top 5 ml of medium was aspirated and replaced with 5 ml of fresh medium and added to a new plate. Cells were pelleted, frozen on dry ice and stored at -80 °C for RNA or DNA analysis or used directly to isolate chromatin.

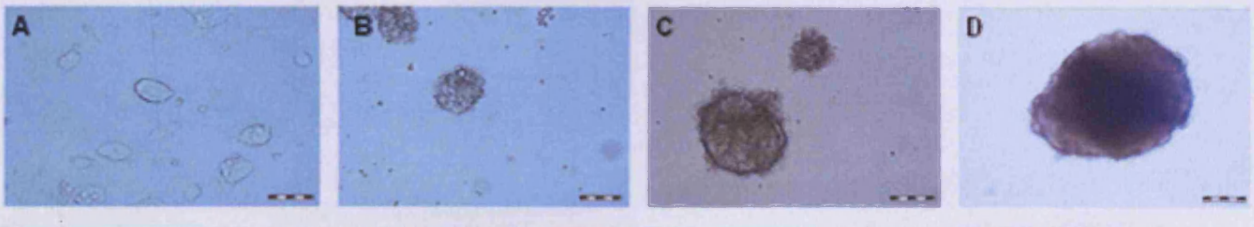


Figure 2.1: ES cells and embryoid bodies in culture.

Mouse ES and EG cells were cultured and differentiated under the same conditions. **A.** ES cells were cultured on an inactivated mouse fibroblast feeder layer in the presence of LIF. **B.** ES cells were differentiated for one day in the absence of the feeder layer and LIF on non-tissue culture plastic to form cell aggregates. **C.** After one week of culture the embryoid bodies appear as hollow balls of cells. **D.** Large embryoid bodies can be seen after two weeks of culture. Scale bars = 50 μM

2.2 Karyotyping

Normal karyotype was confirmed by analysing metaphase chromosomes. Stem cells were grown without feeders on gelatin coated plates until 90 % confluent then treated with ethidium bromide at a final concentration of 1.5 μg/ml two hours prior to harvesting. Colchicine (Sigma) was added to a final concentration of 0.1 μg/ml 80 min before harvesting. Cells were harvested and resuspended in 1 ml pre-warmed 37°C hypotonic solution containing 0.56 % Potassium Chloride. The cells were incubated for 5 min at 37 °C then fixed by the addition of 50 μl of a 3:1 solution of Methanol: Acetic Acid pre-cooled to 4 °C. Cells were centrifuged and resuspended in 1 ml of fixative and left overnight at 4 °C. Cells were centrifuged and fixed with 1 ml of fixative for two more times. The cell suspension was dropped on to clean humidified microscopic slides from a height of 2-3 feet to allow metaphase chromosomes to spread. Slides were air-dried then mounted with Vectorshield containing DAPI. Metaphase chromosomes were viewed by fluorescent microscopy. For each cell line, ten metaphase cells were counted and the average number of chromosomes determined. All cells scored showed the normal number of chromosomes.

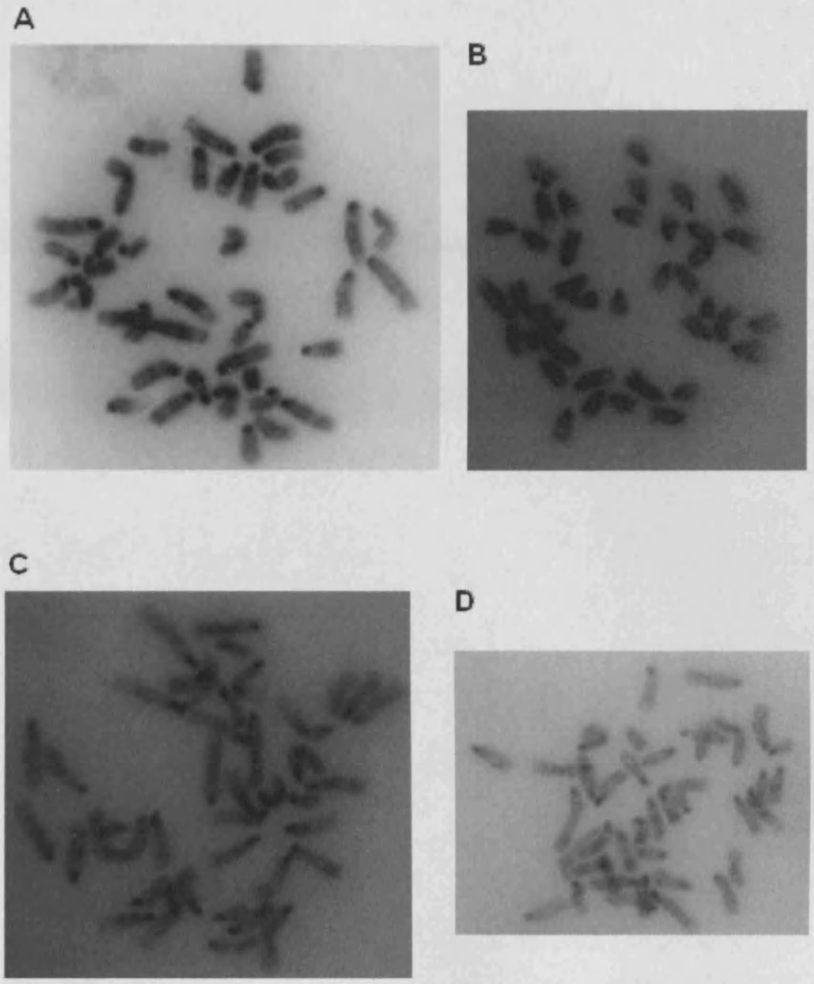


Figure 2.2: Karyotype analysis of ES, AG and EG cells.

Metaphase chromosomes were counted from (**A**) CES3, (**B**) Sv6.1, (**C**) TMAS21G, (**D**) AKR1 cells. Ten metaphase spreads were scored for each cell line and the average number of chromosomes was determined. All cell lines showed the normal complement of chromosomes.

2.3 Real Time Quantitative RT-PCR

RNA was isolated from whole embryos, placentae (homogenised using an eppendorf pestle) or stem cell pellets using Tri reagent (Sigma) according to the manufacturer's protocol. Precipitated RNA was resuspended in RNase free dH_2O and DNase I treated (Promega) at 37 °C for 1 hr according to the manufacturer's instructions. The DNAse I treated RNA was phenol-chloroform extracted and ethanol and sodium acetate precipitated. Purified RNA was quantified by spectrophotometry

and used to make cDNA. Two independent reactions were set up for reverse transcriptase treated and non-treated samples. A known concentration of RNA (5 μg for RT + or 1 μg for RT-) was incubated with 25 μg random primers (Promega) for 10 min then placed on ice. The reaction was then incubated with 5x M-MLV buffer (Promega), 20-40 units of RNase inhibitor (promega) and 0.5 mM dNTP mix (Roche) in 19 μl volume. Samples were incubated at 37 °C for 2 min. To the RT+ sample 100-200 units of M-MLV reverse transcriptase, RNase H minus (Promega) was added and 1 μl of dH_2O to the RT- sample as a negative control then incubated at 37 °C for 1 hr. Reaction was stopped by heating to 70 °C for 15 min. RT+ cDNA was diluted 1 in 20 with 10 mM Tris-HCl (pH7.5) and 1 in 4 for RT- cDNA. A PCR was performed on RT+ and RT- to confirm the absence of contaminating genomic DNA in the cDNA, using primers designed to amplify *Actin* or *GAPDH* cDNA. PCR reaction was performed using Sigma Taq DNA polymerase (SIGMA) according to the manufacturer's instructions. A basic PCR program was used: 95 °C for 3 min then 95 °C for 30 sec, 60 °C for 30 sec, 72 °C for 30 sec repeated for 35 cycles followed by 72 °C for 5 min.

Primers were designed using Primer3 (http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi) for several genes within the mouse distal chromosome 7 imprinting cluster. Primers used are shown in **table 2.1**. 12.5 μl of SYBR green reaction mix (MJ research) was added to 5 μl (5 mM stock of forward and reverse primers) of primers and 7.5 μl of cDNA (1 in 7.5 dilution). Reaction was loaded on to the white 96-well non-skirted plate (Abgene) in triplicate. At least two sets of control primers, either *GAPDH, Actin*, or *18s rRNA* were used for quantification of the samples. Data was collected by the Chromo4 real time machine (MJ research) and analysed using the Opticon monitor software v2.3 (MJ research). Fold changes were

calculated as described in Livak and Schmittgen, 2001 using the $2^{-\Delta\Delta C(T)}$ method (Livak and Schmittgen 2001).

Table 2.1: Real time quantitative RT-PCR primers

Primer Number	**Primer Name**	**Sequence**	**Tm**
EG36	Phlda2 forward	tcagcgctcgagtctgaaa	60
EG37	Phlda2 reverse	tcctgggctcctgtctgat	61
EG57a	Gtl2 forward	atgtgccaagttctgggttc	60
EG58	Gtl2 reverse	caagagctacgcattcacca	60
EG59	Rasgrf1 forward	tgatcgtatccaatccagca	60
EG60	Rasgrf1 reverse	ctgaacaccacctggttcct	60
EG61	Zac1 forward	aagtctcacgcggaagaaaa	60
EG62	Zac1 reverse	ctctgggcacagaactgaca	60
EG63	Nnat forward	agaaaagcagcaccgacaat	60
EG64	Nnat reverse	ggctgttcgatcttcatggt	64
EG65	Gnas ex1A forward	gttgcttcaggtggctggta	64
EG66	Nesp55 forward	gagggcccttagatcagga	59
EG67	Gnasxl forward	atccgagtgtacccgatcac	64
EG68	Gnas Exon2 reverse	gtgcttttgccagactctcc	64
EG71	Kcnq1 3'UTR forward	ggaacatagggatggggagt	60
EG72	Kcnq1 3'UTR reverse	gttccctgatggtctctgga	60
EG73	Rasgrf1A forward	tggagatcacctcctccatc	60
EG74	Rasgrf1A reverse	gtaattgggtgtcccctcct	60
EG75	Zac1 forward	gagcagaggaaggagcagaa	60
EG76	Zac1 reverse	cacaatgggcaacaagaaga	60
EG77	Hymai forward	ctttggccacaaaaattgct	60
EG78	Hymai reverse	aaccagaaactgggcagaga	59
R149	GAPDH forward	cacagtcaaggccgagaatg	62
R150	GAPDH reverse	tctcgtggttcacacccatc	62
R151	B actin forward	cctgtatgcctctggtcgta	59
R152	B actin reverse	ccatctcctgctcgaagtct	59
R161	Cdkn1c exon 3-4 forward	agagaactgcgcaggagaac	60
R162	Cdkn1c exon 3-4 reverse	tctggccgttagcctctaaa	60
R166	Slc22a18 forward	acctgtgtccctgctaccac	60
R167	Slc22a18 reverse	gatcccgaagaaggacatga	60
R168	IGF2 forward	gtcgatgttggtgcttctca	60
R169	IGF2 reverse	aagcagcactcttccacgat	60
R170	IGF2R forward	gttggtgtagggccagtgtt	60
R171	IGF2R reverse	agaaattctgccggggtact	60
R172	Dlk1 forward	cacctgggttctctggaaag	60
R173	Dlk1 reverse	aggggtacagctgttggttg	60
R174	Grb10 forward	ctgcaccacttcttgaggat	60
R175	Grb10 reverse	aactgctggtcttcctcctg	60
R190	LIT1 forward	tggaatcgggtagagattcg	60
R191	LIT1 reverse	agaccatcggaaaacacagg	60
R209	18s rRNA forward	catggccgttcttagttggt	64
R210	18s rRNA reverse	cgctgagccagtcagtgtag	64
R236	Nesp55 forward	ggagagtctggcaaaagcac	60
R237	Nesp55 reverse	tggggtaggacatagcgaag	60
R238	Nespas forward	cactgagtgtcctccaagca	60
R239	Nespas reverse	agaccccagcttctctcctc	60
R244	Ube3a forward	tgctgtcacaaagaatctgg	57
R245	Ube3a reverse	tcctccacaaccaactgaaa	59

2.3 Methylation-Sensitive Restriction Enzyme-Coupled PCR (MSRE-PCR)

Genomic DNA was isolated from stem cell pellets by digesting overnight in 1 ml lysis buffer (100 mM Tris-HCl (pH7.8), 50 mM EDTA, 0.2 % SDS, 200 mM NaCl) and 100 μg/ml proteinase K at 55 °C. DNA was then phenol-chloroform extracted and precipitated with ethanol and sodium acetate (pH6.0). Precipitated DNA was washed in 70 % ethanol and resuspended in 200 μl 10 mM Tris (pH7.5). DNA was stored at 4 °C until required.

Primers were designed to amplify genomic DNA around a methylation-sensitive restriction enzyme site. The list of primers is shown in **table 2.2**. Genomic DNA was digested either with *Eco*RI (NEB) alone or with a combination of *Eco*RI and the methylation-sensitive restriction enzyme *Eag*I or *Not*I (NEB) using the *Eco*R1 buffer recommended by New England Biolabs for double digests. 1 μg of genomic DNA was digested overnight at 37 °C in a total volume of 20 μl. Using the nested primers, 1 μl of digested DNA was amplified using Red Taq DNA polymerase (SIGMA). The PCR conditions were as follows: an initial denaturation step of 95 °C for 4 min followed by 95 °C for 1 min, 55 °C for 2 min and 72 °C for 3 min repeated for 40 cycles.

Table 2.2: Methylation sensitive restriction enzyme coupled PCR primers

Primer Number	**Primer Name**	**Sequence**	**Tm**
EG01	Cdkn1c promoter internal	gcggtgttgttgaaactgaa	61
EG47	Cdkn1c promoter reverse	tgctctcagtcgggccta	61
EG81	Cdkn1c promoter methylation	ggaccgagggaccttcac	60
R89	Cdkn1c intron II methylation	tgagcaggtctctgagcagg	66
EG04	Cdkn1c intron II reverse	cttagctgcacccctaccag	60
R90	Cdkn1c intron II internal	aatccccgagagaggc	60
EG49	KvDMR region 2 methylation	gcgggtttcttctctgagtc	60
EG50	KvDMR region 2 reverse	gcccctcactctcagcatta	60
EG51	KvDMR region 2 internal	ttacagaagcaggggtggtc	60
EG92	Gtl2 DMR methylation	ggaggttcgccggtactaa	60
EG93	Gtl2 DMR internal	agggaagggctgcattatt	60
EG94	Gtl2 DMR reverse	tccatcaggaattccaaagc	60
EG95	H19 DMR methylation	cggttcacctatggcaaact	60
EG96	H19 DMR internal	gatcgatcggttcactctcc	60
EG97	H19 DMR reverse	aatgcctgatccctttgttg	60

2.4 Bisulphite Sequencing

DNA was prepared for bisulphite treatment by digesting 1 µg of genomic DNA in a 50 µl reaction with an appropriate restriction enzyme that cuts around, but not within, the region of interest. 21 µl of the digested DNA was then denatured by heating to 95 °C for 6 min then by the addition of 4 µl of 2 M sodium hydroxide (NaOH), incubating at 37 °C for 30 min. Denatured DNA was mixed with 2 % low melting point agarose (SeaPlaque agarose) and kept at 50 °C. A saturated sodium bisulphite solution (5.20-5.69 M HSO_3^- (pH5.0) (Sigma) and 10 mM Hydroquinone, (Sigma)) was prepared and filter sterilised. 1 ml was added to an eppendorf tube, overlaid with cold mineral oil then chilled on ice for 20 min. 10 µl of agarose/DNA mix was pipetted into the oil to form beads, allowed to harden for 10 min, then pushed into the underlying aqueous solution. Beads were incubated for 30 min on ice, then for 3.5 hrs at 55 °C. At all stages bisulphite solution was protected from light. Beads were equilibrated to remove the bisulphite solution by incubating with 1 x TE buffer (10 mM Tris-HCl (pH 7.4), 1 mM EDTA) for 4 x 15 min at room temperature. Beads were then treated with 0.3 M NaOH for 2 x 15 min to perform the final

desulphonation step. A final equilibration step (3 x 10 min) was performed with 1 x TE buffer to remove all traces of NaOH. Whole beads were used directly in the nested PCR or stored at 4 °C in dH_2O until required.

Semi-nested primers were designed using the Methprimer program (http://www.urogene.org/methprimer/index1.html) to amplify regions within the *Cdkn1c* promoter, *Cdkn1c* intron II, and a CpG island upstream of *Cdkn1c*. Primers are shown in **table 2.3** and primer design is discussed in the results. In the first round PCR the following primers were used on whole beads: *Cdkn1c* promoter EG19 & EG20, *Cdkn1c* intron II EG21 & EG22, upstream CpG island EG45 & EG24. In the second round PCR, 1 µl of the first round reaction was used with the following primers: *Cdkn1c* promoter EG43 & EG20, *Cdkn1c* intron II EG44 & EG 22, upstream CpG island EG23 & EG24. Round one PCR conditions: 94 °C for 2 min, 57 °C for 1 min, 72 °C for 3 min (5 cycles) then 94 °C for 1 min, 57 °C for 1 min, 72 °C for 1 min (35 cycles) and a final extension step of 72 °C for 10 min. Round two PCR conditions: 94 °C for 1 min, 57 °C for 1 min, 72 °C for 1 min (35 cycles) and a final extension step of 72 °C for 30 min. Red Taq DNA polymerase (SIGMA) was used in both of the nested PCRs according to the manufacturer's protocol. The concentration of primer stock was 25 µM per oligonucleotide. 1 µl of primer mix was used per 25 µl reaction.

PCR products were resolved on a 2 % agarose/TAE gel and bands were excised and gel purified using the QIAquick Gel Extraction kit (QIAGEN). PCR products were eluted in 30 µl of 10 mM Tris-HCl (pH7.5) then cloned into the pGEM-Teasy vector (Promega) and transformed into JM109 chemically competent cells (Promega) for blue/white selection as per manufacturer's instructions. Positive clones were checked for the presence of the insert by digesting a sample with *Eco*R1 (NEB) according to the manufacturer's instructions. Confirmed positive plasmids

were sent for sequencing (DNA Sequencing Core, Molecular Biology Unit, Cardiff University). Sequencing was analysed and processed by BiQ analyzer (Bock *et al.* 2005).

Table 2.3: Bisulphite sequencing primers

Primer Number	Primer Name	Sequence	Tm
EG19	Cdkn1c promoter forward round 1	tgggtgtagagggtggatttagtta	62
EG20	Cdkn1c promoter reverse	cccacaaaaaccctaccccc	64
EG21	Cdkn1c intron II forward round 1	gaggaataggtttttgagtaggtttttgag	64
EG22	Cdkn1c intron II reverse	ccttaactacacccctaccaataaaaaaaa	64
EG23	Cdkn1c upstream CpG island Fwd round 2	tgggttttgggtttagttaagttat	59
EG24	Cdkn1c upstream CpG island reverse	caaatctccaactcctaccctatta	59
EG43	Cdkn1c promoter forward round 2	gtattgttaggattaggatttagttggtagtagtag	63
EG44	Cdkn1c intron II forward round 2	ggtgatgagttgggaattgag	63
EG45	Cdkn1c upstream CpG forward round 1	ggggatgtttagtggttttgg	63

2.5 Chromatin immunoprecipitation

Stem cells were cultured on feeder cells until fully confluent. Cells were trypsinised and the feeders removed by panning the cells. Stem cells were collected and the formaldehyde mix (10 x stock: 0.5 mM EGTA, 1 mM EDTA, 50 mM HEPES pH8.0, 100 mM NaCl, 11 % formaldehyde) was added to a final concentration of 1 % and mixed for 10 min at room temperature. The reaction was stopped by the addition of glycine to 0.125 M final concentration. Cells were washed twice in 1 x PBS then harvested in 1 ml of Lysis buffer (1 % SDS, 10 mM EDTA, 10 mM Tris-HCl pH8.1 and protease inhibitors) and incubated at 4 °C for 5 min then stored at -20 °C. Samples were sonicated using the bioruptor (Diagenode) sonicator on high output on cycles of 30 sec on, 30 sec off for 10 min. A 50 µl sample of sonicated chromatin was checked for sonication efficiency by reversing the formaldehyde crosslinks and precipitating the DNA, as described later for the immunoprecipitated samples. Sample DNA was

resuspended in 10 µl of 1 x TE then 10 µl was run out on a 1 % agarose TAE gel at 100 V. DNA concentration was also determined by spectrophotometry. The sonicated chromatin was centrifuged for 20 min at 13,000 rpm at room temperature to prevent precipitation of SDS. Protein concentration was determined by spectrophotometry and 100 µg was diluted in lysis buffer to 1 µg/µl. Samples were diluted 10 fold in dilution buffer (1 % Triton X-100, 2 mM EDTA pH8.0, 150 mM NaCl, 20 mM Tris-HCl pH8.1) containing complete mini protease inhibitors (1 tablet in 10 ml buffer) (Roche). The chromatin lysate was pre-cleared by adding 30 µl protein A sepharose beads (Invitrogen) (100 µg/ml protein A sepharose beads pre-incubated with 100 µg/ml BSA (NEB) and 500 µg/ml sheared salmon sperm DNA (Invitrogen) diluted in 5:1 ratio of dilution buffer and lysis buffer) and rotating for 1-2 hours at 4 °C. The chromatin-bead solution was centrifuged at 4000 rpm for 5 min then the supernatant was transferred to a new tube. Approximately 4 µg of antibody was added to the lysate and incubated overnight with rotation at 4 °C. 30 µl of sepharose beads was added to the IP reaction and incubated with rotation for 1-2 hrs at 4 °C. Beads were washed four times in wash buffer (0.1 % SDS, 1 % Triton X-100, 2 mM EDTA, 150 mM NaCl, 20 mM Tris-HCl pH8.1 with protease inhibitors) centrifuging at low speed for 15 sec between washes, then once in final wash buffer (0.1 % SDS, 1 % Triton X-100, 2 mM EDTA, 500 mM NaCl, 20 mM Tris-HCl pH8.1 with protease inhibitors). Complexes were eluted in 450 µl of elution buffer (1 % SDS, 0.1 M $NaHCO_3$) with 500 µg/ml proteinase K and 500 µg/ml RNase A and incubated at 37 °C for 1-2 hrs. Crosslinks were reversed by heating the samples for at least six hrs at 65 °C. Immunoprecipitated DNA was phenol-chloroform extracted and ethanol and sodium acetone (pH5.2) precipitated with a glycogen carrier. DNA was resuspended in 80 µl

TE and 2 µl were used in a standard PCR reaction. Primers used are shown in **table 2.4**. Antibodies used are shown in **table 2.5**.

The ChIP assays shown in the results were performed using the Orange ChIP kit (Diagenode) according to the manufacturer's instructions. Optimal sonication and fixing conditions were performed as described in the results.

Table 2.4: ChIP primers

Primer Number	Primer Name	Sequence	Tm
EG01	Cdkn1c promoter forward	gcggtgttgttgaaactgaa	60
EG02	Cdkn1c promoter reverse	gtctggatcgcttgtcctgt	60
EG07	Lit1 promoter forward	aagctcacccaatccaaatg	60
EG08	Lit1 promoter reverse	ctcctagcgacaacgggtag	60
EG90	Kcnq1 promoter forward	gctgccttcacctcagctc	65
EG91	Kcnq1 promoter reverse	gctccagtgagaagggacac	64
EG11	Phlda2 promoter forward	cctgcttgggattgagagtg	61
EG12	Phlda2 promoter reverse	atacctggaacaggctgtcg	60

Table 2.5: ChIP antibodies

Catalogue number	Concentration	Volume used/IP	Supplier	Name
Ab9045	1 µg/µl	4 µl	Abcam	Rabbit polyclonal to histone H3 (mono-methyl K9)
Ab8898	0.5 µg/µl	4 µl	Abcam	Rabbit polyclonal to histone H3 (tri-methyl K9)
Ab4441	1 µg/µl	2 µl	Abcam	Rabbit polyclonal to histone H3 (acetyl K9)
Ab7766	0.5 µg/µl	4 µl	Abcam	Rabbit polyclonal to histone H3 (di-methyl K4)
Ab6697	10 µg/µl	0.5 µl	Abcam	Rabbit polyclonal to goat IgG H&L
Kch-403-020	0.5 µg/µl	1 µl	Diagenode	Antibody anti-Histone H3 (K4me3)
Ab6002	1 µg/µl	2 µl	Abcam	Rabbit polyclonal to histone H3K27me3
Ab1791	0.5 µg/µl	4 µl	Abcam	Rabbit polyclonal to histone H3

CHAPTER 3:

Expression of IC2 domain imprinted genes in undifferentiated and differentiated mouse embryonic germ (EG) cells

3.1 Introduction

As described in Chapter 1, parental specific imprints are established during differentiation of PGCs. PGCs can be first identified in mice at around 7.5 dpc as approximately 100 alkaline phosphatase positive cells (Ginsburg *et al.* 1990). These PGCs proliferate as they migrate to the genital ridge and number around 25,000 by around 13.5 dpc (Matsui *et al.* 1992; Tam and Snow 1981). Although *in vivo* studies have been performed on these cells in order to examine their epigenetic status, PGCs are few in number and difficult to extract for analysis. For this reason, it is not practical to isolate large numbers of PGCs from very early stage mouse embryos which would be required to allow an in-depth analysis of specific imprinting domains.

Embryonic Germ (EG) cells are derived from primordial germ cells and are epigenetically similar to the cells from which they were derived, in the same way that Embryonic Stem (ES) cells are similar to the inner cell mass cells from which they are derived (Lee *et al.* 2002; Tada *et al.* 1998). EG cells have been used in many different studies and have been shown to be capable of forming chimeras when injected into mouse blastocysts (Labosky *et al.* 1994; Matsui *et al.* 1992; Stewart *et al.* 1994). EG cells derived from 11.5 and 12.5 dpc PGCs have been shown to colonise both the ICM and the germ line indicating that they have the same developmental potential as ES cells (Durcova-Hills *et al.* 2002; Matsui *et al.* 1992; Tada *et al.* 1998). They are also able to contribute to the extraembryonic lineages in some chimeric embryos and, when subcutaneously injected into nude mice, EG cells are able to give rise to teratomas containing a variety of cell types including neuron ganglia, epithelia and striated muscle (Durcova-Hills *et al.* 2002; Matsui *et al.* 1992). The morphology and behaviour of cultured EG cells is similar to that of ES cells and they are stably

maintained in culture (Matsui *et al.* 1992; Resnick *et al.* 1992). Both EG and ES cells can be expanded in culture to give large numbers of cells for analysis. The differentiation of ES and EG cells *in vitro* as embryoid bodies mimics events that occur in the late pre- and early post-implantation embryo stages and develop cell types derived from all three germ layers (Rohwedel *et al.* 1996; Szabo and Mann 1994).

In our model EG cells derived from 12.5 dpc PGCs were grown in culture along side an ES cell line. In this study we wished to use these EG cells in order to examine the state of non-imprinted chromatin within the mouse IC2 imprinting domain (**figure 3.1**) by comparing gene expression, DNA methylation and chromatin between ES and EG cells as they differentiate. Lee *et al.* (Lee *et al.* 2002) demonstrated the existence of a default state of gene expression in imprint erased PGCs that is common in both male and female germ cells. In imprint erased PGCs (~12.5 dpc), the majority of the maternally expressed genes are silent while most paternally expressed genes are biallelically expressed (Lee *et al.* 2002). The exceptions were *H19* and *Dlk1* leading the authors to suggest that there may be a different regulatory mechanism for these paternally expressed genes. They also observed a complete loss of monoallelic gene expression in clone embryos derived from 12.5 to 13.5 dpc PGCs. These clones were unable to support full-term development possibly due to the lack of imprinting which led to a loss in developmental potential (Yamazaki *et al.* 2005).

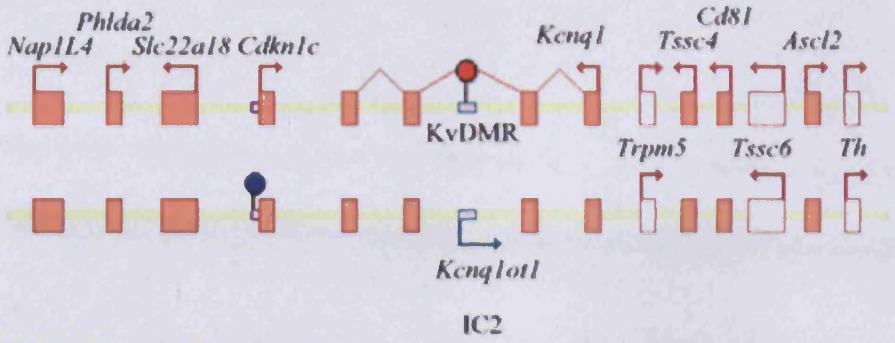


Figure 3.1: Schematic of mouse distal chromosome 7 IC2 domain.

Mouse distal chromosome 7 cluster contains two independently regulated imprinting domains. The centromeric or IC2 domain is regulated by germ line maternal methylation of the promoter of *Kcnq1ot1/Lit1*, at a region known as KvDMR1. This imprinting domain contains the imprinted genes *Cdkn1c, Phlda2* and *Kcnq1*.

The aim of this study was to set up an *in vitro* system using embryonic germ (EG) cells in order to look at the expression states of imprinted genes in non-imprinted chromatin to determine if our model reflects what is seen *in vivo*, that the domain is silent in imprint erased PGCs. We then aimed to examine the changes in expression that occur as ES and EG cells differentiate in culture to see if the IC2 domain genes are expressed or are silenced in the EG cell line.

3.2 Results

Changes in gene expression occur as ES cells differentiate and new cell types appear. The gametic marks required to establish allele-specific gene expression are maintained in ES cells during culture (Szabo and Mann, 1994) making these cells useful for comparing how non-imprinted EG cell gene expression levels differ from levels in ES cells with monoallelic expression. For example, we would predict that, if an imprinted gene was active in the absence of an imprint, expression levels within the EG cells would be approximately twice that in the ES cells. Conversely, if the imprinted gene was silent in the absence of an imprint, expression levels would be

substantially lower in the EG cells. We are only able to look at genes expressed in the embryo using this system since genes expressed only in the extraembryonic material would not be detectable in the ES cells. Within the IC2 domain on mouse distal chromosome 7 there are several genes that are predominantly expressed in the extraembryonic material such as *Ascl2.* It was expected that expression of this gene would be undetectable in ES cells. There are also a number of genes, such as *Obph1, Tssc4* and *Cd81,* that are only imprinted in the placenta (Engemann *et al.* 2000; Paulsen *et al.* 2000). For this reason we decided to look at genes within the IC2 domain that were known to be ubiquitously imprinted in both the placenta and the embryo and expressed in the embryo. Some genes, such as *Kcnq1* are tissue specific and are predominantly expressed later on during development so may not be detectable in the ES cells.

The EG and ES cells can be differentiated in culture as embryoid bodies by removing the supporting feeder layer and culturing in the absence of LIF (as described in chapter 2). At some point during differentiation of the ES cells as embryoid bodies it is expected that levels of the imprinted genes will increase reflecting the emergence of cell types, such as neurons, in which the genes are expressed. For example, *Cdkn1c* is not expressed in stem cells but is upregulated as cells differentiate, indicating that this gene is now expressed (John *et al.* 2001a). We know that both ES and EG cells can contribute to all three germ layers. However, there are differences in expression of imprinted genes which may alter the proportion of the somatic cell types within the embryoid bodies.

3.2.1 Design and optimisation of real time quantitative PCR primers

The primers used in the gene expression study were designed specifically to amplify cDNA. Where possible each primer set was designed to span an intron. An *in silico* PCR was performed to confirm the specificity of the primers (Karolchik *et al.* 2003). Total RNA was extracted from wild type embryos, placentae and embryonic stem cells, treated with DNase I and reverse transcribed to produce cDNA. Each primer set was tested on ES cell cDNA and on either embryonic or placental cDNA in order to demonstrate that they produced a product of the correct size. A list of primers designed to amplify genes within the IC2 domain are shown in **table 3.1**. The primers were designed to have an annealing temperature of around 60 °C and to amplify products that were between 100-300 bp in size. These primers were tested in the real time assay to ensure a product could be detected.

Two sets of primers were designed to amplify *Kcnq1*. The initial primers (EG56/57) amplified a product that was too large for use in the real time assay. Primers were redesigned to give a smaller PCR product (EG71/72). Although a *Kcnq1* could be detected in ES cell cDNA, the expression levels were very low and not ideal for the real time PCR. This suggested that *Kcnq1* was not expressed in ES cells.

Slc22a18 (EG84/85) could not be amplified from either the embryonic or ES cell cDNA samples although we could amplify a product from the placental cDNA indicating that this gene was not highly expressed in the embryo or ES cell samples.

The initial *Kcnq1ot1/Lit1* primers (EG52/53) failed to amplify a product in real time PCR so were redesigned. The second set of primers (R190/191) amplified a product from the ES cell cDNA and the embryonic cDNA. *Kcnq1ot1/Lit1* is expressed in a variety of embryonic, neonatal and adult tissues including brain, kidney, lung, testis, placenta and in whole embryo but is only imprinted in fetal tissues (Mancini-

DiNardo *et al.* 2003; Paulsen *et al.* 1998). It is biallelically expressed in most neonatal tissues suggesting that imprinting of this gene is relaxed later in development (Paulsen *et al.* 1998).

Primers were also designed to amplify *Tssc4* (R248/249) but this produced two PCR products which could possibly be a result of alternative splicing. It was decided not to redesign primers for *Tssc4. Cdkn1c* (R161/162) and *Phlda2* (EG36/37) products could be detected in both embryonic and ES cDNA samples in the real time machine and *Phlda2* could also be detected in placental cDNA.

Table 3.1: List of primers designed for use in real time quantitative RT-PCR for genes within the IC2 imprinting domain.

The primers were designed to amplify a product between 100 – 300 bp. The Pubmed accession numbers are shown for both RNA and DNA. An *in silico* PCR was performed for each primer set to determine the product size in genomic DNA. Primers were either tested on embryonic, placental or ES cell cDNA. (+) PCR product detected (-) PCR product absent.

Primers	Primer number	Sequence	Tm	Product size cDNA	Product size gDNA	Accession numbers: mRNA, gDNA	Tested in Embryo	Tested in Placenta	Tested in ES cells
Cdkn1c	R161 R162	AGAGAACTGCGCAGGAGAAC TCTGGCCGTTAGCCTCTAAA	59.8 60.0	141 bp	231 bp	NM_009876.3 NT_039453.2	Yes +	No	Yes +
Phlda2	EG36 EG37	TCAGCGCTCTGAGTCTGAA TCCTGGGCTCCTGTCTGAT	60.0 60.8	124 bp	364 bp	NM_009434 NT_039453	Yes +	Yes +	Yes +
Slc22a18	EG84 EG85	TGATGTCCAGTGTGCTCCAT AGAGTTCGGGTCAATGGTTG	60.1 60.0	166 bp	1820 bp	NM_001042760.1 NT_039453	Yes -	Yes +	Yes -
Kcnq1	EG56 EG57 EG71 EG72	GATCACCACCCTGTACATTGG CCAGGACTCATCCCATTATCC GGAACATAGGGATGGGGAGT GTTCCCCTGATGGTCTCTGGA	60.0 60.0 60.0 60.0	516 bp 151 bp	10906 bp 151 bp	NM_008434.2 NT_039451.1	Yes + Yes +	No No	Yes + Yes +
Kcnq1ot1 /Lit1	EG52 EG53 R190 R191	CTTACAGAAGCAGGGGTGGTCT CTCAGTTCCACGATACCCTTCC TCCAATCGGGTAGAGATTCG AGACCATCGGAAAACACAGG	61.9 62.1 60.0 60.0	300 bp 245 bp	300 bp 245 bp	NR_001461.3 NT_039453	Yes + Yes +	No No	Yes - Yes +
Tssc4	R248 R249	AAGCCACGAAAATCTGATCG ATCCTAGCGCGCACAGATAC	60.2 60.4	163 bp	404 bp	NM_020285 NT_039451	Yes +	Yes +	No

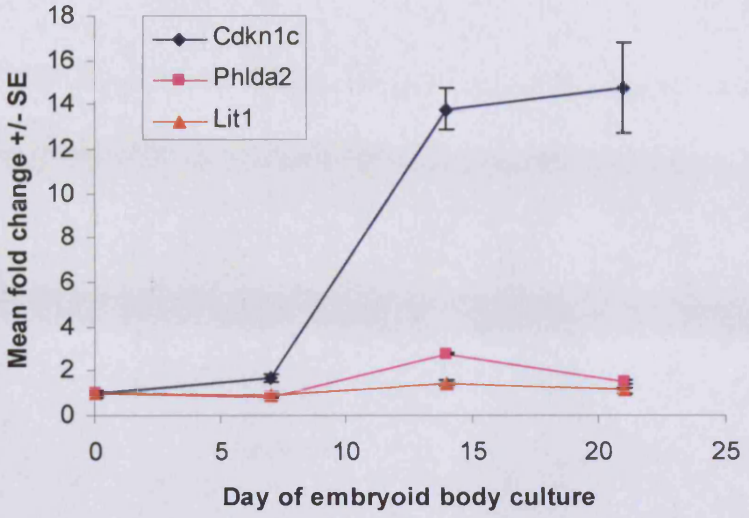
Table 3.1: List of primers continued...

Primers	Primer number	Sequence	Tm	Product size cDNA	Product size gDNA	Accession numbers: mRNA, gDNA	Tested in Embryo	Tested in Placenta	Tested in ES cells
GAPDH	R149 R150	CACAGTCAAGGCCGAGAATG TCTCGTGGTTCACACCCATC	61.8 62.0	242 bp	211-246 bp	NM_008084 NT_03935307	Yes +	Yes +	Yes +
β-Actin	R151 R152	CCTGTATGCCTCTGGTCGTA CCATCTCCTGCTCGAAGTCT	58.7 59.6	262 bp	260 & 262 bp	NM_007393 NT_081055.6	Yes +	Yes +	Yes +
18s rRNA	R209 R210	CATGGCCGTTCTTAGTTGGT CGCTGAGCCAGTCAGTGTAG	60.0 59.8	218 bp	218 bp	NR_003278	Yes +	Yes +	Yes +

3.2.2 Expression profiles of IC2 domain genes in undifferentiated and differentiated ES and EG cells

To determine how the expression levels of imprinted genes change during ES and EG cell differentiation, real time quantitative RT-PCR (qRT-PCR) was performed (as described in chapter 2). In the initial analysis, the relative expression levels of the imprinted genes during differentiation were calculated for the CES3 (ES cell line) and Sv6.1 (EG cell line) cells. For both ES and EG cells there was very little difference in *Cdkn1c* expression between day 0 and day 7 (**figure 3.2**). Between day 7 and 14 the expression of *Cdkn1c* increased indicating that this gene was switched on in the differentiated cells after day 7. The same pattern was observed in both ES and EG cells suggesting that they could be going though similar developmental stages in culture (**figure 3.2 A & B**). *Phlda2* levels remained constant in both EG and ES cells at all time points examined. As this gene is predominantly expressed in embryonic tissues it is likely that this gene was only expressed at low levels in ES cells. *Kcnq1* and *Slc22a18* expression was not analysed by real time quantitative RT-PCR as they were not detectable in undifferentiated and differentiated ES cells. *Kcnq1ot1/Lit1* expression was also examined by real time qRT-PCR. In ES cells the expression of *Kcnq1ot1/Lit1* was similar to that seen for *Phlda2*, the levels were low and constant throughout differentiation (**figure 3.2 A**). In EG cells, *Kcnq1ot1/Lit1* levels steadily increased over time (**figure 3.2 B**).

A: Imprinted gene expression in CES3 ES cells



B: Imprinted gene expression in Sv6.1 EG cells

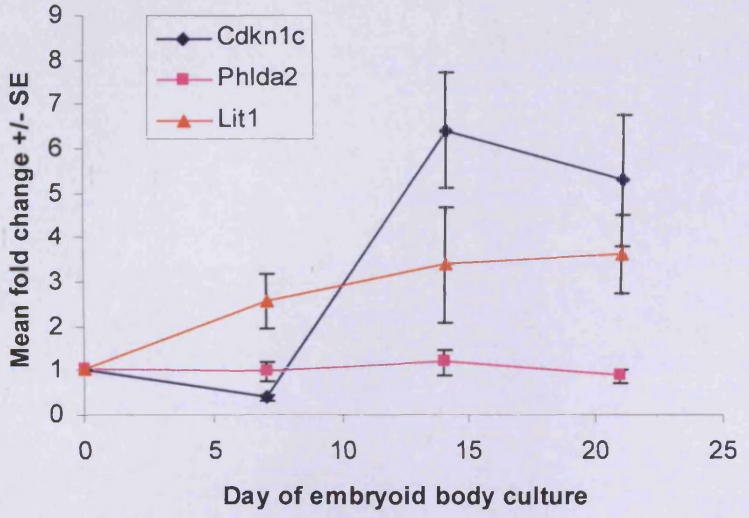


Figure 3.2: Expression of *Cdkn1c*, *Phlda2* and *Kcnq1ot1/Lit1* during differentiation of the ES cell line CES3 and EG stem cell line Sv6.1.

A. Expression of the IC2 domain imprinted genes in CES3 ES stem cells. **B.** Expression in Sv6.1 EG stem cells.
Embryoid bodies were differentiated for 0, 7, 14 and 21 days. Mean fold change was calculated using the delta delta CT method (see chapter 2) comparing the CT values of differentiated ES or EG cells to undifferentiated cells so that expression at each data point is relative to expression at day 0 (i.e. 1) for each gene. Standard error is shown. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

3.2.3 Comparison of gene expression between ES and EG cells

The previous data showed that the expression profiles of *Cdkn1c* and *Phlda2* were similar between the two cell lines. The only difference appeared to be *Kcnq1ot1/Lit1* where the levels increased in the differentiating EG cells. The next step was to compare the expression levels in the EG cell line to the expression levels in the ES cell line (**figure 3.3**). This data indicated that *Cdkn1c* and *Phlda2* levels were suppressed in the EG cell line at all time points but that *Kcnq1ot1/Lit1* levels were higher in the EG cells compared to the ES cells.

To give a comparison of the relative expression in ES versus EG cells, we took the data presented in **figure 3.2A**, showing the expression levels of each gene over time in ES cells, and the data presented in **figure 3.3**, showing the expression in ES versus EG cells, and multiplied the two values together to give the relative expression in EG cells compared with ES cells for each gene.

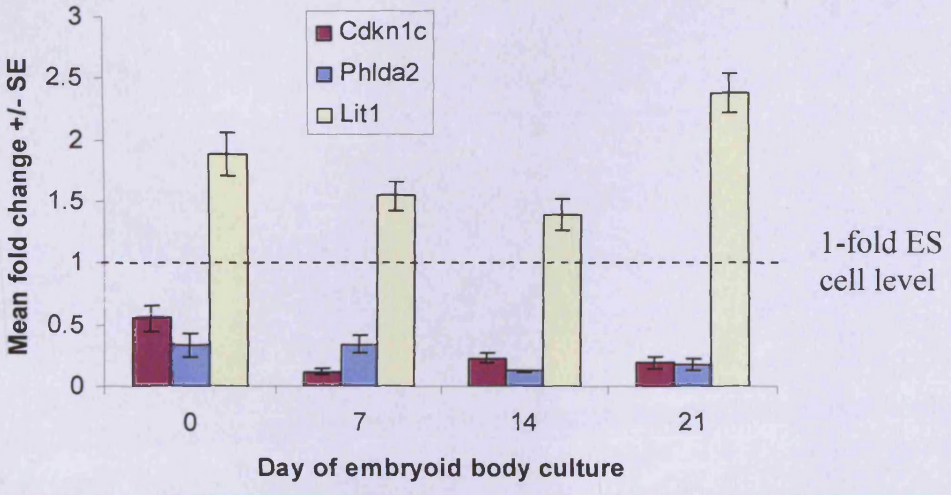


Figure 3.3: Relative expression of *Cdkn1c*, *Phlda2* and *Kcnq1ot1/Lit1* in Sv6.1 EG cells compared to expression in CES3 ES cells at each time point.

Expression in Sv6.1 EG cells compared to expression in CES3 ES cells at each time point. The delta delta CT method (see chapter 2) was used to calculate the fold change in EG cells compared to ES cells. Standard error is shown. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

Although a similarity was seen in the *Cdkn1c* expression profile between ES and EG cells (shown in **figure 3.2**), the absolute levels were reduced in the EG cell line indicating that expression of *Cdkn1c* was somehow repressed (**figure 3.4**). There was a slight increase in *Cdkn1c* expression in the EG cell line at day 14 corresponding to the increase in gene expression also observed in the ES cell line that might indicate that the suppression of *Cdkn1c* in the EG cells was not absolute.

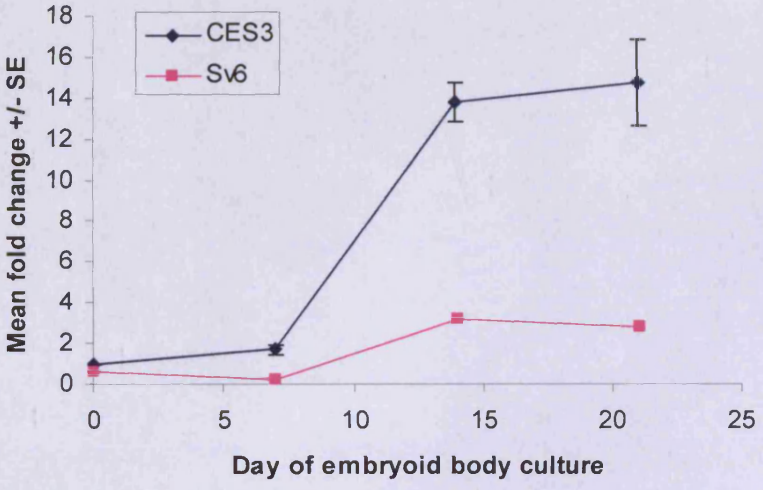


Figure 3.4: Calculated relative expression of *Cdkn1c* in Sv6.1 EG cells compared to expression in CES3 ES cells.

The relative expression of *Cdkn1c* overtime in ES and EG cells was calculated by, firstly, comparing the expression level at each differentiation time point in ES cells to day 0 in ES cells. The expression in EG cells at each time point compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

A similar result was observed for *Phlda2* expression. Levels of *Phlda2* were repressed in EG cells compared to ES cells and showed a low and constant level throughout differentiation (**figure 3.5**). There was an increase in *Phlda2* expression at day 14 in ES cells which may indicate that the gene was expressed at this time point.

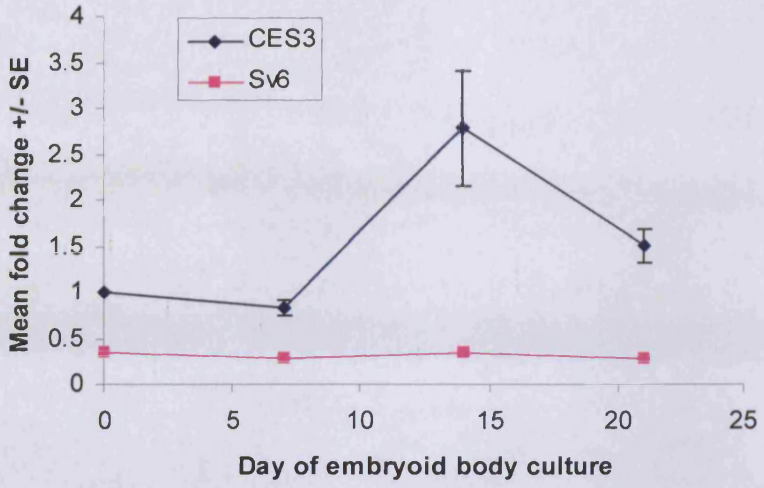


Figure 3.5: Calculated relative expression of *Phlda2* in Sv6.1 EG cells compared to expression in CES3 ES cells.

The relative expression of *Phlda2* overtime in ES and EG cells was calculated by, firstly, comparing the expression level at each time point in ES cells to day 0 in ES cells. The expression in EG cells at each time point compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

Kcnq1ot1/Lit1 expression was approximately 2-fold higher in the EG cells compared to the ES cells suggesting possible expression from both alleles (**figure 3.6**). The level of *Kcnq1ot1/Lit1* expression in the ES cells appeared to be relatively constant throughout differentiation. In contrast, in the EG cells *Kcnq1ot1/Lit1* levels increased slightly as the cells differentiated. Expression was at least 2-fold higher than in the ES cells.

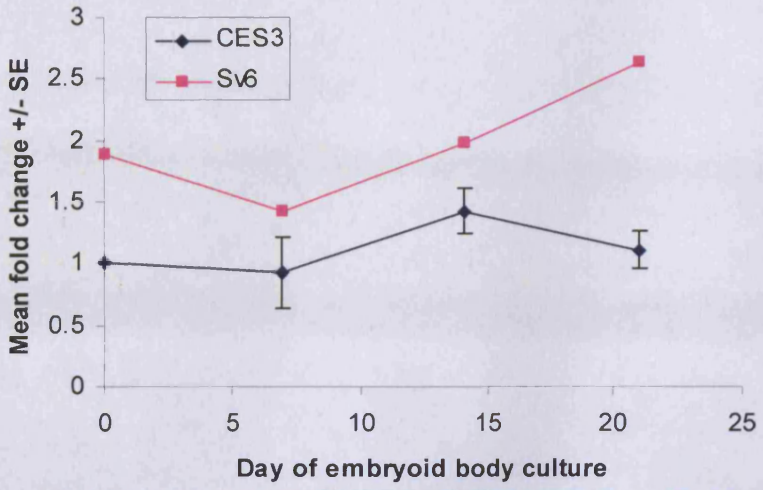


Figure 3.6: Calculated relative expression of *Kcnq1ot1/Lit1* in Sv6.1 EG cells compared to expression in CES3 ES cells.

Relative expression of *Kcnq1ot1/Lit1* overtime in ES and EG cells was calculated by, firstly, comparing the expression level at each time point in ES cells to day 0 in ES cells. The expression in EG cells at each time point compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error of the ES cell data is shown. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

3.2.4 Expression at early time points in Sv6.1 EG cells compared to CES3 ES cells

Our data suggested that *Cdkn1c* and *Phlda2* expression was suppressed in the differentiated Sv6.1 EG cells whereas *Kcnq1ot1/Lit1* expression was not repressed. If *Cdkn1c* and *Phlda2* were actively silenced during differentiation we might have observed an initial increase in *Cdkn1c* and *Phlda2* expression in the EG cells, similar to ES cells, prior to the silencing event. Therefore we examined earlier time points. We also examined *Kcnq1ot1/Lit1* expression at these early time points as this non-coding RNA has been shown to play a role in silencing this domain (Mancini-DiNardo *et al.* 2006; Murakami *et al.* 2007; Shin *et al.* 2008).

Figure 3.7 shows the combined results of the early time point *Cdkn1c* expression levels and the later time points for comparison. Changes to gene expression were calculated as described previously. At early time points in the ES cells we observed a peak in *Cdkn1c* expression on the first day of differentiation which is subsequently lost by day 3. This increase was not observed in the EG cell line where *Cdkn1c* did not appear to be expressed at any time point until after day 7. This suggested that *Cdkn1c* was not expressed at any time point from non-imprinted chromatin. However, we could not exclude the possibility that this lack of expression was due to the absence of a particular cell type in day 1 EG cells that was present in the day 1 ES cells. We also looked at *Phlda2*.

