ULLA AAPOLA

Characterisation and Functional Studies of the DNA Cytosine-5-Methyltransferase 3-like Gene

■

ACADEMIC DISSERTATION

To be presented, with the permission of
the Faculty of Medicine of the University of Tampere,
for public discussion in the auditorium of Finn-Medi 1,
Biokatu 6, Tampere, on September 10th, 2004, at 12 o'clock.

Acta Universitatis Tamperensis 1029

ACADEMIC DISSERTATION
University of Tampere, Institute of Medical Technology
Finland

Supervised by
Docent Pärt Peterson
University of Tampere

Reviewed by
Docent Antero Salminen
University of Kuopio
Professor Heinrich Leonhardt
Ludwig Maximilians University
Munich, Germany

Distribution

University of Tampere
Bookshop TAJU
P.O. Box 617
33014 University of Tampere
Finland

Tel. +358 3 215 6055
Fax +358 3 215 7685
taju@uta.fi
http://granum.uta.fi

Cover design by
Juha Siro

Printed dissertation
Acta Universitatis Tamperensis 1029
ISBN 951-44-6066-9
ISSN 1455-1616

Tampereen Yliopistopaino Oy – Juvenes Print
Tampere 2004

Electronic dissertation
Acta Electronica Universitatis Tamperensis 375
ISBN 951-44-6067-7
ISSN 1456-954X
http://acta.uta.fi

DNMT3L-GEENIN KARAKTERISOINTI JA TOIMINTA

DNA:n metylaatio on tärkeä säätelymekanismi, joka vaikuttaa geenien ekspressioon, alkiokehitykseen, geneettiseen leimautumiseen ja X-kromosomin inaktivaatioon. Syövän synnyssä on usein vaikuttavana osatekijänä viallinen DNA:n metylointi. Geenien ekspressiota voidaan säädellä myös muokkaamalla DNA:n pakkaamiseen osallistuvia histoni-proteiineja. Yleisimmin histoneja muokataan asetyloimalla, metyloimalla ja fosforyloimalla. Nisäkässoluissa DNA:ta metyloivat neljä erilaista DNA-sytosiini-5-metyylitransferaasia, DNMT1, DNMT2, DNMT3A ja DNMT3B. DNMT1 toimii pääasiallisesti replikaation aikana kopioiden olemassaolevat metyyliryhmät replikaatiossa syntetisoituun DNA-juosteeseen. DNA-sytosiini-5-metyylitransferaasi-3 -perheen DNMT3A- ja DNMT3B-proteiinit ovat *de novo* metyylitransferaaseja, jotka liittävät uusia metyyliryhmiä DNA-ketjuun. *De novo* metyylitransferaasit ovat välttämättömiä alkiokehityksen aikana.

Tässä väitöskirjatyössä olemme identifioineet *DNMT3*-geeniperheeseen kuuluvan uuden geenin, *DNMT3L*:n (DNA cytosine-5-methyltransferase 3-like). Olemme selvittäneet sekä ihmisen että hiiren *DNMT3L*/*Dnmt3L*-geenien emäsjärjestykset ja genomiset rakenteet. Tutkimustulostemme mukaan ihmisen ja hiiren geenit ovat hyvin samankaltaisia. RT-PCR-menetelmän avulla osoitimme, että *DNMT3L/Dnmt3L* ilmenee pääasiallisesti kiveksissä, mutta myös munasarjoissa, kateenkorvassa ja alkion kudoksissa. *DNMT3L/Dnmt3L*-geenin koodaaman proteiinin tärkeimmät motiivit ovat tumakohdennussignaali ja sinkkisormialue, joka koostuu kahdesta eri sinkkisormimotiivista: C2C2-tyypin sinkkisormesta sekä PHD-sinkkisormen kaltaisesta rakenteesta. *Dnmt3L*-geenin säätelyä tutkimme promoottorianalyysin avulla. Reportterianalyysia, EMSA:a, kohdennettua mutageneesiä, genomista bisulfidisekvensointia ja reaaliaikaista kvantitatiivista RT-PCR-menetelmää käyttäen tunnistimme *Dnmt3L*:n minimaalipromoottorialueen ja osoitimme, että geeniä säädellään Sp1/Sp3-transkriptiotekijöiden sekä DNA:n metylaation ja histonideasetylaation kautta. Suurin samankaltaisuus DNMT3-perheen jäsenten välillä löytyi sinkkisormialueelta. Aiemmat tutkimukset osoittivat, että Dnmt3a ja Dnmt3b sitoutuvat sinkkisormiensa avulla transkription säätelyssä toimivaan histonideasetylaasi 1 (HDAC1) -proteiiniin hyödyntäen HDAC1:n entsyymiaktiivisuutta estäessään muiden geenien transkriptiota. Tutkimustulostemme mukaan myös DNMT3L sitoutuu sinkkisormiensa avulla HDAC1-proteiiniin ja säätelee näin muiden geenien ilmenemistä. Osoitimme, että PHD-sinkkisormirakenne on DNMT3L:n pääasiallinen repressiomotiivi.

CONTENTS

LIST OF ORIGINAL COMMUNICATIONS

This thesis is based on the following communications, which are referred by their Roman numerals (I-IV) in the text.

I Aapola U, Shibuya K, Scott HS, Ollila J, Vihinen M, Heino M, Shintani A, Kawasaki K, Minoshima S, Krohn K, Antonarakis SE, Shimizu N, Kudoh J and Peterson P (2000): Isolation and initial characterization of a novel zinc finger gene, *DNMT3L*, on 21q22.3, related to cytosine-5-methyltransferase 3 gene family. Genomics 65:293-298.

II Aapola U, Lyle R, Krohn K, Antonarakis SE and Peterson P (2001): Isolation and initial characterization of the mouse *Dnmt3l* gene. Cytogenet Cell Genet 92:122-126.

III Aapola U, Mäenpää K, Kaipia A and Peterson P (2004): Epigenetic modifications affect *Dnmt3L* expression. Biochem J 380:705-713.

IV Aapola U, Liiv I and Peterson P (2002): Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity. Nucleic Acids Res 30:3602-3608.

In addition, some unpublished data is presented in the thesis.

ABBREVIATIONS

aa	amino acid
AdoMet	S-adenosyl-L-methionine
Air	antisense Igf2r RNA
AIRE	autoimmune regulator
APECED	autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy
ATP	adenosine triphoshate
ATRX	alpha-thalassemia/mental retardation syndrome, x-linked
5-aza-CdR	5-aza-2'-deoxycytidine
BORIS	brother of the regulator of imprinted sites
bp	base pair
CAT	chloramphenicol acetyltransferase
CBP	CREB binding protein
CREB	cAMP responsive element binding protein
CTCF	CCCTC binding factor
CTP	cytidine triphosphate
Daxx	death-domain-associated protein
DDM1	decrease in DNA methylation
DMAP	DNMT1 associated protein
DMR	differentially methylated region
DNA	deoxyribonucleic acid
cDNA	complementary deoxyribonucleic acid
DNMT	DNA cytosine-5-methyltransferase
DNMT3L	DNA cytosine-5-methyltransferase 3-like
E2F	E2 promoter binding factor
ELISA	enzyme linked immunosorbent assay
EMSA	electrophoretic mobility shift assay
ES	embryonic stem cell
EST	expressed sequence tag
EZH	enhancer of zeste homolog
FBHM	familial biparental hydatidiform mole
FMR	fragile X mental retardation
Gcn	general control nonrepressed
GFP	green fluorescent protein
G3PDH	glyceraldehyde-3-phosphate dehydrogenase
GST	glutathione S-transferase
HAT	histone acetyltransferse
Hda	histone deacetylase A
HDAC	histone deacetylase
HMTase	histone methyltransferase
HP	heterochromatin protein
ICF	immunodeficiency, centromeric instability, facial anomalies
ICOSL	inducible T-cell co-stimulator ligand
ICR	imprinting control region
Igf	insuline-like growth factor
Igf2r	insuline-like growth factor 2 receptor
ISWI	imitation switch
kb	kilobase

kDa	kilodalton
LOI	loss of imprinting
Lsh	lymphoid-specific helicase
MBD, MeCP	methyl-CpG-binding protein
NAD	nicotinamide adenine dinucleotide
NLS	nuclear localisation signal
NuRD	nucleosome remodelling and deacetylation
OMIM	online mendelian inheritance in man
PCNA	proliferating cell nuclear antigen
PCR	polymerase chain reaction
PGC	primordial germ cell
PHD	plant homeodomain
PIAS	protein inhibitor of activated STAT
PML-RAR	promyelocytic leukaemia-retinoic acid receptor
PRMT	protein arginine methyltransferase
QPCR	quantitative real-time RT-PCR
RACE	rapid amplification of cDNA ends
Rasgrf1	Ras protein-specific guanine nucleotide-releasing factor 1
Rb	retinoblastoma
RP58	repressor protein with a predicted molecular mass of 58 kDa
RD	rhabdomyosarcoma cell
RIZ	retinoblastoma protein-interacting zinc finger
RNA	ribonucleic acid
mRNA	messenger ribonucleic acid
tRNA	total ribonucleic acid
RT-PCR	reverse transcriptase polymerase chain reaction
SDS-PAGE	sodium dodecyl sulphate polyacrylamide gel electophoresis
SET	Suvar3-9, enhancer of zeste, trithorax
Sir	silent information regulator
Sirt	sirtuin
Snrpn	small nuclear ribonucleoprotein particle-associated peptide N
Sp	specificity protein
SWI-SNF	mating type switching/sucrose nonfermenting
STAT	signal transducer and activator of transcription
SUMO	small ubiquitin-like modifier
SUV39H	suppressor of variegation 3-9 homolog, suvar3-9
TAF250	TBP associated factor
TBP	TATA binding protein
TSA	trichostatin A
Ubc9	ubiquitin (SUMO-1) conjugating enzyme
UTR	untranslated region
X-gal	5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside

ABSTRACT

DNA methylation is important in the regulation of gene expression, embryonic development, genomic imprinting and X chromosome inactivation. Abnormal methylation has been associated with carcinogenesis. So far, four separate genes, *DNMT1, DNMT2, DNMT3A* and *DNMT3B* encoding DNA cytosine-5-methyltransferases have been cloned from mammalian cells. DNMT1 is a maintenance methyltransferase, copying existing methylation patterns into newly replicated DNA strand. Dnmt3a and Dnmt3b are required for genome-wide *de novo* methylation and are essential for mammalian development. Gene expression is also regulated through modification of histone proteins that are essential for folding and packing of DNA inside the cells. Usually, histones are modified by acetylation, methylation and phosphorylation.

In this thesis we have identified a novel member of the DNMT3 gene family, DNA cytosine-5-methyltransferase 3-like gene (DNMT3L). DNMT3L protein is similar to DNMT3A/Dnmt3a and DNMT3B/Dnmt3b, sharing the greatest similarity with other DNMT3 family members at the zinc finger region in the N-terminal part of the DNMT3L/Dnmt3L. In order to perform functional studies, the mouse *Dnmt3L* gene was also identified. Human and mouse genes showed high identity with each other and were both strongly expressed in testis and to a lesser extent also in ovary, thymus and fetal tissues. To study the regulation of *Dnmt3L*, we isolated the *Dnmt3L* promoter region and identified a minimal promoter area and regulatory elements in it. We demonstrated that *Dnmt3L* was regulated through Sp1/Sp3 transcription factors and epigenetic chromatin modifications, DNA methylation and histone deacetylation. Since Dnmt3a and Dnmt3b function as transcriptional repressors using zinc finger-domain to interact with histone deacetylase, HDAC1, we tested the repressional capacity of DNMT3L. Our data indicated that DNMT3L represses transcription by binding directly to HDAC1 protein and demonstrated the PHD-like zinc finger as a main repressional domain of DNMT3L.

1. INTRODUCTION

The Human Genome Project began in 1990. The fundamental goal was to determine the sequence of 3 x 10^9 base pairs in human genome and store all the information in databases (http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml). The sequence of the first human chromosome (chromosome 22) was completed in 1999 (Dunham et al. 1999), and was soon followed by the sequence of chromosome 21 (Hattori et al. 2000). A working draft of the entire human genome sequence was announced in 2001 (International Human Genome Sequencing Consortium 2001, Venter et al. 2001) and the Human Genome Project ended in 2003 with an almost complete version of the human genome (see Science April 11, 2003 and Nature April 24, 2003). From approximately 30 000 human genes 15 000 full length human cDNAs were identified during the Human Genome Project (Collins et al. 2003).

The human genome, 3 x 10^9 base pairs of DNA, is folded inside each cell with the help of histone proteins to form a compact organisation called chromatin. Activity of the genes in the genome depends on the structural accessibility of the chromatin. Chromatin structure can be modified via targeted covalent modifications to the tails of histone proteins. At the DNA level, genes can be regulated by modifying cytosines within the CpG dinucleotide. These epigenetic chromatin modifications are heritable through cell divisions and regulate chromatin structure and gene expression without changing the DNA sequence (Khorasanizadeh 2004).

This study was started while the Human Genome Project was still ongoing and the DNA sequence of the human genome was largely unknown. The initial purpose of the study was to identify a novel human gene located on chromosome 21q22.3, a region containing several unidentified disease-related genes. Using chromosome 21 sequencing data provided as part of the Human Genome Project, we isolated the DNA cytosine-5-methyltransferase 3-like (*DNMT3L*) gene. Here we describe the isolation and characterisation of the human and mouse *DNMT3L*/*Dnmt3L* genes and indicate the epigenetic modifications in the *Dnmt3L* promoter regulating gene expression. We moreover demonstrate that DNMT3L protein is a chromatin modifier showing repressional capacity through interaction with histone deacetylase, HDAC1 protein.

2. REVIEW OF THE LITERATURE

2.1 EPIGENETICS AND CHROMATIN MODIFICATIONS

A typical human cell contains 6 x 10^9 base pairs of DNA, which is packed in 46 chromosomes in the cell nucleus. The architecture of the chromosomes is accomplished with the help of proteins that bind to and fold the DNA. The complex, formed by the DNA molecule and associated proteins, is called chromatin. The fundamental repeating unit of chromatin is a nucleosome. Nucleosomes consist of an octamer of histone proteins H2A, H2B, H3 and H4 and of 146 bp of double stranded DNA (2 nm) coiled around the histone octamer. Adjacent nucleosomes are connected with short linker DNA, generating a structure known as a "beads-on-a-string" form of chromatin. With the help of histone H1 proteins, the 10 nm string of beads structure is further folded into 30 nm chromatin fiber. In mitosis chromatin is greatly condensed and the 30 nm chromatin fibers form loops that are attached to a central scaffold consisting of nonhistone acidic proteins. Subsequently, the loop-scaffold chromatin is further coiled to form even more compact 1200 nm mitotic chromosomes (Strachan and Read 1999).

During the interphase, cells perform transcription, translation and DNA replication containing both condensed and decondensed forms of chromatin. Condensed chromatin is called heterochromatin, which represents transcriptionally silent DNA. The decondensed form of chromatin is transcriptionally active and called euchromatin. Compact chromatin structure restricts regulatory proteins from interacting with DNA, thus preventing gene expression, replication and repair. The accessible chromatin structure needed to perform these essential functions is achieved through chromatin modification and chromatin remodelling complexes.

2.1.1 Chromatin remodelling complexes

Mammalian chromatin remodelling complexes can be divided into three different classes based on the structure of their ATPase subunits: SWI/SNF, ISWI and Mi-2/NuRD. All these complexes use the energy of ATP hydrolysis to disrupt the nucleosome structure allowing nucleosomes to slide on the DNA. In other words, chromatin remodelling complexes make the chromatin more accessible for various proteins controlling transcription and other active biological processes (Kadam and Emerson 2002, Burgers et al. 2002, Martens and Winston 2003). For example, three members of the SWI/SNF family, DDM1, ATRX and Lsh have been demonstrated to regulate DNA methylation (Jeddeloh et al. 1999, Gibbons et al. 2000, Dennis et al. 2001).

2.1.2 Histone tail modifications

Epigenetics can be defined as heritable chromatin modifications influencing gene expression without changing the DNA sequence. In addition to chromatin remodelling, chromatin accessibility can be altered by epigenetic labelling of the DNA and core histones. Modification of the histones is focused on the histone tails protruding from the nucleosome, and as a consequence, the activity of chromatin is determined by the combined status of acetylation, phosphorylation, methylation and ubiquitination at the arginine, lysine, serine and threonine residues of the histone tails. Modified amino acids at the histone tails form the so-called histone code that is interpreted by other chromatin modifiers (Strahl and Allis 2000).

2.1.2.1 Histone acetylation and deacetylation

Histones are small basic proteins containing a high proportion of positively charged amino acids, which help to bind negatively charged DNA. The DNA binding capacity of histones can be weakened by acetylating the amino groups of lysine residues at the N-termini of the proteins. Usually, active euchromatin is characterised by acetylated histones, whereas the repressed state is achieved by histone deacetylation (Struhl 1998). Acetyl groups are found at Lys9, 14, 18 and 23 of H3 and at Lys 5, 8, 12 and 16 of H4, and lysines of H2A (K5, K9) and H2B (K5, K12, K15, K20) are also acetylated (Carrozza et al. 2003, Sims et al. 2003). Acetylation is performed by histone acetyltransferases (HATs), which can be divided into five different protein families, including e.g. Gcn5, p300/CBP and TFIID subunit TAF250. HATs are connected to big protein complexes, where other proteins determine the target of acetylation. In addition to transcriptional activation, histone acetylation complexes are connected to several other processes, such as DNA repair and regulation of the cell cycle (Carrozza et al. 2003).

The opposite process, deacetylation of the histones, is performed by histone deacetylases (HDACs). Similar to HATs, HDACs form complexes with other proteins, including targeting, co-repressor and other chromatin modification proteins. For example, HDAC1 forms a complex with the methyl-CpG-binding protein MeCP2 and the co-repressor protein Sin3a to repress transcription in a methylation dependent manner (Nan et al. 1998). There are three different categories of HDACs based on the histone deacetylase homologues found in yeast. Class I HDACs are related to yeast histone deacetylase Rpd3, including HDAC1, 2, 3, 8 and 11. Class II contains yeast deacetylase Hda1-related histone deacetylases HDAC4, 5, 6, 7, 9 and 10. Class III consists of NAD^+-dependent HDACs of the Sir2 family with seven members, Sirt1-7 (Gray and Ekström 2001, de Ruijter et al. 2003).

2.1.2.2 Histone methylation

Histones are methylated at lysine and arginine residues. Arginines are mono- or dimethylated on Arg2, 17 and 26 of H3 and on Arg3 of H4. There are at least five arginine methyltransferases, PRMT1-5, which also methylate arginines of non-histone proteins, such as STAT1 and CBP/p300 (Kouzarides 2002). Methylation of lysines is found on histone H3 Lys4, 9, 27, 36 and 79 and on H4 Lys20 (Sims et al. 2003). Lysine methyltransferases can be divided into four protein families, SET1, SET2, SUV39 and RIZ, according to their conserved SET (Suvar3-9, enhancer of zeste, trithorax) domains, which are essential for methyltransferase activity (Kouzarides 2002). Methylation of histone H3 lysine 9 (H3-K9) has been connected to gene silencing, whereas methylation of H3-K4 has been linked to transcriptional activity (Litt et al. 2001, Nielsen et al. 2001, Noma et al. 2001). Histone methyltransferases (HMTases) have crosstalk with other chromatin modifiers, like HATs and HDACs (Czermin et al. 2001, Daujat et al. 2002, Vandel and Trouche 2001, Chevillard-Briet et al. 2002). Heterochromatin proteins HP1α and HP1β have been shown to bind to methylated H3-K9 and are thought to function downstream of Suv39h HMTases (Bannister et al. 2001, Lachner et al. 2001). Furthermore, a direct connection between DNA methylation and histone methylation has recently been demonstrated, indicating that different chromatin modification complexes are interconnected (Lehnertz et al. 2003, Fuks et al. 2003b).

2.1.2.3 Histone phosphorylation and ubiquitination

Histones H1, H2A, H3 and H4 are also modified by phosphorylation (Sims et al. 2003). Phosphorylation of H3 at serine 10 plays a role in transcriptional activation and in the condensation of chromosomes during mitosis (Nowak and Corces 2004). Phosphorylation of H1 is thought to loosen the chromatin structure (Iizuka and Smith 2003). Lysines at histone H2A, H2B, H3 and H1 tails are targets of ubiquitination. The function of histone ubiquitination is mostly unknown, but it has been connected to meiosis, replication and transcription, and ubiquitination of H2B has been shown to regulate H3 lysine methylation (Kao and Osley 2003).

2.1.3 DNA methylation

The other main type of chromatin modification belonging to the epigenetic inheritance is DNA methylation. In mammalian cells DNA methylation occurs predominantly at the 5th position of cytosine within CpG dinucleotides. Transfer of a methyl group from the methyl donor S-adenosylmethionine is catalysed by DNA methyltransferases (introduced in Chapter

2.3) resulting in 5-methylcytosine (Bestor 2000). 5-methylcytosines constitute 1 % of all DNA bases, affecting about 70 % of the CpG nucleotides of the genome (Ehrlich et al. 1982). However, 5-methylcytosines are chemically unstable and prone to deamination, causing transition from 5-methylcytosine to thymine. Since these mutations are less efficiently repaired, CpG dinucleotides are less frequent than other dinucleotide pairs in the genome. Most of them are located in regions called CpG islands, GC-rich sequences connected to promoter regions of the genes. CpG islands are usually unmethylated and located at the promoter regions of housekeeping genes, leaving most of the methylated CpGs outside CpG islands (Strachan and Read 1999).

Methylated DNA is usually associated with condensed chromatin and inactive transcription. There are two basic ways in which DNA methylation is thought to repress transcription. First, CpG methylation can inhibit gene expression by recruiting methyl-CpG binding proteins and associated HDACs (Nan et al. 1998, Jones et al. 1998, Ng et al. 1999, Feng and Zhang 2001). Second, methylation can interfere with the interaction of transcription factors and other DNA binding proteins with DNA (Tate and Bird 1993). More recently, Fuks et al. (2003a) showed that methyl-CpG binding protein, MeCP2, associates with histone methyltransferase activity, providing a link between two repressive epigenetic modifications, DNA methylation and histone methylation. They also showed that DNA methyltransferases Dnmt1 and Dnmt3a interact directly with histone methyltransferase SUV39H1 and HP1β (Fuks et al. 2003b).

DNA methylation has been shown to function in a diverse range of biological processes. DNA methylation represses the expression of parasitic sequences (endogenous retroviruses and transposable elements), helping to maintain genome stability (Yoder et al. 1997, Walsh et al. 1998). DNA methylation is essential for mammalian development, as Dnmt1 and Dnmt3 deficient mice die during embryogenesis or during postnatal development (Li et al. 1992, Okano et al. 1999). DNA methylation plays an important role in genomic imprinting (Li et al. 1993, Li 2002), X chromosome inactivation (Li 2002) and immune system development and function (Teitell and Richardson 2003). Furthermore, altered methylation patterns, hypermethylation and hypomethylation, have been implicated in tumorigenesis (Feinberg and Tycko 2004).

