University of Alberta

Epigenetic modifications associated with the maternal and paternal alleles of necdin, an imprinted gene in the Prader-Willi Syndrome region

by

Meredith Leah Hanel 

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

In

Medical Sciences – Medical Genetics

Edmonton, Alberta

Spring 2003

Your file Votre référence
ISBN: 0-612-82110-2
Our file Notre référence
ISBN: 0-612-82110-2

University of Alberta

Library Release Form

Name of Author: Meredith Leah Hanel

Title of Thesis: Epigenetic modifications associated with the maternal and paternal alleles of necdin, an imprinted gene in the Prader-Willi Syndrome region

Degree: Doctor of Philosophy

Year this Degree Granted: 2003

128 Tweedsmuir Ave.
London, ON
N5W 5M3

October 10, 2002

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled *Epigenetic modifications associated with the maternal and paternal alleles of necdin, an imprinted gene in the Prader-Willi Syndrome region* by *Meredith Leah Hanel* in partial fulfillment of the requirements for the degree of *Doctor of Philosophy in Medical Sciences-Medical Genetics*.

Dr. Rachel Wevrick

Dr. Michael Walter

Dr. Heather McDermid

Dr. Roseline Godbout

OCT 3, 2002
Date

Dr. Carolyn Brown

Abstract

The human neurobehavioural disorder Prader-Willi Syndrome (PWS) results from loss of the paternal complement of the chromosome region 15q11-q13. An imprinting centre (IC) controls imprinting of a cluster of paternally expressed genes in this region, including necdin (*NDN*). Mouse *Ndn* is located in the region of conserved synteny with the human PWS region on mouse chromosome 7. To investigate how *NDN/Ndn,* a target gene of the IC, is marked for imprinting, I characterized DNA methylation, chromatin structure and protein-DNA interactions in *NDN/Ndn.*

Mouse *Ndn* is expressed in brain, but not in liver or heart. Allele-specific DNA methylation was found in brain and a less distinct allele-specific methylation was found in liver and heart, which were hypomethylated on both alleles. My analysis of mouse *Ndn* in gametes and preimplantation embryos suggested that maternal hypermethylation originates in the gametes, is maintained until the blastocyst stage and then reappears to some extent in adult brain. In human *NDN*, maternal hypermethylation and paternal hypomethylation was evident in fibroblasts and brain (expressed), as well as lymphocytes (not expressed). This suggests that for human *NDN*, DNA methylation serves as a stable epigenetic mark of allele-specific imprinting status.

To investigate chromatin conformation, I analyzed mouse brain and liver nuclei and human fibroblasts for DNase I sensitivity. A large domain (22 kb) of paternal allele-specific sensitivity to DNase I was found in brain on the unmethylated paternal allele. In liver both alleles were resistant to DNase I. Paternal allele-specific sensitivity was also found in fibroblasts for human *NDN*. These studies demonstrate that the paternal allele of

NDN/Ndn is in a large domain of open chromatin conformation in tissues where it is expressed.

To investigate whether allele-specific methylation and allele-specific chromatin structure is correlated with allele-specific binding of proteins to this region, I used *in vivo* footprinting with DNase I, DMS, and UV treatments on fibroblasts and lymphoblasts from PWS and Angelman Syndrome (AS) deletion patients. This identified a potential DNA binding protein that binds only the paternal allele in fibroblasts. These studies provide insight into how genes targeted by the IC adopt epigenetic characteristics that allow allele-specific expression.

Acknowledgements

I would like to thank my supervisor Rachel Wevrick for her constant support. She provided me with a warm learning environment as well as motivated me to strive for high standards in my work during my Ph.D. As a scientist, a teacher and as a friend, she has set an example that I hope to follow.

I would like to thank the other members of our lab, past and present, who were always willing to teach techniques, give advise, discuss ideas, and go for coffee. They all created a nice working environment and made it pleasant to come into the lab each day. I am grateful for all of the friendships I have made with my lab mates, and I am sure we will continue to be friends with for years to come.

I thank my parents, Bill and Helen Munn, who have always supported me throughout my education, and all of my family and friends from back home who have given me encouragement over the years. I would like to especially thank my husband Erich Hanel for giving me comfort, with his constant love and support.

Table of Contents

List of Tables

List of Figures

Glossary

AS Angelman syndrome

Blastocyst The stage of early mammalian development, around implantation, in which an outer layer of cells surrounds an internal cavity and an inner cell mass (ICM).

BWS Beckwith Weidemann syndrome

CTCF CCCTC-binding factor. In vertebrates it is a widely used enhancer-blocking element.

***De novo* methylation** DNA methylation activity that does not need a hemimethylated template.

DMR Differentially methylated region

DMS Dimethylsulfate

Epigenetic A type of modification that can affect the phenotype without affecting the genotype.

Epigenotype The sum of all epigenetic modifications at a given locus.

HDAC Histone deactylase

Imprinting Centre (IC) A single or group of DNA elements that are necessary for the establishment of parent of origin specific imprints of other genes in an imprinted domain.

Maintenance methylation DNA methylation activity that methylates hemimethylated DNA to maintain DNA methylation after DNA replication.

mH_2O milli-Q-H_2O

Morula An embryo at the cleavage stage that has formed a solid cluster of blastomeres resembling a mulberry (around 62 h p.c. in mouse).

PWS Prader-Willi syndrome

SRO Smallest region of deletion overlap

TSA Tricostatin A. An inhibitor of HDAC.

Chapter 1. Introduction

Parts of this chapter were previously published in:

Hanel M.L. and Wevrick, R. (2001) The role of genomic imprinting in human developmental disorders: lessons from Prader-Willi syndrome. Clinical Genetics *59*: 156-164.
Permission was given to republish material.

Genomic Imprinting

Pronuclear transplantation technology brought about experiments that demonstrated the importance of the contributions of the paternal and maternal complements of the embryo. It became known that the two parental genomes are not equivalent from a series of experiments which demonstrated the inability of gynogenetic or parthenogenetic (complete maternal contribution) and androgenetic (complete paternal contribution) embryos to develop properly (Monk *et al.*, 1987; Surani *et al.*, 1984). This is due to a proportion of genes that are imprinted, or expressed at non-equal levels between the maternal and paternal alleles. The first imprinted region was discovered on the distal part of mouse chromosome 7 (Ferguson-Smith *et al.*, 1991). Mouse embryos with two maternal copies of chromosome 7 have growth retardation, which was found to be caused by loss of expression of insulin-like growth factor 2 (*Igf2*), a gene that is repressed on the maternal allele.

Naturally occurring human disorders that are associated with maternal or paternal disomy have also proven useful in the assessment of regions in the genome likely to contain imprinted genes (Engel, 1997). In addition, a few large-scale projects have been used to identify imprinted regions or imprinted genes. One involved the creation of mouse partial uniparental disomies (UPDs), in which both copies of a single chromosome were inherited from the same parent, and the examination of the phenotypic consequences (Beechey and Cattanach, 1995). Another approach used subtractive hybridization of RNA from cell lines derived entirely from either the paternal or maternal genomes (Piras *et al.*, 2000). Since imprinted genes tend to occur in clusters, additional imprinted genes have been identified based on their close proximity to known imprinted

genes (Caspary *et al.*, 1998; Lee and Wevrick, 2000). Some imprinted genes are associated with an Imprinting Centre (IC), which contains elements necessary for establishing parent of origin specific imprints of other genes in the imprinted domain. A schematic diagram showing the resetting of genomic imprints by an IC is shown in Figure 1-1. When an IC is disrupted it leads to improper imprinting of the other imprinted genes in the region (Sutcliffe *et al.*, 1994; Thorvaldsen *et al.*, 1998; Yang *et al.*, 1998; Zwart *et al.*, 2001). A compilation of known imprinted regions in human and mouse is shown in Table 1-1.

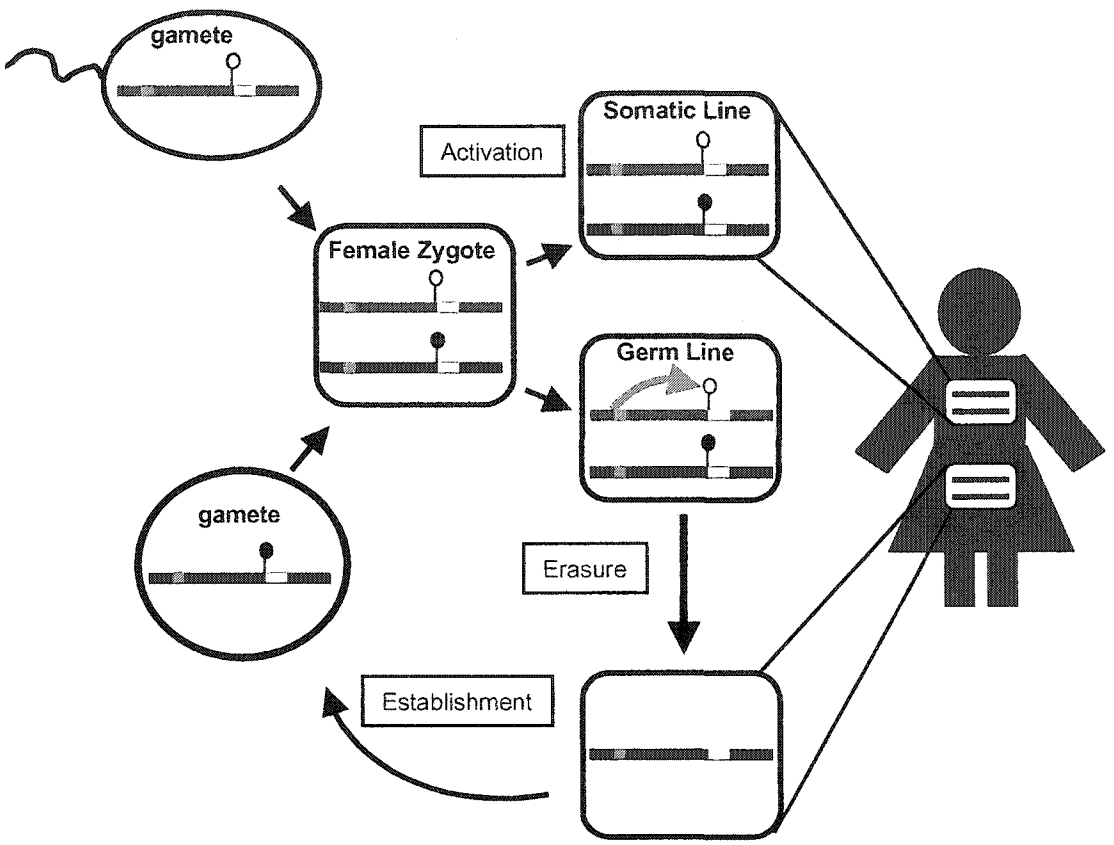


Figure 1-1. Resetting of genomic imprints.

An imprinting centre (IC, green box) controls the imprinting of target genes (white box). An epigenetic mark (filled or open circle), such as methylation of CpG dinucleotides distinguishes the paternal and maternal epigenotypes. The paternal epigenotype (blue) and the maternal epigenotype (pink) are activated and maintained in the somatic cell lineage after fertilization. Imprinted gene expression is indicated by an arrow on the paternally inherited allele. In the female germ line the IC erases and reestablishes the imprint so that both chromosomes are converted to a maternal epigenotype. Conversely, in the male germ line (not shown) both chromosomes are converted to a paternal epigenotype. This figure was taken from Hanel M.L. and Wevrick, R. (2001) Clinical Genetics 59: 156-164.

Table 1-1. Imprinted Genes and Clusters.
Genes within the same cluster are the same colour. A star * indicates this gene is associated with the Imprinting Centre (IC) for the region.

Human Chromosome	Mouse Chromosome	Genes	Associated disease region	Expressed allele
1p31	?	NOEY2	Tumour suppressor	paternal
1p36	?	p73	Oncogene	maternal
5	11	U2AFBPL/U2afbp-rs/U2af1-rs1	Mouse growth effects	paternal (imprinted in mouse only)
6q26	17	M6P/IGF2R	Tumour suppressor	maternal (polymorphic in human)
		Slc22a2		maternal
		Slc22a3		maternal
		Air/IGR2R-AS*		paternal
6q24	10	ZAC/PLAGL1/LOT1	Transient neotatal diabetis mellitus	paternal
		HYMAI		paternal
7q32	6	PEG1/MEST	Silver-Russell syndrome	paternal
		COPG2		maternal
		Cop2as		paternal
		Mit1/Lb9		paternal
7p11.2-p13	11	GRB10/Meg1	Mouse growth affects	maternal
7q21	6	SCGE/Scge	Myoclonus dystonia syndrome	paternal
		PEG10		paternal
11p13		WT1	Wilms tumour	maternal
11p15	7	H19*	Beckwith-Weidemann Syndrome	maternal
		IGF2		paternal
		INS		paternal
11p15	7	KvLQT1/KCNQ1		maternal
		KCNQ1OT1*		paternal
		ASCL2/MASH2		maternal
		p57KIP2/CDKN1C		maternal
		ITM		maternal
		IMPT1/BWR1A/ORCTL2/TSSC5		maternal
		IPL/TSSC3/BWR1C		maternal
		KCNQ1DN		maternal
14q32	12	GTL2/MEG3	human uniparental disomy	maternal
		PEG11		maternal
		MEG8		maternal
		DLK1		paternal
		DAT		paternal
		PEG11-AS		paternal
15q11-13	7	MKRN3/FNZ127	Prader-Willi Syndrome/Angelman Syndrome	paternal
		IPW		paternal
		NDN		paternal
		SNRPN*		paternal
		ZNF127		paternal
		UBE3A antisense		paternal
		UBE3A		maternal
		ATP10C		maternal
18	18	IMPACT	Fetal growth retardation	paternal
19q13.4	7	PEG3	Mouse neonatal lethality	paternal
		ZIM1		maternal
		ZIM3		maternal
		USP29		paternal
		ZNF264/Zfp264		paternal
20	2	GNAS1/Gnas	Mouse neonatal behaviour/lethality	maternal
		Gnasxl		paternal
		NESP55/Nesp		maternal
		Nesp-as		paternal
20q11.2	2	NNAT/Nnat	Leukemia	paternal

Data was compiled from:
Imprinted Gene Databases, Duke University Medical Center, http://www.geneimprint.com/databases/index.html
Mouse imprinted genes, regions and phenotypes, Mammalian Genetics Unit, Harwell, UK, http://www.mgu.har.mrc.ac.uk/imprinting/all_impmaps.html

Prader-Willi Syndrome

Prader-Willi Syndrome (PWS) is a neurobehavioural disorder that is likely due to a basic defect in the hypothalamus of the developing brain. Neonatal hypotonia, hypogonadism, developmental delay and mild dysmorphic features are typically seen and are followed by severe obesity beginning after the first postnatal year (Holm *et al.*, 1993). PWS usually appears sporadically, but in some cases familial transmission has been found (Horsthemke *et al.*, 1997; Orstavik *et al.*, 1992). The most common cause of PWS is a paternal deletion of approximately 3.5 Mb, due to recombination between repeated sequences that flank the common deletion region (Amos-Landgraf *et al.*, 1999; Christian *et al.*, 1999). PWS can also result from maternal UPD for chromosome 15 (Mascari *et al.*, 1992). In either case, patients are missing the paternal contribution of genes in the region. An indistinguishable chromosomal deletion occurring on the maternal allele causes another neurological disorder, distinct from PWS, called Angelman syndrome (AS) (Knoll *et al.*, 1989). AS is characterized by severe mental retardation, seizures and an ataxic gait (Clayton-Smith and Pembrey, 1992). While PWS is thought to be a contiguous gene syndrome involving a number of paternally expressed genes, AS is caused by the inactivation of the single maternally expressed gene *UBE3A* (Kishino *et al.*, 1997; Matsuura *et al.*, 1997; Rougeulle *et al.*, 1997).

A bipartite imprinting centre (IC) controls imprinting of a cluster of genes, including *NDN/Ndn* encoding necdin, in the PWS-AS region. This was discovered from a small group of PWS and AS patients who had no detectable cytogenetic deletion, or uniparental disomy, but had small microdeletions. These microdeletions cause a failure to reset the imprint of chromosome 15 according to the parent of origin during

gametogenesis (Buiting *et al.*, 1998; Saitoh *et al.*, 1996; Sutcliffe *et al.*, 1994). The PWS microdeletions cause an inability to switch to the paternal epigenotype and the AS microdeletions cause an inability to switch to the maternal epigenotype. This means that the PWS and AS microdeletions cause a loss of imprinted expression of genes within this imprinted cluster.

Epigenetic Control of Gene Expression

The most well characterized epigenetic mark is DNA methylation. This is the only chemical modification to the DNA molecule itself that is known to be involved in imprinting. Some other known mammalian epigenetic modifications involve chemical modifications to histones, such as histone acetylation and histone methylation. These histone modifications correlate with differences in the organization of chromatin. Differences in chromatin structure can be detected by allele-specific accessibility to nucleases. All of these modifications seem to act synergistically to produce an epigenotype that determines the level of gene expression. Many epigenetic modifications are not only used for imprinted genes, but also for tissue-specific gene regulation and developmental regulation, so it is important to also understand how they regulate gene expression in general.

DNA methylation in the mammalian genome

In mammals DNA methylation occurs predominantly at CpG dinucleotides and is essential for mammalian development. Because these CpG dinucleotides are susceptible to methylation and methylated cytosine tends to be deaminated to thymine, CpGs are

generally less represented than other dinucleotides in the mammalian genome. In mammals, there are regions called CpG islands, which are defined as a sequence of >200 bp having more than 50% GC content and CpG_{obs}/CpG_{exp} of greater than 0.5 (Gardiner-Garden and Frommer, 1987). These regions are generally protected from methylation, and therefore do not undergo spontaneous deamination. CpG islands are most often located in the promoters of genes. There are also CpG rich regions in repetitive DNA, which are often methylated. This is thought to be the result of an ancient host defense system, where repetitive DNA is recognized as a potential retroposon and methylated by the host to shut down its activity (Walsh and Bestor, 1999). A related function of DNA methylation in repetitive DNA seems to be in genome stability as shown by the chromatin instability syndrome ICF (immunodeficiency, centromeric instability and facial abnormalities), where a loss of the DNA methyltransferase DNMT3B, leads to the formation of multi-radiate chromosomes (Xu *et al.*, 1999). Murine embryonic stem cells that are deficient in the DNA methyltransferase Dnmt1 also show increased genome instability (Chen *et al.*, 1998). Of the CpG dinucleotides in the human genome, 70-80% are methylated (Cross and Bird, 1995). The methylated CpGs are found in repetitive DNA, the 5' regions of X-linked genes on the inactive X chromosome and imprinted genes which may be methylated on one allele. The unmethylated CpG islands constitute 1-2% of the total genome (Colot and Rossignol, 1999).

DNA methylation changes during development

DNA methylation changes dynamically during mammalian development (Figure 1-2). This has been assessed by measuring DNA methylation of specific genes (Kafri *et al.*, 1992; Walsh and Bestor, 1999; Warnecke and Clark, 1999) and of interspersed

repetitive sequences (Howlett and Reik, 1991; Sanford *et al.*, 1987) and by a global analysis of DNA methylation levels (Monk *et al.*, 1987). Together these experiments exemplify an overall pattern of DNA methylation where the levels in sperm are higher than oocytes, but both gametes have lower methylation than somatic cells. Non-CpG island sequences tend to be methylated in sperm, but classical CpG islands are unmethylated in sperm (Walsh and Bestor, 1999). After zygote formation there is a wave of demethylation followed by a wave of remethylation, with the transition stage being around the implantation or blastocyst stage. It is thought that the loss of DNA methylation at the blastocyst stage serves to equalize the two parental genomes. After the blastocyst stage the wave of remethylation occurs, which reestablishes the normal somatic methylation pattern. The establishment of this somatic methylation pattern is essential for survival, as it has been shown that mice without the methyltransferase Dnmt1, do not survive past gastrulation (Lei *et al.*, 1996; Li *et al.*, 1992).

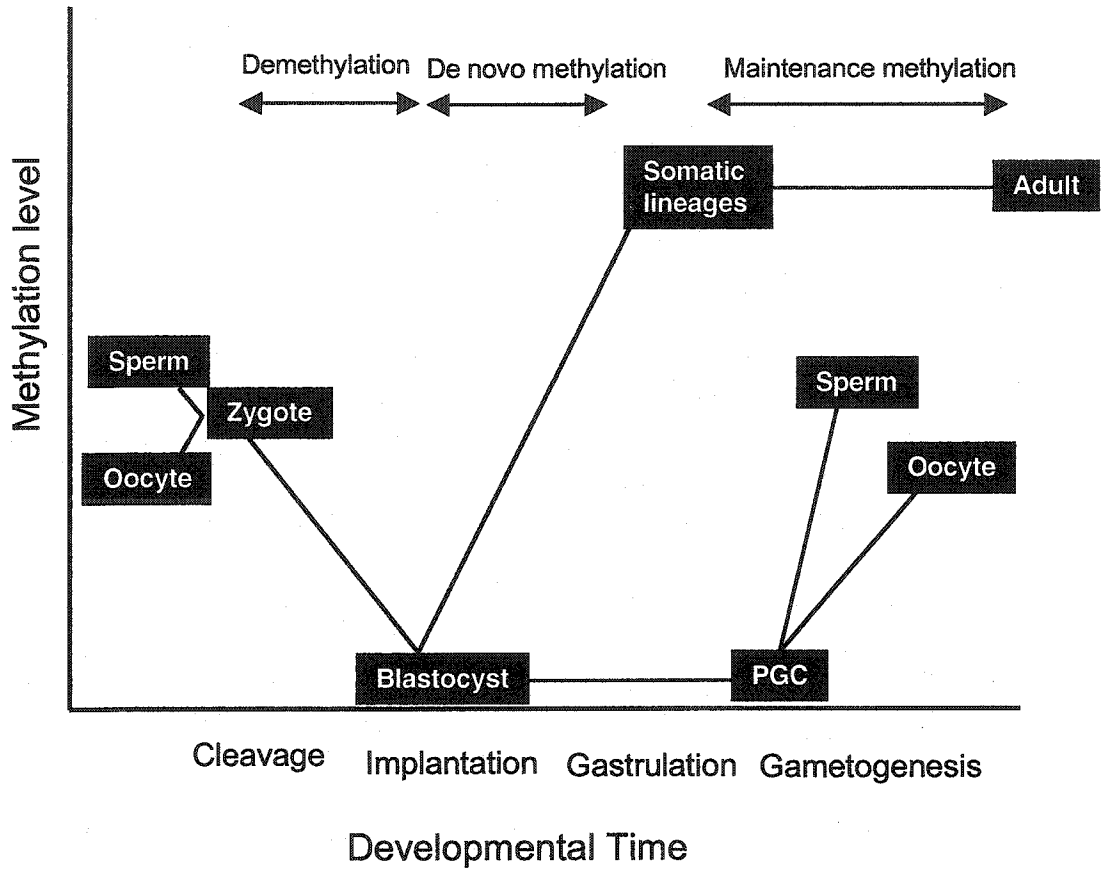


Figure 1-2. Changes in global DNA methylation during development.

(Adapted from Jaenisch, R. (1997) DNA methylation and imprinting: why bother? Trends in Genetics 13: 323-329.)

A global approach to studying methylation during development has been the use of whole cell fluorescence *in situ* hybridization during early development with antibodies to detect methylated CpG (Mayer *et al.*, 2000; Santos *et al.*, 2002). These experiments confirmed that the methylation levels of the two parental genomes are initially unequal with the paternal genome being more methylated than the maternal genome. Furthermore, these experiments demonstrated that the paternal genome undergoes a rapid active demethylation. After the first cell cycle, in the zygote both parental genomes underwent passive demethylation. In contrast to the previous general findings of low levels of methylation at the blastocyst stage, this technique demonstrated that although substantial demethylation had occurred by the morula stage, remethylation had already occurred by the blastocyst stage in the small number of cells that are destined to form the embryo proper, known as the inner cell mass (ICM) (Santos *et al.*, 2002). The authors suggest that this would not be detectable by the previously described methods which analyze bulk material from blastocysts, since the trophectoderm, which constitutes the majority of cells at this stage remains unmethylated.

In the germ line of the developing embryo, DNA methylation levels remain low while the rest of the embryo undergoes a genome wide increase in methylation (Kafri *et al.*, 1992; Monk *et al.*, 1987). This is the time that genomic imprints are thought to be erased (Brandeis *et al.*, 1993; Davis *et al.*, 2000; Labosky *et al.*, 1994). The germ cells then acquire their respective methylation patterns as germ cell maturation ensues.

DNA methyltransferases

There are two different types of DNA methylation activity. There is maintenance methylation activity, which methylates the newly replicated daughter strand

after DNA synthesis to maintain the original methylation pattern. DNA methyltransferase 1 (Dnmt1) appears to perform mainly as a maintenance methyltransferase. It preferentially methylates hemimethylated DNA *in vitro* (Bestor and Ingram, 1983; Gruenbaum *et al.*, 1982), and associates with the replication foci during S-phase (Leonhardt *et al.*, 1992). Mice homozygous for a *Dnmt1* knock out fail to develop properly and die during mid-gestation (Li *et al.*, 1992). Genomic methylation levels in these mice are 3 times lower than normal. Cultured ES cells containing the *Dnmt1* knockout are able to proliferate normally despite the lower levels of genomic methylation, but cannot be induced to differentiate (Tucker *et al.*, 1996).

There is also *de novo* methylation, which is catalyzed by Dnmt3a and Dnmt3b (Hsieh, 1999; Okano *et al.*, 1999). Dnmt3a prefers unmethylated DNA to hemimethylated DNA (Yokochi and Robertson, 2002). Both Dnmt3a and Dnmt3b associate with heterochromatic regions (Bachman *et al.*, 2001). The two *de novo* methyltransferases Dnmt3a and Dnmt3 are also necessary for proper embryonic development. *Dnmt3a* -/- mice survive until shortly after birth, but *Dnmt3b* -/- mice die earlier in embryogenesis, demonstrating that the two *de novo* methyltransferases do not have completely overlapping functions (Okano *et al.*, 1999). A cooperation among all of the methyltransferases has been suggested in repetitive sequences where *Dnmt3a/3b* may compensate for *Dnmt1*, which appears to be less efficient at maintenance in these regions (Liang *et al.*, 2002).

Establishment of allele-specific methylation

For non-imprinted loci, after fertilization, there is a wave of demethylation until the blastocyst stage at implantation, where the two genomes appear to have equally low

levels of DNA methylation. However, imprinted genes are often exempt from following this pattern. Imprinted genes in which allele-specific methylation serves as an imprint, can escape this demethylation and maintain allele-specific methylation at this stage. The unmethylated allele can be protected from *de novo* methylation that follows implantation. Furthermore, while the CpG islands of non-imprinted genes are unmethylated in sperm, in imprinted genes, methylation on the silent paternal allele can originate in the sperm. DNA methylation is suggested to play a role in marking one allele of an imprinted gene either during gametogenesis, or during early embryogenesis, when the two parental genomes are separate. This is usually referred to as the primary imprint. These methylation differences later translate into allele-specific expression. DNA methylation is also thought to be important in maintaining silencing on the inactive allele in somatic cells.

In *Dnmt1* -/- ES cells, expression of the wild type cDNA of *Dnmt1* restores the normal methylation and expression of certain genes and the ability to differentiate. However, for imprinted genes, passage of these ES cells through the germ line is necessary before normal methylation and expression can be restored (Tucker *et al.*, 1996). For many imprinted genes, allele-specific methylation originates in the gametes. This suggests that DNA methyltransferases may act differently in the male and female germ lines. In support of this, a new relative to the Dnmt3a/3b methyltransferases, Dnmt3L, is specifically responsible for the establishment of maternal methylation of a series of imprinted genes (Hata *et al.*, 2002). Interestingly, this methyltransferase, while homologous to Dnmt3a and 3b, does not contain the catalytic domain and has no methyltransferase activity. Its ability to interact with Dnmt3a and Dnmt3b suggests that

it may target these methyltransferases to specific differentially methylated regions (DMRs) during gametogenesis.

Maternal methylation seems to play a much larger role than paternal methylation in genomic imprinting. This comes from the observation that even though a relatively equal proportion of known imprinted genes are repressed on each allele, the majority are methylated on the maternal allele (Reik and Walter, 2001). Many paternally repressed genes are silenced, not by methylation, but by the activity of antisense genes that are active on the paternal allele and repressed on the maternal allele (Reik and Walter, 2001). It has been suggested that the active demethylation of the paternal genome in the zygote (Santos *et al.*, 2002), has lead to this alternate evolutionary strategy (Reik and Walter, 2001).

Some regions of DNA attract or repel DNA methylation

Whether or not a locus becomes methylated may depend on the DNA sequence and the chromatin context. The methylation found in some gene promoters may be caused by spreading of DNA methylation from highly methylated regions, such as transposable elements (Magewu and Jones, 1994; Mummaneni *et al.*, 1995). The p53 gene has a mutation hotspot at codon 248 in exon 7 that is often deaminated at the CpG sites from C to T. This particular site was found to be resistant to demethylation upon treatment with the nucleotide analogue 5-aza-2'-deoxycytidine, that leads to genome-wide hypomethylation, at levels of treatment which caused demethylation in another region of the gene (Magewu and Jones, 1994). Transfection of a DNA segment from this region that included a CpG-rich *Alu* sequence in intron 6 was able to become substantially *de novo* methylated. From this it was suggested that the *Alu* sequence may initiate *de novo*

methylation and act as a methylation centre from which methylation spreads to flanking regions. A common characteristic of imprinted genes is the presence of tandemly repeated sequences, which gives them a resemblance to 'foreign' DNA (Neumann *et al.*, 1995). One theory is that this resemblance to 'foreign' DNA leads to methylation and to inactivation of the imprinted genes in one of the parental gametes. In support of this hypothesis insertion of a defective *intracisternal A-particle* (*IAP*) at an *agouti* locus is able to induce genomic imprinting such that the flanking agouti locus is repressed upon paternal transmission (Morgan *et al.*, 1999). On the other hand, protection from *de novo* methylation can depend on the number of Sp1 binding sites present. For example, a neighbouring methylation centre can inactivate the mouse *aprt* gene when 2 of the 4 Sp1 binding sites are deleted from its promoter (Mummaneni *et al.*, 1995). The imprinted gene *Air/Igf2ras* promoter contains protein-binding elements within the DMR that can promote *de novo* methylation in the female germ cell and protect from methylation in the male germ cell (Birger *et al.*, 1999). Also a region of the imprinting centre for *H19-Igf2* has been shown to contain protein-binding elements that are protective from DNA methylation on the maternal allele (Hori *et al.*, 2002).

DNA methylation and gene expression

In order to understand the role of DNA methylation in the control of imprinted genes, we must first understand how it regulates gene expression in general. DNA methylation is suggested to inactivate genes in two ways: 1) directly blocking binding of transcription factors with CpG sites in their recognition sequence and 2) interaction with CpG binding proteins such as MeCP2 that regulate chromatin structure. In general DNA hypermethylation is correlated with gene repression.

Numerous *in vitro* reporter assays have demonstrated the repressive effect of DNA methylation on transcription (Busslinger *et al.*, 1983; Huq *et al.*, 1997; Keshet *et al.*, 1985). In some cases the inability of transcription factors to bind their recognition sequence when it is methylated was directly shown (Bednarik *et al.*, 1991; Comb and Goodman, 1990; Kanduri *et al.*, 2000; Kovesdi *et al.*, 1987; Prendergast *et al.*, 1991). *In vivo* evidence also supports the link between DNA methylation and gene repression. When Moloney murine leukemia virus (M-MuLV) is integrated into mouse embryos at pre-implantation stage, before *de novo* methylation, it becomes methylated and never expressed, but when integrated after the pre-implantation stage, it is able to be expressed and infect the mice (Jahner *et al.*, 1982). Treatment of embryonic carcinoma (EC) cells with the integrated M-MuLV with 5-azacytidine, a nucleotide analogue and methylation inhibitor, leads to reactivation of the virus (Stewart *et al.*, 1982). There are examples of cancer where loss of expression of tumour suppressor genes is correlated with aberrant hypermethylation of their promoters (Futscher *et al.*, 2002; Myohanen *et al.*, 1998). These are a few examples among a body of evidence that suggests DNA methylation has an important role in gene silencing.

DNA methylation and chromatin structure

In order to really understand the role of DNA methylation in transcriptional control, it needs to be studied in the chromatin context. The link between DNA methylation and chromatin structure was first demonstrated when methylated DNA constructs were integrated into the genome and found to form a closed chromatin structure (Keshet *et al.*, 1986). A more direct link between DNA methylation and chromatin structure was uncovered when a methyl-CpG-binding protein (MeCP2), which

preferentially binds methylated DNA, was shown to recruit histone deacetylase (HDAC) activity (Jones *et al.*, 1998). The number of known methyl-CpG-binding domain proteins (MBD) has expanded to include MBDs 1-4 (Bird and Wolffe, 1999). MBD2 is a component of the MeCP1 complex, and also represses transcription through HDAC activity (Ng *et al.*, 1999). MeCP1, MeCP2 and other MBDs may have specialized roles. For example, MeCP1 prefers regions of dense CpG methylation with 15 or more methylated CpGs (Meehan *et al.*, 1989), whereas MeCP2 can bind a single symmetrically methylated site in 12 nucleotides (Nan *et al.*, 1993). MeCP2 is essential for mammalian development and ES cell differentiation, but not essential for ES cell proliferation (Tate *et al.*, 1996). This is quite similar to the *Dnmt1* deficient mice and demonstrates the cooperation between DNA methylation and MBPs. Interestingly, mutations in the X-linked *MECP2* cause the human disease Rett Syndrome (Dragich *et al.*, 2000), which manifests with a characteristic neural degeneration. It is perplexing that loss of such a general protein would have such a specific effect. A recent examination of MeCP2 protein levels in different tissues suggests that it may normally be more abundant in mature neurons (Shahbazian *et al.*, 2002).

Whether or not DNA methylation is a primary silencing mechanism in many genes is not quite clear. *In vitro* inhibition of transcription by DNA methylation demonstrates the ability of DNA methylation to repress transcription in the skeletal α-actin promoter (Yisraeli *et al.*, 1986). However, DNA methylation of the endogenous skeletal α-actin promoter , and other tissue-specific genes do not seem to correlate with tissue-specific expression *in vivo* (Shani *et al.*, 1984; Walsh and Bestor, 1999; Warnecke and Clark, 1999). In *Dnmt1*-deficient mice some genes previously thought to be regulated

by DNA methylation, did not change in expression, even though they were relatively demethylated compared to normal (Walsh and Bestor, 1999). In some cases DNA demethylation may occur secondary to transcription factor binding or gene expression (Matsuo *et al.*, 1998; Thomassin *et al.*, 2001).

A portion of the hamster *Aprt* CpG island containing 2 Sp1 sites serves to protect a region from the *de novo* methylation that occurs following implantation, and selective removal of the cassette in the adult mouse left the promoter stably protected from DNA methylation (Siegfried *et al.*, 1999). This suggests that the Aprt CpG island element was needed for establishment, but dispensable for maintenance, of a hypomethylated region. Selective removal of this sequence from a transgene before or after implantation allowed the comparison between a methylated and unmethylated promoter respectively. It was found that DNA methylation reduced transcription to varying degrees in different tissues (6-30 fold) compared to the unmethylated transgene (Siegfried *et al.*, 1999). These results suggested that DNA methylation may reduce transcription of genes in the absence of enhancer activity. It has been proposed that DNA methylation may have evolved as a mechanism of reducing background transcriptional activity in non-expressing tissues, without affecting the high level expression in the appropriate expressing cells (Bird, 1992; Bird and Wolffe, 1999). This is supported by findings that DNA methylation alone is often not sufficient to repress transcription in the presence of active enhancers (Boyes and Bird, 1992; Yisraeli *et al.*, 1986).

In cases where DNA methylation is dense, hypermethylation may create a more stable transcriptionally repressed state. A number of studies have analyzed the effects of DNA methylation and histone deacetylation (HDAC) on gene expression by treating cells

with the nucleotide analog 5-aza-2'deoxycytidine (5Aza-dC), which inhibits methylation, and HDAC inhibitors such as Tricostatin A (TSA). For example, TSA alone was unable to activate the tumour suppressor genes *MLH1*, *TIMP3*, *CDKN2B* and *CDKN2A* in a cancer cell line in which they were densely methylated (Cameron *et al.*, 1999). The genes could be activated either by a high dose of 5Aza-dC, or by a low dose of 5Aza-dC, followed by TSA. For these genes it appears that DNA methylation and deacetylation act synergistically, but that demethylation is a necessary first step for "unlocking" the gene for reactivation.

There are tissue-specific genes, which contain particularly dense CpG islands in their 5' regions that appear to use DNA methylation for gene regulation. The MAGE-type genes, of which necdin is a member, are methylated in their 5' ends in all cell types in which they are not expressed (De Smet *et al.*, 1999). The tumour suppressor *SERPINB5* is methylated in non-expressing tissues, aberrantly methylated in tumours, and expression is induced upon treatment with 5Aza-dC (Futscher *et al.*, 2002). This suggests that DNA methylation may be more important for regulation of a specialized groups of genes, such as imprinted genes, and in those with particularly CpG rich promoters.

A series of imprinted genes were tested in androgenetic (Ag) and parthenogenetic (Pg) mouse embryonic fibroblasts (MEFs) when treated with methylation inhibitors and histone deacetylase inhibitors (El Kharroubi *et al.*, 2001). Demethylation alone could reactivate *Peg3*, *Zac1*, $p57^{Kip2}$ and *H19*. DNA demethylation had a heritable effect that lasted several cell cycles after removal of the drug. TSA alone was able to reactivate *Igf2* and $p57^{Kip2}$ but unlike the effect of demethylation on the above genes, the TSA effect was transient and was lost when the drug was removed. A synergistic effect of TSA and

demethylating agents was observed for $p57^{Kip2}$. There were a large group of genes, including *Snrpn*, which were not reactivated by either treatment. These experiments demonstrate the variability in the types of epigenetic modifications that dominate the regulation of different imprinted genes. However, it must be kept in mind that these types of experiments produce a global effect, and epigenetic changes in the surrounding chromatin and or modifier genes may have an effect on whether the gene of interest is reactivated.

Chromatin structure and chromatin domains

One of the earliest measures of chromatin structure was the assessment of nuclease sensitivity, which often used DNase I. This involves the treatment of permeabilized cells or permeabilized nuclei to increasing concentrations of DNase I. The chicken β-globin locus has been extensively studied and was used to demonstrate preferential sensitivity of the active genes to DNase I, by comparing tissues that express and do not express the genes (Weintraub and Groudine, 1976). The size of nuclease accessible domains can be as large as 100 kb as seen in chicken *ovalbumin* (Lawson *et al.*, 1982). For the chicken β-globin locus the entire multigene domain, including intergenic regions, has an increased DNase I sensitivity (Hebbes *et al.*, 1994). The domains of the chicken, human and mouse β-globin loci are bounded by recognition sites for the insulator CTCF (CCCTC-binding factor) (Farrell *et al.*, 2002; Saitoh *et al.*, 2000).