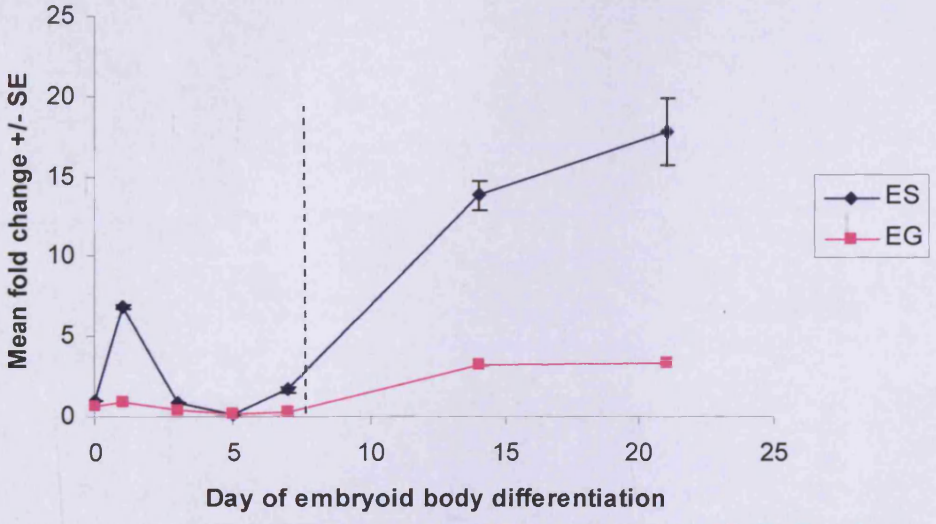


Figure 3.7: Relative expression of *Cdkn1c* in Sv6.1 EG cells compared to expression in CES3 ES cells at both early and late time points.

The relative expression of *Cdkn1c* overtime in ES and EG cells was calculated by, firstly, comparing the expression level at each differentiation time point in ES cells to day 0 in ES cells. The expression in EG cells at each time point compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. Dashed line shows the separation between early and late time points (shown in figure 3.4). At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

A similar result was observed for *Phlda2* (**figure 3.8**). This gene was not expressed at very high levels in differentiated ES cells. However, we did observe the same increase in expression at day one of differentiation for *Phlda2* that we found for *Cdkn1c*. No increase in expression at this time point was observed in the EG cells. *Cdkn1c* and *Phlda2* are not generally expressed in the same cell types. This lack of expression was unlikely to be due to the absence of a specific cell type and more likely to reflect the fact that both genes were silent from the earliest time point in EG cells.

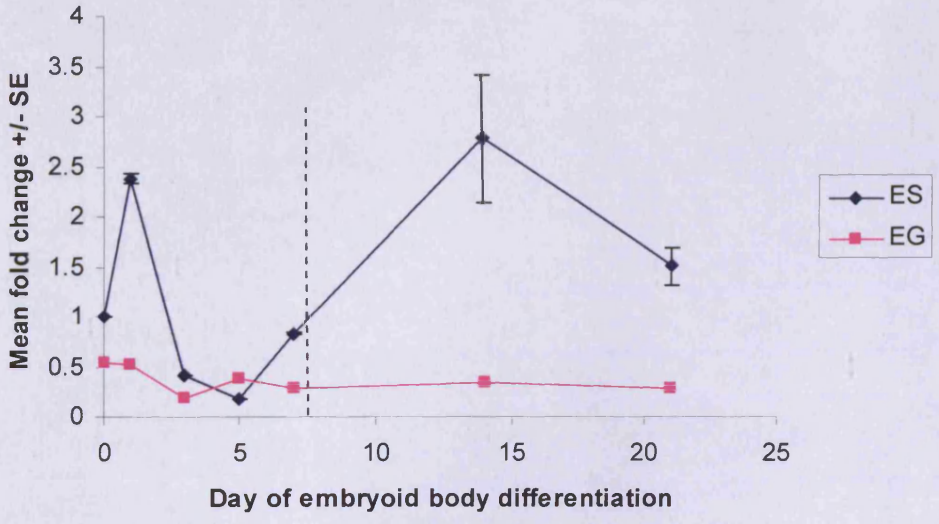


Figure 3.8: Relative expression of *Phlda2* in Sv6.1 EG cells compared to expression in CES3 ES cells at both early and late time points.

The relative expression of *Phlda2* overtime in ES and EG cells was calculated by, firstly comparing the expression level at each differentiation time point in ES cells to day 0 in ES cells. The expression in EG cells at each time point compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. Dashed line shows the separation between early and late time points (shown in figure 3.5). At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

There were dynamic changes in *Kcnq1ot1/Lit1* expression levels at the early time points in both ES and EG cells (**figure 3.9**). Both the ES and the EG cells

displayed roughly the same profile but there were differences in the absolute levels. In the EG cells we observed a 50 % decrease in *Kcnq1ot1/Lit1* levels at day 1 compared with day 0. *Kcnq1ot1/Lit1* levels then appeared to steadily increase to 2-fold levels by day 7. However, in the ES cell line the levels of *Kcnq1ot1/Lit1* decreased sharply to their lowest level at day 5 before returning to 1-fold levels at day 7. It is possible that the different expression patterns in the cell lines represent the different functions of *Kcnq1ot1/Lit1* in these cells. In EG cells, *Kcnq1ot1/Lit1* may be required to silence the IC2 domain whereas, in the ES cells, the domain has already undergone imprinting and no further role for *Kcnq1ot1/Lit1* is required.

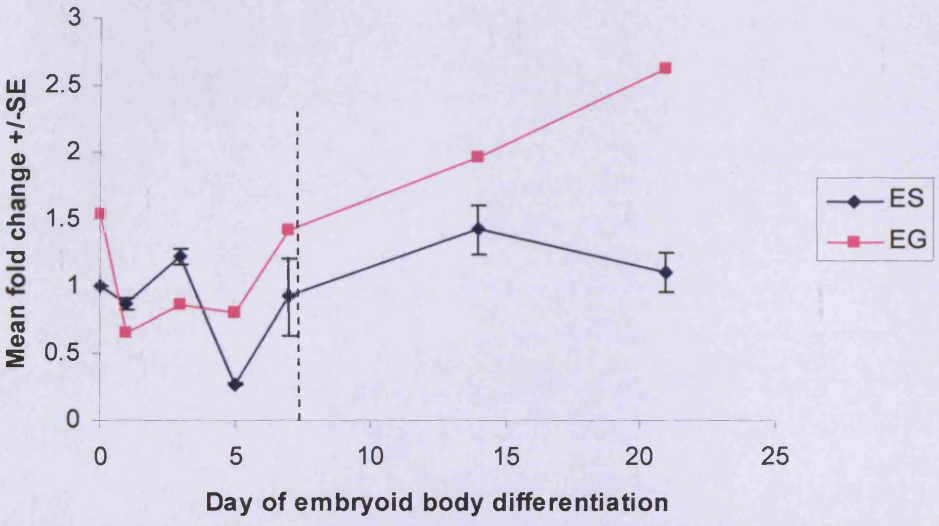


Figure 3.9: Relative expression of *Kcnq1ot1/Lit1* in Sv6.1 EG cells compared to expression in CES3 ES cells at both early and late time points.

The relative expression of *Kcnq1ot1/Lit1* overtime in ES and EG cells was calculated by, firstly comparing the expression level at each differentiation time point in ES cells to day 0 in ES cells. The expression in EG cells at each time point compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. Dashed line shows the separation between early and late time points (shown in figure 3.6). At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

The results so far suggested that the IC2 domain maternally expressed genes were low or silent in the absence of an imprint. This concurs with what was seen in

the Lee *et al.* and Tada *et al.* studies and indicated that there was default silencing of this domain in EG cells (Lee *et al.* 2002; Tada *et al.* 1998). The presence of a somatic DMR on the promoter of *Cdkn1c* suggested that DNA methylation was a likely candidate for establishing silencing of this gene. The cells used in this study are imprint erased which means all of the germ line methylation marks are erased. In the absence of the germ line methylation mark, the IC2 domain is silent. Several studies have shown that KvDMR1 methylation is required for expression of the surrounding imprinted genes. Loss of methylation at KvDMR1 often results in loss of expression of the IC2 domain imprinted genes (Howell *et al.* 2001). Also, deletion of the maternal KvDMR1 results in biallelic expression of the imprinted genes demonstrating the importance of the region for establishing and maintaining imprinted gene expression (Fitzpatrick *et al.* 2002; Horike *et al.* 2000; Mancini-DiNardo *et al.* 2003).

3.2.5 Confirmation that *Kcnq1ot1/Lit1* expression is not the unspliced *Kcnq1* transcript

The *Kcnq1ot1/Lit1* gene overlaps and is antisense to the *Kcnq1* gene. It was therefore possible that the primers designed to amplify the *Kcnq1ot1/Lit1* gene were picking up unspliced *Kcnq1* transcript. To resolve this issue, we looked for a polymorphism within the *Kcnq1ot1/Lit1* transcript. A published polymorphism exists between C57BL/6 and PWK mouse strains and results in G/C change (Yatsuki *et al.* 2002). In 129sv and C57BL/6 mice the polymorphism results in the presence of a *Nci*I cut site, whereas in PWK mice (a highly inbred mouse strain derived from wild mice of mus musculus musculus subspecies) this *Nci*I cut site is absent (Gregorova and Forejt 2000; Yatsuki *et al.* 2002) (**figure 3.10a**). Our aim was to determine if this polymorphism was also present in *M.Spretus* to allow us to identify allele specific

expression in SF1-1 ES cells. SF1-1 ES cells were derived from embryos produced as a result of *in vitro* fertilization of (C57BL/6 x CBA/Ca) F1 eggs with *M. spretus* sperm (Dean *et al.* 1998; Feil *et al.* 1997). If the polymorphism was present then it should be detected in these ES cells as a 195 bp band. Genomic DNA from SF1-1 and CES3 ES cells was amplified using the *Kcnq1ot1/Lit1* real time PCR primers (EG52/53). **Figure 3.10b** shows the result of digesting the PCR product with *Nci*I. There were strong bands at 67 bp, 106 bp and 127 bp indicating the presence of the 106 bp and 173 bp *Nci*I cut sites in both the 129Sv ES cells and the SF1-1 cells. If the polymorphism was present, we would have expected a 195 bp band in the SF1-1 sample.

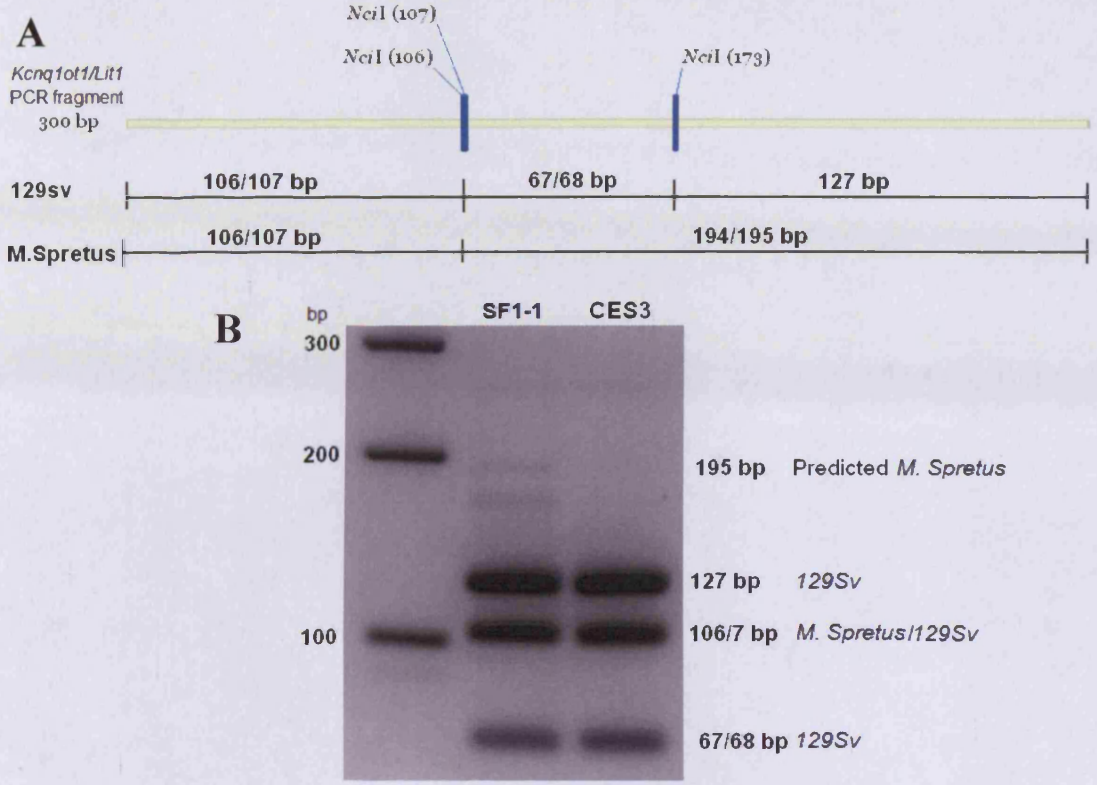


Figure 3.10: Identification of a polymorphism at *Kcnq1ot1/Lit1* within SF1-1 genomic DNA corresponding to the *M.Spretus* allele.

A The 300 bp PCR fragment produced after amplification with primers EG52 and EG53 showing the expected locations of the *Nci*I cut sites and the expected fragment sizes in 129Sv and *M.Spretus* genomic DNA. The 173 bp *Nci*I site is absent in PWK mice (AP001295; 103181 G/C change) and may also be absent from *M. Spretus* mice. **B** Digestion of the 300 bp PCR product by *Nci*I in SF1-1 and CES3 ES cell gDNA.

As we were unable to locate the polymorphism within the SF1-1 cell line, RNA from undifferentiated and differentiated EG cells was sent to Professor Michael Higgins (Department of Cancer Genetics, Roswell Park Cancer Institute, New York) for use in a ribonuclease protection assay (RPA) and a *Kcnq1ot1/Lit1* strand-specific RT-PCR (**figure 3.11**; supplied by Professor Michael Higgins). A protected fragment corresponding to *Kcnq1ot1/Lit1* was detected in both the day 0 and day 5 EG samples and was absent in the negative controls (RNA from KvDMR1 deletion and poly (A) truncation mice). A strand-specific RT-PCR was used to confirm the RPA results.

Reverse transcription was carried out using a strand-specific primer so that RNA corresponding to the primary unspliced *Kcnq1* transcript would not be converted into cDNA and detected. Significant expression of *Kcnq1ot1/Lit1* was detected in both undifferentiated and differentiated EG cells. These results confirmed what was seen in this study, that *Kcnq1ot1/Lit1* was expressed in the undifferentiated and differentiated EG cells.

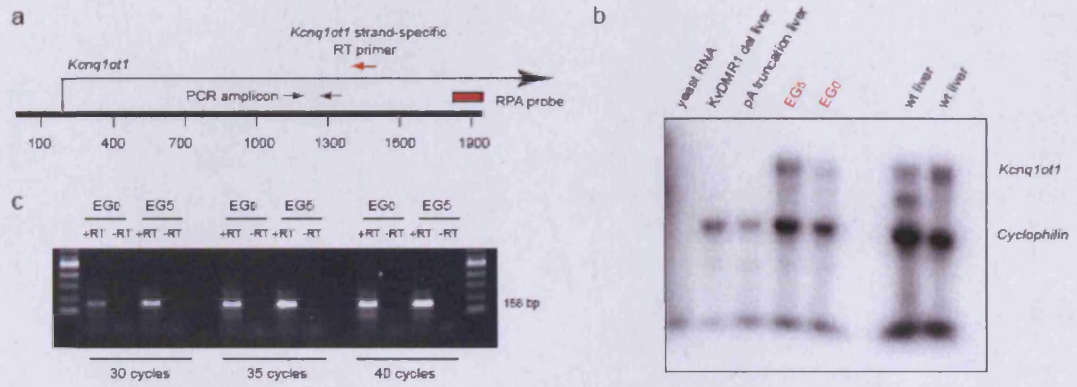


Figure 3.11: Strand specific RT-PCR and RPA on undifferentiated and differentiated Sv6.1 EG cells.

A Schematic of the *Kcnq1ot1/Lit1* gene showing location of the primers used in the strand specific RT-PCR and the location of the RPA probe. **B** RPA assay on yeast RNA, liver RNA from KvDMR1 deletion mice, liver RNA from poly (A) truncation mice, undifferentiated (EG0) and differentiated (EG5) EG cell RNA, and RNA from wild type liver. Cyclophilin was used as a control. **C** Strand-specific RT PCR showing amplification of *Kcnq1ot1/Lit1* in undifferentiated (EG0) and differentiated (EG5) EG cells. Figure from Professor Michael Higgins (Department of Cancer Genetics, Roswell Park Cancer Institute, New York).

3.2.6 Design of primers for the expression analysis of other imprinted genes

The data published by several groups including Lee *et al.* suggested that there are two types of imprinted genes, those that require an imprint to be expressed and those that require an imprint to be silenced (Lee *et al.* 2002). Our imprinted domain is one that requires an imprint for the maternal alleles to be expressed. To identify other

imprinting domains that showed the same pattern and possibly rely on the same silencing mechanisms, we designed primers to amplify a number of imprinted genes and tested them in ES cells. As described in Section 3.2.1, primers were designed to specifically amplify cDNA and were tested on embryonic, placental or ES cell cDNA before use in the real time machine (**table 3.2**). Those that were found to be expressed in ES cells are shown in the results. Primer sets that failed to amplify a product in the ES cell cDNA were not included in the analysis.

Table 3.2: Real time quantitative RT-PCR primers of 'other imprinted genes'.

The primers were designed to amplify a product between 100 – 300 bp. The Pubmed accession numbers are shown for both RNA and DNA. Primers were either tested on embryo, placenta or ES cell cDNA. An *in silico* PCR was performed for each primer set to determine the product size in genomic DNA. (-) in product size gDNA indicates no product detected by *in silico* PCR. (+) PCR product detected (-) PCR product absent.

Primers	Primer number	Sequence	Tm	Product size cDNA	Product size gDNA	Accession numbers cDNA, gDNA	Tested in Embryo	Tested in Placenta	Tested in ES cells
Gtl2	EG57a	ATGTGCCAAGTTCTGGGTTC	60.0	222 bp	-	NR_003633	Yes	No	Yes
(Meg3)	EG58	CAAGAGCTACGCATTCACCA	60.0			NT_114989.1	+		+
Rasgrf1	EG59	TGATCGTATCCAATCCAGCA	60.0	189 bp	3042 bp	NM_011245.2	Yes	No	Yes
	EG60	CTGAACACCACCTGGTTCCT	60.0			NT_039476.7	+		-
	EG73	TGGAGATCACCTCCTCCATC	60.0	223 bp	7508 bp		Yes	No	Yes
	EG74	GTAATTGGGTGTCCCCTCCT	60.0				+		-
Zac1	EG61	AAGTCTCACGCGGAAGAAAA	60.0	158 bp	158 bp	NM_009538	Yes	No	Yes
(Plagl1)	EG62	CTCTGGGCACAGAACTGACA	60.0			NT_039492.7	+		+
	EG75	GAGCAGAGGAAGGAGCAGAA	59.8	178 bp	807 bp		Yes	No	Yes
	EG76	CACAATGGGCAACAAGAAG	57.7				+		+
Nnat	EG63	AGAAAAGCAGCACCGACAAT	60.0	204 bp	204 bp	NM_010923	Yes	No	Yes
	EG64	GGCTGTTCGATCTTCATGGT	60.0			AC_000024	+		+
Dlk1	R172	CACCTGGGTTCTCTGGAAAG	59.7	188 bp	1579 bp	NT_114989.1	Yes	No	Yes
	R173	AGGGGTACAGCTGTTGGTTG	60.0			NM_010052.4	+		+
Grb10	R174	CTGCACCACTTCTTGAGGAT	57.9	151 bp	2888 bp	NT_039515.6	Yes	No	Yes
	R175	AACTGCTGGTCTTCCTCCT	59.4			NM_010345	+		+

Primers	Primer number	Sequence	Tm	Product size cDNA	Product size gDNA	Accession numbers cDNA, gDNA	Tested in Embryo	Tested in Placenta	Tested in ES cells
Igf2r	R170 R171	GTTGGTGTAGGGCCAGTGTT AGAAATTCTGCGGGGTACT	59.9 58.2	181 bp	-	NM_010515 NT_039638.7	Yes +	No	Yes +
Igf2	R168 R169	GTCGATGTTGGTGCTTCTCA AAGCAGCACTCTTCCACGAT	59.8 60.0	195 bp	1400 bp	NM_010514 NT_039437.7	Yes +	No	Yes +
Ube3a	R244 R245	TGCTGTCACAAAGAATCTGG TCCTCCACAACCAACTGAAA	57.4 59.1	288 bp	7953 bp	NM_001033962.1	Yes +	Yes +	Yes -
Hymai	EG77 EG78	CTTTGGCCACAAAAATTGCT AACCAGAAACTGGGCAGAGA	60.1 59.8	227 bp	-	AF_314094	Yes -	No	Yes -
Gnasxl	EG67 EG68	ATCCGAGTGTACCCGATCAC CTGCTTTTGCCAGACTCTCC	60.0 60.0	225 bp	34219 bp	NM_010309.3	Yes +	No	Yes +
Gnas Ex 1A	EG65 EG68	GTTGCTTCAGGTGGCTGGTA CTGCTTTTGCCAGACTCTCC	61.2 60.0	194 bp	6464 bp	NR_003258.1 BC_062654	Yes +	No	Yes +
Nesp55	EG66 EG68 R236 R237	GAGGGCCCTTAGATCAGGA CTGCTTTGCCAGACTCTCC GGAGAGTCTGGCAAAAGCAC TGGGGTAGGACATAGCGAAG	59.2 60.0 60.0 60.1	234 bp 173 bp	- 3464 bp	NM_019690 NM_022000.2	Yes - Yes +	No No	Yes - Yes -
Nespas	R238 R239	CACTGAGTGTCCTCCAAGCA AGACCCCAGCTTCTCTCCTC	60.0 60.0	241 bp	1715 bp	NR_002846.2	Yes +	No	Yes +

3.2.7 Maternally expressed genes *Igf2r* and *Grb10*

Mouse *Igf2r* is a maternally expressed gene that requires a maternal methylation mark to be expressed and is a negative regulator of embryonic growth (Labosky *et al.* 1994; Wutz *et al.* 1997). The model for imprinting of this gene is described in Chapter 1.2.1. Like *Cdkn1c, Igf2r* is only expressed in ES cells after differentiating the cells as embryoid bodies (Szabo and Mann 1994).

Mouse Grb10 is an important inhibitor of embryonic growth and it has been shown that a maternally inherited loss of function mutation in *Grb10* results in embryonic overgrowth in mice (Charalambous *et al.* 2003). The upstream region of *Grb10* contains two CpG islands and is highly conserved between mice and humans (Arnaud *et al.* 2003). In humans, *GRB10* is biallelically expressed in the majority of tissues except for maternal-specific expression in skeletal muscle and paternal-specific expression in the brain (Blagitko *et al.* 2000). In the mouse, *Grb10* is maternally expressed in the majority of tissues except for an isoform specifically expressed in the brain (Blagitko *et al.* 2000; Miyoshi *et al.* 1998; Yamasaki-Ishizaki *et al.* 2007) (**figure 3.12**). The germ line DMR is found within the *Grb10* gene at a maternally methylated CpG island (Arnaud *et al.* 2003). On the paternal allele, CTCF binds the unmethylated DMR and insulates the major isoform promoter of *Grb10* from downstream enhancers. On the maternal allele, methylation of the DMR prevents CTCF binding so *Grb10* is expressed (Hikichi *et al.* 2003; Yamasaki-Ishizaki *et al.* 2007). There are two other CpG islands in the mouse, one of which is located over the promoter of the brain specific isoform of *Grb10* (Arnaud *et al.* 2003; Yamasaki-Ishizaki *et al.* 2007).

The expression of *Igf2r* and *Grb10* was analysed by qRT-PCR in both ES and EG cells as described previously. **Figure 3.13** shows the difference in expression of *Igf2r* and *Grb10* between the ES and EG cell line in undifferentiated and differentiated cells. As seen with the *Cdkn1c* results earlier, *Igf2r* and *Grb10* expression was low in the EG cell lines after differentiating for 14 days. In the undifferentiated ES and EG cells *Igf2r* expression was also low. *Grb10* expression was at levels similar to ES cell levels. The results indicated that, like *Cdkn1c* an imprint is required for these genes to be expressed.

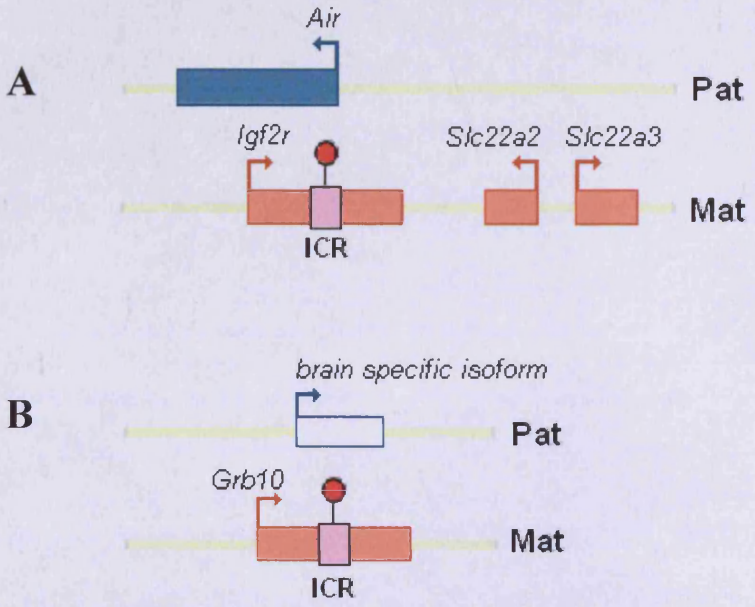


Figure 3.12: Organisation of the mouse *Igf2r* and *Grb10* loci and location of the differentially methylated regions.

A *Igf2r* is predominantly expressed from the maternal allele. Maternal methylation of the intronic DMR prevents expression of the paternally expressed antisense non-coding transcript *Air* and allows expression of *Igf2r*.
B Maternal expression of the predominant isoform *Grb10δ* is found and maintained in all tissues. The germ line DMR is methylated on the maternal allele. A brain specific isoform of *Grb10* is expressed from a second CpG island to give paternal expression of *Grb10*.
Orange boxes are maternally expressed genes, blue boxes are paternally expressed genes and arrows indicate transcription. Pink boxes with lolly pops indicate location of a methylated DMR. Pat = paternal allele, Mat = maternal allele.

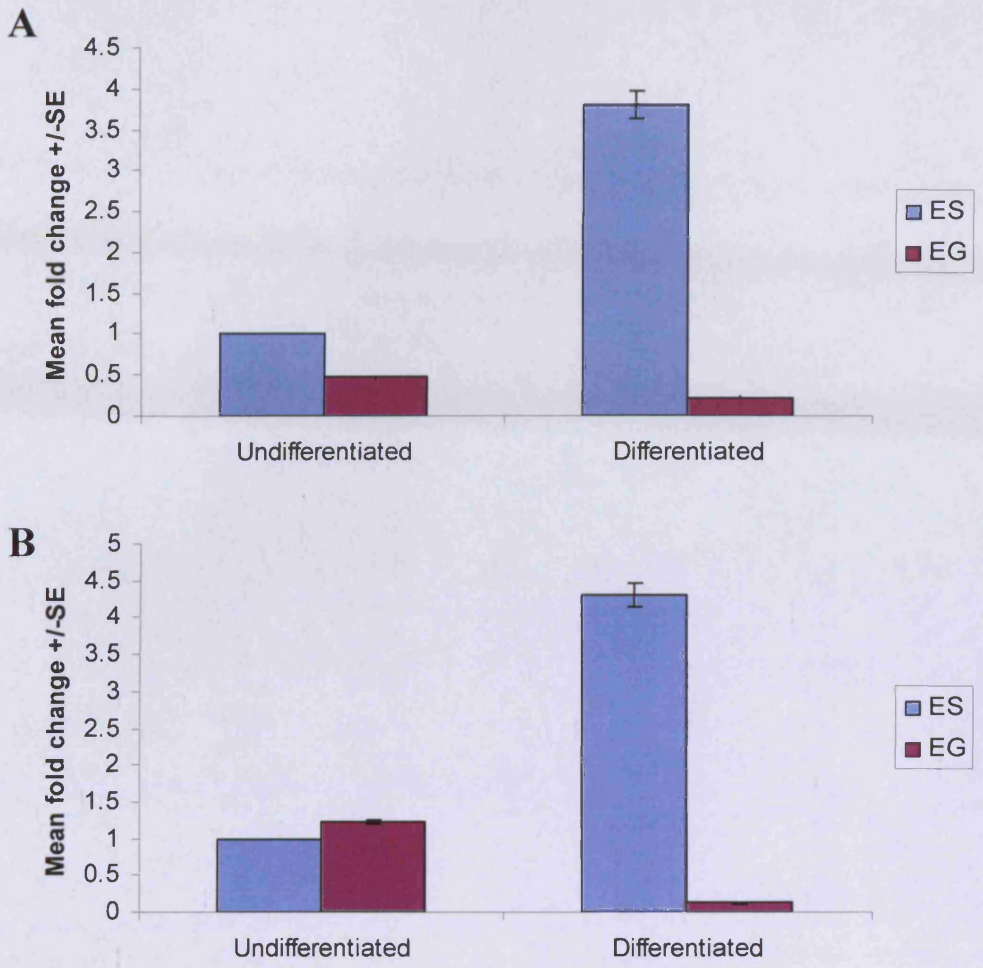


Figure 3.13: Expression of *Igf2r* and *Grb10* in differentiated CES3 ES and Sv6.1 EG cells.

Expression of the imprinted gene **(A)** *Igf2r* and **(B)** *Grb10* in undifferentiated and differentiated ES and EG cells using the delta delta CT method described in chapter 2. The relative expression of each gene at each time point in ES and EG cells was calculated by, firstly comparing the expression level at day 14 in ES cells to day 0 in ES cells. The expression in EG cells at day 0 and day 14 compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

3.2.8 Paternally expressed genes *Nnat, Zac1, Gnas Ex1A, Gnasxl, Dlk1 and Igf2*

Zac1 encodes a zinc finger transcription factor that is found to play a role in embryonic growth, possibly by interacting with other imprinted genes such as *Cdkn1c* (Valente *et al.* 2005). *Zac1* and *Cdkn1c* appear to be expressed in many of the same tissues (Arima *et al.* 2005). There is a zinc-finger motif within the Zac1 protein which suggests that it has DNA binding capacity (Arima *et al.* 2005; Varrault *et al.* 2006). Human *CDKN1C* contains two ZAC1 binding sites and there are eight potential binding sites within the KvDMR1 CpG island, possibly indicating a role for ZAC1 in regulating *CDKN1C* imprinted expression. Human ZAC1 has been shown to bind to unmethylated *KCNQ1OT1/LIT1* promoter and can function as a transcriptional activator acting on the unmethylated paternal allele *in vivo* (Arima *et al.* 2005). *Zac1* is located on mouse chromosome 10 and is expressed in a variety of embryonic and adult mouse tissues. It is found physically linked to one other imprinted gene *Sgce/Hymai*, a non-coding RNA that has no known function (Arima *et al.* 2006; Arima *et al.* 2005; Varrault *et al.* 2006) (**figure 3.14**).

Neuronatin (*Nnat*) is a paternally expressed gene involved in neural development and differentiation (Kikyo *et al.* 1997). It is an isolated imprinted gene found on mouse distal chromosome 2 within the biallelically expressed gene *BC10* (John *et al.* 2001b). From an early stage, *Nnat* expression is restricted to neural tissue in the embryo and loss of *Nnat* results in decreased cerebella folding (Kikyo *et al.* 1997).

Lee *et al.* suggested a model whereby, in the absence of an imprint, some maternally expressed genes are silent and paternally expressed genes are biallelically expressed (Lee *et al.* 2002). *Nnat* was expressed in EG cells at two fold the level

found in ES cells in both undifferentiated and differentiated cells, which suggested that this gene is expressed in the absence of an imprint. *Zac1* appeared to be suppressed in the undifferentiated cells. After differentiation, we observed two fold expression in EG cells compared with ES cells indicating that this gene is also expressed in the absence of an imprint. There was a four fold increase in *Zac1* expression between day 0 and day 14 in the ES cells indicating that, like *Cdkn1c*, *Zac1* is upregulated during differentiation (**figure 3.15**).

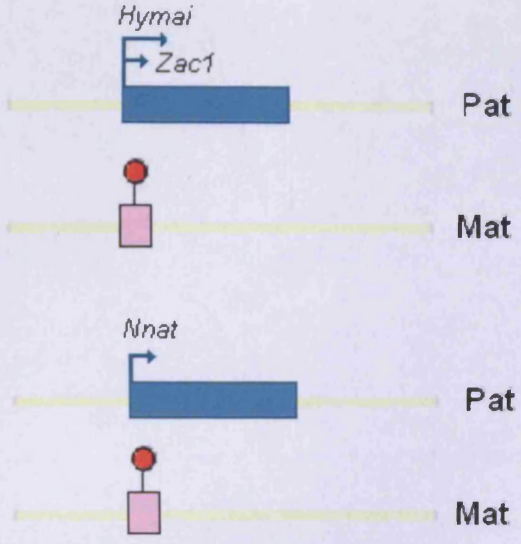


Figure 3.14: Organisation of the mouse *Zac1* and *Nnat* loci and location of the differentially methylated regions.

A *Zac1* and its alternative transcript *Hymai* are predominantly expressed from the paternal allele. Maternal methylation of the DMR located within exon 1 of the non-coding gene *Hymai* is associated with silencing of both *Zac1* and *Hymai.*
B *Nnat* is paternally expressed. Maternal methylation of the DMR is associated with suppression of *Nnat.*
Blue boxes are paternally expressed genes and arrows indicate transcription. Pink boxes with lolly pops indicate location of a methylated DMR. Pat = paternal allele, Mat = maternal allele.

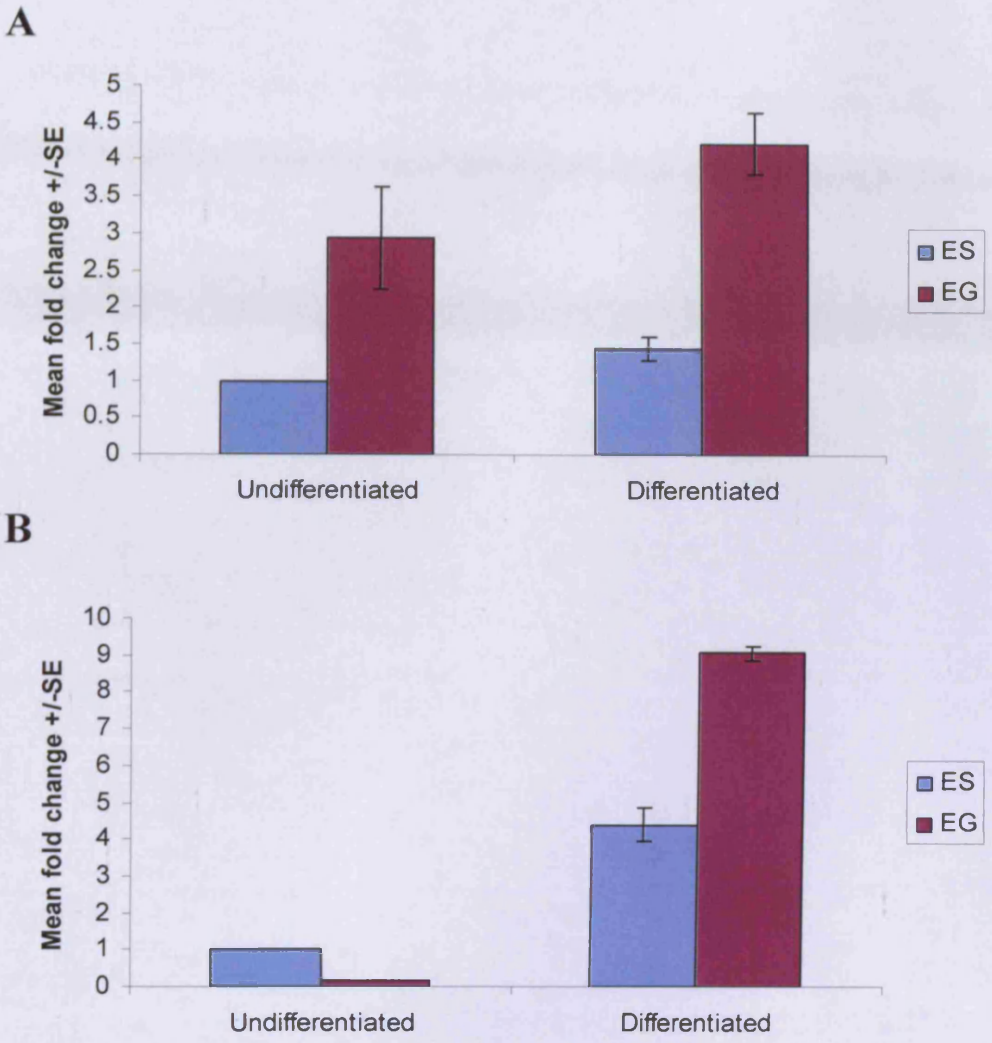


Figure 3.15: Expression of *Nnat* and *Zac1* in undifferentiated and differentiated CES3 ES and Sv6.1 EG cells.

Expression of the imprinted genes *Nnat* (A) and *Zac1* (B) in undifferentiated and differentiated ES and EG cells using the delta delta CT method described in chapter 2. The relative expression of each gene at each time point in ES and EG cells was calculated by, firstly comparing the expression level at day 14 in ES cells to day 0 in ES cells. The expression in EG cells at day 0 and day 14 compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

The *Gnas* cluster, found on mouse distal chromosome 2, has been shown to be involved in regulating neonatal behaviour and growth (Holmes *et al.* 2003) (**figure**

3.16). The cluster contains a number of imprinted genes such as *Nesp55* and *Gnas* which play different roles in embryonic and neonatal development. *Nesp55* also has a role in adult behaviour. The *Gnas* locus consists of several imprinted alternatively spliced transcripts, all differing in the first exon, some of which are non-coding and are possibly involved in the regulation of the imprinting domain (Holmes *et al.* 2003). An antisense transcript (*Nespas*) is also found within the *Nesp55* gene and is possibly involved in regulating *Nesp55* expression (Holmes *et al.* 2003). There are three DMRs found within the promoter regions of *Nesp55, Nespas* and *Exon1A* however only the *Nesp55* DMR is believed to be the ICR for this region (Liu *et al.* 2000). *Gnasxl* is one of the alternative splice forms of *Gnas*, has a role in growth and is paternally expressed. The protein is structurally identical to the $Gnas_{s\alpha}$ protein but has a longer amino terminal extension and is involved in regulating cAMP synthesis (Holmes *et al.* 2003). *Gnas Exon1A* generates an untranslated transcript from the paternal allele and is ubiquitously expressed in all tissues and possibly has a role in controlling imprinting of this domain. However, deletion of the Exon1A DMR has no effect on imprinting of the *Gnas* domain (Holmes *et al.* 2003; Liu *et al.* 2000).

Expression within the *Gnas* cluster is complicated and involves multiple DMRs and non-coding RNA transcripts. The results from two genes within this cluster are shown in **figure 3.17.** *Gnas Ex1A* expression appeared to remain constant in the ES cell line at both time points whereas in the EG cell line expression increased more than 11-fold in the differentiated cells. This dramatic increase in *Gnas Ex1A* expression levels was unexpected but it could be an indication of its involvement in establishing silencing of the *Gnas* cluster in EG cells (**figure 3.17A**). As with *Zac1* we observed low expression of *Gnas Ex1A* in the undifferentiated EG and ES cells. Both *Zac1* and *Gnas Ex1A* appeared to be expressed in the EG cells only after

differentiation. *Gnasxl* expression was more than 2-fold higher in the EG cell line at both time points (**figure 3.17B**). The levels appeared to decrease slightly in differentiated cells yet still maintained the two-fold difference between ES and EG cells and indicated that *Gnasxl* is expressed in the absence of an imprint.

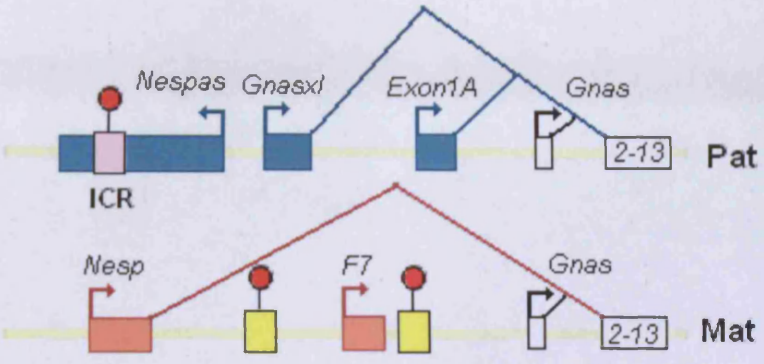


Figure 3.16: Organisation of the *Gnas* cluster and location of the DMRs.

Gnasxl and the two non-coding imprinted genes *Nespas* and *Exon1A* are expressed predominantly from the paternal allele. *Nesp* is the first exon for the Nesp55 protein and is expressed from the maternal allele. *Gnasxl, Exon1A, Gnas* and *Nesp* are alternatively spliced transcripts that share exons 2-13. Differential expression of the imprinted genes is controlled by DMRs located at the promoters of *Nesp*, *Gnasxl* and *Exon1A*. The germ line DMR is found within the *Nesp* promoter. Orange boxes are maternally expressed genes, blue boxes are paternally expressed genes and arrows indicate transcription. White boxes are biallelic expressed genes. Pink boxes with lolly pops indicate location of a methylated DMR. Yellow boxes with lolly pops show location of somatic DMRs.

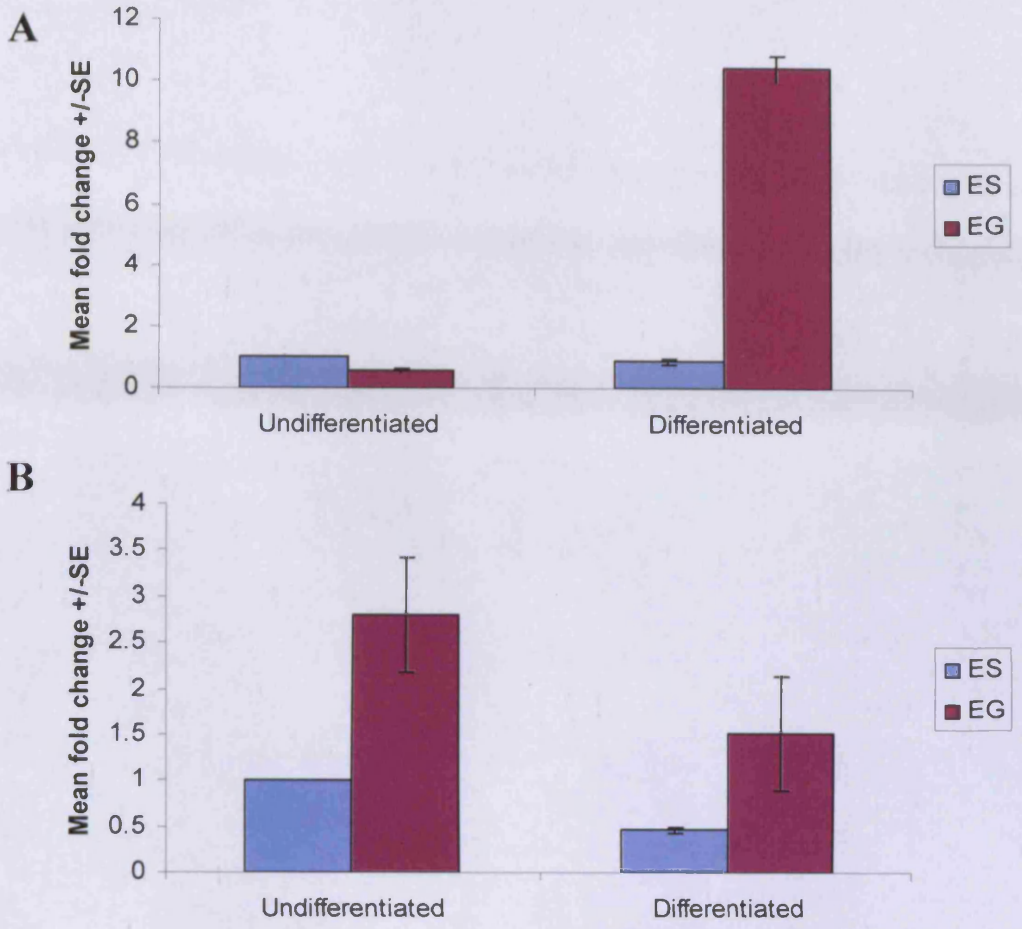


Figure 3.17: Expression of *Gnas Ex1A* and *Gnasxl* in undifferentiated and differentiated CES3 ES and Sv6.1 EG cells.

Expression of the imprinted genes *Gnas Ex1A* (A) and *Gnasxl* (B) in ES and EG cells using the delta delta CT method described in chapter 2. The relative expression of each gene at each time point in ES and EG cells was calculated by, firstly comparing the expression level at day 14 in ES cells to day 0 in ES cells. The expression in EG cells at day 0 and day 14 compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

The *Dlk1* paternally expressed gene is found within an imprinting cluster on mouse distal chromosome 14 and includes the genes *Gtl2, Peg11* and *Dio3* (Takada *et al.* 2002). Reciprocal imprinting of *Gtl2* and *Dlk1* is regulated by a paternally methylated DMR that lies between the two genes (Hiura *et al.* 2007). *Dlk1* is

expressed in numerous embryonic tissues including the placenta. *Gtl2* is a non-coding RNA transcript that possibly plays a role in regulating imprinting of *Dlk1* (Lin *et al.* 2003; Takada *et al.* 2002) (**figure 3.18 A**).

The *H19-Igf2* cluster is found within the mouse distal chromosome 7 imprinting domain and contains a number of imprinted genes involved in embryonic growth. The mechanism of imprinting of this domain is described in detail in chapter 1.2.1. Briefly, a paternally methylated DMR lies within the promoter region of the non-coding RNA transcript *H19* which acts as a transcriptional switch between *H19* and *Igf2* expression (Davis *et al.* 1999; Leighton *et al.* 1995; Thorvaldsen *et al.* 1998). Methylation of the DMR correlates with *H19* expression and suppression of *Igf2* (Thorvaldsen *et al.* 1998) **(figure 3.18 B)**.

The Lee study (Lee *et al.* 2002) demonstrated that *Gtl2, Dlk1* and *H19* were exceptions to the rule that paternally expressed genes are found at biallelic levels and maternally expressed genes are silent in imprint erased PGCs. Our results for *Dlk1* are similar to the Lee study as, in the absence of the paternal imprint, this gene is suppressed compared to that seen in ES cells (**figure 3.19A**). *Igf2* is found within the same cluster of imprinted genes as *H19*. Paternal methylation of the ICR results in expression of *Igf2* and suppression of *H19*. In the absence of an imprint we would expect biallelic levels of *H19* (a maternally expressed gene) and suppression of *Igf2* (a paternally expressed gene). However, Lee *et al.* showed that *H19* was not found at biallelic levels in non-imprinted PGCs. We did not see any suppression of *Igf2* in this study, in fact there was increased *Igf2* expression in the EG cell line (**figure 3.17B**). Both of these genes appeared to be expressed only after differentiation of the stem cells as embryoid bodies.

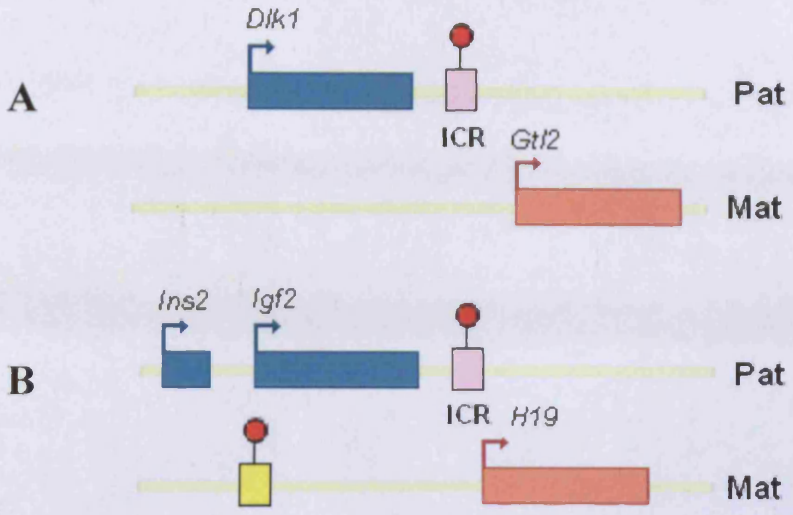


Figure 3.18: Organisation of the *Dlk1-Gtl2* and *H19-Igf2* imprinting clusters and the location of the DMRs.

A Imprinting of the *Dlk1-Gtl2* cluster is regulated by an intragenic DMR located upstream of the *Gtl2* promoter. Paternal methylation of the ICR results in paternal expression of *Dlk1* and silencing of the maternally expressed gene *Gtl2*.
B The ICR for the *Igf2-H19* cluster lies within the promoter region of the *H19* gene. Paternal methylation results in silencing of the maternal *H19* non-coding transcript and expression of paternal *Igf2*. Maternal *Igf2* expression is further suppressed by maternal methylation of the *Igf2* promoter.
Orange boxes are maternally expressed genes, blue boxes are paternally expressed genes and arrows indicate transcription. Pink boxes with lolly pops indicate location of a methylated DMR. Yellow boxes with lolly pops show location of somatic DMRs.

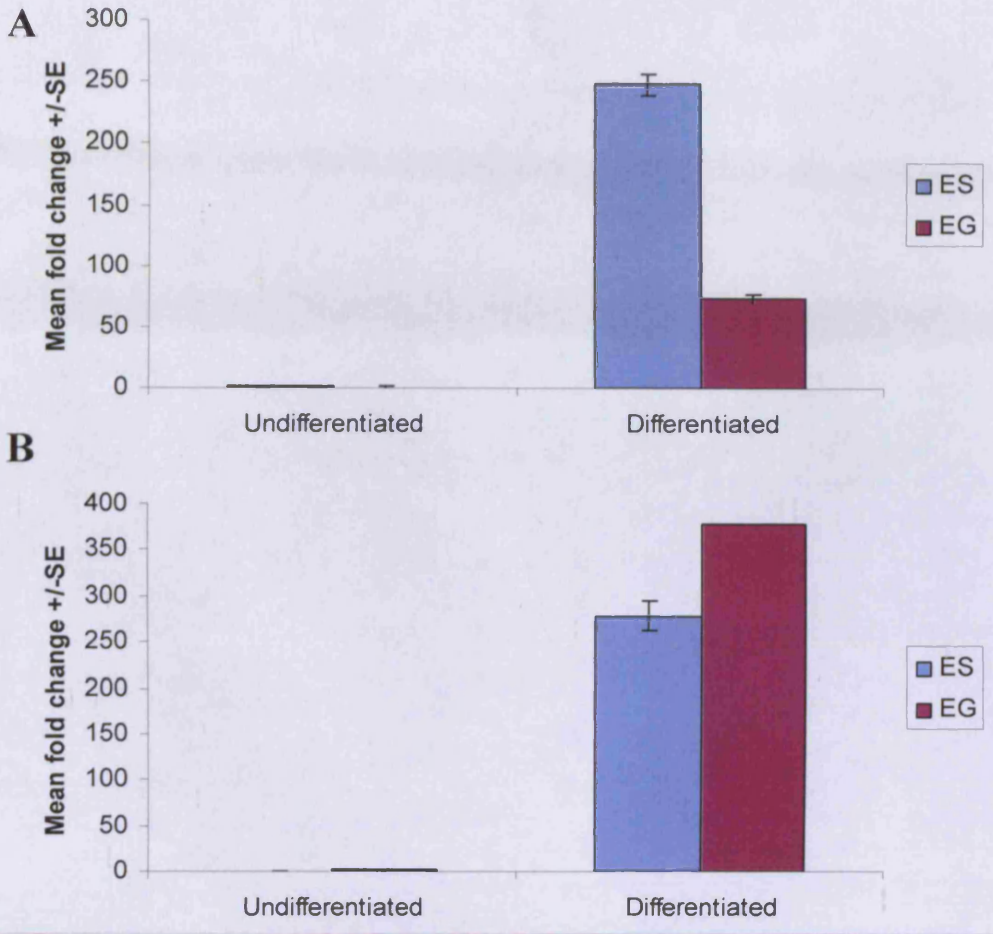


Figure 3.19: Expression of *Dlk1* and *Igf2* in undifferentiated and differentiated CES3 ES and Sv6.1 EG cells.