2.1.3.1 DNA methylation during development

DNA methylation levels change dramatically during mammalian development (Figure 1). Developing germ cells are *de novo* methylated during gametogenesis. After fertilisation, a genome-wide wave of demethylation erases almost all methylation patterns established in germ cells and, finally, in postimplantation embryos global *de novo* methylation generates the methylation patterns seen in somatic cell lineages (Li 2002). The process is best studied in mouse. Germ cells develop from primordial germ cells (PGCs), which are derived from the epiblast of the pregastrulation embryo (McLaren 2003). As PGCs migrate and enter at the site of the developing gonads during E10.5-12.5, genome-wide demethylation of the germ cells occurs erasing all the inherited methylation patterns (Lee et al. 2002, Hajkova et al. 2002). Starting from E15.5 new methylation patterns of the germ cells are established with differential methylation of egg and sperm (Figure 1A) (Sanford et al. 1987, Kafri et al. 1992). This phase of methylation includes the establishment of primary methylation imprints in a process called genomic imprinting, introduced in Chapter 2.2.

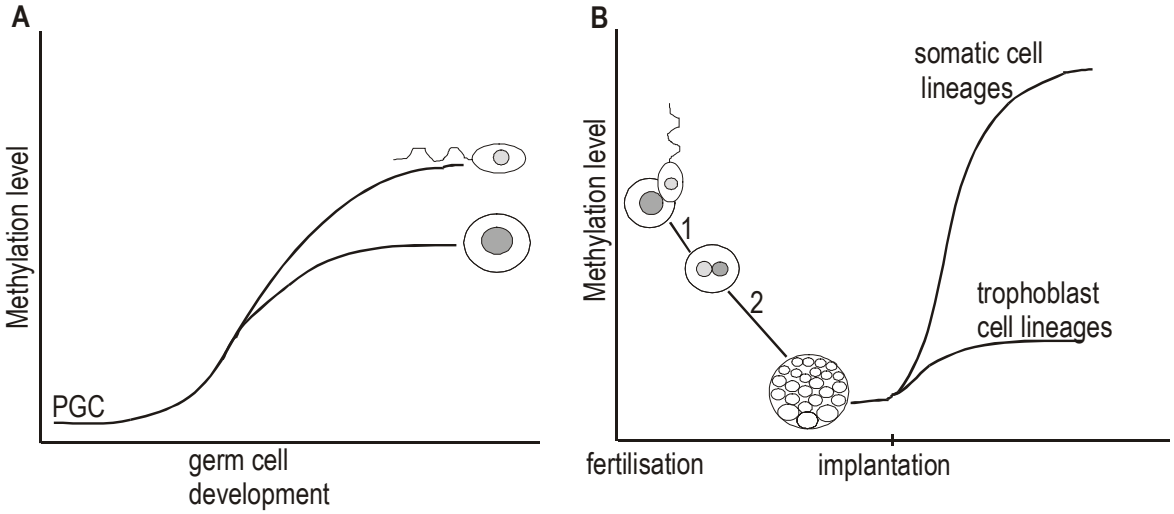

Figure 1. Reprogramming of DNA methylation patterns during development. (A) Inherited methylation patterns are erased at primordial germ cells (PGCs). During gametogenesis oocyte and sperm undergo divergent *de novo* methylations. (B) After fertilisation (1) active and (2) passive waves of demethylation drop the methylation level, with lowest levels of methylation seen at the morula stage of development. After implantation somatic cell lineages become *de novo* methylated, whereas the methylation level of trophoblast cell lineages remains lower. (Modified from Li 2002 and Dean et al. 2003)

After fertilisation both paternal and maternal genomes undergo waves of demethylation (Figure 1B). During preimplantation development most of the methylation patterns arriving with the gametes are erased, imprinted methylation patterns and some repeat sequences, however, are protected from loss of methylation (Mayer et al. 2000, Oswald et al. 2000, Reik et al. 2001, Dean et al. 2003). Just after fertilisation, the fertilised egg contains pronuclei both from sperm and egg. First, genome-wide demethylation occurs in the male pronucleus. The loss of paternal specific DNA methylation is completed within four hours after fertilisation, and is called active demethylation as it occurs before the onset of DNA replication (Mayer et al. 2000, Oswald et al. 2000, Santos et al. 2002). Several mechanisms for active demethylation have been proposed (Dean et al. 2003). According to *in vitro* studies, 5-methyl group alone could be removed from cytosine or 5-methylcytosine base itself could be removed by glycosylation. In the case of nucleotide excision repair a number of nucleotides could be removed locally (Dean et al. 2003). In addition, methylcytosine binding protein MBD2 has been identified as an active demethylase *in vitro* (Bhattacharya et al. 1999), but the *in vivo* data do not support the finding (Santos et al. 2002). After the first cell cycle and formation of the zygote, active demethylation is followed by a wave of passive demethylation in the female pronucleus, erasing remaining methylation patterns from the genome. Passive demethylation is a replication dependent event resulting from the absence of DNA methyltransferase in the nucleus (Howlett and Reik 1991, Rougier et al. 1998). The lowest level of methylation is reached at the morula stage of development (Santos et al. 2002). After implantation, lineage specific *de novo* methylation occurs, beginning from the inner cell mass of the blastocyst and leaving the trophoectoderm of the blastocyst unmethylated (Santos et al. 2002). *De novo* methyltransferases restore and the maintenance methyltransferase maintains the DNA methylation at the common levels of somatic cell lineages. The methylation level in trophoblast lineages giving rise to extra-embryonic tissues, such as placenta, remains lower (Santos et al. 2002).

2.1.3.2 DNA methylation and disease

Many human diseases have been linked to defective DNA methylation. ICF (immunodeficiency, centromeric instability, facial anomalies) syndrome (OMIM 242860) is a rare autosomal recessive disease. ICF patients show DNA hypomethylation and rearrangements of chromosomes 1, 9 and 16 concentrated at the pericentromeric areas (Ehrlich 2003). Patients also have immunoglobulin deficiency, i.e. agammaglobulinemia with B cells, and different facial anomalies like protruding tongue and flat nasal bridge. The ICF

syndrome is caused by mutations in DNA methyltransferase gene, *DNMT3B* (Okano et al. 1999, Hansen et al. 1999, Xu et al. 1999). The neurodevelopmental X-linked disorder, Rett syndrome (OMIM 312750) is caused by mutations in the *MECP2* gene, encoding for a methyl-CpG-binding protein (Amir et al. 1999). Rett syndrome is one of the most common causes of mental retardation in females, occuring with an incidence of ≥1/10000 live female births (Hagberg 1985). Just after birth, Rett syndrome patients develop normally but after 6-18 months they show loss of speech and purposeful hand use and develop microcephaly, seizures, ataxia, autism and repetitive hand movements (Hagberg et al. 1983). Defective *ATRX* (alpha-thalassemia/mental retardation syndrome, x-linked) gene is associated with X-linked ATR-X syndrome (OMIM 301040), characterised by mental retardation, urogenital and facial abnormalities and α-thalassemia (Gibbons and Higgs 2000). Mutations in *ATRX* gene cause abnormal methylation of repeated sequences (Gibbons et al. 2000). ATRX protein belongs to the SWI/SNF2 chromatin remodelling family and has been shown to localise to pericentromeric heterochromatin and interact with proteins related to chromatin modification, including HP1, EZH2 and Daxx (Le Douarin et al. 1996, Cardoso et al. 1998, McDowell et al. 1999, Xue et al. 2003). Fragile X syndrome (OMIM 309550) is the most common inherited mental retardation (1/4000 in males; Turner et al. 1996). In addition to mental retardation, features of the disease include enlarged testicles, abnormal facial features and autism. Fragile X syndrome is caused by massive expansion of CGG repeats within the *FMR1* (fragile X mental retardation 1) gene. The repeat area becomes hypermethylated resulting in epigenetic silencing of *FMR1* (Oberle et al. 1991, Kremer et al. 1991, Coffee et al. 1999). Methylation alterations associated with cancer include hypermethylation, hypomethylation and loss of imprinting (LOI) (Feinberg and Tycko 2004). Promoter hypermethylation can lead to silencing of a tumor suppressor gene and hypomethylation can result in abnormal gene activation. Normally, imprinted genes are expressed only from one allele but in the case of LOI, both maternal and paternal alleles are expressed. For example, LOI of *Igf2* leads to a double dose of growth signal, observed in Wilms tumor and several other cancer types (Ogawa et al. 1993, Feinberg et al. 2002).

2.2 GENOMIC IMPRINTING

At the beginning of the 1980's it became obvious that differential epigenetic modification of maternal and paternal gametes results in two functionally separate genomes, both of which are essential for normal mouse embryogenesis, giving a first clue about the existence of

imprinting (Surani et al. 1984, McGarth and Solter 1984). In genomic imprinting one of the autosomal gene alleles is modified in a parent-of-origin specific manner resulting in monoallelic expression of the gene. Modification of the genomes occurs during germ cell development, when maternal and paternal genomes are separated in developing egg and sperm. The imprinting mechanism involves DNA methylation of differentially methylated regions (DMRs), which are connected to many imprinted genes (Li et al. 1993). According to the web pages of the Mammalian Genetics Unit, Harwell UK, over 70 imprinted genes have been identified in mouse to date (http://www.mgu.har.mrc.ac.uk/imprinting/imprinting.html). Imprinted genes participate in many developmental processes, e.g. in the control of foetal growth, placental function and brain development, as well as in the behaviour of female mice towards their pups (Reik and Walter 2001, http://www.mgu.har.mrc.ac.uk/imprinting/function.htm). Just recently, a Japanese-Korean group was able to develop a viable parthenogenetic mouse by combining two haploid sets of maternal genome (Kono et al. 2004). This was only possible because of uniparental deletion of *H19* mimicing the imprinting.

2.2.1 Establishment of the imprints

After erasure of inherited imprinting methylation patterns in PGCs (see Chapter 2.1.3.1), female germ cells enter meiosis at E13.5 (Lee et al. 2002, Hajkova et al. 2002). Oocytes arrest at the meiotic prophase, and start the growing phase after birth. Establishment of the female imprints begins during the growth phase and is accomplished through co-operation of Dnmt3 family members (Bourc'his et al. 2001, Lucifero et al. 2002, Hata et al. 2002, Chédin et al. 2002, Suetake et al. 2004). In the case of maternally imprinted *Snrpn* gene, methylation imprints are completed by metaphase II (Lucifero et al. 2002). At E13.5, male germ cells arrest to mitosis. After birth, spermatogonia differentiate into spermatocytes starting meiosis. Based on the studies with the paternally imprinted gene *H19*, establishment of the male imprints occurs during the mitotic arrest and is completed by the pachytene stage of meiosis (Davis et al. 2000, Ueda et al. 2000, Lucifero et al. 2002). The testis-specific zinc finger protein, BORIS (Brother Of the Regulator of Imprinted Sites) is involved in epigenetic modifications in male germ cells (Loukinov et al. 2002). BORIS is a paralogue of chromatin insulator protein CTCF (see next chapter), sharing the same zinc finger structure and binding sites with CTCF. CTCF has been shown to protect DNA against methylation in the maternal *H19* allele and it is required for the establishment and maintenance of differential methylation in the *H19* DMR (Pant et al. 2003, Schoenherr et al. 2003, Fedoriw et al. 2004). During

postzygotic changes of methylation, secondary genomic imprints are established. For example, during the *de novo* methylation of postimplanted embryos, methylation in *H19* DMR extends to the promoter region (Ferguson-Smith et al. 1993, Sasaki et al. 1995). Methylated genes are usually transcriptionally silent, however, there are DMRs that are methylated on the active allele, meaning that some imprinted genes are paternally repressed although not paternally methylated (Reik and Walter 2001, Murphy and Jirtle 2003). In point of fact, almost all imprinted genes are methylated during oogenesis; there are only a few known imprinted genes that are methylated during spermatogenesis.

2.2.2 Implementation of the imprints

Once imprints are established they need to be decoded into allele-specific gene expression. Methylation can regulate imprinted trancription in several ways. A common way to regulate expression of imprinted genes is allele-specific promoter methylation (consequences discussed in Chapter 2.1.3) (Reik and Walter 2001). Most of the imprinted genes are clustered in certain chromosomal areas and are regulated through *cis*-acting imprinting control regions (ICR) (Reik and Walter 2001). Expression of the maternally imprinted and expressed *Igf2r* gene is regulated through an ICR called Region 2 (Wutz et al. 1997). Region 2 contains a promoter for the *Igf2r* antisense transcript *Air* that overlaps the *Igf2r* promoter (Lyle et al. 2000). The *Air* promoter is methylated during oogenesis allowing *Air*-antisense transcript expression from the paternal allele but not from the maternal allele. Non-coding *Air* RNA is needed for silencing of *Igf2r/Slc22a2/Slc22a3* gene cluster on the paternal allele, as lack of *Air* results in abnormal paternal expression of the genes (Zwart et al. 2001, Sleutels et al. 2002). The best-characterised imprinting region to date is the *Igf2-H19* loci on mouse chromosome 7. *Igf2-H19* ICR functions as a chromatin boundary and contains a binding site for the chromatin insulator protein CTCF. CTCF acts as an insulator when it binds to unmethylated ICR, preventing distant enhancers from activating *Igf2* in the maternal allele. In the paternal allele the CTCF binding site is methylated, thus preventing CTCF binding and resulting in activation of *Igf2* by enhancers. (Kanduri et al. 2000). The fact that the DMRs of some imprinted genes are methylated on active allele has generated an idea that these DMR sequences are bound by silencers, which cannot bind methylated DMRs. The paternally expressed *Igf2* gene has two paternally methylated DMRs. DMR1 has been shown to function as a silencer. Methylation of DMR1 inhibits the binding of repressor protein and allows trancription from the paternal allele (Constância et al. 2000, Eden et al. 2001). Methylation of Igf2 DMR2 has been shown to increase the level of transcription, thus it is thought to contain

a binding site for a methylation-sensitive activator protein (Murrell et al. 2001). Lopes et al. (2003) showed that the DMRs of *Igf2-H19* loci control each other's methylation, revealing a novel epigenetic regulation system including coordinated long-range chromatin interactions.

2.3 DNA CYTOSINE-5-METHYLTRANSFERASES

Methylated DNA is found in organisms from bacteria to mammals. DNA methylation is accomplished by two major classes of DNA methyltransferases transferring methyl group from S-adenosyl-L-methionine (AdoMet) methyl donor to a specific position in the DNA sequence. C-methyltransferases modify C-5 position of cytosines and N-methyltransferases methylate N-4 and N-6 positions of cysteines and adenines respectively (Lauster et al. 1989a). N-methyltransferases can only be found in prokaryotes, whereas C-methyltransferases have also been isolated from eukaryotes. According to the structures of two prokaryotic C-methyltransferases, M. *Hha*I and *Hae*III, the enzymes consist of two domains, a large catalytic domain and a smaller domain responsible for DNA recognition (Cheng et al. 1993, Klimasauskas et al. 1994, Reinisch et al. 1995). At the C-terminus, DNA cytosine-5-methyltransferases (DNMT) contain ten (I-X) highly conserved motifs (Posfai et al. 1989, Lauster 1989b, Kumar et al. 1994), which form the catalytic centre of the enzymes. Motifs I and X form the binding site for AdoMet and motif IX has a role in target recognition. Motif IV contains the PC dipeptide needed for the nucleophilic attack at the C-6 position of cysteine to form a covalent enzyme-DNA intermediate (Figure 2). Transient protonation of the N3 position through the ENV tripeptide in motif VI facilitates the attack and leads to activation of the C-5 position and subsequent methyl transfer from AdoMet. Finally, elimination of a proton from position 5 leads to the release of the enzyme and methylated DNA (Bestor 2000).

The first mammalian DNA methyltransferase, Dnmt1, was identified with a biochemical assay in 1988 by Bestor and coworkers. The carboxyl-terminus of the Dnmt1 showed high similarity to bacterial C-methyltransferases. Ten years later, EST database searches with the conserved C-terminal regions resulted in the isolation of three other mammalian DNA methyltransferases, Dnmt2 (Yoder and Bestor 1998; Okano et al. 1998a), Dnmt3a and Dnmt3b (Okano et al. 1998b).

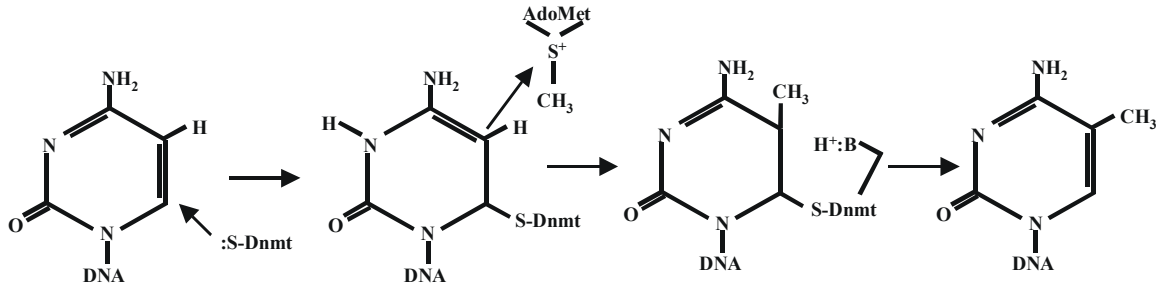

Figure 2. Catalytic mechanism of DNA cytosine-5-methyltransferases (Adapted from Bestor 2000). First, an enzyme cysteinyl thionate attacks the C-6 position and a covalent DNA-enzyme intermediate is formed. Subsequent protonation of N-3 activates the C-5 position allowing methyl transfer from S-adenosyl-L-methionine (AdoMet). Elimination of a proton from the C-5 leads to the release of the enzyme and methylated DNA

2.3.1 DNMT1

The human counterpart of the mouse *Dnmt1* gene (Bestor et al. 1988), *DNMT1*, was isolated in 1992 by Yen et al. *DNMT1* shows ubiquitous expression in somatic cells (Yen et al. 1992, Robertson et al. 1999). The large N-terminus of the DNMT1 protein consists of about 1100 aa and the catalytic C-terminus harbours about 500 aa. DNMT1 is located to the nucleus and during DNA synthesis it is attached to replication foci (Leonhardt et al. 1992). DNMT1 shows a preference for hemimethylated DNA and is therefore considered a maintenance methyltransferase, functioning during DNA replication to correctly copy the existing methylation patterns to newly synthesised DNA strands (Yoder et al. 1997, Pradhan et al. 1999). Dnmt1 deficient ES cells are viable but have reduced levels of DNA methyltransferase activity (Li et al. 1992). Dnmt1 homozygous mutant mice die before midgestation, showing biallelic expression of some imprinted genes and abnormalities in X inactivation (Li et al. 1992, Li et al. 1993, Beard et al. 1995, Panning and Jaenisch 1996).

DNMT1 has several isoforms resulting from alternative splicing of the 5' exons. Isoform DNMT1b contains extra 48 bp between exons 4 and 5 and produces an active methyltransferase (Hsu et al. 1999, Bonfils et al. 2000). In human cells, the level of DNMT1b isoform represents 2-5 % of the DNMT1 protein level (Bonfils et al. 2000). Germ cell-specific, DNMT1o and DNMT1p, isoforms are produced by sex-specific promoters (Mertineit et al. 1998). The oocyte-specific isoform, DNMT1o, is enzymatically active and lacks the first 118 amino acids of DNMT1. DNMT1o substitutes DNMT1 in growing oocytes and preimplantation embryos (Mertineit et al. 1998, Hayward et al. 2003). DNMT1o is a cytoplasmic protein and relocates to the nucleus only at the 8-cell stage of development, where it maintains the DNA methylation patterns of imprinted genes (Howell et al. 2001,

Ratnam et al. 2002). In spermatocytes, pachytene-specific DNMT1p generates several small transcripts, which are thought to prevent translation of DNMT1 mRNA (Mertineit et al. 1998). Although referred to as pachytene-specific isoform, translated DNMT1p has been found in muscle cells, where it may participate in the regulation of myogenesis (Aguirre-Arteta et al. 2000, Robertson 2002).

2.3.2 DNMT2

The *DNMT2/Dnmt2* gene was identifed in 1998 by two independent groups (Yoder and Bestor 1998; Okano et al. 1998a). At low levels *Dnmt2* is expressed in most human and mouse tissues. The Dnmt2 protein of 415 amino acids does not contain the large N-terminal region used for targeting and regulation of Dnmt1 and Dnmt3 proteins, but it contains all conserved domains needed for methyltranferase activity. The structure of Dnmt2 strongly resembles that of prokaryotic C-methyltransferase M.*Hha*I and the protein can bind covalently to DNA (Cheng et al. 1993, Klimasauskas et al. 1994, Dong et al. 2001). In addition, Dnmt2 is the most conserved eukaryotic DNA methyltransferase protein having related proteins in mammals, *Arabidobsis thaliana, Danio rerio, Xenopus laevis* and also in *Schizosaccaromyches pombe* and *Drosophila melanogaster*, which do not have any other Dnmt candidate genes. However, when overexpressed in the baculovirus system, Dnmt2 did not show any methyltransferase activity *in vitro* (Yoder and Bestor 1998; Okano et al. 1998a), and it was suggested that Dnmt2 is not a functional methyltransferase. Recently, several studies have indicated that Dnmt2 proteins in human, mouse and *D. melanogaster* are active *in vitro* and *in vivo*, but the transmethylase activity is low when compared with other Dnmt proteins (Hermann et al. 2003, Liu et al. 2003, Tang et al. 2003, Kunert et al. 2003, Mund et al. 2004). Structural studies have suggested that Dnmt2 has a unique target recognition domain (Dong et al. 2001) and at least in fruit fly Dnmt2 methylates only non-CpG dinucleotides (Kunert et al. 2003, Mund et al. 2004). However, the role of Dnmt2 in methylation remains a mystery. Dnmt2 deficient ES cells are viable and do not show any serious defects in DNA methylation patterns (Okano et al. 1998a). In *D. melanogaster* Dnmt2 is responsible for DNA methylation but has no detectable effect on embryonic development (Kunert et al. 2003).