A nuclease sensitive domain is considered to be a domain of loose chromatin conformation, which is accessible to the transcriptional machinery. Within a domain of generalized DNase I sensitivity, DNase I hypersensitive sites (HSS) are often found. These sites are small regions of about 200 bp that are one or two orders of magnitude

more sensitive to cleavage than an area of generalized sensitivity. These areas tend to be in promoters, enhancers or insulator elements. Hypersensitive sites can correspond to the binding of regulatory proteins to DNA (Bell *et al.*, 1999; Boyes and Felsenfeld, 1996; Felsenfeld *et al.*, 1996) or absent or partially disrupted nucleosomes (Bresnick *et al.*, 1992; McGhee *et al.*, 1981). The ability to form an open chromatin structure can also be intrinsic to the DNA sequence. In particular CpG islands tend to form hypersensitive regions characterized by low amounts of histone H1, increased histone acetylation and nucleosome free regions (Tazi and Bird, 1990). Yeast, which normally lack CpG islands and H1 histones, can still form a hypersensitive site when a human YAC with a CpG island is introduced (Mucha *et al.*, 2000).

Changes in chromatin structure are associated with histone modifications. In general, DNase I sensitivity tends to correlate with acetylated histones (Hebbes *et al.*, 1994). The increased histone acetylation is accomplished by transcriptional regulators that attract histone acetylase (HAT) activity (Utley *et al.*, 1998). Histone hyperacetylation is associated with transcriptional competence but does not necessarily depend on actual transcription (Hebbes *et al.*, 1992). On the other hand histone deacetylases (HDACs) are associated with transcriptionally repressive complexes. In contrast to the developmentally regulated β-globin locus which demonstrates large domains of increased nuclease sensitivity and histone hyperacetylation, other types of genes demonstrate a more localized histone hyperacetylation pattern. For example, both viral induction of *IFN-β* and hormone induction of target genes correlate with histone hyperacetylation confined to the promoter region (Chen *et al.*, 1999; Parekh and Maniatis, 1999). The classic

example of facultative heterochromatin, the human inactive X chromosome, is deficient in histone H3 and H4 acetylation (Belyaev *et al.*, 1996; Jeppesen and Turner, 1993).

Histones can also be modified by methylation of lysines 4 and 9 of histone H3 in active and inactive chromatin domains (Noma *et al.*, 2001). On the inactive X chromosome H3-K9 methylation is found at the promoters of inactivated genes, while H3-K4 methylation is associated with the promoters of active genes (Boggs *et al.*, 2002). The term "histone code" has recently been given to describe the gene regulatory information that can be stored in the combination of histone modifications that exists at any given locus (Jenuwein and Allis, 2001).

Like DNA methylation, chromatin structure is dynamic and chromatin remodelling plays an important role in embryogenesis. After fertilization, changes in linker histone variants are thought to be involved in the formation of the transition from the specialized non-proliferative oocyte cell type to a totipotent embryo (Clarke *et al.*, 1998). It is thought that pluripotent ES cells globally have a relatively loose chromatin structure, which changes to a tighter conformation upon differentiation. This comes from the observations that normally silent retroposons are actively transcribed in ES cells and mouse blastocysts (Packer *et al.*, 1993; Tada and Tada, 2001). The developmental control of β-globin genes seems to involve dynamic histone acetylation and deacetylation (Forsberg *et al.*, 2000). The maturation of chromatin structure during development results in a situation where certain genes cannot be transcribed because they are in a chromatin conformation that is not permissive to transcription even if the required transcription factors are present (Thompson *et al.*, 1995). An interesting example of tissue-specific chromatin remodelling is in the specialization of the B lymphocyte lineage, which

involves activation of the Ikaros protein. The Ikaros protein appears to recruit transcriptionally repressed genes to centromeric heterochromatic compartments within the nucleus (Brown *et al.*, 1997).

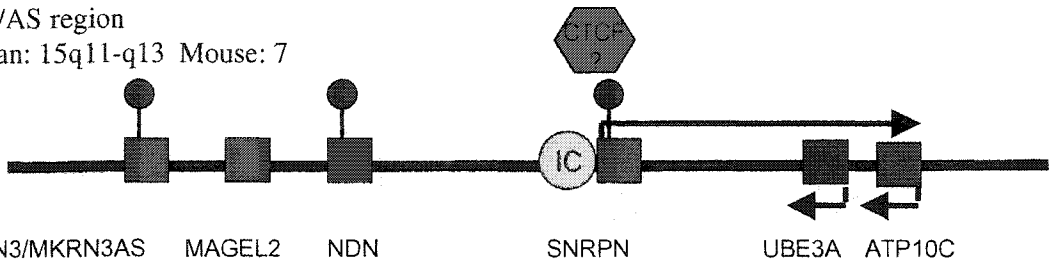
The Genomic Organization and Regulation of Imprinted Genes

A common feature of imprinted genes is that they tend to occur in clusters. Five of the most well studied imprinting clusters are the PWS/AS region, the two clusters in the Beckwith-Wiedemann syndrome Region (BWS), the *DLK-GTL2* gene pair and the *IGF2R* region (Figure 1-3). These clusters all appear to contain an imprinting centre (IC), defined as the minimal region that is necessary for the proper imprinting of all other genes within the imprinted domain. The presence of oppositely imprinted sense and antisense transcripts, allele-specific DNA methylation and putative CTCF binding sites are also shared characteristics between imprinted regions. A common theme in the regulation of imprinted genes that are paternally repressed seems to be the involvement of a paternally expressed antisense transcript. The exception to this rule seems to be in the genes associated with CTCF, like *H19* and *Gtl2*, which show promoter methylation on the paternal allele. On the other hand, imprinted genes that are repressed on the maternal allele more often involve promoter methylation.

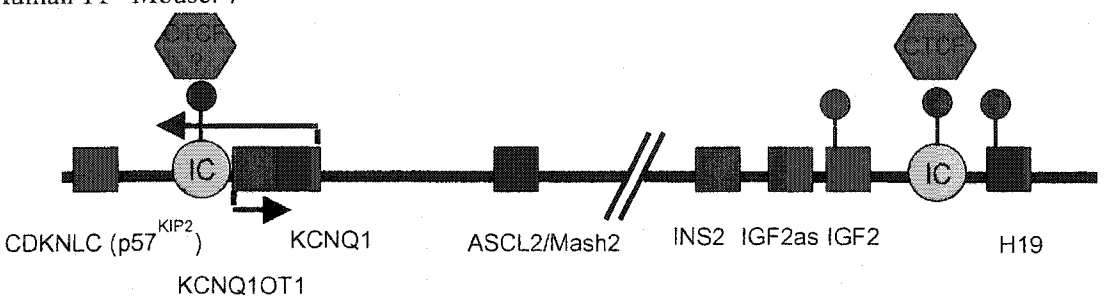
Figure 1-3. Imprinting Clusters.

Paternally expressed transcripts are shown by a blue box and maternally expressed transcripts are shown by a pink box. Blue and pink lollipop symbols represent paternal and maternal methylation, respectively. The transcriptional orientations are shown for those genes with sense and antisense overlapping transcipts. CTCF symbols with question marks are potential, but not proven CTCF binding sites. Not shown in this figure are small nucleolar RNAs which are located between *SNRPN* and *ATP10C* in the PWS region and also downstream of *MEG/Gtl2* in the Human 14/Mouse 12 region .

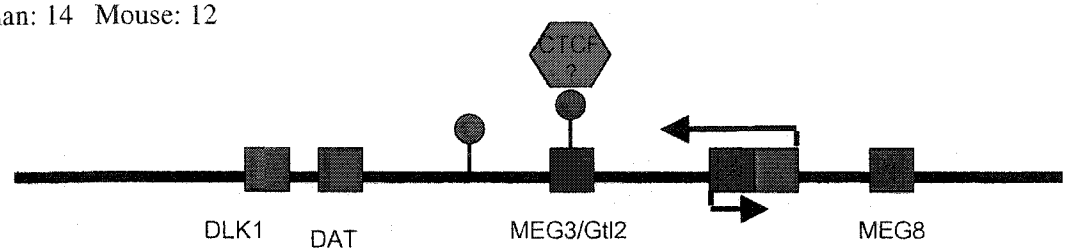
PWS/AS region
Human: 15q11-q13 Mouse: 7



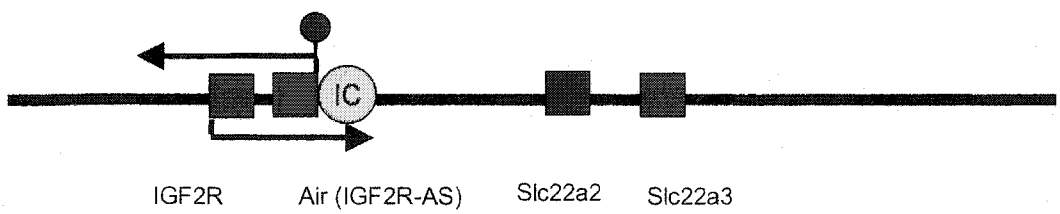
BWS region
Human 11 Mouse: 7



Human: 14 Mouse: 12



Mouse: 17



The insulator CTCF is commonly found at imprinting centres

The most extensively studied imprinted region is the BWS region. One of the IC of the BWS region is an intergenic region between *H19* and *Igf2* (Bartolomei *et al.*, 1993; Leighton *et al.*, 1995). This region is located 2-4 kb upstream of *H19* and is methylated on the paternal allele (Thorvaldsen *et al.*, 1998). *H19* is expressed from the maternal allele and the upstream *Igf2* is oppositely imprinted and expressed from the paternal allele. In the *H19-Igf2* region the dissection of elements necessary for genomic imprinting has been complicated by the fact that *H19* and *Igf2* share a set of enhancers. The IC harbors an insulator element which isolates *Igf2* from the shared enhancers on the maternal allele. The binding of this insulator element by CTCF, a widely used vertebrate enhancer-blocking element (Bell *et al.*, 1999), is allele-specific and regulated by DNA methylation (Bell and Felsenfeld, 2000; Hark *et al.*, 2000). On the methylated paternal allele the absence of CTCF binding allows the enhancers to act on *Igf2*. In DNA methyltransferase deficient mice, both alleles of *Igf2* are silent, while both alleles of *H19* are expressed (Li *et al.*, 1993).

Potential CTCF binding sites have also been found in the *DLK1/GTL2* IC (Wylie *et al.*, 2000), the *Kcnq1* IC (Kanduri *et al.*, 2002; Loukinov *et al.*, 2002) and the PWS IC (Ohta *et al.*, 2001). Interestingly, CTCF also appears to be involved in X-inactivation choice by binding to the unmethylated upstream region of *Tsix* (Chao *et al.*, 2002). Both the *H19-Igf2* and the *Dlk1-Gtl2* CTCF sites are methylated on the paternal allele. A possible important player in the imprinting of these ICs is a male germ line specific protein called BORIS, which binds CTCF recognition sites, and is expressed during a time when CTCF expression is turned off (Loukinov *et al.*, 2002). It is thought that

BORIS-CTCF switching may play a role in setting up the male-germ line derived methylation at the CTCF sites. There must also be specific sequences in the *H19-Igf2* IC that distinguish these imprinted CTCF binding sites from other CTCF general insulators throughout the genome. When the chicken β-globin insulator element, which contains CTCF binding sites was substituted for the *H19-Igf2* IC, the CTCF sites remained unmethylated and displayed full chromatin insulator activity upon both paternal and maternal transmission(Szabo *et al.*, 2002).

Aside from its role as an enhancer-blocking element, CTCF may also play a role in the formation of a transcriptionally repressive domain for *Igf2* and *Ins2* on the maternal allele. Histone deacetylase activity of CTCF has been observed (Lutz *et al.*, 2000) and H4 acetylation is lower on the repressed maternal allele of *Igf2* (Grandjean *et al.*, 2001).

H19-Igf2 in the Beckwith-Wiedemann syndrome region

The upstream *H19* IC is methylated in sperm and not in oocytes, and this allelic methylation is maintained throughout development (Olek and Walter, 1997; Tremblay *et al.*, 1997; Tremblay *et al.*, 1995; Warnecke *et al.*, 1998). Methylation of the *H19* promoter and of *Igf2* is acquired during early development (Brandeis *et al.*, 1993; Tremblay *et al.*, 1997). On the paternal allele, the IC is needed in the zygote for proper paternal allele specific silencing of *H19*, but no longer needed after terminal differentiation (Srivastava *et al.*, 2000). This suggests that a stable epigenetic modification is set up during development on the *H19* gene. This agrees with the findings that the *H19* promoter is not methylated in the gametes, but becomes methylated during early embryogenesis (Tremblay *et al.*, 1997), possibly by methylation spreading from the IC (Olek and Walter, 1997). A part of the *H19-Igf2* IC functions as non-parent of origin

specific silencer in *Drosophila*, an organism that has no DNA methylation (Lyko *et al.*, 1997). Deletion of this silencer element in *H19* reporter transgenes resulted in loss of silencing of the paternal allele, but interestingly, allele-specific methylation at the IC appeared to be established normally (Brenton *et al.*, 1999). These studies demonstrate a separation between silencer and methylation processes in this region. Silencing may involve a combined effect between methylation and chromatin structure. Evidence for this proposal is that the paternal allele-specific silencing of the paternal *H19* allele is mediated by MeCP2 which preferentially binds the methylated allele *in vivo* (Drewell *et al.*, 2002).

DNase I hypersensitive sites are found on both alleles of *Igf2* (Feil *et al.*, 1995; Koide *et al.*, 1994; Sasaki *et al.*, 1992), suggesting that even the inactive maternal allele is in a relatively open chromatin conformation. This lack of allele-specific chromatin structure may explain why *Igf2* requires the *H19-Igf2* IC for continued maintenance of allele-specific expression. The *H19-Igf2* IC and the *H19* promoter show allele-specific chromatin structure (Bartolomei *et al.*, 1993; Khosla *et al.*, 1999). Increased nuclease sensitivity on the maternal allele of *H19* correlated with maternal allele-specific DNase I and dimethylsulfate (DMS) footprints, indicative of protein-DNA interactions (Szabo *et al.*, 1998).

An unusual situation of DNA hypermethylation is found in the upstream *Igf2* (DMR1) region and within the gene (DMR2) on the active paternal allele of *Igf2* (Brandeis *et al.*, 1993; Feil *et al.*, 1994). It has been suggested that these DMRs contain repressor elements that can bind repressor proteins on the hypomethylated allele. For DMR1, it has been shown that DNA methylation abrogates binding of the repressor

protein GCF2 and that deletion of this element in mice causes a 3 fold increase in *Igf2* expression on the repressed maternal allele (Eden *et al.*, 2001). The allele-specific methylation of the *Igf2* DMRs varies greatly between tissues, and correct monoallelic expression still occurs in some tissues where both alleles are hypomethylated (Weber *et al.*, 2001). A model has been proposed for this region where the default state of *Igf2* allows moderate transcription, and a combination of activation on the paternal allele and repression on the inactive allele create the final imprinted state (Eden *et al.*, 2001).

A second imprinting centre in the Beckwith-Wiedemann syndrome region

Another imprinted region is located in the BWS region approximately 300-800 kb telomeric to the *H19-Igf2* region (Figure 1-3). This region is independent of the *H19-Igf2* region (Lee *et al.*, 1999). A DMR located in the promoter region of the paternally expressed *Kcnq1ot1* gene, which is antisense to *Kcnq1*, is methylated on the maternal allele (Cleary *et al.*, 2001; Smilinich *et al.*, 1999). This region constitutes the IC, since deletion of this DMR results in activation of the normally silent paternal alleles of *KCNQ1* and *CDNK1C* (Horike *et al.*, 2000). Like the *H19* IC, the *Kcnq1* IC also acts as a methylation dependent insulator and contains putative CTCF binding sites (Kanduri *et al.*, 2002). One target gene of the *Kcnq1ot1* IC is *Mash2*. *Dnmt1*deficient mice still maintain allele-specific activity of *Mash2*, indicating that methylation-independent mechanisms maintain imprinting of this gene (Tanaka *et al.*, 1999). These authors showed that *Mash2* does contain a CpG island in the coding region, but did not demonstrate that allele-specific methylation is normally present.

The *Igf2r* region

An antisense non-coding transcript to the *Igf2r* gene called *Air* is an imprinting centre for this region on mouse chromosome 17 (Figure 1-3). In mice, the *Igf2r* gene is initially biallellically expressed, but shows paternal allele-specific repression in postimplantation embryos and in adults, correlating with the onset of *Air* transcription on the paternal allele (Sleutels *et al.*, 2002). In humans, *IGF2R* expression is mostly biallelic. In human pre-term embryos, expression is imprinted in some individuals and biallelic in others (Wutz *et al.*, 1998). Unlike the case for *H19*, expression of *Air* is needed for paternal silencing of the genes in this region, even for genes that are upstream of *Air* and would not overlap transcripts (Sleutels *et al.*, 2002). Interestingly *Air* does overlap with the neighbouring gene *Mas1*, a gene which is not imprinted (Lyle *et al.*, 2000). The function of *Air* in the imprinting process is not known, but it may act similar to *Xist* in X-inactivation, which is able to coat and spread along the inactive X chromosome (Brown *et al.*, 1992; Carrel *et al.*, 1996). The function of *Xist* in X-inactivation is thought to be related to its co-localization with the specialized histone macroH2A (Costanzi and Pehrson, 1998).

The *Air* promoter, which also corresponds to an *Igf2r* intron, is methylated on the inactive maternal allele. The maternal methylation at this DMR originates in the oocytes and is maintained throughout development (Stoger *et al.*, 1993). A promoter DMR of *Igf2r* is partially methylated on the repressed paternal allele and this methylation is not found in sperm and does not arise until late in development (Stoger *et al.*, 1993), correlating with the time that *Igf2r* expression becomes imprinted. Within the *Air* DMR, a *de novo* methylation signal has been identified that undergoes *de novo* methylation

when injected into the female pronucleus, but not the male pronucleus (Birger *et al.*, 1999). The lack of *de novo* methylation on the male germ line is due to an allele-discrimination signal, also within the DMR.

The *Dlk1-Glt2* region

The *Dlk1-Gtl2* locus was originally identified in sheep because of the callipyge (muscular hypertrophy in the hindquarters) phenotype associated with a unique imprinted inheritance pattern (Cockett *et al.*, 1996). The organization of this locus is reminiscent of the *H19-Igf2* region in that *Dlk1*, like *Igf2*, is a paternally expressed protein coding gene, while *Gtl2*, like *H19*, is a maternally expressed non coding RNA (Schmidt *et al.*, 2000) (Figure 1-3). At *Dlk1*, a DMR located in a 5' CpG island is more methylated on the expressed paternal allele in most but not all tissues, and methylation does not correlate with expression (Takada *et al.*, 2002), a situation similar to *Igf2*. A more strict hypermethylation on the paternal allele upstream of *Gtl2* was found, which was also present in the sperm in an upstream region, similar to the upstream intergenic DMR of *H19-Igf2*. Another DMR begins at the *Gtl2* promoter and extends into the first intron which contains a CTCF-binding consensus sequence. Unlike the intergenic DMR, DNA methylation in this region is acquired post-fertilization.

The Prader-Willi syndrome region

Similar to the case for the H19-Igf2 IC, a silencing element, able to silence in *Drosophila* has also been identified in the PWS-AS IC (Lyko *et al.*, 1998). The presence of a potential CTCF binding region has also been suggested for the PWS-AS IC (Ohta *et al.*, 2001). Like the *H19-Igf2* IC, the PWS-AS IC also shows allele-specific methylation

that originates in the gametes. The PWS-AS IC is a bipartite structure which was deduced based on microdeletions in patients (smallest region of deletion overlap- SRO) which resulted in either PWS or AS depending on the location (Dittrich *et al.*, 1996). The PWS SRO is co-localized with the *SNRPN* promoter and exon 1, which is the region of maternal allele-specific methylation that originates in the oocyte in mice (Shemer *et al.*, 1997). In humans, allele-specific methylation is acquired post-fertilization at this locus (El-Maarri *et al.*, 2001). Patients with deletions of the PWS SRO are able to switch the maternal epigenotype to paternal in the germ line, but are unable to maintain the paternal epigenotype after fertilization (El-Maarri *et al.*, 2001). This may suggest that the IC on the paternal allele is responsible for protecting this region from *de novo* methylation and chromatin condensation into heterochromatin. The AS SRO co-localizes with alternative upstream exons, 35 kb upstream of the *SNRPN* promoter and exon 1, that are only transcribed from the paternal allele (Dittrich *et al.*, 1996). It is thought that transcription from the AS SRO exons are needed to switch from the paternal to maternal epigenotype in the female germ line. One model suggests that the upstream AS SRO acts as an imprinter and interacts with the PWS SRO, which acts as a site from where the paternal to maternal epigenetic switch initiates (Dittrich *et al.*, 1996). Support for this model of interaction between the PWS and AS SROs comes from AS patients with an inversion in the PWS-AS IC region that places the AS SRO in an inverted orientation and about 1 Mb farther from the PWS SRO (Buiting *et al.*, 2001). As predicted by the model, this inversion is transmitted silently through the paternal germline, indicating that the defect is specific to the paternal to maternal switching.

The PWS-AS region is a domain of paternally expressed transcripts. *UBE3A* and *ATP10C* are thought to be oppositely imprinted in brain only as a consequence of down regulation of the paternally expressed *UBE3A* antisense transcript in this tissue (Chamberlain and Brannan, 2001). In the PWS region, *SNRPN/Snrpn* exon 1 contains a CpG island which displays hypermethylation on the maternal allele (Shemer *et al.*, 1997; Zeschnigk *et al.*, 1997). The equivalent region on the active paternal allele contains prominent nuclease hypersensitive sites. Sites of increased nuclease sensitivity on the maternal allele map to the AS-SRO. Another maternal site of increased nuclease sensitivity correlates with the paternally methylated DMR in intron 7 (Schweizer *et al.*, 1999). Histone acetylation of H3 and H4 is associated with the paternal allele in the exon 1 DMR of *Snrpn* (Gregory *et al.*, 2001). In this study it was shown that the intron 7 DMR is not differentially methylated in liver, but is methylated on both alleles. Paternal allele-specific H4 hyperacetylation was still seen at this site, while H3 was hypoacetylated on both alleles. H3 acetylation correlated well with DNA methylation status, while H4 acetylation correlated with the imprinting status. In the SNRPN exon 1 DMR other histone modifications that correlate with imprinting status have been found (Xin *et al.*, 2001). In this study, H3 lysine 9 methylation was found on the inactive maternal allele and H3 lysine 4 methylation was found on the paternal allele. Histone methylation of K9 is considered to be a mark of facultative heterochromatin (Peters *et al.*, 2002). In yeast, methylation of histone H3 at K4 and K9 mark domains of euchromatin and heterochromatin, respectively (Noma *et al.*, 2001).

Limited analyses of a sense and antisense gene pair in the PWS region, *MKRN3/Mkrn3* (*ZNF127/Zfp127*) and *MKRN3AS/Mkrn3as* (*ZNF127AS/Zfp127as*) have

identified maternal allele-specific methylation within 5' CpG islands. A study in mouse showed that *Mkrn3* is not methylated in oocytes at the single CpG site studied, but that this site became methylated in the zygote (Hershko *et al.*, 1999). The methylation in the zygote was attributed to the maternal allele since androgenetic blastocysts showed no methylation at this site. In adult brain, maternal hypermethylation and paternal hypomethylation was found for human and mouse *MKRN3/Mkrn3*(Jong *et al.*, 1999). In human non-brain tissues, both alleles are hypermethylated even though *MKRN3* is ubiquitously expressed (Jong *et al.*, 1999). The paternal allele-specific hypomethylation correlates with the high expression of *MKRN3/Mkrn3* in the brain. For *MKRN3/Mkrn3* the CpG island is in the 5' region, but in *MKRN3AS/Mkrn3as* this CpG island is within an exon. Analysis of this locus for allele-specific marks is complicated by the presence of sense and antisense genes with different patterns of expression.

Necdin, encoding *NDN/Ndn*, is another gene in the region that contains a 5' CpG island in which allele-specific methylation has been found (Jay *et al.*, 1997; Watrin *et al.*, 1997). While these studies found methylation on the maternal allele, only a single CpG site was assayed, making interpretations limited. H3 K4 methylation was also found in the *NDN* 5' region (Xin *et al.*, 2001). I have performed a detailed analysis of the allele-specific epigenetic modifications at *NDN/Ndn*, including DNA methylation, chromatin structure and protein-DNA interactions. The results of this analysis will be discussed in Chapters 3 and 4.

Parallels with X-inactivation

In general X-inactivation is a good model for epigenetic gene regulation and many aspects of inactivation of genes on the inactive X chromosome parallel imprinted

gene regulation. An X-inactivation centre controls X-inactivation, and is able to inactivate translocated autosomal genes (Brown *et al.*, 1991). *Xist*, which is essential for initiation and establishment of X-inactivation, is not required for maintenance (Brown and Willard, 1994). X-inactivation choice and the expression of *Xist* seem to be regulated by an antisense transcript *Tsix*, which contains CTCF binding sites in its 5' region (Chao *et al.*, 2002). Expression of Xist also appears to be regulated by DNA methylation (Beard *et al.*, 1995; Norris *et al.*, 1994; Panning and Jaenisch, 1996). Stabilization of the *Xist* expression on the future inactive allele appears to set off a series of progressive changes on the inactive X including late replication, gene silencing, histone H4 deacetylation and DNA hypermethylation of X-linked genes in that order (Keohane *et al.*, 1998). Therefore for the X-inactivation process, histone acetylation and DNA methylation appear to have a role in stabilization of the inactive state, rather than in initiating silencing of genes on the inactive X chromosome. Recently, histone methylation of H3-K9 was shown to be a very early event in the process of X inactivation, which closely parallels the onset of *Xist* RNA accumulation (Mermoud *et al.*, 2002). These studies demonstrate the complex layering of epigenetic marks that differentiate the active and inactive X chromosomes, and give insight into the epigenetic characteristics of heterochromatin and euchromatin. The parallels between epigenetic characteristics of genes on the inactive X and genes that are inactivated by genomic imprinting suggest that they share at least some of the same mechanisms.

Necdin is a model target gene of the PWS imprinting centre

The two most studied imprinted loci, *SNRPN* and *H19/IGF2,* are co-localized with their respective IC. The epigenetic marks in these regions may be important in the

regulation of the IC itself, or of the co-localized genes, but it is difficult to distinguish these two possibilities. Imprinted genes are often found in clusters, where one or more IC appears to initiate and maintain allele-specific expression of all the imprinted genes within the domain. The imprinted genes that are not part of the IC can be thought of as "target genes" of the IC. The epigenetic characterization of genes in the cluster that are distant from the IC is important to the understanding of how imprinted genes are marked. In turn this will lead to an understanding of how these allele-specific epigenetic marks may cause allele-specific expression. *NDN/Ndn*, located about 1 Mb from the PWS IC, is a good model target gene of the PWS IC. *NDN/Ndn*, encoding necdin, is a small intronless gene of 1.8 kb (Jay *et al.*, 1997; MacDonald and Wevrick, 1997). The low complexity of the gene structure of *NDN/Ndn* makes it very amenable to a thorough characterization of allele-specific epigenetic marks. Unlike *SNRPN*, *NDN/Ndn* is not ubiquitously expressed, which allows comparisons between epigenetic modifications in expressing and non-expressing tissues. This allows a possible dissection of epigenetic marks that are related to gene expression from those that are more permanent allele-discriminating marks.

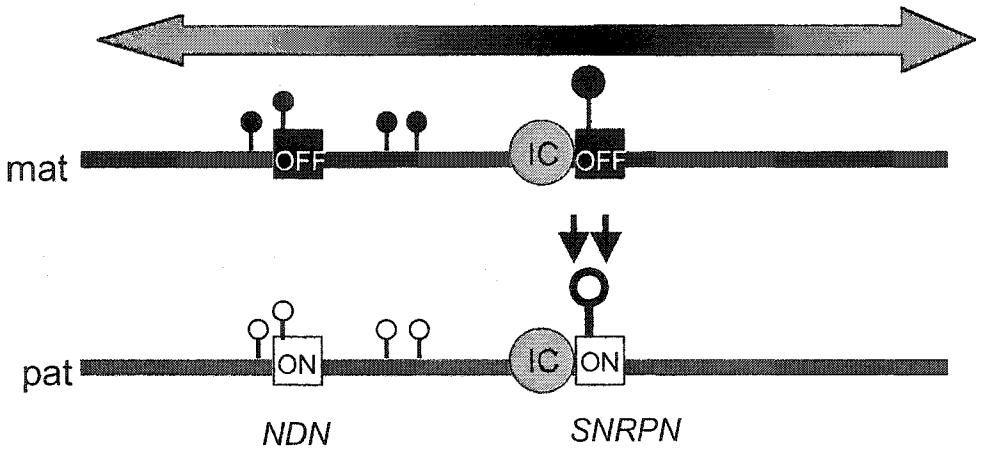
On the maternal allele, the PWS-AS IC is responsible for imprint switching between the maternal and paternal epigenotypes for the PWS-AS region. Like other IC loci, the IC-*SNRPN* locus appears to be regulated by DNA methylation and/or chromatin structural differences thought to be established in the germ line or early zygote. The IC sets up a maternal and paternal epigenotype that results in transcription of the PWS genes only on the paternal allele. It is not known how the IC exerts its effect on distant target genes such as *NDN/Ndn*. Possible mechanisms are shown in Figure 1-4. The experiments

in this thesis were aimed at determining what constitutes the maternal and paternal epigenotypes at the *NDN/Ndn* locus as a first step toward understanding how imprinted target genes are regulated in an allele-specific manner.

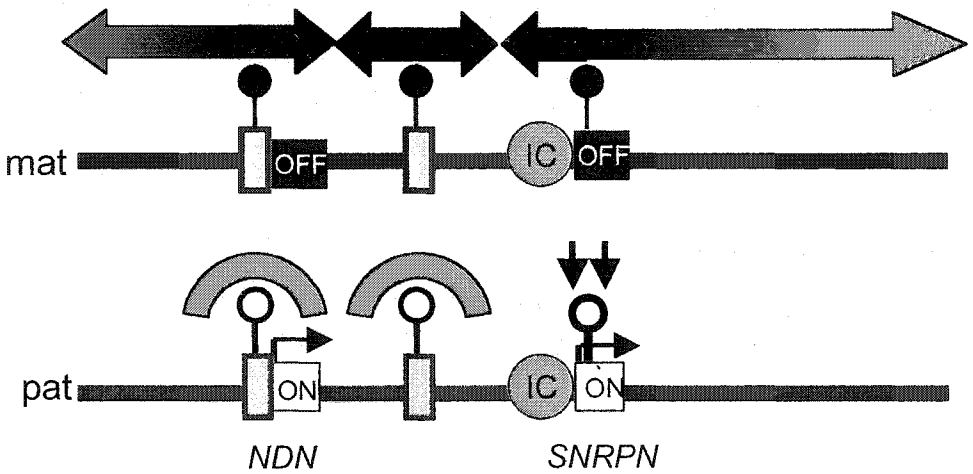
Figure 1-4. Models for targeting of necdin for imprinting by the PWS IC.

The Imprinting Centre (IC), at the 5' end of *SNRPN* is associated with maternal allele specific methylation (filled lollipop) and paternal allele specific sites of nuclease hypersensitivity (vertical arrows). A) Heterochromatin (arrows) spreads from the IC on the maternal allele for a finite distance, reaching neighbouring target genes. DNA methylation outside of the IC is secondary to heterochromatin spreading and may function to reinforce the closed chromatin structure. On the paternal allele heterochromatin spreading is blocked. B) Multiple imprinting response elements (yellow boxes) become methylated and nucleate the formation of heterochromatin and facilitate spreading of a closed chromatin state on the maternal allele. On the paternal allele the unmethylated IC may act as a locus control region and promote a domain of open chromatin structure. Unmethylated imprinting response elements may bind regulatory proteins that block DNA methylation and create a transcriptionally permissive chromatin state at the *NDN* locus.

A



B



Thesis objectives

DNA methylation can be stably maintained, yet can be reversible, and generally has a repressive effect on gene expression. At the time that I began my project, a few imprinted genes had been demonstrated to be allelically marked by DNA methylation in the gametes, early embryo and somatic tissues. Allele-specific DNA methylation was the prime candidate for the mechanism of establishment and maintenance of genomic imprinting. My hypothesis was that DNA methylation is important in the establishment and maintenance of silencing of the maternal allele of *NDN/Ndn*. If this is true, the maternal allele should be more methylated than the paternal allele. A limited analysis of DNA methylation for human and mouse *NDN/Ndn* had demonstrated maternal-allele-specific methylation at one CpG site (Jay *et al.*, 1997; Watrin *et al.*, 1997). I set out to perform a more detailed analysis of the DNA methylation patterns at *NDN/Ndn* in somatic cells. To test whether DNA methylation played a role in the establishment of the imprint on *Ndn*, I analyzed DNA methylation in the gametes and early embryo.

The known epigenetic cross-talk between DNA methylation and chromatin structure lead to my second hypothesis, that other epigenetic modifications, in addition to DNA methylation contribute to the establishment and/or maintenance of the allele-specific expression of *NDN/Ndn*. For example, DNA hypermethylation on the maternal allele could be the cause or the effect of a closed chromatin structure. In addition, the unmethylated paternal allele could contain epigenetic modifications that promote active gene expression such as an open chromatin structure and preferential accessibility to transcription factors. I used DNase I sensitivity mapping and *in vivo* footprinting to

determine whether allele-specific chromatin structure and DNA-protein interactions were allele-specific and to determine how they correlated with DNA methylation.

In the following chapters I demonstrate that allele-specific DNA methylation, chromatin structure and transcription factor binding contribute to the maternal and paternal epigenotypes. How these epigenetic marks may contribute to the establishment and maintenance of allele-specific expression of *NDN/Ndn* will be discussed.

Chapter 2. Materials and Methods

Genomic Structure Analysis

The mouse sequence (GenBank accession AC026388) corresponds to the working draft sequence of *Mus musculus* chromosome 7 clone RP23-426B15. The human sequence (GenBank accession AC006596) corresponds to the complete sequence of human chromosome 15 PAC clone pDJ181P7. Repeatmasker (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker) was used to annotate the sequences for murine and human repeats, respectively. The PipMaker gene analysis program (http://nog.cse.psu.edu/pipmaker/) (Schwartz *et al.*, 2000) was used to compare the sequences of the genomic regions surrounding *Ndn* and *NDN* and to detect regions of high CpG content.

Tissue collection and cell culture

All animal use was approved by the University of Alberta Health Sciences Animal Welfare Committee. *Mus musculus* (C57BL/6), *Mus spretus* (SPRET) and *Mus castaneous* (CAST) mice were from Jackson Labs. Brain, liver and heart were dissected from C57BL/6 x SPRET F1 or C57BL/6 x CAST F1 adult mice at age 6 weeks to 1 year old. Dissected tissues were placed directly into liquid nitrogen and then stored at –80°C until used.

Early embryos and gametes were collected from C57BL/6 mice essentially as described (Hogan *et al.*, 1994). Some of the blastocysts were derived from crosses between C57BL/6 females and SPRET males. Five week-old C57BL/6 females were superovulated with pregnant mare's serum (PMS) and human chorionic gonadotropin (hCG) and mated with 2-7 month old males. Oocytes, 2 cell, 4 cell, 8 cell embryos and

morulae were collected from the oviducts by flushing out through the infundibulum. Oocytes were washed with careful inspection to remove maternal cells. Two, four and eight cell embryos were collected around 32 hours, 38 hours and 56 hours post-coitum (p.c.) respectively. Morulae were collected at about 62 hours p.c. Blastocysts were collected either around 3.5 days p.c. by flushing out the uterus or were collected at the morula stage (2.5-3 days) and cultured for 24 hours in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Sperm was collected from the epididymus of 5 month old males.

Control fibroblasts (GM00650), PWS fibroblasts (University of Miami Brain and Tissue Bank for Developmental Disorders #1889), and AS fibroblasts (deletion line, a generous gift of Dr. Art Beaudet, Baylor College of Medicine) were grown in DMEM supplemented with 10% FBS. Control lymphoblasts (LCL10), PWS lymphoblasts (GM09024B, GM09133) and AS lymphoblasts (GM11515) were grown in RPMI (Roswell Park Memorial Institute) supplemented with 15% fetal bovine serum (FBS). Brain samples were obtained from autopsy materials from the University of Miami Brain and Tissue Bank for Developmental Disorders. Blood was collected from PWS and AS patients and control individuals with informed consent.

DNA extraction

Human and mouse tissues were crushed under liquid nitrogen and DNA was extracted by proteinase K/SDS digestion, phenol/chloroform extraction and ethanol precipitation (Ausubel *et al.*, 1993). Tissue culture cells were pelleted by centrifugation for 5 min. at 500 x g. 1 ml of digestion buffer (same as above) was added for every 10^8 cells and samples were incubated overnight at 50°C. Phenol/chloroform extractions and

DNA precipitation was done as above for tissues. For lymphocytes, red blood cells were lysed by adding 4 volumes of red blood cell lysis buffer (1 mM EDTA, 10 mM $KHCO_3$, 155 mM NH_4Cl, pH 7.4) and rotating for 1 hour at 4°C. Cells were spun down and red blood cell lysate was removed. Cells were resuspended in the red blood lysis cell buffer and rotated for 20 min. at 4°C. Cells were pelleted and washed with phosphate buffered saline (PBS). Genomic DNA was extracted as above for tissue culture cells.

To extract DNA from the oocytes and embryos, 30-200 cells were suspended in ~3 µl of DMEM. Approximately 40 oocytes and 3-37 early embryos were combined for each DNA preparation with the number of embryos pooled being inversely proportional to the age of the embryo because total cell number increases with age. The cell suspension was made up to 18 µl with a 1 mM (0.03%) SDS and 280 µg/ml Proteinase K solution containing 1 or 2 µg salmon sperm carrier DNA in phosphate-buffered saline. This mixture was covered in mineral oil and incubated at 37°C or 50°C for 30-90 min., then at 98°C for 15 min.

Sperm was first treated with 10 mM EDTA, 100 mM NaCl, 2% SDS, 20 µg/ml proteinase K and 10 mM Tris-Cl, pH 8 overnight at 37°C to digest the non-sperm cells. The sperm sample was centrifuged at 600 x g for 10 min. at room temperature and the supernatant was removed. Sperm was then processed as for tissues with the addition of 39 mM DTT to the proteinase K/SDS extraction buffer.

DNA sequencing for detection of mouse Ndn polymorphisms

Polymorphisms between C57BL/6, SPRET and CAST were determined by cycle sequencing PCR products with ^{33}P labelled nucleotide terminators using the Thermo Sequenase Radiolabeled Terminator Cycle Sequencing Kit (United States Biochemical).

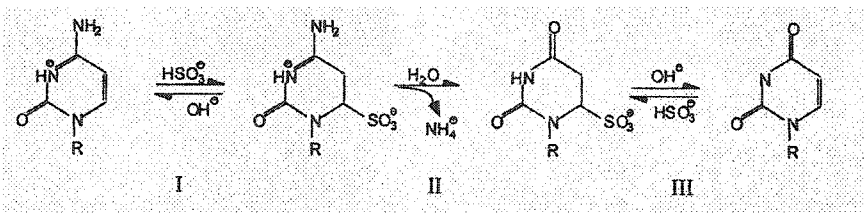
The 5' region of *Ndn* was PCR amplified on C57BL/6 and SPRET DNA with primers NEC11F 5' TCATTCTCCAGGACCTTCAC and NEC12R 5' CTTCGGATCAGAGCAGGAC. PCR was performed in 20 µl reactions with 1.5 mM $MgCl_2$, 1 x PCR buffer (Invitrogen) and 0.5 U Taq polymerase (Invitrogen) as follows: 5 min. at 94°C, (30 s at 94°C, 30 s at 50°C, 30 s at 72°C) cycled 30 times and 10 min. at 72°C, yielding a 401 bp PCR product. Before sequencing, 5-10 µl of each 20 µl PCR product was treated with 10 U of Exonuclease 1 (USB) and 1 U of Shrimp Alkaline Phosphatase (Amersham) and incubated at 37°C for 20 min., followed by heat inactivation for 10 min. at 70°C. This treated PCR product was added to the sequencing reaction. Sequencing was done with the same forward and reverse primers used for the PCR reaction. Reactions were run on a 5% denaturing acrylamide gel made with Long Ranger™ acrylamide (BioProducts). Gels were exposed to film (Kodak) overnight at –80°C.