Expression of the imprinted genes *Dlk1* **(A)** and *Igf2* **(B)** in undifferentiated and differentiated ES and EG cells using the delta delta CT method described in chapter 2. The relative expression of each gene at each time point in ES and EG cells was calculated by, firstly comparing the expression level at day 14 in ES cells to day 0 in ES cells. The expression in EG cells at day 0 and day 14 compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

We have shown here that the maternally expressed genes are suppressed in the EG cell lines whereas the paternally expressed genes are expressed at more than 2-fold levels compared to ES cells with the exceptions of *Igf2* and *Dlk1*. A summary of

the data is shown in **table 3.3**. These results were similar to that reported *in vivo* (Lee *et al.* 2002). This indicated that *in vivo* differentiated EG cells were a good model for determining the expression status of any imprinted gene in the absence of an imprint.

Table 3.3: Summary of the expression of imprinted genes in undifferentiated and differentiated EG cells.

The summary of the expression of a number of imprinted genes in undifferentiated and differentiated EG cells calculated relative to expression in ES cells. Suppressed meant that the expression of the gene was less than that in ES cells. Fold changes are the expression relative to expression in the ES cells so 1-fold = same, 2-fold = approx. biallelic levels.

Gene	Parental origin of expression	Germ line DMR	Expression in imprint-erased EGs	Expression in differentiated EGs	Expressed in absence of imprint
Cdkn1c	Mat	Mat	Suppressed	Suppressed	No
Phlda2	Mat	Mat	Suppressed	Suppressed	No
Igf2r	Mat	Mat	Suppressed	Suppressed	No
Grb10	Mat	Mat	~ 1-fold	Suppressed	No
Kcnq1ot1	Pat	Mat	~ 2-fold	~ 2-fold	Yes
Nnat	Pat	Mat	~3-fold	~3-fold	Yes
Zac1	Pat	Mat	Suppressed	2-fold	Yes
Gnas Ex1A	Pat	Mat	Suppressed	11-fold	Yes
Gnas xl	Pat	Mat	~3-fold	~3-fold	Yes
Dlk1	Pat	Pat	Suppressed	Suppressed	No
Igf2	Pat	Pat	~ 1-fold	~1.4-fold	?

3.2.9 Confirmation of expression results in second EG cell line TMAS21G

A control EG cell line was used to confirm our results. The relative expression of a number of imprinted genes was determined and compared to that seen in the ES and EG cell lines used earlier (**figure 3.14**). All of the imprinted genes that were silent in the Sv6.1 cell line were similarly suppressed in the absence of an imprint after differentiation of the TMAS21G stem cells for 14 days. There were some slight differences in the absolute expression levels between the two cell lines such as a

lower *Cdkn1c* expression in the TMAS21G cell line compared to the Sv6.1 cell line. The TMAS21G cell line was from a different genetic background which may account for the differences between the two cell lines. The TMAS21G cell line also contains a Rosa 26 reporter gene that could somehow affect expression of the imprinted genes.

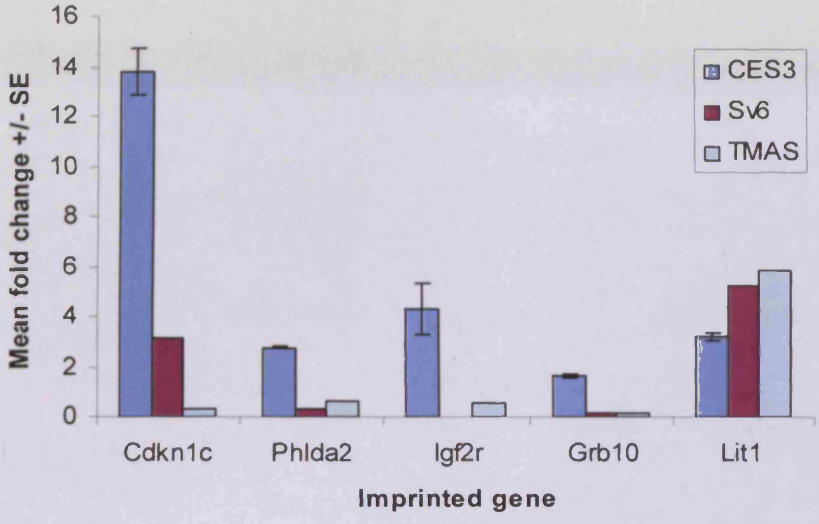


Figure 3.20: Relative expression of *Cdkn1c*, *Phlda2*, *Igf2r* , *Grb10* and *Kcnq1ot1/Lit1* in differentiated CES3, Sv6.1 and TMAS21G EG cells.

Embryoid bodies were differentiated for 0 and 14 days. The fold change difference between ES and EG cells was first calculated using the delta delta CT method for each imprinted gene as described in chapter 2. The relative expression of each gene at each time point in ES and EG cells was calculated by, firstly comparing the expression level at day 14 in ES cells to day 0 in ES cells. The expression in EG cells at day 0 and day 14 compared with ES cells was then calculated. The two values were then multiplied together to give the relative expression in EG cells compared to ES cells. Standard error is shown for the ES cell data. At least two control genes (Actin, GAPDH or 18s rRNA) were included in the calculation.

3.3 Discussion

As described in section 3.1, epigenetic studies of silencing in the germ line usually involve using primordial germ cells taken from early stage mouse embryos. As these PGCs are few in number, it is technically difficult to isolate large numbers in order to perform such in-depth analyses of specific imprinting domains. We, therefore, aimed to determine if *in vitro* differentiated EG cell lines could be a useful model for studying the expression of imprinted genes from non-imprinted chromatin.

3.3.1 Comparing EG and ES cells

EG cells provide a convenient model system in which we can study silencing as they can be grown in culture in large numbers and are developmentally similar to ES cells (Tada *et al.* 1998). Our analysis showed that three genes within the IC2 domain in both ES and EG cells show similarities in the timing of expression during differentiation.

Cdkn1c and *Phlda2* showed similar expression patterns between ES and EG cells implying that, as the cells differentiate, the imprinted genes were expressed at the same time points in both the ES and EG cells indicating that the cells were comparable. In this study we have to presume that the expression of the imprinted genes in the ES cells was from one allele only. As there are no polymorphisms in this cell line we were unable to confirm this. Many studies have reported the expression of imprinted genes in ES cells. Lewis *et al.* have shown that, in ES cells, *Phlda2* and *Cdkn1c* are expressed monoallelically whereas *Osbpl5* and *Tssc4* are expressed biallelically. Only basal levels of *Kcnq1* and *Ascl2* were detected in ES cells (Lewis *et al.* 2006). This concurs with our results where *Kcnq1* expression could not be detected by real time PCR. We also analysed *Kcnq1ot1/Lit1* which appeared to be expressed at

constant levels throughout differentiation in both ES and EG cells. *Kcnq1ot1/Lit1* has been shown to be expressed from the two-cell embryo stage in mice and to be paternally expressed (Lewis *et al.* 2006).

3.3.2 *Cdkn1c* and *Phlda2* expression in EG and ES cells

The only difference between undifferentiated ES cells and undifferentiated EG cells is thought to be the imprinting status of the genome. ES cells show full imprinting of the IC2 domain resulting in allele-specific gene expression. Conversely, the EG cells are derived from imprint erased PGCs so should show no differences between the maternal and paternal genomes that would result in allele-specific gene expression. We predicted that we would see 2-fold levels of gene expression if the imprinted gene was expressed from both alleles. This then allowed us to determine if a particular gene was expressed in the EG cells by comparing its expression relative to that in the ES cell line (equivalent to one-fold for monoallelic expression). Generally, if the gene was silent in the EG cells then expression would be less than the ES cells but if the gene is active then expression will be greater than that in the ES cells.

Our study showed that there was a significant reduction in the level of expression of the *Cdkn1c* gene in the EG cells indicating that this gene was suppressed in some way. This suppression of gene expression was present in the undifferentiated cells and maintained throughout differentiation. A similar result was observed for the *Phlda2* gene however, expression from this gene is relatively stable in ES cells at all time points compared to *Cdkn1c* which showed a dramatic increase in expression from day 7. *Phlda2* is highly expressed in the extraembryonic lineages and is only expressed in the embryonic kidney and liver (Frank *et al.* 2002) whereas *Cdkn1c* plays a role in controlling the cell cycle and in differentiation and is expressed

in a variety of embryonic tissues (Hatada and Mukai 1995; Zhang *et al.* 1997). Our analysis of the early time points showed that *Cdkn1c* and *Phlda2* were suppressed throughout differentiation of EG cells from the very early time points indicating that these genes are silent in the EG cells and remain silent throughout differentiation.

3.3.3 *Kcnq1ot1/Lit1* expression in ES and EG cells

Our analysis showed that *Kcnq1ot1/Lit1* expression was at equivalent to biallelic levels in the EG cells and this was maintained throughout differentiation. However, when we examined the early time points, we observed dynamic changes in *Kcnq1ot1/Lit1* expression in the ES and EG cells. The two different expression patterns in these cells could possibly relate to the potential different functions of *Kcnq1ot1/Lit1* in the two cell lines. For example, establishing silencing in the EG cells and maintaining silencing in the ES cells. Recent data has shown that transcription of full length *Kcnq1ot1/Lit1* is required for normal imprinting in the germ line (Mancini-DiNardo *et al.* 2006; Murakami *et al.* 2007; Shin *et al.* 2008). Deletion of the KvDMR1 imprinting centre or the *Kcnq1ot1/Lit1* promoter results in biallelic expression of at least six imprinted genes within the IC2 domain and loss of *Kcnq1ot1/Lit1* expression, a result similar to that seen in *Dnmt1-/-* stem cells (Fitzpatrick *et al.* 2002; Horike *et al.* 2000; Mancini-DiNardo *et al.* 2006). Deletion of the ICR has no effect on imprinting of the IC1 domain (Horike *et al.* 2000; Murakami *et al.* 2007). Deletion of the human paternal *KCNQ1OT1/LIT1* gene in BWS patients is also associated with loss of imprinting of *CDKN1C* (Niemitz *et al.* 2004).

The *Kcnq1ot1/Lit1* transcript itself is also important as truncation of the full length transcript results in loss of silencing of all KvDMR1-controlled genes in both the placenta and in many embryonic lineages (Kanduri *et al.* 2006; Mancini-DiNardo

et al. 2006; Shin *et al.* 2008). This is similar to the observation by Sleutels *et al.* in which premature termination of the *Air* transcript resulted in biallelic expression of the imprinted genes (Sleutels *et al.* 2002). The length of the *Kcnq1ot1/Lit1* truncation has even been shown to give different phenotypes. Mancini-DiNardo *et al.* showed that a 1.5 kb truncation gave a similar result as deleting the KvDMR1, whereas Shin *et al.* showed that a longer 2.6 kb transcript resulted in loss of imprinting of most of the IC2 domain imprinted genes but *Cdkn1c* maternal-specific expression was maintained in several embryonic tissues including kidney, lung and liver but lost in brain, placenta and heart (Mancini-DiNardo *et al.* 2006; Shin *et al.* 2008). The 2.6 kb truncation mutant also resulted in a less severe growth deficiency at 4 weeks of age compared to the KvDMR1 knockout, possibly as a result of retention of normal *Cdkn1c* expression levels in some tissues (Shin *et al.* 2008). The difference in the two phenotypes could be due to the loss of other sequences in the KvDMR1 that are important for *Cdkn1c* silencing or because the longer transcript is capable of paternal silencing of *Cdkn1c* in some tissues whereas the shorter transcript is not. Shin *et al.* showed loss of methylation at the *Cdkn1c* promoter in the truncation mutants irrespective of whether the gene retained imprinting or not. Their results suggested that the 2.6 kb transcript was sufficient to maintain imprinting of *Cdkn1c* in a subset of tissues in the absence of DNA methylation (Shin *et al.* 2008). Kanduri *et al.* also demonstrated that different lengths of the *Kcnq1ot1/Lit1* transcripts have different effects on silencing of a transgene, particularly in the ability to initiate bi-directional silencing (Kanduri *et al.* 2006). They postulate that the longer time taken for the RNA polymerase to transcribe the full length *Kcnq1ot1/Lit1* gene allows the nascent transcript to interact with the local chromatin machinery, providing an opportunity for the transcript to spread inactive chromatin structures in *cis* (Kanduri *et al.* 2006).

Other studies have shown that *Kcnq1ot1/Lit1* co-localises with the IC2 domain during the cell cycle at least on the *Slc22a18* and *Cdkn1c* regions which are several kb outside of its transcriptional region, possibly indicating a role in silencing these genes (Mancini-DiNardo *et al.* 2006; Murakami *et al.* 2007). This same phenomenon has been observed with another imprinted gene cluster, *Igf2r-Air,* where co-localisation of the non-coding RNA *Air* transcript has been observed along the silent domain (Seidl *et al.* 2006). This is analogous to that seen during X chromosome inactivation where the non-coding RNA *Xist* coats the inactive X chromosome (Clemson *et al.* 1996). These results strongly suggest that non-coding RNAs such as *Air* and *Kcnq1ot1/Lit1* have an active role in establishing silencing within their domains although how the non-coding transcript mediates this silencing in *cis* is still unknown.

Several mechanisms have been proposed for how *Kcnq1ot1/Lit1* is involved in the silencing of the IC2 domain. Firstly, *Kcnq1ot1/Lit1* could participate in a RNA-dependent transcriptional silencing mechanism where the non-coding RNA attracts DNA methyltransferases and repressive chromatin proteins to the domain in a similar way that *Xist* is involved in recruiting DNA methylation and repressive chromatin modifications to the inactive X chromosome (Kanduri *et al.* 2006; Mancini-DiNardo *et al.* 2006). Secondly, as transcription of the full length RNA is required for silencing, it could be that repression of the IC2 domain is occurring because of the movement of RNA polymerase through the gene causing topological changes that attracts repressive chromatin (Kanduri *et al.* 2006; Mancini-DiNardo *et al.* 2006). Finally, it could be that an RNA interference mechanism is responsible for silencing the domain by processing the *Kcnq1ot1/Lit1* transcript into small RNAs that then base pair with the mRNA or DNA of the imprinted genes within the cluster (Mancini-DiNardo *et al.* 2006). The dynamic changes in *Kcnq1ot1/Lit1* expression that we

observe at the early time points could be evidence of this RNA processing involved in setting up the early silencing of *Cdkn1c* and *Phlda2*. Whatever the mechanism, it is likely that DNA methylation and repressive chromatin are involved in propagating silencing along the domain allowing silencing to be stably inherited and maintained.

3.3.4 Establishing silencing in PGCs

We considered the possibility that the suppression in *Cdkn1c* and *Phlda2* expression that we observed could be a result of the normal silencing mechanism already initiated in male PGCs cells and would result in establishing silencing of the IC2 domain. The EG cells used in this study were derived from 12.5 dpc male PGCs and several studies have shown that there are no differences between male and female PGCs at this stage of PGC development (Adams and McLaren 2002; Tada *et al.* 1998). PGCs derived from mouse embryos before 13.5 dpc have the developmental potential to differentiate as either oocytes or as prospermatogonia. After 13.5 dpc, signals from the somatic cells in the male genital ridge direct the PGCs to a spermatogenic fate but in the absence of this masculising signal PGCs develop as oocytes (Adams and McLaren 2002; McLaren and Southee 1997). The differentiation of PGCs as oocytes or prospermatogonia depends, not on the sex of the PGCs, but on the sex of the supporting stromal cells in the gonad as 11.5 dpc female PGCs can develop as prospermatogonia when cultured as an aggregate with a male genital ridge (Adams and McLaren 2002). Also, male PGCs will develop as oocytes when cultured with a female genital ridge, a result similar to that seen when male PGCs are cultured with lung aggregates (McLaren and Southee 1997). However, after 12.5 dpc male PGCs become committed to prospermatogonia development, whereas female PGCs remain sexually dimorphic up to 13.5 dpc (Adams and McLaren 2002). Tada *et al.*

(Tada *et al.* 1998) also demonstrated that male and female EG cell lines were equivalent by creating chimeras then characterising the differences in epigenotype. They showed that the 12.5 dpc EG cells were capable of forming chimeras and showed the same developmental abnormalities irrespective of the sex of the EG cell line. PGCs taken from 13.5 dpc onwards will continue to pursue their germ line pathway of differentiation for several days when cultured as EG cells, suggesting that differentiation as oocytes and prospermatogonia is a cell-autonomous process once the mascularising signal has been received and sex-specific differences have begun to arise. Site-specific methylation in male germ cells, but not female germ cells, is also autonomous and it may be that the number of X chromosomes or the presence of the Y chromosome somehow controls this process (Durcova-Hills *et al.* 2004; Durcova-Hills *et al.* 2006).

From what we know about 12.5 dpc PGCs, the suppression of *Cdkn1c* expression in differentiated EG cells is unlikely to be due to paternal imprinting of the genome as the cells do not receive the required mascularising signal. Instead we suspect that we are observing a default state for silencing that is established in the absence of signals from the genital ridge.

3.3.5 Other imprinted genes

Our observations of the IC2 cluster are not unique. Indeed, we observe a similar pattern for a number of imprinted genes, both maternal and paternal, that are silent in the absence of an imprint. We also identified some genes where the opposite is true, they are expressed in the absence of an imprint and require passage through the germ line to be silent. This data also confirms that the cells are not establishing paternal specific imprints as some of the genes examined, such as *Dlk1,* are paternally

expressed yet their expression is suppressed in differentiated EG cells indicating that they require passage through the male germ line to be expressed. The data also confirms what has been noted before that there are two types of imprinted genes, those that require an imprint to be expressed and those that require an imprint to be silenced.

Several studies have looked at the expression of imprinted genes in imprint-erased PGCs. As mentioned in the introduction Lee *et al.* (Lee *et al.* 2002) demonstrated the existence of a default state of gene expression in imprint erased PGCs where, in the absence of an imprint, maternally expressed genes are silent and most paternally expressed genes are biallelically expressed. The exceptions were *H19, Meg3/Gtl2* and *Dlk1* which require a paternal imprint to be expressed. Other studies have shown similar results. For example, in 11.5 dpc PGCs, *Igf2r, H19, Igf2* and *Snrpn* have been shown to be expressed biallelically (Szabo *et al.* 2002; Szabo and Mann 1995). Analysis of imprinted gene expression has also been conducted at 12.5-15.5 dpc at a stage when sex-specific differences arise and imprinting normally begins to be established (Szabo and Mann 1995). In the ovary and testis at this stage, biallelic expression of *Igf2r* and *Snrpn* is observed but monoallelic expression of *H19* and *Igf2* was seen implying that the *H19* imprint is established differently than at *Igf2r* and *Snrpn* (Szabo and Mann 1995). In pre-migratory PGCs the expression of the imprinted genes is skewed. The majority of the genes appear to be monoallelically expressed as imprints have yet to be erased apart from *Igf2r* which appears to be biallelically expressed at all PGC time points (Szabo *et al.* 2002; Szabo and Mann 1995). However, our data showed that in the absence of an imprint *Igf2r* is suppressed. It is possible that, in the 11.5 dpc PGCs, the imprints have not been fully erased for the *Igf2r* locus. From what we know about this domain, methylation of the

DMR results in silencing of the non-coding RNA transcript *Air* and expression of the surrounding imprinted genes. We would expect from this information that in the absence of an imprint (i.e. the maternal methylation mark) there would be biallelic expression of the *Air* transcript and silencing of *Igf2r*.

H19 and *Igf2* are expressed reciprocally so it seemed unlikely that we would observe monoallelic expression of both imprinted genes in the imprint erased EG cells. From the model of this domain, paternal methylation of the *H19* DMR induces a transcriptional switch to *Igf2* expression. In the absence of paternal methylation of the DMR we could expect that the domain would take on the appearance of the maternal allele. So, in imprint-erased PGCs, we might expect biallelic expression of *H19* and suppression of *Igf2*. However, studies have indicated that *H19* silencing is the default state and only becomes activated after passage through the female germ line (Ainscough *et al.* 1997). In our study we observed levels of *Igf2* that are around 1.4-fold so could be an indication that this gene is biallelically expressed. This would confirm the results shown in Szabo *et al.* 1995 and Szabo *et al.* 2002. As we do not see 2- fold levels of expression of *Igf2,* it is possible that the imprint for this domain has not been completely erased, even in the 12.5 dpc PGCs, or has already begun to be re-established.

What we have noted from our study is that, for a number of imprinted genes, the expression levels in the undifferentiated ES and EG cells are significantly lower than in the differentiated cells indicating that they are not expressed in the stem cells. *Igf2* is one such imprinted gene where the expression in the differentiated cells increases more than 250-fold indicating that this gene is upregulated during differentiation. It is possible that the expression levels in the PGCs, shown by Szabo and Mann 1995 and Szabo *et al.* 2002, could be leaky expression from the gene and in

fact the imprinted gene is not expressed. This demonstrates the usefulness of EG cells for these sorts of studies as they can be differentiated *in vitro* to demonstrate the actual status of the gene in the absence of an imprint even for genes that are not expressed in PGCs or stem cells.

An alternative method for looking at the status of the imprint-erased genes is using parthenogenetic embryos containing the pronuclei from a non-growing (ng) oocyte (imprint-erased) and a pronucleus from a fully grown (fg) oocyte (maternal imprints) (Obata *et al.* 1998). These parthenogenetic embryos survive to around 13.5 dpc, significantly longer than the parthenogenetic embryos created from two fully imprinted maternal pronuclei (Obata *et al.* 1998). Analysis of the imprinted genes has shown that the paternally expressed genes *Peg1* and *Peg3* are expressed from the non-growing oocyte genome allele indicating that these genes are expressed in the absence of an imprint and are silenced after passage through the maternal germ line (Obata *et al.* 1998). The maternally expressed genes *Igf2r* and *Cdkn1c* are expressed from the fully grown oocyte genome allele only confirming our results that, in the absence of an imprint, these genes are not expressed and are only expressed after passage through the maternal genome. *H19* was expressed from both alleles as *H19* is a maternally expressed gene that is silenced by passage through the male germ line (Obata *et al.* 1998). This data corroborates the results shown in our study where *Igf2r* and *Cdkn1c* are not expressed in undifferentiated EG cells and remain suppressed after differentiation of the EG cells as they have not passed through the maternal germ line which is required for their activation. *Peg1* and *Peg3* are paternal expressed genes so we might predict these genes to be at biallelic levels in the differentiated EG cells as they require a maternal imprint to be silenced. This study is more comparable to ours

as it looks at the expression levels of imprinted genes in differentiated cells and not just in the PGCs.

We confirmed that the suppression of maternally expressed genes is not restricted to our cell line by also looking in an alternative cell line that has been extensively studied by others (Tada *et al.* 1998). We observed the same pattern that maternally expressed imprinted genes in the EG cells are suppressed relative to levels in the ES cells.

The majority of the imprinted genes examined in this study required maternal methylation of the DMR to either be expressed or silenced. The two exceptions that we looked at were the *Dlk1-Gtl2* and the *H19-Igf2* clusters which require paternal methylation. Surani *et al.* suggested that the failure of parthenogenetic embryos to develop to term is due to the loss of specific parental imprinting (Surani *et al.* 1984). This is because the majority of the differential DNA methylation marks are maternal resulting in biallelic expression of all the maternally expressed genes (*Cdkn1c, Igf2r, Grb10*) and silencing of all the paternally expressed genes (*Gnasxl, Zac1, Nnat*), except those regulated by a paternal methylation mark (*Dlk1, Igf2*). In the absence of this maternal methylation the EG cells appear to take on a paternal expression profile in the absence of any imprints that may correspond to a default silencing mechanism of the maternally expressed genes. Why silencing of the maternally expressed genes is more important than the paternally expressed genes is not yet known but it may be related to the fact that a number of the maternally expressed genes are negative regulators of growth including *Cdkn1c* and *Grb10* whereas some paternally expressed genes such as *Igf2* are growth promoting.

3.3.6 Conclusion

Our results have shown that the maternally expressed genes *Cdkn1c* and *Phlda2* are suppressed in EG cells and this silencing is maintained throughout differentiation suggesting that silencing of this imprinted domain is a passive process. The dynamic changes in *Kcnq1ot1/Lit1* expression indicated a possible role in establishing or maintaining this silencing. Our EG cell model has been used to determine whether an imprinted gene is suppressed in imprint erased EG cells that re-capitulates what is observed *in vivo*. From this data we can determine what classification an imprinted gene falls into, whether it requires an imprint to be expressed or an imprint to be silenced.

3.3.7 Summary of findings

- The IC2 domain imprinted genes *Cdkn1c* and *Phlda2* are suppressed in the EG cell line compared to the ES cell line at all time points.
- Suppression of *Cdkn1c* is already established in the imprint erased EG cells and is maintained throughout differentiation suggesting that silencing of the IC2 domain is a passive process.
- There are two types of imprinted genes, those that require an imprint to be expressed and those that require an imprint to be silenced. Our model is able to determine which category an imprinted gene belongs too and re-capitulates what is observed *in vivo*.

CHAPTER 4:

Methylation analysis of IC2 domain differentially methylated regions

4.1 Introduction

Methylation of CpG islands is normally associated with gene silencing. These CpG islands are mainly found within the promoters of genes. They are also associated with imprinted clusters where they are thought to be involved in controlling allelic-specific gene expression. While the CpG islands in these clusters can be associated with the promoter region of an imprinted gene, they can also be part of an ICR. When methylated, these differentially methylated CpG islands can either be involved in activating or in silencing imprinted genes.

The mouse distal chromosome 7 IC2 domain is an imprinted domain which has been shown to require maternal methylation of the KvDMR1 for the maternal expression of the surrounding imprinted genes. Conversely, on the paternal allele the lack of DNA methylation correlates with suppression of the surrounding genes. Three paternal specific regions of DNaseI hypersensitivity are located immediately 5' of the transcriptional start site of *Kcnq1ot1/Lit1* that correspond to the location of the KvDMR1 (Mancini-DiNardo *et al.* 2003). The KvDMR1 lies over the promoter region of the non-coding gene *Kcnq1ot1/Lit1*. DNA methylation plays an obvious direct role in silencing of the *Kcnq1ot1/Lit1* gene, by preventing transcription. Methylation of the KvDMR1 must indirectly affect the other genes. This would also involve recruiting silencing trans-factors such as methyl-CpG binding proteins and histone modifying enzymes to modify the local chromatin environment. However, it is still unclear how localised DNA methylation at the KvDMR1 signals to the distantly located genes.

We know that the KvDMR1 is part of the ICR as deleting this region disrupts silencing of six of the imprinted genes within the IC2 domain (Fitzpatrick *et al.* 2002;

Horike *et al.* 2000; Mancini-DiNardo *et al.* 2003). In mouse ES cells, with a targeted paternal deletion of the KvDMR1 CpG island, expression of the *Kcnq1ot1/Lit1* transcript can no longer be detected. Paternal transmission of this deletion results in biallelic expression of the IC2 imprinted genes due to a lack of imprinting in the male germ line (Fitzpatrick *et al.* 2002; Mancini-DiNardo *et al.* 2006). Normal gene expression is detected when the mutation is inherited maternally suggesting that the KvDMR1 CpG island is required for imprinting in the male, but not the female, germ line (Fitzpatrick *et al.* 2002; Mancini-DiNardo *et al.* 2006). Deletion of the KvDMR1 has no impact on the imprinting of genes in the IC1 domain which is regulated by a separate ICR.

Loss of methylation at ICRs is often associated with loss of imprinting of the surrounding genes suggesting that correct methylation of the DMRs is essential for normal regulation of these genes (Diaz-Meyer *et al.* 2003). Loss of methylation at DMRs has been associated with a number of developmental disorders and cancers. A common cause of Beckwith Wiedemann syndrome is loss of imprinting of *IGF2* and *CDKN1C* which can occur as a result of a disruption to the *H19* and KvDMR1 ICRs either through gain or loss of methylation or because of deletions or translocations, demonstrating the importance of these regions in establishing and maintaining imprinting.

Within the IC2 domain, there is a second DMR located within the promoter region and extending into the second intron of the *Cdkn1c* gene. This is a somatic DMR and is not methylated in the germ line (Bhogal *et al.* 2004). Methylation of this DMR in somatic tissues is thought to be required to completely suppress and maintain silencing of *Cdkn1c* from the paternal allele. Human *CDKN1C* does not have any

methylation and this probably corresponds to the small amount of expression observed from the paternal allele in humans (John *et al.* 1999).

Most of the information we have on how imprints in the germ line are established comes from looking at the methylation patterns in sperm and oocytes. In the germ cells, DNA methylation has been established but many of the events that lead to permanent silencing or activation of the target genes, such as histone modifications, have not taken place. There have been some studies that have looked at earlier time points in PGCs but have mainly focused on examining global DNA methylation patterns and histone modifications to try to determine what epigenetic changes occur *in vivo*. A number of these studies are also involved in looking later on in sex-specific PGC development when imprints have already begun to be established. These studies are technically challenging because isolation of the PGCs from the genital ridges of mouse embryos yields very few cells for the analysis, particularly at early embryonic stages.

Germ line and somatic DMRs are erased in EG cell lines derived from 12.5 dpc PGCs (Lee *et al.* 2002; Tada *et al.* 1998). In the Tada study, they showed that the DMR at the *Cdkn1c* promoter was unmethylated in EG cells derived from 12.5 dpc PGCs but also that this region acquired *de novo* methylation after differentiation in chimeras. None of the other DMRs looked at in the Tada study acquired *de novo* methylation when differentiated (Tada *et al.* 1998). This suggested that germ line DMRs require passage through the germ line to become methylated but methylation of somatic DMRs, like *Cdkn1c* promoter, may occur much later as a result of a default silencing mechanism.

In this study we looked at the methylation status of the *Cdkn1c* and KvDMR1 CpG islands in EG cells differentiated *in vitro*. The aim was to confirm that the germ

line methylation of the KvDMR1 and *Cdkn1c* DMRs was absent in the 12.5 dpc Sv6.1 EG cells and to determine if methylation of *Cdkn1c* was acquired during EG cell differentiation *in vitro*. The key differences between our study and that in the Tada paper is that the EG cells would be cultured and differentiated *in vitro*. In the Tada study, EG cells were differentiated in chimeras to mid-gestation, primary embryonic fibroblasts were isolated from these chimeras and these were then grown in culture for several days before DNA methylation was examined. In our EG experiment the cells were differentiated entirely '*in vitro*'. It was necessary to determine whether EG cells would perform in culture as they had done '*in vivo*' i.e. methylate *Cdkn1c*.

4.2 Results

4.2.1 Methylation sensitive restriction enzyme coupled PCR (MSRE-PCR) analysis of the KvDMR1 and *Cdkn1c* DMRs in ES and EG cell lines

The accurate mapping of methylation in the genome is important for understanding how this methylation is established and maintained. There have been many methods designed to do this, the majority of which are based on Southern hybridisation approaches (Ariel 2002). An alternative way to assess DNA methylation at a few critical CpG dinucleotides is using methylation sensitive restriction enzyme coupled PCR (MSRE-PCR) [for review see (Rein *et al.* 1998)]. This PCR-based methylation analysis requires smaller amounts of DNA than conventional methods and can be technically simple and rapid (Ariel 2002). In this protocol, the genomic DNA is digested with a methylation sensitive restriction enzyme followed by PCR spanning the site. If the CpG dinucleotide is unmethylated, the restriction enzyme cuts the DNA and the sequence can no longer be amplified, however, if the CpG is methylated the DNA remains uncut and the PCR product can be amplified. For this

method to work there must be a restriction site available within the region of interest and, to prevent the appearance of false positives, complete cleavage of the DNA is required. Again the design of the primers is critical to the success of this assay. Nested primers were designed so that one set serves as an internal control.

A methylation sensitive restriction site was identified within the KvDMR1 that also corresponded to a potentially methylated CpG dinucleotide (**figure 4.1**). Primers were designed around this site to amplify a product of 525 bp and the control primers were designed to amplify a product outside this site to amplify a product of 249 bp. Both primer sets were tested on genomic DNA to ensure that they amplified the correct sized product when used together. The method described by Ariel (Ariel 2002) was modified to allow for double restriction enzyme digests as an alternative for dialysis of the digested DNA. The location of the primers and cut sites is shown in **figure 4.2**.

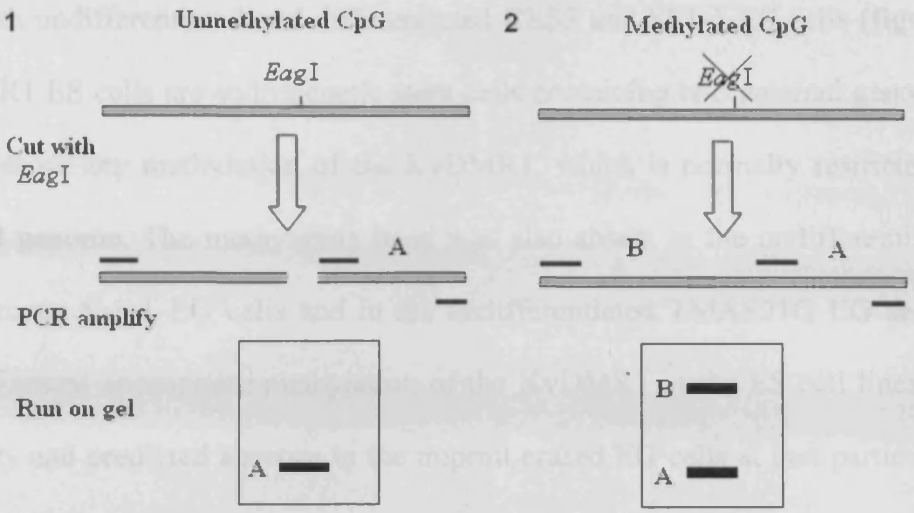


Figure 4.1: MSRE coupled PCR protocol schematic.

Genomic DNA is digested with an enzyme sensitive to the methylation status of the DNA so will not cut when DNA methylation is present. **1** In the absence of DNA methylation the enzyme will cut the DNA and the PCR will fail to amplify product B. **2** If methylation is present then the enzyme will not cut and the PCR will amplify product B. A control PCR (product A) is included that is unaffected by the methylation status of the DNA.

Methylation of KvDMR1 has been studied in depth by a number of groups. Beatty *et al.* showed that KvDMR1 is methylated throughout the maternal CpG island whereas the paternal allele is not methylated (Beatty *et al.* 2006). The MSRE-PCR method was used to determine that the EG cells used in this study had the correct methylation pattern. As a control, ES cells were used which were predicted to be methylated on one allele. The androgenetic cell line was used as a control for the complete absence of maternal DNA methylation. As described in chapter 3 the lack of polymorphisms in our samples meant we could not determine which allele was methylated so we only looked for the presence or absence of methylation in these samples. Undigested samples were included as a further control.

Genomic DNA from ES and EG cells was digested overnight with *Eag*I and used as the template in a PCR reaction. The methylated (undigested) band amplified

from both undifferentiated and differentiated CES3 and SF1-1 ES cells (**figure 4.3**). The AKR1 ES cells are androgenetic stem cells containing two paternal genomes and did not show any methylation of the KvDMR1, which is normally restricted to the maternal genome. The methylation band was also absent in the undifferentiated and differentiated Sv6.1 EG cells and in the undifferentiated TMAS21G EG cells lines. This suggested appropriate methylation of the KvDMR1 in the ES cell lines used in this study and predicted absence in the imprint erased EG cells at that particular CpG dinucleotide sequence.

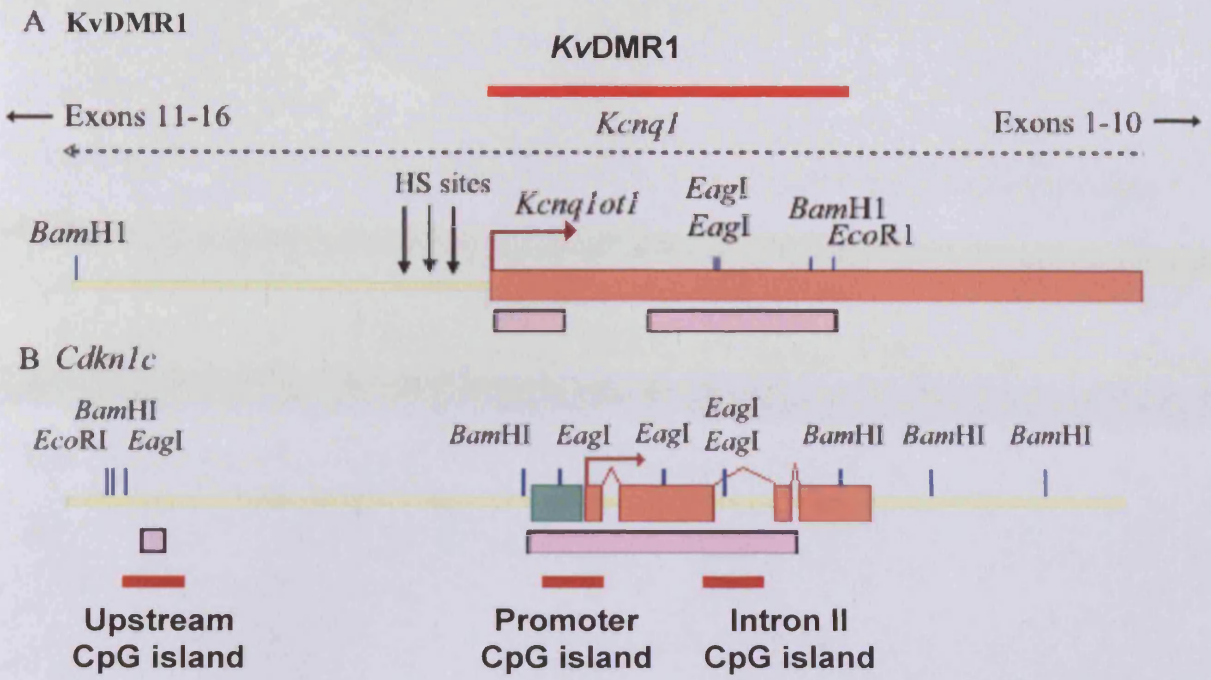


Figure 4.2: Location of CpG islands within *Kcnq1ot1/Lit1* and *Cdkn1c*.

A The differentially methylated imprint control region (KvDMR1) is located within the promoter region of *Kcnq1ot1/Lit1)* and is methylated only on the maternal allele. *Kcnq1ot1/Lit1* is a non-coding RNA transcript expressed only from the paternal allele. Dashed line represents the location of *Kcnq1* intron 10 antisense to *Kcnq1ot1/Lit1*. HS = hypersensitivity sites. Orange box represents the *Kcnq1ot1/Lit1* transcript. Red bar represents the KvDMR1 CpG island region. Blue lines represent the cut sites of restriction enzymes. PubMed accession number NT_039453 from 527668-532398 bp
B The *Cdkn1c* promoter DMR is methylated only on the paternal allele in somatic cells. Differential methylation extends into intron II of *Cdkn1c*. The methylation status of the upstream CpG island is unknown.
Pink boxes show the locations of the CpG islands. Orange boxes are gene exons. The green box indicates the location of the predicted *Cdkn1c* promoter. Regions examined for methylation by bisulphite sequencing or MSRE-PCR are indicated by the red bars. Blue lines represent the cut sites of restriction enzymes. PubMed accession number NT_039437.

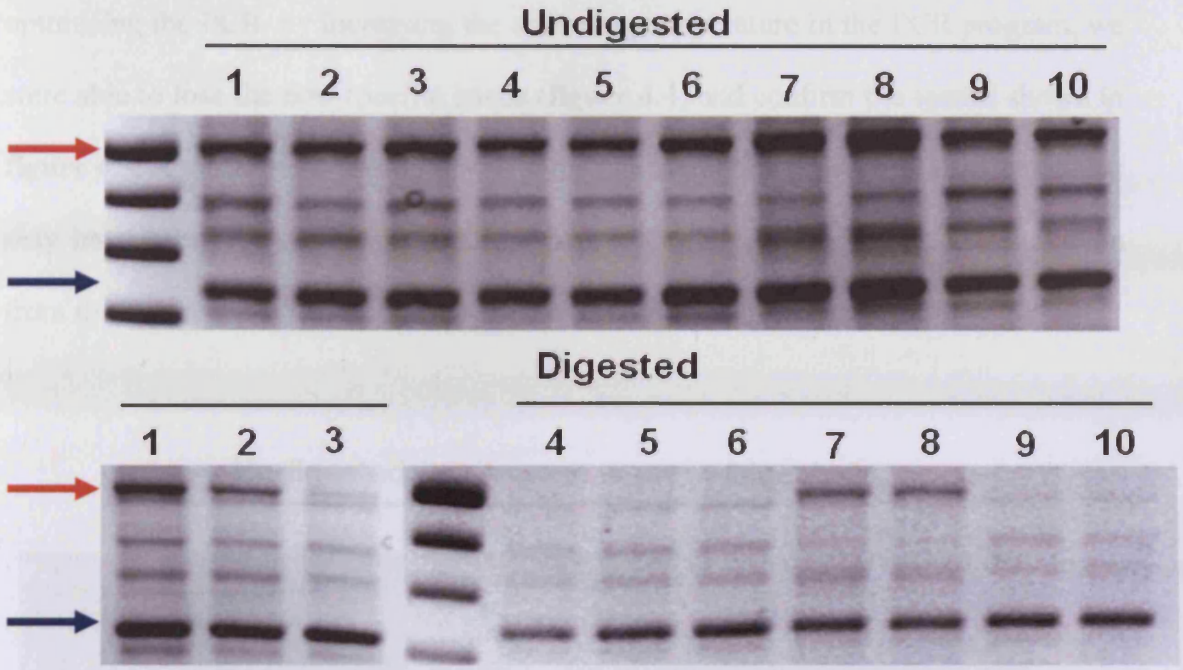


Figure 4.3: Methylation status of the KvDMR1 in a variety of ES and EG cell lines.

In the absence of DNA methylation the methylated band is lost (red arrow). The PCR control is indicated by the blue arrow. Undigested samples were used as a control for methylation. Genomic DNA samples from ES cell lines CES3 and SF1-1, EG cell lines Sv6.1 and TMAS21G and androgenetic cell lines AKR1 were digested with the methylation sensitive restriction enzyme *EagI* before PCR amplification. (1) CES3 d0, (2) CES3 day 21, (3) Sv6.1 day 0, (4) Sv6.1 day 7, (5) Sv6.1 day 14, (6) Sv6.1 day 21, (7) SF1-1 day 0 (8) SF1-1 day 7, (9) TMAS21G day 0, (10) AKR1 day 0.

Methylation of the CpG dinucleotide looked at in this experiment has also been examined using Southern blotting. Previous studies have shown that this CpG dinucleotide is differentially methylated in the germ line and is stably maintained throughout development (Tada *et al.* 1998). The fact that we did not see any methylation of this CpG dinucleotide in the differentiated EG cells confirmed the absence of methylation at this site in undifferentiated cells and demonstrated that DNA methylation was not acquired at this site during *in vitro* differentiation.

A number of non-specific bands are seen in the KvDMR1 methylation analysis, which was probably due to non-specific binding of the primers. After

optimising the PCR, by increasing the annealing temperature in the PCR program, we were able to lose the non-specific bands (**figure 4.4**) and confirm the results shown in **figure 4.3**. A weak band was observed in the undifferentiated Sv6.1 EG sample which may have been due to incomplete digestion of the DNA or possibly contamination from the feeder layer but essentially the result confirmed previous reports.

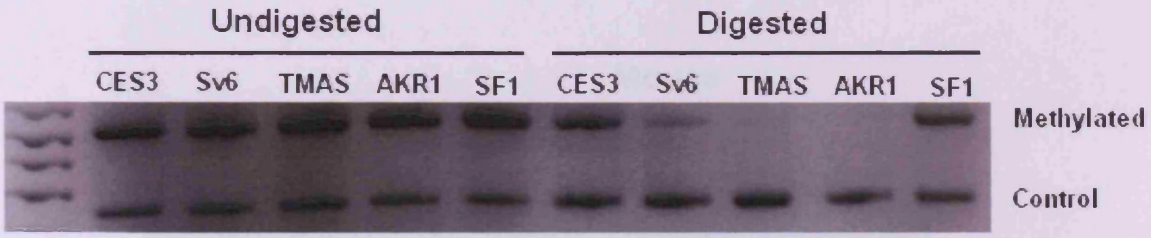


Figure 4.4: Methylation status of the KvDMR1 in undifferentiated ES and EG cell lines.

In the absence of DNA methylation the methylated band is lost. Undigested samples were used as a control. Genomic DNA samples from ES cell lines CES3 and SF1-1, EG cell lines Sv6.1 and TMAS21G and androgenetic cell line AKR1 were digested with the methylation sensitive restriction enzyme *EagI* before PCR amplification.

A set of primers were designed to examine *Cdkn1c* DMR by this method. Initially the primers failed to amplify products of the correct size and repeatedly amplified a number of non-specific bands. This may have been due to the GC rich nature of this particular gene making amplification by PCR more difficult. A number of methods were attempted in order to amplify products of the correct size from around the methylation sensitive restriction site including adding DMSO and Betaine to the PCR reaction and re-designing primers. Using the GC rich kit from Roche we were eventually able to amplify products from the promoter region and intron II of *Cdkn1c* from a bacterial artificial chromosome (BAC) containing the *Cdkn1c* gene

and from genomic DNA isolated from adult mouse brain (**figure 4.5**). An internal primer control was also designed as described earlier. In **figure 4.5 A,** amplification of the undigested genomic sample was inefficient, however a product was amplified in a second analysis using these primers. **Figure 4.5 B** showed that the internal control primers were able to amplify a product in both BAC and genomic DNA samples.

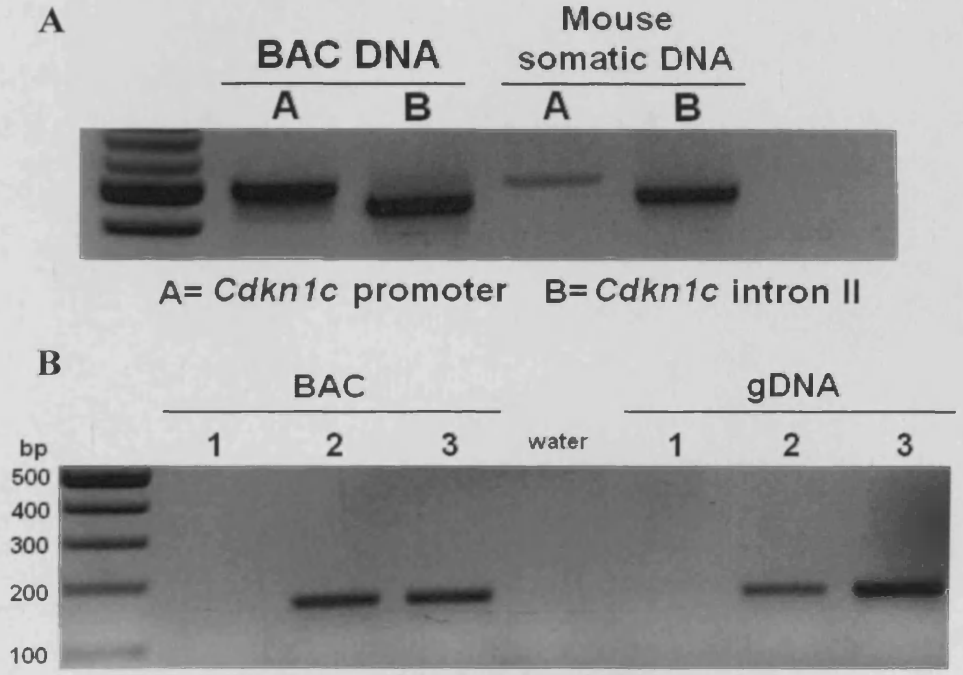


Figure 4.5: Amplification of the mouse *Cdkn1c* promoter and intron II CpG islands from undigested BAC and mouse brain somatic DNA.

A Amplification of the methylated band of *Cdkn1c* promoter (A) and intron II CpG (B) in undigested bacterial artificial chromosome (BAC) containing the *Cdkn1c* gene and undigested mouse brain genomic DNA. **B** A combined PCR for *Cdkn1c* promoter using both the methylated and the control primers in both the undigested BAC and genomic DNA. (1= methylation primers only, 2= internal primers only, 3= methylation primers + internal primers).

The primers were used to look at the methylation status of two potentially methylated CpG dinucleotide located within the promoter of *Cdkn1c* that are located in an *Eag*I site. Studies have shown that these CpGs are differentially methylated in

somatic cells (John *et al.* 1999; Tada *et al.* 1998). We looked at the methylation status of a number of ES and EG cell lines as well as the mouse brain somatic genomic DNA (**figure 4.6**).

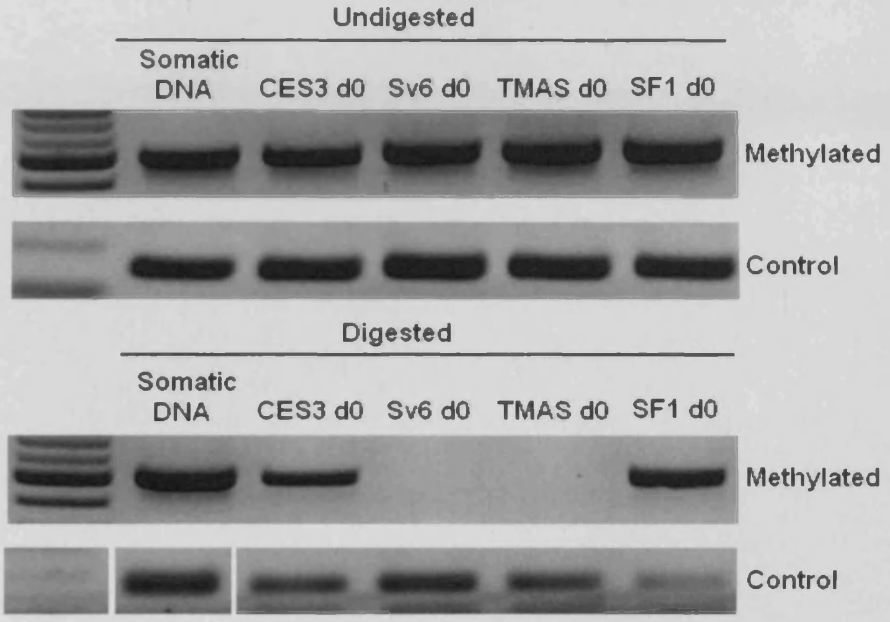


Figure 4.6: Methylation of the *Cdkn1c* promoter DMR in the ES and EG cell lines.