2.3.3 DNMT3A and DNMT3B

Human and mouse DNMT3/Dnmt3 proteins are conserved in evolution, showing significant, 94-98 %, identity with each other (Okano et al. 1998b, Xie et al. 1999). The *DNMT3A* gene

encodes a protein of 912 amino acids and the DNMT3B protein consists of 853 amino acids. Proteins have similar catalytic domains but show only weak sequence identity in their N-terminal regulatory domains. At the N-terminus both proteins have a PWWP-domain that has DNA binding capacity and has been shown to be essential for chromatin targeting of the proteins (Qiu et al. 2002, Ge et al. 2004). DNMT3 proteins also contain a cysteine-rich ATRX-like domain that is used for interactions with other proteins, and conserved methyltransferase domains at the C-termini. DNMT3s are enzymatically active, both *in vitro* and *in vivo*, functioning mainly as *de novo* methyltransferases establishing new methylation patterns in early embryos (Okano et al. 1998b, Hsieh 1999, Lyko et al. 1999, Gowher and Jeltsch 2001, Aoki et al. 2001, Yokochi and Robertson 2002, Mund et al. 2004). However, Dnmt3a and 3b are not only needed to establish, but also to maintain genomic methylation patterns during development (Chen et al. 2003). Dnmt3 proteins have a methylation preference for CpG sites, but they have also been shown to methylate non-CpG sites, mainly CpA and CpT (Ramsahoye et al. 2000, Aoki et al. 2001, Gowher and Jeltsch 2001, Yokochi and Robertson 2002). DNMT3/Dnmt3 transcripts are detected in most foetal tissues and in lower level in adult somatic tissues. The strongest expression is seen in early embryos and undifferentiated ES cells (Okano et al. 1998b, Xie et al. 1999, Okano et al. 1999, Robertson et al. 1999, Watanabe et al. 2002). Subcellularly, Dnmt3 proteins localise in the nucleus, specifically to heterochromatin regions (Bachman et al. 2001, Margot et al. 2001, Hata et al. 2002). Both Dnmt3a and Dnmt3b are essential for development. Dnmt3a deficient (Dnmt3a -/-) mice appear normal at birth but die at four weeks after birth. Dnmt3b -/- mice have multiple developmental defects resulting in death before birth (Okano et al. 1999). ES cell lines homozygous for either Dnmt3a or Dnmt3b mutations show *de novo* methylation activity, whereas in double mutant (Dnmt3a -/-, Dnmt3b -/-) ES cells *de novo* methylation is not detected. In ES cells Dnmt3a and Dnmt3b display overlapping functions, but they have independent targets, at least partly, as demonstrated by different defects in knockout mice (Okano et al. 1999).

2.3.3.1 Isoforms of DNMT3A/Dnmt3a and DNMT3B/Dnmt3b

Several isoforms for human and mouse DNMT3s have been identified, and there are several reports describing alternate 5' UTRs for the genes. Human and mouse DNMT3A proteins have two major isoforms, DNMT3A and DNMT3A2, transcribed by separate promoters (Figure 3) (Chen et al. 2002). Both isoforms show strong transmethylation activity. Dnmt3a2 lacks about 220 amino acids from the N-terminus of Dnmt3a and, unlike Dnmt3a, is localised

to euchromatin. Dnmt3a2 is a major isoform in ES cells, whereas Dnmt3a is expressed at low levels ubiquitously (Chen et al. 2002). In addition, there are several alternate exon 1s (1A, 1B, 1C, 1α and 1β) for DNMT3A/Dnmt3a (Yanagisawa et al. 2002, Weisenberger et al. 2002). Four different alternatively spliced transcripts for human and mouse DNMT3A result from combinations of exons 1α, 1β and intron 4. Mouse ES cells preferentially express Dnmt3a with exon 1β, whereas in mouse and human somatic cells the expression levels of DNMT3a with exon 1α are slightly higher (Weisenberger et al. 2002). Exons 1A, 1B and 1C create four transcriptional start sites for DNMT3A, expressed from three different promoters (Yanagisawa et al. 2002).

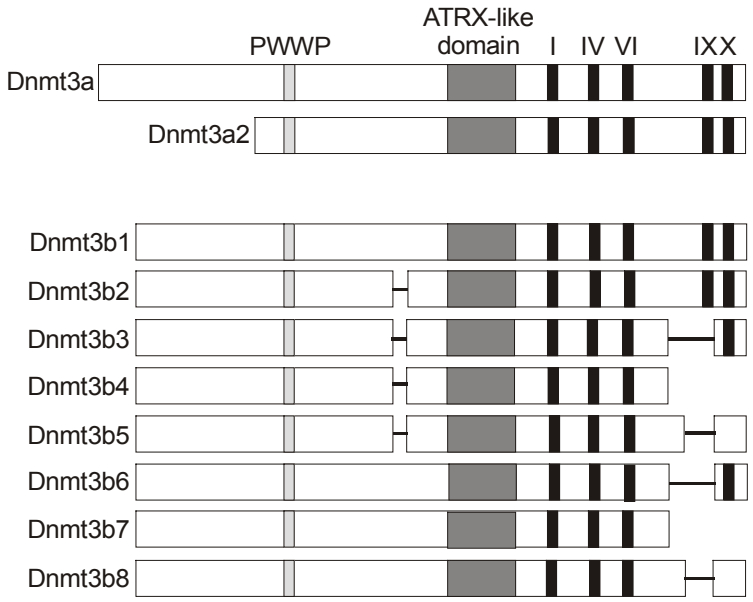

Figure 3. Schematic illustration of mouse Dnmt3a and Dnmt3b protein isoforms. The specific protein domains of each isoform are indicated as boxes and bars. The light grey box represents the PWWP DNA-binding domain. The dark grey box represents the ATRX-like zinc finger domain and the black bars represent the conserved DNA cytosine methyltransferase motifs I, IV, VI, IX and X. 45 amino acids at the C-termini of Dnmt3b5 and Dnmt3b8 are out of frame and do not contain motif X. In addition, there are several alternate first exons and and alternatively spliced 5' UTRs for both Dnmt3 proteins.

Eight different mouse Dnmt3b isoforms (Figure 3) and five different human DNMT3B isoforms have been identified (Okano et al. 1998b, Xie et al. 1999, Robertson et al. 1999, Hansen et al. 1999, Chen et al. 2002, Weisenberger et al. 2004). All these transcripts result from alternative splicing of exons 10, 11, 21 and 22. Both human and mouse somatic cells contain splice variants DNMT3b2-5, without exons 10 and 11 (Weisenberger et al.

2004). Mouse ES cells contain splice variants Dnmt3b1, 6, 7 and 8, which all include exons 10 and 11, suggesting that this region has a specific role during development (Chen et al. 2002, Weisenberger et al. 2004). Similar to DNMT3A, DNMT3B also has alternate first exons (Xu et al. 1999, Yanagisawa et al. 2002), and several transcription initiation sites and alternatively spliced 5' UTRs for mouse Dnmt3b, have been described (Ishida et al. 2003, Weisenberger et al. 2004). Isoforms Dnmt3b1 and Dnmt3b2 are active both *in vitro* and *in vivo* (Okano et al. 1998b, Aoki et al. 2001, Mund et al. 2004). Isoforms Dnmt3b3-8 all lack at least part of the domain IX, or domains IX and X completely. These isoforms have no methyltransferase activity *in vitro* or *in vivo* (Okano et al. 1998b, Aoki et al. 2001, Chen et al. 2003, Mund et al. 2004) and are therefore considered to be inactive. However, results by Weisenberger et al. (2004) suggest that DNMT3B3, Dnmt3b3 and Dnmt3b6 may also bind to DNA and be involved in DNA methylation. Overexpression of DNMT3B4 has been shown to cause hypomethylation of pericentromeric satellite regions in human hepatocarcinogenesis (Saito et al. 2002).

2.3.4 Co-operativity of DNMT proteins

Several studies have indicated that DNA methyltransferases and their isoforms show different targets and activities. In a study using plasmids as targets for methylation, Hsieh (1999) showed that DNMT1, Dnmt3a and Dnmt3b methylate different episomal target sequences. In this system, the activity of Dnmt3b was lower when compared with Dnmt3a. Liu et al. (2003) used a novel ICM (*in vivo* complex of methylase) -technique based on irreversible interaction of DNMT with 5-aza-substituted cytosine in DNA. By using ICM together with DNMT-specific antibodies, they showed that DNMT1, DNMT2 and DNMT3 proteins have different activities and different DNA binding capacities. Different targets and activities of DNMTs were further characterised in studies using a fruit fly system with different combinations of Dnmts (Mund et al. 2004) and mouse ES cells lacking either one or several Dnmt3 isoforms (Chen et al. 2003). In addition to different targets and activities, differential expression patterns of DNMTs and their isoforms suggest that genomic methylation patterns are mediated by co-operation of different DNMT isoforms (Xie et al. 1999, Robertson et al. 1999, Hsieh 1999, Robertson et al. 2000a, Chen et al. 2002, Kim et al. 2002, Chen et al. 2003, Huntriss et al. 2004). In *Drosophila melanogaster* co-expression of Dnmt1 and Dnmt3a leads to the co-operative establishment and maintenance of genomic methylation patterns (Lyko et al. 1999). Co-expression of DNMT1 and DNMT3A or DNMT3B results in methylation spreading in the genome *in vivo* (Kim et al. 2002). Fatemi et al. (2002) showed that Dnmt3a

stimulates Dnmt1 to *de novo* methylation. In human cancer cells co-operation of DNMT1 and DNMT3b is needed to maintain DNA methylation and gene silencing (Rhee et al. 2002).

2.3.5 Interactions between DNMT proteins

Human DNMT3A (aa 1-298) and DNMT3B (aa 199-403) proteins interact with each other via their N-termini (Kim et al. 2002). Using these same N-terminal regions both proteins can also interact with DNMT1. DNMT3A (aa 1-298) and DNMT3B (199-403) can bind DNMT1 simultaneously, since they use different regions of DNMT1 as binding sites (3A: aa 1-148, 3B: aa 149-217). In the complex of DNMT3A and DNMT3B the binding sites of DNMT1 are already engaged. Therefore, DNMT1 first interacts with either DNMT3A or DNMT3B, enabling association of all three methyltransferases (Kim et al. 2002). When expressed together, DNMT1 and DNMT3B are nuclear and DNMT3A shuttles between nucleus and cytoplasm (Kim et al. 2002). Using yeast two-hybrid technique, Margot et al. (2003) demonstrated that the N- and C-terminal domains of Dnmt1 interact with each other, whereas Dnmt3 proteins do not show such an intra-molecular interaction. Unlike the catalytic domains of Dnmt3a and Dnmt3b, the C-terminus of Dnmt1 is not active on its own and needs a large portion of the N-terminus for enzymatic activity (Zimmermann et al. 1997, Margot et al. 2000, Fatemi et al. 2001, Gowher and Jeltsch 2002). These results suggest that the activity of the Dnmt1 C-terminus depends on the folding of the protein, since almost the whole N-terminus is needed for the interaction (Margot et al. 2003).

2.3.6 Interactions of DNMTs with other proteins

All DNMT-associated proteins are somehow involved in transcriptional regulation and chromatin modification (Table 1). Interaction of DNMT1 and proliferating cell nuclear antigen (PCNA) takes place at replication foci, facilitating the methylation of newly replicated DNA by DNMT1 (Chuang et al. 1997, Iida et al. 2002). DNMT1 associated protein, DMAP1, is also attached to DNMT1 in S phase, and during late S phase when heterochromatin regions replicate HDAC2 is recruited to the DMAP1/DNMT1 complex (Rountree et al. 2000). DNMT1 associates with methyl-CpG-binding proteins MBD2/MBD3 and MeCP2, which usually are part of bigger repression complexes containing different corepressors and HDACs (Tatematsu et al. 2000, Kimura and Shiota 2003). Methyl-CpG-binding proteins MBD2 and MBD3 form a complex, which is attached to hemimethylated DNA. At late S phase, this complex is connected to DNMT1 with all three proteins showing co-localisation at the replication foci (Tatematsu et al. 2000). Usually, MeCP2 forms a

complex with mSin3a and HDAC1, but the MeCP2/Dnmt1-complex does not contain HDAC1 (Kimura and Shiota 2003). Similar to MBD2/MBD3/Dnmt1-complex, MeCP2/Dnmt1-complex shows DNA methyltransferase activity to hemimethylated DNA, suggesting that Dnmt1 interacts with MeCP2 in order to perform maintenance methylation (Kimura and Shiota 2003). Interestingly, DNMT1 has also been shown to interact with the hyperphosphorylated C-terminal repeat domain of RNA polymerase II (Carty and Greenleaf 2002).

Table 1. Proteins associating with DNMTs.

DNMT	Interacting protein	Function	Reference
DNMT1	PCNA	auxiliary DNA replication factor	Chuang et al. 1997
Dnmt1	HDAC1	histone deacetylation	Fuks et al. 2000
DNMT1	HDAC2	histone deacetylation	Rountree et al. 2000
DNMT1	DMAP1	co-repressor	Rountree et al. 2000
DNMT1	Rb	tumor suppressor	Robertson et al. 2000b, Pradhan and Kim 2002
DNMT1	MBD2/MBD3	bind methylated DNA	Tatematsu et al. 2000
DNMT1	RNA polymerase II	transcription	Carty and Greenleaf 2002
Dnmt1	PML-RAR	oncogenic transcription factor	Di Croce et al. 2002
DNMT1	DNMT3a	*de novo* DNA methylation	Kim et al. 2002
DNMT1	DNMT3b	*de novo* DNA methylation	Kim et al. 2002
Dnmt1	SUV39H1	histone H3-K9 methylation	Fuks et al. 2003b
Dnmt1	HP1β	heterochromatin protein	Fuks et al. 2003b
Dnmt1	Dnmt1	maintenance DNA methylation	Margot et al. 2003
Dnmt1	p53	tumor suppressor	Peterson et al. 2003
Dnmt1	MeCP2	bind methylated DNA	Kimura and Shiota 2003
Dnmt3a	RP58	co-repressor	Fuks et al. 2001
Dnmt3a	HDAC1	histone deacetylation	Fuks et al. 2001 Bachman et al. 2001
DNMT3a	DNMT1	maintenance DNA methylation	Kim et al. 2002
DNMT3a	DNMT3b	*de novo* DNA methylation	Kim et al. 2002
Dnmt3a	PML-RAR	oncogenic transcription factor	Di Croce et al. 2002
Dnmt3a	SUV39H1	histone H3-K9 methylation	Fuks et al. 2003b
Dnmt3a	HP1β	heterochromatin protein	Fuks et al. 2003b
Dnmt3a	DNMT3L	imprinting regulation	Hata et al. 2002, Margot et al. 2003, Suetake et al. 2004
Dnmt3a	Ubc9/PIAS1/PIASxα	transcriptional regulation	Ling et al. 2004
Dnmt3b	Ubc9, SUMO-1	transcriptional regulation	Kang et al. 2001
Dnmt3b	HDAC1	histone deacetylation	Bachman et al. 2001
DNMT3b	DNMT1	maintenance methylation	Kim et al. 2002
DNMT3b	DNMT3a	*de novo* DNA methylation	Kim et al. 2002
Dnmt3b	DNMT3L	imprinting regulation	Hata et al. 2002, Margot et al. 2003, Suetake et al. 2004
Dnmt3b	HP1α/ HP1β	heterochromatin protein	Lehnertz et al. 2003

DNMT1 forms a complex with Rb, E2F1 and HDAC1 enabling transcriptional repression of E2F-responsive promoters (Robertson et al. 2000b). Interactions of tumor suppressors Rb and p53 with DNMT1 have been shown to modify DNMT1 activity (Prahdan and Kim 2002, Peterson et al. 2003). p53 represses Dnmt1 transcription (Peterson et al. 2003) and Rb-binding prevents the formation of DNMT1-DNA complex (Pradhan and Kim 2002). Furthermore, overexpression of Rb results in hypomethylation of DNA. Interaction with HDAC is a common feature of all DNMTs and is usually a methyltransferase activity-independent process (Fuks et al. 2000, Rountree et al. 2000, Bachmann et al. 2001, Fuks et al. 2001). Dnmt3a and Dnmt3b interact with HDAC1 through ATRX-like domains (Bachman et al. 2001, Fuks et al. 2001). Repression by the Dnmt3a can be targeted through association with sequence-specific co-repressor RP58. In GST pull-down assay Dnmt3b also interacts with RP58 (Fuks et al. 2001). Repression by DNMT1-HDAC1 complex is only partly relieved by the deacetylase inhibitor, TSA, indicating that DNMT1 has both deacetylase-dependent and independent repression mechanisms (Fuks et al. 2000, Robertson et al. 2000b).

Lehnertz et al. (2003) showed that DNMT3b/Dnmt3b interacts with heterochromatin protein HP1α. They also showed that Dnmt1, Dnmt3a and Dnmt3b all associated with H3-K9 HMTase activity, demonstrating a direct connection between DNA methylation and histone methylation. In addition, Fuks et al. (2003b) indicated that Dnmt1 and Dnmt3a interact directly with both HP1β and HMTase SUV39H1. Interactions between leukaemia promoting PML-RAR fusion protein and Dnmt1 or Dnmt3a induce hypermethylation and silencing of RARβ2 gene, promoting its carcinogenic potential (Di Croce et al. 2002). Dnmt3a and Dnmt3b are both sumoylated and thus interact with several components of the sumoylation machinery (Kang et al. 2001, Ling et al. 2004). Sumoylation influences protein interactions (Verger et al. 2003) and has been shown to disturb binding of Dnmt3a to HDAC1. However, sumoylated Dnmt3a can still bind to Dnmt3b (Ling et al. 2004).

2.4 MOUSE SPERMATOGENESIS

The maturation of the male germ cells occurs in the seminiferous tubules, which consist of germ cells at different developmental stages and somatic Sertoli cells that supply structural and nutritional support for the germ cells (Griswold 1998, Cooke and Saunders 2002). In mice, differentiation from diploid spermatogonia to haploid spermatozoa takes 35 days (Oakberg 1956). Spermatogenesis can be divided into three sequential phases: (a) proliferation of spermatogonia, (b) meiosis of primary and secondary spermatocytes, and (c)

spermiogenesis, where spermatids are transformed into spermatozoa (de Rooij and Grootegoed 1998). Spermatogonia are subdivided into spermatogonial stem cells and differentiated type A and B spermatogonia. During the cell divisions type A spermatogonia remain connected by cytoplasmic bridges, which persist throughout spermatogenesis until the late spermatid stages. After several mitotic divisions type A spermatogonia mature into type B spermatogonia (de Rooij and Grootegoed 1998). Primary spermatocytes arise from type B spermatogonia after the last mitotic division (Figure 4).

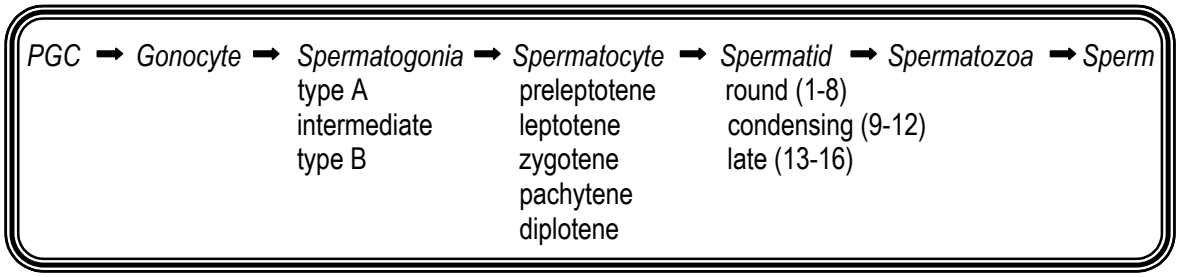

Figure 4. Schematic overview of the development of the male germ cells. Germ cells are derived from the primordial germ cells (PGCs). In males, PGCs are enclosed by precursor Sertoli cells inside seminiferous cords, where they differentiate into gonocytes and begin a developmental arrest, until the start of spermatogenesis after birth. Spermatogonia differentiate into spermatocytes through mitotic divisions. Two meiotic divisions of spermatocytes give rise to haploid spermatids. The development of the spermatids into spermatozoa occurs through 16 steps.

During the meiotic phase of the spermatogenesis, the DNA content is reduced from the diploid to the haploid state through two cell divisions; in the first division primary spermatocytes are transformed into secondary spermatocytes, which after a second meiotic division form the spermatids. The first meiotic division contains a long prophase, which is divided into leptotene, zygotene, pachytene, diplotene and diakinesis stages based on the behaviour of the chromosomes (Page and Hawley 2003). The last DNA replication of spermatogenesis occurs in preleptotene spermatocytes. In leptotene spermatocytes, replicated chromosomes start to condense and find their homologues. In the synapsis of zygotene spermatocytes, chromosomes begin to pair and form the tripartite structure of synaptonemal complex, which is completed at pachytene spermatocytes. During the pachytene stage crossing-over occurs. In diplotene spermatocytes chromosome homologues begin to separate but are held together by chiasmata. During diakinesis chromosome bivalents become more contracted and get ready to align at the metaphase plate (Cook 1997, Scherthan et al. 1998,

Page and Hawley 2003). Meiosis I is continued with breakdown of the nuclear envelope, formation of the meiotic spindle and separation of homologous chromosomes. In the subsequent meiotic division, diploid secondary spermatocytes differentiate to haploid spermatids. Spermatids differentiate further through spermiogenesis, during which the cell is remodelled, e.g. acrosome and tail develop, cytoplasm is eliminated and DNA is compacted (Page and Hawley 2003). In post-meiotic germ cells histones are replaced by transition proteins, which during the spermiogenesis are replaced by protamines (Sassone-Corsi 2002). In spermiation, mature spermatozoa are released into the lumen of the seminiferous tubule. The final maturation of the sperm occurs in the epididymis. In cross sections of seminiferous tubules, germ cells from different developmental phases can be found one upon the other with spermatogonia at the basal lamina and spermatids at the lumen of the tubule.

Spermatogenesis is a cyclic process, which in the mouse can be divided into 12 stages (I-XII), where each stage contains certain combinations of developing cell types (Table 2). The duration of the seminiferous epithelial cycle in mice is 8.5 days (Oakberg 1956). Staggered timing of the cycle in adjacent areas of the epithelium ensures the continuous production of sperm and is called a spermatogenic wave.

Table 2. Diagram representing cellular associations during the 12 stages (I-XII) of the mouse seminiferous epithelial cycle (modified from de Rooij 1998).

I	II	III	IV	V	VI	VII	VIII	IX	X	XI	XII
ES 13	ES 14	ES 14	ES 15	ES 15	ES 15	ES 16	MS				
RS 1	RS 2	RS 3	RS 4	RS 5	RS 6	RS 7	ES 8	ES 9	ES 10	ES 11	ES 12
P	P	P	P	P	P	P	P	P	P	D	MD
SG	SG	SG	SG	SG	PL	PL	PL, L	L	Z	Z	P
SG	SG	SG	SG	SG	SG	SG	SG	SG	SG	SG	SG

SG=spermatogonia, PL=preleptotene, L=leptotene, Z=zygotene, P=pachytene, D=diakinetic, MD=meiotic division, RS=round spermatid, ES=elongating spermatid, MS=mature spermatozoa.

3. AIMS OF THE STUDY

1) To isolate and characterise a novel human gene located on chromosome 21q22.3.

2) To identify the corresponding *Dnmt3L* gene in mouse in order to perform functional studies.

3) To examine the expression pattern of *DNMT3L/Dnmt3L* in human and mouse.

4) To isolate and characterise the *Dnmt3L* promoter and identify the factors affecting *Dnmt3L* expression.

5) To study whether DNMT3L functions as a transcriptional repressor like other members of the DNMT3 gene family.

4. MATERIALS AND METHODS

This chapter briefly summarises the main methods used in the studies. Detailed reaction conditions can be found in the original communications.