Sodium bisulfite sequencing

A diagram of the chemical modification and a schematic of the proceedure for sodium bisulfite sequencing are shown in Figure 2-1. This protocol was modified from Stoger *et al.* (1997) and Tremblay (1998). Early embryo and oocyte DNA samples were denatured by addition of fresh 3M NaOH to a final concentration of 0.3 M, incubated for 20-30 min. at 42°C, followed by 3 min. at 95°C and then placed on ice. DNA from adult tissues or sperm (200 ng) was added to 2 µg of salmon sperm carrier and denatured in 0.3M NaOH at 42°C for 30 min. then 95°C for 3 min. and placed on ice. One or two µg salmon sperm carrier DNA was added to early embryo samples and oocytes samples. To all samples, 255 µl 40.5% (w/v) sodium bisulfite at pH 5, 15 µl 20 mM hydroquinone

and milli-Q-H_2O (mH_2O) to 300 µl were added. The samples were covered in mineral oil and incubated at 55°C for ~16 hours. For purification, the treated DNA was mixed with 100 µl of a purification buffer containing 50 mM KCl, 10 mM Tris-HCl (pH 8.8), 1.5 mM $MgCl_2$ and 0.1% Triton X-100, followed by 1 ml of a resin containing 4 M guanidine thiocyanate, 20 mM EDTA and 20 g/L diatomaceous earth (Sigma D-5384), 50 mM Tris-Cl, pH 7.5. Samples were centrifuged through a Wizard Miniprep column (Promega, Inc.) and eluted in 50 µl mH_2O. To complete the reaction, fresh 3 M NaOH to 0.3 M was added and the samples were incubated at 37°C for 20 min. and placed on ice. The DNA was precipitated in 3 M ammonium acetate and 3 volumes 95% ethanol at 4°C for 15 min. to 1 hour or at –20°C overnight, then centrifuged at 4°C at 12,000 x g in a microcentrifuge for 10 min. DNA pellets were washed twice with 70% ethanol and resuspended in 30 µl mH_2O. For early oocytes and early embryos, an additional 1 or 2 µg of salmon sperm carrier DNA was added at the ethanol precipitation step. Four µl of the sodium bisulfite treated DNA was added to a 20 µl PCR reaction.

A



B

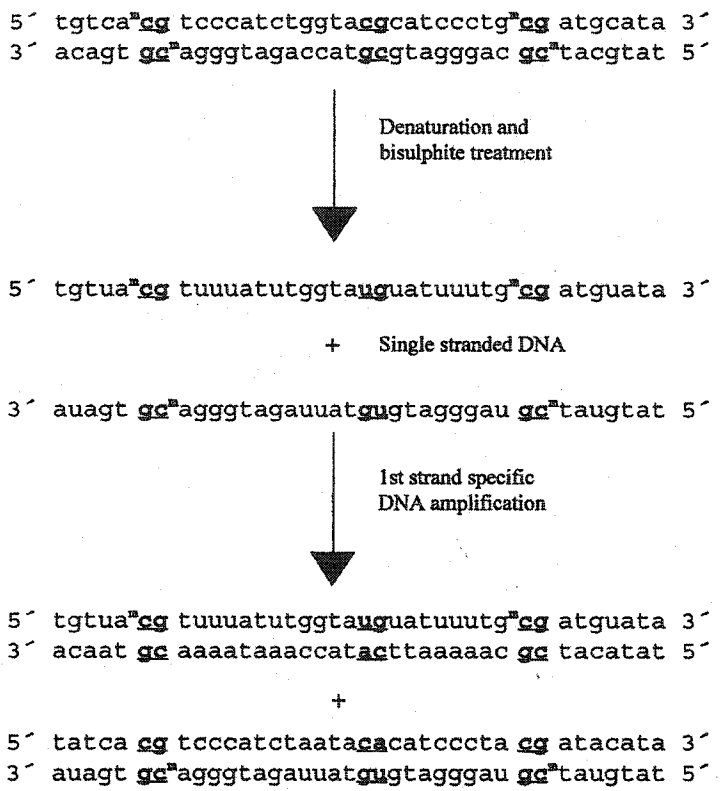


Figure 2-1. An overview of sodium bisulfite sequencing.

Chemistry of sulfonation at position C6 of cytosine, (II) irreversible deamination at position C4, generating 5-sulfonate-uracil, and (III) subsequent desulfonation under alkaline conditions. Methylation at position C5 impedes sulfonation at position C6. (B) Schematic of sodium bisulfite conversion of a DNA sequence. Methylated cytosine are protected from conversion. In subsequent PCR amplification cycles uracil will be replaced by thymine. These figures were taken from Engemann *et al.* Methods in Molecular Biology, *181*: Genomic Imprinting: Methods and Protocols.

Sodium bisulfite PCR on mouse DNA

Primers were designed based on DNA in which all cytosines are converted to thymines and no primers contained CpG sites (See Table 2-1 for primer sequences). Nested PCR was used to amplify a 560 bp *Ndn* product from sodium bisulfite treated DNA. For the majority of clones, the first round PCR forward primer was NEC43F. The reverse primer was NEC45R. For some samples the first round PCR was done with the forward primer NEC78F and the reverse primer NEC79R. Second round nested primers were NEC41F and NEC48R. All PCR reactions were done with Taq (Invitrogen) or Platinum Taq (Invitrogen) and the supplied PCR buffer. First round PCR was performed in 1.5 mM $MgCl_2$ as follows: 5 min. at 94°C, (2 min. at 94°C, 2 min. at 58°C, 2 min. at 72°C) cycled 2 times, (30s at 94°C, 30 s at 58°C, 1 min. at 72°C) cycled 35 times and 10 min. at 72°C. Second round PCR was performed in 1.5 mM $MgCl_2$ as follows: 5 min. at 94°C, (30 s at 94°C, 30 s at 58°C, 1 min. at 72°C) cycled 35 times then 10 min. at 72°C. For some of the samples, a semi-nested PCR reaction was used with the above reaction conditions. The first and second round PCR forward primer was NEC53F and the reverse primers were NEC47R for first round and NEC48R for second round resulting in a 709 bp product. In some semi-nested PCR reactions the first and second round forward PCR primer was NEC62F resulting in a 615 bp product. Control primers amplified a 544 bp product from *H19* region B, a CpG-rich located 2.6 kb upstream of the transcription start site (Warnecke *et al.*, 1998). The first round PCR was performed in 2 mM $MgCl_2$ as follows: 5 min. at 94°C, (30 s at 94°C, 30 s at 55°C, 1 min. at 72°C) cycled 35 times then 10 min. at 72°C. Second round PCR conditions was performed in 2 mM $MgCl_2$ as

follows: 5 min. at 94°C, (30 s at 94°C, 30 s at 58°C, 1 min. at 72°C) cycled 35 times then 10 min. at 72°C.

Sodium bisulfite PCR on human DNA

For the 5' CpG island the first round PCR primers were NEC63F and NEC65R (see Table 2-1 for primer sequences). PCR was performed in 1.5 mM $MgCl_2$ as follows: 5 min. at 94°C, (30 s at 94°C, 30 s at 56°C, 1 min. at 72°C) cycled 35 times, yielding a 842 bp product. 1 µl of the first round PCR was used in the seminested amplification, with PCR primers NEC63F and NEC64R. PCR was performed in 2 mM $MgCl_2$ as follows: 5 min. at 94°C, (30 s at 94°C, 30 s at 56°C, 1 min. at 72°C) cycled 35 times. The final PCR product was 809 bp. Some of the reactions were done as a nested PCR reaction with the first round PCR primers NEC92F and NEC65R. The second round primers were NEC95F and NEC64R. Both rounds were performed at 2 mM $MgCl_2$ as follows: 5 min. at 94C, (30 s at 94°C, 30 s at 57°C, 1 min. at 72°C) cycled 35 times. The sizes of the PCR products for the first and second round amplification were 855 bp and 822 bp, respectively.

For the 3' CpG island the first round amplification primers were NEC126F and NEC128R. PCR was performed in 1.5 mM $MgCl_2$ as follows: 5 min. at 94°C, (30 s at 94°C, 30 s at 58°C, 1 min. at 72°C) cycled 35 times. 1 µl from the first round was added to the second round semi-nested amplification. The second round amplification primers were NEC127F and NEC128R. The PCR was performed as above.

Table 2-1. List of sodium bisulfite primers.

NEC43F	TTTTGTGTTATATAGGAGATTAGGAAATT T/G TTTATA
NEC45R	TCTAACCTACTCCAAAACCTCCCTATATC
NEC78F	TATTTAGTTTTGTGTTATATAGGAGATTAGG
NEC79R	ATTTCTTATAACTACCCATAACCTCTTTCA
NEC41F	TTTTTTAGATTTTAGTGGTTGGGTTTTG
NEC48R	CACCTTCTACACCAACTAAACAAAAGT
NEC53F	ATATTTAATTTGATTTTGTTTAAATTTAGTGTG
NEC47R	CATTCCAAACCACACCCTCTC
NEC62F	TTATTTAGTTTTGTGTTATATAGGAGATTAGGG
NEC63F	TATATTTAAATTAATTTTGGATATATTTAGGTAAG
NEC65R	TCCTACACCACTTCTTATAACTAC
NEC64R	CTTTCACCATATCTAAAAACCAAATAATC
NEC92F	GTTTGTATAAAGTTTATGTTTATAAAGTTGTTTAG
NEC95F	TTTGGGATTTTTATATTTAATTTAATTTTGGATA
NEC126F	GAAGGTAGAAAAATAATGGAAGTAAAAATTTG
NEC128R	CTTCACCCATACTCATCATCACTA
NEC127F	GAGTTTTTTTTTGTTATTTAGGTTGGAG

T/G indicates that the primer was degenerate at this position, where T represents the C57BL/6 sequence and G represents the SPRET sequence.

Cloning and sequencing of sodium bisulfite samples

For sodium bisulfite sequencing samples, PCR products were electrophoresed on 2% agarose gels, gel-purified using the QIAquick Gel Extraction Kit (Qiagen Inc.) and cloned into the pGEM–T vector (Promega Corp.). Recombinant clones were sequenced by cycle sequencing using an Amersham kit with fluorescently labeled M13 forward and reverse primers, and analyzed on a LiCor automated sequencer.

Nuclease sensitivity assay

For the DNase I analysis restriction site polymorphisms between C57BL/6 and SPRET were determined by restriction enzyme digestion, Southern blot analysis and hybridization with probes in the regions of interest (see generation of DNase I probes below).

Isolation of Nuclei

Nuclei were isolated from C57BL/6 x SPRET F1 mice as follows. Two adult mouse brains or one adult mouse liver was used for five DNase I concentrations. Tissues frozen in liquid nitrogen were crushed and one adult mouse brain or one half of an adult mouse liver was placed in each dounce homogenizer and homogenized on ice. I first used the looser-fitting pestle followed by the tighter-fitting pestle in 6 ml of 0.3 M sucrose nuclear buffer (Wu, 1989) with 3.6 ng/ml aprotinin (Sigma). PMSF was added fresh to 0.1 mM. The homogenate was filtered through 8 layers of pre-wet cheesecloth and then centrifuged at 6000 x g for 10 min. at 4°C. The supernatant was removed and the pellet was resuspended in 2.5 ml of 0.3 M sucrose nuclear buffer with 0.1 mM PMSF (Sigma) and 0.2% IGEPAL (Sigma) and put on ice for 5 min. to lyse any remaining cells. The cell

lysate was placed over a 6 ml cushion of 1.2 M sucrose nuclear buffer with 3.6 ng/ml aprotinin and centrifuged at 10,000 x g for 20 min. at 4 C. The supernatant was removed by aspiration and the nuclei were resuspended in DNase I buffer (0.1 M NaOAc, 5 mM $MgCl_2$) to a concentration of 5-10 x 10^6 nuclei/ml.

DNase I treatment

Aliquots (200 µl) of nuclei suspensions were placed in 2 ml microfuge tubes. 2 ml microfuge tubes were preferable to 1.5 ml microfuge tubes because they have a wider bottom. 0 to 1000 U/ml DNase I (Gibco) was added to the samples and then they were incubated at 25°C for 10 min. The reactions were terminated by the addition of an equal volume of 20 mM EDTA (pH 8), 1% SDS solution. Proteinase K was added to 200 µg/ml and digested overnight at 50°C.

After the overnight incubation, an equal volume of mH_2O was added to dilute samples if they were viscous. Genomic DNA was extracted 2 times with phenol/chloroform, and ethanol precipitated as above for isolation of DNA. DNA was resuspended in mH_2O and 10-20 µg of DNA was digested to completion with a restriction enzyme with sites that flank the region of interest. Digested DNA was electrophoresed on a 0.8% agarose gel and blotted to a Hybond-N nylon membrane (Amersham) using standard Southern blotting procedures (Sambrook and Russell, 2001).

For tissue culture cells, cells were permeabilized with IGEPAL (Sigma) and then treated with DNase I (Gibco). Fibroblasts were grown just to confluency in T150 flasks. For FB17 (AS) cells and FB16 (PW) cells, one confluent flask had about 2 x 10^6 cells and 1 x 10^6 cells respectively. For four DNase I treatments, four T150 flasks of FB17 or seven T150 flasks of FB16 cells were used. Cells were trypsinized, combined and

centrifuged at 200 x G for 5 min. The cell pellet was resuspended in phosphate buffered saline (PBS) and aliquoted equally into four 15 ml conical tubes. Cells were centrifuged again as above and resuspended in 1 ml of IGEPAL (NP40) buffer (15 mM Tris-HCl, pH 7.5, 60 mM KCl, 15 mM NaCl, 5 mM $MgCl_2$, 0.5 mM EGTA, 300 mM sucrose, 0.5 mM β-mercaptoethanol, 0.2% IGEPAL) (Aasland and Stewart, 1999). Then 1 ml of IGEPAL buffer was added with the appropriate amount of DNase I to a final concentration of 0 to 500 U/ml and the reaction proceeded at room temperature for 4 min. The reaction was stopped with the addition of 400 µl of 6x lysis buffer (6% SDS, 300 mM Tris-HCl (pH8), 120 mM EDTA). Proteinase K was added to 400 µg/ml and samples were incubated overnight at 50°C. DNA was isolated and processed as above.

Generation of DNase I probes and hybridization

Five probes were used for the analysis of the mouse *Ndn* region (see Figure 4-1 for map of region). The *Ndn* open reading frame probe was produced by digesting a *Ndn* cDNA clone with *Nde* I and *Eag* I, isolating the band from an agarose gel and gel-purified using the QIAquick Gel Extraction Kit (Qiagen Inc.). All other probes were generated by PCR amplification from genomic DNA. The sequences for the primers used to amplify probes for DNase I analysis are in Table 2-2. PCR conditions for each primer pair are as follows. Probe1: 2 mM $MgCl_2$, 5 min. at 94°C, (30 s at 94°C, 30 s at 57°C, 30 s at 72°C) cycled 30 times. Probe 2: 1.2 mM $MgCl_2$, 5 min. at 94°C, (30 s at 94°C, 30 s at 55°C, 30 s at 72°C) cycled 30 times. Probe 3: 1.5 mM $MgCl_2$, 5 min. at 94°C, (30 s at 94°C, 30 s at 56°C, 30 s at 72°C) cycled 30 times.

For human DNase I, primers amplified a probe (H1) in the 5' region starting at position –1335 from the transcription start. The primers were NEC145F and NEC146R.

The PCR was performed at 1.5 mM $MgCl_2$ as follows: 5 min. at 5 min. at 94°C, (30 s at 94°C, 30 s at 55°C, 30 s at 72°C) cycled 30 times.

Table 2-2. List of primers for the generation of DNase I probes.

Probe	Primer		Product Size (bp)
1	NEC74F	GGAGAGTTCGGTGATTCCAA	565
	NEC75R	CATCCCTATTCACAGGCAGA	
2	NEC114F	TAATCTGGGTTTGGCCATTG	606
	NEC115R	TTTCACCTTGAGTGGCTTCC	
3	NEC72F	TTTCTCCAGCTTCACATTT	620
	NEC73R	CAACATCTTCTATCCGTTCTTCG	
4	NEC76F	CGGTTTTTGCCTTTGGAAG	520
	NEC77R	TGAATGCAAAATCCCTTTCC	
H1	NEC145F	AATGGCTCTGCCAAAAGTG	501
	NEC146R	CCACACCCAAGTCTTTCCTG	

PCR products were gel purified as above and cloned into the pGEM–T vector (Promega Corp.). Cloned probes were then PCR amplified and purified as above. Purified probes were labeled by incorporation of ^{32}P dCTP using the Random Primers DNA Labeling System (Invitrogen). Hybridization of Southern blots were performed overnight at 65°C with Church buffer (Molecular Cloning). Blots were washed repeatedly as needed at 65°C with less stringent (2x SSC/0.1% SDS) to more stringent (0.1x SSC/0.1% SDS) conditions.

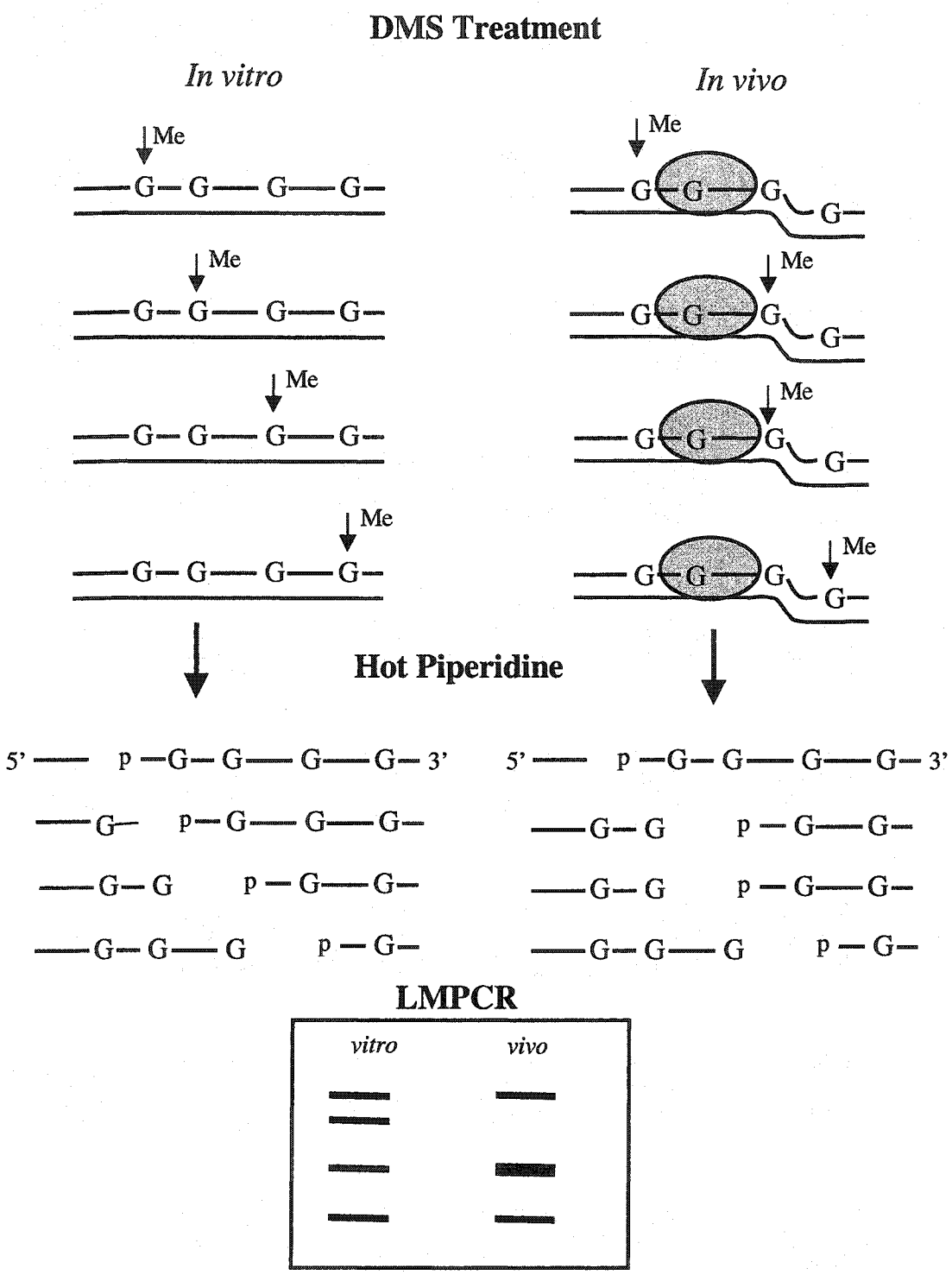
In Vivo DNA footprint analysis

In vivo DNA footprint techniques were performed essentially as described (Drouin *et al.*, 2001). These experiments were performed by me in the laboratory of Dr.

Régen Drouin with technical assistance from Isabelle Paradis, a technician in his laboratory (Université Laval). For each treatment, purified DNA from LCL3 cells was prepared as described (Drouin *et al.*, 2001) and then treated by the same footprint techniques as those used on the *in vivo* samples. This was used as the *in vitro* control DNA for comparison to *in vivo* samples. An example of *in vivo* footprinting using dimethylsulfate (DMS) is shown in Figure 2-2. An overview of the *in vivo* footprinting technique using LMPCR is shown in Figure 2-3.

Figure 2-2. Example of *in vivo* footprinting using DMS treatment.

Vertical arrows indicate the methylation of guanine residues (G) caused by DMS. The G that contacts the bound protein is protected from modification, but the next G residue is more accessible to modification because of the chromatin structure.The methylated G residues are cleaved by hot piperidine to produce phosphorylated 5' ends. Fragmented DNA is amplified by LMPCR (see Figure 2-3) and then separated on a sequencing gel. This figure was modified from Drouin, R., *et al.* (2001) Methods Mol Biol *148*, 175-219.



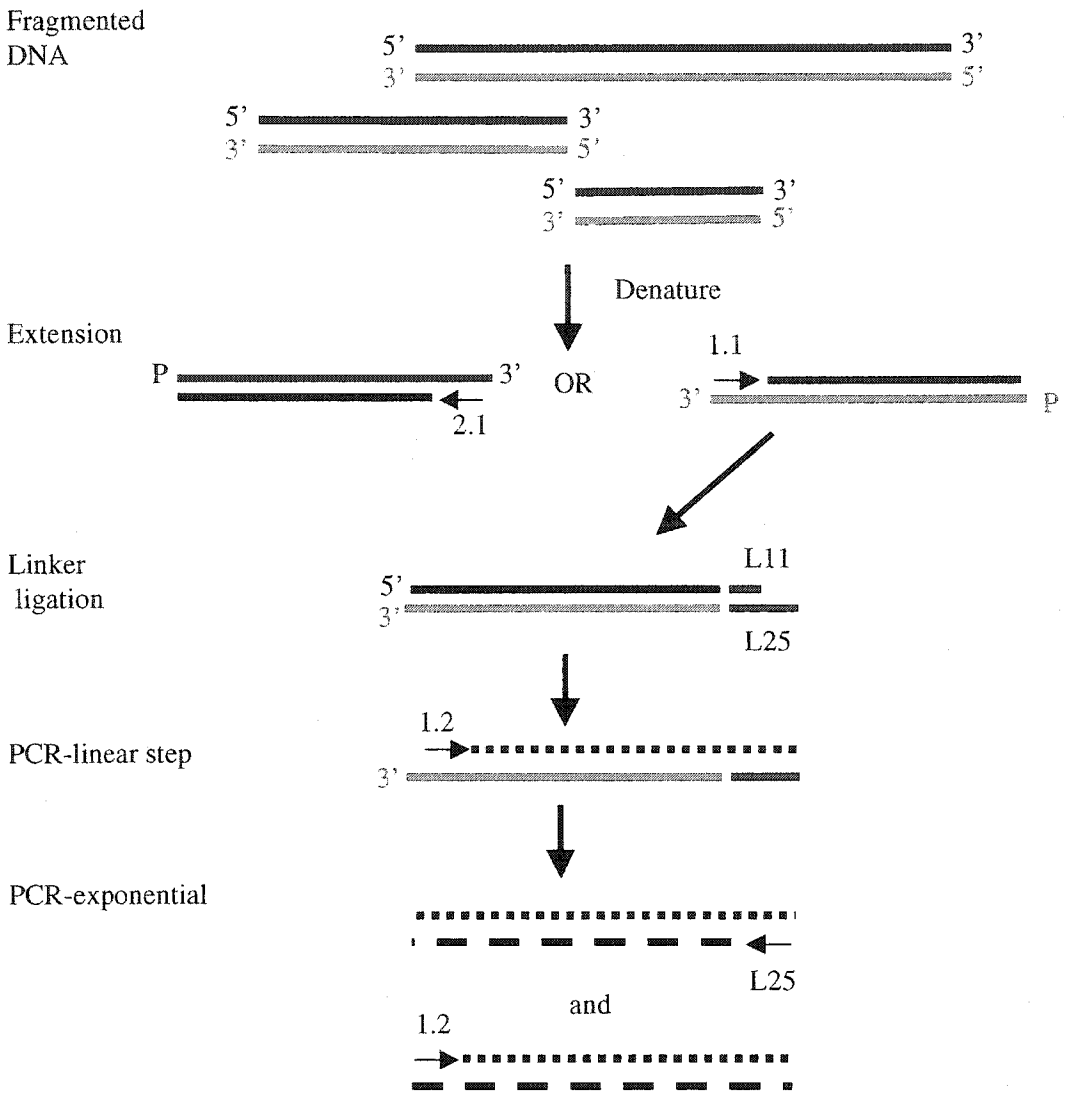


Figure 2-3. Overview of LMPCR.

The DNA is fragmented by DNase I, UV or DMS treatments. In the extension step, an example using my primers NEC2.1 and NEC1.1 (shown as 2.1 and 1.1) extend off of the top and bottom strands respectively. A staggered linker is annealed (L11/L25). The PCR step uses a nested gene-specific primer NEC1.2 (shown as 1.2) and the longer strand of the linker as the other primer. The first step of PCR is linear because only primer NEC1.2 can anneal. After PCR, DNA fragments are separated on a sequencing gel, which is blotted and probed with a single stranded probe amplified from a PCR product using primer NEC1.2.

DNA extraction for LMPCR

For the extraction of high quality DNA for LMPCR, the DNA was extracted as described (Drouin *et al.*, 2001). Briefly, cells were suspended in "buffer A" (300 mM sucrose, 60 mM KCl, 15 mM NaCl, 60 mM Tris-HCl, pH 8, 0.5 mM spermidine, 0.15 mM spermine, and 2 mM EDTA) with 1% Nonidet P40 and incubated on ice for 5 min. The nuclei were sedimented, washed once with "buffer A", resuspended in "buffer B" (150 mM NaCl and 5 mM EDTA, pH 7.8) and then an equivalent volume of "buffer C" (20 mM Tris-HCl, pH 8, 20 mM NaCl, 20 mM EDTA, 1% SDS). Proteinase K to 450 µg/ml was added and samples were incubated at 37°C for 3 h. RNase A was added to a final concentration of 150 µg/ml and incubated for 1 h at 37°C. DNA was extracted by phenol/chloroform and ethanol precipitated using 200 mM NaCl and 2 vol of precooled absolute ethanol. The ethanol was added slowly with gentle rocking of the tube.

DMS Treatment

For lymphoblasts ~7.5 x 10^7 cells were sedimented by centrifugation and diluted in 50 ml Hank's balanced salt solution (HBSS) (1.5 x 10^6 cells/ml) with 0.1% DMS (99+%, Fluka) and aliquoted into 10 ml tubes. Samples were incubated for 6 min. at room temperature. For FB16 and FB17, five 150 mm tissue culture plates (~3 x 10^6 cells) were grown to confluency and treated with DMS as an attached monolayer. For the *in vitro* experiment 50 µg of DNA was ethanol precipitated and pelleted in a microfuge tube. To the DNA pellet, 5 µl of mH_2O, 200 µl of DMS buffer (50 mM sodium cacodylate and 1 mM EDTA, pH 8) and 1 µl of DMS was added and mixed gently by pipetting.

DNA was isolated from the *in vivo* samples as described above and 50 ug of DNA was used in each piperidine (99+%, Fluka) reaction, which converts the modification to a single strand break.

UVC Irradiation

For FB16 and FB17, five 150 mm tissue culture plates of fibroblasts grown to confluency were treated as described (Drouin *et al.*, 2001). Lymphoblasts (~7.5 x 10^7 cells) were centrifuged, then resuspended at a concentration of 1 x 10^6 cells/ml with cold 0.9% NaCl and placed in thin layers in 150 mm Petri dishes. For the *in vitro* experiment 10 µg of DNA was treated as described.

DNA was isolated from the *in vivo* samples and 10 µg of DNA was treated with T4 endonuclease V (Epicentre Technologies) to convert the UVC modified bases to single-strand breaks. This was followed by treatment with Photolyase (from *E. coli*,- Pharmingen) to remove overhanging dimerized bases and leave phosphorylated 5' ends.

DNase I Treatment

In vivo DNase I treatment on FB16, FB17, LCL2 and LCL3 were done by Isabelle Paradis (Université Laval) as described (Drouin *et al.*, 2001) with minor modifications. For lymphoblasts, cells were suspended in solution I (150 mM sucrose, 80 mM KCl, 35 mM HEPES pH 7.4, 5 mM $MgCl_2$, 0.5 mM $CaCl_2$) and aliquoted by 200 µl (3 x 10^6 cells) into microfuge tubes. Cells were permeabilized by the addition of 200 µl solution I, prewarmed to 37°C, with 0.1% lysolecithin (Sigma) and incubating for 3 min. at room temperature. Cells were sedimented by centrifugation and 7 µl DNase I (0.5 mg/ml- Worthington Biochemical Corporation) was added and samples were gently mixed.

Samples were incubated for 15 min. at room temperature, followed by a quick spin to remove supernatant. Cells were resuspended in 0.5 ml of solution B and transferred to a tube with 0.5 ml buffer C and 600 µg/ml proteinase K. DNA was extracted as described above.

For fibroblasts, five 150 mm tissue culture plates were used. 4 ml of prewarmed (37°C) solution I (150 mM sucrose, 80 mM KCl, 35 mM HEPES, pH 7.4, 5 mM $MgCl_2$, 0.5 mM $CaCl_2$) with 0.1% lysolecithin was added to the cells were incubated for 1-2 min. Cells were washed in solution I and 3 ml solution II (150 mM sucrose, 80 mM KCl, 35 mM HEPES pH 7.4, 5 mM MgCl2, 2 mM $CaCl_2$) with 60 µl of DNase I (0.5 mg/ml) was added. Cells were incubated for a total of 20 min. After 8 min. the cells were scraped off with a rubber policeman and put into a 15 ml tube and left for the remainder of the incubation. After 20 min., cells were centrifuged for 1 min. to pellet cells and supernatent was removed. Cells were resuspended in 0.5 ml of solution B and transferred to a tube with 0.5 ml buffer C and 600 µg/ml proteinase K. DNA was extracted as described above.

For *in vitro* sample 40 µg of DNA was digested in 400 µl of solution B with 2.5 µg/ml DNase I. Reaction was stopped with 400 µl phenol, and then samples were extracted once with phenol/chloroform and once with chloroform followed by ethanol precipitation.

Ligation-mediated PCR (LMPCR)

Extension

0.5-2 µg of DNA processed from one of the treatments above was mixed with 3 µl of 10 x Pfu buffer (200 mM Tris-HCl, pH 8.8, 20 mM $MgSO_4$, 100 mM NaCl, 100 mM $(NH_4)SO_4$, 1% (v/v) Triton X-100, 1mg/ml nuclease-free BSA), 1 pmol of the extension primer (see below), 5 µl of cloned Pfu mix (1.5 mM of each dNTP and 1.5 U cloned Pfu DNA polymerase) and mH_2O to a final volume of 30 µl. Reactions were denatured at 98°C for 5 min., incubated at Tm of primer-1°C for 4 min, then at 75°C for 20 min.

Primers NEC1.1, NEC3.1 and NEC5.1 were used for extension off of the bottom (non-transcribed) strand DNA fragments and primers NEC2.1, NEC4.1 and NEC6.1 were used for extension off of the top (transcribed) strand DNA fragments (see Figure 4-2 for map and Table 2-3 for list of primers).

Ligation

Linkers were prepared by annealing the 25-mer, L25 5'GCGGTGACCCGGGAGATCTGAATTC with the 11-mer, L11 5'GAATTCAGATC by members of the laboratory of Dr. Régen Drouin. This was done by heating at 95°C for 3 min., transferring quickly to 70°C, and gradually cooling to 4°C over a period of 3 h. Ligation was performed by addition of 45 µl of ligation mix (30 mM DTT, 1 mM ATP, 80 ng/ml BSA, 2 pmol/µl linker, 3 U T_4 ligase) to the extension reaction, followed by incubation overnight at 18°C. DNA was ethanol precipitated with the addition of 20 µg/ml glycogen and pellet was resuspended in 50 µl mH_2O.

Table 2-3. Primers and T_m for LMPCR.

Primer		Tm
NEC1.1	TCACTGGTTCGCATAAAGC	55
NEC1.2	GCTCATGTTTACAAAGCCGCCCAGACC	65
NEC3.1	CCAGTACGCATCCATCTC	57
NEC3.2	ACTTCTCTCCTGCCCTAGATCTTCTCAGCC	66
NEC5.1	TCCCGCCGCCGCCC	64
NEC5.2	GCCCTGCCCGTCGCTGCGGAAGGC	69
NEC2.1	CGGCACTGCGCCTGCG	64
NEC2.2	CGCGGCCTTGGCCAGCGGCTGG	69
NEC4.1	AGATCCTTACTTTGTTCTGACATG	56
NEC4.2	TCTGCGCCGTCTGGCAAGGGCAGG	69
NEC6.1	CTGGGGAGGCGGCG	61
NEC6.2	GCCTGCGGAGCGGCCGTCGGGC	69

Melting temperatures (T_m) for primers are given in temperature °C.

PCR

Primers NEC1.2, NEC3.2 and NEC5.2 each in conjunction with the linker L25 oligonucleotide were used to amplify off of the bottom (non-transcribed) strand. Primers NEC2.2, NEC4.2 and NEC6.2 each in conjunction with the linker L25 oligonucleotide were used to amplify off the top (transcribed) strand. A list of necdin primers and melting temperatures (T_m) are shown above (Table 2-3). Since the L25 oligonucleotide is overhanging and does not have a complete complementary strand, one round of linear amplification takes place before L25 can anneal.

For primers NEC1.2, NEC2.2, NEC3.2 and NEC4.2 PCR was done with Pfu as follows: To the above ligated samples, 50 µl of cloned Pfu DNA polymerase mix was added (2X cloned Pfu buffer, 0.5 mM of each dNTP, 10 pmol LP25 (linker primer), 10 pmol of gene-specific primer, 3.5 U cloned Pfu DNA polymerase). A final concentration

of 7.5% DMSO was added to the reaction mix for primers NEC2.2, NEC3.2 and NEC4.2. The PCR conditions are shown in Table 2-4.

For primers NEC5.2 and NEC6.2 PCR was done with Taq as follows: To the above ligated samples, 50 µl of Taq polymerase mix was added (4 mM $MgCl_2$, 2 x Taq buffer, 0.5 mM dNTP, 20 pmol LP25, 20 pmol gene-specific primer, 6 U Taq polymerase). See Table 2-4 for PCR conditions.

Table 2-4. PCR conditions for LMPCR.

Cycle	T (°C) for D Denaturation		T (°C) for D Annealing		Polymerization
	Pfu	Taq	Pfu	Taq	
0		93 for 2 m			
1	98 for 4 m	98 for 2:30 m			
2	85 for 1 s	85 for 1 s	Tm for 2 m	Tm for 3 m	3 m
3	98 for 2 m	95 for 1 m	Tm-1 for 2 m	Tm-1 for 2:30 m	3 m
4	98 for 2m	95 for 1 m	Tm-2 for 2 m	Tm-2 for 2 m	3 m
5	98 for 2m	95 for 1 m	Tm-3 for 2 m	Tm-3 for 2 m	3 m
6	98 for 2m	95 for 1 m	Tm-4 for 2 m	Tm-4 for 1:30 m	3 m
7	Repeat cycle 5, 18 more times for Pfu and 13 more times for Taq (for Taq, 5 s were added per cyclefor annealing and polymerization steps)				
8	98 for 2 m	95 for 1 m	Tm-3 for 2 m	Tm-3 for 4 m	3 m
9	98 for 2 m	95 for 1 m	Tm-2 for 2 m	Tm-2 for 3 m	3 m
10	98 for 2 m	95 for 1 m	Tm-1 for 2 m	Tm-1 for 4 m	3 m
11	98 for 2 m	95 for 1 m	Tm for 2 m	Tm for 4 m	10 m

Note: Temperature (T) and duration (D) of the denaturation, annealing and polymerization steps. T is in minutes (m) or seconds (s). Tm is the melting temperature for the primer. This Table was modified from Drouin *et al.* (2001).

Blotting and Hybridization

LMPCR samples were electrophoresed on denaturing acrylamide gels. Sequencing ladders were prepared by members of the laboratory of Dr. Drouin using chemical cleavage as described (Drouin *et al.*, 2001).

The gel was electroblotted to a positively charged nylon membrane (Roche Molecular Biochemicals) and hybridized. The single stranded probes were generated by PCR amplification with a single primer on a purified PCR fragment amplified with NEC1.2 and NEC2.2 or NEC5.2 and NEC6.2. ^{33}P dCTP was incorporated in the PCR amplification. Blots were hybridized and exposed to film for 3 hours to 3 days depending on the intensity of the signal.

Chapter 3. Allele-specific methylation patterns of *NDN/Ndn* CpG islands

The data in this chapter have either been published or submitted.

Hanel, M. L., and Wevrick, R. (2001). Establishment and maintenance of DNA methylation patterns in mouse *Ndn*: implications for maintenance of imprinting in target genes of the imprinting center. Mol Cell Biol *21*, 2384-92. (Permission has been given to republish this material.)

Hanel, M. L.*, Lau, J. C. Y.*, and Wevrick, R. (2002). Epigenetic marks of tissue-specific and imprinted regulation of the human *NDN* gene. (Joint first authorship). Submitted.

* All of the work presented in this chapter was done by Meredith Hanel.

Introduction

DNA methylation is an important form of gene regulation during mammalian development, and has been implicated in such diverse processes as genomic imprinting (Li *et al.*, 1993), X-inactivation (Beard *et al.*, 1995) and differential gene expression (Eden and Cedar, 1994). Molecular control of imprinting requires an epigenetic modification of DNA in the haploid genome leading to hemizygous expression in the diploid embryo. The initial mark which differentiates the two parental alleles likely originates in the gametes or early zygote when the two alleles are still separate. After zygote formation, the parental alleles maintain their identity so that one allele eventually becomes preferentially expressed. Methylation of CpG dinucleotides is proposed to be one mechanism for differentially marking the parental chromosomes, since methylation can be stably inherited in somatic cells, yet can be removed and reset in the next generation according to parent of origin (Jaenisch, 1997; Razin and Cedar, 1994).