MSRE-PCR analysis of the *Cdkn1c* promoter in ES and EG cells. In the absence of DNA methylation the methylated band is lost. A control PCR was also performed on the ES cells that amplifies a product, irrespective of DNA methylation status, located within the methylated primer product. Undigested samples were used as a control for methylation. Genomic DNA samples from ES cell lines CES3 and SF1-1, EG cell lines Sv6.1 and TMAS21G and mouse brain somatic DNA were digested with *Eco*R1 or the methylation sensitive restriction enzyme *Eag*I before PCR amplification. d0 = undifferentiated.

In the undigested (digested with *Eco*R1 only) controls the methylated band and the control band was amplified confirming that there was DNA present in the digested samples and that both primer sets amplified a product. In the samples digested with *Eag*I, there are methylated bands present for the mouse somatic genomic DNA and the two ES cell lines (CES3 and SF1-1) but the two EG cell lines

(Sv6.1 and TMAS21G) appear unmethylated as suggested by the absence of a methylated band on the gel. The control primers showed that the absence of the methylated band was not due to the lack of DNA in the digests. These results confirm what was observed with the KvDMR1 and in the Tada experiments, that methylation is erased in the EG cells but some is present in the ES cell lines. We then repeated the results but looking at the methylation status of the intron II region of the *Cdkn1c* CpG island (**figure 4.7**).

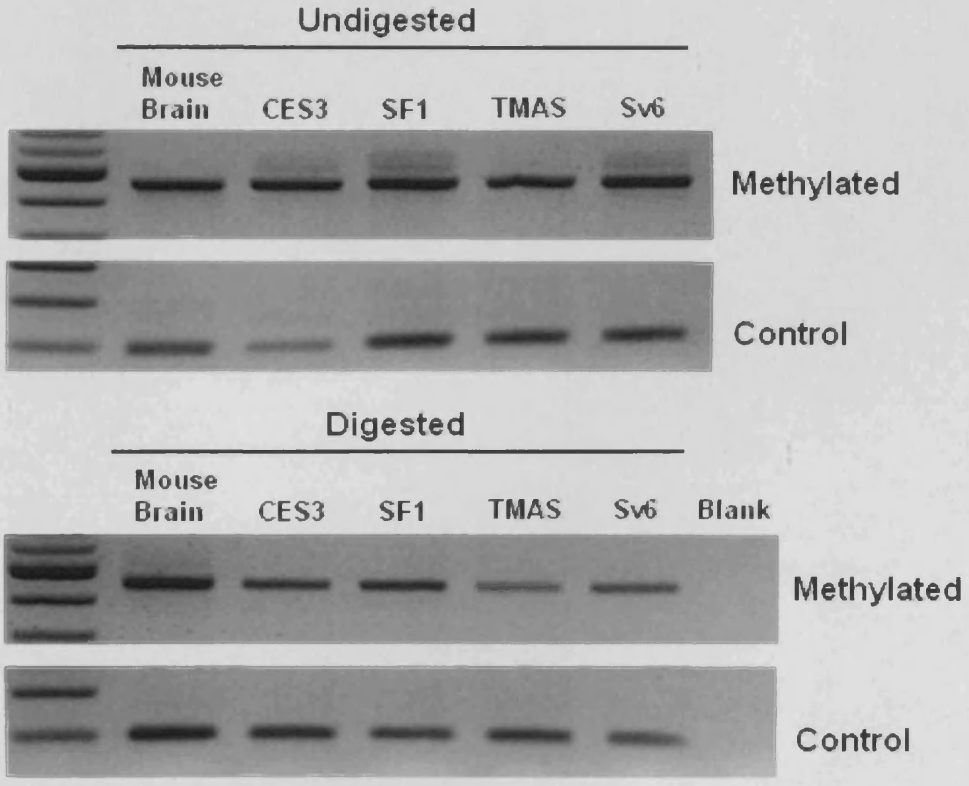


Figure 4.7: Methylation of the *Cdkn1c* intron II DMR in the ES and EG cell lines.

MSRE-PCR analysis of the *Cdkn1c* intron II in undifferentiated ES and EG cells. In the absence of DNA methylation the methylated band is lost. A control PCR was also performed on the ES cells that amplifies a product, irrespective of DNA methylation status, located within the methylated primer product. Undigested samples were used as a control for methylation. Genomic DNA samples from ES cell lines CES3 and SF1-1, EG cell lines Sv6.1 and TMAS21G and mouse brain somatic DNA were digested with *Eco*R1 or the methylation sensitive restriction enzyme *Eag*I.

The results suggested that the intron II CpG island of *Cdkn1c* was weakly methylated in both the ES and EG cell lines at at least one CpG dinucleotide. The extent of the DNA methylation can not be accurately ascertained with this method. It was possible that the methylation observed was residual from incomplete de-methylation of this region in the PGCs. There was also a possibility that re-methylation of the *Cdkn1c* region was initially initiated at a few CpG dinucleotides within intron II of the gene and this then spreads to the promoter as the cells differentiate.

The MSRE-PCR method only tells us the methylation status of a few CpG dinucleotides within the CpG island. It does not indicate the extent of the methylation. Importantly, PCR is a very sensitive method and the smallest amount of undigested DNA would appear as a methylated band. In order to examine additional sites and to test the possibility that partial digestion had occurred, we used a second method that did not rely on methylation sensitive restriction digestion.

4.2.2 Design and optimisation of the bisulphite sequencing protocol

Methylation sensitive restriction enzyme coupled PCR only looks at the methylation of one CpG site within the sequences of interest. It does not show the extent of methylation within these sequences. Methylation at the *Cdkn1c* promoter is a somatic methylation mark that is only present on the paternal allele. Differential DNA methylation at this site is only acquired after the genome wide methylation event that occurs in the early embryo (Bhogal *et al.* 2004; Monk *et al.* 1987). How this methylation mark is established is still unclear. It may be that methylation is initially initiated at a few CpG sites and then spreads across the entire promoter. Data published during the course of this project showed that promoter methylation was

extensive across the paternal allele in somatic cells (Bhogal *et al.* 2004; Yatsuki *et al.* 2002) and that the CpG site we showed to be methylated by MSRE-PCR was methylated in the embryo. Bisulphite sequencing is a more exhaustive method for looking at DNA methylation as it looks at all the CpG sites within a particular region [reviewed in (Hayatsu 2006)]. It would allow us to determine the precise location of 5 methylcytosine (5meC) within CpG islands.

A saturated bisulphite solution is used to convert cytosine to cytosine sulphonate. Cytosine sulphonate is then irreversibly deaminated to uracil sulphonate and finally desulphonated to uracil by the addition of sodium hydroxide. The converted DNA is subsequently amplified by PCR and can then be either directly sequenced or cloned into vectors for colony screening. Methylated cytosine is much more resistant to bisulphite treatment and remains unchanged. The location of the CpG islands within *Cdkn1c* was identified by the Methprimer CpG prediction software (Li and Dahiya 2002), that examines the DNA sequence to identify regions of more than 100 bp with a GC content greater than 50 % (software found at http://www.urogene.org/methprimer/index1.html). This software, which is based on the primer3 program, also allows the design of primers specific for bisulphite converted DNA. Unlike normal primer selection, bisulphite primers have extra constraints, for example they can not contain a CpG site and are selected to eliminate differences in Tm and to remove mononucleotide repeats of more than 5 bases (8 bases for T repeat to account for the bisulphite converted DNA) (Li and Dahiya 2002). Complete conversion of cytosine to uracil is essential for subsequent amplification of the DNA. The bisulphite treatment only works on single stranded DNA and, under the right conditions, 5meC remains unreactive. Amplified products

can then be sequenced to identify the CpG dinucleotides that were methylated in the original sequence by looking for unconverted cytosines.

Cdkn1c contains four exons that are highly conserved between mice and humans (John *et al.* 1999). A CpG island spanning the entire gene has been shown previously to be fully methylated in somatic tissues on the paternal allele (John *et al.* 1999). Using the Methprimer program, it was apparent that this CpG rich region was actually two CpG islands one over the promoter and one in intron II. There was also a small island more than 3 Kb upstream of *Cdkn1c* which was also reported by Yatsuki *et al.* (Yatsuki *et al.* 2002) (**figure 4.8**).

Figure 4.8: Schematic and location of CpG islands within *Cdkn1c*.

Mouse *Cdkn1c* consists of four exons (shown in orange). The location of the predicted promoter sequence is shown in green. A CpG island spans then length of the gene from the predicted promoter region to intron 3. A second CpG island is located approximately 3 Kb upstream of the promoter. Blue diamonds show the location of individual CpG dinucleotides. Pink boxes show the regions of the CpG islands examined. The two CpG islands spanning the promoter and intron II are shown as one in this figure.

Primers were designed against the promoter CpG island, the intron CpG island and the uncharacterised upstream CpG island. Initially the bisulphite sequencing protocol used was based on the method described by Warnecke and colleagues (Warnecke *et al.* 2002). Problems were encountered with PCR amplification of the

bisulphite converted DNA. It was thought to be caused by re-annealing of the DNA prior to the bisulphite conversion. This problem was resolved by embedding denatured genomic DNA in low melting point agarose (LMP) as described in several publications (Engemann *et al.* 2002; Gratchev 2004; Olek *et al.* 1996).

The early bisulphite sequencing results suggested extensive methylation throughout the DMRs in *Cdkn1c*. However, closer analysis revealed inefficient conversion of the DNA. Only around 40-60 % of the non-CpG cytosines within the amplified sequences had converted to uracil. This was thought to be due to the highly GC-rich CpG island spanning the gene which is resistant to heat denaturation. Several approaches were used to rectify this problem including using a variety of restriction enzymes to cut around the sequence of interest to produce shorter fragments, adding betaine during the denaturation step to prevent reannealing of the DNA, heat treatments of various lengths and incubation with sodium hydroxide for up to 30 min. A urea based method for bisulphite conversion was also used in an attempt to get full conversion. To test all of these methods, plasmid DNA containing the cloned *Cdkn1c* promoter was used as a control for conversion. Plasmid DNA normally does not acquire CpG methylation so 5meC would not be present in these sequences. All the cytosines within the sequence should completely convert. The urea based protocol was used along side the previously used agarose bead protocol on the control DNA. In this protocol the denaturing of the DNA was done as with the agarose bead protocol, however, 6.24 M urea was also added to the bisulphite solution (Paulin *et al.* 1998). This method failed to completely convert the DNA however, by varying the denaturing conditions, we were ultimately successful in obtaining complete bisulphite conversion of the plasmid DNA using the bisulphite agarose protocol. The following modifications to the protocol were used. The DNA was digested with an enzyme that

cut around the region of interest then was denatured by heating at 95 °C for 6 min. The DNA solution was incubated on dry ice for three minutes. A final concentration of 0.3 M sodium hydroxide was added and incubated at 37 °C for 30 min to ensure that the DNA was completely denatured. The control plasmid DNA showed complete bisulphite conversion of nearly 100 %.

The same protocol was then applied to genomic DNA of the two undifferentiated EG cell lines (TMAS21G and Sv6.1). Complete bisulphite conversion was observed. Bisulphite conversion was analysed using the BiQanalyzer program (Bock *et al.* 2005) which removes sequences that are less than 90 % converted, show less than 80 % homology to the parent sequence and controls for PCR bias.

4.2.3 Methylation at the *Cdkn1c* promoter

Figure 4.9 shows the results for the bisulphite converted *Cdkn1c* promoter in the control plasmid DNA and in the two undifferentiated EG cell lines. Hypo-methylation of the CpG dinucleotides was observed indicating that *Cdkn1c* promoter methylation was completely erased in the undifferentiated EG cell lines as suggested by southern data (Tada *et al.* 1998) and our MSRE method. The *Cdkn1c* promoter was then analysed in differentiated EG cells to determine if methylation was acquired during differentiation which was reported in the EG cell derived fibroblasts (Tada *et al.* 1998). Methylation at the promoter region of *Cdkn1c* could account for the silencing of this gene observed in our differentiated EG cells and reported in the Tada paper.

Figure 4.10 shows the methylation status of the *Cdkn1c* promoter in differentiated EG cells (Sv6.1). No increase in DNA methylation was apparent.

Therefore, unlike the EG-derived PEFs, which were derived from TMAS cells, no DNA methylation was acquired. Nonetheless, *Cdkn1c* was silenced suggesting that other mechanisms were at work. It is possible that DNA methylation was present at other regions associated with *Cdkn1c*. Intron II and the upstream CpG island were also analysed to determine if DNA methylation was present.

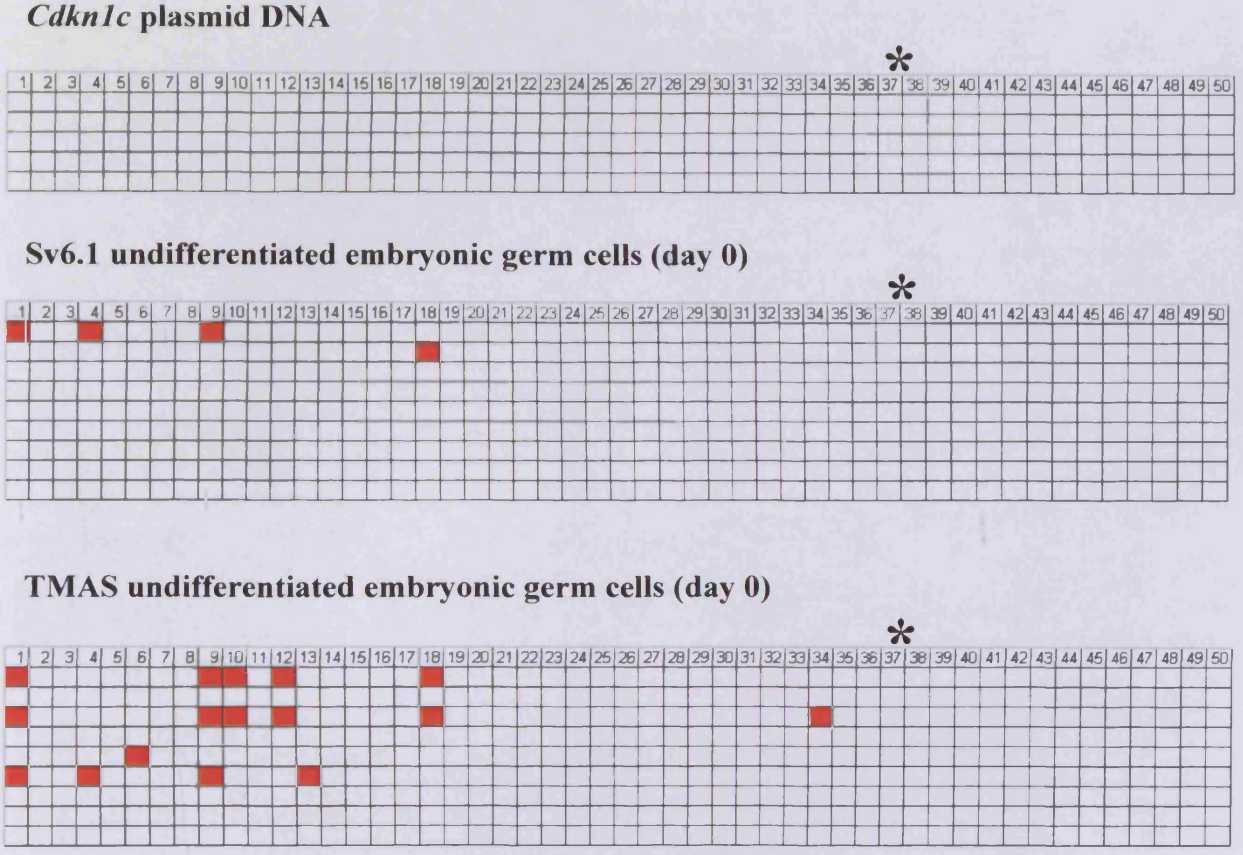


Figure 4.9: Bisulphite sequencing analysis of the *Cdkn1c* promoter region in plasmid DNA and genomic DNA from undifferentiated Sv6.1 and TMAS21G EG cells.

Genomic DNA was extracted from undifferentiated Sv6.1 and TMAS21G cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines. Asterisks (*) represents the location of the *Eag*I site in genomic DNA analysed by MSRE-PCR and southern blot.

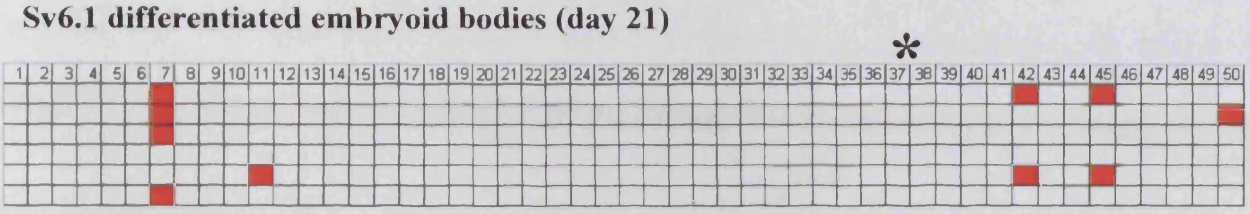


Figure 4.10: Bisulphite sequencing analysis of the *Cdkn1c* promoter region in genomic DNA from differentiated Sv6.1 EG cells.

Sv6.1 cells were differentiated for 21 days by culturing as embryoid bodies. Genomic DNA was extracted differentiated Sv6.1 cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines. Asterisks (*) represents the location of the *Eag*I site in genomic DNA analysed by MSRE-PCR and southern blot.

4.2.4 DNA methylation within other CpG islands in *Cdkn1c*

Differential DNA methylation has now been shown to extend throughout the *Cdkn1c* CpG region as well as located to a region upstream of the transcriptional start site (Bhogal *et al.* 2004). The intron II and the upstream CpG islands were analysed by bisulphite sequencing in order to determine if these regions become methylated as the EG cells differentiate. **Figure 4.11** shows the result for the intron II region. The undifferentiated cells were hypo-methylated as would be expected for imprint erased cells. After differentiation this region was also hypo-methylated indicating that methylation of *Cdkn1c* was not responsible for suppressing *Cdkn1c* expression.

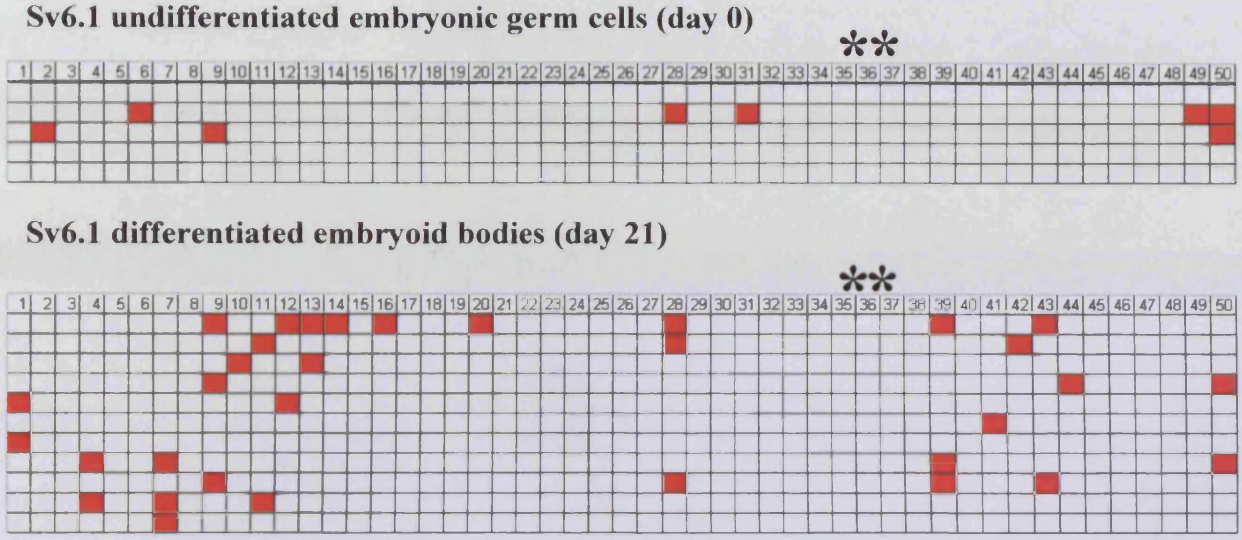


Figure 4.11: Bisulphite sequencing analysis of the *Cdkn1c* intron II region in undifferentiated and differentiated Sv6.1 cells.

Sv6.1 cells were differentiated for 21 days by culturing as embryoid bodies. Genomic DNA was extracted from undifferentiated and differentiated Sv6.1 cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines. Asterisks (*) represents the location of the *Eag*I site in genomic DNA analysed by MSRE-PCR and southern blot.

The upstream CpG island was also examined and was found to be hyper-methylated in undifferentiated EG cells (**figure 4.12**). The function of this region is unknown, however it was possible that erasure of methylation at this site had not occurred. After differentiation this region became hypo-methylated indicating that methylation of this region was lost during differentiation. These data were for line Sv6.1. This line was not the same as the line tested in the Tada paper and it was possible that subtle differences between the Sv6.1 line, which is a pure 129/Sv derived stem cell, and the TMAS21G line, which was a mixed genetic background, accounted

for the absence of DNA methylation after differentiation. We therefore examined TMAS21G, the line in the Tada paper (Tada *et al.* 1998).

Sv6.1 undifferentiated embryonic germ cells (day 0)

Sv6.1 differentiated embryoid bodies (day 21)

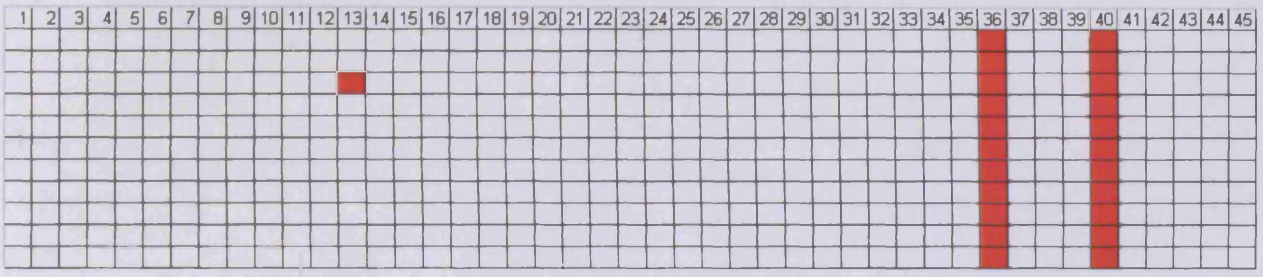


Figure 4.12: Bisulphite sequencing analysis of the *Cdkn1c* upstream CpG region in undifferentiated and differentiated Sv6.1 cells.

Sv6.1 cells were differentiated for 21 days by culturing as embryoid bodies. Genomic DNA was extracted from undifferentiated and differentiated Sv6.1 cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines.

Hypo-methylation of the *Cdkn1c* intron II region was confirmed in the TMAS21G cell line (**figure 4.13**). The upstream CpG island region was also analysed in these cells (**figure 4.14**). Unlike the Sv6.1 EG cells, this cell line did show complete erasure of DNA methylation of this region. After differentiation, some DNA methylation was seen at this region in some of the clones sequenced. The data was not consistant with the Sv6.1 data for this island indicating that there are subtle differences between the two EG cell lines. Recently is has been shown that this island

is methylated on both alleles in somatic cells making it an unlikely candidate for an imprint control element (Bhogal *et al.* 2004).

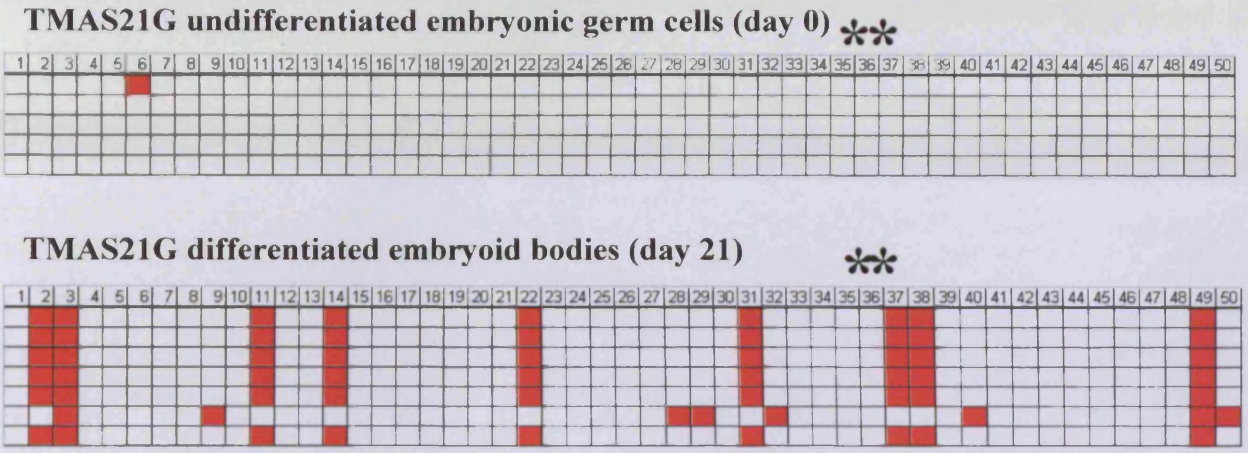
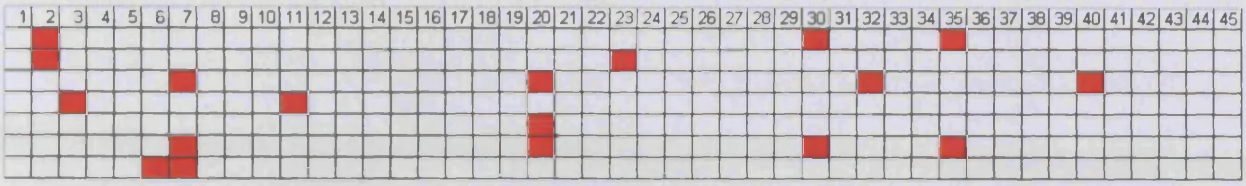


Figure 4.13: Bisulphite sequencing analysis of the *Cdkn1c* intron II region in undifferentiated TMAS21G cells.

TMAS21G cells were differentiated for 21 days by culturing as embryoid bodies. Genomic DNA was extracted from undifferentiated and differentiated TMAS21G and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines. Asterisks (*) represents the location of the *Eag*I site in genomic DNA analysed by MSRE-PCR and southern blot.

TMAS21G undifferentiated embryonic germ cells (day 0)



TMAS21G differentiated embryoid bodies (day 21)

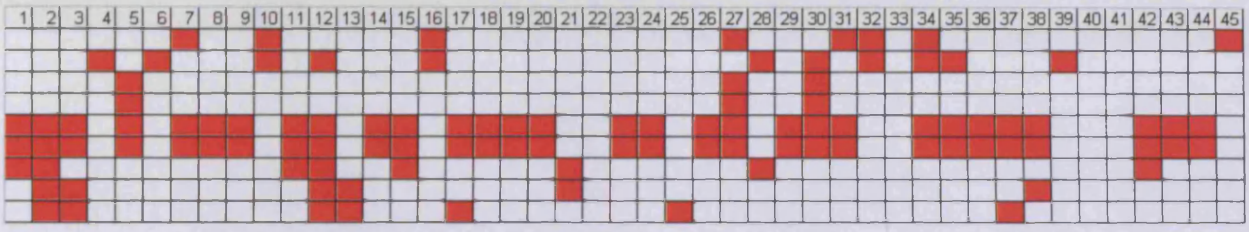


Figure 4.14: Bisulphite sequencing analysis of the *Cdkn1c* upstream CpG region in undifferentiated and differentiated TMAS21G cells.

TMAS21G cells were differentiated for 21 days by culturing as embryoid bodies. Genomic DNA was extracted from undifferentiated and differentiated TMAS21G cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines.

4.2.5 Methylation analysis of *Cdkn1c* CpG islands in ES cells

The *Cdkn1c* promoter CpG island was hypo-methylated in the EG cells. We also examined the methylation pattern in the undifferentiated and differentiated ES cells. These cells were expected to show differential DNA methylation of the *Cdkn1c* promoter as they contain the normal genomic imprints. The cell line we used was from a pure 129sv strain of mouse and so did not contain any polymorphisms that we could use to analyse parental origin of methylation. We would, however, expect a 50:50 ratio of non-methylated clones to methylated clones which would reflect the hypo- and hyper-methylated alleles that should be present in this cell line. DNA methylation was analysed as described previously using bisulphite sequencing. In undifferentiated ES cells there was suprisingly very little methylation at the *Cdkn1c* promoter (**figure 4.15**). There was a lot less methylation than expected in these cells, especially when we compare the levels to that observed in somatic cells in the Bhogal study (Bhogal *et al.* 2004). This could suggest that methylation at this DMR was not stable in culture.

We looked at the methylation status of other regions within the *Cdkn1c* DMR. **Figure 4.16** shows the analysis of the intron II region of *Cdkn1c* which showed the same hypo-methylation in undifferentiated and differentiated ES cells. The same result was seen for the upstream CpG island (**figure 4.17**). Even in the differentiated cells this region appeared hypo-methylated. All of the regions analysed in the CES3 cells were hypo-methylated although there were more methylated CpGs observed in the ES cells than seen in the EG cells.

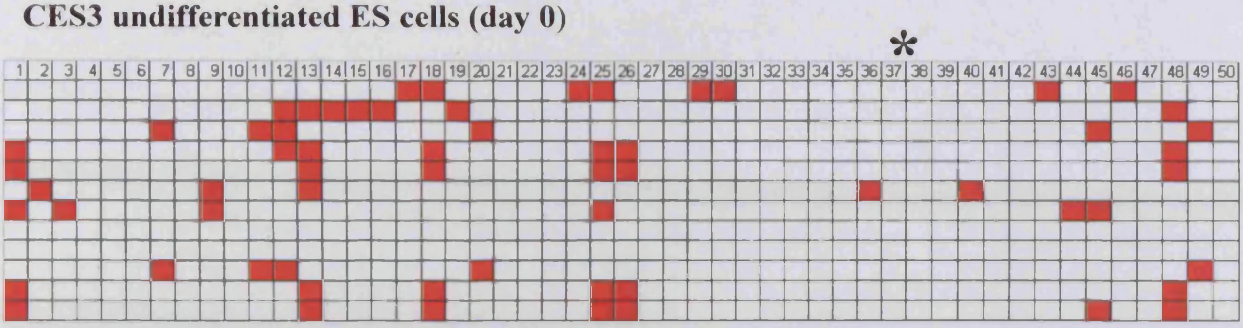


Figure 4.15: Bisulphite sequencing analysis of the *Cdkn1c* promoter domain in undifferentiated CES3 cells.

Genomic DNA was extracted undifferentiated CES3 cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines. Asterisks (*) represents the location of the *Eag*I site in genomic DNA analysed by MSRE-PCR and southern blot.

We also looked at the methylation pattern in an androgenetic cell line. Androgenetic cells contain two paternal genomes. The *Cdkn1c* promoter DMR was first analysed by methylation sensitive restriction enzyme-coupled PCR which showed that the DMR was apparently methylated in undifferentiated androgenetic stem cells (**figure 4.18**). Two different cell lines were tested, those used in these studies and a separate cell line obtained from Dr Nick Allen. As the DMR appears to be methylated we went on to analyse the extent of the methylation by bisulphite sequencing. **Figure 4.19** shows the results from the bisulphite analysis at the *Cdkn1c* promoter and upstream CpG island regions. The androgenetic cell line showed the same result as seen with the ES cells. Both the *Cdkn1c* promoter and upstream CpG island were hypo-methylated. Again this could suggest loss of methylation at the *Cdkn1c* promoter in culture. We therefore designed assays to examine other DMRs that were germ line DMRs and potentially more stable in culture.

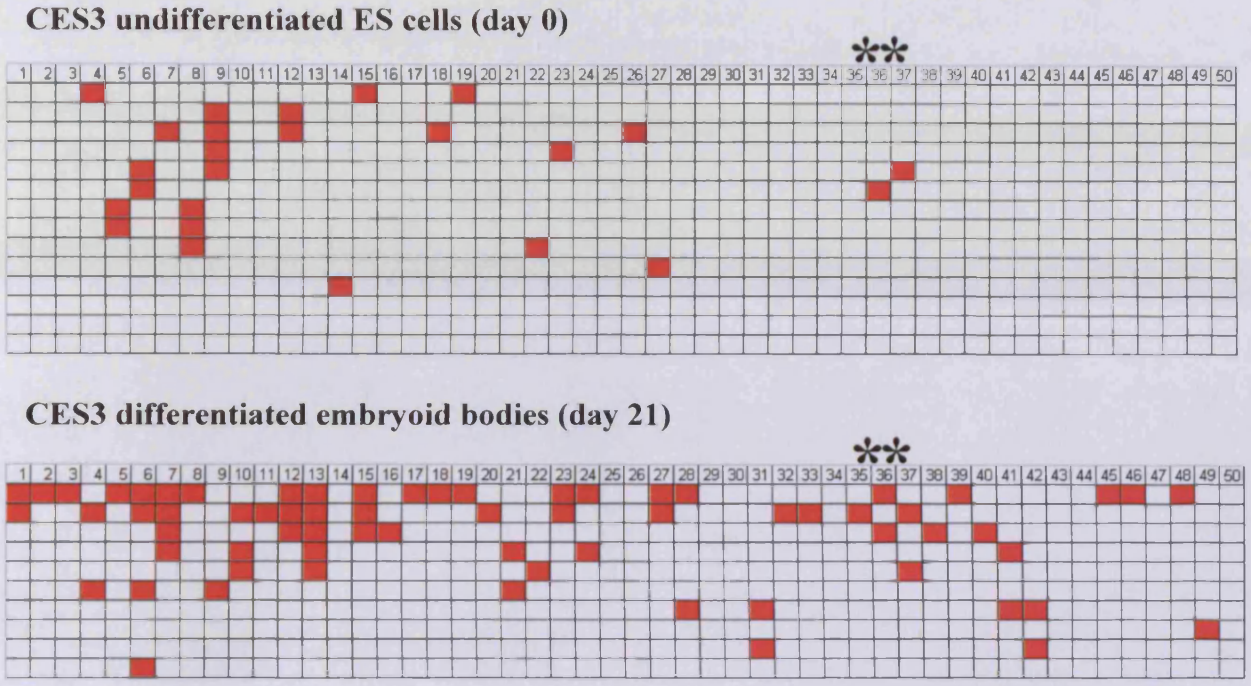


Figure 4.16: Bisulphite sequencing analysis of the *Cdkn1c* intron II domain in undifferentiated and differentiated CES3 cells.

CES3 cells were differentiated for 21 days by culturing as embryoid bodies. Genomic DNA was extracted from undifferentiated and differentiated CES3 cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines. Asterisks (*) represents the location of the *Eag*I site in genomic DNA analysed by MSRE-PCR and southern blot.

The results from the bisulphite and MSRE-PCR data appeared to contradict each other. The MSRE-PCR assay appeared to indicate that there was methylation of the promoter in ES and AG cells. However, the MSRE-PCR is highly sensitive to the presence of residual methylation and, as it is not quantitative, can result in false positive results or an over estimate of the amount of methylation present. It could also be picking up somatic methylation from feeder cells. In contrast bisulphite sequencing can result in false positives if the sequence fails to fully convert. We, however, see a

negative result for all our bisulphite methylation data and as 5meC is unlikely to be converted it is highly likely that the *Cdkn1c* promoter is truly unmethylated.

CES3 undifferentiated embryonic germ cells (day 0)

CES3 differentiated embryoid bodies (day 21)

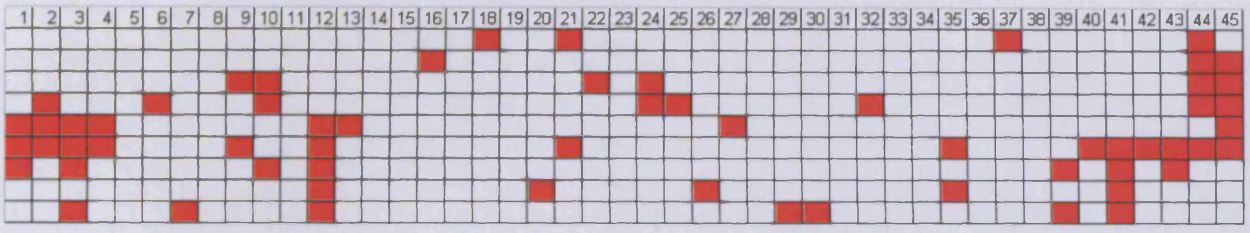


Figure 4.17: Bisulphite sequencing analysis of the *Cdkn1c* upstream CpG region in undifferentiated and differentiated CES3 cells.

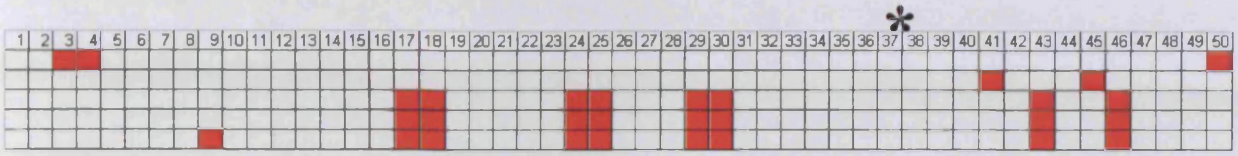
Genomic DNA was extracted undifferentiated CES3 cells and modified by the addition of sodium bisulphite. DNA was amplified by primers specific to bisulphite converted DNA. PCR products were cloned into a vector and sequenced. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines.

Figure 4.18: MSRE analysis of the *Cdkn1c* promoter CpG region in undifferentiated AKR1 cells.

The methylation sensitive restriction enzyme coupled PCR was applied to the androgenetic ES cells to determine if there is methylation at the *Cdkn1c* promoter. Two cell lines are analysed the androgenetic ES cell line used in this study AKR1 and a replacement AKR1 cell line (AKR1 New).

AKR1 undifferentiated embryonic stem cells (day 0)

A



B

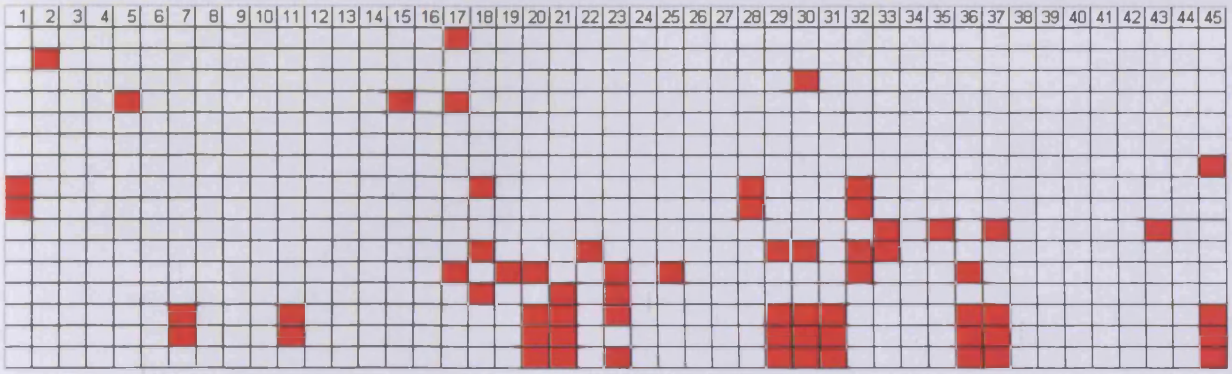


Figure 4.19: Bisulphite sequencing analysis of the *Cdkn1c* promoter and upstream CpG differentially methylated regions in undifferentiated AKR1 stem cells.

Genomic DNA was extracted from undifferentiated AKR1 cells and modified by the addition of sodium bisulphite. **A** *Cdkn1c* promoter, **B** *Cdkn1c* upstream CpG island. Each row represents an individually sequenced clone; each column represents a potentially methylated CpG dinucleotide. Red squares are methylated cytosines. Asterisks (*) represent the location of the *Eag*I site in genomic DNA analysed by MSRE-PCR and southern blot.

4.2.6 Methylation analysis of other germ line DMRs

We designed primers to paternal germ line DMRs by MSRE-PCR. Although this only tests two methylated CpGs it is a convenient tool in order to ascertain the methylation status of these cells. **Figure 4.20** shows that the AKR1 cell line did have methylation at both the *H19* and *Gtl2* germ line DMRs. Methylation was also present in the wild type ES cell line. Conversely the EG cell line showed no methylation at the *Gtl2* DMR, however methylation remained at the *H19* DMR indicating that this

DMR is still methylated in these cells. The methylation status of the *Gtl2*-DMR corresponded to that expected for these cell lines and indicated that the paternal methylation marks were intact for the androgenetic and ES cell lines and with the expected pattern. This then suggested loss of DNA methylation was specific to the *Cdkn1c* locus in the stem cells.

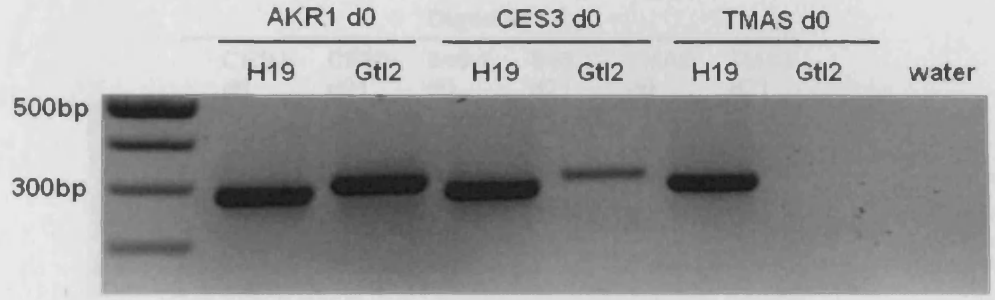


Figure 4.20: Methylation sensitive restriction enzyme coupled-PCR on androgenetic (AKR1), wild type (CES3) and germ line (TMAS21G) embryonic stem cells.

Methylation of the paternal DMRs *H19* and *Gtl2* was confirmed in the androgenetic and wild type ES cell lines. PubMed accession number for *H19* DMR is AF_049091. *H19* primers were designed around a *Hha*I site within the region 3308-3611 bp (Ishizaki *et al.* 2003). PubMed accession number for *Gtl2/Dlk1* IG-DMR is AJ320506. *Gtl2* primers were designed around a *Hha*I site within the region 79751-82320 bp (Takada *et al.* 2002).

The *H19* DMR was further analysed in the different cell lines to confirm the result seen above and to determine if methylation was lost in differentiated cells at this region. **Figure 4.21** shows the methylation status of the *H19* DMR in both undifferentiated and differentiated ES and EG cells. The data appeared to indicate that the *H19* DMR was methylated in all cell lines so may not be completely erased in the EG cells. As we were only looking at a single potentially methylated CpG dinucleotide in this region it could be that this is the only CpG that remains

methylated in these cells. Other studies have shown similar results at this locus (Tada *et al.* 1998).

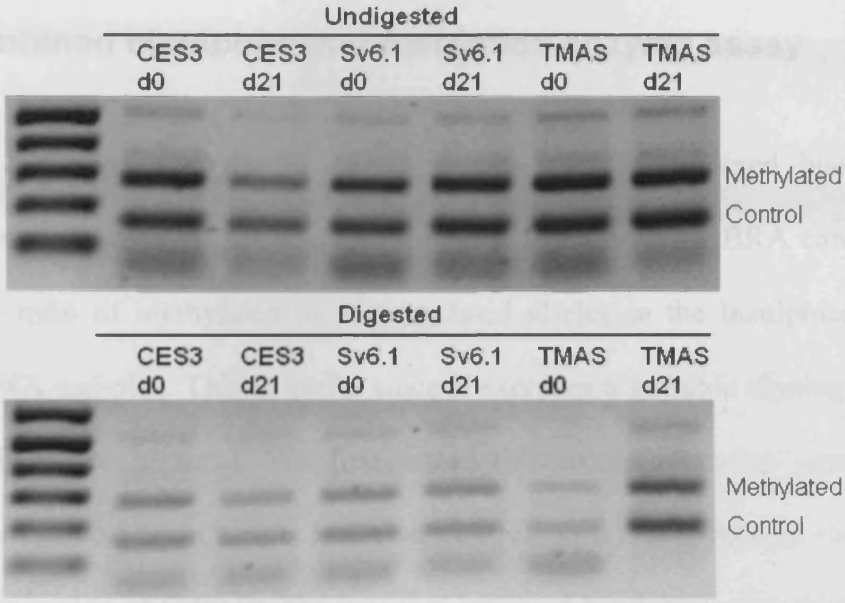


Figure 4.21: Methylation analysis of the H19 DMR.

Methylation at the *H19* DMR was analysed by MSRE-PCR in both ES (CES3) and EG (Sv6.1, TMAS21G) cell lines in undifferentiated and differentiated cells. Undigested samples are shown as a control.

Although the MSRE-PCR assay is a useful assay for quickly determining the methylation status of a CpG dinucleotide, it has several limitations. Firstly, the result depends on the efficient digestion of the DNA by the methylation sensitive restriction enzyme. Secondly, PCR is very sensitive and can amplify low copy number sequences with great efficiency, so if any residual uncut strands are present or contaminating DNA exists in the samples then this may result in a false positive result. Only samples that show a complete absence of the methylated band (but still amplify the internal control) can be confirmed. In **figure 4.20**, the same TMAS21G digested DNA was used to amplify both the *H19* and the *Gtl2* DMRs, one of which appeared positive for DNA methylation and one that appeared negative. In this case because the *Gtl2* DMR did not show a methylated band it suggests that this DNA

sample was fully digested and free from contamination so the *H19* DMR result is likely to be true methylation.

4.2.7 Combined bisulphite and restriction enzyme assay

As a further methylation assay, we designed a combined bisulphite and restriction enzyme assay (COBRA) (Xiong and Laird 1997). COBRA can be used to look at the ratio of methylated to unmethylated alleles in the bisulphite converted genomic DNA samples. This is useful since it excludes a possible cloning bias in the standard bisulphite protocol. We first tested this technique using genomic DNA isolated from mouse brain as this DNA was known to be differentially methylated at the promoter region of *Cdkn1c*. The protocol involved bisulphite converting the DNA, amplifying the *Cdkn1c* intron II DMR, as before, then digesting the DNA with a restriction enzyme that only cuts at an unconverted CpG dinucleotide. Like the MSRE-PCR method, this only shows us if particular CpG dinucleotides are methylated or not but it does have the advantage that we can determine how much methylation there is as a ratio. The results for the mouse brain somatic DNA digest are shown in **figure 4.22**. Within the PCR fragment there are three *Hha*I cut sites which, when the DNA is fully methylated, would produce fragments of various sizes the largest being 248 bp. However, partial methylation of the domain would produce several other larger fragments. The amount of DNA we amplified was low by this protocol. We were unable to quantify the ratio but it was clear that some of the DNA had been digested indicating that there was at least some methylation of the *Cdkn1c* DMR in the somatic DNA.

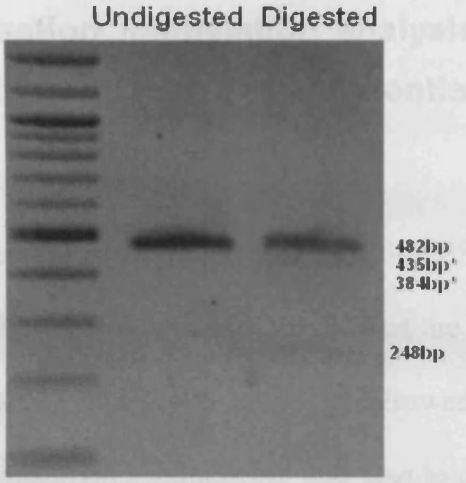


Figure 4.22: Combined bisulphite and restriction enzyme assay (COBRA) of mouse brain somatic genomic DNA.

Somatic genomic DNA was converted using bisulphite solution as described previously. *Cdkn1c* intron II was amplified using the nested bisulphite primers. The DNA was then digested with *Hha*I confirming that the DNA was indeed differentially methylated in somatic cells. The size of the faint bands are indicated which correspond to partially digested and fully digested alleles.

The absence of the DNA methylation at the *Cdkn1c* gene in ES cells, shown by several methods was unexpected. However, during the project a paper was published which suggested that *Cdkn1c* is unmethylated in the early embryo (Bhogal *et al.* 2004) and does not acquire differential methylation until after 8.5 dpc. One possible explanation for our results was that the somatic methylation of the silent paternal allele occurs after the time point at which ES cells are derived. If this were the case, we might expect other ES cell lines to show a similar pattern. In collaboration with Dr Tatyana Nesterova, we used southern blotting to examine this further.

4.2.8 Southern hybridisation methylation analysis of the *Cdkn1c* promoter in undifferentiated and differentiated ES, AG and EG cells

We decided to confirm the methylation status of our ES and EG cells using Southern blotting. Although this technique could only tell us the methylation status of a few CpG dinucleotides within the CpG island, it allowed us to confirm the methylation analysis done by bisulphite sequencing and also to determine the ratio of methylated to unmethylated DNA. A probe was designed to hybridise with a 2.9 kb *Bam*H1 digested genomic DNA fragment within the *Cdkn1c* DMR. The *Bam*H1-digested genomic DNA fragment contains several *Eag*I site (**figure 4.23**). This methylation sensitive enzyme will not cut the DNA if the binding site overlaps a methylated CpG dinucleotide. Ten micrograms of genomic DNA from undifferentiated and differentiated ES and EG cells was digested with *Bam*H1 and *Eag*I, run out on a TBE gel, transferred to a nitrocellulose membrane then probed with the radioactive DNA probe (southern blotting was performed by Tatyana Nesterova). If the *Eag*I site was methylated then a 2.9 kb band would be visible on the blot. An unmethylated product would be visible as a 332 bp fragment. The ratio between the two bands would allow us to determine the extent, if any, of the differential DNA methylation.

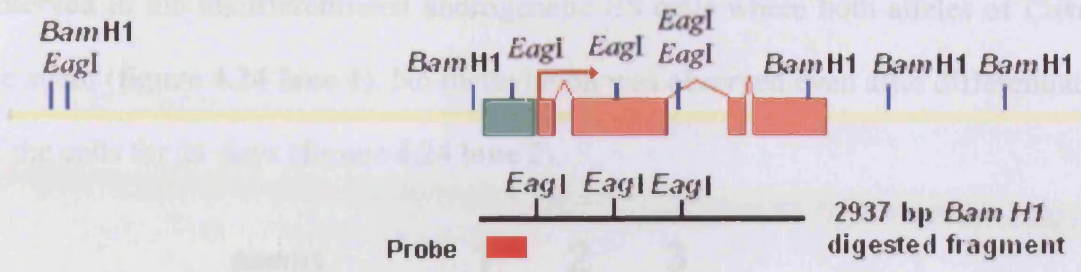


Figure 4.23: Location of the Southern blot probe.

Schematic of the *Cdkn1c* gene showing locations of the *Bam*H1 and *Eag*I cut sites. A 2.9 Kb fragment is released by digesting with *Bam*H1 which contains the site the probe hybridises with. Within the fragment there are several *Eag*1 cut sites. The methylation status of this site can be determined from the size of the hybridised fragment. A methylated site will appear as a 2.9 Kb hybridised band on the gel whereas unmethylated sites appear as a 332 bp hybridised fragment.

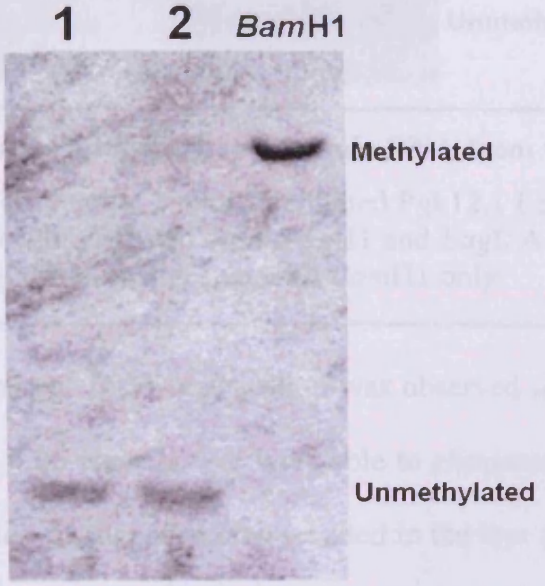


Figure 4.24: Southern blotting on genomic DNA from androgenetic ES cells.