4.1 IN SILICO ANALYSIS (Studies I, II, III)

Computer programs Genscan (Burge and Karlin 1997) and NIX were used for gene and exon predictions during *DNMT3L* cDNA isolation. NIX is a tool for running and viewing results from many prediction programs simultaneously, including GRAIL, Fex, Hexon, MZEF, Genemark, Genefinder, FGene, BLAST, Polyah, RepeatMasker, and tRNAscan. Different BLAST programs (Altschul et al. 1990, Altschul et al. 1997) were used for nucleotide and protein similarity searches. Protein patterns of Dnmt3L were searched with ScanProsite (Gattiker et al. 2002). PeptideMass calculated predicted molecular mass for DNMT3L/Dnmt3L protein (Wilkins et al. 1997) and subcellular localisation of protein was predicted by PSORTII (Horton and Nakai 1997). Multiple sequence alignments were performed with ClustalW and printed with BOXSHADE program. The predicted three-dimensional model of the zinc finger of DNMT3L protein was built using the programs InsightII and Discover (Accelrys, San Diego, CA). The model was evaluated with program PROCHECK (Laskowski et al. 1993, Morris et al. 1992). MatInspector (Quandt et al. 1995) was used for searching transcription factor binding sites in the *Dnmt3L* promoter sequence. Table 3 shows the www-addresses of the programs used in the studies.

Table 3. Www-addresses of different analysis and prediction programs used in the studies.

program	www-address
Genscan	http://genes.mit.edu/GENSCAN.html
NIX	http://www.hgmp.mrc.ac.uk/Registered/Webapp/nix/
BLAST	http://www.ncbi.nlm.nih.gov/BLAST/
ScanProsite	http://us.expasy.org/tools/scanprosite/
PeptideMass	http://us.expasy.org/tools/peptide-mass.html
PSORTII	http://www.psort.org/
ClustalW	http://www.ch.embnet.org/software/ClustalW.html
BOXSHADE	http://www.ch.embnet.org/software/BOX_form.html
MatInspector	http://www.genomatix.de/

4.2 HUMAN AND MOUSE DNMT3L/Dnmt3L cDNA ISOLATIONS (Studies I, II)

The genomic sequencing of the human chromosome 21q22.3 region *MX1-D21S171* was performed at the Laboratory of Genomic Medicine, Keio University School of Medicine, Tokyo. The exon trapping experiments from HC21 specific cosmids have been previously described (Chen et al. 1996, Kudoh et al. 1997). The computer software programs Genscan and NIX were employed for gene modelling and exon prediction of the genomic sequence. *DNMT3L* cDNA fragments were cloned by RT-PCR and RACE techniques from testis, foetal liver and placenta mRNAs and thymus cDNA library (Clontech) using primers designed according to predicted and trapped exon sequences. Primers for the mouse *Dnmt3L* cDNA cloning were designed according to the mouse EST sequences identified with the BLAST program using human *DNMT3L* cDNA as a query sequence. 5'- and 3'- *Dnmt3L* cDNA fragments were amplified by RT-PCR using mRNAs extracted from testis and thymus of a Balb/C mouse. RNAs were extracted using RNAgents Total RNA Isolation System kit (Promega) and converted to cDNA with the First-Strand cDNA Synthesis Kit (Amersham Biosciences). The *DNMT3L/Dnmt3L* cDNA fragments were gel-purified, subcloned into either a pBluescriptII SK+ (Stratagene) or a pCRII-TOPO vector (Invitrogen), and sequenced by the dye deoxy terminator cycle-sequencing method with AmpliTaq/FS DNA polymerase (Perkin Elmer) and then analysed by an automatic DNA sequencer (Applied Biosystems, 310 or 377). The cDNA sequences were compared against public databases with the BLAST program (Altschul et al. 1990).

4.3 EXPRESSION ANALYSIS (Studies I, II)

Mouse RNAs used in the expression analysis were isolated from frozen Balb/C mouse tissues, including newborn mouse muscle, kidney, liver, lung and thymus, and adult mouse thymus, testis and ovary. Total RNA isolations were performed using the RNAgents Total RNA Isolation System kit (Promega), after which mRNAs were converted to cDNAs with the First-Strand cDNA Synthesis Kit (Amersham Biosciences).

4.3.1 Northern blot analysis

For *DNMT3L* expression analysis Northern blots (Human MTN blot1-4 and Human Fetal MTN Blot) were purchased from Clontech and hybridised according to manufacturer's instructions. In the mouse *Dnmt3L* expression analysis, twenty micrograms of total RNA extracted from different Balb/C mouse tissues were electrophoresed in a denaturing formaldehyde gel, blotted and hybridised using standard protocols (Sambrook et al. 1989).

Human and mouse cDNA fragments used as probes were labelled with [α-^{32}P]dCTP either by PCR or by using Prime-a-Gene Labeling System (Promega), and purified with MicroSpin G-50 Columns (Amersham Biosciences). Filters were exposed to a Phosphor Screen (Molecular Dynamics) or to an Image Analyser FLA-3000's image plate (Fuji Film) for 3-4 days.

4.3.2 RT-PCR and Southern blot analysis

Expression patterns of the human and mouse *DNMT3L*/*Dnmt3L* genes were also studied by RT-PCR, using Human Multiple Tissue cDNA (MTC) panels (I, II, foetal and immune system panels; Clontech), Mouse MTC Panel I (Clontech) and cDNAs extracted from different newborn and adult Balb/C mouse tissues as templates. As a control, RT-PCR was performed with the glyceraldehyde-3-phosphate dehydrogenase (G3PDH) primers. Subsequently, nested PCRs or Southern blot analysis were performed in mouse *Dnmt3L* expression studies. In the Southern blot analysis the *Dnmt3L* first-round PCR products were characterised by agarose gel electrophoresis, the DNA was transferred to a nylon membrane and the filter was hybridized with the *Dnmt3L*-specific cDNA fragment labelled with [α-^{32}P]dCTP as described above (Section 4.3.1). The filter was exposed to a Phosphor Screen (Molecular Dynamics) overnight.

4.4 CELL CULTURE (Studies II, III, IV)

Mouse NIH3T3 fibroblasts, human rhabdomyosarcoma (RD) cells (American Type Culture Collection), Cos-7 and Cos-1 cells were cultured in Dulbecco's Modified Eagle's Medium supplemented with 10 % foetal calf serum and antibiotics. THP-1 monocytes were grown in RPMI 1640 medium and mouse thymic epithelial cells Tec427.1 (Vukmanovic et al. 1992) were cultured in Minimum Essential Medium Eagle, both supplemented with 10 % foetal calf serum, 2 mM L-glutamine and antibiotics. All media and supplements were obtained from Bio Whittaker Europe.

4.5 DNA CONSTRUCTS (Studies I, II, III, IV)

All DNMT3/Dnmt3 plasmids generated for the studies are shown in Table 4. The mammalian expression vector pcDNA3.1 with myc-his epitope tag was purchased from Invitrogen. The expression vector pSI was purchased from Promega. In pM vector *DNMT3L* cDNA fragments were fused with Gal4 DNA binding domain and in the pEGFP-C3 vector cDNA fragments were cloned downstream of the green fluorescent protein. Both, pM and pEGFP-C3,

Table 4. DNA constructs cloned for the studies

plasmid	vector	cloning	cloning sites	assay	study
pcDnmt3L	pcDNA3.1	PCR	*Eco*RI/*Hind*III	immunofluorescence	II
pLuc1345	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pLuc900	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pLuc440	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pLuc226	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pLucmut262	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pLucmut290	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pLucmut300	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pLucmut318	pLTR-Luc	PCR	*Not*I/*BamH*I	luciferase assay	III
pM-DNMT3L 1-84	pM	PCR	*Eco*RI/*Sal*I	CAT assay	IV
pM-DNMT3L 45-84	pM	PCR	*Eco*RI/*Sal*I	CAT assay	IV
pM-DNMT3L 88-153	pM	PCR	*Eco*RI/*Sal*I	CAT assay	IV
pM-DNMT3L 1-195	pM	digestion/subcloning	*Eco*RI/*Hind*III	CAT assay	IV
pM-DNMT3L	pM	digestion/subcloning	*Eco*RI/*Hind*III	CAT assay	IV
pM-DNMT3L 164-387	pM	digestion/subcloning	*Eco*RI/*Hind*III	CAT assay	IV
pM-DNMT3L 88-387	pM	PCR	*Eco*RI/*Hind*III	CAT assay	IV
pcDNMT3L 1-195	pcDNA3.1	PCR	*Eco*RI/*Hind*III	*in vitro* translation	IV
pcDNMT3L	pcDNA3.1	PCR	*Eco*RI/*Hind*III	immunofluorescence, WB	I
pGST-DNMT3L 1-84	pGEX-1λT	digestion/subcloning	*Eco*RI/*Sal*I,*Xho*I	GST pull-down, deacetylation	IV
pGST-DNMT3L 45-84	pGEX-1λT	digestion/subcloning	*Eco*RI/*Sal*I,*Xho*I	GST pull-down, deacetylation	IV
pGST-DNMT3L 88-153	pGEX-1λT	digestion/subcloning	*Eco*RI/*Sal*I,*Xho*I	GST pull-down, deacetylation	IV
pGST-DNMT3L 1-195	pGEX-1λT	digestion/subcloning	*Eco*RI/*Hind*III	GST pull-down, deacetylation	IV
pGST-DNMT3L 88-387	pGEX-1λT	digestion/subcloning	*Eco*RI/*Hind*III	GST pull-down, deacetylation	IV
pGST-DNMT3L	pGEX-1λT	digestion/subcloning	*Eco*RI/*Hind*III	GST pull-down, deacetylation	IV
pGST-DNMT3L 164-387	pGEX-1λT	PCR	*Eco*RI/*Xho*I	GST pull-down, deacetylaton	IV
pGFP-NLS (142-173)	pEGFP-C3	PCR	*Eco*RI/*Sal*I	immunofluorescence	IV
pGFP-DNMT3L	pEGFP-C3	PCR	*Eco*RI/*Sal*I	immunofluorescence	IV
pSI-DNMT3B 220-753	pSI	PCR	*Eco*RI/*Not*I	*in vitro* translation	

vectors were purchased from Clontech. The pGEX-1λT vector with modified multiple cloning site and pLTR-Luc, pLacZ and pBL-KS vectors used in luciferase assays were donated by Dr. K. Saksela (University of Tampere, IMT, Finland). The pING14AHDAC1 and GST-HDAC1 vectors were provided by Drs G. Akusjärvi and T. Punga (University of Uppsala, Sweden). pcHDAC1 was a generous gift from Dr. T. Kouzarides (University of Cambridge, UK) and Gal4TKCAT reporter plasmid was provided by Dr. Y. Shi (Harvard Medical School, USA).

4.6 WESTERN BLOTTING AND IMMUNOFLUORESCENCE (Studies I, II, IV)

The full-length cDNA of *DNMT3L* was cloned into pcDNA3.1 Myc-His vector (Invitrogen). Transfections of the pcDNMT3L and pcDNA constructs were performed using Qiagen's SuperFect Transfection Reagent. Cos-1 cells were transfected according to the manufacturer's instructions, and after 48 hours the transfected cells were harvested and cell extracts were separated by electrophoresis on a 10 % SDS-PAGE gel, transferred to nitrocellulose, and probed with an anti-myc antibody (Invitrogen).

For subcellular localisation of the DNMT3L/Dnmt3L protein, the pcDNMT3L and pcDnmt3L constructs were transfected into Cos-1 cells using either Qiagen's SuperFect Transfection Reagent or standard calcium phosphate transfection protocol (Ausubel et al. 1995), and stained with anti-myc antibody (Invitrogen). Cos-7 cells were transfected with pEGFP-C3, pGFP-NLS and pGFP-DNMT3L plasmids using ExGen500 transfection reagent (Fermentas) according to the instructions provided by the manufacturer. Immunofluorescence data were acquired using an Olympus IX70 microscope. Images were captured with a digital CCD camera (Wallac) and UltraVIEW Ver 4.0.15 software for PC.

4.7 LUCIFERASE ASSAY (Study III)

NIH3T3, Tec427.1 or Cos-7 cells were transfected with pBL-KS, pLuc226, pLuc440, pLuc900 and pLuc1345 constructs together with the pLacZ plasmid using ExGen500 transfection reagent (Fermentas) according to the manufacturer’s instructions. Cells were lysed in Promega’s Reporter lysis buffer and luciferase activities were measured with the Luciferase Assay System (Promega) using luminometer Luminoskan Ascent (Thermo Labsystems). All results were normalised against β-galactosidase activity.

4.8 ELECTROPHORETIC MOBILITY SHIFT ASSAY (Study III)

NIH3T3 and Cos-7 nuclear extracts were prepared as described by Murumägi et al. (2003). First, nuclear extracts were preincubated with 0.05 mg/ml poly(dI-dC). After 30 min preincubation on ice, double-stranded oligonucleotides end-labelled by T4 polynucleotide kinase using [γ-^{32}P]dATP were added and incubation was continued for 30 min on ice. Reactions were resolved in a 4.5 % nondenaturing polyacrylamide gel, followed by autoradiography. For competition assays a 100-fold excess of the corresponding unlabelled oligonucleotide was added, and for supershift assay anti-Sp1 and/or anti-Sp3 antibodies (Santa Cruz Biotechnology) were added to the preincubation reactions.

4.9 SITE-DIRECTED MUTAGENESIS (Study III)

Site-directed mutagenesis was performed using two sequential PCRs. During the first PCR, mutations were introduced into Sp1 transcription factor binding sites using two sets of primers including mutated reverse and forward primers. In the second PCR, the products from the first PCR were annealed together to produce a fragment with the Sp1-mutation in the middle of the fragment. After the second PCR, fragments were cloned into a luciferase vector resulting in constructs pLucmut262, pLucmut290, pLucmut300, and pLucmut318. Mutations were verified by DNA sequencing.

4.10 IN VITRO METHYLATION (Study III)

In vitro methylation of Sp1-318 and Sp1-262 oligonucleotides and pLuc226, pLuc440, pLuc900 and pLuc1345 constructs was performed with CpG methylase M.*SssI* (New England Biolabs) following the manufacturer's instructions. *In vitro* methylated Sp1 oligonucleotides were tested in EMSA. Methylated luciferase constructs were purified with phenol/chloroform extraction and ethanol precipitation and, subsequently, used for transfections.

4.11 BISULPHITE SEQUENCING (Study III)

For the bisulphite sequencing, genomic DNAs from mouse ES, NIH3T3 and Tec427.1 cell lines were isolated according to the standard protocols (Sambrook et al. 1989). Genomic DNA extractions from frozen C57BL/6 mouse tissues (testis, ovary, thymus, spleen, liver, heart, lungs) were performed with Trizol reagent (Life Technologies Inc.) according to the manufacturer's instructions. The mouse seminiferous tubule samples used in Study III were isolated by using transillumination-assisted microdissection (Parvinen and Vanha-Perttula

1972), immediately frozen with liquid nitrogen and stored at -80°C until genomic DNA extractions with Trizol reagent (Life Technologies Inc.). Genomic DNAs were denatured by incubating with 0.3 M NaOH for 15 min at 42°C. For deamination 0.5 mM hydroquinone and 2.6 M sodium bisulphite (pH 5.0) were added and incubated in the dark for 16 hours at 50°C. The DNA was desalted with the Wizard DNA Clean-up System (Promega) and desulfonated with 0.3 M NaOH for 15 min at 37°C. The DNA was neutralised with 2 M ammonium acetate, precipitated with ethanol and eluted in 15-100 μl TE. The bisulphite treated DNA was amplified by PCR using primers specific to *Dnmt3L* promoter and subsequently, the PCR products were subcloned into a CR2.1-TOPO vector (Invitrogen) and sequenced.

4.12 QUANTITATIVE REAL-TIME RT-PCR (Study III)

Total RNAs from frozen C57BL/6 mouse tissues (testis, ovary, thymus, spleen, liver, heart, lung) and ES, NIH3T3 and Tec427.1 cell lines were extracted using the Trizol reagent (Life Technologies Inc.), and five micrograms of total RNAs were used for first-strand cDNA synthesis with the First Strand cDNA Synthesis Kit (Fermentas). The PCR amplifications were performed with the LightCycler instrument (Roche Applied Sciences) using the QuantiTect SYBR Green Kit (Qiagen). To control the specificity of the reaction, melting-curve analysis was performed after amplification. For normalisation, the amount of mouse *G3pdh* mRNA was measured in each cDNA sample. Standard curves were generated using cloned *Dnmt3L* and *G3pdh* cDNAs.

4.13 5-AZA-2'-DEOXYCYTIDINE AND TRICHOSTATIN A TREATMENTS (Study III)

Tec427.1 and NIH3T3 cells were plated 24 hours before adding the chemicals. For luciferase assays NIH3T3 cells were transfected with pLuc440 and pLuc1345 constructs using the ExGen500 transfection reagent (Fermentas). Varying concentrations (0-10 μM) of DNA methyltransferase inhibitor 5-azaCdR (Sigma) were added to the media for 48 h. Deacetylase inhibitor, TSA (100 nM, Sigma) was added for 24 hours together with 5-azaCdR where indicated. TSA alone (100nM) was added to Tec427.1 cells for 12 hours and to NIH3T3 cells for 24 hours. Each day cells were treated with fresh medium supplemented with the corresponding chemicals. Subsequently, the total RNA was extracted from the cells using the Trizol reagent (Life Technologies Inc.) according to the manufacturer's instructions and, after the cDNA conversion, used for QPCR. In addition, cells transfected with different luciferase

constructs were treated with 5-azaCdR or TSA as described above and used for luciferase activity assay.

4.14 GLUTATHIONE S-TRANSFERASE PULL-DOWN ASSAY (Study IV)

GST fusion proteins were purified from XL-1 Blue cultures with Glutathione Sepharose 4B (Amersham Biosciences) as described by Frangioni and Neel (1993). For *in vitro* binding assay pcDNMT3L, pcDNMT3L 1-195, pING14AHDAC1 and pSI-DNMT3B 220-753 were *in vitro* translated using the TNT Coupled Reticulocyte Lysate System (Promega) and labelled with ^{35}S-Cys according to the manufacturer's protocol. *In vitro* translated, ^{35}S-radiolabelled proteins were incubated together with purified GST fusion proteins as described by Fuks et al. (2000). Bound proteins were separated in SDS-PAGE and subjected to autoradiography.

In GST pull-downs performed with RD cell lysate, Flag-epitope tagged pcHDAC1 was transfected into RD cells using ExGen500 transfection reagent (Fermentas). After 48 hours cells were lysed and equivalent amounts of purified GST-DNMT3L, GST-DNMT3L 1-195 and GST proteins were incubated with the cell lysate. Pull-down products were separated in SDS-PAGE and analysed by immunoblotting with anti-Flag antibody (Sigma).

4.15 CHLORAMPHENICOL ACETYLTRANSFERASE ASSAY (Study IV)

RD cells were transfected with pM, pM-DNMT3L 1-195, pM-DNMT3L 88-387, pM-DNMT3L 164-387, pM-DNMT3L 1-84, pM-DNMT3L 45-84, pM-DNMT3L 88-153 and pM-DNMT3L constructs together with the Gal4TKCAT reporter using Lipofectamine (Life Technologies) according to the manufacturer's protocol. After 46 hours the cells were lysed and the CAT activity assay was performed using the CAT ELISA Kit (Roche) following the manufacturer's instructions. The deacetylase inhibitor, 100 nM TSA (Sigma) was included in the media for 24 hours where indicated. Transfections were normalised against total protein measured from lysates with the BioRad D_C Protein Assay System (Bio-Rad Laboratories).

4.16 DEACETYLATION ASSAY (Study IV)

THP-1 monocytes were lysed in NET0.2 lysis buffer (50 mM Tris-HCl pH 7.5, 150 mM NaCl, 5 mM EDTA, 2 μg/ml aprotinin, 500 ng/ml leupeptin) on ice and the total protein concentration of the lysate was measured using BioRad D_C Protein Assay system (Bio-Rad Laboratories). Equal amounts of the THP-1 protein lysate were incubated either with Glutathione Sepharose 4B-purified GST-DNMT3L fusion proteins or with anti-HDAC1 and

anti-GAL4 antibodies (Santa Cruz Biotechnology) and Protein A Sepharose beads (Amersham Biosciences). Sepharose beads with attached proteins were used for deacetylation assay performed with the Histone Deacetylation Assay Kit (Upstate Biotechnology). Sepharose-bound proteins were incubated with [^{3}H]-acetylated peptide substrate and deacetylation was measured as released [^{3}H]-acetate using a scintillation counter.

5. RESULTS

5.1 ISOLATION AND CHARACTERISATION OF THE HUMAN AND MOUSE DNMT3L/Dnmt3L GENES

5.1.1 Cloning of the *DNMT3L* and *Dnmt3L* genes (Studies I, II)

In Study I, the genomic sequence of the cosmid clone KU21D28B11 from the chromosome region 21q22.3 was analysed *in silico* using the programs BLAST, Genscan and NIX. Gene modelling programs predicted a novel gene showing similarity with six previously trapped exons (Chen et al. 1996, Kudoh et al. 1997). By screening human thymus cDNA library and placenta, testis and foetal liver mRNAs with primers designed according to predicted and trapped exon sequences, four overlapping cDNA fragments were isolated by RT-PCR. The 5'- and 3'- cDNA fragments were amplified using RACE technique. The isolated cDNA sequence contained 1705 bp including 12 exons distributed over 16 kb of genomic sequence. The cDNA sequence coded for a protein of 387 amino acids, with the translation initiation codon ATG in exon 2. All exons followed the rule of canonical splice sites. A splice variant for the gene lacking exon 8 was also identified. This alternatively spliced transcript caused a frameshift in exon 9, resulting in 21 unrelated amino acid residues followed by a stop codon. In addition, an EST clone lacking serine 333 at the 5' of the exon 12 was detected. Protein sequence analysis revealed that the novel protein was similar to DNA methyltransferases DNMT3A/Dnmt3a and DNMT3B/Dnmt3b (Okano et al. 1998b, Xie et al. 1999) and, consequently, we named the novel gene *DNMT3L* for DNA cytosine-5-methyltranferase 3-like.

A BLAST similarity search performed with human *DNMT3L* cDNA sequence revealed several overlapping mouse EST sequences that were used as templates when designing primers for *Dnmt3L* cloning (Study II). The 5'- *Dnmt3L* cDNA fragment was amplified from testis and the 3'-RACE cDNA product was isolated from thymus. The *Dnmt3L* cDNA contained 1642 bp coding for a protein of 421 amino acids.

Human and mouse genes showed 74 % identity and contained twelve exons with almost identical sizes. The 5' and 3' ends of the genes were slightly different with *Dnmt3L* having the initiation codon ATG in exon 1, whereas *DNMT3L* had the ATG codon in exon 2. The genomic organisation of human and mouse genes was similar, *Dnmt3L* occupying less genomic sequence because of smaller intron sizes. Human and mouse genes are located

between autoimmune regulator, *AIRE/Aire* (Nagamine et al. 1997, Mittaz et al. 1999), and inducible T-cell co-stimulator ligand, *ICOSL/Icosl* (Yoshinaga et al. 2000, Ling et al. 2000) genes in chromosomes 21 and 10 respectively.