The developmental stages prior to blastocyst formation are of particular importance in genomic imprinting (Solter, 1998). Genome-wide, oocyte DNA tends to be hypomethylated while sperm DNA tends to be hypermethylated (Monk *et al.*, 1987). During preimplantation development, the overall level of methylation decreases. Most methylation moieties present on the original parental chromosomes are removed from the DNA by the morula stage, giving rise to a predominantly unmethylated genome which remains this way at least through blastulation. A wave of *de novo* methylation follows, leading to an overall increase in genome methylation levels as the newly implanted embryo develops and differentiates. In imprinted genes, gamete-derived methylated CpG

sequences are predicted to be preserved during pre-implantation development and must therefore be specifically recognized and protected from the global generalized demethylation that takes place at these stages (Olek and Walter, 1997; Tremblay *et al.*, 1995). Allelic differences present in germ cells could be rapidly expanded during preimplantation development and therefore serve as a primary imprinting signal. Such methylated sites may or may not be retained in all somatic cells as a marker of parental identity.

All imprinted genes analyzed to date that have CpG islands have displayed allele-specific DNA methylation patterns (Neumann *et al.*, 1995; Razin and Cedar, 1994). For some of these genes, differentially methylated regions (DMRs) in the gametes precede allele-specific methylation differences in adult tissues. However, studies of methylation in early embryogenesis have been limited to a small number of imprinted genes (Bartolomei *et al.*, 1993; Birger *et al.*, 1999; Feil *et al.*, 1994; McDonald *et al.*, 1998; Shemer *et al.*, 1997; Stoger *et al.*, 1993; Tremblay *et al.*, 1995; Warnecke *et al.*, 1998). The interpretation of these studies is complicated because the DMRs have been implicated in imprinting control of other genes. For example, the DMR upstream of *H19* is implicated in reciprocal imprinting of *Igf2* (Bartolomei *et al.*, 1993; Bell and Felsenfeld, 2000; Hark *et al.*, 2000; Thorvaldsen *et al.*, 1998), and the upstream region of *Snrpn* contains an essential regional imprinting element (Yang *et al.*, 1998). Although DNA methylation is essential for normal development, its exact role in imprinting at all developmental stages remains controversial, and other epigenetic events have been invoked to explain the development of the final imprinted expression pattern (Jaenisch, 1997; Tilghman, 1999). For genes that are not involved in the imprinting of other genes

but are the target of an imprinting center, it is not known whether the establishment of allele-specific methylation in the gametes is required for allele-specific expression of genes. For these genes, it is also not known whether maintenance of gamete-specific methylation patterns throughout development is necessary for proper imprinted gene expression.

To better understand the role of DNA methylation in the maintenance of imprinted gene expression, I have analyzed allele-specific DNA methylation patterns in the imprinted mouse *Ndn* gene and the human homolog *NDN*. *Ndn/NDN* is an ideal gene for imprinting study because of its simple, intronless gene structure, and because it is a target of the imprinting centre rather than being intimately involved in imprinting other genes. Mouse *Ndn* is located on mouse Chromosome 7 in a region of conserved synteny with human 15q11-q13, the region commonly deleted in Prader-Willi syndrome (PWS) (Nicholls, 2000). PWS results from the loss of expression of genes, including *NDN*, that are normally active on the paternal copy of chromosome 15q11-13 (Nicholls, 2000). Imprinting in the PWS region is controlled by the imprinting center (IC), located in the 5' end of the *SNRPN/Snrpn* gene, about one Mb from the *NDN/Ndn* locus.

DNA methylation is a good candidate for the establishment of a primary epigenetic mark that distinguishes the two alleles of *NDN/Ndn* and in maintenance of imprinted gene expression. The experiments in this chapter were aimed at elucidating the role of DNA methylation in the imprinted regulation of necdin. I hypothesized that if DNA methylation is important in the imprinting of *NDN/Ndn*, the two parental alleles would be differentially methylated. I also expected that if allele-specific methylation was important in the imprinting process, this epigenetic characteristic would be conserved

between human and mouse. If DNA methylation serves as the primary epigenetic mark that distinguished the two alleles of necdin, differential methylation was expected to be present in the gametes or early embryo.

Mouse and human *Ndn/NDN* contain a CpG-rich region that extends from immediately upstream of the transcription start site into the first half of the open reading frame. Human *NDN* contains an additional 3' CpG island not found in mouse. In this study, I have analyzed the adult methylation profiles of the 5' end of the mouse and human *Ndn/NDN* gene and the human *NDN* 3' CpG island by sodium bisulfite sequencing (SBS). SBS enables methylation analysis of individual DNA strands, even in limited tissue samples (Frommer *et al.*, 1992). In mouse *Ndn*, I have also analyzed gametic and early embryo methylation patterns.

Results

The genomic structure of human and mouse *NDN/Ndn*

The mouse and human *Ndn/NDN* genes are both composed of single exons containing small open reading frames of 325 and 321 amino acids, respectively. The genomic sequence of a 15,522 bp region surrounding *Ndn* and a 104 kb region surrounding *NDN* was retrieved from GenBank. I first analyzed the genomic sequence surrounding *Ndn* (15,522 bp) and *NDN* (20,000 bp) for the presence of repetitive elements using the RepeatMasker web-based software (http://ftp.genome.washington.edu/cgi-bin/RepeatMasker) (Figure 3-1). With the repeat-masked sequence as input, I used the PipMaker web-based software (http://nog.cse.psu.edu/pipmaker/) (Schwartz *et al.*, 2000) to compute percentage identity

between mouse and human *Ndn/NDN* (Figure 3-1). As expected, significant nucleotide identity extends through the open reading frame. Short regions immediately flanking the open reading frame, and isolated regions several kb on either side of the gene which may correspond to gene regulatory elements also show some sequence similarity.

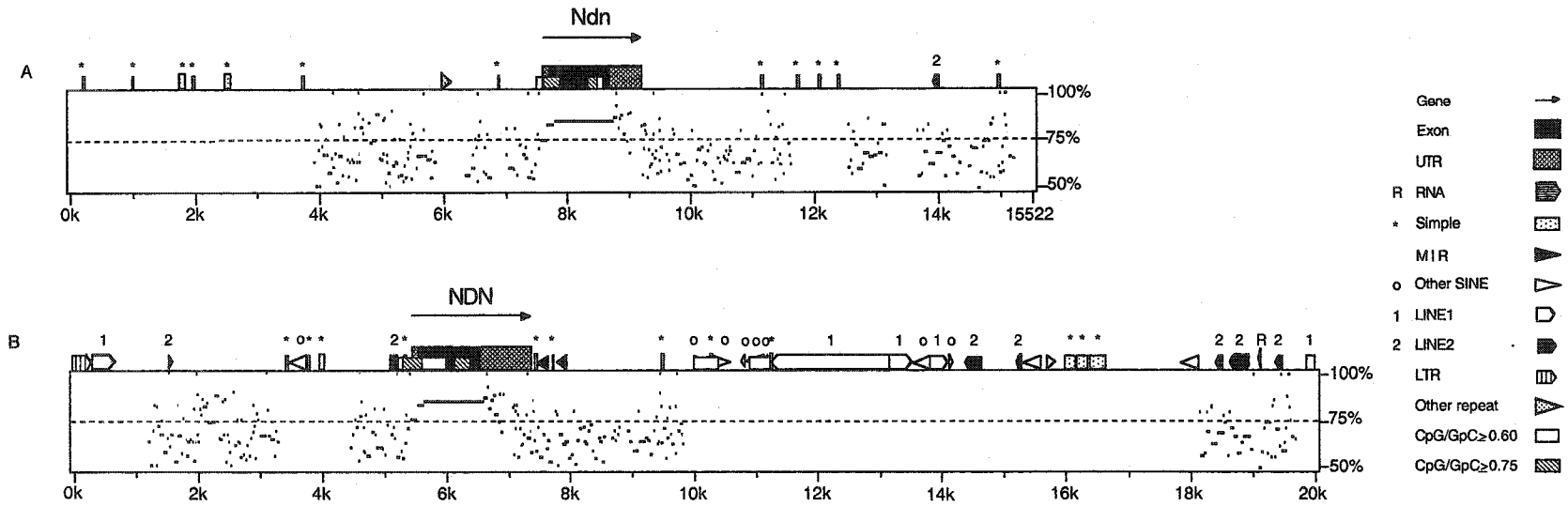


Figure 3-1. PipMaker sequence alignments.

A). 15,522 bp of sequence surrounding mouse *Ndn* was compared to 20 kb of sequence surrounding human *NDN*. The mouse *Ndn* transcript and surrounding sequence elements are shown. B). 20 kb of sequence surrounding human *NDN* was compared to mouse *Ndn*. The human *NDN* transcript and surrounding sequence elements are shown. The arrow and full height box indicate the position of *Ndn/NDN*, with the open reading frame shaded black and the untranslated regions shaded with cross hatchings. Half-height boxes and arrows outside *Ndn/NDN* indicate repetitive elements. Low boxes within *Ndn/NDN* and downstream of *NDN* indicate CpG-rich regions, within which white shading has CpG/GpC content ≥0.60 and diagonal hatching has CpG/GpC≥0.75. On the plot below each sequence schematic, regions of homology are shown by black dashes and percent nucleotide identity is indicated on the right. See http://nog.cse.psu.edu/pipmaker/ for details of mouse and human repeat types.

PipMaker analysis indicated the locations of CpG islands, defined as having a CpG/GpC ratio of greater than or equal to 0.6. PipMaker analysis also distinguished between denser CpG islands with a ratio of at least 0.75 from islands meeting the 0.60 criteria. Two CpG islands, located in equivalent positions in human and mouse *NDN/Ndn* were found (Figure 3-1). The 5' CpG island begins in the putative promoter and extends into the open reading frame, while a 3' CpG island is located in the 3' portion of the open reading frame. I predicted that if a gamete-specific methylation imprint exists in *NDN/Ndn*, then the 5' CpG island represents the most probable location for the imprint. This prediction is based on two recent gene-targeting experiments that created null mutations of *Ndn* in the mouse (Gerard *et al.*, 1999; Tsai *et al.*, 1999). In both cases, the *lacZ* gene replaced the *Ndn* gene starting at the *Bam* HI site located between CpG positions 23 and 24 (Figure 3-2) and terminating at the stop codon. Mice inheriting the deleted allele maintain proper parent- of-origin gene expression of the reporter gene, and in one case show an imprinted phenotype due to loss of necdin gene expression on germline transmission (Gerard *et al.*, 1999). This implies that the region downstream of the *Bam* HI site in the open reading frame is not necessary for proper imprinting of *Ndn*. Therefore I focused on the 5' portion of this promoter CpG for our analysis of mouse and human *Ndn/NDN*.

I found another CpG island about 4.4 kb downstream of the end of the *NDN* transcription start. Unlike the promoter CpG island, no equivalent 3' CpG island was found in the mouse sequence for 30 kb downstream of *Ndn* (Genbank #AC027298). It co-localizes with repetitive elements in a region that have no sequence similarity to the murine sequence. However the presence of this CpG island in such close proximity to an

imprinted gene still made it a candidate for a regulatory element in the imprinting process. Other imprinted genes have been known to have CpG islands in introns or in the flanking 5' or 3' regions. The *Igf2-H19* locus contains an intergenic differentially methylated region that acts as a boundary element which binds the chromatin remodelling factor CTCF only when this region is unmethylated.

Strategy for sodium bisulfite sequencing

To analyze the developmental pattern of CpG methylation within the 5' CpG island, I used sodium bisulfite sequencing (SBS), a sensitive technique for single CpG dinucleotide methylation analysis. Sodium bisulfite treatment of DNA converts all unmethylated cytosines to uracil (Figure 2-1). Upon DNA sequencing of cloned PCR products, bases only appear as cytosine if they were methylated in the original DNA sample, and thus protected from bisulfite moderated conversion. Mouse brain, liver, heart and testes samples were prepared from 6 week old C57BL/6 x SPRET and C57BL/6 x CAST interspecific F1 mice. Gametes and early embryos were isolated from C57BL/6 mice and F1 interspecific crosses. The parental alleles of human *NDN* were analysed in isolation by using tissues and cell lines from PWS and AS patients that are deleted for the 15q11-q13 region from the paternal and maternal alleles respectively. DNA isolated from all samples was then processed for sodium bisulfite treatment. Bisulfite-modified DNA was amplified by nested PCR and cloned products were sequenced from both ends.

Because of previous reports of difficulties in obtaining complete conversion of sodium bisulfite treated DNA, particularly with samples derived from small numbers of cells, I performed control experiments using primers specific to the *H19* gene (Warnecke *et al.*, 1998). We amplified sodium bisulfite treated DNA from sperm and adult brain

with *H19* Region B primers and sequenced the cloned PCR products. These experiments demonstrated that I had obtained a high conversion efficiency (>99%), based on the complete conversion of C residues not present in CpG dinucleotides. In addition, I replicated the observation that sperm DNA is almost completely methylated at each CpG in *H19* Region B. These experiments validated the sodium bisulfite sequencing method in our hands.

Allele-specific methylation in mouse *Ndn* is tissue-specific

PCR primers were designed to amplify the 5' CpG-rich region in *Ndn* (Figure 3-1). Single site polymorphisms among C57BL/6, SPRET and CAST mice were determined by direct sequencing of PCR products generated from genomic DNA. All experiments on adult tissues were done on samples from F1 interspecific mice, and the single site polymorphisms were used to identify the parent of origin of cloned PCR products in these experiments.

We analyzed 39 CpG sites in the mouse promoter CpG island. The CpG sites are approximately evenly distributed over this interval. The predicted transcriptional start site is between CpG sites 10 and 11, and the translational start site is between CpG sites 16 and 17 (Figure 3-2) (Uetsuki *et al.*, 1996). DNA samples were derived from adult brain, in which *Ndn* is transcribed, and liver and heart tissues in which *Ndn* is not active. In order to obtain a representative sample, at least 7 clones from each parental allele were sequenced. We noted a parental bias in the distribution of clones obtained, with maternally derived alleles appearing about five times more frequently than paternally derived clones. A similar bias in parental alleles had also been noted for studies of *Xist* (McDonald *et al.*, 1998), but in our case the presence of an interspecies polymorphism

under the forward primer for the first round of PCR may be the cause of the bias. To enrich for paternally derived clones, some PCR reactions were performed with a first round forward primer that contained only the SPRET (paternal) allele of this polymorphism.

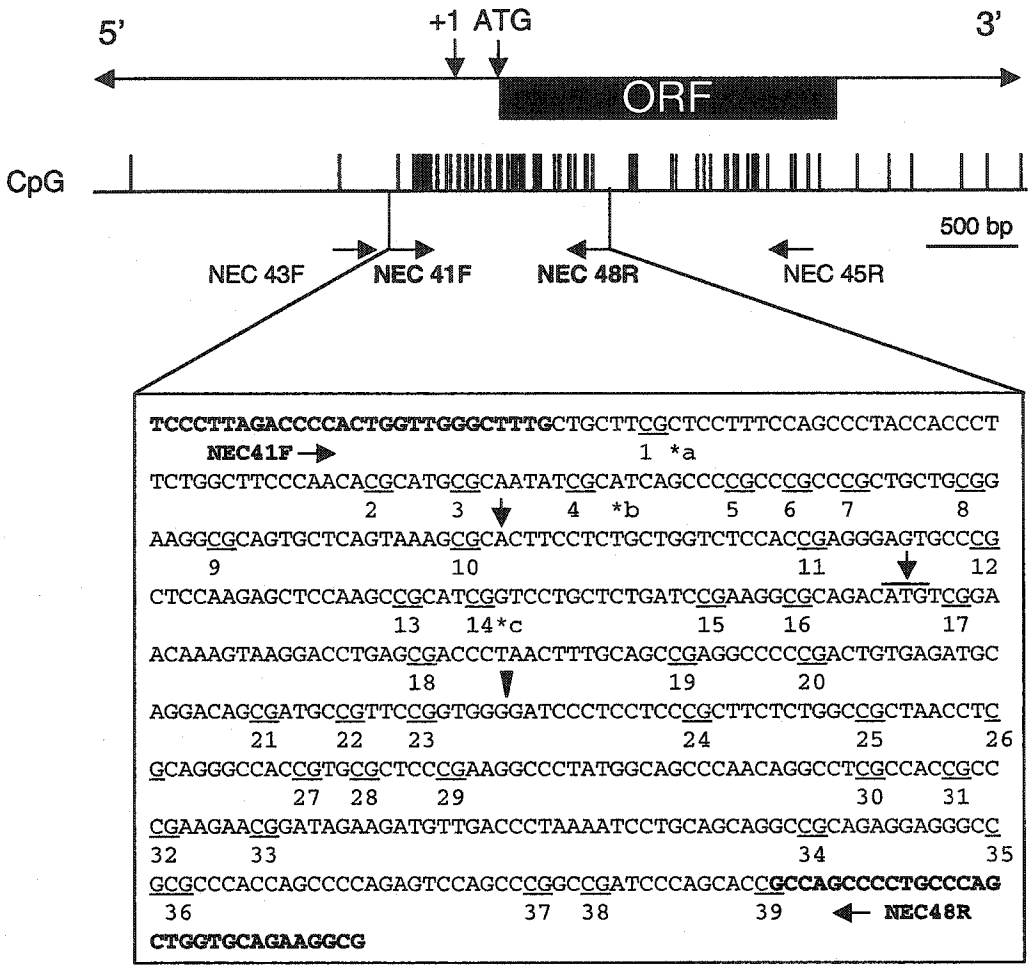


Figure 3-2. Map of the region of mouse necdin analyzed for CpG methylation status.

The gene structure is shown at the top, where the black box represents the open reading frame and the +1 and ATG the transcriptional and translational start sites respectively. CpG dinucleotides are shown as vertical bars below the map. The triangular arrow points to the *Bam* HI restriction site used to insert the *lacZ* gene into the *Ndn* knock-out mouse. The relative positions of the primers used for nested PCR are shown, and the second set of primers (**NEC41F/NEC48R**) are in bold. The 560 bp region analyzed covers CpG sites 1-39 and are underlined in the C57BL/6 type sequence. For some of the samples other primers were used to amplify the same region (see Methods). Single nucleotide polymorphisms between C57BL/6 and SPRET or CAST are indicated by an asterisk and are: a, CAST CGGTC, and b, SPRET and CAST, CGCGT. A five bp insertion was found in SPRET at c: SPRET CGCATCGCATCG. The changes from the C57BL/6 sequence are in bold typeface.

Sodium bisulfite sequencing analysis of adult brain revealed a mosaic pattern of DNA methylation on both parental alleles (Figure 3-3A). Additional clones, not shown in Figure 3-3A, derived from PCR products that contained only CpG dinucleotides 1-20 were also included in subsequent analysis in Figure 3-4 which were consistent with the rest of the data. Most sites showed variable methylation between clones, and no sites were consistently methylated, although a few sites (numbered 1, 12 and 39) were unmethylated in all or almost all clones on both parental alleles. The average level of methylation was 22% in the maternal clones, and 17% in the paternal clones. The maternal clones in liver appear to be more sparsely methylated (Figure 3-3B), with the heaviest methylation being 20%, compared to brain where the heaviest methylation was 62% and a substantial portion of clones were between 20-62% methylated. The average level of methylation in liver and heart were lower than in brain. For example, the average levels of methylation in maternal and paternal heart clones were 10% methylated and 0.4% methylated respectively.

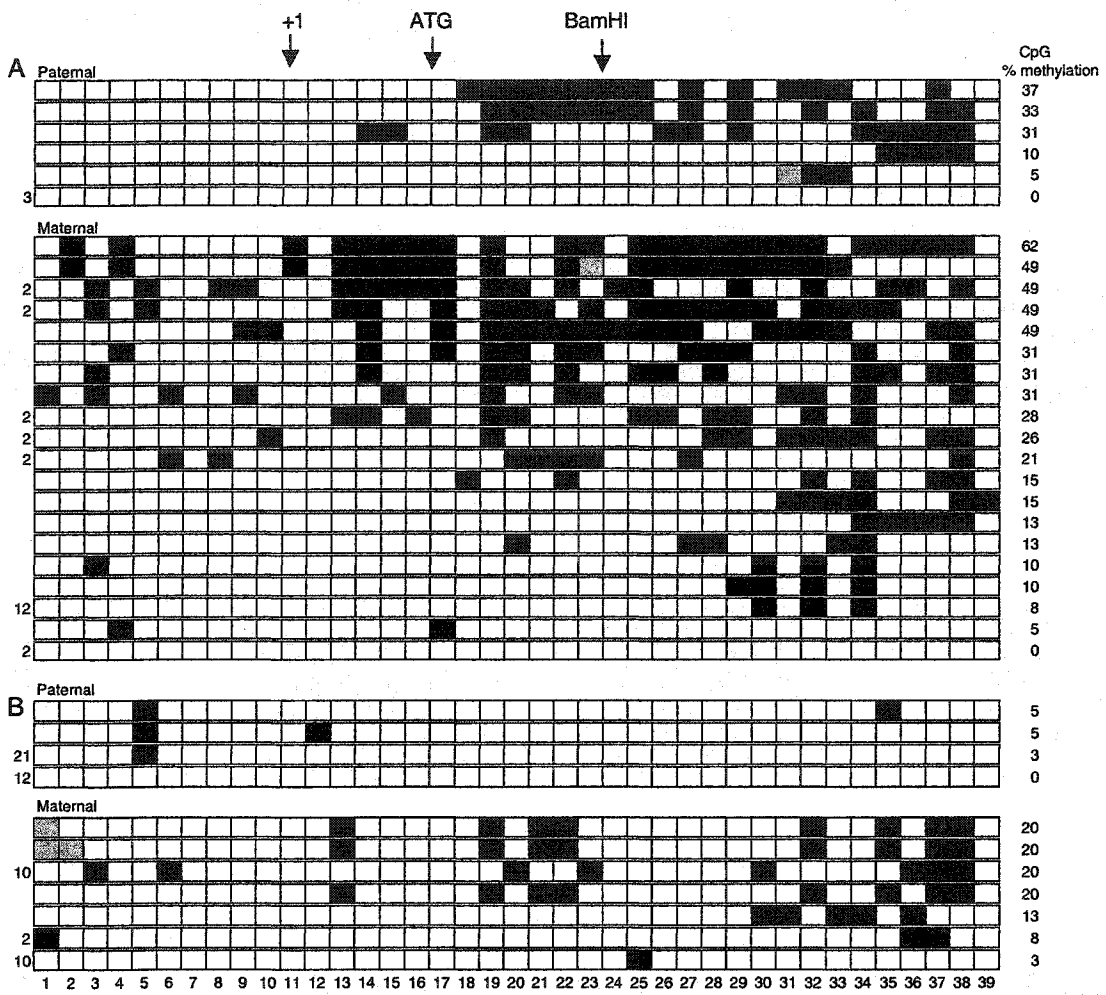


Figure 3-3. Heterogeneity in methylation analysis.

(A) Brain. (B) Liver. The CpG dinucleotides are numbered across the bottom. Each row represents a single clone derived from a C57BL/6 X CAST F1 mouse or a C57BL/6 X SPRET F1 mouse. Black boxes represent methylated CpG sites, white boxes are unmethylated CpG sites, and gray boxes were ambiguous in the sequence analysis. In cases where clones with the same profile appeared more than once, the number of times it was found is indicated at the left. Paternally-derived clones are at the top, with maternally derived clones below, as determined by polymorphism analysis. The clones are ordered by the % methylation for each clone, shown to the right.

The proportion of clones methylated at each individual CpG dinucleotide in liver, heart and brain was calculated for each parental allele (Figure 3-4). No CpG site was methylated in all maternally derived clones. In the adult brain, paternally inherited CpG sites 1-17 were hypomethylated compared to the maternal allele, while in sites 18 through 39, methylation levels were more equivalent (Figure 3-4A). In the adult heart and liver, the maternal allele was more highly methylated than the paternal alleles throughout the region studied but the methylation levels were much lower overall than in brain (Figure 3-4B,C). A lack of methylation in the more 5' end of this region is consistent between all three tissues examined. The apparent high methylation at CpG site 5 in liver (Figure 3-4B) is probably not significant since there were a high number of identical clones with this methylation pattern (Figure 3-3B). This can arise from clonal PCR amplification when the amount of sodium bisulfite treated DNA is too low and the sequenced clones do not represent the true variation in methylation found in the original DNA sample.

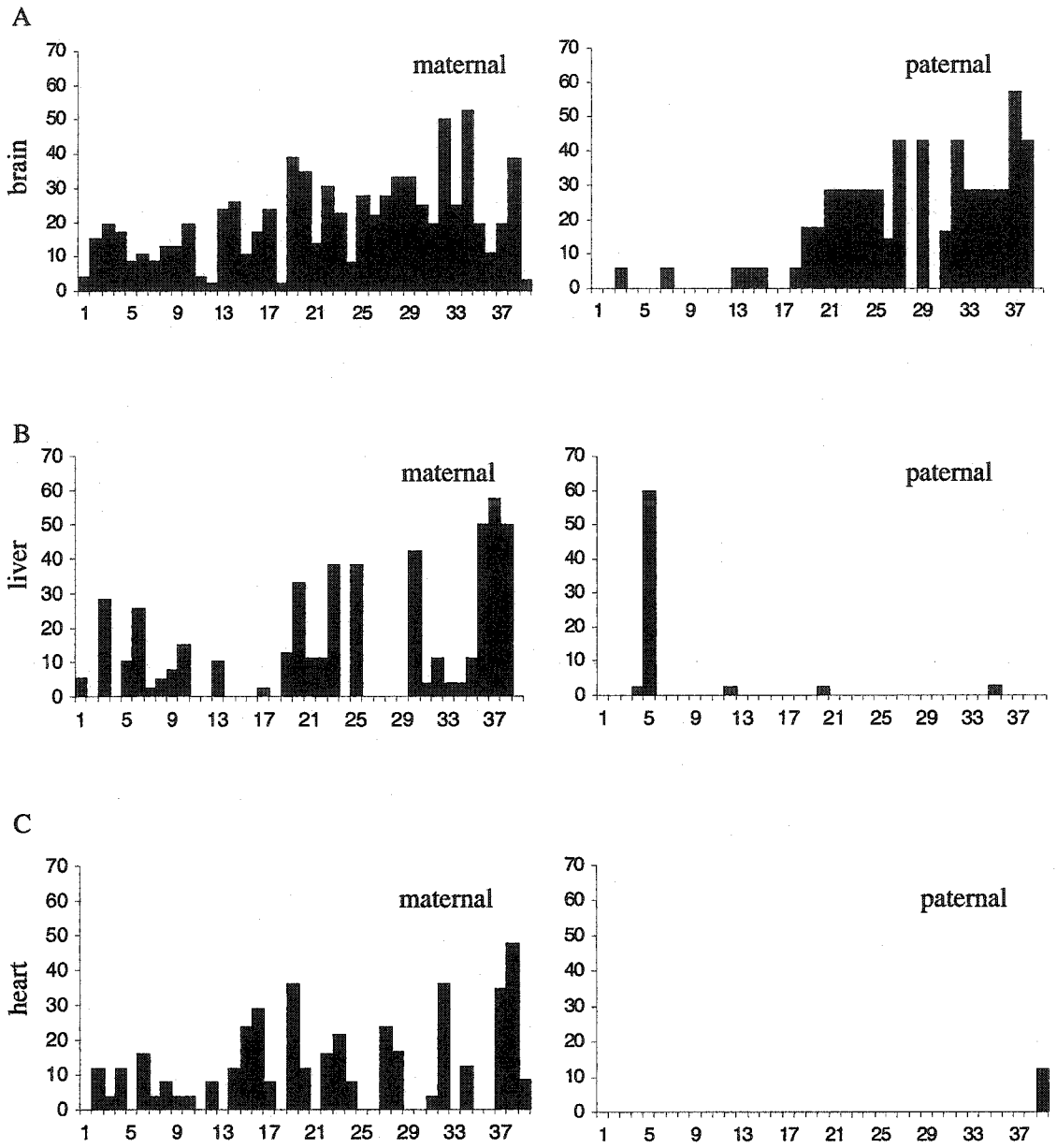


Figure 3-4. Methylation profile of 39 CpG dinucleotides in mouse adult brain, liver and heart.

The percentage of clones methylated at each CpG dinucleotide position was calculated and plotted against the CpG position (1 to 39). Profiles are of (A) maternal and paternal clones from adult brain, (B) maternal and paternal clones from adult liver and (C) maternal and paternal clones from adult heart.

Allele-specific methylation in the human *NDN* 5' CpG island is not correlated with gene expression

I designed PCR primers to analyze the 73 CpG sites from the promoter CpG island and 19 CpG sites from a 350 bp segment in the 3' end of the downstream CpG island. I analyzed the CpG island in the *NDN* promoter by sodium bisulfite sequencing. In order to analyze the maternal and paternal alleles in isolation, I used cell lines and tissues lacking either the maternal allele, derived from AS individuals with maternal deletion of 15q11-q13, or the paternal allele, derived from PWS individuals with a paternal deletion of 15q11-q13. Previous studies had shown that the *NDN* gene is not expressed in control lymphoblasts or lymphocytes but is expressed in brain and fibroblasts (MacDonald and Wevrick, 1997). We therefore compared the methylation patterns of *NDN* in these cell types. Individual DNA strands produced by nested PCR of sodium bisulfite treated DNA were cloned and sequenced.

In all cell/tissue types the inactive maternal allele was hypermethylated compared to the paternal allele (Figure 3-5). In lymphocytes (*NDN* not expressed) the maternal methylation levels were similar to that seen in brain. The levels of methylation were relatively uniform among clones within a cell/tissue type but no CpG sites were methylated in all maternal clones or unmethylated in all paternal clones. In lymphoblasts the levels of methylation on the maternal and paternal alleles were distinguishable, but the levels of methylation on both alleles were higher than in lymphocytes. Notably, the paternal allele shows a considerable amount of methylation compared to that found in other cell types.

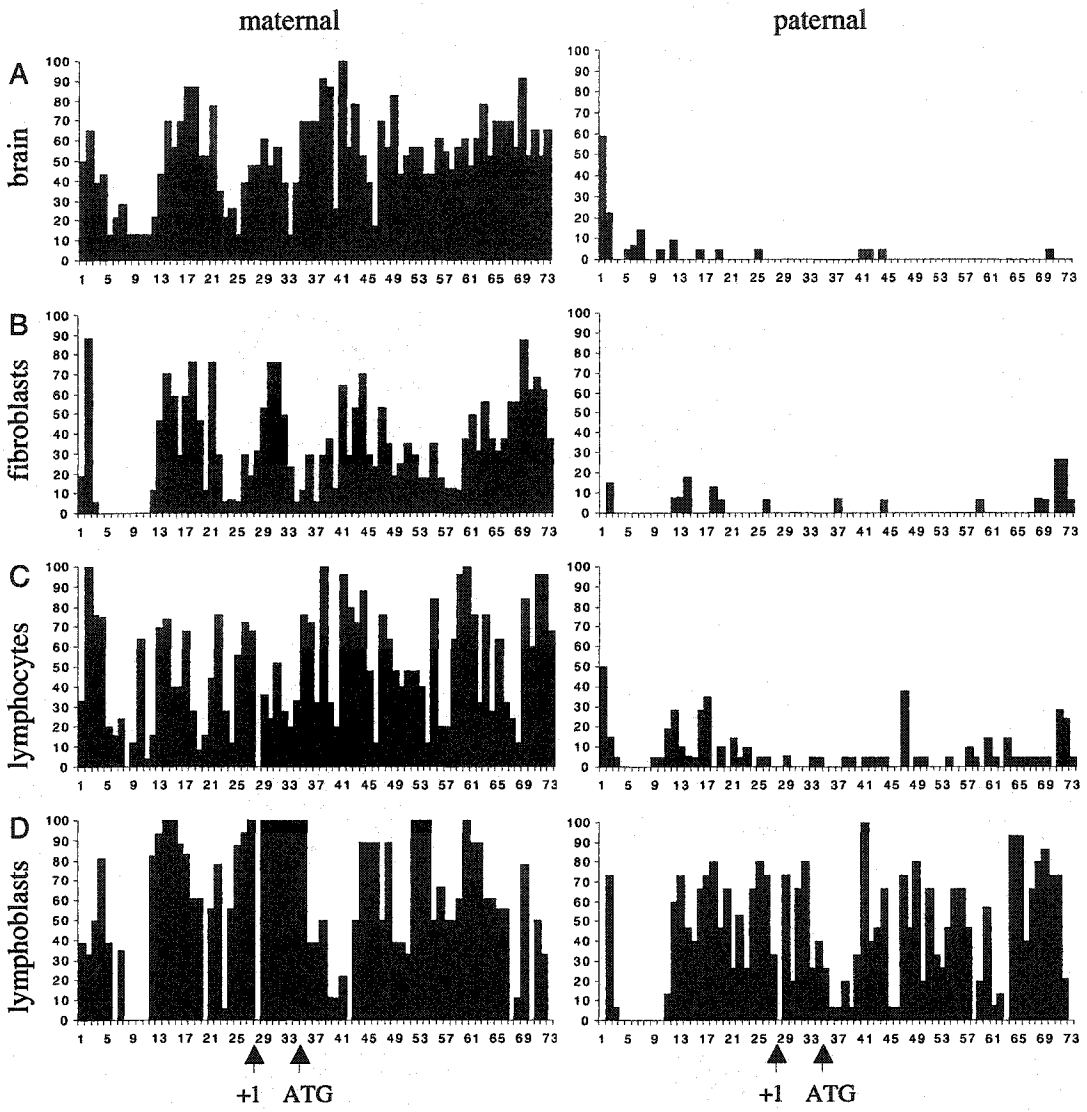


Figure 3-5. Methylation profile of 73 CpG sites in the promoter CpG island of human *NDN*.

The percentage of clones methylated at each CpG dinucleotide position was calculated and plotted against the CpG position (1 to 73). Profiles are of (A) brain, (B) fibroblasts, (C) lymphocytes and (D) lymphoblasts. Data are from the maternal allele (left) and paternal allele (right), derived from PWS and AS deletion cells.

The human *NDN* 3' CpG is hypermethylated on both alleles

I analyzed a 480 bp region with 17 CpG sites within this island with sodium bisulfite sequencing and found that both alleles were hypermethylated and indistinguishable from each other in all cell/tissue types examined (Figure 3-6). Therefore this CpG island does not have a methylation-dependent regulatory role in the allele-specific expression of *NDN*. The first 5 CpG sites were consistently not completely converted by sodium bisulfite as determined by a lack of cytosine to thymine conversion at non CpG sites. This was not restricted to either parental allele or to any cell or tissue type. Denaturation is a required step in the sodium bisulfite conversion, and it is possible that this region does not denature well because of the high CG composition. However I cannot exclude the possibility that this represents rare non-CpG methylation (Ramsahoye *et al.*, 2000).

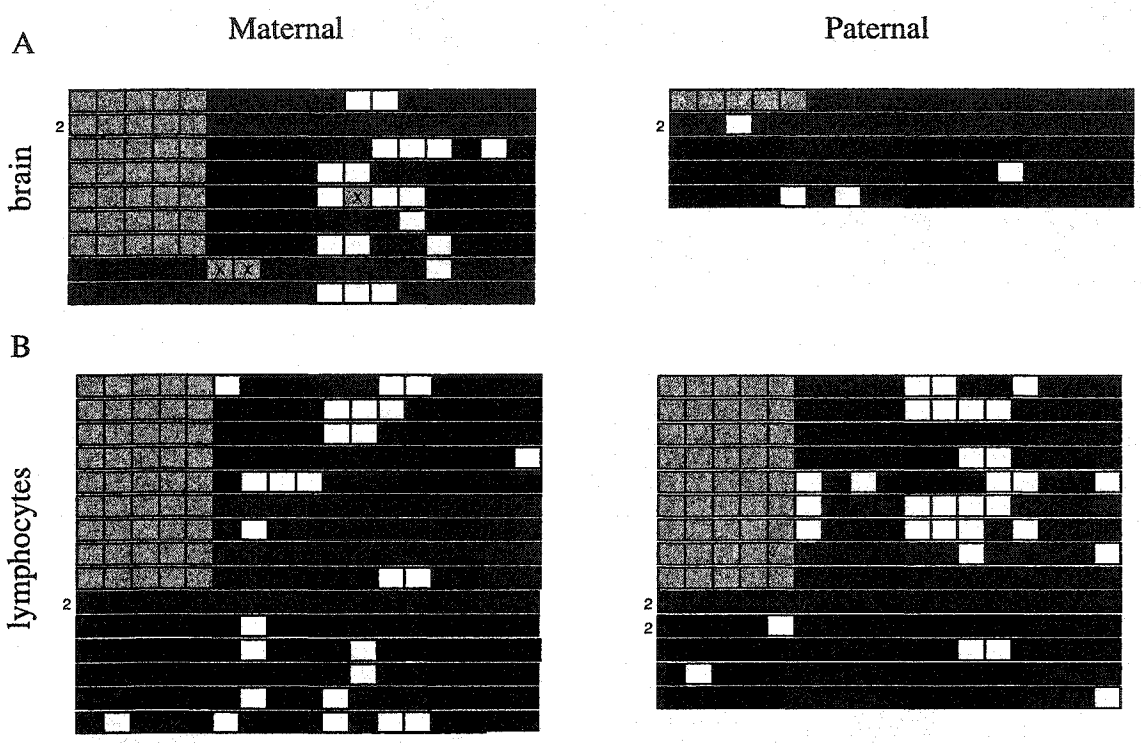


Figure 3-6. DNA hypermethylation at the 3' CpG island of *NDN*.

Data from (A) brain and (B) lymphocytes are shown.Black boxes represent methylated CpG sites, white boxes are unmethylated CpG sites, and gray boxes with an X were ambiguous in the sequence analysis. The grey boxes without the X were regions of incomplete reactivity with sodium bisulfite. In cases where clones with the same profile appeared more than once, the number of times it was found is indicated at the left. These experiments were done by Aliyah Kanji, a summer student under my supervision.