1 undifferentiated AKR1 ES cells and **2** differentiated (21 day) AKR1 ES cells digested with *Bam*H1 and *Eag*I. A control digest is included where genomic DNA is digested with *Bam*H1 only.

The Southern blotting was performed on our samples by Tatyana Nesterova. The data supported our findings by bisulphite sequencing that the *Cdkn1c* locus was unmethylated in ES cells. Only the fully digested (unmethylated) fragment was

observed in the undifferentiated androgenetic ES cells where both alleles of *Cdkn1c* are silent (**figure 4.24 lane 1**). No methylation was observed even after differentiation of the cells for 21 days (**figure 4.24 lane 2**).

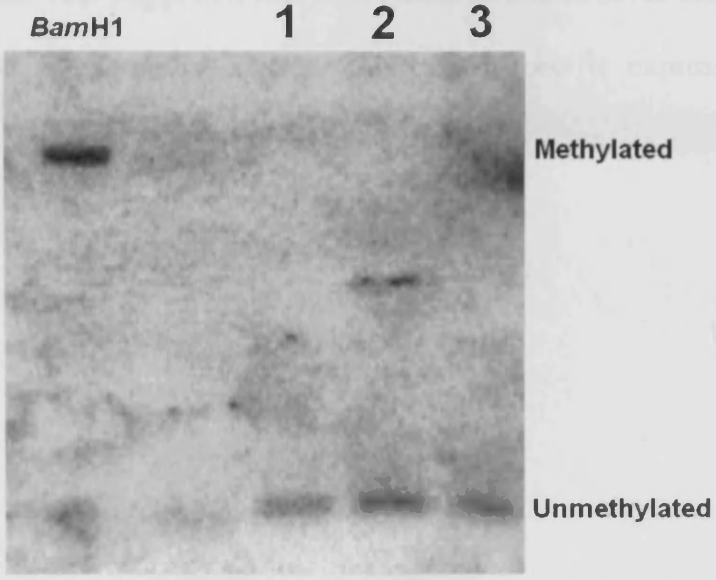


Figure 4.25: Southern blotting on genomic DNA from wild type ES cells.

1 undifferentiated 129/1 ES cells; **2** undifferentiated Pgk12.1 ES cells and **3** differentiated CES3 ES cells digested with *Bam*H1 and *Eag*I. A control digest is included where genomic DNA is digested with *Bam*H1 only.

Similarly, no apparent DNA methylation was observed in the differentiated ES cell line CES3 (**figure 4.25 lane 3**). We were able to eliminate the possibility that it was our culturing conditions that somehow resulted in the loss of *Cdkn1c* methylation as the two independent cell lines cultured by Dr Nesterova also showed a lack of DNA methylation at *Cdkn1c* (**figure 4.25 lanes 1 and 2**). Unfortunately, we were unable to determine the methylation status of the EG cells lines in this experiment due to high background staining overlapping the location of the 2.9 kb fragment. However, the lack of DNA methylation in even the differentiated androgenetic ES cells strongly supports our conclusion that *Cdkn1c* is unmethylated in stem cells and, unlike the *in vivo* situation, DNA methylation is not acquired as the cells differentiate.

This could suggest that stem cells lack a critical trans-factor that is required for the further epigenetic modification of this locus after silencing by histone modifications.

This experiment convincingly confirmed the absence of DNA methylation in all three ES cell lines. This suggested that DNA methylation is never fully established at the *Cdkn1c* locus in ES cells. Despite this, allele specific expression of *Cdkn1c* occurs.

4.3 Discussion

The association of differentially methylated regions and ICRs makes DNA methylation a good candidate for controlling differential gene expression of imprinted genes. DNA methylation at germ line DMRs is established in the germ line during gametogenesis and is thought to be the primary imprinting mark. During post-implantation development, DNA methylation is also established at some somatic DMRs (Santos *et al.* 2002), such as that found at the promoter of *Cdkn1c* which is believed to completely repress expression of this gene. In EG cells, derived from 12.5 dpc PGCs, the methylation marks at both the KvDMR1 and the *Cdkn1c* DMRs have been shown to be completely erased (Tada *et al.* 1998). In the Tada *et al.* study, EG cells were differentiated *in vivo* as PEFs in chimeras and the methylation status of DMRs at several imprinted loci was examined (Tada *et al.* 1998). Most of the germ line DMRs remained unmethylated after differentiation; however they observed that the somatic DMR at the *Cdkn1c* promoter became *de novo* methylated in differentiated PEFs. The aim of this study was to determine if *de novo* methylation of the *Cdkn1c* DMR would occur after differentiating EG cells *in vitro* and to determine the extent and timing of this methylation mark in more detail than that reported in the Tada paper.

4.3.1 Methylation analysis of the KvDMR1 in EG cells

The results for the methylation analysis of the KvDMR1 by MSRE-PCR indicated that methylation was absent in undifferentiated EG cells reflecting the methylation status of the imprint-erased PGCs from which they were derived. The KvDMR1 was also un-methylated even after differentiating the EG cells as embryoid

bodies. This was expected for the KvDMR1 as methylation of this region only occurs after passage through the female germ line.

Methylation of the KvDMR1 was observed in both undifferentiated and differentiated ES cells. We have to assume that methylation in the cells is maternal in origin as we have no polymorphisms in which we could study allele specific methylation patterns. However, evidence from other studies has shown that germ line DNA methylation marks, like KvDMR1, are epigenetically stable in culture (Humphreys *et al.* 2001; Szabo and Mann 1994) and it is unlikely that the paternal KvDMR1 would become *de novo* methylated in culture. Methylation at the KvDMR1 was maintained even after differentiating the cells for 21 days confirming that this germ line DNA methylation mark was stable. The specificity of the methylation assay was confirmed in androgenetic ES cells, which contain two paternal genomes and showed no methylation of the KvDMR1.

4.3.2 Methylation analysis of *Cdkn1c* promoter and intron II DMRs in EG cells

Methylation of the somatic DMR at the *Cdkn1c* promoter was shown by Tada *et al.* to occur in differentiated PEFs isolated from chimeras created from imprint erased EG cells (Tada *et al.* 1998). In the Tada study, methylation of the *Cdkn1c* promoter was examined by southern blotting which only examined the methylation status of a few CpG dinucleotides within the DMR. They showed that in undifferentiated EG cells, derived from at 11.5 and 12.5 dpc PGCs, the *Cdkn1c* promoter was unmethylated, however the cells showed full methylation of both *Cdkn1c* alleles after differentiation *in vivo* as PEFs (Tada *et al.* 1998). In our initial analysis we looked at the methylation status of the same CpG dinucleotide in undifferentiated EG cells and confirmed what was seen in the Tada study, that *Cdkn1c*

was unmethylated in imprint erased EG cells. Some weak methylation was observed in the ES cell line samples and was absent from both EG cell lines looked in this study. It was possible that the single CpG dinucleotide examined in this study and the Tada study is the only CpG dinucleotide within the whole promoter CpG island that was not methylated in the undifferentiated EG cells. For this reason we examined the methylation status of approximately 50 potentially methylated CpG dinucleotides within the promoter of *Cdkn1c*, one of which corresponds to the differentially methylated CpG examined in the MSRE-PCR experiment. Our analysis showed that *Cdkn1c* promoter was hypomethylated in the undifferentiated and in the differentiated EG cells, indicating that the suppression of *Cdkn1c* expression observed in chapter 3 was not due to *de novo* DNA methylation of the *Cdkn1c* promoter. There was always a possibility that DNA methylation at other regions within the *Cdkn1c* CpG island was causing the *Cdkn1c* suppression. In fact, when we used MSRE-PCR to look at the methylation of a CpG dinucleotide within the intron II region of the DMR, this region did appear to be methylated, albeit to a lesser extent in the EG cell lines, possibly indicating that DNA methylation of the DMR is initially initiated at a few discrete dinucleotides which then spreads to the entire CpG island later in development. However, when we looked at this region by bisulphite sequencing we did not see methylation in either of the EG cell lines. This region also appeared to be relatively hypomethylated in the ES cells and in the androgenetic cell line which had been shown by the MSRE-PCR method to be methylated.

We now have an explanation for these findings. Essentially, in collaboration with Dr Tatyana Nesterova, we have shown that three ES cells lines are unmethylated at *Cdkn1c* despite maintaining differential methylation at other DMRs. A paper by Bhogal *et al.* determined that the DNA methylation at *Cdkn1c* is not acquired until 8.5

dpc, several days after ES cells are derived (Bhogal *et al.* 2004). It seems likely that the ES cells have not lost DNA methylation in culture – they have not yet acquired it. Nonetheless, in the case of the Bhogal study and as suggested by our qRT-PCR analysis, the *Cdkn1c* locus is still silenced i.e. DNA methylation is not required to keep *Cdkn1c* off, at least at these time points. The methylation we detected by the MSRE assay may reflect a very low level of DNA methylation in some cells but it seems more likely that it was due to partial digestion of the sample or, potentially, contamination from the feeder layer, highlighting the importance of using several approaches to study DNA methylation.

No other studies have reported the methylation status of the *Cdkn1c* DMR in ES cells so we have no idea of what the extent of the methylation is supposed to be. Bhogal *et al.* (Bhogal *et al.* 2004) have analysed part of this region and did show extensive methylation on the paternal allele in neonatal liver and adult tail somatic tissue derived from a cross between C57BL/6J (B6) and either *Mus musculus castaneus* (CAST) or B6(CAST7) mice.

ES and AG cells are both derived from the ICM of blastocyst stage embryos so should be epigenetically similar to the cells from which they were derived i.e. relatively de-methylated except at germ line DMRs. As the androgenetic ES cells were derived at the same time point as ES cells it is therefore likely that, like ES cells, the two paternal *Cdkn1c* alleles had not yet acquired DNA methylation. As with the ES cells, the two paternal alleles would be predicted to acquire *de novo* methylation at *Cdkn1c* during post-implantation embryonic development.

Our ES and AG cells are derived from the inner cell mass of blastocyst stage embryos so have yet to undergo the post-implantation re-methylation event. Several studies have shown that CpG islands are generally unmethylated in the early embryo.

Prissette *et al.* looked at the methylation of the *DxPas34* locus, a controlling element that plays a role in X inactivation. They showed that in the early embryo up to the blastocyst stage this gene was unmethylated and did not acquire *de novo* methylation until around 7.5 dpc (Prissette *et al.* 2001). In ES cells, derived from 129sv male mice, there were only low levels of methylation at the locus which increased during differentiation of the cells as embryoid bodies but the methylation remained limited and heterogeneous (Prissette *et al.* 2001). This is similar to what we observed at the *Cdkn1c* DMR.

The data indicated that methylation of the *Cdkn1c* DMR was not important for suppressing *Cdkn1c* expression in ES and AG cells. *Cdkn1c* was not expressed in undifferentiated ES cells, so methylation of the paternal allele of *Cdkn1c* was probably not essential at this stage. Tucker *et al.* demonstrated that ES cell lines derived from embryos with a targeted deletion in Dnmt1 (Dnmt1-/-) could be maintained in culture and proliferate normally (Tucker *et al.* 1996). The authors indicated that methylation of the genome was only an issue when the cells were differentiated in culture as they showed that ES cells lacking Dnmt1 die upon differentiation. Dnmt1-/- embryos are embryonic lethal and death is likely due to mis-regulation of gene expression as expression of Dnmt1 cDNA in mutant ES cells was able to restore normal global methylation patterns and these cells were able to differentiate normally in culture. However, DNA methylation at imprinted loci was not re-established and imprinting was lost at several loci (Tucker *et al.* 1996).

Differentiating the ES cells as embryoid bodies mimics post-implantation events. After only 3-8 days of culture, the embryoid bodies differentiate into tissues that represent all three germ layers and are morphologically similar to the 6-8 dpc egg cylinder stage (Doetschman *et al.* 1985). By 8-10 days of differentiation as embryoid

bodies, approximately half of the embryoid bodies have been shown to develop into large cystic structures that show similarities to the visceral yolk sac of post-implantation embryos (Doetschman *et al.* 1985). The promoter region of the *Cdkn1c* DMR was shown to be hypomethylated until the 7.5 dpc stage of embryonic development then slowly becomes fully methylated until at least 9.5 dpc (Bhogal *et al.* 2004). We differentiated Sv6.1 EG cells for 21 days in culture but *Cdkn1c* did not become methylated. It may be that the embryoid bodies do not reach the same stage in development at which *de novo* methylation of the *Cdkn1c* DMR would occur *in vivo*. In our study we did observe a slight increase in methylation in differentiated ES cells compared to undifferentiated ES cells which may reflect that there is some methylation activity occurring. Our differentiation system produced a heterogeneous population of embryoid bodies as they were not seeded out as hanging drops. This resulted in production of embryoid bodies of various sizes which may represent different differentiation stages and so may have differing methylation potentials.

In the Tada study (Tada *et al.* 1998) *Cdkn1c* was shown to become methylated on both parental alleles indicating that in the imprint-erased EG cells there is an epigenetic mark present on both chromosomes that is recognised in somatic cells as a region in to which *de novo* DNA methylation is attracted. The difference between our study and the Tada study is that our cells do not undergo full embryonic development whereas the Tada EG cells do, as part of the chimeric mouse. The interesting result of this study is that DNA methylation does not appear to be the reason that *Cdkn1c* is repressed in the EG cells. Therefore perhaps repressive histone marks were responsible for this silencing. These repressive marks might also be responsible in marking out the silent allele so that in the post-implantation embryo, *de novo* DNA methylation is recruited to the *Cdkn1c* promoter as a late event.

4.3.3 Methylation analysis of the *Cdkn1c* upstream DMR in ES and EG cells

We observed hypomethylation of both the promoter and the intron II DMR of the *Cdkn1c* gene in both the Sv6.1 and TMAS21G cell lines indicating that methylation was not acquired during *in vitro* differentiation of these cells. We did, however, observe differences in the methylation status of the upstream CpG island which was hypermethylated in the undifferentiated Sv6.1 EG cells and hypomethylated after differentiation. The opposite was observed for the TMAS21G EG cells. Bhogal *et al.* also looked at a region upstream of the *Cdkn1c* promoter (Bhogal *et al.* 2004). The Bhogal study showed that outside of the *Cdkn1c* gene CpG island, both parental alleles were equally hyper-methylated in somatic tissue, this analysis extended up to 2 Kb upstream of the transcriptional start site. We analysed the methylation status of a CpG island that was more than 3 Kb upstream of the *Cdkn1c* start of transcription so was out of the range examined in the Bhogal *et al.* study. Even in ES and AG cells, this region was only partially methylated with no clear pattern in which CpG dinucleotides were methylated. The methylation of this region appeared to be sporadic in both cell lines. The differences between the two EG cell lines could be due to differences in genotype as the TMAS21G cells are genetically modified, however we do not know why we would see de-methylation of this region in one cell line and *de novo* methylation in another except perhaps if they were at a different developmental stage. For example, migrating PGCs do not arrive at the developing gonadal ridge at exactly the same time point so consist of a heterogeneous population of cells possibly resulting in epigenetic differences between cells. As methylation of the upstream DMR appears so inconsistent it is unlikely to play a role in silencing of *Cdkn1c* expression in the EG cell lines.

4.3.4 Methylation at other germ line DMRs

We examined other DMRs in our stem cells to see if they were stable. We had already analysed the KvDMR1 maternal methylation mark in these cells. Very few germ line DMRs are known to be methylated in the male germ line, however those established at the *Igf2-H19,* and *Dlk1-Gtl2* ICRs have been well characterised. Primers were designed for MSRE-PCR to amplify the *H19* and *Gtl2* DMRs in ES cells. These regions were expected to be differentially methylated in ES cells but methylated on both alleles in AG cells and absent from both alleles in EG cells. The lack of polymorphisms in our cell lines meant that we could not determine allele-specific methylation, we could only detect the presence or absence of methylation at the DMR. We confirmed that the *Gtl2* DMR was indeed methylated in both the AG and ES cells but was absent in the EG cells indicating that the cells had maintained the methylation profile we expected. We also expected to see the same pattern for the *H19* DMR however in all cell lines examined, methylation appeared to be retained. These results were confirmed in both undifferentiated and differentiated ES and EG cells. The data indicated that the *H19* DMR is not erased in EG cells confirming what is seen in the Tada study. Tada *et al.* showed that the *H19* DMR is partially methylated in the 12.5 dpc male EG cells possibly indicating that this germ line DMR is not completely erased in these cells or that they have already begun to acquire parental specific methylation at this DMR (Tada *et al.* 1998).

In the previous chapter we observed that expression of *Igf2* was higher than expected. Several studies have shown that *Igf2* is suppressed in the absence of an imprint (Kato *et al.* 1999; Lee *et al.* 2002). As the EG cells appeared to be at least partially methylated at the *H19* DMR then this could account for the high levels of *Igf2* expression observed in the differentiated EG cells. Methylation of the *H19* DMR

blocks binding of CTCF allowing *Igf2* to access its downstream enhancers whereas, in the absence of methylation, CTCF can bind the DMR and *Igf2* is repressed. As we observe levels of *Igf2* expression that are around 1.4 fold of that seen in ES cells it is likely that the methylation of the *H19* DMR that we observe is only partial in EG cells so may not be able to completely block CTCF binding. The *H19* and *Gtl2* DMRs were analysed using the MSRE-PCR assay which, as has been previously discussed, is at risk from displaying false positives. This assay could, therefore, be picking up a subset of methylated strands within the gDNA that represent a minority so the *H19* DMR could be generally unmethylated. To confirm this data we would need to perform another assay, such as bisulphite sequencing or southern blotting, that could accurately quantify the methylation status of this DMR.

The timing of imprint erasure and establishment has been studied extensively in mouse PGCs. The exact timing of erasure appears to vary between imprinted gene clusters and within PGC populations. Genomic imprints are erased in migrating PGCs and is usually completed by the time the cells colonise the genital ridge at around 11.5-12.5 dpc. Several studies have examined the timing of imprint erasure in germ cells and striking differences in the timing of the erasure process between genes and between PGCs have been observed. In 9.5 dpc PGCs the *H19* DMR is predominantly methylated on the paternal allele whereas the maternal allele is still relatively unmethylated in both male and female germ cells indicating that the imprints for this imprinting cluster have yet to be erased (Yamazaki *et al.* 2005). The *Snrpn* DMR also appears to be fully methylated in these cells, however the DMRs of *Igf2r* and *Cdkn1c* were shown to be completely unmethylated indicating that these DMRs are probably de-methylated earlier in PGCs (Durcova-Hills *et al.* 2001; Yamazaki *et al.* 2005). There appears to be some discrepancy in the timing of erasure of the KvDMR1

methylation mark which, in one study, was suggested to occur before 9.5 dpc (Durcova-Hills *et al.* 2001) yet another study suggested that this DMR was still methylated by 10.5 dpc (Hajkova *et al.* 2002). The *H19* and *Snrpn* DMRs were shown to be methylated in 10.5 dpc PGCs albeit to a lesser extent compared to that observed in 9.5 dpc PGCs (Li *et al.* 2004a; Yamazaki *et al.* 2005). The *Gtl2*-DMR also appears to have retained its imprints at this stage (Li *et al.* 2004a). By 11.5 dpc, the methylation at the *H19* DMR appears to vary between studies. Yamazaki *et al.* showed that the *H19* DMR was more than 50 % methylated on only around 25 % of the clones analysed compared with 88 % observed in the 9.5 dpc PGCs (Yamazaki *et al.* 2005). This data appeared to indicate that the *H19* DMR is progressively de-methylated overtime in PGCs. However, this data contrasts with that observed in other studies which showed that *H19* is in fact more methylated at this stage than in 9.5 dpc PGCs (Durcova-Hills *et al.* 2001). Lee *et al.* has also shown that there is significant variation in the methylation status of the *H19* gene in clones derived from 11.5 dpc PGCs (Lee *et al.* 2002). This is probably due to the fact that some PGCs initiate de-methylation earlier during migration and not all PGCs arrive at the genital ridge at the same time, possibly resulting in epigenetic differences between PGC clones. At other DMRs within 11.5 dpc PGCs, methylation appears to be completely erased at least from *Igf2r, Cdkn1c, Kcnq1ot1/Lit1, Peg1, Nnat* and *Peg3* (Durcova-Hills *et al.* 2001; Tada *et al.* 1998) indicating that there are differences in timing of erasure between imprinted genes as well as between PGCs. Partial methylation of the *H19* and *Snrpn* DMRs was also observed by Li *et al.* and Ueda *et al.* in 12.5 dpc PGCs (Li *et al.* 2004a; Ueda *et al.* 2000). This partial methylation appears to persist as it is also found at the *H19* DMR in 13.5-14.5 dpc PGCs (Davis *et al.* 2000; Ueda *et al.* 2000). This partial methylation is mainly restricted to the paternal allele in male PGCs

(some methylation is observed on the paternal allele in female PGCs and this is subsequently lost before entry into meiosis) and this methylation may still be present in male germ cells when re-methylation of the male genome begins. In contrast, Iwahashi *et al.* showed that both the *Snrpn* and *H19* DMRs were completely erased in 13.5 dpc PGCs (Iwahashi *et al.* 2007).

The *H19* DMR is a very well studied ICR and its timing of erasure during germ line reprogramming appears to indicate that de-methylation of this region is progressive and may not be completely erased in 12.5 dpc PGCs. These 'imprint-erased' PGCs may in fact possess residual DNA methylation at the *H19* DMR either due to incomplete erasure of the imprint or as a result of early re-methylation of the DMR in male PGCs. This could account for the methylation of the *H19* DMR that we observed in our EG cells however, without using bisulphite sequencing, we can not determine the extent of the methylation at this DMR.

4.3.5. Conclusion

In this study we wanted to determine if *de novo* methylation of the *Cdkn1c* promoter occurred during differentiation of imprint erased EG cells, similar to that observed when EG cells are differentiated *in vivo* (Tada *et al.* 1998). Our studies indicated that the *Cdkn1c* promoter remains unmethylated in the EG cells even after prolonged differentiation in culture. This suggested that DNA methylation was not responsible for the repression of *Cdkn1c* expression observed in chapter 3. Our data also indicated that the *Cdkn1c* DMR was relatively hypomethylated even in ES and AG cells, an unexpected finding. Recent data suggests this may be because methylation of this DMR occurs during post-implantation development in the embryo.

As *Cdkn1c* expression is imprinted in ES cells, which has been confirmed in many studies, then monoallelic gene expression must be maintained through other factors. Tada *et al.* observed *de novo* methylation of the *Cdkn1c* DMR after *in vivo* differentiation, indicating that there is a default sequence of epigenetic events that is either already pre-set or become established in the imprint erased PGCs during differentiation and is required to direct *de novo* methylation to the silent *Cdkn1c* allele in the developing embryo (Tada *et al.* 1998). As DNA methylation does not appear to be acquired then histone modifications and other transfactors must be responsible for silencing.

4.3.6 Summary of findings

- KvDMR1 showed the correct methylation status in all of the cell lines used in this study indicating that this germ line methylation mark was stable in culture and that the EG cells were 'imprint erased'.
- *Cdkn1c* DMR was unmethylated in imprint erased EG cells and did not become methylated even after differentiation *in vitro,* contrary to that observed in EG cells differentiated *in vivo.* This implied that chromatin modifications and other DNA binding proteins are probably responsible for repressing *Cdkn1c* expression in differentiated EG cells.
- *Cdkn1c* DMRs were relatively unmethylated in differentiated ES and AG cells supporting the idea that methylation of this DMR was a late event during embryogenesis.
- The CpG island located upstream of the *Cdkn1c* DMR showed variable methylation patterns in EG, ES and AG cells so was probably not involved in establishing silencing of the *Cdkn1c* gene.

- The *Gtl2* DMR was unmethylated in the EG cells but methylated in the AG and ES cells confirming that the EG cells are 'imprint erased' and that some germline DMRs were stable in culture.
- Methylation of the *H19* DMR was observed in all cell lines indicating that this DMR was either not completely erased in EG cells or was being re-established implying that, at at least one germ line DMR, 12.5 dpc PGCs can not be considered to be completely 'imprint erased'.

CHAPTER 5:

Analysis of the histone modifications at IC2 domain promoters

5.1 Introduction

Many imprinted genes are found within large imprinting domains that can span several kilobases. These domains contain reciprocally imprinted genes as well as non-imprinted genes implying that there is a more complex regulation system that can not be explained from discrete DNA methylation at an ICR alone. The spreading of silencing or activation could be perpetuated by a spreading of open or closed chromatin from the ICR to the promoters and enhancer elements of the surrounding imprinted genes. Allele specific histone modifications have been shown to regulate imprinting at several domains (Carr *et al.* 2007; Li *et al.* 2004b; Umlauf *et al.* 2004; Yamasaki-Ishizaki *et al.* 2007). Within the IC2 domain of mouse distal chromosome 7, DNA methylation at KvDMR1 correlates with silencing of the non-coding transcript *Kcnq1ot1/Lit1*, and expression of the surrounding genes. It is possible that methylation of KvDMR1 recruits histone modifying enzymes to alter the chromatin structure in female PGCs, and are passed on through the maternal germ line to the somatic cell. Or it could be that KvDMR1 methylation is only required to silence *Kcnq1ot1/Lit1* and that it is the non-coding RNA that recruits histone modifying enzymes to silence the domain in the male germ line. Parallels have been drawn with dosage compensation that exists in females and results in imprinted X chromosome inactivation, particularly as a number of features on the silent X chromosome, such as DNA methylation, repressive histone modifications, replication asynchrony and the presence of non-coding RNA are all found within imprinted domains (Boggs and Chinault 1994; Jeppesen and Turner 1993; McDonald *et al.* 1998; Ohhata *et al.* 2008).

The final chromatin structure for the IC2 domain of mouse distal chromosome 7 is well documented and has been shown to be similar in both mice and humans

(Umlauf *et al.* 2004). A variety of histone modifications have been shown to be associated with the promoter regions of the imprinted genes to give either an open or closed chromatin conformation. A number of histone modifying enzyme complexes have also been shown to bind to imprinted genes within this domain and to the KvDMR1 (Fitzpatrick *et al.* 2007; Umlauf *et al.* 2004; Wagschal *et al.* 2008). Having shown that silencing of the maternally expressed genes was maintained in *in vitro* differentiated EG cells, we next sought to characterise the histone modifications at the KvDMR1 locus and at several key promoters; *Cdkn1c, Phlda2* and *Kcnq1*. There were two possibilities. The silencing histone marks would be present in both the undifferentiated and differentiated cells, in which case the domain could be regarded as pre-silenced in EG cells. Alternatively, histone marks could be recruited to the silent genes during differentiation, which would suggest that silencing of the genes was an active, dynamic process that takes place during differentiation of the EG cells (**figure 5.1**).

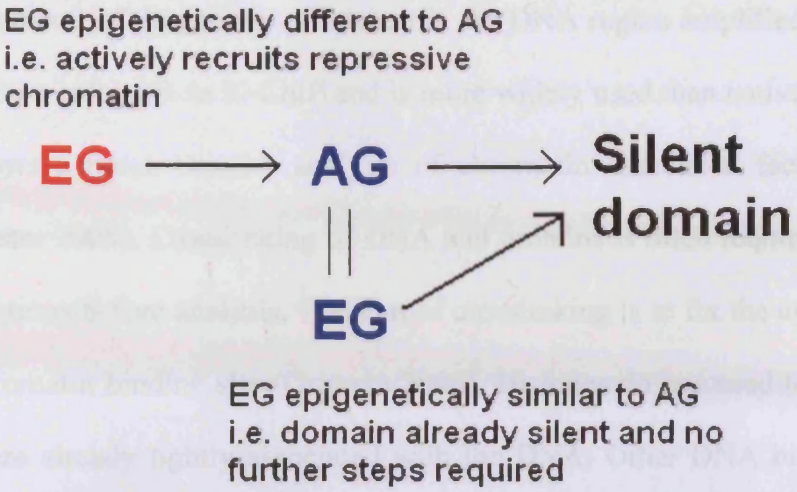


Figure 5.1: Comparison of the IC2 domain in AG and EG cells.

The analysis of the histone modifications associated with the IC2 domain in AG and EG cells would allow the comparison of the histone marks associated with silent chromatin to determine if this silencing was an active or passive process. If AG and EG cells had similar patterns of histone modifications then the domain could be considered to be already silent in undifferentiated EG cells. If, however, repressive histone modification marks were absent in the undifferentiated EG cells but present in the differentiated EG cells and both undifferentiated and differentiated AG cells then it could be concluded that silencing is an active process that occurs during differentiation of the EG cells.

5.2 Design and optimisation of the Chromatin Immunoprecipitation (ChIP) assay

5.2.1 Optimisation of crosslinking and sonication parameters

Chromatin immunoprecipitation (ChIP) can be used to identify the location of various histone modifications along the imprinted domain. The protocol involves crosslinking proteins to DNA by the addition of formaldehyde to the cultured cells. The crosslinked DNA is then sonicated to yield fragments of approximately 100-500 bp. The DNA-protein complexes are incubated with an antibody that binds the protein (or in this case histone modification) of interest, purified by protein sepharose A beads and the antibody bound DNA is eluted. A PCR reaction can then be performed on the eluate and on the input chromatin fraction to determine the relative

'enrichment' of the protein of interest in the DNA region amplified. This crosslinking procedure is known as X-ChIP and is more widely used than native ChIP (N-ChIP) as it allows a much broader analysis of chromatin associated factors (Cuthbert and Bannister 2005). Crosslinking of DNA and proteins is often required to stabilise their interactions before analysis. The aim of crosslinking is to fix the antigen of interest in its chromatin binding site (Orlando 2000). Histones do not need to be crosslinked as they are already tightly associated with the DNA. Other DNA binding proteins that are not as tightly bound to DNA or histones may need to be crosslinked. Crosslinking holds them in place and prevents the proteins from dissociating or moving from the chromatin binding site (Orlando 2000). The further away from the DNA the interaction of interest lies, then the less effective ChIP will be without crosslinking. Formaldehyde is used as the links it forms are reversible with mild heat treatment. After crosslinking, the chromatin can be either enzymatically digested or can be mechanically sheered to generate chromatin fragments of around 300 bp (Cuthbert and Bannister 2005).

Crosslinking of proteins to DNA is an important step in the ChIP protocol. Variables include the percentage of formaldehyde used, the duration of incubation and the temperature. Formaldehyde is a tight crosslinking agent that efficiently produces both protein-protein and protein-DNA complexes *in vivo* within minutes after it is added to living cells (Orlando 2000). In the standard fixation protocol for cell lines, formaldehyde is added to the cell culture media at a final concentration of 1 % for 10 min at room temperature. Other variations on the protocol include varying the formaldehyde concentration from 0.75 % to 1.5 %, shortening the fixation time and performing the fixation reaction at 37 °C. Crosslinking is a time critical procedure and should generally be only carried out for a few minutes. Too much crosslinking can

lead to an increase in the amount of non-specific protein bound to the DNA and even loss of antibody recognition of the epitope (O'Neill and Turner 2003; Orlando 2000). Normally, a time-course experiment is performed to optimise crosslinking conditions (Orlando 2000).

Fragmentation of the chromatin is necessary to break down the DNA into small fragments in order to increase the resolution of the ChIP assay and to make the crosslinked proteins accessible to the antibodies. With Micrococcal endonuclease digestion it is possible to generate single monosomes (~175 bp). Nucleosomes, however, are dynamic and it is possible that they may rearrange during the enzymatic digestion. Enzymatic cleavage will not produce 'random' chromatin fragments but will digest certain areas of genome sequence preferentially over others. Sonication is necessary for X-ChIP as formaldehyde crosslinking restricts the access of endonuclease enzymes to their targets so is inefficient on crosslinked samples (Orlando 2000). Sonication generates randomly sized DNA. The fragments obtained are usually larger than those obtained with enzymatic digestion resulting in a lower resolution, however, sonication is the best way to fragment crosslinked chromatin efficiently. Over-fixed cells are resistant to sonication and can result in the loss of up to 90 % chromatin (Orlando 2000).

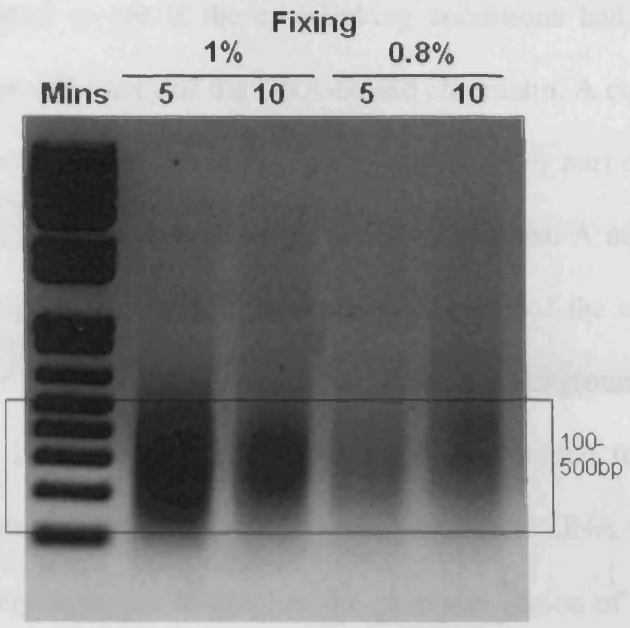


Figure 5.2: Optimisation of formaldehyde concentration on sonication efficiency.

SNL cells were fixed with either 1 % or 0.8 % formaldehyde for either 5 or 10 minutes. Lysate was then sonicated for cycles of 30 sec on, 1 min off for 15 min using the Bioruptor (Diagenode). Crosslinking was then reversed and the DNA extracted and runout on a 1 % agarose TAE gel at 100 V as described in chapter 2. The desired size range for the sonicated DNA was between 100 and 500 bp as indicated by the box.

The first step was to optimise the crosslinking of the chromatin. The embryonic fibroblast cell line SNL was used to optimise the formaldehyde crosslinking as they could be grown quickly and in large numbers. The cells were treated with either 1 % or 0.8 % formaldehyde for five or ten minutes then sonicated using the bioruptor (diagenode) using the standard protocol of 10 cycles of 30 sec of sonication and 1 min off for 15 min. Samples of the sheared chromatin were run on a 1 % agarose TAE gel in order to determine the efficiency of shearing. The results indicated that none of the fixing conditions adversely affected the chromatin shearing efficiency as the range of fragments was concentrated around 100-500 bp (**figure 5.2**).

We then tested to see if the crosslinking conditions had any effect on the immunoprecipitation efficiency of the DNA-bound chromatin. A control antibody that binds Histone H3 was used as a positive control as it is a key part of nucleosome core proteins and should be present at all of the promoter regions. A negative control was also included. Anti-goat IgG should fail to recognise any of the mouse proteins and will be an indicator of any background signal. Another background control using no antibody was also included to check for non-specific binding to the beads. Fixed chromatin was immunoprecipitated and the resulting bound DNA was used in a PCR reaction with primers designed to amplify the promoter region of *Cdkn1c* and the *β-actin* primers used in qPCR. All the crosslinking conditions showed efficient binding of the H3 antibody and little or no background noise was observed in the two negative controls (**figure 5.3**). As a comparison for the efficiency of enrichment, the input DNA was included. 1 % formaldehyde treatment worked better at five minutes incubation time rather than ten minutes as a higher background binding signal was observed in these samples. 0.8 % formaldehyde for both five and ten minutes gave a much higher H3 binding efficiency compared to background possibly indicating that the higher concentration of formaldehyde may somehow be inhibiting binding of the Histone H3 antibody. These initial results indicated that the optimal condition for fixing the chromatin seemed to be 0.8 % formaldehyde for 10 min at room temperature. We did not optimise the temperature for this experiment as the majority of the protocols available incubate the cells at room temperature while fixing and the efficiency of crosslinking at this temperature appeared sufficient for our needs. If we were to look at bound protein complexes it may be necessary to optimise the temperature, however, as we are focusing on histone modifications it does not seem necessary and a higher temperature would likely result in over-crosslinking.

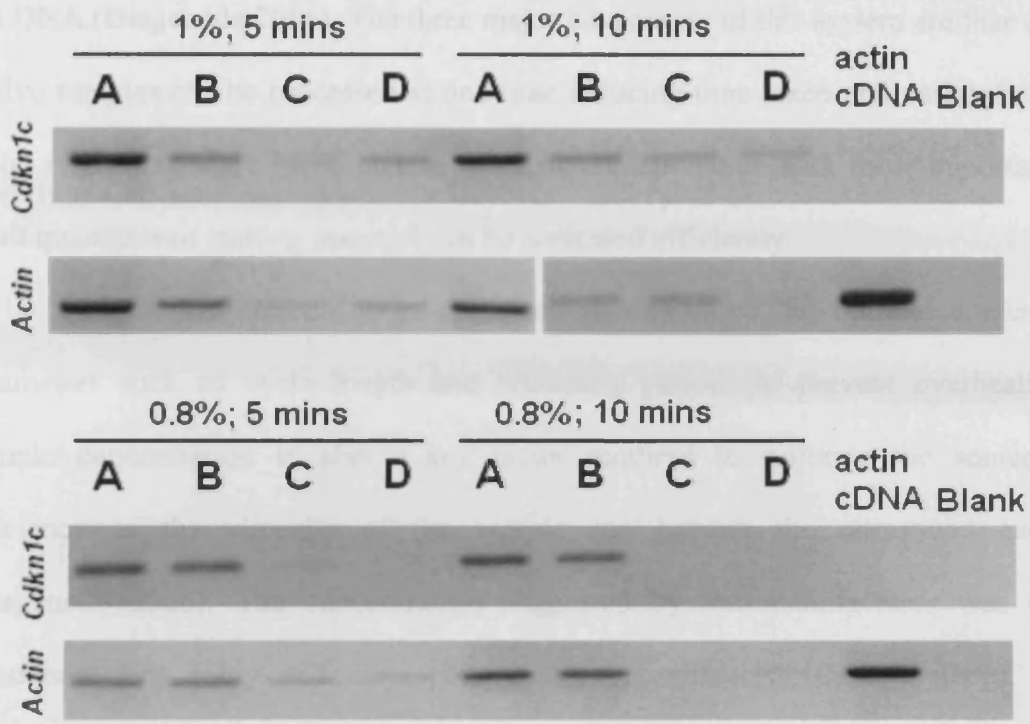


Figure 5.3: Chromatin immunoprecipitation on sonicated lysate from formaldehyde treated SNL cells.

ChIP was performed on chromatin lysate fixed with either 0.8 % or 1 % formaldehyde for either five or ten minutes. Chromatin was sonicated for 10 cyles of 30 sec on/ 1 min off using the bioruptor from diagenode. ChIP was performed on the sonicated chromatin and the DNA was extracted as described in chapter 2. **(A)** Input chromatin, **(B)** histone H3 IP, **(C)** anti-goat IgG IP, **(D)** no antibody IP. PCR was performed to amplify the *Cdkn1c* promoter and *Actin* genes to determine the efficiency of enrichment using the different fixing parameters.

The bioruptor (Diagenode) sonicator system was used to sonicate all the samples within this study. There is a basic program suggested by the manufacturers which involves sonicating for ten minutes at 30 second intervals. All of the chromatin samples sonicated well on this basic program, however, sonication must be optimised to ensure efficient shearing of the DNA. The Bioruptor system is a water bath with a high power ultrasound generator located below the tank. The ultrasound energy causes vibration and alteration of pressure within the water to produce small bubbles from the dissolved gases. These bubbles vibrate and are sufficient to break up cells

and DNA (Diagenode 2006). The three major advantages of this system are that up to twelve samples can be processed at one time reducing time taken and variability, no probe is used so there is no chance of cross contamination and, most importantly, small quantities of starting material can be sonicated efficiently.

Initially the protocol was optimised to determine the optimal sonication parameters such as cycle length and refractory period (to prevent overheating). Sample concentration is also a key factor required to calibrate the sonication efficiency as the viscosity of the sample can hamper the ultrasound energy (Diagenode 2006). The concentration suggested by the manufacturer was used throughout this study and corresponded to approximately 10,000 cells/µl. The sonication parameters should be reproducible to give an optimal and uniformal fragment size.

SNL chromatin was crosslinked for 10 minutes with 1 % formaldehyde as this concentration was shown previously not to impede sonication efficiency. Sonication time and refractory period was varied to determine the effects on shearing efficiency. **Figure 5.4** shows sonication of crosslinked DNA with increasing sonication time. All of the sonication parameters produced DNA fragments in the desired range of 100-500 bp including the sonication parameters suggested by the manufacturer. A sonication time of 15 minutes with intervals of 30 sec on/ 30 sec off repeated after 30 min on ice seemed to be the most optimal as it gave a fragment size within 100-500 bp without any higher molecular weight smears. A smear at high molecular weights can indicate that the DNA has not been sonicated effectively due to over-crosslinking of the DNA to protein. Over sonicating the DNA, however, can result in degradation of the DNA. All of the sonication protocols used here produced fragments in the correct range. We chose to use a lower formaldehyde concentration with the standard

sonication parameters suggested by the manufacturer. The concentration used in these optimisation protocols correlated to approximately 2 x 10^6 SNL cells in 200 µl (10,000 cells/µl) of sonication buffer.

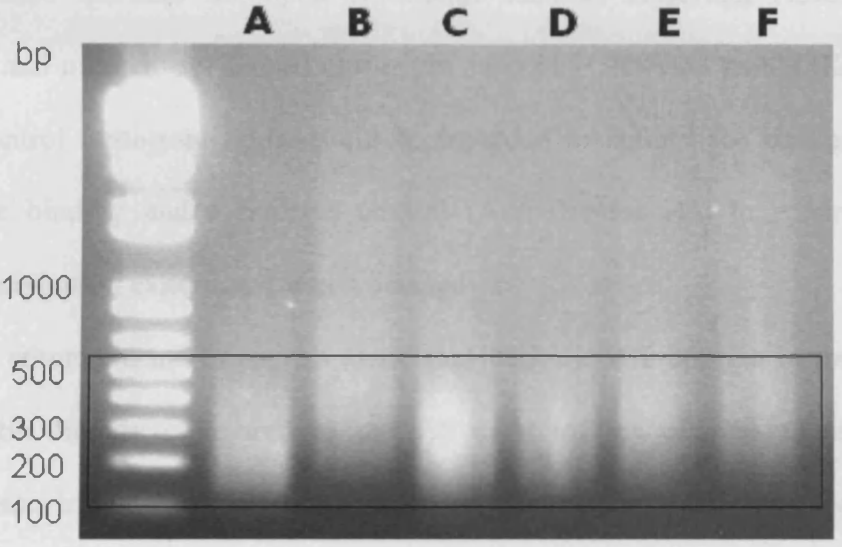


Figure 5.4: Optimisation of sonication parameters using the Bioruptor from Diagenode.

SNL cells were treated with 1 % formaldehyde for 10 min. Crosslinked chromatin was harvested then six 200 µl samples were sonicated using the bioruptor. Parameters were varied to identify the optimum conditions for sonication. **(A)** Sonicated for 30 sec on/ 30 sec off for 15 min, then 30 min on ice followed by 30 sec on/ 30 sec off for another 15 min. **(B)** Sonicated for 30 sec on/ 30 sec off for 15 min. **(C)** Sonicated for 2 min on/ 30 sec off, followed by 30 min on ice then 2 min on/ 30 sec off. **(D)** 2 min on/ 30 sec off for 15 min. **(E)** 2 min on/ 2 min off for 15 min. **(F)** 45 sec on/ 45 sec off for 15 min. Crosslinking was then reversed and the DNA extracted and runout on a 1 % agarose TAE gel at 100 V as described in chapter 2. The box indicates the desired fragment sizes for the sonicated chromatin in the range 100-500 bp.

5.2.2 Design of the chromatin immunoprecipitation protocol

There are several reasons why we chose to use crosslinking rather than native ChIP. Firstly, the amount of chromatin required per native ChIP experiment was a lot higher than that required using crosslinked chromatin and secondly, we wanted the option of looking at chromatin associated factors such as Eed in the future. The

chromatin immunoprecipitation protocol used in the crosslinking and optimisation experiments was designed from a number of protocols from the literature (protocol described in detail in Chapter 2). The plan was to look at a range of histone modifications that had been reported at the promoter regions of imprinted genes. These included markers for open chromatin such as H3K9Ac, H3K4Me2 and H3K4me3, and markers for closed chromatin such as H3K9Me3 and H3K27Me3. A negative control (anti-goat IgG) would be included to control for background and non-specific binding and a positive control (Anti-Histone H3) to ensure that the immunoprecipitation experiment was working.

The chromatin modifications associated with the IC2 domain in the EG cells could not be directly compared to our ES cells in this case as there were no polymorphisms in which to determine allele-specific differences. Instead, we decided to compare the chromatin structure in EG cells to that seen in the AKR1 cell line which contains two paternal genomes. This should be representative of the chromatin modifications found at the paternal allele only and so should be enriched for markers of silent chromatin.

We examined the promoter regions of *Cdkn1c, Phlda2, Kcnq1ot1/Lit1* and *Kcnq1* to allow us to look at the spread of chromatin across the domain. Primers were designed using the Primer3 program to amplify the promoter regions of the IC2 domain genes. There are many published primers available that have been shown to work efficiently in ChIP. However, most of these are designed to amplify DNA over polymorphisms in order to determine allelic specific differences in enrichment. The regions amplified do not necessarily correlate with the promoter regions of the imprinted genes and are often several hundred base pairs from the transcriptional start sites of the genes. We designed primers specifically for the promoters so that we

could correlate the expression status of these genes with the enrichment of histone marks at their promoters. The primers were first tested on SNL input chromatin to determine whether they amplified a product of the correct size in a normal PCR reaction. Primers were also designed to amplify the upstream CpG island of *Cdkn1c* and the intron II region of *Cdkn1c* to determine if there were changes to these regions over time. However, the primer sets designed for these regions failed to amplify a product possibly due to the GC rich nature of the DNA associated with these regions. Primers that were shown to work on SNL input DNA were then tested on immunoprecipitated DNA from crosslinked undifferentiated TMAS21G chromatin. **Figure 5.5** shows the results from testing the PCR primers using sonicated undifferentiated TMAS21G EG cells in input samples, H3 immunoprecipitated samples and negative control IgG IP samples. The results showed that the PCR primers amplified products of the correct size. The results also showed that there was an enrichment of the DNA in the immunoprecipitated samples. The differences in the intensities of the bands relates to the difference in the amplification efficiencies of the primer sets.

When designing and optimising this protocol, a number of problems were identified. Firstly, the number of ChIP experiments that can be done was limited to three per 2 x 10^6 cells. This is the maximum number of cells that we were able to yield from one confluent 10 cm plate of stem cells. At least six IPs per time point were required which would correspond to two full confluent plates of ES cells. Input samples of chromatin would also be required. The first protocol used was designed from a variety of sources (see chapter 2 materials and methods). However, there were problems with efficiency with this method. Due to time constraints we looked into using ChIP kits in order to boost efficiency and save time (**table 5.1**).

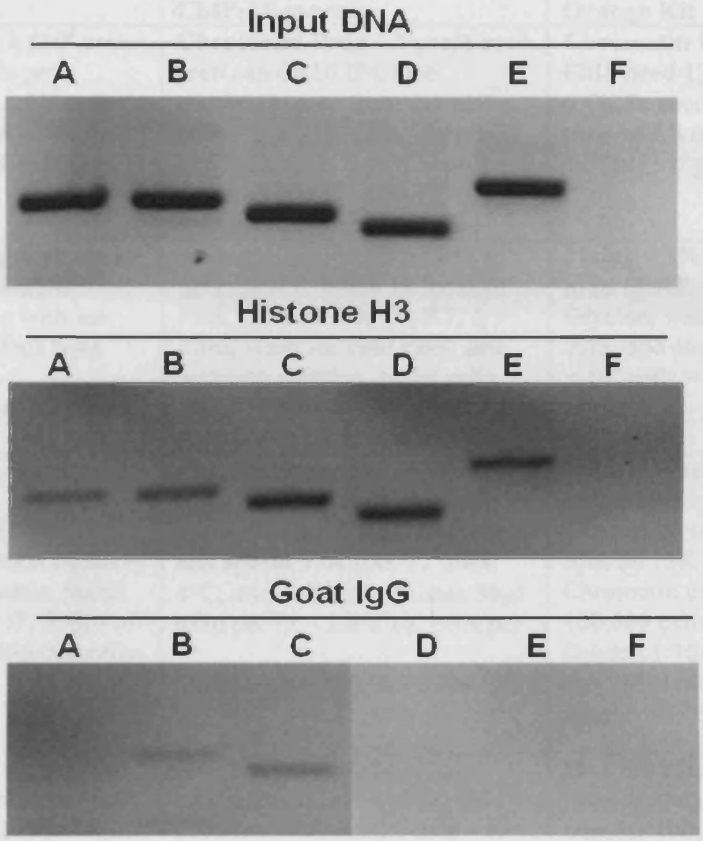


Figure 5.5: Chromatin immunoprecipitation on undifferentiated EG cells to show that the primers amplify the immunoprecipitated and input DNA.

(A) *Cdkn1c,* (B) *Phlda2*, (C) *Kcnq1ot1*, (D) *Kcnq1*, (E) *Actin.* TMAS21G cells were cultured in the presence of a feeder layer and LIF until confluent. The feeder layer was removed and the single cell suspension of stem cells was then treated with 0.8 % formaldehyde for 10 min. The cells were sonicated for 10 cycles of 30 sec on/ 30 sec off for 10 min and the sonicated chromatin was used in a ChIP assay using the histone H3 positive control antibody and the Goat IgG negative control antibody. Pulled down and input DNA was amplified using primers designed for the promoter regions of the IC2 domain genes.

Table 5.1: Comparison of X-ChIP methods.