5.1.2 Expression pattern (Studies I, II, III)

The expression pattern of *DNMT3L/Dnmt3L* was first studied using Northern blot analysis (Studies I and II). For human *DNMT3L* expression studies commercial Human MTN Blots 1-4 and Human Fetal MTN Blot (Clontech) were used. Mouse *Dnmt3L* expression was investigated using total RNAs extracted from newborn mouse muscle, kidney, liver, lung and thymus, and adult mouse thymus, testis and ovary. No detectable *DNMT3L/Dnmt3L* expression was found in Northern blots even after prolonged exposure, suggesting a low expression level of the gene. Expression patterns were further studied by RT-PCR. The *DNMT3L* 5'- cDNA fragment was amplified from ovary, testis, thymus and foetal thymus, with testis showing the strongest expression level (Study I). However, *DNMT3L* mRNA at the lower level was detected by PCR using increased amount of mRNA or cDNA from other tissues such as foetal liver and placenta. Similar to *DNMT3L*, the mouse *Dnmt3L* transcript was strongly expressed in testis and to a lesser extent in ovary, thymus and foetal tissues (Study II). *Dnmt3L* expression was very high in mouse ES cells and the transcripts were also detected in heart and lung (Study III).

5.1.3 DNMT3L/Dnmt3L proteins

5.1.3.1 ATRX-like zinc finger domain of DNMT3L/Dnmt3L (Studies I, II)

At the protein level DNMT3L and Dnmt3L showed 61 % identity and 75 % similarity. The greatest similarity with other DNMT3/Dnmt3 family members was found at the cysteine-rich region in the N-terminal part of the DNMT3L/Dnmt3L proteins. The DNMT3/Dnmt3 family members shared this cysteine-rich region with ATRX/Atrx (Villard et al. 1997) protein and related fruit fly, zebrafish and *Arabidobsis thaliana* EST sequences were present in GenBank revealing the conserved nature of cysteins and histidines in this region. The novel zinc-finger domain contained a C2C2-type of zinc finger and a PHD-like zinc finger structure, where the histidine within the C4HC3 motif was replaced by a cysteine (C4C4). A schematic representation of the DNMT3L protein is shown in Figure 5A.

5.1.3.2 DNMT3L is not an active methyltransferase

Although sharing significant similarity with DNMT3/Dnmt3 proteins, DNMT3L/Dnmt3L did not contain the conserved FGG, PC or ENV DNA methylase motifs found in the catalytic region of all functional methyltransferases (Lauster et al. 1989a, Kumar et al. 1994). Human DNMT3L was found to be enzymatically inactive in a test performed using a protocol based on the incorporation of 3H methyl groups in poly(dI-dC) (Bestor 1992) (unpublished data in collaboration with Jean Margot and Heinrich Leonhardt of the Franz-Volhard-Klinik, Berlin, Germany).

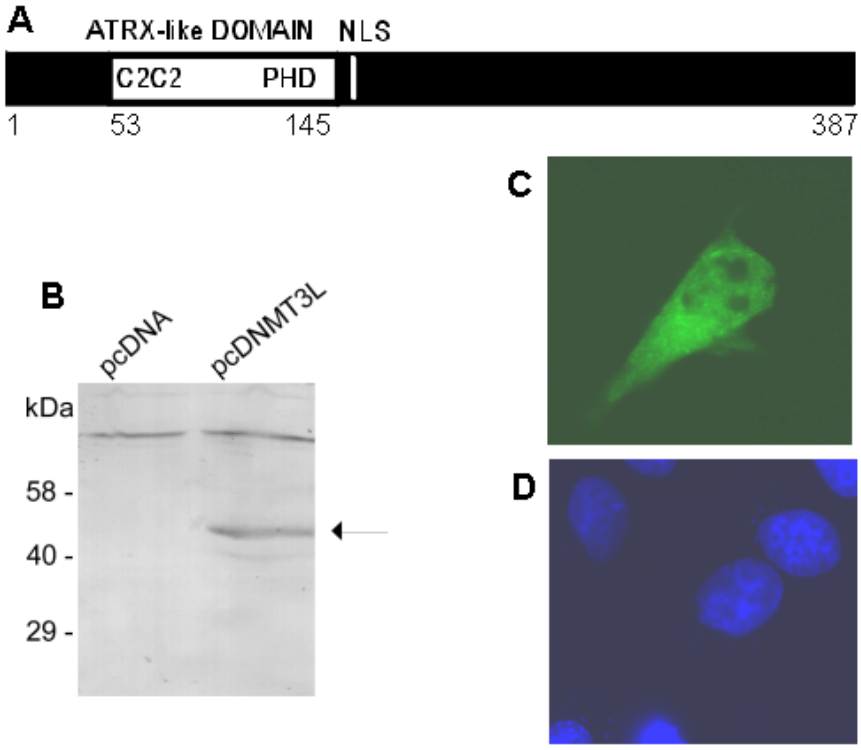

Figure 5. Western blot analysis and subcellular localization of DNMT3L. (A) Schematic representation of the DNMT3L protein. Numbers indicate the amino acids of DNMT3L. The nuclear localisation signal (NLS) is coded by amino acids 156-159. (B) Western blot analysis of the DNMT3L protein. Cos-1 cells were transfected with pcDNMT3L construct containing a myc-his epitope tag downstream to DNMT3L cDNA. DNMT3L protein was detected from the cell extract using monoclonal antibody against the myc tag. The arrow indicates the expressed DNMT3L protein. Protein markers are in kDa. (C) Subcellular localisation of the DNMT3L protein. Cos-1 cells were transfected with pcDNMT3L and immunofluorescence staining was performed with anti-myc antibody. (D) The DAPI staining of the same cells.

5.1.3.3 Western blotting and subcellular localisation of the DNMT3L/Dnmt3L (Studies II, IV)

The *DNMT3L* cDNA was cloned into eukaryotic expression vector under control of the CMV promoter and, subsequently, Cos-1 cells were transfected with the resulting pcDNMT3L construct containing a myc-his epitope tag. Western blot analysis of the cell lysate with the anti-myc tag antibody showed a DNMT3L protein with an approximate molecular mass of 45 kDa (Figure 5B, unpublished data), which is in agreement with the molecular mass of 43.7 kDa predicted *in silico*.

To determine the subcellular localisation of the DNMT3L, Cos-1 cells were transiently transfected with the pcDNMT3L construct. In immunofluorescence staining with anti-myc tag antibody the DNMT3L was not found to locate to any specific cellular structures as the staining was observed diffusely distributed throughout the cytoplasm and cell nucleus (Figure 5C, unpublished data). Similar results were obtained with the mouse pcDnmt3L (Study II). The PSORT II program predicted a potential nuclear localisation signal (NLS), *RRRK* at DNMT3L amino acids 156-159 (see Figure 5A; *RRKR* at Dnmt3L amino acids 190-193). In experiments with green fluorescent protein (GFP) construct containing human NLS, the GFP-NLS fusion protein localised to nucleus demonstrating that the nuclear localisation signal of DNMT3L was functional (Study IV).

5.2 CHARACTERISATION OF THE DNMT3L PROMOTER (Study III)

5.2.1 Minimal promoter area of *Dnmt3L*

A genomic DNA fragment containing 1345 bp upstream from the Dnmt3L translational start site was cloned in front of the luciferase gene and the resulting pLuc1345 construct was transiently transfected into Cos-7 cells. After showing strong luciferase activation by this *Dnmt3L* promoter fragment, several 5' deletion constructs were created. Results from the luciferase assays performed with pLuc1345, pLuc900, pLuc440 and pLuc226 constructs located the *Dnmt3L* minimal promoter area within 440 bp upstream from the translation start site. Luciferase transfections were performed using three different cell lines, mouse fibroblasts NIH3T3, mouse thymic epithelial cells Tec427.1 and Cos-7 cells, all of which yielded similar results. In addition, a repressive effect on the transcriptional activity was found which was mediated by the sequences upstream of the minimal promoter.

5.2.2 The *Dnmt3L* promoter contains functional Sp1/Sp3 transcription factor binding sites

MatInspector program predicted five putative GC boxes in the *Dnmt3L* minimal promoter sequence. EMSAs performed with the corresponding oligonucleotides [Sp1-374, Sp1-318, Sp1-303 (covering both Sp1-290 and Sp1-300 sites), and Sp1-262] showed three separate band-shifts with oligonucleotides Sp1-262, Sp1-303, and Sp1-318. Specificities of the band-shifts were confirmed with the competing unlabelled oligonucleotides and with the oligonucleotides with mutated binding sites. Supershift assays performed using Sp1 and Sp3 specific antibodies indicated the presence of Sp1 protein in the uppermost complex, whereas two lower complexes contained Sp3 proteins. EMSAs performed with NIH3T3 and Cos-7 nuclear extracts yielded identical results. *In vitro* methylation of the oligonucleotides Sp1-262 and Sp1-318 did not affect the formation of the band-shift complexes in EMSA, indicating that CpG-methylation does not interfere with Sp1/Sp3-binding to the *Dnmt3L* promoter.

Luciferase assays performed using pLuc440 constructs with mutations in four different Sp1 binding sites (Sp1-262, Sp1-290, Sp1-300, Sp1-318) showed decreased transcriptional activities (30-79 %) when compared with control pLuc440 activity, demonstrating that the minimal promoter area contained at least four functional Sp1/Sp3 binding sites.

5.2.3 The *Dnmt3L* promoter and epigenetic chromatin modifications

In vitro methylated GC-rich pLucDnmt3L promoter constructs were completely negative in the luciferase assay, suggesting that *Dnmt3L* gene could be regulated through the methylation of its promoter. To test this further, bisulphite sequencing analysis of the minimal promoter area was performed using genomic DNAs extracted from different mouse cell lines and tissues. To find out whether there was a correlation between the promoter methylation and *Dnmt3L* mRNA expression levels, the RNAs extracted from the same samples were used to detect the expression level of *Dnmt3L* by QPCR. Results obtained with mouse ES, NIH3T3 and Tec427.1 cell lines showed a correlation between the expression and promoter methylation. In ES cells, the *Dnmt3L* promoter was completely unmethylated and the *Dnmt3L* expression level was very high. NIH3T3 and Tec427.1 cells did not express *Dnmt3L* and promoter methylation level was 80 %. The correlation was less obvious with tissue samples from testis, ovary, thymus, spleen, liver, heart, lung, most likely due to the heterogeneous sample material. For example, *Dnmt3L* transcripts were detected both in testis and ovary but

in testis the promoter was not methylated (5 %) and in ovary the methylation level was 60 %. In ovary, *Dnmt3L* is expressed in growing oocytes (Bourc'his et al. 2001, Hata et al. 2002) with the rest of the tissue lacking the *Dnmt3L* transcript. *Dnmt3L* expression in other tissues, except in lungs, was very weak with methylation levels of 53-78 % (lungs 31 %).

NIH3T3 and Tec427.1 cells did not express *Dnmt3L*. Both cell lines were treated with DNA methyltranferase inhibitor, 5-aza-CdR and/or deacetylase inhibitor, TSA and tested for *Dnmt3L* expression. QPCR results demonstrated that the inhibitors could restore *Dnmt3L* expression in non-expressive cell lines, suggesting inactivation of *Dnmt3L* promoter by methylation and histone deacetylation. To conclude, the results showed that *Dnmt3L* expression is regulated through the reversible epigenetic chromatin modifications, DNA methylation and histone deacetylation.

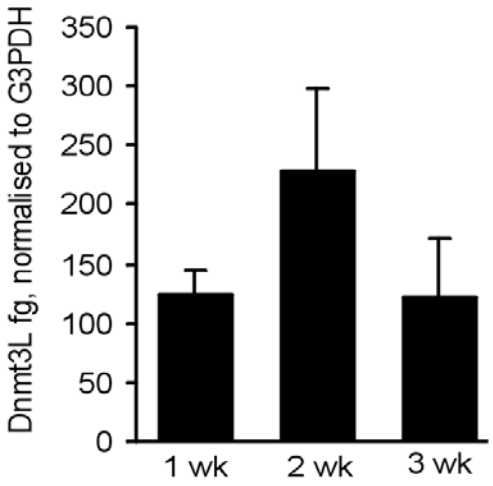

Figure 6. Quantitative RT-PCR analysis of *Dnmt3L* mRNA extracted from 1, 2, and 3-week-old mouse testis. Results are shown as the amounts of *Dnmt3L* mRNA (fg) in each sample. The bars represent the mean ±SD of two separate experiments, each performed in duplicate and normalised against the amount of *G3pdh* mRNA in each sample.

5.2.4 *Dnmt3L* and mouse spermatogenesis

Dnmt3L deficient mice were sterile due to the defect in spermatogenesis resulting in azoospermia, indicating that Dnmt3L is required for normal spermatogenesis (Bourc'his et al. 2001, Hata et al. 2002). Bisulphite sequencing of the *Dnmt3L* minimal promoter DNA and QPCR analysis of *Dnmt3L* mRNA were performed with genomic DNA and total RNA

extracted from adult mouse seminiferous tubules at the defined stages (II-VI, VII-VIII, IX-I) of the mouse seminiferous epithelial cycle. The results showed that the *Dnmt3L* promoter was unmethylated and *Dnmt3L* transcripts were detected at all developmental stages. The highest expression level was seen during stages VII-VIII when type B spermatogonia enter the preleptotene stage and more mature spermatocytes are at the pachytene stage of meiosis (Oakberg 1956). The expression result was confirmed with samples from 1, 2, and 3-week-old mouse testes, representing the first wave of spermatogenesis. Consistently, *Dnmt3L* was expressed at all stages with the highest expression level at two weeks, at the beginning of meiosis (Figure 6).

5.3 DNMT3L IS A TRANSCRIPTIONAL REPRESSOR (Study IV)

5.3.1 DNMT3L represses transcription through a PHD-like zinc finger domain

To study whether DNMT3L, like Dnmt3a and Dnmt3b (Bachman et al. 2001, Fuks et al. 2001), functions as a transcriptional repressor, the full-length DNMT3L cDNA was cloned downstream from the Gal4 DNA binding domain of the pM expression vector. RD cells were transiently transfected with the resulting pM-DNMT3L 1-387 construct, together with a CAT reporter gene containing five Gal4 binding sites upstream of the strong thymidine kinase promoter. The construct pM-DNMT3L 1-387 caused a 30 % reduction in reporter activity, when compared with the transcriptional activity of the reporter seen with the pM vector alone. The repressive capacity of DNMT3L was further studied using five different pM-DNMT3L constructs containing different DNMT3L protein domain combinations (pM-DNMT3L 1-84, pM-DNMT3L 1-195, pM-DNMT3L 88-387, pM-DNMT3L 164-387, pM-DNMT3L 1-387). The results demonstrated that the repressor activity was mediated by the PHD-like zinc finger region of DNMT3L protein. However, some repression effect was also found with the construct containing the C-terminal part of DNMT3L lacking the PHD-like zinc finger. The level of repression was dose-dependent since increasing amounts (from 200 ng to 1 μg) of DNMT3L PHD-like zinc-finger construct (pM-DNMT3L 1-195) decreased the CAT activity fourfold.

5.3.2 DNMT3L repression is TSA sensitive

For some time histone deacetylation has been associated with transcriptional repression (Davie 1998). To test whether the repressive effect of DNMT3L is HDAC-dependent, RD

cells were transiently transfected with the pMDNMT3L 1-195 construct and, subsequently, the cells were treated with the HDAC-specific inhibitor trichostatin A, TSA. TSA-treatment relieved the repressive state of CAT activity, suggesting that the repressional activity of DNMT3L could be HDAC-dependent.

5.3.3 DNMT3L interacts with HDAC1 and associates with deacetylase activity

All functional DNA methyltransferases act as transcriptional repressors through interacting with HDAC1 (Fuks et al. 2000, Bachman et al. 2001, Fuks et al. 2001). *In vitro* GST pull-down experiments performed using GST-DNMT3L 1-195 and *in vitro* translated HDAC1 showed that DNMT3L also interacts with HDAC1. The same results were obtained with GST-HDAC1 and *in vitro* translated DNMT3L. Interaction was further tested by performing a GST-DNMT3L pull-down assay from RD cell lysate containing transfected Flag-epitope tagged pcHDAC1, followed by Western blotting with anti-Flag antibody.

Association of DNMT3L with histone deacetylase activity was tested by GST-DNMT3L pull-down assays using THP-1 total protein lysate as a HDAC1 source. Pull-down products were incubated with [^{3}H]-acetylated peptide substrate and the deacetylase activity was measured as released [^{3}H]-acetate. Consistent with the finding that the PHD-like zinc finger is a main repression domain of DNMT3L in CAT assay, these results showed that deacetylase activity was linked to the PHD-like zinc finger containing GST-DNMT3L fusion proteins.

5.4 DNMT3L FORMS DIMERS AND INTERACTS WITH DNMT3B

This chapter contains our unpublished results.

5.4.1 DNMT3L forms dimers through its zinc fingers

To further characterise the role of the zinc finger region (amino acids 53-145) at the N-terminus of the DNMT3L, the protein capability for intramolecular interaction was tested by GST pull-down assay, using *in vitro* translated pcDNMT3L and full-length GST-DNMT3L or GST-DNMT3L 1-195 with zinc fingers. As shown in Figure 7A, both GST-DNMT3L and GST-DNMT3L 1-195 interacted with *in vitro* translated DNMT3L protein. Figure 7B indicates that GST-DNMT3L and GST-DNMT3L 1-195 bound DNMT3L 1-195, whereas GST alone or the C-terminal part of the DNMT3L protein GST-DNMT3L 164-387 (without zinc fingers) did not. It was concluded that DNMT3L could form homodimers *in vitro* using the zinc finger region for interaction.

5.4.2 DNMT3L interacts with DNMT3B

Since DNMT3L shares the zinc finger region with other DNMT3 family members, the obvious question was whether DNMT3L could interact with DNMT3A or DNMT3B through the zinc finger domain. We cloned a partial DNMT3B cDNA sequence from an IMAGE EST clone BE278296 into the expression vector pSI and repeated the *in vitro* GST pull-down assay using *in vitro* translated pSI-DNMT3B and GST-DNMT3L 1-195 proteins. Figure 7C shows that DNMT3L interacts with DNMT3b and most likely uses zinc fingers for interaction.

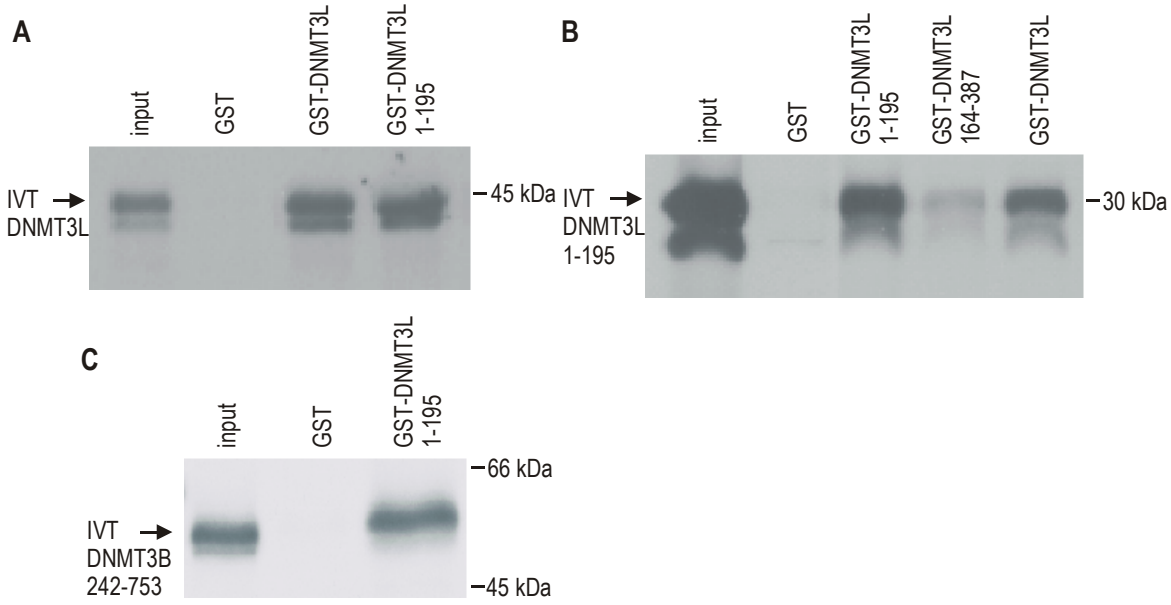

Figure 7. DNMT3L forms homodimers and interacts with DNMT3B through its zinc fingers. (A) GST pull-down assay using *in vitro* translated (IVT) pcDNMT3L and GST-DNMT3L or GST-DNMT3L 1-195. (B) pcDNMT3L 1-195 does not interact with the C-terminal part of the DNMT3L protein (residues 164-387, without zinc fingers). (C) *In vitro* translated DNMT3B protein (amino acids 242-753) was tested in GST pull-down experiment for binding to GST-DNMT3L 1-195. IVT proteins are indicated by an arrow on the left and the molecular weight is indicated on the right.

6. DISCUSSION

This chapter presents a general discussion of Dnmt3L function based on the previously published reports. Detailed discussions of the results can be found in the original communications.

6.1 ATRX-LIKE DOMAIN OF DNMT3L

The most distinct domain in DNMT3L/Dnmt3L protein is a combined zinc finger motif consisting of a C2C2-type and an imperfect PHD-type of zinc fingers (Studies I and II). In an imperfect PHD zinc finger, the histidine within C4HC3 motif is replaced by cysteine (C4C4). In addition to DNMT3 family members, the combined zinc finger motif is an important structural feature of the ATRX protein (Villard et al. 1997) and as a consequence, the zinc finger domain of the DNMT3 family proteins is generally called an ATRX-like domain. About 65 % of the mutations leading to ATRX syndrome have been located in the zinc finger domain of the ATRX, underlining the functional importance of the region (Gibbons et al. 1997). The zinc finger-domain of ATRX has been shown to bind DNA *in vitro* and mutations in zinc fingers have been demonstrated to decrease the DNA binding efficiency and alter the subcellular localisation of the ATRX protein (Cardoso et al. 2000). However, the specific target sequence of ATRX is unknown and there is no further evidence of DNA binding capacity for any other ATRX-like domain. Instead, all DNMT3 family members use ATRX-like domain for interaction with HDAC1 protein (Bachman et al. 2001, Fuks et al. 2001, Deplus et al. 2002, Study IV). Furthermore, Dnmt3a uses the ATRX-like domain for interactions with HMTase SUV39H1, HP1 and transcription factor RP58 (Fuks et al. 2001, Fuks et al. 2003b). We showed with a GST pull-down assay that DNMT3L could form dimers through its zinc fingers.