Allele-specific methylation of mouse *Ndn* originates in the gametes

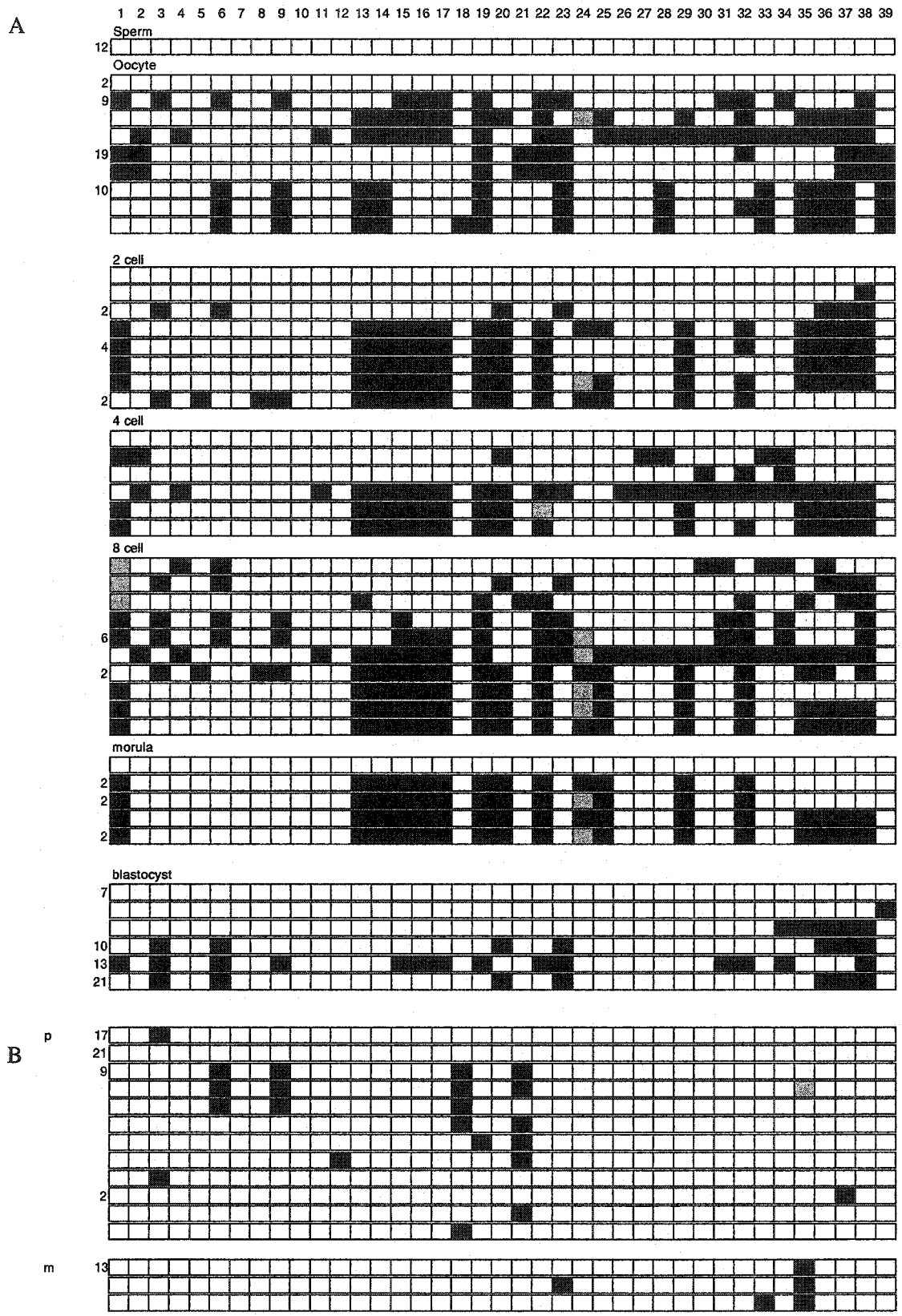
I next performed sodium bisulfite sequencing analysis on gametes and early embryos. The parental origin of the cloned PCR products from blastocysts was inferred from interspecific polymorphisms, although analysis of earlier stages was performed on inbred C57BL/6 crosses where parent-of origin could not be determined. Because of previous reports that individual PCR reactions may not contain clones that adequately represent the initial DNA strands when amplifying from limited starting material (Warnecke *et al.*, 1998), multiple PCRs were carried out with each sample. We sequenced between 6 and 25 clones from at least two PCR reactions from each of the gamete samples and each developmental stage. The only exception was the morula data which came from a single PCR reaction.

The sperm and oocyte methylation patterns define the initial methylation patterns which are subsequently remodeled in the early embryo. In sperm, all 39 CpG sites within the region were unmethylated in all clones (Figure 3-7). In testis, which contains a substantial proportion of germ cells, I detected clones that were completely unmethylated with both maternally and paternally derived alleles represented (data not shown). The lack of partially methylated clones is either because the somatic cell DNA is unmethylated, or because in the sample of clones sequenced there were no somatic cell DNA clones represented. In contrast, oocyte DNA did not display a single pattern of methylated CpG sites. Furthermore, there were no CpG sites consistently methylated in all oocyte clones. Two of the 45 oocyte clones were completely unmethylated, and were distinct from the other clones. This may represent a true diversity in the oocyte population; alternatively, these clones may represent contamination from adhering

maternal cells despite precautions taken to eliminate those cells. If these two clones are excluded, the only methylated site conserved among the remaining 43 oocyte clones is at position 19. Other CpG sites that are methylated in a consistent pattern in most clones are at positions 13 and 14, 15-17, 22, 23, 32, and 35-39.

Figure 3-7. Methylation in gametes and embryos.

The presentation is the same as in Figure 3-3. The CpG dinucleotides are numbered across the top. Paternal (p) and maternal (m) designations are given in the clone sets that came from interspecific crosses. (A) The clones are grouped according to the tissue source. CpG sites 1 to 39 are numbered along the top. (B) Clones from blastocysts derived from C57BL/6 x SPRET crosses.



Analysis of 2, 4 and 8-cell embryos, morulae and blastocyst DNA derived from C57BL/6 inbred crosses was then performed. The sites noted as consistently methylated in most oocyte clones were in general highly methylated at all stages up to the morula, in the presumed maternally derived clones. Overall, methylation at CpG positions 13 to 17 was found at all stages except for the blastocyst stage (Figure 3-7). Methylation at positions 15 to 17 was found in blastocysts and methylation at positions 35 to 38 was found in all early developmental stages. Completely or almost completely unmethylated clones, were present in the 2 cell, 4 cell, morula and blastocyst stages. These likely represent paternal clones since they are unmethylated as in the male gametes. It is also possible that some unmethylated maternal DNA strands exist at these embryonic stages. Many clones with low levels of methylation could not be assigned a putative parent-of-origin because they could be derived either from *de novo* methylation of the paternal allele or demethylation of the maternal allele.

The overall methylation levels were low in the blastocyst samples and the putative maternal alleles retained less of the oocyte specific methylation than in the earlier developmental stages. To gain more information on the parent-of-origin of the partially methylated clones, I next analyzed blastocysts from a C57BL/6J x SPRET cross (Figure 3-7B). Maternal and paternal clones were both sparsely methylated, indicating that *de novo* methylation occurs in the paternal allele and demethylation occurs on the maternal allele leading to equivalent levels of methylation on both parental alleles. No single CpG dinucleotide was consistently methylated in all preimplantation stages.

Discussion

Paternal hypomethylation is common between mouse and human *Ndn/NDN*

Many imprinted genes studied so far have the same allele-specific methylation pattern in all tissues irrespective of gene expression. In *Ndn*, I noted tissue-specific differences. Mainly, relative hypermethylation of the maternal allele in the first 17 CpG dinucleotides is apparent in brain, where *Ndn* is expressed. In the liver and heart, which do not express *Ndn*, the maternal allele is relatively hypermethylated compared to the paternal allele throughout the region analyzed and overall methylation levels are much lower than in the brain. In general, differences between tissues in methylation levels are not always correlated to gene expression (Shani *et al.*, 1984; Warnecke and Clark, 1999). The methylation differences between tissues may be due to the binding of tissue-specific repressors and/or differences in chromatin structure between tissues which may change accessibility to methyltransferases.

In mouse the paternal allele has become methylated to about the same level as the maternal allele in brain DNA, in the 3' end of the region analyzed. This suggests that there may be tissue-specific hypermethylation of *Ndn* in the brain, with the 5' end (CpG dinucleotides 1-17) being protected from this hypermethylation on the expressed paternal allele. The paternal allele-specific hypomethylation of the 5' region may be a consequence of transcription since promoter demethylation may come secondary to transcription (Matsuo *et al.*, 1998; Thomassin *et al.*, 2001). Alternatively, hypermethylation on the maternal allele may be a remnant of the initial gametic pattern or a reflection of other allele-specific factors such as chromatin structure. While promoter

methylation is usually correlated with gene repression, in some cases the downstream regions of genes are more methylated when the gene is active (Jones, 1999).

In human *NDN*, allele-specific DNA methylation is maintained both in tissues that express and do not express *NDN*. In lymphocytes the maternal allele is hypermethylated compared to the paternal allele and levels of methylation are similar to those seen in brain and fibroblasts. The hypermethylation found on both alleles in lymphoblasts may be due to an inherent increased methylation at CpG islands that is often found in cell lines (Antequera *et al.*, 1990; Jones *et al.*, 1990). This suggests that these lymphoblast cell lines do not represent the natural epigenetic state. In mouse *Ndn*, allele-specific methylation levels were more variable between tissues than for human *NDN*. These differences may be related to general tissue-specific methylation of necdin since the non-expressing tissues that I analyzed were not the same in human and mouse necdin. In addition, the differences in methylation patterns between mouse and human necdin could reflect some species specific differences in gene regulation. Human *NDN* is much more widely expressed (MacDonald and Wevrick, 1997) than mouse *Ndn* which is expressed preferentially in post mitotic neurons (Uetsuki *et al.*, 1996). The greater heterogeneity in methylation levels between clones in mouse brain may be related to the restricted expression pattern of *Ndn* because whole brain was used, which is composed of different cell types that do not uniformly express *Ndn*.

An alternative explanation for the differences between the human and mouse methylation patterns may be a higher stability of the maternal allele-specific hypermethylation in human *NDN*. The increased density of the human 5' CpG island compared to the mouse CpG island, in addition to possible differences in the DNA

sequence and chromatin context between human and mouse necdin may contribute to differences in CpG methylation levels between human and mouse *NDN/Ndn*.

The other difference between mouse and human was that in mouse, both alleles were relatively hypermethylated in the 3' end of the promoter proximal CpG island, with the allele-specific methylation being confined to the 5' end. In human *NDN*, the allele-specific methylation extends the entire length of the CpG island. Mouse and human necdin have in common allele-specific methylation in the 5' region of the promoter proximal CpG island. It seems that either maternal specific methylation or paternal specific protection from methylation or a combination of both are essential to maintaining the imprinted expression.

The absence of DNA methylation on the expressed paternal allele may reflect some epigenetic mark such as chromatin conformation or specific protein-DNA interactions that may serve as an imprint carried from the male germ line. For example, *SNRPN* and *H19* have regions of nuclease hypersensitivity on the unmethylated allele in regions that are methylated on the other allele (Khosla *et al.*, 1999; Schweizer *et al.*, 1999). Evidence to support the presence of the primary imprint on the unmethylated paternal allele for *NDN* is that, in both human and mouse, sperm DNA is completely unmethylated (This study, (El-Maarri *et al.*, 2001)). In human and possibly in mouse oocytes, the methylation pattern is quite variable with some unmethylated oocyte DNA strands. *SNRPN* also demonstrates a consistent unmethylated paternal allele in the gametes, but only mouse oocytes, and not human ooctyes are methylated in *SNRPN*. However, some imprinted genes in mouse acquire allele-specific methylation after fertilization (Brandeis *et al.*, 1993; Tremblay *et al.*, 1995). Recent evidence for spatial

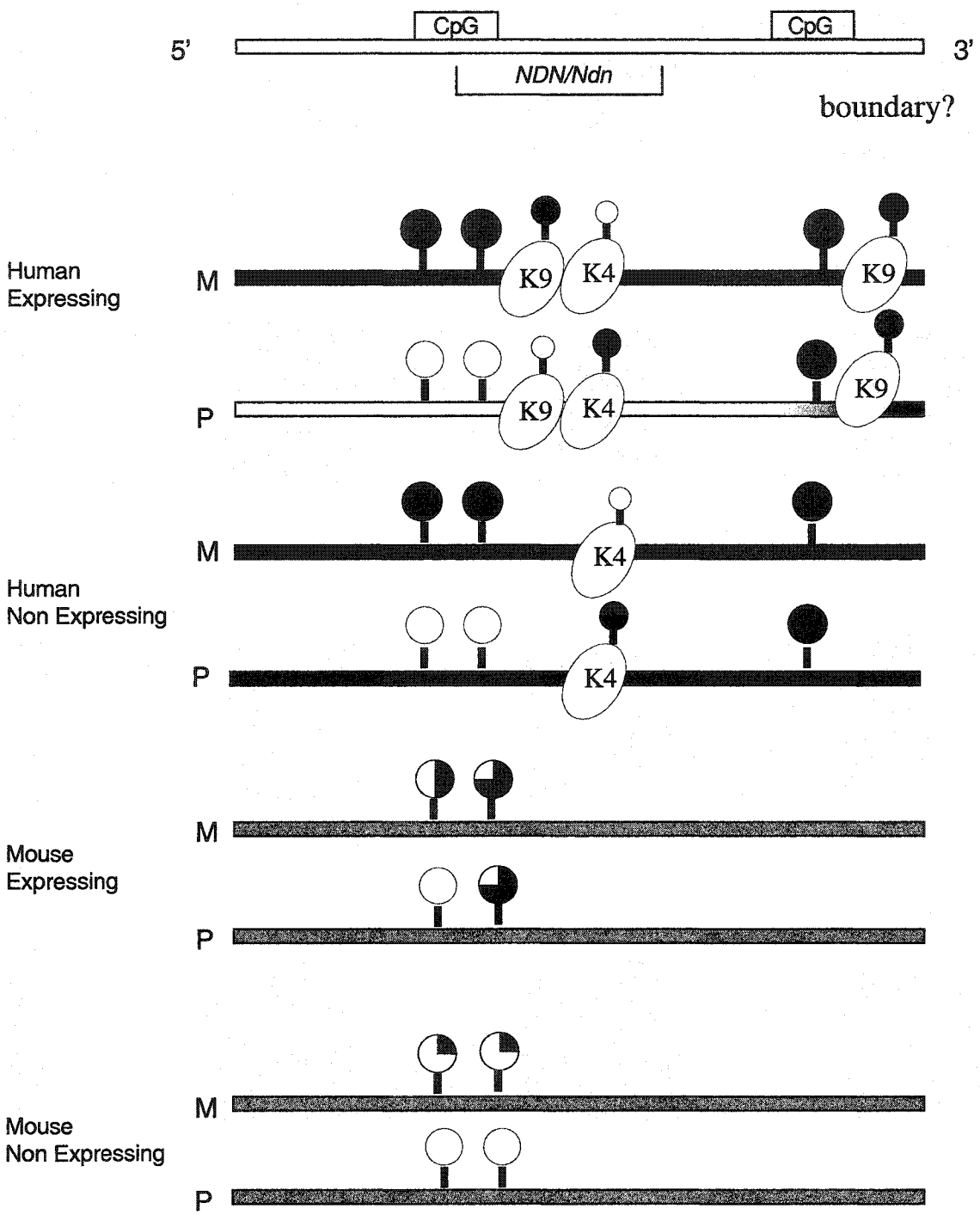
separation of the two parental genomes in the zygote suggests the possibility that methylation serves as the initial imprint even if methylation is not present at the gamete stage (Mayer *et al.*, 2000).

Interplay between DNA methylation and histone modifications

There is evidence that the epigenetic marks of DNA methylation and histone acetylation could serve to differentiate the paternal allele of *NDN*. Some recent work in our laboratory has demonstrated the presence of histone hyperacetylation and histone methylation of H3 lysine 4 on the active paternal allele (Jason Lau,(Hanel *et al.*, 2002)) (Figure 3-8). These are both marks of an open chromatin structure. The mechanism which maintains these epigenetic marks on the paternal allele may exclude DNA methylation on the paternal allele. Interestingly, histone methylation on lysine 9 has also been found on the inactive maternal allele of human *NDN*, which is a mark of heterochromatin. It is tempting to speculate that this could represent a lasting mark on the maternal allele of mouse *Ndn* in the blastocyst when allele-specific DNA methylation is lost.

Figure 3-8. Summary of DNA methylation, histone methylation and histone acetylation data on *NDN/Ndn*.

Large lollipops represent a group of CpG sites and the proportion of black fill indicates the relative level of DNA methylation in that region. The underlying bars represent the gene region and are white, black or gray to indicate hyperacetylated, hypoacetylated and unknown acetylation of the histones in the region. Methylation of lysines K4 and K9 are indicated by small lollipops on the K4 or K9 ovals. M and P are the maternal and paternal alleles respectively. The histone acetylation data were obtained by Jason Lau (Hanel *et al.* 2002) and the histone methylation data are from Jason Lau and Xin *et al.* (2001).



The 3' CpG island in human *NDN* may act as a boundary

The lack of allele-specific DNA methylation or histone acetylation of the 3' CpG island provided evidence against it having a role in the imprinting of *NDN* gene expression. This is not surprising since there was no equivalent CpG rich region in mouse. The sequence that composes the CpG island is internally repetitive and has homology to sequence in many other places in the genome. In general the methylation of repetitive DNA is not uncommon. Some examples are hypermethylation of multicopy transgenes (Garrick *et al.*, 1998; Lau *et al.*, 1999), expansion at the *FMR1* locus (Oberle, I *et al.* 1991, Hansen, R.S. *et al.* 1992) and methylation of satellite repeats (Ji, W *et al* 1997, Jeanpierre, M *et al.* 1993, Schuffenhauer, S *et al.* 1995). Interestingly, this CpG island co-localizes with the 3' end of a large domain of paternal histone hyperacetylation around the *NDN* locus (Jason Lau, (Hanel *et al.*, 2002)), suggesting that it may act as a domain boundary to the hyperacetylated domain on the paternal allele (Figure 3-8).

The role of DNA methylation in establishing the necdin imprint

In this chapter I have analyzed the developmental patterns of CpG methylation in the imprinted murine *Ndn* gene. Previous analysis of non-imprinted genes by restriction enzyme analysis (Kafri *et al.*, 1992) or SBS (Warnecke and Clark, 1999) suggested that partial removal of gametic methylation was found by the morula stage, and by the blastocyst stage little methylation was detectable. During this time frame, allele-specific methylation of a core DMR 2-4 kb upstream of the imprinted *H19* gene remains distinct until the blastocyst stage, and is only somewhat remodeled in midgestation embryos (Tremblay *et al.*, 1997; Warnecke *et al.*, 1998). The promoter proximal region of *H19*,

showed more variability in methylation levels and a lack of differential methylation at the blastocyst stage (Tremblay *et al.*, 1997).

Our studies showed that the 5' CpG-rich region of *Ndn* is unmethylated in sperm, but the paternal allele becomes gradually methylated resulting in partial mosaic methylation of the majority of DNA strands in adult tissues. The maternal allele is partially methylated in the majority of oocytes, and remains at about the same level of methylation at least to the morula stage. A reduced level of methylation of the maternal allele is seen in blastocysts, and *de novo* methylation occurs subsequently resulting in a mosaic pattern of methylation in adult tissues. Analysis of the maternally derived adult brain clones reveals that about 35% have the same or similar methylation patterns as the major pattern in oocyte, with overall higher levels of methylation at positions 19, 32, and 38 (Figure 3-3 and 3-7). Many paternal clones are also highly methylated in positions 19-38. The distinction between overall parental methylation levels is confined to sites 1-17 in the more 5' region. However, no well-defined pattern emerges to suggest that CpG methylation of either allele is the signal that carries the imprint through from the blastocyst stage to adult tissues. Thus, the methylation patterns seen in *Ndn* resemble the *H19* promoter proximal region in that no core region of differential methylation is maintained to the blastocyst stage. Differential methylation at a distant but linked site may play an important role in the imprinting of necdin.

One interpretation of our data is that the DNA methylation patterns originating in the oocyte are important early in imprinting, but that these methylation patterns are only required to initiate a series of additional epigenetic events leading to the silencing of the maternal allele. Interestingly, the most prominent oocyte pattern, the methylation of CpG

sites 13 through 22, covers 115 bp of DNA, in two segments of 36 (CpG sites 13 to 17) and 39 bp (CpG sites 19 to 22), with an intervening unmethylated space of 40 bp. Hypermethylation on the maternal allele in preimplantation embryos may block binding of a transiently expressed transcriptional activator or chromatin remodelling factor which is free to bind to the unmethylated paternal allele and maintain an active chromatin state (Figure 3-9B). Alternatively, a methylated maternal allele in the early embryo may bind methyl-CpG binding proteins and promote subsequent epigenetic events (Figure 3-9A). Methyl-CpG binding proteins have been shown to preferentially bind methylated CpG sites and recruit histone deacetylases (Wade *et al.*, 1999). If the two prominently methylated oocyte DNA segments were located on the outer part of the nucleosome as it winds around the core particle twice, the methylated DNA could potentially be a recognizable structural target for subsequent epigenetic events (Figure 3-9C). Alternatively, the positions that consistently lack methylation may reflect specific DNA-protein contact points. Studies of chromatin structure in early embryos are limited by difficulties in obtaining sufficient experimental material, but would help to prove or disprove this hypothesis.

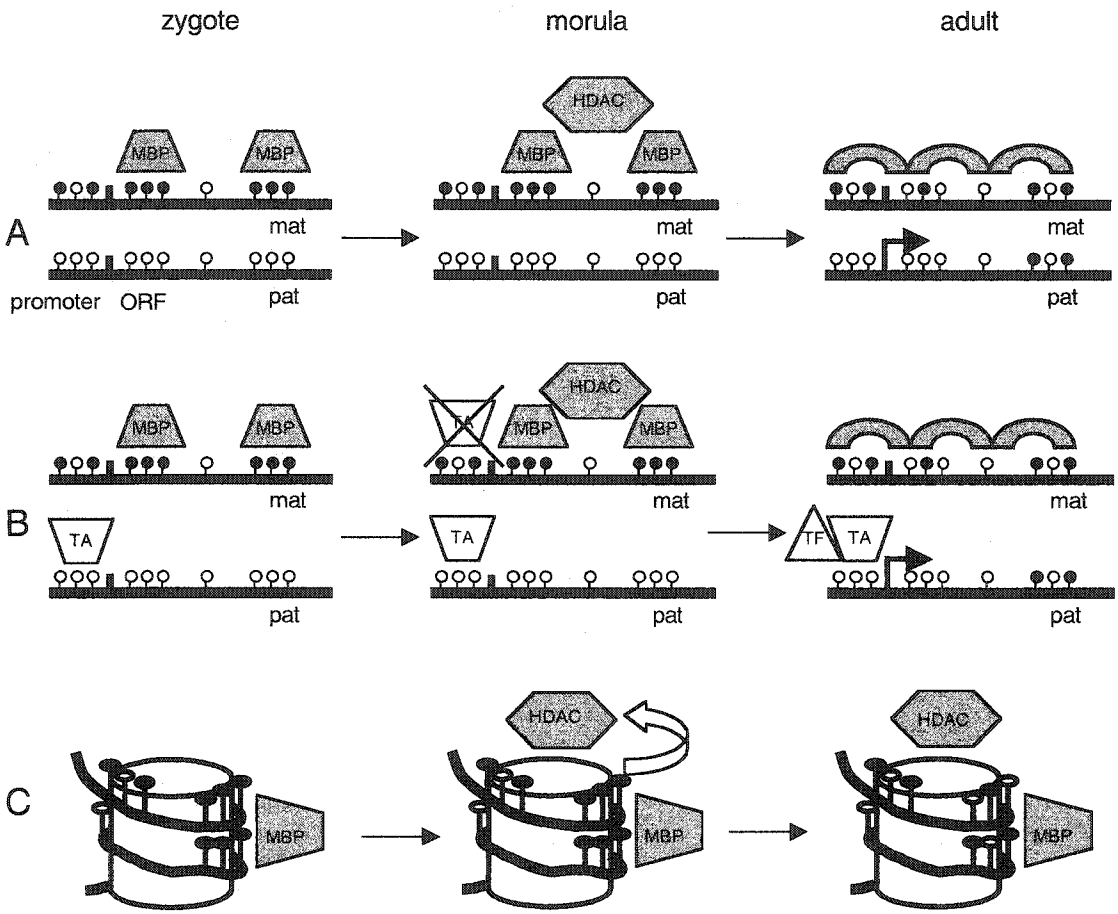


Figure 3-9. Models for imprinting of target genes.

The gametic methylation represents a simplified version of the average methylation seen at these stage. Black and white lollipops represent methylated and unmethylated CpG dinucleotides respectively. (A) The remodeling of the maternal allele to a closed chromatin conformation may involve methyl-binding proteins (MBPs) that recognize a target of densely methylated CpG sites which then recruit histone deacetylases (HDAC). Since the closed chromatin conformation is maintained by a methylation independent mechanism, the methylation state seen in the adult tissues can be variable. (B) The hypermethylation on the maternal allele blocks the binding of a transiently expressed transcriptional activator (TA) which binds only to the paternal allele. The TA keeps the paternal allele in a transcriptionally permissive state and allows transcription factors (TFs) to bind. In the absence of this transcriptional activator, the maternal allele is remodeled to a closed chromatin state. (C) The positions of methylated CpG sites 13 to 22 found in the early embryo samples may become closely situated on the surface of a nucleosome and act as a target for MBPs.

Our study contrasts with previous studies of the *Snrpn* and *H19* upstream DMRs in which maintenance of differential DNA methylation is proposed to be important throughout development. For mouse *Ndn*, no other CpG-rich clusters are found upstream or downstream of the region analyzed over a distance of about 15 kb, although it is possible that isolated CpG sites outside of the region studied may play a role in the imprinting process. For *Snrpn* and *H19*, their DMRs are presumed to have a direct role in the imprinting process, rather than being a target of a separate imprinting center (Barlow, 1997). In addition, the *Snrpn* imprinting center and its human equivalent have recently been shown to be important in imprint maintenance as well as establishment, which could explain why they show a stable differential methylation throughout development (Bielinska *et al.*, 2000). Another imprinted gene, *Mash2*, was found to be unaffected by loss of methylation in mice deficient for the DNA methyltransferase gene (*Dnmt*) (Caspary *et al.*, 1998; Tanaka *et al.*, 1999) and like *Ndn*, is not known to contain an imprinting center for other genes. *Mash2* and *Ndn* may fall into a category of genes whose imprinted expression is not maintained by methylation, but by some other imprinting control mechanism such as allele-specific chromatin structure. In support of this hypothesis, *Ndn* gene expression does not respond to demethylation during treatment of androgenetic and parthenogenetic mouse fibroblasts with aza-cytidine (C. Stewart, personal communication)(El Kharroubi *et al.*, 2001). Further developmental characterization of imprinted genes that are targets of imprinting centers will be an important step towards understanding the role of DNA methylation in imprinting and how genes under the control of an imprinting center are regulated.

Chapter 4. Allele-specific chromatin structure and DNA binding proteins in the 5' region of necdin

Hanel, M.L., Paradis, I., Drouin, R., and Wevrick, R. Allele-specific chromatin structure and DNA binding proteins in the 5' region of necdin (manuscript in preparation).

Introduction

Some imprinted genes show differences in chromatin structure, where nuclease hypersensitive sites are usually found on the active allele in the equivalent region that is methylated on the inactive allele (Bartolomei *et al.*, 1993; Hark and Tilghman, 1998; Khosla *et al.*, 1999; Schweizer *et al.*, 1999). Regions of DNase I hypersensitivity often coincide with disrupted nucleosome positioning in regulatory regions of genes (Bresnick *et al.*, 1992). This can correlate with binding of a regulatory protein to the DNA. For example, with *H19-Igf2*, a region in the IC is methylated on the paternal allele and is bound by the methylation sensitive insulator, CTCF, on the active and unmethylated maternal allele (Bell and Felsenfeld, 2000; Hark *et al.*, 2000; Kanduri *et al.*, 2000). Not all protein bound regions of DNA will produce nuclease sensitive sites. It may depend on the number and types of DNA binding proteins present (Boyes and Felsenfeld, 1996). For example, the promoter region of the active maternal allele of *H19* does not have any hypersensitive sites in adult liver, but there are promoter hypersensitive sites in neonatal liver where *H19* is expressed at high levels (Khosla *et al.*, 1999). In addition to nuclease hypersensitive sites, some genes can contain extended regions of generalized increased sensitivity (Hebbes *et al.*, 1994). The imprinted *U2af1-rs1* gene contains paternal allele-specific hypersensitive sites in the 5' untranslated region, but also displays a generalized increased DNase I sensitivity on the paternal allele that extends past the open reading frame (Feil *et al.*, 1997).

Allele-specific epigenetic modifications on imprinted genes are responsible for imparting allele-specific expression. Allele-specific modifications, such as DNA

methylation may attract additional epigenetic modifications such as methyl-binding proteins (MBPs), which attract histone deacetylase activity (Nan *et al.*, 1998). It is thought that some DNA binding proteins can protect a region from DNA methylation, which would be important to keep the active allele from being inactivated by DNA methylation linked mechanisms. A transcriptional activator has been discovered that contains a CpG binding domain and specifically recognizes unmethylated CpG sites (Voo *et al.*, 2000). Some imprinted regions contain specific DNA elements that protect a region from DNA methylation, which is likely to occur through DNA-protein interactions. A dyad of Oct-binding sequences in the *H19-Igf2* imprinting control region seems to protect CTCF sequences from *de novo* methylation (Hori *et al.*, 2002). The *Igf2r/Air* imprinted locus also contains protein binding elements that are protective against DNA methylation (Birger *et al.*, 1999). It is likely that allele-specific modifications on the parental alleles create a situation where transcription factors can only bind to one allele. Preferential binding of putative transcription factors to the promoter region of *H19* on the active maternal allele has been shown (Szabo *et al.*, 1998).

I previously analyzed the 5' CpG islands of *NDN/Ndn* for allele-specific methylation and found that the maternal allele was hypermethylated compared to the paternal allele in adult tissues. DNase I sensitivity mapping is a technique that allows the long range mapping of the presence of a generalized open or closed chromatin structure. It can also detect specific nuclease hypersensitive sites, which may represent gene regulatory elements. I used the technique of DNase I sensitivity mapping to analyze *NDN/Ndn* for differences in nuclease sensitivity.

In vivo footprinting with ligation mediated PCR (LMPCR), is a highly sensitive technique that can identify possible protein-binding sites and also identify subtle differences in chromatin structure. To search for allele-specific DNA binding proteins and more specific differences in chromatin structure in the *NDN* promoter, I performed *in vivo* footprinting on Prader-Willi Syndrome (PWS) and Angelman Syndrome (AS) cell lines, which are deleted for the paternal and maternal copies of chromosome 15 respectively. I used DNase I treatment, dimethyl sulfate (DMS) treatment and UVC irradiation to identify footprints within the 5' region of *NDN*, which includes a CpG island and a previously defined neuron-restrictive promoter (Nakada *et al.*, 1998).

Allele-specific DNA binding proteins identified by these methods may be important in the imprinting process. In addition, allele-specific epigenetic modifications such as DNA methylation or a closed chromatin structure may exclude transcription factors from binding to the maternal allele of *NDN*. I investigated epigenetic modifications such as chromatin structure and DNA binding proteins and found that these mark the active paternal allele of *NDN/Ndn*.

Results

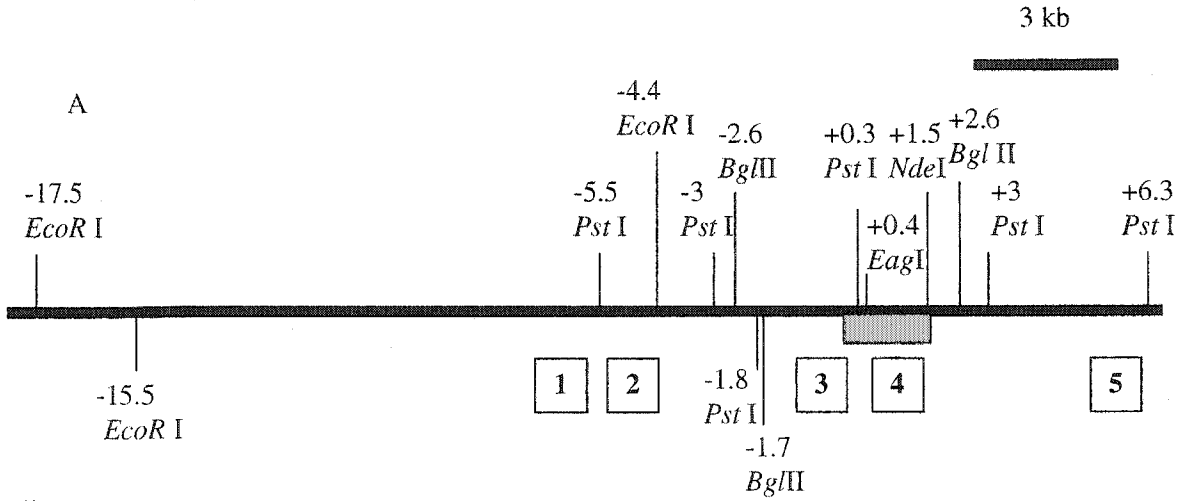
Increased DNase I sensitivity on the paternal allele of mouse *Ndn*

I first identified restriction site polymorphisms between C57BL/6 and SPRET mice within the region (Figure 4-1). I analyzed mouse *Ndn* using tissues from interspecific F1 crosses between C57BL/6 and *M. spretus* (SPRET) in which I was able to distinguish the parent of origin of the two alleles. The use of polymorphic restriction sites allowed a comparison of DNase I sensitivity between the two parental alleles within

the same tissue sample. After treatment of isolated nuclei with increasing amounts of DNase I, I isolated the DNA and digested to completion with a restriction enzyme that when probed, produces separate bands for the maternal and paternal alleles. I analyzed the putative promoter region in nuclei isolated from brain and liver. Allele-specific increased sensitivity on the paternal allele compared to the maternal allele was observed in brain, but not in liver. In liver, both alleles were resistant to digestion indicating that for *NDN*, DNase I sensitivity correlated with gene expression. I then expanded my analysis in brain to determine whether the allele-specific DNase I sensitivity was localized to the promoter. In brain I found a general DNase I sensitivity over a large region of about 22 kb (Figure 4-1). There were no obvious hypersensitive sites, and the entire region seemed to have the same relative sensitivity on the paternal allele. The active paternal allele of *NDN* appears to be regulated as a large domain.

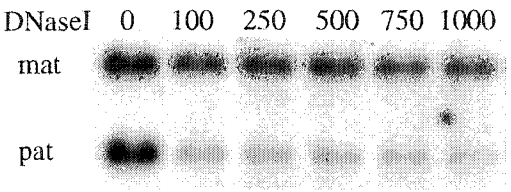
Figure 4-1. DNase I analysis of mouse *Ndn*.

(A)*Ndn* (open box) and relative positions of restriction enzyme sites to the transcription start site are shown. Restriction sites common to both mouse species are on top of the line and shown in black. Restriction sites found only in C57BL/6 are on top and in red and restriction sites found only in SPRET are on the bottom and in blue. Boxes numbered 1 to 5 indicate the positions of probes. (B) Southern blots for brain and liver DNA isolated from DNase I treated nuclei, probed with the open reading frame probe. The maternal C57BL/6 fragment and the paternal SPRET fragments are labeled mat and pat. (C) Southern blots for brain DNA isolated from DNase I treated nuclei. DNase I probed with 5' and 3' probes. Numbers indicate units of DNase I/ml.

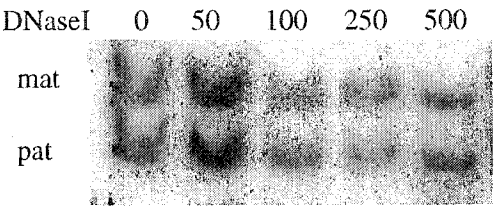


B

Probe 4. *Bgl* II



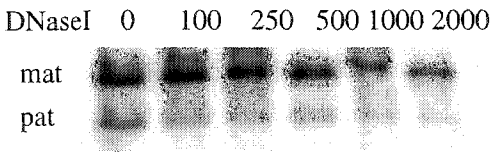
Brain



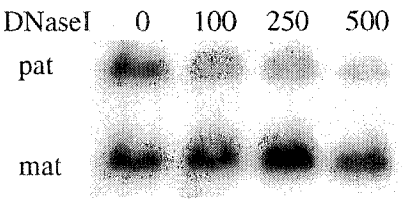
Liver

C

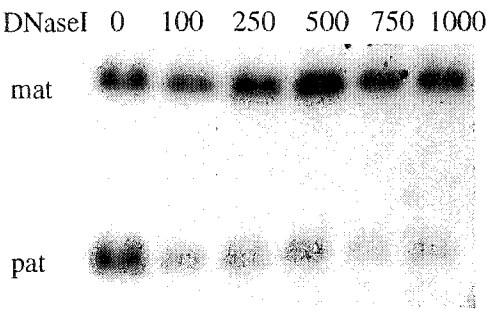
Probe 1 *EcoR* I



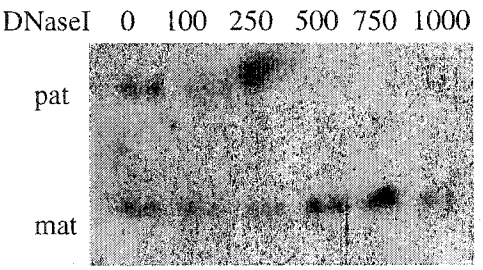
Probe 2. *Pst* I



Probe 3. *Bgl* II and *Pst* I



Probe 5. *Pst* I



Increased sensitivity of the paternal allele in human *NDN*

I also performed a less extensive analysis on human *NDN*. I analyzed the maternal and paternal alleles by using PWS and AS deletion fibroblasts that are deleted for the paternal and maternal 15q11-q13 region respectively. Since I treated the maternal and paternal alleles in separate cell lines, I controlled for equal DNase I digestion and equal loading between samples by probing all blots with a heterochromatin probe for satellite 2 (sat2) (Hansen *et al.*, 2000). In particular, at the 100 U/ml concentration of DNase I, sat2 bands are of relatively equal intensity for both alleles, but the maternal allele of *NDN* is clearly more resistant than the paternal allele (Figure 4-2A). The hybridization bands were measured by densitometry and ratios of sat2/*NDN* band intensity for 0, 50 and 100 U/ml were calculated as shown graphically in Figure 4-2B, where I have subtracted the ratio at 0 U/ml in each data set from the subsequent ratios so that the graph starts at 0. In fibroblasts the sat2 band was lost in the 250 U/ml DNase I lanes for PWS and AS cell lines, while the *NDN* band was still visible. In this lane the sat2 probe detected a slightly lower molecular weight smear, indicating that the satellite 2 DNA was highly degraded. This could be due to the fact that the sat2 probe detects a very high molecular weight fragment (>12 kb), while the *NDN* probe detects a smaller fragment (~2.6 kb) and the higher molecular weight DNA becomes quite digested at this concentration of DNase I. My DNase I sensitivity analysis confirmed that human *NDN* had a general difference between the two alleles in chromatin structure (Figure 4-2). This is consistent with the recent characterization of histone acetylation, that demonstrated a large domain of hyperacetylation on the paternal allele of *NDN*, which is also a mark of open chromatin (Chapter 3-experiments done by Jason Lau) (Hanel *et al.*, 2002).