My Protocol	ChIP-IT express	Orange Kit
Chromatin prep – 2 x10^6 cells for 3 IPs. Need 6 IPs per chromatin prep = 4 x 10^6 cells per cell line per time point (~ 2 plates @ 2 x10^6 cells/plate) = 600,000 cells/IP	**Chromatin Prep** – 5 preps each prep can do 16 IPs. One chromatin prep from 4.5 x10^7 cells = 2.8 x10^6 cells/IP (~ 1-3 x10^6 cells/IP). Total ES cells required = at least 6 x10^6 cells ~ 4 plates/ timepoint/ cell line	**Chromatin Prep** – For one ChIP need 125,000 cells/IP. For 6 ChIPs need 750,000 cells. One plate of ES cells will do at least 12 ChIPs if get 2 x10^6 cells.
Fixing – 0.8% formaldehyde 10 mins @ RT. 1/10^{th} 1.25M Glycine, wash twice with ice cold 1xPBS, add 200µl lysis buffer. = 10,000 cells/µl buffer.	**Fixing** – 1% formaldehyde 10 mins @ RT. Wash 1x ice cold PBS, Glycine stop fix RT, 5 mins, wash ice cold PBS. 2ml scraping solution, pellet cells, add 1ml lysis buffer = 45,000 cells/µl buffer	**Fixing** – 1% formaldehyde 10 mins @ RT. 1/10^{th} 1.25M Glycine, wash twice ice cold 1x PBS, add 60µl buffer A for 0.75 x 10^6 cells = 12, 500 cells/µl buffer
Lysis & Shearing – On ice 10 mins, sonicate with bioruptor 30s on/ 1min off 15 mins. Dilute in 1.8ml dilution buffer = 1000 cells/µl chromatin. 600µl chromatin used per IP. 200µl of input chromatin = 200,000 cells.	**Lysis & Shearing** - Split into three aliquots, Sonicated with a probe sonicator. Pool chromatin and spin at 13K rpm 12 mins, 4°C, aliquoted into 4 tubes 50µl used per IP ~ 2.8 x 10^6 cells per IP. 10µl Input = 560,000 cells.	**Lysis & Shearing** – Sonicate with a bioruptor 30s on/ 30s off, 10 mins up to 300 µl per tube. Spin @ 13K rpm, 5 mins. 10µl Chromatin used per IP ~ 100,000 cells/IP Diluted 1/10 in IP incubation mix. 10µl diluted chromatin is input = 10,000 cells.
IP – Pre-cleared with pre-blocked agarose A slurry (60µl) @ 4°C 2 hours. 15K rpm 1 min, divide into 3 tubes, add 1-3 µg antibody, o/n @ 4°C, 20µl agarose slurry, 4°C 2 hours, 15k rpm 1 min	**IP** – 50µl chromatin, 25µl beads, 10µl ChIP buffer, 1µl protease inhibitors, 1-3µg antibody, up to 100µl water, o/n @ 4°C. spin briefly, pellet with magnet, remove supernatant.	**IP** – Pre-clear 20 µl beads, 1 hour @ 4°C, spin 2500rpm, transfer 100µl to new tube, add 0.5µg Ab/IP (0.5-2µg) o/n @ 4°C. Add 20µl beads incubate @ 4°C 1 hour. Spin 2500 rpm 2 mins
Washes – once low salt, once high salt, once LiCl, twice TE	**Washes** – once ChIP buffer 1, twice ChIP buffer 2	**Washes** – once buffer 1, twice buffer 2, twice buffer 3
Elution – 200µl elution buffer with 8.5µl NaCl rotation RT for 15 mins; o/n @ 65°C stored input treated for crosslinking. Dilute samples in 200µl Tris, 20µg RNase A @ 37°C, 30 mins, 30µg PK @ 55°C 1 hour, Phenol chloroform extract. ChIP = 30µl, Input = 50µl 10mM Tris. IP= 20,000 cells/µl	**Elution** – 100µl elution buffer, add 88µl ChIP buffer 2 + 2µl 5M NaCl to input DNA. o/n @ 65°C, 2µl PK 37°C 1 hour, 2µl PK stop solution, Pellet beads with magnet. Use in PCR. Input must be phenol chloroform extracted. IP= 27,000 cells/µl	**Elution** – Add 400µl buffer C, 16µl 5M NaCl, RT with rotation. Spin 2500rpm @RT. Transfer to new tube, add 390µl buffer C to diluted input. 16µl 5M NaCl thermoshaker 4 hours @ 65°C to reverse crosslinking, cool, phenol-chloroform extract resuspend in 50µl. IP= 2000 cells/µl

Two ChIP kits were examined to determine which would be the most efficient (**table 5.1**). The Orange ChIP by Diagenode came out the best. It was designed specifically for histone modifications and used the least amount of cells. The ChIP-IT express kit would be more useful for examining bound proteins such as Eed and other

transfactors. Using the conditions determined in the optimising sonication section, we analysed the sonication efficiency of the EG and ES cell lysates using the Orange ChIP kit. All the undifferentiated stem cell lysates sonicated efficiently between 100 and 500 bp (**figure 5.6 & 5.7**). Actin DNA could be amplified from each sonicated lysate indicating that the chromatin was intact.

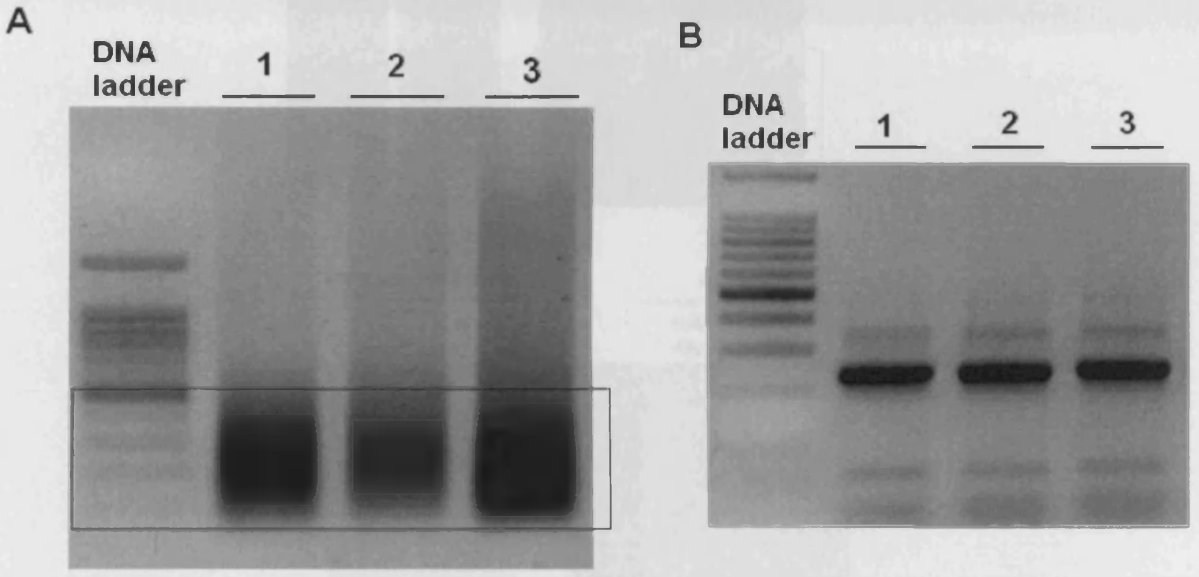


Figure 5.6: Undifferentiated EG and AG cell chromatin lysate sonication efficiency.

A (1) SNL control (163 ng/µl), (2) Sv6 d0 (141 ng/µl), (3) AKR1 d0 (195 ng/µl). Two confluent plates of undifferentiated ES cells were harvested and treated with 0.8 % formaldehyde for 10 min at room temperature. Chromatin was harvested using the orange ChIP kit from diagenode and sonicated as described previously in figure 5.5. Crosslinking was then reversed and the DNA extracted and runout on a 1 % agarose TAE gel at 100 V as described in chapter 2. The concentration of the DNA was determined by spectrophotometry. The box highlights the location of the majority of the sonicated fragments which are found in a region of less than 500 bp. **B** (1) SNL control, (2) Sv6.1 d0, (3) AKR1 d0. Approximately 100 ng of DNA was used in a PCR reaction to amplify *β-actin*.

The sonicated lysate from the differentiated EG cells was sonicated to the correct DNA range. However, the amount of DNA recovered after reversing crosslinks was considerably lower than expected (**figure 5.7**). This was likely a result of the trypsinisation process which can cause cell death.

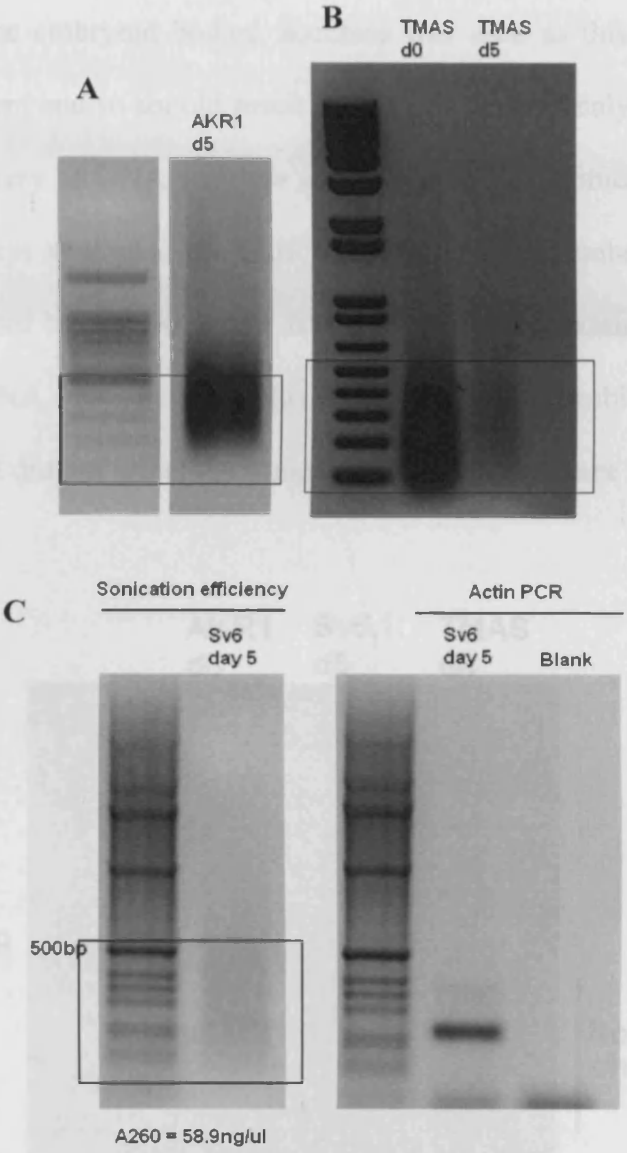


Figure 5.7: Sonication efficiency of the AKR1, TMAS21G and Sv6.1 day 5 lysates after digesting the embryoid bodies with trypsin.

(A) AKR1, (B) TMAS21G and (C) Sv6.1 ES cell sonication efficiencies. Cells were grown and differentiated for 5 days as embryoid bodies. Embryoid bodies were harvested and trypsinised to produce a single cell suspension. Cells were treated as described in figure 5.5. Crosslinking was then reversed and the DNA extracted and runout on a 1 % agarose TAE gel at 100 V as described in chapter 2. DNA concentration was checked by spectrophotometry and are as follows: AKR1 d5 (91.8 ng/µl), TMAS21G d0 (96.7 ng/µl), TMAS21G d5 (8.2 ng/µl) and Sv6.1 d5 (8.9ng/µl). TMAS21G d0 chromatin sonication efficiency is shown as a comparison. A PCR reaction was performed on the extracted Sv6.1 d5 chromatin to amplify *β-actin*. Boxes indicate the location of the majority of the sonicated fragments around 100-500 bp.

The experiment was repeated for the day 5 lysates but instead of using trypsin to disaggregate the embryoid bodies, accutase was used as this was a more gentle enzymatic treatment and so should result in less cell death. Analysis of the chromatin showed that recovery of DNA was low again (**figure 5.8**). Sonication efficiency was unaffected. We next attempted the ChIP protocol on whole embryoid bodies. As the size of the embryoid bodies was small, it should not inhibit fixation by formaldehyde. The amount of DNA obtained using this method was comparable to that seen in the day 0 samples and did not affect the sonication efficiency (**figure 5.9**).

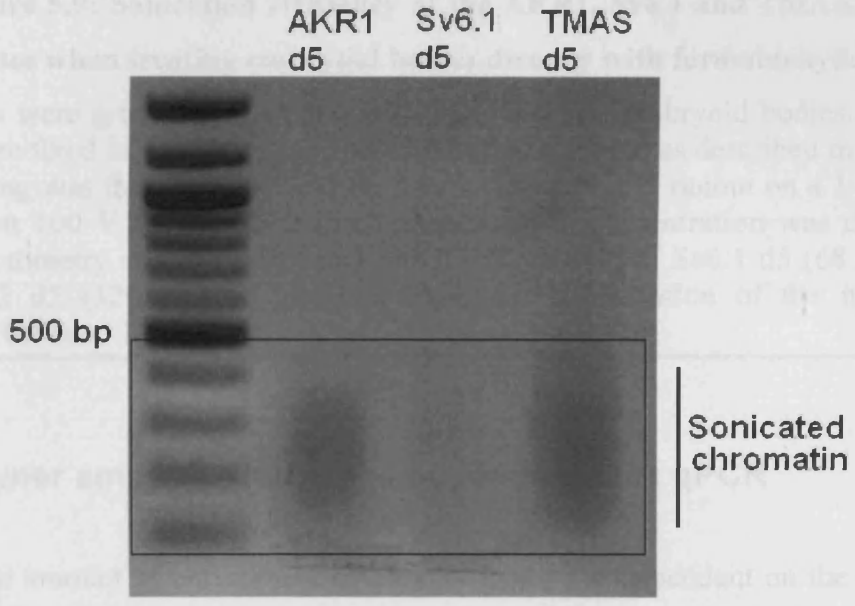


Figure 5.8: Sonication efficiency of the AKR1, Sv6.1 and TMAS21G day 5 lysates after digesting the embryoid bodies with accutase.

Stem cells were grown and differentiated for 5 days as embryoid bodies. Embryoid bodies were harvested and treated with accutase to produce a single cell suspension. Chromatin was harvested as described in figure 5.5. Crosslinking was then reversed and the DNA extracted and runout on a 1 % agarose TAE gel at 100 V as described in chapter 2. DNA concentration was checked by spectrophotometry and are as follows: AKR1d5 (18.7 ng/µl), Sv6.1 d5 (7.6 ng/µl) and TMAS d5 (15 ng/µl). The box highlights the location of the majority of fragments between 100-500 bp.

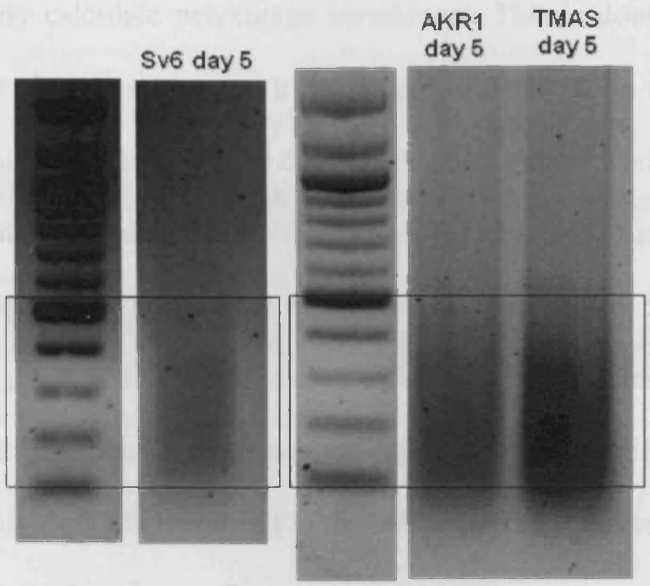


Figure 5.9: Sonication efficiency of the AKR1, Sv6.1 and TMAS21G day 5 lysates when treating embryoid bodies directly with formaldehyde.

Stem cells were grown and differentiated for 5 days as embryoid bodies. Embryoid bodies were fixed in formaldehyde and chromatin harvested as described in figure 5.5. Crosslinking was then reversed and the DNA extracted and runout on a 1 % agarose TAE gel at 100 V as described in chapter 2. DNA concentration was checked by spectrophotometry and are as follows: AKR1d5 (65 ng/µl), Sv6.1 d5 (68 ng/µl) and TMAS21G d5 (126 ng/µl). The box highlights the location of the majority of fragments between 100-500 bp.

5.2.3 Primer amplification efficiency in real time qPCR

The amount of enrichment of a target is not just dependent on the amount of antigen available. Immunoprecipitation is affected by the surrounding chromatin environment, such as the nearby histone modifications and attached proteins so that, for each gene and each antibody, there will be a difference in the actual enrichment of the target DNA. For this reason, absolute levels of different antigens present at the same sequence cannot be directly compared. We can only compare the enrichment of one histone modification at the different locations and is always expressed as a ratio of bound sequence over input or as percentage enrichment over input (Cuthbert and Bannister 2005). The primer amplification efficiencies first need to be determined in

order to accurately calculate percentage enrichment. This is done by making serial dilutions of input chromatin then performing real time qPCR on the DNA using the ChIP primers. This also allowed us to confirm that the primers worked in the real time PCR. The CT value at which the DNA is amplified above the base line was plotted against natural log of the dilution for each set of primers (**figure 5.10**). The gradient of the resulting graph is the amplification efficiency of the primer set. In an ideal setting this would be around 2. However we constantly saw amplification efficiencies of around 1.4. This figure is then used in the calculations for percentage amplification efficiency as described in chapter 2.

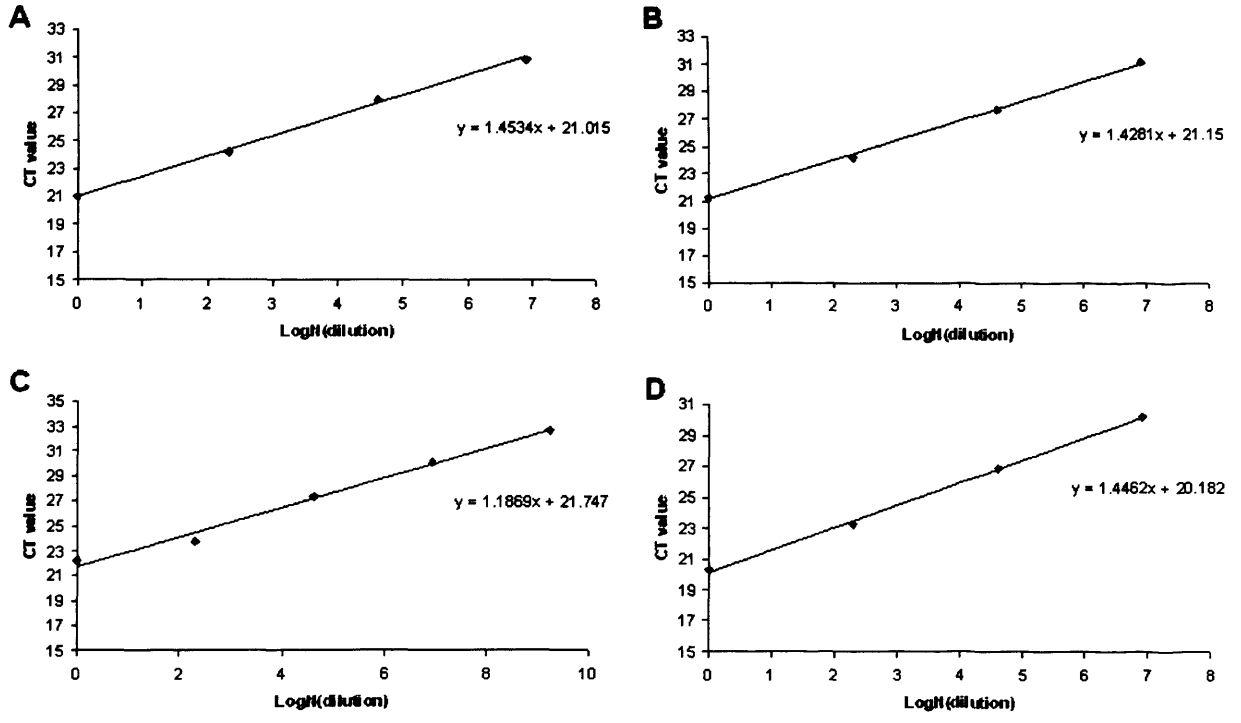


Figure 5.10: Primer amplification efficiencies using real time PCR machine.

Serial dilutions of 0.8 % formaldehyde fixed input TMAS21G chromatin was used to test the amplification efficiency of the ChIP primers. (A) *Phlda2,* (B) *Cdkn1c,* (C) *Kcnq1ot1/Lit1,* (D) *Kcnq1.* Average CT was plotted against the natural log of the dilution factor. The gradient of the trendline is the amplification efficiency of the primers and is used as part of the percentage enrichment calculations. Percentage enrichment = A^(Input CT-IP CT)*100 where A is the amplification efficiency.

The percentage enrichment shown in the results takes into account background levels of enrichment due to non-specific binding to the plastic-ware and sepharose beads. Each antibody has a different enrichment efficiency so can not be directly compared to each other. From the data it would appear that levels of H3K4me2 and H3K9me3 are at found at much higher levels at all the promoters as the percentage enrichment is often similar to H3 levels of enrichment; whereas the enrichment of all other modifications are extremely low. This is likely related to the efficiency of pull down by the antibody rather than actual levels of the histone modifications. To quantify actual levels we would need to compare our levels of enrichment to genes that are known to be depleted for the modifications of interest. For example, our analysis looked at the changes in the histone modifications across the entire IC2 domain and we are comparing the enrichment at the repressed promoters to the active promoters within this domain. In domains where the expression status of the genes is unknown or there are no other genes in which to compare enrichment, we would have to compare the enrichment at our target gene to a repressed gene known to be highly enriched for our histone modification. This is usually a tissue specific gene that is silent in the cells of interest, such as *Myoglobin*. Enrichment of active chromatin markers are normally compared to constitutively expressed genes such as *β-actin*. Our expression analysis data (chapter 3) already showed that *Kcnq1ot1/Lit1* is expressed at levels that appeared biallelic in undifferentiated and differentiated EG cells and it has been shown by other studies that this gene is also biallelically expressed in androgenetic cells (Lee *et al.* 1999; Ogawaa *et al.* 2006; Ruf *et al.* 2006). The other genes are known to be silent so, by comparing the pattern of histone modifications across the promoters of the IC2 domain genes, we can look at the association of gene expression to histone modifications.

5.3 Results

5.3.1 Analysis of the enrichment of active histone modifications in undifferentiated and differentiated EG and AG cells

In ES cells, the paternally expressed *Kcnq1ot1/Lit1* gene is associated with enrichment of active histone modifications whereas the paternally silenced *Cdkn1c* and *Phlda2* are associated with repressive histone modifications (Umlauf *et al.* 2004). From the previous chapters we observed that *Kcnq1ot1/Lit1* was expressed in undifferentiated EG cells so was likely to be enriched for active modifications. *Cdkn1c* and *Phlda2* were not expressed in undifferentiated EG cells so may be associated with repressive histone modifications.

Three histone modifications associated with 'open' chromatin and two histone modifications associated with 'closed' chromatin were examined in undifferentiated and differentiated EG cells using the chromatin immunoprecipitation protocol. The enrichment of these histone modifications was also examined in AKR1 androgenetic cells which contain two paternal genomes. As there are no available polymorphisms within our ES cell line, the AKR1 cell line was used to examine the final chromatin modifications associated with the repressed paternal allele in ES cells. This allowed us to compare 'imprinted' chromatin to 'non-imprinted' chromatin.

Figure 5.11 shows the enrichment of markers of 'open' chromatin across the IC2 domain in Sv6.1 EG cells. Enrichment of H3K9Ac was highest at the *Kcnq1ot1/Lit1* promoter in undifferentiated cells (**figure 5.11 A**). Lower levels of enrichment were associated with the promoters of the silent genes within this domain. After differentiation of the cells, there was a similar pattern of enrichment with the

highest levels of H3K9Ac at the *Kcnq1ot1/Lit1* and *Phlda2* promoters and the lowest levels found at the promoters of *Cdkn1c* and *Kcnq1* (**figure 5.11 B**).

A similar pattern was seen for H3K4me3 where the greatest enrichment of this histone modification was associated with the promoters of *Kcnq1ot1/Lit1, Phlda2* and *Kcnq1* in undifferentiated Sv6.1 EG cells (**figure 5.11 E**). After differentiation of the EG cells, the pattern remained relatively the same with the highest levels at *Kcnq1ot1/Lit1* and *Phlda2* (**figure 5.11 f**). H3K4me2 levels, in contrast, appeared to be enriched only at *Phlda2* and *Kcnq1* promoters in undifferentiated cells (**figure 5.11 C**). After differentiation, the levels of H3K4me2 appeared to be relatively constant across the entire domain (**figure 5.11 D**).

In the TMAS21G EG cells we observed similar patterns of enrichment in undifferentiated and differentiated cells as we saw with the Sv6.1 EG cells (**figure 5.12**). H3K9Ac levels appeared to be specifically enriched at the *Kcnq1ot1/Lit1* promoter in undifferentiated cells (**figure 5.12 A**). After differentiation, the pattern of enrichment was maintained although levels of H3K9Ac were also enriched at the *Phlda2* promoter (**figure 5.12 B**). Contrary to what was seen with the Sv6.1 EG cells, the highest enrichment of H3K4me2 was observed at the promoter of *Kcnq1ot1/Lit1* and *Kcnq1* in undifferentiated cells (**figure 5.12 C**). The lowest levels were observed at the promoters of *Cdkn1c* and *Phlda2*. After differentiation, there was no specific enrichment of H3K4me2 at any of the promoters (**figure 5.12 D**). The greatest enrichment of H3K4me3 was observed at the *Kcnq1ot1/Lit1* promoter in undifferentiated cells (**figure 5.12 E**) and this pattern was maintained in differentiated cells (**figure 5.12 F**).

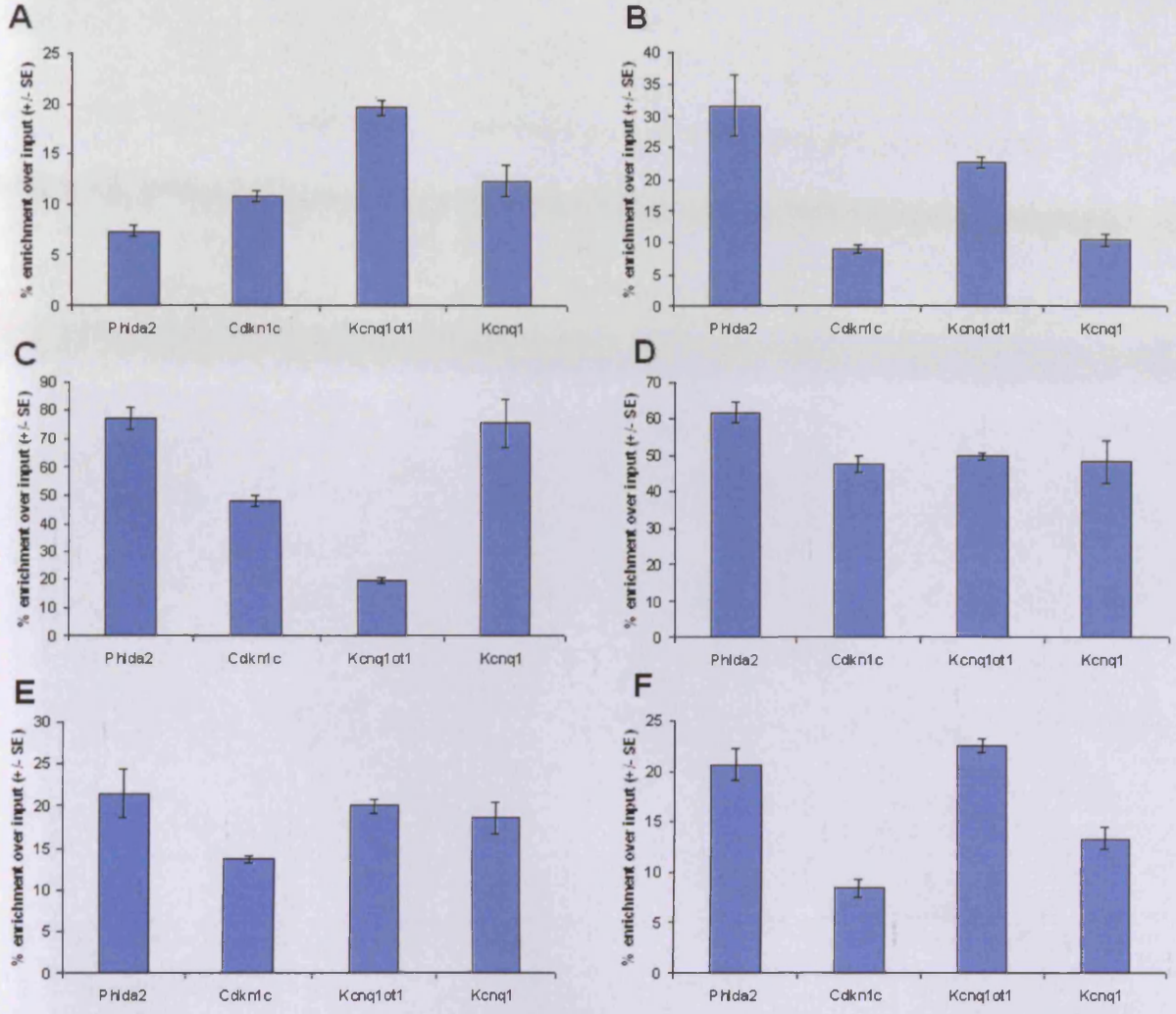


Figure 5.11: Enrichment of active histone modifications at the promoter regions of *Phlda2, Cdkn1c, Kcnq1ot1/Lit1* and *Kcnq1* in the undifferentiated (day 0) and differentiated (day 5) Sv6.1 EG cell line.

The percentage enrichment of each target over input levels is shown for each promoter (calculated as described in figure 5.10). ChIP experiments were carried out using the Orange ChIP kit (Diagenode) using the optimised sonication and fixing conditions. **A** and **B** Enrichment of H3K9Ac, **C** and **D** enrichment of H3K4me2, **E** and **F** enrichment of H3K4me3. **A, C** and **E** results from undifferentiated cells; **B, D** and **F** results from differentiated cells. *Kcnq1ot1/Lit1* promoter is also the location of the ICR.

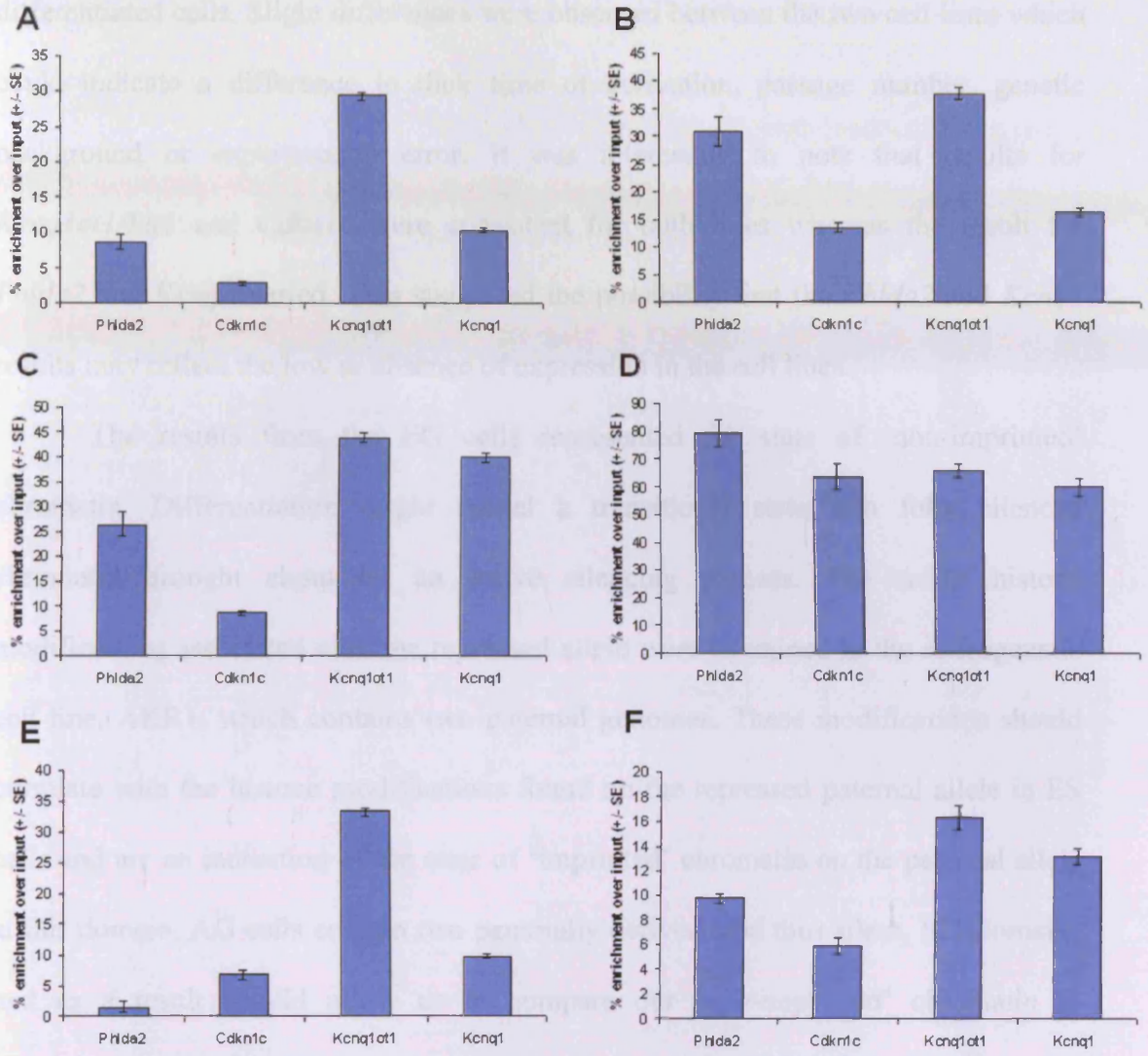


Figure 5.12: Enrichment of active histone modifications at the promoter regions of *Phlda2, Cdkn1c, Kcnq1ot1/Lit1* and *Kcnq1* in the undifferentiated (day 0) and differentiated (day 5) TMAS21G EG cell line.

The percentage enrichment of each target over input levels is shown for each promoter (calculated as described in figure 5.10). ChIP experiments were carried out using the Orange ChIP kit (Diagenode) using the optimised sonication and fixing conditions. **A** and **B** Enrichment of H3K9Ac, **C** and **D** enrichment of H3K4me2, **E** and **F** enrichment of H3K4me3. **A, C** and **E** results from undifferentiated cells; **B, D** and **F** results from differentiated cells. *Kcnq1ot1/Lit1* promoter is also the location of the ICR.

Our analysis indicated that histone modifications associated with actively expressed genes were enriched at the *Kcnq1ot1/Lit1* promoter yet depleted at the *Cdkn1c* promoter in both EG cell lines. This pattern was essentially maintained in the

differentiated cells. Slight differences were observed between the two cell lines which could indicate a difference in their time of derivation, passage number, genetic background or experimental error. It was interesting to note that results for *Kcnq1ot1/Lit1* and *Cdkn1c* were consistant for both lines whereas the result for *Phlda2* and *Kcnq1* varied. This suggested the possibility that the *Phlda2* and *Kcnq1* results may reflect the low or absence of expression in the cell lines.

The results from the EG cells represented the state of 'non-imprinted' chromatin. Differentiation might model a transitional state into fully silenced chromatin brought about by an active silencing process. The active histone modifications associated with the repressed allele were examined in the androgenetic cell line, AKR1, which contains two paternal genomes. These modifications should correlate with the histone modifications found on the repressed paternal allele in ES cells and are an indication of the state of 'imprinted' chromatin on the paternal allele at this domain. AG cells contain two paternally derived, and thus silent, IC2 domains and as a result should allow us to compare our 'non-imprinted' chromatin to chromatin that has passed through the paternal germ line and undergone full silencing of the domain. Differentiation of the AG cells would also allow us to observe changes in chromatin that are related to differentiation of the cells. **Figure 5.13 A** and **B** show the enrichment of H3K9Ac before and after differentiation, respectively. H3K9Ac levels were highest at the *Kcnq1ot1/Lit1* promoter in undifferentiated cells. However, after differentiation the highest levels of enrichment were observed at the *Phlda2* promoter. **Figure 5.13 C** and **D** showed the enrichment of H3K4me2. As with the differentiated EG cells, levels of H3K4me2 were level across the entire domain however, after differentiation, the highest level of H3K4me2 enrichment was observed at *Phlda2*. The promoters of *Phlda2, Kcnq1ot1/Lit1* and *Kcnq1* were

enriched for H3K4me3 in undifferentiated AG cells (**figure 5.13 E** and **F**). After differentiation, this pattern was essentially the same.

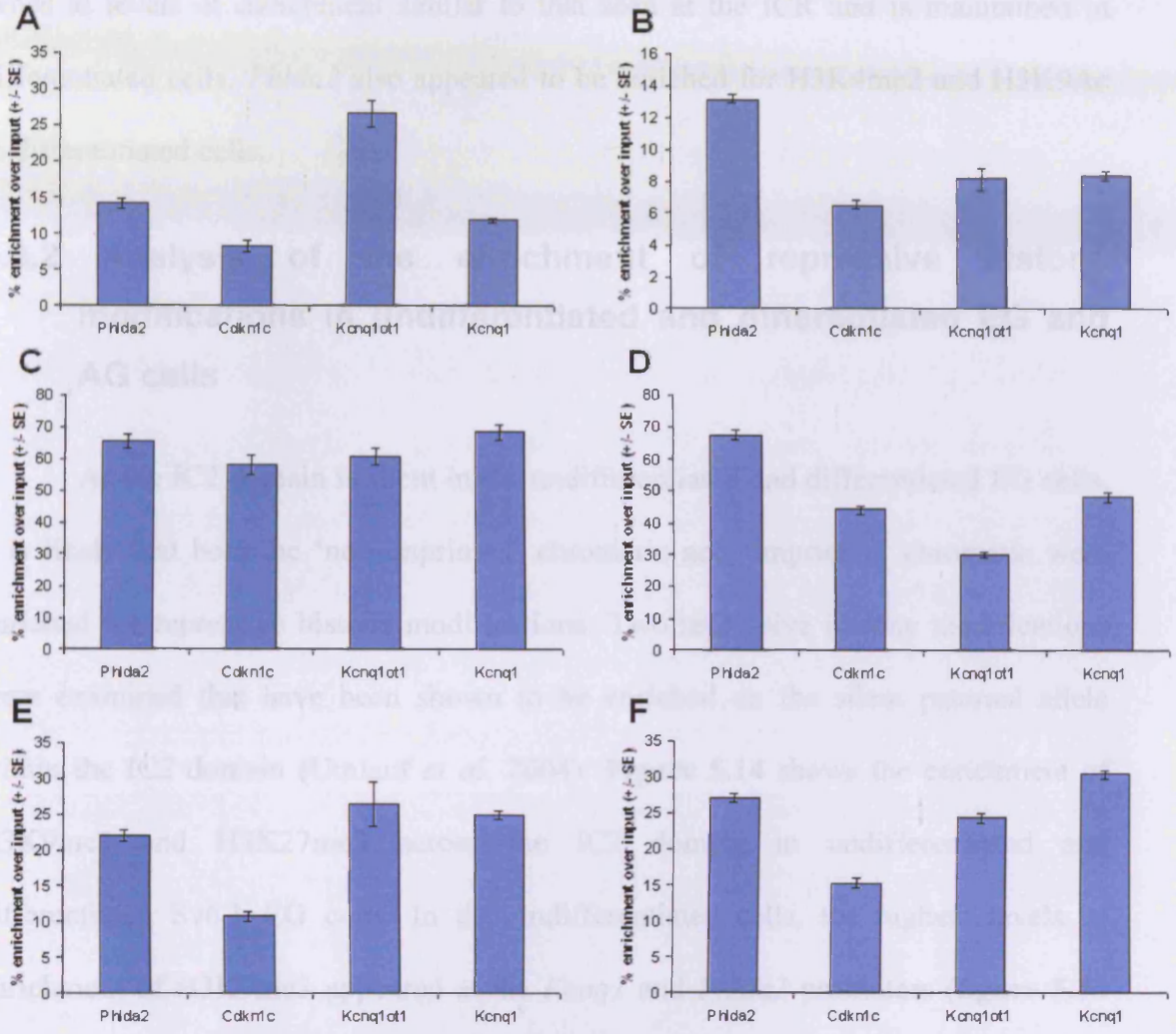


Figure 5.13: Enrichment of active histone modifications at the promoter regions of *Phlda2, Cdkn1c, Kcnq1ot1/Lit1* and *Kcnq1* in the undifferentiated (day 0) and differentiated (day 5) AKR1 AG cell line.

The percentage enrichment of each target over input levels is shown for each promoter (calculated as described in figure 5.10). ChIP experiments were carried out using the Orange ChIP kit (Diagenode) using the optimised sonication and fixing conditions. **A** and **B** Enrichment of H3K9Ac, **C** and **D** enrichment of H3K4me2, **E** and **F** enrichment of H3K4me3. **A, C** and **E** results from undifferentiated cells; **B, D** and **F** results from differentiated cells. *Kcnq1ot1/Lit1* promoter is also the location of the ICR.

These results indicated that, in undifferentiated AG cells, *Kcnq1ot1/Lit1* is marked with H3K9Ac and H3K4me3. H3K4me3 also marks the *Phlda2* and *Kcnq1* genes at levels of enrichment similar to that seen at the ICR and is maintained in differentiated cells. *Phlda2* also appeared to be enriched for H3K4me2 and H3K9Ac in differentiated cells.

5.3.2 Analysis of the enrichment of repressive histone modifications in undifferentiated and differentiated EG and AG cells

As the IC2 domain is silent in the undifferentiated and differentiated EG cells, it is likely that both the 'non-imprinted' chromatin and 'imprinted' chromatin were enriched for repressive histone modifications. Two repressive histone modifications were examined that have been shown to be enriched on the silent paternal allele within the IC2 domain (Umlauf *et al.* 2004). **Figure 5.14** shows the enrichment of H3K9me3 and H3K27me3 across the IC2 domain in undifferentiated and differentiated Sv6.1 EG cells. In the undifferentiated cells, the highest levels of enrichment of H3K9me3 appeared at the *Kcnq1* and *Phlda2* promoters (**figure 5.14 A**). After differentiation this pattern was maintained (**figure 5.14 B**). The levels of H3K9me3 enrichment were similar between the promoters of *Cdkn1c* and *Kcnq1ot1/Lit1* at both time points.

H3K27me3 is known to be associated with the repressed alleles of *Cdkn1c* and *Kcnq1* in somatic cells (Umlauf *et al.* 2004). In undifferentiated cells, we observed enrichment of H3K27me3 at the *Phlda2* and *Kcnq1* promoters but an absence of enrichment at the *Cdkn1c* and *Kcnq1ot1/Lit1* promoters (**figure 5.14 C**). After differentiation of the cells there was an increase in enrichment at the *Cdkn1c* promoter but levels remained low at the *Kcnq1ot1/Lit1* promoter, possibly indicating that this

repressive histone modification was recruited to the *Cdkn1c* promoter. This modification also appeared to be lost from the *Kcnq1* promoter in differentiated cells.

The enrichment of the repressive histone modifications was also examined in the second EG cell line, TMAS21G. This cell line showed a different pattern of H3K9me3 enrichment to that observed in the undifferentiated Sv6.1 cells (**figure 5.15 A**). H3K9me3 enrichment was highest at the *Kcnq1ot1/Lit1* promoter. After differentiation, the levels of H3K9me3 appeared to be constant across the domain with the highest levels at the *Phlda2* and *Kcnq1* promoters (**figure 5.15 B**). The levels of H3K27me3 across the IC2 domain showed the same pattern as observed in Sv6.1 EG cells (**figure 5.15 C** and **D**). In undifferentiated cells, H3K27me3 was enriched on the promoter region of *Kcnq1* and was depleted at all other promoters. After differentiation of the EG cells we observed enrichment of H3K27me3 at the *Cdkn1c* and *Phlda2* promoters to levels similar to that observed at the *Kcnq1* promoter.

The data from the EG cells indicated that there were no differences in the levels of H3K9me3 across the IC2 domain in differentiated cells. In undifferentiated cells, there were some differences in H3K9me3 enrichment between the cell lines. H3K27me3 was unchanged after differentiation at the promoter region of *Kcnq1,* and *Phlda2* was enriched for H3K27me3 only in differentiated TMAS21G cells. However, the *Cdkn1c* promoter became enriched for H3K27me3 in both cell lines, suggesting that further silencing of *Cdkn1c* had occurred which did not occur at the *Phlda2* and *Kcnq1* promoters. One key difference between *Cdkn1c, Kcnq1* and *Phlda2* was their expression level in stem cells as we did not detect *Kcnq1* by qRT-PCR and *Phlda2* was only expressed at low levels.

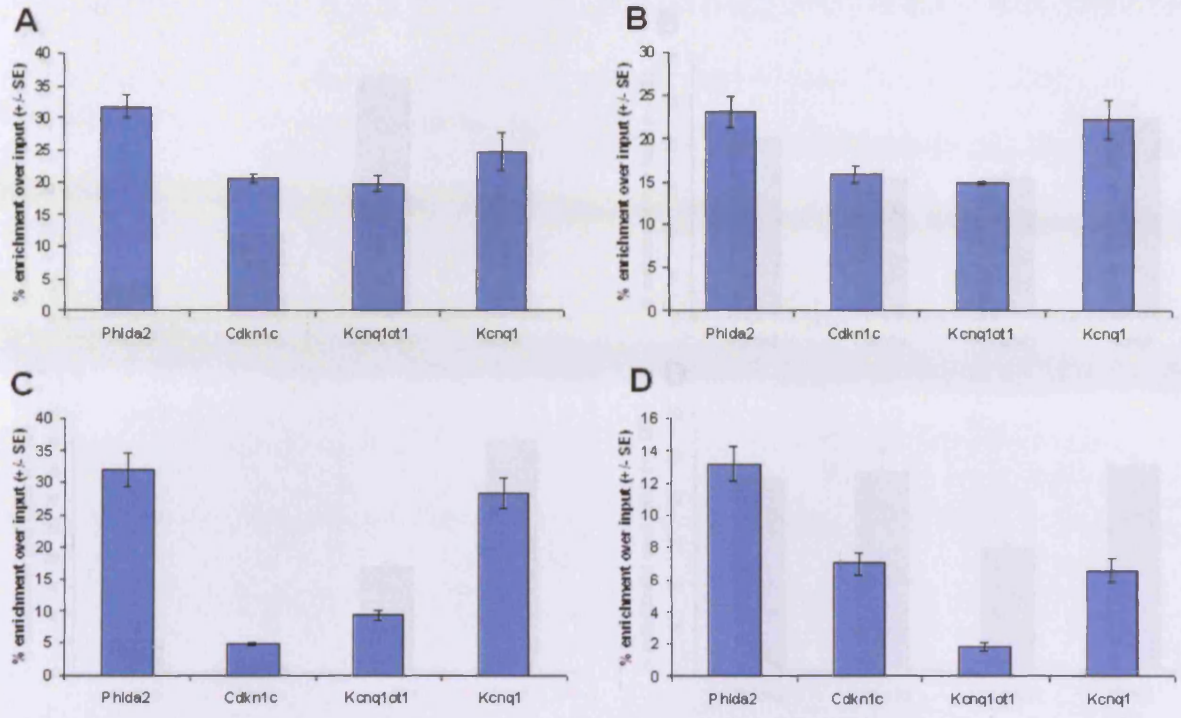


Figure 5.14: Enrichment of repressive histone modifications across the IC2 domain in undifferentiated (day 0) and differentiated (day 5) Sv6.1 EG cells.

Enrichment of repressive histone modifications, H3K9me3 (**A** and **B**) and H3K27me3 (**C** and **D**), across the IC2 domain in undifferentiated (**A** and **C**) and differentiated EG cells (**B** and **D**). The percentage enrichment of each target over input levels is shown for each promoter (calculated as described in figure 5.10). ChIP experiments were carried out using the Orange ChIP kit (Diagenode) using the optimised sonication and fixing conditions. *Kcnq1ot1/Lit1* promoter is also the location of the ICR.

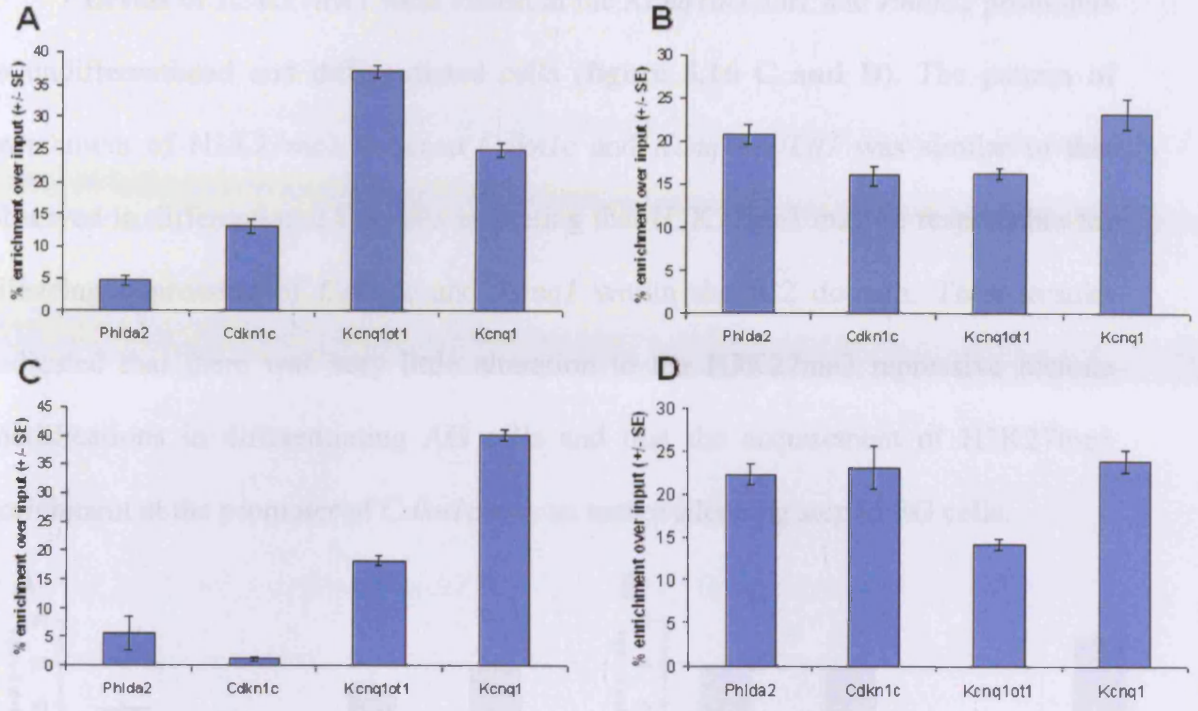


Figure 5.15: Enrichment of repressive histone modifications across the IC2 domain in undifferentiated (day 0) and differentiated (day 5) TMAS21G EG cells.

Enrichment of repressive histone modifications, H3K9me3 (**A** and **B**) and H3K27me3 (**C** and **D**), across the IC2 domain in undifferentiated (**A** and **C**) and differentiated EG cells (**B** and **D**). The percentage enrichment of each target over input levels is shown for each promoter (calculated as described in figure 5.10). ChIP experiments were carried out using the Orange ChIP kit (Diagenode) using the optimised sonication and fixing conditions. *Kcnq1ot1/Lit1* promoter is also the location of the ICR.

The repressive histone modifications were also examined in the AKR1 AG cells. In this cell line we observed constant levels of H3K9me3 across the domain in undifferentiated cells (**figure 5.16 A**). This is similar to what was observed in the differentiated EG cells, hinting that this histone modification was absent in both forms of 'silent' chromatin. However, after differentiation this modification was depleted at *Kcnq1ot1/Lit1* which could indicate that this modification is responsible for silencing in differentiated ES cells (**figure 5.16 B**).

Levels of H3K27me3 were lowest at the *Kcnq1ot1/Lit1* and *Phlda2* promoters in undifferentiated and differentiated cells (**figure 5.16 C and D**). The pattern of enrichment of H3K27me3 between *Cdkn1c* and *Kcnq1ot1/Lit1* was similar to that observed in differentiated EG cells indicating that H3K27me3 may be responsible for silencing expression of *Cdkn1c* and *Kcnq1* within the IC2 domain. These results indicated that there was very little alteration to the H3K27me3 repressive histone modifications in differentiating AG cells and that the acquirement of H3K27me3 enrichment at the promoter of *Cdkn1c* was an active silencing step in EG cells.