Similar to ATRX, several other disease-related genes contain mutated or deleted PHD zinc finger domains (Nagamine et al. 1997, Jacobson and Pillus 1999, Gunduz et al. 2000). The first and so far the only example of naturally occurring mutations in a mammalian DNA methyltransferase gene was described in 1999, when it was shown that the chromosome instability and immunodeficiency syndrome (ICF) is caused by the mutations in *DNMT3B* gene (Okano et al. 1999, Hansen et al. 1999, Xu et al. 1999). However, no ICF-causing mutations have been found in the ATRX-like domain of the DNMT3B protein (Lappalainen

and Vihinen 2002). It is possible that a mutated ATRX-domain is embryonic lethal and ICF-patients with mutated zinc fingers are never even born.

Interestingly, within the imperfect PHD zinc finger, DNMT3L shares a similarity to its neighbouring gene product *AIRE*, which has two PHD zinc finger domains (Nagamine et al. 1997). Some of the mutations in *AIRE*, causing a syndrome called autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED, OMIM 240300), result in truncated AIRE proteins lacking the PHD domain(s). The 5' regions of *DNMT3L/Dnmt3L* and *AIRE/Aire* genes lie next to each other on human chromosome 21 and on the corresponding mouse chromosome 10. In Study III, we indicated that the first exon of an alternative splice variant of *Dnmt3L* is located in the middle of the *Aire* gene. Altogether, the data suggests that these two genes may share some regulatory regions and can be expressed co-ordinately. *DNMT3L/Dnmt3L* appears to be expressed predominantly in the gonads and thymus and expression in thymus seems to be restricted to only certain cell types (Studies I, II and III). Similarly, *AIRE* is expressed predominantly in thymus and the expression is limited to an epithelial set of cells in thymic medulla (Zuklys et al. 2000). Like DNMT3L, AIRE also functions as a transcriptional regulator (Pitkänen et al. 2000) and it has been shown to have a central role in the maintenance of immune tolerance (reviewed in Pitkänen and Peterson 2003). Given the structural and functional similarities, it remains to be studied whether the DNMT3L and AIRE proteins are involved in common cellular mechanisms.

6.2 DNMT3L AS A TRANSCRIPTIONAL REPRESSOR

In addition to their methyltransferase activity, Dnmt1, Dnmt3a and Dnmt3b can act as transcriptional repressors by associating with histone deacetylase activity (Fuks et al. 2000, Robertson et al. 2000b, Rountree et al. 2000, Fuks et al 2001, Bachman et al. 2001). Dnmt3a and Dnmt3b repress transcription through the ATRX-like domain, which they use for interaction with HDAC1 (Fuks et al. 2001, Bachman et al. 2001). DNA-binding transcription factors, interacting with the Dnmt-repression complexes, specify the targets of repression. For example, the repressional effect of Dnmt3a is accomplished through direct interaction with transcription factor RP58 (Fuks et al. 2001).

In Study IV we showed that DNMT3L, like Dnmt3a and Dnmt3b, can repress transcription by binding directly to HDAC1 protein. Our findings were confirmed by Deplus et al. (2002), who published their results in a subsequent issue of the same journal. We identified the PHD zinc finger of the ATRX-like domain as a strong repressional domain of DNMT3L (see Figure 5A). In addition, we identified the C-terminal part without PHD finger

as an additional repressional region of DNMT3L, although the repressional effect was not as strong as with PHD zinc finger domains. Dnmt1 has two subdomains (aa 653-730 and 686-812) associating with repressional activity (Fuks et al. 2000). The other domain (Dnmt1 653-730) represses transcription in an HDAC-independent way, showing no interaction with HDAC1 or sensitivity towards deacetylase inhibitor, TSA, whereas the opposite is seen with Dnmt1 686-812. An HDAC-independent repressive effect is also seen with the N-terminal region (without ATRX-domain) of the Dnmt3a (aa 1-493) and Dnmt3b (aa 1-389) proteins (Bachman et al. 2001). The C-terminus of DNMT3L does not interact with HDAC1 (Deplus et al. 2002) or associate with deacetylase activity (Study IV), but it does show repressive activity (Study IV), suggesting that DNMT3L may also have an HDAC-independent way to repress transcription. A recent study by Suetake and coworkers (2004) suggests that DNMT3L does not have DNA binding capacity on its own. Whether some DNA-binding factor guides DNMT3L to specific target sequences, like with Dnmt1 and Dnmt3a, remains to be studied.

6.3 DNMT3L COFUNCTIONS WITH DNMT3A AND DNMT3B

Co-immunoprecipitation data by Hata et al. (2002) indicated that Dnmt3L interacts with Dnmt3a and Dnmt3b. Thereafter, the interaction has been demonstrated in more detail in several other studies (Chédin et al. 2002, Margot et al. 2003, Suetake et al. 2004). Margot et al. (2003) used full-length human DNMT3L in a yeast two-hybrid system showing that Dnmt3a and Dnmt3b interact with DNMT3L through their C-terminal domains. Recently it was shown by gel shift assay that DNMT3L also uses the C-terminus (aa 160-387) for interaction with Dnmt3a and Dnmt3b (Suetake et al. 2004). Our GST pull-down result (Chapter 5.4.2) with DNMT3L and DNMT3B demonstrated the interaction between human proteins. Contradictory to recent data obtained with human DNMT3L and mouse Dnmt3b (Suetake et al. 2004), our result indicates that DNMT3L uses the N-terminal region (1-195) for interaction with DNMT3B. We did not study DNMT3L/DNMT3B interaction using the C-terminal region of DNMT3L. In both studies GST-DNMT3L was used, but we used ^{35}S-labelled *in vitro* translated DNMT3B for detection and the gel shift assay was based on detection of DNA in the DNMT3L/Dnmt3b/DNA complex formed in reaction conditions favouring DNA methylation. Further *in vivo* studies are needed to confirm the findings and demonstrate if both interaction domains are functional under different circumstances.

We showed that in Cos-1, Cos-7 and RD cells, DNMT3L/Dnmt3L is distributed both, in the cytoplasm and the nucleus (Figure 5C, Studies II and IV). Co-transfections of Dnmt3L

with Dnmt3a or Dnmt3b in NIH3T3 cells caused re-localisation of diffusely distributed Dnmt3L to the nucleus (Hata et al. 2002), which was most likely caused by the interaction between Dnmt3 proteins. However, we demonstrated in Study IV that nuclear localisation of the DNMT3L can also occur through the functional nuclear localisation signal, *RRRK*. The NLS locates C-terminal to ATRX-like domain (Figure 5A) and can also be found in other DNMT3 proteins. Bachman et al. (2001) have previously shown that nuclear targeting of Dnmt3a and Dnmt3b occurs through the N-terminus of the proteins and, therefore, the function of the novel NLS in Dnmt3a and Dnmt3b is still unclear.

Recently it has been shown that DNMT3L strongly stimulates *de novo* methylation by Dnmt3a. In the study by Chédin et al. (2002), DNMT3L was co-transfected into human cells together with Dnmt3a or DNMT3B and subsequently the methylation status of replicating minichromosomes carrying various imprinting control regions (ICRs) was investigated. DNMT3L stimulated Dnmt3a-mediated methylation of maternally methylated *SNRPN*, *Snrpn* and *Igf2r* ICRs. Stimulation was also observed in non-imprinted episomal sequences and endogenous genomic sequences, indicating that DNMT3L is a general stimulator of Dnmt3a (Chédin et al. 2002). At first, the stimulatory effect was not observed when DNMT3L was co-transfected with DNMT3B (Chédin et al. 2002), but now it has been shown that in different reaction conditions DNMT3L also stimulates DNMT3B-mediated methylation (mentioned in Suetake et al. 2004 as personal communication with Dr. F. Chédin). 1.5 to 3-fold stimulation in methylation activity is seen when DNMT3L is incubated with Dnmt3a or Dnmt3b *in vitro* (Suetake et al. 2004). As in a previous study, *in vitro* results indicated that DNMT3L did not have any specific DNA target functioning as a general stimulator of Dnmt3a and Dnmt3b.

6.4 DNMT3L DEFICIENT MICE ARE INFERTILE

The first report describing *Dnmt3L* knockout mice was released in November 2001 by the group of Timothy Bestor of Columbia University, New York (Bourc'his et al. 2001). The data was soon confirmed by another group (Hata et al. 2002). In both mouse models, *Dnmt3L* was knocked out by replacing exons coding for the ATRX-like domain of the protein. *Dnmt3L* targeting vector contained the β-galactosidase reporter gene, which was used for X-gal stainings to study the expression pattern of *Dnmt3L*. Consistent with our expression data showing *Dnmt3L* expression in ovary and testis (Studies II, III), *Dnmt3L* was expressed in growing oocytes of females and in gonocytes and differentiating spermatocytes in males (Bourc'his et al. 2001, Hata et al. 2002). In addition, E7.5 and E8.5 embryos showed high

Dnmt3L expression in chorion (Hata et al. 2002), which could explain our successful PCR amplifications using placenta mRNA as template during *DNMT3L* cloning (Study I).

Homozygous *Dnmt3L* mutant mice (*Dnmt3L* -/-) were viable with an apparently normal phenotype, but both sexes were infertile (Bourc'his et al. 2001, Hata et al. 2002). Oogenesis of *Dnmt3L* -/- females was normal and heterozygous progeny of *Dnmt3L* -/- females developed normally at E8.5, but showed reduced size at E9.5 and died around E9.5-10.5. These embryos showed neural tube defects and abnormalities of extraembryonic tissues (Bourc'his et al. 2001, Hata et al. 2002). The data indicated that maternally expressed *Dnmt3L* is essential for normal embryonic development. *Dnmt3L* -/- male mice had normal complements of germ cells at birth and one week, but showed severe hypogonadism and Sertoli-cell-only phenotype by adulthood (Bourc'his et al. 2001, Hata et al. 2002). At four weeks of age, seminiferous tubules contained hardly any differentiated spermatocytes (Hata et al. 2002). The results indicated that Dnmt3L is involved in mouse spermatogenesis and is probably needed during the differentiation of spermatogonia into spermatocytes. Spermatogonia differentiate into spermatocytes through meiosis. During the prophase of the meiosis chromosomes start to condense and pair with their homologues. Crossing-over occurs in pachytene spermatocytes (Page and Hawley 2003). In Study III, we showed that *Dnmt3L* is expressed at all stages of spermatogenesis with the highest expression levels seen at two weeks, when pachytene spermatocytes appear. It is possible that Dnmt3L plays a role in meiosis, during the processes involving chromosome condensation, homologous chromosome pairing and recombination.

6.5 DNMT3L IS NEEDED FOR ESTABLISHMENT OF METHYLATION IMPRINTS

Embryos lacking maternal *Dnmt3L* showed loss of maternal methylation imprints, resulting in abnormal expression of maternally imprinted genes (Bourc'his et al. 2001, Hata et al. 2002). Paternally methylated *H19* and *Rasgrf1* genes and other common targets of methylation were not affected in heterozygous embryos of *Dnmt3L* -/- females, indicating that *Dnmt3L* is especially required for the establishment of maternal methylation imprints during oogenesis. Since *Dnmt3L* -/- males lacked spermatocytes and did not produce any sperm but did contain spermatogonia, the role of *Dnmt3L* in the establishment of paternal methylation imprints remained unclear (Bourc'his et al. 2001, Hata et al. 2002).

Paternal imprinting occurs mainly in gonocytes and is completed by the onset of meiosis as spermatocytes start to appear (Davis et al. 2000, Ueda et al. 2000, Lucifero et al. 2002). Thus, spermatogonia should already include most of the methylation imprints. Chédin

et al. (2002) did not study the effect of DNMT3L, Dnmt3a and DNMT3B on paternal ICRs. However, *in vitro* studies indicated that DNMT3L stimulates methylation of the paternally methylated *H19* gene (Suetake et al. 2004). In addition, expression profiles of *Dnmt3a* and *Dnmt3L* support the function in paternal imprinting. Both *Dnmt3a* and *Dnmt3L* transcripts are upregulated before birth, at the initiation of methylation pattern establishment (La Salle et al. 2004). Interestingly, *Dnmt3a* -/- mice also have defects in spermatogenesis, although these defects are not as severe as in *Dnmt3L* -/- mice (Hata et al. 2002).

In mouse maternal methylation imprints are established during oocyte growth with the help of Dnmt3L, Dnmt3a and Dnmt3b. In fact, Dnmt3a -/- mice also showed defective maternal methylation (Hata et al. 2002). Maternal methylation is acquired asynchronously at different genes and establishment of the methylation is dependent on oocyte diameter and accumulation of *Dnmt3a*, *Dnmt3b* and *Dnmt3L* transcripts (Lucifero et al. 2004). Several studies have indicated that all *Dnmt3*s are expressed during various stages of oocyte development (Bourc'his et al. 2001, Hata et al. 2002, Huntriss et al. 2004, Lucifero et al. 2004, La Salle et al. 2004). In concordance with this, the amount of *Dnmt3L* is remarkably upregulated during oocyte growth (Lucifero et al. 2004, La Salle et al. 2004).

The *DNMT3L* expression pattern and timing of the establishment of maternal imprints in human are still obscure. We have shown that similar to mouse *Dnmt3L*, *DNMT3L* is expressed in ovary (Studies I and II). However, the exact place and cell type remained unknown. A recent study by Huntriss et al. (2004) indicated that *DNMT3L* is expressed only after fertilisation, as *DNMT3L* transcript was never seen during oocyte development and did not appear until the blastocyst stage of development. A discrepancy in the timing of methylation acquisition has also been demonstrated. A study by El-Maarri et al. (2001) indicated that in the case of *SNRPN*, methylation imprints are established during or after fertilisation since *SNRPN* ICR was completely unmethylated in human oocytes. Recently, it was shown that *SNRPN* ICR methylation imprints are established already at the germinal vesicle (GV)-stage oocytes, well before fertilisation (Geuns et al. 2003).

Familial biparental hydatidiform mole (FBHM, OMIM 231090) is a disorder where affected women repeatedly have molar pregnancies even with different partners (Judson et al. 2002). The condition is characterised by the absence of an embryo and abnormal proliferation of extra-embryonic tissue, caused by defects in the establishment of maternal methylation imprints at multiple loci. It has been suggested that FBHM is caused by a defective, *trans*-acting regulator of maternal imprinting, which is mutated in affected women (Judson et al. 2002). Although all *DNMT*s, especially *DNMT3L*, seemed to be excellent candidate genes for

FBHM, no disease-causing mutations were found in these genes (Hayward et al. 2003). These results suggested that there are other still unknown *trans*-acting factors required for the establisment of maternal imprints in humans.

6.6 REGULATION OF DNMT3L GENE

In Study III, we demonstrated that *Dnmt3L* expression is regulated by Sp1 and Sp3 transcription factors and through epigenetic modification of the promoter sequence. *Dnmt3a*, *Dnmt3b* and *Dnmt3L* are all highly expressed in undifferentiated ES cells and downregulated in differentiated embryoid bodies (Okano et al. 1999, Hata et al. 2002, Study III). In agreement with this, we showed that in ES cells the *Dnmt3L* promoter is unmethylated, whereas in non-expressive somatic cell lines the promoter was heavily methylated (Study III). Previous studies have shown that multiple Sp1 sites can target local demethylation specifically in embryonic cells (Frank et al. 1991). However, further studies are needed to ascertain the role of Sp1-mediated *Dnmt3L* regulation in specific cell-types. *Dnmt3L* mRNA was upregulated when non-expressive cells were treated with methyltransferase and deacetylase inhibitors, indicating that the repressed state of *Dnmt3L* transcription was achieved through reversible epigenetic chromatin modifications, DNA methylation and histone deacetylation (Study III). Human *DNMT3A* and *DNMT3B* contain several promoters that are either CpG-rich or CpG-poor in sequence and typically lack TATA consensus sequences (Yanagisawa et al. 2002). The human *DNMT3L* promoter has not been studied but in Study III we showed that mouse *Dnmt3L* promoter is CpG-rich and TATA-less. The mouse *Dnmt3b* minimal promoter upstream sequences tended to repress transcription (Ishida et al. 2003). The same effect was seen with the *Dnmt3L* promoter (Study III) and the human *DNMT3A* and *DNMT3B* promoters (Yanagisawa et al. 2002). Interestingly, the repression effect of the Dnmt3b promoter upstream region was not seen in ES cells, only in somatic cells (Ishida et al. 2003). Active transcription of Dnmt3a and Dnmt3b is required for normal embryonic development, as Dnmt3a -/-, Dnmt3b -/- ES cells lack *de novo* methylation activity and retain an undifferentiated morphology (Okano et al. 1999). *Dnmt3L* -/- ES cell lines grow and differentiate normally, indicating that the zygotic function of Dnmt3L is not essential for embryonic development (Hata et al. 2002). However, the function of maternally expressed *Dnmt3L* is essential for embryonic development as seen with the developmental defect in heterozygous embryos of Dnmt3L -/- females (Bourc'his et al. 2001, Hata et al. 2002).

DNMT3A/Dnmt3a and *DNMT3B/Dnmt3b* genes are regulated through several promoters and isoforms (Chapters 2.3.3.1 and 2.3.4). We identified three different splice

variants for human *DNMT3L* (Study I) and in Study III alternative 5' exons of *Dnmt3L* were discussed (GenBank sequences AJ404467, NM_019448, CA559020). The functional significance of *DNMT3L/Dnmt3L* promoters and isoforms needs to be addressed in the future.

6.7 FUTURE PROSPECTS

Usually genes are expressed in biallelic manner from both parental alleles. Imprinted genes are expressed only from one allele, in a parent of origin-specific manner and Dnmt3L has been shown to regulate this event. This kind of allelic exclusion resulting in monoallelic expression is also seen with some non-imprinted genes (Ohlsson et al. 1998). Allelic exclusion of immunoglobulin, T-cell receptor, interleukin IL2 and IL4, olfactory receptor and certain X-linked genes occurs independent of the parent of origin. Since methylation is considered as one of the mechanisms of allelic exclusion and Dnmt3L has been determined as a common regulator of Dnmt3a and Dnmt3b, it is possible that Dnmt3L also has an important role in allelic exclusion of non-imprinted genes. Recent results indicate that different chromatin modification complexes are interconnected and have cross talk with each other (Chapters 2.1.2.2, 2.1.3, 2.3.6). Direct interaction between HMTase SUV39H1 and Dnmt3a is mediated by the ATRX-like domain of Dnmt3a (Fuks et al. 2003b). Furthermore, targeted disruption of *Suv39h1* and *Suv39h2* HMTsases results in similar meiotic defect in spermatogenesis as seen with *Dnmt3L -/-* mice (Peters et al. 2001, Bourc'his et al. 2001, Hata et al. 2002). This data suggests that Dnmt3L may be associated with HMTase activity. The role of Dnmt3L in the regulation of paternal imprinting and in the development of cancer also remains to be studied.

7. SUMMARY AND CONCLUSIONS

In this thesis we described the identification of the human and mouse *DNMT3L/Dnmt3L* genes sharing a similarity with *de novo* methyltransferases Dnmt3a and Dnmt3b, which are essential for normal mammalian development. Human and mouse DNMT3L/Dnmt3L genes showed 74 % identity and contained a similar genomic organisation on human chromosome 21 and on mouse chromosome 10. *DNMT3L*/*Dnmt3L* was strongly expressed in testis and to a lesser extent in ovary, thymus and foetal tissues. *Dnmt3L* expression was very high in mouse ES cells. At the protein level DNMT3L and Dnmt3L showed 61 % identity and 75 % similarity. The highest similarity with other DNMT3/Dnmt3 family members was found at the cysteine-rich region in the N-terminal part of the DNMT3L/Dnmt3L proteins. In addition to the cysteine-rich zinc finger region, consisting of C2C2 and PHD-like zinc fingers, DNMT3L protein contained a functional nuclear localisation signal. Subcellularly, DNMT3L/Dnmt3L was located both in cytoplasm and nucleus.

To study the regulation of *Dnmt3L*, we isolated the *Dnmt3L* promoter region. By using luciferase reporter assay, EMSA and site-directed mutagenesis we identified the *Dnmt3L* minimal promoter area containing four functional Sp1/Sp3 transcription factor binding sites. By using bisulphite sequencing and quantitative RT-PCR analysis we demonstrated that the *Dnmt3L* expression level correlates with the methylation level of the promoter. Regulation of *Dnmt3L* was further studied using methyltransferase and deacetylase inhibitors in cell culture. The results confirmed that *Dnmt3L* was regulated through epigenetic chromatin modifications, DNA methylation and histone deacetylation.

Earlier studies indicated that Dnmt3a and Dnmt3b are transcriptional repressors using a zinc finger-domain to interact with histone deacetylase, HDAC1. Using a Gal4-based DNA binding system together with CAT reporter gene, we showed that DNMT3L represses trancription and that repression is TSA sensitive, indicating that repression is HDAC-dependent. We identified the PHD-like zinc finger as a main repressional domain of DNMT3L, although some repressional activity was also detected in the C-terminal part of the DNMT3L. GST pull-down assays showed a direct interaction between DNMT3L and HDAC1. The association of DNMT3L with deacetylase activity was demonstrated by using GST pull-down assay and an [^{3}H]-acetylated peptide substrate.

In conclusion, we have identified and characterised human and mouse *DNMT3L*/*Dnmt3L* genes and shown that the expression of *Dnmt3L* is regulated by epigenetic

chromatin modifications. Although DNMT3L was found enzymatically inactive being unable to methylate DNA *in vitro*, we demonstrated that DNMT3L protein participates in chromatin modification. We identified DNMT3L as a transcriptional repressor interacting with HDAC1 to repress transcription from other genes.

ACKNOWLEDGEMENTS

This study was carried out in the Autoimmunity Study Group, at the Institute of Medical Technology, University of Tampere.

I want to thank my supervisor, Pärt Peterson, for giving me a chance to work with the *Dnmt3L* gene. His excellent guidance and support have helped me to solve the problems raised during the study. I also want to thank our former group leader and the head of the institute, Kai Krohn, for admitting me to his research group and for teaching me the importance of making decisions and putting them into action to get results.

I am grateful for the financial support and security provided by the graduate school position during the 4-year period. It has also been interesting to have an opportunity to be a student member on the Tampere Graduate School in Biomedicine and Biotechnology-board and influence the developing graduate school system on the Kauppi campus.

My sincere thanks go to my collaborators and co-authors for their contribution to this study. The excellent technical assistance, organising ability and common sense of Ulla Kiiskinen especially is highly appreciated – she made life in the lab much easier.

I wish to thank Heinrich Leonhardt and Antero Salminen, the external reviewers of the manuscript and language consultant Virginia Mattila for their criticism and suggestions on improving the thesis.

The colleagues and other staff of the IMT are acknowledged for their willingness to help in different technical problems or questions. Special thanks go to the members and visitors of our group Jukka, Maarit, Astrid, Niko, John and Ingrid for their help and friendship. The technical assistance of Marianne Karlsberg in the mutation screening experiments is also appreciated. The help of Vladimir Ovod with α-DNMT3L antibody production is acknowledged, likewise our interesting discussions in the buss on the way to work. These and other people in our coffee room (Aulikki, Tuula, Mariitta, Anne, Merja, Anssi, Heli, Mari, Marja, Auni, Anita, Jorma, Minna, Yinka, Ulla, Sari, Päivi, Eeva, Synnöve and others) on the 4th floor served as a rescue team on lousy days and we shared many absolutely fabulous moments during the coffee breaks.