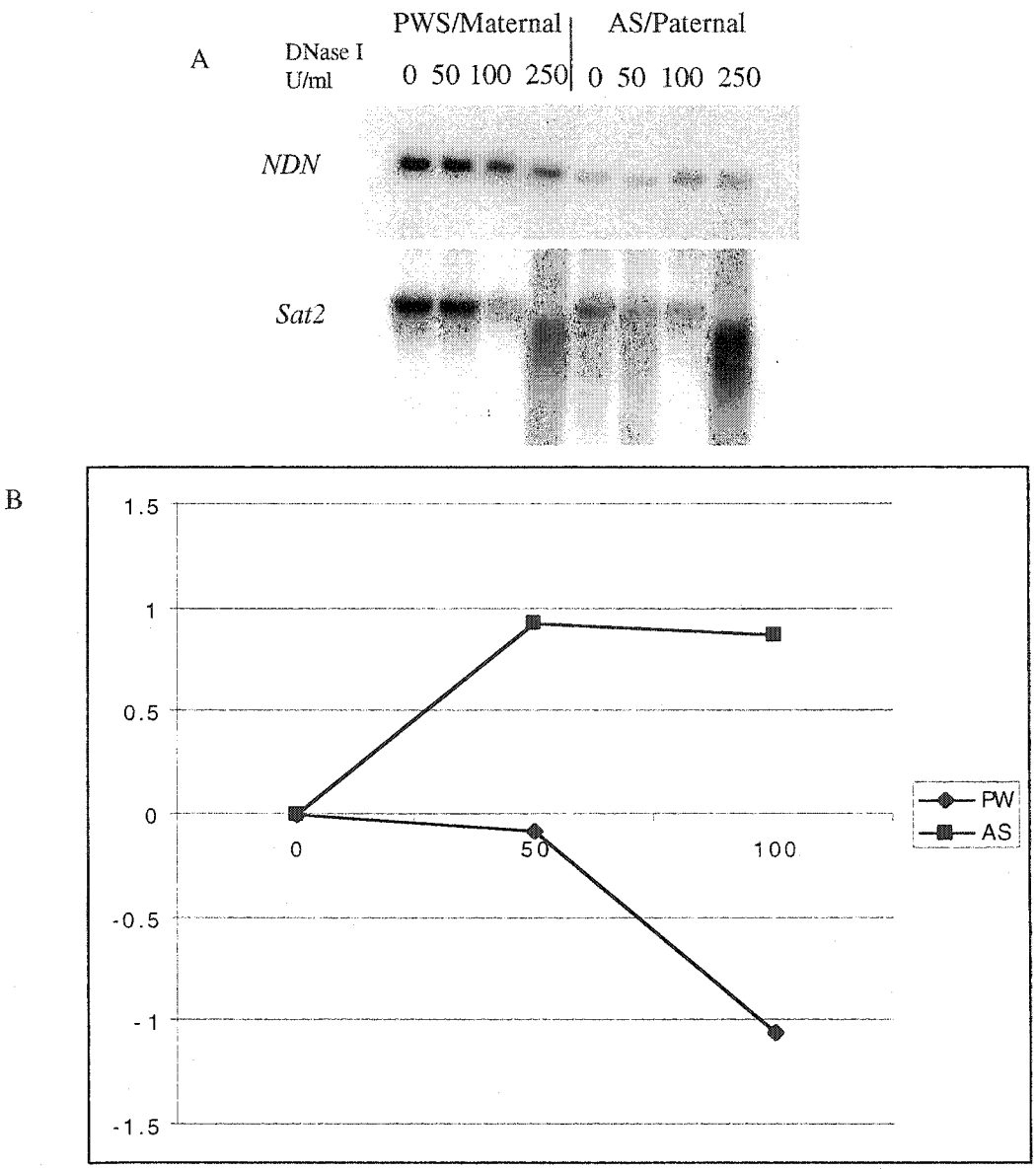


Figure 4-2. DNase I analysis of human *NDN*.

(A) Southern blot of DNA from fibroblasts. The DNA was digested with *Bgl* II and probed with a human *NDN* probe yielding a 2.6 kb band. (B) Graphical representation of densitometry analysis of the above Southern blots. Ratios of densitometry values Sat2/*NDN* were calculated. The value at zero was subtracted from each.

In vivo footprinting in the *NDN* promoter region

In the human and mouse DNase I sensitivity analysis I did not detect any hypersensitive sites. In order to look at the more specific aspects of allele-specific chromatin structure, and to identify possible binding sites for regulatory proteins, I used the highly sensitive technique of *in vivo* footprinting analysis with LMPCR. For this analysis, I used the PWS and AS fibroblast and lymphoblast cell lines. I focused on the putative promoter region which contains the upstream part of the 5'CpG island (Figure 4-3).

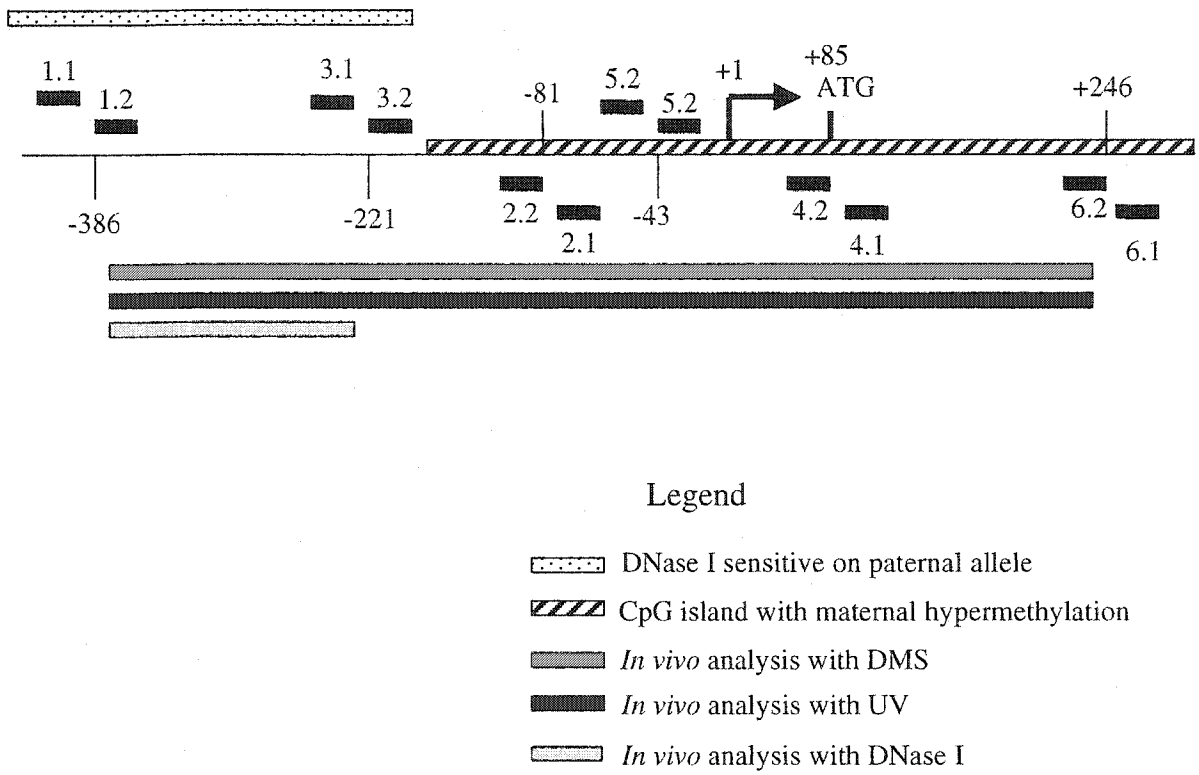


Figure 4-3. The 5' region of *NDN* analyzed by DNase I sensitivity and *in vivo* footprinting.

The black lines represent primer positions with the primer names (ie NEC1.1 shown as 1.1) shown above each line. Each pair of closely spaced primers consist of the more 5' extension step primer and the PCR primer. The relative positions of each of the PCR primers to the transcription start site are shown. The primers on top amplify the sense strand primers on the bottom amplify the antisense strand. The striped bar represents the CpG island.

The *in vivo* analysis included DNase I, UV and DMS treatments. It was important to use multiple methods to detect footprints, since each method has advantages and disadvantages. DNase I is not sequence-specific, and therefore can detect all regions bound by protein. It can also identify regional changes in chromatin structure, but does not detect the exact nucleotide positions of protein-DNA contact points. DMS and UVC can detect specific nucleotide bases involved in DNA-protein interactions, but can only detect footprints at guanine residues and at pyrimidine dimers respectively. However DNase I treatment is more disruptive to the cells than DMS and UVC because it is a large molecule that requires permeabilization of the cells.

The cell lines were treated with one of the three agents, then the DNA was extracted and amplified by LMPCR (Figures 2-2 and 2-3). The entire 632 bp region was analyzed by UV and DMS. I attempted to analyze the entire region by DNase I as well, but due to the high CpG content, it was not possible. There are potentially two reasons why DNase I was more difficult in the CG rich region than the other treatments. One reason may be that DNase I is the harshest treatment of the three treatments and would result in a lower quality of DNA to use for the subsequent LMPCR, which is already challenged by high CG content DNA. Another challenge may be that DNase I generates small 10-bp fragments that may also interfere with the LMPCR (Drouin, R., personal communication). Therefore only the 5' region covered by primer NEC1.2, which is 5' of the CpG island, was analyzed by DNase I. Using primer NEC1.2 I identified DNase I footprints on both alleles of *NDN* in fibroblasts and lymphoblasts (Figure 4-4). Using primer NEC3.2, located just upstream of the CpG island and upstream of the transcription start site, I identified footprints by DMS and UV that were present only on the active

paternal allele in fibroblasts (Figure 4-5). Recognition sites for transcription factors were identified using MatInspector (http://transfac.gbf.de/cgi-bin/matSearch/matsearch.pl). A summary of the footprints and nearby potential transcription factor binding sites are shown in Figure 4-6. The CpG sites that are near or within the footprinted region are CpG sites 15, 16 and 17 in Figure 3-5.

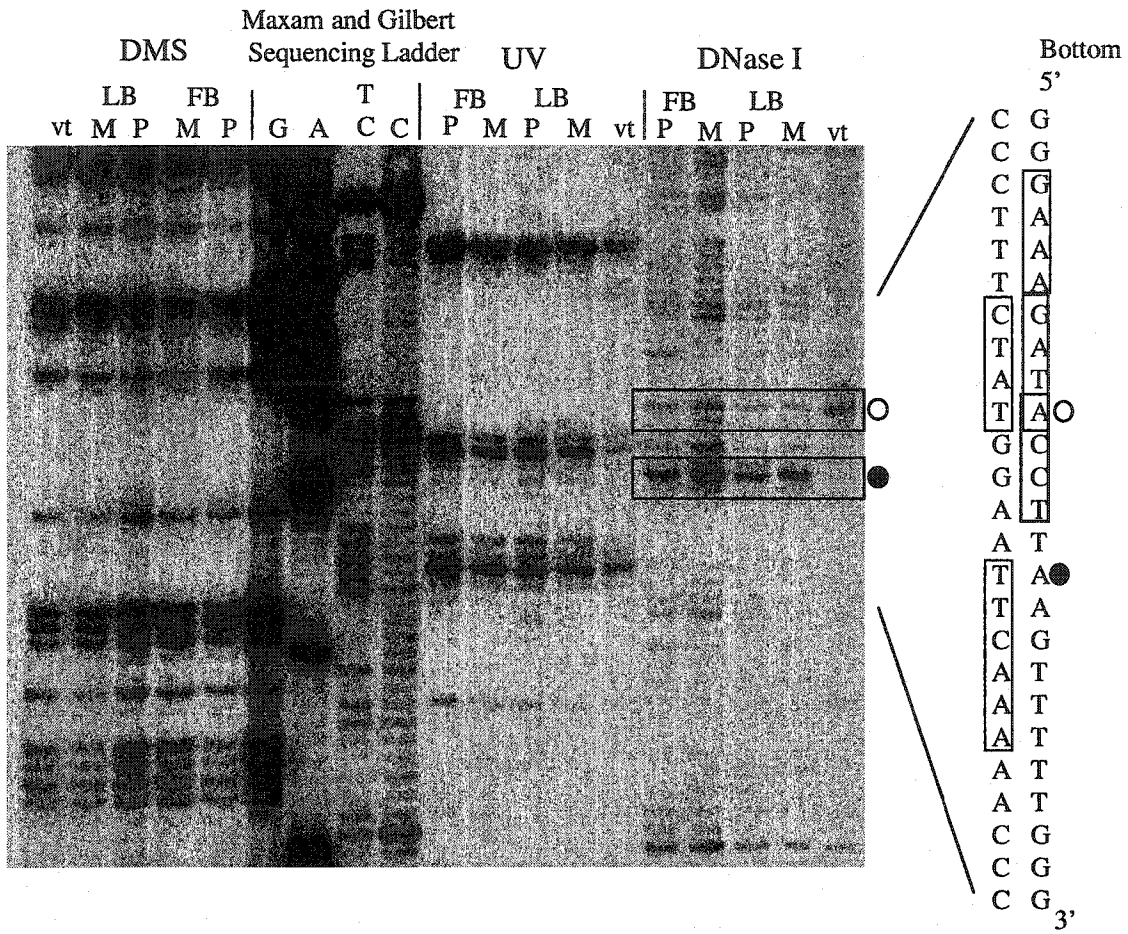


Figure 4-4. Detection of DNase I footprints on the bottom strand in all four cell lines.

Extension was off the bottom strand with the top strand primer NEC1.1. The gel was transferred to membrane and then probed with top strand primer NEC1.2. LB Lymphoblasts, FB Fibroblasts, M maternal allele, P paternal allele, vt *in vitro* DNA. Both footprints affect A residues in all four cell types. An open circle represents a protected base and a black circle represents a more exposed base compared to the *in vitro* lane DNA. The sequence is shown to the right. It is read from 3' to 5' on the bottom strand. Boxes highlight core recognition sequences for known transcription factors (see Figure 4-6 for details).

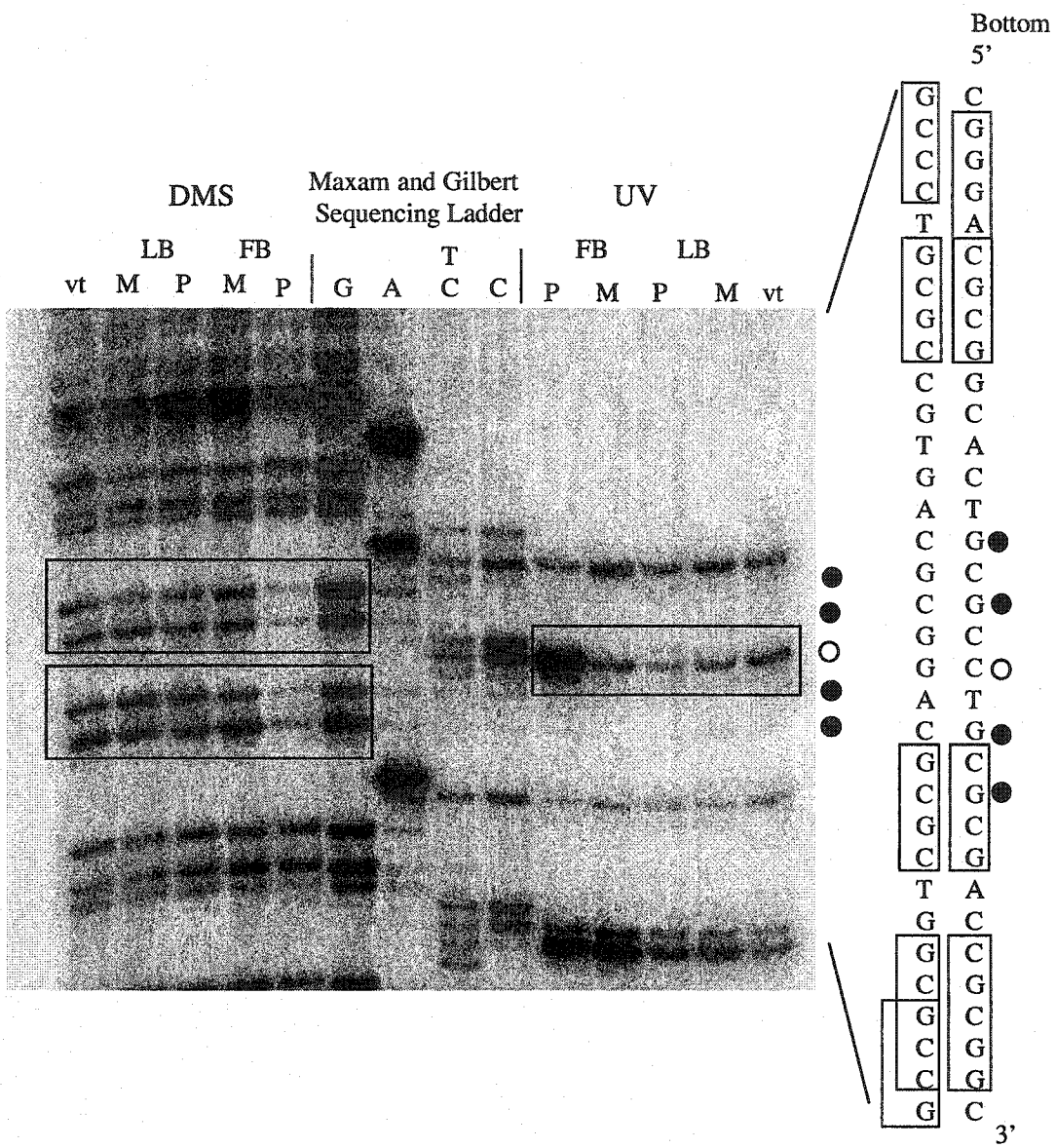


Figure 4-5. Detection of DMS and UV footprints on the bottom strand of the paternal allele in fibroblasts.

Symbols are the same as for Figure 4-4. Extension was off the bottom strand with the top strand primer NEC3.1 and the blot was probed with NEC3.2.

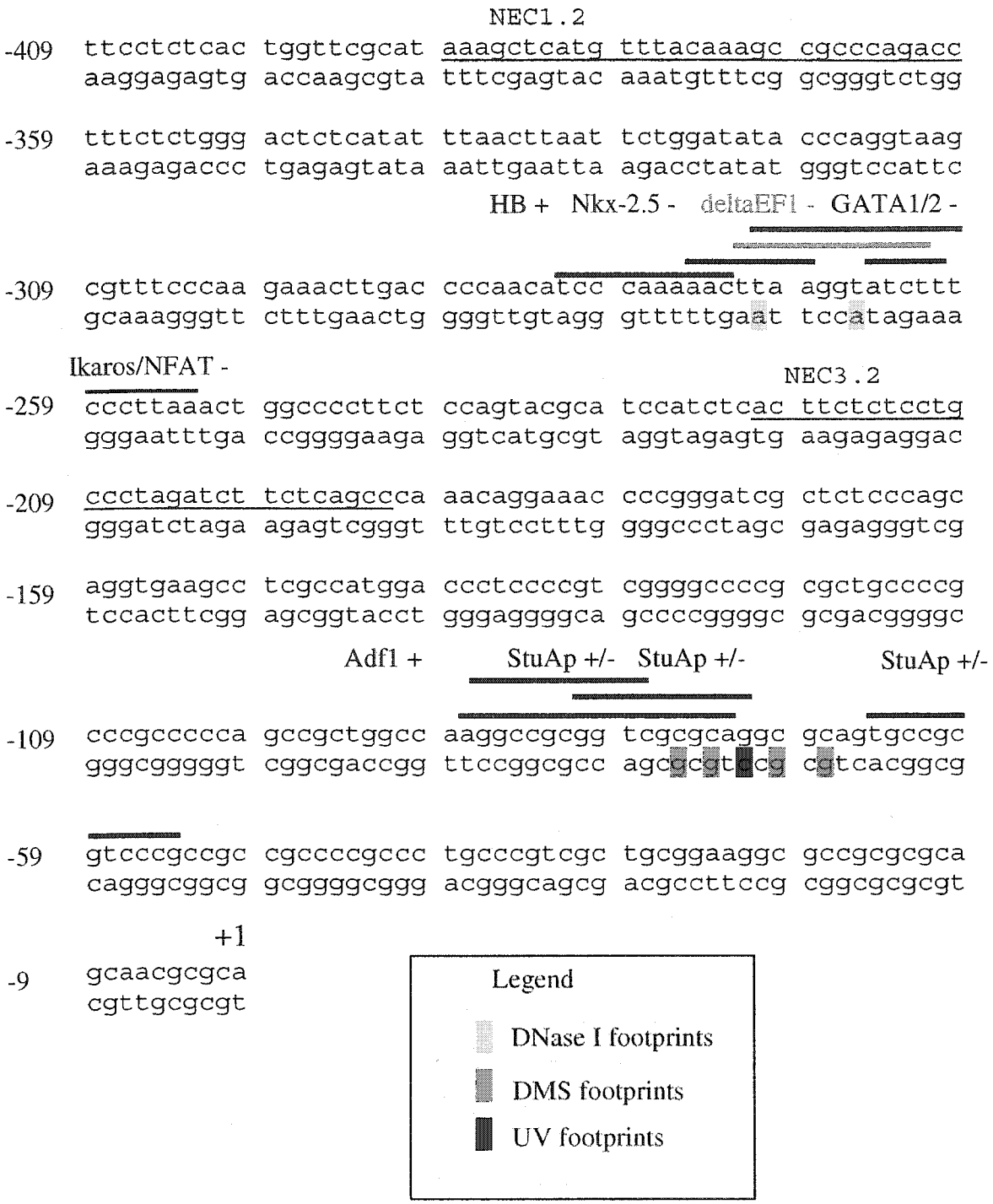


Figure 4-6. Summary of *in vivo* footprint data.

Upstream sequence to transcription start site is shown. Primers NEC1.2 and NEC3.2, which were used to detect these footprints, are underlined. Footprints are highlighted. The recognition sites for transcription factors in the regions of the footprints are shown. + or – indicate transcribed or non-transcribed strand.

Potential regulatory proteins bound to both alleles in fibroblasts and lymphoblasts

Without further experiments, there is no way to know which, if any of these potential DNA binding proteins actually bind to the *NDN* promoter. In the region of the footprints identified with primer NEC1.2 (Figure 4-4), recognition sequences for GATA-1 and/or GATA-2 completely overlap the footprinted nucleotide bases. The GATA family is named for containing the nucleotides G A T A in the consensus binding sequence (Merika and Orkin, 1993). GATA-1 is a zinc finger transcription factor known to be important in erythropoiesis (Cantor and Orkin, 2002). GATA-2 is more widely expressed (Merika and Orkin, 1993). For example, GATA-2 is expressed in neuronal cell types including a gonadotrophin-releasing hormone (GnRH)-secreting hypothalamic neuron derived cell line (Lawson *et al.*, 1996). GATA-2 also seems to play a role in pituitary cell type determination (Dasen *et al.*, 1999). A potential binding site for the zinc finger, basic-helix-loop-helix (bHLH), homeodomain protein, dEF1 was also found to overlap the two DNase I footprints. dEF1 has been shown to be expressed mainly in mesoderm, neural tissues and the lens in mouse (Funahashi *et al.*, 1993). It has been shown to be a repressor which down regulates collagen I and II promoters and a chondrocyte protein (Davies *et al.*, 2002; Murray *et al.*, 2000; Terraz *et al.*, 2001). Other potential binding proteins are the transcription factor Nkx2-5 homeoprotein, involved in cardiac development (Kasahara *et al.*, 2001) and the *Drosophila* transcription factor hunchback (HB), which is involved in the development of anterior structures (Schmidt-Ott, 2001). Another widely expressed transcriptional regulator, nuclear factor of activated T cells (NFAT), which is controlled by calcineurin is also located around these DNase I

footprints. Members of the NFAT family seem to have a role in a wide variety of cell types including adipocytes, chondrocytes and muscle (Horsley and Pavlath, 2002). The binding of Ikaros would be consistent with the absence of *NDN* expression in lymphocytes since it is associated with transcriptionally silenced genes in lymphocytes (Brown *et al.*, 1997; Kim *et al.*, 1999).

Potential regulatory proteins that preferentially bind to the paternal allele in fibroblasts

The DMS and UV footprints identified with the primer NEC3.2 (Figure 4-5) are more interesting from the point of view of genomic imprinting since they were found only on the paternal allele in fibroblasts. No known mammalian transcription factor recognition sites were found to co-localize with the footprints. However multiple binding sites for the protein called stunted in the fungus *Aspergillus nidulans* (StuAp) were found. This bHLH transcription factor regulates developmental and cell cycle genes in *A. nidulans* (Dutton *et al.*, 1997). Interestingly, it acts as a transcriptional repressor in *A. nidulans*, but as an activator in budding yeast (Dutton *et al.*, 1997). The other recognition sequence in this vicinity was for the *Drosophila* transcription factor Adf*1*. Adf1is involved in synapse formation and olfactory memory (DeZazzo *et al.*, 2000). Our footprints in the AS fibroblasts may represent the binding of the human homologues to these transcription factors, or other as yet unidentified human transcription factors may share similarities in their recognition sequence.

Discussion

Higher order regulation of *NDN/Ndn*

I have demonstrated that human and mouse *NDN/Ndn* are both more sensitive to DNase I on the active paternal allele. In mouse *NDN*, this encompasses a large domain of increased DNase I sensitivity. In human *NDN*, a large domain of open chromatin has also been demonstrated by increased H3 and H4 histone acetylation in cells that express *NDN* (acetylation experiments were done by Jason Lau) (Hanel *et al.*, 2002). The large domain of open chromatin structure on the paternal allele of *NDN/Ndn* suggests a higher order of gene regulation. Domain-wide histone acetylation has only been found for other tissue-specific genes. DNase I sensitivity and histone acetylation in *NDN/Ndn* seem to correlate with gene expression and may be related to tissue-dependent gene regulation. If *NDN/Ndn* is regulated by enhancers and domain insulators, the IC may regulate imprinting of *NDN/Ndn* by targeting these other elements in addition to the promoter region.

It is possible that on the active paternal allele, an open chromatin conformation encompasses the entire PWS region and that in tissues where *NDN/Ndn* is not expressed a localized closed chromatin structure is attained. The β-globin multi-gene locus displays a similar large domain of open chromatin, which is organized in a domain bounded by recognition sites for the insulator CTCF (Farrell *et al.*, 2002; Saitoh *et al.*, 2000) and containing a locus control region (LCR), which enhances transcription within the region (Bulger *et al.*, 2002). The fact that clusters of imprinted genes depend on imprinted regulation from single imprinting centres, which often contain CTCF sites, suggests a similar domain-wide control in imprinted regions. Recently it has been found that the

H19-Igf2 region contains an additional CTCF binding insulator element at the 3' boundary to the imprinted domain (Ishihara and Sasaki, 2002). The region between these two CTCF insulators contains the maternally expressed *H19* gene and one biallelically expressed trascriptional unit called *Nctc1*. This may suggest regional control of a predominantly maternally expressed domain with *Nctc1* escaping imprinting. It would be interesting to extend the analysis of chromatin structure around the *NDN* locus in the PWS region to include the neighbouring imprinted *Magel2* gene to establish if the PWS-IC controls imprinting as multiple domains, or as a single domain.

Potential regulatory proteins bind both alleles of *NDN* in fibroblasts and lymphoblasts

The promoter region was previously defined using a reporter assay in neurally differentiated P19 cells (Nakada *et al.*, 1998). These authors defined the minimal postmitotic neuron-restrictive promoter elements to be located between -177 and -33 from the transcription start. The biallelic DNase I footprints that I found in fibroblasts and lymphoblasts were located about 260 bp upstream of the transcription start site. The potential transcription factor recognition sites that co-localized with these footprints all correspond to candidate DNA-binding proteins that could regulate *NDN*. There may be some relevance to GATA-2 regulating *NDN*, which is primarily involved in neuron differentiation. The expression of GATA-2 in a hypothalamus cell line and the role of GATA-2 in pituitary development are of particular interest since the hypothalamus and pituitary are affected tissues in Prader-Willi Syndrome (Burman *et al.*, 2001). A specific role for necdin in the development of these cell types is likely since a *NDN* knock-out mouse was shown to have reduced numbers of specific hormone releasing cells in the

hypothalamus (Muscatelli *et al.*, 2000). One possible link between the dEF1 factor, involved in bone and cartilage development, to Prader-Willi Syndrome is that affected individuals have short stature and typical facies (Gunay-Aygun *et al.*, 2001). The possible role of NFAT in adipocytes is interesting since *NDN* could have a role in obesity, the major characteristic of Prader-Willi Syndrome individuals. One paper reported expression of *NDN* in proliferating brown preadipocytes (Boeuf *et al.*, 2001). The possible significance of the other recognition sites near the DNase I footprints to the regulation of *NDN* is unknown.

Allele-specific DNA-protein interactions in the differentially methylated domain

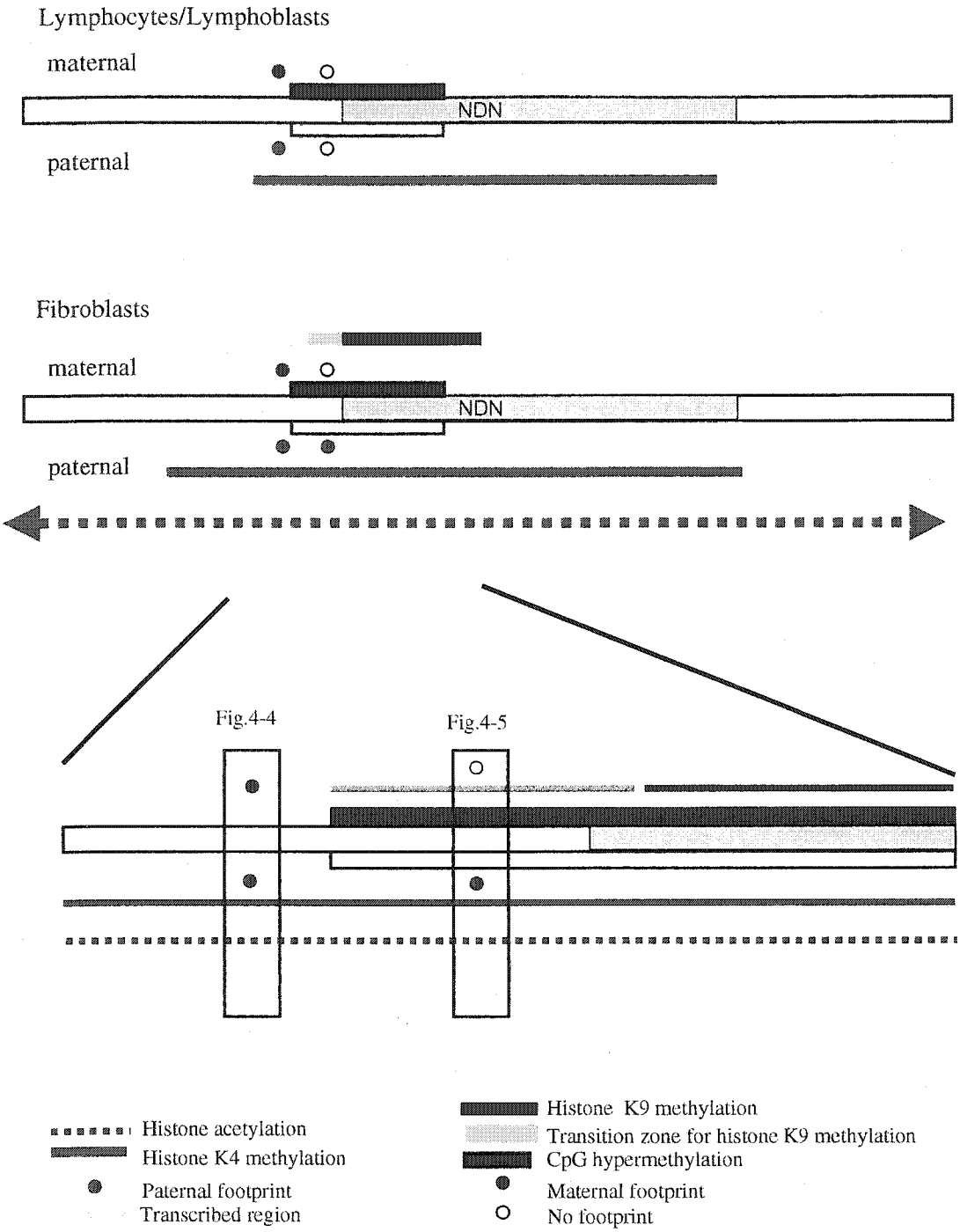
The paternal allele-specific footprints were located at about 80 bp upstream of the transcription start site, within the maternally methylated CpG island. This may suggest that the DNA binding proteins represented by these footprints are methylation sensitive. In support of this, DNA methylation of the *NDN* promoter on reporter constructs caused transcriptional repression (Nakada *et al.*, 1998). The recognition sequences that co-localized with the footprints did not correspond to any known mammalian transcription factors. The footprints could thus represent a novel transcriptional regulator or a human homologue to Adf1 and/or StuAp.

Future experiments, such as ligand screening of an expression cDNA library could identify proteins that bind to my footprinted region. If human homologues of Adf1 and /or StuAp could be found, antibodies for these could be used in electromobility shift assays to prove that the transcription factors actually bind.

How the presence of an allele-specific DNA binding protein(s) fits into the epigenotype of *NDN* is not clear, but some interesting correlations can be made. Figure 4-7 shows a summary of all of the allele-specific marks in the *NDN* promoter region. In fibroblasts, the biallelic footprinted region is outside the CpG island, but within the paternal-specific histone K4 acetylation, histone acetylation and DNase I sensitivity domains. This suggests that at least some DNA regulatory proteins can access both parental alleles with equal efficiency even in the presence of allele-specific epigenetic modifications.

Figure 4-7. Summary of all epigenetic analysis in the 5' region of *NDN*.

Epigenetic modifications on the maternal allele and paternal allele are shown on the top and bottom of each representation of the *NDN* locus. The two footprinted areas are marked by boxes to focus on the layers of epigenetic modifications at each position. Blue and red symbols indicate paternal or maternal-specific or biased epigenetic marks respectively. Pink designates a transition zone between maternally biased K9 methylation and unbiased K9 methylation.



One hypothesis is that allele-specific marks that are present in a tissue that does not express *NDN* would more likely represent true imprint marks than allele-specific marks in a tissue that expresses *NDN*. The paternal allele-specific marks in a tissue that expresses *NDN* may be related to active gene expression or to the imprinting process. I have analyzed fibroblasts (expressing) and lymphoblasts (non-expressing) for *in vivo* footprints and did not find any allele-specific differences in lymphoblasts. We have shown that cultured lymphoblasts have altered allele-specific epigenetic marks from those present in lymphocytes, at least in the form of DNA methylation and histone K4 methylation (Chapter 3) (Hanel *et al.*, 2002). The paternal allele in lymphoblasts does not contain any marks of open chromatin as it does in lymphocytes. It is possible that the paternal allele-specific footprints that I have found here in fibroblasts may be present in lymphocytes. It would be interesting to test this, but it may be difficult to obtain sufficient quantity of blood from PWS and AS patients for *in vivo* footprinting on lymphocytes.

The paternal allele-specific DNA binding protein could have an early role in the establishment of the imprint on the active paternal allele by binding in the sperm or early embryo, and it could protect this region from the maternal-specific DNA methylation that is set up in the zygote. Alternatively, this DNA binding protein may be a trancriptional activator that is blocked from binding to the maternal allele by the DNA methylation. The presence of this activator may subsequently protect this region from *de novo* methylation during embryogenesis. The domain of open chromatin represented by histone K4 methylation also overlaps with the binding site of this paternal allele-specific factor. The

binding of this paternal allele-specific DNA binding protein may lead to K4 methylation or K4 methylation may facilitate binding of the protein to the paternal allele.

One way that the inactive allele of an imprinted gene may be maintained inactive would be through blocking access to transcription factors. One interpretation of my data is that the binding of the paternal allele-specific binding protein is not a part of setting up imprinting, but is the result of the other epigenetic modifications, such as DNA methylation and chromatin structure, that are already established. For example, a transcription factor may bind to the paternal allele to activate transcription in cells where *NDN* is expressed. The maternal allele could be inaccessible to this transcription factor due to the DNA methylation and/or histone K9 methylation. This is consistent with the biparental footprints being located upstream of the CpG island and the K9 methylation.

The DNase I sensitivity and *in vivo* footprinting analysis presented in this chapter add important information about the tissue-specific and imprinting-specific regulation of *NDN/Ndn*. The pattern emerging from the analysis of different epigenetic marks at the *NDN/Ndn* locus has revealed a complex layering of epigenetic modifications.

Chapter 5. Discussion

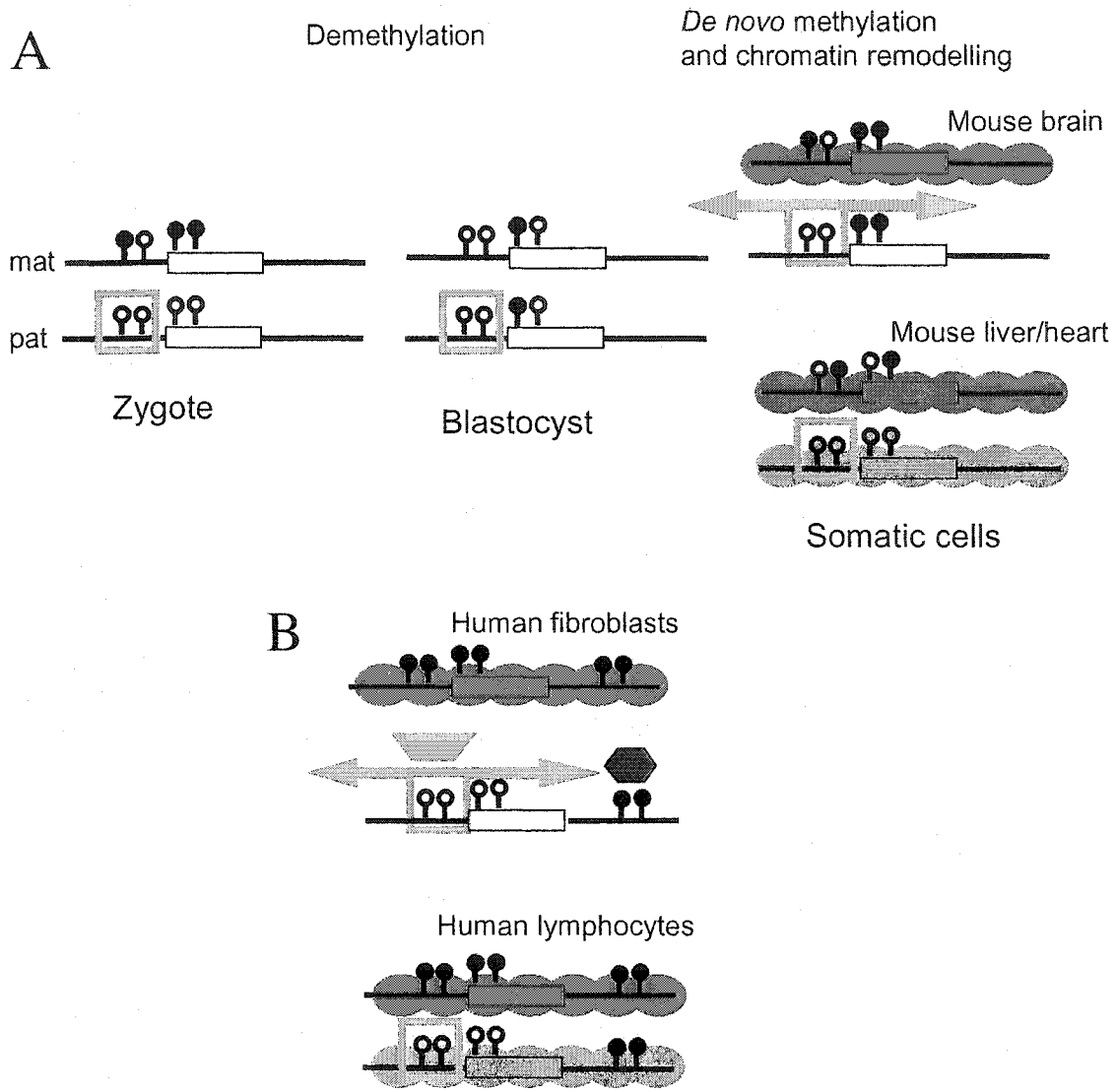
In this thesis I have characterized the epigenetic modifications of the maternal and paternal alleles of *NDN/Ndn*, a target gene of the Prader-Willi Syndrome imprinting centre (IC), in order to better understand how imprinted target genes are regulated. To date the most well studied imprinted genes co-localize with an IC. It is becoming clear that these IC associated genes only make up a small proportion of the total existing imprinted genes. In many cases, the IC appears to regulate the imprinted expression of a large domain of target genes. The experiments in this thesis represent the first detailed analysis of the epigenetic marks on an imprinted gene that is not intimately associated with an imprinting centre. I characterized human and mouse *NDN/Ndn* for three types of epigenetic modifications that I hypothesized would be important in the imprinting process: DNA methylation, chromatin structure and DNA-protein interaction sites. My results suggest that the establishment of the allele-specific expression of *NDN/Ndn* may be a multi-step process involving several layers of epigenetic modification.