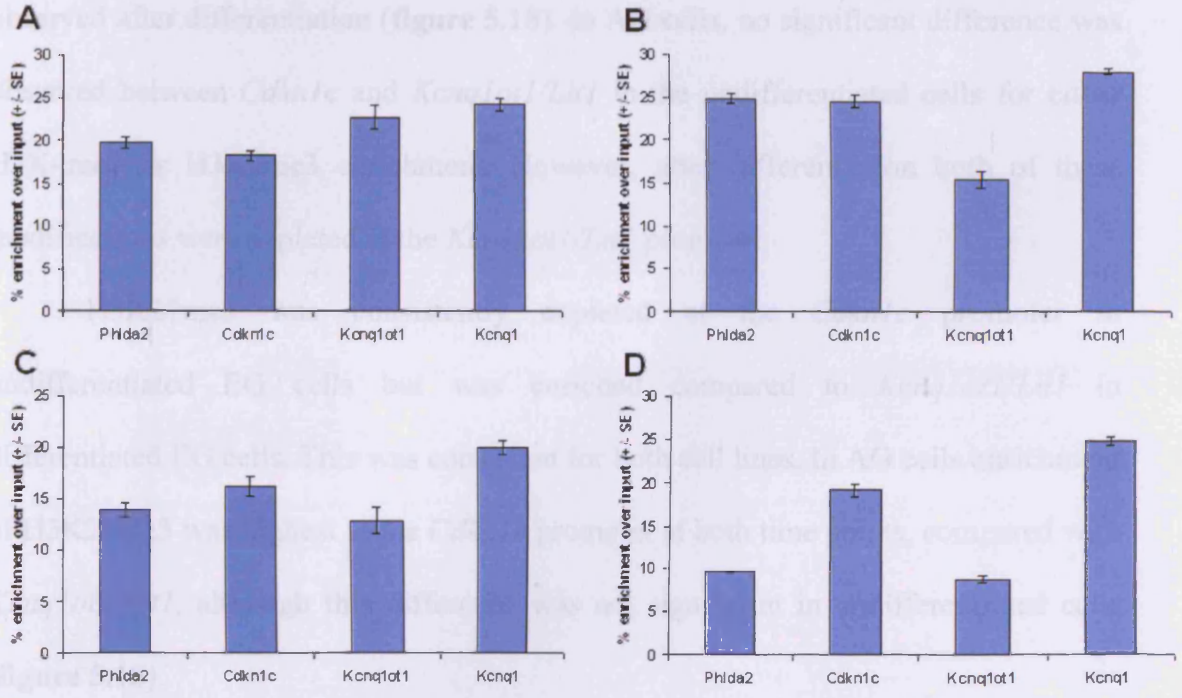


Figure 5.16: Enrichment of repressive histone modifications across the IC2 domain in undifferentiated (day 0) and differentiated (day 5) AKR1 AG cells

Enrichment of repressive histone modifications, H3K9me3 (**A** and **B**) and H3K27me3 (**C** and **D**), across the IC2 domain in undifferentiated (**A** and **C**) and differentiated EG cells (**B** and **D**). The percentage enrichment of each target over input levels is shown for each promoter (calculated as described in figure 5.10). ChIP experiments were carried out using the Orange ChIP kit (Diagenode) using the optimised sonication and fixing conditions. *Kcnq1ot1/Lit1* promoter is also the location of the ICR.

In summary, we found that the active histone marks H3K9Ac and H3K4me3 were consistently enriched at the *Kcnq1ot1/Lit1* promoter compared to *Cdkn1c* in all three cell lines (**figure 5.17**). No change in the status of these marks was observed after differentiation. H3K4me2 enrichment varied between EG cell lines in undifferentiated cells. However, after differentiation there was no significant difference in enrichment at the *Cdkn1c* and *Kcnq1ot1/Lit1* promoters in either of the EG cell lines. The same result was observed for the H3K9me3 levels where there were differences between cell lines in undifferentiated cells but no differences were observed after differentiation (**figure 5.18**). In AG cells, no significant difference was observed between *Cdkn1c* and *Kcnq1ot1/Lit1* in the undifferentiated cells for either H3K4me2 or H3K9me3 enrichment. However, after differentiation both of these modifications were depleted at the *Kcnq1ot1/Lit1* promoter.

H3K27me3 was consistently depleted at the *Cdkn1c* promoter in undifferentiated EG cells but was enriched compared to *Kcnq1ot1/Lit1* in differentiated EG cells. This was consistent for both cell lines. In AG cells enrichment of H3K27me3 was highest at the *Cdkn1c* promoter at both time points, compared with *Kcnq1ot1/Lit1*, although this difference was not significant in undifferentiated cells (**figure 5.18**).

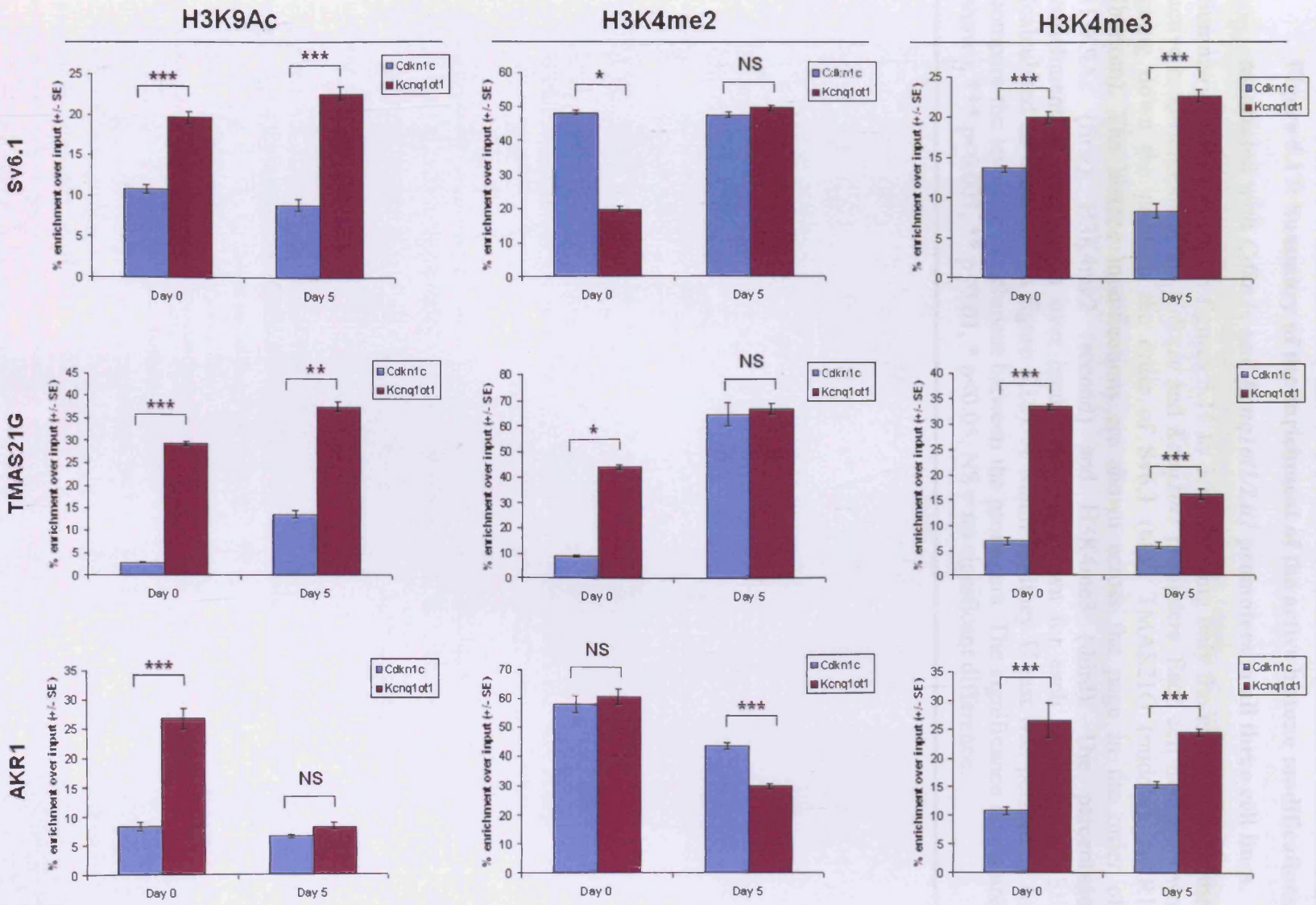

H3K9Ac
H3K4me2
H3K4me3
Sv6.1
TMAS21G
AKR1
% enrichment over input (+/- SE)
Cdkn1c
Kcnq1ot1
Day 0
Day 5
***
**
*
NS

Figure 5.17: Summary of the enrichment of the active histone modifications associated with *Cdkn1c* and *Kcnq1ot1/Lit1* promoters in all three cell lines.

Summary of the data from figures 5.11 to 5.16 showing only the enrichment of the active modifications at the *Cdkn1c* and *Kcnq1ot1* promoters. Each cell line is shown going down the page in the order of Sv6.1 (top), TMAS21G (middle), AKR1 (bottom). The histone modifications are shown across the page in the order of H3K9Ac (first), H3K4me2 (second) and H3K4me3 (third). The percentage enrichment of each target over input levels is shown for each promoter +/- SE (calculated as described in figure 5.10). A Mann-Whitney U test was performed to compare the levels of enrichment between the promoters. The significance levels are shown. *** $p<0.001$, ** $p<0.01$, * $p<0.05$, NS = no significant difference.

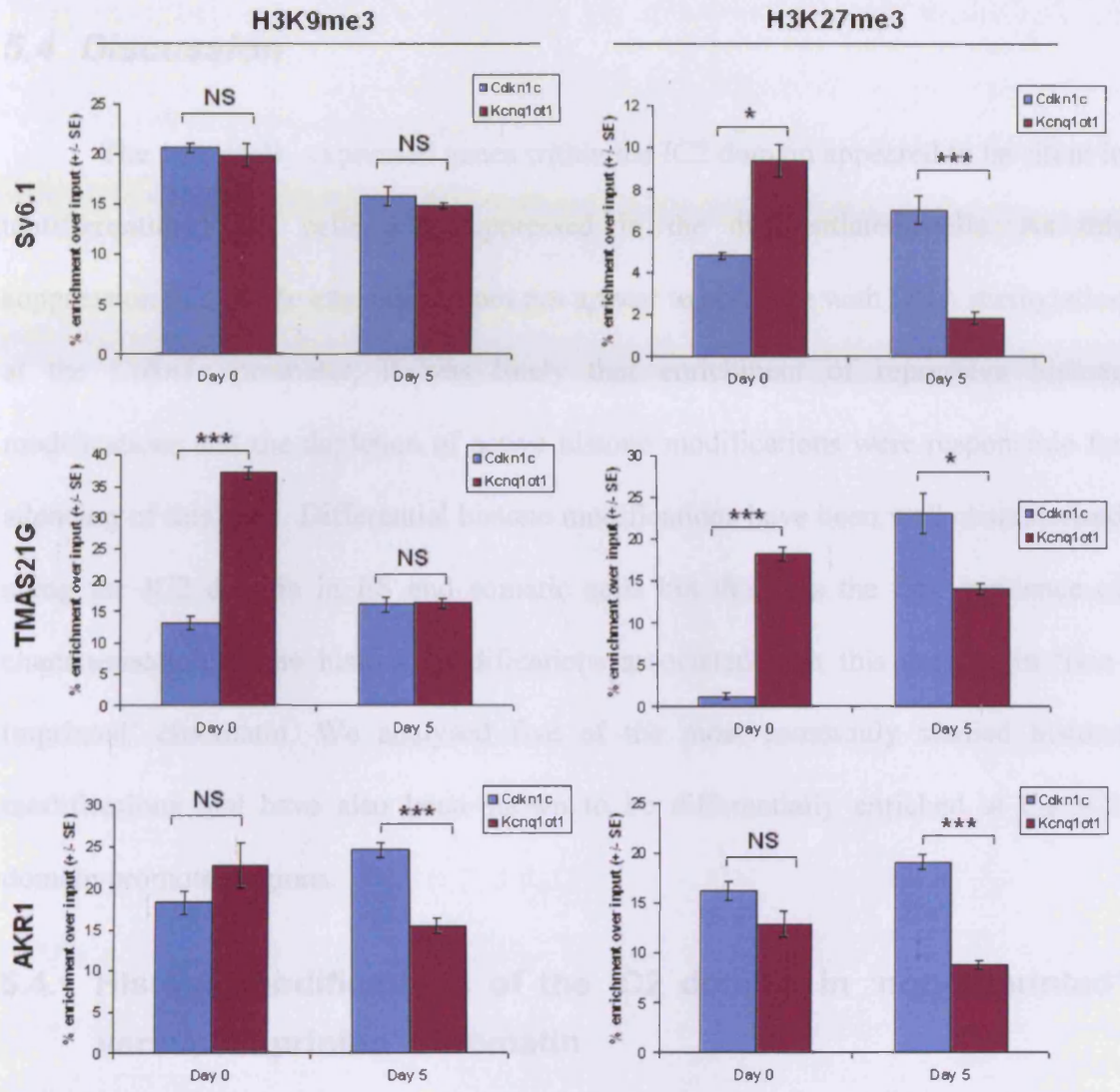


Figure 5.18: Summary of the enrichment of the repressive histone modifications associated with *Cdkn1c* and *Kcnq1ot1/Lit1* promoters in all three cell lines.

Summary of the data from figures 5.11 to 5.16 showing only the enrichment of the repressive modifications at the *Cdkn1c* and *Kcnq1ot1* promoters. Each cell line is shown going down the page in the order of Sv6.1 (top), TMAS21G (middle), AKR1 (bottom). The histone modifications are shown across the page in the order of H3K9me3 (first) and H3K27me3 (second). The percentage enrichment of each target over input levels is shown for each promoter +/- SE (calculated as described in figure 5.10). A Mann-Whitney U test was performed to compare the levels of enrichment between the promoters. The significance levels are shown. *** $p<0.001$, ** $p<0.01$, * $p<0.05$, NS = no significant difference.

5.4 Discussion

The maternally expressed genes within the IC2 domain appeared to be silent in undifferentiated EG cells and suppressed in the differentiated cells. As this suppression of *Cdkn1c* expression does not appear to correlate with DNA methylation at the *Cdkn1c* promoter, it was likely that enrichment of repressive histone modifications and the depletion of active histone modifications were responsible for silencing of this gene. Differential histone modifications have been well characterised along the IC2 domain in ES and somatic cells but this was the first incidence of characterisation of the histone modifications associated with this domain in 'non-imprinted' chromatin. We analysed five of the most commonly studied histone modifications that have also been shown to be differentially enriched at the IC2 domain promoter regions.

5.4.1 Histone modifications of the IC2 domain in 'non-imprinted' versus 'imprinted' chromatin

Our analysis of the IC2 domain indicated that there was an enrichment of active modifications on the *Kcnq1ot1/Lit1* promoter, compared with the other promoters in both EG cell lines. The enrichment of these active modifications correlated with the expression status of this gene in the EG cells. In the AG cells, the *Kcnq1ot1/Lit1* promoter also showed enrichment of the active histone modifications indicating that in 'non-imprinted' chromatin, histone modifications associated with gene transcription are established at the KvDMR1 and are maintained through the male germ line.

Cdkn1c appears to be depleted for all active histone modifications in the undifferentiated EG and AG cells which correlated with the silencing of this gene.

These modifications were also absent in the differentiated cells. Surprisingly, in undifferentiated cells, the *Cdkn1c* promoter also appeared to be depleted for H3K27me3. This was in contrast to the levels observed at the promoter regions of *Phlda2* and *Kcnq1*. However, after differentiation there was an enrichment of H3K27me3 at this promoter in both EG cell lines that then matched the pattern of enrichment observed in both the undifferentiated and differentiated AG cells. This data indicated that H3K27me3 was actively acquired at the *Cdkn1c* promoter during differentiation of the EG cells. This is potentially interesting as H3K27me3 is a target of the histone methyltransferase complex PRC2. This complex contains the protein Eed which is required for the stability of the complex. Knock out studies of Eed have shown that the PRC2 complex is required to methylate H3K27me3 specifically at a selection of imprinted genes including *Cdkn1c, Gtl2* and *Grb10* (Mager *et al.* 2003). In the Mager *et al.* study, loss of Eed resulted in biallelic expression of these genes with no loss of methylation at the DMRs, however the expression of *Cdkn1c* was not quantified and does, in fact, appear to be a partial loss of silencing (Mager *et al.* 2003). Several other imprinted genes examined in the Mager *et al.* study maintained monoallelic expression, including *Phlda2*, indicating that the PRC2 complex acts to silence expression of specific imprinted genes. Other studies have also shown that in the placenta, Eed is enriched at the paternally derived *Kcnq1* and *Cdkn1c* promoters indicating that it may be required to maintain silencing of these genes in somatic cells (Umlauf *et al.* 2004). There is also evidence to suggest that the methylation of H3K27me3 by the PRC2 complex facilitates binding of polycomb (pc), a component of the PRC1 complex, acting to further condense the chromatin and inhibit gene expression (Cao *et al.* 2002).

Phlda2 and *Kcnq1* seemed to be regulated in slightly different ways. The *Phlda2* gene appeared to acquire enrichment of some active histone modifications without any increase in gene expression whereas the *Kcnq1* gene lost enrichment of active histone modifications upon differentiation although this gene was not expressed in our undifferentiated EG cells. In both cases, levels of H3K27me3 appeared to remain enriched in both undifferentiated and differentiated cells indicating that this modification is most likely responsible for suppressing transcription. The enrichment of both active and repressive histone modifications has been noted at the promoter regions of developmentally important genes in stem cells where H3K4me3 and H3K27me3 have been shown to be co-ordinately associated with the silent promoters of genes in undifferentiated stem cells (Bernstein *et al.* 2006). After differentiating the cells, one or another of the histone marks is lost and leads to either increased gene expression or the maintenance of silencing. This occurred for the *Kcnq1* gene in our ES and AG cells with the loss of active histone modifications in the differentiated cells and maintenance of silencing, probably due to the presence of H3K27me3. For *Phlda2,* however, we actually see these 'bivalent' modifications only in the differentiated cells. Expression of *Kcnq1,* although ubiquitously imprinted, is restricted to certain tissues in the embryo including the heart and kidney. The bivalent chromatin modifications may reflect its status as a tissue-specific gene. Even after differentiation, the promoter of this gene still retained high levels of both repressive and active histone modifications. A switch to silencing or active marks may not necessarily occur until later on in development.

Although we did not observe any significant changes in H3K9me3 patterns at any of the promoters in our EG cell lines, paternal silencing of *Kcnq1* has been shown to correlate with H3K9 methylation in the placenta. The SET-domain protein G9a is a

histone methyltransferase involved in H3K9 and K27 methylation at euchromatic regions (Ikegami *et al.* 2007; Wagschal *et al.* 2008). G9a deficiency results in partial CpG hypomethylation at several loci in ES cells, however G9a deficient mice show normal DNA methylation at DMRs and imprinted gene expression is unchanged (Ikegami *et al.* 2007; Wagschal *et al.* 2008). There is some evidence to suggest that placental-specific imprinting is impaired in the absence of G9a, correlating with reduced levels of H3K9 di- and tri-methylation. There seems to be a moderate effect on trophoblastic differentiation which could be as a result of the relaxation of imprinting of some placental-specfic genes within the IC2 domain such as *Osbpl5, Ascl2* and *Cd81* (Wagschal *et al.* 2008). Consequently, G9a has also been shown to be enriched on *Ascl2*. This data implies that H3K9 methylation is part of a multi-layered silencing mechanism. Loss of one of the components could result in partial derepression of the placental-specific genes (Wagschal *et al.* 2008). It has also been shown that G9a specifically binds Dnmt1 so may be involved in maintenance of silencing at tissue specific methylated regions (Estève *et al.* 2006). We did not observe any differences in H3K9me3 enrichment between *Cdkn1c* and *Kcnq1ot1/Lit1* in any of the EG cell lines indicating that this modification is probably not contributing to the differences in expression between these two genes. We do see enrichment at *Cdkn1c* in the differentiated AG cells which may indicate that this modification is an extra silencing step in differentiated ES cells.

5.4.2 Comparison of the histone modifications associated with the IC2 domain in ES, EG and AG cells

Our analysis of the gene expression levels showed that *Cdkn1c* was repressed in undifferentiated ES cells. This data correlates with the histone modifications associated with the promoter in ES cells observed by Umlauf *et al.* (Umlauf *et al.*

2004). Umlauf *et al.* did not see any alleleic differences in *Cdkn1c* histone methylation patterns yet, after differentiation, the *Cdkn1c* promoter became enriched for H3K4me2 on the maternal allele and enriched for H3K27me3 on the paternal allele (Umlauf *et al.* 2004). Our results for the AG cells indicated that the H3K27me3 mark is present in undifferentiated cells and is then maintained upon differentiation. It could be that the maternal allele is also marked with H3K27me3 in undifferentiated ES cells and that the allelic specific differences observed after differentiation is due to loss of H3K27me3 from the maternal allele and not due to enrichment on the paternal allele. Although we also observed an increase in H3K4me2 after differentiation of the AG cells so, rather than the different parental alleles gaining enrichment of either H3K27me3 or H3K4me2, it could be that the allelic specific differences observed in the ES cells is due to loss of these modifications. This is consistant with the observation that, in ES cells, 'bivalent' chromatin marks are associated with developmentally important genes (Bernstein *et al.* 2006). If *Cdkn1c* is marked by both H3K4me2 and H3K27me3 then this gene would be repressed and would also account for the lack of allele specific differences observed by Umlauf *et al.* in undifferentiated ES cells. After differentiation, H3K4me2 would then be lost from the paternal allele and H3K27me3 lost from the maternal allele to give the differential histone modifications seen in the differentiated cells. This would be consistant with our results. Umlauf *et al.* also showed that *Cdkn1c* is expressed in both undifferentiated and differentiated ES cells although this was not done quantitively (Umlauf *et al.* 2004). In our analysis we have showed that, although *Cdkn1c* expression could be detected in undifferentiated ES cells at very low levels, this gene was only strongly expressed after differentiating the stem cells.

Phlda2 and *Kcnq1* are also not expressed in undifferentiated ES cells. Umlauf *et al.* did not analyse the enrichment of histone modifications at the *Phlda2* or *Kcnq1* promoters in ES cells. They did examine the enrichment in the placenta where there appeared to be enrichment of H3K9me2 and H3K27me3 on the repressed paternal alleles of both genes. The maternal allele appeared to be enriched for H3K4me2 and H3K9Ac. In the AG cells, the *Phlda2* promoter appeared enriched for the active histone modifications which did not correlate with expression of this gene. There was also no enrichment of H3K27me3 in the differentiated cells indicating that other modifications may be involved in silencing *Phlda2* in these cells.

Our data showed that *Kcnq1ot1/Lit1* was expressed at levels two-fold higher in EG cells than in ES cells. This gene is expressed from a promoter found within the KvDMR1 and depends on the methylation status of the DMR. In ES cells, the KvDMR1 is enriched for H3K9Ac and H3K4me2 on the paternal allele and for H3K9me2 and H3K27me3 on the maternal allele (Umlauf *et al.* 2004). Enrichment of H3K9Ac, H3K4me2 and H3K4me3 has been observed in somatic cells at actively expressed genes and are also found on the unmethylated paternal KvDMR1 (Delaval *et al.* 2007). This data is consistant with our results as we also observed enrichment of H3K9Ac and H3K4me3 and a depletion of H3K27me3. Our results are consistant with that observed on the paternal allele in ES cells indicating that the paternal repressive histone modifications are the default pattern at the IC2 domain genes.

5.4.3 Conclusion

The histone modifications associated with the IC2 domain genes are consistant with their transcriptional status in EG cells. Active histone modifications are associated with the actively transcribed *Kcnq1ot1/Lit1* gene whereas the silenced

Kcnq1 and *Phlda2* are marked with the repressive histone modification H3K27me3. *Kcnq1* appears to display bivalent histone modifications consistant with its role as a tissue specific imprinted gene. *Cdkn1c* appeared to acquire an additional repressive modification, H3K27me3, in differentiated EG cells. It is possible that this is mediated by the polycomb repressive complex PRC2 as knockout studies have shown that *Cdkn1c* is a target of this complex.

5.4.4 Summary of findings

- The active *Kcnq1ot1/Lit1* promoter in EG and AG cells is enriched with active histone modifications.
- The repressed genes *Phlda2* and *Kcnq1* are associated with both repressive and active histone modifications.
- *Cdkn1c* shows enrichment of H3K27me3 only in differentiated EG cells mimicking the enrichment seen at this locus in both undifferentiated and differentiated AG cells.
- Active acquirement of H3K27me3 suggests the suppression of *Cdkn1c* expression involves additional modifications which could be mediated by the Eed containing complex PRC2.

Chapter 6:

General Discussion and Summary of Findings

6 General discussion and summary of findings

The work in this thesis initially follows on from the finding by Tada *et al.*, that the *Cdkn1c* locus was silent in *in vivo* differentiated EG cells and acquired *de novo* DNA methylation (Tada *et al.* 1998). Our aim was to replicate these results *in vitro* and examine the epigenetic events that led to the silencing and methylation of *Cdkn1c*.

6.1.1 *Cdkn1c* expression in undifferentiated and differentiated EG cells

In our study, we found that the levels of *Cdkn1c* expression in undifferentiated EG cells and ES cells were very low (late CT). When we differentiated ES cells *in vitro,* we detected high levels of *Cdkn1c* expression from day 7 onwards. This was consistent with previous data demonstrating that this gene is only expressed after cells start to differentiate. In our differentiated EG cells, *Cdkn1c* expression remained at a very low level even after several days of differentiation. This suggested that, as in *in vivo* differentiated EG cells, the *Cdkn1c* locus was suppressed in the absence of a germ line imprint. The *Kcnq1ot1/Lit1* transcript was detectable in both ES and EG cells. In EG cells the levels of expression was approximately two-fold greater than in ES cells which suggested that the transcript was expressed from both alleles. However, we were not able to demonstrate this directly as no polymorphisms were present in the EG cell lines.

At earlier timepoints, the expression of *Kcnq1ot1/Lit1* was dynamic in both ES and EG cells which could indicate a role in the silencing process. Silencing must have occurred at a very early time point as we were not able to detect *Cdkn1c* or *Phlda2*

expression even at day one of differentiation when we observed a modest increase in expression in ES cells.

6.1.2 Methylation of the *Cdkn1c* promoter DMR

One of our aims was to determine if *de novo* DNA methylation of the *Cdkn1c* promoter occurred when the EG cells were differentiated *in vitro.* Methylation of the *Cdkn1c* promoter could be involved in the suppression of *Cdkn1c* expression that we observed in the differentiated EG cells. Methylation of the paternal *Cdkn1c* is present in somatic cells (Bhogal *et al.* 2004; John *et al.* 1999) and it may function as an extra layer of silencing to prevent leaky expression from the paternal allele. However, we did not detect any significant DNA methylation of the *Cdkn1c* promoter in either undifferentiated or differentiated EG cells. Unexpectedly, we also did not observe *Cdkn1c* methylation in the undifferentiated and differentiated ES or AG cells. Since *Cdkn1c* does appear to be suppressed in differentiated EG cells, this suggests that silencing of *Cdkn1c* does not require DNA methylation. This is interesting because in humans there is no methylation of *Cdkn1c* yet the gene is still imprinted (Chung *et al.* 1996; John *et al.* 1999; Onyango *et al.* 2000). There is also imprinted expression in the early embryo prior to the acquisition of DNA methylation (Bhogal *et al.* 2004).

6.1.3 H3K27me3 enrichment at the *Cdkn1c* promoter in differentiated EG cells

Our expression data suggested that silencing of *Cdkn1c* was established in differentiated EG cells from a very early time point. We wanted to determine if this silencing of *Cdkn1c* occurred in a stepwise fashion. For example, silencing could have occurred by the acquisition of repressive histone modifications during differentiation of the EG cells or it could have been that the silent epigenetic state of the locus was

present in undifferentiated cells. We compared 'non-imprinted' chromatin to 'imprinted' chromatin in undifferentiated and differentiated cells to determine if there were differences in histone modifications associated with *Cdkn1c* at the two time points. Our data showed that *Kcnq1ot1/Lit1* promoter was associated with active histone marks whereas these marks were depleted at the *Cdkn1c* promoter in both EG and AG cells. No changes in these 'active' histone modifications were observed upon differentiation. In undifferentiated AG cells and differentiated AG and EG cells, the *Cdkn1c* promoter was shown to be enriched for the repressive H3K27me3 modification yet this mark was absent in undifferentiated EG cells indicating that H3K27me3 at the *Cdkn1c* promoter was specifically acquired during differentiation of the EG cells *in vitro*. This might involve the action of the PRC2 complex which has been shown to specifically methylate H3K27 at the *Cdkn1c* locus and several other imprinted loci (Mager *et al.* 2003). This was not observed in the AG cells where this mark was present in both undifferentiated and differentiated cells. Therefore, despite the similarities between the paternal allele and the imprint erased allele (i.e. no methylation at KvDMR1, expression of *Kcnq1ot1/Lit1* and no expression of *Cdkn1c, Phlda2, Slc22a18,* and *Kcnq1*) the epigenetic status of the chromatin in the undifferentiated cells was not the same.

H3K27me3 has been shown to be enriched on the paternal allele of *Grb10.* Yamasaki-Ishisaki *et al.* showed that H3K27me3 enrichment was localised to the promoter of the major-type transcript on the paternal allele (Yamasaki-Ishizaki *et al.* 2007). As with *Cdkn1c*, H3K27me3 is involved in silencing of *Grb10* as loss of the PRC2 complex component Eed results in biallelic expression of *Grb10* (Mager *et al.* 2003). However, unlike the *Cdkn1c* situation, the CpG island overlapping the *Grb10* major-type promoter remains unmethylated in somatic cells (Arnaud *et al.* 2003;

Hikichi *et al.* 2003; Yamasaki-Ishizaki *et al.* 2007). It may be that methylation of this CpG island is not needed to suppress paternal *Grb10,* as seen with *Cdkn1c*, because the germ line DMR, in this case, functions as an insulator element. The CTCF protein binds the unmethylated DMR and insulates *Grb10* from its downstream enhancers and is analogous to the IC1 domain model of imprinting. The fact that loss of Eed results in biallelic expression of *Grb10* indicates that H3K27me3 is important for silencing *Grb10* expression. It would be interesting to determine the sequence of epigenetic events that lead to *Grb10* silencing in the germ line particularly if, like *Cdkn1c,* H3K27me3 is specifically enriched on the *Grb10* promoter in differentiated EG cells. However, it could be that CTCF binding is the main factor responsible for the suppression of *Grb10* expression in differentiated EG cells. As CTCF binding sites have also been found in *Cdkn1c* and KvDMR1 it may be that this protein also has some role in establishing silencing at *Cdkn1c* (Du *et al.* 2003; Fitzpatrick *et al.* 2007).

6.1.4 Factors establishing silencing of *Cdkn1c* in differentiated EG cells

We did not see any DNA methylation of the *Cdkn1c* promoter in the undifferentiated ES cells, which is consistent with this being a late event in embryogenesis i.e. the ES cells are derived from embryos (4.5 dpc) before DNA methylation has taken place (Bhogal *et al.* 2004). However, it might be expected that DNA methylation of the promoter would occur in the differentiated ES and AG cells as differentiating the cells as embryoid bodies mimics the events that occur in the pre- and post-implantation embryo (Doetschman *et al.* 1985). It was possible that a key *de novo* methyltransferase or some other factor required for establishing methylation of the *Cdkn1c* promoter was not expressed in the differentiated stem cells or that the cells did not reach the appropriate differentiated state in which these factors would be

expressed. DNA methylation of the *Cdkn1c* DMR is acquired during post-implantation development from 7.5 dpc onwards and likely involves transfactors such as Lyphoid specific helicase (Lsh) as well as the *de novo* DNA methyltransferases (Bhogal *et al.* 2004; Fan *et al.* 2005; Zhu *et al.* 2006). It has been shown that acquisition of DNA methylation in the mouse depends on the presence of Lsh and has been shown to be important for the formation of normal heterochromatin in the mouse (Dennis *et al.* 2001; Geiman *et al.* 1998; Yan *et al.* 2003). Targeted deletion of *Lsh* results in a substantial loss of DNA methylation in the mouse, particularly at repetitive elements and single copy sequences (Dennis *et al.* 2001). *Cdkn1c* expression is specifically affected by the loss of Lsh and there is hypomethylation of the paternal somatic DMR (Fan *et al.* 2005). This suggests that DNA methylation is required for silencing *Cdkn1c* but perhaps at a later stage of development. No other imprinted genes have been shown to be affected. Lsh is known to associate with the *de novo* DNA methyltransferases Dnmt3a and Dnmt3b but not Dnmt1 indicating that Lsh has a role in establishing methylation patterns but not in their maintenance (De La Fuente *et al.* 2006). In addition to Eed, other factors may be required to establish silencing before DNA methylation occurs, such as the deposition of the histone variant macroH2A or the binding of non-histone proteins, such as HP1, which are known to be deposited at silent imprinting domains and DMRs (Choo *et al.* 2006; Nesterova *et al.* 2002; Smallwood *et al.* 2007).

Many of the epigenetic processes that are thought to be involved in allele-specific silencing are also involved in imprinted X inactivation in female mammals (Okamoto *et al.* 2004). This includes transcription of the non-coding RNAs *Xist* and *Tsix* and the recruitment of DNA methylation. Repressive histone modifications are also recruited to the silent X chromosome by chromatin remodelling complexes such

as PRC2 (Rougeulle *et al.* 2004). This remodelling of chromatin may be a stepwise process with certain modifications occurring first providing a base for extra layers of silencing, such as ubiquitination of H2A which occurs on the silent X chromosome and correlates with recruitment of the PRC1 complex (de Napoles *et al.* 2004; Nesterova *et al.* 2002). Future work could involve examining the expression of proteins that are potentially involved in establishing silencing of *Cdkn1c* in differentiating EG and ES cells to determine if a key element is absent. ChIP could also be performed to determine if these trans-factors localise to the IC2 domain during differentiation of the EG cells, such as Eed and Lsh. Sequential ChIP (SeqChIP) could then be used to address whether two proteins occupy the same genomic region and could also determine whether this co-occupency is dependent on a particular protein binding (Geisberg and Struhl 2004). This would allow us to determine if there is a particular sequence of binding of transfactors to *Cdkn1c* during differentiation that leads to the formation of heterochromatin and methylation of *Cdkn1c.* It would also be interesting to identify other somatic DMRs that are absent in ES cells and only become *de novo* methylated in the developing embryo, such as at *Nesp55.*

6.1.5 Other histone modifications involved in silencing of *Cdkn1c*

We only examined six of the most common histone modifications associated with the IC2 domain genes and, in fact, there are numerous others that have been shown to be involved in establishing allele specific differences in chromatin between the two parental alleles - such as H3K20 methylation and H3K9 mono- and di-methylation. In some cases of BWS there is down regulation of *CDKN1C* expression yet there is normal methylation at KvDMR1 and the *CDKN1C* promoter remains unmethylated. In these cases the loss of *CDKN1C* expression has been shown to be

associated with the loss of the active histone mark H3K4me2 and enrichment of markers of heterochromatin, H3K9me2 and HP1γ (Diaz-Meyer *et al.* 2005). To get a full understanding of the histone modifications that are responsible for silencing the IC2 domain genes it may be necessary to perform ChIP on the EG cells for a wider range of histone modifications. A distinct advantage would be to use 'ChIP on chip' where ChIP is combined with microarray analysis and could be used to analyse the changes in histone modifications across multiple sites within the IC2 domain, increasing the resolution of our data (Ren *et al.* 2000) ['ChIP on chip' reviewed in (Buck and Lieb 2004)]. We could also use our system to examine earlier time points to determine at what time point H3K27me3 appears at the *Cdkn1c* promoter. Our expression data indicated that *Cdkn1c* was already silenced by day one of differentiation so it is possible that acquisition of H3K27me3 occurred very early on during differentiation. There could also be other repressive histone marks that we haven't yet examined that could be acquired at this time point before H3K27me3. Our analysis looked at the relative levels of the histone modifications within the IC2 domain comparing the enrichment of modifications at an actively expressed gene (*Kcnq1ot1/Lit1*) and a suppressed gene (*Cdkn1c*). It would be interesting to compare this enrichment to control genes that are known to be silent in both undifferentiated and differentiated EG cells and genes that are expressed at both time points. We could then determine whether the enrichment that we observe at our imprinted genes is comparable to that of a fully silenced gene. There are several other imprinted genes within this cluster, although the majority are tissue specific. With ChIP microarray analysis it might be possible to look at the promoter regions of all the imprinted genes within this cluster to determine if the entire domain is silent from the very beginning

or if there is a spread of repressive chromatin occurring in the absence of methylation of KvDMR1.

6.1.6 Role of the non-coding RNA in *Cdkn1c* silencing

Kcnq1ot1/Lit1 appeared to be expressed throughout the differentiation of the EG and ES cells, although dynamic changes in expression were observed at early time points. Recent evidence suggests that transcription of the full length non-coding RNA is required for silencing of the paternal allele in somatic cells indicating that, in our EG cells, *Kcnq1ot1/Lit1* may somehow be responsible for the active silencing of *Cdkn1c* (Mancini-DiNardo *et al.* 2006; Shin *et al.* 2008). As the different length truncations of *Kcnq1ot1/Lit1* appeared to have different effects on the silencing of *Cdkn1c,* it may be that the transcription of this non-coding RNA is required to recruit the PRC2 complex to establish H3K27me3 at distant loci. This then may be the initiation mark for the recruitment of other silencing mechanisms such as the PRC1 complex, HP1γ, Lsh and ultimately DNA methylation to the silent promoter (**figure 6.1**). Shin *et al.* showed that a 2.6 kb truncation of *Kcnq1ot1/Lit1* resulted in loss of methylation at *Cdkn1c* but imprinting was still maintained in some tissues. It would be interesting to see if we still get methylation of H3K27 in the absence of *Kcnq1ot1/Lit1* in the differentiated EG cells or if this repressive histone mark is lost in AG cells when *Kcnq1ot1/Lit1* is no longer expressed. This would allow us to determine if this non-coding transcript and H3K27me3 were required for establishing silencing or maintaining it. It would also be interesting to see what the effect of the different truncation lengths had on the chromatin structure around *Cdkn1c.* How the non-coding RNA mediates silencing of the IC2 domain genes is still unclear. It may involve altering the local chromatin environment by recruiting complexes such as the

PRC2 complex. Future work could include using RNA interference (RNAi) technology to knock down components of the PRC2 complex in EG cells to confirm that this complex is involved in establishing H3K27me3 at *Cdkn1c*. Preliminary experiments, not described in this thesis, were performed with this intention. This included cloning *Eed* oligos to create short hairpin RNA (shRNA) constructs within pSUPER-RED in order to knock down *Eed* expression in EG and ES cells. This technology could also be used to identify other transfactors that may be involved in silencing and heterochromatin formation. Knockdown of macroH2A1 by RNAi has already been shown to reduce expression of some of the *Peg3* domain genes without altering DNA methylation at the DMR so is likely involved in establishing heterochromatin at some imprinted domains (Choo *et al.* 2007). Alternatively, expressing factors such as Lsh or the DNA methyltransferases (if they are found not to be expressed in the ES or EG cells) could allow us to determine if it is the absence of expression of these proteins that is responsible for the hypomethylation of the *Cdkn1c* DMR or if it is because the chromatin structure is not permissive for *de novo* methylation until later in development.

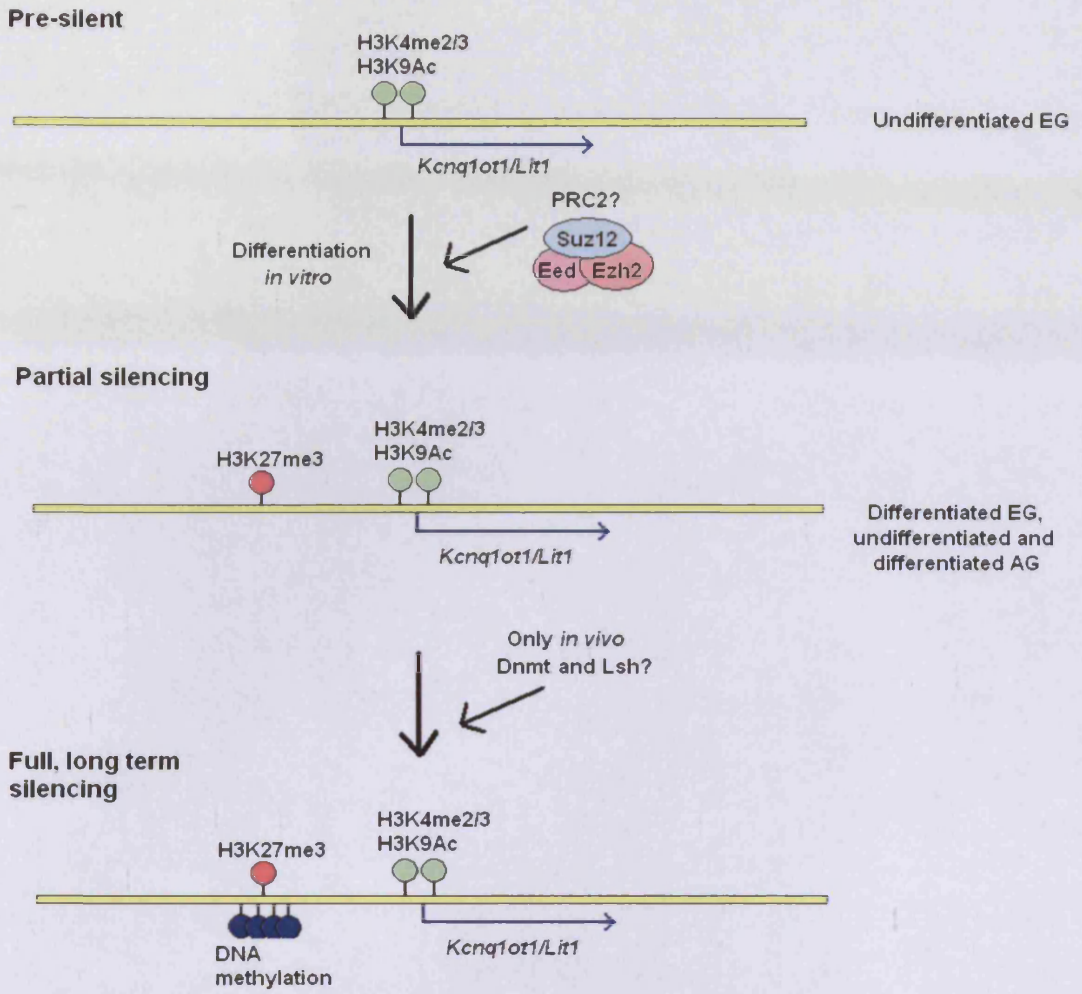


Figure 6.1: Model of the epigenetic events leading to silencing of *Cdkn1c*.

EG cells differentiated *in vitro* showed enrichment of the repressive histone modification H3K27me3 which was absent in undifferentiated EG cells. Recruitment was likely mediated by the action of the PRC2 complex. The *Cdkn1c* promoter was hypo-methylated in differentiated EG cells so DNA methylation is probably a late epigenetic event required for full long term silencing of paternal *Cdkn1c* in the embryo.

6.1.7 Conclusion

In somatic cells the maternal and paternal alleles of *Cdkn1c* are marked with different epigenetic modifications that result in monoallelic expression from the maternal allele only. These epigenetic marks are established in the germ line. Our results indicated that paternal silencing of *Cdkn1c* is a default silencing process that occurs in the male germ line and involves the recruitment of the repressive histone mark H3K27me3. In the maternal germ line this default pathway appears to be inhibited by methylation of the germ line DMR, KvDMR1, preventing transcription of *Kcnq1ot1/Lit1*. In BWS, loss of imprinting of *KCNQ1OT1/LIT1* is one of the most common epigenetic mutations and can be caused by hypomethylation of KvDMR1, the maternal methylation mark established in the developing oocyte (Bliek *et al.* 2001; Lee *et al.* 1999). If this methylation mark is not established in the maternal germ line then the IC2 domain undergoes a default pathway resulting in silencing of *CDKN1C*. Loss of *CDKN1C* expression and biallelic expression of *IGF2* are both causes of BWS in humans, either through the action of genetic or epigenetic mutations. Studying the epigenetic marks that lead to alterations in expression of these genes in the germ line could lead to a better understanding of the causes of the disease and to possible therapeutic interventions.

In addition to our study of the IC2 domain genes, we identified several other imprinted genes that also appeared to be suppressed in the absence of imprint. The majority of the germ line DMRs are maternally methylated with only a subset that are methylated in the male germ line. It is likely that methylation of the germ line DMR is the primary imprinting mark that, as in the case of the IC2 domain, switches the domain from the default pathway to a maternal specific pathway. In fact recently a

study has shown that, in the absence of paternal methylation at the IG-DMR, both alleles of the *Gtl2-Dlk1* imprinted domain establish a maternal histone acetylation pattern at the *Gtl2* DMR (Carr *et al.* 2007). Our EG cell model could be used to study other imprinting domains, particularly those that, like *Cdkn1c*, are suppressed in differentiated EG cells, in order to identify the key epigenetic marks required to establish default silencing of these genes.

6.1.8 Summary

Our EG system provides a unique model for studying the early epigenetic events that lead to heritable silencing of imprinted genes. We were able to show that, despite a lack of DNA methylation at the somatic and germ line DMRs, *Cdkn1c* was still suppressed in differentiated EG cells and that this was likely due to the active acquirement of repressive histone modification marks such as H3K27me3. This analysis could be applied to other imprinted domains in order to ascertain the key epigenetic modifications that contribute to monoallelic expression in the embryo.

Reference List

Aapola, U. *et al.* 2002. Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity. *Nucleic Acids Research* 30(16), pp. 3602-3608.

Aapola, U. *et al.* 2004. Epigenetic modifications affect *Dnmt3L* expression. *Biochemical Journal* 380, pp. 705-713.

Aapola, U. *et al.* 2000. Isolation and initial characterization of a novel zinc finger gene, *DNMT3L*, on 21q22.3, related to the cytosine-5-methyltransferase 3 gene family. *Genomics* 65, pp. 293-298.

Abdellah, Z. *et al.* 2004. Finishing the euchromatic sequence of the human genome. *Nature* 431(7011), pp. 931-945.

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

Ainscough, J. F.-X. *et al.* 2000. A skeletal muscle-specific mouse *Igf2* repressor lies 40kb downstream of the gene. *Developmental Biology* 127, pp. 3923-3930.

Ainscough, J. F. *et al.* 1997. Imprinting of *Igf2* and *H19* from a 130 kb YAC transgene. *Development* 124(18), pp. 3621-3632.

Allen, N. D. *et al.* 1994. A functional analysis of imprinting in parthenogenetic embryonic stem cells. *Development* 120, pp. 1473-1482.

Amir, R. E. *et al.* 1999. Rett syndrome is caused by mutations in X-linked *MECP2*, encoding methyl-CpG-binding protein 2. *Nature Genetics* 23(2), pp. 185-188.

Andrews, S. *et al.* 2007. *Cdkn1c (p57Kip2)* is the major regulator of embryonic growth within its imprinted domain on mouse distal chromosome 7. *BMC Developmental Biology* 7(1), pp. 53-67.

Arents, G. *et al.* 1991. The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. *Proceedings of the National Academy of Science USA* 88(22), pp. 10148-10152.

Ariel, M. 2002. A PCR-based method for studying DNA methylation. In: Ward, A. ed. *Genomic Imprinting: Methods and Protocols*. Vol. 181. Totowa, NJ: Humana Press Inc., pp. 205-216.

Arima, T. *et al.* 2006. Loss of the maternal imprint in *Dnmt3L*mat−/− mice leads to a differentiation defect in the extraembryonic tissue. *Developmental Biology* 297, pp. 361-373.

Arima, T. *et al.* 2005. *ZAC, LIT1 (KCNQ1OT1)* and *p57^{KIP2} (CDKN1C)* are in an imprinted gene network that may play a role in Beckwith-Wiedemann syndrome. *Nucleic Acids Research* 33(8), pp. 2650-2660.

Arnaud, P. *et al.* 2006. Stochastic imprinting in the progeny of *Dnmt3L-/-* females. *Human Molecular Genetics* 15(4), pp. 589-598.

Arnaud, P. *et al.* 2003. Conserved methylation imprints in the human and mouse *GRB10* genes with divergent allelic expression suggests differential reading of the same mark. *Human Molecular Genetics* 12(9), pp. 1005-1019.

Barton, S. *et al.* 1984. Role of paternal and maternal genomes in mouse development. *Nature* 311, pp. 374-376.

Beatty, L. G. *et al.* 2006. Detailed analysis of the methylation patterns of the KvDMR1 imprinting control region of human chromosome 11. *Genomics* 87, pp. 46-56.

Beaudet, A. L. and Jiang, Y.-h. 2002. A rheostat model for a rapid and reversible form of imprinting-dependent evolution. *American Journal of Human Genetics* 70, pp. 1389-1397.

Beckwith, J. B. 1969. Macroglossia, omphalocoele, adrenal cytomegaly, gigantism and hyperplastic visceromegaly. *Birth Defects* 5, pp. 188-196.

Beechey, C. V. *et al.* 2008. *MRC Harwell, Oxfordshire. World Wide Web Site - Mouse Imprinting Data and References* [Online]. Available at: http://www.har.mrc.ac.uk/research/genomic_imprinting/ [Accessed: 27/05/2008].

Bell, A. C. and Felsenfeld, G. 2000. Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. *Nature* 405(6785), pp. 482-485.

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

Bestor, T. *et al.* 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. The carboxyl-terminal domain of the mammalian enzymes is related to bacterial restriction methyltransferases. *Journal of Molecular Biology* 203(4), pp. 971-983.

Bhogal, B. *et al.* 2004. Methylation at mouse *Cdkn1c* is acquired during postimplantation development and functions to maintain imprinted expression. *Genomics* 84(6), pp. 961-970.

Bird, A. 2002. DNA methylation patterns and epigenetic memory. *Genes and Development* 16, pp. 6-21.

Bird, A. *et al.* 1985. A fraction of the mouse genome that is derived from islands of nonmethylated, CpG-rich DNA. *Cell* 40, pp. 91-99.

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

Bird, A. P. and Wolffe, A. P. 1999. Methylation-Induced Repression - Belts, Braces, and Chromatin. *Cell* 99, pp. 451-454.

Blagitko, N. *et al.* 2000. Human *GRB10* is imprinted and expressed from the paternal and maternal allele in a highly tissue- and isoform-specific fashion. *Human Molecular Genetics* 9(11), pp. 1587-1595.

Bliek, J. *et al.* 2001. Increased tumour risk for BWS patients correlates with aberrant *H19* and not *KCNQ1OT1* methylation: occurrence of *KCNQ1OT1* hypomethylation in familial cases of BWS *Human Molecular Genetics* 10(5), pp. 467-476.

Bock, C. *et al.* 2005. BiQ Analyzer: visualization and quality control for DNA methylation data from bisulfite sequencing. *Bioinformatics* 21(21), pp. 4067-4068.

Boggs, B. A. and Chinault, A. C. 1994. Analysis of replication timing properties of human X-chromosomal loci by fluorescence in situ hybridization. *Proceedings of the National Academy of Science USA* 91(13), pp. 6083-6087.

Bouniol, C. *et al.* 1995. Endogenous transcription occurs at the 1-cell stage in the mouse embryo. *Experimental Cell Research* 218(1), pp. 57-62.

Bourc'his, D. *et al.* 2001. Dnmt3L and the establishment of the maternal genomic imprints. *Science* 294, pp. 2536-2539.

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

Boyer, L. A. *et al.* 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. *Nature* 411, pp. 349-353.

Brandeis, M. *et al.* 1994. Sp1 elements protect a CpG island from *de novo* methylation. *Nature* 371, pp. 435-438.