Finally, I want to thank my family and friends for their love and support and most of all for the refreshing moments outside work. My dear husband, Lauri, has experienced my scientific ups and downs in our everyday life, and he deserves special thanks for the encouragement and his realistic attitude towards my temperament.

This study was financially supported by the Research Funds of the Tampere University Hospital, the University of Tampere and the Finnish Cultural Foundation.

Tampere, August 2004

Ulla Aapola

REFERENCES

Altschul SF, Gish W, Miller W, Myers EW and Lipman DJ (1990): Basic local alignment search tool. J Mol Biol 215:403-410.

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W and Lipman DJ (1997): Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402.

Amir RE, Van Den Veyver IB, Wan M, Trans CQ, Francke U and Zoghi HY (1999): Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. Nat Genet 23:185-188.

Aoki A, Suetake I, Miyagawa J, Fujio T, Chijiwa T, Sasaki H and Tajima S (2001): Enzymatic properties of de novo-type mouse DNA (cytosine-5) methyltransferases. Nucleic Acids Res 29:3506-3512.

Aquirre-Arteta AM, Grunewald I, Cardoso MC and Leonhardt H (2000): Expression of an alternative Dnmt1 isoform during muscle differentiation. Cell Growth Differ 11:551-559.

Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA and Struhl K (1995): Short Protocols in Molecular Biology, Third edition. John Wiley & Sons Inc., New York, USA.

Bachman KE, Rountree MR and Baylin SB (2001): Dnmt3a and Dnmt3b are transcriptional repressors that exhibit unique localization properties to heterochromatin. J Biol Chem 276:32282-32287.

Bannister AJ, Zegerman P, Partridge JF, Miska EA, Thomas JO, Allshire RC and Kouzarides T (2001): Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. Nature 410:120-124.

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

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

Bestor TH (1992): Activation of mammalian DNA methyltransferase by cleavage of a Zn binding regulatory domain. EMBO J 11:2611-2617.

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

Bhattacharya SK, Ramchandani S, Cervoni N and Szyf M (1999): A mammalian protein with specific demethylase activity for mCpG DNA. Nature 397:579-583.

Bonfils C, Beaulieu N, Chan E, Cotton-Montpetit J and MacLeod AR (2000): Characterization of the human DNA methyltransferase splice variant Dnmt1b. J Biol Chem 275:10754-10760.

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

Burge C and Karlin S (1997): Prediction of complete gene structures in human genomic DNA. J Mol Biol 268:78-94.

Burgers WA, Fuks F and Kouzarides T (2002): DNA methyltransferases get connected to chromatin. Trends Genet 18:275-277.

Carrozza MJ, Utley RT, Workman JL and Côté J (2003): The diverse functions of histone acetyltransferase complexes. Trends Genet 19:321-329.

Cardoso C, Timsit S, Villard L, Khrestchatisky M, Fontes M and Colleaux L (1998): Specific interaction between the XNP/ATR-X gene product and the SET domain of the human EZH2 protein. Hum Mol Genet 7:679-684.

Cardoso C, Lutz Y, Mignon C, Compe E, Depetris D, Mattei M-G, Fontes M and Colleaux L (2000): ATR-X mutations cause impaired nuclear location and altered DNA binding properties of the XNP/ATR-X protein. J Med Genet 37:746-751.

Carty SM and Greenleaf AL (2002): Hyperphosphorylated C-terminal repeat domain-associating proteins in the nuclear proteome link transcription to DNA/chromatin modification and RNA processing. Mol Cell Proteomics 1:598-610.

Chédin F, Leiber M and Hsieh C-L (2002): The DNA methyltransferase-like protein DNMT3L stimulates de novo methylation by Dnmt3a. Proc Natl Acad Sci 99:16916-16921.

Chen HM, Chrast R, Rossier C, Morris MA, Lalioti MD and Antonarakis SE (1996): Cloning of 559 potential exons of genes of human chromosome 21 by exon trapping. Genome Res 6:747-760.

Chen T, Ueda Y, Xie S and Li E (2002): A novel Dnmt3a isoform produced from an alternative promoter localizes to euchromatin and its expression correlates with active de novo methylation. J Biol Chem 277:38746-39754

Chen T, Ueda Y, Dodge J, Wang Z and Li E (2003): Establishment and maintenance of genomic methylation patterns in mouse embryonic stem cells by Dnmt3a and Dnmt3b. Mol Cell Biol 23:2294-5605.

Cheng X, Kumar S, Posfai J, Pflugrath JW and Roberts RJ (1993): Crystal structure of the HhaI DNA methyltransferase complexed with S-adenosyl-L-methionine. Cell 74:299-307.

Chevillard-Briet M, Trouche D and Vandel L (2002): Control of CBP co-activating activity by arginine methylation. EMBO J 21:5457-5466.

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

Coffee B, Zhang F, Warren ST and Reines D (1999): Acetylated histones are associated with FMR1 in normal but not fragile X-syndrome cells. Nat Genet 22:98-101.

Collins FS, Morgan M and Patrinos A (2003): The human genome project: lessons from large-scale biology. Science 300:286-290.

Constância M, Dean W, Lopes S, Moore T, Kelsey G and Reik W (2000): Deletion of a silencer element in Igf2 results in loss of imprinting independent of H19. Nat Genet 26:203-206.

Cook PR (1997): The transcriptional basis of chromosome pairing. J Cell Sci 110:1033-1040.

Cooke HJ and Saunders PTK (2002): Mouse models of male infertility. Nat Rev Genet 3:790-801.

Czermin B, Schotta G, Hülsmann BB, Brehm A, Becker PB, Reuter G and Imhof A (2001): Physical and functional association of SU(VAR)3-9 and HDAC1 in Drosophila. EMBO Rep 2:915-919.

Daujat S, Bauer U-M, Shah V, Turner B, Berger S and Kouzarides T (2002): Crosstalk between CARM1 methylation and CBP acetylation on histone H3. Curr Biol 12:2090-2097.

Davie JR (1998): Covalent modifications of histones: expression from chromatin templates. Curr Opin Genet Dev 8:173-178.

Davis TL, Yang GJ, McCarrey JR and Bartolomei MS (2000): The H19 methylation imprint is erased and re-established differentially on the parental alleles during male germ cell development. Hum Mol Genet 9:2885-2894.

Dean W, Santos F and Reik W (2003): Epigenetic reprogramming in early mammalian development and following somatic nuclear transfer. Semin Cell Dev Biol 14:93-100.

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

Deplus R, Brenner C, Burgers W, Putmans P, Kouzarides T, de Launoit Y and Fuks F (2002): Dnmt3L is a transcriptional repressor that recruits histone deacetylase. Nucleic Acids Res 30:3831-3838.

de Rooij DG (1998): Stem cells in the testis. Int J Exp Path 79:67-80.

de Rooij DG and Grootegoed JA (1998): Spermatogonial stem cells. Curr Opin Cell Biol 10:694-701.

de Ruijter AJM, van Gennip AH, Kemp S and van Kuilenburg ABP (2003): Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem J 370:737-749.

Di Croce L, Raker VA, Corsaro M, Fazi F, Fanelli M, Faretta M, Fuks F, Lo Coco F, Kouzarides T, Nervi C, Minucci S and Pelicci PG (2002): Methyltransferase recruitment and DNA hypermethylation of target promoters by an oncogenic transcription factor. Science 295:1079-1082.

Dong A, Yoder JA, Zhang X, Zhou L, Bestor, TH and Cheng X (2001): Structure of human DNMT2, an enigmatic DNA methyltransferase homolog that displays denaturant-resistant binding to DNA. Nucleic Acids Res 29:439-448.

Dunham I, Hunt AR, Collins JE, Bruskiewich R, Beare DM, Clamp M, Smink LJ, Ainscough R, Almeida JP, Babbage A, Bagguley C, Bailey J, Barlow K, Bates KN, Beasley O, Bird CP, Blakey S, Bridgeman AM, Buck D, Burgess J, Burrill WD, Burton J, Carder C, Carter NP, Chen Y, Clark G, Clegg SM, Cobley V, Cole CG, Collier RE, Connor RE, Conroy D, Corby N, Coville GJ, Cox AV, Davis J, Dawson E, Dhami PD, Dockree C, Dodsworth SJ, Durbin RM, Ellington A, Evans KL, Fey JM, Fleming K, French L, Garner AA, Gilbert JGR, Goward ME, Grafham D, Griffiths MN, Hall C, Hall R, Hall-Tamlyn G, Heathcott RW, Ho S, Holmes S, Hunt SE, Jones MC, Kershaw J, Kimberley A, King A, Laird GK, Langford CF, Leversha MA, Lloyd C, Lloyd DM, Martyn ID, Mashreghi-Mohammadi M, Matthews L, McCann OT, McClay J, McLaren S, McMurray AA, Milne AS, Mortimore BJ, Odell CN, Pavitt R, Pearce AV, Pearson D, Phillimore BJ, Phillips SH, Plumb RW, Ramsay H, Ramsey Y, Rogers L, Ross MT, Scott CE, Sehra HK, Skuce CD, Smalley S, Smith ML, Soderlund C, Spragon L, Steward CA, Sulston JE, Swann RM, Vaudin M, Wall M, Wallis JM, Whiteley MN, Willey D, Williams L, Williams S, Williamson H, Wilmer TE, Wilming L, Wright L, Hubbard T, Bentley DR, Beck S, Rogers J, Shimizu N, Minoshima S, Kawasaki K, Sasaki T, Asakawa S, Kudoh J, Shintani A, Shibuya K, Yoshizaki Y, Aoki N, Mitsuyama S, Roe BA, Chem F, Chu L, Crabtree J, Deschamps S, Do A, Do T, Dorman A, Fang F, Fu Y, Hu P, Hua A, Kenton S, Lai H, Lao HI, Lewis J, Lewis S, Lin S-P, Loh P, Malaj E, Nguyen T, Pan H, Phan S, Qi S, Qian Y, Ray L, Ren Q, Shaull S, Sloan D, Song L, Wang Q, Wang Y, Wang Z, White J, Willingham D, Wu H, Yao Z, Zhang G, Chissoe S, Murray J, Miller N, Minx P, Fulton R, Johnson D, Bemis G, Bentley D, Bradshaw H, Bourne S, Cordes M, Du Z, fulton L, Goela D, Graves T, Hawkins J, Hinds K, Kemp K, Latreille P, Layman D, Ozersky P, Rohlfing T, Scheet P, Walker C, Wamsley A, Wohldmann P, pepin K, Nelson J, Korf I, Bedell JA, Hillier L, Mardis E, Waterston R, Wilson R, Emanuel BS, Shaikh T, Kurahashi H, Saitta S, Budarf ML, McDermid hE, Johnson A, Wong ACC, morrow BE, Edelmann L, Kim UJ, Shizuya H, Simon MI, Dumanski JP, Peyrard M, Kedra D, Seroussi E, Fransson I, Tapia I, Bruder CE and O'Brien KP (1999): The DNA sequence of human chromosome 22. Nature 402:489-495.

Eden S, Constancia M, Hashimshony T, Dean W, Goldstein B, Johnson AC, Keshet I, Reik W and Cedar H (2001): An upstream repressor element plays a role in Igf2 imprinting. EMBO J 20:3518-3525.

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

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

El-Maarri O, Buiting K, Peery EG, Kroisel PM, Balaban B, Wagner K, Urman B, Heyd J, Lich C, Brannan CI, Walter J and Horsthernke B (2001): Maternal methylation imprints on human chromosome 15 are established during or after fertilization. Nature 27:341-344.

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

Fatemi M, Hermann A, Gowher H and Jeltsch A (2002): Dnmt3a and Dnmt1 functionally cooperate during de novo methylation of DNA. Eur J Biochem 269:4981-4984.

Fedoriw AM, Stein P, Svoboda P, Schultz RM and Bartolomei MS (2004): Transgenic RNAi reveals essential function for CTCF in H19 gene imprinting. Science 303:238-40.

Feinberg AP, Cui H and Ohlsson R (2002): DNA methylation and genomic imprinting: insights from cancer into epigenetic mechanisms. Semin Cancer Biol 12:389-398.

Feinberg AP and Tycko B (2004): The history of cancer epigenetics. Nat Rev Cancer 4:1-11.

Feng Q and Zhang Y (2001): The MeCP1 complex represses transcription through preferential binding, remodeling, and deacetylating methylated nucleosomes. Genes Dev 15:827-832.

Ferguson-Smith AC, Sasaki H, Cattanach BM and Surani MA (1993): Parent-origin-specific epigenetic modification of the mouse H19 gene. Nature 362:751-755.

Fragnioni JV and Neel BG (1993): Solibilization and purification of enzymatically active glutathione S-transferase (pGEX) fusion proteins. Anal Biochem 210:179-187.

Frank D, Keshet I, Shani M, Levine A, Razin A and Cedar H (1991): Demethylation of CpG islands in embryonic cells. Nature 351:239-241.

Fuks F, Burgers W, Brehm A, Hughes-Davies L and Kouzarides T (2000): DNA methyltransferase Dnmt1 associates with histone deacetylase activity. Nat Genet 24:88-91.

Fuks F, Burgers WA, Godin N, Kasai M and Kouzarides T (2001): Dnmt3a binds deacetylases and is recruited by a sequence-specific repressor to silence transcription. EMBO J 20:2536-2544.

Fuks F, Hurd PJ, Wolf D, Nan X, Bird AP and Kouzarides T (2003a): The methyl-CpG-binding protein MeCP2 links DNA methylation to histone methylation. J Biol Chem 278:4035-4040.

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

Gattiker A, Gasteiger E and Bairoch A (2002): ScanProsite: a reference implementation of a PROSITE scanning tool. Applied Bioinformatics 1:107-108.

Ge YZ, Pu MT, Gowher H, Wu HP, Ding JP, Jeltsch A and Xu GL (2004): Chromatin targeting of de novo DNA methyltransferases by the PWWP domain. J Biol Chem 279:25447-25454

Geuns E, De Rycke M, Van Steirteghem A and Liebaers I (2003): Methylation imprints of the imprint control region of the SNRPN-gene in human gametes and preimplantation embryos. Hum Mol Genet 12:28733-2879.

Gibbons RJ, Bachoo S, Picketts DJ, Aftimos S, Asenbauer B, Bergoffen J, Berry SA, Dahl N, Fryer A, Keppler K, Kurosawa K, Levin ML, Masuno M, Neri G, Peirpoint M, Slaney SF and Higgs DR (1997): Mutations in transcriptional regulator ATRX establish the functional significance of a PHD-like domain. Nat Genet 17:146-148.

Gibbons RJ and Higgs DR (2000): Molecular-clinical spectrum of the ATR-X syndrome. Am J Med Genet 97:204-212.

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

Gowher H and Jeltsch A (2001): Enzymatic properties of recombinant Dnmt3a DNA methyltransferase from mouse: the enzyme modifies DNA in a non-processive manner and also methylates non-CpG sites. J Mol Biol 309:1201-1208.

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

Gray SG and Ekström TJ (2001): The human histone deacetylase family. Exp Cell Res 262:75-83.

Griswold MD (1998): The central role of Sertoli cells in spermatogenesis. Semin Cell Dev Biol 9:411-416.

Gunduz M, Ouchida M, Fukushima K, Hanafusa H, Etani T, Nishioka S, Nishizaki K and Shimizu K (2000): Genomic structure of the human ING1 gene and tumor-specific mutations detected in head and neck squamous cell carcinomas. Cancer Res 60:3143-3146.

Hagberg B, Aicardi J, Dias K and Ramos O (1983): A progressive syndrome of autism, dementia, ataxia, and loss of purposeful hand use in girls: Rett's syndrome: report of 35 cases. Ann Neurol 14:471-479.

Hagberg B (1985): Rett's syndrome: prevalence and impact on progressive severe mental retardation in girls. Acta Peadiatr Scand 74:405-408.

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

Hansen RS, Wijmenga C, Luo P, Staned AM, Canfield TK, Weemaes CMR and Gartler SM (1999): The DNMT3B DNA methylransferse gene is mutated in the ICF immunodeficiency syndrome. Proc Natl Acad Sci 96:14412-14417.

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 129:1983-1993.

Hattori M, Fujiyama A, Taylor TD, Watanabe H, Yada T, Park HS, Toyoda A, Ishii K, Totoki Y, Choi DK, Groner Y, Soeda E, Ohki M, Takagi T, Sakaki Y, Taudien S, Blechschmidt K, Polley A, Menzel U, Delabar J, Kumpf K, Lehmann R, Patterson D, Reichwald K, Rump A, Schillhabel M, Schudy A, Zimmermann W, Rosenthal A, Kudoh J, Schibuya K, Kawasaki K, Asakawa S, Shintani A, Sasaki T, Nagamine K, Mitsuyama S, Antonarakis SE, Minoshima S, Shimizu N, Nordsiek G, Hornischer K, Brant P, Scharfe M, Schon O, Desario A, Reichelt J, Kauer G, Blocker H, Ramser J, Beck A, Klages S, Hennig S, Riesselmann L, Dagand E, Haaf T, Wehrmeyer S, Borzym K, Gardiner K, Nizetic D, Francis F, Lehrach H, Reinhardt R and Yaspo ML; Chromosome 21 mapping and sequencing consortium (2000): The DNA sequence of human chromosome 21. Nature 405:311-319.

Hayward BE, De Vos M, Judson H, Hodge D, Huntriss J, Picton HM, Sheridan E and Bonthron DT (2003): Lack of involvement of known DNA methyltranferases in familial hydatidiform mole implies the involvement of other factors in establishment of imprinting in the human female germline. BMC Genet 4:2

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

Horton P and Nakai K (1997): Better prediction of protein cellular localization sites with the k nearest neighbor classifier. Proc Int Conf Intell Syst Mol Biol 5:147-152.

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

Howlett SK and Reik W (1991): Methylation levels of maternal and paternal genomes during preimplantation development. Development 113:119-127.

Hsieh C-L (1999): In vivo activity of murine de novo methyltransferases, Dnmt3a and Dnmt3b. Mol Cell Biol 19:8211-8218.

Hsu D-W, Lin M-J, Lee T-L, Wen S-C, Chen X and Shen JC-K (1999): Two major forms of DNA (cytosine-5) methyltransferase in human somatic tissues. Proc Natl Acad Sci 96:9751-9756.

Huntriss J, Hinkins M, Oliver B, Harris SE, Beazley JC, Rutherford AJ, Gosden RG, Lazendorf SE and Picton HM (2004): Expression of mRNAs for DNA metyltransferases and methyl-CpG-binding proteins in the human female germ line, preimplantation embryos and embryonic stem cells. Mol Reprod Dev 67:323-336.

Iida T, Suetake I, Tajima S, Morioka H, Ohta S, Obuse C and Tsurimoto T (2002): PCNA clamp facilitates action of DNA cytosine methyltransferase 1 on hemimethylated DNA. Genes Cells 7:997-1000.

Iizuka M and Smith MM (2003): Functional consequences of histone modifications. Curr Opin Genet Dev 13:154-160.

International Human Genome Sequencing Consortium (2001): Initial sequencing and analysis of the human genome. Nature 409:860-921.

Ishida C, Ura K, Hirao A, Sasaki H, Toyoda A, Sakaki Y, Niwa H, Li E and Kaneda Y (2003): Genomic organisation and promoter analysis of Dnmt3b gene. Gene 310:151-159.

Jacobson S and Pillus L (1999): Modifying chromatin and concepts of cancer. Curr Opin Genet Dev 9:175-184.

Jeddeloh JA, Stokes TL and Richards EJ (1999): Maintenance of genomic methylation requires a SWI/SNF2-like protein. Nat Genet 22:94-97.

Jones PL, Veenstra GJ, Wade PA, Vermaak D, Kass SU, Landsberger N, Strouboulis J and Wolffe AP (1998): Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription Nat Genet 19:187-191.

Judson H, Hayward BE, Sheridan E and Bonthron DT (2002): A global disorder of imprinting in the human female germ line. Nature 416:539-542.

Kadam S and Emerson BM (2002): Mechanism of chromatin assembly and transcription. Curr Opin Cell Biol 14:262-268.

Kafri T, Ariel M, Brandeis M, Shemer R, Urven L, McCarrey J, Cedar R and Razin A (1992): Developmental pattern of gene-specific DNA methylation in the mouse embryo and germ line. Genes Dev 6:705-714.

Kanduri C, Pant V, Loukinov D, Pugacheva E, Qi CF, Wolffe A, Ohlsson R and Lobanenkov VV (2000): Functional association of CTCF with the insulator upstream of the H19 gene is parent of origin-specific and methylation-sensitive. Curr Biol 10:853-856.

Kang ES, Park CW and Chung JH (2001): Dnmt3b, de novo DNA methyltransferase, interacts with SUMO-1 and Ubc9 through its N-terminal region and is subject to modification by SUMO-1. Biochem Biophys Res Commun 289:862-868.

Kao C-F and Osley MA (2003): In vivo assays to study histone ubiquitylation. Methods 21:59-66.

Khorasanizahed S (2004): The nucleosome: from genomic organization to genomic regulation. Cell 116:259-272.

Kim G-D, Ni J, Kelesoglu N, Roberts RJ and Pradhan S (2002): Co-operation and communication between the human maintenance and de novo DNA (cytosine-5) methyltransferases. EMBO J 21:4183-4195.

Kimura H and Shiota K (2003): Methyl-CpG-binding protein MeCP2, is a target molecule for maintenance DNA methyltransferase, Dnmt1. J Biol Chem 278:4806-4812.

Klimasauskas S, Kumar S, Roberts RJ and Cheng X (1994): HhaI methyltransferase flips its target base out of the DNA helix. Cell 76:357-369.

Kono T, Obata Y, Wu Q, Niwa K, Ono Y, Yamamoto Y, Park ES, Seo J-S and Ogawa H (2004): Birth of parthenogenetic mice that can develop to adulthood. Nature 428:860-864.

Kouzarides T (2002): Histone methylation in transcriptional control. Curr Opin Genet Dev 12:198-209.

Kremer EJ, Pritchard M, Lynch M, Yu S, Holman K, Baker E, Warren ST, Schlessinger D, Sutherland GR and Richards RI (1991): Mapping of DNA instability at the fragile X to a trinucleotide repeat sequence p(CCG)n. Science 252:1711-1714.

Kudoh J, Nagamine K, Asakawa S, Abe I, Kawasaki K, Maeda H, Tsujimoto S, Minoshima S, Ito F and Shimizu N (1997): Localization of 16 exons to a 450-kb region involved in the autoimmune polyglandular disease type I (APECED) on human chromosome 21q22.3. DNA Res 4:45-52.

Kumar S, Cheng X, Klimasauskas S, Mi S, Posfai J, Roberts RJ and Wilson GG (1994): The DNA (cytosine-5) methyltransferases. Nucleic Acids Res 22:1-10.

Kunert N, Marhold J, Stanke J, Stach, D and Lyko F (2003): A Dnmt2-like protein mediates DNA methylation in Drosophila. Development 130:5083-5090.

Lachner M, O'Carrol D, Rea S, Machler K and Jenuwein T (2001): Methylation of histone H3 lysine 9 creates binding site for HP1 proteins. Nature 410:110-120.