A summary of the epigenetic marks on the maternal and paternal alleles of *NDN/Ndn* are shown in Figure 5-1. Initially in the zygote, the maternal allele is hypermethylated and the paternal allele has no DNA methylation. This maternal allele-specific methylation pattern is maintained until the blastocyst stage, where both alleles become demethylated. Despite this loss of allele-specific methylation, the two parental alleles are still marked differentially so that allele-specific expression follows in the somatic cells. This implies that some other epigenetic modifications distinguish the two alleles at this stage. The paternal allele was consistently marked by an absence of DNA methylation in the promoter, and at least in the human adult cells, by increased histone

K4 methylation. In adult tissues the maternal allele was consistently marked by a relatively increased DNA methylation and a closed chromatin structure. In tissues that express *NDN/Ndn*, additional modifications were present on the active paternal allele such as histone hyperacetylation (human), increased sensitivity to DNase I and preferential binding of a protein in the promoter (human). These findings demonstrate that multiple epigenetic modifications differentiate the maternal and paternal alleles of *NDN/Ndn*. The possible significance of these epigenetic modifications to the imprinting process are discussed below.

Figure 5-1. Summary of the maternal and paternal epigenotypes of *NDN/Ndn*.

(A) Mouse *Ndn* changes in DNA methylation during development. The green box encloses the 5' end of the CpG island that is protected from methylation on the paternal allele. In the adult tissues, maternal allele-specific closed chromatin is represented in pink and tissue-specific closed chromatin is represented in grey. We were not able to distinguish between these two types of closed chromatin in mouse. An open chromatin conformation is represented by arrows coming from the green box. (B) Human *NDN* epigenetic marks in adult tissues. In fibroblasts the methylated 3' CpG island seems to act as a boundary (red octagon) to the domain of open chromatin. A transcription factor (green trapezoid) binds only to the paternal allele. In lymphocytes the maternal and paternal alleles show the same closed chromatin structure when assayed for histone acetylation, but the paternal allele is distinguishable by increased histone methylation at H3 lysine 4 compared to the maternal allele.



DNA methylation may help establish the maternal epigenotype

The distinct maternal hypermethylation pattern that originated in the mouse oocytes and was maintained to the morula stage, was lost at the blastocyst stage, where both parental alleles were hypomethylated. Only a handful of imprinted genes have been analyzed for DNA methylation during early preimplantation development, but it appears that putative target site differentially methylated regions (DMRs) outside imprinting control regions undergo changes in DNA methylation, whereas IC DMRs display consistent differential methylation. For example, the *H19* promoter region, which can be considered a target of the *H19-Igf2* IC, loses the DNA hypermethylation on the paternal allele by the blastocyst stage (Tremblay *et al.*, 1997). The *Igf2r* promoter region does not become differentially methylated until late in development after implantation (Stoger *et al.*, 1993). *Igf2* becomes methylated on the expressed paternal allele at the morula stage (Brandeis *et al.*, 1993; Shemer *et al.*, 1996). In contrast, all of the IC associated DMRs have allele-specific methylation in the gametes or zygote and maintain allele-specific DNA methylation through the blastocyst stage (Shemer *et al.*, 1997; Stoger *et al.*, 1993; Tremblay *et al.*, 1997). Perhaps only the IC needs to maintain allele-specific methylation and differences in DNA methylation at target genes are able to be established or reestablished after implantation as a result of spreading of the maternal or paternal epigenetic state from the IC.

The fact that allele-specific DNA methylation is not maintained throughout development for *NDN* and for other imprinted target genes does not necessarily mean that it is unimportant. It may have key roles at specific stages of early development in the establishment of the final epigenetic state of each allele. The importance of CpG

methylation in imprinting of target genes is evidenced by the observation that CpG islands seem to correlate with genes that are subject to imprinting while most of those that escape imprinting, but reside in an imprinted region, lack CpG islands (Engemann *et al.*, 2000; Paulsen *et al.*, 2000). CpG islands at target genes could facilitate spreading of DNA hypermethylation and gene inactivation along an imprinted domain (Mummaneni *et al.*, 1995). This may be related to the ability of methylated CpG rich regions to recruit repressive complexes. The paternally methylated *H19-Igf2* IC DMR and the maternally methylated *U2af1-rs1* DMR bind the methyl binding protein MeCP2, which is known to recruit histone deacetylase activity (Drewell *et al.*, 2002; Gregory *et al.*, 2001). It is possible that the DNA methylation on the maternal allele of *NDN* in early embryogenesis binds MeCP2 or other methyl CpG binding proteins and plays a role in the establishment of a repressive chromatin structure.

Does DNA methylation maintain maternal silencing of NDN/Ndn in somatic cells?

The low level and heterogeneity of the maternal-specific hypermethylation in mouse adult tissue makes it unlikely that DNA methylation maintains the maternal allele in a repressed state in somatic cells. This is supported by the finding that the maternal allele of *NDN* does not become reactivated on treatment of parthenogenetic mouse embryonic fibroblasts with demethylating agents (C. Stewart, personal communication) (El Kharroubi *et al.*, 2001). It is possible that the more extensive maternal hypermethylation on human *NDN* is due to a more significant role for DNA methylation in maintenance of imprinted gene repression in human cells. Some experiments on *SNRPN/Snrpn* may suggest some differences in the role of DNA methylation in

maintenance of imprinted gene repression between mouse and human. In parthenogenetic mouse embryonic fibroblasts *Snrpn* cannot be reactivated upon treatment with demethylating agents (El Kharroubi *et al.*, 2001). In contrast, human PWS cell lines treated with demethylating agents are able to reactivate *SNRPN* (Fulmer-Smentek and Francke, 2001; Saitoh and Wada, 2000). Alternatively, the differences between human and mouse *NDN/Ndn* in maternal methylation levels may be due to human *NDN* having a more dense CpG island or additional DNA methylation signal elements that could possibly make it more resistant to demethylation during early development or more susceptible to higher *de novo* methylation.

The paternal allele is in an open chromatin conformation that is protected from DNA methylation

The side by side epigenetic analysis of tissues/cells that express and do not express *NDN/Ndn* has allowed the dissection of epigenetic marks that are due to the imprinting process from those related to transcriptional activity. For example, if DNA methylation was simply used to regulate gene expression, I would expect both alleles to be hypermethylated in tissues where *NDN/Ndn* is not expressed. Instead I found that the promoter region of the paternal allele was virtually devoid of methylation in all tissues examined for mouse and human *NDN/Ndn*. Given that gene inactivity and a closed chromatin conformation are often associated with DNA hypermethylation, it is interesting that the paternal allele in a non-expressing tissue was not methylated. This suggests the possibility that DNA methylation is being blocked from these regions. An allele-specific DNA binding protein, or certain histone modifications could be protective against DNA methylation. A consistent allele-specific methylation pattern in expressing

and non-expressing tissues has also been found for the *H19* promoter region (Bartolomei *et al.*, 1993).

In the mouse adult tissues, the maternal methylation was less distinct than in the oocyte and early embryo. It was also variable to the extent that many maternal clones in the adult tissues were just as hypomethylated as the paternal clones. The high levels of DNA methylation on the paternal allele in the adult brain revealed a region protected from DNA methylation in the *NDN* promoter. The area on the paternal allele that appears to be protected from DNA methylation could represent a specific open chromatin conformation. This may be a domain wide property of the paternal allele of the PWS region since the paternal allele of *Snrpn* has the ability to resist *de novo* methylation in embryonic stem cells with *Dnmt1* over expressed at levels where *Igf2* and *H19* become methylated on the normally unmethylated maternal allele (Biniszkiewicz *et al.*, 2002). Therefore the paternal allele could contain the allele distinguishing mark that carries through the blastocyst stage when the maternal allele-specific methylation is lost. If this is the case, allele-specific DNA methylation in somatic cells may be a secondary consequence of allele-specific chromatin remodelling and/or gene repression. The allele-specific methylation in mouse adult tissues may be the result of tissue-specific DNA methylation layered on top of allele-specific chromatin conformation, such that the promoter of the paternal allele is protected from *de novo* methylation.

Increased histone acetylation in human *NDN* and increased DNase I sensitivity in *NDN/Ndn* were found on the active paternal allele. These measurements of chromatin structure seem to correlate with gene activity since differences were not observed in non-expressing tissues/cells. Since the maternal allele was in a closed chromatin conformation

in expressing tissue/cells, silencing of the maternal allele must involve a closed chromatin structure that is impervious to transcriptional activation. In tissues that do not express *NDN/Ndn*, the paternal allele was in a closed chromatin conformation that was indistinguishable from the closed chromatin on the maternal allele when assayed by DNase I sensitivity and histone acetylation (human). Interestingly the analysis of histone methylation detected methylation at H3 lysine 4 on the paternal allele in both expressing and non-expressing tissues (Hanel *et al.*, 2002). Therefore this lysine 4 methylation may specifically mark the paternal allele as having the potential for gene activity. An opportunity for the paternal allele to acquire this type of chromatin modification would be just after fertilization when the male genome is repackaged into histones. The development of new techniques that would be suitable for analysis of chromatin structure in a small number of cells would be needed for the assessment of chromatin structure in early embryos.

The *in vivo* footprint analysis detected a paternal allele-specific footprint in the promoter region in fibroblasts which overlapped 2 CpG sites within the CpG island of the promoter region (CpG 16 and 17 in Figure 3-5). This suggests that CpG methylation could directly or indirectly affect binding of this factor on the maternal allele. Alternatively, binding of this protein on the paternal allele could be an allele-specific imprint that shields the paternal allele from methylation and keeps this region in an open chromatin conformation. Such an allele-specific imprint does not necessarily have to be found in all tissue types. Abnormal epigenetic remodelling could occur in lymphoblasts due to the immortalization process and no longer represent the natural state. Indeed I have found abnormal hypermethylation in lymphoblasts compared to lymphocytes. It is

also possible that the natural tissue-specific chromatin remodelling in the lymphocyte cell lineage was incompatible with the binding of this DNA binding factor.

How does NDN/Ndn respond to the IC?

The PWS-IC is responsible for post-zygotic maintenance of the paternal epigenotype for this chromosomal region by an as yet unknown mechanism. Post-zygotic mutations in the PWS-IC result in a conversion to the maternal epigenotype (Bielinska *et al.*, 2000). Therefore the PWS-IC must actively maintain an open chromatin structure on the paternal allele for all of the paternally expressed target genes in the region during the critical chromatin remodelling that occurs in the early embryo. It is not known if the PWS-IC plays a role in maintenance of proper imprinting later on in somatic cells. The *H19-Igf2* IC is needed for proper establishment of the *H19* promoter methylation and silencing on the paternal allele which occurs early in development (Srivastava 2000). However the *H19-Igf2* IC is no longer needed to maintain imprinted expression of *H19* at a later stage after implantation (Srivastava 2000). This suggests that the *H19-Igf2* IC directs the long range establishment of the allele-specific modifications but that once they are established the epigenotypes that distinguish the two alleles are maintained autonomously. Similarly the X-inactivation centre (Xic) sets up a stably inactivated X chromosome during a critical stage in development but is also no longer needed to maintain the inactive X in somatic cells (Wutz and Jaenisch, 2000). It has been shown that the chromatin structure of the preimplantation embryo undergoes a chromatin maturation process (Thompson *et al.*, 1995). It is believed that around the implantation stage, epigenetic modifications progress further as cells differentiate and lose totipotency (Packer *et al.*, 1993; Tada and Tada, 2001). For imprinted regions, it is likely that the IC

sets up allele-specific cues early in development that lead to further epigenetic programming and results in permanent allelic expression or repression.

The characterization of the maternal and paternal epigenotype at the *NDN/Ndn* locus makes it possible to generate models as to how the IC may exert its effect on the target gene of *NDN/Ndn*. I propose that the *NDN/Ndn* promoter CpG island is an important response element for the imprinting process. The differences in DNase I sensitivity in mouse brain and the allele-specific histone acetylation in human fibroblasts extended far beyond the promoter region containing the DMR. These large domains of open chromatin structure suggest that in addition to the promoter there may be domain boundaries that could potentially be targeted by the IC to control gene expression of *NDN/Ndn*. The 3' CpG island in human *NDN* seems to act as a boundary to the spread of an open chromatin conformation, but this CpG island does not appear to be a target of imprinting since no allele-specific modifications were found.

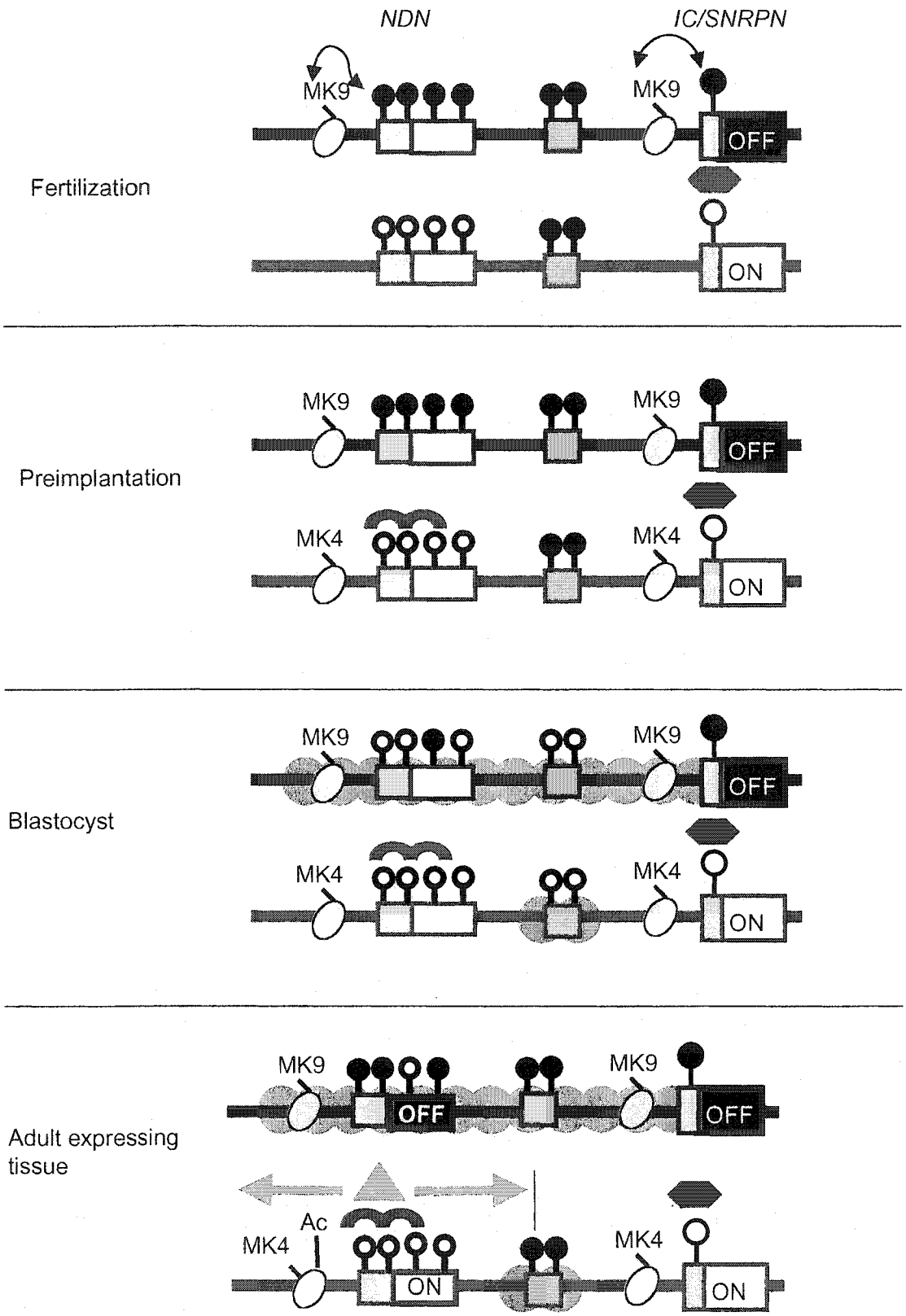
A model for the multi-step establishment of imprinting at the NDN/Ndn locus

In the model below, the maternal and paternal alleles of *NDN/Ndn* are modified in different ways (Figure 5-2). DNA methylation of the maternal allele of *SNRPN/Snrpn* and *NDN/Ndn* are established in the oocyte or shortly after fertilization. DNA methylation may be acquired secondary to H3 K9 histone methylation or DNA methylation may recruit a histone methyltransferase to methylate the H3 K9 residue. I propose that, during preimplantation development, the heterochromatin structure could spread from the IC region and methylated CpG islands on the maternal allele may facilitate the spreading. On the paternal allele, spreading of heterochromatin from the IC

may be blocked by binding of some insulator element to the unmethylated CpG island. During protamine exchange the paternal allele may be packaged into potentially active chromatin, which is marked by H3 K4 histone methylation. The CpG island at the *NDN/Ndn* promoter is protected from *de novo* methylation on the paternal allele, possibly by binding to some methylation sensitive binding protein which gives the paternal allele the potential to be active later on in tissues where *NDN/Ndn* expression is turned on. A transcriptional activator that specifically binds unmethylated CpG sites without sequence specificity could be a candidate for this role (Voo *et al.*, 2000). Recently it has been demonstrated that two closely positioned Oct-binding sequence (DOS) maintain an unmethylated state on the maternal allele of the *H19-Igf2* IC in a cell type line that is equivalent to an early post-implantation stage (Hori *et al.*, 2002). Once the closed chromatin structure is established at the *NDN/Ndn* locus on the maternal allele, it is permanent and does not depend on DNA methylation for maintenance. During differentiation various tissue-specific epigenetic modifications occur but the promoter of the paternal allele is always protected from DNA methylation. Regional H3 K4 histone methylation is maintained even in tissues where the paternal allele is inaccessible to DNase I and hypoacetylated. The protection of the promoter region from DNA methylation and the H3 K4 histone methylation make it possible for transcription factors to bind the paternal allele. Transcription factors could induce an open chromatin structure on the paternal allele which is detectable by increased DNase I sensitivity and histone hyperacetylation. The distance of spreading of the open chromatin structure on the paternal allele may be limited by boundary elements such as the 3' CpG island found on the maternal allele of human *NDN*.

Figure 5-2. A model for the multi-step establishment of imprinting at the *NDN/Ndn* locus.

The maternal allele is red and the paternal allele is blue. Filled and unfilled lollipops represent methylated and unmethylated CpG regions. Histones are represented by yellow ovals and are modified by methylation of the lysine 9 residue of histone H3 (MK9), methylation of lysine 4 of H3 (MK4) or acetylation (Ac). Promoter regions are shown as yellow boxes and the 3' CpG island (possible boundary element) is shown as a green box. The blue octagon at the IC represents a hypothetical boundary or insulator element which prevents spreading of heterochromatin on the paternal allele. The blue arc represents a methylation sensitive protein that protects the 5' promoter CpG island from DNA methylation. The gray cloud along the maternal allele is the mature heterochromatin structure on the maternal allele. On the paternal allele in adult expressing tissue a green triangle demonstrates the binding of a transcription factor which establishes or maintains a region of open chromatin structure.



Future directions: NDN/Ndn and beyond

I have characterized the maternal and paternal epigenotypes at the *NDN/Ndn* locus. This has revealed layers of allele-specific epigenetic modifications. Initially, DNA methylation was thought to be the most likely candidate for the establishment and maintenance of allele-specific gene expression. It has become clear that other allele-specific modifications relating to chromatin structure are also found at imprinted loci. In many cases, chromatin structure and DNA methylation seem to be linked processes, but the order of the establishment of these modifications is not clear. Allele-specific DNA methylation is still the only allele-specific mark known to be found in the gametes or early embryo of imprinted genes. This does not mean that it is the only epigenetic modification at these stages. It would be of interest to find out if other epigenetic marks such as chromatin structure, or DNA binding proteins exist in the gametes or early stages of development. However this would require the development of new techniques for the analysis of very small samples which have not yet been well established.

A method of analyzing chromatin structure using DNase I treatment and SSCP (single strand conformation polymorphism) to distinguish parental alleles has been shown to be feasible for the analysis of a small number of cells (Gregory and Feil, 1999). The development of this technique would be useful to compare the chromatin structure of the maternal and paternal alleles in the mouse preimplantation embryo. This could determine whether DNA methylation on the maternal allele preceeds chromatin structural differences, and also whether chromatin structural differences are present in the blastocyst when both parental alleles are hypomethylated.

It is likely that some elements of the *NDN/Ndn* target locus are important for responding to the cues from the IC to create allele-specific expression. The techniques described in this thesis could be applied to other target imprinted genes from the PWS region or from other imprinting clusters. A comparison between imprinted target gene promoters for allele-specific epigenetic modifications could reveal common elements that make these genes able to respond to an imprinting centre. For example, one could identify common DNA sequences that are protected from DNA methylation on one allele. *In vivo* footprinting assays may identify regions that are commonly bound by proteins in an allele-specific manner. If the DNA binding protein responsible for producing footprints could be identified, the sensitivity of the DNA binding sites to DNA methylation could be tested by electromobility shift assays (EMSA) using unmethylated and *in vitro* methylated DNA fragments.

The regulation of imprinting may also depend on elements distant to the promoter. Therefore, it would be of interest to determine the boundaries to allele-specific chromatin structure at *NDN/Ndn* and whether there is a boundary between *NDN/Ndn* and the neighbouring *MAGEL2/Magel2*. Boundaries to allele-specific chromatin structure could harbour insulators or boundary elements which could be important in maintaining domains of allele-specific chromatin structure. It is not known whether imprinted genes within a cluster are controlled individually or if multiple imprinted genes are controlled as a single imprinted domain.

Targeted deletions or mutations of elements suspected to be important for responding to the IC regulation, could be tested in mice to determine if imprinted regulation is lost. Alternatively, transgenes could be constructed with a target gene linked

to an IC to determine minimal elements necessary for the target gene to respond to the IC.

These types of studies would help us to better understand the biological principles underlying imprinted gene regulation. In addition, knowledge gained about how epigenetic modifications relate to gene expression in imprinted genes is important to the study of gene regulation in general.

References

Aasland, R., and Stewart, A. F. (1999). Analysis of DNaseI hypersensitive sites in chromatin by cleavage in permeabilized cells. Methods Mol Biol *119*, 355-62.

Amos-Landgraf, J. M., Ji, Y., Gottlieb, W., Depinet, T., Wandstrat, A. E., Cassidy, S. B., Driscoll, D. J., Rogan, P. K., Schwartz, S., and Nicholls, R. D. (1999). Chromosome breakage in the Prader-Willi and Angelman syndromes involves recombination between large, transcribed repeats at proximal and distal breakpoints. Am J Hum Genet *65*, 370-86.

Antequera, F., Boyes, J., and Bird, A. (1990). High levels of *de novo* methylation and altered chromatin structure at CpG islands in cell lines. Cell *62*, 503-14.

Ausubel, F. M., Brent, R. E., Kingston, J. G., Seidman, J. G., Smith, J. A., and Struhl, K. (1993). Current protocols in molecular biology, Volume 1 (New York, N.Y.: Greene Publishing Associates and Wiley-Interscience).

Bachman, K. E., Rountree, M. R., and Baylin, S. B. (2001). Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. J Biol Chem *276*, 32282-7.

Barlow, D. P. (1997). Competition--a common motif for the imprinting mechanism? Embo J *16*, 6899-905.

Bartolomei, M. S., Webber, A. L., Brunkow, M. E., and Tilghman, S. M. (1993). Epigenetic mechanisms underlying the imprinting of the mouse *H19* gene. Genes Dev *7*, 1663-73.

Beard, C., Li, E., and Jaenisch, R. (1995). Loss of methylation activates Xist in somatic but not in embryonic cells. Genes Dev *9*, 2325-34.

Bednarik, D. P., Duckett, C., Kim, S. U., Perez, V. L., Griffis, K., Guenthner, P. C., and Folks, T. M. (1991). DNA CpG methylation inhibits binding of NF-kappa B proteins to the HIV- 1 long terminal repeat cognate DNA motifs. New Biol *3*, 969-76.

Beechey, C. V., and Cattanach, B. M. (1995). Genetic imprinting map. Mouse Genome *93*, 89-91.

Bell, A. C., and Felsenfeld, G. (2000). Methylation of a CTCF-dependent boundary controls imprinted expression of the *Igf2* gene. Nature *405*, 482-5.

Bell, A. C., West, A. G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-96.

Belyaev, N., Keohane, A. M., and Turner, B. M. (1996). Differential underacetylation of histones H2A, H3 and H4 on the inactive X chromosome in human female cells. Hum Genet *97*, 573-8.

Bestor, T. H., and Ingram, V. M. (1983). Two DNA methyltransferases from murine erythroleukemia cells: purification, sequence specificity, and mode of interaction with DNA. Proc Natl Acad Sci U S A *80*, 5559-63.

Bielinska, B., Blaydes, S. M., Buiting, K., Yang, T., Krajewska-Walasek, M., Horsthemke, B., and Brannan, C. I. (2000). *De novo* deletions of SNRPN exon 1 in early human and mouse embryos result in a paternal to maternal imprint switch. Nat Genet *25*, 74-8.

Biniszkiewicz, D., Gribnau, J., Ramsahoye, B., Gaudet, F., Eggan, K., Humpherys, D., Mastrangelo, M. A., Jun, Z., Walter, J., and Jaenisch, R. (2002). Dnmt1 overexpression causes genomic hypermethylation, loss of imprinting, and embryonic lethality. Mol Cell Biol *22*, 2124-35.

Bird, A. (1992). The essentials of DNA methylation. Cell *70*, 5-8.

Bird, A. P., and Wolffe, A. P. (1999). Methylation-induced repression--belts, braces, and chromatin. Cell *99*, 451-4.

Birger, Y., Shemer, R., Perk, J., and Razin, A. (1999). The imprinting box of the mouse *Igf2*r gene. Nature *397*, 84-8.

Boeuf, S., Klingenspor, M., Van Hal, N. L., Schneider, T., Keijer, J., and Klaus, S. (2001). Differential gene expression in white and brown preadipocytes. Physiol Genomics *7*, 15-25.

Boggs, B. A., Cheung, P., Heard, E., Spector, D. L., Chinault, A. C., and Allis, C. D. (2002). Differentially methylated forms of histone H3 show unique association patterns with inactive human X chromosomes. Nat Genet *30*, 73-6.

Boyes, J., and Bird, A. (1992). Repression of genes by DNA methylation depends on CpG density and promoter strength: evidence for involvement of a methyl-CpG binding protein. Embo J *11*, 327-33.

Boyes, J., and Felsenfeld, G. (1996). Tissue-specific factors additively increase the probability of the all- or-none formation of a hypersensitive site. Embo J *15*, 2496-507. abs.html.

Brandeis, M., Kafri, T., Ariel, M., Chaillet, J. R., McCarrey, J., Razin, A., and Cedar, H. (1993). The ontogeny of allele-specific methylation associated with imprinted genes in the mouse. Embo J *12*, 3669-77.

Brenton, J. D., Drewell, R. A., Viville, S., Hilton, K. J., Barton, S. C., Ainscough, J. F., and Surani, M. A. (1999). A silencer element identified in Drosophila is required for imprinting of *H19* reporter transgenes in mice. Proc Natl Acad Sci U S A *96*, 9242-7.

Bresnick, E. H., Rories, C., and Hager, G. L. (1992). Evidence that nucleosomes on the mouse mammary tumor virus promoter adopt specific translational positions. Nucleic Acids Res *20*, 865-70.

Brown, C. J., Hendrich, B. D., Rupert, J. L., Lafreniere, R. G., Xing, Y., Lawrence, J., and Willard, H. F. (1992). The human XIST gene: analysis of a 17 kb inactive X-specific RNA that contains conserved repeats and is highly localized within the nucleus. Cell *71*, 527-42.

Brown, C. J., Lafreniere, R. G., Powers, V. E., Sebastio, G., Ballabio, A., Pettigrew, A. L., Ledbetter, D. H., Levy, E., Craig, I. W., and Willard, H. F. (1991). Localization of the X inactivation centre on the human X chromosome in Xq13. Nature *349*, 82-4.

Brown, C. J., and Willard, H. F. (1994). The human X-inactivation centre is not required for maintenance of X- chromosome inactivation. Nature *368*, 154-6.

Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M., and Fisher, A. G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845-54.

Buiting, K., Barnicoat, A., Lich, C., Pembrey, M., Malcolm, S., and Horsthemke, B. (2001). Disruption of the bipartite imprinting center in a family with Angelman syndrome. Am J Hum Genet *68*, 1290-4.

Buiting, K., Dittrich, B., Gross, S., Lich, C., Farber, C., Buchholz, T., Smith, E., Reis, A., Burger, J., Nothen, M. M., Barth-Witte, U., Janssen, B., Abeliovich, D., Lerer, I., van den Ouweland, A. M., Halley, D. J., Schrander-Stumpel, C., Smeets, H., Meinecke, P., Malcolm, S., Gardner, A., Lalande, M., Nicholls, R. D., Friend, K., Horsthemke, B., and *et al.* (1998). Sporadic imprinting defects in Prader-Willi syndrome and Angelman syndrome: implications for imprint-switch models, genetic counseling, and prenatal diagnosis. Am J Hum Genet *63*, 170-80.

Bulger, M., Sawado, T., Schubeler, D., and Groudine, M. (2002). ChIPs of the beta-globin locus: unraveling gene regulation within an active domain. Curr Opin Genet Dev *12*, 170-7.

Burman, P., Ritzen, E. M., and Lindgren, A. C. (2001). Endocrine dysfunction in Prader-Willi syndrome: a review with special reference to GH. Endocr Rev *22*, 787-99.

Busslinger, M., Hurst, J., and Flavell, R. A. (1983). DNA methylation and the regulation of globin gene expression. Cell *34*, 197-206.

Cameron, E. E., Bachman, K. E., Myohanen, S., Herman, J. G., and Baylin, S. B. (1999). Synergy of demethylation and histone deacetylase inhibition in the re- expression of genes silenced in cancer. Nat Genet *21*, 103-7.

Cantor, A. B., and Orkin, S. H. (2002). Transcriptional regulation of erythropoiesis: an affair involving multiple partners. Oncogene *21*, 3368-76.

Carrel, L., Clemson, C. M., Dunn, J. M., Miller, A. P., Hunt, P. A., Lawrence, J. B., and Willard, H. F. (1996). X inactivation analysis and DNA methylation studies of the ubiquitin activating enzyme E1 and PCTAIRE-1 genes in human and mouse. Hum Mol Genet *5*, 391-401.

Caspary, T., Cleary, M. A., Baker, C. C., Guan, X. J., and Tilghman, S. M. (1998). Multiple mechanisms regulate imprinting of the mouse distal chromosome 7 gene cluster. Mol Cell Biol *18*, 3466-74.

Cavaille, J., Buiting, K., Kiefmann, M., Lalande, M., Brannan, C. I., Horsthemke, B., Bachellerie, J. P., Brosius, J., and Huttenhofer, A. (2000). Identification of brain-specific and imprinted small nucleolar RNA genes exhibiting an unusual genomic organization. Proc Natl Acad Sci U S A *97*, 14311-6.

Cavaille, J., Seitz, H., Paulsen, M., Ferguson-Smith, A. C., and Bachellerie, J. P. (2002). Identification of tandemly-repeated C/D snoRNA genes at the imprinted human 14q32 domain reminiscent of those at the Prader-Willi/Angelman syndrome region. Hum Mol Genet *11*, 1527-38.

Chamberlain, S. J., and Brannan, C. I. (2001). The Prader-Willi syndrome imprinting center activates the paternally expressed murine Ube3a antisense transcript but represses paternal Ube3a. Genomics *73*, 316-22.

Chao, W., Huynh, K. D., Spencer, R. J., Davidow, L. S., and Lee, J. T. (2002). CTCF, a candidate trans-acting factor for X-inactivation choice. Science *295*, 345-7.

Chen, H., Lin, R. J., Xie, W., Wilpitz, D., and Evans, R. M. (1999). Regulation of hormone-induced histone hyperacetylation and gene activation via acetylation of an acetylase. Cell *98*, 675-86.

Chen, R. Z., Pettersson, U., Beard, C., Jackson-Grusby, L., and Jaenisch, R. (1998). DNA hypomethylation leads to elevated mutation rates. Nature *395*, 89-93.

Christian, S. L., Fantes, J. A., Mewborn, S. K., Huang, B., and Ledbetter, D. H. (1999). Large genomic duplicons map to sites of instability in the Prader- Willi/Angelman syndrome chromosome region (15q11-q13). Hum Mol Genet *8*, 1025-37.

Clarke, H. J., McLay, D. W., and Mohamed, O. A. (1998). Linker histone transitions during mammalian oogenesis and embryogenesis. Dev Genet *22*, 17-30.

Clayton-Smith, J., and Pembrey, M. E. (1992). Angelman syndrome. J Med Genet *29*, 412-5.

Cleary, M. A., van Raamsdonk, C. D., Levorse, J., Zheng, B., Bradley, A., and Tilghman, S. M. (2001). Disruption of an imprinted gene cluster by a targeted chromosomal translocation in mice. Nat Genet *29*, 78-82.

Cockett, N. E., Jackson, S. P., Shay, T. L., Farnir, F., Berghmans, S., Snowder, G. D., Nielsen, D. M., and Georges, M. (1996). Polar overdominance at the ovine callipyge locus. Science *273*, 236-8.

Colot, V., and Rossignol, J. L. (1999). Eukaryotic DNA methylation as an evolutionary device. Bioessays *21*, 402-11.

Comb, M., and Goodman, H. M. (1990). CpG methylation inhibits proenkephalin gene expression and binding of the transcription factor AP-2. Nucleic Acids Res *18*, 3975-82.

Costanzi, C., and Pehrson, J. R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599-601.

Cross, S. H., and Bird, A. P. (1995). CpG islands and genes. Curr Opin Genet Dev *5*, 309-14.

Dasen, J. S., O'Connell, S. M., Flynn, S. E., Treier, M., Gleiberman, A. S., Szeto, D. P., Hooshmand, F., Aggarwal, A. K., and Rosenfeld, M. G. (1999). Reciprocal interactions of Pit1 and GATA2 mediate signaling gradient- induced determination of pituitary cell types. Cell *97*, 587-98.

Davies, S. R., Sakano, S., Zhu, Y., and Sandell, L. J. (2002). Distribution of the Transcription Factors Sox9, AP-2, and [Delta]EF1 in Adult Murine Articular and Meniscal Cartilage and Growth Plate. J Histochem Cytochem *50*, 1059-65.

Davis, T. L., Yang, G. J., McCarrey, J. R., and Bartolomei, M. S. (2000). The *H19* methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. Hum Mol Genet *9*, 2885-94.

de los Santos, T., Schweizer, J., Rees, C. A., and Francke, U. (2000). Small evolutionarily conserved RNA, resembling C/D box small nucleolar RNA, is transcribed from PWCR1, a novel imprinted gene in the Prader- Willi deletion region, which Is highly expressed in brain. Am J Hum Genet *67*, 1067-82.

De Smet, C., Lurquin, C., Lethe, B., Martelange, V., and Boon, T. (1999). DNA methylation is the primary silencing mechanism for a set of germ line- and tumor-specific genes with a CpG-rich promoter. Mol Cell Biol *19*, 7327-35.

DeZazzo, J., Sandstrom, D., de Belle, S., Velinzon, K., Smith, P., Grady, L., DelVecchio, M., Ramaswami, M., and Tully, T. (2000). nalyot, a mutation of the Drosophila myb-related Adf1 transcription factor, disrupts synapse formation and olfactory memory. Neuron *27*, 145-58.

Dittrich, B., Buiting, K., Korn, B., Rickard, S., Buxton, J., Saitoh, S., Nicholls, R. D., Poustka, A., Winterpacht, A., Zabel, B., and Horsthemke, B. (1996). Imprint switching on human chromosome 15 may involve alternative transcripts of the SNRPN gene. Nat Genet *14*, 163-70.

Dragich, J., Houwink-Manville, I., and Schanen, C. (2000). Rett syndrome: a surprising result of mutation in MECP2. Hum Mol Genet *9*, 2365-75.

Drewell, R. A., Goddard, C. J., Thomas, J. O., and Surani, M. A. (2002). Methylation-dependent silencing at the *H19* imprinting control region by MeCP2. Nucleic Acids Res *30*, 1139-44.

Drouin, R., Therrien, J. P., Angers, M., and Ouellet, S. (2001). In vivo DNA analysis. Methods Mol Biol *148*, 175-219.

Dutton, J. R., Johns, S., and Miller, B. L. (1997). StuAp is a sequence-specific transcription factor that regulates developmental complexity in Aspergillus nidulans. Embo J *16*, 5710-21.

Eden, S., and Cedar, H. (1994). Role of DNA methylation in the regulation of transcription. Curr Opin Genet Dev *4*, 255-9.

Eden, S., Constancia, M., Hashimshony, T., Dean, W., Goldstein, B., Johnson, A. C., Keshet, I., Reik, W., and Cedar, H. (2001). An upstream repressor element plays a role in *Igf2* imprinting. Embo J *20*, 3518-25.

El Kharroubi, A., Piras, G., and Stewart, C. L. (2001). DNA demethylation reactivates a subset of imprinted genes in uniparental mouse embryonic fibroblasts. J Biol Chem *276*, 8674-80.

El-Maarri, O., Buiting, K., Peery, E. G., Kroisel, P. M., Balaban, B., Wagner, K., Urman, B., Heyd, J., Lich, C., Brannan, C. I., Walter, J., and Horsthemke, B. (2001). Maternal methylation imprints on human chromosome 15 are established during or after fertilization. Nat Genet *27*, 341-4.

Engel, E. (1997). Uniparental disomy (UPD). Genomic imprinting and a case for new genetics (prenatal and clinical implications: the "Likon" concept). Ann Genet *40*, 24-34.

Engemann, S., Strodicke, M., Paulsen, M., Franck, O., Reinhardt, R., Lane, N., Reik, W., and Walter, J. (2000). Sequence and functional comparison in the Beckwith-Wiedemann region: implications for a novel imprinting centre and extended imprinting. Hum Mol Genet *9*, 2691-706.

Farrell, C. M., West, A. G., and Felsenfeld, G. (2002). Conserved CTCF insulator elements flank the mouse and human beta-globin loci. Mol Cell Biol *22*, 3820-31.

Feil, R., Boyano, M. D., Allen, N. D., and Kelsey, G. (1997). Parental chromosome-specific chromatin conformation in the imprinted U2af1-rs1 gene in the mouse. J Biol Chem *272*, 20893-900.

Feil, R., Handel, M. A., Allen, N. D., and Reik, W. (1995). Chromatin structure and imprinting: developmental control of DNase-I sensitivity in the mouse insulin-like growth factor 2 gene. Dev Genet *17*, 240-52.

Feil, R., Walter, J., Allen, N. D., and Reik, W. (1994). Developmental control of allelic methylation in the imprinted mouse *Igf2* and *H19* genes. Development *120*, 2933-43.

Felsenfeld, G., Boyes, J., Chung, J., Clark, D., and Studitsky, V. (1996). Chromatin structure and gene expression. Proc Natl Acad Sci U S A *93*, 9384-8.

Ferguson-Smith, A. C., Cattanach, B. M., Barton, S. C., Beechey, C. V., and Surani, M. A. (1991). Embryological and molecular investigations of parental imprinting on mouse chromosome 7. Nature *351*, 667-70.

Forsberg, E. C., Downs, K. M., Christensen, H. M., Im, H., Nuzzi, P. A., and Bresnick, E. H. (2000). Developmentally dynamic histone acetylation pattern of a tissue-specific chromatin domain. Proc Natl Acad Sci U S A *97*, 14494-9.

Frommer, M., McDonald, L. E., Millar, D. S., Collis, C. M., Watt, F., Grigg, G. W., Molloy, P. L., and Paul, C. L. (1992). A genomic sequencing protocol that yields

a positive display of 5- methylcytosine residues in individual DNA strands. Proc Natl Acad Sci U S A *89*, 1827-31.