Breiling, A. *et al.* 2004. Epigenome changes in active and inactive polycomb-group-controlled regions. *EMBO Reports* 5(10), pp. 976-981.

Buck, M. J. and Lieb, J. D. 2004. ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics* 83, pp. 349-360.

Buiting, K. *et al.* 1995. Inherited microdeletions in the Angelman and Prader–Willi syndromes define an imprinting centre on human chromosome 15. *Nature Genetics* 9, pp. 395-400.

Cao, R. *et al.* 2002. Role of histone H3 lysine 27 methylation in polycomb-group silencing. *Science* 298, pp. 1039-1043.

Carr, M. S. *et al.* 2007. Allele-specific histone modifications regulate expression of the *Dlk1-Gtl2* imprinted domain. *Genomics* 89(2), pp. 280-290.

Caspary, T. *et al.* 1998. Multiple mechanisms regulate imprinting of the mouse distal chromosome 7 gene cluster. *Molecular and Cellular Biology* 18(6), pp. 3466-3474.

Caspary, T. *et al.* 1999. Oppositely imprinted genes *p57 (Kip2)* and *Igf2* interact in a mouse model for Beckwith-Wiedemann syndrome. *Genes and Development* 13, pp. 3115-3124.

Cattanach, B. M. and Kirk, M. 1985. Differential activity of maternally and paternally derived chromosome regions in mice. *Nature* 315(6019), pp. 496-498.

Cerrato, F. *et al.* 2005. The two-domain hypothesis in Beckwith-Wiedemann syndrome: autonomous imprinting of the telomeric domain of the distal chromosome 7 cluster. *Human Molecular Genetics* 14(4), pp. 503-511.

Chan, M. F. *et al.* 2001. Reduced rates of gene loss, gene silencing, and gene mutation in *Dnmt1*-deficient embryonic stem cells. *Molecular and Cellular Biology* 21(22), pp. 7587-7600.

Charalambous, M. *et al.* 2003. Disruption of the imprinted *Grb10* gene leads to disproportionate overgrowth by an *Igf2*-independent mechanism. *Proceedings of the National Academy of Science U S A* 100(14), pp. 8292-8297.

Chedin, F. *et al.* 2002. The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. *Proceedings of the National Academy of Science USA* 99(26), pp. 16916-16921.

Chen, H. Y. *et al.* 1998. Ubiquitination of histone H3 in elongating spermatids of rat testes. *Journal of Biological Chemistry* 273(21), pp. 13165-13169.

Chen, T. *et al.* 2003. Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. *Molecular Cell Biology* 23(16), pp. 5594-5605.

Chen, Z.-X. *et al.* 2005. Physical and functional interactions between the human DNMT3L protein and members of the de novo methyltransferase family. *Journal of Cellular Biochemistry* 2005(95), pp. 902-917.

Choo, J. H. *et al.* 2006. Allele-specific deposition of macroH2A1 in imprinting control regions. *Human Molecular Genetics* 15(5), pp. 717-724.

Choo, J. H. *et al.* 2007. MacroH2A1 knockdown effects on the Peg3 imprinted domain. *BMC Genomics* 8, p. 479.

Chung, W. Y. *et al.* 1996. Chromosome 11p15.5 regional imprinting: comparative analysis of *KIP2* and *H19* in human tissues and Wilms' tumors. *Human Molecular Genetics* 5(8), pp. 1101-1108.

Cirio, M. C. *et al.* 2008. Preimplantation expression of the somatic form of *Dnmt1* suggests a role in the inheritance of genomic imprints. *BMC Developmental Biology* 8, pp. 9-22.

Clemson, C. M. *et al.* 1996. XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. *Journal of Cellular Biology* 132(3), pp. 259-275.

Cooper, D. N. and Krawczak, M. 1989. Cytosine methylation and the fate of CpG dinucleotides in vertebrate genomes. *Human Genetics* 83(2), pp. 181-188.

Cooper, W. N. *et al.* 2005. Molecular subtypes and phenotypic expression of Beckwith-Wiedemann syndrome. *European Journal of Human Genetics* 1, pp. 1-8.

Cuthbert, G. and Bannister, A. 2005. *Guide to chromatin immunoprecipitation methods* [Online]. Available at: www.abcam.com [Accessed: 01/10/05].

Dao, D. *et al.* 1998. *IMPT1*, an imprinted gene similar to polyspecific transporter and multi-drug resistance genes. *Human Molecular Genetics* 7(4), pp. 597-608.

Davis, T. L. *et al.* 1999. Acquisition of the *H19* methylation imprint occurs differentially on the parental alleles during spermatogenesis. *Genomics* 58(1), pp. 18-28.

Davis, T. L. *et al.* 2000. The *H19* methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. *Human Molecular Genetics* 9(19), pp. 2885-2894.

de la Cruz, C. C. *et al.* 2005. Developmental regulation of Suz12 localization. *Chromosoma* 114, pp. 183-192.

De La Fuente, R. *et al.* 2006. Lsh is required for meiotic chromosome synapsis and retrotransposon silencing in female germ cells. *Nature Cell Biology* 8(12), pp. 1448-1454.

de Napoles, M. *et al.* 2004. Polycomb group proteins ring1A/B link ubiquitylation of histone H2A to heritable gene silencing and X inactivation. *Developmental Cell* 7, pp. 663-676.

Dean, W. *et al.* 1998. Altered imprinted gene methylation and expression in completely ES cell-derived mouse fetuses: association with aberrant phenotypes. *Development* 125, pp. 2273-2282.

Delaval, K. and Feil, R. 2004. Epigenetic regulation of mammalian genomic imprinting. *Current Opinion in Genetics and Development* 14(2), pp. 188-195.

Delaval, K. *et al.* 2007. Differential histone modifications mark mouse imprinting control regions during spermatogenesis. *EMBO Journal* 26, pp. 720-729.

Dennis, K. *et al.* 2001. Lsh, a member of the SNF2 family, is required for genome-wide methylation. *Genes and Development* 15(22), pp. 2940-2944.

Deplus, R. *et al.* 2002. Dnmt3L is a transcriptional repressor that recruits histone deacetylase. *Nucleic Acids Research* 30(17), pp. 3831-3838.

Diagenode. 2006. *Bioruptor manual* [Online]. Available at: www.diagenode.com [Accessed: 21/06/06].

Diaz-Meyer, N. *et al.* 2003. Silencing of *CDKN1C (p57KIP2)* is associated with hypomethylation at KvDMR1 in Beckwith-Wiedemann syndrome. *Journal of Medical Genetics* 40(11), pp. 797-801.

Diaz-Meyer, N. *et al.* 2005. Alternative mechanisms associated with silencing of *CDKN1C* in Beckwith-Wiedemann syndrome. *Journal of Medical Genetics* 42(8), pp. 648-655.

Doetschman, T. *et al.* 1985. The *in vitro* development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *Journal of Embryology and Experimental Morphology* 87, pp. 27-45.

Du, M. *et al.* 2003. Insulator and silencer sequences in the imprinted region of human chromosome 11p15.5. *Human Molecular Genetics* 12(15), pp. 1927-1939.

Durcova-Hills, G. *et al.* 2001. Pluripotent stem cells derived from migrating primordial germ cells. *Differentiation* 68, pp. 220-226.

Durcova-Hills, G. *et al.* 2004. Analysis of sex differences in EGC imprinting. *Developmental Biology* 268, pp. 105-110.

Durcova-Hills, G. *et al.* 2006. Influence of sex chromosome constitution on the genomic imprinting of germ cells. *Proceedings of the National Academy of Science USA* 103(30), pp. 11184-11188.

Durcova-Hills, G. *et al.* 2002. Developmental fate of embryonic germ cells (EGCs), *in vivo* and *in vitro*. *Differentiation* 71, pp. 135-141.

Ehrlich, M. *et al.* 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues or cells. *Nucleic Acids Research* 10(8), pp. 2709-2721.

Ehrlich, M. *et al.* 2006. Immunodeficiency, centromeric region instability, facial anomalies syndrome (ICF). *Orphanet Journal of Rare Diseases* 1(2).

Engemann, S. *et al.* 2002. Bisulfite-based methylation analysis of imprinted genes. In: Ward, A. ed. *Genomic Imprinting Methods and Protocols*. Vol. 181. Totowa: Humana Press, pp. 217-228.

Engemann, S. *et al.* 2000. Sequence and functional comparison in the Beckwith-Wiedemann region: implications for a novel imprinting centre and extended imprinting. *Human Molecular Genetics* 9(18), pp. 2691-2706.

Estève, P.-O. *et al.* 2006. Direct interaction between DNMT1 and G9a coordinates DNA and histone methylation during replication. *Genes and Development* 20(22), pp. 3089-3103.

Evans, M. J. and Kaufman, M. H. 1981. Establishment in culture of pluripotential cells from mouse embryos *Nature* 292(5819), p. 154.

Fan, T. *et al.* 2005. Lsh controls silencing of the imprinted *Cdkn1c* gene. *Development* 132(4), pp. 635-644.

Feil, R. *et al.* 1997. Parental chromosome-specific chromatin conformation in the imprinted *U2af1-rs1* gene in the mouse. *Journal of Biological Chemistry* 272(33), pp. 20893-20900.

Finch, J. T. *et al.* 1977. Structure of nucleosome core particles of chromatin. *Nature* 269(5623), pp. 29-36.

Fitzpatrick, G. V. *et al.* 2007. Allele-specific binding of CTCF to the multipartite imprinting control region KvDMR1. *Molecular and Cellular Biology* 27(7), pp. 2636-2647.

Fitzpatrick, G. V. *et al.* 2002. Regional loss of imprinting and growth deficiency in mice with a targeted deletion of KvDMR1. *Nature Genetics* 32, pp. 426-431.

Fournier, C. *et al.* 2002. Allele-specific histone lysine methylation marks regulatory regions at imprinted mouse genes. *EMBO Journal* 21(23), pp. 6560-6570.

Frank, D. *et al.* 2002. Placental overgrowth in mice lacking the imprinted gene *Ipl*. *Proceedings of the National Academy of Science USA* 99(11), pp. 7490-7495.

Fuks, F. *et al.* 2000. DNA methyltransferase Dnmt1 associates with histone deacetylase activity. *Nature Genetics* 24(1), pp. 88-91.

Fulmer-Smentek, S. B. and Uta, F. 2001. Association of acetylated histones with paternally expressed genes in the Prader-Willi deletion region. *Human Molecular Genetics* 10(6), pp. 645-652.

Gabory, A. *et al.* 2006. The *H19* gene: regulation and function of a non-coding RNA. *Cytogenetic Genome Research* 113, pp. 188-193.

Gardiner-Garden, M. and Frommer, M. 1987. CpG islands in vertebrate genomes. *Journal of Molecular Biology* 196(2), pp. 261-282.

Geiman, T. M. *et al.* 1998. Characterization of gene expression, genomic structure, and chromosomal localization of Hells (Lsh). *Genomics* 54(3), pp. 477-483.

Geisberg, J. V. and Struhl, K. 2004. Quantitative sequential chromatin immunoprecipitation, a method for analyzing co-occupancy of proteins at genomic regions *in vivo*. *Nucleic Acids Research* 32(19), pp. e151-159.

Gicquel, C. and Le Bouc, Y. 2001. *Beckwith-Wiedemann syndrome* [Online]. Orphanet encyclopedia. Available at: http://www.orpha.net/data/patho/GB/uk-BWS.pdf [Accessed: 24/01/05].

Ginsburg, M. *et al.* 1990. Primordial germ cells in the mouse embryo during gastrulation. *Development* 110(2), pp. 521-528.

Goll, M. G. *et al.* 2006. Methylation of $tRNA^{Asp}$ by the DNA methyltransferase homolog Dnmt2. *Science* 311, pp. 395-398.

Gould, T. D. and Pfeifer, K. 1998. Imprinting in mouse *Kvlqt1* is developmentally regulated. *Human Molecular Genetics* 7(3), pp. 483-487.

Grandjean, V. *et al.* 2000. Increased IGF-II protein affects $p57^{kip2}$ expression *in vivo* and *in vitro*: implications for Beckwith-Wiedemann syndrome. *Proceedings of the National Academy of Science USA* 97(10), pp. 5279-5284.

Gratchev, A. 2004. *Analysis of methylation using bisulphite sequencing* [Online]. Methods.info. Available at: http://www.methods.info/Methods/DNA_methylation/Bisulphite_sequencing.html [Accessed: 3 November].

Gregorova, S. and Forejt, J. 2000. PWD/Ph and PWK/Ph inbred mouse strains of Mus m. musculus subspecies--a valuable resource of phenotypic variations and genomic polymorphisms. *Folia Biol (Praha)* 46(1), pp. 31-41.

Gregory, R. I. *et al.* 2001. DNA methylation is linked to deacetylation of histone H3, but not H4, on the imprinted genes *Snrpn* and *U2af1-rs1* *Molecular and Cellular Biology* 21(16), pp. 5426-5436.

Haig, D. 1997. Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proceedings of the Royal Society B: Biological Sciences* 264, pp. 1657-1662.

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

Haig, D. and Westoby, M. 1989. Parent-specific gene expression and the triploid endosperm. *American Naturalist* 134(1), p. 147.

Hajkova, P. *et al.* 2002. Epigenetic reprogramming in mouse primordial germ cells. *Mechanisms of Development* 117, pp. 15-23.

Hansen, R. S. *et al.* 1999. The *DNMT3B* DNA methyltransferase gene is mutated in the ICF immunodeficiency syndrome. *Proceedings of the National Academy of Science USA* 96(25), pp. 14412-14417.

Harrington, M. A. *et al.* 1988. Cytosine methylation does not affect binding of transcription factor Sp1. *Proceedings of the National Academy of Science USA* 85, pp. 2066-2070.

Hata, K. *et al.* 2002. Dnmt3L cooperates with the Dnmt3 family of de novo DNA methyltransferases to establish maternal imprints in mice. *Development* 129, pp. 1983-1993.

Hatada, I. and Mukai, T. 1995. Genomic imprinting of $p57^{KIP2}$, a cyclin-dependent kinase inhibitor, in mouse. *Nature Genetics* 11, pp. 204-206.

Hatada, I. *et al.* 1996. An imprinted gene *p57KIP2* is mutated in Beckwith–Wiedemann syndrome. *Nature Genetics* 14, pp. 171-173.

Hayatsu, H. 2006. Bisulfite modifcation of cytosine and 5-methyltcytosine as used in epigenetic studies. *Genes and Environment* 28(1).

Hebbes, T. R. *et al.* 1988. A direct link between core histone acetylation and transcriptionally active chromatin. *Cell* 74, pp. 281-289.

Henry, I. *et al.* 1991. Uniparental paternal disomy in a genetic cancer-predisposing syndrome. *Nature* 351, pp. 665-667.

Higashimoto, K. *et al.* 2003. Loss of CpG methylation is strongly correlated with loss of histone H3 lysine 9 methylation at DMR-LIT1 in patients with Beckwith-Widedmann syndrome. *American Journal of Human Genetics* 73, pp. 948-956.

Hikichi, T. *et al.* 2003. Imprinting regulation of the murine *Meg1/Grb10* and human *GRB10* genes; roles of brain-specific promoters and mouse-specific CTCF-binding sites. *Nucleic Acids Research* 31(5), pp. 1398-1406.

Hiura, H. *et al.* 2007. DNA methylation imprints on the IG-DMR of the *Dlk1-Gtl2* domain in mouse male germline. *FEBS letters* 581, pp. 1255-1260.

Hiura, H. *et al.* 2006. Oocyte growth-dependent progression of maternal imprinting in mice. *Genes to Cells* 11(4), pp. 353-361.

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

Holmes, R. *et al.* 2003. A Comprehensive transcript map of the mouse *Gnas* imprinted complex. *Genome Research* 13, pp. 1410-1415.

Hoovers, J. M. *et al.* 1995. Multiple genetic loci within 11p15 defined by Beckwith-Wiedemann syndrome rearrangement breakpoints and subchromosomal transferable fragments. *Proceedings of the National Academy of Science USA* 92(26), pp. 12456-12460.

Horii, T. *et al.* 2008. Loss of genomic imprinting in mouse parthenogenetic embryonic stem cells. *Stem Cells* 26(1), pp. 79-88.

Horike, S.-i. *et al.* 2000. Targeted disruption of the human *LIT1* locus defines a putative imprinting control elelment playing an essential role in Beckwith-Wiedemann syndrome. *Human Molecular Genetics* 9(14), pp. 2075-2083.

Howell, C. Y. *et al.* 2001. Genomic imprinting disrupted by a maternal effect mutation in the Dnmt1 gene. *Cell* 104, pp. 829-838.

Hsieh, C. L. 1999. *In vivo* activity of murine *de novo* methyltransferases, Dnmt3a and Dnmt3b. *Molecular Cell Biology* 19(12), pp. 8211-8218.

Humphreys, D. *et al.* 2001. Epigenetic instablity in ES cells and cloned mice. *Science* 293, pp. 95-97.

Ikegami, K. *et al.* 2007. Genome-wide and locus-specific DNA hypomethylation in G9a deficient mouse embryonic stem cells. *Genes to Cells* 12, pp. 1-11.

Ishihara, K. *et al.* 2006. CTCF-dependent chromatin insulator is linked to epigenetic remodeling. *Molecular Cell* 23, pp. 733-742.

Ishizaki, T. *et al.* 2003. Loss of *Igf2* imprinting in monoclonal mouse hepatic tumor cells is not associated with abnormal methylation patterns for the *H19*, *Igf2*, and *Kvlqt1* differentially methylated regions. *Journal of Biological Chemistry* 278(8), pp. 6222-6228.

Iwahashi, K. *et al.* 2007. Autonomous regulation of sex-specific developmental programming in mouse fetal germ cells. *Biology of Reproduction* 77(4), pp. 697-706.

Jeffery, L. and Nakielny, S. 2004. Components of the DNA methylation system of chromatin control are RNA-binding proteins. *The Journal of Biological Chemistry* 279(47), pp. 49479-49487.

Jeltsch, A. *et al.* 2006. Two substrates are better than one: dual spefificities for Dnmt2 methyltransferases. *Trends in Biochemical Sciences* 31(6), pp. 306-308.

Jenuwein, T. and Allis, C. D. 2001. Translating the histone code. *Science* 293, pp. 1074-1079.

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

Jia, D. *et al.* 2007. Structure of Dnmt3a bound to Dnmt3L suggests a model for de novo DNA methylation. *Nature* 449(7159), pp. 248-251.

John, R. M. *et al.* 2001a. Distant *cis*-elements regulate imprinted expression of the mouse $p57^{Kip2}$ (*Cdkn1c*) gene: implications for the human disorder, Beckwith-Wiedemann syndrome. *Human Molecular Genetics* 10(15), pp. 1601-1609.

John, R. M. *et al.* 2001b. Imprinted expression of *Neuronatin* from modified BAC transgenes reveals regulation by distinct and distant enhancers. *Developmental Biology* 236, pp. 387-389.

John, R. M. *et al.* 1999. A human $p57^{KIP2}$ transgene is not activated by passage through the maternal mouse germline. *Human Molecular Genetics* 8(12), pp. 2211-2219.

Jones, B. K. *et al.* 1998a. *Igf2* imprinting does not require its own DNA methylation or *H19* RNA *Genes and Development* 12(14), pp. 2200-2207.

Jones, P. A. and Takai, D. 2001. The role of DNA methylation in mammalian epigenetics. *Science* 293, pp. 1068-1070.

Jones, P. A. *et al.* 1983. Inhibition of DNA methylation by 5-azacytidine. *Recent Results in Cancer Research* 84, pp. 202-211.

Jones, P. L. *et al.* 1998b. Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. *Nature Genetics* 19, pp. 187-191.

Kafri, T. *et al.* 1993. Mechanistic aspects of genome-wide demethylation in the preimplantation mouse embryo. *Proceedings of the National Academy of Science USA* 90(22), pp. 10558-10562.

Kalantry, S. and Magnuson, T. 2006. The Polycomb group protein EED is dispensable for the initiation of random X-chromosome inactivation. *PLoS Genetics* 2(5), p. e66.

Kalantry, S. *et al.* 2006. The polycomb group protein Eed protects the inactive X-chromosome from differentiation-induced reactivation. *Nature Cell Biology* 8(2), pp. 195-202.

Kalscheuer, V. M. *et al.* 1993. The insulin-like growth factor type-2 receptor gene is imprinted in the mouse but not in humans. *Nature Genetics* 5(1), pp. 74-78.

Kanduri, C. *et al.* 2002. A differentially methylated imprinting control region within the *Kcnq1* locus harbors a methylation sensitive chromatin insulator. *The Journal of Biological Chemistry* 277(20), pp. 18106-18110.

Kanduri, C. *et al.* 2006. The length of the transcript encoded from the *Kcnq1ot1* antisense promoter determines the degree of silencing. *EMBO Journal* 25, pp. 2096-2106.

Kaneda, M. *et al.* 2004. Essential role for *de novo* DNA methyltransferase Dnmt3a in paternal and maternal imprinting. *Nature* 429, pp. 900-903.

Karolchik, D. *et al.* 2003. The UCSC Genome Browser Database. *Nucleic Acids Research* 31(1), pp. 51-54.

Kato, Y. *et al.* 2007. Role of the Dnmt3 family in *de novo* methylation of imprinted and repetitive sequences during male germ cell development in the mouse. *Human Molecular Genetics* 16(19), pp. 2272-2280.

Kato, Y. *et al.* 1999. Developmental potential of mouse primordial germ cells. *Development* 126, pp. 1823-1832.

Kerjean, A. *et al.* 2000. Establishment of the paternal methylation imprint of the human *H19* and *MEST/PEG1* genes during spermatogenesis. *Human Molecular Genetics* 9(14), pp. 2183-2187.

Keverne, E. B. *et al.* 1996. Genomic imprinting and the differential roles of parental genomes in brain development. *Developmental Brain Research* 92(1), pp. 91-100.

Kikyo, N. *et al.* 1997. Genetic and functional analysis of *neuronatin* in mice with maternal or paternal duplication of distal chr 2. *Developmental Biology* 190, pp. 66-77.

Killian, J. K. *et al.* 2000. *M6P/IGF2R* imprinting evolution in mammals. *Molecular Cell* 5(4), pp. 707-716.

Killian, J. K. *et al.* 2001. Monotreme *IGF2* expression and ancestral origin of genomic imprinting. *Journal of Experimental Zoology* 291(2), pp. 205-212.

Kimura, H. and Shiota, K. 2003. Methyl-CpG-binding protein, MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. *Journal of Biological Chemistry* 278(7), pp. 4806-4812.

Kohlmaier, A. *et al.* 2004. A chromosomal memory triggered by *Xist* regulates histone methylation in X inactivation. *PLoS Biology* 2(7), pp. 991-1003.

Kornberg, R. D. 1974. Chromatin structure: a repeating unit of histones and DNA. *Science* 184(139), pp. 868-871.

Koufos, A. *et al.* 1989. Familial Wiedemann-Beckwith syndrome and a second wilms tumor locus both map to 11p15.5. *American Journal of Human Genetics* 44, pp. 711-719.

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

Labosky, P. A. *et al.* 1994. Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (*Igf2r*) gene compared with embryonic sterm (ES) cell lines. *Development* 120, pp. 3197-3204.

Lam, W. W. K. *et al.* 1999. Analysis of germline *CDKN1C (p57KIP2)* mutations in familial and sporadic Beckwith-Wiedemann syndrome (BWS) provides a novel genotype-phenotype correlation *Journal of Medical Genetics* 36, pp. 518-523.

Lan, F. *et al.* 2007. A histone H3 lysine 27 demethylase regulates animal posterior development. *Nature* 449(7163), pp. 689-694.

Lander, E. S. *et al.* 2001. Initial sequencing and analysis of the human genome. *Nature* 409(860-921).

Lee, J. *et al.* 2002. Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. *Development* 129, pp. 1807-1817.

Lee, J. T. 2003. Molecular links between X-inactivation and autosomal imprinting: X-inactivation as a driving force for the evolution of imprinting? *Current Biology* 13, pp. R242-R254.

Lee, M. H. *et al.* 1995. Cloning of *p57KIP2*, a cyclin-dependent kinase inhibitor with unique domain structure and tissue distribution. *Genes and Development* 9(6), pp. 639-649.

Lee, M. P. *et al.* 1999. Loss of imprinting of a paternally expressed transcript, with antisense orientation to *KvLQT1*, occurs frequently in Beckwith-

Wiedemann syndrome and is independent of *insulin-like growth factor II* imprinting. *Proceedings of the National Academy of Science USA* 96, pp. 5203-5208.

Lee, M. P. *et al.* 1997. Low frequency of *p57^{KIP2}* mutation in Beckwith-Wiedemann syndrome. *American Journal of Human Genetics* 61, pp. 304-309.

Lehnertz, B. *et al.* 2003. Suv39h-mediated histone H3 lysine 9 methylation directs DNA methylation to major satellite repeats at pericentric heterochromatin. *Current Biology* 13, pp. 1192-1200.

Lei, H. *et al.* 1996. De novo DNA cytosine methyltransferase activities in mouse embryonic stem cells. *Development* 122, pp. 3195-3205.

Leighton, P. A. *et al.* 1995. Disruption of imprinting caused by deletion of the H19 gene region in mice. *Nature* 375, pp. 34-39.

Lerchner, W. and Barlow, D. P. 1997. Paternal repression of the imprinted mouse *Igf2r* locus occurs during implantation and is stable in all tissues of the post-implantation mouse embryo. *Mechanisms of Development* 61(1-2), pp. 141-149.

Leu, Y. W. *et al.* 2003. Double RNA interference of DNMT3b and DNMT1 enhances DNA demethylation and gene reactivation. *Cancer Research* 63(19), pp. 6110-6115.

Lewis, A. *et al.* 2006. Epigenetic dynamics of the *Kcnq1* imprinted domain in the early embryo. *Development* 133(21), pp. 4203-4210.

Li, E. *et al.* 1993. Role for DNA methylation in genomic imprinting. *Nature* 366, pp. 362-365.

Li, E. *et al.* 1992. Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. *Cell* 69, pp. 915-926.

Li, J.-Y. *et al.* 2004a. Timing of establishment of paternal methylation imprints in the mouse. *Genomics* 84, pp. 952-960.

Li, L. C. and Dahiya, R. 2002. MethPrimer: designing primers for methylation PCRs. *Bioinformatics* 18(11), pp. 1427-1431.

Li, T. *et al.* 2004b. Activating and silencing histone modifications form independent allelic switch regions in the imprinted *Gnas* gene. *Human Molecular Genetics* 13(7), pp. 741-750.

Lin, S.-P. *et al.* 2003. Asymmetric regulation of imprinting on the maternal and paternal chromosomes at the *Dlk1-Gtl2* imprinted cluster on mouse chromosome 12. *Nature Genetics* 35(1), pp. 97-102.

Liu, J. *et al.* 2000. Identification of a methylation imprint mark within the mouse *Gnas* locus. *Molecular and Cellular Biology* 20(16), pp. 5808-5817.

Livak, K. J. and Schmittgen, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-deltadeltaCT}$ method. *Methods* 25, pp. 402-408.

Lopes, S. *et al.* 2003. Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. *Human Molecular Genetics* 12(3), pp. 295-305.

Lucifero, D. *et al.* 2007. Coordinate regulation of DNA methyltransferase expression during oogenesis. *BMC Developmental Biology* 7, p. 36.

Lucifero, D. *et al.* 2004. Gene-specific timing and epigenetic memory in oocyte imprinting. *Human Molecular Genetics* 13(8), pp. 839-849.

Luger, K. *et al.* 1997. Crystal structure of the nucleosome core particle at 2.8 A resolution. *Nature* 389(6648), pp. 251-260.

Lyle, R. *et al.* 2000. The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nature Genetics* 25, pp. 19-21.

Macleod, D. *et al.* 1994. Sp1 sites in the mouse *aprt* gene promoter are required to prevent methylation of the CpG island. *Genes and Development* 8(19), pp. 2282-2292.

Mager, J. *et al.* 2003. Genome imprinting regulated by the mouse polycomb group protein Eed. *Nature Genetics* 33, pp. 502-507.

Maher, E. R. and Reik, W. 2000. Beckwith-Wiedemann syndrome: imprinting in clusters revisited. *Journal of Clinical Investigation* 105(3), pp. 247-252.

Makedonski, K. *et al.* 2005. MeCP2 deficiency in Rett syndrome causes epigenetic aberrations at the PWS/AS imprinting center that affects UBE3A expression. *Human Molecular Genetics* 14(8), pp. 1049-1058.

Mancini-DiNardo, D. *et al.* 2003. A differentially methylated region within the gene *Kcnq1* function as as an imprinted promoter and silencer. *Human Molecular Genetics* 12(3), pp. 283-294.

Mancini-DiNardo, D. *et al.* 2006. Elongation of the *Kcnq1ot1* transcript is required for genomic imprinting of neighbouring genes. *Genes and Development* 20, pp. 1268-1282.

Mann, M. R. *et al.* 2004. Selective loss of imprinting in the placenta following preimplantation development in culture. *Development* 131(15), pp. 3727-3735.

Margot, J. B. *et al.* 2003. Interactions within the mammalian DNA methyltransferase family. *BMC Molecular Biology* 4, p. 7.

Martin, G. R. 1981. Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proceedings of the National Academy of Science U S A* 78(12), pp. 7634-7638.

Matsui, Y. *et al.* 1992. Derivation of pluripotential embryonic stem cells from murine primordial germ cells in culture. *Cell* 70, pp. 841-847.

Matsuoka, S. *et al.* 1995. p57KIP2, a structurally distinct member of the p21CIP1 Cdk inhibitor family, is a candidate tumor suppressor gene. *Genes and Development* 9(6), pp. 650-662.

Matsuoka, S. *et al.* 1996. Imprinting of the gene encoding a human cyclin-dependent kinase inhibitor, *p57KIP2*, on chromosome 11p15. *Proceedings of the National Academy of Science USA* 93(7), pp. 3026-3030.

McDonald, L. E. *et al.* 1998. Bisulfite genomic sequencing-derived methylation profile of the *Xist* gene throughout early mouse development. *Genomics* 54, pp. 379-386.

McGhee, J. D. *et al.* 1983. Higher order structure of chromatin: orientation of nucleosomes within the 30 nm chromatin solenoid is independent of species and spacer length. *Cell* 33(3), pp. 831-841.

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

McLaren, A. and Southee, D. 1997. Entry of mouse embryonic germ cells into meiosis. *Developmental Biology* 187(1), pp. 107-113.

Mertineit, C. *et al.* 1998. Sex-specific exons control DNA methyltransferase in mammalian germ cells. *Development* 125(5), pp. 889-897.

Misteli, T. *et al.* 2000. Dynamic binding of histone H1 to chromatin in living cells. *Nature* 408(6814), pp. 877-881.

Miyoshi, N. *et al.* 1998. Identification of the *Meg1/Grb10* imprinted gene on mouse proximal chromosome 11, a candidate for the Silver-Russell syndrome gene. *Proceedings of the National Academy of Science U S A* 95(3), pp. 1102-1107.

Mohandas, T. *et al.* 1981. Reactivation of an inactive human X chromosome: evidence for X inactivation by DNA methylation. *Science* 211(4480), pp. 393-396.

Monk, D. *et al.* 2006. Limited evolutionary conservation of imprinting in the human placenta. *Proceedings of the National Academy of Science USA* 103(17), pp. 6623-6628.

Monk, M. *et al.* 1987. Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. *Development* 99, pp. 371-382.

Montgomery, N. D. *et al.* 2005. The murine polycomb group protein Eed is required for global histone H3 Lysine-27 methylation. *Current Biology* 15, pp. 942-947.

Moore, T. and Haig, D. 1991. Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in Genetics* 7(2), pp. 45-49.

Moore, T. and Reik, W. 1996. Genetic conflict in early development: parental imprinting in normal and abnormal growth. *Reviews in Reproduction* 1(2), pp. 73-77.

Morison, I. M. and Reeve, A. E. 1998. A catalogue of imprinted genes and parent-of-origin effects in humans and animals. *Human Molecular Genetics* 7(10), pp. 1599-1609.

Murakami, K. *et al.* 2007. Suggestive evidence for chromosomal localization of non-coding RNA from imprinted *LIT1*. *Journal of Human Genetics* 52, pp. 926-933.

Nan, X. *et al.* 1998. Transcriptional repression by the methyl CpG binding protein MeCP2 involves a histone deacetylase complex. *Nature* 393, pp. 386-389.

Nesterova, T. B. *et al.* 2002. *Xist* expression and macroH2A1.2 localisation in mouse primordial and pluripotent embryonic germ cells. *Differentiation* 69, pp. 216-225.

Niemitz, E. L. *et al.* 2004. Microdeletion of *LIT1* in familial Beckwith-Wiedemann syndrome. *American Journal of Human Genetics* 75, pp. 844-849.

Nimura, K. *et al.* 2006. Dnmt3a2 targets endogenous Dnmt3L to ES cell chromatin and induces regional DNA methylation. *Genes to Cells* 11, pp. 1225-1237.

Nolan, C. M. *et al.* 2001. Imprint status of *M6P/IGF2R* and *IGF2* in chickens. *Development Genes and Evolution* 211(4), pp. 179-183.

O'Neill, L. P. and Turner, B. M. 2003. Immunoprecipitation on native chromatin: NChIP. *Methods* 31, pp. 76-82.

O'Sullivan, F. *et al.* 2007. Imprinted expression of the canine IGF2R, in the absence of an anti-sense transcript or promoter methylation. *Evolution and Development* 9(6), pp. 579-589.

Obata, Y. *et al.* 1998. Disruption of primary imprinting during oocyte growth leads to the modified expression of imprinted genes during embryogenesis. *Development* 125, pp. 1553-1560.

Ogawaa, H. *et al.* 2006. Disruption of parental-specific expression of imprinted genes in uniparental fetuses *FEBS letters* 580(22), pp. 5377-5384.

Ohhata, T. *et al.* 2008. Crucial role of antisense transcription across the *xist* promoter in *Tsix*-mediated *Xist* chromatin modifications. *Development* 135, pp. 227-235.

Okamoto, I. *et al.* 2004. Epigenetic dynamics of imprinted X inctivation during early mouse development. *Science* 303, pp. 644-649.

Okano, M. *et al.* 1999. DNA methyltransferases Dnmt3a and Dnmt3b are essential for de novo methylation and mammalian development. *Cell* 99(3), pp. 247-257.

Okano, M. *et al.* 1998a. Cloning and characterization of a family of novel mammalian DNA (cytosine-5) methyltransferases. *Nature Genetics* 19(3), pp. 219-220.

Okano, M. *et al.* 1998b. Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells. *Nucleic Acids Research* 26(11), pp. 2536-2540.

Olek, A. *et al.* 1996. A modified and improved method for bisulphite base cytosine methylation analysis. *Nucleic Acids Research* 24(24), pp. 5064-5066.

Olins, A. L. and Olins, D. E. 1974. Spheroid chromatin units (v bodies). *Science* 183(122), pp. 330-332.

Onyango, P. *et al.* 2000. Sequence and comparative analysis of the mouse 1-megabase region orthologous to the human 11p15 imprinted domain. *Genome Research* 10, pp. 1697-1710.

Orlando, V. 2000. Mapping chromosomal proteins *in vivo* by formaldehyde-crosslinked-chromatin immunoprecipitation. *Trends in Biochemical Sciences* 25, pp. 99-104.

Oswald, J. *et al.* 2000. Active demethylation of the paternal genome in the mouse zygote. *Current Biology* 10, pp. 475-478.

Pasini, D. *et al.* 2004. Suz12 is essential for mouse development and for EZH2 histone methyltransferase activity. *EMBO Journal* 23, pp. 4061-4071.

Paulin, R. *et al.* 1998. Urea improves efficiency of bisulphite-mediated sequencing of 5'-methylcytosine in genomic DNA. *Nucleic acids research* 26(21), pp. 5009-5010.

Paulsen, M. *et al.* 1998. Syntenic organization of the mouse distal chromosome & imprinting cluster and the Beckwith-Wiedemann syndrome region in chromosome 11p15.5. *Human Molecular Genetics* 7(7), pp. 1149-1159.

Paulsen, M. *et al.* 2000. Sequence conservation and variability of imprinting within the Beckwith-Wiedemann syndrome gene cluster in human and mouse. *Human Molecular Genetics* 9(12), pp. 1829-1841.

Pennings, S. *et al.* 1994. Linker histones H1 and H5 prevent the mobility of positioned nucleosomes. *Proc Natl Acad Sci U S A* 91(22), pp. 10275-10279.

Ping, A. J. *et al.* 1989. Genetic linkage of Beckwith-Wiedemann syndrome to 11p15. *American Journal of Human Genetics* 44, pp. 720-723.

Plagge, A. *et al.* 2005. Imprinted Nesp55 influences behavioral reactivity to novel environments. *Molecular and Cellular Biology* 25(8), pp. 3019-3026.

Prissette, M. *et al.* 2001. Methylation profiles of *DXPas34* during the onset of X inactivation. *Human Molecular Genetics* 10(1), pp. 21-38.

Ram, P. T. and Schultz, R. M. 1993. Reporter gene expression in G2 of the 1-cell mouse embryo. *Developmental Biology* 156(2), pp. 552-556.

Ratnam, S. *et al.* 2002. Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. *Developmental Biology* 245(2), pp. 304-314.

Razin, A. and Cedar, H. 1977. Distribution of 5-methylcytosine in chromatin. *Proceedings of the National Academy of Science USA* 74, pp. 2725-2728.

Reese, K. J. *et al.* 2007. Maintenance of paternal methylation and repression of the imprinted *H19* gene requires MBD3. *PLoS Genetics* 3(8), p. e137.

Reik, W. *et al.* 2001. Epigenetic reprogramming in mammalian development. *Science (Washington D C)* 293(5532), pp. 1089-1093.

Reik, W. and Walter, J. 2001. Genomic imprinting: parental influence on the genome. *Nature Reviews Genetics* 2, pp. 21-32.

Rein, T. *et al.* 1998. Identifying 5-methylcytosine and related modifications in DNA genomes. *Nucleic Acids Research* 26(10), pp. 2255-2264.

Ren, B. *et al.* 2000. Genome-wide location and function of DNA binding proteins. *Science* 290(5500), pp. 2306-2309.

Resnick, J. L. *et al.* 1992. Long -term proliferation of mouse primordial germ cells in culture. *Nature* 359, pp. 550-551.

Rhee, I. *et al.* 2002. DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. *Nature* 416(6880), pp. 552-556.

Riesewijk, A. M. *et al.* 1998. Absence of an obvious molecular imprinting mechanism in a human fetus with monoallelic IGF2R expression. *Biochemical and Biophysical Research Communications* 245(1), pp. 272-277.

Robertson, K. D. *et al.* 2000. DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. *Nature Genetics* 25(3), pp. 338-342.

Roemer, I. *et al.* 1997. Epigenetic inheritance in the mouse. *Current Biology* 7(4), pp. 277-280.

Rohwedel, J. *et al.* 1996. Primordial germ cell-derived mouse embryonic germ (EG) cells *in vitro* resemble undifferentiated stem cells with respect to differentiation capacity and cell cycle distribution. *Cell Biology International* 20(8), pp. 579-587.

Rougeulle, C. *et al.* 2004. Differential Histone H3 Lys-9 and Lys-27 methylation profiles on the X chromosome. *Molecular and Cellular Biology* 24(12), pp. 5475-5484.

Ruf, N. *et al.* 2006. Expression profiling of uniparental mouse embryos is inefficient in identifying novel imprinted genes. *Genomics* 87, pp. 509-519.

Rugg-Gunn, P. J. *et al.* 2005. Epigenetic status of human embryonic stem cells. *Nature Genetics* 37(6), pp. 585-587.

Sado, T. *et al.* 2000. X inactivation in the mouse embryo deficient for *Dnmt*1: distinct effect of hypomethylation on imprinted and random X inactivation. *Developmental Biology* 225, pp. 294-303.

Salas, M. *et al.* 2004. Placental growth retardation due to loss of imprinting of *Phlda2*. *Mechanisms of Development* 121, pp. 1199-1210.

Salido, E. C. *et al.* 1992. Expression of the X-inactivation-associated gene XIST during spermatogenesis. *Nature Genetics* 2(3), pp. 196-199.

Santos, F. *et al.* 2002. Dynamic reprogramming of DNA methylation in the early mouse embryo. *Developmental Biology* 241, pp. 172-182.

Santos, K. F. *et al.* 2005. The prima donna of epigenetics: the regulation of gene expression by DNA methylation. *Brazilian Journal of Medical and Biological Research* 38, pp. 1531-1541.

Schoenherr, C. J. *et al.* 2003. CTCF maintains differential methylation at the Igf2/H19 locus. *Nature Genetics* 33, pp. 66-69.

Seidl, C. I. M. *et al.* 2006. The imprinted *Air* ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *The EMBO Journal* 25, pp. 3565-3575.

Shin, J.-Y. *et al.* 2008. Two distinct mechanisms of silencing by the KvDMR1 imprinting control region. *EMBO Journal* 27(1), pp. 168-178.

Siegfried, Z. *et al.* 1999. DNA methylation represses transcription *in vivo*. *Nature Genetics* 22(2), pp. 203-206.

Sleutels, F. *et al.* 2002. The non-coding *Air* RNA is required for silencing autosomal imprinted genes. *Nature* 415, pp. 810-813.

Smallwood, A. *et al.* 2007. Functional cooperation between HP1 and DNMT1 mediates gene silencing. *Genes and Development* 21, pp. 1-10.

Smilinich, N. J. *et al.* 1999. A maternally methylated CpG island in *KvLQT1* is associated with an antisense paternal transcript and loss of imprinting in Beckwith-Wiedemann syndrome. *Proceedings of the National Academy of Science USA* 96, pp. 8064-8069.

Soejima, H. *et al.* 2004. Silencing of imprinted CDKN1C gene expression is associated with loss of CpG and histone H3 lysine 9 methylation at DMR-LIT1 in esophageal cancer. *Oncogene* 23(5), pp. 4380-4388.

Sparago, A. *et al.* 2007. Mechanisms causing imprinting defects in familial Beckwith-Wiedemann syndrome with Wilms' tumour. *Human Molecular Genetics* 16(3), pp. 254-264.

Stein, A. *et al.* 1977. Crosslinked histone octamer as a model of the nucleosome core. *Proceedings of the National Academy of Science USA* 74(7), pp. 2780-2784.

Stewart, C. L. *et al.* 1994. Stem cells from primordial germ cells can reenter the germ line. *Developmental Biology* 161, pp. 626-628.

Strahl, B. D. *et al.* 1999. Methylation of histone H3 at lysine 4 is highly conserved and correlates with transcriptionally active nuclei in Tetrahymena. *Proceedings of the National Academy of Science USA* 96(26), pp. 14967-14972.

Surani, M. A. and Barton, S. C. 1983. Development of gynogenetic eggs in the mouse: implications for parthenogenetic embryos. *Science* 222(4627), pp. 1034-1036.

Surani, M. A. *et al.* 1984. Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. *Nature* 308, pp. 548-550.

Sutcliffe, J. *et al.* 1994. Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. *Nature Genetics* 8(1), pp. 52-58.

Suzuki, S. *et al.* 2005. Genomic imprinting of *IGF2*, *p57^{KIP2}* and *PEG1/MEST* in a marsupial, the tammar wallaby. *Mechanisms of Development* 122, pp. 213-222.

Szabo, P. and Mann, J. R. 1994. Expression and methylation of imprinted genes during *in vitro* differentiation of mouse parthenogenetic and androgenetic embryonic stem cell lines. *Development* 120, pp. 1651-1660.

Szabo, P. E. *et al.* 2002. Allele-specific expression of imprinted genes in mouse migratory primordial germ cells. *Mechanisms of Development* 115, pp. 157-160.

Szabo, P. E. and Mann, J. R. 1995. Biallelic expression of imprinted genes in the mouse germ line: implications for erasure, establishment, and mechanisms of genomic imprinting. *Genes and Development* 9, pp. 1857-1868.

Tada, M. *et al.* 1997. Embryonic germ cells induce epigenetic reprogramming of somatic nucleus in hybrid cells. *EMBO Journal* 16(21), pp. 6510-6520.

Tada, T. *et al.* 1998. Epigenotype switching of imprintable loci in embryonic germ cells. *Development, Genes and Evolution* 207, pp. 551-561.

Takada, S. *et al.* 2002. Epigenetic analysis of the *Dlk1-Gtl2* imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with *Igf2-H19*. *Human Molecular Genetics* 11(1), pp. 77-86.

Takahashi, K. *et al.* 2000. Mice lacking a CDK inhibitor, *p57Kip2,* exhibit skeletal abnormalities and growth retardation. *Journal of Biochemistry* 127(1), pp. 73-83.

Tam, P. P. and Snow, M. H. 1981. Proliferation and migration of primordial germ cells during compensatory growth in mouse embryos. *Journal of Embryology and Experimental Morphology* 64, pp. 133-147.

Thorvaldsen, J. L. *et al.* 1998. Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. *Genes and Development* 12, pp. 3693-3702.

Tremblay, K. D. *et al.* 1997. A 5' 2-kilobase-pair region of the imprinted mouse H19 gene exhibits exclusive paternal methylation throughout development. *Molecular and Cellular Biology* 17(8), pp. 4322-4329.

Tucker, K. L. *et al.* 1996. Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of non-imprinted genes. *Genes and Development* 10, pp. 1008-1020.

Turner, B. M. 2005. Reading signals on the nucleosome with a new nomenclature for modified histones. *Nature Structural and Molecular Biology* 12(2), pp. 110-112.

Ueda, T. *et al.* 2000. The paternal methylation imprint of the mouse *H19* locus is acquired in the gonocyte stage during foetal testis development. *Genes to Cells* 5, pp. 649-659.

Umlauf, D. *et al.* 2004. Imprinting along the *Kcnq1* domain on mouse chromosome 7 involves repressive histone methylation and recruitment of Polycomb group complexes. *Nature Genetics* 36(12), pp. 1296-1300.

Valente, T. *et al.* 2005. *Zac1* is expressed in progenitor/stem cells of the neuroectoderm and mesoderm during embryogenesis: differential phenotype of the *Zac1*-expressing cells during development. *Developmental Dynamics* 233, pp. 667-679.

Varmuza, S. and Mann, M. 1994. Genomic imprinting - defusing the ovarian time bomb. *Trends in Genetics* 10(4), pp. 118-123.

Varrault, A. *et al.* 2006. Zac1 regulates an imprinted gene network critically involved in the control of embryonic growth. *Developmental Cell* 11, pp. 711-722.

Vasques, L. R. *et al.* 2005. XIST Repression in the Absence of DNMT1 and DNMT3B. *DNA research* 12(5), pp. 373-378.

Waddington, C. H. ed. 1942. *Organisers and Genes.* Cambridge, UK: Cambridge University Press.

Wagschal, A. *et al.* 2008. G9a histone methyltransferase contributes to imprinting in the mouse placenta. *Molecular Cell Biology* 28(3), pp. 1104-1113.

Walsh, C. P. *et al.* 1998. Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. *Nature Genetics* 20(2), pp. 116-117.

Warnecke, P. M. *et al.* 2002. Identification and resolution of artifacts in bisulfite sequencing. *Methods* 27, pp. 101-107.

Waterston, R. H. *et al.* 2002. Initial sequencing and comparative analysis of the mouse genome. *Nature* 420(6915), pp. 520-562.

Webster, K. E. *et al.* 2005. Meiotic and epigenetic defects in *Dnmt3L*-knockout mouse spermatogenesis. *Proceedings of the National Academy of Science USA* 102(11), pp. 4068-4073.

Wei, Y. *et al.* 1999. Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. *Cell* 97(1), pp. 99-109.

Weidman, J. R. *et al.* 2006. Imprinting of opossum *Igf2r* in the absence of differential methylation and *Air*. *Epigenetics* 1(1), pp. 49-54.

Weksberg, R. *et al.* 2001. Tumor development in the Beckwith-Wiedemann syndrome is associated with a variety of constitutional molecular 11p15 alterations including imprinting defects of *KCNQ10T1*. *Human Molecular Genetics* 10(26), pp. 2989-3000.

Wilkins, J. F. and Haig, D. 2003. What good is genomic imprinting: The function of parent-specific gene expression. *Nature Reviews Genetics* 4(5), pp. 359-368.

Wutz, A. *et al.* 1997. Imprinted expression of the *Igf2r* gene depends on an intronic CpG island. *Nature* 389, pp. 745-749.

Xie, S. *et al.* 1999. Cloning, expression and chromosome locations of the human DNMT3 gene family. *Gene* 236(1), pp. 87-95.

Xie, Z. H. *et al.* 2006. Mutations in DNA methyltransferase DNMT3B in ICF syndrome affect its regulation by DNMT3L. *Human Molecular Genetics* 15(9), pp. 1375-1385.

Xiong, Z. and Laird, P. W. 1997. COBRA: a sensitive and quantitative DNA methylation assay. *Nucleic Acids Research* 25(12), pp. 2532-2534.

Xu, G. L. *et al.* 1999. Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. *Nature* 402(6758), pp. 187-191.

Xu, Y. *et al.* 1993. Functional polymorphism in the parental imprinting of the human *IGF2R* gene. *Biochemical and Biophysical Research Communications* 197(2), pp. 747-754.

Yamasaki-Ishizaki, Y. *et al.* 2007. Role of DNA methylation and histone H3 lysine 27 methylation in tissue-specific imprinting of mouse *Grb10*. *Molecular and Cellular Biology* 27(2), pp. 732-742.

Yamashita, R. *et al.* 2005. Genome-wide analysis reveals strong correlation between CpG islands with nearby transcription start sites of genes and their tissue specificity. *Gene* 350, pp. 129-136.

Yamazaki, Y. *et al.* 2005. Adult mice cloned from migrating primordial germ cells. *Proceedings of the National Academy of Science USA* 102(32), pp. 11361-11366.

Yan, Q. *et al.* 2003. Lsh, a modulator of CpG methylation, is crucial for normal histone methylation. *EMBO Journal* 22(19), pp. 5154-5162.

Yan, Y. *et al.* 1997. Ablation of the CDK inhibitor p57Kip2 results in increased apoptosis and delayed differentiation during mouse development. *Genes and Development* 11(8), pp. 973-983.

Yang, Y. *et al.* 2003. Epigenetic regulation of Igf2/H19 imprinting at CTCF insulator binding sites. *Journal of Cellular Biochemistry* 90, pp. 1038-1055.

Yatsuki, H. *et al.* 2002. Domain regulation of imprinting cluster in *Kip2/Lit1* subdomain on mouse chromosome 7F4/ F5: large-scale DNA methylation analysis reveals that DMR-Lit1 is a putative imprinting control region. *Genome Research* 12, pp. 1860-1870.

Yoder, J. A. *et al.* 1997. Cytosine methylation and the ecology of intragenomic parasites. *Trends in Genetics* 13, pp. 335-340.

Yokomine, T. *et al.* 2006. Evolution of the vertebrate DNMT3 gene family: a possible link between existence of *DNMT3L* and genomic imprinting. *Cytogenetic Genome Research* 113, pp. 75-80.

Zhang, P. *et al.* 1997. Altered cell differentiation and proliferation in mice lacking $p57^{Kip2}$ indicates a role in Beckwith-Wiedemann syndrome. *Nature* 387, pp. 151-158.

Zhang, Y. *et al.* 1999. Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. *Genes and Development* 13, pp. 1924-1935.

Zhu, H. *et al.* 2006. Lsh is involved in de novo methylation of DNA. *EMBO Journal* 25(2), pp. 335-345.

Zvetkova, I. *et al.* 2005. Global hypomethylation of the genome in XX embryonic stem cells. *Nature Genetics* 37(11), pp. 1274-1279.

Zwart, R. *et al.* 2001. Bidirectional action of the *Igf2r* imprint control element on upstream and downstream imprinted genes. *Genes and Development* 15(18), pp. 2361-2366.