Lappalainen I and Vihinen M (2002): Structural basis of ICF-causing mutations in the methyltransferase domain of DNMT3B. Protein Eng 15:1005-1014.

La Salle S, Mertineit C, Taketo T, Moens PB, Bestor TH and Trasler JM (2004): Windows for sex-specific methyaltion marked by DNA methyltransferase expression profiles in mouse germ cells. Dev Biol 268:403-415.

Laskowski RA, MacArthur MW, Moss DS and Thornton JM (1993): PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst 26:283-291.

Lauster R, Trautner TA and Noyer-Weidner M (1989a): Cytosine-specific type II DNA methyltransferases. A conserved enzyme core with variable target-recognizing domains. J Mol Biol 206:305-312

Lauster R (1989b): Evolution of type II DNA methyltransferases. A gene duplication model. J Mol Biol 206:313-321.

Lee J, Inoue K, Ono R, Ogonuki N, Kohda T, Kaneko-Ishino T, Ogurea A and Ishino F (2002): Erasing genomic imprinting memory in mouse clone embryos produced from day 11.5 primordial germ cells. Development 129:1807-1817.

Le Douarin B, Nielsen AL, Garnier J-M, Ichinose H, Jeanmougin F, Losson R and Chambon P (1996): A possible involvement of TIF1 alpha and TIF1 beta in the epigenetic control of transcription by nuclear receptors. EMBO J 15:6701-6715.

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

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

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

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

Li E (2002): Chromatin modification and epigenetic reprogramming in mammalian development. Nat Rev Genet 3:662-673.

Ling V, Wu PW, Finnerty HF, Bean KM, Spaulding V, Fouser LA, Leonard JP, Hunter SE, Zollner R, Thomas JL, Miyashiro JS, Jacobs KA and Collins M (2000): Cutting edge: identification of GL50, a novel B7-like protein that functionally binds to ICOS receptor. J Immunol 164:1653-1657.

Ling Y, Sankpal UT, Robertson AK, McNally JG, Karpova T and Robertson KD (2004): Modification of de novo DNA methyltransferase 3a (Dnmt3a) by SUMO-1 modulates its interaction with histone deacetylases (HDACs) and its capacity to repress transcription. Nucleic Acids Res 32:598-610.

Litt MD, Simpson M, Gaszner M, Allis CD and Feisenfeld G (2001): Correlation between histone lysine methylation and developmental changes at the chicken β-globin locus. Science 293:2453-2455.

Liu K, Wang Y, Cantemir C and Muller MT (2003): Endogenous assays of DNA methyltransferases: evidence for differential activities of DNMT1, DNMT2, and DNMT3 in mammalian cells in vivo. Mol Cell Biol 23:2709-2719.

Lopes S, Lewis A, Hajkova P, Dean W, Oswald J, Forné T, Murrell A, Constância M, Bartolomei M, Walter J and Reik W (2003): Epigenetic modifications in an imprinting cluster are controlled by a hierarchy of DMRs suggesting long-range chromatin interactions. Hum Mol Genet 12:295-305.

Loukinov DI, Pugacheva E, Vatolin S, Pack SD, Moon H, Chernukhin I, Mannan P, Larsson E, Kanduri C, Vostrov AA, Cui H, Niemitz EL, Rasko JEJ, Docquier FM, Kistler M, Breen JJ, Zhuang Z, Quitschke WW, Renkawitz R, Klenova EM, Feinberg AP, Ohlsson R, Morse HC and Lobanenkov VV (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 99:6806-6811.

Lucifero D, Mertineit C, Clarke HJ, Bestor TH and Trasler JM (2002): Methylation dynamics of imprinted genes in mouse germ cells. Genomics 79:530-538.

Lucifero D, Mann MR, Bartolomei MS and Trasler JM (2004): Gene-specific timing and epigenetic memory in oocyte imprinting. Hum Mol Genet 13:839-849.

Lyko F, Ramsahoye BH, Kahevsky H, Tudor M, Mastrangelo M-A, Orr-Weaver TL and Jaenisch R (1999): Mammalian (cytosine-5) methyltransferase cause genomic DNA methylation and lethality in Drosophila. Nat Genet 23:363-366.

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

Margot JB, Aguirre-Arteta AM, di Giacco BV, Pradhan S, Roberts RJ, Cardoso MC and Leonhardt H (2000): Structure and function of the mouse DNA methyltransferase gene: Dnmt1 shows a tripartite structure. J Mol Biol 297: 293-300.

Margot JB, Cardoso MC and Leonhardt H (2001): Mammalian DNA methyltransferses show different subnuclear distributions. J Cell Biochem 83:373-379.

Margot JB, Ehrenhofer-Murray AE and Leonhardt H (2003): Interactions within the mammalian DNA methyltransferase family. BMC Mol Biol 4:7.

Martens J and Winston F (2003): Recent advances in understanding chromatin remodeling by Swi/Snf complexes. Curr Opin Genet Dev 13:136-142.

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

McDowell TL, Gibbons RJ, Sutherland H, O'Rourke DM, Bickmore WA, Pombo A, Turley H, Gatter K, Picketts DJ, Buckle VJ, Chapman L, Rhodes D and Higgs DR (1999): Localization of a putative transcriptional regulator (ATRX) at pericentromeric heterochromatin and the short arms of acrocentric chromosomes. Proc Natl Acad Sci 96:13983-13988.

McGrath J and Solter D (1984): Completion of mouse embryogenesis requires both maternal and paternal genomes. Cell 37:179-184.

McLaren A (2003): Primordial germ cells in the mouse. Dev Biol 262:1-15.

Mertineit C, Yoder JA, Taketo T, Laird DW, Trasler JM and Bestor TH (1998): Sex-specific exons control DNA methyltransferase in mammalian cells. Development 125:889-897.

Mittaz L, Rossier C, Heino M, Peterson P, Krohn KJE, Gos A, Morris MA, Kudoh J, Shimizu N, Antonarakis SE and Scott HS (1999): Isolation and characterization of mouse *Aire* gene. Biochem Biophys Res Commun 255:483-490.

Morris AL, MacArthur MW, Hutchinson EG and Thornton JM (1992): Stereochemical quality of protein structure coordinates. Proteins 12:345-364.

Mund C, Musch T, Strödike M, Assmann B, Li E and Lyko F (2004): Comparative analysis of DNA methylation patterns in transgenic Drosophila overexpressing mouse DNA methyltransferases. Biochem J 378:763-768.

Murphy SK and Jirtle RL (2003): Imprinting evolution and the price of silence. BioEssays 25:577-588.

Murrell A, Heeson S, Bowden L, Constância M, Dean W, Kelsey G and Reik W (2001): An intragenic methylated region in the imprinted Igf2 gene auments transcription. EMBO Rep 2:1101-1106.

Murumägi A, Vähämurto P and Peterson P (2003): Characterization of regulatory elements and methylation pattern of the autoimmune regulator (AIRE) promotor. J Biol Chem 278:19784-19790.

Nagamine K, Peterson P, Scott HS, Kudoh J, Minoshima S, Heino M, Krohn K, Lalioti MD, Mullis PE, Antonarakis SE, Kawasaki K, Asakawa S, Ito F and Shimizu N (1997): Positional cloning of the APECED gene. Nat Genet 17:393-398.

Nan X, Ng HH, Johnson CA, Laherty CD, Turner BM, Eisenman RN and Bird A (1998): Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature 393:386-389.

Ng HH, Zhang Y, Hendrich B, Johnson CA, Turner BM, 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.

Nielsen SJ, Schneider R, Bauer U-M, Bannister AJ, Morrison A, O'Carrol D, Firestein R, Cleary M, Jenuwein T, Herrera RE and Kouzarides T (2001): Rb targets histone H3 methylation and HP1 to promoters. Nature 412:561-565.

Noma K-I, Allis CD and Grewal SIS (2001): Transitions in distinct histone H3 methylation patterns are the heterochromatin domain boundaries. Science 293:1150-1155.

Nowak SJ and Corces VG (2004): Phosphorylation of histone H3: a balancing act between chromosome condensation and transcriptional activation. Trends Genet 20:214-220.

Oakberg E (1956): A description of spermatogenesis in the mouse and its use in analysis of the cycle of the seminiferous epithelium and germ cell renewal. Am J Anat 99:507-515.

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

Ogawa O, Eccles MR, Szeto J, McNoe LA, Yun K, Maw MA, Smith PJ and Reeve AE (1993): Relaxation of insulin-like growth factor II gene imprinting implicated in Wilms' tumour. Nature 362:749-751.

Ohlsson R, Tycko B and Sapienza C (1998): Monoallelic expression: 'there can only be one'. Trends Genet 14:435-438.

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

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

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

Oswald J, Engemann S, Lane N, Mayer W, Olek A, Fundele R, Dean W, Reik W and Walter J (2000): Active demethylation of the paternal genome in the mouse zygote. Curr Biol 10:475-478.

Page SL and Hawley RS (2003): Chromosome choreography: the meiotic ballet. Science 301:785-789.

Panning B and Jaenisch R (1996): DNA hypomethylation can activate Xist expression and silence X-linked genes. Genes Dev 15:1991-2002.

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

Parvinen M and Vanha-Perttula T (1972): Identification and enzyme quantitation of the stages of the seminiferous epithelial wave in the rat. Anat Rec 174:345-350.

Peters AHFM, O'Carrol D, Scherthan H, Mechtler K, Sauer S, Schöfer C, Weipoltshammer K, Pagani M, Lachner M, Kohlmaier A, Opravil S, Doyle M, Sibilia M and Jenuwein T (2001): Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. Cell 107:323-337.

Peterson EJ, Bogler O and Taylor SM (2003): p53-mediated repression of DNA methyltransferase 1 expression by specific DNA binding. Cancer Res 63:6579-6582.

Pitkänen J, Doucas V, Sternsdorf T, Nakajima T, Aratani S, Jensen K, Will H, Vähämurto P, Ollila J, Vihinen M, Scott HS, Antonarakis SE, Kudoh J, Shimizu N, Krohn K and Peterson P (2000): The autoimmune regulator protein has transcriptional transactivating properties and interacts with the common coactivator CREB-binding protein. J Biol Chem 275:16802-16809.

Pitkänen J and Peterson P (2003): Autoimmune regulator: from loss of function to autoimmunity. Genes Immun 4:12-21.

Posfai J, Bhagwat AS, Posfai G and Roberts RJ (1989): Predictive morifs derived from cytosine methyltranferases. Nucleic Acids Res 17:2421-2435.

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

Pradhan S and Kim G-D (2002): The retinoblastoma gene product interacts with maintenance human DNA (Cytosine-5) methyltransferase and modulates its activity. EMBO J 21:779-788.

Qiu C, Sawada K, Zhang X and Cheng X (2002): The PWWP domain of mammalian DNA methyltransferase Dnmt3b defines a new family of DNA-binding folds. Nat Struct Biol 9:217-224.

Quandt K, Frech K, Karas H, Wingender E and Werner T (1995): MatInd and MatInspector – new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res 23:4878-4884.

Ramsahoye BH, Biniszkiewicz D, Lyko F, Clark V, Bird AP 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 97:5237-5242.

Ratnam S, Mertineit C, Ding F, Howell CY, Clarke HJ, Bestor TH, Chaillet JR and Trasler JM (2002): Dynamics of Dnmt1 methyltransferase expression and intracellular localization during oogenesis and preimplantation development. Dev Biol 245:304-314.

Reik W and Walter J (2001): Genomic imprinting: parental influence on the genome. Nat Rev Genet 2:21-32.

Reik W, Dean W and Walter J (2001): Epigenetic reprogramming in mammalian development. Science 293:1089-1093.

Reinisch KM, Chen L, Verdine GL and Lipscomb WN (1995): The crystal structure of HaeIII methyltransferase covalently complexed to DNA: an extrahelical cytosine and rearranged base pairing. Cell 82:143-153.

Rhee I, Bachman KE, Park BH, Jair K-W, Yen R-WC, Schuebel KE, Cui H, Feinberg AP, Lengauer C, Kinzler KW, Baylin SB and Vogelstein B (2002): DNMT1 and DNMT3b cooperate to silence genes in human cancer cells. Nature 416:552-556.

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

Robertson KD, Keyomarsi K, Gonzales FA, Velicescu M and Jones PA (2000a): Differential mRNA expression of the human DNA methyltransferases (DNMTs) 1, 3a and 3b during the G_o/G_1 to S phase transition in normal and tumour cells. Nucleic Acids Res 28:2108-2113.

Robertson KD, Ait-Si-Ali S, Yokochi T, Wade PA, Jones PL and Wolffe AP (2000b): DNMT1 forms a complex with Rb, E2F1 and HDAC1 and represses transcription from E2F-responsive promoters. Nat Genet 25:338-342.

Robertson KD (2002): DNA methylation and chromatin - unraveling the tangled web. Oncogene 21:5361-5379.

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

Rountree MR, Bachman KE and Baylin SB (2000): DNMT1 binds HDAC2 and a new co-repressor, DMAP1, to form a complex at replication foci. Nat Genet 25:269-277.

Saito Y, Kanai Y, Sakamoto M, Saito H, Ishii H and Hirohashi S (2002): Overexpression of a splice variant of DNA methyltransferase 3b, DNMT3b4, associated with DNA hypomethylation on pericentromeric satellite regions during human hepatocarcinogenesis. Proc Natl Acad Sci 99:10060-10065.

Sambrook J, Fritsch EF and Maniatis T (1989): Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, USA.

Sanford JP, Clark HJ, Chapman VM and Rossant J (1987): Differencies in DNA methylation during oogenesis and spermatogenesis and their persistence during early embryogenesis in the mouse. Genes Dev 1:1039-1046.

Santos F, Hendrich B, Reik W and Dean W (2002): Dynamic reprogramming of DNA methylation in the early mouse embryo. Dev Biol 241:172-182.

Sasaki H, Ferguson-Smith AC, Shum ASW, Barton SC and Surani MA (1995): Temporal and spatial regulation of H19 imprinting in normal and uniparental mouse embryos. Development 121:4195-4202.

Sassone-Corsi P (2002): Unique chromatin and transcriptional regulation in spermatogenesis. Science 296:2176-2178.

Scherthan H, Eils R, Trelle-Sticken E, Dietzel S, Cremer T, Walt H and Jauch A (1998): Aspects of three-dimensional chromosome reorganisation during the onset of human male meiotic prophase. J Cell Sci 111:2337-2351.

Schoenherr CJ, Levorse JM and Tilghman SM (2003): CTCF maintains differential methylation at the Igf2/H19 locus. Nat Genet 33:66-69.

Sims III RJ, Nishioka K and Reinberg D (2003): Histone lysine methylation: a signature for chromatin function. Trends Genet 19:629-639.

Sleutels F, Zwart R and Barlow DP (2002): The non-coding Air RNA is required for silencing autosomal imprinted genes. Nature 415:810-813.

Strachan T and Read AP (1999): Human Molecular Genetics 2, John Wiley & Sons Inc., New York, USA.

Strahl BD and Allis D (2000): The language of covalent histone modifications. Nature 403:41-45.

Struhl K (1998): Histone acetylation and transcriptional regulatory mechanisms. Genes Dev 12:599-606.

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

Surani MAH, Barton SC and Norris ML (1984): Development of reconstituted mouse eggs suggests imprinting of the genome during gametogenesis. Nature 308:548-550.

Tang L-Y, Reddy MN, Rasheva V, Lee T-L, Lin M-J, Hung M-S and Shen C-KJ (2003): The eukaryotic DNMT2 genes encode a new class of cytosine-5 DNA methyltransferases. J Biol Chem 278:33613-33616.

Tate PH and Bird AP (1993): Effects of DNA methylation on DNA-binding proteins and gene expression. Curr Opin Genet Dev 3:226-231.

Tatematsu K, Yamazaki T and Ishikawa F (2000): MBD2-MBD3 complex binds to hemi-methylated DNA and forms a complex containing DNMT1 at the replication foci in late S phase. Genes Cells 5:677-688.

Teitell M and Richardson B (2003): DNA methylation in the immune system. Clin Immunol 109:2-5.

Turner G, Webb T, Wake S and Robinson H (1996): Prevalence of fragile X syndrome. Am J Med Genet 64:196-197.

Ueda T, Abe K, Miura A, Yuzuriha M, Zubair M, Noguchi M, Niwa K, Kawase Y, Kono T, Matsuda Y, Fujimoto H, Shibata H, Hayashizaki Y and Sasaki H (2000): The paternal methylation imprint of the mouse H19 locus is acquired in the gonocyte stage during foetal testis development. Genes Cells 5:649-659.

Vandel L and Trouche D (2001): Physical association between the histone acetyl transferase CBP and a histone methyl transferase. EMBO Rep 2:21-26.

Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hannenhalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman

TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov GV, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferriera S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigo R, Campbell MJ, Sjolander KV, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A and Zhu X (2001): The sequence of the human genome. Science 291:1304-1351.

Verger A, Perdomo J, Crossley M (2003): Modification with SUMO. A role in transcriptional regulation. EMBO Rep 4:137-142.

Villard L, Lossi AM, Cardoso C, Proud V, Chiaroni P, Colleaux L, Schwartz C and Fontes M. (1997): Determination of the genomic structure of the XNP/ATRX gene encoding a potential zinc finger helicase. Genomics 43:149-55.

Vukmanovic S, Grandea A, Faas S, Knowles B and Beva M (1992): Positive selection of T-lymphocytes induced by intrathymic injection of a thymic epithelial cell line. Nature 359:729-732.

Walsh CP, Chaillet JR and Bestor TH (1998): Transcription of IAP endogenous retroviruses is constrained by cytosine methylation. Nat Genet 20:116-117.

Watanabe D, Suetake I, Tada T and Tajima S (2002): Stage- and cell-specific expression of Dnmt3a and Dnmt3b during embryogenesis. Mech Dev 118:187-190.

Weisenberger DJ, Velicescu M, Preciado-Lopez MA, Gonzales FA, Tsai YC, Linag G and Jones PA (2002): Identification and characterization of alternatively spliced variants of DNA methyltransferse 3a in mammalian cells. Gene 298:91-99.

Weisenberger DJ, Velicescu M, Cheng JC, Gonzales FA, Liang G and Jones PA (2004): Role of the DNA methyltransferase variant DNMT3b3 in DNA methylation. Mol Cancer Res 2:62-72.

Wilkins MR, Lindskog I, Gasteiger E, Bairoch A, Sanchez J-C, Hochstrasser DF and Appel RD (1997): Detailed peptide characterisation using PEPTIDEMASS - a World-Wide Web accessible tool. Electrophoresis 18:403-408.

Wutz A, Smrzka OW, Schweifer N, Schellander K, Wagner EF and Barlow DP (1997): Imprinted expression of the Igf2r gene depends on an intronic CpG island. Nature 389:745-749.

Xie S, Wang Z, Okano M, Nogami M, Li Y, He WW, Okumura K and Li E (1999): Cloning, expression and chromosome locations of the human DNMT3 gene family. Gene 236:87-95.

Xu GL, Bestor TH, Bourc'his D, Hsieh CL, Tommerup N, Bugge M, Hulten M, Qu X, Russo JJ and Viegas-Pequignot E (1999): Chromosome instability and immunodeficiency syndrome caused by mutations in a DNA methyltransferase gene. Nature 402:187-191.

Xue Y, Gibbons R, Yan Z, Yang D, McDowell TL, Sechi S, Qin J, Zhou S, Higgs D and Wang W (2003): The ATRX syndrome protein forms a chromatin-remodeling complex with Daxx and localizes in promyelocytic leukemia nuclear bodies. Proc Natl Acad Sci 100:10635-10640.

Yanagisawa Y, Ito E, Yuasa Y and Maruyama K (2002): The human DNA methyltrasferases DNMT3A and DNMT3B have two types of promoters with different CpG contents. Biochim Biophys Acta 1577:457-465.

Yen R-WC, Vertino PM, Nelkin BD, Yu JJ, El-Deiry W, Cumarawamy A, Lennon GG, Transk BJ, Celano P and Baylin SB (1992): Isolation and characterization of the cDNA encoding human DNA methyltransferase. Nucleic Acids Res 20:2287-2291.

Yoder JA, Soman NS, Verdine GL and Bestor TH (1997): DNA (cytosine-5)-methyltransferases in mouse cells and tissues. Studies with a mechanism-based probe. J Mol Biol 270:385-395.

Yoder JA and Bestor TH (1998): A candidate mammalian DNA methyltranferase related to pmt1p of fission yeast. Hum Mol Genet 7:279-284.

Yokochi T and Robertson KD (2002): Preferential methylation of unmethylated DNA by mammalian de novo methyltransferase Dnmt3a. J Biol Chem 277:11735-11745.

Yoshinaga SK, Zhang M, Pistillo J, Horan T, Khare SD, Miner K, Sonnenberg M, Boone T, Brankow D, Dai T, Delaney J, Han H, Hui A, Kohno T, Manoukian R, Whoriskey JS and Coccia MA (2000): Characterization of a new human B7-related protein: B7RP-1 is the ligand to the co-stimulatory protein ICOS. Int Immunol 12:1439-1447.

Zimmermann C, Guhl E and Graessmann A (1997): Mouse DNA methyltransferase (Mtase) deletion mutants that retain the catalytic domain display neither de novo nor maintenance methylation activity in vivo. Biol Chem 378:393-405.

Zuklys S, Balciunaite G, Agarwal A, Fasler-Kan E, Palmer E and Hollander GA (2000): Normal thymic architecture and negative selection are associated with Aire expression, the

gene defective in the autoimmune-polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED). J Immunol 165:1976-1983.

Zwart R, Sleutels F, Wutz A, Schinkel AH and Barlow DP (2001): Bidirectional action of the Igf2r imprint control element on upstream and downstream imprinted genes. Genes Dev 15:2361-2366.

ORIGINAL COMMUNICATIONS

The following copyright owners are acknowledged for permission to reprint the original communications:

I Reprinted from Genomics, 65, Aapola U, Shibuya K, Scott HS, Ollila J, Vihinen M, Heino M, Shintani A, Kawasaki K, Minoshima S, Krohn K, Antonarakis SE, Shimizu N, Kudoh J and Peterson P, Isolation and initial characterization of a novel zinc finger gene, *DNMT3L*, on 21q22.3, related to cytosine-5-methyltransferase 3 gene family, 293-298, 2000, with permission from Elsevier.

II Reprinted from Cytogenet Cell Genet, 92, Aapola U, Lyle R, Krohn K, Antonarakis SE and Peterson P, Isolation and initial characterization of the mouse *Dnmt3l* gene, 122-126, 2001, with permission from S. Karger AG, Basel.

III Reproduced with permission, from Aapola U, Mäenpää K, Kaipia A and Peterson P (2004), (Biochem J), (**380)**, (705-713). © the Biochemical Society.

IV Reprinted from Aapola U, Liiv I and Peterson P, Imprinting regulator DNMT3L is a transcriptional repressor associated with histone deacetylase activity, Nucleic Acids Res, 2002, 30(16), 3602-3608, by permission of Oxford University Press