Fulmer-Smentek, S. B., and Francke, U. (2001). Association of acetylated histones with paternally expressed genes in the Prader--Willi deletion region. Hum Mol Genet *10*, 645-52.

Funahashi, J., Sekido, R., Murai, K., Kamachi, Y., and Kondoh, H. (1993). Delta-crystallin enhancer binding protein delta EF1 is a zinc finger- homeodomain protein implicated in postgastrulation embryogenesis. Development *119*, 433-46.

Futscher, B. W., Oshiro, M. M., Wozniak, R. J., Holtan, N., Hanigan, C. L., Duan, H., and Domann, F. E. (2002). Role for DNA methylation in the control of cell type specific maspin expression. Nat Genet *31*, 175-9.

Gardiner-Garden, M., and Frommer, M. (1987). CpG islands in vertebrate genomes. J Mol Biol *196*, 261-82.

Garrick, D., Fiering, S., Martin, D. I., and Whitelaw, E. (1998). Repeat-induced gene silencing in mammals. Nat Genet *18*, 56-9.

Gerard, M., Hernandez, L., Wevrick, R., and Stewart, C. L. (1999). Disruption of the mouse necdin gene results in early post-natal lethality. Nat Genet *23*, 199-202.

Grandjean, V., O'Neill, L., Sado, T., Turner, B., and Ferguson-Smith, A. (2001). Relationship between DNA methylation, histone H4 acetylation and gene expression in the mouse imprinted *Igf2-H19* domain. FEBS Lett *488*, 165-9.

Gregory, R. I., and Feil, R. (1999). Analysis of chromatin in limited numbers of cells: a PCR-SSCP based assay of allele-specific nuclease sensitivity. Nucleic Acids Res *27*, e32.

Gregory, R. I., Randall, T. E., Johnson, C. A., Khosla, S., Hatada, I., O'Neill, L. P., Turner, B. M., and Feil, R. (2001). DNA methylation is linked to deacetylation of histone H3, but not H4, on the imprinted genes Snrpn and U2af1-rs1. Mol Cell Biol *21*, 5426-36.

Gruenbaum, Y., Cedar, H., and Razin, A. (1982). Substrate and sequence specificity of a eukaryotic DNA methylase. Nature *295*, 620-2.

Gunay-Aygun, M., Schwartz, S., Heeger, S., O'Riordan, M. A., and Cassidy, S. B. (2001). The changing purpose of Prader-Willi syndrome clinical diagnostic criteria and proposed revised criteria. Pediatrics *108*, E92.

Hanel, M. L., and Wevrick, R. (2001). Establishment and maintenance of DNA methylation patterns in mouse *Ndn*: implications for maintenance of imprinting in target genes of the imprinting center. Mol Cell Biol *21*, 2384-92.

Hanel, M. L., Lau, J. C. Y., and Wevick, R. (2002). Epigenetic marks of tissue-specific and imprinted regulation of the human *NDN* gene. (Joint first authorship). Submitted.

Hansen, R. S., Stoger, R., Wijmenga, C., Stanek, A. M., Canfield, T. K., Luo, P., Matarazzo, M. R., D'Esposito, M., Feil, R., Gimelli, G., Weemaes, C. M., Laird, C. D., and Gartler, S. M. (2000). Escape from gene silencing in ICF syndrome: evidence for advanced replication time as a major determinant. Hum Mol Genet *9*, 2575-87.

Hark, A. T., Schoenherr, C. J., Katz, D. J., Ingram, R. S., Levorse, J. M., and Tilghman, S. M. (2000). CTCF mediates methylation-sensitive enhancer-blocking activity at the *H19/Igf2* locus. Nature *405*, 486-9.

Hark, A. T., and Tilghman, S. M. (1998). Chromatin conformation of the *H19* epigenetic mark. Hum Mol Genet *7*, 1979-85.

Hata, K., Okano, M., Lei, H., and Li, E. (2002). Dnmt3L cooperates with the Dnmt3 family of *de novo* DNA methyltransferases to establish maternal imprints in mice. Development *129*, 1983-93.

Hebbes, T. R., Clayton, A. L., Thorne, A. W., and Crane-Robinson, C. (1994). Core histone hyperacetylation co-maps with generalized DNase I sensitivity in the chicken beta-globin chromosomal domain. Embo J *13*, 1823-30.

Hebbes, T. R., Thorne, A. W., Clayton, A. L., and Crane-Robinson, C. (1992). Histone acetylation and globin gene switching. Nucleic Acids Res *20*, 1017-22.

Hershko, A., Razin, A., and Shemer, R. (1999). Imprinted methylation and its effect on expression of the mouse Zfp127 gene. Gene *234*, 323-7. .

Hogan, B., Beddinton, R., Costanini, F., and Lacy, E. (1994). Manipulating the mouse embryo (Plainview, N.Y.: Cold Springs Harbour Laboratory Press).

Holm, V. A., Cassidy, S. B., Butler, M. G., Hanchett, J. M., Greenswag, L. R., Whitman, B. Y., and Greenberg, F. (1993). Prader-Willi syndrome: consensus diagnostic criteria. Pediatrics *91*, 398-402.

Hori, N., Nakano, H., Takeuchi, T., Kato, H., Hamaguchi, S., Oshimura, M., and Sato, K. (2002). A Dyad Oct-binding Sequence Functions as a Maintenance Sequence for the Unmethylated State within the *H19/Igf2*-imprinted Control Region. J Biol Chem *277*, 27960-27967.

Horike, S., Mitsuya, K., Meguro, M., Kotobuki, N., Kashiwagi, A., Notsu, T., Schulz, T. C., Shirayoshi, Y., and Oshimura, M. (2000). Targeted disruption of the human LIT1 locus defines a putative imprinting control element playing an essential role in Beckwith- Wiedemann syndrome. Hum Mol Genet *9*, 2075-83.

Horsley, V., and Pavlath, G. K. (2002). NFAT: ubiquitous regulator of cell differentiation and adaptation. J Cell Biol *156*, 771-4.

Horsthemke, B., Dittrich, B., and Buiting, K. (1997). Imprinting mutations on human chromosome 15. Hum Mutat *10*, 329-37.

Howlett, S. K., and Reik, W. (1991). Methylation levels of maternal and paternal genomes during preimplantation development. Development *113*, 119-27.

Hsieh, C. L. (1999). In vivo activity of murine *de novo* methyltransferases, Dnmt3a and Dnmt3b. Mol Cell Biol *19*, 8211-8.

Huq, A. H., Sutcliffe, J. S., Nakao, M., Shen, Y., Gibbs, R. A., and Beaudet, A. L. (1997). Sequencing and functional analysis of the SNRPN promoter: *in vitro* methylation abolishes promoter activity. Genome Res *7*, 642-8.

Ishihara, K., and Sasaki, H. (2002). An evolutionarily conserved putative insulator element near the 3' boundary of the imprinted *Igf2/H19* domain. Hum Mol Genet *11*, 1627-36.

Jaenisch, R. (1997). DNA methylation and imprinting: why bother? Trends Genet *13*, 323-9.

Jahner, D., Stuhlmann, H., Stewart, C. L., Harbers, K., Lohler, J., Simon, I., and Jaenisch, R. (1982). *De novo* methylation and expression of retroviral genomes during mouse embryogenesis. Nature *298*, 623-8.

Jay, P., Rougeulle, C., Massacrier, A., Moncla, A., Mattei, M. G., Malzac, P., Roeckel, N., Taviaux, S., Lefranc, J. L., Cau, P., Berta, P., Lalande, M., and Muscatelli, F. (1997). The human necdin gene, NDN, is maternally imprinted and located in the Prader-Willi syndrome chromosomal region. Nat Genet *17*, 357-61.

Jenuwein, T., and Allis, C. D. (2001). Translating the histone code. Science *293*, 1074-80.

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*, 281-9.

Jones, P. A. (1999). The DNA methylation paradox. Trends Genet *15*, 34-7.

Jones, P. A., Wolkowicz, M. J., Rideout, W. M., 3rd, Gonzales, F. A., Marziasz, C. M., Coetzee, G. A., and Tapscott, S. J. (1990). *De novo* methylation of the MyoD1 CpG island during the establishment of immortal cell lines. Proc Natl Acad Sci U S A *87*, 6117-21.

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat Genet *19*, 187-91.

Jong, M. T., Gray, T. A., Ji, Y., Glenn, C. C., Saitoh, S., Driscoll, D. J., and Nicholls, R. D. (1999). A novel imprinted gene, encoding a RING zinc-finger protein, and overlapping antisense transcript in the Prader-Willi syndrome critical region. Hum Mol Genet *8*, 783-93.

Kafri, T., Ariel, M., Brandeis, M., Shemer, R., Urven, L., McCarrey, J., Cedar, H., and Razin, A. (1992). Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. Genes Dev *6*, 705-14.

Kanduri, C., Fitzpatrick, G., Mukhopadhyay, R., Kanduri, M., Lobanenkov, V., Higgins, M., and Ohlsson, R. (2002). A differentially methylated imprinting control region within the Kcnq1 locus harbors a methylation-sensitive chromatin insulator. J Biol Chem *277*, 18106-10.

Kanduri, C., Pant, V., Loukinov, D., Pugacheva, E., Qi, C. F., Wolffe, A., Ohlsson, R., and Lobanenkov, V. V. (2000). Functional association of CTCF with the insulator upstream of the *H19* gene is parent of origin-specific and methylation-sensitive. Curr Biol *10*, 853-6.

Kasahara, H., Usheva, A., Ueyama, T., Aoki, H., Horikoshi, N., and Izumo, S. (2001). Characterization of homo- and heterodimerization of cardiac Csx/Nkx2.5 homeoprotein. J Biol Chem *276*, 4570-80.

Keohane, A. M., Lavender, J. S., O'Neill, L. P., and Turner, B. M. (1998). Histone acetylation and X inactivation. Dev Genet *22*, 65-73.

Keshet, I., Lieman-Hurwitz, J., and Cedar, H. (1986). DNA methylation affects the formation of active chromatin. Cell *44*, 535-43.

Keshet, I., Yisraeli, J., and Cedar, H. (1985). Effect of regional DNA methylation on gene expression. Proc Natl Acad Sci U S A *82*, 2560-4.

Khosla, S., Aitchison, A., Gregory, R., Allen, N. D., and Feil, R. (1999). Parental allele-specific chromatin configuration in a boundary- imprinting-control element upstream of the mouse *H19* gene. Mol Cell Biol *19*, 2556-66.

Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R., and Georgopoulos, K. (1999). Ikaros DNA-binding proteins direct formation of chromatin *remodelling* complexes in lymphocytes. Immunity *10*, 345-55.

Kishino, T., Lalande, M., and Wagstaff, J. (1997). UBE3A/E6-AP mutations cause Angelman syndrome. Nat Genet *15*, 70-3.

Knoll, J. H., Nicholls, R. D., Magenis, R. E., Graham, J. M., Jr., Lalande, M., and Latt, S. A. (1989). Angelman and Prader-Willi syndromes share a common chromosome 15 deletion but differ in parental origin of the deletion. Am J Med Genet *32*, 285-90.

Koide, T., Ainscough, J., Wijgerde, M., and Surani, M. A. (1994). Comparative analysis of Igf-2/*H19* imprinted domain: identification of a highly conserved intergenic DNase I hypersensitive region. Genomics *24*, 1-8.

Kovesdi, I., Reichel, R., and Nevins, J. R. (1987). Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. Proc Natl Acad Sci U S A *84*, 2180-4.

Labosky, P. A., Barlow, D. P., and Hogan, B. L. (1994). Mouse embryonic germ (EG) cell lines: transmission through the germline and differences in the methylation imprint of insulin-like growth factor 2 receptor (*Igf*2r) gene compared with embryonic stem (ES) cell lines. Development *120*, 3197-204.

Lau, S., Jardine, K., and McBurney, M. W. (1999). DNA methylation pattern of a tandemly repeated LacZ transgene indicates that most copies are silent. Dev Dyn *215*, 126-38.

Lawson, G. M., Knoll, B. J., March, C. J., Woo, S. L., Tsai, M. J., and O'Malley, B. W. (1982). Definition of 5' and 3' structural boundaries of the chromatin domain containing the ovalbumin multigene family. J Biol Chem *257*, 1501-7.

Lawson, M. A., Whyte, D. B., and Mellon, P. L. (1996). GATA factors are essential for activity of the neuron-specific enhancer of the gonadotropin-releasing hormone gene. Mol Cell Biol *16*, 3596-605.

Lee, M. P., DeBaun, M. R., Mitsuya, K., Galonek, H. L., Brandenburg, S., Oshimura, M., and Feinberg, A. P. (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. Proc Natl Acad Sci U S A *96*, 5203-8.

Lee, S., and Wevrick, R. (2000). Identification of novel imprinted transcripts in the Prader-Willi syndrome and Angelman syndrome deletion region: further evidence for regional imprinting control. Am J Hum Genet *66*, 848-58.

Lei, H., Oh, S. P., Okano, M., Juttermann, R., Goss, K. A., Jaenisch, R., and Li, E. (1996). *De novo* DNA cytosine methyltransferase activities in mouse embryonic stem cells. Development *122*, 3195-205.

Leighton, P. A., Ingram, R. S., Eggenschwiler, J., Efstratiadis, A., and Tilghman, S. M. (1995). Disruption of imprinting caused by deletion of the *H19* gene region in mice. Nature *375*, 34-9.

Leonhardt, H., Page, A. W., Weier, H. U., and Bestor, T. H. (1992). A targeting sequence directs DNA methyltransferase to sites of DNA replication in mammalian nuclei. Cell *71*, 865-73.

Li, E., Beard, C., and Jaenisch, R. (1993). Role for DNA methylation in genomic imprinting. Nature *366*, 362-5.

Li, E., Bestor, T. H., and Jaenisch, R. (1992). Targeted mutation of the DNA methyltransferase gene results in embryonic lethality. Cell *69*, 915-26.

Liang, G., Chan, M. F., Tomigahara, Y., Tsai, Y. C., Gonzales, F. A., Li, E., Laird, P. W., and Jones, P. A. (2002). Cooperativity between DNA methyltransferases in the maintenance methylation of repetitive elements. Mol Cell Biol *22*, 480-91.

Loukinov, D. I., Pugacheva, E., Vatolin, S., Pack, S. D., Moon, H., Chernukhin, I., Mannan, P., Larsson, E., Kanduri, C., Vostrov, A. A., Cui, H., Niemitz, E. L., Rasko, J. E., Docquier, F. M., Kistler, M., Breen, J. J., Zhuang, Z., Quitschke, W. W., Renkawitz, R., Klenova, E. M., Feinberg, A. P., Ohlsson, R., Morse, H. C., 3rd, and Lobanenkov, V. V. (2002). BORIS, a novel male germ-line-specific protein associated with epigenetic reprogramming events, shares the same 11-zinc-finger domain with CTCF, the insulator protein involved in reading imprinting marks in the soma. Proc Natl Acad Sci U S A *99*, 6806-11.

Lutz, M., Burke, L. J., Barreto, G., Goeman, F., Greb, H., Arnold, R., Schultheiss, H., Brehm, A., Kouzarides, T., Lobanenkov, V., and Renkawitz, R. (2000). Transcriptional repression by the insulator protein CTCF involves histone deacetylases. Nucleic Acids Res *28*, 1707-13.

Lyko, F., Brenton, J. D., Surani, M. A., and Paro, R. (1997). An imprinting element from the mouse *H19* locus functions as a silencer in Drosophila. Nat Genet *16*, 171-3.

Lyko, F., Buiting, K., Horsthemke, B., and Paro, R. (1998). Identification of a silencing element in the human 15q11-q13 imprinting center by using transgenic Drosophila. Proc Natl Acad Sci U S A *95*, 1698-702.

Lyle, R., Watanabe, D., te Vruchte, D., Lerchner, W., Smrzka, O. W., Wutz, A., Schageman, J., Hahner, L., Davies, C., and Barlow, D. P. (2000). The imprinted antisense RNA at the *Igf2r* locus overlaps but does not imprint Mas1. Nat Genet *25*, 19-21.

MacDonald, H. R., and Wevrick, R. (1997). The necdin gene is deleted in Prader-Willi syndrome and is imprinted in human and mouse. Hum Mol Genet *6*, 1873-8.

Magewu, A. N., and Jones, P. A. (1994). Ubiquitous and tenacious methylation of the CpG site in codon 248 of the p53 gene may explain its frequent appearance as a mutational hot spot in human cancer. Mol Cell Biol *14*, 4225-32.

Mascari, M. J., Gottlieb, W., Rogan, P. K., Butler, M. G., Waller, D. A., Armour, J. A., Jeffreys, A. J., Ladda, R. L., and Nicholls, R. D. (1992). The frequency of uniparental disomy in Prader-Willi syndrome. Implications for molecular diagnosis. N Engl J Med *326*, 1599-607.

Matsuo, K., Silke, J., Georgiev, O., Marti, P., Giovannini, N., and Rungger, D. (1998). An embryonic demethylation mechanism involving binding of transcription factors to replicating DNA. Embo J *17*, 1446-53.

Matsuura, T., Sutcliffe, J. S., Fang, P., Galjaard, R. J., Jiang, Y. H., Benton, C. S., Rommens, J. M., and Beaudet, A. L. (1997). *De novo* truncating mutations in E6-AP ubiquitin-protein ligase gene (UBE3A) in Angelman syndrome. Nat Genet *15*, 74-7.

Mayer, W., Niveleau, A., Walter, J., Fundele, R., and Haaf, T. (2000). Demethylation of the zygotic paternal genome. Nature *403*, 501-2.

Mayer, W., Smith, A., Fundele, R., and Haaf, T. (2000). Spatial separation of parental genomes in preimplantation mouse embryos. J Cell Biol *148*, 629-34.

McDonald, L. E., Paterson, C. A., and Kay, G. F. (1998). Bisulfite genomic sequencing-derived methylation profile of the xist gene throughout early mouse development. Genomics *54*, 379-86.

McGhee, J. D., Wood, W. I., Dolan, M., Engel, J. D., and Felsenfeld, G. (1981). A 200 base pair region at the 5' end of the chicken adult beta-globin gene is accessible to nuclease digestion. Cell *27*, 45-55.

Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L., and Bird, A. P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell *58*, 499-507.

Merika, M., and Orkin, S. H. (1993). DNA-binding specificity of GATA family transcription factors. Mol Cell Biol *13*, 3999-4010.

Mermoud, J. E., Popova, B., Peters, A. H., Jenuwein, T., and Brockdorff, N. (2002). Histone H3 lysine 9 methylation occurs rapidly at the onset of random X chromosome inactivation. Curr Biol *12*, 247-51.

Monk, M., Boubelik, M., and Lehnert, S. (1987). Temporal and regional changes in DNA methylation in the embryonic, extraembryonic and germ cell lineages during mouse embryo development. Development *99*, 371-82.

Morgan, H. D., Sutherland, H. G., Martin, D. I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. Nat Genet *23*, 314-8.

Mucha, M., Lisowska, K., Goc, A., and Filipski, J. (2000). Nuclease-hypersensitive chromatin formed by a CpG island in human DNA cloned as an artificial chromosome in yeast. J Biol Chem *275*, 1275-8.

Mummaneni, P., Walker, K. A., Bishop, P. L., and Turker, M. S. (1995). Epigenetic gene inactivation induced by a cis-acting methylation center. J Biol Chem *270*, 788-92.

Murray, D., Precht, P., Balakir, R., and Horton, W. E., Jr. (2000). The transcription factor deltaEF1 is inversely expressed with type II collagen mRNA and can repress Col2a1 promoter activity in transfected chondrocytes. J Biol Chem *275*, 3610-8.

Muscatelli, F., Abrous, D. N., Massacrier, A., Boccaccio, I., Le Moal, M., Cau, P., and Cremer, H. (2000). Disruption of the mouse Necdin gene results in hypothalamic and behavioral alterations reminiscent of the human Prader-Willi syndrome. Hum Mol Genet *9*, 3101-10.

Myohanen, S. K., Baylin, S. B., and Herman, J. G. (1998). Hypermethylation can selectively silence individual p16ink4A alleles in neoplasia. Cancer Res *58*, 591-3.

Nakada, Y., Taniura, H., Uetsuki, T., Inazawa, J., and Yoshikawa, K. (1998). The human chromosomal gene for necdin, a neuronal growth suppressor, in the Prader-Willi syndrome deletion region. Gene *213*, 65-72.

Nan, X., Meehan, R. R., and Bird, A. (1993). Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. Nucleic Acids Res *21*, 4886-92.

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-9.

Neumann, B., Kubicka, P., and Barlow, D. P. (1995). Characteristics of imprinted genes. Nat Genet *9*, 12-3.

Ng, H. H., Zhang, Y., Hendrich, B., Johnson, C. A., Turner, B. M., Erdjument-Bromage, H., Tempst, P., Reinberg, D., and Bird, A. (1999). MBD2 is a transcriptional repressor belonging to the MeCP1 histone deacetylase complex. Nat Genet *23*, 58-61.

Nicholls, R. D. (2000). The impact of genomic imprinting for neurobehavioral and developmental disorders. J Clin Invest *105*, 413-8.

Noma, K., Allis, C. D., and Grewal, S. I. (2001). Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science *293*, 1150-5.

Norris, D. P., Patel, D., Kay, G. F., Penny, G. D., Brockdorff, N., Sheardown, S. A., and Rastan, S. (1994). Evidence that random and imprinted Xist expression is controlled by preemptive methylation. Cell *77*, 41-51.

Ohta, T., Khadake, J. R., Rodriquez-Jato, S., Mione, C., Knepper, J. L., Svoboda, P., McCarrey, J. R., Schultz, R. M., Yang, T. P., and Nicholls, R. D. (2001). Imprint Establishment In Prader-Willi/Angelman Syndromes Is Controlled By Nrf-1 And Ctcf Transcription Factors At Alternative Promoters Of Snurf-Snrpn. American Journal of Human Genetics *69*.

Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999). DNA methyltransferases Dnmt3a and Dnmt3b are essential for *de novo* methylation and mammalian development. Cell *99*, 247-57.

Olek, A., and Walter, J. (1997). The pre-implantation ontogeny of the *H19* methylation imprint. Nat Genet *17*, 275-6.

Orstavik, K. H., Tangsrud, S. E., Kiil, R., Hansteen, I. L., Steen-Johnsen, J., Cassidy, S. B., Martony, A., Anvret, M., Tommerup, N., and Brondum-Nielsen, K. (1992). Prader-Willi syndrome in a brother and sister without cytogenetic or detectable molecular genetic abnormality at chromosome 15q11q13. Am J Med Genet *44*, 534-8.

Packer, A. I., Manova, K., and Bachvarova, R. F. (1993). A discrete LINE-1 transcript in mouse blastocysts. Dev Biol *157*, 281-3.

Panning, B., and Jaenisch, R. (1996). DNA hypomethylation can activate Xist expression and silence X-linked genes. Genes Dev *10*, 1991-2002.

Parekh, B. S., and Maniatis, T. (1999). Virus infection leads to localized hyperacetylation of histones H3 and H4 at the IFN-beta promoter. Mol Cell *3*, 125-9.

Paulsen, M., El-Maarri, O., Engemann, S., Strodicke, M., Franck, O., Davies, K., Reinhardt, R., Reik, W., and Walter, J. (2000). Sequence conservation and variability of imprinting in the Beckwith- Wiedemann syndrome gene cluster in human and mouse. Hum Mol Genet *9*, 1829-41.

Peters, A. H., Mermoud, J. E., O'Carroll, D., Pagani, M., Schweizer, D., Brockdorff, N., and Jenuwein, T. (2002). Histone H3 lysine 9 methylation is an epigenetic imprint of facultative heterochromatin. Nat Genet *30*, 77-80.

Piras, G., El Kharroubi, A., Kozlov, S., Escalante-Alcalde, D., Hernandez, L., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Stewart, C. L. (2000). Zac1 (Lot1), a potential tumor suppressor gene, and the gene for epsilon-sarcoglycan are maternally imprinted genes: identification by a subtractive screen of novel uniparental fibroblast lines. Mol Cell Biol *20*, 3308-15.

Prendergast, G. C., Lawe, D., and Ziff, E. B. (1991). Association of Myn, the murine homolog of max, with c-Myc stimulates methylation-sensitive DNA binding and ras cotransformation. Cell *65*, 395-407.

Ramsahoye, B. H., Biniszkiewicz, D., Lyko, F., Clark, V., Bird, A. P., and Jaenisch, R. (2000). Non-CpG methylation is prevalent in embryonic stem cells and may be mediated by DNA methyltransferase 3a. Proc Natl Acad Sci U S A *97*, 5237-42.

Razin, A., and Cedar, H. (1994). DNA methylation and genomic imprinting. Cell *77*, 473-6.

Reik, W., and Walter, J. (2001). Evolution of imprinting mechanisms: the battle of the sexes begins in the zygote. Nat Genet *27*, 255-6.

Rougeulle, C., Glatt, H., and Lalande, M. (1997). The Angelman syndrome candidate gene, UBE3A/E6-AP, is imprinted in brain. Nat Genet *17*, 14-5.

Runte, M., Huttenhofer, A., Gross, S., Kiefmann, M., Horsthemke, B., and Buiting, K. (2001). The IC-SNURF-SNRPN transcript serves as a host for multiple small nucleolar RNA species and as an antisense RNA for UBE3A. Hum Mol Genet *10*, 2687-700.

Saitoh, N., Bell, A. C., Recillas-Targa, F., West, A. G., Simpson, M., Pikaart, M., and Felsenfeld, G. (2000). Structural and functional conservation at the boundaries of the chicken beta-globin domain. Embo J *19*, 2315-22.

Saitoh, S., Buiting, K., Rogan, P. K., Buxton, J. L., Driscoll, D. J., Arnemann, J., Konig, R., Malcolm, S., Horsthemke, B., and Nicholls, R. D. (1996). Minimal definition of the imprinting center and fixation of chromosome 15q11-q13 epigenotype by imprinting mutations. Proc Natl Acad Sci U S A *93*, 7811-5.

Saitoh, S., and Wada, T. (2000). Parent-of-origin specific histone acetylation and reactivation of a key imprinted gene locus in Prader-Willi syndrome. Am J Hum Genet *66*, 1958-62.

Sambrook, J., and Russell, D. W. (2001). Molecular Cloning: A laboratory manual, Volume 1 (Cold Spring Harbour, N.Y.: Cold Spring Harbour Laboratory Press).

Sanford, J. P., Clark, H. J., Chapman, V. M., and Rossant, J. (1987). Differences in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. Genes Dev *1*, 1039-46.

Santos, F., Hendrich, B., Reik, W., and Dean, W. (2002). Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol *241*, 172-82.

Sasaki, H., Jones, P. A., Chaillet, J. R., Ferguson-Smith, A. C., Barton, S. C., Reik, W., and Surani, M. A. (1992). Parental imprinting: potentially active chromatin of the repressed maternal allele of the mouse insulin-like growth factor II (*Igf2*) gene. Genes Dev *6*, 1843-56.

Schmidt, J. V., Matteson, P. G., Jones, B. K., Guan, X. J., and Tilghman, S. M. (2000). The Dlk1 and Gtl2 genes are linked and reciprocally imprinted. Genes Dev *14*, 1997-2002.

Schmidt-Ott, U. (2001). Different ways to make a head. Bioessays *23*, 8-11.

Schwartz, S., Zhang, Z., Frazer, K. A., Smit, A., Riemer, C., Bouck, J., Gibbs, R., Hardison, R., and Miller, W. (2000). PipMaker--a web server for aligning two genomic DNA sequences. Genome Res *10*, 577-86.

Schweizer, J., Zynger, D., and Francke, U. (1999). In vivo nuclease hypersensitivity studies reveal multiple sites of parental origin-dependent differential chromatin conformation in the 150 kb SNRPN transcription unit. Hum Mol Genet *8*, 555-66.

Shahbazian, M. D., Antalffy, B., Armstrong, D. L., and Zoghbi, H. Y. (2002). Insight into Rett syndrome: MeCP2 levels display tissue- and cell- specific differences and correlate with neuronal maturation. Hum Mol Genet *11*, 115-24.

Shani, M., Admon, S., and Yaffe, D. (1984). The methylation state of 2 muscle-specific genes: restriction enzyme analysis did not detect a correlation with expression. Nucleic Acids Res *12*, 7225-34.

Shemer, R., Birger, Y., Dean, W. L., Reik, W., Riggs, A. D., and Razin, A. (1996). Dynamic methylation adjustment and counting as part of imprinting mechanisms. Proc Natl Acad Sci U S A *93*, 6371-6.

Shemer, R., Birger, Y., Riggs, A. D., and Razin, A. (1997). Structure of the imprinted mouse Snrpn gene and establishment of its parental-specific methylation pattern. Proc Natl Acad Sci U S A *94*, 10267-72.

Siegfried, Z., Eden, S., Mendelsohn, M., Feng, X., Tsuberi, B. Z., and Cedar, H. (1999). DNA methylation represses transcription *in vivo*. Nat Genet *22*, 203-6.

Sleutels, F., Zwart, R., and Barlow, D. P. (2002). The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature *415*, 810-3.

Smilinich, N. J., Day, C. D., Fitzpatrick, G. V., Caldwell, G. M., Lossie, A. C., Cooper, P. R., Smallwood, A. C., Joyce, J. A., Schofield, P. N., Reik, W., Nicholls, R. D., Weksberg, R., Driscoll, D. J., Maher, E. R., Shows, T. B., and Higgins, M. J. (1999). A maternally methylated CpG island in KvLQT1 is associated with an antisense paternal transcript and loss of imprinting in Beckwith- Wiedemann syndrome. Proc Natl Acad Sci U S A *96*, 8064-9.

Solter, D. (1998). Imprinting. Int J Dev Biol *42*, 951-4.

Srivastava, M., Hsieh, S., Grinberg, A., Williams-Simons, L., Huang, S. P., and Pfeifer, K. (2000). *H19* and *Igf2* monoallelic expression is regulated in two distinct ways by a shared cis acting regulatory region upstream of *H19*. Genes Dev *14*, 1186-95.

Stewart, C. L., Stuhlmann, H., Jahner, D., and Jaenisch, R. (1982). *De novo* methylation, expression, and infectivity of retroviral genomes introduced into embryonal carcinoma cells. Proc Natl Acad Sci U S A *79*, 4098-102.

Stoger, R., Kajimura, T. M., Brown, W. T., and Laird, C. D. (1997). Epigenetic variation illustrated by DNA methylation patterns of the fragile-X gene FMR1. Hum Mol Genet *6*, 1791-801.

Stoger, R., Kubicka, P., Liu, C. G., Kafri, T., Razin, A., Cedar, H., and Barlow, D. P. (1993). Maternal-specific methylation of the imprinted mouse *Igf2*r locus identifies the expressed locus as carrying the imprinting signal. Cell *73*, 61-71.

Surani, M. A., Barton, S. C., and Norris, M. L. (1984). Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature *308*, 548-50.

Sutcliffe, J. S., Nakao, M., Christian, S., Orstavik, K. H., Tommerup, N., Ledbetter, D. H., and Beaudet, A. L. (1994). Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. Nat Genet *8*, 52-8.

Szabo, P. E., Pfeifer, G. P., and Mann, J. R. (1998). Characterization of novel parent-specific epigenetic modifications upstream of the imprinted mouse *H19* gene. Mol Cell Biol *18*, 6767-76.

Szabo, P. E., Tang, S. H., Reed, M. R., Silva, F. J., Tsark, W. M., and Mann, J. R. (2002). The chicken beta-globin insulator element conveys chromatin boundary activity but not imprinting at the mouse *Igf2/H19* domain. Development *129*, 897-904.

Tada, T., and Tada, M. (2001). Toti-/pluripotential stem cells and epigenetic modifications. Cell Struct Funct *26*, 149-60.

Takada, S., Paulsen, M., Tevendale, M., Tsai, C. E., Kelsey, G., Cattanach, B. M., and Ferguson-Smith, A. C. (2002). Epigenetic analysis of the Dlk1-Gtl2 imprinted domain on mouse chromosome 12: implications for imprinting control from comparison with *Igf2-H19*. Hum Mol Genet *11*, 77-86.

Tanaka, M., Puchyr, M., Gertsenstein, M., Harpal, K., Jaenisch, R., Rossant, J., and Nagy, A. (1999). Parental origin-specific expression of Mash2 is established at the time of implantation with its imprinting mechanism highly resistant to genome-wide demethylation. Mech Dev *87*, 129-42.

Tate, P., Skarnes, W., and Bird, A. (1996). The methyl-CpG binding protein MeCP2 is essential for embryonic development in the mouse. Nat Genet *12*, 205-8.

Tazi, J., and Bird, A. (1990). Alternative chromatin structure at CpG islands. Cell *60*, 909-20.

Terraz, C., Toman, D., Delauche, M., Ronco, P., and Rossert, J. (2001). delta Ef1 binds to a far upstream sequence of the mouse pro-alpha 1(I) collagen gene and represses its expression in osteoblasts. J Biol Chem *276*, 37011-9.

Thomassin, H., Flavin, M., Espinas, M. L., and Grange, T. (2001). Glucocorticoid-induced DNA demethylation and gene memory during development. Embo J *20*, 1974-83.

Thompson, E. M., Legouy, E., Christians, E., and Renard, J. P. (1995). Progressive maturation of chromatin structure regulates HSP70.1 gene expression in the preimplantation mouse embryo. Development *121*, 3425-37.

Thorvaldsen, J. L., Duran, K. L., and Bartolomei, M. S. (1998). Deletion of the *H19* differentially methylated domain results in loss of imprinted expression of *H19* and *Igf2*. Genes Dev *12*, 3693-702.

Tilghman, S. M. (1999). The sins of the fathers and mothers: genomic imprinting in mammalian development. Cell *96*, 185-93.

Tremblay, K. D. (1998). Bisulfite methylation analysis of single DNA strands. Technical TIps Online *1*, P01242.

Tremblay, K. D., Duran, K. L., and Bartolomei, M. S. (1997). A 5' 2-kilobase-pair region of the imprinted mouse *H19* gene exhibits exclusive paternal methylation throughout development. Mol Cell Biol *17*, 4322-9.

Tremblay, K. D., Saam, J. R., Ingram, R. S., Tilghman, S. M., and Bartolomei, M. S. (1995). A paternal-specific methylation imprint marks the alleles of the mouse *H19* gene. Nat Genet *9*, 407-13.

Tsai, T. F., Armstrong, D., and Beaudet, A. L. (1999). Necdin-deficient mice do not show lethality or the obesity and infertility of Prader-Willi syndrome. Nat Genet *22*, 15-6.

Tucker, K. L., Beard, C., Dausmann, J., Jackson-Grusby, L., Laird, P. W., Lei, H., Li, E., and Jaenisch, R. (1996). Germ-line passage is required for establishment of methylation and expression patterns of imprinted but not of nonimprinted genes. Genes Dev *10*, 1008-20.

Uetsuki, T., Takagi, K., Sugiura, H., and Yoshikawa, K. (1996). Structure and expression of the mouse necdin gene. Identification of a postmitotic neuron-restrictive core promoter. J Biol Chem *271*, 918-24.

Utley, R. T., Ikeda, K., Grant, P. A., Cote, J., Steger, D. J., Eberharter, A., John, S., and Workman, J. L. (1998). Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. Nature *394*, 498-502.

Voo, K. S., Carlone, D. L., Jacobsen, B. M., Flodin, A., and Skalnik, D. G. (2000). Cloning of a mammalian transcriptional activator that binds unmethylated CpG motifs and shares a CXXC domain with DNA methyltransferase, human trithorax, and methyl-CpG binding domain protein 1. Mol Cell Biol *20*, 2108-21.

Wade, P. A., Gegonne, A., Jones, P. L., Ballestar, E., Aubry, F., and Wolffe, A. P. (1999). Mi-2 complex couples DNA methylation to chromatin remodelling and histone deacetylation. Nat Genet *23*, 62-6.

Walsh, C. P., and Bestor, T. H. (1999). Cytosine methylation and mammalian development. Genes Dev *13*, 26-34.

Warnecke, P. M., and Clark, S. J. (1999). DNA methylation profile of the mouse skeletal alpha-actin promoter during development and differentiation. Mol Cell Biol *19*, 164-72.

Warnecke, P. M., Mann, J. R., Frommer, M., and Clark, S. J. (1998). Bisulfite sequencing in preimplantation embryos: DNA methylation profile of the upstream region of the mouse imprinted *H19* gene. Genomics *51*, 182-90.

Watrin, F., Roeckel, N., Lacroix, L., Mignon, C., Mattei, M. G., Disteche, C., and Muscatelli, F. (1997). The mouse Necdin gene is expressed from the paternal allele only and lies in the 7C region of the mouse chromosome 7, a region of conserved synteny to the human Prader-Willi syndrome region. Eur J Hum Genet *5*, 324-32.

Weber, M., Milligan, L., Delalbre, A., Antoine, E., Brunel, C., Cathala, G., and Forne, T. (2001). Extensive tissue-specific variation of allelic methylation in the *Igf2* gene during mouse fetal development: relation to expression and imprinting. Mech Dev *101*, 133-41.

Weintraub, H., and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. Science *193*, 848-56.

Wu, C. (1989). Analysis of hypersensitive sites in chromatin. Methods Enzymol *170*, 269-89.

Wutz, A., and Jaenisch, R. (2000). A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. Mol Cell *5*, 695-705.

Wutz, A., Smrzka, O. W., and Barlow, D. P. (1998). Making sense of imprinting the mouse and human IGF2R loci. Novartis Found Symp *214*, 251-9.

Wylie, A. A., Murphy, S. K., Orton, T. C., and Jirtle, R. L. (2000). Novel imprinted DLK1/GTL2 domain on human chromosome 14 contains motifs that mimic those implicated in IGF2/*H19* regulation. Genome Res *10*, 1711-8.

Xin, Z., Allis, C. D., and Wagstaff, J. (2001). Parent-specific complementary patterns of histone H3 lysine 9 and H3 lysine 4 methylation at the Prader-Willi syndrome imprinting center. Am J Hum Genet *69*, 1389-94.

Xu, G. L., Bestor, T. H., Bourc'his, D., Hsieh, C. L., Tommerup, N., Bugge, M., Hulten, M., Qu, X., Russo, J. J., and Viegas-Pequignot, E. (1999). Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature *402*, 187-91.

Yang, T., Adamson, T. E., Resnick, J. L., Leff, S., Wevrick, R., Francke, U., Jenkins, N. A., Copeland, N. G., and Brannan, C. I. (1998). A mouse model for Prader-Willi syndrome imprinting-centre mutations. Nat Genet *19*, 25-31.

Yisraeli, J., Adelstein, R. S., Melloul, D., Nudel, U., Yaffe, D., and Cedar, H. (1986). Muscle-specific activation of a methylated chimeric actin gene. Cell *46*, 409-16.

Yokochi, T., and Robertson, K. D. (2002). Preferential methylation of unmethylated DNA by Mammalian *de novo* DNA methyltransferase dnmt3a. J Biol Chem *277*, 11735-45.

Zeschnigk, M., Schmitz, B., Dittrich, B., Buiting, K., Horsthemke, B., and Doerfler, W. (1997). Imprinted segments in the human genome: different DNA methylation patterns in the Prader-Willi/Angelman syndrome region as determined by the genomic sequencing method. Hum Mol Genet *6*, 387-95.

Zwart, R., Sleutels, F., Wutz, A., Schinkel, A. H., and Barlow, D. P. (2001). Bidirectional action of the *Igf*2r imprint control element on upstream and downstream imprinted genes. Genes Dev *15*, 2361-6.